Research in Microbiology 169 (2018) 608–617

Contents lists available at ScienceDirect

Research in Microbiology

journal homepage: www.elsevier.com/locate/resmic

Insights into the biology of acidophilic members of the *Acidiferrobacteraceae* family derived from comparative genomic analyses

Francisco Issotta [a, b], Ana Moya-Beltrán [a, b], Cristóbal Mena [c], Paulo C. Covarrubias [a, b, 1], Christian Thyssen [d], Sören Bellenberg [d], Wolfgang Sand [d, e, f], Raquel Quatrini [a, **, 2], Mario Vera [c, g, *]

[a] *Fundación Ciencia y Vida, Avenida Zañartu 1482, Ñuñoa, 7780272, Santiago, Chile*
[b] *Facultad de Ciencias de la Vida, Universidad Andres Bello, Santiago, Chile*
[c] *Instituto de Ingeniería Biológica y Médica, Escuelas de Ingeniería, Medicina y Ciencias Biológicas, Pontificia Universidad Católica de Chile, Av. Vicuña Mackenna 4860, Macul, 7820486, Santiago, Chile*
[d] *Universität Duisburg Essen, Biofilm Centre, Aquatische Biotechnologie, Universitätsstr. 5, 45141, Essen, Germany*
[e] *College of Environmental Science and Engineering, Donghua University, 2999 North Ren Min Rd., Song Jiang District, Shanghai, 201620, PR China*
[f] *Technische Universität Bergakademie Freiberg, Institut für Biowissenschaften, Leipziger Str. 29, 09599, Freiberg, Germany*
[g] *Departamento de Ingeniería Hidráulica y Ambiental, Escuela de Ingeniería, Pontificia Universidad Católica de Chile, Av. Vicuña Mackenna 4860, Macul, 7820486, Santiago, Chile*

ARTICLE INFO

Article history:
Received 8 April 2018
Accepted 8 August 2018
Available online 22 August 2018

Keywords:
Acidiferrobacteraceae
Bioleaching
Genome analyses
Sulfur oxidation
Rusticianin

ABSTRACT

The family *Acidiferrobacteraceae* (order *Acidiferrobacterales*) currently contains Gram negative, neutrophilic sulfur oxidizers such as *Sulfuricaulis* and *Sulfurifustis*, as well as acidophilic iron and sulfur oxidizers belonging to the *Acidiferrobacter* genus. The diversity and taxonomy of the genus *Acidiferrobacter* has remained poorly explored. Although several metagenome and bioleaching studies have identified its presence worldwide, only two strains, namely *Acidiferrobacter thiooxydans* DSM 2932^T, and *Acidiferrobacter* spp. SP3/III have been isolated and made publically available.

Using 16S rRNA sequence data publically available for the *Acidiferrobacteraceae*, we herein shed light into the molecular taxonomy of this family. Results obtained support the presence of three clades *Acidiferrobacter*, *Sulfuricaulis* and *Sulfurifustis*. Genomic analyses of the genome sequences of *A. thiooxydans*^{T} and *Acidiferrobacter* spp. SP3/III indicate that ANI relatedness between the SPIII/3 strain and *A. thiooxydans*^{T} is below 95–96%, supporting the classification of strain SP3/III as a new species within this genus. In addition, approximately 70% of *Acidiferrobacter* sp. SPIII/3 predicted genes have a conserved ortholog in *A. thiooxydans* strains. A comparative analysis of iron, sulfur oxidation pathways, genome plasticity and cell-cell communication mechanisms of *Acidiferrobacter* spp. are also discussed.

1. Introduction

The monospecies genus *Acidiferrobacter* was proposed in 2011 by Hallberg and colleagues [1], with *Acidiferrobacter thiooxydans* m-1 (DSM 2392^T) as type strain. It corresponded to an anomalous assignation of a strain previously described as *Thiobacillus ferrooxidans* m-1^T [2]. This strain troubled phylogenetic reconstruction of the acidithiobacilli for a long time. Among its differentiating characteristics is a much higher chromosomal G + C content (8–10 mol% higher than *Acidithiobacillus ferrooxidans*). *A. thiooxydans*^{T} is able to grow chemolithoautotrophically by

Abbreviations: RISC, reduced inorganic sulfur compounds; MGEs, mobile genetic elements; ICEs, integrative conjugative elements.

* Corresponding author. Departamento de Ingeniería Hidráulica y Ambiental, Escuela de Ingeniería, Pontificia Universidad Católica de Chile, Av. Vicuña Mackenna 4860, Macul, 7820486, Santiago, Chile.
** Corresponding author. Fundación Ciencia y Vida, Avenida Zañartu 1482, Ñuñoa 7780272, Santiago, Chile.
E-mail addresses: rquatrini@cienciavida.org (R. Quatrini), mariovera@uc.cl (M. Vera).
[1] Present address: uBiome Chile, SpA, Santiago, Chile.
[2] Present address: Millennium Nucleus in the Biology of the Intestinal Microbiota, Facultad de Ciencias de la Vida, Universidad Andres Bello, Santiago, Chile.

https://doi.org/10.1016/j.resmic.2018.08.001

oxidation of ferrous iron or elemental sulfur, sulfide and tetrathionate, using either oxygen or ferric iron as terminal electron acceptors. Strain m-1^T was also shown to be diazotrophic and exhibited tolerance to heavy metals such as aluminum, manganese, iron, zinc and nickel. Minimal inhibitory concentration (MIC) values were reported to be higher than 200 mM for several of these metals. In the case of copper tolerance, the m-1 ^{T} strain was shown to be much more sensitive, with MIC values of 20 mM. Its optimal growth temperature was found to be 38 °C, but it exhibited a certain thermotolerance, since growth was shown to occur up to a maximum of 47 °C. *A. thiooxydans*^{T} is also a moderate osmophile, since growth in liquid media required an external osmotic potential of 2 bar, and was optimal at 5 bar. As a reference, typical culture medium for acidophiles has an osmotic potential of ~2 bars when amended with 4 g/L ferrous iron sulfate. The addition of magnesium sulfate to 100 mM, increases it to ~4.6 bars. This becomes critical in elemental sulfur amended cultures, where *Acidiferrobacter* strains do not grow well, unless magnesium sulfate is added. Under these conditions, a medium with 100 mM magnesium sulfate possess an osmotic potential of ~2.8 bars.

At the time of proposal of the *Acidiferrobacter* genus, it was allocated into the *Ectothiorhodospiraceae* family within the *Gammaproteobacteria*, order *Chromatiales* [1]. Further analyses suggested that *Acidiferrobacter* is not closely affiliated to this family [3]. In 2015 the family *Acidiferrobacteraceae* of the order *Acidiferrobacterales* was proposed [4] currently accommodating three genera of chemolithoautotrophs: *Sulfuricaulis* [4], *Sulfurifustis* [5] and *Acidiferrobacter* [1]. The first two genera are neutrophilic sulfur oxidizers isolated from lake sediments in Japan. *Sulfurifustis variabilis* (DSM 100313^T) is pleomorphic and grows optimally at 42–45 °C and pH values between 6.8 and 8.2 [5], while *Sulfuricaulis limicola* (DSM 100373^T) is not pleomorphic and grows optimally between 28 and 32 °C, at pH values between 6.1 and 9.2 [4]. As mentioned, members of the genus *Acidiferrobacter* are obligate acidophiles and moderately osmophilic iron/sulfur oxidizers. Several strains and sequence clones assigned to this genus have been identified in different natural and man-made environments around the world, including mine waters [6], mine tailings [7], biomining reactors [8], volcanic ashes [9] and macroscopic streamers [10]. A recent study highlighted the importance of several uncultivated genera in performing dark carbon fixation in coastal sediments. This process is light-independent and driven by chemolithotrophs. Several samples from tidal sediments from Europe and Australia were analyzed by combining 16S rDNA sequencing, single cell genomics, transcriptomics and metagenomics coupled to ^{14}C fixation experiments [11]. It was found that three clades were responsible for more than 50% of the dark-carbon fixed. One of these clades was affiliated to *Acidiferrobacter*. Its presence in several ecological niches suggests an important ecological role of this genus, supporting the need for further studies.

Acidiferrrobacter sp. strain SP-III/3 (DSM 27195), was obtained from an acid mine drainage in Cartagena (Murcia, Spain) [7]. It differs from *A. thiooxydans*^{T} in its cell morphology, possessing higher resistance to copper, and a lower optimal growth temperature of 30 °C in comparison to 38 °C for *A. thiooxydans*^{T} (Thyssen et al., in preparation). Recently, 16S rRNA and MLSA-concatenate phylogenetic reconstruction [12], as well as DNA-DNA hybridization analysis, suggested the assignment of strain SP-III/3 to a new species within the *Acidiferrobacter*, different from *A. thiooxydans*^{T}. In addition, another clade that emerged from the 16S rRNA-based phylogeny had values of 16S rRNA sequence divergence to be recognized as a third species [12].

In this work we have sequenced and analyzed the genome sequences of *A. thiooxydans*^{T} and *Acidiferrobacter* sp. SP3/III. In addition, all available draft genome sequences of members of *Acidiferrobacteraceae* have been considered in order to perform a comprehensive comparative genomic study. This will not only help to refine the taxonomic structure of the *Acidiferrobacter*, but also facilitate the identification of metabolic features and further differentiating characteristics of this genus.

2. Materials and methods

2.1. Sequence sources

Data was collected from public databases as of January 2018, including Silva (https://www.arb-silva.de), Nucleotide, WGS, Pubmed and PMC from NCBI (https://www.ncbi.nlm.nih.gov). The genome sequences used in this study were *A. thiooxydans*^{T} (PSYR00000000), *Acidiferrobacter* spp. strain DSM 27195 (SPIII/3) (CP027663), *Acidiferrobacter* spp. ZJ strain (GCF_001705075.1, draft status), *Sf. variabilis* DSM 100313/skN76^T (NZ_AP014936.1) and *Sc. limicola* DSM 100373/HA5^T (NZ_AP014879.1). In addition, public metagenome-derived genomes assigned to *Acidiferrobacteraceae,* from the Tara Oceans circumnavigation expedition [13] available at the WGS database in various contigs (57–208 contigs per genome), were included in the analysis. In house developed scripts were used to retrieve the data by keyword searches and sequence alignments (identity cut-off of 85%; minimal coverage 70%). Small subunit ribosomal RNA genes (16S rRNA) were predicted from genomic and metagenomic assemblies using ribosomal RNA predictors as Barrnap (V0.7) (https://github.com/tseemann/barrnap) and RNAmmer (1.2) [14].

2.2. Genome annotation

Protein coding genes (CDS) were predicted using MetaGeneMark [15]. The functional annotation was done with RPS-BLAST v. 2.7.1 (cutoff E-value of 10^{E-5}) and multiple search-profiles (COG, Pfam, SMART, cd, KEGG, PRK, TIGRFAM) were retrieved from the NCBI Conserved Domain Database (CDD v3.16). To provide a non-redundant annotation, rpsbproc v0.1 was used. Genes in internal clusters were detected using CD-Hit, with thresholds of 70% covered length and 30% sequence identity [16]. Complete genomes were analyzed with additional resources, including SignalP v4.1 and TMHMM v2.0 (http://www.cbs.dtu.dk/services/) for prediction of signal peptides and transmembrane helices, respectively. The algorithms, tRNAscan-SE (http://lowelab.ucsc.edu/tRNAscan-SE) and RNAmmer (http://www.cbs.dtu.dk/services) were used for rRNA gene prediction. CRISPRFinder (http://crispr.i2bc.paris-saclay.fr/Server/) and CRISPR Target (http://bioanalysis.otago.ac.nz/CRISPRTarget/crispr_analysis.html) were used for identification of Clusters of Regularly Interspaced Short Palindromic Repeats (CRISPR) and Cas proteins, respectively. Results from all predictions were curated and parsed using in house developed protocols and complemented with the systems-based curated information available at the RAST annotator server [17]. Prediction of mobile genetic elements MGEs in sequenced *Acidiferrobacteracea*e was performed as in Gonzalez et al. [18]. Genome completeness was inferred as in Raes et al. [19]. Artemis and DNAplotter were used for genome visualization [20].

2.3. Phylogenetic tree reconstruction

The CD-HIT-est webserver (http://cd-hit.org) was used to reduce the redundancy in the 16S rRNA gene sequence dataset.

Gene sequences were aligned with MAFFT v7.229 software using the L-INS-I method [21]. The resulting alignments were trimmed and masked (>100%) with MEGA7 [22]. Phylogenetic tree reconstruction was done by Bayesian inference with Mr. Bayes (http://mrbayes.sourceforge.net) using 10^6 generations and sampling every 100 generations. The protocol was run twice and the results were summarized with SumTrees for further support (http://dendropy.org). The resulting trees were visualized in FigTree (http://tree.bio.ed.ac.uk/software/figtree). Oligotyping was performed with the "Oligo typing" (V2.1) software (http://merenlab.org/software/oligotyping), as described [23].

2.4. Comparative genomics

Shared and exclusive genes were identified using GET_HOMOLOGUES v.07112016 (https://github.com/eead-csic-compbio/get_homologues). Orthology was determined based on all-versus-all Best Bidirectional BlastP Hits, using COGtriangles v2.1 [24] as clustering algorithm. Pairwise alignment cutoffs were set at 75% coverage and E-value of 10^{E-5}. The phylogenomic relationships between the *Acidiferrobacter* strains were inferred from the Average Nucleotide Identity (ANI) values assessed by BLASTn and the TETRA indexes using the pyani software v0.2.7 (https://github.com/widdowquinn/pyani).

2.5. Biogeographycal pattern analysis

In house scripts were used to parse the data retrieved from different databases and analyses performed in this study. The world map was created in R v3.4.2 (https://www.r-project.org) using the tools and libraries ggmap [25], maptools [26], and maps [27], to display the parsed data.

3. Results and discussion

3.1. Occurrence and distribution of the Acidiferrobacteraceae family acidophilic members

To gain insights into the global genetic diversity of the acidophilic members of the *Acidiferrobacteraceae* family, the geographic distribution of cognate 16S rRNA gene sequences derived from public gene, genome and metagenome databases was assessed and analyzed. In order to uncover meaningful differences between available strains and sequence clones at the 16S rRNA gene level, sequences were aligned, used to build phylogenetic trees and oligotyped. A total of 171 representatives assigned to the family or matching the queries used in sequence similarity searches were retrieved from GenBank (Supplementary Table 1), and plotted in the map to derive biogeographical information on the taxon (Fig. 1). This set included the 16S rRNA gene sequences derived from genomic drafts obtained for isolates and from metagenome-assembled genomes, termed MAGs. Yet, most 16S rRNA gene data available correspond to uncharacterized sequence clones. Several of these have been only partially sequenced.

For phylogenetic analysis, a set of filtering criteria was applied to these data (sequence redundancy, length, masks, etc., see Materials and Methods). A final set of 84 sequences was obtained (Supplementary Table 2). This set encompassed 1228 bp of the full 16S rRNA gene sequence spanning the V2–V8 region and contained 411 variable sites and 284 parsimony informative sites. Within this data set, 9.3% of the sequences had taxonomic assignment to the species level. The phylogenetic tree built for this dataset using maximum likelihood inference is shown in Supplementary Fig. 1. The phototrophic sulfur bacterium *Chromatium okenii* DSM 169 was used as outgroup.

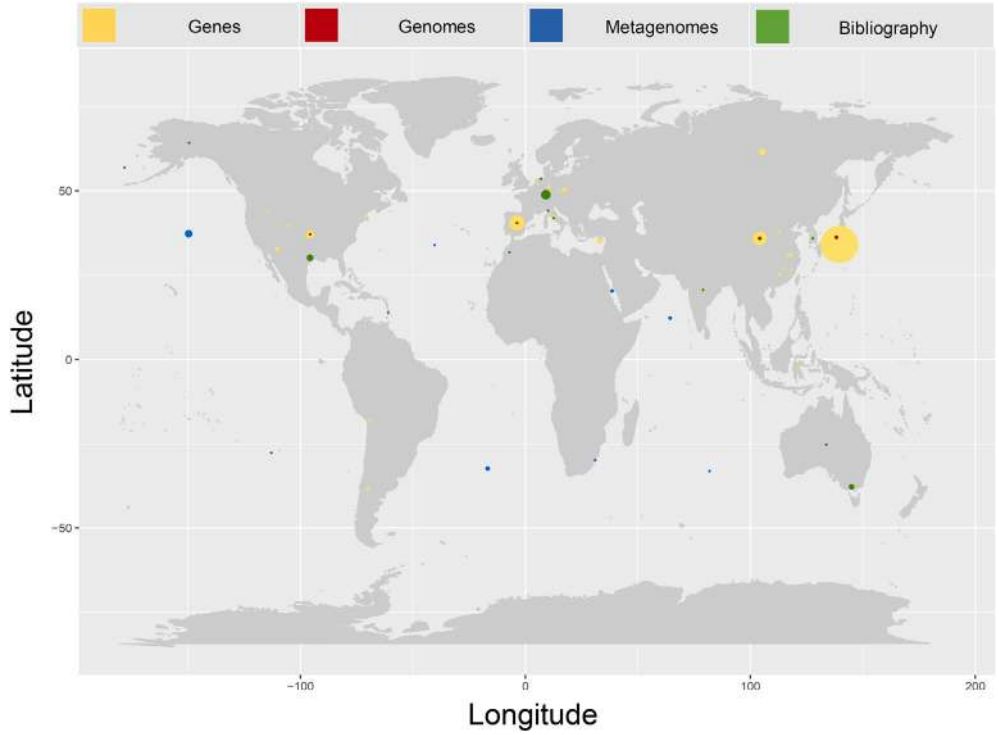

Fig. 1. Geographical distribution and relative abundance of *Acidiferrobacter* representatives based on 16S rRNA gene sequence data. Sequence data used in the analysis is summarized in Supplementary Table 1, and includes 16S rRNA genes retrieved from Genes (119 sequences), Genome (5 sequences) and Metagenome (17 sequences) databases, as well as from bibliographic references (30 sequences). Color-coding is indicated in the bar above the map.

The phylogenetic analysis showed that *A. thiooxydans*^{T} (Clade 1), *Sc. limicola*^{T} (Clade 2) and *Sf. variabilis*^{T} (Clade 3) branch apart in the tree, in three well-defined clades. The analysis also revealed inherent sequence variability within each of the main clades, as determined by the presence of several subclades (or sister clades). To further characterize this variability all sequences were oligotyped. Oligotyping analysis resulted in twelve information-rich positions, spanning variable regions of the 16S rRNA gene V2 to V8. A total of 38 oligotypes (OTs) were derived from the dataset (Supplementary Table 2; Supplementary Fig. 1). Of these, three OT-defined groups matched exactly the tree branches defining currently recognized *Acidiferrobacteraceae*, i.e. *A. thiooxydans*^{T} (OT-1), *Sc. limnicola*^{T} (OT-20) and *Sf. variabilis*^{T} (OT30). However, additional OTs matching unassigned sequence representatives clustering in coherent subclades or sister clades within the main tree branches, were also identified. This is the case of subclade 1.b in the 1st branch (Clade 1), which accommodates sequences representing 2 sequence clones bearing a distinctively different oligotype (OT2) and one cultivated strain, (*Acidiferrobacter* sp. SPIII/3). The OT 1 related to the type species of the family, *A. thiooxydans*^{T} (Sc.1a), was found to be present in pyrite, copper and coal mining sites worldwide as well as in the acidic river Rio Tinto in Spain (Fig. 1; Supplementary Table 2). Similarly, the OT associated to SPIII/3 (OT 2; Sc.1b), was found to occur in a distant location with respect to the isolation origin of strain SPIII/3 (in Murcia, Spain), of clearly acidic nature. A third subclade within the family (OT 4; Sc.1c), of presently unclear taxonomic affiliation (to the species level), was found to occur in volcanic acidic ash deposits in the Japanese island of Miyake. Neutrophilic members of the family have been found in all types of benthic habitats ranging from intertidal sediments to deep-sea hydrothermal chimneys [11], and references therein. According to our 16S rRNA-based oligotyping analyses, acidophilic *Acidiferrobacteraceae* appear to be similarly ubiquitous (ash, mud, rivers of volcanic origin; ore and AMD from mining sites; sulfidic caves), spanning a wide diversity of anthropogenic and natural habitat types.

3.2. *Genomic relatedness of sequenced* Acidiferrobacteraceae *family members*

The genome of sulfur oxidizing strains belonging to *Acidiferrobacteraceae* have been made publically available recently [28]. In this study we have included also three additional sequences of acidophilic representatives (see Materials & Methods). Together with these, there are a number of metagenome-derived draft genomic assemblies of *Acidiferrobacteraceae* available, all recovered from the Tara Oceans Consortium project (Supplementary Table 3). Using this information, the genomic relatedness between acidophilic and neutrophilic representatives of the taxon was inspected. Average Nucleotide Identity was calculated using BLAST-based sequence similarity analysis (ANIb) and the genomic divergence between strains and emergent strain-clusters was measured through the TETRA index. Results are summarized in Table 1. Four strain-clusters were recovered from these analyses, showing coherent genomic similarity levels. Both indexes show important levels of divergence for the great majority of genomes in the data set, suggesting that a taxonomic revision of the whole family, after thorough resampling and sequencing should be considered. In the case of the acidophilic representatives, the ANI relatedness between the SPIII/3 strain and *A. thiooxydans*^{T} was below the established 95–96% ANI, used for prokaryotic species delimitation [29]. The same holds true for the rest of the strains and strain-clusters in the comparison (Table 1). These results indicate that at genome level there is sufficient divergence to consider the subclade 1b (represented by strain SPIII/3) as a new *Acidiferrobacter* species.

Table 1
Genome relatedness indexes calculated for sequenced isolates and metagenome-assembled genomes (MAGs) of the *Acidiferrobacteraceae* family.

TETRA/ANIB	M-1	ZJ	SPIII/3	HA5	skN76	NP54	NP97	ARS46	NP122	IN47	SAT1416	RS446	NAT194	SAT1475	NP79	EAC671	NP959	SAT1331	NP24	ARS29	RS826	SP295
m-1	100	0,99	0,91	0,72	0,72	0,70	0,69	0,70	0,69	0,69	0,69	0,69	0,69	0,69	0,69	0,68	0,68	0,68	0,69	0,69	0,67	0,67
ZJ	1,00	100	0,91	0,72	0,72	0,69	0,69	0,69	0,69	0,69	0,69	0,69	0,69	0,69	0,69	0,68	0,68	0,69	0,68	0,68	0,67	0,67
SPIII/3	0,99	0,99	100	0,72	0,72	0,70	0,69	0,70	0,69	0,69	0,69	0,69	0,69	0,69	0,69	0,68	0,68	0,68	0,69	0,68	0,68	0,67
HA5	0,74	0,75	0,78	100	0,74	0,70	0,69	0,70	0,69	0,70	0,69	0,69	0,69	0,69	0,69	0,68	0,68	0,68	0,68	0,68	0,68	0,67
skN76	0,71	0,71	0,75	0,87	100	0,70	0,69	0,70	0,69	0,70	0,69	0,69	0,69	0,69	0,69	0,68	0,68	0,68	0,68	0,68	0,67	0,67
NP54	0,65	0,66	0,66	0,77	0,64	100	0,72	0,77	0,71	0,71	0,71	0,71	0,71	0,70	0,73	0,70	0,70	0,70	0,69	0,71	0,69	0,68
NP97	0,65	0,65	0,65	0,71	0,64	0,93	100	0,73	0,72	0,71	0,70	0,70	0,70	0,70	0,70	0,70	0,70	0,70	0,69	0,69	0,69	0,68
ARS46	0,63	0,63	0,63	0,72	0,61	0,95	0,92	100	0,71	0,76	0,75	0,75	0,75	0,75	0,74	0,70	0,69	0,70	0,70	0,70	0,69	0,68
NP122	0,57	0,57	0,56	0,64	0,51	0,88	0,93	0,89	100	0,70	0,70	0,70	0,69	0,70	0,70	0,70	0,70	0,71	0,69	0,69	0,71	0,69
IN47	0,53	0,53	0,52	0,57	0,46	0,84	0,82	0,93	0,87	100	0,96	0,98	0,98	0,89	0,72	0,69	0,70	0,69	0,70	0,69	0,69	0,68
SAT1416	0,51	0,51	0,50	0,58	0,47	0,85	0,82	0,91	0,88	0,98	100	0,95	0,96	0,88	0,85	0,69	0,69	0,69	0,69	0,69	0,69	0,68
RS446	0,51	0,51	0,49	0,53	0,42	0,83	0,82	0,92	0,88	0,99	0,98	100	0,97	0,88	0,87	0,69	0,69	0,69	0,69	0,69	0,69	0,68
NAT194	0,51	0,50	0,49	0,52	0,41	0,83	0,82	0,91	0,88	0,99	0,97	0,99	100	0,88	0,87	0,69	0,69	0,69	0,69	0,69	0,68	0,68
SAT1475	0,50	0,50	0,48	0,52	0,40	0,83	0,80	0,90	0,86	0,99	0,98	0,99	0,98	100	0,87	0,69	0,69	0,69	0,69	0,69	0,68	0,69
NP79	0,52	0,52	0,51	0,54	0,41	0,87	0,88	0,89	0,89	0,86	0,72	0,72	0,72	0,71	100	0,69	0,69	0,70	0,70	0,69	0,69	0,68
EAC671	0,43	0,43	0,41	0,49	0,33	0,80	0,77	0,81	0,85	0,86	0,87	0,86	0,86	0,87	0,78	100	0,94	0,98	0,68	0,68	0,69	0,68
NP959	0,42	0,43	0,40	0,50	0,35	0,81	0,78	0,82	0,86	0,86	0,89	0,87	0,86	0,87	0,79	0,99	100	0,94	0,69	0,69	0,69	0,68
SAT1331	0,42	0,42	0,40	0,47	0,33	0,79	0,77	0,81	0,85	0,85	0,86	0,85	0,85	0,86	0,78	1,00	0,99	100	0,68	0,68	0,69	0,68
NP24	0,41	0,42	0,41	0,53	0,48	0,64	0,69	0,65	0,71	0,69	0,72	0,68	0,67	0,68	0,56	0,66	0,69	0,64	100	0,98	0,68	0,68
ARS29	0,37	0,38	0,36	0,45	0,38	0,62	0,67	0,63	0,70	0,71	0,72	0,69	0,69	0,69	0,58	0,69	0,71	0,67	0,98	100	0,69	0,68
RS826	0,37	0,37	0,37	0,56	0,46	0,64	0,58	0,63	0,62	0,65	0,72	0,64	0,62	0,66	0,56	0,68	0,73	0,67	0,71	0,68	100	0,71
SP295	0,34	0,34	0,33	0,52	0,45	0,59	0,54	0,57	0,59	0,62	0,70	0,61	0,60	0,65	0,52	0,68	0,73	0,67	0,71	0,67	0,94	100

Indexed used include: the Average Nucleotide Identity calculated using BLAST-based sequence identity analysis (ANIb, in black) and the tetranucleotide index (TETRA, in blue). Values obtained for each pairwise comparison are highlighted in color from the maximum (in green) to minimum (in red). Threshold values used for prokaryotic species delimitation were 95-96 % and 99 %, respectively [29,47,48]. Information on the genomes and MAGs sequences used in the analysis is detailed further in Table 2 and Supplementary Table 3.

3.3. General genomic features of the acidophilic Acidiferrobacter spp.

Relevant genome attributes and aspects of the annotation of *A. thiooxydans* strains (m-1^T and ZJ) and the only sequenced representative of the *Acidiferrobacter* SPIII/3 strain-cluster are summarized in Table 2. Neutrophilic *Acidiferrobacteraceae* family members, *Sf. variabilis*^{T} and *Sc. limnicola*^{T}, are included for comparative purposes. The total size of the genome of *A. thiooxydans* strains m-1^T and ZJ is similar (3.25 Mbp), while the *Acidiferrobacter* SPIII/3 genome is on average 3.1% larger (~3.4Mbp). Both *Acidiferrobacter* genomes are smaller than the ~3.96 Mb genome of *Sf. variabilis* and larger than that of *Sc. limnicola* [28]. The average G + C content of the SPIII/3 strain genome (64.2 mol %) is also higher than those of the *A. thiooxydans* (63.6 mol %) and *Sc. limnicola*^{T} (63 mol %) but lower than *Sf. variabilis*^{T} (69 mol%).

In addition to the global genomic differences uncovered in the preceding section (section 3.2) between SPIII/3 and the two *A. thiooxydans*, strains (m-1^T and ZJ), significant differences in gene content were also identified through comparative genomic analyses (Fig. 2A). Approximately 70% of *Acidiferrobacter* sp. SPIII/3 predicted genes have a conserved ortholog in both *A. thiooxydans* strains, while an additional 2.6%–3.7% of the SPIII/3 genes are conserved only in the m-1^T or in the ZJ strain, respectively. Average amino acidic sequence level divergence (100 − identity of amino-acidic sequence percent) between protein coding genes shared by all three strains is 3.7%. In turn, partially shared gene sets have higher average levels of divergence (36.5%–26.9%), suggesting that these are probably part of the itinerant mobilome. Functional assignment distribution and abundance (gene counts in each COG category) for the shared and partially shared gene orthologs, and the exclusive gene complement of *Acidiferrobacter* SPIII/3 is shown in Fig. 2B. *Acidiferrobacter* sp. strain SPIII/3 exclusive gene complement spans most COG categories. However, the largest amount of SPIII/3 exclusive genes fall into categories R, S, L and C. Categories R (General function prediction only) and S (Function unknown) group poorly characterized genes, and category L groups genes predicted to have a role in replication, recombination and repair. All three categories are typically enriched in mobile genetics elements [30]. The fourth category (C), appearing to be highly represented in the exclusive gene complement of the SP3/III strain, contains genes involved in energy production and conversion.

3.4. The exclusive gene complement of acidophilic Acidiferrobacter SPIII/3

To derive further insights into the distribution of shared and exclusive genes in the genome of interest (*Acidiferrobacter* SPIII/3), its predicted proteome was crossed compared against the other acidophilic *Acidiferrobacteraceae* and the neutrophilic sequenced representatives available (see Materials and Methods 2.4). Mobile genetic element (MGE) prediction tools were used to discriminate the mobilome (gene complement linked to MGEs) from the rest of the genome. As detailed bellow, several genomic segments of strain SPIII/3 (Fig. 3, segments A-M) are predicted to have features typical of foreign sequences (Supplementary Table 4). A number of these segments contain genes totally missing (e.g. segment A), or only partially conserved (e.g. segment D), in the genomes of *A. thiooxydans*^{T}, *Sf. variabilis*^{T} and *Sc. limnicola*^{T}. These exclusive genomic segments entail putative transposons (e.g. segment A; predicted to transpose through the action of a transposase), integrative mobile elements (e.g. segment E; predicted to integrate by site specific recombination mediated by a cognate integrase), and mobile retroelements (e.g. segment H [31]; predicted to retro-transpose via a RNA intermediate that is reverse-transcribed to

Table 2
Genome sequence attributes of sequenced *Acidiferrobacteraceae*.

	Acidiferrobacter sp.	*Acidiferrobacter thiooxydans*	*Acidiferrobacter thiooxydans*	*Sulfurifustis variabilis*	*Sulfuricaulis limicola*
Strain ID	SPIII/3 (ACM)	DSM 2392/m-1^T (ACT)	ZJ (ACT)	DSM 100313/skN76T (SFV)	DSM 100373/HA5T (SCL)
Origin	Spain, Murcia	USA, Missouri	China, Fujian	Japan, Lake Mizugaki	Japan, Lake Mizugaki
Sample	Acid mine drainage	Coal spoil refuse	Acid mine drainage	Sediment	Sediment
pH range	1.0–2.5	1.2–2.5	–	6.3–8.9	6.1–9.2
T^o range	15–30 °C	20–47 °C	–	28–46 °C	8–37 °C
BioSample ID	SAMN08523059	SAMN08523060	SAMN05526184	SAMD00036173	SAMD00031266
Seq. Technol.	PacBio; Illumina MiSeq	PacBio	Illumina MiSeq	PacBio RSII	PacBio RSII
WGS	CP027663	PSYR00000000	NZ_MDCF00000000.1	NZ_AP014936.1	NZ_AP014879.1
Assembly	–	–	ASM170507v1	ASM235541v1	ASM235573v1
Scaffolds	1	4	257	1	1
N50	3,398,398	1,759,369	28.166	–	–
Depth coverage	150×	400×	130.0×	–	–
Finishing quality	High quality draft	High quality draft	Draft	Finished	Finished
Completeness (%)[a]	100%	100%	100%	100%	100%
Genbank ID	CP027663	PSYR00000000	MDCF00000000.1	AP014936	AP014879
G + C content (%)	64.2	63.7	63.6	67.5	61.4
Genome size	3,398,398	3,252,820	3,252,820	3,958,814	2,864,672
Total genes	3403	3163	2938	3849	2775
Protein	3307	3067	2886	3760	2708
RNA genes	48	48	52	51	53
rRNA	3	3	3	3	3
tRNA	45	45	45	44	45
Other RNA	0	0	4	4	5
Pseudogenes	0	0	175	38	14
Genes in internal clusters	2689	2562	2750	3387	2519
Genes with CDD hits	2724	3067	3149	3316	2405
Genes assigned to COGs	2239	2196	2147	2801	2099
Genes with Pfam domains	2360	2263	2196	2962	2176
Genes with signal peptides	124	126	119	440	276
Genes with transmembrane helices	646	619	649	1040	766
CRISPR repeats	5	0	1	0	0

[a] Completeness percentage by identification universal housekeeping genes.

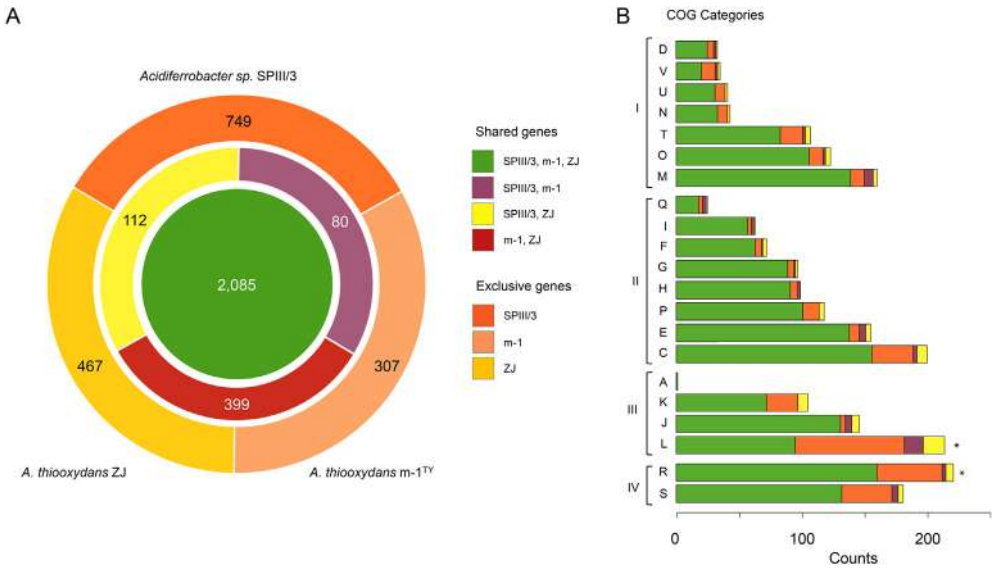

Fig. 2. Comparative analysis of the predicted gene content of *Acidiferrobacter* sp. SPIII/3 against *A. thiooxydans* sequenced strains m-1^T and ZJ. (A) Number of shared, partially shared and exclusive genes for the three strains. (B) Functional assignment distribution of the SPIII/3 related gene complements. Color-coding is as follows: SPIII/3 exclusive genes (orange), genes shared in all three strains (green), partially shared genes between SPIII/3 and m-1^T (purple), partially shared genes between SPIII/3 and ZJ (yellow). COG categories are displayed as follows: I) Cellular processes and signaling: [D], Cell cycle control, cell division, chromosome partitioning, [V] Defense mechanisms, [U] Intracellular trafficking, secretion, and vesicular transport, [N] Cell motility, [T] Signal transduction mechanisms, [O] Posttranslational modification, protein turnover, chaperones, [M] Cell wall/membrane/envelope biogenesis. II) Metabolism: [Q] Secondary metabolites biosynthesis, transport and catabolism, [I] Lipid transport and metabolism, [F] Nucleotide transport and metabolism, [G] Carbohydrate transport and metabolism, [H] Coenzyme transport and metabolism, [P] Inorganic ion transport and metabolism, [E] Amino acid transport and metabolism, Nucleotide transport and metabolism, [C] Energy production and conversion; III) Information storage and processing: [A] RNA processing and modification, [K] Transcription, [J] Translation, ribosomal structure and biogenesis, [L] Replication, recombination and repair; IV) Poorly characterized: [R] General function prediction only, [S] Function unknown. Functional categories grouping the largest counts of exclusive genes are indicated with an asterix.

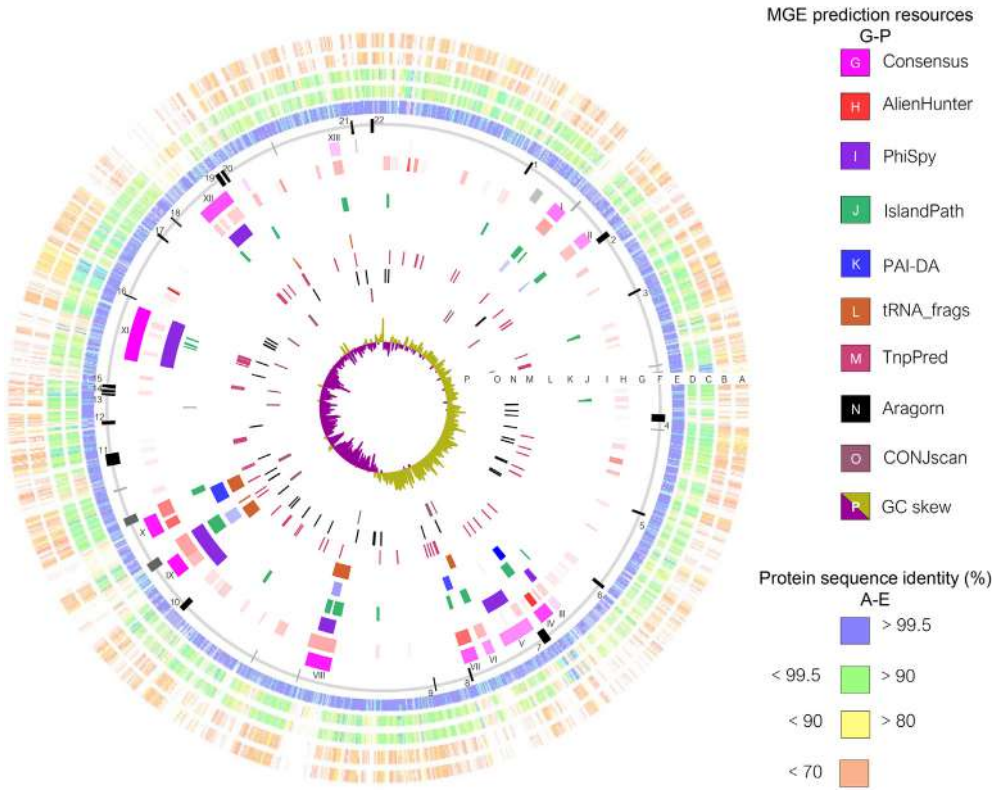

Fig. 3. *Acidiferrobacter* SPIII/3 whole-genome sequence comparative analysis. External rings display conserved orthologs between *Acidiferrobacter* SPIII/3 and the genomes of *Sc. limnicola* (A), *Sf. variabilis* (B), *A. thiooxydans* ZJ (C), *A. thiooxydans* m-1^T (D), and a second sequencing run of *Acidiferrobacter* SPIII/3 (E). Internal rings (from outside to inside) display the genomic location of predicted key energy metabolism genes and gene clusters (numbered 1 to 22 on F ring), and predicted MGEs (numbered I to XIII). Resources used in the delimitation of MGEs are color-coded in the right side bar.

DNA by a reverse transcriptase encoded within the element). A few of these segments could be delimited by flanking direct repeats (*attL* and *attR*) or duplicated tRNAs and truncated tRNAs (e.g. segment E). Seven of the predicted MGEs encode repeat-adjacent integrases (int, COG 4974), predicted to mediate unidirectional site-specific recombination of the mobile element in the host genome. Another four MGEs encode components of type IV secretion systems (T4SS) for conjugative transfer, including the VirD2 relaxase and the VirD4 coupling protein, typically present in non-conjugative mobilizable plasmids. The essential components of a functional T4SS (VirB1 to VirB11) secretion system are absent for the SPIII/3 genome, suggesting that these MGEs are better defined as an Integrative Mobilizable Element (IME). Another two segments (H and I) encode group II intron reverse transcriptases, predicted to drive mobilization of a cDNA copy of the mobile elements into a new location of the target DNA [32]. Two others segments encode plasmid-type replicases (J and K), suggesting that these entail integrated plasmids or their remnants. All the genomic regions alluded above encode defense systems against foreign DNA invasion (CRISPRs and restriction modification systems), copper tolerance genes, surface modification genes and a vast amount of hypothetical genes of unknown function (Supplementary Table 4). A few genes encoding key metabolic features playing a role in energy production and conversion were identified between the predicted MGEs or in their close vicinity (e.g. dissimilatory sulfite reductase *dsr* enzyme complex; see below), suggesting these could also have been the object of recent horizontal transfer.

3.5. Acidiferrobacter *spp. core genome attributes: energy acquisition mechanisms*

Genome-based metabolic reconstruction of gene features of the core genome of acidophilic *Acidiferrobacter* spp. confirms them to be chemolithotrophic autotrophs capable of using a diversity of electron donors including ferrous iron, as well as a variety of reduced sulfur compounds (RISC), polythionates and thiocyanate. It also highlighted that several electron acceptors can be used, such as O_2, Fe^{3+}, NO_3^-, as well as some RISC. All three *Acidiferrobater* strains analyzed encode in their genomes genes and pathways for iron oxidation, yet the most prominent aspect of their metabolism is the oxidation of reduced sulfur compounds (Fig. 4; Supplementary Table 5). As the non-acidophilic *Acidiferrobacteraceae* sequenced members, *Sc. limnicola*^{T} and *Sf. variabilis*^{T}, the acidophilic *Acidiferrobater* SPIII/3, m-1^T and ZJ genomes contain a full repertoire of *dsr* genes, involved in the oxidation of elemental sulfur to sulfite, two sets of *hdr* genes, involved in the oxidation of disulfide intermediaries (from sulfur oxidation) to sulfite and two sets of *apr* genes, involved in oxidation of sulfite to sulfate. All three genomes also encode the dissimilatory thiosulfate oxidation Sox multienzyme system [33,34]. Also, all five *Acidiferrobacteraceae* sequenced representatives encode a membrane bound polysulfide reductase-like Fe–S molybdoprotein (*soeABC*), presumably involved in sulfite oxidation in the cytoplasm [28]. A gene ortholog of the sulfur oxygenase reductase (SOR), potentially mediating elemental sulfur oxidation to sulfite, thiosulfate and hydrogen sulfide, is also found in acidophilic and non-acidophilic *Acidiferrobacteraceae* (79% similarity to *Acidithiobacillus ferrivorans* AEM48683). In turn, the tetrathionate intermediate pathway S4I [35] present in acidithiobacilli [36,37], which in *At. ferrooxidans*^{T} is dependent on a tetrathionate hydrolase (*tetH*, AFE_0029), a thiosulfate quinone reductase (*doxDA*, AFE_0044) and a putative sulfide quinone reductase (*sqr*, AFE_1792) among other proteins, is entirely missing in *Acidiferrobacter*.

The dissimilatory sulfite reductase *dsr* enzyme complex [38] is the second part of the Branched Thiosulfate Oxidation pathway (BTO) [39,40], mediating the cytoplasmic transport and respiratory

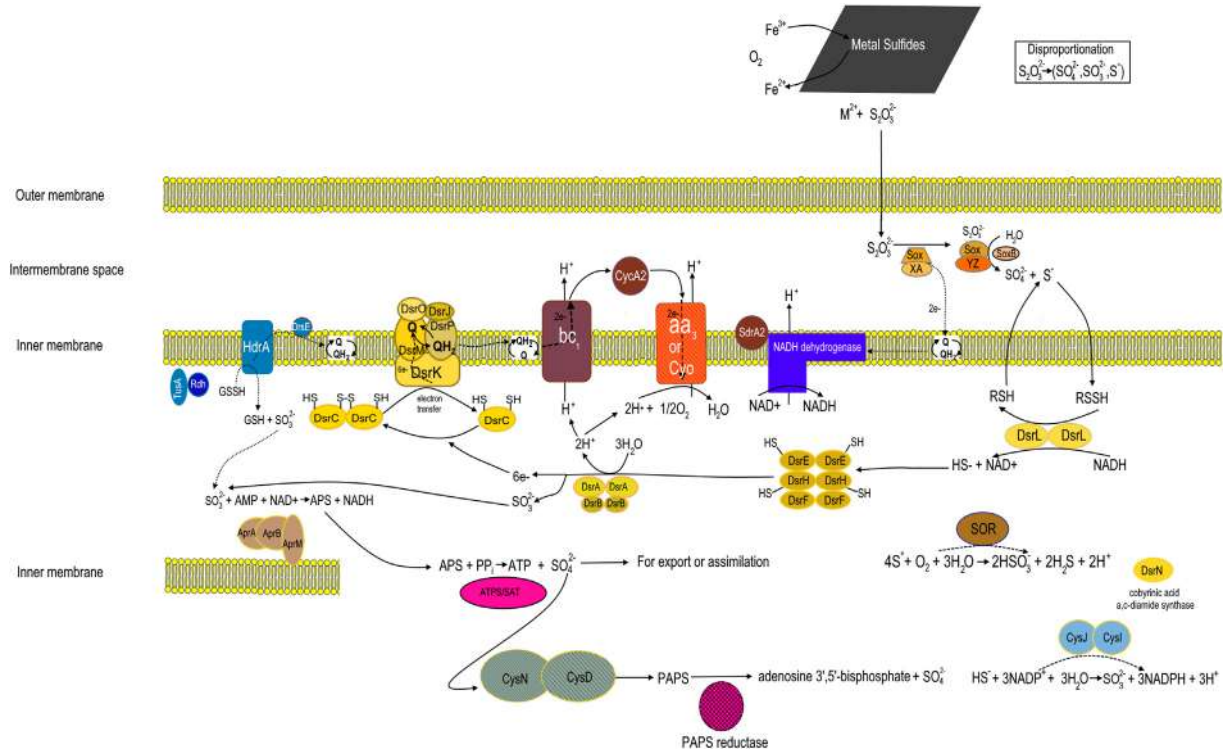

Fig. 4. Model for sulfur oxidation in *Acidiferrobacter* species. Potential transference of electrons is represented by the thick dotted lines. Proteins encoded by *soxXYZAB* are in the intermembrane space/periplasm. Proteins encoded by the *Dsr* cluster in yellow. In other colors, proteins encoded by other clusters and *loci* predicted to be involved in elemental sulfur and RISC oxidation.

oxidation of sulfur into sulfite with the transference of electrons to the quinone pool. In *Acidiferrobacter* species this complex is encoded in a gene cluster consisting of thirteen *dsr* subunits *dsrABEFHCMKLJOPN* (Supplementary Table 5), a conserved protein of unknown function, a two-component response regulator/histidine-kinase similar to the *narQ/narL* that could mediate sulfur homeostasis (COG3850 and COG2197, e-values $2.0e^{-56}$ and $1.13e^{-67}$ respectively), and a metallophosphoesterase-like protein of unknown function (cd07424, e-value $1.41e^{-46}$). No ortholog of the *dsrD* subunit that allows for the usage of sulfate as an electron acceptor was predicted in sequenced *Acidiferrobacter* genomes [39]. The complete *dsr* operon is located in between some transposases, suggesting its origin as a mobile element of horizontal genetic transfer.

The heterodisulfide reductase complex, *hdrABC*, is encoded in two different loci: a) the first one is similar in sequence and organization to the *hdr* cluster found in *At. ferrooxidans*^{T} [36], including genes encoding for the sulfur relay system formed by a rhodanese-like protein, a sulfur transferase *tusA*, a *dsrE*-like protein and hypothetical protein of unknown function and b) a second smaller locus encoding only a partial HdrA subunit, followed by the full *hdrABC* subunits. Fifteen orthologs of *dsrE*, thought to deliver the collected electrons to the membrane quinol pool, are found distributed in the genome of *Acidiferrobacter* sp. SPIII/3 and similar numbers in the other *Acidiferrobacter* strains.

Two genetic *loci* (*aprMBA* and *tauE/sat_aprBA_ykgJ*), far distant in the genome of *Acidiferrobacter* sp. SPIII/3 encode for the adenylyl-sulfate reductase composed of subunits alfa AprA and beta AprB. One of the *loci*, also in the same gene cluster, encodes the anchor protein AprM, (homolog to *Allochromatium vinosum* own *aprM*, 55% identity), while the other cluster encodes for a Fe−S-cluster containing protein and a sulfate adenylyl transferase ATP-S (*sat*). The latter could hypothetically transform APS into ATP and sulfate by substrate level phosphorylation (EC 2.7.7.4.). All genomes analyzed also carry a 9-ORF long "*cysB operon*" that encodes all necessary components for assimilatory sulfate reduction via phosphoadenylyl-sulfate/phosphoadenosine phosphosulfate PAPS.

Two independent *loci* are also predicted to encode for the proteins of the Sox periplasmic enzyme complex. The first *locus* comprises an 8-gene cluster encoding besides the Sox subunits X, Y, Z, A and B, a conserved hypothetical protein of unknown function and a Zn^{2+} dependent protease of the YfgC superfamily (with at least two tetratricopeptide repeats TPR), potentially involved in the assembly and maturation of the Sox complex. The entire cluster is flanked by transposases, suggesting its foreign origin. In turn, the second *locus* encodes for identical copies of two subunits of the Sox complex SoxY and SoxZ, which are also conserved in *S. limnicola*^{T} and *S. variabilis*^{T}. The presence of the periplasmic Sox system, that oxidizes thiosulfate ions ($S_2O_3^{2-}$) and the cytoplasmic/inner membrane bound Dsr system, coupling the oxidation of elemental sulfur with the production of NADH, by means of a respiratory chain, together with the inability to oxidize tetrathionate is characteristically found in organisms that use the Branched Thiosulfate Oxidation pathway (BTO), such as purple sulfur bacteria and chemotropic sulfur oxidizers [39,40]. Microorganisms using the BTO pathway are generally anaerobic or facultative aerobic ones, and possess systems for sulfur deposition (intra or extracellular) [41]. No evidence for the formation of sulfur globules has been reported to date in *Acidiferrobacter* spp. Concomitantly, no homologous genes encoding proteins related to the formation of sulfur globules complexes (such as *sgpABC*) [42], have been identified in the analyzed genomes. Cytoplasmic elemental sulfur could likely be aerobically disproportionated to sulfite and sulfide by the sulfur oxygenase/reductase SOR, found encoded in the genomes of the *Acidiferrobacteraceae*.

Acidophilic *Acidiferrobacter* spp. differs from the non-acidophilic members of the family in their ability to derive energy through the oxidation of iron. *Acidiferrobacter* spp. SPIII/3 and *A. thiooxydans*^{T} and ZJ strains encode in their genomes respiratory/dissimilatory iron oxidation genes and pathways described for model iron-oxidizing acidophiles of the *Acidithiobacillus* genus [43]. Iron oxidation in *Acidiferrobacter* spp. is predicted to proceed through a "downhill" (favorable redox potential difference between ferrous iron and oxygen) and an "uphill" (unfavorable redox potential difference between ferrous iron and NAD) pathway very similar to those of *A. ferroxidans* and related iron oxidizing species [44]. *Acidiferrobacter* spp. SPIII/3 and *A. thiooxydans* m-1^T and ZJ possess a gene cluster of similar gene content to the *rus* operon (Fig. 5). The *rus* operon in *At. ferrooxidans*^{T} encodes a multiprotein system driving electrons "downhill", from iron to oxygen [36,45]. *Acidiferrobacter* putative *rus* operons encode for a high molecular weight cytochrome-C Cyc1, a c4 cytochrome Cyc2, an aa_3 cytochrome oxidase subunit, a rusticyanin (Rus1) and a cupredoxin-like protein, as in *At. ferrooxidans*^{T}, though they are not entirely syntenous (Fig. 5; Supplementary Table 5). *Acidiferrobacter* spp. strain ZJ lacks an ORF encoding for a Cyc2 homolog, although this could be due to the draft status of its genome sequence. Cytochrome C oxidase biosynthesis and assembly genes *ctaABTRUS* and regulatory *regBA* genes, neighboring the *rus* operon in *At. ferrooxidans*^{T}, are organized differently in *Acidiferrobacter* species (Fig. 5) and/or only partially conserved. Regulatory *ctaRUS* and the *regB/regA* sensor/regulator two-component signal transduction system are entirely missing in *Acidiferrobacter* genomes. In addition to the rusticyanin (Rus1), encoded in the *rus* operon, two extra rusticyanin-like proteins named Rus2 and Rus3 (54 and 36% identical to *At. ferrooxidans*^{T} rusticyanin A) are encoded in *Acidiferrobacter* genomes. Both, *rus2* and *rus3* possess the conserved type 1 copper binding sites $HX_nCX_4HX_{4\text{-}5}M$ (cd04231). These rusticyanin-like proteins are present in the three *Acidiferrobacter* strains analyzed, with Rus2 being 53 and 32% identical to Rus1 and Rus3 respectively; and Rus3 being 36% identical to Rus1. It is currently unknown if Rus2 and Rus3 may have a role in iron oxidation, or if these are interchangeable with Rus1.

Uphill electron flow, from ferrous iron to NAD, in *Acidiferrobacter* species is predicted to occur as in *At. ferrooxidans*^{T}, through a cytochrome A (CycA), an ubiquinol cytochrome c complex (PetABC) and a NADH dehydrogenase complex. Two well-conserved *pet* operons (*petI* and *petII*) have been described in *At. ferrooxidans*^{T}, *petI* is transcriptionally upregulated during growth with ferrous iron and *petII* during growth in sulfur [36]. *Acidiferrobacter* genomes encode two syntenous and nearly identical *pet* operons (over 94% identity in all proteins), neither of which is conclusively more similar to *At. ferrooxidans petI* or *petII*. As a distinctive feature, *At. ferrooxidans*^{T} *petII* possess at its end a gene encoding for a high potential iron-sulfur protein of unknown function *HiPiP* (AFE_2732), that has not been found encoded in the studied *Acidiferrobacter* genomes.

3.6. Biofilm formation & Quorum Sensing

Biofilm formation is an essential aspect during bioleaching. *Acidiferrobacter* sp. SP3/III is able to form monolayer biofilms on metal sulfides such as pyrite (FeS_2) and chalcopyrite ($CuFeS_2$) (Bellenberg, Sand & Vera, unpublished). Biofilm formation processes are controlled by Quorum Sensing, among other factors. All *Acidiferrobacter* strains sequenced possess a *luxR* family transcriptional regulator-like ORF (cd06170, E-value $6.89e^{-18}$), with an antisense 3' overlapping *luxI* autoinducer biosynthetic protein encoding ORF. Both share 28% identity with the LuxR/LuxI proteins from *Aliivibrio fischeri*. No ORFs encoding for LuxM, LuxS, or LuxP were found. It has previously been shown that *Acidiferrobacter* SP3/

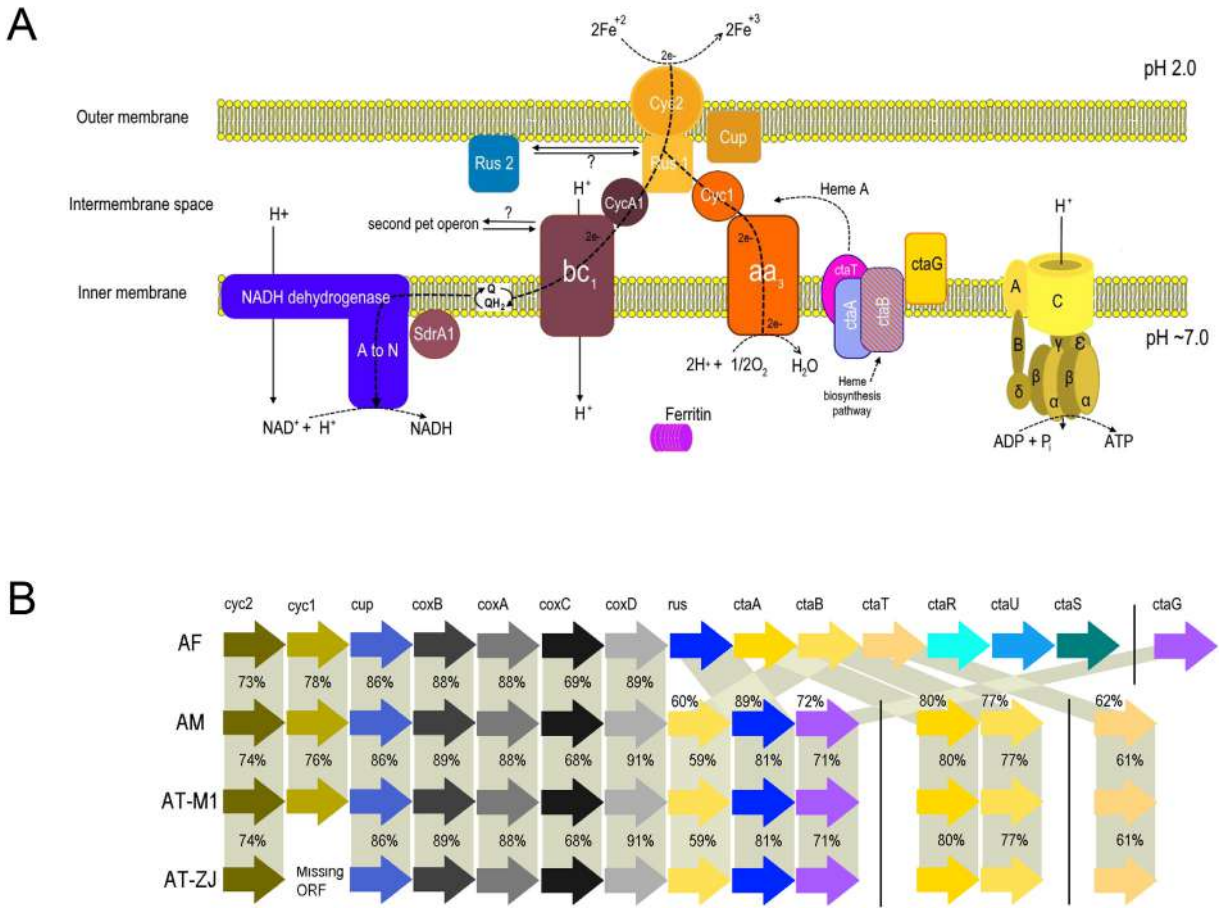

Fig. 5. Model for iron oxidation in *Acidiferrobacter* species (A) Potential Transference of electrons is represented by the thick dotted lines. Genes encoded by the *rus* operon are shown in orange, in burgundy the proteins encoded in the *pet* cluster, in navy blue the NADH dehydrogenase, and in yellow ATP synthase. In other colors, proteins encoded by different clusters and operons. Shown here in magenta in the cytoplasm is also bacterioferritin. (B) Comparison of the *rus* operon structures between *At. ferrooxidans* ^{T} (**AF**, NC_011761.1), *Acidiferrobacter* spp. SPIII/3 (**AM**), *Af. thiooxydans*^{T} (**AT-M1**) and ZJ (**AT-ZJ**) strains. Thick lines represent separation of *loci*. Percentage of amino-acid similarity against AF is indicated over the arrows.

III strain, as well as *A. thyooxidans*^{T} produced a great diversity of N-acyl homoserine lactones (AHL): C10-AHL, C12-AHL, 3-hydroxy-C12-AHL, C14-AHL, 3-hydroxy-C14-AHL and C16-AHL, suggesting the presence of active Type I cell-cell communication systems in *Acidiferrobacter* [46]. Interestingly, AHLs were detected in supernatants from pyrite-grown cultures, but not in supernatants from iron-grown cultures of both strains. In addition, pyrite leaching assays showed a mutualistic inhibition between cells of *Leptospirillum ferrooxidans* DSM 2391 and *Acidiferrobacter* sp. SPIII/3 [46]. Recent experiments have confirmed the ability of *Acidiferrobacter* strains to negatively influence pyrite-leaching behavior in mixed cultures with other bioleaching species (Mena & Vera, unpublished). The reasons for these phenomena are still unknown.

4. Concluding remarks

Despite the finding of genome sequences clustering within the *Acidiferrobacteraceae* in several locations, and that some strains may play major ecological roles [11], few strains from this interesting family have been isolated. A future need to improve cultivation techniques or media for *Acidiferrobacter* is envisaged. Further studies in order to determine the potential utilization of *Acidiferrobacter* spp. for influencing bioleaching processes are in progress. These studies include the classification of new species (Thyssen et al., in preparation), cell-cell communication studies, as well as microscopy, cellular and molecular interactions of *Acidiferrobacter* spp. in mixed species biofilms with other acidophiles.

Conflict of interest

The authors declare no conflict of interest.

Acknowledgements

Raquel Quatrini and Mario Vera acknowledge support from Fondecyt Grants 1181251, 1140048 (RQ) and 1161007 (MV), respectively. Francisco Issotta and Ana Moya-Beltrán acknowledge support from Conicyt PhD scholarships.

Appendix A. Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.resmic.2018.08.001.

References

[1] Hallberg KB, Hedrich S, Johnson DB. *Acidiferrobacter thiooxydans*, gen. nov. sp. nov.; an acidophilic, thermo-tolerant, facultatively anaerobic iron- and sulfur-oxidizer of the family *Ectothiorhodospiraceae*. Extremophiles 2011;15:271–9.
[2] Harrison AP. Genomic and physiological diversity amongst strains of *Thiobacillus ferrooxidans*, and genomic comparison with *Thiobacillus thiooxidans*. Arch Microbiol 1982;131:68–76.

[3] Oren A. The family *Ectothiorhodospiraceae*. The prokaryotes: Gammaproteobacteria. Berlin, Heidelberg: Springer Berlin Heidelberg; 2014. p. 199–222.
[4] Kojima H, Shinohara A, Fukui M. *Sulfurifustis variabilis* gen. nov., sp. nov., a sulfur oxidizer isolated from a lake, and proposal of *Acidiferrobacteraceae* fam. nov. and *Acidiferrobacterales* ord. nov. Int J Syst Evol Microbiol 2015;65:3709–13.
[5] Kojima H, Watanabe T, Fukui M. *Sulfuricaulis limicola* gen. nov., sp. nov., a sulfur oxidizer isolated from a lake. Int J Syst Evol Microbiol 2016;66:266–70.
[6] He Z, Xie X, Xiao S, Liu J, Qiu G. Microbial diversity of mine water at Zhong Tiaoshan copper mine, China. J Basic Microbiol 2007;47:485–95.
[7] Mitchell D, Harneit K, Meyer G, Sand W, Stackebrandt E. Systematic analysis of our culture collection for "genospecies" of *Acidithiobacillus ferrooxidans, Acidithiobacillus thiooxidans* and *Leptospirillum ferrooxidans*. In: Tsezos M, Hatzikoseyian A, Remoundaki E, editors. Biohydrometallurgy: a sustainable technology in evolution, Part II. Proceedings of the international biohydrometallurgy symposium, IBS 2003, held in Athens, Hellas, september 14-19, 2003. Athens, Greece: National Technical University of Athens; 2004. p. 1369–78.
[8] Bulaev A, Belyi A, Panyushkina A, Solopova N, Pivovarova T. Microbial population of industrial biooxidation reactors. Solid State Phenom 2017;262:48–52.
[9] Fujimura R, Sato Y, Nishizawa T, Nanba K, Oshima K, Hattori M, et al. Analysis of early bacterial communities on volcanic deposits on the island of Miyake (Miyake-jima), Japan: a 6-year study at a fixed site. Microbes Environ 2012;27:19–29.
[10] Garcia-Moyano A, Gonzalez-Toril E, Aguilera A, Amils R. Prokaryotic community composition and ecology of floating macroscopic filaments from an extreme acidic environment, Rio Tinto (SW, Spain). Syst Appl Microbiol 2007;30:601–14.
[11] Dyksma S, Bischof K, Fuchs BM, Hoffmann K, Meier D, Meyerdierks A, et al. Ubiquitous *Gammaproteobacteria* dominate dark carbon fixation in coastal sediments. ISME J 2016;10:1939–53.
[12] Issotta F, Covarrubias PC, Moya-Beltrán A, Bellenberg S, Thyssen C, Sand W, et al. 16S rRNA and multilocus phylogenetic analysis of the iron oxidizing acidophiles of the *Acidiferrobacteraceae* family. Solid State Phenom 2017;262:339–43.
[13] Tully BJ, Graham ED, Heidelberg JF. The reconstruction of 2,631 draft metagenome-assembled genomes from the global oceans. Sci Data 2018;5:170203.
[14] Lagesen K, Hallin P, Rodland EA, Staerfeldt HH, Rognes T, Ussery DW. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. Nucleic Acids Res 2007;35:3100–8.
[15] Besemer J, Lomsadze A, Borodovsky M. GeneMarkS: a self-training method for prediction of gene starts in microbial genomes. Implications for finding sequence motifs in regulatory regions. Nucleic Acids Res 2001;29:2607–18.
[16] Fu L, Niu B, Zhu Z, Wu S, Li W. CD-HIT: accelerated for clustering the next-generation sequencing data. Bioinformatics 2012;28:3150–2.
[17] Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, et al. The RAST Server: rapid annotations using subsystems technology. BMC Genom 2008;9:75.
[18] Gonzalez C, Yanquepe M, Cardenas JP, Valdes J, Quatrini R, Holmes DS, et al. Genetic variability of psychrotolerant *Acidithiobacillus ferrivorans* revealed by (meta)genomic analysis. Res Microbiol 2014;165:726–34.
[19] Raes J, Korbel JO, Lercher MJ, von Mering C, Bork P. Prediction of effective genome size in metagenomic samples. Genome Biol 2007;8:R10.
[20] Rutherford K, Parkhill J, Crook J, Horsnell T, Rice P, Rajandream MA, et al. Artemis: sequence visualization and annotation. Bioinformatics 2000;16:944–5.
[21] Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol 2013;30:772–80.
[22] Kumar S, Stecher G, Tamura K. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol Biol Evol 2016;33:1870–4.
[23] Nunez H, Moya-Beltrán A, Covarrubias PC, Issotta F, Cardenas JP, Gonzalez M, et al. Molecular systematics of the genus *Acidithiobacillus*: insights into the phylogenetic structure and diversification of the taxon. Front Microbiol 2017;8:30.
[24] Kristensen DM, Kannan L, Coleman MK, Wolf YI, Sorokin A, Koonin EV, et al. A low-polynomial algorithm for assembling clusters of orthologous groups from intergenomic symmetric best matches. Bioinformatics 2010;26:1481–7.
[25] Kahle D, Wickham H. ggmap: spatial Visualization with ggplot2. The R J 2013;5:144–61.
[26] Bivand R, Lewin-Koh N. Maptools: tools for reading and handling spatial objects. 2017.
[27] Brownrigg R. Maps: draw geographical maps. R package version 3.1.1. Original S code by richard A. Becker, Allan R. Wilks. Enhancements by Thomas P. Minka and Alex Deckmyn. 2017.
[28] Umezawa K, Watanabe T, Miura A, Kojima H, Fukui M. The complete genome sequences of sulfur-oxidizing Gammaproteobacteria *Sulfurifustis variabilis* skN76$^{(T)}$ and *Sulfuricaulis limicola* HA5$^{(T)}$. Stand Genomic Sci 2016;11:71.
[29] Meier-Kolthoff JP, Auch AF, Klenk HP, Goker M. Genome sequence-based species delimitation with confidence intervals and improved distance functions. BMC Bioinformatics 2013;14:60.
[30] Acuna LG, Cardenas JP, Covarrubias PC, Haristoy JJ, Flores R, Nunez H, et al. Architecture and gene repertoire of the flexible genome of the extreme acidophile *Acidithiobacillus caldus*. PLoS One 2013;8, e78237.
[31] Toro N, Jimenez-Zurdo JI, Garcia-Rodriguez FM. Bacterial group II introns: not just splicing. FEMS Microbiol Rev 2007;31:342–58.
[32] Lambowitz AM, Zimmerly S. Group II introns: mobile ribozymes that invade DNA. Cold Spring Harb Perspect Biol 2011;3:a003616.
[33] Kelly DP. Physiology and biochemistry of unicellular sulfur bacteria. In: Schlegel HG, Bowien B, editors. Autotrophic bacteria, science tech publishers, Madison, WI. Berlin: Springer-Verlag; 1989. p. 193–217.
[34] Friedrich CG, Rother D, Bardischewsky F, Quentmeier A, Fischer J. Oxidation of reduced inorganic sulfur compounds by bacteria: emergence of a common mechanism? Appl Environ Microbiol 2001;67:2873–82.
[35] Kelly DP, Shergill JK, Lu WP, Wood AP. Oxidative metabolism of inorganic sulfur compounds by bacteria. Antonie Van Leeuwenhoek 1997;71:95–107.
[36] Quatrini R, Appia-Ayme C, Denis Y, Jedlicki E, Holmes DS, Bonnefoy V. Extending the models for iron and sulfur oxidation in the extreme acidophile *Acidithiobacillus ferrooxidans*. BMC Genom 2009;10:394.
[37] Valdes J, Pedroso I, Quatrini R, Dodson RJ, Tettelin H, Blake 2nd R, et al. *Acidithiobacillus ferrooxidans* metabolism: from genome sequence to industrial applications. BMC Genom 2008;9:597.
[38] Sander J, Engels-Schwarzlose S, Dahl C. Importance of the DsrMKJOP complex for sulfur oxidation in *Allochromatium vinosum* and phylogenetic analysis of related complexes in other prokaryotes. Arch Microbiol 2006;186:357–66.
[39] Grimm F, Franz B, Dahl C. Thiosulfate and sulfur oxidation in purple sulfur bacteria. Berlin, Heidelberg: Springer Berlin Heidelberg; 2008. p. 101–16.
[40] Ghosh W, Dam B. Biochemistry and molecular biology of lithotrophic sulfur oxidation by taxonomically and ecologically diverse bacteria and archaea. FEMS Microbiol Rev 2009;33:999–1043.
[41] Dahl C, Schulte A, Stockdreher Y, Hong C, Grimm F, Sander J, et al. Structural and molecular genetic insight into a widespread sulfur oxidation pathway. J Mol Biol 2008;384:1287–300.
[42] Pattaragulwanit K, Brune DC, Truper HG, Dahl C. Molecular genetic evidence for extracytoplasmic localization of sulfur globules in *Chromatium vinosum*. Arch Microbiol 1998;169:434–44.
[43] Nitschke W, Bonnefoy V. Energy acquisition in low pH environments. In: Quatrini R, Johnson DB, editors. Acidophiles: life in extremely acidic environments. Caister Academic Press; 2016. p. 19–48.
[44] Ilbert M, Bonnefoy V. Insight into the evolution of the iron oxidation pathways. Biochim Biophys Acta 2012;1827:161–75.
[45] Appia-Ayme C, Bengrine A, Cavazza C, Giudici-Orticoni MT, Bruschi M, Chippaux M, et al. Characterization and expression of the co-transcribed cyc1 and cyc2 genes encoding the cytochrome c4 (c552) and a high-molecular-mass cytochrome c from *Thiobacillus ferrooxidans* ATCC 33020. FEMS Microbiol Lett 1998;167:171–7.
[46] Bellenberg S, Diaz M, Noel N, Sand W, Poetsch A, Guiliani N, et al. Biofilm formation, communication and interactions of leaching bacteria during colonization of pyrite and sulfur surfaces. Res Microbiol 2014;165:773–81.
[47] Goris J, Konstantinidis KT, Klappenbach JA, Coenye T, Vandamme P, Tiedje JM. DNA-DNA hybridization values and their relationship to whole-genome sequence similarities. Int J Syst Evol Microbiol 2007;57:81–91.
[48] Richter M, Rossello-Mora R. Shifting the genomic gold standard for the prokaryotic species definition. Proc Natl Acad Sci USA 2009;106:19126–31.