Castro et al. Standards in Genomic Sciences (2017) 12:77
DOI 10.1186/s40793-017-0282-y

Standards in Genomic Sciences

REPORT Open Access

Draft genome sequence of the type strain of the sulfur-oxidizing acidophile, *Acidithiobacillus albertensis* (DSM 14366)

Matías Castro[1†], Ana Moya-Beltrán[1†], Paulo C. Covarrubias[1,5], Mónica Gonzalez[1], Juan Pablo Cardenas[1,5], Francisco Issotta[1], Harold Nuñez[1,5], Lillian G. Acuña[1], Gonzalo Encina[2], David S. Holmes[1,3], D. Barrie Johnson[4] and Raquel Quatrini[1*]

Abstract

Acidithiobacillus albertensis is an extremely acidophilic, mesophilic, obligatory autotrophic sulfur-oxidizer, with potential importance in the bioleaching of sulfidic metal ores, first described in the 1980s. Here we present the draft genome sequence of *Acidithiobacillus albertensis* DSM 14366^T, thereby both filling a long-standing gap in the genomics of the acidithiobacilli, and providing further insight into the understanding of the biology of the non iron-oxidizing members of the *Acidithiobacillus* genus. The assembled genome is 3,1 Mb, and contains 47 tRNAs, tmRNA gene and 2 rRNA operons, along with 3149 protein-coding predicted genes. The Whole Genome Shotgun project was deposited in DDBJ/EMBL/GenBank under the accession MOAD00000000.

Keywords: Acidithiobacilli, Extreme acidophile, Sulfur oxidizer, Bioleaching, Sulfidic metal ores, Phylogenomics

Introduction

The genus *Acidithiobacillus* [1] comprises a group of obligatory acidophilic chemolithotrophic bacteria that derive energy from the oxidation of reduced sulfur compounds, thereby contributing to the bioleaching of ores and to the formation of polluting mine drainage waters. Although they were considered until relatively recently as members of the Gamma-proteobacteria, multi-protein phylogenetic analysis of concatenated ribosomal proteins re-categorized the order *Acidithiobacillales* as a new class of proteobacteria, now known as *Acidithiobacillia* [2]. Currently, seven species are recognized: *Acidithiobacillus thiooxidans* [3], *A. ferrooxidans* [4], *A. albertensis* [5], *A. caldus* [6], *A. ferrivorans* [7], *A. ferridurans* [8], *A. ferriphilus* [9], four of which also catalyze the dissimilatory oxidation of ferrous iron while three (*A. thiooxidans*, *A. albertensis* and *A. caldus)* do not.

Being capable of biogenic acid production and oxidation of reduced sulfur compounds, most species of the taxon have been exploited industrially in the recovery of valuable metals such as copper and gold and other relevant elements from ores and wastes ([10] and references therein). Not only are they frequent members of most analyzed bioleaching consortia, but tend also to be numerically relevant ([11] and references therein). Due to their biotechnological relevance most species of the taxon have been the object of intensive research since the early 1900's [12]. Yet, despite compelling evidence regarding the widespread occurrence of *A. albertensis* [13–16] and its potential for chalcopyrite and sphalerite bioleaching [13, 17], *A. albertensis* remains the least studied species of all acidithiobacilli.

Whole genome sequences of a number of representative strains of four species of *Acidithiobacillus* (*A. thiooxidans, A. ferrooxidans, A. caldus* and *A. ferrivorans*) have been reported to date [18] and genome comparisons have been performed both between and within species [19–23]. However, no representative genome sequence is yet available for *A. albertensis*. Given that *A. albertensis* resembles *A. thiooxidans* in several aspects of their biology and physiology [5, 24], and that presence of either species in the natural and industrial environments tend to be confounded due to the high similarity between species at the 16S rRNA level

* Correspondence: rquatrini@cienciavida.org
†Equal contributors
[1]Fundación Ciencia & Vida, Av. Zañartu 1482, Santiago 7780272, Chile
Full list of author information is available at the end of the article

[25], further characterization of the former is required to shed light into the species-specific processes. Availability of the whole-genome of the type strain of *A. albertensis* represents a first necessary step in this direction.

Here we present a description of the first draft of the genome sequence and annotation of the type strain of *A. albertensis* (DSM 14366^T) along with relevant genomic indices of the taxon. The data presented fill a long-standing gap in the understanding of the genomic landscape of the acidithiobacilli and of the biology of *A. albertensis* and paves the way for more encompassing phylogenomic analyses of the species complex of these fascinating model acidophiles.

Organism information

Classification and features

Originally described by Bryant and colleagues [5], *A. albertensis* (formerly *Thiobacillus albertis*) was recognized as a new species in 1988 [26]. The species epithet derives from the Latin (al.ber.ten'sis. M.L. adj. *albertensis* Albertan), meaning pertaining to Alberta, a province of Canada, from where it was first isolated. The type strain is DSM 14366/ATCC 35403. *A. albertensis* was described as a mesophilic, obligatory autotrophic sulfur-oxidizer that did not oxidize iron. Differentiating characteristics from other members of the acidithiobacilli include forming yellowish colonies on solid sulfur-containing media, a slightly larger cellular size, a tuft of polar flagella, a glycocalyx and a number of large intracellular sulfur globules [5, 17]. *A. albertensis* was reported to have a more confined pH range for growth (2–4.5) and a slightly higher temperature growth optimum with respect to other members of the genus [1], although these features may vary between strains [17]. Additional properties of *A. albertensis* are listed in Table 1.

Phylogenetic analysis of the 16S rRNA gene sequence of *A. albertensis* DSM 14366^T places the type strain close to a few other cultivated members of the species and several uncultured clones deposited in GenBank, all of which are 100% identical at the16S rRNA gene level (Fig. 1). The *A. albertensis* type strain and its closest relatives branch apart from *A. thiooxidans*^{T}.

Table 1 Classification and general features of *A. albertensis* strain^T [22]

MIGS ID	Property	Term	Evidence code[a]
	Classification	Domain *Bacteria*	TAS [1]
		Phylum *Proteobacteria*	TAS [1]
		Class *Acidithiobacillia*	TAS [2]
		Order *Acidithiobacillales*	TAS [47, 48]
		Family *Acidithiobacillaceae*	TAS [47, 49]
		Genus *Acidithiobacillus*	TAS [1]
		Species *Acidithiobacillus albertensis*	TAS [5, 26]
		(Type) strain: *Strain^T (DSM 14366)*	
	Gram stain	Negative	TAS [5]
	Cell shape	Rod	TAS [5]
	Motility	Motile	TAS [5]
	Sporulation	Not reported	NAS
	Temperature range	10–40 °C	TAS [5]
	Optimum temperature	25–30 °C	TAS [5]
	pH range; Optimum	2.0–4.5; 3.5–4.0	TAS [5]
	Carbon source	CO_2	TAS [5]
MIGS-6	Habitat	Acidic mineral-sulfur rich environments	TAS [5]
MIGS-6.3	Salinity	Not reported	NAS
MIGS-22	Oxygen requirement	Aerobic	TAS [5]
MIGS-15	Biotic relationship	Free-living	NAS
MIGS-14	Pathogenicity	Non-pathogen	NAS
MIGS-4	Geographic location	Canada/Alberta	TAS [5]
MIGS-5	Sample collection	1983	TAS [5]
MIGS-4.1	Latitude	Not reported	NAS
MIGS-4.2	Longitude	Not reported	NAS
MIGS-4.4	Altitude	Not reported	NAS

[a]Evidence codes – *IDA* Inferred from Direct Assay, *TAS*: Traceable Author Statement (i.e., a direct report exists in the literature), *NAS* Non-traceable Author Statement (i.e., not directly observed for the living, isolated sample, but based on a generally accepted property for the species, or anecdotal evidence). These evidence codes are from the Gene Ontology project [50]

Genome sequencing information

Genome project history

The organism was selected for sequencing on the basis of its phylogenetic position and 16S rRNA similarity to members of the genus *Acidithiobacillus*. This represents the first draft genome sequence of an *A. albertensis* strain. The Whole Genome Shotgun project has been deposited at GenBank under the accession MOAD00000000. The version described in this paper consists of 1 scaffold (2.7 > X Mbp) and 140 smaller contigs and is the first version, MOAD01000000. Table 2 presents the project information and its association with MIGS (version 2.0) compliance [27].

Growth conditions and genomic DNA preparation

A. albertensis strain DSM 14366^T was obtained from the DSMZ collection and grown in DSMZ 71 medium at 30 °C. DNA isolation and routine manipulations were carried out following standard protocols [28].

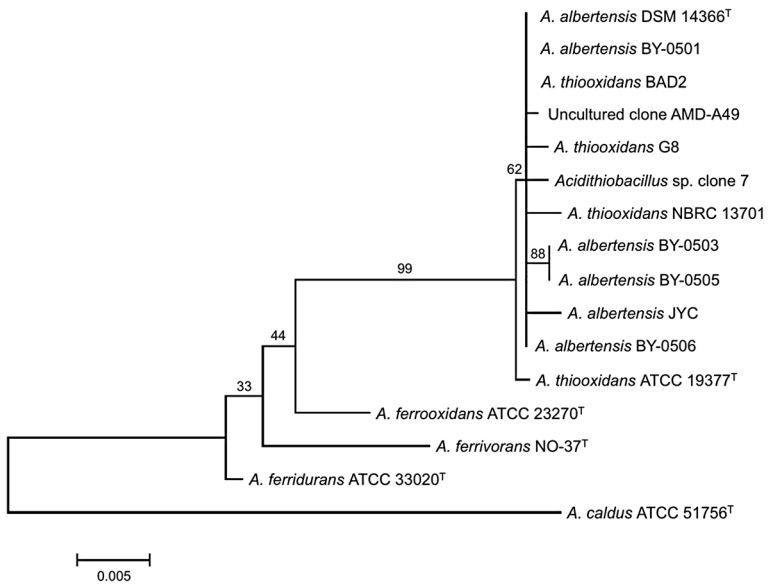

Fig. 1 Phylogenetic tree based on 16S rDNA sequence information position of *A. albertensis* strain DSM 14366^T (type strain = ^{T}) relative to other type and non-type strains within the acidithiobacilli. The strains and their corresponding GenBank accession numbers for 16S rRNA genes are: *A. albertensis* DSM 14366^T, NR_028982; *A. albertensis* BY0501, FJ032185; *A. albertensis* BY0503, FJ032186; *A. albertensis* BY0505, FJ032187; *A. albertensis* BY0506, GQ254658; *A. albertensis* JYC, FJ172635; *A. thiooxidans* ATCC 19377^T, Y11596; *A. thiooxidans* BAD2, KC902821; *A. thiooxidans* G8, KC902819; *A. thiooxidans* NBRC13701, AY830902, AMD uncultured clone c7, JX989232; *A. ferrooxidans* ATCC 23270^T, NR_074193; *A. ferrivorans* NO-37, NR_114620; *A. ferridurans* ATCC 33020 ^{T}, NR_117036; *A. caldus* ATCC 51756 ^{T}, CP005986. The tree was inferred using the Neighbor-Joining method [51]. The optimal tree with the sum of branch length = 0.08720008 is shown. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) are shown next to the branches. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Maximum Composite Likelihood method [52] and are in the units of the number of base substitutions per site. The analysis involved 34nucleotide sequences. There were a total of 1314 positions in the final dataset. Evolutionary analyses were conducted in MEGA6 [53]

Genome sequencing and assembly

The genome of *A. albertensis* DSM 14366^T was sequenced using Illumina sequencing technology (MiSeq platform) and paired-end libraries. Duplicate high quality libraries with insert sizes of ~460 bp were prepared using Nextera™ DNA Sample Preparation kit (Nextera, USA). Raw sequencing reads were preprocessed using Trimmomatic v0.32 [29]. Only reads with a quality score > Q30 (corresponding to less than 1 error per 1000 bp) and a read length > 35 nt were retained. High quality reads were assembled de novo using Velvet (v1.2.10) [30] and a k-mer length of 151, with an N50 of 39,225. Contig segments with at least 37 fold coverage were further scaffolded. The final draft assembly contained 1 scaffold (2.7 > X Mbp) and 140 smaller contigs. The total size of the draft genome is ~3.1 Mbp and the final assembly is based on 3.1 Gbp of Illumina data.

Table 2 Project information

MIGS ID	Property	Term
MIGS 31	Finishing quality	Draft
MIGS-28	Libraries used	Nextera 2.1
MIGS 29	Sequencing platforms	Illumina MiSeq
MIGS 31.2	Fold coverage	64 x
MIGS 30	Assemblers	Velvet v 1.2.10
MIGS 32	Gene calling method	Glimmer 3.02
	Locus Tag	BLW97
	Genbank ID	MOAD00000000
	GenBank Date of Release	FEB 15, 2017
	GOLD ID	Gp0225628
	BIOPROJECT	PRJNA351776
MIGS 13	Source Material Identifier	DSM 14366
	Project relevance	Biomining, Tree of life

Genome annotations

Genes were identified using Glimmer 3.02 [31] as part of the RAST annotation pipeline [32]. The tRNA and tmRNA predictions were made using ARAGORN v1.2.36 [33] and the rRNA prediction was carried out via HMMER3 [34]. Additional gene prediction analysis and manual functional annotation curation was performed using in house resources. The predicted CDSs were used to search the National Center for Biotechnology Information non-redundant database, UniProt, TIGRFam,

Pfam, PRIAM, KEGG, COG and InterPro databases. Protein coding genes were analyzed for signalpeptides using SignalP v4.1 [35] and transmembrane helices using TMHMM v2.0 [36]. The circular map was drawn with CGView [37]. Single nucleotide polymorphisms were called using SNAP v2.1.1 [www.hiv.lanl.gov/content/sequence/SNAP/SNAP.html]. Non-synonymous substitution rates were calculated as the proportion between the number of observed synonymous substitutions in pairwise gene alignments and the size of the each alignment, and are expressed in percent.

Genome comparisons were performed using the GET_HOMOLOGUES software package (version 07112016). Orthology was determined based on all-versus-all Best Bidirectional BlastP Hit and COGtriangles v2.1 as clustering algorithm. Pairwise alignment cutoffs were set at 75% coverage and *E*-value of 10E-5. The phylogenomic relationships between the *A. albertensis*^{T} and other *Acidithiobacillus* strains were inferred from the average nucleotide identity (ANI) values assessed by BLASTn [38] and the in silico DNA-DNA hybridization indexes (DDH) assessed using the Genome-to-Genome Distance Calculator with recommended formula 2 [39]. Species cutoff limits were those defined by Meier-Kolthoff and colleagues [40].

Genome properties

The 3.5 Mbp draft genome of *A. albertensis*^{T} is currently arranged into one high quality scaffold (Fig. 2) and 140 smaller contigs, most of which correspond to fragments of plasmids and other mobile genetic elements. According to the criteria of conservation of universal housekeeping genes [41], the genome is predicted to be 99.9% complete. Its average G + C content is 52.5% (Table 3). From a total of 3202 predicted genes, 3149 were protein-coding genes and 53 were RNA genes. A total of 63.4% of the CDSs were assigned a putative function while the remainders were annotated as hypotheticals. A total of 53 RNA genes partitioned into 47 tRNAs, 1 tmRNA and 2 rRNA operons (Table 3). The presence of two rRNA operons has recently been experimentally

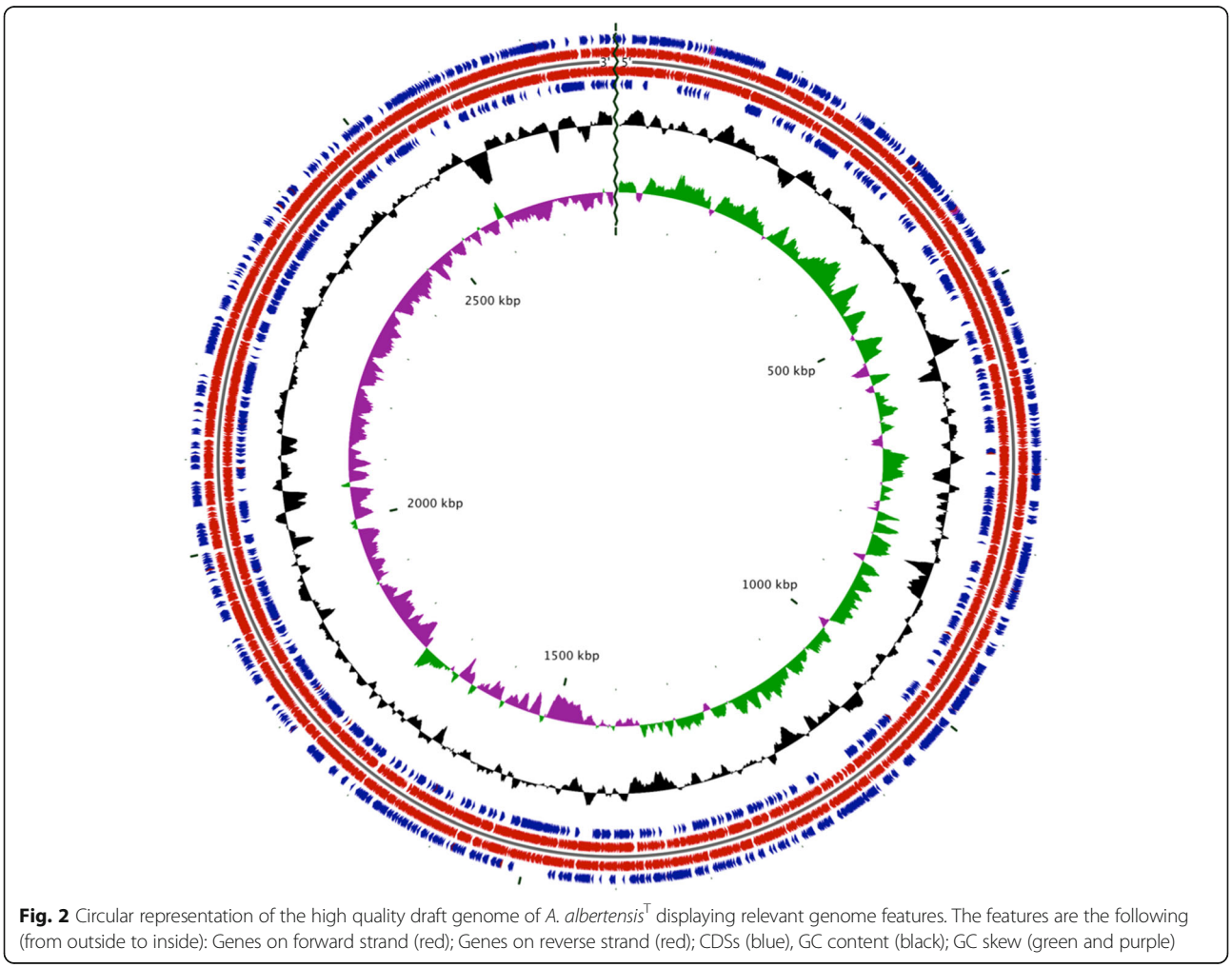

Fig. 2 Circular representation of the high quality draft genome of *A. albertensis*^{T} displaying relevant genome features. The features are the following (from outside to inside): Genes on forward strand (red); Genes on reverse strand (red); CDSs (blue), GC content (black); GC skew (green and purple)

Table 3 Genome statistics

Attribute	Value	% of Total[a]
Genome size (bp)	3,497,418	100.00
DNA coding (bp)	2,930,787	83.80
DNA G + C (bp)	1,836,144	52.50
DNA scaffolds	141	100.00
Total genes[b]	3202	100.00
Protein coding genes	3149	98.34
RNA genes[c]	53	1.66
Pseudo genes	n.d	n.d
Genes in internal clusters	n.d	n.d
Genes with function prediction	1967	61.43
Genes assigned to COGs	2322	72.52
Genes with Pfam domains	2152	67,21
Genes with signalpeptides	374	11.68
Genes with transmembrane helices	727	22.70
CRISPR repeats	0	0

[a]The total is based on either the size of the genome in base pairs or the total number of genes in theannotated genome
[b]Includes tRNA, tmRNA, rRNA
[c]Includes 23S, 16S and 5S rRNA

validated [25]. According to the genomic sequence information, the two operons are 100% identical. The distribution of genes into COGs functional categories is presented in Table 4.

Insights from the genome sequence

Metabolic reconstruction analysis revealed a complete suite of genes for sulfur oxidation, including those encoding the SOX complex (*soxYZB-AX* and *soxYZA-B, soxH*), tetrathionate hydrolase (*tetH, doxD*) and heterodisulfide reductase (*hdrBC and hdrABC*) previously found in *A. thiooxidans*^{T} and *A. caldus*^{T} [42, 43]. Multiple copies of cytochrome d (*cydAB*) and cytochrome o (*cyoACBD*) terminal oxidases found in professional sulfur-oxidizing acidithiobacilli [19], also occur in *A. albertensis*^{T}. Genes for carbon dioxide fixation are well conserved, but no genes for nitrogen fixation were detected in the draft genome. Instead, genes for nitrate/nitrite assimilation and urea hydrolysis, both resulting in the production of ammonia, were found in the genome of the *A. albertensis*^{T}, along with a number of ammonia transporters.

Gene clusters for the biosynthesis and assembly of flagella, which is a differential morphologic trait between this species and *A. thiooxidans*, are conserved with respect to those encoded in the latter, in both general architecture and gene content. The pairwise identity between the predicted protein products of the flagellar genes of both type strains ranges from 87 to 100%, suggesting as well, the common ancestry of the operons.

Table 4 Number of genes associated with general COG functional categories

Code	Value	%age	Description
J	135	4.22	Translation
A	1	0.03	RNA processing and modification
K	124	3.87	Transcription
L	181	5.65	Replication, recombination and repair
B	1	0.03	Chromatin structure and dynamics
D	29	0.91	Cell cycle control, mitosis and meiosis
Y	0	0.00	Nuclear structure
V	52	1.62	Defense mechanisms
T	127	3.97	Signal transduction mechanisms
M	203	6.34	Cell wall/membrane biogenesis
N	66	2.06	Cell motility
Z	0	0.00	Cytoskeleton
W	0	0.00	Extracellular structures
U	102	3.19	Intracellular trafficking and secretion
O	104	3.25	Posttranslational modification, protein turnover, chaperones
C	169	5.28	Energy production and conversion
G	113	3.53	Carbohydrate transport and metabolism
E	156	4.87	Amino acid transport and metabolism
F	53	1.66	Nucleotide transport and metabolism
H	102	3.19	Coenzyme transport and metabolism
I	57	1.78	Lipid transport and metabolism
P	109	3.40	Inorganic ion transport and metabolism
Q	37	1.16	Secondary metabolites biosynthesis, transport and catabolism
R	222	6.93	General function prediction only
S	179	5.60	Function unknown
–	880	27,48	Not in COGs

The total is based on the total number of predicted protein coding genes in the annotated genome

Yet, a relevant number of SNPs (single nucleotide polymorphisms) producing non-synonymous amino acidic substitutions of presently unclear relevance were uncovered in nine genes of the *A. albertensis*^{T} flagellar cluster (Fig. 3), namely: *flaB2*, *flhF*, *flhG*, *fliH*, *fliK*, *fliR*, *fliS2*, *fleS* and *fleQ1*. All these genes are well conserved between *A. thiooxidans* strains (Fig. 3). The gene variants identified in *A. albertensis* were validated by read recruitment on a one-to-one basis, and are supported by more than 75 fold average (deep) coverage. These genes encode the flagellins FlaB2, the hook-length control protein FliK, the biosynthesis proteins FlhF, FliR and FliS, the biosynthesis regulator FlhG, also known as FleN, the assembly protein FliH, the sensor histidine kinase FleS and the regulator FleQ. Among these proteins, FlhF and FlhG/FleN encode proteins that have

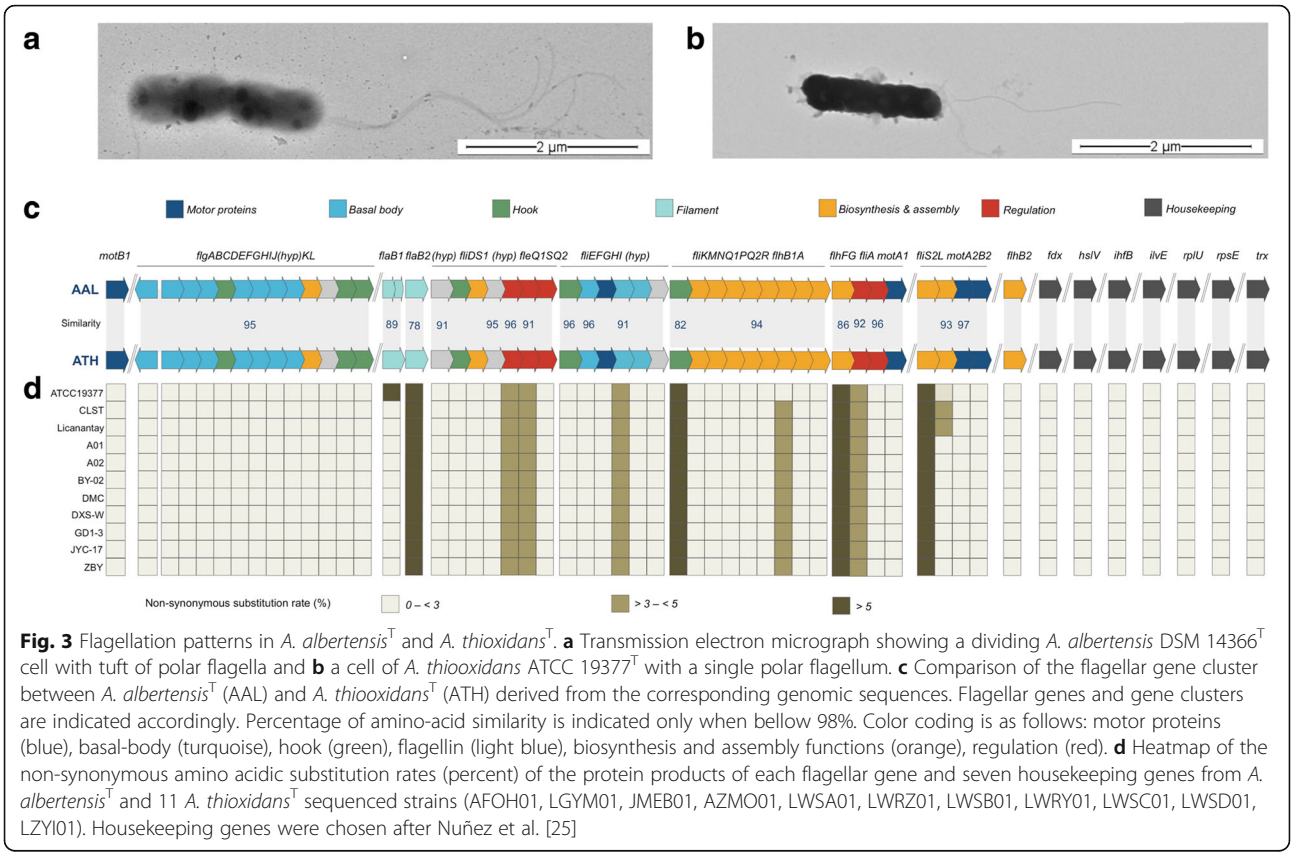

Fig. 3 Flagellation patterns in *A. albertensis*^{T} and *A. thioxidans*^{T}. **a** Transmission electron micrograph showing a dividing *A. albertensis* DSM 14366^{T} cell with tuft of polar flagella and **b** a cell of *A. thiooxidans* ATCC 19377^{T} with a single polar flagellum. **c** Comparison of the flagellar gene cluster between *A. albertensis*^{T} (AAL) and *A. thiooxidans*^{T} (ATH) derived from the corresponding genomic sequences. Flagellar genes and gene clusters are indicated accordingly. Percentage of amino-acid similarity is indicated only when bellow 98%. Color coding is as follows: motor proteins (blue), basal-body (turquoise), hook (green), flagellin (light blue), biosynthesis and assembly functions (orange), regulation (red). **d** Heatmap of the non-synonymous amino acidic substitution rates (percent) of the protein products of each flagellar gene and seven housekeeping genes from *A. albertensis*^{T} and 11 *A. thioxidans*^{T} sequenced strains (AFOH01, LGYM01, JMEB01, AZMO01, LWSA01, LWRZ01, LWSB01, LWRY01, LWSC01, LWSD01, LZYI01). Housekeeping genes were chosen after Nuñez et al. [25]

been shown to be relevant in the control flagellation patterns in other model bacteria [44], suggesting that differences in flagellation between *A. albertensis* (lophotrichous) and *A. thiooxidans* (monotrichous) shown in Fig. 3 might be partially attributed to divergence in these genes (6–14%). For the rest of the flagellar genes the rate of SNPs conductive to amino acidic substitutions between *A. albertensis* and other *A. thiooxidans* sequenced strains is low (<3) and similar to the rate observed in well conserved housekeeping genes. Further studies should be pursued to clarify the relevance of the uncovered substitutions in the flagellation patterns of the acidithiobacilli. Also, a larger number of chemotaxis genes were predicted in the *A. albertensis*^{T} genome sequence with respect to those in *A. thiooxidans*. This latter set of genes is organized in a cluster that includes *mcp1-cheYSA-mcp2-cheWRDB*, and encodes proteins participating in sensory adaptation to changing environmental signals rather than flagellar motor control [45].

Differences between the *A. albertensis*^{T} genome and the pangenome of 10 other sequenced *A. thiooxidans* (recently reported by [22]) can be attributed to little over 1000 genes (1066 genes). Nearly half of these genes pertain to at least 10 integrated mobile genetic elements and a presently unclear number of plasmids, representing up to 16.2% of the *A. albertensis*^{T} genome. In these genomic segments 54.5% of the genes are hypotheticals but a number of relevant functions were also detected, including among others: a) four orthologs of the sulfur oxygenase reductases (*sor1–4*), b) the gene cluster encoding the assimilatory nitrate and nitrite reductases, c) the urea carboxylase/allophanate hydrolase and the urea ABC transporter encoding genes, d) the spermidine/putrescine ABC transporter *potABC* and e) the three-gene operon associated with rubrerythrin, recently described by Cárdenas et al. [46]. All of these functions could confer adaptive advantages to *A. albertensis*^{T} over *A. thiooxidans* strains under nitrogen and oxygen limitation and/or under extremely low pH.

Differences in gene dosage have also been observed between the two mesophilic sulfur-oxidizing/non iron-oxidizing species based on the comparison of the two type strains. *A. albertensis*^{T} has more copies or gene variants (2 to more than 30) of the following: a) transposases and inactivated derivatives, b) thiol:disulfide interchange protein DsbG precursor, c) methyl-accepting chemotaxis receptor proteins, d) Crp/Fnr, LysR and MerR family transcriptional regulators, e) cytochrome d ubiquinol oxidases and e) SOR sulfur oxygenase reductases. The latter occur in four copies in the *A. albertensis*^{T} genome, being completely absent in *A. thiooxidans*^{T}. Also more than 30 predicted protein products with GGDEF/EAL domains, likely involved in

nucleotide driven signaling pathways, control and modulate gene expression and/or activity in *A. albertensis*^{T}, 40% of which seem to be exclusive to this species. Significant quantitative and qualitative differences in gene content have been reported before between strains of *A. thiooxidans* obtained from industrial processes [21, 22].

Despite the above mentioned differences between the type strains of *A. albertensis* and *A. thiooxidans*, the average nucleotide identity value assessed by BLASTn (97,4%) and the in silico DNA-DNA hybridization index assessed by GGDC (82.9%) are bellow the currently recognized species cutoff limits [39], implying that *A. albertensis* and *A. thiooxidans* probably comprise a single genospecies.

Conclusions

Altogether, the evidence presented herein suggests that validity of *A. albertensis* as an independent species should be reconsidered. In this respect, genomic approaches are crucial for understanding evolutionary processes and the origins of microbial biodiversity. The availability of the first high quality draft genome sequence of an *A. albertensis* strain will certainly enable more comprehensive comparative genomic studies and contribute to the resolution of the taxonomy and phylogeny of the genus. From a genomic standpoint, further analyses should be performed to assess if existing differences between the two type strains extend to other strains of each 'presumed species'.

Abbreviations

ATCC: American Type Culture Collection; DSMZ: Deutsche Sammlung von Mikroorganismen und Zellkulturen; NCBI: National Center for Biotechnology Information

Funding

This work was performed under the auspices of the projects FONDECYT 1140048 and FONDECYT 1130683 and Basal PFB-16. We would like to acknowledge also FONDECYT Post Doctoral research funds (FONDECYT 3140003 and 3130376) and CONICYT and UNAB scholarships.

Authors' contributions

RQ, DBJ and DSH conceived and supervised the study. PCC and GE prepared the libraries and ran the sequencing. MC and AM carried out the sequence processing and the bioinformatic analysis. JPC did the phylogenetic analysis. LA grew the cultures and prepared high quality DNA. MG prepared and maintained the strains. FI and HN supported the sequence analyses. MC, AM and RQ analyzed and interpreted the data and wrote the paper. All authors read and approved the final manuscript.

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

Author details

[1]Fundación Ciencia & Vida, Av. Zañartu 1482, Santiago 7780272, Chile. [2]Genómica S.p.A, Santiago, Chile. [3]Facultad de Ciencias Biologicas, Universidad Andres Bello, Santiago, Chile. [4]College of Natural Sciences, Bangor University, Bangor, UK. [5]Current Address: uBiome Chile, SpA, Santiago, Chile.

Received: 21 April 2017 Accepted: 21 November 2017
Published online: 15 December 2017

References

1. Kelly DP, Wood AP. Reclassification of some species of *Thiobacillus* to the newly designated genera *Acidithiobacillus* gen. nov., *Halothiobacillus* gen. nov. and *Thermithiobacillus* gen. nov. Int J Syst Evol Microbiol. 2000; doi:10.1099/00207713-50-2-511.
2. Williams KP, Kelly DP. Proposal for a new class within the phylum *Proteobacteria*, *Acidithiobacillia* classis nov., with the type order *Acidithiobacillales*, and emended description of the class *Gammaproteobacteria*. Int J Syst Evol Microbiol. 2013; doi:10.1099/ijs.0.049270-0.
3. Waksman SA, Joffe JS. Microorganisms concerned in the oxidation of sulfur in the soil. II. *Thiobacillus thiooxidans*, a new sulfur-oxidizing organism isolated from the soil. J Bacteriol. 1922; PMC http://www.ncbi.nlm.nih.gov/pmc/articles/PMC378965. Accessed 1922.
4. Temple KL, Colmer AR. The autotrophic oxidation of iron by a new bacterium: *Thiobacillus ferrooxidans*. J Bacteriol. 1951; PMC http://www.ncbi.nlm.nih.gov/pmc/articles/PMC386175. Accessed 1951.
5. Bryant RD, McGroarty KM, Costerton JW, Laishley EJ. Isolation and characterization of a new acidophilic *Thiobacillus* species (*T. albertis*). Can J Microbiol. 1983; doi:10.1139/m83-178.
6. Hallberg KB, Lindström EB. Characterization of *Thiobacillus caldus* sp. nov., a moderately thermophilic acidophile. Microbiology. 1994; doi:10.1099/13500872-140-12-3451.
7. Hallberg KB, González-Toril E, Johnson DB. *Acidithiobacillus ferrivorans* sp. nov.; facultatively anaerobic, psychrotolerant iron-, and sulfur-oxidizing acidophiles isolated from metal mine-impacted environments. Extremophiles. 2010; doi:10.1007/s00792-009-0282-y.
8. Hedrich S, Johnson DB. *Acidithiobacillus ferridurans* sp. nov., an acidophilic iron-, sulfur- and hydrogen-metabolizing chemolithotrophic gammaproteobacterium. Int J Syst Evol Microbiol. 2013; doi:10.1099/ijs.0.049759-0.
9. Falagán C, Johnson DB. *Acidithiobacillus ferriphilus* sp. nov., a facultatively anaerobic iron- and sulfur-metabolizing extreme acidophile. Int J Syst Evol Microbiol. 2016; doi:10.1099/ijsem.0.000698.
10. Johnson DB. Biomining-biotechnologies for extracting and recovering metals from ores and waste materials. Curr Opin Biotechnol. 2014; doi:10.1016/j.copbio.2014.04.008.
11. Nuñez H, Covarrubias PC, Moya-Beltrán A, Issotta F, Atavales J, Acuña LG, Johnson DB, Quatrini R. Detection, identification and typing of species and strains: a review. Res Microbiol. 2016; doi:10.1016/j.resmic.2016.05.006.
12. Johnson DB, Quatrini R. Acidophile microbiology in space and time. In: Quatrini R, Johnson DB, editors. Acidophiles: life in extremely acidic environments. UK: Caister Academic Press; 2016. p. 3–16. 10.21775/9781910190333.01.
13. Kelly DP, Harrison AH. Genus *Thiobacillus*. In: Staley JT, Bryant MP, Pfennig N, Holt JG, editors. Bergey's manual of systematic bacteriology. Baltimore: Williams & Wilkins; 1989. p. 1842–58.
14. Puhakka AJ, Kaksonen AH, Riekkola-Vanhanen M. Heap leaching of black schist. In: Rawlings DE, Johnson BD, editors. Biomining. Berlin, Heidelberg, New York: Springer; 2007. p. 139–50.
15. Urbieta MS, González-Toril E, Aguilera A, Giaveno MA, Donati E. First prokaryotic biodiversity assessment using molecular techniques of an acidic river in Neuquén, Argentina. Microb Ecol. 2012; doi:10.1007/s00248-011-9997-2.
16. Smeulders MJ, Pol A, Zandvoort MH, Jetten MS, Op den Camp HJ. Diversity and ecophysiology of new isolates of extremely acidophilic CS2-converting *Acidithiobacillus* strains. Appl Environ Microbiol. 2013; doi:10.1128/AEM.02167-13.
17. Xia JL, Peng AA, He H, Yang Y, Liu XD, Qiu GZ. A new strain *Acidithiobacillus albertensis* BY-05 for bioleaching of metal sulfides ores. Trans Nonferrous Metal Soc China. 2007; doi:10.1016/S1003-6326(07)60067-3.
18. Cárdenas JP, Quatrini R, Holmes DS. Progress in acidophile genomics. In: Quatrini R, Johnson DB, editors. Acidophiles: life in extremely acidic environments. UK: Caister Academic Press; 2016a. p. 179–97.
19. Valdés J, Cárdenas JP, Quatrini R, Esparza M, Osorio H, Duarte F, et al. Comparative genomics begins to unravel the ecophysiology of bioleaching. Hydrometallurgy. 2010; doi:10.4028/www.scientific.net/AMR.71-73.143.
20. Acuña LG, Cárdenas JP, Covarrubias PC, Haristoy JJ, Flores R, Nuñez H, et al. Architecture and gene repertoire of the flexible genome of the extreme

acidophile *Acidithiobacillus caldus*. PLoS One. 2013; doi:10.1371/journal.pone.0078237.
21. Travisany D, Cortés MP, Latorre M, Di Genova A, Budinich M, Bobadilla-Fazzini RA, et al. A new genome of *Acidithiobacillus thiooxidans* provides insights into adaptation to a bioleaching environment. Res Microbiol. 2014; doi:10.1016/j.resmic.2014.08.004.
22. Zhang X, Feng X, Tao J, Ma L, Xiao Y, Liang Y, et al. Comparative genomics of the extreme Acidophile *Acidithiobacillus thiooxidans* reveals intraspecific divergence and niche adaptation. Int J Mol Sci. 2016a; doi:10.3390/ijms17081355.
23. Zhang X, She S, Dong W, Niu J, Xiao Y, Liang Y, et al. Comparative genomics unravels metabolic differences at the species and/or strain level and extremely acidic environmental adaptation of ten bacteria belonging to the genus *Acidithiobacillus*. Syst Appl Microbiol. 2016b; doi:10.1016/j.syapm.2016.08.007.
24. Bryant RD, Costerton JW, Laishley EJ. The role of *Thiobacillus albertis* glycocalyx in the adhesion of cells to elemental sulfur. Can J Microbiol. 1984; doi:10.1139/m84-015.
25. Nuñez H, Moya-Beltrán A, Covarrubias PC, Issotta F, Cárdenas JP, González M, et al. Molecular systematics of the genus *Acidithiobacillus*: insights into the phylogenetic structure and diversification of the taxon. Front Microbiol. 2017; doi:10.3389/fmicb.2017.00030.
26. Anon. Validation of the publication of new names and new combinations previously effectively published outside the IJSB. List No. 25. Int J Syst Bacteriol. 1988; doi:10.1099/00207713-38-2-220.
27. Field D, Garrity G, Gray T, Morrison N, Selengut J, Sterk P, et al. The minimum information about a genome sequence (MIGS) specification. Nat Biotechnol. 2008; doi:10.1038/nbt1360.
28. Nieto PA, Covarrubias PC, Jedlicki E, Holmes DS, Quatrini R. Selection and evaluation of reference genes for improved interrogation of microbial transcriptomes: case study with the extremophile *Acidithiobacillus ferrooxidans*. BMC Mol Biol. 2009; doi:10.1186/1471-2199-10-63.
29. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014; doi:10.1093/bioinformatics/btu170.
30. Zerbino DR, Birney E. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. Genome Res. 2008; doi:10.1101/gr.074492.107.
31. Delcher AL, Bratke KA, Powers EC, Salzberg SL. Identifying bacterial genes and endosymbiont DNA with glimmer. Bioinformatics. 2007; doi:10.1093/bioinformatics/btm009.
32. Overbeek R, Olson R, Pusch GD, Olsen GJ, Davis JJ, Disz T, et al. The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). Nucleic Ac Res. 2014; doi:10.1093/nar/gkt1226.
33. Laslett D, Canback B. ARAGORN, a program to detect tRNA genes and tmRNA genes in nucleotide sequences. Nucleic Acids Res. 2004; doi:10.1093/nar/gkh152.
34. Huang Y, Gilna P, Li W. Identification of ribosomal RNA genes in metagenomic fragments. Bioinformatics. 2009; doi:10.1093/bioinformatics/btp161.
35. Petersen TN, Brunak S, von Heijne G, Nielsen H. SignalP 4.0: discriminating signal peptides from transmembrane regions. Nat Methods. 2011; doi:10.1038/nmeth.1701.
36. Krogh A, Larsson B, von Heijne G, Sonnhammer EL. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. J Mol Biol. 2001; doi:10.1006/jmbi.2000.4315.
37. Stothard P, Wishart DS. Circular genome visualization and exploration using CGView. Bioinformatics. 2005; doi:10.1093/bioinformatics/bti054.
38. 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; doi:10.1099/ijs.0.64483-0.
39. Meier-Kolthoff JP, Auch AF, Klenk HP, Göker M. Genome sequence-based species delimitation with confidence intervals and improved distance functions. BMC Bioinformatics. 2013; doi:10.1186/1471-2105-14-60.
40. Meier-Kolthoff JP, Hahnke RL, Petersen J, Scheuner C, Michael V, Fiebig A, et al. Complete genome sequence of DSM 30083^T, the type strain (U5/41^T) of *Escherichia coli*, and a proposal for delineating subspecies in microbial taxonomy. Stand Genomic Sci. 2014; doi:10.1186/1944-3277-9-2.
41. Raes J, Korbel JO, Lercher MJ, von Mering C, Bork P. Prediction of effective genome size in metagenomic samples. Genome Biol. 2007; doi:10.1186/gb-2007-8-1-r10.
42. Valdés J, Ossandon F, Quatrini R, Dopson M, Holmes DS. Draft genome sequence of the extremely acidophilic biomining bacterium *Acidithiobacillus thiooxidans* ATCC 19377 provides insights into the evolution of the *Acidithiobacillus* genus. J Bacteriol. 2011; doi:10.1128/JB.06281-11.
43. Valdés J, Quatrini R, Hallberg K, Dopson M, Valenzuela PD, Holmes DS. Draft genome sequence of the extremely acidophilic bacterium *Acidithiobacillus caldus* ATCC 51756 reveals metabolic versatility in the genus *Acidithiobacillus*. J Bacteriol. 2009; doi:10.1128/JB.00843-09.
44. Kazmierczak BI, Hendrixson DR. Spatial and numerical regulation of flagellar biosynthesis in polarly flagellated bacteria. Mol Microbiol. 2013; doi:10.1111/mmi.12221.
45. Parkinson JS, Hazelbauer GL, Falke JJ. Signaling and sensory adaptation in *Escherichia coli* chemoreceptors: 2015 update. Trends Microbiol. 2015; doi:10.1016/j.tim.2015.03.003.
46. Cárdenas JP, Quatrini R, Holmes DS. Aerobic lineage of the oxidative stress response protein rubrerythrin emerged in an ancient microaerobic, (hyper)thermophilic environment. Front Microbiol. 2016b; doi:10.3389/fmicb.2016.01822.
47. Euzéby J. Validation List No. 106. Validation of publication of new names and new combinations previously effectively published outside the IJSEM. Int J Syst Evol Microbiol. 2005; doi:10.1099/ijs.0.64108-0.
48. Garrity GM, Bell JA, Lilburn T. Order II. *Acidithiobacillales* ord. nov. In: Brenner DJ, Krieg NR, Staley JT, Garrity GM, editors. Bergey's manual of systematic bacteriology, vol. 2, part B. 2nd ed. New York: Springer; 2005. p. 60.
49. Garrity GM, Bell JA, Lilburn T. Family I. *Acidithiobacillaceae* fam. nov. In: Brenner DJ, Krieg NR, Staley JT, Garrity GM, editors. Bergey's manual of systematic bacteriology, vol. 2, part B. 2nd ed. New York: Springer; 2005. p. 60.
50. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, et al. Gene ontology: tool for the unification of biology. The gene ontology consortium. Nat Genet. 2000; doi:10.1038/75556.
51. Saitou N, Nei M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol. 1987;4:406–25.
52. Tamura K, Nei M, Kumar S. Prospects for inferring very large phylogenies by using the neighbor-joining method. Proc Natl Acad Sci U S A. 2004; doi:10.1073/pnas.0404206101.
53. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. MEGA6: molecular evolutionary genetics analysis version 6.0. Mol Biol Evol. 2013; doi:10.1093/molbev/mst197.