Iron homeostasis strategies in acidophilic iron oxidizers: Studies in *Acidithiobacillus* and *Leptospirillum*

Hector Osorio [a], Verónica Martínez [b], Felipe A. Veloso [a], Inti Pedroso [a], Jorge Valdés [a], Eugenia Jedlicki [b], David S. Holmes [a], Raquel Quatrini [a,*]

[a] *Center for Bioinformatics and Genome Biology, Life Science Foundation, MIFAB & Andrés Bello University, Santiago, Chile*
[b] *ICBM, Faculty of Medicine, University of Chile, Santiago, Chile*

Keywords:
Fe(II)-oxidizers
Acidithiobacilli
Leptospirilli
Iron management strategies
Comparative genomics

A B S T R A C T

An understanding of the physiology and metabolic complexity of microbial consortia involved in metal solubilization is a prerequisite for the rational improvement of bioleaching technologies. Among the most challenging aspects that remain to be addressed is how aerobic acidophiles, especially Fe(II)-oxidizers, contend with the paradoxical hazards of iron overload and iron deficiency, each with deleterious consequences for growth. Homeostatic mechanisms regulating the acquisition, utilization/oxidation, storage and intracellular mobilization of cellular iron are deemed to be critical for fitness and survival of bioleaching microbes. In an attempt to contribute to the comprehensive understanding of the biology and ecology of the microbial communities in bioleaching econiches, we have used comparative genomics and other bioinformatic tools to reconstruct the iron management strategies in newly sequenced *Acidithiobacilli* and other biomining genomes available in public databases.

Species specific genes have been identified with distinctive functional roles in iron management as well as genes shared by several species in biomining consortia. Their analysis contributes to our understanding of the general survival strategies in acidic and iron loaded environments and suggests functions for genes with currently unknown roles that might reveal novel aspects of iron response in acidophiles. Comprehensive examination of the occurrence and conservation of regulatory functions and regulatory sites also allowed the prediction of the metal regulatory networks for these biomining microbes.

1. Introduction

In acidic environments, such as biomining biotopes, dramatic changes in the concentration of protons, Fe(III)/Fe(II) and heavy metals occur as mineral leaching proceeds. Understanding microbial cycling of iron under these conditions has implications for the leaching of ores and the development of remediation techniques for sites affected by acid mine drainage or contaminated with metals from industrial wastes (Malki et al., 2006). In this context, knowledge of iron management mechanisms in acidophiles is essential for generating a comprehensive description of the role that these microorganisms play in the development and maintenance of the biological assemblages and for understanding their relative contributions to the cycling of iron.

Acidophiles are known to cope with the highest levels of soluble iron in nature (Cornell and Schwertmann, 2003), but many unanswered questions remain regarding the nature and ecological distribution of the genetic determinants underlying their tolerance/resistance mechanisms. The need to understand these issues calls for a comprehensive investigation of the strategies to a) acquire iron in acidic conditions, b) coordinate iron acquisition with utilization, storage and oxidation of iron through metal responsive regulation and/or to c) mitigate oxidative stress caused by iron overload.

In this work we analyze aspects of the iron homeostasis response of some of the major contributors to microbial bioleaching through multiple genomic comparisons of currently available completed (Valdés et al., submitted for publication) and draft genome sequences (CBGB, Tyson et al., 2004) and contrast them with what is already known in *Acidithiobacillus ferrooxidans* (Quatrini et al., 2005a; Quatrini et al., 2007). Using bioinformatics and comparative genomic strategies, we aim to establish a) the genetic determinants for iron management in *Acidithiobacilli* and *Leptospirilli* species through functional gene identification, b) to reconstruct the Fur-dependent genetic regulatory network and c) to compare the gene complements and relevant aspects of the iron homeostasis response of the available biomining genomes.

The examination of the coding potential, the conservation, the organization and the distribution of iron management functions within this "conceptual consortia of acidophiles", is starting to provide the background for predictions about how their gene products mediate microbial survival in iron-rich acidic environments.

* Corresponding author. Tel.: +56 2 237 5523; fax: +56 2 237 2259.
E-mail addresses: hosoriou@gmail.com (H. Osorio), rquatrini@yahoo.com.ar (R. Quatrini).

2. Methods

2.1. Sequence data

Complete genome sequences of *A. ferrooxidans* ATCC 23270 was obtained from the Institute for Genomic Research database (TIGR; Valdés et al., in preparation). Partially sequenced genomes of *Acidithiobacillus thiooxidans* ATCC 19377 and *Acidithiobacillus caldus* ATCC 51756 were obtained from the Center for Bioinformatics and Genome Biology (CBGB). Gapped genomic sequences of *Leptospirillum ferriphilum* type II and "*Leptospirillum ferrodiazotrophum*" type III were redrawn from the Joint Genome Institute (JGI) and the Genome database from the National Center for Biotechnology Information (NCBI).

2.2. ORF prediction

ORFs likely to encode proteins were predicted by GLIMMER (Salzberg et al., 1998). This program, based on interpolated Markov models, was trained with ORFs larger than 600 bp from the proper genes available in GenBank and our private databases. All predicted proteins larger than 100 aa were searched against a nonredundant protein database as described (Fleischmann et al., 1995). Manual curation of the predicted genes was performed to correct errors in start site prediction and identify missing candidate genes. The 5′ and 3′ regions of each ORF were inspected to define initiation codons using homologies, position of ribosomal binding sites, and transcriptional terminators.

2.3. Gene identification

The following bioinformatic programs were used to further characterize candidate genes and their predicted protein products: BlastP and PsiBlast (Altschul et al., 1990), the suite of protein characterization programs available in InterproScan (Mulder, 2007), Blocks (Henikoff et al., 2000) and ClustalW (Larkin et al., 2007). Three sets of hidden Markov models were used to determine ORF membership in families and superfamilies: PFAM V5.5 (Bateman et al., 2000), TIGRFAMS 1.0 H (Haft et al., 2001) and COGs (Tatusov et al., 1997). The annotated genomes were displayed in the interactive format of Artemis (Rutherford et al., 2000).

2.4. Fur box identification

A set of 66 experimentally confirmed Fur boxes from *Escherichia coli*, *Salmonella typhimurium*, *Pseudomonas aeruginosa* and *Staphylococcus aureus* was used to generate an alignment matrix and a weight matrix by the information content method (Schneider, 1997). The weight matrix used to search the all complete and partial genomic sequences included in this study using a 19-bp sliding window as described previously (Quatrini et al., 2007). Genes carrying candidate Fur boxes in their upstream regions were retained as putative iron related functions directly targeted by the Fur regulator and further used as search queries.

2.5. Comparative genomics

A comprehensive search in the NCBI public database was performed to identify all proteins that are related to iron homeostasis in bacteria using textmining strategies. Amino acid sequences for the iron homeostasis related genes identified were then searched in the genome sequence of *A. ferrooxidans* and draft genome sequences of *A. caldus*, *A. thiooxidans* and the *Leptospirilli* using wu-BLAST (Altschul and Gish, 1996) and candidate genes were then compared against each other. Orthologous and paralogous families were derived by performing all-versus-all searches on the remaining protein sequences by using a modified version of a previously described method (Nierman et al., 2001). Matching hits had to have an expectation cutoff value of at least 10^{-20}, 30% sequence identity, and an alignment overlap length of at least 100 amino acids to the search querys. Each sequence was inspected to verify that it contained appropriately positioned catalytic residues, an N-terminal signal peptide if corresponding (Nielsen et al., 1997) and/or membrane spanning domains (TOPPRED, von Heijne, 1992; TMPRED, Hofmann and Stoffel, 1993). For further analysis annotations were compared to known metabolic models using perl scripts written in our laboratory. Model metabolic pathways were obtained from BIOCYC (Karp et al., 2005), KEGG (Kanehisa et al., in press), and ERGO (Overbeek et al., 2003).

3. Results and discussion

3.1. Bioinformatics pipeline

We have analyzed the draft genomic sequences of the strict iron oxidizers *Leptospirillum ferriphilum* – type II (2.66 Mbp) and "*L. ferrodiazotrophum*" – type III (2.23 Mbp) (Tyson et al., 2004), the complete genome of the iron and sulfur oxidizer *A. ferrooxidans* ATCC 23270 (2.98 Mbp) (Valdés et al., in preparation) and the two newly sequenced draft genomes of the sulfur oxidizers *A. thiooxidans* ATCC 19377 (2.90 Mbp) and *A. caldus* ATCC 51756 (2.94 Mbp) (CBGB) using

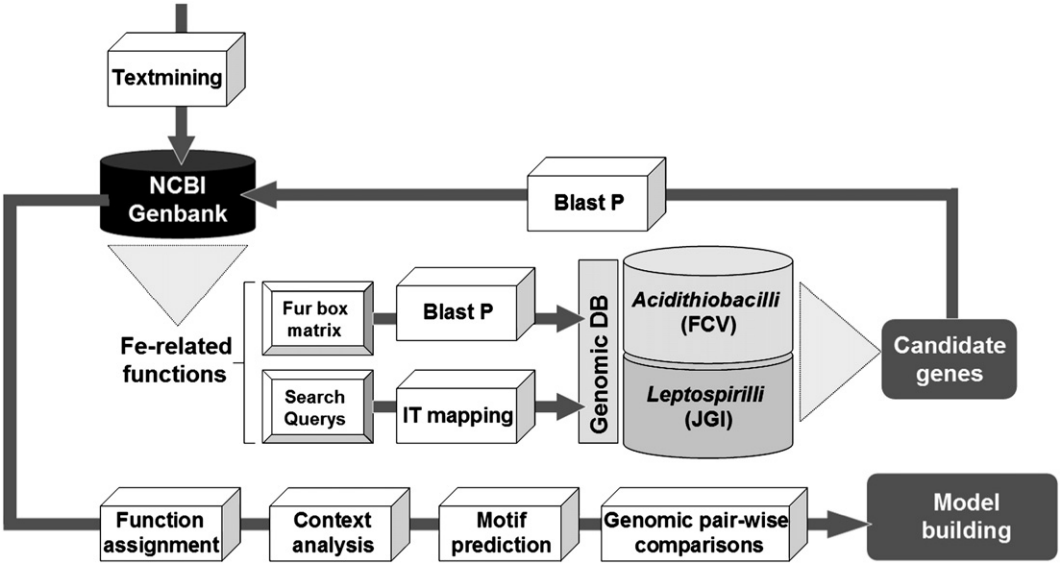

Fig. 1. Methods used: Bioinformatic tools and work flow pipeline.

Table 1
Abundance and distribution of determinants driving iron homeostasis in biomining acidophiles

	FeIII-siderophore uptake				FeII uptake		Regulation
	OMRs	TonB	Fe-ABC	Sid uptake	FeoB	MntH	Fur-like
Acidithiobacilli							
A. ferrooxidans	11	8	2	NF	1	2	3
A. thiooxidans	8	6	1	NF	1	2	3
A. caldus	6	4	2	NF	1	1	3
Leptospirilli							
L. ferriphilum II	3	3	1	NF	NF	NF	3
"*L. ferrodiazotrophum*" III	2	2	1	NF	NF	NF	2
Other bacteria							
E. coli	6	2	3	2	1	1	2

OMR: Outer membrane siderophore receptor; TonB: TonBExbBD biopolymer transport system; Fe-ABC: Fe-siderophore inner membrane ABC transporter; Sid uptake: siderophore biosynthesis operon; FeoB: Fe(II) uptake; Mn(II)/Fe(II) NRAMP transporter; NF: Not found.

bioinformatic analysis and comparative genomic tools as summarized in Fig. 1.

3.2. Genetic determinants for iron management in Acidithiobacilli and Leptospirilli species

Despite having a presumed abundant supply of iron, all acidophiles analyzed encode candidates for various distinct Fe(III)-siderophore TonB-dependent outer membrane receptors (OMRs), typical of iron scavenging microorganisms (Table 1). In the three *Acidithiobacilli*, the number of OMRs rivals the complexity found in well-studied neutrophilic bacteria and exceeds the diversity found in *Leptospirilli*. Except for *A. ferrooxidans*, which also encodes OMRs with affinity for hydroxamate- and catecholate-type siderophores (Quatrini et al., 2005a), all biomining microorganisms analyzed have OMRs with predicted affinity for Fe(III)-dicitrate. Besides typical siderophores, simple chemical compounds such as citrate are known to chelate Fe (III). Notably, the acid tolerant bacterium *Helicobacter pylori* carries only Fe(III)-dicitrate-type siderophore receptors (Velayudhan et al., 2000). This might be related to siderophore chelating capacities at low pH or to its ubiquity and abundance in soils, roots exudates and plant debris. In agreement with this finding, none of the *Acidithiobacilli* or *Leptospirlli* species examined encoded functions involved in standard siderophore production (e.g. bactins, chelins, etc), while all share the capacity to synthesize citrate through TCA cycling suggesting that alternative strategies for Fe(III) acquisition are being employed.

While both the iron-oxidizing *Acidithiobacilli* and *Leptospirilli* share conserved Fe(III) uptake system components, Fe(II) uptake functions are absent in *Leptospirilli* (Table 1). We favor the hypothesis that this absence is a functional adaptation of the *Leptospirilli* to growth at lower pH (<1) and higher iron concentrations (500 mM) compared to *A. ferrooxidans* (200 mM). This provides a strategy to evade imminent oxidative stress associated with transient intracellular Fe(II) overloads. In addition, this hypothesis is consistent with the observation that the *Leptospirilli* employ only Fe(III) uptake systems which are known to be amenable to tight regulation at several levels. These features might explain why *Leptospirilli* dominate the microbial population as oxidation proceeds and the Fe(II)/Fe(III) ratio decreases (Rawlings et al., 1999) and why the presence of high concentrations of Fe(III) selects against *A. ferrooxidans*. In addition, in the absence of iron biooxidation, soluble and highly available Fe(II), persisting in the environment, would be the preferred form of iron as a micronutrient for sulfur oxidizing *Acidithiobacilli* as suggested by the iron management genes observed (Table 1).

3.3. Transcriptional regulation of iron management functions

Both *Acidithiobacilli* and *Leptospirilli* encode orthologs of the iron dependent ferric uptake regulator (Fur) implicated in the transcriptional control of genes related to iron metabolism in well defined microbial model systems and shown to be functional in *A. ferrooxidans* (Quatrini et al., 2005b). The sequence and genomic context of Fur are highly conserved in the three *Acidithiobacilli* (Fig. 2). In addition, although there is gene context diversity, key functional residues of Fur are also conserved in the *Leptospirilli*, suggesting that an iron dependant Fur controls iron related functions in the two groups of microorganisms.

In addition, the co-occurrence of several predicted Fur target gene functions between *A. ferrooxidans* and the other species, suggests functional linkage and similar metal dependent regulation (e.g. Fig. 2). Reconstruction of the Fur-dependent genetic regulatory networks for these genomes in a comparative context substantiates the identification of individual target promoters in *A. ferrooxidans* and outline differences with respect to *Leptospirilli* (Quatrini et al., 2007).

3.4. New insights into iron management in acidophiles

New and unusual gene arrangements have been detected for iron uptake functions in *Acidithiobacilli* and mechanistic variations for iron uptake appear to have evolved for the acquisition of chelated Fe(III). For example, it is suggested that a malate dehydrogenase, working together with an acetyltransferase and a dicarboxylate efflux pump,

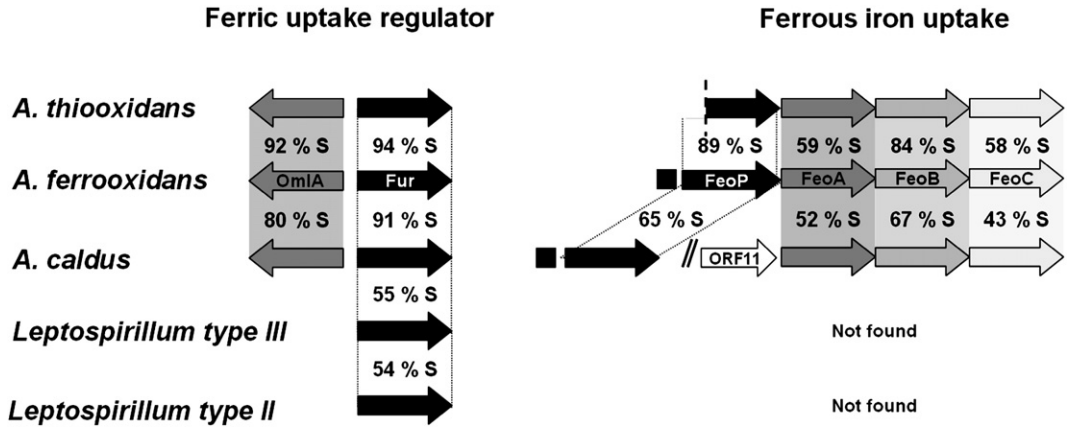

Fig. 2. Sequence similarity and gene positional analysis between *Acidithiobacilli* and *Leptospirilli*. S = percent similarity of each gene with reference to *A. ferrooxidans*. OmlA: outer membrane lipoprotein A, Fur: Fe-dependent Fur regulator, FeoP: porin, FeoABC: Fe(II) uptake transporter. Black box: Fur binding site.

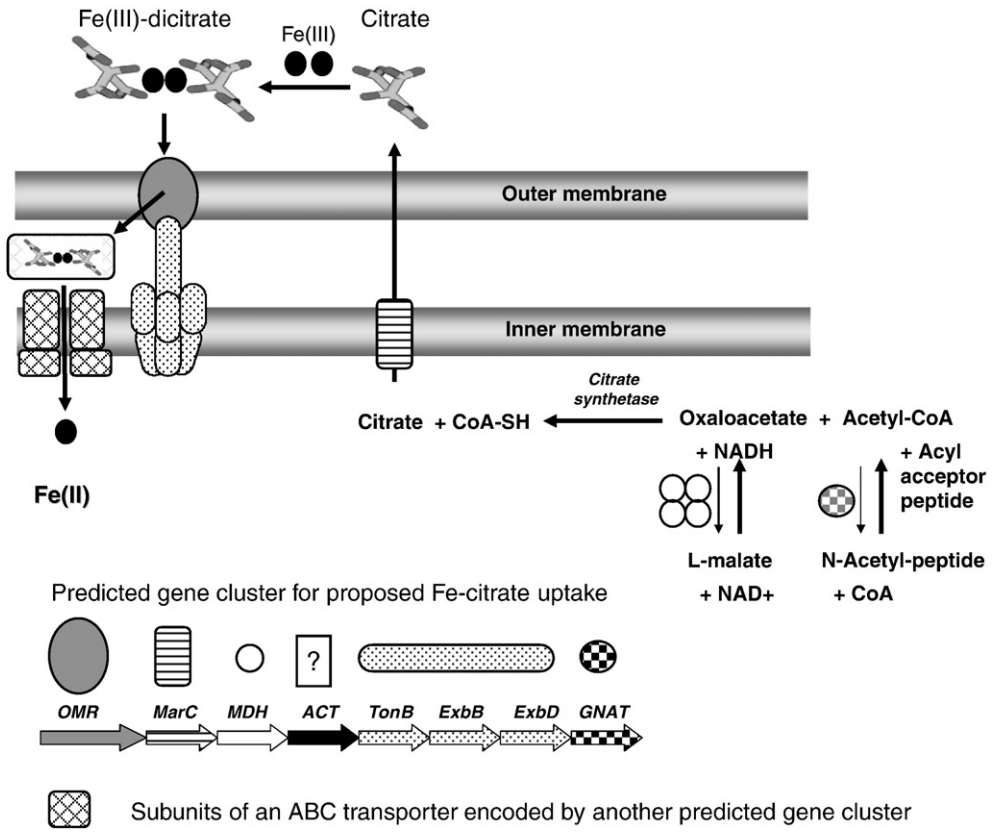

Fig. 3. Model for the acquisition of citrate chelated Fe(III) in acidophiles. Inset: Gene positional analysis in *Acidithiobacilli*. OMR = Ferric dicitrate type outer membrane receptor, MarC = dicarboxylate efflux pump, MDH = malate dehydrogenase, ACT = ACT domain protein of unknown function, TonBExbBD = TonB system for energy transduction, GNAT = Gcn5-related N-acetyltransferase family protein.

synthesize and secrete citrate, in order to acquire citrate chelated Fe (III) by cognate Fec type OMRs (Fig. 3). A gene locus that includes all these functions is fully conserved between *A. ferrooxidans* and *A. thiooxidans*, but was not detected in *A. caldus* or the *Leptospirilli* (Fig. 3, inset). Interestingly, all biomining microorganisms have OMRs with predicted affinity for Fe(III)-dicitrate regardless of their predicted ability to synthesize citrate. Being common to all biomining microorganisms, this mechanism of iron acquisition seems to have been under strong selection pressure and to be essential for function and survival in this kind of environmental conditions.

3.5. Contributions of iron management insights to biomining

During bioleaching, significant variations in pH, oxygen availability and salt concentrations occur. All three parameters directly affect iron solubility. It is likely that niche partitions and ecological successions between biomining microorganisms can be at least partially explained by changes in iron bioavailability. Mechanisms that help the microorganisms to deal with such changes are thus deemed to be critical for fitness and survival of bioleaching microbes and their understanding might contribute to improving the capacity to control the bioleaching processes.

More iron transport pathways are present in *Acidithiobacilli* than in *Leptospirilli*, providing the former with more versatility to acquire various environmental forms of iron. This characteristic may partially explain why *A. ferrooxidans* is more abundant in heap leaching environments than *Leptospirilli* (Personal communication, C. Demergasso). Differential iron uptake capacities might also help explain *Leptospirillum* strains dominance over *A. ferrooxidans* during bio-oxidation of minerals in tank reactors that build up very high concentrations ferric iron.

On the other hand, sulfate produced during S^0 oxidation strongly interacts with Fe(III) and forms complex iron oxides precipitates. Thus, S^0 oxidizing *Acidithiobacilli*, might compromise the bioavailability of Fe(III) and thereby have a greater need for high affinity Fe(III) transporters. The greater iron uptake genetic potential predicted in *A. ferrooxidans* raises important questions regarding potential alternative life styles of this bacterium and greater adaptability compared to other acidophiles (e.g. growth at higher pHs).

Acknowledgements

Work supported by Fondecyt 1050063, Fondecyt 11060164, DI-UNAB 34-06 and a Microsoft Sponsored Research Award.

References

Altschul, S.F., Gish, W., 1996. Local alignment statistics. Ed. R. Doolittle. Meth.Enzimol 266, 460–480.
Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. J. Mol. Biol 215, 403–410.
Bateman, A., Birney, E., Durbin, R., Eddy, S.R., Howe, K.L., Sonnhammer, E.L., 2000. The Pfam protein families database. Nucleic Acids Res. 28, 263–266.
CBGB: http://www.cienciavida.cl/cbgb.htm/.
Cornell, R.M., Schwertmann, U., 2003. The Iron Oxides: Structure, Properties, Reactions, Occurrences and Uses, 2nd Ed. Wiley-VCH, Weinheim.
Fleischmann, R.D., Adams, M.D., White, O., Clayton, R.A., Kirkness, E.F., Kerlavage, A.R., Bult, C.J., Tomb, J.F., Dougherty, B.A., Merrick, J.M., et al., 1995. Whole-genome random sequencing and assembly of *Haemophilus influenzae*. Rd. Sci. 269, 496–512.
Haft, D.H., Loftus, B.J., Richardson, D.L., Yang, F., Eisen, J.A., Paulsen, I.T., White, O., 2001. TIGRFAMs: a protein family resource for the functional identification of proteins. Nucleic Acids Res. 29, 41–43.s.
Henikoff, J.G., Greene, E.A., Pietrokovski, S., Henikoff, S., 2000. Increased coverage of protein families with the blocks database servers. Nucleic Acids Res. 28, 228–230.
Hofmann, K., Stoffel, W., 1993. TMbase — a database of membrane spanning proteins segments. Biol. Chem. Hoppe-Seyler 374, 166–170.
JGI: http://genome.jgi-psf.org/microbiomes/acidm/acidm.home.html.

Kanehisa, M., Araki, M., Goto, S., Hattori, M., Hirakawa, M., Itoh, M., Katayama, T., Kawashima, S., Okuda, S., Tokimatsu, T., Yamanishi, Y. in press. KEGG for linking genomes to life and the environment. Nucleic Acids Res. 36.

Karp, P.D., Ouzounis, C.A., Moore-Kochlacs, C., Goldovsky, L., Kaipa, P., Ahren, D., Tsoka, S., Darzentas, N., Kunin, V., Lopez-Bigas, N., 2005. Expansion of the BioCyc collection of pathway/genome databases to 160 genomes. Nucleic Acids Res. 19, 6083–6089.

Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H., Valentin, F., Wallace, I.M., Wilm, A., Lopez, R., Thompson, J.D., Gibson, T.J., Higgins, D.G., 2007. ClustalW2 and ClustalX version 2. Bioinformatics 23 (21), 2947–2948.

Malki, M., González-Toril, E., Sanz, J.L., Gómez, F., Rodríguez, N., Amils, R., 2006. Importance of the iron cycle in biohydrometallurgy. Hydrometallurgy 83, 223–228.

Mulder, N.J., 2007. New developments in the InterPro database. Nucleic Acids Res. D224–D228.

NCBI : http://www.ncbi.nlm.nih.gov/sites/entrez.

Nielsen, H., Engelbrecht, J., Brunak, S., von Heijne, G., 1997. Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. Protein Eng. 10, 1–6.

Nierman, W.C., Feldblyum, T.V., Laub, M.T., Paulsen, I.T., Nelson, K.E., Eisen, J., Heidelberg, J.F., Alley, M.R., Ohta, N., Maddock, J.R., et al., 2001. Complete genome sequence of *Caulobacter crescentus*. Proc. Natl. Acad. Sci. U. S. A. 98, 4136–4141.

Overbeek, R., Larsen, N., Walunas, T., D'Souza, M., Pusch, G., Selkov Jr., D., Liolios, K., Joukov, V., Kaznadzey, D., Anderson, I., Bhattacharyya, A., Burd, H., Gardner, W., Hanke, P., Kapatral, V., Mikhailova, N., Vasieva, O., Osterman, A., Vonstein, V., Fonstein, M., Ivanova, N., Kyrpides, N., 2003. The ERGOTM genome analysis and discovery system. Nucleic Acids Res. 31 (1), 164–171.

Quatrini, R., Jedlicki, E., Holmes, D.S.J., 2005a. Genomic insights into the iron uptake mechanisms of the biomining microorganism *Acidithiobacillus ferrooxidans*. Indust. Microbiol. Biotech. 32, 606–614.

Quatrini, R., Lefimil, C., Holmes, D.S., Jedlicki, E., 2005b. The ferric iron uptake regulator (Fur) from the extreme acidophile *Acidithiobacillus ferrooxidans*. Microbiol 151, 2005–2015.

Quatrini, R., Lefimil, C., Veloso, F.A., Pedroso, I., Holmes, D.S., Jedlicki, E., 2007. prediction and experimental verification of Fur-regulated genes in the extreme acidophile *Acidithiobacillus ferrooxidans*. Nucleic Acids Res. 35, 2153–2166.

Rawlings, D.E., Tributsch, H., Hansford, G.S., 1999. Reasons why 'Leptospirillum'-like species rather than *Thiobacillus ferrooxidans* are the dominant iron-oxidizing bacteria in many commercial processes for the biooxidation of pyrite and related ores. Microbiology 145, 5–13.

Rutherford, K., Parkhill, J., Crook, J., Horsnell, T., Rice, P., Rajandream, M.A., Barrell, B., 2000. Artemis: sequence visualization and annotation. Bioinformatics 16, 944–945.

Salzberg, S.L., Delcher, A.L., Kasif, S., White, O., 1998. Microbial gene identification using interpolated Markov models. Nucleic Acids Res. 26, 544–548.

Schneider, T.D., 1997. Information content of individual genetic sequences. J. Theor. Biol 189, 427–441.

TIGR: http://www.tigr.org/.

Tatusov, R.L., Koonin, E.V., Lipman, D.J., 1997. A genomic perspective on protein families. Science 278, 631–637.

Tyson, G.W., Chapman, J., Hugenholtz, P., Allen, E.E., Ram, R.J., Richardson, P.M., Solovyev, V.V., Rubin, E.M., Rokhsar, D.S., Banfield, J.F., 2004. Community structure and metabolism through reconstruction of microbial genomes from the environment. Nature 428, 37–43.

Valdés, J., Pedroso, I., Quatrini, R., Tettelin, H., Eisen, J.A., Holmes, D.S., submitted for publication. Modeling *Acidithiobacillus ferrooxidans* Metabolism: from Genome Sequence to Industrial Applications.

Velayudhan, J., Hughes, N.J., McColm, A.A., Bagshaw, J., Clayton, C.L., Andrews, S.C., Kelly, D.J., 2000. Iron acquisition and virulence in *Helicobacter pylori*: a major role for FeoB, a high-affinity ferrous iron transporter. Mol Microbiol 37, 274–286.

von Heijne, G., 1992. Membrane protein structure prediction: hydrophobicity analysis and the 'positive inside' rule. J.Mol.Biol 225, 487–494.