Genomic and metagenomic challenges and opportunities for bioleaching: a mini-review

Juan Pablo Cárdenas a, Raquel Quatrini a, David S. Holmes a,b,*

a Fundación Ciencia & Vida, Santiago, Chile
b Facultad de Ciencias Biológicas, Universidad Andres Bello, Santiago, Chile

Received 5 April 2016; accepted 29 June 2016
Available online 7 July 2016

Abstract

High-throughput genomic technologies are accelerating progress in understanding the diversity of microbial life in many environments. Here we highlight advances in genomics and metagenomics of microorganisms from bioleaching heaps and related acidic mining environments. Bioleaching heaps used for copper recovery provide significant opportunities to study the processes and mechanisms underlying microbial successions and the influence of community composition on ecosystem functioning. Obtaining quantitative and process-level knowledge of these dynamics is pivotal for understanding how microorganisms contribute to the solubilization of copper for industrial recovery. Advances in DNA sequencing technology provide unprecedented opportunities to obtain information about the genomes of bioleaching microorganisms, allowing predictive models of metabolic potential and ecosystem-level interactions to be constructed. These approaches are enabling predictive phenotyping of organisms many of which are recalcitrant to genetic approaches or are unculturable. This mini-review describes current bioleaching genomic and metagenomic projects and addresses the use of genome information to: (i) build metabolic models; (ii) predict microbial interactions; (iii) estimate genetic diversity; and (iv) study microbial evolution. Key challenges and perspectives of bioleaching genomics/metagenomics are addressed.

Keywords: Bioleaching; Genomics; Metagenomics; Bioinformatics; Acidophile; Biomining

1. Introduction

Bioleaching involves the chemical microbe-assisted solubilization of sulfidic minerals for metal recovery at an industrial scale [1−3]. At the present time, copper is the principle primary metal recovered, although other metals such as nickel can be recovered from mixed-metal ores [4]. In general, the heap bioleaching process consists of crushing ore to the size of gravel, piling the crushed ore in a heap and then applying sulfuric acid to the surface of the heap promoting the growth of acidophilic microorganisms (optimal pH for growth <3). The acidophilic microorganisms carry out biochemical reactions that, coupled with chemical reactions, solubilize the copper [2].

The key aspects of bioleaching that are relevant for this mini-review are that: (i) it involves consortia of acidophilic microorganisms that include chemolithoautotrophic iron and sulfur oxidizers that fix CO₂ from air and heterotrophs that scavenge fixed carbon from the excretions or dead remains of the chemolithoautotrophs; (ii) it involves oxidative processes that use oxygen from the air as a terminal electron acceptor; (iii) several of the key metabolic and chemical reactions involved are exothermic, driving the temperature of the bioleaching heap from ambient temperatures at the beginning to as high as 70 °C over a period of weeks to months [5]; and (iv) the sources of ferrous iron and sulfur compounds that support microbial metabolism are solids, and microbial attack of these substrates often involves cell adhesion and biofilm formation.
2. Defining the focus of the mini-review

This review focuses on the genomics and metagenomics of acidophilic microorganisms from bioleaching heaps or closely related mining environments. The definition of what constitutes a closely related mining environment is moot. This is an especially important point to clarify because much of the genomic and metagenomic data that has been used to propose models of bioleaching are derived from microorganisms that inhabit other environments such as acid mine drainage (AMD) and acidic (hot) springs. One argument that can be put forward is that many of the microorganisms found in the latter environments are similar, according to 16S rDNA data, to those that have been detected in bioleaching heaps. Therefore, a case can be made that, in the absence of data from authentic bioleaching-derived genomes, related acidophiles serve as credible surrogates for developing genetic and metabolic models of individual species and for suggesting ecophysiological interactions that could occur during bioleaching. But nagging questions arise that need attention; how credible are these models and to what extent can ecophysiological interactions be predicted from such (potentially inadequate or even incorrect) data? Unfortunately, similarity of 16S rDNA sequences between two microorganisms is not sufficient to state that these microorganisms contain the same complement of genes. This concept is developed more fully in a later section (Comparative genomics: estimating genetic and metabolic diversity).

We have included in this mini-review a discussion of genomic and metagenomic information derived from bioleaching heaps and from related mining environments including AMD. We have excluded a discussion of genomics and metagenomics of thick streamer biofilms in AMDs, such as those found in Iron Mountain, because this has been reviewed elsewhere [6,7]. However, we have included some genomes from AMDs that seep out of mines (water-column) and bioleaching heaps. These AMDs share important environmental properties with bioleaching heap environments that help determine microbial composition, such as low pH, high metal concentrations and availability of iron and sulfur. Also, with exceptions such as AMD from coal and lignite mines, both environments are practically depleted of organic matter, a characteristic that promotes the growth of chemolithoautotrophs as drivers of primary biological production [8-10]. However, bioleaching econiches are much more variable than AMDs in several of the environmental cues highlighted above, with concentrations of metals and protons building up dramatically during much shorter mineral leaching cycles. Hence, although the microbial biodiversity of bioleaching heaps might be expected to exhibit some similarity to the assemblages of microorganisms from AMD, it displays important differences.

We have also included in this mini-review some genomes from mining-related environments that are not bioleaching heaps. We justify their inclusion because these genomes include the type strains of \textit{Acidithiobacillus ferrooxidans} (coal waste), \textit{Acidithiobacillus caldus} (coal waste) and \textit{Acidithiobacillus thiooxidans} (Kimmeridge clay), considered to be major players in bioleaching, and their genomic analyses have produced some of the more advanced models of genetic and metabolic prediction.

3. Current status of genome projects from bioleaching heaps and related mining environments

As of March 2016, there are 157 genomes of acidophiles deposited in public databases. Of these, 29 (20%) are derived from microorganisms associated with bioleaching heaps or related biomining environments (Fig. 1). A list of these genomes is provided in Table 1. Three metagenome studies have been carried out on bioleaching heaps [11,12] (Table 1), whereas ten metagenomic studies of other acidic environments have been published (reviewed in Ref. [13]).

4. Bioinformatic prediction of genetic and metabolic potential

Genomics has allowed unprecedented insights into the genetic and metabolic potential of acidophiles of bioleaching microorganisms and their close relatives. Many of these microorganisms are recalcitrant to genetic manipulation, and bioinformatic analysis of genome information has been a major route for gaining insight into their biology. Bearing the caveat in mind that much of this model building has come from analyses of genomes not directly derived from bioleaching heaps, we address how genome models have been used to predict genes and metabolism and the ecophysiological interactions that are hypothesized to occur during bioleaching.

![Image](https://example.com/image.png)

**Fig. 1.** Distribution of 157 (archaeal and bacterial) acidophile genomes sorted by environmental location: bioleaching heaps and related mining environments, AMD (biofilm streamers), hot springs and other acidic environments. The chart was constructed using information derived from 151 genomes compiled in March 2015 [13] and 6 additional genomes published between March 2015 and March 2016 (NCBI Accession numbers: LPVJ00000000, LJWX00000000, LRRD00000000, LQZA00000000, JFH00000000, JXYS00000000).
### Table 1
List of available genomes and metagenomes from microorganisms from bioleaching heaps and related mining environments.

<table>
<thead>
<tr>
<th>Organism</th>
<th>NCBI Accession</th>
<th>Source</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acidiphilium angustum ATCC 39503&lt;sup&gt;T&lt;/sup&gt;</td>
<td>JN9H000000000</td>
<td>Waste coal mine waters, USA</td>
<td>N/A</td>
</tr>
<tr>
<td>Acidiphilium cryptum JF-5</td>
<td>NC_009484</td>
<td>Acidic coal mine lake sediment, Germany</td>
<td>N/A</td>
</tr>
<tr>
<td>Acidiphilium sp. JA12-A1</td>
<td>JIF0H000000000</td>
<td>Pilot treatment plant water, Germany</td>
<td>[97]</td>
</tr>
<tr>
<td>Acidithiobacillus caldus ATCC 51756&lt;sup&gt;T&lt;/sup&gt;</td>
<td>CP005986</td>
<td>Coal spoil enrichment culture, UK</td>
<td>[98]</td>
</tr>
<tr>
<td>Acidithiobacillus caldus SM-1</td>
<td>NC_015850</td>
<td>Pilot bioleaching reactor, China</td>
<td>[99]</td>
</tr>
<tr>
<td>Acidithiobacillus ferrivorans CF27</td>
<td>CCCS000000000000</td>
<td>Abandoned copper/cobalt mine drainage, USA</td>
<td>[69]</td>
</tr>
<tr>
<td>Acidithiobacillus ferrivorans SS3</td>
<td>NC_015942</td>
<td>Enrichment culture from mine-impacted soil samples, Russia</td>
<td>[100]</td>
</tr>
<tr>
<td>Acidithiobacillus ferrivorans ATCC 23270&lt;sup&gt;T&lt;/sup&gt;</td>
<td>NC_011761</td>
<td>Acid, bituminous coal mine effluent, USA</td>
<td>[27]</td>
</tr>
<tr>
<td>Acidithiobacillus ferrivorans ATCC 53993</td>
<td>NC_011206</td>
<td>Copper deposits, Armenia</td>
<td>N/A</td>
</tr>
<tr>
<td>Acidithiobacillus sp. GGI-221</td>
<td>AEF0B000000000</td>
<td>Mine water, India</td>
<td>N/A</td>
</tr>
<tr>
<td>Acidithiobacillus thiooxidans A01</td>
<td>AZM000000000</td>
<td>Wastewater of coal dump, China</td>
<td>[101]</td>
</tr>
<tr>
<td>Acidithiobacillus thiooxidans ATCC 19377&lt;sup&gt;T&lt;/sup&gt;</td>
<td>AFOH000000000</td>
<td>Kimeridge clay, UK</td>
<td>[30]</td>
</tr>
<tr>
<td>Acidithiobacillus thiooxidans Licanatay</td>
<td>JMEB000000000</td>
<td>Copper mine, Chile</td>
<td>[39]</td>
</tr>
<tr>
<td>Acidithrix ferroxidans DSM 28176&lt;sup&gt;T&lt;/sup&gt;</td>
<td>JXY500000000000</td>
<td>Acidic stream draining in abandoned copper mine, UK</td>
<td>[102]</td>
</tr>
<tr>
<td>Ferrimicrobium acidiphilum DSM 19497&lt;sup&gt;T&lt;/sup&gt;</td>
<td>JQK900000000000</td>
<td>Mine water, UK</td>
<td>[103]</td>
</tr>
<tr>
<td>Leptospirillum ferripilum DSM 14647&lt;sup&gt;T&lt;/sup&gt;</td>
<td>JPGK000000000000</td>
<td>Enrichment culture, Peru</td>
<td>[104]</td>
</tr>
<tr>
<td>Leptospirillum sp. Sp-Cl</td>
<td>LGSSH000000000000</td>
<td>Industrial bioleaching solution, Chile</td>
<td>[105]</td>
</tr>
<tr>
<td>“Ferrovum myxofaciens” P3G&lt;sup&gt;T&lt;/sup&gt;</td>
<td>JPOQQ000000000000</td>
<td>Stream draining an abandoned copper mine, UK</td>
<td>[87]</td>
</tr>
<tr>
<td>Ferrovum sp. JA12</td>
<td>LJWX000000000000</td>
<td>Pilot treatment plant water, Germany</td>
<td>[41]</td>
</tr>
<tr>
<td>Ferrovum sp. Z-31</td>
<td>LRRD000000000000</td>
<td>AMD water, Germany</td>
<td>N/A</td>
</tr>
<tr>
<td>Ferrovum sp. PN-J185</td>
<td>LQZ400000000000</td>
<td>AMD water, Germany</td>
<td>N/A</td>
</tr>
<tr>
<td>“Acidbacillus ferroxidans” DSM 5130&lt;sup&gt;T&lt;/sup&gt;</td>
<td>LPJW000000000000</td>
<td>Neutral drainage from copper mine, Brazil</td>
<td>[106]</td>
</tr>
<tr>
<td>Sulfolobacillus acidophilus DSM 10332&lt;sup&gt;T&lt;/sup&gt;</td>
<td>NC_016884</td>
<td>Coal spoil heap, UK</td>
<td>[107]</td>
</tr>
<tr>
<td>Sulfolobacillus acidocaldarius DSM 42317&lt;sup&gt;T&lt;/sup&gt;</td>
<td>LGRO000000000000</td>
<td>Percolate solution of a bioleaching heap in copper mine, Chile</td>
<td>N/A</td>
</tr>
<tr>
<td>Sulfolobacillus acidocaldarius Cutipay</td>
<td>ALWJ000000000000</td>
<td>Naturally mining environment, Chile</td>
<td>[108]</td>
</tr>
<tr>
<td>Sulfolobus acidocaldarius DSM 9293&lt;sup&gt;T&lt;/sup&gt;</td>
<td>(2506210005)&lt;sup&gt;*&lt;/sup&gt;</td>
<td>Spontaneously heated ore deposit, Kazakhstan</td>
<td>N/A</td>
</tr>
<tr>
<td>Bioleaching heap surface Metagenome</td>
<td>(4664533.3)&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Dexting Copper Mine, China</td>
<td>[11]</td>
</tr>
<tr>
<td>Bioleaching heap PLS sample Metagenome</td>
<td>(4554868.3)&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Dexting Copper Mine, China</td>
<td>[10]</td>
</tr>
<tr>
<td>Bioleaching heap sample Metagenome</td>
<td>(4554867.3)&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Dexting Copper Mine, China</td>
<td>[10]</td>
</tr>
</tbody>
</table>

Abbreviations used: <sup>T</sup> = type strain; N/A = not available; * = sequence only available in IMG-JGI where the IMG Taxon ID value is provided; a = sequence only available in MG-RAST where the correspondent ID value is provided. Excluded from this table are genomes and metagenomes from biofilm streamers in AMD, hot springs and some other miscellaneous acidic environments. A full list of acidophile genomes can be found in Ref. [13], plus there are 6 additional genomes whose accession numbers appear in the legend to Fig. 1.

The first draft genome of a bioleaching-related microorganism (A. ferrooxidans ATCC 23270, isolated from a coal waste heap) was published in 2000 and the information was used to partially reconstruct its amino acid metabolism [14] and, subsequently, to predict sulfur assimilation [15]. Fur regulation and iron homeostasis [16,17], quorum sensing [18–20], extracellular polysaccharide formation [21], CO₂ fixation [22], iron and sulfur oxidation [23,24] and other metabolic aspects [25,26]. The first complete genome sequence, also of A. ferrooxidans, was published 8 years later and formed the basis for comparative genomics based prediction of its genetic and metabolic potential that consolidated and extended earlier models, including iron and sulfur oxidation, iron reduction, CO₂ fixation, nitrogen metabolism, ecophysiology [2,9,24,27–30] and antisense RNA [31,32], as well as studies in heavy metal resistance [33,34] and quorum sensing [35–37].

A. ferrooxidans has become the best studied model for a bioleaching microorganism. However, it was recognized early on that A. ferrooxidans was only one member of a consortium of microorganisms involved in bioleaching and, since 2008, permanent draft genomes from other bioleaching microorganisms have become available, extending the prediction of genetic and metabolic potential of bioleaching microorganisms, including models for energy metabolism in A. caldus [38] and A. thiooxidans [39], in addition to models for overall metabolism for Sulfolobus spp. [40] and “Ferrovum” sp. [41].

### 5. Origin and ecological succession of acidophiles in bioleaching heaps

While it appears that an increasing number of bioleaching heaps are being inoculated with starter cultures of microorganisms by companies, many heaps are probably not inoculated; rather, the initial pioneering microorganisms are presumably derived from the original ore, and occasional seeps of microbial activity can be observed in the walls of mines. However, seeding of heaps could also be achieved via eolian dispersal [42]. Since the ore deposited in bioleaching heaps is depleted in organic matter and low in fixed nitrogen,
one might expect the pioneering microorganisms to be enriched in chemolithotrophic archaea using energy sources and to contain diazotrophs that fix nitrogen from the air. *A. ferrooxidans* fulfills these requirements, which could help explain why it is an early major colonizer of bioleaching heaps [43]. Similar questions have been posed regarding the initial colonization of fresh basaltic ash [44] and basaltic lava flows [45] that also do not contain organic matter or fixed nitrogen. Metagenomic analyses of these environments carried out shortly after volcanic eruption found that pioneering microorganisms are enriched in chemolithotrophs, that can oxidize iron and sulfur, and also in diazotrophs. Further study of the microbial pioneers and ecological succession in volcanic environments could provide insight into similar processes in bioleaching heaps.

Most, if not all, of the ore sources used to prepare bioleaching heaps are not acidic, but are rich in sulfides. When mechanically exposed, these sulfide minerals can react with water and air to produce sulfuric acid. If not neutralized by accompanying basic minerals present in the ore, the end result is a dramatic lowering of the pH [46], which may trigger the resuscitation of reservoirs of inactive acidophiles hypothesized to be present in a dormant state. Such a scenario has been described by the rare biosphere concept, coined to describe the observation that most microbial taxa are extremely uncommon in natural environments [47] and that the relative abundance of these taxa varies through time due to fluctuations in environmental conditions [48], with transitions between active and inactive metabolic states [49].

The concept of dormancy may also be important to consider when addressing the question of where the thermophiles come from that colonize bioleaching heaps after exothermic reactions have driven the temperature of the heaps up to over 55 °C. It has been shown that the more robust enzymes of thermophiles and hyperthermophiles are stabilized over a broader temperature range than in mesophiles and psychrophiles (discussed in Refs. [50,51]), and it has been shown that thermophiles can exist in a state of dormancy in moderate and even in cold conditions [52]. However, their ability to enter a dormant state in acidic conditions has not yet been evaluated, and this may present a challenge, as the external pH could endanger the proton gradient across the membrane. With this caveat, thermophiles may be present in the initial ore surviving in an inactive state in moderate temperatures. The reverse is not necessarily true. Generally, mesophiles do not survive high temperatures [53], although there may be exceptions [54]. The death and recycling of mesophiles may be an important source of fixed carbon and, perhaps, the only source of fixed nitrogen for the thermophilic bioleaching microbial community, since no nitrogen-fixing thermophiles have been reported from bioleaching heaps.

Studies linking dormancy-resuscitation dynamics to environmental drivers (e.g. Ref. [55]) in bioleaching niches are considered very relevant to improve our mechanistic understanding and modeling capacity [56] of bioleaching processes. Insights into these dynamics can be obtained from time-course experiments using targeted metagenomics and metatranscriptomic approaches. Understanding these dynamics may provide clues to the biodiversity of different bioprocesses, the patterns of microbial succession during initial and advanced stages of bioleaching and the changes in microbial community composition that control specific ecological processes within the bioleaching ecosystem.

Little is known about how the capacity to oxidize iron is “handed over” from one species to the next as temperatures rise in a bioleaching heap [43]. Are there temperatures that are compatible for more than one species that oxidize iron by different processes and, if so, are the microorganisms synergistic, commensal or competitive? Is the change from one process to the next rather abrupt or does iron oxidation segue from one process to the next? Similar questions and issues regarding sulfur oxidation can be raised.

Sulfur oxidizers use reduced inorganic sulfur compounds (RISCs) as electron sources, which are finally converted to sulfate, usually in an oxygen-dependent reaction. The protons generated in these reactions produce a decrease in pH, playing an important role in bioleaching. Sulfur is a potentially rich electron source, with up to 8 electrons available for release when a sulfur atom from a sulfide is oxidized to sulfate and the potential amount of energy that can be made available in this process is considerable [1]. RISCs are naturally present wherever sulfide-containing minerals are exposed to the surface [1]. In addition, some RISCs are released as a result of the chemical reaction of sulfide minerals with water, oxygen and ferric iron and in environments where sulfate reduction occurs [57,58]. A multiplicity of electron donors that includes sulfide, elemental sulfur, thiosulfate, etc., provides diverse options for sulfur oxidizers to use alternative energetic pathways, generating opportunities for the development of synergistic co-existence. Although comparative genomics predicts differences in sulfur oxidation pathways (e.g. Ref. [59]), little information is available regarding how they impact bioleaching, for example in the passivation, kinetics and temperature optimization of chalcopyrite leaching. A transcriptomic study of a mutant of *A. caldus*, lacking the canonical sulfur oxygenease reductase (SOR) gene, revealed a complex network of compartmentalized reactions in the sulfur oxidation superpathway that suggests that some of these pathways can be differentially expressed according to environmental conditions [38].

6. Ecophysiology of bioleaching: individual genome to community predictions

Using predictions derived from single genome analyses, preliminary models have been built that suggest the roles of different microorganisms and how they could interact during bioleaching with their dynamically changing environment, including with abiotic drivers (Fig. 2A) and with each other (Fig. 2B) [5,60]. Analyses of predicted metabolic functions across all members of the bioleaching community have generated integrated models predicting how microorganisms could share the workload of maintaining nutrient and energy budgets, including how iron and sulfur oxidation/reduction
Fig. 2. A. Schematic representation of postulated interactions between substrates (air includes CO₂, N₂ and O₂), abiotic drivers, biodiversity and ecosystem functions in a bioleaching heap. B. Model of predicted ecosystem functions and interactions and their respective microorganisms that relate Fe and S transformations with C and N fixation/cycling during the sequential development of a bioleaching heap from the start at ambient temperature to maturity at >55°C. The diagram is divided along the horizontal axis into oxic (aerobic) and microaerobic conditions. Hypothetical interactions were derived from [5,94,95].
and carbon and nitrogen fixation capacities are shared. Such studies can potentially pinpoint bottlenecks in energy production and resource allocations. For example, the prediction of the use of carboxysome-mediated CO₂ fixation in mesophiles and the reverse TCA pathway in thermophiles provided explanations [67] for CO₂ limitations in some locations within the bioleaching heap or at certain stages during the process, such as at high temperature, when the solubility of CO₂ is reduced [61–63].

A related issue involves the bioavailability of oxygen during bioleaching. Although air is pumped into a typical bioleaching heap, it still might be limiting in parts of the heap that even forced air does not penetrate, such as inside biofilms or where intensive consumption of oxygen might generate local microaerobic conditions, situations that might be exacerbated when temperatures rise, decreasing the solubility of oxygen. In such zones, iron and sulfur reduction may occur, impeding metal solubilization but, on the other hand, nitrogen fixation might be facilitated.

Nitrogen fixation is predicted to be carried out by only a very limited number of microorganisms, notably by *A. ferrooxidans* at mesophilic temperatures and *Leptospirillum* spp. at moderately thermophilic temperatures [5,64]. Although other diazotrophs might yet be discovered, at the moment, it would appear that entire microbial bioleaching communities rely on these two groups for primary production of fixed nitrogen. Interestingly, no extreme thermophiles capable of nitrogen fixation have been identified in bioleaching heaps. Does this mean that extremely thermophilic communities rely on excreted fixed nitrogen from living or dead microorganisms that grew at lower temperatures? If so, could that represent a potential limitation in biological mass production and possibly a bottleneck in metal solubilization?

7. Comparative genomics: estimating genetic and metabolic diversity

Classical microbiological studies use 16S rDNA typing to estimate the diversity of microorganisms in bioleaching heaps (e.g. Refs. [43,65,66]). While this strategy paints a broad brush picture of qualitative and quantitative biodiversity, it significantly underestimates the gene content and metabolic potential of the system. Instead, comparative genomics analyses, relying on a few sequenced strains, have begun to uncover unsuspected levels of genetic diversity between bioleaching acidophiles. Such studies also highlight the contribution of mobile genetic elements and their gene cargo to the adaptation of bioleaching acidophiles to the diverse abiotic and biotic stresses of this challenging environment [34,67,68].

For example, using complete genome comparisons, it was found that two strains of *A. ferrooxidans* (ATCC 23270 and ATCC 53993) that are 100% identical at the 16S rDNA level share only about 80% of their genes [5]. A significant number of unique genes of each strain have been shown to map to large genomic regions that fit the criteria of genomic islands. This is the case of a 300 kb genome segment in ATCC 23270 [67,37] and a 160 kb genomic island in ATCC 53993 [34], exclusive of each strain. The genomic island in ATCC 23270 has been shown to be an actively exciting integrative and conjugative element hypothesized to participate in CRISPR (clustered regularly interspaced short palindromic repeats) propagation and the establishment of acquired immunity against bacteriophages and other mobile genetic elements [70]. The genomic island in *A. ferrooxidans* ATCC 53993 has been associated with copper resistance [34]. Frequent occurrence of additional sets of tRNAs within these genomic islands has also been reported [37,69,70]; these are thought to provide a selective advantage to bioleaching acidophiles bearing these elements, possibly related to translation efficiency due to increased concentrations of core tRNAs [71]. Similarly, whole genome comparisons between two strains of *A. caldus* (ATCC 51756 and strain SM-1) detected important differences in gene content, including those encoding a conjugative system and flagellum, among other functions [68]. Also, a comparison of the genome of a cultured strain of *Acidithiobacillus ferrivorans* with partially reconstructed *A. ferrivorans* genomes in an AMD metagenome detected differences in many genes, including those associated with metal efflux and pH homeostasis [72].

Therefore, constructing models of metabolic potential and predicting ecophysiological interactions based on analyses of genomes derived from non-bioleaching environments, even if the microorganisms in question are defined as the same species based on 16S rDNA typing, probably provides an incomplete picture of the biology of bioleaching.

8. Eukaryotic genome information is poorly represented

Despite increasing knowledge of the diversity and important role of eukaryotes in acidic environments [73], including AMD ecosystems [74,75], where 4 genomes have been published [13], little is known about their presence and function in bioleaching heaps and no genomes have been sequenced thus far.

9. Progress in viral genomics and metagenomics

One of the aspects of bioleaching microbial ecology that has been significantly overlooked is that of virus–host interaction. Viruses affect genome stability and host fitness and thereby shape the structure and impact the function of microbial communities. As in other microbiologically-driven biotechnological processes, viruses could potentially cause die-offs of specific bioleaching microorganisms, perturb the successional dynamics inside microbial consortia or even cause the crash of biologically-driven mineral leaching operations. Despite their potential importance, very little is known about the influence of these biological agents on the performance of the bioleaching consortia.

Support for the existence of viruses capable of infecting bioleaching acidophiles has been obtained during the last 3 decades, thanks to the generation of genomic and metagenomic data [76]. Fuselloviruses, bicaudaviruses, ampullavirus, rudiviruses, and other bizarre types of viruses inserted...
in the genomes of the *Sulfolobales* crenarchaeae, have been described and characterized in varying degrees [77,78]. A 59-kb inducible temperate Myoviridae-like prophage located in the *srrA* tmRNA of the *A. caldus* type strain genome and a 56-kb prophage productively infecting *Thiennesia* sp. 3A have been identified and partially characterized [79,80]. Also, blocks of genes with anomalous G + C contents of putative viral origin have been found in the Leptospirillli datasets from the Iron Mountain AMD system [81,82].

On the other hand, CRISPRs known to protect microbial cells from invasion by phages and other foreign mobile genetic elements, have been shown to occur in the genomes of bioleaching microorganisms [68,70,83] and AMD community genomic datasets. Widespread presence of CRISPRs in these microorganisms supports the view that they constitute a relevant antiviral mechanism in biomining ecologies and also provides indirect evidence that viruses and other foreign elements are important factors structuring bioleaching communities. Access to diverse viral metagenomes and to the population of CRISPR spacers sequences will provide an opportunity to begin defining the links between viruses and their specific host and to model the underlying effect of viral predation or infection on bioleaching community endurance and performance.

10. “Microbial dark matter” in bioleaching heaps

The study of metagenomes from environmental samples has repeatedly provided evidence for sequences that have no hits in homology searches other than sequences from their own metagenome, suggesting that phylogenetic diversity inferences based on cultivation studies are limited and that a vast “microbiological dark matter” deserves to be critically analyzed.

Three metagenomic studies of bioleaching heaps have been carried out (Table 1). These studies are beginning to provide insight into the structuring of the microbial community and the partitioning of functions, including those necessary to maintain the carbon and nitrogen cycle during bioleaching [11,12]. Biodiversity was dominated by just a few types of microorganism, but other microbial groups were detected at low taxonomic abundance levels, in agreement with microbial diversity estimates using classical detection procedures [2].

Recently, extremely small (<0.2 µ) microorganisms with reduced genomes have been recovered from metagenomic studies of ground water, where they constitute more than 15% of the taxonomic diversity [84,85]. These microbes have been missed in previous metagenomic studies because they pass through <0.2 µm filters used to collect microorganisms from environmental samples, they are largely uncultivable and are not detected by 16S rDNA typing. There are no studies yet of such microorganisms in bioleaching heaps, but it is likely that they are present there and may be contributing to bioleaching in unknown ways.

Microorganisms that cannot be cultured or are difficult to culture have been discussed elsewhere [60], and these and others that remain to be found are easily missed in microbial surveys of bioleaching heaps. The importance of co-culture for the growth of some species can reflect syntrophic properties inside a microbial consortium [86,87]. Opportunities presented by single-cell genomics [88] and microfluidic manipulation of co-cultured microbes [89] provide a novel portal to find new organisms with unexpected roles in the bioleaching community.

11. Using genomics to find genes of biotechnological interest

Acidophiles, including bioleaching microorganisms, provide opportunities for the detection and isolation of genes encoding novel biomolecules (DNA, lipids, enzymes, osmolytes, etc.) that might be useful for biotechnological applications; scanning genomes offers a way to predict these genes. For example, licanantase, a secreted protein capable of improving copper recovery in the bioleaching process, was detected in a strain of *A. thiooxidans* [90] and bioactive compounds have been isolated from euakaryotes growing in extremely low pH drainage [91,92]. Also, a strain of *A. ferroxidans* has been engineered to produce isobutyric acid or heptadecane from CO2 and the oxidation of Fe3+ [93].

12. Conclusions and perspectives

Only a few genomes from acidophilic microorganisms derived from bioleaching heaps have been published. The majority of acidophile genomes come from AMD and other low pH environments, and these have been used as surrogates for constructing models of genetic and metabolic potential of bioleaching microorganisms and ecophysiology of bioleaching. Even though many of these non-bioleaching genomes are related at the 16S rDNA level to bioleaching microorganisms, a concern is that such model building from surrogates probably portrays an incomplete picture of the metabolic potential and interaction dynamics operating within a bioleaching heap. Clearly, more genomes need to be derived directly from bioleaching heaps, and from different stages of bioleaching and locations within a heap. However, with a few exceptions, a major problem has been the difficulty that the research community has had in gaining access to commercial bioleaching heaps to sample for microorganisms. Let us hope that this situation will improve in the future.

Very few metagenomes of bioleaching heaps are available. More are required to deepen our understanding of the microbial complexity of bioleaching processes in the context of diverse ore mineralogies and operational settings and their temporal dynamics.

Genomes and metagenomes only allow predictions of genetic and metabolic potential to be assessed. Transcriptomic, proteome and metabolic evidence is needed to actually determine what part of this potential is expressed during bioleaching and how it varies in time and space. Such data will be hard to obtain given the difficulties in isolating RNA and proteins from bioleaching heaps.
Very little metadata is available that links physico-chemical parameters within the bioleaching heap, such as mineral composition, Eh, pH, temperature, and CO2 and O2 levels, to the distribution and function of microorganisms. Such data, if collected, is often not available to the research community.

Conflict of interest

None declared.

Acknowledgments

Conicyt Basal CCTE PFB16 and Fondecyt 1130683 and 1140048.

References

J.P. Cárdenas et al. / Research in Microbiology 167 (2016) 529–538

[75] Aguilera A, Olsson S, Puente-Sánchez F. Physiological and phyloge


