

1 Analysis of co-regulated abundance of genes associated with arsenic and 2 phosphate metabolism in Andean Microbial Ecosystems

3
4 Saona L.A.¹, Valenzuela-Diaz S.², Kurth D.¹, Contreras M.³, Meneses C.^{4,5}, Castro-Nallar E.² and Farías M.E.^{1*}
5

6 ¹*Laboratorio de Investigaciones Microbiológicas de Lagunas Andinas (LIMLA), Planta Piloto de Procesos Industriales*
7 *Microbiológicos (PROIMI), CCT, CONICET, San Miguel de Tucumán, Tucumán, Argentina.*

8 ²*Facultad de Ciencias de la Vida, Center for Bioinformatics and Integrative Biology, Universidad Andrés Bello,*
9 *Santiago, Chile.*

10 ³*Centro de Ecología Aplicada, Santiago, Chile.*

11 ⁴*Facultad de Ciencias de la Vida, Centro de Biotecnología Vegetal, Universidad Andres Bello, Avenida República 330,*
12 *8370035 Santiago, RM, Chile*

13 ⁵*FONDAP Center for Genome Regulation, Santiago, Chile*

14
15 *corresponding author: mefarias2009@gmail.com

16 **Abstract**

17
18 Phosphate and arsenate are very similar compounds, and there is great interest in studying their
19 relationship and their interaction with biological systems. Despite having no apparent biological
20 function, specific genes regulate arsenic interaction with cells and can be located in regions of the
21 genome called arsenic islands, where phosphate metabolism genes are also present. Although they
22 are neighboring genes, the nature of their relationship and how they have been selected is still
23 unknown.
24

25 In this work, we analyzed the metagenomes of the four microbial ecosystems inhabiting hypersaline
26 lakes of the Argentine Puna and the Atacama salt flat in Chile and have evaluated the presence and
27 abundance of both arsenic and phosphate metabolism genes. The samples analyzed included
28 microbialites, biofilms and microbial mats; all of them established under high arsenic
29 concentrations, high UV radiation and high temperature fluctuation, among others.
30

31 The results show great differences in the dispersion and abundance of genes related to both
32 phosphate and arsenic metabolism in the analyzed samples. The main difference is given in the
33 Diamante Lake, located in the crater of the Galan volcano characterized by being one of the lakes
34 with the highest arsenic concentration (2.34 mM). Correlating genes abundance with the
35 physicochemical parameters of the lakes studied, our results suggest that arsenic and phosphate
36 metabolism are intricately co-regulated in environmental conditions.
37

38 **Keywords:** arsenate and phosphate, Pst transport system, arsenic respiration, Andean Microbial
39 Ecosystem.
40

41 **Introduction**

42
43 Phosphorus is typically found in nature as inorganic phosphate (Pi). Together with carbon,
44 hydrogen, nitrogen, oxygen, and sulfur, Pi is one of the six chemical elements essential to life as we
45 know it (Karl 2000; Rasuk et al. 2016). Pi is involved in many cellular functions as it is part of
46 DNA and RNA, ATP and polyphosphates (polyP) (Kornberg 1999), membrane phospholipids and

47 proteins. In addition, the metabolism of both prokaryotes and eukaryotes depends on reactions
48 where Pi is essential, e.g., phosphorylation/dephosphorylation reactions, which regulate protein
49 functions in intracellular signaling pathways.

50

51 Arsenic is an element structurally similar to phosphorus, they both share characteristics such as
52 atomic radius, and identical electronegativity and oxidation states. In their most characteristic
53 oxidized form (+5), arsenate [As(V)] as well as phosphate (PO_4^{3-}) are negatively charged at
54 physiological pH. Indeed, both molecules present similar speciation at different pH, as well as
55 similar pK_a values (Wolfe-Simon et al. 2008, 2011; Tawfik and Viola 2011) and thermochemical
56 radii, differing by only 4% (Kish and Viola 1999). Further, inorganic arsenate is able to form key
57 biological ester bonds which are analogous to those formed by Pi, including genetic material. This
58 occurs spontaneously between the 5'-hydroxyl group of ribose sugars and arsenate, giving rise to
59 the formation of the mononucleotides 5'-arsenate (Lagunas et al. 1984; Wolfe-Simon et al. 2008).

60

61 In spite of such similarities, arsenic is one of the most toxic elements for most lifeforms (Parke
62 2013). Biomolecules based on As(V) have a much higher rate of hydrolysis than those based on
63 phosphate (Lagunas et al. 1984; Fekry et al. 2011; Parke 2013), and –unlike phosphate– arsenic
64 presents redox reactions in the physiological range of redox potentials, thus, gradually transforming
65 As(V) into As(III) (Schoepp-Cothenet et al. 2011) in the prokaryotic cytoplasm.

66

67 In 2010 a super arsenic-resistant bacterium belonging to the Halomonadaceae family, denominated
68 GFAJ-1, was isolated by Wolfe-Simon (2011). According to the authors, the bacterium –isolated
69 from Mono Lake, which is characterized by high arsenic concentrations (200 μM) (Wolfe-Simon et
70 al. 2011; Elias et al. 2012)– might replace the phosphorus in its DNA with arsenic, enabling it to
71 grow in culture media with arsenate and without the addition of phosphate. That article challenged
72 the role of arsenic in biology, and even the concept of life we had until that moment. Although
73 arsenate as an essential component of DNA is totally discredited (Schoepp-Cothenet et al. 2011;
74 Erb et al. 2012; Reaves et al. 2012; Kim and Rensing 2012), microbial cells exhibit specific cellular
75 mechanisms that modulate the As intracellular concentration. Due to its similarities with phosphate,
76 arsenate can enter cells through phosphate transporters; consequently, various organisms possess
77 enzymes specifically related to arsenic resistance and arsenic metabolism, and the genes that code
78 for these enzymes are distributed along the three domains: Bacteria, Archaea and Eukarya (Jackson
79 and Dugas 2003).

80

81 Phosphate uptake occurs mainly by two pathways: through a phosphate specific transport (Pst)
82 system or a phosphate inorganic transport (Pit) system (Willsky and Malamy 1980b, a; Guo et al.
83 2011). Depending on extracellular phosphate concentrations, one transport system is activated over
84 the other, a regulation process that has been reported in different microorganisms such as
85 Proteobacteria, Cyanobacteria, Algae, and Archaea (Rosenberg et al. 1977, 1982; Elias et al. 2012;
86 Hudek et al. 2016; Yan et al. 2017). In presence of arsenic, phosphate uptake is also modified.
87 Under these conditions, arsenate competes for phosphate transporters and enters cells preferably
88 through the non-specific Pit transporter, which fails to distinguish between phosphate and arsenate
89 (Cleiss-Arnold et al. 2010). On the other hand, under arsenate stress, the Pst system increases its
90 specific phosphate affinity more than 100-fold for some microorganisms (Guo et al. 2011) and may
91 even exceed 4500-fold for extremophile microorganisms such as GFAJ-1 (Elias et al. 2012).

92 Moreover, genes related to arsenic metabolism (e.g., *ai*o and *arr*) are contiguous to phosphate

93 metabolism genes (like *pst*) in arsenic genomic islands (Li et al. 2013), indicating that there could
94 be functional and/or evolutionary associations to allow for the adaptation under different arsenic
95 and phosphate concentration conditions (Muller et al. 2007; Li et al. 2013). Nonetheless, this
96 potential relationship, if any, between *aii/arr* genes and *pst* genes is still unknown. Considering this
97 proximity of genes related to arsenic metabolism and phosphate transport genes in the arsenic
98 genomic islands (Li et al. 2013), we investigated how the presence and/or abundance of these genes
99 varies at different chemical conditions mainly referred to the arsenic and phosphate concentration.

100 To perform this goal, we deliver a comparative metagenomic study using samples from four High
101 Altitude Andean Lakes (HAAL): Tebenquiche and Brava in Chile (Atacama salt flat), Diamante
102 and Socompa in Argentina (Argentine Puna region). These lakes harbor polyextremophilic Andean
103 Microbial Ecosystems (AMEs) -such as biofilms, microbial mats, microbialites and endoevaporites,
104 among others- that tolerate high UV solar radiation, low oxygen pressure, extreme temperature
105 fluctuation and high concentrations of heavy metals and metalloids such as arsenic (Farías et al.
106 2013, 2014, 2017; Albarracín et al. 2015; Fernandez et al. 2016; Rasuk et al. 2016; Rascovan et al.
107 2016).

108 **Materials and Methods**

109

110 **Sampling**

111

112 Tebenquiche Lake and Brava Lake microbial mat samples were obtained in November 2012.
113 Sample coordinates: Tebenquiche (S 23°08' 18.5" W 068°14'49.9", Fig. 1B) and Brava mat (S
114 23°43'48.2" W 068°14'48.7", Fig. 1C). For DNA analyses, triplicate cores (2 cm² each) were taken
115 to a depth of 3 cm of the mat and were pooled prior to homogenizing in order to obtain
116 representative samples. Homogenates used for DNA extraction were stored at -20 °C in the dark and
117 processed within a week (Farías et al. 2014; Fernandez et al. 2016).

118

119 Stromatolites can be found along the southern shore of Socompa Lake (coordinates:
120 24°32.168'S, 68°12.165'W, Fig. 1D). Samples were taken in February 2001, sampling
121 methodology was already described by Farías (2013). These stromatolites are rounded domal
122 structures that present a clear stratification that appears in vertical sections. Stromatolite samples
123 were collected in sterile plastic bags and were fragmented into 2.5 cm-diameter cylinders that
124 measured 5 cm deep from the top of the stromatolite. They were stored at -20 °C in the dark and
125 processed within a week (Farías et al. 2013).

126

127 Biofilm samples were collected in November 2016 from Diamante Lake, located inside the Galan
128 Volcano crater, located at 4589 masl (coordinates: 26 °00'09.71"S, 67 °02'34.91"W, Fig. 1E).
129 Galan Volcano lies in the province of Catamarca, Argentina and reaches 5912 masl. Three
130 independent samples were taken from Diamante Lake, and each was carried out by pooling together
131 three representative samples. The samples were collected in centrifuge tubes, stored at -20 °C in the
132 dark until they were processed (Rascovan et al. 2016).

133

134 For phosphate and arsenic content analysis, all water samples (1 L) were collected immediately
135 over the sampling sites of the corresponding sediment systems and stored in acid-cleaned bottles on
136 ice in the dark.

137

138 **Water physicochemical analysis**

139

140 Tebenquiche and Brava water samples were stored in acid-cleaned bottles, on ice and in the dark
141 until analyses in the laboratory were carried out within 48 h. Dissolved oxygen, salinity,
142 conductivity, total P, NO₃⁻, NO₂⁻, dissolved Si, Ca²⁺, Mg²⁺, K⁺, SO₄²⁻, and Na⁺, were
143 measured according to the methodology described by Eaton et al. (2005). NH₄⁺, orthophosphates,
144 and Total Organic Nitrogen (TON) were analyzed using a Merck Nova 60 Spectro Photometer by
145 following standard methods, as described by the American Public Health Association (1998).

146

147 For the Socompa Lake and Diamante Lake samples: nutrients, ions and general chemical analyses
148 (arsenic and phosphate) were performed in a chemical IRAM-certified laboratory at Estación
149 Experimental Obispo Colombes, Tucumán, Argentina (<http://www.eeaoc.org.ar/>).

150

151 **DNA extraction and sequencing**

152

153 For all samples, total metagenomic DNA used for sequencing was extracted from a mixture of three
154 replicate extractions using Power Biofilm DNA Isolation Kit (MO BIO Laboratories, Inc.) from
155 0.5g of collected biofilm/stromatolites/microbial mat which was processed according to
156 manufacturer's indications. From the biofilm of Diamante Lake, 3 independent extractions from 3
157 different replicates were made. In the case of stromatolites samples, first they were frozen and then
158 were lyophilized and homogenized into a powder that was processed (Farías et al. 2013).

159

160 Total DNA was sheared by sonication with a target fragment size of ~400 bp. DNA fragment
161 distribution was confirmed in a Fragment Analyzer (Advanced Analytical) and libraries were
162 constructed using the Illumina TruSeq DNA kit 2x300 bp as per manufacturer's instructions.
163 Libraries were multiplexed and run in an Illumina MiSeq instrument. Raw data for Diamante Lake's
164 red biofilm is available from NCBI's SRA SRP136179 under BioProject PRJNA438526; for
165 Tebenquiche and La Brava is available from ENA European Nucleotide Archive (ENA) under
166 accession number: PRJEB25599; and for Socompa's stromatolites is available from NCBI's SRA
167 SRP072938 under BioProject PRJNA317551.

168

169 **Metagenomic analyses**

170

171 Metagenome raw data from each lake were filtered at PHRED > Q20 using Prinseq version 0.20.4
172 (Schmieder and Edwards 2011). Then, reads were assembled using 3.9.0 SPAdes version
173 (Bankevich et al. 2012) with the --meta parameter to call the metaSPAdes module. The four
174 assemblies were annotated using Prokka version 1.11 (Seemann 2014), adding the --metagenome
175 parameter to improve gene prediction. Since many genes involved in arsenic and phosphate
176 metabolism are known, we quantified their difference in abundance among the four metagenomes
177 studied. Phosphate metabolism genes included in the analysis were: *ppk*, *ppx*, *pitA*, *phoB*, *phoR*,
178 *pstS*, *pstC*, *pstA*, *pstB* and *phoU*; and the genes related to arsenic were: *acr3*, *aioA*, *aioB*, *arrA*, *arsR*,
179 *arsD*, *arsA*, *arsB* and *arsC*. Relative abundances of these genes were obtained by mapping the reads
180 against the contigs of each lake metagenome using bowtie2 (Langmead and Salzberg 2012). The
181 raw counts of each gene were calculated with a htseq-count, a module of HTSeq (Anders et al.
182 2015). We used DESeq2 (Love et al. 2014) to normalize sequence coverage among the

183 metagenomes and to perform inter-sample comparisons (negative binomial normalization;
184 parametric Wald test at $FDR < 0.01$; Benjamini-Hochberg correction for multiple testing). Intra-
185 sample comparisons for gene abundance analysis were performed using the FPKM metric. The
186 taxonomic profile of the lakes was obtained by recruiting the small ribosomal sub unit gene
187 sequence as implemented in metaxa2 (Bengtsson-Palme et al. 2015) by aligning the filtered reads
188 against archaea, bacteria and eukaryote databases (parameter -t a,b,e of metaxa2). The resulting
189 OTUs were analyzed in R using the phyloseq package (McMurdie and Holmes 2013).

190

191 **Results**

192

193 **Characteristics of High-Altitude Andean Lakes**

194

195 The HAAL studied here are characterized as shallow saline or thalassic lakes located in the Chilean
196 Atacama Desert and in the Argentine Puna (Fig. 1A). All of them represent practically unexplored
197 environments that have a set of unique extreme characteristics. Tebenquiche and Brava Lakes (Fig.
198 1B and 1C, respectively) are located in the lower region of the Atacama basin (Chile). Lake water
199 comes from underground sources of tertiary and quaternary volcanic origin (Risacher and Alonso
200 1996) and they are characterized by a high concentration of lithium, boron and arsenic (Lara et al.
201 2012; Farías et al. 2014). Specifically, Tebenquiche Lake is one of the largest water bodies in the
202 Atacama salt flat (Demergasso et al. 2008). Socompa Lake (Fig. 1D) is located at 3.800 meters
203 above sea level in a basin at the base of the active Socompa Volcano (Argentina). As with the rest
204 of the lakes, arsenic concentration is very high (Farías et al. 2013). Finally, Diamante Lake
205 (Argentina) (Fig. 1E) is located at 4.570 meters above sea level within the crater of the Galan
206 Volcano, near a hydrothermal spring effluent. Apart from sharing all the HAAL characteristics, its
207 arsenic concentration is extremely high (Fig. 2), more than 10 times the concentration of Mono
208 Lake, a habitat that features high arsenic content (Wolfe-Simon et al. 2011; Rascovan et al. 2016).

209

210 In order to determine arsenic and phosphate concentrations in the four Andean lakes, a
211 physicochemical analysis was performed which revealed that Tebenquiche, Brava and Socompa
212 (TBS) lakes have low concentrations of both phosphate and arsenic in comparison to Diamante
213 Lake. Tebenquiche showed the lowest concentrations with 0.011 mM and 0.053 mM of phosphate
214 and arsenic, respectively. On the other hand, the concentration of phosphate and arsenic in
215 Diamante Lake was 5.100 and 2.340 mM, respectively (Fig. 2), which shows that, compared to
216 Tebenquiche, arsenic concentration is 40-fold higher.

217

218

219 **Taxonomic compositions of metagenomes**

220

221 We found significant taxonomic similarities between TBS lakes, whereas Diamante Lake displayed
222 differences regarding the dominant phyla, mainly due to the high abundance of the Euryarchaeota
223 phyla. This observation agrees with previous reports pertaining Diamante Lake (Rascovan et al.
224 2016). On the other hand, Fig. 3B shows the most abundant taxonomic groups per sample classified
225 at the family level, revealing large differences among all the lakes with a predominant
226 Halobacteriaceae family in Diamante Lake. While TBS are dominated by bacteria, Diamante is
227 dominated by archaea, where the most abundant phylum is Euryarchaeota. Indeed, this phylum
228 represents 94% of the total microbial community (Rascovan et al. 2016). In Diamante, the only

229 phylum belonging to the bacteria domain is Proteobacteria represented in more than 5%. At the
230 family level, the taxonomic similarity between TBS is not as evident as it is at the phylum level
231 with the 5 most abundant families that represent 49.7%, 43.5% and 39.7%, respectively.
232 Spirochaetaceae is the unique family shared by Tebenquiche and Socompa (Fig. 3B). There are also
233 no major taxonomic similarities between the studied lakes when examining the most abundant
234 families (Fig. S1). Alpha diversity consistently showed that Diamante is less diverse compared to
235 the other lakes (Fig. 3C; using 4 diversity metrics), which in turn concurs with the taxonomic
236 profile at the Phylum and Family levels (Fig. 3A and 3B). Altogether, this suggests that Diamante
237 Lake imposes harsh conditions for microbial life to thrive.

238

239 **Phosphate metabolism genes.**

240

241 *PhoR/phoB* genes belong to the phosphate (Pho) regulon, and are implicated in bacterial/archaeal
242 phosphate regulation (Wanner and Chang 1987; Santos-Beneit 2015). Through *phoR* and *phoB*
243 genes, the Pho regulon is involved in phosphate concentration sensing on the periplasm (Lamarche
244 et al. 2008). TBS lakes have much less phosphate than Diamante Lake (Fig. 2) and exhibit higher
245 *phoR* and *phoB* gene abundance. The *phoR* gene is 7.26, 16.57 and 3.46 times less abundant in
246 Diamante than in TBS, respectively. Likewise, the abundance values for the *phoB* gene are 13.54,
247 25.82 and 21.26. On the other hand, another key molecular player in phosphate metabolism is the
248 PhoU enzyme –a metal binding protein– that is involved in the Pst phosphate transport system and
249 belongs to the *pstSCAB-phoU* operon (Steed and Wanner 1993; Lamarche et al. 2008; McCleary
250 2017). Unlike *phoR* and *phoB*, the *phoU* gene is 280.15, 689.69 and 146.02 times more abundant in
251 Diamante than in TBS lakes, respectively (Fig. 4), suggesting that Diamante might favor the
252 transcription of genes related to specific phosphate transporters (Pst) by means of PhoU, whereas,
253 the synthesis of these transporters in the other lakes might be very well regulated by PhoB and
254 PhoR.

255

256 The genes of the *pst* operon (*pstS*, *pstC*, *pstA* and *pstB*) encode for a high-affinity ABC type
257 phosphate transporter. Specifically, the *pstS* gene codes for a periplasmic phosphate-binding
258 protein, whose function is to bind phosphate from the external medium and transfer it to another
259 protein complex (PstCAB) that will enter the cell cytoplasm (Blus-Kadosh et al. 2013). Our results,
260 unexpectedly, show that the *pstS* gene is 1.90, 1.81 and 1.43 times more abundant in Diamante than
261 in TBS. The *pstC* and *pstA* genes coding for inner-membrane channel proteins for Pst transport
262 (Lamarche et al. 2008) are more abundant in Diamante than in the other lakes, with the exception of
263 Brava Lake that has 1.49 times more *pstC* than Diamante. The *pstC* gene is 3.27 and 5.66 times
264 more abundant in Diamante than in Tebenquiche and Socompa, respectively; whereas *pstA* is 5.10,
265 5.66 and 5.78 times more abundant in Diamante than in TBS, respectively (Fig. 4).

266

267 Regarding the Pit system –the low-affinity phosphate transport–, the *pitA* gene is 1.18, 3.29 and
268 1.15 times more abundant in Diamante Lake than in TBS, respectively (Fig. 4). The intra-sample
269 comparison reveals a very low relevance of the Pit transporter in all lakes, moreover, the abundance
270 of the *pitA* gene is the lowest among the phosphate metabolism genes evaluated (Fig. S2), which
271 suggests that phosphate uptake would take place mostly by means of the Pst system.

272

273 Finally, as to polyphosphate metabolism genes –*ppk* and *ppx*–, both show very similar abundance
274 in TBS lakes, while at Diamante abundance is very low. Specifically, the *ppk* gene is 18.89, 20.83

275 and 18.01 times less abundant in Diamante than in TBS lakes, respectively; in the same order, the
276 values for the *ppx* gene are: 1.53, 1.50 and 2.48.

277

278 **Arsenic resistance and metabolism genes**

279

280 The analysis of *ars* operon genes reveals great differences between the four metagenomes studied.
281 The *arsR* gene has a similar abundance in Diamante, Tebenquiche and Brava lakes, while in
282 Socompa it is 2.33, 1.34 and 2.89 times more abundant, respectively. The *arsD* and *arsA* genes are
283 very well distributed in Diamante compared to the other lakes; *arsD* is 4.86, 17.39 and 1.56 times
284 more abundant in Diamante than in Tebenquiche, Brava and Socompa, while in the same order, the
285 values for the *arsA* gene are: 7.68, 9.72 and 6.50. The *arsC* gene is significantly more abundant in
286 the lakes where arsenic concentration is low, as it is 24.30, 34.30 and 15.78 times more abundant in
287 Tebenquiche, Brava and Socompa, respectively, than in Diamante. The arsenate reductase enzyme,
288 ArsC, is responsible for reducing As(V) into As(III) which is subsequently expelled from the cell by
289 the ArsB efflux pump. Though the *arsB* gene was not found in Diamante Lake, intra-sample
290 comparisons made using the FPKM method show that the *arsB* gene is not abundant in the other
291 three lakes either (Fig. S3).

292

293 In addition to the ArsB efflux pump, there is another arsenic efflux pump widely distributed among
294 microorganisms: the Acr3. The *acr3* gene is 3.27, 3.21 and 3.21 times more abundant in Diamante
295 than in Tebenquiche, Brava and Socompa, respectively (Fig. 5). Meanwhile, the abundance of the
296 *acr3* gene is equivalent among Tebenquiche, Brava and Socompa, and its abundance in each
297 individually is much higher than the abundance of the *arsB* gene (Fig. S3). This suggests that the
298 expulsion of arsenic from the cytoplasm might be carried out mainly by the Acr3 pump, which is
299 common to the 4 lakes.

300

301 Finally, the genes corresponding to arsenic respiration –*aioAB* and *arrA*– were analyzed (Rascovan
302 et al. 2016). The *aioA* and *aioB* genes that encode an arsenite oxidase have a much greater relative
303 abundance in Diamante Lake than in the others. Diamante and Brava show the largest difference,
304 where *aioA* and *aioB* are 101.15 and 44.95 times more abundant in Diamante, respectively. On the
305 other hand, the *arrA* gene coding for arsenate reductase presents a great difference as it is much
306 more abundant in Diamante Lake and *arrA* is 30.91 and 90.51 times more abundant in there than in
307 Tebenquiche and Brava. For its part, Socompa does not show significant differences with Diamante
308 in terms of the *arrA* gene (Fig. 5).

309

310 **Discussion**

311

312 Our work characterizes the microbial communities and analyzes the abundance of genes related to
313 arsenic and phosphate metabolisms found in four salt lakes distributed along the Argentine Puna
314 and the Atacama Desert in Chile. The results reveal that the greatest differences in gene abundance
315 appear at Diamante Lake in contrast with TBS lakes (Fig. 4 and Fig. 5). These differences are also
316 reflected in the distribution of species and biodiversity (Fig. 3); While archaea comprises 94% of
317 Diamante (Rascovan et al. 2016), in the other lakes the main species are bacteria (Fariás et al. 2013;
318 Fernandez et al. 2016; Rasuk et al. 2016), and have a much higher biodiversity than Diamante (Fig.
319 3D). Finally, in terms of phosphate and arsenic concentration, physicochemical conditions also
320 evidence considerable differences between Diamante and the other lakes (Fig 2).

321

322 The *aioA/B* genes encode proteins that use As(III) as an electron donor, oxidizing it to As(V).
323 Because these enzymes have been described as transmembrane proteins (Andres and Bertin 2016),
324 this process occurs within the extracellular environment in the presence of arsenic. As(V), on the
325 other hand, is a substrate for another family of transmembrane enzymes, the ArrA/B, which use
326 As(V) as the final acceptor of the electron transport chain and reduce it to As(III). The theorization
327 of this process, has shown that cells can *breathe* arsenic and thereby obtain energy from its
328 oxidation and reduction (Rascovan et al. 2016; Andres and Bertin 2016). The *aio* and *arr* genes are
329 very well distributed in Diamante Lake and their abundance is much higher than in the other lakes
330 analyzed (those with the lowest arsenic concentration) (Fig. 2). In summary, the higher the arsenic
331 concentration, the higher *aio* and *arr* gene abundance.

332

333 Regarding the *arsRDABC* operon, both *arsA* and *arsD* genes are highly abundant in Diamante
334 compared to the other lakes, yet the same does not apply to *arsR*, *arsB* and *arsC* genes. The *arsRBC*
335 operon confers basal tolerance to arsenic, mainly by means of the arsenate reductase ArsC, which is
336 in charge of the electrochemical transformation of As(V) to As(III), subsequently expelled from the
337 cell by the ArsB efflux pump (Dey and Rosen 1995; Lin et al. 2006; Andres and Bertin 2016).
338 These genes belong to the *arsRBC* operon and their abundance in Diamante is very low, particularly
339 *arsC*. As a matter of fact, *arsB* is simply not present, even when mapping the reads against a
340 Hidden Markov Model (HMM) of the *arsB* gene, no matches are found. Although *arsB* did not turn
341 out to be a particularly abundant gene in TBS lakes, it is very likely that it is not present in
342 Diamante. The Diamante metagenome was sequenced in triplicate from three different samples to
343 minimize the risk of not finding a gene due to lack of sequencing (BioSamples accessions
344 SAMN08719551, SAMN08719552 and SAMN08719553).

345

346 Unexpectedly, arsenic resistance genes from *arsRBC* operon were (in the case of *arsC*) up to 35
347 time more abundant in lakes with low arsenic concentration, while in Diamante, where arsenic
348 concentration is up to 40 times greater than in the other lakes, only genes related to oxidation and
349 reduction of extracellular arsenic are more abundant. Previous works have demonstrated that *arsC*
350 activity is modified in the presence of phosphate (Zhang et al. 2014, 2017); it may be the case that
351 despite high arsenic concentration, there is low arsenic uptake by the cells, rendering an arsenate
352 reductase (ArsC) not vital for survival. In addition, the expulsion of arsenic from the cell -in all
353 lakes- might not be performed by ArsB but by Acr3, a pump that does not belong to the *arsRBC*
354 operon. An intra-sample analysis shows that the *acr3* gene is very well distributed in all lakes,
355 especially in Diamante (Fig. S3). Previous findings have shown *acr3*-gene abundance in
356 *Exiguobacterium* strain genomes (Ordoñez et al. 2015; Castro-Severyn et al. 2017), nevertheless,
357 further analyses are necessary to understand the relationship between the *arsB* and *acr3* genes, as
358 well as their abundance according to the physicochemical conditions of different environments.

359

360 In the case of phosphate transport systems, both the Pst and Pit, are distributed in the four
361 metagenomes studied. However, while the Pit system does not show remarkable differences among
362 the 4 lakes, the Pst system is high abundant in Diamante Lake in comparison with TBS lakes (Fig.
363 4). The Pst is a high affinity and a low speed transport system, while the Pit transport system is
364 characterized as a constitutive shock-resistant system which does not depend on membrane-bound
365 proteins like the Pst (Lamarche et al. 2008; McCleary 2017). Previous studies have shown that the
366 Pst system is activated in conditions of phosphate deprivation (Sprague et al. 1975; Richards and

367 Vanderpool 2012; Elias et al. 2012; Reistetter et al. 2013), however in presence of high phosphate
368 concentration it remains active playing a key role (Guo et al. 2011; Hudek et al. 2016). Based on
369 this, apparently, the Pst system would be key to the phosphate uptake at low and high phosphate
370 concentrations (Bergwitz and Jüppner 2011). This is consistent with our results, why in the
371 Diamante Lake, where the phosphate concentration is excessively high, is there a greater abundance
372 of Pst genes compared to the TBS lakes?

373 On the other hand, studies with *Halomonas* GFAJ-1 have discovered the presence of two *pst*
374 operon, where the PstS phosphate binding protein of one of them selects phosphate at least 500-fold
375 over arsenate (Elias et al. 2012), whereas the PstS belonging to the Pst2 operon shows 4,500 times
376 more affinity (Wolfe-Simon et al. 2011; Yan et al. 2017). As mentioned by Schoepp-Cothenet
377 (2011), it seems that the *Halomonas* GFAJ-1 has evolved to incorporate enough small amounts of
378 phosphate to survive while intoxicating with arsenic (Schoepp-Cothenet et al. 2011); this is why
379 GFAJ-1 would be able to grow with trace phosphate concentrations (3 μ M) even in high arsenic
380 concentration. Our results support that hypothesis. In Diamante Lake where arsenic concentration is
381 extremely high (2.34 mM) and despite the high phosphate concentration, the Pst system is more
382 abundant in comparison with the other lakes. The intra-sample analysis in Diamante Lake reveals
383 the *pstSCAB-phoU* operon genes are the most relevant genes belonging to phosphate metabolism,
384 and their abundances are even greater than the *dnaG* housekeeping gene (Fig. S2).

385 Under arsenic stress, cells need to have more specific phosphate uptake systems (Li et al. 2013),
386 besides, through the transcriptional profile of *H. arsenicoxydans* it was observed that under
387 conditions of As(V) exposition, phosphate uptake occurs preferentially through the specific Pst
388 system rather than the Pit general transport, in order to reduce the entry of As(V) (Cleiss-Arnold et
389 al. 2010). This phenomenon could mean that evolutionarily the microbial life in Diamante Lake had
390 to select specific phosphate transporter system to avoid arsenic uptake. By means of *in silico*
391 analyses of bacterial genomes, arsenic islands have been discovered where *pst* genes are next to *aio*
392 genes (Li et al. 2013; Chen et al. 2015). The relation of the *aio* genes with the *pst* operon around the
393 arsenic islands is not yet elucidated (Lebrun et al. 2003; Kim and Rensing 2012), however, as other
394 authors have suggested, it is possible that through the arsenic islands, arsenic induces production of
395 Pst transporters that allows the cells to obtain the phosphate they need in arsenic laden hypersaline
396 lakes (Willisky and Malamy 1980a; Cleiss-Arnold et al. 2010; Li et al. 2013).

397 Regarding the *pst* operon regulator, Diamante lake shows a huge abundance of *phoU* gene and a
398 very low presence of *phoR/phoB* genes (Fig. 4). This is consistent with expectations since
399 *phoR/phoB* genes are expressed under conditions of low phosphate concentration (Lamarche et al.
400 2008; Chen et al. 2015; McCleary 2017), while *phoU* are expressed under conditions of high
401 phosphate concentration (Steed and Wanner 1993; Lamarche et al. 2008). In addition, PhoU is
402 essential to avoid an uncontrolled uptake of Pi that could become toxic to the cells (Surin et al.
403 1986; Santos-Beneit 2015), which is what could be happening in Diamante Lake. However, the
404 whole functions of PhoU are not elucidated yet, even less its relationship with arsenic metabolism.
405 On the other hand, the PhoB protein is essential for polyP accumulation, in fact, *E. coli* mutant
406 strains lacking *phoB* do not have the ability to accumulate intracellular polyP (Rao et al. 1998). The
407 polyP metabolism genes (*ppk* and *ppk*) are not well distributed in Diamante Lake (Fig. S2), instead
408 both genes are very abundant in TBS Lakes (Fig. 4).

409 High expression of phosphate metabolism-related genes (Pst system and Pho regulon mainly)
410 influenced by presence of arsenic in microorganisms has been previously verified (Cleiss-Arnold et
411 al. 2010; Kang et al. 2012a; Li et al. 2013). In the same way, the arsenic metabolism (*arsC* mainly)
412 affected by the phosphate concentration has been also documented (Markley and Herbert 2010;
413 Slaughter et al. 2012; Kang et al. 2012b; Wang et al. 2013; Zhang et al. 2014). In accordance with
414 this, our results show that in the face of high phosphate and arsenic concentration: (i) the *aio* and
415 *arr* genes are very abundant suggesting that obtaining energy from arsenic substrates would be a
416 well distributed process, (ii) high specificity of phosphate transporters (Pst) would possibly restrict
417 arsenic uptake, (iii) *phoU* gene is extremely abundant, this would act by regulating phosphate
418 homeostasis avoiding high phosphate uptake (iv) *arsC* gene involving in the reduction of arsenate is
419 not abundant suggesting that arsenic uptake would be restricted in this conditions. On the other
420 hand, in face of low phosphate concentration (and more arsenic than phosphate), our results show:
421 (i) *aio* and *arr* genes are not abundant, (ii) *phoB/phoR* genes involving in the activation of Pst
422 system in phosphate starvation are very abundant and finally, (iii) *arsC* is highly abundant, maybe
423 because in these conditions the arsenic uptake is higher.

424

425 These results expand our understanding of how the arsenic and phosphate metabolism would be
426 intricately co-regulated and the way that the environmental factors would influence on the microbial
427 metabolism through genes selection. With this manuscript, we attempt to contribute to the current
428 discussion about As and P metabolism, specifically in polyextremophile conditions.

429

430 **Acknowledgements**

431

432 ECN was funded by “CONICYT-FONDECYT de iniciación en la investigación 11160905”. ECN
433 would like to thank George Washington University’s high-performance computing facility,
434 Colonial One, for providing data storage, support, and computing power for metagenomic analyses
435 (colonialone.gwu.edu/).

436

437 **Author contributions statement**

438

439 LAS contributed with the principal idea of this work, participated in the study design, performed
440 data analysis, interpreted data and wrote the paper. SVD contributed with design methodology for
441 data analysis, performed data analysis, and contributed with article writing. DK contributed with the
442 work proposal. MC obtained funding for the original project idea and performed the
443 physicochemical analysis. CM contributed with sequencing of some metagenomes. ECN
444 participated in the study design, choice of methodology and data analysis, sequenced some of the
445 metagenomes, and helped write the manuscript. MEF obtained funding for the original project idea,
446 contributed with the work proposal, with the sampling and sequencing of the metagenomes. All
447 authors read and approved this manuscript.

448

449 **Conflict of Interest**

450

451 The authors declare no conflict of interest.

452

453 **References**

454

- 455 Albarracín VH, Kurth D, Ordoñez OF, et al (2015) High-Up: A Remote Reservoir of Microbial
456 Extremophiles in Central Andean Wetlands. *Front Microbiol* 6:1404. doi: 10.3389/fmicb.2015.01404
- 457 Anders S, Pyl PT, Huber W (2015) HTSeq--a Python framework to work with high-throughput sequencing
458 data. *Bioinformatics* 31:166–169. doi: 10.1093/bioinformatics/btu638
- 459 Andres J, Bertin PN (2016) The microbial genomics of arsenic. *FEMS Microbiol Rev* 40:299–322. doi:
460 10.1093/femsre/fuv050
- 461 Bankevich A, Nurk S, Antipov D, et al (2012) SPAdes: A New Genome Assembly Algorithm and Its
462 Applications to Single-Cell Sequencing. *J Comput Biol* 19:455–477. doi: 10.1089/cmb.2012.0021
- 463 Bengtsson-Palme J, Hartmann M, Eriksson KM, et al (2015) metaxa 2: improved identification and
464 taxonomic classification of small and large subunit rRNA in metagenomic data. *Mol Ecol Resour*
465 15:1403–1414. doi: 10.1111/1755-0998.12399
- 466 Bergwitz C, Jüppner H (2011) Phosphate Sensing. *Adv Chronic Kidney Dis* 18:132–144. doi:
467 10.1053/j.ackd.2011.01.004
- 468 Blus-Kadosh I, Zilka A, Yerushalmi G, et al (2013) The Effect of *pstS* and *phoB* on Quorum Sensing and
469 Swarming Motility in *Pseudomonas aeruginosa*. *PLoS One* 8:e74444. doi:
470 10.1371/journal.pone.0074444
- 471 Castro-Severyn J, Remonsellez F, Valenzuela SL, et al (2017) Comparative Genomics Analysis of a New
472 Exiguobacterium Strain from Salar de Huasco Reveals a Repertoire of Stress-Related Genes and
473 Arsenic Resistance. *Front Microbiol* 8:456. doi: 10.3389/fmicb.2017.00456
- 474 Chen F, Cao Y, Wei S, et al (2015) Regulation of arsenite oxidation by the phosphate two-component system
475 PhoBR in *Halomonas* sp. HAL1. *Front Microbiol* 6:
- 476 Cleiss-Arnold J, Koechler S, Proux C, et al (2010) Temporal transcriptomic response during arsenic stress in
477 *Herminiimonas arsenicoxydans*. *BMC Genomics* 11:709. doi: 10.1186/1471-2164-11-709
- 478 Demergasso C, Escudero L, Casamayor EO, et al (2008) Novelty and spatio-temporal heterogeneity in the
479 bacterial diversity of hypersaline Lake Tebenquiche (Salar de Atacama). *Extremophiles* 12:491–504
- 480 Dey S, Rosen BP (1995) Dual mode of energy coupling by the oxyanion-translocating ArsB protein. *J*
481 *Bacteriol* 177:385–9
- 482 Elias M, Wellner A, Goldin-Azulay K, et al (2012) The molecular basis of phosphate discrimination in
483 arsenate-rich environments. *Nature* 491:134–137. doi: 10.1038/nature11517

- 484 Erb TJ, Kiefer P, Hattendorf B, et al (2012) GFAJ-1 Is an Arsenate-Resistant, Phosphate-Dependent
485 Organism. *Science* (80-) 337:
- 486 Farías ME, Contreras M, Rasuk MC, et al (2014) Characterization of bacterial diversity associated with
487 microbial mats, gypsum evaporites and carbonate microbialites in thalassic wetlands: Tebenquiche and
488 La Brava, Salar de Atacama, Chile. *Extremophiles* 18:311–329
- 489 Farías ME, Rascovan N, Toneatti DM, et al (2013) The Discovery of Stromatolites Developing at 3570 m
490 above Sea Level in a High-Altitude Volcanic Lake Socompa, Argentinean Andes. *PLoS One* 8:e53497.
491 doi: 10.1371/journal.pone.0053497
- 492 Farías ME, Rasuk MC, Gallagher KL, et al (2017) Prokaryotic diversity and biogeochemical characteristics
493 of benthic microbial ecosystems at La Brava, a hypersaline lake at Salar de Atacama, Chile. *PLoS One*
494 12:. doi: 10.1371/journal.pone.0186867
- 495 Fekry MI, Tipton PA, Gates KS (2011) Kinetic Consequences of Replacing the Internucleotide Phosphorus
496 Atoms in DNA with Arsenic. *ACS Chem Biol* 6:127–130. doi: 10.1021/cb2000023
- 497 Fernandez AB, Rasuk MC, Visscher PT, et al (2016) Microbial Diversity in Sediment Ecosystems
498 (Evaporites Domes, Microbial Mats, and Crusts) of Hypersaline Laguna Tebenquiche, Salar de
499 Atacama, Chile. *Front Microbiol* 7:1284. doi: 10.3389/fmicb.2016.01284
- 500 Guo P, Gong Y, Wang C, et al (2011) Arsenic speciation and effect of arsenate inhibition in a *Microcystis*
501 *aeruginosa* culture medium under different phosphate regimes. *Environ Toxicol Chem* 30:1754–1759.
502 doi: 10.1002/etc.567
- 503 Hudek L, Premachandra D, Webster WAJ, Bräu L (2016) Role of Phosphate Transport System Component
504 PstB1 in Phosphate Internalization by *Nostoc punctiforme*. *Appl Environ Microbiol* 82:6344–6356.
505 doi: 10.1128/AEM.01336-16
- 506 Jackson CR, Dugas SL (2003) Phylogenetic analysis of bacterial and archaeal *arsC* gene sequences suggests
507 an ancient, common origin for arsenate reductase. *BMC Evol Biol* 3:18. doi: 10.1186/1471-2148-3-18
- 508 Kang YS, Bothner B, Rensing C, McDermott TR (2012a) Involvement of RpoN in regulating bacterial
509 arsenite oxidation. *Appl Environ Microbiol* 78:5638–5645. doi: 10.1128/AEM.00238-12
- 510 Kang YS, Heinemann J, Bothner B, et al (2012b) Integrated co-regulation of bacterial arsenic and
511 phosphorus metabolisms. *Environ Microbiol* 14:3097–3109. doi: 10.1111/j.1462-2920.2012.02881.x
- 512 Karl DM (2000) Aquatic ecology: Phosphorus, the staff of life. *Nature* 406:31–33. doi: 10.1038/35017683

- 513 Kim E-H, Rensing C (2012) Genome of halomonas strain GFAJ-1, a blueprint for fame or business as usual.
514 J Bacteriol 194:1643–5. doi: 10.1128/JB.00025-12
- 515 Kish MM, Viola RE (1999) Oxyanion Specificity of Aspartate- β -semialdehyde Dehydrogenase. Inorg Chem
516 38:818–820. doi: 10.1021/ic981082j
- 517 Kornberg A (1999) Inorganic Polyphosphate: A Molecule of Many Functions. Springer, Berlin, Heidelberg,
518 pp 1–18
- 519 Kurth D, Amadio A, Ordoñez OF, et al (2017) Arsenic metabolism in high altitude modern stromatolites
520 revealed by metagenomic analysis. Sci Rep 7:1024. doi: 10.1038/s41598-017-00896-0
- 521 Lagunas R, Pestana D, Diez-Masa JC (1984) Arsenic mononucleotides. Separation by high-performance
522 liquid chromatography and identification with myokinase and adenylate deaminase. Biochemistry
523 23:955–960. doi: 10.1021/bi00300a024
- 524 Lamarche MG, Wanner BL, Crépin S, et al (2008) The phosphate regulon and bacterial virulence: a
525 regulatory network connecting phosphate homeostasis and pathogenesis. FEMS Microbiol Rev 32:461–
526 473. doi: 10.1111/j.1574-6976.2008.00101.x
- 527 Langmead B, Salzberg SL (2012) Fast gapped-read alignment with Bowtie 2. Nat Methods 9:357–359
- 528 Lara J, González LE, Ferrero M, et al (2012) Enrichment of arsenic transforming and resistant heterotrophic
529 bacteria from sediments of two salt lakes in Northern Chile. Extremophiles 16:523–538
- 530 Lebrun E, Brugna M, Baymann F, et al (2003) Arsenite Oxidase, an Ancient Bioenergetic Enzyme. Mol Biol
531 Evol 20:686–693. doi: 10.1093/molbev/msg071
- 532 Li H, Li M, Huang Y, et al (2013) In silico analysis of bacterial arsenic islands reveals remarkable synteny
533 and functional relatedness between arsenate and phosphate. Front Microbiol 4:347. doi:
534 10.3389/fmicb.2013.00347
- 535 Lin Y-F, Walmsley AR, Rosen BP (2006) An arsenic metallochaperone for an arsenic detoxification pump.
536 Proc Natl Acad Sci U S A 103:15617–22. doi: 10.1073/pnas.0603974103
- 537 Love MI, Huber W, Anders S (2014) Moderated estimation of fold change and dispersion for RNA-seq data
538 with DESeq2. Genome Biol 15:550. doi: 10.1186/s13059-014-0550-8
- 539 Markley CT, Herbert BE (2010) Modeling Phosphate Influence on Arsenate Reduction Kinetics by a
540 Freshwater Cyanobacterium. Environ Model Assess 15:361–368. doi: 10.1007/s10666-009-9212-8
- 541 McCleary WR (2017) Molecular Mechanisms of Phosphate Homeostasis in *Escherichia coli*. In: *Escherichia*

- 542 *coli* - Recent Advances on Physiology, Pathogenesis and Biotechnological Applications. InTech
- 543 McMurdie PJ, Holmes S (2013) phyloseq: an R package for reproducible interactive analysis and graphics of
544 microbiome census data. PLoS One 8:e61217
- 545 Muller D, Médigue C, Koechler S, et al (2007) A Tale of Two Oxidation States: Bacterial Colonization of
546 Arsenic-Rich Environments. PLoS Genet 3:e53. doi: 10.1371/journal.pgen.0030053
- 547 Ordoñez OF, Lanzarotti E, Kurth D, et al (2015) Genome comparison of two Exiguobacterium strains from
548 high altitude andean lakes with different arsenic resistance: identification and 3D modeling of the Acr3
549 efflux pump. Front Environ Sci 3:50. doi: 10.3389/fenvs.2015.00050
- 550 Parke EC (2013) What could arsenic bacteria teach us about life? Biol Philos 28:205–218. doi:
551 10.1007/s10539-012-9338-y
- 552 Rao NN, Liu S, Kornberg A (1998) Inorganic polyphosphate in Escherichia coli: the phosphate regulon and
553 the stringent response. J Bacteriol 180:2186–93
- 554 Rascovan N, Maldonado J, Vazquez MP, Eugenia Farías M (2016) Metagenomic study of red biofilms from
555 Diamante Lake reveals ancient arsenic bioenergetics in haloarchaea. ISME J 10:299–309. doi:
556 10.1038/ismej.2015.109
- 557 Rasuk MC, Fernández AB, Kurth D, et al (2016) Bacterial Diversity in Microbial Mats and Sediments from
558 the Atacama Desert. Microb Ecol 71:44–56. doi: 10.1007/s00248-015-0649-9
- 559 Reaves ML, Sinha S, Rabinowitz JD, et al (2012) Absence of Detectable Arsenate in DNA from Arsenate-
560 Grown GFAJ-1 Cells. Science (80-) 337:
- 561 Reistetter EN, Krumhardt K, Callnan K, et al (2013) Effects of phosphorus starvation versus limitation on
562 the marine cyanobacterium *Prochlorococcus* MED4 II: gene expression. Environ Microbiol 15:2129–
563 2143. doi: 10.1111/1462-2920.12129
- 564 Richards GR, Vanderpool CK (2012) Induction of the Pho regulon suppresses the growth defect of an
565 Escherichia coli sgrS mutant, connecting phosphate metabolism to the glucose-phosphate stress
566 response. J Bacteriol 194:2520–30. doi: 10.1128/JB.00009-12
- 567 Risacher F, Alonso H (1996) Geoquímica del Salar de Atacama, parte 2: evolución de las aguas. Andean
568 Geol 23:123–134
- 569 Rosenberg H, Gerdes RG, Chegwidden K (1977) Two systems for the uptake of phosphate in Escherichia
570 coli. J Bacteriol 131:505–11

- 571 Rosenberg H, Russell LM, Jacomb PA, Chegwiddden K (1982) Phosphate exchange in the pit transport
572 system in *Escherichia coli*. *J Bacteriol* 149:123–30
- 573 Santos-Beneit F (2015) The Pho regulon: a huge regulatory network in bacteria. *Front Microbiol* 6:402. doi:
574 10.3389/fmicb.2015.00402
- 575 Schmieder R, Edwards R (2011) Quality control and preprocessing of metagenomic datasets. *Bioinformatics*
576 27:863–864. doi: 10.1093/bioinformatics/btr026
- 577 Schoepp-Cothenet B, Nitschke W, Barge LM, et al (2011) Comment on “A bacterium that can grow by
578 using arsenic instead of phosphorus”; author reply 1149. *Science* 332:1149; author reply 1149. doi:
579 10.1126/science.1201438
- 580 Seemann T (2014) Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 30:2068–2069. doi:
581 10.1093/bioinformatics/btu153
- 582 Slaughter DC, Macur RE, Inskip WP (2012) Inhibition of microbial arsenate reduction by phosphate.
583 *Microbiol Res* 167:151–156. doi: 10.1016/j.micres.2011.05.007
- 584 Sprague GF, Bell RM, Cronan JE (1975) A mutant of *Escherichia coli* auxotrophic for organic phosphates:
585 evidence for two defects in inorganic phosphate transport. *Mol Gen Genet* 143:71–7
- 586 Steed PM, Wanner BL (1993) Use of the rep technique for allele replacement to construct mutants with
587 deletions of the *pstSCAB-phoU* operon: evidence of a new role for the PhoU protein in the phosphate
588 regulon. *J Bacteriol* 175:6797–6809
- 589 Surin B, Dixon N, Rosenberg H (1986) Purification of the *phoU* protein, a negative regulator of the *pho*
590 regulon of *Escherichia coli* K-12. *J Bacteriol*
- 591 Tawfik DS, Viola RE (2011) Arsenate Replacing Phosphate: Alternative Life Chemistries and Ion
592 Promiscuity. *Biochemistry* 50:1128–1134. doi: 10.1021/bi200002a
- 593 Wang NX, Li Y, Deng XH, et al (2013) Toxicity and bioaccumulation kinetics of arsenate in two freshwater
594 green algae under different phosphate regimes. *Water Res* 47:2497–2506. doi:
595 10.1016/j.watres.2013.02.034
- 596 Wanner BL, Chang BD (1987) The *phoBR* operon in *Escherichia coli* K-12. *J Bacteriol* 169:5569–74. doi:
597 10.1128/JB.169.12.5569-5574.1987
- 598 Willsky GR, Malamy MH (1980a) Characterization of two genetically separable inorganic phosphate
599 transport systems in *Escherichia coli*. *J Bacteriol* 144:356–65

- 600 Willsky GR, Malamy MH (1980b) Effect of arsenate on inorganic phosphate transport in *Escherichia coli*. *J*
601 *Bacteriol* 144:366–74
- 602 Wolfe-Simon F, Davies PCW, Anbar A (2008) Did nature also choose arsenic? *Nat Preced*. doi:
603 10.1038/NPRE.2008.1482.1
- 604 Wolfe-Simon F, Switzer Blum J, Kulp TR, et al (2011) A bacterium that can grow by using arsenic instead
605 of phosphorus. *Science* 332:1163–6. doi: 10.1126/science.1197258
- 606 Yan Y, Ding K, Yu X-W, et al (2017) Ability of Periplasmic Phosphate Binding Proteins from
607 *Synechocystis* sp. PCC 6803 to Discriminate Phosphate Against Arsenate. *Water, Air, Soil Pollut*
608 228:148. doi: 10.1007/s11270-017-3334-4
- 609 Zhang S-Y, Su J-Q, Sun G-X, et al (2017) Land scale biogeography of arsenic biotransformation genes in
610 estuarine wetland. *Environ Microbiol* 19:2468–2482. doi: 10.1111/1462-2920.13775
- 611 Zhang S, Rensing C, Zhu Y-G (2014) Cyanobacteria-Mediated Arsenic Redox Dynamics Is Regulated by
612 Phosphate in Aquatic Environments. *Environ Sci Technol* 48:994–1000. doi: 10.1021/es403836g
613

614 **Fig. 1.- Location and photographs of the High-Altitude Andean Lakes under study.** (A) Geographical location of
615 the four lakes under study: Tebenquiche and Brava in Chile, Diamante and Socompa in Argentina. (B) Tebenquiche
616 Lake; (C) Brava Lake; (D) Socompa Lake; (E) Diamante Lake. Photo credits to Farías ME and Saona LA.

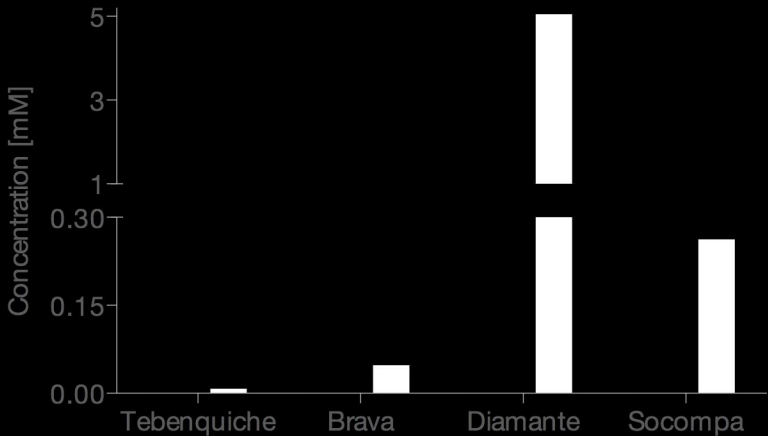
617
618 **Fig. 2.- Chemical analysis of samples used for metagenomic studies.** Phosphate and arsenic concentration of the
619 samples obtained from Tebenquiche, Brava, Diamante and Socompa lakes were quantified. Black bars represent arsenic
620 concentration and white bars phosphate concentration.

621
622 **Fig. 3.- Comparison of microbial diversity in Tebenquiche, Brava, Diamante, and Socompa lakes.** (A) Stacked bar
623 chart representing relative abundance of dominant phyla in the four lakes. (B) Stacked bar chart representing relative
624 abundance at the family level. (C) Alpha diversity with four measures: Observed, Chao1, Shannon and InvSimpson.

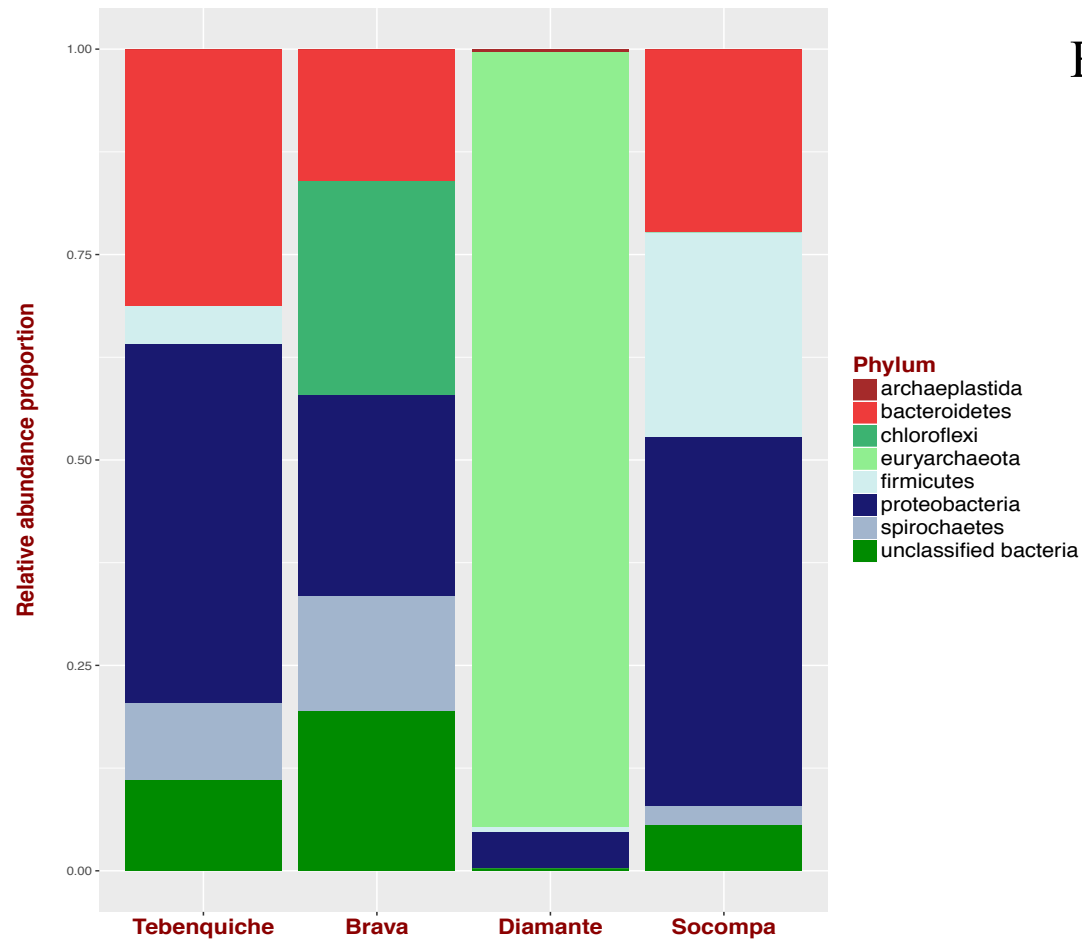
625
626 **Fig. 4.- Differential abundance analysis of genes associated with phosphate metabolism.** The differential
627 abundance of the *ppk*, *ppx*, *pitA*, *phoB*, *phoR*, *pstS*, *pstC*, *pstA*, *pstB* and *phoU* genes in all the lakes studied was
628 calculated. The results were plotted by heatmap when comparing one lake against another. The biggest differences are
629 in the Diamante Lake.

630
631 **Fig. 5.- Differential abundance analysis of genes associated with arsenic metabolism.** The differential abundance of
632 the *acr3*, *aioA*, *aioB*, *arrA*, *arsA*, *arsC*, *arsD* and *arsR* genes in all the lakes studied was calculated. The results were
633 plotted by heatmap when comparing one lake against another. The biggest differences are in the Diamante Lake.
634

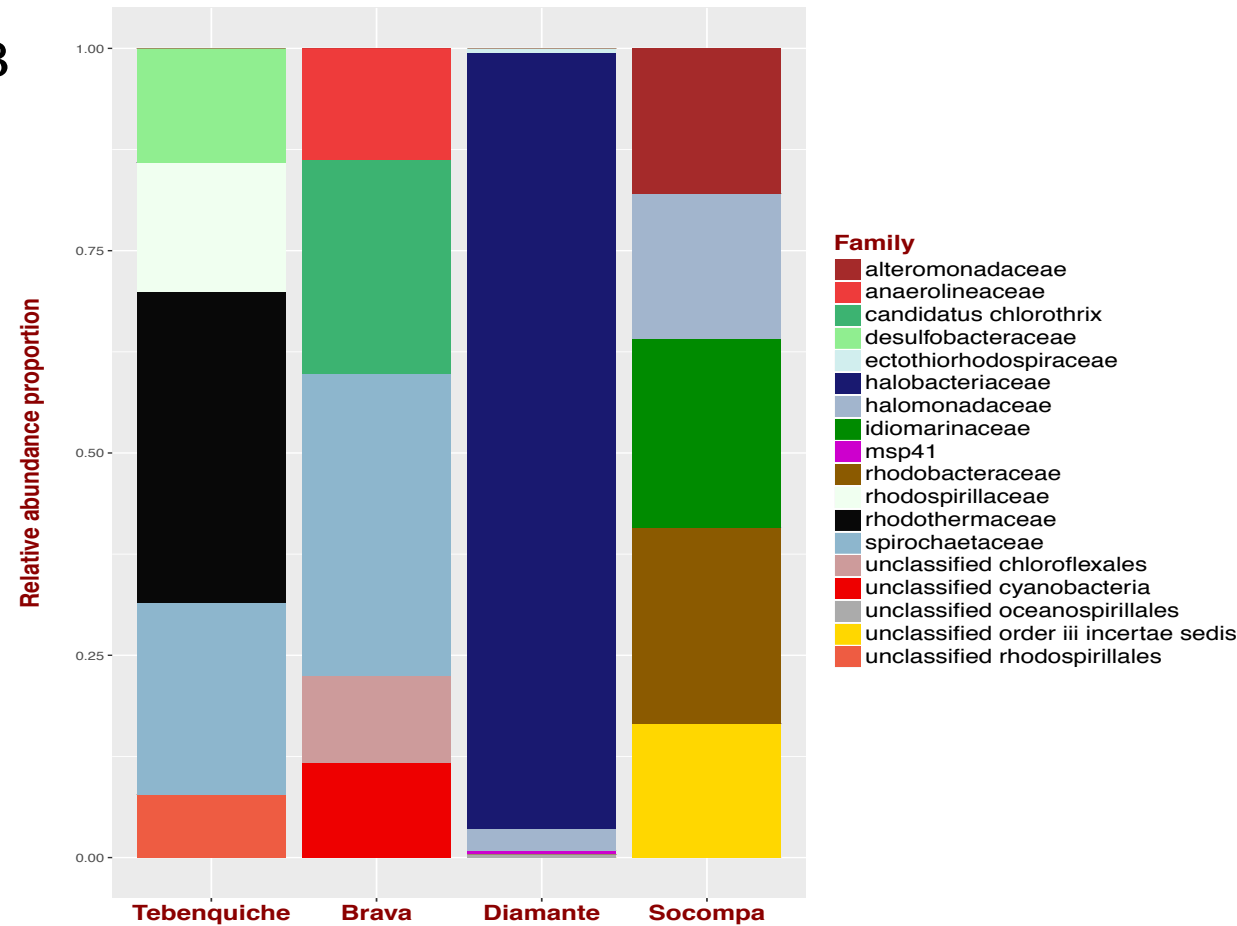




A



B



C

