



# A meta-taxonomic investigation of the prokaryotic diversity of water bodies impacted by acid mine drainage from the São Domingos mine in southern Portugal

Sara Ettamimi<sup>1</sup> · Jorge D. Carlier<sup>2</sup> · Cymon J. Cox<sup>2</sup> · Youssef Elamine<sup>3</sup> · Khalil Hammani<sup>1</sup> · Hassan Ghazal<sup>4</sup> · Maria C. Costa<sup>2,5</sup>

Received: 16 June 2019 / Accepted: 23 September 2019 / Published online: 9 October 2019  
© Springer Japan KK, part of Springer Nature 2019

## Abstract

The prokaryotic communities of water bodies contaminated by acid mine drainage from the São Domingos mining area in southern Portugal were analyzed using a meta-taxonomics approach with 16S rRNA gene sequences. Samples were collected in two seasonal sampling campaigns (summer and winter of 2017) from the most contaminated sites from where the water flows downstream to the freshwater reservoir of the river Chança. The physicochemical data indicate a trend of decreasing acid mine drainage contamination downstream of the mining area to the Chança's reservoir. The most contaminated sites (pH = 2.3–3.1) are distinguished by prokaryotic diversity with high abundances of operational taxonomics units related to acidophiles (genera *Metallibacterium*, *Acidibacter*, *Leptospirillum*, *Acidobacterium*, *Thiomonas*, *Acidicapsa*, *Acidocella*, *Acidiphilium*; family *Acidobacteriaceae*, order CPla-3 termite group). Likewise, in the transition zone in the mouth of the contaminated water flow into the Chança's reservoir (pH = 6.4), a specific prokaryotic flora exists with some acidophiles, but notably with a cyanobacteria bloom and a high abundance of the genus *Sediminibacterium* (family I; order Subsection III). Moreover, the strong correlation between the abundance of acidophiles and characteristic physiochemical parameters (metals, acidity, and sulfate) confirm their potential as biomarkers of acid mine drainage pollution.

**Keywords** AMD · Heavy metal contamination · Microbiome · Biomarkers

Communicated by H. Atomi.

**Electronic supplementary material** The online version of this article (<https://doi.org/10.1007/s00792-019-01136-1>) contains supplementary material, which is available to authorized users.

✉ Jorge D. Carlier  
jcarlier@ualg.pt

- <sup>1</sup> Natural Resources and Environment Laboratory, Sidi Mohammed Ben Abdellah University, Polydisciplinary Faculty of Taza, Taza, Morocco
- <sup>2</sup> Centre of Marine Sciences (CCMAR), Gambelas Campus, University of the Algarve, 8005-139 Faro, Portugal
- <sup>3</sup> Laboratory of Physiology-Pharmacology and Environmental Health, Faculty of Sciences, Sidi Mohammed Ben Abdellah University, Dhar El Mehraz, Fez, Morocco
- <sup>4</sup> Laboratory of Physiology and Genetics, Faculty of Sciences, University Mohammed Premier, Oujda, Morocco
- <sup>5</sup> Faculty of Sciences and Technologies, Gambelas Campus, University of the Algarve, 8005-139 Faro, Portugal

## Introduction

Located in the Beja district of the Alentejo region, the São Domingos mine is one of the most famous abandoned metalworking mines in Portugal. The site of the mine is part of a vast geographical area with particular geological features called the Iberian Pyrite Belt which extends from Rio Tinto in Spain along the southern region of Portugal, and is considered as one of the largest metallogenetic provinces of volcanogenic massive sulphides (VMS ores) in the world (Alvarenga et al. 2012; Álvarez-Valero et al. 2008). The São Domingos mine was extensively exploited during the Roman and Islamic occupations of the Iberian Peninsula, and between 1857 and 1966 the mine was the largest mine operating in Europe. Massive pyrite was the main mineral ore extracted, but extraction also included ores associated to other metals such as copper (Cu), aluminium (Al), arsenic (As), lead (Pb), antimony (Sb), zinc (Zn) and mercury (Hg) (Tavares et al. 2008). Since mining activity ceased in the 1960s, the mine has been left without any local surveillance

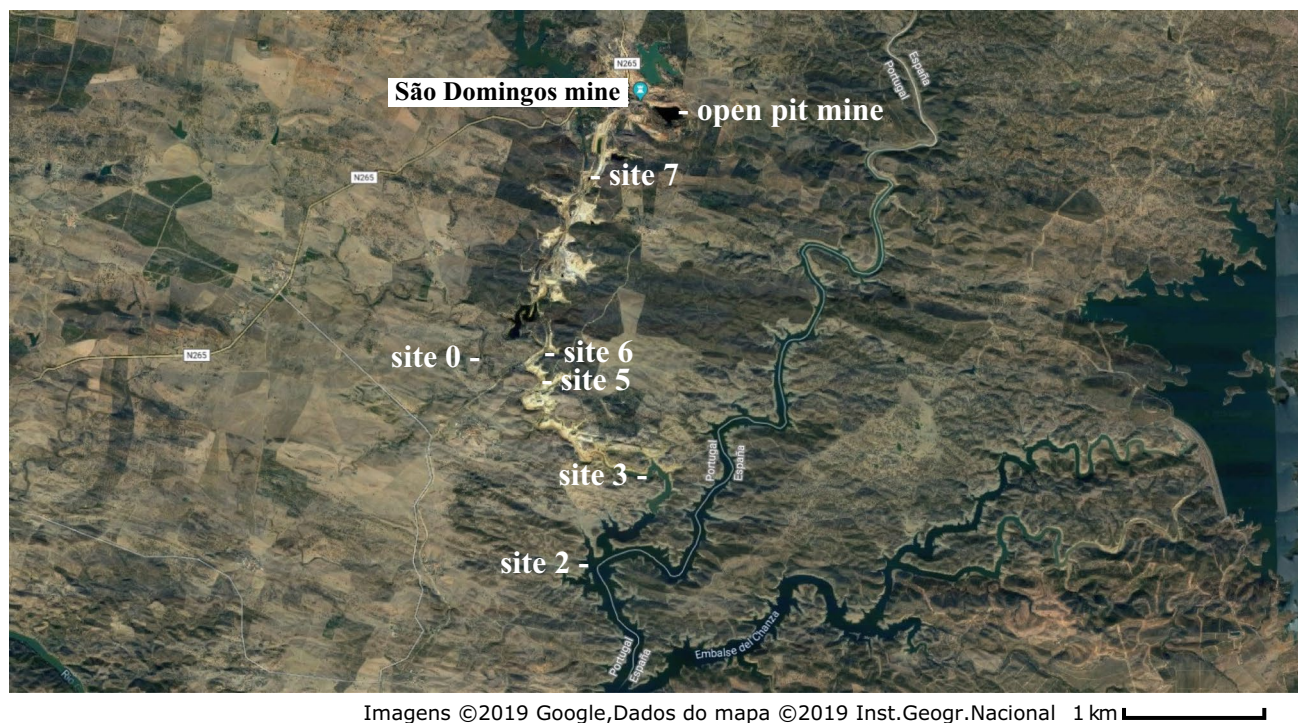
or containment, which has led to a serious environmental deterioration of the region. A Portuguese public company (Empresa de Desenvolvimento Mineiro, S.A.) is responsible for the environmental rehabilitation of the mine, and although this process was initiated in 2005, it has been suspended for a long time and most of the work is still to be executed (Dias-Sardinha et al. 2013). Considerable amounts of mining waste is exposed to air and water producing acid mine drainage (AMD) which is characterized by low pH values, usually close to 2, and high concentrations of sulphate and metals [mainly Al, iron (Fe), Zn and Cu]. Indeed, AMD is one of the most important sources of contamination in abandoned mining areas, and the São Domingos mine significantly affected (Alvarenga et al. 2012).

AMD affects various water bodies within several kilometers of the São Domingos mine, including the São Domingos stream which passes near the open-pit mine lake and flows along the mining area valley, and the dam lakes that are scattered throughout this area (Fig. 1). Water, therefore, promotes the transportation of metals and other mining wastes throughout a large area and likely impacts the microbial diversity of the water and sediments (Pereira et al. 2015). Downstream of the mine site, the São Domingos stream joins the uncontaminated Mosteirão stream at a confluence called “Telheiro” and the merged contaminated water flows to the Chança reservoir before joining the Guadiana River near the Spanish borders close to where the Chança dam

was built by the Spanish government in 1985. On the Portuguese side of São Domingos stream, a system of channels and dams were constructed during recent years which were intended to promote the evaporation the acidic water and thereby control its discharge into the Chança reservoir (Abreu et al. 2008). However, in the rainy season, when the streams are full, the contaminated water still reaches the Chança reservoir, although the uncontaminated Mosteirão stream reduces the concentration of AMD.

Phylogenetic and taxonomic studies of Bacteria and Archaea are typically based on the 16S rRNA gene due to its relatively slow rate of evolution and its ubiquitous presence in prokaryotes (Janda and Abbott 2007; Jovel et al. 2016). A recent approach, known as meta-taxonomics (Marchesi and Ravel 2015), uses high-throughput sequencing of millions of 16S rRNA gene sequences from environmental DNA samples. The sequences are then clustered into groups according to sequence similarities and taxonomically classify as OTUs based on their similarities to existing sequences in databases.

The AMD microbial ecosystem has been studied extensively (e.g. Schrenk et al. 1998; Bond et al. 2000; Bruneel et al. 2006). The typical acidophilic microbial population of AMD is difficult to identify using traditional methods as they require highly specific conditions for cultivation. Indeed, to date, no study has been conducted to investigate the communities of prokaryotes present in the water bodies of the São Domingos mining area. Thus, the main objective of this



**Fig. 1** Google Maps image (images © 2019 Google, Map data © Instituto Geográfico Nacional) showing the São Domingos mine area and the sampling sites

study was to characterize the prokaryotic diversity of the water bodies within the São Domingos mining area by identifying their constituent bacteria and archaea using a meta-taxonomic approach, and to evaluate the degree to which these populations are affected by AMD. The correlation between the most abundant prokaryotic microbiomes of the studied water bodies and the physicochemical characteristics of the sampled waters was analyzed to assess the effects of pollution on prokaryotic diversity. These analyses enabled us to explore the potential value of specific microbial taxa as bioindicators of mine pollution and provide baseline data to establish monitoring strategies for future restoration plans of abandoned mines.

## Material and methods

### Sample collection and processing

Samples of water were collected from six different sites located in the mining area of São Domingos in southern Portugal (Fig. 1). Each site was selected based on its location and distance from the pit lake and surrounding mine-waste piles, which were assumed to be the main source of AMD contamination. Longitude and latitude for each site were defined using a Map 330 GPS. All the samples were collected in sterile acid-washed bottles from the surface of the water body and immediately transported to the laboratory (2 h distance), and processed next day. Approximately, 5 L of water was collected at each site and a volume of 50 mL of sodium azide solution (50 mMol) was added to the samples

to inhibit bacterial growth during the transportation to the laboratory and the overnight storage.

The samples names, the sampling site coordinates, and their brief descriptions are summarized in Table 1. Site 0, located on the Mosteirão stream before it merges with the São Domingos stream was selected as a control sample. Site 2, located at the Chança reservoir, was chosen as it is likely a site where the AMD contamination is very low. Site 5 was selected due to its location in the area called “Telheiro” in the confluence of the contaminated São Domingos stream and the non-contaminated Mosteirão stream, while sites 3 and 6 were selected due to their locations downstream and upstream to the confluence point, respectively. Site 7, a small dam at São Domingos stream located downstream the dam called “Tapadinha”, is highly impacted by AMD.

Sampling was performed during winter and summer seasons, in February and September 2017, respectively, aiming to study the impact of seasonal factors on the chemical characterization and on the bacterial diversity of the water samples. However, the sites S0s, S5s and S6s were dry in summer and, thus, could not be assessed.

### Physicochemical characterization

Water temperature, pH, electro-conductivity (EC), and dissolved oxygen (DO) were determined in situ during sampling using a multiparameter probe (YSI Professional Plus. ProPlus). Sulphate and phosphate concentrations were analyzed in the laboratory by molecular UV/visible spectrophotometry (HACH DR2800), at 520 nm and 880 nm, using the Sulphate Ver Reagent and the Total High Rang

**Table 1** Designation, location and description of the São Domingos waterbodies sampling sites

Site ID	Winter designation	Summer designation	Site name	Latitude/longitude	Brief description of the site
Site 0	S0w	–	<i>Mosteirão stream</i>	37.635849, –7.516451	Mosteirão stream—a small brook (dry in the summer). Not impacted by AMD until it merges with the São Domingos stream, thus used as reference
Site 2	S2w	S2s	<i>Chança reservoir</i>	37.612376, –7.494921	Large reservoir with dammed water from Chanca’s river into which the contaminated water flows. Low AMD contamination
Site 3	S3w	S3s	<i>Upstream Chança reservoir</i>	37.624352, –7.514251	Transition zone between the contaminated stream and Chança’s reservoir. Low to moderate AMD contamination
Site 5	S5w	–	<i>Telheiro</i>	37.636529, –7.513189	Site where the water coming from Site 0 is mixed with the water coming from Site 6 (dry in the summer). A mixture of uncontaminated and AMD contaminated water
Site 6	S6w	–	<i>São Domingos stream</i>	37.635274, –7.513640	A small stream downstream site 7, after merging with several non-contaminated water streams (dry in the summer). Moderate to high AMD contamination
Site 7	S7w	S7s	<i>Tapadinha</i>	37.659461, –7.505471	Medium-sized dammed reservoir containing reddish brown acidic water. Very high AMD contamination

Phosphorus Test, respectively (both from HACH). Fe, Zn, Al, and Cu concentrations were analyzed by microwave plasma atomic emission spectroscopy (MP-AES) with an Agilent 4200 equipment, using viewing positions at 0, read times of 3 s, nebulizer flows of 0.5, 0.7, 0.95 and, 0.45 L/min for Fe, Zn, Al and Cu, respectively, using the following emission wavelengths: Fe—373.486 nm, Zn—213.857 nm, Cu—324.754 nm, Al—396.152 nm, and using calibrations curves with R2 regression values higher than 0.999 build with the following standards: Fe—0.5, 1, 5, 10, 25, 50, 75, 100 mg/L, Zn—0.5, 1, 5, 10, 25, 50, 75, 100 mg/L, Al—0.5, 1, 5, 10, 25, 50, 75, 100 mg/L, Cu—0.5, 1, 5, 10, 25, 50 mg/L. Metal concentrations were recorded as the mean and standard deviations among the analyses of five emission reads (Table 2).

### DNA extraction

A total volume between 1 and 1.75 L of water sample was filtered (until the filter was clogged) using a 0.22 µm filter (Nylon filters, Whatman® GE Healthcare Life Sciences, UK). The filter was cut into small pieces (around 3 mm) and used for DNA extraction using the PowerSoil® DNA Isolation Kit (Mo Bio Laboratories, Carlsbad, CA, USA) as described in the manufacturer's protocol. The concentration and quality of eluted DNA were determined using a spectrophotometer (NanoDrop3300, Thermo Fisher Scientific, USA).

### Library preparation for high-throughput sequencing

Prokaryotic 16S rRNA gene sequence (V4 region) libraries were prepared by a custom protocol based on an Illumina protocol (Illumina 2015). Up to 10 ng of extracted DNA was used as a template for PCR amplification of the V4 region of the 16S rRNA gene. Each PCR (25 µL) contained dNTPs (100 µM of each), MgSO<sub>4</sub> (1.5 mM), Platinum®Taq DNA polymerase HF (1 U), 1X Platinum® High Fidelity buffer (Thermo Fisher Scientific, USA) and tailed primer mix (400 nM of each forward and reverse). PCR was run with the following cycle conditions: initial denaturation at 95 °C for 2 min, 35 cycles of amplification (95 °C for 20 s, 50 °C for 30 s, 72 °C for 60 s) and a final elongation at 72 °C for 5 min. Duplicate PCRs were performed for each sample and the duplicates were pooled after PCR. The forward and reverse tailed primers were designed as described by Illumina's Support Center (2015) and targeted to the bacterial and archaeal V4 region of the 16S rRNA gene (Caporaso et al. 2011): 5'-GTGCCAGCMGCCGCGGTAA (515F) and 5'-GGACTACHVGGGTWTCTAAT (806R). The primer tails enable attachment of Illumina Nextera adaptors for sequencing in a subsequent PCR. The resulting amplicon

**Table 2** Chemical characterization of waterbody samples collected at the São Domingos mining area

	Sample ID	T (°C)	pH	EC (µs/cm)	DO (mg/L)	SO <sub>4</sub> <sup>2-</sup> (mg/L)	PO <sub>4</sub> <sup>3-</sup> (mg/l)	Fe (mg/L) <sup>a</sup>	Zn (mg/L) <sup>a</sup>	Cu (mg/L) <sup>a</sup>	Al (mg/L) <sup>a</sup>
Winter	S0w	16.6	7.59	369	8.04	61	12.64	<LOD	<LOD	<LOD	<LOD
	S2w	14.1	6.65	173	8.73	36	4.85	<LOD	<LOD	<LOD	0.32±0.03
	S3w	14.5	6.40	205	9.35	35	1.22	<LOD	0.26±0.01	0.121±0.002	0.34±0.01
	S5w	16.8	5.63	557	8.55	122	0.17	3.2±0.4	0.78±0.01	0.781±0.005	6.6±0.1
	S6w	17.1	3.06	795	8.39	514	0.28	14.6±0.7	2.98±0.03	2.80±0.03	22.9±0.5
	S7w	17.7	2.25	2883	9.00	3477	0.10	144±1	31.8±0.2	23.15±0.04	202±2
	S0s	Dry	Dry	Dry	Dry	Dry	Dry	Dry	Dry	Dry	Dry
Summer	S2s	26.9	7.65	267	7.50	34	0.07	0.265±0.001	<LOD	<LOD	<LOD
	S3s	28.2	7.50	267	8.10	38	0.04	0.299±0.003	<LOD	<LOD	<LOD
	S5s	Dry	Dry	Dry	Dry	Dry	Dry	Dry	Dry	Dry	Dry
	S6s	Dry	Dry	Dry	Dry	Dry	Dry	Dry	Dry	Dry	Dry
	S7s	32.1	2.26	6750	7.48	6256	0.61	770±4	194.2±0.8	81±1	810±8

<sup>a</sup>Limits of detection (LOD) for metal analysis: copper = 0.034 mg/L; iron = 0.211 mg/L; zinc = 0.122 mg/L; aluminum = 0.098 mg/L

libraries were purified using the standard protocol for Agencourt Ampure XP Bead (Beckman Coulter, USA) with a modified bead to sample ratio of 4:5. The DNA was eluted in 33  $\mu$ L of nuclease-free water (Qiagen, Germany). DNA concentration was measured using Qubit HS DNA Assay kit (Thermo Fisher Scientific, USA).

Sequencing libraries were prepared from the purified amplicon libraries using a second PCR. Each PCR (25  $\mu$ L) contained 1 $\times$  PCR BIO HiFi buffer (PCR Biosystems, UK), PCR BIO HiFi Polymerase (1U) (PCRBiosystems, UK), adaptor mix (400 nM of each forward and reverse) and up to 10 ng of amplicon library template. PCR was run with the following cycle conditions: initial denaturation at 95  $^{\circ}$ C for 2 min, eight cycles of amplification (95  $^{\circ}$ C for 20 s, 55  $^{\circ}$ C for 30 s, 72  $^{\circ}$ C for 60 s) and a final elongation at 72  $^{\circ}$ C for 5 min. The resulting sequencing libraries were purified using the standard protocol for Agencourt Ampure XP Bead (Beckman Coulter, USA) with a modified bead to sample ratio of 4:5. The DNA was eluted in 33  $\mu$ L of nuclease-free water (Qiagen, Germany). DNA concentration was measured using Qubit HS DNA Assay kit (Thermo Fisher Scientific, USA). Gel electrophoresis using TapeStation 2200 and D1000 screentapes (Agilent, USA) was used to check the product size and purity of a subset of sequencing libraries.

### 16S rRNA gene amplicon sequencing

The purified sequencing libraries were pooled in equimolar concentrations and diluted to 2 nM. The samples were sequenced (251–301 bp) on an MiSeq (Illumina, USA) using an MiSeq Reagent kit v3 following the standard guidelines for preparing and loading samples on the MiSeq. The Phix control library was used as an in-run control (20% spiked-in) for run quality monitoring to overcome the issue of low complexity that is often observed with amplicon samples. Sequencing was performed by DNASense Company (Aalborg, Denmark). Amplicon sequencing data reported in this study have been submitted to the Sequences Read Archive SRA database with the accessing number: PRJNA472343.

### Bioinformatics and statistical analysis

Forward reads were trimmed to 225 bp, de-replicated, and formatted for use in the UPARSE workflow (Edgar 2013). The de-replicated reads were clustered, using the “cluster\_otus” command of USEARCH (vers. 7.0.1090; Edgar 2013) with default settings. Sequences were clustered into OTUs at a 97% sequence identity using the “usearch\_global” command with parameter-id 0.97. Taxonomy was assigned using the RDP classifier (vers. 11; Wang et al. 2007) as implemented in the parallel\_assign\_taxonomy\_rdp.py script in QIIME (vers. 1.7.0; Caporaso et al. 2010), using the MiDAS database (vers.2.1.2; McIlroy et al. 2017). The results were

analyzed in R (vers. 1.0.153; R Core Team 2017) through the RStudio IDE using the “ampvis” package (vers. 2.2.6; Albertsen et al. 2015). Shannon and Simpson diversity indices were calculated with raw and normalized OTUs tables, using the R package “RAM” (Chen et al. 2018). To extract the discriminatory information from the data, principal component analysis (PCA) was used to analyze the inter-correlation between the variables describing the given observation and concentrate the information in new orthogonal variables, or the principal components (Abdi and Williams 2010). The plotting of the data was focused on the two first principal components. The “ampvis” package was used to run the PCA with OTU abundances higher than 0.1%, while the MultBiplot package (vers.16.0430; Vicente-Villardón 2015) of MATLAB (MATLAB and Statistics Toolbox Release 2016) was used to run the PCA with the physicochemical parameters, which were normalized to eliminate possible effects of the variability between their scales.

## Results and discussion

### Physicochemical characterization of the water bodies

Physicochemical parameters are the traditional and the most common indicators of water quality. Mining pollution directly affects the quality of a water resource by dissolution and transportation of mining contaminants such as heavy metals and sulphate. The toxicity of these elements depends essentially on their oxidation state and their solubility in the environment. These characteristics are strongly related to other physicochemical parameters in water such as EC, DO and pH. While all metals can conduct electricity, the EC of mine-contaminated water has typically been used as marker of metal pollution in addition to the pH (Salomons 1995).

In this study, the physicochemical characteristics of the sampling sites show a general trend of increasing mining contamination (from sites 2–7), except at the control site (S0w) which is the only sample with all metal concentrations below the limits of detection (LOD). The most contamination was found at sites 6 and 7, the closest sampling sites to the sources of contamination at the pit lake and to the debris from mining activity. Both sites 6 and 7 had high values of EC, high sulphate and metal concentrations, and very low pH values: all characteristics of extremely contaminated environments. The chemical characterization of the samples collected in São Domingos mine during the winter and summer seasons is shown in Table 2.

The highly acidic pH value (2.26) measured at site 7 in summer (S7s) is similar to the values previously reported for AMD samples from São Domingos mine (Martins et al. 2011; Pérez-López et al. 2010), as well as for AMD

collected at the Rio Tinto mine in Spain, also located on the Iberian Pyrite Belt (Hubbard et al. 2009). The highest EC readings also occurred in the most contaminated samples S7w and S7s and are in the range of values previously reported in this mining area (Martins et al. 2011; Pérez-López et al. 2010) and much higher than the values reported for example in Zarza-Perrunal stream (12.99  $\mu\text{S}/\text{cm}$ ), another impacted environment located in the Iberian Pyrite Belt in Spain (González-Toril et al. 2011). As for the pH and EC, the concentration of sulphate is directly related to the contamination degree of the samples and results from the oxidation of sulphide mineral wastes, such as pyrite. Indeed, sulphate concentrations increased considerably between site 2 and site 7. Again, these concentrations of sulphate were in the range of values reported for this mining area, and for example lower than values recorded in a similar Chinese mine ( $3787 \pm 2129$  mg/L) (Kuang et al. 2013) and higher than in the Guryong mine in South Korea (Kim et al. 2009).

DO concentration is a complex parameter to interpret with regard to water contamination as it is known to depend on both chemical and biological factors, namely on the occurrence of microbially-mediated reactions, as well as the movement of the water. The values of DO concentration recorded at the sites were similar and show that all samples were collected from well aerated sites, as was to be expected from surface waters. Thus, in this case, the contamination seems not to affect the aerobic conditions in the superficial water bodies contaminated with AMD. The DO values reported in this study reveal that the sampled sites were more oxygenated than superficial water samples taken from similar contaminated waters (Kim et al. 2009; González-Toril et al. 2011; Kuang et al. 2013).

The concentrations of the metals known to be the most abundant in the AMD of São Domingos mine (Fe, Zn, Al, and Cu) were measured in the collected samples. Samples from site 7 (S7w and S7s) had much higher concentrations of metals than the other samples. These high concentrations at site 7 are because it is the closest site to the pit lake and debris from mining activity and, therefore, leaching of metal sulphide waste is much more pronounced, leading to highly acidic pH and increased solubility of sulphate and metals. As the water flows downstream from the source of contamination, other uncontaminated streams converge with it diluting the contaminants and promoting metal precipitation due to pH neutralization.

The highest phosphate concentration was reported in samples with the least metal contamination, S0w, followed by S2w, and S3w. All the other samples showed very low phosphate concentrations. This is largely explained by the high absorption of phosphate to metal surfaces which promote its precipitation (Mekki and Sayadi 2017).

Seasonal variations in the physicochemical characteristics of water bodies can influence the microbial density

and diversity of the aquatic ecosystems (Gilbert et al. 2009; Clarke et al. 1996). It should be emphasised that in the present study the investigation of the effect of seasonal variations on the physicochemical characteristics was limited to samples from sites 2, 3 and 7 collected in winter (February 2017) and summer (September 2017), because sites 0, 5 and 6 were dry in the summer. The São Domingos mine region has a dry and sunny climate in the summer, with an average temperature of 32.5 °C during the summer of 2017 and a cold climate with precipitation in the winter, with 5.4 °C average temperature in the winter of 2017 and an average precipitation between 2.54 and 83.82 mm in the same year. Water samples collected in September after a prolonged dry summer showed considerably higher values of temperature in comparison to those collected in February (Table 2), with differences of 12.8, 13.7 and 14.4 °C, for sites 2, 3 and 7, respectively. Major differences between summer and winter samples were observed in sites 7 and 3 with respect to EC, sulphate and metal concentrations. Hence, the variations between samples S7w and S7s are likely due to changes from a wet to a dry climate. Metals and sulphate are more concentrated in the summer as expected, due to strong evaporation of water in the reservoir corresponding to site 7 and to the deficiency of rain in the months that preceded sampling. June, July and August were the warmest months of year 2017, with monthly averages for the maximum temperatures of 32, 34 and 34 °C, respectively, and total rain precipitation below 5 mm in the São Domingos mine region (<https://www.ipma.pt/pt/oclima/monitorizacao/>). On the other hand, because the stream passing through sites 5 and 6 was dry during the summer, there was no transportation of contaminants to the downstream sites 2 and 3 in that season. This may explain why Zn, Cu and Al concentrations in the sample collected at site S3 in winter (S3w) are above the respective LODs, while in the sample collected at the same site in summer (S3s) their concentrations were below the LODs.

## Prokaryotic richness and diversity of water bodies

In this study, the measurement of the diversity of prokaryotic communities is based on relative abundances of 16S rRNA gene sequences. These are not absolute abundances of each microbe in the community but have been widely used as an indirect measure in such studies.

## Sequencing results

A total of 297,826 single-end Illumina sequence reads were generated from partial 16S rRNA gene sequences in the nine water samples from the São Domingos mine. The length of the sequence reads was 225 bp after quality trimming. The number of reads retrieved from each sample as well as their

number of OTUs is listed in Table ESM\_1.1 on the Online Resource 1.

### Prokaryotic community richness and evenness

The diversity of bacteria and archaea in each sample was studied using Shannon and Simpson diversity indices. Both are calculated using the number of OTUs and their respective abundances, thereby taking into consideration the richness and evenness of OTUs within communities. However, while the Shannon index is more greatly influenced by species richness and by rare species, the Simpson index gives more weight to evenness and common species (Shannon 1948; Simpson 1949; Spellerberg and Fedor 2003; <https://www.davidzeleny.net/anadat-r/doku.php/en:div-ind>). Moreover, since normalization of the OTUs table (equalising the number of reads in each sample) can distort alpha diversity comparisons in some datasets, the indices were calculated using both the raw OTUs tables and normalized OTUs tables.

The Shannon and Simpson indices, either calculated using raw or normalized OTUs tables, revealed similar results (Table ESM\_1.2 on the Online Resource 1). In winter, when AMD was flowing all the way down to the Chança's reservoir, the prokaryotic diversity indices were much lower in the site closest to the São Domingos mining area (site S7w) than inside the large downstream reservoir (site S2w), while in the intermediate zones (sites S3w, S5w and S6w) the indices indicated intermediate diversities. This indicates a decrease of the prokaryotic richness and evenness is correlated with an increase of AMD contamination. Interestingly, the diversity indices in the reference water stream not affected by AMD (site S0w) also have intermediate values compared to sites S7w and S2w. This may be due to different causes, but the results reported here are not sufficient to provide an explanation for this observation.

Observations of seasonal variations in the study parameters are limited to sites 2, 3, and 7. Nevertheless, it is possible to see that from winter to summer the diversity at site S3 increased, becoming more similar to the diversity observed at site S2, and the diversity at site S7 fell during the summer when compared to winter. These differences are related to changes in physicochemical parameters of these sites. As noted above, the stream that carries the contaminated water from the mine area to the Chança's reservoir was dry at the time of the summer sampling, therefore, in summer site S3 no longer acted as a transition zone, being instead a sample similar to that taken from the Chança's reservoir. On the other hand, at site 7 water evaporation in the dry season concentrated metals and sulphate, creating conditions that few microorganisms could withstand, thereby reducing species richness at this site in the summer. Previous investigations have also documented similar seasonal effects

on the microbial diversity in AMD impacted environments (González-Toril et al. 2011; Edwards et al. 1999).

### Prokaryotic community composition

#### Phylum level

The prokaryotic community structure was determined at the phylum level for all the samples according to the classification of the OTUs identified by the V4 region of the 16S rRNA gene sequences. 97% and 2% of the total reads corresponded to bacterial and archaeal sequences. 1% of the reads were identified as 16S rRNA gene sequences but were assigned “unclassified taxa” because of insignificant similarity values to those in public database. The graphs in Online Resource 2 show the relative abundances of dominant phyla ( $\geq 1\%$  of the dataset) at the different sites sampled in winter and summer, with the label “others” representing all taxa that scored a relative abundance below 1%.

Microbial communities were in general dominated by taxa from the phylum *Proteobacteria* (minimum: 36% in S2w; maximum: 62% in S7s), with the next most abundant phyla being the *Bacteroidetes*, *Actinobacteria*, *Acidobacteria*, *Cyanobacteria*, or unclassified phyla, depending on the level of AMD contamination. These results are consistent with other studies that reported the dominance of these phyla in comparable habitats (Kim et al. 2009; Okabayashi et al. 2005).

At the phylum level, besides *Proteobacteria* with relative abundances above 35% in all samples, the other phyla reveal clear distinctions between groups of samples. The most pronounced difference between samples is the relative abundance of phylum *Bacteroidetes*, which is lower than 1% in samples S6w, S7w and Ss7, with more acidic pHs (between 2 and 3), and between 18 and 39% in all other samples where the pH values are closer to 7 (between 5 and 8). Furthermore, four types of samples can be distinguished:

- (i) the reference sample of the water stream not affected by AMD (S0w) and the sample taken at the confluence between this stream and the contaminated water course from the mine zone (S5w), both collected in the winter, have high abundances of *Bacteroidetes* (36% and 39%) and of *Actinobacteria* (both 14%) and abundances between 1 and 5% for *Verrucomicrobia* and for “others”;
- (ii) the samples collected in winter and summer at the transition site where the contaminated water stream drains into the Chança's reservoir (S3w and S3s) and those collected in this reservoir but at a location further away from this mouth (S2w and S2s) have high abundances of *Bacteroidetes* (minimum: 18%, maximum 26%) and *Actinobacteria* (minimum:

6%, maximum 31%), relatively high abundances of *Cyanobacteria* (minimum: 2%, maximum 15%) and *Verrucomicrobia* (minimum 2%, maximum 4%) and abundances of 1–2% for the “others”—in addition, there are subgroups of 2 or 3 of these samples that have abundances between 1 and 4% of *Planctomycetes*, *Chloroflexi*, *Acidobacteria*, *Armatimonadetes* and of unclassified taxa;

- (iii) the samples with high contamination collected in the winter at the two locations successively closer to the mine (S6w and S7w) have both high abundances of *Acidobacteria* (21% and 19%) and *Actinobacteria* (7% in both) and relatively high abundances of *Planctomycetes*, *Nitrospirae*, “others” and unclassified taxa (2–7%);
- (iv) the sample with most extreme contamination, collected in the summer at the closest site to the mine area (S7s), is distinguished by having a high abundance of unclassified taxa (28%) and by having in addition to *Proteobacteria* only one other phylum with more than 1% abundance: *Planctomycetes* (7%)—the sum of “others” in this sample is 3%.

## Lower taxonomic levels

OTUs were assigned to lower taxonomic classification levels to conduct a more detailed analysis of the composition of the prokaryotic communities and further understand the differences in dominant taxa between the samples. The 30 most abundant prokaryotes among all sites sampled in winter and in summer were classified separately and compared (Tables 3 and 4).

In winter, there were nine prokaryotic taxa (first nine lines of Table 3) with abundances varying between 1.3% and 21.7% in the most contaminated samples (S6w and S7w), which had abundances of just 0.2% or less in the non- or less contaminated samples (S0w, S2w, S3w and S5w), while on the contrary, there were nineteen taxa (last 19 lines of Table 3) with varying abundances in these non- or less contaminated samples, yet which were not detected in both the highly contaminated samples S6w and S7w. Moreover, two taxa (lines 10 and 11 of Table 3) had relatively high abundances (> 5%) in one or both the highly contaminated samples S6w and S7w as well as in at least one of the two samples from transition zones: at the mouth of the contaminated stream with the Chança’s reservoir (S3w) and at the confluence of the uncontaminated stream

**Table 3** Relative abundances (%) of the 30 most abundant prokaryotes among all water samples collected at São Domingos mining area during winter

Classification	S0w	S2w	S3w	S5w	S6w	S7w
<i>Proteobacteria; Metallibacterium</i>	0.0	0.0	0.0	0.0	3.3	21.7
<i>Proteobacteria; Acidibacter</i>	0.0	0.1	0.1	0.1	4.5	13.5
<i>Acidobacteria; f_Acidobacteriaceae (Subgroup 1)_OTU_16</i>	0.0	0.0	0.0	0.0	1.6	7.4
<i>Nitrospirae; Leptospirillum</i>	0.0	0.0	0.0	0.1	2.5	6.2
<i>Acidobacteria; Acidobacterium</i>	0.0	0.0	0.0	0.0	4.3	2.6
<i>Proteobacteria; Thiomonas</i>	0.0	0.0	0.0	0.1	3.3	3.4
<i>Acidobacteria; Acidicapsa</i>	0.0	0.0	0.0	0.0	2.4	3.1
<i>Planctomycetes; o_CPla-3 termite group_OTU_43</i>	0.0	0.0	0.0	0.0	1.3	3.2
<i>Proteobacteria; Acidocella</i>	0.0	0.1	0.0	0.2	2.8	1.5
<i>Proteobacteria; Acidiphilium</i>	0.0	0.0	0.1	7.0	12.2	5.6
<i>Proteobacteria; c_Betaproteobacteria_OTU_3</i>	0.0	0.6	7.0	3.0	5.9	0.0
<i>Bacteroidetes; Flavobacterium</i>	20.5	4.7	2.1	15.7	0.0	0.0
<i>Bacteroidetes; Sediminibacterium</i>	2.9	7.8	14.4	2.8	0.0	0.0
<i>Bacteroidetes; Pseudarcicella</i>	7.8	1.8	3.3	13.3	0.0	0.0
<i>Proteobacteria; Limnohabitans</i>	12.8	2.4	1.0	4.1	0.0	0.0
<i>Cyanobacteria; f_FamilyI_OTU_8</i>	0.0	2.8	14.1	0.0	0.0	0.0
<i>Proteobacteria; f_Comamonadaceae_OTU_7</i>	0.8	2.3	12.1	0.2	0.0	0.0
<i>Proteobacteria; Polynucleobacter</i>	5.3	1.4	2.8	6.5	0.0	0.0
<i>Actinobacteria; f_Sporichthyaceae_OTU_15</i>	0.5	8.9	2.0	0.8	0.0	0.0
<i>Actinobacteria; hgcI clade</i>	4.0	2.7	0.7	3.2	0.0	0.0
<i>Actinobacteria; f_Sporichthyaceae_OTU_24</i>	1.2	6.6	1.2	0.7	0.0	0.0
<i>Proteobacteria; f_Comamonadaceae_OTU_177</i>	5.0	1.7	0.7	2.7	0.0	0.0
<i>Actinobacteria; f_Sporichthyaceae_OTU_35</i>	4.7	1.5	0.5	3.1	0.0	0.0
<i>Bacteroidetes; Fluviicola</i>	5.6	1.2	0.2	2.7	0.0	0.0
<i>Proteobacteria; Methylothera</i>	4.2	0.2	0.1	2.4	0.0	0.0
<i>Proteobacteria; LD28 freshwater group</i>	0.3	4.4	1.5	0.1	0.0	0.0
<i>Proteobacteria; f_Comamonadaceae_OTU_25</i>	4.1	0.1	0.1	2.2	0.0	0.0
<i>Proteobacteria; f_Comamonadaceae_OTU_91</i>	2.4	1.4	0.9	0.9	0.0	0.0
<i>Proteobacteria; f_Comamonadaceae_OTU_104</i>	4.4	0.3	0.3	0.7	0.0	0.0
<i>Actinobacteria; f_Sporichthyaceae_OTU_46</i>	2.6	0.5	0.2	1.7	0.0	0.0

The names are shown based on the phylum and the genus assignment or the phylum and the lowest possible taxonomic assignment (class (c\_\_), order (o\_\_) or family (f\_\_)) together with the OTU number



**Table 4** Relative abundances (%) of the 30 most abundant prokaryotes among all water samples collected at São Domingos mining area during summer

Classification	S2s	S3s	S7s
Proteobacteria; <i>Acidiphilium</i>	0.0	0.0	51.6
Unclassified_OTU_12	0.0	0.0	27.5
Planctomycetes; o__CPla-3 termite group_OTU_38	0.0	0.0	6.2
Proteobacteria; <i>Acidibacter</i>	0.9	0.8	7.7
Actinobacteria; hgcI clade	27.6	20.1	0.0
Proteobacteria; f__LD12 freshwater group_OTU_11	10.8	6.0	0.0
Proteobacteria; <i>Polynucleobacter</i>	3.6	9.1	0.0
Bacteroidetes; <i>Candidatus Aquirestis</i>	2.5	4.6	0.0
Proteobacteria; f__Comamonadaceae_OTU_22	3.8	3.0	0.0
Proteobacteria; LD28 freshwater group	2.4	2.1	0.0
Proteobacteria; f__Comamonadaceae_OTU_42	1.4	3.0	0.0
Proteobacteria; f__Burkholderiaceae_OTU_50	1.8	2.4	0.0
Proteobacteria; f__Comamonadaceae_OTU_66	1.9	2.2	0.0
Bacteroidetes; f__NS11-12 marine group_OTU_27	1.7	2.1	0.0
Bacteroidetes; f__NS11-12 marine group_OTU_62	1.3	2.4	0.0
Bacteroidetes; f__NS11-12 marine group_OTU_40	1.5	1.8	0.0
Bacteroidetes; <i>Flavisolibacter</i>	1.6	1.6	0.0
Bacteroidetes; <i>Fluviicola</i>	1.6	1.5	0.0
Actinobacteria; f__Sporichthyaceae_OTU_80	1.3	1.6	0.0
Proteobacteria; MWH-UniP1 aquatic group	1.3	1.5	0.0
Proteobacteria; f__Oxalobacteraceae_OTU_82	1.7	1.0	0.0
Bacteroidetes; <i>Sediminibacterium</i>	1.8	0.4	0.0
Bacteroidetes; f__NS11-12 marine group_OTU_1269	0.8	0.9	0.0
Armatimonadetes; <i>Armatimonas</i>	0.8	0.9	0.0
Chloroflexi; f__Roseiflexaceae_OTU_77	0.5	1.3	0.0
Actinobacteria; <i>Candidatus Planktophila</i>	1.0	0.7	0.0
Proteobacteria; f__Comamonadaceae_OTU_468	0.9	0.7	0.0
Actinobacteria; f__Sporichthyaceae_OTU_360	0.5	1.0	0.0
Verrucomicrobia; c__OPB35 soil group_OTU_93	0.6	0.8	0.0
Cyanobacteria; <i>Synechococcus</i>	0.7	0.7	0.0

The names are shown based on the phylum and the genus assignment or the phylum and the lowest possible taxonomic assignment [class (c\_\_), order (o\_\_) or family (f\_\_)] together with the OTU number

with the contaminated stream (S5w). Indeed, sample S3w had a unique distribution of relative abundances among the most abundant taxa, and had a characteristic high percentage of Cyanobacteria (14.1%)—which is in accordance with the blue-green colour of the water at that site—and a high percentage of genus *Sediminibacterium*. By contrast, sample S5w, though being slightly more acidic and with higher sulphate and metal contamination than sample S3w, had a similar composition of prokaryotes (except for the already mentioned two taxa in lines 9 and 10 of Table 3) to the non-contaminated reference sample S0w, both included high percentages of the genera *Flavobacterium*, *Pseudarcicella*, *Limnohabitans* and *Polynucleobacter*. This can be explained by a very weak water flow in the transition zone in site S3 at the entry of the Chança's reservoir, thus creating conditions for the establishment of new equilibria in the microbial population, while by contrast, the two streams entering in the confluence zone at site S5 had relatively strong flows. Therefore, the microbial population at S3 was the result of an equal mixing of the microbial populations from the two streams from sites S0 and S6. However, the distribution of microbial populations seems

to indicate that sample S5w is mostly composed by water coming from site S0, and this observation is corroborated by its pH, sulphate and metal concentrations (Tables 2, 3).

During summer, the same pattern was observed among the 30 most abundant prokaryotic taxa. Three taxa (first 3 lines of Table 4) were representative of the extremely contaminated sample collected in the mine area (S7s) and 26 (last 26 lines of Table 4) were characteristic of both samples collected in Chança's reservoir (S2s and S3s) uncontaminated with AMD in summer due to prolonged drought. In addition, one taxa (line 4 of Table 4) was common to all the three samples, though with low abundances (<1%) in samples S2s and S3s and a relatively high abundance (7.7%) in sample S7s. Three of the four taxa found to be distinctive for such extreme conditions in summer were also among the nine taxa characteristic of the most contaminated samples collected in winter and one remained unclassified.

These three taxa present in samples S6w, S7w and S7s (genera *Acidiphilium* and *Acidibacter* and the OTU\_38 from order “CPla-3 termite group” but not classified for the family and genus levels) have already been correlated with extremely acidic and highly metallic environments (Barns et al. 1999). Due to their acidophilic character, they can survive in extreme acidity, with pHs of 1.5–2.5 (Kishimoto et al. 1991). They have been isolated in a variety of environments including soils, oceans, metal-contaminated waters, and acidic biofilms and were previously reported in AMD samples from the Rio Tinto mine in Spain (González-Toril et al. 2010), and the Nanshan deposits on Mountain Xiang in China (Hao et al. 2007). The other seven taxa which were among the most abundant in samples S6w and S7w, but not in S7s (genera *Metalibacterium*, *Leptospirillum*, *Acidobacterium*, *Thiomonas*, *Acidicapsa*, *Acidocella* and the OTU\_16 from family *Acidobacteriaceae* (Subgroup 1) but unclassified at the generic level) were also previously reported in AMD samples from several mining areas (e.g. García-Moyano et al. 2015; Ziegler et al. 2013; Zhang et al. 2019). The results obtained in this study, together with reports from other authors, reinforce the possibility of using these microorganisms as bioindicators of AMD pollution.

Relatively high percentages of Cyanobacteria in transition zones (as observed in this study) have also been reported in other mining areas (e.g. Podda et al. 2000; Zhang et al. 2019). Indeed, the potential use of Cyanobacteria strains, isolated from booms occurring in such transition zones, for the polishing of heavy metals in mine waste waters has already been demonstrated (e.g. Podda et al. 2000), thereby highlighting the interest in further studies focusing on isolating and describing the Cyanobacteria strain blooming at site 3 in winter.

## Winter and summer differences

Due to drought, sites 0, 5 and 6 were not sampled in summer and, therefore, differences in prokaryotic diversity between winter and summer were based on data only from sites 2, 3 and 7. Although similar temperature raises were registered at these three sites, their profiles in terms of phyla with the highest relative abundances were variable: the greater the variation in levels of AMD contamination (between the seasons, more pronounced were the changes in the prokaryotic communities.

At site 2, the major physicochemical differences from winter to summer were the decay of Al and the rise in pH from 6.65 to 7.65 (Table 2). These changes, caused by the interruption of AMD arrival to the Chança reservoir due to the drought, were enough to cause a reduction of *Acidobacteria* from 2% in winter to less than 1% in summer (Online Resource 2). At lower taxonomic levels, the major differences at site 2 were the reduction of representatives from the family *Sporichthyaceae* from 17.6% in winter to 2.6% in summer and the rise of representatives from the genus “hgcI clade” from 2.7 to 27.6% (Tables 3, 4). This also reflects the flow of AMD in to the Chança reservoir in winter but not in summer. It is known that the family *Sporichthyaceae* includes species adapted to moderately acidic environments, such as compost and human skin (Normand 2006; Lee et al. 2018) while the genus “hgcI clade” is usually among the most abundant prokaryotes in freshwater reservoirs (Llirós et al. 2014; Keshri et al. 2018). At site 3, the changes were similar to those at site 2, but with more pronounced changes of concentrations of Zn and Cu (Table 2). In this case, the major changes at the phylum level were the reduction of *Cyanobacteria* (from 15 to 2%) and the increase of *Actinobacteria* (from 6 to 25%). At a lower axonomic level, the major changes at site 3 were the reduction of genus *Sediminibacterium* (from 14.4% to 0.4%), the reduction of a *Cyanobacteria* OTU classified in Subsection III, Family I (from 14.1 to <0.1%) and the increase of genus “hgcI clade” (from 0.7 to 20.1%). These changes also reflect the discharge of AMD into the Chança reservoir in the winter and the interruption of this discharge in the summer. A relatively high abundancy of *Sediminibacterium* bacteria was previously observed in water sampled from a Canadian lake affected by AMD with a similar level of pollution as at site 3 in winter (Laplante et al. 2013). As noted above, blooms of *Cyanobacteria* in such transition zones may occur; however, the genus “hgcI clade” is typically highly represented in samples from freshwater reservoirs not affected by AMD. At site 7, the pH was extremely acidic in both seasons ( $\approx 2.3$ ) and there was a large rise in the concentrations of sulphate and metals from winter to summer (Table 2). This might have led to the drastic decrease in the relative abundances of *Acidobacteria*, *Actinobacteria* and *Nitrospirae* from 19%, 7% and 6%,

respectively, in winter to less than 1% in summer, and a large rise of unclassified taxa from 2% in winter to 28% in summer. Interestingly, the relative abundance of *Planctomycetes* was not affected by the physicochemical changes at this site (6% in winter and 7% in summer). *Planctomycetes* is part of the PVC superphylum, a grouping of distinct phyla of bacteria which contains organisms adapted to sharply contrasting habitats. Indeed, bacteria from this phylum have been previously reported in similar metal-rich environments (e.g. Kuang et al. 2013; Reis et al. 2016).

## Correlations between physicochemical characteristics and prokaryotic diversity

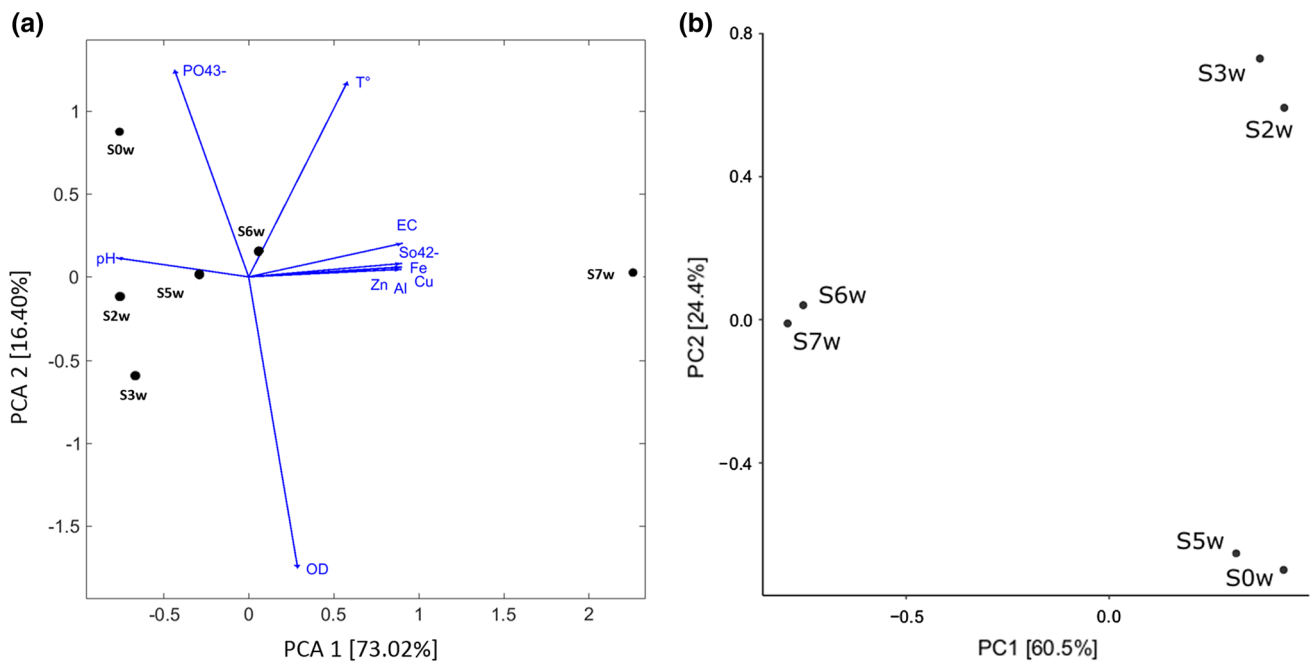
The limited number of samples (three) collected in the summer season due to drought led to the decision to conduct this study with only the six winter samples.

## Principal component analysis

To identify groups of samples with similar physicochemical profiles and groups of samples with similar prokaryotic communities, independent PCAs were performed for these two types of data.

The PCA of the physicochemical parameters reveals that the first two components allow the representation of 89.42% of data variability. The first component (PCA1) of the physicochemical data accounts for 73.02% of the variability in the data and confirms a strong positive correlation between EC and sulphate ( $\text{SO}_4^{2-}$ ) and metals (Fe, Zn, Al, and Cu) concentrations, all with a negative correlation to pH. These parameters clearly discriminate the sample from the site most affected by AMD (S7w) from the other samples; all positioned in the graph along the  $x$ -axis according to the level of AMD pollution (Fig. 2a). The second component (PC2) represents just 16.40% of data variability and shows a negative correlation between the phosphate concentration and DO, according to which the sample S0w is distinguished from the others. In addition, the analysis shows that the temperature has little influence on the physicochemical discrimination of samples.

A PCA graph displaying the coefficient vectors for all OTUs with abundances not lower than 0.1% would be illegible; therefore, only the points representing the distribution of samples according to the level of similarity regarding their prokaryotic communities are shown (Fig. 2b). The PCA shows that the first two components account for 84.9% of data variability and clearly separates the samples in three groups: the samples from the two sites most affected by AMD (S6w and S7w) cluster together in one group; the sample from the Chança’s reservoir (S2w) together with the sample from the transition zone of the AMD flow to this reservoir (S3w) form a second group, and the sample from



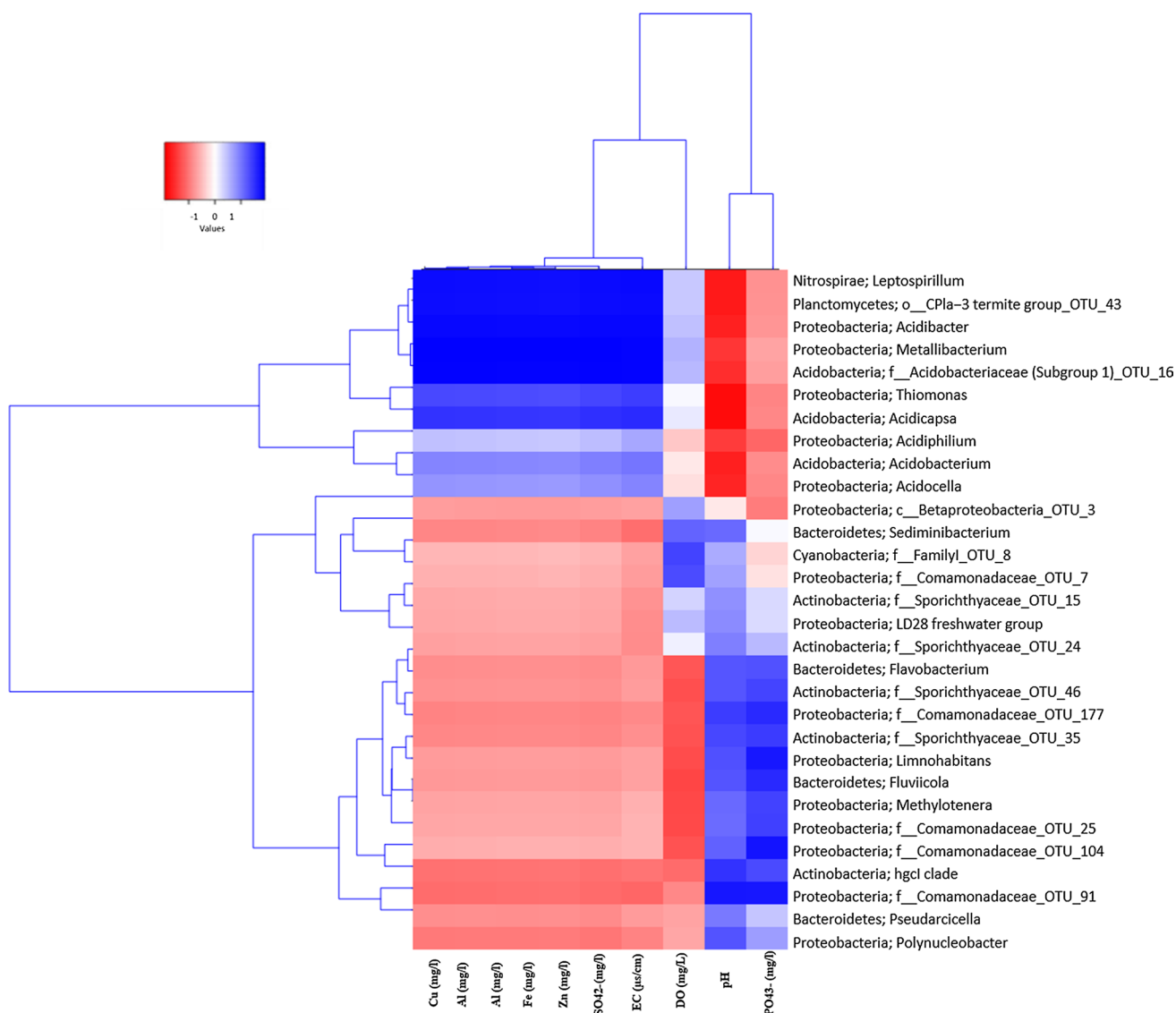
**Fig. 2** **a** Principal component analysis of normalized physicochemical parameter data. **b** Principal component analysis of taxonomic profiles with prokaryotic OTU abundances higher than 0.1%, in samples collected at São Domingos mining area in winter

the water stream not affected with AMD (S0w) together with the sample from the confluence zone of this stream with the water flow affected by AMD (S5w) form a third distinct cluster. Except for sample S5w, this grouping of samples coincides with the distribution resulting from the PCA with the physicochemical data, confirming the strong influence of AMD pollution on the prokaryotic communities in the water bodies. The exceptional grouping of sample S5w (moderately contaminated with AMD according to the physicochemical parameters) with the reference sample S0w (not affected by AMD) was already evident in the analysis of the 30 most abundant prokaryotes and was discussed above in “Prokaryotic community composition” and “Lower taxonomic levels”.

### Correlations between abundance and physicochemical characteristics

A Pearson correlation matrix was calculated between the relative abundances of the 30 most abundant taxa and the physicochemical parameters. The results are displayed in a combined figure showing the clustering dendrograms and the  $r$  values of the correlation matrix illustrated as a heat map (Fig. 3). The dendrogram obtained for the physicochemical parameters has three major clusters: cluster 1—gathers the sulphate ( $\text{SO}_4^{2-}$ ) and the metals (Cu, Al, Fe, Zn) concentrations together with the EC; cluster 2—has the DO concentration alone; cluster 3—has the pH

grouped with the concentration of phosphate ( $\text{PO}_4^{3-}$ ). The dendrogram of the most abundant taxa also reveals three major clusters: cluster 1—the first ten taxa of Fig. 3, which according to the heat map have a strong positive correlation with cluster 1 of the analysis of physicochemical parameters and a strong negative correlation with cluster 3 of the physicochemical parameters, especially pH, but without or with weak correlations with cluster 2, indicating that the abundances of these taxa are strongly correlated with the level of AMD pollution but not with the observed low variation of DO; cluster 2—the next seven taxa of Fig. 3, which according to the heat map have a moderate negative correlation with cluster 1 of the analysis of physicochemical parameters of and no, or weak correlations (neither positive nor negative) with the physicochemical parameters of both clusters 2 and 3; cluster 3—the remaining 13 taxa, which according to the heat map have moderate negative correlations with the clusters 1 and 2 of the analysis of physicochemical parameters, and moderate to high correlations with the parameters of cluster 3. The ten taxa of cluster 1 are the same taxa in the first ten lines of Table 3, indicated in the “Prokaryotic community composition” as being distinctive biomarkers for water bodies contaminated with AMD. The findings presented in Fig. 3 imply that the relative abundances of these taxa are directly correlated to the degree of AMD contamination. Thus, the evidence suggest that these taxa can be used as quantitative biomarkers for mine-drainage pollution.



**Fig. 3** Heat map and dendrograms showing correlations between physicochemical parameters and the 30 most abundant prokaryotes in samples collected at São Domingos mining area in winter. Colours

indicate *R* values correlations between physicochemical parameters (columns) and bacterial taxa (rows)

## Conclusions

As anticipated the physicochemical parameters analyzed in this study confirm a trend of decreasing AMD contamination with increasing distance from the São Domingos mining area, with the least contaminated site being the freshwater reservoir of the Chança river, to where the AMD flows.

The species richness of the prokaryotic community of the water bodies affected by AMD coming from the São Domingos mining area increases as the acidity and metals and sulphate concentrations become more diluted as they flow to the Chança's reservoir. Indeed, prokaryotic diversity distinguishes the sites highly contaminated with AMD (pH=2.3–3.1), with high abundances of acidophiles adapted

to these conditions (genera *Metallibacterium*, *Acidibacter*, *Leptospirillum*, *Acidobacterium*, *Thiomonas*, *Acidicapsa*, *Acidocella*, *Acidiphilium*; family *Acidobacteriaceae* (Subgroup 1); order “CPla-3 termite group”), and also to distinguish the transition zone (pH=6.4) at the mouth of the contaminated water flow into the freshwater reservoir, where samples still contain some acidophiles, but have especially high abundances of *Cyanobacteria* and genus *Sediminibacterium*. In conclusion, there is a demonstrable correlation between the level of AMD contamination (acidity and metals and sulphate concentrations) and the diversity of prokaryotic communities and indicates that the presence of the acidophiles identified in this study are good quantitative biomarkers for this type of pollution.

**Acknowledgements** Sampling and samples preparation received Portuguese national funds from the Foundation for Science and Technology (FCT) through project UID/Multi/04326/2019. The bioinformatics and interpretation of the results were framed in the project METALCHEM-BIO (no. 29251) financed by national funds through the FCT and co-financed by the Algarve's Regional Operational Program (CRESC Algarve 2020), through Portugal 2020 and European Regional Development Fund (FEDER).

## References

- Abdi H, Williams LJ (2010) Principal component analysis. *WIREs Comp. Stat* 2:433–459. <https://doi.org/10.1002/wics.101>
- Abreu MM, Tavares MT, Batista MJ (2008) Potential use of *Erica andevalensis* and *Erica australis* in phytoremediation of sulphide mine environments: São Domingos, Portugal. *J Geochem Explor* 96:210–222. <https://doi.org/10.1016/j.gexplo.2007.04.007>
- Albertsen M, Karst SM, Ziegler AS, Kirkegaard RH, Nielsen PH (2015) Back to basics—the Influence of DNA extraction and primer choice on phylogenetic analysis of activated sludge communities. *PLoS ONE* 10(7):e0132783. <https://doi.org/10.1371/journal.pone.0132783>
- Alvarenga P, Palma P, de Varennes A, Cunha-queda AC (2012) A contribution towards the risk assessment of soils from the São Domingos Mine (Portugal): chemical, microbial and ecotoxicological indicators. *Environ Pollut* 161:50–56. <https://doi.org/10.1016/j.envpol.2011.09.044>
- Álvarez-Valero AM, Pérez-López R, Matos J, Capitán MA, Nieto JM, Sáez R, Delgado J, Caraballo M (2008) Potential environmental impact at São Domingos mining district (Iberian Pyrite Belt, SW Iberian Peninsula): evidence from a chemical and mineralogical characterization. *Environ Geol* 55:1797–1809. <https://doi.org/10.1007/s00254-007-1131-x>
- Barns SM, Takala SL, Kuske CR (1999) Wide distribution and diversity of members of the bacterial kingdom acidobacterium in the environment. *Appl Environ Microbiol* 65:1731–1737
- Bond PL, Smriga SP, Banfield JF (2000) Phylogeny of microorganisms populating a thick, subaerial, predominantly lithotrophic biofilm at an extreme acid mine drainage site. *Appl Environ Microbiol* 66:3842–3849. <https://doi.org/10.1128/AEM.66.9.3842-3849.2000>
- Bruneel O, Duran R, Casiot C, Elbaz-Poulichet F, Personné J-C (2006) Diversity of microorganisms in Fe–As-rich acid mine drainage waters of Carnoulès, France. *Appl Environ Microbiol* 72:551–556. <https://doi.org/10.1128/AEM.72.1.551-556.2006>
- Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, Fierer N, Peña AG et al (2010) QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* 7:335–336. <https://doi.org/10.1038/nmeth.f.303>
- Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Lozupone CA, Turnbaugh PJ, Fierer N, Knight R (2011) Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proc Natl Acad Sci USA* 108:4516–4522. <https://doi.org/10.1073/pnas.100080107>
- Chen W, Simpson J, Levesque CA (2018) RAM: R for amplicon-sequencing-based microbial-ecology. R package version 1.2.1.7. <https://cran.r-project.org/package=RAM>
- Clarke A, Raymond J, Leakey RJG (1996) The seasonal cycle of phytoplankton, macronutrients, and the microbial community in a nearshore antarctic marine ecosystem. *Limnol Oceanogr* 41:1281–1294. <https://doi.org/10.4319/lo.1996.41.6.1281>
- Cowan DA, Ramond J-B, Makhalanyane TP, De Maayer P (2015) Metagenomics of extreme environments. *Curr Opin Microbiol* 25:97–102. <https://doi.org/10.1016/j.mib.2015.05.005>
- Dias-Sardinha I, Craveiro D, Milheiras S (2013) A sustainability framework for redevelopment of rural brownfields: stakeholder participation at SÃO DOMINGOS mine, Portugal. *J Cleaner Prod* 57:200–208. <https://doi.org/10.1016/j.jclepro.2013.05.042>
- Edgar RC (2013) UPARSE: highly accurate OTU sequences from microbial amplicon reads. *Nat Methods* 10:996–998. <https://doi.org/10.1038/nmeth.2604>
- Edwards KJ, Gihring TM, Banfield JF (1999) Seasonal variations in microbial populations and environmental conditions in an extreme acid mine drainage environment. *Appl Environ Microbiol* 65:3627–3632
- García-Moyano A, Austnes AE, Lanzén A, González-Toril E, Aguilera A, Øvreås L (2015) Novel and unexpected microbial diversity in acid mine drainage in Svalbard (78°N), revealed by culture-independent approaches. *Microorganisms* 3:667–694. <https://doi.org/10.3390/microorganisms3040667>
- Gilbert JA, Field D, Swift P, Newbold L, Oliver A, Smyth T, Somerfield PJ, Huse S, Joint I (2009) The seasonal structure of microbial communities in the Western English Channel. *Environ Microbiol* 11:3132–3139. <https://doi.org/10.1111/j.1462-2920.2009.02017.x>
- González-Toril E, Aguilera A, Rodriguez N, Fernández-Remolar D, Gómez F, Diaz E, García-Moyano A, Sanz JL, Amils R (2010) Microbial ecology of Río Tinto, a natural extreme acidic environment of bihydrometallurgical interest. *Hydrometallurgy* 104:329–333. <https://doi.org/10.1016/j.hydromet.2010.01.011>
- González-Toril E, Aguilera A, Souza-Egipsy V, Pamo EL, España JS, Amils R (2011) Geomicrobiology of La Zarza-Perrunal acid mine effluent (Iberian Pyritic Belt, Spain). *Appl Environ Microbiol* 77:2685–2694. <https://doi.org/10.1128/AEM.02459-10>
- Hao C-B, Zhang H-X, Bai Z-H, Hu Q, Zhang B-G (2007) A novel acidophile community populating waste ore deposits at an acid mine drainage site. *J Environ Sci* 19:444–450. [https://doi.org/10.1016/S1001-0742\(07\)60074-6](https://doi.org/10.1016/S1001-0742(07)60074-6)
- Hubbard CG, Black S, Coleman ML (2009) Aqueous geochemistry and oxygen isotope compositions of acid mine drainage from the Río Tinto, SW Spain, highlight inconsistencies in current models. *Chem Geol* 265:321–334. <https://doi.org/10.1016/J.CHEMGEO.2009.04.009>
- Illumina I (2015) 16S metagenomic sequencing library preparation. Part # 15044223 Rev. B. [https://support.illumina.com/downloads/16s\\_metagenomic\\_sequencing\\_library\\_preparation.html](https://support.illumina.com/downloads/16s_metagenomic_sequencing_library_preparation.html)
- Janda JM, Abbott SL (2007) 16S rRNA gene sequencing for bacterial identification in the diagnostic laboratory: pluses, perils, and pitfalls. *J Clin Microbiol* 45:2761–2764. <https://doi.org/10.1128/JCM.01228-07>
- Jovel J, Patterson J, Wang W, Hotte N, O'Keefe S, Mitchel T, Perry T, Kao D, Mason AL, Madsen KL, Wong GK-S (2016) Characterization of the Gut microbiome using 16S or shotgun metagenomics. *Front Microbiol* 7:459. <https://doi.org/10.3389/fmicb.2016.00459>
- Kennedy J, O'Leary ND, Kiran GS, Morrissey JP, O'Gara F, Selvin J, Dobson ADW (2011) Functional metagenomic strategies for the discovery of novel enzymes and biosurfactants with biotechnological applications from marine ecosystems. *J Appl Microbiol* 111:787–799. <https://doi.org/10.1111/j.1365-2672.2011.05106.x>
- Keshri J, Pradeep Ram AS, Sime-Ngando T (2018) Distinctive patterns in the taxonomical resolution of bacterioplankton in the sediment and pore waters of contrasted freshwater lakes. *Microb Ecol* 75:662–673. <https://doi.org/10.1007/s00248-017-1074-z>
- Kim J, Koo S-Y, Kim J-Y, Lee EH, Lee S-D, Ko K-S, Dong-Chan K, Cho K-S (2009) Influence of acid mine drainage on microbial communities in stream and groundwater samples at Guryong

- Mine. S Korea Environ Geol 58:1567. <https://doi.org/10.1007/s00254-008-1663-8>
- Kishimoto N, Kosako Y, Tano T (1991) *Acidobacterium capsulatum* gen. nov., sp. nov.: an acidophilic chemoorganotrophic bacterium containing menaquinone from acidic mineral environment. *Curr Microbiol* 22:1–7. <https://doi.org/10.1007/BF02106205>
- Kuang J-L, Huang L-N, Chen L-X, Hua Z-S, Li S-J, Hu M, Li JT, Shu W-S (2013) Contemporary environmental variation determines microbial diversity patterns in acid mine drainage. *ISME J* 7:1038–1050. <https://doi.org/10.1038/ismej.2012.139>
- Laplante K, Sébastien B, Derome N (2013) Parallel changes of taxonomic interaction networks in lacustrine bacterial communities induced by a polymetallic perturbation. *Evol Appl* 6:643–659. <https://doi.org/10.1111/eva.12050>
- Lee D-G, Trujillo ME, Kang S, Nam J, Kim Y (2018) *Epidermidibacterium keratini* gen. nov., sp. nov., a member of the family Sporichthyaceae, isolated from keratin epidermis. *Int J Syst Evol Microbiol* 68:745–750. <https://doi.org/10.1099/ijsem.0.002579>
- Llirós M, Inceoğlu Ö, García-Armisen T, Anzil A, Leporcq B, Pigneur L-M, Viroux L, Darchambeau F, Descy J-P, Servais P (2014) Bacterial community composition in three freshwater reservoirs of different alkalinity and trophic status. *PLoS ONE* 9:e116145. <https://doi.org/10.1371/journal.pone.0116145>
- Marchesi JR, Ravel J (2015) The vocabulary of microbiome research: a proposal. *Microbiome* 3:31. <https://doi.org/10.1186/s40168-015-0094-5>
- Martins M, Santos ES, Leonor M, Chaves S, Tenreiro R, Barros RJ, Barreiros A, Costa MC (2011) Performance and bacterial community shifts during bioremediation of acid mine drainage from two Portuguese mines. *Int Biodeterior Biodegrad* 65:972–981. <https://doi.org/10.1016/j.ibiod.2011.07.006>
- MATLAB and Statistics Toolbox Release (2016) The MathWorks, Inc., Natick, Massachusetts, United States
- McIlroy SJ, Kirkegaard RH, McIlroy B, Nierychlo M, Kristensen JM, Karst SM, Albertsen M, Nielsen PH (2017) MiDAS 2.0: an ecosystem-specific taxonomy and online database for the organisms of wastewater treatment systems expanded for anaerobic digester groups. *Database* 2017:1–9. <https://doi.org/10.1093/database/bax016>
- Mekki A, Sayadi S (2017) Study of heavy metal accumulation and residual toxicity in soil saturated with phosphate processing wastewater. *Water Air Soil Pollut* 228:215. <https://doi.org/10.1007/s11270-017-3399-0>
- Normand P (2006) The Families Frankiaceae, Geodermatophilaceae, Acidothermaceae and Sporichthyaceae. In: Dworkin M, Falkow S, Rosenberg E, Schleifer KH, Stackebrandt E (eds) *The prokaryotes*. Springer, New York, pp 669–681. <https://doi.org/10.1007/0-387-30743-5-26>
- Okabayashi A, Wakai S, Kanao T, Sugio T, Kamimura K (2005) Diversity of 16S ribosomal DNA-defined bacterial population in acid rock drainage from Japanese pyrite mine. *J Biosci Bioeng* 100:644–652. <https://doi.org/10.1263/jbb.100.644>
- Pereira R, Ribeiro R, Gonçalves F (2015) Plan for an Integrated human and environmental risk assessment in the S. Domingos Mine Area (Portugal). *Hum Ecol Risk Assess* 10:543–578. <https://doi.org/10.1080/10807030490452197>
- Pérez-López R, Delgado J, Nieto JM, Márquez-García B (2010) Rare earth element geochemistry of sulphide weathering in the São Domingos mine area (Iberian Pyrite Belt): a proxy for fluid–rock interaction and ancient mining pollution. *Chem Geol* 276:29–40. <https://doi.org/10.1016/j.chemgeo.2010.05.018>
- Podda F, Zuddas P, Minacci A, Pepi M, Baldi F (2000) Heavy metal coprecipitation with hydrozincite  $[Zn_5(CO_3)_2(OH)_6]$  from mine waters caused by photosynthetic microorganisms. *Appl Environ Microbiol* 66:5092–5098. <https://doi.org/10.1128/AEM.66.11.5092-5098.2000>
- R Core Team (2017) R version 1.0.153: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>
- Reis MP, Dias MF, Costa PS, Ávila MP, Leite LR, de Araújo FMG, Salim ACM, Bucciarelli-Rodríguez M et al (2016) Metagenomic signatures of a tropical mining-impacted stream reveal complex microbial and metabolic networks. *Chemosphere* 161:266–273. <https://doi.org/10.1016/j.chemosphere.2016.06.097>
- Salomons W (1995) Environmental impact of metals derived from mining activities: processes, predictions, prevention. *J Geochem Explor* 52:5–23. [https://doi.org/10.1016/0375-6742\(94\)00039-E](https://doi.org/10.1016/0375-6742(94)00039-E)
- Schrenk MO, Edwards KJ, Goodman RM, Hamers RJ, Banfield JF (1998) Distribution of *Thiobacillus ferrooxidans* and *Leptospirillum ferrooxidans*: implications for generation of acid mine drainage. *Science* 279:1519–1522. <https://doi.org/10.1126/science.279.5356.1519>
- Shannon CE (1948) A mathematical theory of communication. *Bell Syst Tech J* 27:379–423. <https://doi.org/10.1002/j.1538-7305.1948.tb01338.x>
- Simpson EH (1949) Measurement of diversity. *Nature* 163:688. <https://doi.org/10.1038/163688a0>
- Spellerberg IF, Fedor PJ (2003) A tribute to Claude Shannon (1916–2001) and a plea for more rigorous use of species richness, species diversity and the ‘Shannon–Wiener’ Index. *Global Ecol Biogeogr* 12:177–179. <https://doi.org/10.1046/j.1466-822X.2003.00015.x>
- Stewart EJ (2012) Growing unculturable bacteria. *J Bacteriol* 194:4151–4160. <https://doi.org/10.1128/JB.00345-12>
- Tavares MT, Sousa AJ, Abreu MM (2008) Ordinary kriging and indicator kriging in the cartography of trace elements contamination in São Domingos mining site (Alentejo, Portugal). *J Geochem Explor* 98:43–56. <https://doi.org/10.1016/j.gexplo.2007.10.002>
- Thomas T, Gilbert J, Meyer F (2012) Metagenomics—a guide from sampling to data analysis. *Microb Inform Exp* 2:3. <https://doi.org/10.1186/2042-5783-2-3>
- Vicente-Villardón JL (2015) MULTBIPLLOT version 16.0430: a package for multivariate analysis using biplots Departamento de Estadística. Universidad de Salamanca. Spain. <https://biplot.usal.es/ClassicalBiplot/index.html>
- Wang Q, Garrity GM, Tiedje JM, Cole JR (2007) Naïve Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl Environ Microbiol* 73:5261–5267. <https://doi.org/10.1128/AEM.00062-07>
- Zhang X, Tang S, Wang M, Sun W, Xie Y, Peng H, Zhong A, Liu H et al (2019) Acid mine drainage affects the diversity and metal resistance gene profile of sediment bacterial community along a river. *Chemosphere* 217:790–799. <https://doi.org/10.1016/j.chemosphere.2018.10.210>
- Ziegler S, Waidner B, Itoh T, Schumann P, Spring S, Gescher J (2013) *Metallibacterium scheffleri* gen. nov., sp. nov., an alkalizing gammaproteobacterium isolated from an acidic biofilm. *Int J Syst Evol Microbiol* 63:1499–1504. <https://doi.org/10.1099/ijss.0.042986-0>

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.