Title

Bioreactor microbial ecosystems for thiocyanate and cyanide degradation unravelled with genome-resolved metagenomics

Journal

Environmental Microbiology, 17(12)

ISSN

1462-2912

Authors

Kantor, Rose S

van Zyl, A Wynand

van Hille, Robert P

et al.

Publication Date

2015-12-01

DOI

10.1111/1462-2920.12936

Peer reviewed

Bioreactor microbial ecosystems for thiocyanate and cyanide degradation unravelled with genome-resolved metagenomics

Rose S. Kantor,[1] A. Wynand van Zyl,[2] Robert P. van Hille,[2] Brian C. Thomas,[3] Susan T. L. Harrison[2] and Jillian F. Banfield[3,4]*

Departments of [1] Plant and Microbial Biology, [3] Earth and Planetary Sciences, and [4] Environmental Science, Policy, and Management, University of California, Berkeley, CA, USA. [2] Center for Bioprocess Engineering Research, Department of Chemical Engineering, University of Cape Town, Cape Town, South Africa.

*For correspondence. E-mail jbanfield@berkeley.edu

Summary

Gold ore processing uses cyanide (CN^-), which often results in large volumes of thiocyanate- (SCN^-) contaminated wastewater requiring treatment. Microbial communities can degrade SCN^- and CN^-, but little is known about their membership and metabolic potential. Microbial-based remediation strategies will benefit from an ecological understanding of organisms involved in the breakdown of SCN^- and CN^- into sulfur, carbon and nitrogen compounds. We performed metagenomic analysis of samples from two laboratory-scale bioreactors used to study SCN^- and CN^- degradation. Community analysis revealed the dominance of *Thiobacillus* spp., whose genomes harbour a previously unreported operon for SCN^- degradation. Genome-based metabolic predictions suggest that a large portion of each bioreactor community is autotrophic, relying not on molasses in reactor feed but using energy gained from oxidation of sulfur compounds produced during SCN^- degradation. Heterotrophs, including a bacterium from a previously uncharacterized phylum, compose a smaller portion of the reactor community. Predation by phage and eukaryotes is predicted to affect community dynamics. Genes for ammonium oxidation and denitrification were detected, indicating the potential for nitrogen removal, as required for complete remediation of wastewater. These findings suggest optimization strategies for reactor design, such as improved aerobic/anaerobic partitioning and elimination of organic carbon from reactor feed.

Introduction

The use of microbes in biodegradation is widely studied as an alternative to traditional methods for remediating environmentally harmful waste. Thiocyanate (SCN^-), a cyanide derivative with lower toxicity, is produced at high concentrations in gold mining effluents and other industrial processes, including coal gasification and steel processing. SCN^- also occurs naturally, both as a product of cyanide (CN^-) detoxification in many organisms (Cipollone *et al*., 2007) and as a degradation product of cyanogenic glucosinolates, found in plants and insects (Jensen *et al*., 2011). However, SCN^- at high concentrations can be harmful to human health (Erdogan,

2003) and aquatic life (Watson and Maly, 1987; Speyer and Raymond, 1988), necessitating SCN^- removal from mining wastewater. Several chemical methods have been used for SCN^- destruction, but these can be costly and inefficient, or can produce other toxic chemicals (Gould *et al*., 2012). In contrast, biological degradation is relatively inexpensive and can completely remove SCN^- and co-contaminants, such as CN^-. Known degradation pathways for SCN^- include an autotrophic pathway in which SCN^- is a source of energy, sulfur and nitrogen. This pathway is reported to produce ammonium and carbonyl sulfide (OCS), which is subsequently converted to sulfide (Sorokin *et al*., 2001; 2004; Arakawa *et al*., 2007; Ogawa *et al*., 2013). A heterotrophic pathway has been proposed for organisms capable of growth on organic carbon with SCN^- as the sole nitrogen source (Stratford *et al*., 1994). This pathway may produce sulfide and cyanate (OCN^-), which can be converted to ammonium via the enzyme cyanase (Sung and Fuchs, 1988). The principal end-products of microbial community SCN^- degradation are typically sulfate and ammonium or nitrate (Boucabeille *et al*., 1994), consistent with either degradation pathway, coupled to ammonium and sulfide oxidation.

SCN^- biodegradation has been studied with a focus on reactor design and optimization for industrial applications (Boucabeille *et al*., 1994; Dictor *et al*., 1997; Hung and Pavlostathis, 1999; du Plessis *et al*., 2001; Stott *et al*., 2001; van Zyl *et al*., 2015), and some demonstration-scale plants have shown biodegradation to be successful (e.g. Activated Sludge Tailings Effluent Removal: ASTER™ process, Biomin Limited; Consort Mine in Barberton, South Africa; Suzdal Mine, Kazakhstan; and Homestake Mine, SD, USA) (van Buuren *et al*., 2011). The microbiological and molecular study of SCN^- and CN^- degradation has focused on isolates lacking sequenced genomes (Katayama *et al*., 1998; Wood *et al*., 1998; du Plessis *et al*., 2001; Sorokin *et al*., 2004; 2007; 2012; Arakawa *et al*., 2007; Hussain *et al*., 2013). Molecular fingerprinting of SCN^- degrading consortia (Quan *et al*., 2006; Felföldi *et al*., 2010; Huddy *et al*., 2015) provides the only knowledge at the community level to date. Hence, little is known about the metabolic potential of SCN^- degrading consortia, although this is critical to understanding the degradation process.

In this study, we sampled biofilm and supernatant from two long-running laboratory-scale continuous flow bioreactors, treating either SCN^- ('SCN-only') or a mixture of CN^- and SCN^- ('CN–SCN'). We used high-throughput metagenomic sequencing to reconstruct microbial draft and curated genomes from these two communities. We describe the community structure and outline potential nutrient flow paths, revealing a complex community that includes chemoautotrophs, heterotrophs and possible predators.

Results

Reactor chemistry and observations

Prior to sampling for metagenomic analysis, the SCN-only reactor degradation performance was stable and highly efficient. Ammonium in reactor effluent accounted for between 35% and 82% of nitrogen input from SCN^- (Fig. 1A), suggesting the possibility of nitrogen uptake or removal. Reactor biomass accumulated visibly in the form of thick biofilms (Fig. 1C), beginning when SCN^- loading reached 1.72 mmol h^{-1} (100 mg h^{-1}), 173 days prior to sampling (not shown).

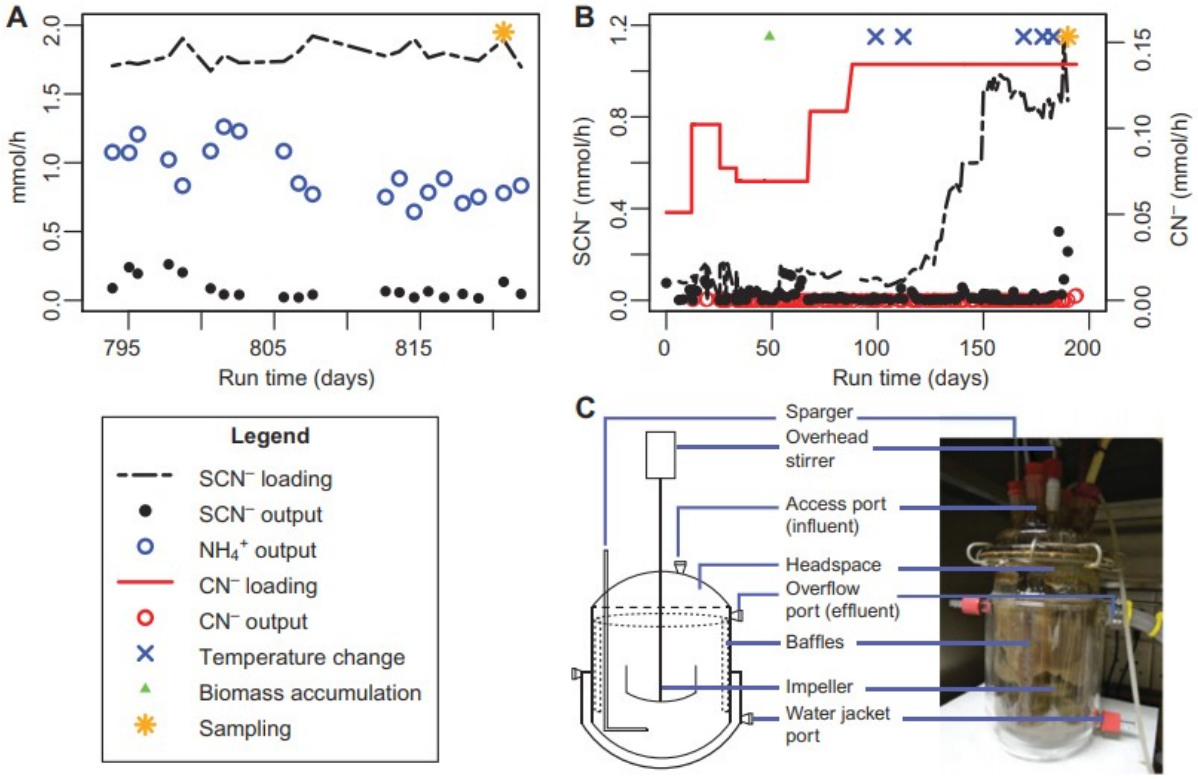

Fig. 1. Conditions and set-up for two aerobic continuously stirred reactors treating (A) only SCN^- or (B) CN^- and SCN^-. Loadings indicate rates of compound input to reactors via media, while outputs indicate rates of compound produced in effluent. Biomass was observed and sampling for metagenomics was performed where indicated. In the CN–SCN reactor, temperature was varied, beginning at 25°C and shifting at times marked to 40 (and back to 25), 52 (and back to 25), 20 (held), 18 (held) and 15°C (held) respectively. Diagram and photograph of the SCN-only reactor in operation (C) shows mechanisms of stirring and aeration. Biofilm (dark colour) occurred on glass reactor walls, impeller, sparger and baffles.

The CN–SCN reactor mimicked conditions of industrial reactors processing mining wastewater, including the presence of CN^- in the feed and temperature manipulations. Brief fluctuations in degradation efficiency were observed in the CN–SCN reactor following such manipulations, after which stable operation was resumed (Fig. 1B). Biomass accumulated as biofilm in this reactor, similar to the SCN-only reactor, and samples were collected from both reactors for extraction of community genomic DNA (Fig. 1B).

Genome recovery and community structure

High-throughput sequencing and assembly of short-read data into longer scaffolds yielded metagenomes of 97 Mbp for the SCN-only sample and 295 Mbp for the CN–SCN sample (on scaffolds ≥ 5 Kbp; Table 1). Scaffolds were

binned into genomes based on their coverage by reads in each sample and their di- and trinucleotide frequencies, generating 29 genome bins for the SCN-only reactor and 64 for the CN–SCN reactor (Table S1). These bins accounted for essentially all organisms sampled, such that all scaffolds containing > 8 key ribosomal proteins (RPs) (Fig. S2) and nearly all scaffolds (≥ 5 Kbp) containing single copy genes were binned. The overall taxonomic compositions of both reactors were similar, as the inoculum for the CN–SCN reactor was taken from the SCN-only reactor. More genomes were recovered from the CN–SCN reactor sample, perhaps owing to the greater depth of sequencing for this sample (Table 1), to sampling bias, or to community shifts due to the different conditions in each reactor.

Table 1. Reactor conditions at time of sampling and sequencing data acquired.

	SCN stock reactor	SCN–CN reactor
Reactor volume	1 l	1 l
Hydraulic retention time	12 h	14 h
Thiocyanate loading (as KSCN)	1.9 mmol h^{-1} (110.0 mg h^{-1})	0.9 mmol h^{-1} (50.6 mg h^{-1})
Cyanide loading (as NaCN)	0 mmol h^{-1}	0.14 mmol h^{-1} (3.57 mg h^{-1})
Feed		
Thiocyanate (as KSCN)	22.7 mM (1320 mg l^{-1})	12.2 mM (708.4 mg l^{-1})
Cyanide (as NaCN)	0 mM	1.9 mM (50.0 mg l^{-1})
Phosphate (as KH_2PO_4)	0.28 mM (27 mg l^{-1})	0.28 mM (27 mg l^{-1})
Molasses	150 mg l^{-1}	150 mg l^{-1}
Temperature	25°C	15°C
Reactor pH (with NaOH)	8.5	8.5
SCN degradation efficiency	93%	76%
CN degradation efficiency	NA	98%
Total sequence	5.5 Gbp	34.7 Gbp
Length of assembly in contigs ≥ 5 kb	97 Mbp	295 Mbp
Scaffolds ≥ 5 kb	3 811	13 707
Genome bins	29	64

Rank abundance analysis based on binned genomes demonstrated the dominance of *Thiobacillus* spp. in both samples (Fig. 2). As the assemblies of these genomes were highly fragmented, we assembled subsamples of the reads to recover the two highest abundance *Thiobacillus* spp. genomes present in the SCN-only reactor. These, and a third lower abundance genome from the original assembly, accounted for 27.2% of all reads in the SCN dataset. Within the CN–SCN dataset, *Thiobacillus* spp. were similarly abundant, but high coverage and strain variation produced fragmented genomes that could not be satisfactorily resolved with subsampling of reads and reassembly.

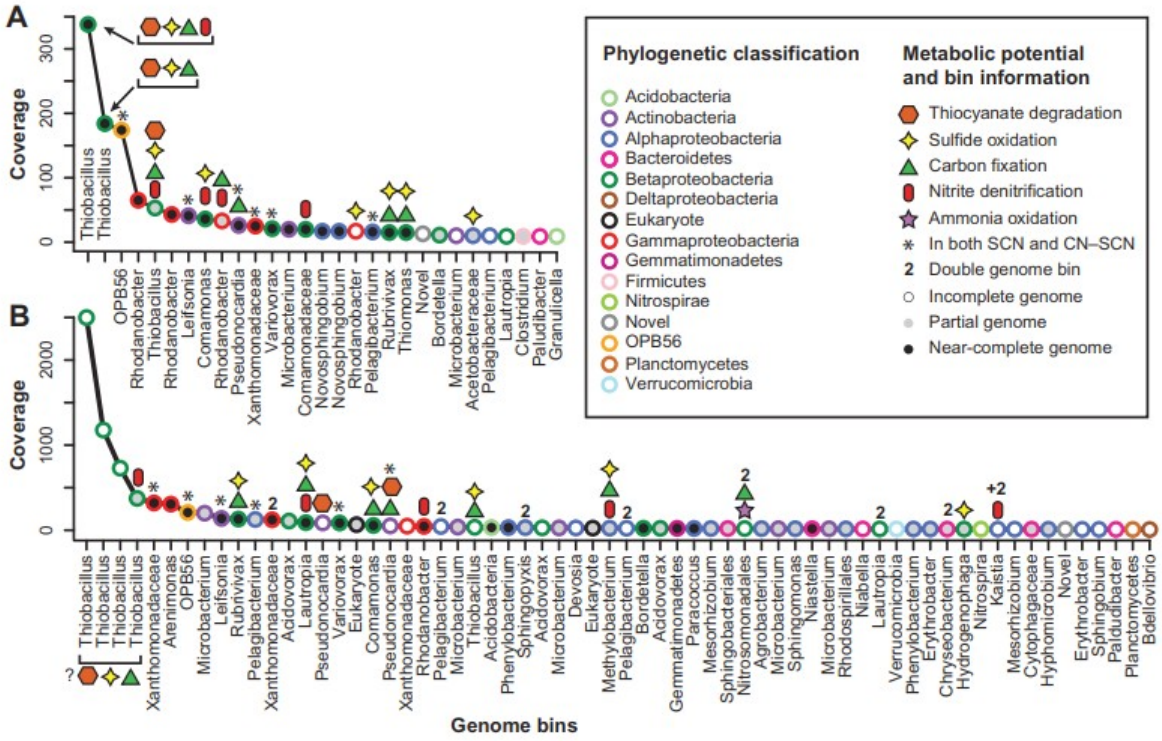

Fig. 2. Rank abundance and metabolic potential for organisms in (A) the SCN-only or (B) the CN–SCN reactor. For each reactor, genome bins are ordered from highest to lowest coverage (the average read depth over all scaffolds in the bin). Outer circle colour indicates phylogenetic affiliation, and circle fill indicates genome completeness measured by the presence/absence of 51 single copy genes: white (0–25 genes), grey (26–47) and black (48–51) (see Appendix S1).

In addition to *Thiobacillus* spp., six other organisms were also found at high relative abundance in both samples (Fig. 2; Fig. S2). The reconstructed sequences for these organisms in the two samples shared > 98% nucleotide identity across 93–99% of the genomes' length. One of these six organisms was a member of the uncultivated OPB56 radiation within the Bacteroidetes-Ignavibacterium-Chlorobi superphylum (Figs S2 and S3). Near-identical draft genomes for this organism were independently assembled and binned from both bioreactor metagenomic datasets. Notably, the two sequences were broken in the same regions due to repeated elements. Manual correction of local assembly errors and closure of internal gaps using reads from both datasets generated a high-quality, near-complete genome sequence. Another genome found in both samples belongs to an Actinobacteria most closely related to *Leifsonia* spp. This genome was also manually curated to near-complete status. The other four genomes, *Pelagibacterium* sp., *Variovorax* sp., *Pseudonocardia* sp. and a member of the Xanthomonadaceae, remain as draft genomes in both datasets.

Also included in both reactors were multiple members of the genera *Rhodanobacter*, *Microbacterium*, *Pseudonocardia* and *Pelagibacterium*, as well as members of the orders Sphingomonadales, Burkholderiales and Xanthomonadales (Fig. 2; Figs S2 and S3). Near-complete genomes were recovered for several organisms found in one but not both reactors. These

included members of the phyla Acidobacteria and Gemmatimonadetes that are distantly related to previously genomically sampled organisms (Fig. S2). Several other genomes remained partial, and in a few instances individual genomes could not be resolved.

In addition to bacterial genomes, we detected microbial eukaryotes in both reactors via assembly of partial mitochondrial genomes, and in two cases, nuclear genomes (Table S1). One of these nuclear genomes was identified as belonging to the Rhizaria, while the other fell within Opisthokonta, sibling to the slime mold *Fonticula alba* (Fig. S4). Many phage, plasmid and transposon sequences were also recovered.

Metabolic analysis and nutrient flow

In order to identify potential roles for organisms in this system and their interactions, we examined metabolisms predicted from sequences within each genome bin. Of particular interest were genes and pathways potentially involved in the breakdown of SCN^- and CN^- and their degradation products: sulfur, carbon and nitrogen compounds (Fig. 2). These findings, by pathway, are detailed below.

Thiocyanate degradation via SCNase operon in Thiobacillus spp. and Pseudonocardia spp

We searched both datasets for the genes encoding the cobalt-coordinating enzyme thiocyanate hydrolase (SCNase), known to convert SCN^- to carbonyl sulfide (COS) and ammonium in *Thiobacillus thioparus* THI115 (Kataoka *et al*., 2006; Arakawa *et al*., 2007). These genes did not assemble well in the CN–SCN dataset, but were present on *Thiobacillus* spp. scaffolds in some subassemblies. The genes were complete in three *Thiobacillus* spp. genomes in the SCN-only reactor. Here, the SCNase genes were colocated in a conserved operon (Fig. 3; Fig. S5A), which also contained cyanase, the enzyme responsible for cyanate (OCN^-) metabolism to ammonium. We searched for carbonyl sulfide hydrolase (COSase), the carbonic anhydrase family enzyme responsible for the degradation of COS produced from SCN^- in *T. thioparus* THI115 (Ogawa *et al*., 2013), but we did not detect this gene in either dataset. Other genes in the SCN^- operon, including three with possible roles in sulfur metabolism (Fig. 3; Fig. S5A), may compensate for the lack of a COSase.

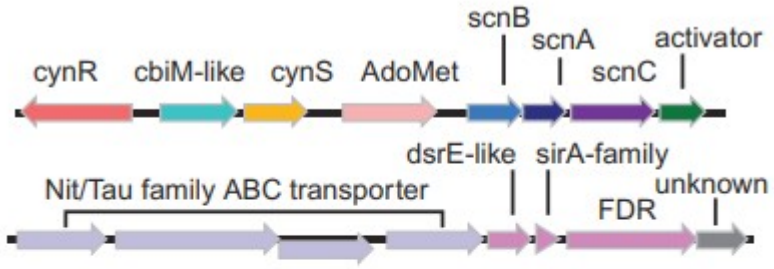

Fig. 3. The thiocyanate operon conserved in *Thiobacillus* spp. Genes included are cyanase (*cynS*) and the cyanase transcriptional regulator (*cynR*), *cbiM* (membrane component of cobalt transporter), a methyltransferase (adomet), thiocyanate hydrolase subunits (*scnBAC*) and the activator (P15K), a four-subunit Nit/Tau family ABC transporter, three genes possibly related to sulfur oxidation (in pink; FDR is an FAD-dependent pyridine nucleotide-disulfide oxidoreductase family protein), and an unknown protein (conserved in all SCN operons analysed, see Fig. S5A).

A genomic region similar to the SCN^- operon was found in both of the *Pseudonocardia* spp. genomes from the CN–SCN reactor (Fig. S5B). The region contains a nitrile hydratase whose metal-coordinating alpha subunit has high sequence similarity to the SCNase gamma subunit, including conserved cobalt-coordinating residues and residues that dictate substrate specificity for SCN^- rather than nitriles (Yamanaka *et al*., 2013). Additionally, an unrelated type of SCNase, recently identified in *Afipia* sp. TH201 (Hussain *et al*., 2013), appears to be conserved in two of the *Thiobacillus* spp. genomes in the SCN-only reactor (Table S2). This protein is found in a region of the *Thiobacillus* spp. genomes that contains genes for SoxA, SoxX and a thioredoxin-like protein, and is not associated with the main *sox* operon (see below).

Capacity for cyanide degradation and tolerance may be widespread within the community

Rhodaneses are enzymes that convert CN^- and thiosulfate to SCN^- *in vitro* (Cipollone *et al*., 2004), and confer improved growth with CN^- *in vivo* (Cipollone *et al*., 2007). Because rhodanese domains occur in a variety of proteins with functions unrelated to CN^- metabolism, we confined our search to two-domain rhodaneses *sensu stricto* (rhdA, EC: 2.8.1.1) (Cipollone *et al*., 2007). A majority of genomes from both reactors were found to contain at least one of these rhodaneses (Table S2).

Other candidates for enzymatic CN^- degradation in the bioreactor system include homologues of fungal cyanide hydratase (EC: 4.2.1.66) (Wang and VanEtten, 1992; Cluness *et al*., 1993) and bacterial cyanide dihydratase (cynD, EC: 3.5.5.-) (Jandhyala *et al*., 2003; 2005; Fernandez and Kunz, 2005). These proteins are class I nitrilases (Pace and Brenner, 2001), for which the current understanding of structure–function relationship is limited (Thuku *et al*., 2009). We searched for homologues to the bacterial and fungal proteins

and identified several putative cyanide (di)hydratases in both datasets (Table S2). Additionally, genes annotated as cytochrome *bd* ubiquinol oxidases (*cydAB*) were sometimes identified adjacent to class I nitrilases in *Rhodanobacter* sp. and some Alphaproteobacteria (Table S2). These cytochromes could act as alternative terminal reductases when cytochrome C oxidases are inhibited by CN^-, although the cyanide insensitivity of *cydAB* cytochrome oxidases is not predictable based on protein sequence alone (Borisov *et al*., 2011). Lastly, in some CN-degrading pathways, CN^- may first be converted to a nitrile, which is suggested to be subsequently degraded by nitrile hydratases (Luque-Almagro *et al*., 2011). Nitrile hydratases were identified in several genome bins in both datasets (Table S2).

Reduced sulfur as an energy source in the reactor system

As SCN^- is a reduced form of sulfur and sulfate is known to accumulate in the bioreactor system, we hypothesized that oxidation of some intermediate sulfur compounds (such as sulfide or thiosulfate) may serve as a key energy-generating process for some organisms in the reactors. We identified pathways for sulfur-compound oxidation in at least eight genomes within the SCN-only dataset and nine in the CN–SCN dataset (Fig. 2 and Table S2). These pathways include the Sox enzymes, which can perform complete oxidation of sulfide to sulfate, and reverse dissimilatory sulfite reductase (rDsr) enzymes, which can oxidize sulfide to sulfite. We also identified genes for APS reductase and ATP sulfurylase, which may convert sulfite to sulfate (Table S2), completing the oxidation.

The *Thiobacillus* genomes each contains partial *sox* operons (Table S2) as previously observed for *T. denitrificans* (Beller *et al*., 2006), and complete operons containing rDsr were detected in all three *Thiobacillus* genomes and a *Rubrivivax sp.* genome from the SCN-only reactor. Additionally, several sulfide-quinone reductase-like genes (*sqr*), which can oxidize sulfide to sulfur, were identified within genomes in both datasets (Table S2). Although active sites and binding sites were confirmed for these predicted proteins (Marcia *et al*., 2009; Cherney *et al*., 2010), the SQR proteins are difficult to distinguish from proteins of other functions by sequence or structural prediction (Marcia *et al*., 2010), and therefore are not included in Fig. 2.

Carbon flow: dominance of chemolithoautotrophy

Carbon input for the system comes from both SCN^- and molasses in the reactor feed, but due to a substantial accumulation of biomass observed as SCN^- loading was increased, we expected that some community members were capable of carrying out carbon fixation. Concordantly, we identified genes for the complete Calvin–Benson–Bassham (CBB) cycle in seven genome bins in the SCN-only reactor and 10 genome bins in the CN–SCN reactor (Table S2). Key genes for the Wood–Ljungdahl pathway (anaerobic CODH/ACS) and reverse TCA cycle (citrate lyase) were absent from all genomes. The *Thiobacillus* spp. (Fig. 4A) and *Thiomonas* sp. genomes carry both type I and type II RuBisCO. The high-abundance *Thiobacillus* sp.

genome in the SCN-only dataset also contains a putative type ID RuBisCO located in the same operon as the gene encoding type II RuBisCO. Only *Thiobacillus* spp., *Thiomonas* sp. and *Pseudonocardia* spp. genomes possessed genes encoding carboxysome proteins. As expected, genomes belonging to the dominant *Thiobacillus* spp. and several other relatively abundant organisms encoded both the CBB cycle and sulfur oxidation pathways, suggesting that chemolithoautotrophy is important in this system (Fig. 2 and Table S2).

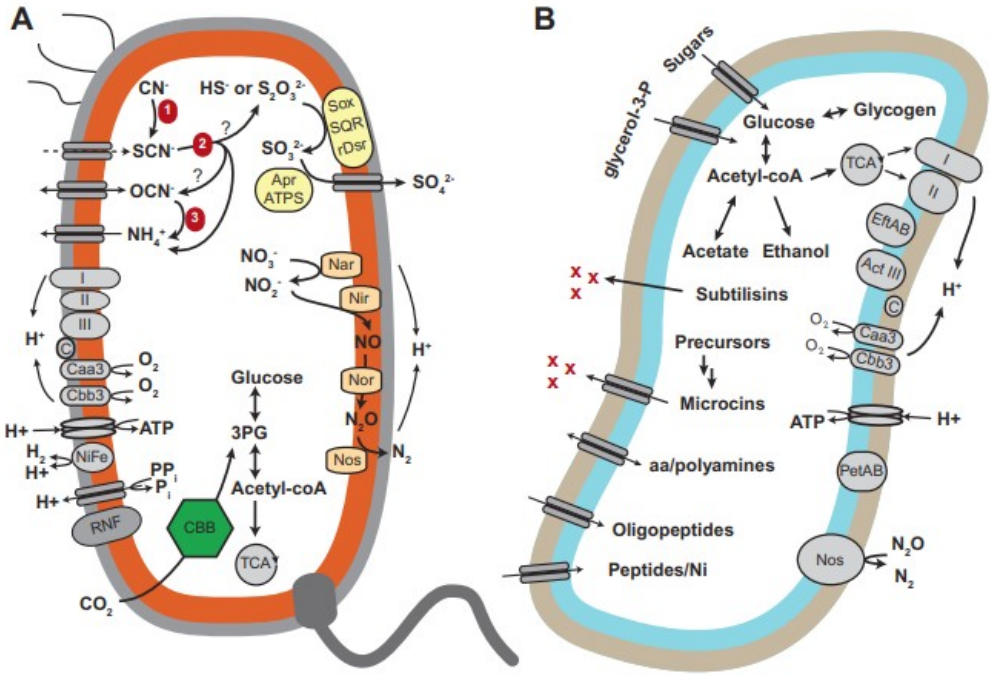

Fig. 4. Cell diagrams depicting predicted metabolic potential for two abundant community members. The dominant, autotrophic *Thiobacillus* sp. in the SCN-only reactor (A) contains genes involved in sulfur oxidation, denitrification and carbon fixation. Key genes for SCN^- and CN^- metabolism in this genome include (1) rhodanese, (2) scnase and (3) cyanase. The newly described OPB56 organism (B) is a predicted heterotroph, present at high abundance in both reactors. This genome encodes proteins for scavenging and bactericidal activity. It also harbours genes for alternative electron transport proteins eftab, alternative complex III and petab, as well as nitrous oxide reductase.

Carbon flow: an abundant, novel heterotroph in both reactors

Each reactor sample contained many genome bins belonging to predicted heterotrophs, including five relatively abundant genomes common to both reactors (Fig. 2). Metabolic reconstruction focused on the OPB56 genome, representing the first sequenced member of this phylum. Analysis suggests this bacterium is Gram-negative and capable of aerobic heterotrophic metabolism (Fig. 4B). Notably, it possesses a full suite of genes for production of antibacterial microcins. These include five genes encoding microcin precursors with near-identical leader sequences, as described by Haft and colleagues (2010). The OPB56 genome also encodes at least six predicted extracellular subtilisin proteases and predicted transporters for complex organics, including sugars, peptides and amino acids.

Nitrogen removal: capacity for ammonium oxidation and denitrification

The principal source of nitrogen for the bioreactor systems is SCN^-, which yields ammonium upon degradation (Fig. 1). Tracing all genes responsible for ammonium oxidation and denitrification revealed that the capacity for nitrite production (at low abundance) and denitrification likely exists in both reactor communities (Table S2). Sequences for ammonium monooxygenase and hydroxylamine oxidoreductase were found on unbinned scaffolds in both datasets, corresponding to relatives of *Nitrosospira multiformis*. Genes for nitrite oxidation, anaerobic ammonium oxidation and dissimilatory nitrate reduction to ammonium were not detected.

There were five organisms with predicted capacity for denitrification of nitrite within the SCN-only reactor and five different organisms with this potential in the CN–SCN reactor (Fig. 2). Other genome bins were missing one gene in the pathway, possibly due to incomplete genome recovery or because the pathway is incomplete in these organisms. Some members of the Xanthomonadaceae appear to denitrify to N_2O, as has been observed in certain Xanthomonadaceae isolates (Finkmann *et al*., 2000; Chen *et al*., 2002). Interestingly, two other genome bins (Gemmatimonadetes and OPB56) possess only genes encoding nitrous oxide reductase, suggesting they could scavenge this intermediate to complete denitrification. A *nirK* sequence was identified in the Rhizaria (eukaryote) genome bin within the CN-treated reactor (Table S2), suggesting this eukaryote may be able to use nitrite as an electron acceptor under anaerobic conditions, as observed in the fungus *Fusarium oxysporum* (Kim *et al*., 2009).

Of the *Thiobacillus spp.* present in the SCN-only community, two possess complete denitrification pathways (Figs 2 and 4). The third does not encode any denitrification genes despite containing all single-copy genes. A complete denitrification pathway, including NirS instead of NirK, was recovered from one *Thiobacillus* sp. bin in the CN–SCN reactor. The other *Thiobacillus* bins in this dataset encoded some denitrification genes. The denitrification pathways may be incomplete in these organisms, but more likely the remaining genes were not recovered due to poor assembly of these genomes.

Cyanate as a possible nitrogen source

Several genomes in each reactor encoded predicted active cyanases (Sung and Fuchs, 1988; 1992; Guilloton *et al*., 1993; Walsh *et al*., 2000) (Table S2). In most of these genomes, the *cynS* gene was located immediately downstream of genes encoding a putative cyanate transporter, as described previously for *Pseudomonas pseudoalcaligenes* (Luque-Almagro *et al*., 2008). Possession of cyanase may allow access to nitrogen, in the form of OCN^-, for organisms without the capacity to degrade SCN^- directly; however, cyanase has been shown to be unnecessary for cyanide degradation (Luque-Almagro *et al*., 2008).

Discussion

Community structure based on energy and nutrient flow predictions

The two SCN-degrading bioreactor communities described here are enriched in several genera that have previously been detected in or isolated from SCN^- and/or CN-degrading microbial communities. These included *Thiobacillus*, *Mesorhizobium*, *Sphingomonas*, *Sphingopyxis*, *Comamonas*, *Rhodanobacter* and *Microbacterium* (du Plessis *et al*., 2001; Quan *et al*., 2006; Felföldi *et al*., 2010; Huddy *et al*., 2015). The survival and proliferation of multiple *Thiobacillus* spp. and six other organisms under both high SCN^- and CN–SCN conditions indicates tolerance to and/or the capacity to use SCN^-. Overall, the communities in both reactors were diverse, but some functional capacities were shared (Fig. 2). This may provide resilience to fluctuations in environmental factors, such as temperature, access to oxygen, pH (lowered due to sulfate production) and the presence of CN^-.

From our genome-based metabolic analysis, we can divide both communities into trophic groups (Fig. 5). Of the several predicted sulfur-oxidizing chemolithoautotrophs (Fig. 2), only the *Thiobacillus* spp. possess known genes for SCN^- degradation. The other chemolithoautotrophs may benefit from the production of reduced sulfur compounds from SCN^-. Several of these autotrophs also encode complete denitrification pathways (Table S2). Additional organisms with the capacity for carbon fixation include *Pseudonocardia* spp. and the ammonium-oxidizing *Nitrosomonas* spp.

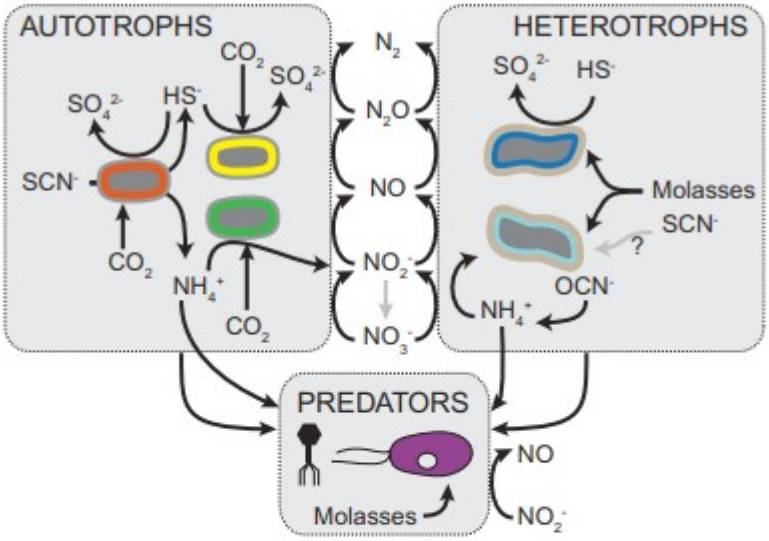

Fig. 5. Schematic diagram of SCN^- degradation in a molasses-fed bioreactor community describing potential flow of sulfur, nitrogen and carbon through the reactor. Specific organisms are represented in different colours: *Thiobacillus* spp. (orange), other sulfur-oxidizing autotrophs (yellow), *Nitrosomonas* spp. (green), sulfur-oxidizing mixotrophs (dark blue), other heterotrophs (light blue), phage (black) and eukaryotes (purple). Grey boxes represent trophic groups as a whole. Black arrows indicate chemical transformation by predicted proteins found in the metagenomes, while grey arrows indicate genes that were either undetected (e.g. *Nxr*) or unknown.

Heterotrophs may utilize molasses in the reactor feed, ammonium produced from SCN^- degradation, and accumulated biomass or extracellular organic compounds produced by autotrophs (Fig. 5). Additionally, some heterotrophic organisms may be able to use SCN^- directly as a source of nitrogen, as has been observed for *Burkholderia phytofirmans* and *Methylobacterium thiocyanatum* (Wood *et al*., 1998; Vu *et al*., 2013). Others are predicted mixotrophs, oxidizing sulfur compounds for energy while using organic carbon. The OPB56 organism may be considered a heterotroph or a predator, based on its predicted capacity to produce bactericidal compounds.

Other predators include phage and eukaryotes, which were present in both samples. Eukaryotes may be capable of utilizing sugars in the molasses media and nitrogen products of SCN^- degradation, but they likely also consume bacterial biomass (Fig. 5). In fact, light microscopy revealed the presence of unidentified grazers in the SCN-only reactor (data not shown), and a variety of eukaryotes have been observed in this system based on 18S rRNA gene clone libraries and isolations (Huddy *et al*., 2015). To our knowledge, the eukaryotic genomes presented here are some of the few reported from complex metagenomic data and the only ones from a high-SCN^- environment. No genome annotator assessed was able to predict genes for these genomes accurately, highlighting the need for improvement

in methodology in this area and for further work to obtain transcriptomic data.

Overall, the apparent dominance of autotrophic SCN^- degraders leads us to speculate that it may be possible to eliminate molasses from the reactor feed, reducing operating costs while maintaining functionality. It may also be possible to selectively remove eukaryotes from the system, although such alterations could affect reactor stability.

Thiocyanate degradation pathways: conservation and horizontal gene transfer

While several bacteria with SCNase activity have been reported in the literature, only two SCNase genes have been identified (Katayama *et al*., 1998; Hussain *et al*., 2013). We detected both of these in the bioreactor systems, and notably, the genes for cobalt-coordinating SCNase were found in a newly identified operon in *Thiobacillus* spp. We found the same operon in the genomes of *Thiobacillus thioparus* DSM 505, an aerobe, and *Thioalkalivibrio thiocyanodenitrificans* ARhD 1, a Gammaproteobacteria known to be able to perform SCN degradation under denitrifying conditions (Sorokin *et al*., 2004) (Fig. S5A). The genes in the SCN^- operon of *T. thiocyanodenitrificans* are conserved relative to the *Thiobacillus* spp. genomes, except for the substitution of the ABC-type cobalt transporter genes, *cbtAB*, in place of *cbiM*. High amino acid similarity between all genes in the *T. thiocyanodenitrificans* operon to those in *Thiobacillus* spp. and lack of SCNase in other bacteria with sequenced genomes suggests possible horizontal gene transfer, although gene order has been only partially conserved (Fig. S5A). Studies examining diverse bacterial species have suggested that cyanase is coexpressed with SCNase and/or is highly expressed under SCN-degrading conditions (Wood *et al*., 1998; Sorokin *et al*., 2004; Bezsudnova *et al*., 2007). The presence of a cyanase gene within the SCN operon could account for its observed coexpression with SCNase.

Previous work suggests the existence of a heterotrophic pathway for SCN^- degradation, which produces OCN^- to be used as a nitrogen source; however, the proteins involved in this pathway are unknown. Within the bioreactor systems, only the *Thiobacillus* spp., and possibly *Pseudonocardia* spp. genome bins, possess the SCN^- operon or alternative SCNase. Thus, while a small number of species may be responsible for most SCN^- degradation here, expression data and isolation experiments are needed to provide insight into other possible pathways.

Nitrogen cycling in aerobic/anaerobic reactor phases

Given that genes for both ammonium oxidation and denitrification were detected, it may be possible to achieve near-complete removal of nitrogen compounds from SCN-contaminated wastewater if reactor operation were modified to increase the rates of these processes. The presence of thick biofilm in the reactors (Fig. 1C) may already provide microaerobic or

anaerobic environments suitable for denitrification, possibly accounting for some of the nitrogen released through SCN^- degradation. As performed by Kraft and colleagues (2014), mass balances for nitrogen (and sulfur), expression data, and short-term reactor manipulations could provide insight into the dynamics of nitrogen and sulfur cycling within this complex system.

Accurate reconstruction of the genome of a novel bacterium

The genome for the novel OPB56 organism is the first near-complete sequence from this group within the Bacteroidetes–Ignavibacterium–Chlorobi radiation. Given that this organism has been maintained in the bioreactor mixed culture and identical draft genomes were recovered independently from both reactors, we propose the name Candidatus '*Kapabacteria thiocyanatum*' for this organism. The name, 'Kapa', reflects the location of cultivation, University of Cape Town. Further, we propose the name Candidatus Kapabacteria to replace OPB56 as the designation for this phylum.

Conclusion

Metagenomics has been used to understand nutrient flow in relatively low-diversity bioreactor communities degrading contaminants such as chlorinated organics (Hug *et al*., 2012) and terephthalate (Lykidis *et al*., 2010; Wu *et al*., 2012). This is the first application of genome-resolved metagenomics to characterize SCN^- and CN^- bioreactors, revealing a complex community containing novel organisms and genes. The analysis identified members with potentially important roles in the sulfur and nitrogen cycles, providing a framework for understanding a microbial community performing SCN^- degradation. Further manipulations of the reactor system will be required to elucidate the factors controlling this process and subsequent nitrogen removal.

Experimental procedures

Reactor operation and sampling

SCN-only reactor

A 1 l water-jacketed reactor (Glass Chem, Stellenbosch, South Africa) was inoculated with sludge from a demonstration-scale bioreactor treating SCN-containing mining effluent (Consort Mine, Barberton, South Africa). The original source of inoculum for the demonstration-scale reactor was obtained from a mixture of sludge from an SCN-contaminated tailings pond and sludge from a domestic wastewater treatment plant. The 1 l reactor was operated in continuous flow, with a 12 h hydraulic retention time under continuous aeration (0.8 l min^{-1}) and stirring (270 r.p.m.). The feed SCN^- loading was increased incrementally from 0.22–1.72 mmol h^{-1} over 150 days. Microbial biomass was maintained in the reactor by means of a clarifier, which allowed settling and underflow sludge recycling. The set-up of this reactor, depicted in Fig. 1C, is described further by van Zyl *et al*. (2011). The reactor was

sampled at day 820 for metagenomic analysis (Fig. 1A and Table 1). The sample consisted of biofilm with planktonic cells in the associated fluid.

CN–SCN reactor

A second 1 l reactor was inoculated with biomass from the SCN-only reactor and operated at a residence time of 14 h degrading 0.12 mmol h^{-1} SCN^- and 0.5 mmol h^{-1} CN^-. Following steady-state operation for multiple residence times, the loadings were gradually increased to 0.92 mmol h^{-1} SCN^- and 0.14 mmol h^{-1} CN^- respectively. Reactor temperature was decreased step-wise from 25°C to 18°C to 15°C, at which point (day 190) a reactor sample, containing biofilm and planktonic cells in associated fluid, was collected for metagenomic analysis (Fig. 1B and Table 1).

Reactor effluents were monitored for SCN^-, CN^- and NH_4^+ via high performance liquid chromatography (HPLC) and cyanide probe, further described in Appendix S1.

DNA extraction and sequencing

Genomic DNA was extracted from the samples using the High Pure PCR Template Preparation Kit (Roche Applied Sciences) with the following modifications: samples were mixed with 200 µl tissue lysis buffer, vortexed and stored at −20°C overnight. Upon thawing, 50 µl of proteinase K and 250 µl of binding buffer were added, and extraction proceeded via the manufacturer's protocol. Illumina library preparation and sequencing were performed at the UC Berkeley Vincent J. Coates Genomics Sequencing Laboratory (Berkeley, CA), using an insert size of 500 bp and read length of 100 bp. Sequencing results are shown in Table 1.

Read processing, assembly and annotation

Both read sets were trimmed for quality using Sickle with default parameters (https://github.com/najoshi/sickle) and assembled independently using idba_ud with default parameters (Peng *et al*., 2012). To improve assemblies of high-abundance organisms in the SCN-only sample, two subassemblies were performed with idba_ud using subsets of randomly selected reads representing either 1/5th or 1/20th of the full dataset. For all assemblies, open reading frames (ORFs) were predicted with Prodigal, run in metagenome mode (Hyatt *et al*., 2010; 2012). Annotation was performed using USEARCH (Edgar, 2010) against the Uniref100 (Suzek *et al*., 2007), UniProt and KEGG databases to identify the single best hit and the phylogenetic affiliation of the best hit for each ORF. Uniref100 hits were used to assign a tentative phylogenetic affiliation to each scaffold, to the lowest taxonomic level possible, based on majority rules.

Binning of the metagenomes

Bins were assigned based on the coverage, GC content and phylogenetic best-hit profile of each scaffold ≥ 5 Kbp. For full assemblies, bins were confirmed using emergent self-organizing maps (ESOMs) based on di- and

trinucleotide frequencies, and differential coverage across the two samples (Appendix S1) (Dick *et al.*, 2009; Sharon *et al.*, 2013).

Bin information was superimposed onto the ESOMs as class files (Fig. S1), and bins were checked manually for chimeric and mis-binned scaffolds (using the Databionix GUI, esomana). Chimeras were resolved using paired-read mapping with BOWTIE2 (Langmead and Salzberg, 2012) and manual alignment editing in Geneious (Biomatters) to generate the correct consensus. Some genomes were manually curated using paired-read mapping to resolve assembly errors, extend scaffolds and join scaffolds.

Phylogenetic analysis

The ribosomal proteins (RPs) and 16S rRNA gene sequences were collected from each dataset. Nearest neighbour sequences were acquired from the National Center for Biotechnology Information (NCBI) Genome database, the Joint Genome Institute-Integrated Microbial Genomes database, or for 16S rRNA genes NCBI non-environmental sequences. Genes for 16 RPs (listed in Hug *et al.*, 2013), including those from a custom reference set, were independently aligned with MUSCLE v3.8.31 (Edgar, 2004). The RP alignments were then concatenated and used to construct a maximum likelihood phylogeny (Fig. S2). The 16S rRNA gene sequences and nearest neighbours were aligned with a custom reference set using SSU-ALIGN (Nawrocki, 2009). This alignment was used to generate a maximum likelihood phylogeny (see Appendix S1).

Metabolic analysis

In order to identify metabolic pathways, gene annotations were searched by name. Additionally, sequences of biochemically and/or structurally characterized proteins of interest were used to create databases for BLAST+ (Camacho *et al.*, 2009). The SCN-only and CN–SCN datasets were used as queries with a bit-score cut-off ≥ 60. In order to confirm conservation of active residues in the proteins identified, hits were aligned to reference sequences using MUSCLE (Edgar, 2004) and visualized in Geneious (Biomatters). The web-based tool PSORTB v3.0.2 was used to predict subcellular localization for proteins of interest (Yu *et al.*, 2010).

Access to data online

Sequence data are publicly available through the online database ggKbase at http://genegrabber.berkeley.edu/SCN-stock/organisms and http://genegrabber.berkeley.edu/CN-SCN/organisms. Read datasets are available at NCBI (http://www.ncbi.nlm.nih.gov/) under BioProject ID PRJNA279279 with BioSample identifiers SAMN03445100 (SCN-only) and SAMN03445079 (CN–SCN).

Acknowledgements

Funding was provided by the Joint Genome Institute's Emerging Technologies Opportunity Program, National Science Foundation Graduate Research

Fellowship Program and the South African Research Chairs Initiative (GUN 64778) through the Department of Science and Technology and National Research Foundation, South Africa. A financial contribution from Goldfields SA (Ltd) and Biomin Ltd, supporting research focused on the ASTER process, facilitated this work. We gratefully acknowledge assistance from Andrea Singh, Laura Hug, Chris Brown, Itai Sharon and Anders Norman (University of California, Berkeley), and from Nathan Van Wyk and Rob Huddy (University of Cape Town). The authors declare no conflict of interest.

References

Arakawa, T., Kawano, Y., Kataoka, S., Katayama, Y., Kamiya, N., Yohda, M., and Odaka, M. (2007) Structure of thiocyanate hydrolase: a new nitrile hydratase family protein with a novel five-coordinate cobalt(III) center. *J Mol Biol* 366: 1497– 1509.

Beller, H.R., Chain, P.S.G., Letain, T.E., Chakicherla, A., Larimer, F.W., Richardson, P.M., *et al*. (2006) The genome sequence of the obligately chemolithoautotrophic, facultatively anaerobic bacterium *Thiobacillus denitrificans*. *J Bacteriol* 188: 1473– 1488.

Bezsudnova, E.Y., Sorokin, D.Y., Tikhonova, T.V., and Popov, V.O. (2007) Thiocyanate hydrolase, the primary enzyme initiating thiocyanate degradation in the novel obligately chemolithoautotrophic halophilic sulfur-oxidizing bacterium *Thiohalophilus thiocyanoxidans*. *Biochim Biophys Acta* 1774: 1563– 1570.

Borisov, V.B., Gennis, R.B., Hemp, J., and Verkhovsky, M.I. (2011) The cytochrome *bd* respiratory oxygen reductases. *Biochim Biophys Acta* 1807: 1398– 1413.

Boucabeille, C., Bories, A., Ollivier, P., and Michel, G. (1994) Microbial degradation of metal complexed cyanides and thiocyanate from mining wastewaters. *Environ Pollut* 84: 59– 67.

van Buuren, C., Makhotla, N., and Olivier, J.W. (2011) The ASTER process: technology development through to piloting, demonstration, and commercialization. Proceedings of the ALTA 2011 Nickel-Cobalt-Copper, Uranium and Gold Conference, Perth, Australia.

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., and Madden, T.L. (2009) BLAST+: architecture and applications. *BMC Bioinformatics* 10: 421.

Chen, M.-Y., Tsay, S.-S., Chen, K.-Y., Shi, Y.-C., Lin, Y.-T., and Lin, G.-H. (2002) *Pseudoxanthomonas taiwanensis* sp. nov., a novel thermophilic, N_2O-producing species isolated from hot springs. *Int J Syst Evol Microbiol* 52: 2155– 2161.

Cherney, M.M., Zhang, Y., Solomonson, M., Weiner, J.H., and James, M.N.G. (2010) Crystal structure of sulfide:quinone oxidoreductase from

Acidithiobacillus ferrooxidans: insights into sulfidotrophic respiration and detoxification. *J Mol Biol* 398: 292– 305.

Cipollone, R., Bigotti, M.G., Frangipani, E., Ascenzi, P., and Visca, P. (2004) Characterization of a rhodanese from the cyanogenic bacterium *Pseudomonas aeruginosa*. *Biochem Biophys Res Commun* 325: 85– 90.

Cipollone, R., Ascenzi, P., and Visca, P. (2007) Common themes and variations in the rhodanese superfamily. *IUBMB Life* 59: 51– 59.

Cipollone, R., Frangipani, E., Tiburzi, F., Imperi, F., Ascenzi, P., and Visca, P. (2007) Involvement of *Pseudomonas aeruginosa* rhodanese in protection from cyanide toxicity. *Appl Environ Microbiol* 73: 390– 398.

Cluness, M.J., Turner, P.D., Clements, E., Brown, D.T., and O'Reilly, C. (1993) Purification and properties of cyanide hydratase from *Fusarium lateritium* and analysis of the corresponding *chy1* gene. *J Gen Microbiol* 139: 1807– 1815.

Dick, G.J., Andersson, A.F., Baker, B.J., Simmons, S.L., Thomas, B.C., Yelton, A.P., and Banfield, J.F. (2009) Community-wide analysis of microbial genome sequence signatures. *Genome Biol* 10: R85.

Dictor, M.C., Battaglia-Brunet, F., Morin, D., Bories, A., and Clarens, M. (1997) Biological treatment of gold ore cyanidation wastewater in fixed bed reactors. *Environ Pollut* 97: 287– 294.

Edgar, R.C. (2004) Muscle: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 32: 1792– 1797.

Edgar, R.C. (2010) Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* 26: 2460– 2461.

Erdogan, M.F. (2003) Thiocyanate overload and thyroid disease. *Biofactors* 19: 107– 111.

Felföldi, T., Székely, A.J., Gorál, R., Barkács, K., Scheirich, G., András, J., *et al*. (2010) Polyphasic bacterial community analysis of an aerobic activated sludge removing phenols and thiocyanate from coke plant effluent. *Bioresour Technol* 101: 3406– 3414.

Fernandez, R.F., and Kunz, D.A. (2005) Bacterial cyanide oxygenase is a suite of enzymes catalyzing the scavenging and adventitious utilization of cyanide as a nitrogenous growth substrate. *J Bacteriol* 187: 6396– 6402.

Finkmann, W., Altendorf, K., Stackebrandt, E., and Lipski, A. (2000) Characterization of N_2O-producing Xanthomonas-like isolates from biofilters as *Stenotrophomonas nitritireducens* sp. nov., *Luteimonas mephitis* gen. nov., sp. nov. and *Pseudoxanthomonas broegbernensis* gen. nov., sp. nov. *Int J Syst Evol Microbiol* 50 (Part 1): 273– 282.

Gould, W.D., King, M., Mohapatra, B.R., Cameron, R.A., Kapoor, A., and Koren, D.W. (2012) A critical review on destruction of thiocyanate in mining effluents. *Miner Eng* 34: 38– 47.

Guilloton, M.B., Lamblin, A.F., Kozliak, E.I., Gerami-Nejad, M., Tu, C., Silverman, D., *et al*. (1993) A physiological role for cyanate-induced carbonic anhydrase in *Escherichia coli*. *J Bacteriol* 175: 1443– 1451.

Haft, D.H., Basu, M.K., and Mitchell, D.A. (2010) Expansion of ribosomally produced natural products: a nitrile hydratase- and Nif11-related precursor family. *BMC Biol* 8: 70.

Huddy, R.J., van Zyl, A.W., van Hille, R.P., and Harrison, S.T.L. (2015) Characterisation of the complex microbial community associated with the ASTER™ thiocyanate biodegradation system. *Miner Eng* 76: 65– 71.

Hug, L.A., Beiko, R.G., Rowe, A.R., Richardson, R.E., and Edwards, E.A. (2012) Comparative metagenomics of three Dehalococcoides-containing enrichment cultures: the role of the non-dechlorinating community. *BMC Genomics* 13: 327.

Hug, L.A., Castelle, C.J., Wrighton, K.C., Thomas, B.C., Sharon, I., Frischkorn, K.R., *et al*. (2013) Community genomic analyses constrain the distribution of metabolic traits across the Chloroflexi phylum and indicate roles in sediment carbon cycling. *Microbiome* 1: 22.

Hung, C.-H., and Pavlostathis, S.G. (1999) Kinetics and modeling of autotrophic thiocyanate biodegradation. *Biotechnol Bioeng* 62: 1– 11.

Hussain, A., Ogawa, T., Saito, M., Sekine, T., Nameki, M., Matsushita, Y., *et al*. (2013) Cloning and expression of a gene encoding a novel thermostable thiocyanate-degrading enzyme from a mesophilic alphaproteobacteria strain THI201. *Microbiology* 159: 2294– 2302.

Hyatt, D., Chen, G.-L., Locascio, P.F., Land, M.L., Larimer, F.W., and Hauser, L.J. (2010) Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* 11: 119.

Hyatt, D., Locascio, P.F., Hauser, L.J., and Uberbacher, E.C. (2012) Gene and translation initiation site prediction in metagenomic sequences. *Bioinformatics* 28: 2223– 2230.

Jandhyala, D., Berman, M., Meyers, P.R., Sewell, B.T., Willson, R.C., and Benedik, M.J. (2003) CynD, the cyanide dihydratase from *Bacillus pumilus*: gene cloning and structural studies. *Appl Environ Microbiol* 69: 4794– 4805.

Jandhyala, D.M., Willson, R.C., Sewell, B.T., and Benedik, M.J. (2005) Comparison of cyanide-degrading nitrilases. *Appl Microbiol Biotechnol* 68: 327– 335.

Jensen, N.B., Zagrobelny, M., Hjernø, K., Olsen, C.E., Houghton-Larsen, J., Borch, J., *et al*. (2011) Convergent evolution in biosynthesis of cyanogenic defence compounds in plants and insects. *Nat Commun* 2: 273.

Kataoka, S., Arakawa, T., Hori, S., Katayama, Y., Hara, Y., Matsushita, Y., *et al*. (2006) Functional expression of thiocyanate hydrolase is promoted by its activator protein, P15K. *FEBS Lett* 580: 4667– 4672.

Katayama, Y., Matsushita, Y., Kaneko, M., Kondo, M., Mizuno, T., and Nyunoya, H. (1998) Cloning of genes coding for the three subunits of thiocyanate hydrolase of *Thiobacillus thioparus* THI 115 and their evolutionary relationships to nitrile hydratase. *J Bacteriol* 180: 2583– 2589.

Kim, S.-W., Fushinobu, S., Zhou, S., Wakagi, T., and Shoun, H. (2009) Eukaryotic *nirK* genes encoding copper-containing nitrite reductase: originating from the protomitochondrion? *Appl Environ Microbiol* 75: 2652– 2658.

Kraft, B., Tegetmeyer, H.E., Sharma, R., Klotz, M.G., Ferdelman, T.G., Hettich, R.L., *et al*. (2014) The environmental controls that govern the end product of bacterial nitrate respiration. *Science* 345: 676– 679.

Langmead, B., and Salzberg, S.L. (2012) Fast gapped-read alignment with Bowtie 2. *Nat Methods* 9: 357– 359.

Luque-Almagro, V.M., Huertas, M.-J., Sáez, L.P., Luque-Romero, M.M., Moreno Vivián, C., Castillo, F., *et al*. (2008) Characterization of the *Pseudomonas pseudoalcaligenes* CECT5344 Cyanase, an enzyme that is not essential for cyanide assimilation. *Appl Environ Microbiol* 74: 6280– 6288.

Luque-Almagro, V.M., Merchan, F., Blasco, R., Igeno, M.I., Martínez-Luque, M., Moreno-Vivián, C., *et al*. (2011) Cyanide degradation by *Pseudomonas pseudoalcaligenes* CECT5344 involves a malate: quinone oxidoreductase and an associated cyanide-insensitive electron transfer chain. *Microbiology* 157: 739– 746.

Lykidis, A., Chen, C.-L., Tringe, S.G., McHardy, A.C., Copeland, A., Kyrpides, N.C., *et al*. (2010) Multiple syntrophic interactions in a terephthalate-degrading methanogenic consortium. *ISME J* 5: 122– 130.

Marcia, M., Ermler, U., Peng, G., and Michel, H. (2009) The structure of *Aquifex aeolicus* sulfide:quinone oxidoreductase, a basis to understand sulfide detoxification and respiration. *Proc Natl Acad Sci USA* 106: 9625– 9630.

Marcia, M., Ermler, U., Peng, G., and Michel, H. (2010) A new structure-based classification of sulfide:quinone oxidoreductases. *Proteins* 78: 1073– 1083.

Nawrocki, E.P. (2009) Structural RNA homology search and alignment using covariance models. Ph.D. Thesis. St Louis, MO, USA: Washington University, School of Medicine.

Ogawa, T., Noguchi, K., Saito, M., Nagahata, Y., Kato, H., Ohtaki, A., *et al*. (2013) Carbonyl sulfide hydrolase from *Thiobacillus thioparus* strain THI115 is one of the β-carbonic anhydrase family enzymes. *J Am Chem Soc* 135: 3818– 3825.

Pace, H.C., and Brenner, C. (2001) The nitrilase superfamily: classification, structure and function. *Genome Biol* 2: REVIEWS0001.

Peng, Y., Leung, H.C.M., Yiu, S.M., and Chin, F.Y.L. (2012) IDBA-UD: a de novo assembler for single-cell and metagenomic sequencing data with highly uneven depth. *Bioinformatics* 28: 1420– 1428.

du Plessis, C.A., Barnard, P., Muhlbauer, R.M., and Naldrett, K. (2001) Empirical model for the autotrophic biodegradation of thiocyanate in an activated sludge reactor. *Lett Appl Microbiol* 32: 103– 107.

Quan, Z.X., Rhee, S.K., Bae, J.W., Baek, J.H., Park, Y.H., and Lee, S.T. (2006) Bacterial community structure in activated sludge reactors treating free or metal-complexed cyanides. *J Microbiol Biotechnol* 16: 232– 239.

Sharon, I., Morowitz, M.J., Thomas, B.C., Costello, E.K., Relman, D.A., and Banfield, J.F. (2013) Time series community genomics analysis reveals rapid shifts in bacterial species, strains, and phage during infant gut colonization. *Genome Res* 23: 111– 120.

Sorokin, D.Y., Tourova, T.P., Lysenko, A.M., and Kuenen, J.G. (2001) Microbial thiocyanate utilization under highly alkaline conditions. *Appl Environ Microbiol* 67: 528– 538.

Sorokin, D.Y., Tourova, T.P., Antipov, A.N., Muyzer, G., and Kuenen, J.G. (2004) Anaerobic growth of the haloalkaliphilic denitrifying sulfur-oxidizing bacterium *Thialkalivibrio thiocyanodenitrificans* sp. nov. with thiocyanate. *Microbiology* 150: 2435– 2442.

Sorokin, D.Y., Tourova, T.P., Bezsoudnova, E.Y., Pol, A., and Muyzer, G. (2007) Denitrification in a binary culture and thiocyanate metabolism in *Thiohalophilus thiocyanoxidans* gen. nov. sp. nov. – a moderately halophilic chemolithoautotrophic sulfur-oxidizing Gammaproteobacterium from hypersaline lakes. *Arch Microbiol* 187: 441– 450.

Sorokin, D.Y., Muntyan, M.S., Panteleeva, A.N., and Muyzer, G. (2012) *Thioalkalivibrio sulfidiphilus* sp. nov., a haloalkaliphilic, sulfur-oxidizing gammaproteobacterium from alkaline habitats. *Int J Syst Evol Microbiol* 62: 1884– 1889.

Speyer, M.R., and Raymond, P. (1988) The acute toxicity of thiocyanate and cyanate to rainbow trout as modified by water temperature and pH. *Environ Toxicol Chem* 7: 565– 571.

Stott, M.B., Franzmann, P.D., Zappia, L.R., Watling, H.R., Quan, L.P., Clark, B.J., *et al*. (2001) Thiocyanate removal from saline CIP process water by a rotating biological contactor, with reuse of the water for bioleaching. *Hydrometallurgy* 62: 93– 105.

Stratford, J., Dias, A.E., and Knowles, C.J. (1994) The utilization of thiocyanate as a nitrogen source by a heterotrophic bacterium: the degradative pathway

involves formation of ammonia and tetrathionate. *Microbiology* 140: 2657– 2662.

Sung, Y.C., and Fuchs, J.A. (1988) Characterization of the *cyn* operon in *Escherichia coli* K12. *J Biol Chem* 263: 14769– 14775.

Sung, Y.C., and Fuchs, J.A. (1992) The *Escherichia coli* K-12 *cyn* operon is positively regulated by a member of the *lysR* family. *J Bacteriol* 174: 3645– 3650.

Suzek, B.E., Huang, H., McGarvey, P., Mazumder, R., and Wu, C.H. (2007) UniRef: comprehensive and non-redundant UniProt reference clusters. *Bioinformatics* 23: 1282– 1288.

Thuku, R.N., Brady, D., Benedik, M.J., and Sewell, B.T. (2009) Microbial nitrilases: versatile, spiral forming, industrial enzymes. *J Appl Microbiol* 106: 703– 727.

Vu, H.P., Mu, A., and Moreau, J.W. (2013) Biodegradation of thiocyanate by a novel strain of *Burkholderia phytofirmans* from soil contaminated by gold mine tailings. *Lett Appl Microbiol* 57: 368– 372.

Walsh, M.A., Otwinowski, Z., Perrakis, A., Anderson, P.M., and Joachimiak, A. (2000) Structure of cyanase reveals that a novel dimeric and decameric arrangement of subunits is required for formation of the enzyme active site. *Structure* 8: 505– 514.

Wang, P., and VanEtten, H.D. (1992) Cloning and properties of a cyanide hydratase gene from the phytopathogenic fungus *Gloeocercospora sorghi*. *Biochem Biophys Res Commun* 187: 1048– 1054.

Watson, S.J., and Maly, E.J. (1987) Thiocyanate toxicity to *Daphnia magna*: modified by pH and temperature. *Aquat Toxicol* 10: 1– 8.

Wood, A.P., Kelly, D.P., McDonald, I.R., Jordan, S.L., Morgan, T.D., Khan, S., *et al*. (1998) A novel pink-pigmented facultative methylotroph, *Methylobacterium thiocyanatum* sp. nov., capable of growth on thiocyanate or cyanate as sole nitrogen sources. *Arch Microbiol* 169: 148– 158.

Wu, J.-H., Wu, F.-Y., Chuang, H.-P., Chen, W.-Y., Huang, H.-J., Chen, S.-H., and Liu, W.-T. (2012) Community and proteomic analysis of methanogenic consortia degrading terephthalate. *Appl Environ Microbiol* 79: 105– 112.

Yamanaka, Y., Arakawa, T., Watanabe, T., Namima, S., Sato, M., Hori, S., *et al*. (2013) Two arginine residues in the substrate pocket predominantly control the substrate selectivity of thiocyanate hydrolase. *J Biosci Bioeng* 116: 22– 27.

Yu, N.Y., Wagner, J.R., Laird, M.R., Melli, G., Rey, S., Lo, R., *et al*. (2010) PSORTb 3.0: improved protein subcellular localization prediction with refined localization subcategories and predictive capabilities for all prokaryotes. *Bioinformatics* 26: 1608– 1615.

van Zyl, A.W., Harrison, S.T.L., and van Hille, R.P. (2011) Biodegradation of thiocyanate by a mixed microbial population. Proceedings of the International Mine Water Association Conference, Aachen, Germany, 2011, pp. 119–124.

van Zyl, A.W., Huddy, R.J., Harrison, S.T.L., and van Hille, R.P. (2015) Evaluation of the ASTER™ process in the presence of suspended solids. *Miner Eng* 76: 72– 80.