"*Candidatus* Siderophilus nitratireducens": a psychrophilic, *nap*-dependent nitrate-reducing iron oxidizer within the new order Siderophiliales

Francesc Corbera-Rubio[a], Gerben R. Stouten[a], Jantinus Bruins[b], Simon F. Dost[c], Alexander Y. Merkel[d], Simon Müller[a], Mark C. M. van Loosdrecht[a], Doris van Halem[a], Michele Laureni[a,*]

* corresponding author: m.laureni@tudelft.nl

[a] Delft University of Technology, Delft, the Netherlands

[b] WLN, Glimmen, the Netherlands

[c] WMD Water company Drenthe, Assen, the Netherlands

[d] Winogradsky Institute of Microbiology, Research Center of Biotechnology, Russian Academy of Sciences, Moscow, Russia

Competing Interests

The authors declare no competing financial interests

Abstract

Nitrate leaching from agricultural soils is increasingly found in groundwater, a primary source of drinking water worldwide. This nitrate influx can potentially stimulate the biological oxidation of iron in anoxic groundwater reservoirs. Nitrate-reducing iron-oxidizing (NRFO) bacteria have been extensively studied in laboratory settings, yet their ecophysiology in natural environments remains largely unknown. To this end, we established a pilot-scale filter on nitrate-rich groundwater to elucidate the structure and metabolism of nitrate-reducing iron-oxidizing microbiomes under oligotrophic conditions mimicking natural groundwaters. The enriched community stoichiometrically removed iron and nitrate consistently with NRFO metabolism. Genome-resolved metagenomics revealed the underlying metabolic network between the dominant iron-dependent denitrifying autotrophs and the less abundant organoheterotrophs. The most abundant genome belonged to a new *Candidate* order, named Siderophiliales. This new species, "*Candidatus* Siderophilus nitratireducens", carries central genes to iron oxidation (cytochrome *c cyc2*), carbon fixation (*rbc*), and for the sole periplasmic nitrate reductase (*nap*). To our knowledge, this is the first report of *nap*-based lithoautotrophic growth, and we demonstrate that iron oxidation coupled to dissimilatory reduction of nitrate to nitrite is thermodynamically favourable under realistic Fe^{3+}/Fe^{2+} and NO_3^-/NO_2^- concentration ratios. Ultimately, by bridging the gap between laboratory investigations and real-world conditions, this study provides insights into the intricate interplay between nitrate and iron in groundwater ecosystems, and expands our understanding of NRFOs taxonomic diversity and ecological role.

1. Introduction

Globally, approximately one third of the nitrogen applied to agricultural soils is lost via leaching to the surrounding waterbodies (1). This has led to elevated nitrate (NO_3^-) levels in anoxic groundwaters, a primary source of drinking water worldwide (2). Owing to population growth and agriculture intensification, nitrate concentrations in subsurface waters are expected to continue increasing (3). Besides its direct impact on human health (4), nitrate can significantly alter the biogeochemistry of groundwater reservoirs (5). Nitrate promotes the oxidation of sulfide and in particular of iron (Fe) – the most prevalent groundwater contaminant – leading to the mobilisation of heavy metals and the emission of greenhouse gases (6). Despite these implications, the consequences of nitrate-iron interactions on ecosystems and drinking water production systems remain largely unexplored. A detailed understanding of the underlying principles is paramount for anticipating and mitigating current and future challenges, as well as for exploring potential synergies and biotechnological opportunities.

Nitrate-reducing iron-oxidizing (NRFO) bacteria couple the anoxic reduction of nitrate to the oxidation of Fe^{2+} (eq.1). Since their discovery in 1996 by Straub et al. (1996), NRFO microorganisms have been the focus of extensive research both in pure and mixed cultures (reviewed in (8)), and several complete genomes are already publicly available (9,10). The metabolic versatility of NRFO bacteria spans from lithoautotrophic to mixotrophic growth (11), to partial denitrification using nitric oxide (NO) (12) and nitrous oxide (N_2O) (9) as terminal electron acceptors. At the same time, due to the inherently low energetic yield of iron oxidation, NRFO bacteria live close to the thermodynamic edge (13). Their fitness is highly dependent on environmental factors such as substrate and product availability, pH and temperature (14), and chemical reactions - such as the quasi-instantaneous precipitation of the biologically formed Fe^{3+} - can play a pivotal role by modulating iron and nitrogen concentrations (15). However, our current understanding is largely based on laboratory settings, and does not necessarily reflect the complexity of natural and engineered

ecosystems where several (a)biotic reactions occur simultaneously at temperatures significantly lower than tested to date (16).

$$10Fe^{2+} + 2NO_3^- + 24\,H_2O \rightarrow 10Fe(OH)_3 + N_2 + 18H^+ \qquad \text{eq. 1}$$

To address these knowledge gaps, we established a pilot-scale filter on anoxic groundwater containing both Fe^{2+} and NO_3^-. The emulated groundwater conditions allowed for the establishment of a microbial enrichment that simultaneously removed Fe^{2+} and NO_3^-. In depth metagenomic analysis of the steady-state community revealed a new order-level NRFO lineage, deepening our understanding of their taxonomic diversity and ecological roles. Overall, our study bridges the gap between laboratory studies and real-world conditions, and offers a nuanced view on the intricate interplay between nitrate and iron in groundwater ecosystems.

2. Results

2.1. Nitrate-dependent iron removal irrespective of the limiting nutrient

Stable anoxic nitrate and iron removal was achieved after less than three weeks of operation and maintained for over 100 days in a pilot-scale filter fed with nitrate-rich anoxic groundwater (Figure S3). With nitrate as the limiting nutrient at both groundwater (8.9 ± 2.8 µM) and nitrate-amended concentrations (13.5 ± 1.5 and 20.2 ± 2.4 µM), effluent nitrate concentrations were consistently below detection limit (1 µM). Throughout the nitrate-limiting period, NO_3^- and Fe^{2+} were consumed at a 7.1 ± 1.4 Fe^{2+}:NO_3^- molar ratio (Figure 1Figure 1). Oxygen was always below the quantification limit of 3 µM, and roughly 80 µC-mol dissolved organic carbon (DOC·l^{-1}) was consistently removed from the influent, likely due to the formation of Fe^{2+}-DOC complexes formation owing to the non-biodegradable nature of organic matter in Dutch groundwater (17). Ammonia consumption was negligible (<0.1 µM). Effluent nitrite concentrations stayed below the detection limit (<0.2 µM), while other denitrification intermediates – nitric oxide and nitrous oxide – were not measured. The observed consistent stoichiometric coupling between nitrate and iron removals strongly suggests Fe^{2+} oxidation to be primarily driven by microbial nitrate-reducing iron oxidation.

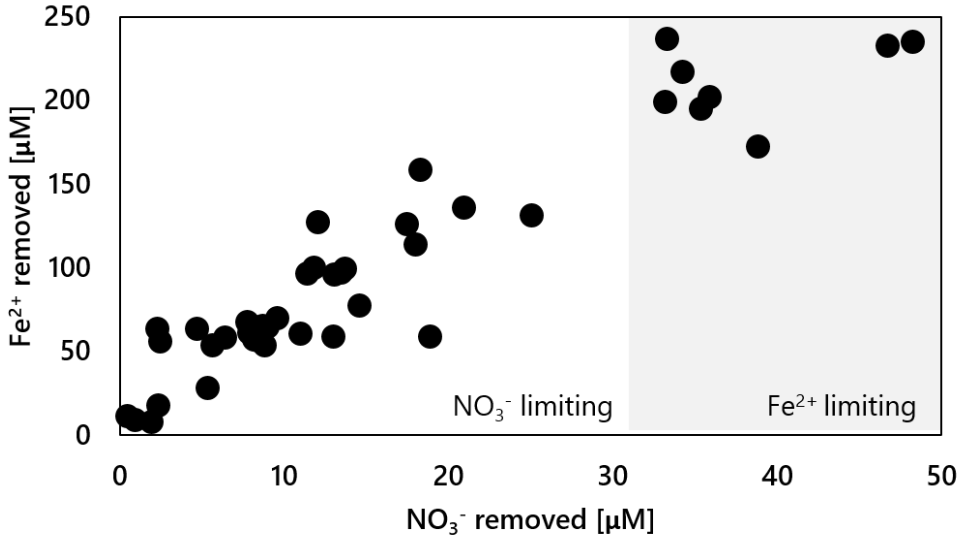

Figure 1. Simultaneous NO_3^- and Fe^{2+} removals in the groundwater-fed pilot-scale filter during the 120 days of continuous operation. The groundwater Fe^{2+} concentration was constant throughout the experiment (236 ± 4 µM). NO_3^- was dosed in the influent to step-wise increase the natural groundwater concentration from 8.1 ± 2.1 to 20.2 ± 2.4 µM (NO_3^- limitation), and up to 83.8 ± 0.6 µM (Fe^{2+} limitation).

2.2. Microbial community dominated by iron oxidizers and denitrifiers

Metagenomic DNA sequencing yielded a total of 107,512 and 8,754,261 quality filtered short and long reads, respectively. After assembly and polishing, this resulted in 19,127 contigs with an N50 value of 15,927. Contigs binning resulted in 13 high and medium quality metagenome assembled genomes (MAGs) as defined by (18) with a relative abundance exceeding 0.5% of the quality filtered long reads. Collectively, these 13 most abundant genomes accounted for 66.9% of the total quality filtered reads, and belonged to four phyla: *Proteobacteria* (51.6%), *Actinobacteria* (8.3%), *Bacteroidetes* (5.6%) and *Chloroflexi* (1.4%) (Table 1). Every genome in the community contained at least one denitrification enzyme genes, while five featured the genetic potential for iron oxidation. Notably, all putative iron oxidizers also possessed the genetic repertoire for carbon fixation. The most abundant MAG (MAG.13), accounting for 19.3% of the community, could only be taxonomically classified at class level (*Gammaproteobacteria*). Given its high abundance and potential metabolic relevance, the taxonomy and metabolic potential of MAG.13 was further investigated (Table S2).

Table 1. General characteristics of the MAGs recovered from the pilot-scale filter. The last three columns indicate the presence/absence of the essential genes for iron oxidation, denitrification and carbon fixation via the reductive phosphate pentose phosphate cycle. Cov_il and cov_np are the coverage with Illumina short reads and Nanopore long reads, respectively.

MAG ID	Phyloegnetic affiliation	Phylum	Relative abundance (%)	Scaffolds	N50	GC (%)	Completeness (%)	Contamination (%)	cov_Il	cov_np	Coding genes	MAG quality	Iron oxidation[1]	Denitrification[2]	Carbon fixation[3]
MAG.13	** "*Candidatus* Siderophilus nitratireducens"	Proteobacteria	19.3	9	1,897,007	56	98.59	1.73	298.27	300.86	3217	High	Yes	NO_3^+ -> NO_2^-	Yes
MAG.26	f__Gallionellaceae	Proteobacteria	10.0	10	569,264	55	95.81	0.95	234.92	155.81	3749	High	Yes	NO_3^+ -> NO ; N_2O -> N_2	Yes
MAG.18	f__Anaeromyxobacteraceae	Proteobacteria	8.6	21	518,414	73	98.06	0.65	186.12	131.88	4018	High	No	N_2O -> N_2	No
MAG.19	g__*Devosia*	Proteobacteria	6.5	27	4,138,465	64	99.28	0.14	24.53	100.33	4029	High	No	NO_2^- -> NO	No
MAG.10	f__Chitinophagaceae	Bacteroidetes	5.6	4	3,646,631	40	97.54	0.9	304.83	87.01	3193	High	No	NO -> N_2	No
MAG.03	o__Nanopelagicales	Actinobactiria	4.3	11	779,605	69	99	2.37	140.41	67.49	4065	High	No	NO_3^- -> NO	No
MAG.27	o__Nanopelagicales	Actinobactiria	3.9	9	1,166,278	71	98.3	1.98	140.56	60.77	3910	High	No	NO_3^- -> NO	No
MAG.34	g__*Gallionella*	Proteobacteria	3.5	112	46,607	57	95.91	3.25	64.39	52.65	2516	High	Yes	NO_3^- -> NO	Yes
MAG.00	o__Anaerolineales	Chloroflexi	1.4	100	114,338	55	97.27	2.91	21.93	20.04	4828	High	No	NO_2^- -> NO	No
MAG.08	g__*Rhodoferax*	Proteobacteria	1.2	121	69,171	63	88.95	2.37	33.64	17.42	3870	Medium	No	NO_3^- -> NO	Yes
MAG.04	f__Rhizobiaceae	Proteobacteria	0.9	115	80,809	60	95.12	4.1	25.97	13.93	4768	High	No	NO_3^- -> N_2O	No
MAG.16	f__Gallionellaceae	Proteobacteria	0.9	165	27,226	56	88.51	3.94	52.76	13.55	3002	Medium	Yes	NO_3^- -> N_2O	Yes
MAG.29	g__*Rhizobacter*	Proteobacteria	0.7	257	25,511	66	94.54	4.54	21.73	9.94	4225	Medium	Yes	NO_3^+ -> NO_2^-	Yes

[1]*Based on FeGenie* (54). *'Yes' if either cyc*2, *cyc*1, *fox*ABC, *fox*EYZ, *sulfocyanin*, *pio*ABC or *mto*AB *are present*

[2]*Based on GhostKOALA* (53) *and the presence of nap*AB or *nar*GHI (K02567, K02568 or K00370, K00371, K00374; $NO_3^- \rightarrow NO_2^-$), *nir*K or *nir*S (K00368, K15864; $NO_2^- \rightarrow NO$), *nor*BC (K04561, K02305; $NO \rightarrow N_2O$), *nos*Z (K00376; $N_2O \rightarrow N_2$) from denitrification (M00529).

[3]*Based on GhostKOALA* (53). *'Yes' if either rbcL* (K01601) *or rbcS* (K01602) *and prk*B (K00855) *from reductive phosphate pentose phosphate cycle (Calvin Cycle) (*M00165).

*** Initially classified as c_Gammaproteobacteria.*

2.3. *"Candidatus* Siderophilus *nitratireducens" represents a new order within Gammaproteobacteria*

Our phylogenomic analysis based on the concatenated amino acid sequences of 120 bacterial single copy conservative marker genes revealed that MAG.13 belongs to a bacterium forming a new order-level lineage Ga0077554 (GTDB release 08-RS214) within the class *Gammaproteobacteria,* with no known closely related pure-culture representatives (Figure 2). We propose to name the new species "*Candidatus* Siderophilus nitratireducens" gen.nov., sp.nov., a member of the *Candidate* order and family Siderophiliales and Siderophiliaceae, respectively. This lineage, along with several other MAGs from similar groundwater habitats (19), is mostly related to lineages including lithoautotrophic sulfur-ozidizing bacteria from the genera *Sulfuriflexus*, *Thioalbus* and the members of the order *Thiohalomonadales*, including *Thiohalomonas*, *Sulfurivermis* and *Thiohalophilus*. Average nucleotide identity and DNA-DNA hybridization comparison between *Ca.* Siderophilus nitratireducens and its closest relative, a MAG from a drinking water treatment plant (GCA_001464965.1), indicated that the two organisms belong to the same genus but different species (ANI = 89%, DDH = 36.6%).

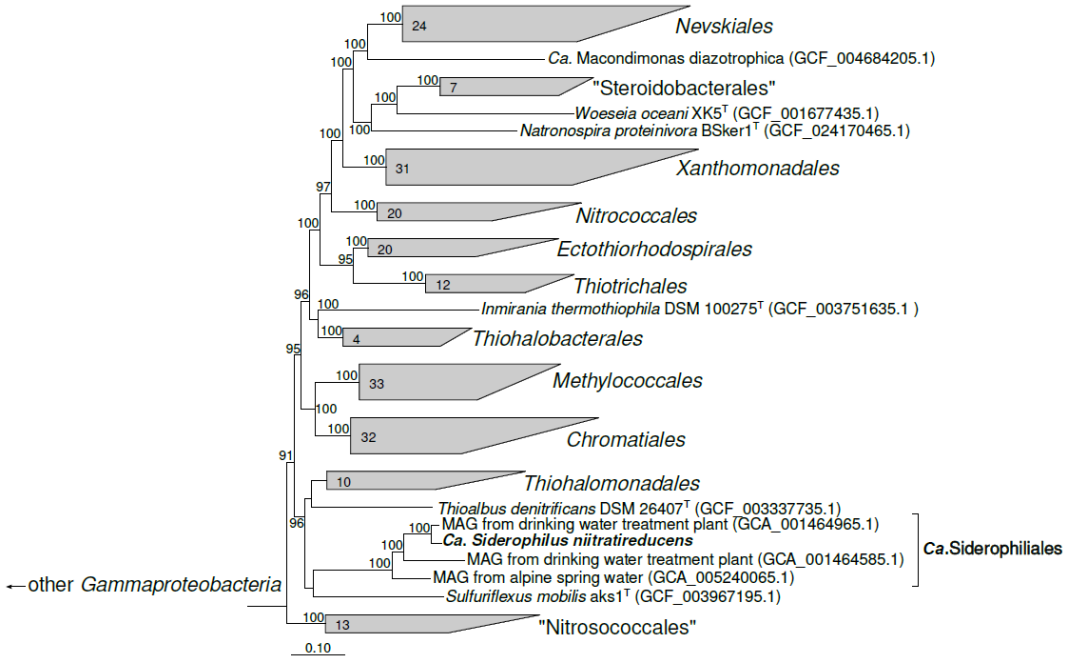

*Figure 2. Phylogenetic position of the "*Ca. Siderophilus *niitratireducens" based on sequence analyses of concatenated alignment of 120 single-copy conserved bacterial protein markers* (49) *- taxonomic designations correspond to the Genome Taxonomy DataBase 207). The trees were built using IQ-TREE2* (ref. 23) *with approximate likelihood-ratio test for branches* (62). *Bootstrap consensus tree is shown with values above 90% placed at the nodes. Bar, 0.10 changes per position.*

2.4. *Autotrophy in "Ca.* Siderophilus nitratireducens*"*

To resolve the main anabolic and catabolic pathways of "*Ca*. Siderophilus nitratireducens", open reading frames (ORF) were predicted and annotated (Table 3, detailed version in Table S1). The genome contains marker genes coding for two key proteins of autotrophic CO_2 fixation via the reductive pentose phosphate (Calvin–Benson–Bassham; CBB) cycle, including the large and small subunits of the ribulose-1,5-bisphosphate carboxylase-oxygenase (*rbc*LS form I) and the phosphoribulokinase (*prk*). Genes encoding for carboxysomal shell proteins and carbonic anhydrase were also present, further supporting the inorganic

carbon uptake ability of "*Ca.* Siderophilus nitratireducens". The absence of phosphofructokinase (*pfk*) indicates a modified glycolytic pathway initiating at the glyceraldehyde 3-phosphate level. All tricarboxylic acid (TCA) cycle genes were identified except for fumarate hydratase (*fh*). However, the glyoxylate shunt enzymes malate synthase (*glc*B) and isocitrate lyase (*ace*A) were present. Taken together, these findings suggest the capability for full autotrophic growth of "*Ca.* Siderophilus nitratireducens".

2.5. Iron oxidation in "Ca. Siderophilus *nitratireducens"*

The presence of a monoheme *c* cytochrome *cyc2*, a primary iron oxidation gene, suggests that "*Ca.* Siderophilus nitratireducens" can use Fe^{2+} as an electron donor. Other common Fe^{2+} oxidases, namely the diheme *c* cytochrome cyc1 and the multiheme c cytochromes MtoA and MtoB, were not annotated. Despite the close phylogenetic proximity to lithoautotrophic sulfur-oxidizing bacteria, the genes of sulfide dehydrogenases Sqr and FccAB and sulfite dehydrogenases SorAB and SoeABC were not identified.

In terms of potential catabolic electron acceptors, the genes for a periplasmic nitrate reductase (*nap*ABCD and its membrane ferrodoxins *nap*GH) and a cbb_3-type cytochrome c oxidase (*cco*NOP) were annotated. However, genes encoding for other known denitrification reductases, namely membrane-bound nitrate reductase (*nar*GHI) and nitrite, nitric oxide and nitrous oxide reductases, *nir*K/*nir*S, *nor*BC and *nos*Z respectively, were not found (Table 3, detailed version in Table S2). Additionally, alternative oxidases, such as the cytochrome *bd* ubiquinol oxidase (*cyd*AB) or the aa_3-type cytochrome c oxidase (*cox*ABCD) were not identified. Also, genes of dissimilatory sulfate reduction (*apr*AB and *dsr*ABC) and the *sox* complex, responsible for sulfate reduction could not be identified. These findings suggest that "*Ca.* Siderophilus *nitratireducens"* relies exclusively on nitrate and oxygen as electron acceptors.

171 *Table 2. Key enzyme of the main catabolic and anabolic pathways of "Ca.* Siderophilus *nitratireducens*". "+" and "-" indicate
172 presence or absence in the genome. CBB, Calvin-Benson-Bassham cycle; TCA, Tricarboxylic acid cycle/Krebs cycle.

Electron acceptor			Electron donor			Carbon fixation			Carbon metabolism		
Element	*Key genes*	*Presence*	*Element*	*Key genes*	*Presence*	*Pathway*	*Key genes*	*Presence*	*Pathway*	*Key genes*	*Presence*
Nitrogen	*nap* ABCD	+	Iron	*cyc2*	+	CBB	*rbc* LS form I	+	Gycolysis	*pfk*	-
	nar GHI	-		*cyc1*	-		*prk*	+		*gpml*	+
	nir KS/*nor* BC/*nos* Z	-		*mto* AB	-		Carboxysome	+	TCA cycle	*cs / gltA*	+
Sulfur	*apr* AB	-	Sulfur	*sqr*	(+)					*suc* AB	+
	sox XABYZ(CD)	-		*fcc* AB	-					*fh*	-
Oxygen	*cco* NOP	+								*mdh*	+
	qox ABCD	-							Glyxoylate shunt	*glc* B	+
	cyd ABX	-								*ace* A	+

173
174

175 ## *2.6. Complete denitrification: a collaborative effort of iron-oxidizing autotrophs and*
176 ## *organoheterotrophs*

177 Iron oxidation genes were identified in five MAGs, namely MAG.13 ("*Ca.* Siderophilus nitratireducens"),
178 MAG.29 (g_*Rhizobacter*) and MAG.26, MAG.34 and MAG.16 (f_*Gallionellaceae,* commonly associated with
179 autotrophic iron oxidation). These MAGs also encoded for the central enzymes of the carbon dioxide
180 fixation via the CBB cycle. While all 13 MAGs contained genes encoding for at least one denitrification
181 enzyme (Table 2), none of them possessed a comprehensive gene set to fully reduce nitrate to dinitrogen
182 gas. Dissimilatory nitrate reduction to nitrite, the first denitrification step, was present in 9 MAGs, while
183 the final step, nitrous oxide reduction to nitrogen gas, was only found in MAG.26 (g_*Rhizobacter*), MAG.18
184 (f_*Anaeromyxobacteraceae*) and MAG.10 (f_*Chitinophagaceae*). The five most abundant MAGs (> 5 %)
185 alone accounted for up to 50 % of the community and covered the full denitrification process (Figure 3).
186 Interestingly, two distinct potential niches were identified. The autotrophic iron oxidizers, *"Ca.* Siderophilus
187 *nitratireducens*" and MAG.26 (f_*Gallionelleceae*), performed the initial denitrification reductions, while the
188 lower-abundant organoheterotrophs complemented the reduction of (at least) NO to nitrous oxide and,
189 possibly took advantage from the autotrophically fixed carbon excreted by the iron oxidizers. Noticeably,

due to the absence of sufficient biodegradable organic matter in the influent, a portion of the biologically generated NO must have been chemically reduced to N_2O with Fe^{2+}(20).

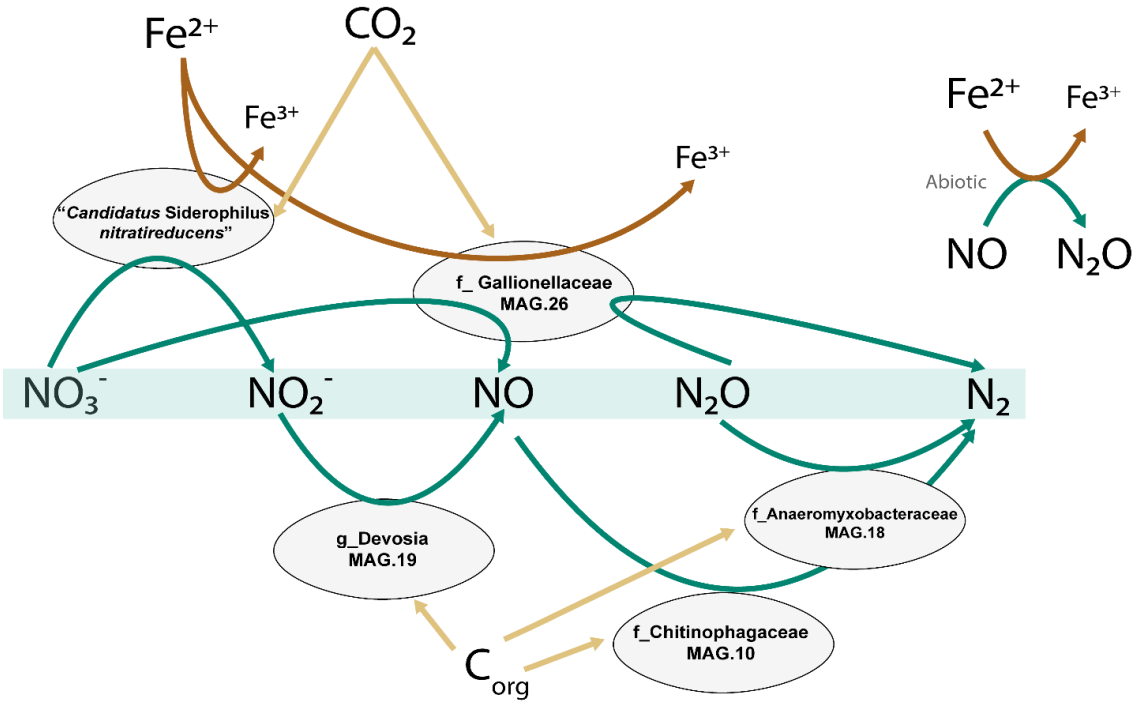

Figure 3. Genome-based conceptual model of substrates fluxes within the microbial community represented by the five most abundant MAGs. Putative autotrophic iron oxidizers perform the upstream part of denitrification, while flanking communities reduce the toxic intermediates to innocuous dinitrogen gas. The putative autotrophic metabolism was inferred based on the presence of ribulose-1,5-biphosphate carboxylase/oxygenase (RuBisCO) and phosphofructokinase (pfk).

3. Discussion

We established a pilot-scale filter on nitrate-rich groundwater to elucidate the structure and metabolism of nitrate-reducing iron-oxidizing microbial communities under oligotrophic conditions mimicking natural groundwater. The enriched community stoichiometrically removed iron and nitrate, and was dominated by a genome belonging to a new *Candidate* order, named Siderophiliales. The genome of this new species, "*Ca.* Siderophilus *nitratireducens*", encoded the genes for iron oxidation (cytochrome *c cyc2*) and the periplasmic nitrate reductase (*nap*), thereby supporting the hypothesis of nitrate-driven iron oxidation being its primary metabolism under the restricted availability of alternative substrates. The absence of other denitrification genes suggests a short catabolic path, which might offer a kinetic advantage in terms of higher iron oxidation rates (21) especially under N limitation (22). "*Ca.* Siderophilus *nitratireducens*" was identified as a putative-autotroph, adding the additional challenge of energy and electrons needs for anabolic CO_2 fixation to the growth on iron, a weak electron-donor at standard conditions (14). However, thermodynamic evaluations do suggest, that *nap*-dependent iron oxidation can sustain growth, given realistic Fe^{3+}/Fe^{2+} and NO_3^-/NO_2^- concentration ratios (SI 4). An essential factor in this context is the quasi-instantaneous precipitation of the biologically formed Fe^{3+} as iron oxides under circum-neutral pH (23). Biochemically, by accepting electrons from the quinol pool, *nap* consumes cytoplasmic protons effectively generating a proton gradient. While the specific mechanism by which this thermodynamic potential is harnessed for carbon fixation remains to be fully elucidated, to the best of our knowledge this is the first report of *nap*-based lithoautotrophic growth. The presence of the sole *nap* also broadens the biochemical diversity of NRFOs currently believed to rely exclusively on the membrane-bound nitrate reductase (*nar*) (9,24–27).

The subsequent reduction of the produced nitrite resulted from the concerted activity of putative autotrophic iron-oxidizers and organoheterotrophs. Within the microbial community, the second most abundant genome, MAG.26 (f_*Gallionellaceaea*), featured the genetic potential for iron oxidation and most

222 denitrification steps, with the exception of nitrous oxide reductase (*nor*). Interestingly, this genome
223 contained genes for CO_2 fixation, a trait mirrored in all other less abundant genomes with the ability to
224 oxidize iron. This suggests that autotrophy may represent an essential trait for NRFOs in anoxic
225 groundwaters where the dissolved organic carbon is largely non-biodegradable (17). The three second most
226 abundant genomes, MAG.18 (f_*Anaeromyxobacteraceae*), MAG.19 (g_*Devosia*) and MAG.10
227 (f_*Chitinophagaceae*) were found to lack the genes for iron oxidation and CO_2 assimilation. Yet, these
228 genomes encompassed the full denitrification pathway starting from nitrite. Besides the likely occurrence
229 of chemical reduction of NO to N_2O (28), we speculate that these heterotrophs complemented the NRFOs
230 for at least the reduction of NO using autotrophically fixed organic carbon as substrate. A similar metabolic
231 network was also recently observed in mesophilic NRFO communities (12). Overall, the measured iron and
232 nitrate consumption yield of 7.1 mol Fe^{2+}: mol NO_3^- is consistent with the expected 5.6, *i.e.* considering the
233 theoretical catabolism (eq.1) and the recently estimated 12 % of electrons used for growth (29), but higher
234 than the experimentally observed range of 3.8 - 4.7 (7,24,30). At first, we hypothesized nitrate
235 ammonification to be the reason for the slight excess in iron oxidation, yet none of the putative iron-
236 oxidizing genomes encoded for the common *nrf* nor for the newly reported octaheme complex (31). Also,
237 the oxygen sporadically detected in the influent was always significantly below the quantification limit of 3
238 µM, a conservative concentration that alone would explain less than 15 % of the total iron consumption via
239 chemical oxidation. As no Fe^{3+} was detected in the reactor effluent, all iron necessarily accumulated inside
240 the reactor either as Fe^{2+} or Fe^{3+} precipitates. X-ray diffraction and Mossbauer spectroscopy identified over
241 94% of the Fe in solids as amorphous ferrihydrite, an Fe^{3+} oxide, with less than 6% of the solids attributed
242 to magnetite, an Fe^{2+}-Fe^{3+} oxide typically formed under anaerobic conditions (SI 5 and 6). Consequently,
243 the Fe^{2+} unaccounted for was likely continuously adsorbed onto the newly-formed Fe^{3+} oxides, a well-
244 studied phenomenon (32), yet the extent to which this occurred was not investigated. In conclusion,

pending experimental validation, we surmise that NRFO microorganisms may not only contribute to iron removal by direct oxidation but also by continuously providing newly-formed iron oxides for its adsorption.

***Description of "Ca.* Siderophilus nitratireducens" gen. nov., sp. nov.**

Siderophilus (Si.de.ro'phi.lus Gr. masc.n. *sidêros* iron; Gr. masc. adj. *philos* loving; N.L. masc. n. *Siderophilus* (loving iron).

ni.tra.ti.re.du'.cens (N.L. masc. n. *nitras (gen. nitratis)*, nitrate; L. pres. part. *reducens*, converting to a different state; N.L. part. adj. *nitratireducens*, reducing nitrate).

Autotrophic nitrate-reducing iron-oxidizing bacterium isolated from a filtration unit fed with anaerobic groundwater with iron(II) and nitrate.

4. Materials and Methods

4.1. Groundwater and pilot-scale filter characteristics

An iron reducing microbial community was enriched anoxically on the granular activated carbon of a 10-L pilot-scale filter in Emmen (the Netherlands) (Figure S1, Figure S2 and Table S1). The media was devoid of any previously formed biofilm. The anoxic, nitrate-rich groundwater (-75,2 ± 28.4 mV) featured constant Fe^{2+} and NO_3^- concentrations, 236 ± 4 µM and 8.1 ± 2.1 µM respectively (Table 3). Oxygen was consistently below quantification limit (3 µM). The groundwater pH and temperature were 6.7 ± 0.2 and 10.5 ± 0.1 °C, respectively. The filter was operated at a filtration flowrate of 3.8 $m \cdot h^{-1}$ during 120 days. After 75 days of steady-state operation, the influent nitrate concentration was manually increased in four steps up to 83.8 ± 0.6 µM, when the system changed from nitrate ($NO_3^- < 1$ µM) to iron limiting conditions ($Fe^{2+} < 4$ µM).

Table 3. Influent and effluent water characteristics corresponding to average and standard deviation of daily measurements of days 21-77, during the nitrate-limited steady-state. Fe^{3+} was calculated as described in the following section.

Parameter	Units	Value	Value
		Influent	***Effluent***
pH	-	6.9 ± 0.4	6.7 ± 0.1
T	˚C	10.5 ± 0.1	10.9 ± 0.6
ORP	mV	-64 ± 17	-58 ± 18
O_2	µmol·L^{-1}	<3*	<3*
NH_4^+	µmol·L^{-1}	11 ± 1.0	11 ± 8.1
NO_2^-	µmol·L^{-1}	<0.2*	<0.2*
NO_3^-	µmol·L^{-1}	8.1 ± 2.1	<1
Fe^{2+}	µmol·L^{-1}	236 ± 4	178 ± 5
Fe^{3+}	µmol·L^{-1}	$2 \cdot 10^{-12}$	-
DOC	mg·L^{-1}	3.1 ± 0.1	2.0 ± 0.2

** Below detection limit*

4.2. Dissolved Fe^{3+} estimation

At pH 7, Fe^{3+} has a markedly low solubility and precipitates as iron oxyhydroxide ($Fe(OH)_3$). Thermodynamically, this phase transition favours the oxidation of Fe^{2+} to Fe^{3+}, and the resulting low Fe^{3+} concentration is the primary driving force of eq. 1 (ref. 36). In our filter, the dissolved concentration of the Fe^{3+}, resulting from Fe^{2+} oxidation, was always below detection limit (0.01mg·l^{-1}). To discuss the thermodynamic feasibility of the NRFO process, we estimated the steady-state [Fe^{3+}]/[Fe^{2+}] ratio following the method proposed by Gorski et al. (2016), which assumes thermodynamic equilibrium between [Fe^{2+}] – [Fe^{3+}] – [FeOx] phases based on the fact that the hydroxylation of dissolved Fe^{3+} is quasi-instantaneous at pH > 3 (ref. 37). Consequently, the following equation can be used to determine the Fe^{3+} concentration as function of pH and the solid solubility constant.

$$\{Fe^{3+}_{(aq)}\}\{OH^-\}^3 = K_{sp}$$

281

282 The most abundant iron oxide in the sand filter was amorphous ferrihydrite (SI 5 and 6), with a K_{sp} of 10^{-39}
283 (35). Therefore:

284 $$\{Fe_{aq}^{3+}\} = \frac{10^{-39}}{(10^{-7.1})^3} = 2 \cdot 10^{-18}\ \mathrm{M}$$

285 *4.3. Analytic procedures*

286 Samples for ammonium, nitrite, and nitrate quantification were immediately filtered through a 0.2 µm
287 nanopore filter and measured within 12 h using photometric analysis (Gallery Discrete Analyzer, Thermo
288 Fischer Scientific, Waltham, Massachusetts, USA). Samples for dissolved iron were filtered through a 0.2
289 µm nanopore filter and quantified by ICP-MS (Analytik Jena, Jena, Germany). Temperature, pH, oxidation-
290 reduction potential (ORP), and dissolved oxygen concentration (DO) were monitored daily using a HI9829-
291 01042 multiparameter analyser (Hanna Instruments, Smithfield, RI, USA) in the raw water, after nitrate
292 dosage and in the effluent.

293

294 *4.4. DNA extraction and quality control*

295 Nucleic acid extraction was carried out using DNeasy PowerSoil Pro Kit (QIAGEN, Hilden, Germnay)
296 following manufacturer instructions. To improve DNA recovery and avoid the interference of carbon with
297 the extraction, 25 µL of 20 $g \cdot l^{-1}$ autoclaved (20 min, 121° Celsius, 2 bar) skimmed milk solution (Sigma
298 Aldrich, Saint Louis, MO, USA) were added to the extraction tube. After extraction, DNA was concentrated
299 to 7.68 $ngDNA \cdot \mu^{-1}$ using Microcon centrifugal filter units YM-100 (MilliporeSigma, Burlington, MA, USA)
300 following the manufacturer's instructions. DNA was quantified with the Qubit 4 Fluorometer and Qubit

dsDNA HS assay kit (Invitrogen, Waltham, MA, USA) following the manufacturer's instructions. DNA purity was determined using a NanoDrop One Spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA).

4.5. Library preparation, sequencing and reads processing

Long-read and short-read DNA sequencing were carried out independently. Long-read library preparation was carried out using the ligation sequencing kit SQK-LSK 109 (Oxford Nanopore Technologies, Oxford, UK). R.9.4.1 flowcells on a GridION were used for sequencing. Raw data was basecalled in super-accurate mode using Guppy v.5.0.16 (https://nanoporetech.com). Raw reads were quality-filtered and trimmed using Filtlong (https://github.com/rrwick/Filtlong) to remove reads below 4000 kb and mean quality score below 80. Adapters were removed using Porechop v.0.2.3 (https://github.com/rrwick/Porechop).

Short-read library preparation was performed using the Nextera XT kit (Illumina, San Diego, CA, U.S.A.) according to the manufacturer's instructions. The libraries were pooled, denatured, and sequenced with Illumina MiSeq (San Diego, CA, USA). Paired end sequencing of 2 x 300 base pairs was performed using the MiSeq Reagent Kit v3 (San Diego, CA, USA) according to the manufacturer's instructions. Raw sequencing data was quality-filtered and trimmed using Trimmomatic v0.39 (HEADCROP:16 LEADING:3 TRAILING:5 SLIDINGWINDOW:4:10 CROP:240 MINLEN:35) (36). Sequencing data quality was analyzed using FastQC v0.11.7 before and after trimming (37).

4.6. Reads assembly and binning

Reads assembly and binning were done as in (38) with minor modifications. Long-reads were assembled using Flye v. 2.9-b1768 (39) with the '–meta' setting enabled and the '–nano-hq' option. Polishing was carried out with Minimap2 v.2.17 (40), Racon v. 1.3.3 (41), Medaka v.1.4.4 (two rounds) (https://github.com/nanoporetech/medaka). At the end, short-reads were incorporated in a final round of polishing with Racon. Both long- and short- raw reads were independently mapped back to the assembled

contigs using BWA-MEM2 (42). SAMtools v1.14 was used to determine contig coverage and for indexing with default settings (43) .

Automated binning was carried out with the long-reads assembly (polished with short-reads) using MetaBAT2 v. 2.12.1 (ref. 46) with '-s 500000', MaxBin2 v. 2.2.7 (ref. 47), and Vamb v. 3.0.2 (ref. 48) with '-o C–minfasta 500000'. Additionally, contig coverage from the short-reads assembly was provided as input to the three binners to improve binning. Output integration and refinement was done in DAS Tool v. 1.1.2 (ref. 49). CoverM v. 0.6.1 (https://github.com/wwood/CoverM) was applied to calculate the bin coverage (using the '-m mean' setting) and the relative abundance ('-m relative_abundance'). Additional manual bin polishing was done in R using mmgenome (https://github.com/MadsAlbertsen/mmgenome).

4.7. Assembly processing and gene annotation

The completeness and contamination of the genome bins were estimated using CheckM v. 1.1.2 (ref. 50). The bins were classified using GDTB-Tk v. 1.5.0 (ref. 22) 202 database. Barrnap v 0.9 (https://github.com/tseemann/barrnap) and structRNAfinder (50) were used to predict 23S, 16S and 5S ribosomal RNA sequences, and transfer RNA sequences were determined using tRNAscan-SE v.20 (51) with *default* search mode. Bins were classified using the Minimum Information about a Metagenome-Assembled Genome (MIMAG) standards (18): high-quality bins were > 90 % complete and < 5 % contaminated, and contained full-length 23S, 16S and 5S ribosomal RNA genes and ≥ 18 transfer RNA genes. Bins with completeness > 50 % and contamination < 10 % were classified as medium-quality, and bins with completeness <50% and <10% contamination as low-quality bins. The remaining ones were considered contamination. 68% of the filtered reads rendered high- and medium-quality bins, 11% low-quality bins and 21.1% were unbinned.

The open reading frames (ORF) of the 10 resulting high-quality and 3 medium-quality bins with relative abundance > 0.5 % were predicted using Prodigal v2.6.3 (ref. 53) and functionally annotated with

346 GhostKoala v2.2 (53) (Kyota Enciclopedia of Genes and Genomes; accessed March 2022). FeGenie (54) was
347 used to improve the annotation of the iron metabolism using the metagenomics ('-meta') settings. To refine
348 the annotation for MAG.13 (*Candidatus* Siderophilus nitratireducens), the genome was uploaded to the
349 National Center for Biotechnology Information (NCBI) database Prokaryotic Genome Annotation Pipeline
350 v6.1 (ref. 54). Additionally, manual annotation of genes potentially relevant but not automatically
351 annotated was done by aligning a set of manually selected sequences from UniProtKB against the translated
352 ORFs from MAG.13 with local blastp v2.13 (ref. 55). After annotation, all the predicted genes of interest
353 (manually and automatically annotated) were translated and aligned against the Non-redundant protein
354 sequences (nr) database from NCBI using blastp (accessed June 2022) and accepted only if the coverage
355 was > 70 % (54) and the identity > 35 % (57).

356 RStudio v1.4.1106 was used for data analysis and visualization.

357 ### *4.8. Phylogenetic tree construction*

358 Genome-based phylogenetic reconstruction was done by using 120 bacterial single copy conservative
359 marker genes, as described previously (58). The trees were built using the IQ-TREE 2 (ref. 23) with fast
360 model selection via ModelFinder (60) and ultrafast approximation for phylogenetic bootstrap (61), as well
361 as approximate likelihood-ratio test for branches (62). Whole genome comparison was conducted by using
362 two different methods: Average Nucleotide Identity (ANI), using JSpeciesWS web server and DNA-DNA
363 Hybridization (DDH) by the Genome-to-Genome Distance Calculator 2.1 online tool
364 (https://ggdc.dsmz.de/ggdc.php) (63).

365

5. Competing interest

The authors declare no competing financial interests

6. Acknowledgements

The authors would like to thank Dimitry Sorokin (TU Delft and Russian Academy of Sciences) for the thorough discussions and invaluable insights, Martin Pabst and Dita Heikens (TU Delft) for the proteomic data, and Mantas Sereika, Thomas Nielsen and Mads Albertsen (Aalborg University) for their help with Nanopore Sequencing and genome assembly and binning.

This work was financed by the NWO partnership program 'Dunea–Vitens: Sand Filtration' (project 17830) of the Dutch Research Council (NWO) and the drinking water companies Vitens NV and Dunea Duin & Water. ML was supported by NWO (VI.Veni.192.252).

7. Data availability statement

Raw reads and MAGs have been deposited in the National Center for Biotechnology Information (NCBI) website (https://www.ncbi.nlm.nih.gov/bioproject/) under BioProject PRJNA834785. The BioSample accession numbers for the raw reads and the five most abundant MAGs are SAMN28600298, SAMN28058410 ("*Ca*. Siderophilus nitratireducens), SAMN36381704 (*f_Gallionelaceae*), SAMN36381891 (*f_Anaeromyxobacteraceae*), SAMN36382736 (*g_Devosia*), SAMN36401011 (*f_Chitinophagaceae*).

8. Authors contributions

FCR, JB, MvL, DvH and ML conceived the study. JB and SD designed the filter and performed the on-site experiments. FCR run the metagenomics analysis, and AM the phylogenetic classification. FCR and GS conducted the thermodynamics calculations. SM performed the solid characterization. FCR, GS, MvL, DvH

and ML performed data analysis and/or helped interpret the results. FCR wrote the manuscript, with contributions from GS, MvL, DvH and ML. All co-authors critically reviewed the manuscript and approved the final version.

9. References

1. Matassa S, Batstone DJ, Hülsen T, Schnoor J, Verstraete W. Can direct conversion of used nitrogen to new feed and protein help feed the world? Environ Sci Technol. 2015;49(9):5247–54.

2. Giordano M. Global Groundwater? Issues and Solutions. Annu Rev Environ Resour [Internet]. 2009 Oct 15 [cited 2022 Jun 2];34(34):153–78. Available from: https://www.annualreviews.org/doi/abs/10.1146/annurev.environ.030308.100251

3. Ward MH, Jones RR, Brender JD, de Kok TM, Weyer PJ, Nolan BT, et al. Drinking water nitrate and human health: An updated review. Vol. 15, International Journal of Environmental Research and Public Health. 2018.

4. WHO. Guidelines for drinking-water quality. Resuscitation. 2010. 26–28 p.

5. McAllister SM, Vandzura R, Keffer JL, Polson SW, Chan CS. Aerobic and anaerobic iron oxidizers together drive denitrification and carbon cycling at marine iron-rich hydrothermal vents. ISME J [Internet]. 2021 [cited 2022 Nov 2];15(5):1271–86. Available from: https://doi.org/10.1038/s41396-020-00849-y

6. Schaedler F, Lockwood C, Lueder U, Glombitza C, Kappler A, Schmidt C. Microbially mediated coupling of Fe and N cycles by nitrate-reducing Fe(II)-oxidizing bacteria in littoral freshwater sediments. Appl Environ Microbiol. 2018;84(2):1–14.

7. Straub KL, Benz M, Schink B, Widdel F. Anaerobic, nitrate-dependent microbial oxidation of ferrous iron. Appl Environ Microbiol. 1996;62(4):1458–60.

8. Bryce C, Blackwell N, Schmidt C, Otte J, Huang YM, Kleindienst S, et al. Microbial anaerobic Fe(II) oxidation – Ecology, mechanisms and environmental implications [Internet]. Vol. 20, Environmental Microbiology. Blackwell Publishing Ltd; 2018 [cited 2020 Apr 10]. p. 3462–83. Available from: http://doi.wiley.com/10.1111/1462-2920.14328

9. Huang YM, Jakus N, Straub D, Konstantinidis KT, Blackwell N, Kappler A, et al. 'Candidatus ferrigenium straubiae' sp. nov., 'Candidatus ferrigenium bremense' sp. nov., 'Candidatus ferrigenium altingense' sp. nov., are autotrophic Fe(II)-oxidizing bacteria of the family Gallionellaceae. Syst Appl Microbiol [Internet]. 2022 [cited 2022 Nov 1];45(3). Available from: http://busco.ezlab.org/v3/datasets/bac-

10. Price A, Macey MC, Miot J, Olsson-Francis K. Draft Genome Sequences of the Nitrate-Dependent Iron-Oxidizing Proteobacteria Acidovorax sp. Strain BoFeN1 and Paracoccus pantotrophus Strain KS1. Microbiol Resour Announc [Internet]. 2018 Sep 13 [cited 2022 Nov 1];7(10). Available from: https://journals-asm-org.tudelft.idm.oclc.org/doi/10.1128/MRA.01050-18

11. Bryce C, Blackwell N, Schmidt C, Otte J, Huang Y-M, Kleindienst S, et al. Microbial anaerobic Fe(II) oxidation - Ecology, mechanisms and environmental implications. Environ Microbiol [Internet]. 2018 Oct 1 [cited 2021 Mar 4];20(10):3462–83. Available from: http://doi.wiley.com/10.1111/1462-2920.14328

12. Huang Y, Straub D, Blackwell N, Kappler A, Kleindienst S. Meta-omics Reveal Gallionellaceae and Rhodanobacter Species as Interdependent Key Players for Fe(II) Oxidation and Nitrate Reduction in the Autotrophic Enrichment Culture KS. Appl Environ Microbiol. 2021 Jul 13;87(15).

13. Weber KA, Achenbach LA, Coates JD. Microorganisms pumping iron: Anaerobic microbial iron oxidation and reduction [Internet]. Vol. 4, Nature Reviews Microbiology. 2006 [cited 2022 Dec 5]. p. 752–64. Available from: https://digitalcommons.unl.edu/bioscifacpub

14. Hedrich S, Schlömann M, Barrie Johnson D. The iron-oxidizing proteobacteria [Internet]. Vol. 157, Microbiology. Microbiology Society; 2011 [cited 2022 Oct 31]. p. 1551–64. Available from: https://www.microbiologyresearch.org/content/journal/micro/10.1099/mic.0.045344-0

15. Kleerebezem R, Van Loosdrecht MCM. Thermodynamic and kinetic characterization using process dynamics: Acidophilic ferrous iron oxidation by Leptospirillum ferrooxidans. Biotechnol Bioeng. 2008 May 1;100(1):49–60.

16. Agudelo-Vera C, Avvedimento S, Boxall J, Creaco E, de Kater H, Nardo A Di, et al. Drinking water temperature around the globe: Understanding, policies, challenges and opportunities. Water (Switzerland). 2020;12(4).

17. Caltran I, Rietveld LC, Shorney-Darby HL, Heijman SGJ. Separating NOM from salts in ion exchange brine with ceramic nanofiltration. Water Res. 2020 Jul 15;179.

18. Bowers RM, Kyrpides NC, Stepanauskas R, Harmon-Smith M, Doud D, Reddy TBK, et al. Minimum information about a single amplified genome (MISAG) and a metagenome-assembled genome (MIMAG) of bacteria and archaea. Nat Biotechnol. 2017 Aug 8;35(8):725–31.

19. Pinto AJ, Marcus DN, Ijaz UZ, Bautista-de lose Santos QM, Dick GJ, Raskin L. Metagenomic Evidence for the Presence of Comammox Nitrospira -Like Bacteria in a Drinking Water System. mSphere. 2016 Feb 25;1(1).

20. Hayhurst AN, Lawrence AD. The reduction of the nitrogen oxides NO and N2O to molecular nitrogen in the presence of iron, its oxides, and carbon monoxide in a hot fluidized bed. Combust

452 Flame. 1997 Aug 1;110(3):351–65.

453 21. Costa E, Pérez J, Kreft JU. Why is metabolic labour divided in nitrification? Trends Microbiol.
454 2006;14(5):213–9.

455 22. Pessi IS, Viitamäki S, Virkkala AM, Eronen-Rasimus E, Delmont TO, Marushchak ME, et al. In-depth
456 characterization of denitrifier communities across different soil ecosystems in the tundra. Environ
457 Microbiomes [Internet]. 2022;17(1):1–17. Available from: https://doi.org/10.1186/s40793-022-
458 00424-2

459 23. Gorski CA, Edwards R, Sander M, Hofstetter TB, Stewart SM. Thermodynamic Characterization of
460 Iron Oxide-Aqueous Fe2+ Redox Couples. Environ Sci Technol. 2016;50(16):8538–47.

461 24. Huang Y-M, Straub D, Kappler A, Smith N, Blackwell N, Kleindienst S. A Novel Enrichment Culture
462 Highlights Core Features of Microbial Networks Contributing to Autotrophic Fe(II) Oxidation
463 Coupled to Nitrate Reduction. Microb Physiol [Internet]. 2021 [cited 2021 Aug 13];1–16. Available
464 from: https://www.karger.com/Article/FullText/517083

465 25. Cheng B, Wang Y, Hua Y, Heal K V. The performance of nitrate-reducing Fe(II) oxidation processes
466 under variable initial Fe/N ratios: The fate of nitrogen and iron species. Front Environ Sci Eng
467 [Internet]. 2021 Aug 15 [cited 2021 Aug 27];15(4):73-. Available from:
468 https://journal.hep.com.cn/fese/EN/10.1007/s11783-020-1366-2

469 26. Jamieson J, Prommer H, Kaksonen AH, Sun J, Siade AJ, Yusov A, et al. Identifying and Quantifying
470 the Intermediate Processes during Nitrate-Dependent Iron(II) Oxidation. 2018 [cited 2020 Dec 16];
471 Available from: https://pubs.acs.org/sharingguidelines

472 27. He S, Tominski C, Kappler A, Behrens S, Roden EE. Metagenomic analyses of the autotrophic Fe(II)-
473 oxidizing, nitrate-reducing enrichment culture KS. Appl Environ Microbiol. 2016 May

1;82(9):2656–68.

28. Van Cleemput O. Subsoils: Chemo- and biological denitrification, N2O and N2 emissions. Nutr Cycl Agroecosystems [Internet]. 1998 [cited 2023 Jul 10];52(2–3):187–94. Available from: https://link.springer.com/article/10.1023/A:1009728125678

29. Huang J, Mellage A, Garcia JP, Glöckler D, Mahler S, Elsner M, et al. Metabolic Performance and Fate of Electrons during Nitrate-Reducing Fe(II) Oxidation by the Autotrophic Enrichment Culture KS Grown at Different Initial Fe/N Ratios. Appl Environ Microbiol. 2023;(Ii).

30. Tominski C, Heyer H, Lösekann-Behrens T, Behrens S, Kappler A. Growth and population dynamics of the anaerobic Fe(II)-oxidizing and nitrate-reducing enrichment culture KS. Appl Environ Microbiol [Internet]. 2018 May 1 [cited 2021 Mar 4];84(9):2173–90. Available from: http://aem.asm.org/

31. Sorokin DY, Tikhonova T V, Koch H, Berg EM Van Den, Hinderks RS, Pabst M, et al. Trichlorobacter ammonificans , a dedicated acetate-dependent ammoni fi er with a novel module for dissimilatory nitrate reduction to ammonia. 2023;(December 2022):1–10.

32. Fredrickson JK, Zachara JM, Kennedy DW, Dong H, Onstott TC, Hinman NW, et al. Biogenic iron mineralization accompanying the dissimilatory reduction of hydrous ferric oxide by a groundwater bacterium. Geochim Cosmochim Acta. 1998;62(19–20):3239–57.

33. H. Lees, Kwok SC, Suzuki I. The thermodynamics of iron oxidation by the ferrobacilli. 1968;(12):6–8.

34. Jolivet JP, Chanéac C, Tronc E. Iron oxide chemistry. From molecular clusters to extended solid networks. Chem Commun. 2004;4(5):477–83.

35. Vlek PLG, Blom TJM, Beek J, Lindsay WL. Determination of the Solubility Product of Various Iron

Hydroxides and Jarosite by the Chelation Method. Soil Sci Soc Am J. 1974;38(3):429–32.

36. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics [Internet]. 2014 Aug 1 [cited 2022 Mar 16];30(15):2114–20. Available from: https://pubmed.ncbi.nlm.nih.gov/24695404/

37. Andrews S. FastQC: A Quality Control Tool for High Throughput Sequence Data [Online] [Internet]. 2010. Available from: http://www.bioinformatics.babraham.ac.uk/projects/fastqc/

38. Sereika M, Kirkegaard RH, Karst SM, Michaelsen TY, Sørensen EA, Wollenberg RD, et al. Oxford Nanopore R10.4 long-read sequencing enables the generation of near-finished bacterial genomes from pure cultures and metagenomes without short-read or reference polishing. Nat Methods 2022 197 [Internet]. 2022 Jul 4 [cited 2023 May 4];19(7):823–6. Available from: https://www.nature.com/articles/s41592-022-01539-7

39. Lin Y, Yuan J, Kolmogorov M, Shen MW, Chaisson M, Pevzner PA. Assembly of long error-prone reads using de Bruijn graphs. Proc Natl Acad Sci U S A [Internet]. 2016 Dec 27 [cited 2023 May 4];113(52):E8396–405. Available from: https://www.pnas.org/doi/abs/10.1073/pnas.1604560113

40. Li H. Minimap2: pairwise alignment for nucleotide sequences. Bioinformatics [Internet]. 2018 Sep 15 [cited 2023 May 4];34(18):3094–100. Available from: https://academic.oup.com/bioinformatics/article/34/18/3094/4994778

41. Vaser R, Sović I, Nagarajan N, Šikić M. Fast and accurate de novo genome assembly from long uncorrected reads. Genome Res [Internet]. 2017 May 1 [cited 2023 May 4];27(5):737–46. Available from: https://genome.cshlp.org/content/27/5/737.full

42. Vasimuddin M, Misra S, Li H, Aluru S. Efficient architecture-aware acceleration of BWA-MEM for multicore systems. In: Proceedings - 2019 IEEE 33rd International Parallel and Distributed

Processing Symposium, IPDPS 2019. Institute of Electrical and Electronics Engineers Inc.; 2019. p. 314–24.

43. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics [Internet]. 2009 Aug 15 [cited 2022 Mar 16];25(16):2078–9. Available from: https://academic-oup-com.tudelft.idm.oclc.org/bioinformatics/article/25/16/2078/204688

44. Kang DD, Li F, Kirton E, Thomas A, Egan R, An H, et al. MetaBAT 2: An adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies. PeerJ [Internet]. 2019 [cited 2023 May 4];2019(7). Available from: https://kbase.us/applist/apps/metabat/run_metabat/

45. Wu YW, Simmons BA, Singer SW. MaxBin 2.0: An automated binning algorithm to recover genomes from multiple metagenomic datasets. Bioinformatics [Internet]. 2016 [cited 2023 May 4];32(4):605–7. Available from: www.jbei.org

46. Nissen JN, Johansen J, Allesøe RL, Sønderby CK, Armenteros JJA, Grønbech CH, et al. Improved metagenome binning and assembly using deep variational autoencoders. Nat Biotechnol [Internet]. 2021 [cited 2023 May 4];39(5):555–60. Available from: https://doi.org/10.1038/s41587-020-00777-4

47. K Sieber CM, Probst AJ, Sharrar A, Thomas BC, Hess M, Tringe SG, et al. Recovery of genomes from metagenomes via a dereplication, aggregation and scoring strategy. 2018 [cited 2023 May 4]; Available from: https://doi.org/10.1038/s41564-018-0171-1

48. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. Genome Res

[Internet]. 2015 Jul 1 [cited 2022 Aug 25];25(7):1043–55. Available from: https://genome-cshlp-org.tudelft.idm.oclc.org/content/25/7/1043.full

49. Parks DH, Chuvochina M, Waite DW, Rinke C, Skarshewski A, Chaumeil PA, et al. A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. Nat Biotechnol 2018 3610 [Internet]. 2018 Aug 27 [cited 2022 Aug 25];36(10):996–1004. Available from: https://www-nature-com.tudelft.idm.oclc.org/articles/nbt.4229

50. Arias-Carrasco R, Vásquez-Morán Y, Nakaya HI, Maracaja-Coutinho V. StructRNAfinder: An automated pipeline and web server for RNA families prediction. BMC Bioinformatics [Internet]. 2018 Feb 17 [cited 2022 Aug 25];19(1):1–7. Available from: https://bmcbioinformatics.biomedcentral.com/articles/10.1186/s12859-018-2052-2

51. Chan PP, Lin BY, Mak AJ, Lowe TM. tRNAscan-SE 2.0: improved detection and functional classification of transfer RNA genes. Nucleic Acids Res [Internet]. 2021 Sep 20 [cited 2022 Aug 25];49(16):9077–96. Available from: https://academic-oup-com.tudelft.idm.oclc.org/nar/article/49/16/9077/6355886

52. Hyatt D, Chen GL, LoCascio PF, Land ML, Larimer FW, Hauser LJ. Prodigal: Prokaryotic gene recognition and translation initiation site identification. BMC Bioinformatics [Internet]. 2010 Mar 8 [cited 2022 Aug 16];11(1):1–11. Available from: https://bmcbioinformatics.biomedcentral.com/articles/10.1186/1471-2105-11-119

53. Kanehisa M, Sato Y, Morishima K. BlastKOALA and GhostKOALA: KEGG Tools for Functional Characterization of Genome and Metagenome Sequences. J Mol Biol. 2016 Feb 22;428(4):726–31.

54. Garber AI, Nealson KH, Okamoto A, McAllister SM, Chan CS, Barco RA, et al. FeGenie: A Comprehensive Tool for the Identification of Iron Genes and Iron Gene Neighborhoods in Genome

and Metagenome Assemblies. Front Microbiol [Internet]. 2020;11:37. Available from: https://github.com/Arkadiy-Garber/FeGenie.

55. Tatusova T, Dicuccio M, Badretdin A, Chetvernin V, Nawrocki EP, Zaslavsky L, et al. NCBI prokaryotic genome annotation pipeline. Nucleic Acids Res [Internet]. 2016 Aug 19 [cited 2022 Aug 16];44(14):6614–24. Available from: https://pubmed-ncbi-nlm-nih-gov.tudelft.idm.oclc.org/27342282/

56. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res [Internet]. 1997 Sep 1 [cited 2022 Aug 16];25(17):3389–402. Available from: https://academic-oup-com.tudelft.idm.oclc.org/nar/article/25/17/3389/1061651

57. Rost B. Twilight zone of protein sequence alignments. Protein Eng Des Sel [Internet]. 1999 Feb 1 [cited 2022 Jun 6];12(2):85–94. Available from: https://academic-oup-com.tudelft.idm.oclc.org/peds/article/12/2/85/1550637

58. Parks DH, Chuvochina M, Chaumeil P-A, Rinke C, Mussig AJ, Hugenholtz P. A complete domain-to-species taxonomy for Bacteria and Archaea. Nat Biotechnol [Internet]. 2020;38. Available from: https://doi.org/10.1038/s41587-020-0501-8

59. Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, Von Haeseler A, et al. IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. Mol Biol Evol [Internet]. 2020 May 1 [cited 2022 Aug 29];37(5):1530–4. Available from: https://pubmed-ncbi-nlm-nih-gov.tudelft.idm.oclc.org/32011700/

60. Kalyaanamoorthy S, Minh BQ, Wong TKF, Von Haeseler A, Jermiin LS. ModelFinder: fast model selection for accurate phylogenetic estimates. Nat Methods [Internet]. 2017 May 30 [cited 2022

584 Aug 29];14(6):587–9. Available from: https://pubmed-ncbi-nlm-nih-
585 gov.tudelft.idm.oclc.org/28481363/

586 61. Thi Hoang D, Chernomor O, von Haeseler A, Quang Minh B, Sy Vinh L, Rosenberg MS. UFBoot2:
587 Improving the Ultrafast Bootstrap Approximation. Mol Biol Evol [Internet]. 2017;35(2):518–22.
588 Available from: http://www.iqtree.org.

589 62. Anisimova M, Gascuel O. Approximate likelihood-ratio test for branches: A fast, accurate, and
590 powerful alternative. Syst Biol [Internet]. 2006 Aug 1 [cited 2022 Aug 29];55(4):539–52. Available
591 from: https://pubmed-ncbi-nlm-nih-gov.tudelft.idm.oclc.org/16785212/

592 63. Richter M, Rosselló-Móra R, Oliver Glöckner F, Peplies J. JSpeciesWS: A web server for prokaryotic
593 species circumscription based on pairwise genome comparison. Bioinformatics [Internet]. 2016
594 [cited 2022 Nov 2];32(6):929–31. Available from: http://jspecies.ribohost.com/

595

596