


RESEARCH ARTICLE

Open Access



# Genome analysis of *Pseudomonas* sp. OF001 and *Rubrivivax* sp. A210 suggests multicopper oxidases catalyze manganese oxidation required for cylindrospermopsin transformation

Erika Berenice Martínez-Ruiz<sup>1\*</sup>, Myriel Cooper<sup>1\*</sup> , Jimena Barrero-Canosa<sup>1</sup>, Mindia A. S. Haryono<sup>2</sup>, Irina Bessarab<sup>2</sup>, Rohan B. H. Williams<sup>2</sup> and Ulrich Szewzyk<sup>1</sup>

## Abstract

**Background:** Cylindrospermopsin is a highly persistent cyanobacterial secondary metabolite toxic to humans and other living organisms. Strain OF001 and A210 are manganese-oxidizing bacteria (MOB) able to transform cylindrospermopsin during the oxidation of  $Mn^{2+}$ . So far, the enzymes involved in manganese oxidation in strain OF001 and A210 are unknown. Therefore, we analyze the genomes of two cylindrospermopsin-transforming MOB, *Pseudomonas* sp. OF001 and *Rubrivivax* sp. A210, to identify enzymes that could catalyze the oxidation of  $Mn^{2+}$ . We also investigated specific metabolic features related to pollutant degradation and explored the metabolic potential of these two MOB with respect to the role they may play in biotechnological applications and/or in the environment.

**Results:** Strain OF001 encodes two multicopper oxidases and one haem peroxidase potentially involved in  $Mn^{2+}$  oxidation, with a high similarity to manganese-oxidizing enzymes described for *Pseudomonas putida* GB-1 (80, 83 and 42% respectively). Strain A210 encodes one multicopper oxidase potentially involved in  $Mn^{2+}$  oxidation, with a high similarity (59%) to the manganese-oxidizing multicopper oxidase in *Leptothrix discophora* SS-1. Strain OF001 and A210 have genes that might confer them the ability to remove aromatic compounds via the catechol meta- and ortho-cleavage pathway, respectively. Based on the genomic content, both strains may grow over a wide range of  $O_2$  concentrations, including microaerophilic conditions, fix nitrogen, and reduce nitrate and sulfate in an assimilatory fashion. Moreover, the strain A210 encodes genes which may convey the ability to reduce nitrate in a dissimilatory manner, and fix carbon via the Calvin cycle. Both MOB encode CRISPR-Cas systems, several predicted genomic islands, and phage proteins, which likely contribute to their genome plasticity.

\* Correspondence: [erika.b.martinezruiz@campus.tu-berlin.de](mailto:erika.b.martinezruiz@campus.tu-berlin.de); [myriel.cooper@tu-berlin.de](mailto:myriel.cooper@tu-berlin.de)

<sup>1</sup>Chair of Environmental Microbiology, Technische Universität Berlin, Institute of Environmental Technology, Straße des 17. Juni 135, 10623 Berlin, Germany  
Full list of author information is available at the end of the article



© The Author(s). 2021 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

**Conclusions:** The genomes of *Pseudomonas* sp. OF001 and *Rubrivivax* sp. A210 encode sequences with high similarity to already described MCOs which may catalyze manganese oxidation required for cylindrospermopsin transformation. Furthermore, the analysis of the general metabolism of two MOB strains may contribute to a better understanding of the niches of cylindrospermopsin-removing MOB in natural habitats and their implementation in biotechnological applications to treat water.

**Keywords:** Metabolic potential, Manganese-oxidizing bacteria, Biotransformation, Cyanotoxins

## Background

Cylindrospermopsin (CYN) is a secondary metabolite produced by several cyanobacteria, toxic for humans and other living organisms [1]. The two bacterial strains OF001 and A210 transform the cyanotoxin CYN during the oxidation of  $Mn^{2+}$  [2, 3]. Strain OF001 belongs to the gammaproteobacteria and was isolated from the effluent of an experimental fixed-bed biofilm reactor established for the removal of recalcitrant substances from wastewater. Strain A210 belongs to the betaproteobacteria and was isolated from an iron manganese-depositing biofilm in a freshwater pond in the Lower Oder Valley National Park, Germany.

The removal of CYN by both strains required the active oxidation of  $MnCO_3$  whereas no or low CYN removal was observed with  $MnSO_4$  or in setups without manganese. Sterile biogenic oxides formed by the strains did not show any influence on CYN removal, highlighting the importance of the active manganese oxidation. Both strains are able to remove 100% of CYN at the highest rates reported for biological CYN removal so far [2, 3]. Furthermore, analysis of CYN transformation products revealed that the same seven transformation products were formed by both strains corroborating the important role of manganese oxidation. However, strain OF001 and A210 showed important differences. *Pseudomonas* sp. strain OF001 degraded CYN within 3 days. Whereas strain A210 degraded CYN within 14 to 28 days when cultivated under the same conditions. Moreover, strain OF001 required yeast extract as additional carbon source for the removal of CYN. In contrast, strain A210 was able to transform CYN in mineral media [2].

So far, little is known about biological CYN removal [2–6]. Even though several bacterial strains have been reported to remove CYN, to date, no enzymes or defined metabolic pathway for the transformation of CYN have been identified [7, 8]. Moreover, for biological CYN removal, no transformation products have been identified except for CYN transformed by MOB [3].

MOB are present in terrestrial [9], marine and freshwater environments [10–13], but they also occur in drinking water systems and reactors aiming at the removal of manganese and other pollutants [13–16]. MOB belong to diverse phylogenetic lineages with a broad

physiological diversity (e.g. autotrophs and mixotrophs) [10, 17, 18]. Through the oxidation of  $Mn^{2+}$ , MOB form water-insoluble biogenic manganese oxides, which are one of the strongest natural oxidants [17, 19]. Biogenic manganese oxides often interact with other compounds and thus play an important role in the biogeochemical cycle of manganese and other elements [17, 18, 20, 21].

The physiological role of manganese oxidation is not fully understood. Manganese oxidation was proposed to provide energy to support the growth of bacteria. However, no conclusive results were shown [22]. Other proposed functions are the protection against the toxicity of organic compounds, and reactive oxygen species [23, 24], the breakdown of organic matter into utilizable substrates [25, 26], and the use as a carbon reservoir [27]. Nevertheless, the precise physiological role of manganese oxidation remains unknown [18]. Different manganese oxidation mechanisms have been described including non-enzymatic pathways based on a pH increase, the oxidation through superoxide production, or an anaerobically photo-driven reaction; and enzymatic reactions generally associated to the activity of multicopper oxidases (MCO) and haem peroxidases [11, 18, 28].

Besides CYN, MOB transform different organic and inorganic pollutants, including diclofenac, benzotriazole, 17  $\alpha$ -ethinylestradiol, bisphenol A, As(III), and Sb(III) [2, 3, 29–34]. The mechanism of pollutant transformation was proposed to be based on unspecific oxidation by reactive manganese  $Mn^{3+}/Mn^{4+}$  species that are formed through the oxidation of  $Mn^{2+}$  [29]. For CYN transformation, a similar mechanism was assumed based on the requirement of active  $Mn^{2+}$  oxidation for efficient CYN removal and the formation of the same transformation products among all tested MOB, including strain OF001 and A210 [3]. Thus, it is suggested that MOB act as suppliers of biogenic oxides that indirectly oxidize the pollutants. However, the intrinsic capacity of MOB to remove organic compounds has not been deeply investigated [18].

The whole genome sequences of some MOB have been analysed previously to gain a better insight into the mechanism of manganese oxidation [35, 36]. However, so far, no reported pollutant-removing MOB strains were analyzed on a genomic level. Besides, for strain

OF001 and A210, information about the metabolic potential, including also about genes potentially involved in manganese oxidation, was missing. The genomic analysis of strain OF001 and strain A210 might allow to identify enzymes potentially involved in the oxidation of manganese based on the comparison with manganese oxidizing enzymes reported to date. Furthermore, metabolic differences between the two MOB strains became evident during cultivation experiments in presence of CYN [2], and further dissimilarities could be assumed. Such metabolic differences could be relevant for the application of the strains for the removal of pollutants from water in technical systems including but not limited to wastewater or drinking water treatment plants, and for the understanding of the niche they may occupy in natural environments.

Therefore, in this study, we analysed the draft genomes of the MOB strains OF001 and A210, both of which are able to transform CYN during oxidation of  $MnCO_3$ . We aim to provide further insight into i) manganese oxidation mechanism, ii) other metabolic pathways relevant for pollutant removal, iii) energy harvesting processes such as respiration, iv) their metabolic potential in comparison with their closest described phylogenetic relatives, and v) genome plasticity related to horizontal gene transfer mechanisms.

## Results and discussion

### General genome features

Genome quality estimation determined with CheckM showed that both genomes are of high quality (> 90% completeness and < 5% contamination). Genomes of strains OF001 and A210 have a completeness of 99.59 and 99.38%, respectively, with a contamination level of 2.12 and 0.35%.

The genome sequence of strain OF001 contains 4,476,686 bp in 65 contigs with a N50 contig length of 147,742 bp, and a GC content of 68.01%. The genome of strain OF001 encodes 4845 genes of which 4720 are protein coding sequences (CDS). Furthermore, one 16S, one 23S, and six 5S rRNA genes were identified in the genome of OF001, as well as, sixty-seven tRNA genes that enable recognition of codons for all 20 amino acids.

The genome sequence of strain A210 contains 5,371,534 bp in 72 contigs with a N50 contig length of 327,374 bp, and a GC content of 69.54%. The genome of strain A210 encodes 5184 genes from which 5112 are CDS. In addition, one 5S–23S–16S rRNA operon, and 52 tRNA genes were identified in the genome of strain A210. Genome quality estimation and general genomic features are summarized in Table 1.

Genome Taxonomy Database tool kit (GTDB-tk) was used to classify the bacterial genomes. GTDB-tk analysis classified strain OF001 as a member of the

**Table 1** Genomic features of strains OF001 and A210

	OF001	A210
N50	147,742	327,374
Number of contigs	65	72
CheckM completeness	99.59%	99.38%
CheckM contamination	2.12%	0.35%
Complete genome size (bp)	4,476,686	5,371,534
Undetermined bases	6400	7100
% GC	68.01	69.54
% Protein coding density	90.64	93.09
Pseudogene	6	2
CDS	4720	5112
Genes assigned to:		
COG	3436 (72.70%)	3798 (74.27%)
KEGG	2490 (52.7%)	2632 (51.5%)
Hypothetical proteins / unknown function	1327 (28.08%)	1492 (29.17%)
Fragment CDS	6	2
tRNA	67	52
rRNA	8	3
misc_RNA	43	14
tmRNA	1	1

CDS Coding sequences, COG Cluster of Orthologous Groups, KEGG Kyoto Encyclopedia of Genes and Genomes, tRNA transfer RNA, rRNA Ribosomal RNA, misc\_RNA Miscellaneous RNA, tmRNA Transfer-messenger RNA

*Pseudomonas\_K* group. According to the GTDB (May, 2020) *P. oryzae*, *P. sagittaria*, *P. linyingensis*, and *P. guangdongensis* belong to the *Pseudomonas\_K* group. The genus status of the strain OF001 in the *Pseudomonas\_K* group was supported by the genetic relatedness determined by whole-genome analysis and 16S rRNA phylogeny (Additional file 1: Fig. S1).

To determine the species affiliation of *Pseudomonas* sp. OF001 average nucleotide identity (ANI), and tetra-nucleotide frequencies (TETRA) analysis were done. The analysis revealed highest similarity between strain OF001 and *P. oryzae* KCTC 32247 with an ANI based on BLAST (ANiB) value of 89.06%, an ANI based on MUMmer (ANIm) value of 90.98%, and a TETRA value of 0.998 (Fig. 1a). Organisms with an ANI value above 95%, and a TETRA value above 0.99 are suggested to delineate the same species level [38–40]. TETRA values should be in agreement with ANI values to support the species assignment [39]. TETRA values of strain OF001 and the organisms of the *Pseudomonas\_K* group were higher than 0.99, but ANI values were below the species limit. Together, the data suggest strain OF001 is a potential new species of the *Pseudomonas\_K* group.

GTDB-tk analysis classified strain A210 as a member of the *Rubrivivax* genus. According to the GTDB database this genus belongs to the order Burkholderiales and has so far only three described species: *R. benzoatilyticus*, *R. gelatinosus*, and *R. albus* [41–43].

Based on the phylogenetic analysis using the whole 16S rRNA gene sequence (Additional file 1: Fig. S2a), strain A210 could not be classified at genus level. Organisms with high similarity to the 16S rRNA gene sequence of strain A210 were mainly bacteria of the *genera incertae sedis* from the *Comamonadaceae* family (*Aquabacterium*, *Ideonella*, *Leptothrix*, *Roseateles*, *Rubrivivax*, *Sphaerotilus*). However, phylogenomic analysis of strain A210 done with TYGS platform affiliated A210

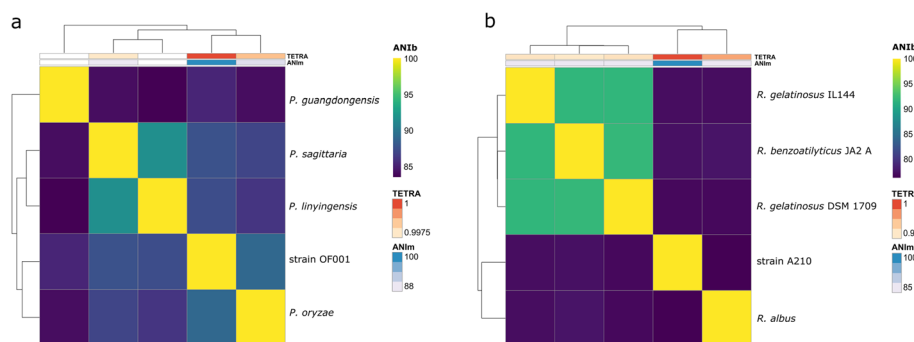
with organisms of the genus *Rubrivivax* (Additional file 1: Fig. S2b), supporting the results obtained with the GTDB-tk.

ANI, and TETRA analysis were done with the genome of A210 to analyze species affiliation. The analysis showed the highest similarity of *Rubrivivax* sp. A210 with *R. benzoatilyticus* JA2 with an ANiB value of 76.69%, an ANIm value of 84.45%, and a TETRA value of 0.913 (Fig. 1b). Thus suggesting, strain A210 is a potential new species of the genus *Rubrivivax*.

### Pan and core genome

The pan-genome of the *Pseudomonas\_K* group genomes comprised 20,296 genes belonging to 6805 Microscope gene Families (MICFAM) [44, 45]. The core-genome comprised 11,985 genes that correspond to 1957 MICFAM, and the variable-genome contained 8311 corresponding to 4848 MICFAM. The *Pseudomonas* sp. OF001 genome contains 1091 strain-specific genes from 1052 MICFAM that correspond to 24.08% strain-specific coding sequences. With this, strain OF001 contains the highest number of CDS from the *Pseudomonas\_K* group genomes analyzed (Additional file 1: Fig. S3a). Among the strain-specific genes in OF001, we found genes related to mercury resistance, transport, and foreign DNA (see also section 2.5 *Elements potentially acquired by horizontal gene transfer*).

Pan- and core-genome size evolutions were estimated with the four available genomes of the *Pseudomonas\_K* group and the genome of strain OF001. The curve of the pan-genome of strain OF001 and *Pseudomonas\_K* group did not reach the plateau, suggesting that the pan-genome of *Pseudomonas\_K* group is open and the sequences of other genomes from this group might increase the gene pool of novel genes (Additional file 1: Fig. S4a). The plateau of the core-genome is reached



**Fig. 1** Heatmap representing the degree of similarity of the MOB genomes. **a** *Pseudomonas* sp. OF001, and **b** *Rubrivivax* sp. A210. Heatmaps were derived from the average nucleotide identity (ANI) matrix based on BLAST (ANiB). Dendrogram directly reflects the degree of identity between genomes. ANIm: ANI based on MUMmer; TETRA: correlation indexes of the tetra-nucleotide frequencies; DDH  $d_4$ : DDH calculated with the formula  $d_4$ , which is the non-logarithmic version of formula  $d_5$  (used for the Fig. S1 and S2). Formula  $d_4$  is highly recommended when using draft genomes to assure confident results [37]

with the five genomes selected and is composed of approximately 2000 MICFAM (Additional file 1: Fig. S4b).

The pan-genome of *Rubrivivax* genomes comprised 23,140 genes belonging to 9974 MICFAM. The core-genome comprises 10,154 genes that correspond to 1629 MICFAM, and the variable-genome contains 12,986 genes corresponding to 8345 MICFAM. The *Rubrivivax* sp. A210 genome contains 2123 strain-specific genes from 2035 MICFAM that correspond to 42.98% strain-specific coding sequences (Additional file 1: Fig. S3b). Among the strain-specific genes in A210, we found genes related to transport like ABC transporters, and cytochromes (see also section 2.4.3 *Aerobic respiration*).

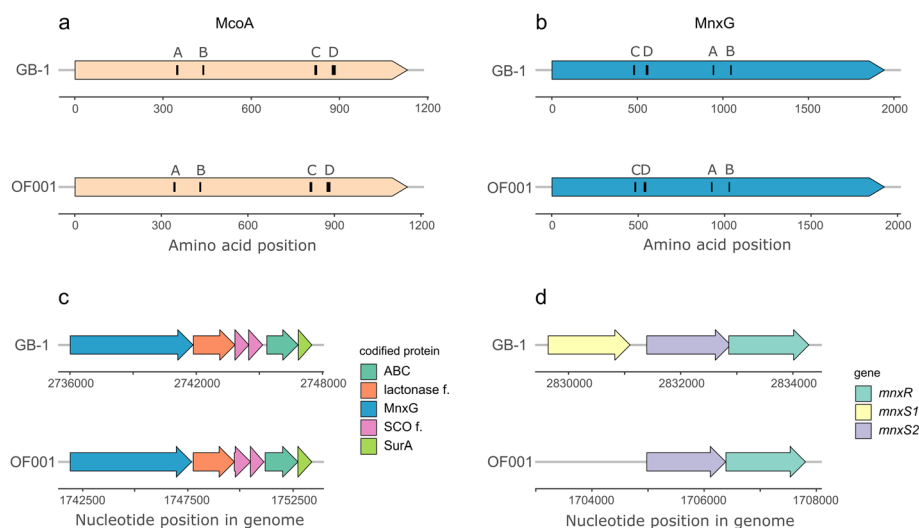
Pan- and core-genome size evolutions were estimated according to the genomes selected for the A210 analysis (Additional file 1: Fig. S4c-d). The core-genome plateau is apparently reached with the analyzed genomes and is composed of approximately 1600 MICFAM.

### Genes potentially involved in manganese oxidation

In *Pseudomonas* sp. OF001, we detected three different homologues of manganese-oxidizing multicopper oxidases (MO-mco's) (OF001\_u20185, OF001\_u60094, and OF001\_u90046). Gene name, accession number, locus tag in the evaluated genomes, E-value, and percent similarity of amino acid alignments are shown in Additional file 1: Table S1. All three MO-mco's homologues of strain OF001 belong to the homologous cupredoxin superfamily (IPR008972), according to the InterPro-

based analysis. The amino acid sequences encoded by OF001\_u20185 and OF001\_u60094 exhibit the four characteristic motifs found in multicopper oxidases, in the same order and in a similar position as observed in McoA and MnxG from *P. putida* GB-1 (Fig. 2a, b). In addition, MO-mco's homologues OF001\_u20185 and OF001\_u60094, showed the highest similarity to the Mn<sup>2+</sup> oxidases *mnxG* (80%) and *mcoA* (83%) from *P. putida* GB-1 [47], whereas OF001\_u90046 showed the highest similarity to *moxA* (51%) from *Pedomicrobium* sp. ACM 3067 [48]. According to the InterPro-based analysis, all three MO-mco's contain non-cytoplasmic domain regions of membrane-bound proteins that cover more than 94% of the whole protein sequence. These regions are predicted to be outside the membrane in the extracellular region. Moreover, OF001\_u90046 and OF001\_u60094 contain transmembrane helices. The presence of non-cytoplasmic and transmembrane domains suggests that these enzymes are loosely bound to the outer membrane, which is in agreement with the localization of MO-mco's in other MOB [47, 49, 50]. Functional domains of the proteins and the ontology classification are shown in Additional file 1: Table S2.

OF001\_u60094 in *Pseudomonas* sp. OF001 is located in a predicted operon similar to *mnxG* in *P. putida* GB-1 [51]. The operon is composed out of five additional genes with high similarity to those located in the *mnxG* operon of *P. putida* GB-1, (63–76% according to blastp analysis, Fig. 2c). Expression of the MO-mco's in *P.*



**Fig. 2** Genetic organization of regulatory system and MO-mco's in *P. putida* GB-1 and *Pseudomonas* sp. OF001. **a** McoA protein of strain GB-1, and the putative homolog found in strain OF001, **b** MnxG protein of strain GB-1, and the putative homolog found in strain OF001, **c** predicted operon organization in which *mnxG* (MO-mco) is found in strain GB-1, and putative homologues found in a predicted operon in strain OF001, and **d** regulatory system for Mn<sup>2+</sup> oxidation of strain GB-1, and putative homologues found in strain OF001. Capital letters (A-D) in **a**, and **b**) represent the multicopper oxidase motifs [46]. *mnxR*: response regulator; *mnxS1* and *mnxS2*: sensor histidine kinases; ABC: ABC transporter; lactonase f.: beta-propeller fold lactonase family protein; *mnxG*: MO-mco, SCO f. SCO family protein; SurA: SurA N-terminal domain-containing protein, McoA: MO-mco, MnxG: MO-mco

*putida* GB-1 is regulated by a two-component pathway, *mnxS1/mnxS2/mnxR* [51]. In the genome of strain OF001, we found putative homologues to the *mnxS2* histidine kinase, and to the *mnxR* regulator, arranged in a similar operon structure as in *P. putida* GB-1 (Fig. 2d) [51]. Our results suggest that the regulation of the MO-mco's of strain OF001 follows a similar regulation to the one observed in *P. putida* GB-1.

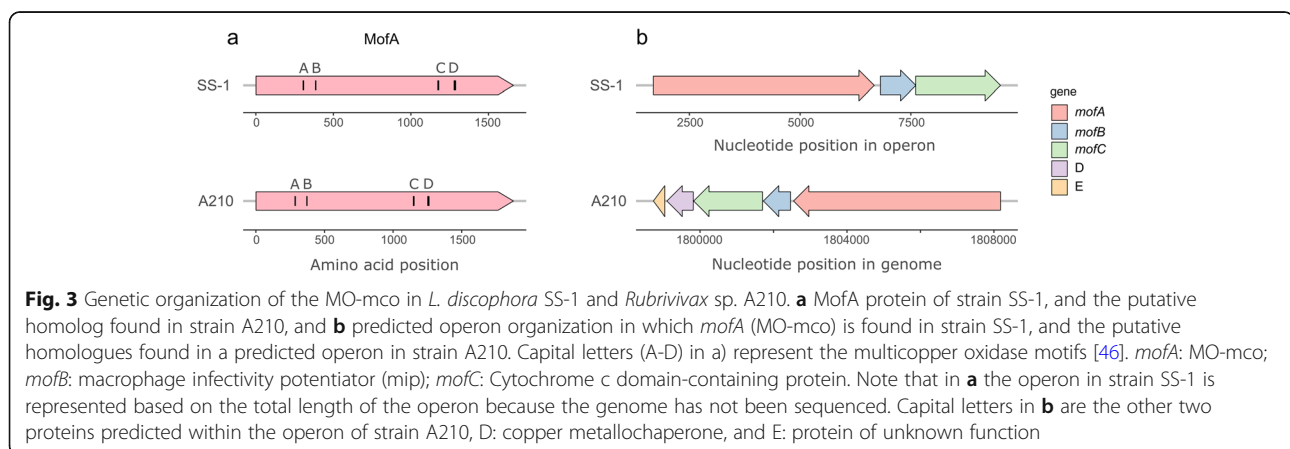
Furthermore, two homologues of manganese-oxidizing haem peroxidases (MO-hpox's) (OF001\_u100035, and OF001\_u220048) were identified in strain OF001. The putative MO-hpox homologue OF001\_u100035 showed the highest similarity with the Mn<sup>2+</sup> oxidase *mopA* of *P. putida* GB-1 (42%). Together with the MO-hpox of *A. manganoxydans* SI85-9A1, they belong to the haem peroxidase superfamily (IPR010255). The MO-hpox homologue OF001\_u220048, showed highest similarity to *mopA* of *Erythrobacter* sp. SD-21 (38%) and neither of the two belong to the haem peroxidase superfamily (Additional file 1: Table S2). No cytoplasmic or non-cytoplasmic domains could be identified for the putative MO-hpox's homologues of strain OF001. Therefore, we evaluated the probable subcellular localization with Loc-Tree3 [52]. According to this analysis, both putative MO-hpox's of OF001 are likely secreted to the media (accuracy percentage 88%), similar as previously described for several MO-hpox's of other MOB [47–50, 53, 54].

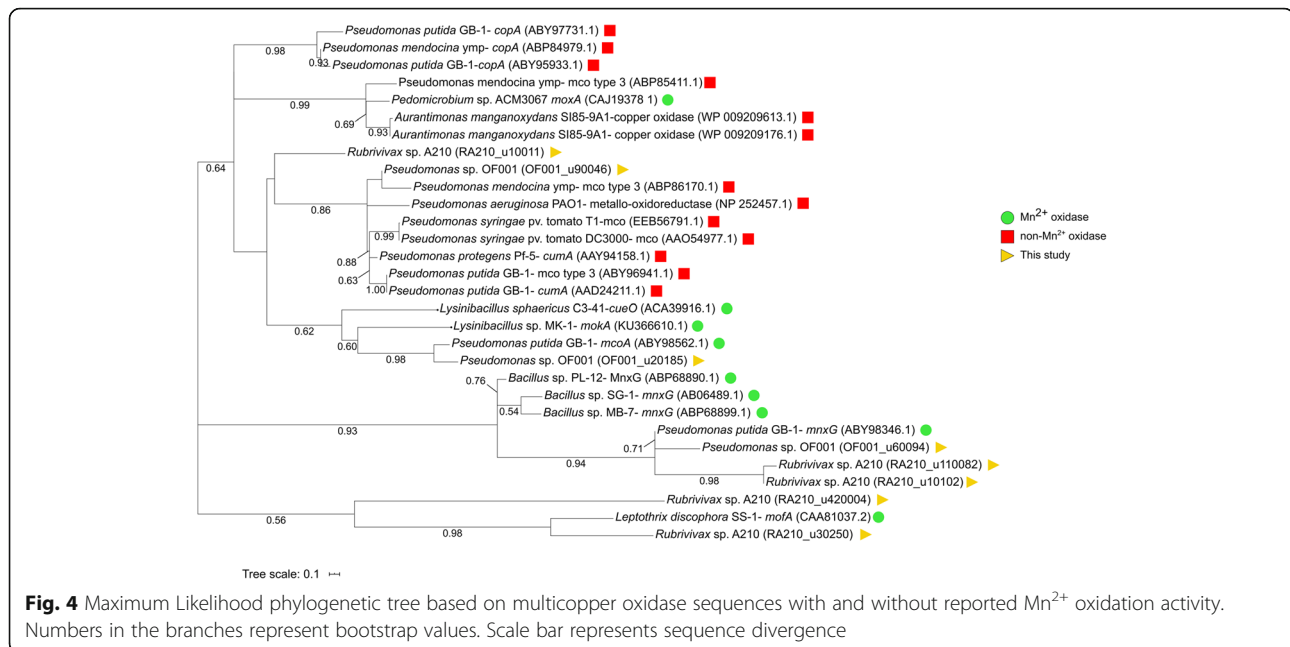
In the genome of *Rubrivivax* sp. A210, five MO-mco's homologues were identified (RA210\_u420004, RA210\_u30250, RA210\_u110082, RA210\_u10102, and RA210\_u100111) (Additional file 1: Table S1). Two MO-hpox's homologues (RA210\_u10091, and RA210\_u140033) were identified, but were discarded for further analysis due to very low coverage of the query sequences (Additional file 1: Table S1).

All MO-mco's homologues of strain A210 belong to the homologous cupredoxin superfamily (IPR008972), according to the InterPro-based analysis. They contain

non-cytoplasmic domains which cover more than 84% of the whole protein sequence, and possess either a transmembrane domain or a transmembrane helix, except for RA210\_u420004. This suggests that these enzymes are loosely bound to the outer membranes, similar as previously reported for other MO-mco's [47, 49, 50]. RA210\_u30250 shows highest similarity (59%) to the *mofA* gene of *L. discophora* SS-1 [49]. In addition, the amino acid sequence of RA210\_u30250 encodes the four characteristic motifs found in multicopper oxidases in the same order and in a similar position than those found in MofA from *L. discophora* SS-1 (Fig. 3a). In addition, it is located in a predicted operon similar to *mofA* in *L. discophora* SS-1 (Fig. 3b). The *mof* operon in *L. discophora* SS-1 is composed out of *mofA*, *mofB* and *mofC* [55]. The putative *mof* operon in strain A210 encodes five genes, including the putative *mofA* homologue, and two genes with high similarity to *mofB* (68%) and *mofC* (60%), together with a putative metallochaperon, and an exported protein of unknown function (RA210\_u30246 – RA210\_u30250).

In spite of the low homology between MO-mco's from different organisms, we attempted to gain further evidence for the Mn<sup>2+</sup> oxidation activity of the suggested multicopper oxidases by using a phylogenetic approach. For this purpose, a phylogenetic tree was constructed with sequences of MO-mco and non-MO-mco retrieved from the NCBI database excluding the newly identified putative MO-mco homologues (Additional file 1: Table S3), to discard the possibility that the new sequences were the main factor driving the topology of the tree (Additional file 1: Fig. S5). Subsequently, the putative MO-mco homologues of the strains OF001 and A210 were added. Phylogenetic analysis revealed one cluster of all MO-mco sequences and one cluster of non-MO-mco (Fig. 4). The only identified outlier was *moxA* from *Pedomicrobium* sp. ACM 3067, a reported MO-mco, affiliated with the non-MO-mco. Possibly, this is due to an uncertain assignment as suggested previously by





Anderson et al. (2009). In contrast to OF001\_u20185, OF001\_u60094, and RA210\_u30250, the proteins encoded by OF001\_u90046, and RA210\_u100111 did not cluster with the MO-mco (Fig. 4). This result suggests that the two annotated multicopper oxidases OF001\_u90046 and RA210\_u100111 in *Pseudomonas* sp. OF001 and *Rubrivivax* sp. A210, respectively, do not possess the Mn<sup>2+</sup> oxidation activity. Collectively, the data suggest that the best candidates for Mn<sup>2+</sup> oxidation are MO-mco OF001\_u20185, OF001\_u60094, and MO-hpox OF001\_u100035 in strain OF001 and MO-mco RA210\_u30250 in strain A210.

Our results indicate that both MOB strains, OF001 and A210, oxidize manganese through enzyme-mediated mechanisms. In spite of the evidences found based on the genomic analysis, further experiments are required to determine which enzymes are involved in the oxidation of Mn<sup>2+</sup> in *Pseudomonas* sp. OF001 and *Rubrivivax* sp. A210.

## General metabolism

### Organic carbon metabolism

*Pseudomonas* sp. OF001 and *Rubrivivax* sp. A210 possess all genes necessary for commonly found central carbohydrate metabolism in aerobic organism including glycolysis (Embden-Meyerhof-Parnas), gluconeogenesis, tricarboxylic acid cycle (Krebs cycle), and the non-oxidative branch of the pentose phosphate pathway, to support basic growth. In both MOB, genes involved in the oxidative branch of the pentose phosphate pathway were incomplete which is in accordance with its absence in many aerobic and thermophilic organisms [56].

## CO<sub>2</sub> fixation

*Pseudomonas* sp. OF001 possesses several genes encoding enzymes related to CO<sub>2</sub> fixation via the Calvin cycle, however the key enzyme D-ribulose-1,5-bisphosphate carboxylase/oxygenase (RuBisCO) is missing (Additional file 1: Table S4). This is in accordance with our previous study which demonstrated the growth of strain OF001 only in presence of an organic carbon source [2].

In contrast, *Rubrivivax* sp. A210 has the complete repertoire of genes required for CO<sub>2</sub> fixation via the Calvin cycle, including the RuBisCO, which is supported by previous studies of our group showing that A210 was able to grow in mineral media [2]. The *cbb* operon in strain A210 has all genes predicted to be encoded together with the RuBisCO small (*cbxSP*) and large (*cbbL*) subunits (*gpx*, *cbbYP*, *prkB*, *fbp*, *cbxXC*). The presence of genes coding for enzymes of the Calvin cycle in strain A210 is in accordance with their detection in the three described species of the *Rubrivivax* genus *R. albus* [42], *R. gelatinosus* [43] and *R. benzoatilyticus* [41].

## Aerobic respiration

All genes for oxidative phosphorylation and aerobic respiration were present in *Pseudomonas* sp. OF001. Among them, twenty genes annotated as cytochromes, including 14 *c*-type, 5 *b*-type, and 1 *d*-type cytochromes were found (Additional file 1: Table S5). Also, several predicted terminal oxidases are present including cytochrome *bd*-type quinol oxidase, cytochrome *c* oxidases, and *cbb*<sub>3</sub>-type cytochrome *c* oxidases.

*cbb*<sub>3</sub>-type cytochrome *c* oxidases in the genome of strain OF001 are predicted to be organized in two

operons, one operon containing *cbbP* and *cbbQON* genes, similar as reported for other bacteria [57–60], and one operon containing a copy of *cbbPO* and a gene of unknown function. Next to the *cbbPQON* operon, a predicted operon with three (*ccoSIG*) genes encoding the enzymes responsible of the assembly of *cbb<sub>3</sub>*-type cytochrome *c* oxidases was observed [61]. The *ccoH* assembly factor for the *cbb<sub>3</sub>*-type cytochrome oxidase is missing in strain OF001, suggesting that it follows a *ccoH*-independent assembly mechanism, similar as described for *H. pylori*, and *R. gelatinosus* (Durand et al. 2018).

Likewise, all genes for oxidative phosphorylation and aerobic respiration were found in *Rubrivivax* sp. A210. Thirty-nine genes were annotated as cytochromes, including thirty *c*-type, and nine *b*-type cytochromes. Predicted terminal oxidases are present, including cytochrome *c* oxidases and *cbb<sub>3</sub>*-type cytochrome *c* oxidases (Additional file 1: Table S5). *cbb<sub>3</sub>*-type cytochrome *c* oxidases and the enzymes responsible of their assembly in the genome of strain A210 are predicted to be organized in a single operon *ccoISNOQPG*. Similar as for OF001, the *ccoH* assembly factor for the *cbb<sub>3</sub>*-type cytochrome oxidase is missing in strain A210, which suggest that it also follows a *ccoH*-independent assembly mechanism likewise to strain OF001.

The presence of diverse cytochrome oxidases, with high O<sub>2</sub> affinity, rather than only cytochrome *c* oxidases, indicate the potential of strain OF001 and A210 to grow under a wide range of O<sub>2</sub> concentrations.

### Nitrogen metabolism

*Pseudomonas* sp. OF001 possesses genes predicted to participate in ammonium uptake, including specific transporters like *amtB* and genes involved in the regulation of the process such as *glnA*, *glnL*, and *glnK* [62–67]. The genes are predicted to be arranged within different operons, with *glnA* as a single regulated gene, located immediately downstream from the operon encoding *glnG* and *glnL* (Additional file 1: Table S4).

In addition, the genome of *Pseudomonas* sp. OF001 encodes the genes *nifDKH*, implicated in nitrogen fixation. Nitrogenase genes in strain OF001 are not predicted to form an operon, but cluster together in the genome. The detection of the *nifDKH* genes is in accordance with their detection in the genomes of the two taxonomically closest organisms to OF001, *P. linyingensis* [68] and *P. sagittaria* [69].

Furthermore, strain OF001 encodes genes related with assimilatory nitrate reduction, including nitrate transport, the ammonium-forming nitrite reductase small subunit *nasD*, and the nitrate reductase *nasA* [70–73]. Strain OF001 also possesses two nitrite reductases, one in a predicted operon together with the nitrate reductase

*nasA*, and the other as an independent regulated gene (Additional file 1: Table S4). Genes involved in dissimilatory nitrate reduction were missing. This is in agreement with the absence of genes involved in dissimilatory nitrate reduction in the closest related *Pseudomonas\_K* group.

Similar to strain OF001, the genome of *Rubrivivax* sp. A210 encodes genes predicted to participate in ammonium uptake in five predicted operons, including specific transporters like *amtB* and genes involved in the regulation of the process such as *glnA*, *glnL*, and *glnK* [62–67] (Additional file 1: Table S4). However, in contrast to strain OF001, the *glnB* gene is encoded in the genome of strain A210. *GlnB* is a PII signal transcription protein, homologue to *GlnK* [74]. Both are key for the metabolic regulation of ammonium uptake. The presence of the *glnK* and *glnB* genes in *Rubrivivax* sp. A210 suggests that ammonium uptake in strain A210 follows a similar regulation as described for *Escherichia coli* [64, 75]. *GlnB* found in Proteobacteria is commonly associated with glutamine synthetase genes [76], and likewise, the *glnB* gene of strain A210 is located in an operon structure next to the glutamine synthetase *nadE*.

The genes *nifDKH*, implicated in nitrogen fixation, are encoded in the genome of *Rubrivivax* sp. A210 similarly as observed for all three known *Rubrivivax* species [41–43]. The *nifDKH* genes are located in a predicted operon structure together with a putative ferredoxin and a conserved protein of unknown function. Ferredoxin may mediate nitrogenase activity when ammonium is available for uptake [77–80].

Moreover, strain A210 encodes genes related with assimilatory nitrate reduction, including nitrate transport, the ammonium-forming nitrite reductase small subunit *nirB*, and the nitrate reductase *nasA* [70–73].

In contrast to strain OF001, in the genome of strain A210 genes related with dissimilatory nitrate reduction were detected, including nitrate reductase *narGHI*, and nitrite reductase *nirBD*. Noteworthy, in the three described species of the *Rubrivivax* genus [41–43], genes coding for enzymes related to dissimilatory nitrate reduction were not detected. The enzymes related with the dissimilatory and assimilatory nitrate reduction are organized in two predicted operons.

The genomic data indicate that both MOB strains have the ability to assimilate ammonium. In addition, it seems likely that OF001 can use nitrate in an assimilatory but not dissimilatory pathway. In contrary, the genomic data suggest that *Rubrivivax* sp. A210 has not only the genomic potential to assimilate nitrate, but also to perform anaerobic respiration using nitrate as final electron acceptor. This characteristic may confer a higher flexibility to strain A210, compared to OF001, to adapt to changing conditions in technical and natural environments.



### Sulfur metabolism

*Pseudomonas* sp. OF001 harbours all genes required for assimilatory sulfate reduction, which are organised in several predicted operons, and some as single regulated genes (Additional file 1: Table S4). Strain OF001 also possesses different sulfate transporters including ABC type UWA [81], the proton: sulfate symporter or putative sulfate: bicarbonate antiporter SulP [82, 83], and the high affinity sulfate transporter CysZ, essential for sulfate uptake at low concentrations [84, 85]. No genes required for dissimilatory sulfate reduction were detected in the genome of strain OF001.

*Rubrivivax* sp. A210 harbours several genes required for assimilatory sulfate reduction, organised in three predicted operons (Additional file 1: Table S4). However, A210 lacks the adenylyl sulfate kinase *cysC*, responsible of the transformation of adenosine 5'-phosphosulfate (APS) to 3'-phosphoadenosine-5'-phosphosulfate (PAPS), which is an essential step in the assimilatory sulfate reduction [86]. Nevertheless, other organisms like *P. aeruginosa*, *Sinorhizobium meliloti*, and *Burkholderia cenocepacia* lacking *cysC*, reduce APS via the phosphoadenosine phosphosulfate reductase *cysH* to sulphite [87–89]. Strain A210 also possesses different sulfate transporters like ABC type UWA [81]. Similar to strain OF001 and other *Rubrivivax* species, not all genes involved in dissimilatory sulfate reduction were identified in strain A210.

### Iron metabolism

*Pseudomonas* sp. OF001 possess genes predicted to participate in the transport and storage of iron, including ferrous and ferric iron transporters and bacterioferritin [90] (Additional file 1: Table S4). Strain OF001 also possesses different genes related to heme uptake, such as heme-binding protein and periplasmic heme chaperone [91]. AntiSmash analysis could not detect gene clusters related to siderophores synthesis in strain OF001 genome. Nonetheless, the search with blastp revealed the presence of thirty-one genes related to the synthesis and transport of siderophores. Twenty-one out of the thirty-one genes detected in the genome of strain OF001, correspond to siderophores transport (Additional file 1: Table S4).

Similar to strain OF001, the genome of *Rubrivivax* sp. A210 encodes genes predicted to participate in the transport and storage of iron, such as ferric iron transporters and bacterioferritin [90] (Additional file 1: Table S4). Strain A210 also possesses different genes related to heme uptake, such as heme-binding protein and periplasmic heme chaperone [91]. AntiSmash could not detect gene clusters related to siderophores synthesis in strain A210 genome. However, the search with blastp revealed the presence of thirty-six genes related to the

synthesis and transport of siderophores. Twenty-two out of the thirty-six genes detected in A210 genome, correspond to siderophores transport (Additional file 1: Table S4).

### Cell motility and biofilm formation

Proteins for motility, including genes related to chemotaxis, and flagellar proteins were present in the genome of *Pseudomonas* sp. OF001 (Additional file 1: Table S6). Genes encoding flagellar proteins in strain OF001 belong to the *flg*, and *fli* family, which are part of the core set of flagellar genes [92]. Among the genes related to the central signal transduction pathway for chemotaxis, we found *cheAWYBR* genes, and the transmembrane chemoreceptors, methyl-accepting chemotaxis proteins (MCPs) in the genome of strain OF001. This is in agreement with the description of the closest relatives of strain OF001, *P. oryzae*, *P. sagittaria*, *P. guangdongensis*, and *P. linyingensis*, as motile bacteria [68, 69, 93, 94]. Sixty-four genes associated with biofilm formation were found in the genome of *Pseudomonas* sp. OF001, including *siaD*, *bifA*, and *fleQ* (Additional file 1: Table S6).

In the genome of *Rubrivivax* sp. A210 genes required for motility were present, including genes related to chemotaxis, and genes encoding flagellar proteins (Additional file 1: Table S6). Similar as in strain OF001, in strain A210 found genes encoding flagellar proteins belong to the *flg*, and *fli* family. Strain A210 possesses *cheAWYBR* genes, and the transmembrane chemoreceptors MCPs, which are related to the central signal transduction pathway for chemotaxis. The taxonomically closest bacteria to strain A210, *R. gelatinosus*, and *R. benzoatilyticus*, are also motile bacteria [41, 43]. However, despite the presence of motility-related genes in *R. albus*, the absence of motility was experimentally evidenced [42]. Therefore, further experiments are required to verify motility of strain A210. One hundred twenty one genes associated with biofilm formation were found in the genome of *Rubrivivax* sp. A210, including *sadC*, *pill*, and *pslH*. (Additional file 1: Table S6).

As discussed above, both strains contain genes that encode proteins involve in biofilm formation, which is in agreement with the isolation of both MOB from biofilms. Moreover, both strains form biofilm in pure culture.

### Organic compound degradation

Genes for the aerobic degradation of aromatic compounds via the catechol meta-cleavage pathway, and for the specific degradation of benzoate, phenol, and benzene, including phenol/toluene 2-monooxygenases (NADH), benzoate/toluene 1,2-dioxygenases, and catechol 2,3-dioxygenases, were detected in *Pseudomonas* sp. OF001 (Additional file 1: Table S7). In addition,

strain OF001 may have the potential to transform other compounds like 2-, 3- and 4-fluorobenzoate, toluene, steroids, citalopram, trinitrotoluene, p-methylbenzoate, trans-cinnamate, phenylpropanoate, and 4-hydroxyphenylacetate. This is in accordance with the ability of several *Pseudomonas* spp. to transform diverse organic pollutants such as benzoate, toluene, phenol, and poly-chlorobiphenyls (PCBs) [95].

Because strain OF001 is able to degrade the cyanotoxin CYN we were also interested in the potential of the strains to degrade other cyanobacterial toxins. However, specific enzymes for the degradation of cyanotoxins are described only for microcystin, the most studied cyanotoxin [7, 8]. No genes involved in microcystin degradation were found in the genome of OF001. Although biodegradation of CYN is considered one of the main natural attenuation processes [96], no specific genes involved in their transformation are known yet [3].

*Rubrivivax* sp. A210 harbors genes involved in the aerobic degradation of aromatic compounds via the catechol ortho-cleavage pathway, and for the specific degradation of benzoate, and 3- and 4-fluorobenzoate, including benzoate/toluene 1,2-dioxygenases, muconate cycloisomerase, and catechol 1,2-dioxygenase (Additional file 1: Table S7). Similarly, the closely-related strain *R. bezoatilyticus* JA2 catabolizes different aromatic compounds including benzoate [41]. Strain A210 also has the potential to transform other compounds such as 4-methylcatechol, acrylonitrile, 2-fluorobenzoate, and trinitrotoluene.

Because strain A210 is able to degrade CYN similarly as strain OF001, we searched the genome for genes related to cyanotoxin transformation. We could not find genes associated with the transformation of microcystin.

#### Elements potentially acquired by horizontal gene transfer

##### Genomic islands

Genomic islands are genomic regions potentially obtained by horizontal gene transfer that can drive strain differentiation and support adaptation. Analysis of *Pseudomonas* sp. OF001 genome with IslandViewer 4 led to the identification of at least 12 genomic islands with size ranges from 4.2 to 70.5 Kb (Additional file 1: Table S8 and Fig. S6a). Genomic islands in strain OF001 include genes associated with transposases, phage proteins, CRISPR systems, 147 proteins of unknown function, toxin-antitoxin systems, metal-related proteins, and mercury resistance. Metal resistance genes are related with environmental pollution, and specifically, mercury resistance genes are the genes most frequently associated with genomic islands [97].

Analysis of *Rubrivivax* sp. A210 genome with IslandViewer 4 led to the identification of at least 8 genomic islands with size ranges from 3.8 to 99.5 Kb (Additional

file 1: Table S8 and Fig. S6b). Genomic islands in strain A210 included genes associated with transposases, phage proteins, 136 proteins of unknown function, toxin-antitoxin systems, transporters, and nitrate reduction. Genes related to nitrogen metabolism associated to genomic island have been previously reported [98, 99].

##### Prophages

Using PHASTER (PHAge Search Tool Enhanced Release), we detected four incomplete and three intact prophage regions (Score  $\geq 100$ ) in *Pseudomonas* sp. OF001 genome (Additional file 1: Table S9 and Fig. S7a). The three intact prophages were named OF001 region 2, OF001 region 5, and OF001 region 7 based on the genome location retrieved by PHASTER. A summary of the distribution and genetic features of these prophages is shown in Additional file 1: Fig. S7b. All prophages in OF001 exhibited structural proteins, including major capsid, fiber, and tail proteins.

Based on the proteomic tree generated with the VIP-Tree server, all complete prophages in OF001 belong to the order Caudovirales (Additional file 1: Fig. S8). OF001 region 2 and OF001 region 5 were classified in the family Siphoviridae, while OF001 region 7 was classified in the family Myoviridae. Interestingly, OF001 region 2 and OF001 region 7 display putative site-specific integrases and excisionases, indicating site-specific recombination [100]. Multiple prophages have already been observed in other members of the genus *Pseudomonas* [101–104].

Within A210 genome, two incomplete prophages of 28.8 and 9.9 kb were detected (Additional file 1: Table S9 and Fig. S7c). No complete prophage was identified.

##### CRISPR-Cas systems

Using the CRISPRCas finder tool, we identified one complete class 1 CRISPR-Cas system, with a level of confidence of 4 (levels from 1 to 4, representing level 4 the most confident identification [105]) in the genome of OF001 (Additional file 1: Table S10 and S11). The seven *cas* genes are downstream of the repeat/spacer region. The repeats and spacers were compared with the CRISPRCas database [106] and were highly similar to sequences found in other bacteria including some *Pseudomonas* spp. like *P. stutzeri*, *P. aeruginosa*, and *Pseudomonas* sp. phDV1 (Additional file 1: Table S12 and S13). From the taxonomically closest organisms to strain OF001, only *P. guandongensis* has one confirmed class 1 CRISPR/Cas system.

In the genome of A210, one complete class 1 CRISPR-Cas system, with a level of confidence of 4, and one CRISPR region without *cas* genes associated, with a level of confidence of 2 were identified (Additional file 1: Table S10 and S11). The three *cas* genes associated to

the complete CRISPR-Cas loci are downstream of the repeat/spacer region. The repeats and spacers of the CRISPR region with a level of confidence 4 were compared with the CRISPRCas database [106] and, for the majority of them, no matching were found (Additional file 1: Table S14 and S15). For the three spacers and two repeats, only results with a similarity around 50% were found. The organisms were these repeats and spacers were found are *Verrucomicrobium spinosum*, *Raphidiopsis curvata*, *Pectobacterium carotovorum* and *Opitutaceae bacterium*. The taxonomically closest organisms to strain A210, *R. gelatinosus* and *R. benzoatilyticus*, have two CRISPRs without associated *cas* genes, and 4 incomplete and 2 complete CRISPR-Cas systems, respectively.

Class 1 CRISPR-Cas systems, as the one found in both MOB, are the most abundant class in Beta and Gamma-proteobacteria, and in general in archaea and bacteria [107].

The presence of CRISPR-Cas systems in strain OF001 and A210 might represent protection from phage infections, but could represent a disadvantage if useful genes for competitive adaptation cannot be acquired via external DNA [108].

Together the presence of genomic islands, including phage material and CRISPR-Cas systems in *Pseudomonas* sp. OF001 and *Rubrivivax* sp. A210 suggest that both MOB have undergone diverse genetic changes related to different horizontal gene transfer mechanisms which likely contribute to their genome plasticity.

#### Implications of the metabolic potential of strains OF001 and A210

In this study, we aimed at a better understanding of the metabolic capacities of the two CYN removing MOB which could potentially contribute to the biotechnological use of MOB for the removal of pollutants from water. In agreement with the genomes of other MOB with so far uncharacterized degradation ability [35, 36], the content of the genomes of strain OF001 and strain A210 suggests a potential metabolic versatility and thus, a broader application potential.

The genomic potential of MOB strains OF001 and A210 for the degradation of different organic compounds via specific enzymatic pathways might complement the unspecific transformation pathways of substances like diclofenac [109] and CYN [2, 3], via manganese oxidation. Our results suggest that strain OF001 and A210 might be able to remove different organic pollutants by a coupled mechanism involving specific enzymatic activity and unspecific oxidation by the reactive manganese species, as has been observed for the removal of phenolic compounds, which are common wastewater pollutants [110]. Moreover, it seems likely that the MOB described in this study transform other

organic compounds like carbofuran, ciprofloxacin, and 17 $\alpha$ -ethinylestradiol, similar to other MOB [30, 111, 112].

Both analyzed MOB, strain OF001 and A210, transform CYN indirectly through the oxidation of Mn<sup>2+</sup> [3], and according to the results of this study most likely mediated by the activity of multicopper oxidases and haem peroxidases. The unspecific transformation of CYN by MOB does not require an adaptation phase or a pre-conditioned towards the toxin as it is known for many enzymatically catalyzed processes. Therefore, the use of MOB to remove the only periodically occurring CYN molecule, might represent an advantage in comparison to other biological removal processes that require a pre-conditioning with the toxin to remove it. Moreover, the unspecific oxidation of organic pollutants via reactive manganese species might allow for the removal of other cyanotoxins, however further studies are required.

The different metabolic pathways encoded in the genome of strain OF001 and strain A210 also suggest different fields of application aiming at the removal of pollutants. For instance, the ability of strain A210 to thrive and degrade CYN in the absence of an organic carbon source suggests that it is more suitable for an application in settings, in which readily degradable organic carbon sources are depleted, such as reactors for the removal of pollutants from secondary wastewater. Also, due to the metabolic potential of strain A210, it may adapt within the reactor or the biofilm to varying oxygen concentrations or even the depletion of oxygen by a shift to nitrate respiration [113]. Moreover, both strains are able to form biofilms which may allow them to establish and be retained on fixed bed reactors.

*Pseudomonas* sp. OF001 showed the highest CYN removal efficiency and the fastest growth from all tested MOB [2]. Furthermore, it was isolated from a fixed-bed reactor system, however, the genome of strain OF001 encodes less diverse metabolic pathways to adapt to changing environments. Together, this data suggests that studies investigating degradation potential of MOB should consider the phylogenetic and metabolic diversity of MOB to identify the most suitable organisms that fulfil the requirements of the removal system.

The metabolic diversity of strains OF001 and A210 also suggests an important role of MOB in the removal of CYN in different habitats. For instance, strain A210 was isolated from a freshwater lake in the National Park Lower Oder Valley in Germany. This strain has the metabolic potential to dissimilatory reduced nitrate, which is an important mechanism to control nitrogen loading in aquatic environments [114–116]. Dissimilatory nitrate reduction to ammonium has been related to the promotion of eutrophic conditions in water systems, due to the release of ammonium that could be used

preferentially by cyanobacteria, and therefore favouring cyanobacterial blooms [116]. MOB strains with the ability to denitrify and degrade CYN may be therefore tightly interconnected with the production and removal of the cyanotoxin. Furthermore, the metabolic versatility of MOB may allow them to inhabit sediments and water columns. Therefore, MOB might contribute to the removal of CYN produced by benthic organisms in sediments, but also might transform CYN produced by planktonic cyanobacteria in the water column. However, further studies on the occurrence and distribution of MOB in CYN contaminated environments are required.

## Conclusions

In summary, this study provides an insight into the molecular basis of  $Mn^{2+}$  oxidation, and into the metabolic potential of two CYN-transforming MOB strains. We identified sequences in *Pseudomonas* sp. OF001 and *Rubrivivax* sp. A210 that show high similarity to already described MCOs which may catalyze manganese oxidation required for CYN transformation. Furthermore, considering the mechanism proposed for the removal of other pollutants by MOB the multicopper oxidases found in both strains and the haem peroxidase identified in strain OF001 might convey the ability to both strains to transform also other pollutants susceptible to reactive Mn species. Both MOB share the potential to grow over a wide range of  $O_2$  concentrations, to fix nitrogen, and reduce nitrate and sulfate via the assimilatory pathway. Both strains encode pathways that might enable them to remove different aromatic compounds such as benzoate, benzene, and phenol. However, while strain A210 harbors the genomic potential to fix  $CO_2$  and to reduce nitrate as final respiratory electron acceptor, strain OF001 requires additional organic carbon sources and lacks the ability for dissimilatory nitrate reduction. The analysis of the general metabolism of two MOB strains able to remove organic pollutants such as CYN and DCF might help to implement MOB in biotechnological applications and contributes to a better understanding of the natural niches of CYN-removing MOB in natural habitats.

## Methods

### Strains, culturing conditions and genomic DNA extraction

*Pseudomonas* sp. OF001 and *Rubrivivax* sp. A210 were obtained from the culture collection of the Laboratory of Environmental Microbiology from the TU Berlin, Germany [2]. Bacteria were routinely cultivated in a medium that was originally developed for *Leptothrix* strains [117], which was modified by our research group and is known as LSM2.

Cells from a pure, fresh 50 mL liquid culture from each strain were harvested by centrifugation at 15,000  $\times g$  for 3 min and washed three times with sterile Milli Q

water under sterile conditions. Total genomic DNA was extracted using the GeneMATRIX Soil DNA Purification Kit (EUR<sub>x</sub> Gdańsk, Poland) following the manufacturer's instructions. Quality and quantity of the extracted DNA was determined using Qubit<sup>TM</sup> fluorometric quantitation and NanoDrop 2000 (both Thermo Fisher Scientific, Bremen, Germany).

### Genome sequencing, assembly and annotation

The genome of both MOB strains was sequenced on an Illumina MiSeq platform with a read length of 301 bp (paired end). The genome of each isolate was assembled using SPAdes 3.10.1 and draft genomes obtained using manual binning procedures based on coverage-GC plots performed in R 3.6.1 (Additional file 1: Fig. S9). Genome quality estimation based on completeness and contamination was determined with CheckM [118]. Genome annotation was performed with the interface Magnifying Genomes (MaGE) of the MicroScope web-based service from GenoScope [119]. Protein coding genes were classified based on the annotation into Cluster of Orthologous Groups (COG) functional categories [120] with the automatic classification COG tool at Microscope platform. Function and pathway analysis were performed using BlastKOALA web tool of KEGG (Kyoto Encyclopedia of Genes and Genomes) database according to the KEGG groups of orthologs [121], and using MicroCyc tool of the MicroScope web-based service from GenoScope [119] which is a collection of microbial Pathway/Genome databases (PGDBs). PGDBs within MicroScope are generated by comparing the genome annotations to the metabolic reference database MetaCyc [122]. In the present work, metabolic potential will refer to the possibility of the strains to follow a specific metabolic pathway based only on their genome information, without being so far experimentally corroborated.

The data for this study have been deposited in the European Nucleotide Archive (ENA) at EMBL-EBI under project number PRJEB40009 with accession numbers GCA\_904426495 and GCA\_904426505 for strain OF001 and A210, respectively (<https://www.ebi.ac.uk/ena/browser/view/PRJEB40009>).

### Genomes comparison

First classification of the genomes was determined according to the Genome Taxonomy Database (GTDB) using the GTDB-tool kit (GTDB-tk) v.1.1.0 integrated in the MicroScope web-based service [44, 123, 124]. GTDB-tk provides a taxonomic classification of bacterial and archaeal genomes based on the combination of the GTDB reference tree, the relative evolutionary divergence and the ANI value against reference genomes [123]. GTDB proposed a bacterial taxonomy based on the phylogeny inferred from the concatenation of 120

ubiquitous single-copy proteins that normalizes taxonomic ranks by using the relative evolutionary divergence [124]. Therefore, it is considered that the analysis performed by GTDB-tk has an advantage over other phylogenies currently in use [124].

Genomes sequences were uploaded to the Type strain genome server (TYGS), a free bioinformatics platform (<https://tygs.dsmz.de>) for a whole genome-based taxonomic analysis [125]. TYGS platform runs automatically all the analysis. Briefly, TYGS performed first a determination of closely related type strain genomes, comparing the query genome against all available genomes in the TYGS database with the MASH algorithm [126] and selecting ten type strains. Then, additionally ten close related type strains were determined based on the 16S rRNA sequence extracted from the query genome using RNAmmer [127]. 16S rRNA sequences were compared with BLAST [128] against the TYGS database. The best 50 matching types were used to calculate precise distances using the Genome BLAST distance phylogeny approach (GBDP) [37]. The distances calculated by GBDP were then used to determine the ten closest type strain genomes for each query. Afterwards, GBDP conducted all pairwise comparisons among the set of genomes selected in the previous steps, and inferred accurate intergenomic distances under the algorithm “trimming” and distance formula  $d_5$  [37]. One hundred distance replicates were calculated each. In silico DNA-DNA hybridization (DDH) analysis were calculated using the recommended settings of the Genome-to-genome distance calculator (GGDC) 2.1 [37]. The resulting intergenomic distances were used to infer a balanced minimum evolution tree with branch support via FASTME 2.1.4 including subtree pruning and regrafting (SPR) postprocessing [129]. Branch support was inferred from 100 pseudo-bootstrap replicates each.

JSpeciesWS [40] was used to calculate the average nucleotide identity (ANI) values [40] based on BLAST (ANiB) [38, 128] and MUMmer (ANIm) [130], and to calculate the correlation indexes of the tetra-nucleotide frequencies (TETRA) [131].

For the ANI and TETRA analysis, the genome of strain OF001 was compared to the Pseudomonads belonging to the *Pseudomonas*\_K group: *P. oryzae* (GCA\_900104805.1), *P. sagittaria* (GCA\_900109175.1), *P. guangdongensis* (GCA\_900105885.1), and *P. liyingensis* (GCA\_900115715.1).

For the ANI and TETRA analysis, the genome of strain A210 was compared to the genomes of the three species of the genus *Rubrivivax*: *R. benzoatilyticus* JA2 (GCA\_000420125.1), *R. gelatinosus* IL144 (GCA\_000284255.1), *R. gelatinosus* DSM 1709 (GCA\_00430905.1), and *R. albus* ICH-03 (GCA\_004016515.1).

### Core- and pan-genome

Determination of the core- and pan-genome analysis was performed with the Pan/Core-genome tool from the MicroScope web-based service [119]. The analysis is based on the computation of Microscope gene families (MICFAM) using a single linkage clustering algorithm of homologous genes sharing an amino-acid alignment coverage and identity above the defined threshold [45]. This analysis considered i) any MICFAM associated with at least one gene from every genome used for the comparison as a part of the core-genome, ii) any MICFAM associated with at least 2 compared genomes as a part of the variable- genome, and iii) the sum of the core-genome and variable-genome as the pan-genome [44]. Parameter of 50/80 was selected (50% amino-acid identity, 80% amino-acid alignment coverage). All bacterial genomes used for the comparison with the genomes of strain OF001 or strain A210 that were not available in the MicroScope database were also annotated with MaGe from GenoScope [119].

For the pan- and core-genome analysis, the same strains as for the ANI and TETRA analysis, were used.

### Manganese-oxidation genes

We used the blastp function on the Microscope web server [128, 132] to identify potential  $Mn^{2+}$  oxidases in *Pseudomonas* sp. OF001 and *Rubrivivax* sp. A210, using experimentally verified  $Mn^{2+}$  oxidases of other manganese-oxidizing bacteria. Nine sequences of multicopper oxidases and three sequences of haem peroxidases related with the oxidation of  $Mn^{2+}$  in other MOB were used for the search (Table S1). Multicopper oxidases and haem peroxidases with  $Mn^{2+}$  oxidation activity will be referred as MO-mco and MO-hpox, respectively. We considered as homologue any protein with an E-value lower than  $10^{-10}$ .

$Mn^{2+}$  oxidases and putative homologues found in OF001 and A210 were functionally analyzed with the InterPro web server [133]. InterPro web server classifies proteins into families, and predicts functional domains and important sites of the proteins, integrating protein signatures from 13 different databases. We predicted the sub-cellular localization with LocTree3 [52] of those putative homologues without a predicted cytoplasmic or non-cytoplasmic domain according to the InterPro analysis.

To determine a possible phylogenetic relationship between manganese-oxidizing multicopper oxidases (MO-mco) and non-manganese oxidizing multicopper oxidases (non-MO-mco), we created a dataset sequences of multicopper oxidases with experimental evidence of  $Mn^{2+}$  oxidation [47–49, 53, 134, 135] and multicopper oxidases with experimental evidence of non- $Mn^{2+}$  oxidation activity [47, 54] and included our sequences. They

were aligned using MUSCLE [136] in MEGA v7.0.25. A phylogenetic tree was constructed with Maximum Likelihood method in MEGA v7.0.25. A bootstrap analysis was performed with 1000 replicates for the Maximum Likelihood tree.

### Operon prediction

Operon prediction was done using the FGENESB program [137]. FGENESB gene prediction algorithm is based on Markov chain models of coding regions, start of translation, and termination sites. Predicted genes are then used for the operon models using distances between ORFs frequencies of neighboring genes in known bacterial genomes, and positions of predicted promoters and terminators [137].

### Siderophores

Identification of siderophore biosynthesis gene clusters was performed with AntiSMASH tool [138] from the MicroScope web-based service. In addition, we used the blastp function on the Microscope web server [128, 132] to search for genes previously reported for the biosynthesis of siderophores pyoverdine, enterobactin, yersinia-bactin, ornibactin and pyochelin [139–142].

### Elements potentially acquired by horizontal gene transfer Genomic islands

Genomic islands were predicted with the IslandPath-DIMOB [143] and SIGI-HMM [144] method included in the IslandViewer 4 tool using the default settings [145]. Among the prediction methods included in IslandViewer 4 tool, SIGI-HMM has the highest precision and overall accuracy [146].

### Prophages

Putative phages from *Pseudomonas* sp. OF001 and *Rubrivivax* sp. A210 were predicted with PHASTER (PHAge Search Tool Enhanced Release) web server [147]. PHASTER classifies genome regions with a score below 70 as incomplete, between 70 to 90 as questionable, and greater than 90 as complete prophages [147].

The resulting complete prophage genomes were annotated with multiPhATE v.1.0 (multiple-genome Phage Annotation Toolkit and Evaluator) [148] using Phanotate to predict ORFs [149]. PhAnToMe (Phage Annotation Tools and Methods), pVOGs [150], and SwissProt [151] databases were used for the identification of the homologs of the input genomes and its predicted gene and peptide sequences. Additionally, highly divergent structural proteins were detected with iVireons [152] and confirmed with VIRALPro [153].

To classify the complete prophages, a whole proteomic tree based on genome-wide similarities was computed by tBLASTx, using the VIPTree web server v.1.9 [154].

### CRISPR-Cas systems

The presence of Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) and their associated genes (*cas*) was evaluated with CRISPRCasFinder [105]. CRISPRCasFinder include a rating system which classifies the detected CRISPRs to differentiate between CRISPR-like elements and true CRISPRs. Evidence levels from 1 to 4 are assigned, with 1 representing the lowest evidence classification and 4 the most confident identification [105].

Spacers and repeat regions detected in both genomes were searched with BLAST (blastn) against the CRISPRCasdb to identify their presence in other organisms. CRISPRCasdb contains CRISPR arrays and *cas* genes from complete genome sequences [106].

### Data graphics

Figures were made with the R packages ggplot2 (Wickham, 2016), gridExtra (Auguie, 2017), pheatmap [155], VennDiagram [156], and gggenes [157], using Viridis [158] and RcolorBrewer [159] packages for colouring in RStudio version 1.0.153 [160, 161]. Genomic maps of prophages were generated using the Snapgene® software (GSL Biotech). Trees generated with the TYGS tool and the multicopper oxidase tree were visualized and annotated with the online server iTOL [162].

### Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12864-021-07766-0>.

**Additional file 1: Fig. S1.** Phylogenetic tree based on 16S rDNA sequences and whole genome sequences including strain OF001 sequence. Tree inferred with FastME 2.1.6.1 [129] from GBDP distances calculated from a) 16S rDNA gene sequences and b) genome sequences. The branch lengths are scaled in terms of GBDP distance formula  $d_s$ . The numbers above branches are GBDP pseudo-bootstrap support values > 60% from 100 replications, with an average branch support of a) 68.8% and b) 92.5%. Tree was rooted at the midpoint [163]. Bold text represent the sequences generated in the present work. Scale bar represent sequence divergence. **Fig. S2.** Phylogenetic tree based on 16S rDNA sequences and whole genome sequences including strain A210 sequence. Tree inferred with FastME 2.1.6.1 [129] from GBDP distances calculated from a) 16S rDNA gene sequences and b) genome sequences. The branch lengths are scaled in terms of GBDP distance formula  $d_s$ . The numbers above branches are GBDP pseudo-bootstrap support values > 60% from 100 replications, with an average branch support of a) 76.8% and b) 83.4%. Tree was rooted at the midpoint [163]. Bold text represent the sequences generated in the present work. Scale bar represent sequence divergence. **Fig. S3.** Pan- and core genome overview. Venn diagram shows the number of shared and specific Microscope gene families (MICFAM) a) among *Pseudomonas* sp. OF001 and the members of the *Pseudomonas\_K* group, and b) among *Rubrivivax* sp. A210 and the members of the *Rubrivivax* genus. MICFAM grouping was based on 50% amino acid identity cut-off and at least 80% amino-acid alignment coverage. **Fig. S4.** Pan- and core- genome sizes estimated evolution. a, c) Number of MICFAM families in the pan-genome size by the number of genomes, and b, d) number of MICFAM families in the core-genome by the number of genomes. a, b) Including *Pseudomonas* sp. OF001, and c, d) including *Rubrivivax* sp. A210. **Fig. S5.** Maximum Likelihood phylogenetic tree based on multicopper oxidases sequences with and without reported

Mn<sup>2+</sup> oxidation activity. Sequences of the studied strains in the present study are not included. Numbers in the branches represent bootstrap value. Scale bar represent sequence divergence. **Fig. S6.** Putative genomic islands harbored by the studied MOB. a) *Pseudomonas* sp. OF001, and b) *Rubrivivax* sp. A210. Outer circle represents the genome size in Mbps. Genomic islands obtained by different prediction methods are highlighted in color. Integrated represent those islands detected by at least one method. **Fig. S7.** Distribution and genetic features of prophages detected in *Pseudomonas* sp. OF001 and *Rubrivivax* sp. A210. a) Circular genome map of strain OF001, b) genetic features of the complete prophages in strain OF001, and b) circular genome map of strain A210. In the genome maps location of prophages are highlighted in colors depending on the completeness of the prophages (Table S9). Number assigned to each prophage region is based on the genome location retrieved by PHASTER [147]. **Fig. S8.** Whole proteomic tree of *Pseudomonas* sp. OF001 prophages based on genome-wide similarities computed by tBLASTx. The tree was constructed using the VIPTree web server v.1.9 [154]. Numbers in brackets in the figure legend represent the number of virus genomes. Red stars represent the three complete prophages of strain OF001. Scale bar represent sequence divergence. **Fig. S9.** Coverage-GC plots of contig properties for strain OF001 and A210. a, c) Coverage vs GC content, and b, d) coverage vs length. Both samples has a primary, high abundance cluster of contigs, with GC centered around 0.6 with arise form the primary culture populations. For strain OF001 a second contig cluster with GC centered around 0.35 at a much lower abundance was detected, which might represent a slight DNA contamination. **Table S1.** List of genes in strains OF001 and A210 with homology to putative Mn<sup>2+</sup> oxidases from other MOB. **Table S2.** Functional domains and ontology classification of MO-mco and MO-hpox from MOB. **Table S3.** List of sequences of MO-mco and non- Mn<sup>2+</sup> oxidases. **Table S4.** List of genes related to the metabolic potential of *Pseudomonas* sp. OF001 and *Rubrivivax* sp. A210. **Table S5.** List of cytochrome genes within *Pseudomonas* sp. OF001 and *Rubrivivax* sp. A210. **Table S6.** List of genes related to cell motility and biofilm formation within *Pseudomonas* sp. OF001 and *Rubrivivax* sp. A210. **Table S7.** List of genes related to degradation of organic compounds in *Pseudomonas* sp. OF001 and *Rubrivivax* sp. A210. **Table S8.** Genomic islands identified within *Pseudomonas* sp. OF001 and *Rubrivivax* sp. A210 genome sequences. **Table S9.** Characteristics of prophage regions identified in *Pseudomonas* sp. OF001 and *Rubrivivax* sp. A210 genome. **Table S10.** CRISPR-Cas systems detected within *Pseudomonas* sp. OF001 and *Rubrivivax* sp. A210 genome. **Table S11.** Characteristics of the *cas* genes detected within *Pseudomonas* sp. OF001 and *Rubrivivax* sp. A210 genome. **Table S12.** Comparison of repeats in the CRISPR of confidence level 4 found in strain OF001 to CRISPRCasdb. **Table S13.** Comparison of spacers in the CRISPR of confidence level 4 found in strain OF001 to CRISPRCasdb. **Table S14.** Comparison of repeats in the CRISPR of confidence level 4 found in strain A210 to CRISPRCasdb. **Table S15.** Comparison of spacers in the CRISPR of confidence level 4 found in strain A210 to CRISPRCasdb.

#### Acknowledgments

The LABGeM (CEA/Genoscope & CNRS UMR8030), the France Génomique and French Bioinformatics Institute national infrastructures (funded as part of Investissement d'Avenir program managed by Agence Nationale pour la Recherche, contracts ANR-10-INBS-09 and ANR-11-INBS-0013) are acknowledged for support within the MicroScope annotation platform.

#### Authors' contributions

EBMR discussed the conceptualization and methodology of the project, extracted the DNA and performed all the bioinformatic analysis except those related with the phages and the assembly of the genomes, and wrote the original draft including all tables and figures. MC supervised and discussed the conceptualization and methodology of the project, and participated in the revision and edition of the manuscript. JBC performed the analysis related with the phages, and participated in the revision of the manuscript. IB coordinated the sequencing submission, and participated in the revision of the manuscript. MASH performed the assembly of the genomes, and participated in the revision of the manuscript. RW provided resources for the research, and participated in the revision of the manuscript. US supervised, provided resources for the research, and participated in the revision and

edition of the manuscript. The author(s) read and approved the final manuscript.

#### Funding

Erika B. Martínez-Ruiz was supported by a research scholarship from the DAAD (Deutscher Akademischer Austauschdienst). Contributions of Mindia A. S. Haryono, Irina Bessarab, and Rohan B. H. Williams are funded by the Singapore National Research Foundation and Ministry of Education under the Research Centre of Excellence Programme. We acknowledge support by the German Research Foundation and the Open Access Publication Fund of TU Berlin. Open Access funding enabled and organized by Projekt DEAL.

#### Availability of data and materials

The data for this study have been deposited in the European Nucleotide Archive (ENA) at EMBL-EBI under project number PRJEB40009 with accession numbers GCA\_904426495 and GCA\_904426505 for strain OF001 and A210, respectively.

#### Declarations

##### Ethics approval and consent to participate

Not applicable.

##### Consent for publication

Not applicable.

##### Competing interests

The authors declare that they have no competing interests.

#### Author details

<sup>1</sup>Chair of Environmental Microbiology, Technische Universität Berlin, Institute of Environmental Technology, Straße des 17. Juni 135, 10623 Berlin, Germany. <sup>2</sup>Singapore Centre for Environmental Life Sciences Engineering, National University of Singapore, Singapore 119077, Singapore.

Received: 9 October 2020 Accepted: 3 June 2021

Published online: 22 June 2021

#### References

- Poniedzialek B, Rzymyński P, Kokociński M. Cylindrospermopsin: water-linked potential threat to human health in Europe. *Environ Toxicol Pharmacol.* 2012;34(3):651–60. <https://doi.org/10.1016/j.etap.2012.08.005>.
- Martínez-Ruiz EB, Cooper M, Fastner J, Szewzyk U. Manganese-oxidizing bacteria isolated from natural and technical systems remove cylindrospermopsin. *Chemosphere.* 2020;238:124625. <https://doi.org/10.1016/j.chemosphere.2019.124625>.
- Martínez-Ruiz EB, Cooper M, Al-Zeer MA, Kurreck J, Adrian L, Szewzyk U. Manganese-oxidizing bacteria form multiple cylindrospermopsin transformation products with reduced human liver cell toxicity. *Sci Total Environ.* 2020;729:138924. <https://doi.org/10.1016/j.scitotenv.2020.138924>.
- Mohamed ZA, Alamri SA. Biodegradation of cylindrospermopsin toxin by microcystin-degrading bacteria isolated from cyanobacterial blooms. *Toxicon.* 2012;60(8):1390–5. <https://doi.org/10.1016/j.toxicon.2012.10.004>.
- Dziga D, Kokociński M, Maksylewicz A, Czaja-Prokop U, Barylski J. Cylindrospermopsin biodegradation abilities of *Aeromonas* sp. isolated from Rusalka Lake. *Toxins (Basel).* 2016;8:55.
- Nybom SMK, Salminen SJ, Meriluoto JAO. Specific strains of probiotic bacteria are efficient in removal of several different cyanobacterial toxins from solution. *Toxicon.* 2008;52:214–20.
- Kumar P, Hegde K, Brar SK, Cledon M, Kermanshahi-pour A. Potential of biological approaches for cyanotoxin removal from drinking water: a review. *Ecotoxicol Environ Saf.* 2019;172:488–503. <https://doi.org/10.1016/j.ecoenv.2019.01.066>.
- Kormas KA, Lymeropoulou DS. Cyanobacterial toxin degrading bacteria: who are they? *Biomed Res Int.* 2013;2013:463894.
- Yang W, Zhang Z, Zhang Z, Chen H, Liu J, Ali M, et al. Population structure of manganese-oxidizing bacteria in stratified soils and properties of manganese oxide aggregates under manganese-complex medium enrichment. *PLoS One.* 2013;8(9):e73778. <https://doi.org/10.1371/journal.pone.0073778>.

10. Tebo BM, Johnson HA, McCarthy JK, Templeton AS. Geomicrobiology of manganese(II) oxidation. *Trends Microbiol.* 2005;13:421–8.
11. Hansel CM, Learman DR. Geomicrobiology of manganese. In: Ehrlich H, Newman D, Kappler A, editors. *Ehrlich's Geomicrobiology*. 6th ed. Boca Raton: CRC Press; 2015. p. 403–33.
12. Schmidt B, Sánchez LA, Fretschner T, Kreps G, Ferrero MA, Siñeriz F, et al. Isolation of *Sphaerotilus-Leptothrix* strains from iron bacteria communities in Tierra del Fuego wetlands. *FEMS Microbiol Ecol.* 2014;90:454–66.
13. Szewzyk U, Szewzyk R, Schmidt B, Braun B. Neutrophilic iron-depositing microorganisms. In: Flemming H-C, Wingender J, Szewzyk U, editors. *Biofilm highlights*. Springer: Berlin Heidelberg; 2011. p. 63–79. <https://doi.org/10.1007/978-3-642-19940-0>.
14. Cerrato JM, Falkinham JO, Dietrich AM, Knocke WR, McKinney CW, Pruden A. Manganese-oxidizing and -reducing microorganisms isolated from biofilms in chlorinated drinking water systems. *Water Res.* 2010;44:3935–45.
15. Tobiason JE, Bazilio A, Goodwill J, Mai X, Nguyen C. Manganese removal from drinking water sources. *Curr Pollut Reports.* 2016;2(3):168–77. <https://doi.org/10.1007/s40726-016-0036-2>.
16. Zhang Y, Zhu H, Szewzyk U, Geissen SU. Removal of pharmaceuticals in aerated biofilters with manganese feeding. *Water Res.* 2015;72:218–26. <https://doi.org/10.1016/j.watres.2015.01.009>.
17. Tebo BM, Bargar JR, Clement BG, Dick GJ, Murray KJ, Parker D, et al. Biogenic manganese oxides: properties and mechanisms of formation. *Annu Rev Earth Planet Sci.* 2004;32(1):287–328. <https://doi.org/10.1146/annurev.ea.rth.32.101802.120213>.
18. Zhou H, Fu C. Manganese-oxidizing microbes and biogenic manganese oxides: characterization, Mn(II) oxidation mechanism and environmental relevance. *Rev Environ Sci Bio/Technol.* 2020;1(3):489–507. <https://doi.org/10.1007/s11157-020-09541-1>.
19. Hennebel T, De Gussem B, Boon N, Verstraete W. Biogenic metals in advanced water treatment. *Trends Biotechnol.* 2009;27:90–8.
20. Lee S, Xu H. XRD and TEM studies on nanophase manganese oxides in freshwater ferromanganese nodules from Green Bay, Lake Michigan. *Clays Clay Miner.* 2016;64:523–36.
21. Stein LY, La Duc MT, Grund TJ, Nealon KH. Bacterial and archaeal populations associated with freshwater ferromanganese micronodules and sediments. *Environ Microbiol.* 2001;3:10–8.
22. Kepkay PE, Nealon KH. Growth of a manganese oxidizing *Pseudomonas* sp. in continuous culture. *Arch Microbiol.* 1987;148:63–7.
23. Banh A, Chavez V, Doi J, Nguyen A, Hernandez S, Ha V, et al. Manganese (Mn) oxidation increases intracellular Mn in *Pseudomonas putida* GB-1. *PLoS One.* 2013;8:e77835.
24. Zerfuß C, Christie-Oleza JA, Soyer OS. Manganese oxide biomineralization provides protection against nitrite toxicity in a cell-density-dependent manner. *Appl Environ Microbiol.* 2019;85:5–8.
25. Jones ME, Nico PS, Ying S, Regier T, Thieme J, Keiluweit M. Manganese-driven carbon oxidation at oxic-anoxic interfaces. *Environ Sci Technol.* 2018;52:12349–57.
26. Sunda WG, Kieber DJ. Oxidation of humic substances by manganese oxides yields low-molecular-weight organic substrates. *Nature.* 1994;367:62–4.
27. Estes ER, Andeer PF, Nordlund D, Wankel SD, Hansel CM. Biogenic manganese oxides as reservoirs of organic carbon and proteins in terrestrial and marine environments. *Geobiology.* 2017;15:158–72.
28. Geszvain K, Butterfield C, Davis RE, Madison AS, Lee S-W, Parker DL, et al. The molecular biogeochemistry of manganese(II) oxidation. *Biochem Soc Trans.* 2012;40(6):1244–8. <https://doi.org/10.1042/BST20120229>.
29. Meerburg F, Hennebel T, Vanhaecke L, Verstraete W, Boon N. Diclofenac and 2-anilinophenylacetate degradation by combined activity of biogenic manganese oxides and silver. *Microb Biotechnol.* 2012;5(3):388–95. <https://doi.org/10.1111/j.1751-7915.2011.00323.x>.
30. Tran TN, Kim DG, Ko SO. Synergistic effects of biogenic manganese oxide and Mn(II)-oxidizing bacterium *Pseudomonas putida* strain MnB1 on the degradation of 17  $\alpha$ -ethinylestradiol. *J Hazard Mater.* 2018;344:350–9. <https://doi.org/10.1016/j.jhazmat.2017.10.045>.
31. Watanabe J, Tani Y, Chang J, Miyata N, Naitou H, Seyama H. As(III) oxidation kinetics of biogenic manganese oxides formed by *Acremonium strictum* strain KR21-2. *Chem Geol.* 2013;347:227–32. <https://doi.org/10.1016/j.chemgeo.2013.03.012>.
32. Wang H, Lv Z, Song Y, Wang YN, Zhang D, Sun Y, et al. Adsorptive removal of Sb(III) from wastewater by environmentally-friendly biogenic manganese oxide (BMO) materials: efficiency and mechanisms. *Process Saf Environ Prot.* 2019;124:223–30. <https://doi.org/10.1016/j.psep.2019.02.022>.
33. Wang R, Wang S, Tai Y, Tao R, Dai Y, Guo J, et al. Biogenic manganese oxides generated by green algae *Desmodesmus* sp. WR1 to improve bisphenol a removal. *J Hazard Mater.* 2017;339:310–9. <https://doi.org/10.1016/j.jhazmat.2017.06.026>.
34. Sochacki A, Felis E, Bajkacz S, Kalka J, Michalska JK. Removal and transformation of benzotriazole in manganese-oxide biofilters with Mn(II) feeding. *Chemosphere.* 2018;212:143–51. <https://doi.org/10.1016/j.chemosphere.2018.08.092>.
35. Dick GJ, Podell S, Johnson HA, Rivera-Espinoza Y, Bernier-Latmani R, McCarthy JK, et al. Genomic insights into Mn(II) oxidation by the marine alphaproteobacterium *Aurantimonas* sp. strain S185-9A1. *Appl Environ Microbiol.* 2008;74:2646–58.
36. Wang X, Yu M, Wang L, Lin H, Li B, Xue CX, et al. Comparative genomic and metabolic analysis of manganese-oxidizing mechanisms in *Celeribacter manganoxidans* DY25T: its adaptation to the environment of polymetallic nodules. *Genomics.* 2020;112(2):2080–91. <https://doi.org/10.1016/j.ygeno.2019.12.002>.
37. Meier-Kolthoff JP, Auch AF, Klenk HP, Göker M. Genome sequence-based species delimitation with confidence intervals and improved distance functions. *BMC Bioinformatics.* 2013;14(1):60. <https://doi.org/10.1186/1471-2105-14-60>.
38. Goris J, Konstantinidis KT, Klappenbach JA, Coenye T, Vandamme P, Tiedje JM. DNA-DNA hybridization values and their relationship to whole-genome sequence similarities. *Int J Syst Evol Microbiol.* 2007;57:81–91.
39. Richter M, Rosselló-Móra R. Shifting the genomic gold standard for the prokaryotic species definition. *Proc Natl Acad Sci U S A.* 2009;106:19126–31.
40. Richter M, Rosselló-Móra R, Oliver Glöckner F, Peplies J. JSpeciesWS: a web server for prokaryotic species circumscription based on pairwise genome comparison. *Bioinformatics.* 2016;32(6):929–31. <https://doi.org/10.1093/bioinformatics/btv681>.
41. Ramana CV, Sasikala C, Arunasri K, Anil Kumar P, Srinivas TNR, Shivaji S, et al. *Rubrivivax benzoatilyticus* sp. nov., an aromatic hydrocarbon-degrading purple betaproteobacterium. *Int J Syst Evol Microbiol.* 2006;56:2157–64.
42. Sheu SY, Li ZH, Young CC, Chen WM. *Rubrivivax albus* sp. nov., isolated from a freshwater pond. *Int J Syst Evol Microbiol.* 2020;70:805–13.
43. Nagashima S, Kamimura A, Shimizu T, Nakamura-Isaki S, Aono E, Sakamoto K, et al. Complete genome sequence of phototrophic betaproteobacterium *Rubrivivax gelatinosus* IL144. *J Bacteriol.* 2012;194:3541–2.
44. Vallenet D, Calteau A, Cruveiller S, Gachet M, Lajus A, Josso A, et al. MicroScope in 2017: an expanding and evolving integrated resource for community expertise of microbial genomes. *Nucleic Acids Res.* 2017;45:D517–28.
45. Miele V, Penel S, Duret L. Ultra-fast sequence clustering from similarity networks with SiLiX. *BMC Bioinformatics.* 2011;12:116.
46. Spiro TG, Bargar JR, Sposito G, Tebo BM. Bacteriogenic manganese oxides. *Acc Chem Res.* 2008;43:2–9.
47. Geszvain K, McCarthy JK, Tebo BM. Elimination of manganese(II, III) oxidation in *Pseudomonas putida* GB-1 by a double knockout of two putative multicopper oxidase genes. *Appl Environ Microbiol.* 2013;79:357–66. <https://doi.org/10.1128/AEM.01850-12>.
48. Ridge JP, Lin M, Larsen EI, Fegan M, Mcewan AG, Sly LI. A multicopper oxidase is essential for manganese oxidation and laccase-like activity in *Pedomicrobium* sp. ACM 3067. *Environ Microbiol.* 2007;9(4):944–53. <https://doi.org/10.1111/j.1462-2920.2006.01216.x>.
49. Corstjens PLAM, Devrind JPM, Goosen T, de Vrind-de Jong EW. Identification and molecular analysis of the *Leptothrix discophora* SS-1 *mofA* gene, a gene putatively encoding a manganese-oxidizing protein with copper domains. *Geomicrobiol J.* 1997;14(2):91–108. <https://doi.org/10.1080/01490459709378037>.
50. Dick GJ, Torpey JW, Beveridge TJ, Tebo BM. Direct identification of a bacterial manganese(II) oxidase, the multicopper oxidase MnxG, from spores of several different marine *Bacillus* species. *Appl Environ Microbiol.* 2008;74:1527–34.
51. Geszvain K, Tebo BM. Identification of a two-component regulatory pathway essential for Mn(II) oxidation in *Pseudomonas putida* GB-1. *Appl Environ Microbiol.* 2010;76:1224–31.
52. Goldberg T, Hecht M, Hamp T, Karl T, Yachdav G, Ahmed N, et al. LocTree3 prediction of localization. *Nucleic Acids Res.* 2014;42:350–5.



53. Geszvain K, Smesrud L, Tebo BM. Identification of a third Mn(II) oxidase enzyme in *Pseudomonas putida* GB-1. *Appl Environ Microbiol.* 2016;82(13):3774–82. <https://doi.org/10.1128/AEM.00046-16>.
54. Anderson CR, Johnson HA, Caputo N, Davis RE, Torpey JW, Tebo BM. Mn(II) oxidation is catalyzed by heme peroxidases in *Aurantimonas manganoxydans* strain S185-9A1 and *Erythrobacter* sp. strain SD-21. *Appl Environ Microbiol.* 2009;75:4130–8.
55. Brouwers GJ, Corstjens PLAM, De Vrind JPM, Verkamman A, De Kuyper M, De Vrind-De Jong EW. Stimulation of Mn<sup>2+</sup> oxidation in *Leptothrix discophora* SS-1 by Cu<sup>2+</sup> and sequence analysis of the region flanking the gene encoding putative multicopper oxidase MofA. *Geomicrobiol J.* 2000;17:25–33.
56. Stincone A, Prigione A, Cramer T, Wmeling MM, Campbell K, Cheung E, et al. The return of metabolism: biochemistry and physiology of the pentose phosphate pathway. *Biol Rev Camb Philos Soc.* 2015;90(3):927–63. <https://doi.org/10.1111/brv.12140>.
57. Ekcici S, Pawlik G, Lohmeyer E, Koch HG, Daldal F. Biogenesis of *cbb*<sub>3</sub>-type cytochrome *c* oxidase in *Rhodobacter capsulatus*. *Biochim Biophys Acta.* 1817;2012(2):898–910. <https://doi.org/10.1038/jid.2014.371>.
58. Koch HG, Hwang O, Daldal F. Isolation and characterization of *Rhodobacter capsulatus* mutants affected in cytochrome *cbb*<sub>3</sub> oxidase activity. *J Bacteriol.* 1998;180:969–78.
59. Kulajita C, Thumfart JO, Haid S, Daldal F, Koch HG. Multi-step assembly pathway of the *cbb*<sub>3</sub>-type cytochrome *c* oxidase complex. *J Mol Biol.* 2006;355:989–1004.
60. Preisig O, Zufferey R, Hennecke H. The *Bradyrhizobium japonicum* *fixGHIS* genes are required for the formation of the high-affinity *cbb*<sub>3</sub>-type cytochrome oxidase. *Arch Microbiol.* 1996;165(5):297–305. <https://doi.org/10.1007/s002030050330>.
61. Durand A, Bourbon ML, Steunou AS, Khalfaoui-Hassani B, Legrand C, Legrand A, et al. Biogenesis of the bacterial *cbb*<sub>3</sub> cytochrome *c* oxidase: active subcomplexes support a sequential assembly model. *J Biol Chem.* 2018;293:808–18.
62. Atkinson MR, Ninfa AJ. Characterization of the GlnK protein of *Escherichia coli*. *Mol Microbiol.* 1999;32(2):301–13. <https://doi.org/10.1046/j.1365-2958.1999.01349.x>.
63. Blauwkamp TA, Ninfa AJ. Physiological role of the GlnK signal transduction protein of *Escherichia coli*: survival of nitrogen starvation. *Mol Microbiol.* 2002;46:203–14.
64. Gosztolai A, Schumacher J, Behrends V, Bundy JG, Heydenreich F, Bennett MH, et al. GlnK facilitates the dynamic regulation of bacterial nitrogen assimilation. *Biophys J.* 2017;112:2219–30.
65. Magasanik B. Genetic control of nitrogen assimilation in bacteria. *Annu Rev Genet.* 1982;16:135–68.
66. Pahel G, Rothstein DM, Magasanik B. Complex *glnA-glnL-glnG* operon of *Escherichia coli*. *J Bacteriol.* 1982;150:202–13.
67. Schreier HJ, Sonenshein AL. Altered regulation of the *glnA* gene in glutamine synthetase mutants of *Bacillus subtilis*. *J Bacteriol.* 1986;167(1):35–43. <https://doi.org/10.1128/JB.167.1.35-43.1986>.
68. He WH, Wang YN, Du X, Zhou Y, Jia B, Bian J, et al. *Pseudomonas linyingensis* sp. nov.: a novel bacterium isolated from wheat soil subjected to long-term herbicides application. *Curr Microbiol.* 2012;65:595–600.
69. Lin SY, Hameed A, Liu YC, Hsu YH, Lai WA, Chen WM, et al. *Pseudomonas sagittaria* sp. nov., a siderophore-producing bacterium isolated from oil-contaminated soil. *Int J Syst Evol Microbiol.* 2013;63:2410–7.
70. Mohan SB, Cole JA. The dissimilatory reduction of nitrate to ammonia by anaerobic bacteria. In: Bothe H, Ferguson SJ, Newton WE, editors. *Biology of the Nitrogen Cycle*. Elsevier B.V.; 2007. p. 93–106.
71. Kawasaki S, Arai H, Kodama T, Igarashi Y. Gene cluster for dissimilatory nitrite reductase (*nir*) from *Pseudomonas aeruginosa*: sequencing and identification of a locus for heme *d*<sub>1</sub> biosynthesis. *J Bacteriol.* 1997;179:235–42.
72. Ramos F, Blanco G, Gutiérrez JC, Luque F, Tortolero M. Identification of an operon involved in the assimilatory nitrate-reducing system of *Azotobacter vinelandii*. *Mol Microbiol.* 1993;8(6):1145–53. <https://doi.org/10.1111/j.1365-2958.1993.tb01659.x>.
73. Yukioka Y, Tanahashi T, Shida K, Oguchi H, Ogawa S, Saito C, et al. A role of nitrite reductase (NirBD) for NO homeostatic regulation in *Streptomyces coelicolor* A3(2). *FEMS Microbiol Lett.* 2017;364:1–6.
74. Arcondéguy T, Jack R, Merrick M. PII signal transduction proteins, pivotal players in microbial nitrogen control. *Microbiol Mol Biol Rev.* 2001;65:80–105.
75. Yurgel SN, Rice J, Mulder M, Kahn ML. GlnB/GlnK PII proteins and regulation of the *Sinorhizobium meliloti* Rm1021 nitrogen stress response and symbiotic function. *J Bacteriol.* 2010;192(10):2473–81. <https://doi.org/10.1128/JB.01657-09>.
76. Sant'Anna FH, Trentini DB, De Souto WS, Cecagno R, Da Silva SC, Schrank IS. The PII superfamily revised: a novel group and evolutionary insights. *J Mol Evol.* 2009;68:322–36.
77. Egner T, Martin DE, Sarkar A, Reinhold-Hurek B. Role of a ferredoxin gene cotranscribed with the *nifHDK* operon in N<sub>2</sub> fixation and nitrogenase “switch-off” of *Azoarcus* sp. strain BH72. *J Bacteriol.* 2001;183:3752–60.
78. Martin AE, Burgess BK, Iismaa SE, Smartt CT, Jacobson MR, Dean DR. Construction and characterization of an *Azotobacter vinelandii* strain with mutations in the genes encoding flavodoxin and ferredoxin I. *J Bacteriol.* 1989;171(6):3162–7. <https://doi.org/10.1128/jb.171.6.3162-3167.1989>.
79. Poudel S, Colman DR, Fixen KR, Ledbetter RN, Zheng Y, Pence N, et al. Electron transfer to nitrogenase in different genomic and metabolic backgrounds. *J Bacteriol.* 2018;200:1–19.
80. Yakunin AF, Gennaro G, Hallenbeck PC. Purification and properties of a *nif*-specific flavodoxin from the photosynthetic bacterium *Rhodobacter capsulatus*. *J Bacteriol.* 1993;175:6775–80.
81. Sirko A, Zatyka M, Sadowy E, Hulanicka D. Sulfate and thiosulfate transport in *Escherichia coli* K-12: evidence for a functional overlapping of sulfate- and thiosulfate-binding proteins. *J Bacteriol.* 1995;177(14):4134–6. <https://doi.org/10.1128/JB.177.14.4134-4136.1995>.
82. Loughlin P, Sheldon MC, Tierney ML, Howitt SM. Structure and function of a model member of the SulP transporter family. *Cell Biochem Biophys.* 2002;36(2-3):183–90. <https://doi.org/10.1385/CBB:36:2-3:183>.
83. Van De Kamp M, Pizzini E, Vos A, Van Der Lende TR, Schuurs TA, Newbert RW, et al. Sulfate transport in *Penicillium chrysogenum*: cloning and characterization of the *sutA* and *sutB* genes. *J Bacteriol.* 1999;181:7228–34.
84. Parra F, Britton P, Castle C, Jones-Mortimer MC, Kornberg HL. Two separate genes involved in sulphate transport in *Escherichia coli* K12. *J Gen Microbiol.* 1983;129:357–8.
85. Rückert C, Koch DJ, Rey DA, Albersmeier A, Mormann S, Pühler A, et al. Functional genomics and expression analysis of the *Corynebacterium glutamicum* *fpr2-cyclXHDNYZ* gene cluster involved in assimilatory sulphate reduction. *BMC Genomics.* 2005;6:1–18.
86. Kushkevych I, Cejnar J, Tremel J, Dordević D, Kollar P, Vítězová M. Recent advances in metabolic pathways of sulfate reduction in intestinal bacteria. *Cells.* 2020;9:698.
87. Bick JA, Dennis JJ, Zylstra GJ, Nowack J, Leustek T. Identification of a new class of 5'-adenylylsulfate (APS) reductases from sulfate-assimilating bacteria. *J Bacteriol.* 2000;182(1):135–42. <https://doi.org/10.1128/JB.182.1.135-142.2000>.
88. Abola AP, Willits MG, Wang RC, Long SR. Reduction of adenosine-5'-phosphosulfate instead of 3'-phosphoadenosine-5'-phosphosulfate in cysteine biosynthesis by *Rhizobium meliloti* and other members of the family *Rhizobiaceae*. *J Bacteriol.* 1999;181:5280–7.
89. Iwanicka-Nowicka R, Zielak A, Cook AM, Thomas MS, Hryniewicz MM. Regulation of sulfur assimilation pathways in *Burkholderia cenocepacia*: identification of transcription factors CysB and SsuR and their role in control of target genes. *J Bacteriol.* 2007;189:1675–88.
90. Rivera M. Bacterioferritin: structure, dynamics, and protein-protein interactions at play in iron storage and mobilization. *Acc Chem Res.* 2017;50(2):331–40. <https://doi.org/10.1021/acs.accounts.6b00514>.
91. Tong Y, Guo M. Bacterial heme-transport proteins and their heme-coordination modes. *Arch Biochem Biophys.* 2009;481(1):1–15. <https://doi.org/10.1016/j.abb.2008.10.013>.
92. Liu R, Ochman H. Stepwise formation of the bacterial flagellar system. *Proc Natl Acad Sci U S A.* 2007;104:7116–21.
93. Yu Z, Chang M, Wu M, Yang G, Zhou S, Zhuang L. *Pseudomonas oryzae* sp. nov. isolated from a paddy soil in South China. *Arch Microbiol.* 2013;195(12):815–22. <https://doi.org/10.1007/s00203-013-0930-6>.
94. Yang G, Han L, Wen J, Zhou S. *Pseudomonas guangdongensis* sp. nov., isolated from an electroactive biofilm, and emended description of the genus *Pseudomonas* Migula 1894. *Int J Syst Evol Microbiol.* 2013;63 PART 12:4599–605.
95. Nogales J, García JL, Díaz E. Degradation of aromatic compounds in *Pseudomonas*: a systems biology view. In: Rojo F, editor. *Aerobic utilization of hydrocarbons, oils and lipids. Handbook of hydrocarbon and lipid microbiology*. Springer International Publishing AG; 2017. [https://doi.org/10.1007/978-3-319-39782-5\\_32-1](https://doi.org/10.1007/978-3-319-39782-5_32-1).

96. de la Cruz AA, Hiskia A, Kaloudis T, Chernoff N, Hill D, Antoniou MG, et al. A review on cylindrospermopsin: the global occurrence, detection, toxicity and degradation of a potent cyanotoxin. *Environ Sci Process Impacts*. 2013; 15(11):1979–2003. <https://doi.org/10.1039/c3em00353a>.
97. Reva ON, Bezuidt O. Distribution of horizontally transferred heavy metal resistance operons in recent outbreak bacteria. *Mob Genet Elements*. 2012; 2:96–100.
98. Loux V, Mariadassou M, Almeida S, Chiapello H, Hammani A, Buratti J, et al. Mutations and genomic islands can explain the strain dependency of sugar utilization in 21 strains of *Propionibacterium freudenreichii*. *BMC Genomics*. 2015;16:1–19.
99. Zhu B, Zhang X, Zhao C, Chen S, Yang S. Comparative genome analysis of marine purple sulfur bacterium *Marichromatium gracile* YL28 reveals the diverse nitrogen cycle mechanisms and habitat-specific traits. *Sci Rep*. 2018; 8:1–11. <https://doi.org/10.1038/s41598-018-36160-2>.
100. Grindley NDF, Whiteson KL, Rice PA. Mechanisms of site-specific recombination. *Annu Rev Biochem*. 2006;75:567–605.
101. Martínez-García E, Jatsenko T, Kivisaar M, de Lorenzo V. Freeing *Pseudomonas putida* KT2440 of its proviral load strengthens endurance to environmental stresses. *Environ Microbiol*. 2014;17:76–90.
102. Dziejewit L, Radlinska M. Two inducible prophages of an antarctic *Pseudomonas* sp. ANT\_H14 use the same capsid for packaging their genomes - characterization of a novel phage helper-satellite system. *PLoS One*. 2016;11:1–27.
103. Paulsen IT, Press CM, Ravel J, Kobayashi DY, Myers GSA, Mavrodj DV, et al. Complete genome sequence of the plant commensal *Pseudomonas fluorescens* pf-5. *Nat Biotechnol*. 2005;23:873–8.
104. de Sousa LP. Mobile genetic elements in *Pseudomonas stutzeri*. *Curr Microbiol*. 2020;77:179–84. <https://doi.org/10.1007/s00284-019-01812-7>.
105. Couvin D, Bernheim A, Toffano-Nioche C, Touchon M, Michalik J, Néron B, et al. CRISPRCasFinder, an update of CRISPRFinder, includes a portable version, enhanced performance and integrates search for Cas proteins. *Nucleic Acids Res*. 2018;46(W1):W246–51. <https://doi.org/10.1093/nar/gky425>.
106. Pourcel C, Touchon M, Villeriot N, Vernadet JP, Couvin D, Toffano-Nioche C, et al. CRISPRCasdb a successor of CRISPRdb containing CRISPR arrays and *cas* genes from complete genome sequences, and tools to download and query lists of repeats and spacers. *Nucleic Acids Res*. 2020;48:D535–44.
107. Makarova KS, Wolf YI, Iranzo J, Shmakov SA, Alkhnbashi OS, Brouns SJJ, et al. Evolutionary classification of CRISPR–Cas systems: a burst of class 2 and derived variants. *Nat Rev Microbiol*. 2020;18:67–83.
108. Bhaya D, Davison M, Barrangou R. CRISPR-Cas Systems in Bacteria and Archaea: versatile small RNAs for adaptive defense and regulation. *Annu Rev Genet*. 2011;45:273–97.
109. Nega M. Effects of micropollutants on the Panke River microbial communities and degradation of selected pharmaceuticals by manganese-oxidizing bacteria. *Technische Universität Berlin*; 2020. <http://dx.doi.org/10.14279/depositone-10452>.
110. Wang G, Liu Y, Wu M, Zong W, Yi X, Zhan J, et al. Coupling the phenolic oxidation capacities of a bacterial consortium and in situ-generated manganese oxides in a moving bed biofilm reactor (MBBR). *Water Res*. 2019;166:115047. <https://doi.org/10.1016/j.watres.2019.115047>.
111. Tu J, Yang Z, Hu C, Qu J. Characterization and reactivity of biogenic manganese oxides for ciprofloxacin oxidation. *J Environ Sci (China)*. 2014; 26(5):1154–61. [https://doi.org/10.1016/S1001-0742\(13\)60505-7](https://doi.org/10.1016/S1001-0742(13)60505-7).
112. Liu Z, Wang J, Qian S, Wang G, Wang J, Liao S. Carbofuran degradation by biogenic manganese oxides. *Bull Environ Contam Toxicol*. 2017;98:420–5. <https://doi.org/10.1007/s00128-016-1940-2>.
113. Wu Y, Shukal S, Mukherjee M, Cao B. Involvement in denitrification is beneficial to the biofilm lifestyle of *Comamonas testosteroni*: a mechanistic study and its environmental implications. *Environ Sci Technol*. 2015;49(19): 11551–9. <https://doi.org/10.1021/acs.est.5b03381>.
114. Burgin AJ, Hamilton SK, Gardner WS, McCarthy MJ. Nitrate reduction, denitrification, and dissimilatory nitrate reduction to ammonium in wetland sediments; 2015. p. 519–37.
115. Wang S, Pi Y, Song Y, Jiang Y, Zhou L, Liu W, et al. Hotspot of dissimilatory nitrate reduction to ammonium (DNRA) process in freshwater sediments of riparian zones. *Water Res*. 2020;173:115539. <https://doi.org/10.1016/j.watres.2020.115539>.
116. Li X, Song C, Zhou Z, Xiao J, Wang S, Yang L, et al. Comparison of community and function of dissimilatory nitrate reduction to ammonium (DNRA) bacteria in Chinese Shallow Lakes with different eutrophication degrees. *Water (Switzerland)*. 2020;12:1–15.
117. Atlas R. Handbook of microbiological media, fourth edition. Fourth edi. Washington, D. C: Taylor & Francis Group, LLC; 2010. <https://doi.org/10.1201/EBK1439804063>.
118. 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*. 2015;25:1043–55.
119. Vallenet D, Belda E, Calteau A, Cruveiller S, Engelen S, Lajus A, et al. MicroScope - an integrated microbial resource for the curation and comparative analysis of genomic and metabolic data. *Nucleic Acids Res*. 2013;41(D1):D636–47. <https://doi.org/10.1093/nar/gks1194>.
120. Tatusov RL. The COG database: a tool for genome-scale analysis of protein functions and evolution. *Nucleic Acids Res*. 2000;28:33–6.
121. Kanehisa M, Sato Y, Morishima K. BlastKOALA and GhostKOALA: KEGG tools for functional characterization of genome and metagenome sequences. *J Mol Biol*. 2016;428:726–31. <https://doi.org/10.1016/j.jmb.2015.11.006>.
122. Caspi R, Billington R, Fulcher CA, Keseler IM, Kothari A, Krummenacker M, et al. The MetaCyc database of metabolic pathways and enzymes. *Nucleic Acids Res*. 2018;46(D1):D633–9. <https://doi.org/10.1093/nar/gkx935>.
123. Chaumeil P-A, Mussig AJ, Hugenholtz P, Parks DH. GTDB-Tk: a toolkit to classify genomes with the genome taxonomy database. *Bioinformatics*. 2019;36:1925–7.
124. 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;36(10):996–1004. <https://doi.org/10.1038/nbt.4229>.
125. Meier-Kolthoff JP, Göker M. TYGS is an automated high-throughput platform for state-of-the-art genome-based taxonomy. *Nat Commun*. 2019;10. <https://doi.org/10.1038/s41467-019-10210-3>.
126. Ondov BD, Treangen TJ, Melsted P, Mallonee AB, Bergman NH, Koren S, et al. Mash: fast genome and metagenome distance estimation using MinHash. *Genome Biol*. 2016;17:1–14. <https://doi.org/10.1186/s13059-016-0997-x>.
127. Lagesen K, Hallin P, Rødland EA, Stærfeldt HH, Rognes T, Ussery DW. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Res*. 2007;35:3100–8.
128. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, et al. BLAST+: architecture and applications. *BMC Bioinformatics*. 2009;10:1–9.
129. Lefort V, Desper R, Gascuel O. FastME 2.0: a comprehensive, accurate, and fast distance-based phylogeny inference program. *Mol Biol Evol*. 2015;32: 2798–800.
130. Kurtz S, Phillippy A, Delcher AL, Smoot M, Shumway M, Antonescu C, et al. Versatile and open software for comparing large genomes. *Genome Biol*. 2004;5(2):R12. <https://doi.org/10.1186/gb-2004-5-2-r12>.
131. Teeling H, Meyerdiets A, Bauer M, Amann R, Glöckner FO. Application of tetranucleotide frequencies for the assignment of genomic fragments. *Environ Microbiol*. 2004;6:938–47.
132. Vallenet D, Calteau A, Dubois M, Amours P, Bazin A, Beuvin M, et al. MicroScope: an integrated platform for the annotation and exploration of microbial gene functions through genomic, pangenomic and metabolic comparative analysis. *Nucleic Acids Res*. 2020;48:D579–89.
133. Mitchell AL, Attwood TK, Babbitt PC, Blum M, Bork P, Bridge A, et al. InterPro in 2019: improving coverage, classification and access to protein sequence annotations. *Nucleic Acids Res*. 2019;47(D1):D351–60. <https://doi.org/10.1093/nar/gky1100>.
134. Francis CA, Tebo BM. Enzymatic manganese(II) oxidation by metabolically dormant spores of diverse *Bacillus* species. *Appl Environ Microbiol*. 2002;68: 874–80.
135. Tang W, Liu Y, Gong J, Chen S, Zeng X. Analysis of manganese oxidase and its encoding gene in *Lysinibacillus* strain MK-1. *Process Saf Environ Prot*. 2019;127:299–305. <https://doi.org/10.1016/j.psep.2019.04.002>.
136. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res*. 2004;32(5):1792–7. <https://doi.org/10.1093/nar/gkh340>.
137. Solovyev V, Salamov A. Automatic annotation of microbial genomes and metagenomic sequences. In: Li RW, editor. *Metagenomics and its applications in agriculture, biomedicine and environmental studies*: Nova Science Publishers; 2011. p. 61–78. [https://www.researchgate.net/publication/259450599\\_V\\_Solovyev\\_A\\_Salamov\\_2011\\_Automatic\\_Annotation\\_of\\_Microbial\\_Genomes\\_and\\_Metagenomic\\_Sequences\\_In\\_Metagenomics\\_and\\_its\\_Applications\\_in\\_Agriculture\\_Biomedicine\\_and\\_Environmental\\_Studies\\_Ed\\_RW\\_Li\\_Nova\\_Sc](https://www.researchgate.net/publication/259450599_V_Solovyev_A_Salamov_2011_Automatic_Annotation_of_Microbial_Genomes_and_Metagenomic_Sequences_In_Metagenomics_and_its_Applications_in_Agriculture_Biomedicine_and_Environmental_Studies_Ed_RW_Li_Nova_Sc).

138. Blin K, Shaw S, Steinke K, Villebro R, Ziemert N, Lee SY, et al. AntiSMASH 5.0: updates to the secondary metabolite genome mining pipeline. *Nucleic Acids Res.* 2019;47(W1):W81–7. <https://doi.org/10.1093/nar/gkz310>.
139. Schalk IJ, Rigouin C, Godet J. An overview of siderophore biosynthesis among fluorescent pseudomonads and new insights into their complex cellular organization. *Environ Microbiol.* 2020;22:1447–66.
140. Crosa JH, Walsh CT. Genetics and assembly line enzymology of siderophore biosynthesis in bacteria. *Microbiol Mol Biol Rev.* 2002;66(2):223–49. <https://doi.org/10.1128/MMBR.66.2.223-249.2002>.
141. Khan A, Singh P, Srivastava A. Synthesis, nature and utility of universal iron chelator – Siderophore: a review. *Microbiol Res.* 2018;212–213:103–11. <https://doi.org/10.1016/j.micres.2017.10.012>.
142. Visser MB, Majumdar S, Hani E, Sokol PA. Importance of the ornibactin and pyochelin siderophore transport systems in *Burkholderia cenocepacia* lung infections. *Infect Immun.* 2004;72:2850–7.
143. Hsiao W, Wan I, Jones SJ, Brinkman FSL. IslandPath: aiding detection of genomic islands in prokaryotes. *Bioinformatics.* 2003;19:418–20.
144. Waack S, Keller O, Asper R, Brodag T, Damm C, Fricke WF, et al. Score-based prediction of genomic islands in prokaryotic genomes using hidden Markov models. *BMC Bioinformatics.* 2006;7:1–12.
145. Bertelli C, Laird MR, Williams KP, Lau BY, Hoad G, Winsor GL, et al. IslandViewer 4: expanded prediction of genomic islands for larger-scale datasets. *Nucleic Acids Res.* 2017;45:W30–5.
146. Langille MGI, Hsiao WWL, Brinkman FSL. Evaluation of genomic island predictors using a comparative genomics approach. *BMC Bioinformatics.* 2008;9:1–10.
147. Arndt D, Grant JR, Marcu A, Sajed T, Pon A, Liang Y, et al. PHASTER: a better, faster version of the PHAST phage search tool. *Nucleic Acids Res.* 2016; 44(W1):W16–21. <https://doi.org/10.1093/nar/gkw387>.
148. Ecalle Zhou CL, Malfatti S, Kimbrel J, Philipson C, McNair K, Hamilton T, et al. multiPhATE: bioinformatics pipeline for functional annotation of phage isolates. *Bioinformatics.* 2019;35:4402–4.
149. McNair K, Zhou C, Dinsdale EA, Souza B, Edwards RA, Hancock J. PHANOTATE: a novel approach to gene identification in phage genomes. *Bioinformatics.* 2019; 35(22):4537–42. <https://doi.org/10.1093/bioinformatics/btz265>.
150. Graziotin AL, Koonin EV, Kristensen DM. Prokaryotic virus orthologous groups (pVOGs): a resource for comparative genomics and protein family annotation. *Nucleic Acids Res.* 2017;45:D491–8.
151. Bairoch A, Apweiler R. The SWISS-PROT protein sequence database and its supplement TrEMBL in 2000. *Nucleic Acids Res.* 2000;28:45–8.
152. Seguritan V, Alves N, Arnoult M, Raymond A, Lorimer D, Burgin AB, et al. Artificial neural networks trained to detect viral and phage structural proteins. *PLoS Comput Biol.* 2012;8:e1002657.
153. Galiez C, Magnan CN, Coste F, Baldi P. VIRALpro: a tool to identify viral capsid and tail sequences. *Bioinformatics.* 2016;32(9):1405–7. <https://doi.org/10.1093/bioinformatics/btv727>.
154. Nishimura Y, Yoshida T, Kuronishi M, Uehara H, Ogata H, Goto S. ViPTree: the viral proteomic tree server. *Bioinformatics.* 2017;33:2379–80.
155. Kolde R. pheatmap: Pretty heatmaps. R package version 1.0.8; 2015.
156. Chen H, Boutros PC. VennDiagram: a package for the generation of highly-customizable Venn and Euler diagrams in R. *BMC Bioinformatics.* 2011;12:35.
157. Wilkins D. gggenes: Draw gene arrow maps in “ggplot2.” R package version 0.4.0. 2019. <https://cran.r-project.org/package=gggenes>.
158. Garnier S. viridis: Default color maps from “matplotlib.” R package version 0.5.1. 2018. <https://cran.r-project.org/package=viridis>.
159. Neuwirth E. RColorBrewer: ColorBrewer palettes. R package version 1.1–2. 2014. <https://cran.r-project.org/package=RColorBrewer>.
160. R Core Team. R: A language and environment for statistical computing. 2020. <https://www.r-project.org/>.
161. R Core Team. RStudio: Integrated development for R. v1.0.153. 2016. <http://www.rstudio.com/>.
162. Letunic I, Bork P. Interactive tree of life (iTOL) v4: recent updates and new developments. *Nucleic Acids Res.* 2019;47:256–9.
163. Farris JS. Estimating phylogenetic trees from distance matrices. *Am Nat.* 1972;106(951):645–68. <https://doi.org/10.1086/282802>.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Ready to submit your research? Choose BMC and benefit from:**

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

**At BMC, research is always in progress.**

Learn more [biomedcentral.com/submissions](https://www.biomedcentral.com/submissions)

