Enrichment of Novel Actinomycetales and the Detection of Monooxygenases during Aerobic 1,4-Dioxane Biodegradation with Uncontaminated and Contaminated Inocula

Vidhya Ramalingam and Alison M. Cupples*

Department of Civil and Environmental Engineering, Michigan State University, East Lansing, Michigan, USA

*Corresponding Author:

Alison M. Cupples

A135, 1449 Engineering Research Court, Michigan State University, East Lansing, MI 48824,

cupplesa@egr.msu.edu

KEYWORDS: 1,4-dioxane, soluble di-iron monooxygenases, *Nocardioides, Gordonia*, *Kribbella, prmA, tomA3*

Abstract

1,4-dioxane, a co-contaminant at many chlorinated solvent sites, is a problematic groundwater pollutant because of risks to human health and characteristics which make remediation challenging. *In situ* 1,4-dioxane bioremediation has recently been shown to be an effective remediation strategy. However, the presence/abundance of 1,4-dioxane degrading species across different environmental samples is generally unknown. Here, the objectives were to identify which 1,4-dioxane degrading functional genes are present and which genera may be using 1,4-dioxane and/or metabolites to support growth across different microbial communities. For this, laboratory sample microcosms and abiotic control microcosms (containing media) were inoculated with four uncontaminated soils and sediments from two contaminated sites. Live control microcosms were treated in the same manner, except 1,4-dioxane was not added. 1,4-dioxane decreased in live microcosms with all six inocula, but not in the abiotic controls, suggesting biodegradation occurred. A comparison of live sample microcosms and live controls (no 1,4-dioxane) indicated nineteen genera were enriched following exposure to 1,4-dioxane, suggesting a growth benefit for 1,4-dioxane biodegradation. The three most enriched were *Mycobacterium*, *Nocardioides*, *Kribbella* (classifying as *Actinomycetales*). There was also a higher level of enrichment for *Arthrobacter, Nocardia* and *Gordonia* (all three classifying as *Actinomycetales*) in one soil, *Hyphomicrobium* (*Rhizobiales*) in another soil, *Clavibacter* (*Actinomycetales*) and *Bartonella* (*Rhizobiales*) in another soil and *Chelativorans* (*Rhizobiales*) in another soil. Although *Arthrobacter, Mycobacterium* and *Nocardia* have previously been linked to 1,4-dioxane degradation, *Nocardioides, Gordonia* and *Kribbella* are potentially novel degraders. The analysis of the functional genes associated with 1,4-dioxane demonstrated three genes were present at higher relative abundance values, including *Rhodococcus* sp. RR1 *prmA*, *Rhodococcus jostii* RHA1 *prmA* and *Burkholderia cepacia* G4 *tomA3*. Overall, this study provides novel insights into the identity of the multiple genera and functional genes associated with aerobic degradation of 1,4-dioxane in mixed communities.

Introduction

1,4-dioxane, a probable human carcinogen (DeRosa et al. 1996), was commonly used as a stabilizer in 1,1,1-trichloroethane formulations and is now frequently detected at sites where the chlorinated solvents are present (Adamson et al. 2015; Adamson et al. 2014; ATSDR 2012; Mohr et al. 2010). For example, 1,4-dioxane was found at 195 sites in California with 95% containing one or more of the chlorinated solvents (Adamson et al. 2014). 1,4-dioxane has been classified as a probable carcinogen (Group 2B) by the U.S. EPA and a possible human carcinogen (B2) by the International Agency for Research on Cancer based on animal studies (IARC 1999; USEPA 2017). No federal maximum contaminant level for 1,4-dioxane in drinking water has been established (EPA 2017), however, several states have set low advisory action levels (e.g. California, Florida, Michigan and North Carolina have levels <5 ppb). A major challenge to 1,4-dioxane remediation concerns chemical characteristics that result in migration and persistence (Adamson et al. 2015; Mohr et al. 2010). A low organic carbon partition coefficient (log K_{OC} = 1.23) and Henry's Law Constant (5 X 10^{-6} atm. m^3mol^{-1}), make traditional remediation methods such as air stripping or activated carbon largely ineffective (Mahendra and Alvarez-Cohen 2006; Steffan et al. 2007; Zenker et al. 2003). *Ex situ* oxidation methods including ozone and hydrogen peroxide (Adams et al. 1994) or hydrogen peroxide and ultraviolet light (Stefan and Bolton 1998) have been commercially applied, however these can be expensive at high concentrations (Steffan et al. 2007).

Many bacteria have been linked to the aerobic metabolic and co-metabolic degradation of 1,4-dioxane. Currently, *Pseudonocardia dioxanivorans* CB1190 (Parales et al. 1994), *Rhodococcus ruber* 219 (Bock et al. 1996), *Pseudonocardia benzennivorans* B5 (Kämpfer and Kroppenstedt 2004), *Mycobacterium sp.* PH-06 (Kim et al. 2008), *Afipia sp.* D1, *Mycobacterium sp*. D6, *Mycobacterium sp*. D11, *Pseudonocardia sp.* D17 (Sei et al. 2013), *Acinetobacter baumannii* DD1 (Huang et al. 2014), *Rhodanbacter* AYS5 (Pugazhendi et al. 2015), *Xanthobacter flavus* DT8 (Chen et al. 2016) and *Rhodococcus aetherivorans* JCM 14343 (Inoue et al. 2016) are known to degrade 1,4-dioxane metabolically. A large number of microorganisms are known to co-metabolically degrade this contaminant. For example, *Pseudonocardia tetrahydrofuranoxydans sp.* K1 (Kohlweyer et al. 2000)*, Pseudonocardia sp.* ENV478 (Vainberg et al. 2006)*, Rhodococcus ruber* T1, *Rhodococcus ruber* T5 (Sei et al. 2013)*, Rhodococcus ruber* ENV 425 (Steffan et al. 1997)*, Rhodococcus* RR1 (Stringfellow and Alvarez-Cohen 1999)*,*

Flavobacterium sp. (Sun et al. 2011), *Mycobacterium vaccae* (Burback and Perry 1993), *Mycobacterium sp.* ENV 421 (Masuda et al. 2012b), *Pseudomonas mendocina* KR1 (Whited and Gibson 1991), *Ralstonia pickettii* PKO1 (Kukor and Olsen 1990), *Burkholderia cepacia* G4 (Nelson et al. 1986), *Methylosinus trichosporium* OB3b (Whittenbury et al. 1970), *Pseudonocardia acacia* JCM (Inoue et al. 2016) and *Pseudonocardia asaccharolytica* JCM (Inoue et al. 2016) are among those linked to co-metabolic degradation of 1,4-dioxane. Co-metabolic 1,4-dioxane degradation has previously been reported with growth supporting substrates such as tetrahydrofuran, methane, propane, toluene, ethanol, sucrose, lactate, yeast extract and 2-propanol (Burback and Perry 1993; Hand et al. 2015; Kohlweyer et al. 2000; Mahendra and Alvarez-Cohen 2006; Vainberg et al. 2006).

The initiation of 1,4-dioxane biodegradation has been associated with various groups of soluble di-iron monooxygenases (SDIMOs) (He et al. 2017). Monooxygenases are enzymes that facilitate bacterial oxidation through the introduction of oxygen. SDIMOs have been classified into 6 groups based on their preferred substrate and sequence similarity (Coleman et al. 2006). SDIMOs associated with metabolic and co-metabolic 1,4-dioxane degradation include [as summarized in (He et al. 2017)] *Burkholderia cepacia* G4 *tomA3* (Group 1) (Mahendra and Alvarez-Cohen 2006; Newman and Wackett 1995), *Pseudomonas pickettii* PKO1 *tbuA1* (Group 2) (Fishman et al. 2004; Mahendra and Alvarez-Cohen 2006), *Pseudomonas mendocina* KR1 *tmoA* (Group 2) (Mahendra and Alvarez-Cohen 2006; Yen et al. 1991), *Methylosinus trichosporium* OB3b *mmoX* (Group 3) (Mahendra and Alvarez-Cohen 2006; Oldenhuis et al. 1989), *Pseudonocardia dioxanivorans* CB1190 *prmA* (Group 5) (Parales et al. 1994; Sales et al. 2013; Sales et al. 2011), *Pseudonocardia tetrahydrofuranoxydans* K1 *thmA* (Group 5) (Kampfer et al. 2006; Thiemer et al. 2003), *Pseudonocardia* sp. strain ENV478 *thmA* (Group 5) (Masuda et al. 2012a), *Rhodococcus* sp. strain YYL *thmA* (Group 5)(Yao et al. 2009), *Rhodococcus jostii* RHA1 *prmA* (Group 5) (Hand et al. 2015; Sharp et al. 2007), *Rhodococcus* sp. RR1 *prmA* (Group 5) (Sharp et al. 2007), *Mycobacterium* sp. ENV421 *prmA* (Group 6) (Masuda 2009) and *Mycobacterium dioxanotrophicus* PH-06 *prmA* (Group 6) (He et al. 2017).

As the success of natural attenuation or biostimulation often depends on the population of native degraders present at the contaminated site, several studies have developed methods targeting

these functional genes (Gedalanga et al. 2014; He et al. 2018; Li et al. 2013a; Li et al. 2013b). For example, methods have been developed for the functional genes associated with *Pseudonocardia* and *Mycobacterium* (Deng et al. 2018; Gedalanga et al. 2014; He et al. 2017). Another study focused specifically on detecting functional genes of four 1,4-dioxane degraders (*Pseudonocardia dioxanivorans* CB1190*, Pseudonocardia* sp. strain ENV478, *Pseudonocardia tetrahydrofuranoxydans* K1*, Rhodococcus* sp. strain YYL). A larger number of functional genes were investigated with microarray-based technology (GeoChip 4.0) and denaturing gradient gel electrophoresis (Li et al. 2013b). More recently, high throughput shotgun sequencing was used to evaluate the presence of the functional genes listed above in impacted and non-impacted groundwater (Dang et al. 2018) . This approach has the added advantage of enabling taxonomic as well as functional analysis of microbial communities. The current study adopted a similar approach to examine the microbial communities involved in 1,4-dioxane degradation in contaminated and uncontaminated sediment and soil inoculated microcosms.

In the current study, the objectives were 1) to identify which 1,4-dioxane degrading functional genes are present across different microbial communities and 2) to determine which genera may be using 1,4-dioxane and/or metabolites to support growth. The research focused on laboratory microcosms inoculated with four uncontaminated soils and sediment samples from two 1,4-dioxane contaminated sites. The media selected for the experiments followed the approach used to enrich *Pseudonocardia dioxanivorans* CB1190 from industrial sludge (Parales et al. 1994). The work is novel as it combines taxonomic and functional data to generate a more complete picture of the multiple microorganisms and genes linked to 1,4-dioxane degradation in mixed communities.

Methods

Chemicals and Inocula

1,4-dioxane was purchased from Sigma-Aldrich (MO, USA) with 99.8% purity. All stock solutions and dilutions were prepared using DI water. The agricultural samples were collected from two locations on the campus of Michigan State University, East Lansing, Michigan (herein called soils F and G) and two locations at the Kellogg Biological Station, Hickory Corners, Michigan (soils 1 and 2). The characteristics of the agricultural soils have been summarized

147 (Table S1). The contaminated site samples were obtained from California (contaminated with
148 trichloroethene, 1,1-dichloroethene and 1,4-dioxane, herein called C7A) and Maine
149 (contaminated with traces of 1,4-dioxane, herein called M10A). All samples were stored in the
150 dark at 6 °C until use.
151
152 **Experimental Setup, DNA Extraction, 1,4-Dioxane Analysis**
153 Laboratory microcosms were established with soil or sediment (5g wet weight) and 25 mL of
154 media in 30 mL serum bottles. For each of the six inocula (four uncontaminated soils or two
155 contaminated sediments), the experiment design included triplicate sample microcosms,
156 triplicate live control microcosms and triplicate abiotic control microcosms (autoclaved daily for
157 three consecutive days). The triplicate live control microcosms were treated in the same manner
158 as the sample microcosms except no 1,4-dioxane was added. This treatment was included to
159 enable comparisons to the microbial communities exposed to 1,4-dioxane. Following the
160 approach used to enrich *Pseudonocardia dioxanivorans* CB1190 from industrial sludge, each
161 liter of the final media contained 100 mL of a buffer stock [K_2HPO_4 (32.4 g/L), KH_2PO_4 (10
162 g/L), NH_4Cl (20 g/L)] and 100 mL of a trace metal stock [nitrilotriacetic acid (disodium salt)
163 (1.23 g/L), $MgSO_4.7H_2O$ (2 g/L), $FeSO_4.7H_2O$ (0.12 g/L), $MnSO_4.H_2O$ (0.03 g/L), $ZnSO_4.7H_2O$
164 (0.03 g/L) and $CoCl_2.6H_2O$ (0.01 g/L)] (Parales et al., 1994). The nitrilotriacetic acid (within the
165 trace metal stock solution) represents an additional carbon source. The live sample microcosms
166 were re-spiked with 1,4-dioxane two additional times.
167
168 A GC/MS with Agilent 5975 GC/single quadrupole MS (Agilent Technologies, CA, USA)
169 equipped with a CTC Combi Pal autosampler was used to determine 1,4-dioxane concentrations.
170 Sterile syringes (1 mL) and needles (22 Ga 1.5 in.) were used to collect samples (0.7 mL) from
171 each microcosm. The samples were filtered (0.22 µm nylon filter) before being injected into an
172 amber glass vial (40 mL) for GC/MS analysis. A method was developed to analyze 1,4-dioxane
173 using solid phase micro extraction (SPME). The SPME fiber was inserted in the headspace of the
174 vial and exposed to the analyte for 1 minute before being injected into the GC for thermal
175 desorption. The fiber coating can adsorb the analytes in the vapor phase. Splitless injection was
176 executed and the vials were maintained at 40 °C. The SPME fiber assembly involved 50/30µm
177 Divinylbenzene/ Carboxen/ Polydimethylsiloxane (DVB/CAR/PDMS) and 24 Ga needle for

injection. The initial oven temperature was 35 °C and was programmed to increase at a rate of 20 °C/min to 120 °C. Once it reached 120 °C, it increased at a rate of 40 °C/min to 250 °C, which was maintained for 3 min. A VF5MS column was used with helium as the carrier gas in constant flow mode at a flow rate of 1 ml/min. The conditioning of the SPME fiber was at 270 °C for 60 min at the beginning of each sequence.

DNA was extracted from the soil inoculated sample microcosms and live control microcosms (1.2 mL and 0.4 g soil) using QIAGEN DNeasy PowerSoil kit as per the manufacturer's instructions. The QIAGEN DNeasy Powermax Soil kit was used to extract DNA from the microcosms inoculated with the two contaminated sediments. For this, the entire content of each microcosm was sacrificed for DNA extraction. The DNA concentrations were determined using QUBIT dsDNA HS kit. The DNA extracts with the highest DNA yields were selected for shotgun sequencing.

Library Preparation, Sequencing, MG-RAST and DIAMOND analysis

Twenty-six samples were submitted for library generation and shotgun sequencing to the Research Technology Support Facility Genomics Core at Michigan State University. Libraries were prepared using the Takara SMARTer ThruPLEX DNA Library Preparation Kit following manufacturer's recommendations. Completed libraries were QC'd and quantified using a combination of Qubit dsDNA HS and Agilent 4200 TapeStation HS DNA1000 assays. Eight samples did not generate libraries of sufficient concentration for sequencing and were removed from further analysis. The remaining eighteen libraries were pooled in equimolar amounts for multiplexed sequencing. The pool was quantified using the Kapa Biosystems Illumina Library Quantification qPCR kit and loaded onto one lane of an Illumina HiSeq 4000 flow cell. Sequencing was performed in a 2x150 bp paired end format using HiSeq 4000 SBS reagents. Base calling was done by Illumina Real Time Analysis (RTA) v2.7.7 and output of RTA was demultiplexed and converted to FastQ format with Illumina Bcl2fastq v2.19.1.

The Meta Genome Rapid Annotation using Subsystem Technology (MG-RAST) (Meyer et al. 2008) version 4.0.2. was used for the taxonomic analysis of the metagenomes. The processing pipeline involved merging paired end reads, SolexaQA (Cox et al. 2010) to trim low-quality

209 regions and dereplication to remove the artificial duplicate reads. Gene calling was performed
210 using FragGeneScan (Rho et al. 2010). For the taxonomic profiles, the best hit classification at a
211 maximum e-value of $1e^{-5}$, a minimum identity of 60% and a minimum alignment length of 15
212 against the ReqSeq database (Pruitt et al. 2005) were used. MG-RAST ID numbers and
213 sequencing data have been summarized (Table S2) and the datasets are publicly available on
214 MG-RAST. The number of sequences generated post quality control per sample was 4.7 ± 2.0
215 million (ranging from ~1.2 to ~11 million) and the average length was 237.3 ± 2.9 bp (averages
216 ranging from 233 to 243 bp). The sequencing data was submitted to NCBI under Bioproject
217 PRJNA590532 (accession numbers SAMN13332566 to SAMN13332583).
218
219 The MG-RAST data files were downloaded and analyzed in Microsoft Excel 2016, STAMP
220 (Statistical Analyses of Metagenomic Profiles, software version 2.1.3.) (Parks et al. 2014) and
221 MEGAN6 (version 6.11.7) (Huson et al. 2016). STAMP was used to detect differences in the
222 relative proportions of the taxonomic profiles between the live controls (no 1,4-dioxane) and the
223 samples for each soil. This analysis included Welch's two sided t-test for two groups (samples
224 and live controls) ($p<0.05$) to generate extended error bar figures for each soil. The same
225 Welch's test was performed to compare the profiles of all samples ($n= 9$) to all live controls ($n=$
226 8). MEGAN6 was used to generate two phylograms. One phylogram illustrates the eighteen
227 metagenomes classified to the Class Level. The other phylogram represents the most common
228 genera (ranked by average relative abundance, then selected if average values >0.5%) across all
229 metagenomes.
230
231 The relative abundance of 1,4-dioxane degrading functional genes was determined using the
232 alignment tool DIAMOND (double index alignment of next-generation sequencing data)
233 (Buchfink et al. 2015). Specifically, reads aligning to the twelve genes previously associated
234 with aerobic degradation (metabolic and co-metabolic) of 1,4-dioxane, as summarized previously
235 (He et al. 2017), were determined. First, low quality sequences and Illumina adapters sequences
236 were removed using Trimmomatic in the paired end mode (Bolger et al. 2014). The two paired
237 output files were used for gene alignments in DIAMOND. Following alignment, the DIAMOND
238 files were analyzed within Excel, which included combining the data from the two paired files
239 and deleting duplicated data. The sort function was used to select reads that exhibited an identity

of ≥ 60% and an alignment length ≥ 49 amino acids. For each gene, the relative abundance values were calculated using the number of aligned reads divided by the total number of sequences for each sample (determined by Trimmomatic).

The above analysis indicated two functional genes (*Rhodococcus jostii* RHA1 *prmA* and *Rhocodoccus* sp. RR1 *prmA*) were dominant in the soil and sediment metagenomes. Therefore, BLASTP 2.9.0+ (Altschul et al. 1997) (protein-protein BLAST) from the NCBI website was used to search for similar protein sequences to these two genes. The sequences obtained (>94.8% similar to the two query sequences) were used to create a phylogenetic tree in MEGAN7 using the Maximum Likelihood method (Jones et al. 1992; Kumar et al. 2016).

Results

Biodegradation of 1,4-Dioxane

The concentration of 1,4-dioxane declined in all of the live microcosms (inoculated with all four uncontaminated soils and with two contaminated site sediments), but not in the abiotic controls, indicating biological removal (Figure 1). All of the live microcosms, except M10A (contaminated site soil), demonstrated >50% removal in 1,4-dioxane in approximately 40 days. Following the initial biodegradation of the chemical, the microcosms were reamended with 1,4-dioxane twice. A steady depletion of 1,4-dioxane occurred in all four agricultural soils and the two contaminated site soil samples after each reamendment, while no removal was noted in the corresponding abiotic controls. However, limited biodegradation (only one replicate decreased) was observed for soil F after the last amendment. Overall, between approximately 220 and 245 days was required to reduce the majority of the amended 1,4-dioxane.

Taxonomic Analysis of Metagenomes

Although DNA was extracted from all microcosms and live controls (no 1,4-dioxane), in some cases insufficient DNA was extracted (and was not submitted for library generation) or did not generate libraries of sufficient concentration for sequencing. Unfortunately, this included all DNA extracts for the microcosms inoculated with sediment from one of the two 1,4-dioxane contaminated sites (M10A three samples and three live controls). Also, only one sample from the other contaminated site (C7A) generated enough DNA for sequencing. Overall, eighteen libraries

were sequenced, which included two samples and two live controls for each soil (except soil F which included three samples) and one sample from one contaminated site (C7A).

The phylogenetic analysis of the eighteen soil and sediment metagenomes indicated the majority of the microorganisms classified within the classes *Acidobacteria, Alpha-, Beta-, Gamma-, Deltaproteobacteria, Actinobacteria, Bacilli* and *Clostridia* (Figure S1). The most abundant genera, averaged across all metagenomes, included *Candidatus Solibacter, Bradyrhizobium, Mesorhizobium, Burkholderia, Pseudomonas, Stenotrophomonas, Xanthomonas, Mycobacterium* and *Streptomyces* (Figure S2). The relative abundance (%) of the most abundant genera (25 most abundant) for each soil analyzed separately is also shown (Figure S3). The most abundant genera in all four soils were similar and included *Xanthomonas*, *Streptomyces*, *Mesorhizobium, Bradyrhizobium* and *Burkholderia.* In contrast, *Pseudomonas, Rhodococcus, Arthrobacter, Mycobacterium* and *Corynebacterium* were the most abundant genera in the contaminated site microcosms.

Genera Associated with 1,4-Dioxane Degradation

The metagenomes of the samples were compared to the live controls (no 1,4-dioxane) to determine which genera were positively influenced by 1,4-dioxane degradation. First, all of the samples (n=9) were compared to all of the live controls (n=8) (Figure 2). Overall, fifteen genera were statistically significantly enriched in the live samples compared to the controls. The greatest differences between the means were noted for *Mycobacterium* (0.304%, p=0.0029), followed by *Nocardioides* (0.127%, p=0.023), and *Kribbella* (0.079%, p=0.017). The trends suggest these genera are obtaining a growth benefit from the presence of 1,4-dioxane. The relative abundance of these genera in the contaminated site microcosm is also shown (Figure 2B, insert). Except for *Ureaplasma*, the enriched genera all classify within the order *Actinomycetales* (Table 1).

The metagenomes of the samples and controls were also compared for each soil individually. Two (*Clavibacter, Bartonella*) and seven genera (*Arthrobacter, Nocardia, Gordonia, Kocuria, Brevibacterium, Rothia, Erysipelothrix)* were statistically significantly enriched in the samples compared to the controls in soils 1 and 2, respectively (Figure 3). Three genera (*Hyphomicrobium, Acetobacter, Veillonella*) and one genus (*Chelativorans*) were statistically significantly enriched in the samples compared to the controls in soils F and G, respectively (Figure 4). Seven of the thirteen listed above classify within the *Actinomycetales*

(Table 1). The differences between the means (for the individual soil analysis) were the highest (≥0.033%) for *Arthrobacter, Nocardia*, *Hyphomicrobium* and *Chelativorans* (Table 1). The relative abundance of the thirteen enriched genera in the samples compared to the controls (for the individual soil analysis) and for the contaminated site sample is shown (Figure 5). The contaminated site sample is shown with a different scale as it involved higher relative abundance values compared to the other microcosms (Figure 5 insert). Both *Arthrobacter* and *Nocardia* have relative abundance values of >1% in the contaminated site sample.

Relative Abundance of Genera Associated with 1,4-dioxane Bioegradation

The metagenomes were also investigated to determine the relative abundance (%) of fifteen genera previously associated with metabolic or co-metabolic 1,4-dioxane degradation (Figure 6). All except *Pseudonocardia* and *Rhodanbacter* were present in the samples and controls. *Burkholderia, Mycobacterium, Pseudomonas* and *Rhodococcus* were present at the highest relative abundance levels (0.84-2.45%). Only *Mycobacterium* was statistically significantly (p<0.05) enriched in the samples compare to the live controls. *Pseudonocardia* and *Rhodanbacter* were also absent in the contaminated site sample (Figure 6, insert). In the contaminated site metagenome, the four most abundant genera were *Pseudomonas* (49.0%), *Rhodococcus* (5.9%), *Mycobacterium* (3.9%) and *Nocardia* (1.2%). The contaminated site sample indicates a 25-fold higher relative abundance in *Pseudomonas* and almost a 6-fold higher relative abundance in *Rhodococcus* compared to the samples from agricultural sites.

Genes Associated with 1,4-Dioxane Degradation

The reads aligning to the genes previously associated with 1,4-dioxane degradation were determined using DIAMOND and the data were analyzed using Excel. Only the reads with ≥ 60% identity for ≥ 49 amino acids were included in the analysis. Among the twelve genes previously associated with 1,4-dioxane degradation, the majority were present in all the samples including the contaminated site sample (Figure 7).

The analysis demonstrates a uniform trend of higher relative abundance values for *Rhodococcus* sp. RR1 *prmA* and *Rhodococcus jostii* RHA1 *prmA* in all four agricultural samples and the contaminated site sample compared to the other genes. The contaminated site sample demonstrates more than twice the relative abundance of these two genes compared to the

samples from the agricultural sites. A high relative abundance of *Burkholderia cepacia G4 tomA3* was also noted in all of the samples, with higher abundance (10-fold increase) in the contaminated site sample compared to other agricultural samples (Figure 7). Only six and eleven metagenomes contained reads aligning with *Methylosinus trichosporium* OB3b *touA* and *Pseudomonas pickettii PKO1 tbuA1*, respectively. Seven functional genes (*Pseudomonas mendocina* KR1 *tmoA, Rhodococcus* sp. YYL *thmA, Pseudonocardia* sp. ENV478 *thmA, Mycobacterium* sp. ENV421 *prmA, Pseudonocardia tetrahydrofuranoxydans thmA*, *Pseudonocardia dioxanivorans CB1190 thmA, Mycobacterium dioxanotrophicus PH-06 prmA*) were present in between fourteen and eighteen metagenomes. All four soils generated similar trends for the functional genes and no statistically significant differences were noted between the live controls and samples. The contaminated site sample generated the same trend for the three most abundant genes.

Following the discovery of the dominance of *Rhodococcus jostii* RHA1 *prmA* and *Rhodococcus* sp. RR1 *prmA* in the soil metagenomes, a BLASTP search was performed to find the closest matching sequences in the NCBI database. The matching protein sequences, with number of microorganisms shown in parenthesis, belonged to the genera *Rhodococcus* (60), *Kribbella* (16), *Gordonia* (10), *Mycolicibacterium* (10), *Mycobacterium* (8), *Nocardia* (7), *Nocardioides* (6), *Hoyosella* (3), *Intrasporanqium* (2), *Millisia* (1), *Cryptosporangium* (1) and *Acidobacteria* (1). Interestingly, five of these genera (*Mycobacterium, Nocardioides*, *Kribbella, Gordonia* and *Nocardia*) were enriched in the samples compared to the live controls (as discussed above, Table 1). A phylogenetic tree was generated to illustrate the evolutionary relationships between the two query sequences and the enriched genera (Figure 8). *Rhodococcus jostii* RHA1 *prmA* clustered closest to *Nocardia* sequences and *Rhodococcus* sp. RR1 *prmA* clustered closest to *Kribbella* sequences.

Discussion

The remediation of sites contaminated with 1,4-dioxane is challenging because of the physical and chemical properties of this chemical (Mohr et al. 2010). Although bioremediation can be a

363 viable option at some sites, it is still unclear which microorganisms and functional genes are
364 linked to 1,4-dioxane degradation in mixed communities.
365 To address this knowledge gap, the current study identified which genera could obtain a growth
366 benefit from 1,4-dioxane biodegradation. For this, the sample microcosms were supplied with
367 media and 1,4-dioxane and the live control microcosms were supplied with the same media, but
368 no 1,4-dioxane. Consequently, an increase in the relative abundance of any microorganism
369 between the samples and live controls could be attributed to the presence of 1,4-dioxane. From
370 this, a reasonable hypothesis would be that the enriched microorganisms are being exposed to
371 growth supporting substrates from 1,4-dioxane degradation. Here, nineteen genera increased in
372 abundance following 1,4-dioxane degradation compared to the live controls (no 1,4-dioxane).
373 The three most enriched across all four soils were *Mycobacterium*, *Nocardioides*, *Kribbella* (all
374 classifying as *Actinomycetales*). There was also a higher level of enrichment for *Arthrobacter,*
375 *Nocardia* and *Gordonia* (*Actinomycetales*), *Hyphomicrobium* (*Rhizobiales*), *Clavibacter*
376 (*Actinomycetales*) and *Bartonella* (*Rhizobiales*) and *Chelativorans* (*Rhizobiales*) in individual
377 soils.
378
379 There are at least two hypotheses on why these genera increased in abundance in 1,4-dioxane
380 amended samples compared to the live controls. One hypothesis being that these microorganisms
381 are obtaining a growth benefit from consuming 1,4-dioxane biodegradation products. Several
382 studies have examined 1,4-dioxane biodegradation pathways (Grostern et al. 2012; Huang et al.
383 2014; Kim et al. 2009; Mahendra et al. 2007; Sales et al. 2013; Vainberg et al. 2006). A study
384 with *Pseudonocardia dioxanivorans* CB1190 provided evidence that carbon from 1,4-dioxane
385 enters central metabolism via glyoxlate (Grostern et al. 2012). In contrast, *Pseudonocardia* sp.
386 strain ENV478 produces 2-hydroxyethoxyacetic acid (HEAA) as a terminal product of 1,4-
387 dioxane biodegradation (Vainberg et al. 2006). Conversely, 1,4-dioxane biodegradation by
388 *Pseudonocardia dioxanivorans* CB1190 (metabolic 1,4-dioxane degrader), *Mycobacterium*
389 *vaccae* JOB5, *Pseudomonas mendocina* KR1, *Pseudonocardia tetrahydrofuranoxydans* K1 (co-
390 metabolic 1,4-dioxane degraders) produced HEAA transiently, but the chemical did not
391 accumulate. They identified ethylene glycol, glycolic acid, glyoxylic acid and oxalic acid as 1,4-
392 dioxane biodegradation intermediates by these isolates (Mahendra et al. 2007). Others have also
393 identified ethylene glycol (Huang et al. 2014; Kim et al. 2009), oxalic acid (Huang et al. 2014)

and ethane-1,2-diol (Kim et al. 2009) during 1,4-dioxane degradation. The enriched genera may have benefited from funneling these degradation intermediates into central metabolism.
A second hypothesis being that the enriched genera are responsible for both the initial attack on 1,4-dioxane and for the consumption of degradation products. Evidence for this concerns the similarity of genes belonging to the enriched genera (*Mycobacterium*, *Nocardioides*, *Kribbella, Nocardia* and *Gordonia)* to *Rhodococcus jostii* RHA1 *prmA* and *Rhodococcus* sp. RR1 *prmA* (as shown in the phylogenetic tree). Although *Rhodococcus jostii* RHA1 and *Rhodococcus* sp. RR1 co-metabolically degrade 1,4-dioxane, the enriched genera may also contain genes downstream in the pathway enabling growth on 1,4-dioxane. *Arthrobacter* did not contain genes similar to the *Rhodococcus* strains, although others have reported that *Arthrobacter* (ATCC 27779) can co-metabolically degrade 1,4-dioxane (Chu et al. 2009). *Arthrobacter, Mycobacterium* and *Nocardia* have previously been linked to 1,4-dioxane degradation (Chu et al. 2009; Lan et al. 2013; Masuda 2009), whereas *Nocardioides, Gordonia* and *Kribbella* are potentially novel degraders. Certain species of *Gordonia* such as *G. terrae* are known to aid in degrading certain chemicals, including ethyl tertiary butyl ether (ETBE) metabolically, methyl tertiary butyl ether (MTBE) co-metabolically (Hernandez-Perez et al. 2001) as well as long chain hydrocarbons (Kubota et al. 2008). Overall, both hypotheses in this work suggests many genera (almost all classifying with the *Actinomycetales*) are likely involved in the degradation of 1,4-dioxane and/or 1,4-dioxane metabolites in the soil microcosms studied.

In the current study, reads from all of the 1,4-dioxane degrading function genes were observed in soil metagenomes. Consistent with the current study, others have detected SDIMOs from the majority (five from six groups) of SDIMO groups (Li et al. 2013b). In that research, the authors examined Arctic groundwater impacted by 1,4-dioxane using high-throughput microarrays and denaturing gradient gel electrophoresis and found an enrichment of *thmA*-like genes near the source zone (Li et al. 2013b). Also similar to the current work, a 1,4-dioxane degrading consortia contained a high percentage of group five SDIMOs (*Rhodococcus jostii* RHA1 *prmA* and *Rhodococcus* sp. RR1 *prmA* are group five SDIMOs), although the specific genes were not determined (He et al. 2018). Another study noted a correlation between *dxmA/thmA* (designed based on *Rhodococcus* sp. YYL *thmA, Pseudonocardia* sp. ENV478 *thmA, Pseudonocardia tetrahydrofuranoxydans K1 thmA* and *Pseudonocardia dioxanivorans CB1190 thmA*) and the

425 amount of 1,4-dioxane degraded in groundwater inoculated microcosms (Li et al. 2013a). These
426 genes were also present in the soil metagenomes (between fourteen and sixteen) of the current
427 study.
428
429 Recently, shotgun sequencing was used to examine 1,4-dioxane degrading genes in groundwater
430 from multiple chlorinated solvent sites (previously bioaugmented with SDC-9) (Dang et al.
431 2018). From the twelve genes examined, only six were found in the groundwater metagenomes.
432 The six included the three most abundant genes in the current study; *Rhodococcus sp.* RR1
433 *prmA*, *Rhodococcus jostii* RHA1 *prmA* and *Burkholderia cepacia* G4 *tomA3*. From these, the
434 *Rhodococcus* genes were both found in a only small number of metagenomes (~18%) and *B.*
435 *cepacia* G4 *tomA3* was found in the majority (~68%). The occurrence of the three genes in both
436 studies could suggest their importance across different environments (soil vs. groundwater,
437 aerobic vs. oxygen depleted). Unlike the current study, the groundwater metagenomes contained
438 high relative abundance values for *Methylosinus trichosporium* OB3b *touA* (up to 0.0031%)
439 followed by *Pseudomonas mendocina* KR1 *tmoA* (up to 0.00022%) and *Pseudomonas pickettii*
440 PKO1 *tbuA1* (up to 0.0013%). The different results between the two studies are likely due to
441 variations in the conditions (redox potential, carbon availability, nutrient availability, soil vs.
442 groundwater) from which the samples were obtained.
443
444 In summary, several key findings highly relevant for 1,4-dioxane bioremediation were generated
445 here. Shotgun sequencing enabled both taxonomic and functional analyses to be performed on
446 multiple mixed microbial communities. Multiple genera classifying (including novel and
447 previously identified degraders) within the *Actinomycetales* were enriched during 1,4-dioxane
448 degradation and may be associated with growth linked 1,4-dioxane degradation.
449 The three most enriched were *Mycobacterium*, *Nocardioides*, *Kribbella* (classifying as
450 *Actinomycetales*). There was also a higher level of enrichment of other genera in individual soils.
451 The current research found that both previously reported genera as well as novel genera (e.g.
452 *Nocardioides, Gordonia* and *Kribbella*) were linked to 1,4-dioxane degradation. However, it is
453 unknown if these microorganisms are benefiting from the complete degradation of the chemical
454 or from the consumption of 1,4-dioxane degradation products, such as HEAA, ethylene glycol,
455 glycolic acid, glyoxylic acid or oxalic acid. Finally, all of the functional genes associated with

1,4-dioxane were found in the soil and sediment metagenomes. Reads aligning to *Rhodococcus jostii* RHA1 *prmA* and *Rhodococcus* sp. RR1 *prmA* illustrated the highest relative abundance values and were present in all eighteen metagenomes. Future research should be directed towards similar molecular analyses of groundwater and sediment samples from 1,4-dioxane contaminated sites as well as comparisons to 1,4-dioxane removal rates for propane amended samples.

Acknowledgements

Thanks to Dr. Dan Jones and Dr. Scott Smith at the Mass Spectrometry Laboratory at the Research Technology Support Facility (MSU) for 1,4-dioxane analytical methods support. Also, our thanks to Dr. Anthony Danko (Naval Facilities Engineering Command) for providing the contaminated site sediments. Support for this research was also provided by the NSF Long-term Ecological Research Program (DEB 1832042) at the Kellogg Biological Station and by Michigan State University AgBioResearch.

Ethical Statement

This work is supported by a grant awarded to Dr. Cupples from Strategic Environmental Research and Development Program (SERDP). Contract Number: W912HQ-17-C-0006. All authors declare no conflict of interest. This article does not contain any studies with human participants or animals performed by any of the authors.

References

Adams CD, Scanlan PA, Secrist ND (1994) Oxidation and biodegradability enhancement of 1,4-dioxane using hydrogen peroxide and ozone. Environ Sci Technol 28(11):1812-1818

Adamson DT, Anderson RH, Mahendra S, Newell CJ (2015) Evidence of 1,4-dioxane attenuation at groundwater sites contaminated with chlorinated solvents and 1,4-dioxane. Environ Sci Technol 49(11):6510-6518

Adamson DT, Mahendra S, Walker KL, Rauch SR, Sengupta S, Newell CJ (2014) A multisite survey to identify the scale of the 1,4-dioxane problem atcontaminated groundwater sites. Environ Sci Tech Let 1(5):254-258

Altschul SF, Madden TL, Schaffer AA, Zhang JH, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25(17):3389-3402 doi:DOI 10.1093/nar/25.17.3389

ATSDR AfTSaDR (2012) Toxicological profile for 1,4-dioxane. In: Registry AfTSaD (ed). Agency for Toxic Substances and Disease Registry. https://www.atsdr.cdc.gov/toxprofiles/TP.asp?id=955&tid=199, Atlanta, Georgia

Bock C, Kroppenstedt RM, Diekmann H (1996) Degradation and bioconversion of aliphatic and aromatic hydrocarbons by *Rhodococcus ruber* 219. Appl Microbiol Biotechnol 45(3):408-410 doi:10.1007/s002530050704

Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30(15):2114-2120 doi:10.1093/bioinformatics/btu170

Buchfink B, Xie C, Huson DH (2015) Fast and sensitive protein alignment using DIAMOND. Nat Methods 12(1):59-60

Burback BL, Perry JJ (1993) Biodegradation and biotransformation of groundwater pollutant mixtures by *Mycobacterium vaccae*. Appl Environ Microb 59(4):1025-1029

Chen D-Z, Jin X-J, Chen J, Ye J-X, Jiang N-X, Chen J-M (2016) Intermediates and substrate interaction of 1,4-dioxane degradation by the effective metabolizer *Xanthobacter flavus* DT8. Inter Biodeter & Biodeg 106:133-140 doi:10.1016/j.ibiod.2015.09.018

Chu MYJ, Bennett P, Dolan M, Hyman M, Anderson R, Bodour A, Peacock A (2009) Using aerobic cometabolic biodegradation and groundwater recirculation to treat 1,4-dioxane and co-contaminants in a dilute plume. Paper presented at the Tenth International Conference on the Remediation of Chlorinated and Recalcitrant Compounds, Baltimore, Maryland,

Coleman NV, Bui NB, Holmes AJ (2006) Soluble di-iron monooxygenase gene diversity in soils, sediments and ethene enrichments. 8(7):1228-1239 doi:10.1111/j.1462-2920.2006.01015.x

Cox MP, Peterson DA, Biggs PJ (2010) SolexaQA: At-a-glance quality assessment of Illumina second-generation sequencing data. BMC Bioinform 11:485:1-6

Dang HY, Kanitkar YH, Stedtfeld RD, Hatzinger PB, Hashsham SA, Cupples AM (2018) Abundance of chlorinated solvent and 1,4-dioxane degrading microorganisms at five chlorinated solvent contaminated sites determined via shotgun sequencing. Environ Sci Technol 52(23):13914-13924 doi:10.1021/acs.est.8b04895

Deng DY, Li F, Li MY (2018) A novel propane monooxygenase initiating degradation of 1,4-dioxane by *Mycobacterium dioxanotrophicus* PH-06. Environ Sci Tech Let 5(2):86-91 doi:10.1021/acs.estlett.7b00504

528 DeRosa CT, Wilbur S, Holler J, Richter P, Stevens YW (1996) Health evaluation of 1,4-dioxane. Toxicol &
529 Indust Health 12(1):1-43
530 EPA (2017) Techincal fact sheet - 1,4-dioxane. vol EPA 505-F-17-011. EPA Office of Land and Emergency
531 Management
532 Fishman A, Tao Y, Wood TK (2004) Toluene 3-monooxygenase of *Ralstonia pickettii* PKO1 is a para-
533 hydroxylating enzyme. J Bacteriol 186(10):3117-3123 doi:10.1128/Jb.186.12.3117-3123.2004
534 Gedalanga PB, Pornwongthong P, Mora R, Chiang SYD, Baldwin B, Ogles D, Mahendra S (2014)
535 Identification of biomarker genes To predict biodegradation of 1,4-dioxane. Appl Environ Microb
536 80(10):3209-3218 doi:10.1128/Abm.04162-1
537 Grostern A, Sales CM, Zhuang WQ, Erbilgin O, Alvarez-Cohen L (2012) Glyoxylate metabolism is a key
538 feature of the metabolic degradation of 1,4-dioxane by *Pseudonocardia dioxanivorans* strain
539 CB1190. Appl Environ Microb 78(9):3298-3308
540 Hand S, Wang BX, Chu KH (2015) Biodegradation of 1,4-dioxane: Effects of enzyme inducers and
541 trichloroethylene. Sci Total Environ 520:154-159
542 He Y, Mathieu J, da Silva MLB, Li MY, Alvarez PJJ (2018) 1,4-Dioxane-degrading consortia can be enriched
543 from uncontaminated soils: prevalence of *Mycobacterium* and soluble di-iron monooxygenase
544 genes. Microb Biotechnol 11(1):189-198 doi:10.1111/1751-7915.12850
545 He Y, Mathieu J, Yang Y, Yu P, da Silva MLB, Alvarez PJJ (2017) 1,4-Dioxane biodegradation by
546 *Mycobacterium dioxanotrophicus* PH-06 is associated with a Group-6 soluble di-Iron
547 monooxygenase. Environ Sci & Tec Lett 4(11):494-499 doi:10.1021/acs.estlett.7b00456
548 Hernandez-Perez G, Fayolle F, Vandecasteele J-P (2001) Biodegradation of ethyl t-butyl ether (ETBE),
549 methyl t-butyl ether (MTBE) and t-amyl methyl ether (TAME) by *Gordonia terrae*. Appl Micro &
550 Biotec 55(1):117-121 doi:10.1007/s002530000482
551 Huang H, Shen D, Li N, Shan D, Shentu J, Zhou Y (2014) Biodegradation of 1,4-dioxane by a novel strain
552 and its biodegradation pathway. Water Air Soil Poll 225(9) doi:10.1007/s11270-014-2135-2
553 Huson DH, Beier S, Flade I, Gorska A, El-Hadidi M, Mitra S, Ruscheweyh HJ, Tappu R (2016) MEGAN
554 Community Edition - Interactive exploration and analysis of large-scale microbiome sequencing
555 data. Plos Comput Biol 12(6) doi:ARTN e1004957 10.1371/journal.pcbi.1004957
556 IARC (1999) Re-evaluation of some organic chemicals, hydrazine and hydrogen peroxide. Proceedings of
557 the IARC Working Group on the Evaluation of Carcinogenic Risks to Humans. Lyon, France, 17-24
558 February 1998. 71 Pt 1:1-315
559 Inoue D, Tsunoda T, Sawada K, Yamamoto N, Saito Y (2016) 1,4-Dioxane degradation potential of
560 members of the genera *Pseudonocardia* and *Rhodococcus*. Biodegrad 27(4-6):277-286
561 doi:10.1007/s10532-016-9772-7
562 Jones DT, Taylor WR, Thornton JM (1992) The rapid generation of mutation data matrices from protein
563 sequences. Comput Appl Biosci 8(3):275-282
564 Kampfer P, Kohlweyer U, Thiemer B, Andreesen JR (2006) *Pseudonocardia tetrahydrofuranoxydans* sp
565 nov. Int J Syst Evol Micr 56:1535-1538 doi:10.1099/ijs.0.64199-0
566 Kämpfer P, Kroppenstedt RM (2004) *Pseudonocardia benzenivorans* sp. nov. Int J Sys Evol Micro
567 54(3):749-751 doi:10.1099/ijs.0.02825-0
568 Kim Y-M, Jeon J-R, Murugesan K, Kim E-J, Chang Y-S (2008) Biodegradation of 1,4-dioxane and
569 transformation of related cyclic compounds by a newly isolated *Mycobacterium* sp. PH-06.
570 Biodegrad 20(4):511 doi:10.1007/s10532-008-9240-0
571 Kim YM, Jeon JR, Murugesan K, Kim EJ, Chang YS (2009) Biodegradation of 1,4-dioxane and
572 transformation of related cyclic compounds by a newly isolated *Mycobacterium* sp PH-06.
573 Biodegrad 20(4):511-519

574 Kohlweyer U, Thiemer B, Schrader T, Andreesen JR (2000) Tetrahydrofuran degradation by a newly
575 isolated culture of *Pseudonocardia* sp strain K1. Fems Microbiol Lett 186(2):301-306 doi:DOI
576 10.1111/j.1574-6968.2000.tb09121.x
577 Kubota K, Koma D, Matsumiya Y, Chung SY, Kubo M (2008) Phylogenetic analysis of long-chain
578 hydrocarbon-degrading bacteria and evaluation of their hydrocarbon-degradation by the 2,6-
579 DCPIP assay. Biodegrad 19(5):749-57 doi:10.1007/s10532-008-9179-1
580 Kukor JJ, Olsen RH (1990) Molecular cloning, characterization, and regulation of a *Pseudomonas pickettii*
581 PKO1 gene encoding phenol hydroxylase and expression of the gene in *Pseudomonas*
582 *aeruginosa* PAO1c. J Bac doi:10.1128/jb.172.8.4624-4630.1990
583 Kumar S, Stecher G, Tamura K (2016) MEGA7: Molecular evolutionary genetics analysis version 7.0 for
584 bigger datasets. Mol Biol Evol 33(7):1870-1874 doi:10.1093/molbev/msw054
585 Lan RS, Smith CA, Hyman MR (2013) Oxidation of cyclic ethers by alkane-grown *Mycobacterium vaccae*
586 JOB5. Remediation 23(4):23-42 doi:10.1002/rem.21364
587 Li M, Mathieu J, Liu Y, Van Orden E.T., Yang Y, Fiorenza S, Alvarez PJ (2013a) The abundance of
588 tetrahydrofuran/dioxane monooxygenase genes (*thmA/dxmA*) and 1,4-dioxane degradation
589 activity are significantly correlated at various impacted aquifers. Env Sci Tech Lett 1:122-127
590 Li MY, Mathieu J, Yang Y, Fiorenza S, Deng Y, He ZL, Zhou JZ, Alvarez PJJ (2013b) Widespread distribution
591 of soluble di-iron monooxygenase (SDIMO) genes in Arctic groundwater impacted by 1,4-
592 dioxane. Environ Sci Technol 47(17):9950-9958 doi:10.1021/es402228x
593 Mahendra S, Alvarez-Cohen L (2006) Kinetics of 1,4-dioxane biodegradation by monooxygenase-
594 expressing bacteria. Environ Sci Technol 40(17):5435-5442 doi:10.1021/es060714v
595 Mahendra S, Petzold CJ, Baidoo EE, Keasling JD, Alvarez-Cohen L (2007) Identification of the
596 intermediates of in vivo oxidation of 1,4-dioxane by monooxygenase-containing bacteria.
597 Environ Sci Technol 41(21):7330-7336 doi:10.1021/es0705745
598 Masuda H (2009) Identification and characterization of monooxygenase enzymes involved in 1,4-
599 dioxane degradation in *Pseudonocardia* sp. strain ENV478, *Mycobacterium* sp. strain ENV421,
600 and *Nocardia* sp. strain ENV425. Rutgers The State University of New Jersey and University of
601 Medicine and Dentistry of New Jersey
602 Masuda H, McClay K, Steffan RJ, Zylstra GJ (2012a) Biodegradation of tetrahydrofuran and 1,4-dioxane
603 by soluble diiron monooxygenase in *Pseudonocardia* sp. strain ENV478. J Mol Microb Biotech
604 22(5):312-316
605 Masuda H, McClay K, Steffan RJ, Zylstra GJ (2012b) Characterization of three propane-inducible
606 oxygenases in Mycobacterium sp. strain ENV421. Lett in App Micro 55(3):175-181
607 doi:10.1111/j.1472-765x.2012.03290.x
608 Meyer F, Paarmann D, D'Souza M, Olson R, Glass EM, Kubal M, Paczian T, Rodriguez A, Stevens R, Wilke
609 A, Wilkening J, Edwards RA (2008) The metagenomics RAST server - a public resource for the
610 automatic phylogenetic and functional analysis of metagenomes. BMC Bioinformatics 9 doi:Artn
611 386 10.1186/1471-2105-9-386
612 Mohr T, Stickney J, Diguiseppi B (2010) Environmental investigation and remediation : 1,4-dioxane and
613 other solvent stabilizers. CRC Press/Taylor & Francis, Boca Raton, FL
614 Nelson MJ, Montgomery SO, O'Neill EJ, Pritchard PH (1986) Aerobic metabolism of trichloroethylene by
615 a bacterial isolate. Appl Environ Microbiol 52(2):383-384
616 Newman LM, Wackett LP (1995) Purification and characterization of toluene 2-monooxygenase from
617 *Burkholderia cepacia* G4. Biochemistry-Us 34(43):14066-14076 doi:DOI 10.1021/bi00043a012
618 Oldenhuis R, Vink RLJM, Janssen DB, Witholt B (1989) Degradation of chlorinated aliphatic hydrocarbons
619 by *Methylosinus trichosporium* OB3b expressing soluble methane monooxygenase. Appl Environ
620 Microb 55(11):2819-2826

621 Parales RE, Adamus JE, White N, May HD (1994) Degradation of 1,4-dioxane by an actinomycete in pure
622 culture. Appl Environ Microb 60(12):4527-4530
623 Parks DH, Tyson GW, Hugenholtz P, Beiko RG (2014) STAMP: statistical analysis of taxonomic and
624 functional profiles. Bioinformatics 30(21):3123-3124
625 Pruitt KD, Tatusova T, Maglott DR (2005) NCBI Reference Sequence (RefSeq): a curated non-redundant
626 sequence database of genomes, transcripts and proteins. Nucleic Acids Res 33:D501-D504
627 Pugazhendi A, Rajesh Banu J, Dhavamani J, Yeom IT (2015) Biodegradation of 1,4-dioxane by
628 Rhodanobacter AYS5 and the role of additional substrates. Annal of Micro 65(4):2201-2208
629 doi:10.1007/s13213-015-1060-y
630 Rho MN, Tang HX, Ye YZ (2010) FragGeneScan: predicting genes in short and error-prone reads. Nucleic
631 Acids Res 38(20)
632 Sales CM, Grostern A, Parales JV, Parales RE, Alvarez-Cohen L (2013) Oxidation of the cyclic ethers 1,4-
633 dioxane and tetrahydrofuran by a monooxygenase in two *Pseudonocardia* species. Appl Environ
634 Microb 79(24):7702-7708 doi:10.1128/Aem.02418-13
635 Sales CM, Mahendra S, Grostern A, Parales RE, Goodwin LA, Woyke T, Nolan M, Lapidus A, Chertkov O,
636 Ovchinnikova G, Sczyrba A, Alvarez-Cohen L (2011) Genome sequence of the 1,4-dioxane-
637 degrading *Pseudonocardia dioxanivorans* Strain CB1190. J Bacteriol 193(17):4549-4550
638 doi:10.1128/Jb.00415-11
639 Sei K, Miyagaki K, Kakinoki T, Fukugasako K, Inoue D, Ike M (2013) Isolation and characterization of
640 bacterial strains that have high ability to degrade 1,4-dioxane as a sole carbon and energy
641 source. Biodegrad 24(5):665-674
642 Sharp JO, Sales CM, LeBlanc JC, Liu J, Wood TK, Eltis LD, Mohn WW, Alvarez-Cohen L (2007) An inducible
643 propane monooxygenase is responsible for N-nitrosodimethylamine degradation by
644 *Rhodococcus* sp strain RHA1. Appl Environ Microb 73(21):6930-6938 doi:10.1128/Aem.01697-07
645 Stefan MI, Bolton JR (1998) Mechanism of the degradation of 1,4-dioxane in dilute aqueous solution
646 using the UV hydrogen peroxide process. Env Sci Tech 32(11):1588-1595
647 Steffan RJ, McClay K, Vainberg S, Condee CW, Zhang D (1997) Biodegradation of the gasoline oxygenates
648 methyl tert-butyl ether, ethyl tert-butyl ether, and tert-amyl methyl ether by propane-oxidizing
649 bacteria. Appl Environ Microbiol 63(11):4216-4222
650 Steffan RJ, McClay KR, Masuda H, Zylstra GJ (2007) Biodegradatdion of 1,4-dioxane. Strategic
651 Environmental Research and Development Program: ER-1422 Final Report
652 Stringfellow WT, Alvarez-Cohen L (1999) Evaluating the relationship between the sorption of PAHs to
653 bacterial biomass and biodegradation. Wat Res 33(11):2535-2544 doi:10.1016/s0043-
654 1354(98)00497-7
655 Sun B, Ko K, Ramsay JA (2011) Biodegradation of 1,4-dioxane by a *Flavobacterium*. Biodegrad 22(3):651-
656 659 doi:10.1007/s10532-010-9438-9
657 Thiemer B, Andreesen JR, Schrader T (2003) Cloning and characterization of a gene cluster involved in
658 tetrahydrofuran degradation in *Pseudonocardia* sp strain K1. Arch Microbiol 179(4):266-277
659 doi:10.1007/s00203-003-0526-7
660 USEPA (2017) Technical Fact Sheet 1,4-Dioxane. In: Management USEOoLaE (ed).
661 Vainberg S, McClay K, Masuda H, Root D, Condee C, Zylstra GJ, Steffan RJ (2006) Biodegradation of ether
662 pollutants by *Pseudonocardia* sp strain ENV478. Appl Environ Microb 72(8):5218-5224
663 Whited GM, Gibson DT (1991) Separation and partial characterization of the enzymes of the toluene-4-
664 monooxygenase catabolic pathway in *Pseudomonas mendocina* KR1. J Bacteriol 173(9):3017-
665 3020 doi:10.1128/jb.173.9.3017-3020.1991
666 Whittenbury R, Phillips KC, Wilkinson JF (1970) Enrichment, isolation and some properties of methane-
667 utilizing bacteria. J of Gen Micro 61(2):205-218 doi:10.1099/00221287-61-2-205

668 Yao YL, Lv ZM, Min H, Lv ZH, Jiao HP (2009) Isolation, identification and characterization of a novel
669 *Rhodococcus* sp strain in biodegradation of tetrahydrofuran and its medium optimization using
670 sequential statistics-based experimental designs. Bioresource Technol 100(11):2762-2769
671 doi:10.1016/j.biortech.2009.01.006
672 Yen KM, Karl MR, Blatt LM, Simon MJ, Winter RB, Fausset PR, Lu HS, Harcourt AA, Chen KK (1991)
673 Cloning and characterization of a *Pseudomonas mendocina* KR1 gene cluster encoding toluene-
674 4-monooxygenase. J Bacteriol 173(17):5315-5327 doi:DOI 10.1128/jb.173.17.5315-5327.1991
675 Zenker MJ, Borden RC, Barlaz MA (2003) Occurrence and treatment of 1,4-dioxane in aqueous
676 environments. Env Eng Sci 20(5):423-432

677

678

679 **Table and Figure Legends**

680 **Table 1.** Classification of genera statistically significantly enriched (p<0.05) in the samples
681 compared to the controls (no 1,4-dioxane) following the degradation of 1,4-dioxane in all soils
682 collectively and when the soils were analyzed individually. The last column also illustrates the
683 difference in means between the controls and the samples for each genera. Genera in bold were
684 identified in the BLASTP search as containing genes similar to *Rhodococcus jostii* RHA1 *prmA*
685 and *Rhocodoccus* sp. *prmA* (as discussed in the results section for the functional gene analysis)

686 **Figure 1.** Average 1,4-dioxane concentrations (mg/L) in triplicate samples and abiotic controls
687 with different inocula, including four agricultural soils and sediments from two contaminated
688 sites (bars represent standard deviations). 1,4-dioxane was reamended to the samples
689 microcosms twice (arrows).
690
691 **Figure 2.** Extended error bar plot illustrating genera statistically significantly different in relative
692 abundance (Welch's two sided t-test, p <0.05) between the samples (n=9) and the live controls
693 (no 1, 4-dioxane, n=8) following 1,4-dioxane degradation (A). The symbols to the left of the
694 dashed line (yellow) indicate a higher relative abundance in the samples compared to the controls
695 and the symbols to the right (blue) indicate the reverse. The figure was created with the software
696 STAMP. A comparison of the relative abundance values (%) for the genera enriched in the
697 samples is also shown in a box plot format (B). The insert illustrates the relative abundance of
698 these enriched genera in the contaminated site sample (C7A) with a different y-axis scale.
699
700 **Figure 3.** Extended error bar plots illustrating genera statistically significantly different in
701 relative abundance (Welch's two sided t-test, p <0.05) between the samples and the live controls
702 following 1,4-dioxane degradation in soil 1 (A) and 2 (B). The symbols to the left of the dashed
703 line (in yellow) indicate a higher relative abundance in the samples compared to the controls and
704 the symbols to the right (in blue) indicate the reverse.
705
706 **Figure 4.** Extended error bar plots illustrating genera statistically significantly different in
707 relative abundance (Welch's two sided t-test, p <0.05) between the samples and the live controls
708 following 1,4-dioxane degradation in soil F (A) and G (B). The symbols to the left of the dashed

line (in yellow) indicate a higher relative abundance in the samples compared to the controls and the symbols to the right (in blue) indicate the reverse.

Figure 5. Summary of the relative abundance of statistically significantly enriched genera in the samples compared to the controls (no 1,4-dioxane) for soils 1, 2, F and G. The insert illustrates the relative abundance of these genera in the contaminated site sample (C7A) with a different scale on the y-axis.

Figure 6. Relative abundance (%) of genera associated with metabolic and co-metabolic degradation of 1,4-dioxane in live controls (n=8) and samples (n=9) in four soils and one contaminated site sample (C7A). The value "a" indicates a significant difference (p<0.05) in a two tailed student's t-test between the samples and controls. The insert illustrates the same data with a different y-axis.

Figure 7. Relative abundance (%) of reads aligning (≥60% identity for ≥ 49 amino acids) to genes previously associated with the metabolic and co-metabolic degradation of 1,4- dioxane in Soil F and C7A (A), Soil G (B), Soil 1 (C) and Soil 2 (D).

Figure 8. Phylogenetic tree of *Rhodococcus jostii* RHA1 *prmA* and *Rhocodoccus* sp. *prmA* and BLASTP results (>94.8% similar to the two query sequences). Only genera that were enriched following 1,4-dioxane degradation (compared to the controls) are shown (Table 1). The evolutionary history was inferred by using the Maximum Likelihood method based on the Jones-Taylor-Thornton (JTT) matrix-based model. The tree with the highest log likelihood (-2731.06) is shown. Initial tree(s) for the heuristic search were obtained automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using a JTT model, and then selecting the topology with superior log likelihood value. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. The analysis involved 48 amino acid sequences. All positions containing gaps and missing data were eliminated. There were a total of 439 positions in the final dataset. Evolutionary analyses were conducted in MEGA7.

Table 1.

Phylum	Class	Order	Family	Genus	Difference in Means (%)
All Soils: All samples (n=9) compared to all controls (n=8)					
Actinobacteria	*Actinobacteria*	*Actinomycetales*	*Mycobacteriaceae*	***Mycobacterium***	0.304
Actinobacteria	*Actinobacteria*	*Actinomycetales*	*Nocardioidaceae*	***Nocardioides***	0.127
Actinobacteria	*Actinobacteria*	*Actinomycetales*	*Nocardioidaceae*	***Kribbella***	0.079
Actinobacteria	*Actinobacteria*	*Actinomycetales*	*Pseudonocardiaceae*	*Amycolatopsis*	0.042
Actinobacteria	*Actinobacteria*	*Actinomycetales*	*Cellulomonadaceae*	*Cellulomonas*	0.035
Actinobacteria	*Actinobacteria*	*Actinomycetales*	*Actinosynnemataceae*	*Actinosynnema*	0.027
Actinobacteria	*Actinobacteria*	*Actinomycetales*	*Beutenbergiaceae*	*Beutenbergia*	0.025
Actinobacteria	*Actinobacteria*	*Actinomycetales*	*Sanguibacteraceae*	*Sanguibacter*	0.023
Actinobacteria	*Actinobacteria*	*Actinomycetales*	*Pseudonocardiaceae*	*Saccharomonospora*	0.019
Actinobacteria	*Actinobacteria*	*Actinomycetales*	*Promicromonosporaceae*	*Xylanimonas*	0.018
Actinobacteria	*Actinobacteria*	*Actinomycetales*	*Glycomycetaceae*	*Stackebrandtia*	0.015
Actinobacteria	*Actinobacteria*	*Actinomycetales*	*Gordoniaceae*	***Gordonia***	0.014
Actinobacteria	*Actinobacteria*	*Actinomycetales*	*Nocardioidaceae*	*Aeromicrobium*	0.011
Actinobacteria	*Actinobacteria*	*Actinomycetales*	*Tsukamurellaceae*	*Tsukamurella*	0.008
Tenericutes	*Mollicutes*	*Mycoplasmatales*	*Mycoplasmataceae*	*Ureaplasma*	0.0002
Soil 1: Samples (n=3) compared to controls (n=2)					
Actinobacteria	*Actinobacteria*	*Actinomycetales*	*Microbacteriaceae*	*Clavibacter*	0.017
Proteobacteria	*Alphaproteobacteria*	*Rhizobiales*	*Bartonellaceae*	*Bartonella*	0.010
Soil 2: Samples (n=2) compared to controls (n=2)					
Actinobacteria	*Actinobacteria*	*Actinomycetales*	*Micrococcaceae*	*Arthrobacter*	0.276
Actinobacteria	*Actinobacteria*	*Actinomycetales*	*Nocardiaceae*	***Nocardia***	0.049
Actinobacteria	*Actinobacteria*	*Actinomycetales*	*Gordoniaceae*	***Gordonia***	0.019
Actinobacteria	*Actinobacteria*	*Actinomycetales*	*Micrococcaceae*	*Kocuria*	0.017
Actinobacteria	*Actinobacteria*	*Actinomycetales*	*Brevibacteriaceae*	*Brevibacterium*	0.015
Actinobacteria	*Actinobacteria*	*Actinomycetales*	*Micrococcaceae*	*Rothia*	0.005
Firmicutes	*Erysipelotrichi*	*Erysipelotrichales*	*Erysipelotrichaceae*	*Erysipelothrix*	0.0005
Soil F: Samples (n=2) compared to controls (n=2)					
Proteobacteria	*Alphaproteobacteria*	*Rhizobiales*	*Hyphomicrobiaceae*	*Hyphomicrobium*	0.033
Proteobacteria	*Alphaproteobacteria*	*Rhodospirillales*	*Acetobacteraceae*	*Acetobacter*	0.003
Firmicutes	*Negativicutes*	*Selenomonadales*	*Veillonellaceae*	*Veillonella*	0.002
Soil G: Samples (n=2) compared to controls (n=2)					
Proteobacteria	*Alphaproteobacteria*	*Rhizobiales*	*Phyllobacteriaceae*	*Chelativorans*	0.055

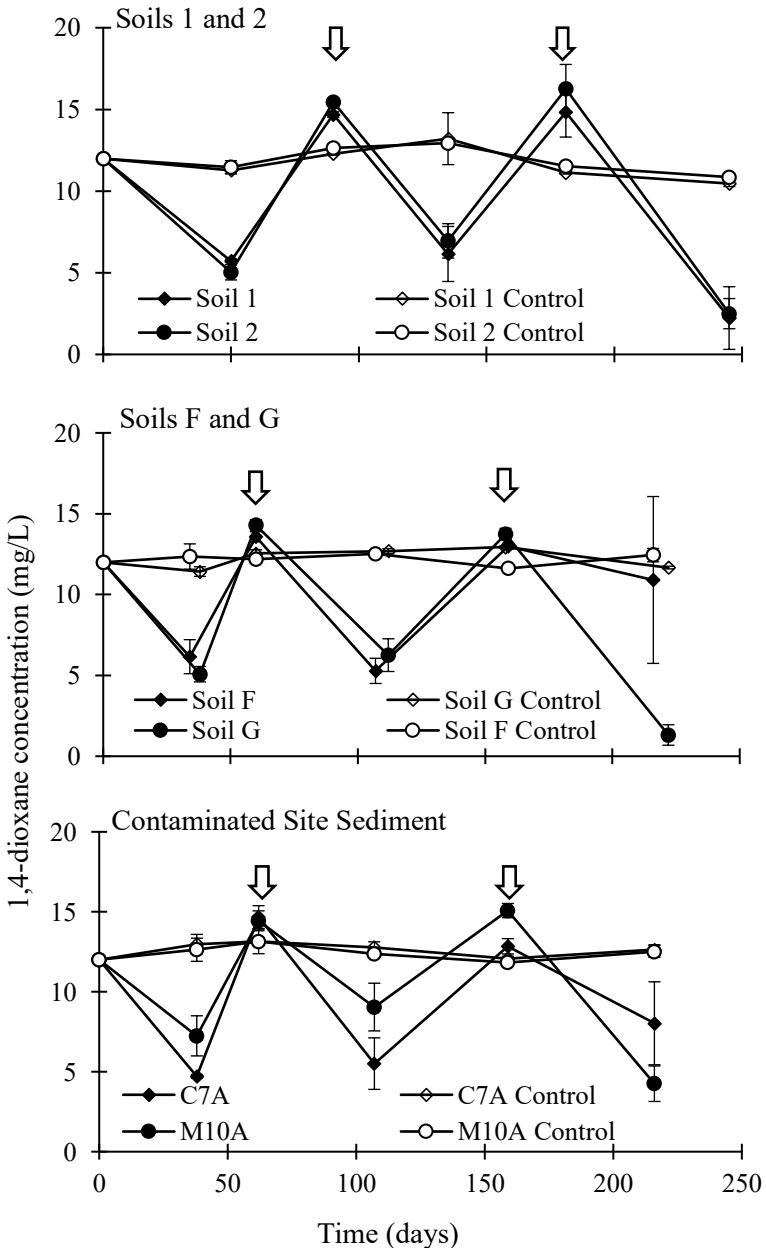

Soils 1 and 2
Soil 1
Soil 2
Soil 1 Control
Soil 2 Control
Soils F and G
Soil F
Soil G
Soil G Control
Soil F Control
Contaminated Site Sediment
C7A
M10A
C7A Control
M10A Control
1,4-dioxane concentration (mg/L)
Time (days)
0
5
10
15
20
0
50
100
150
200
250

Figure 1

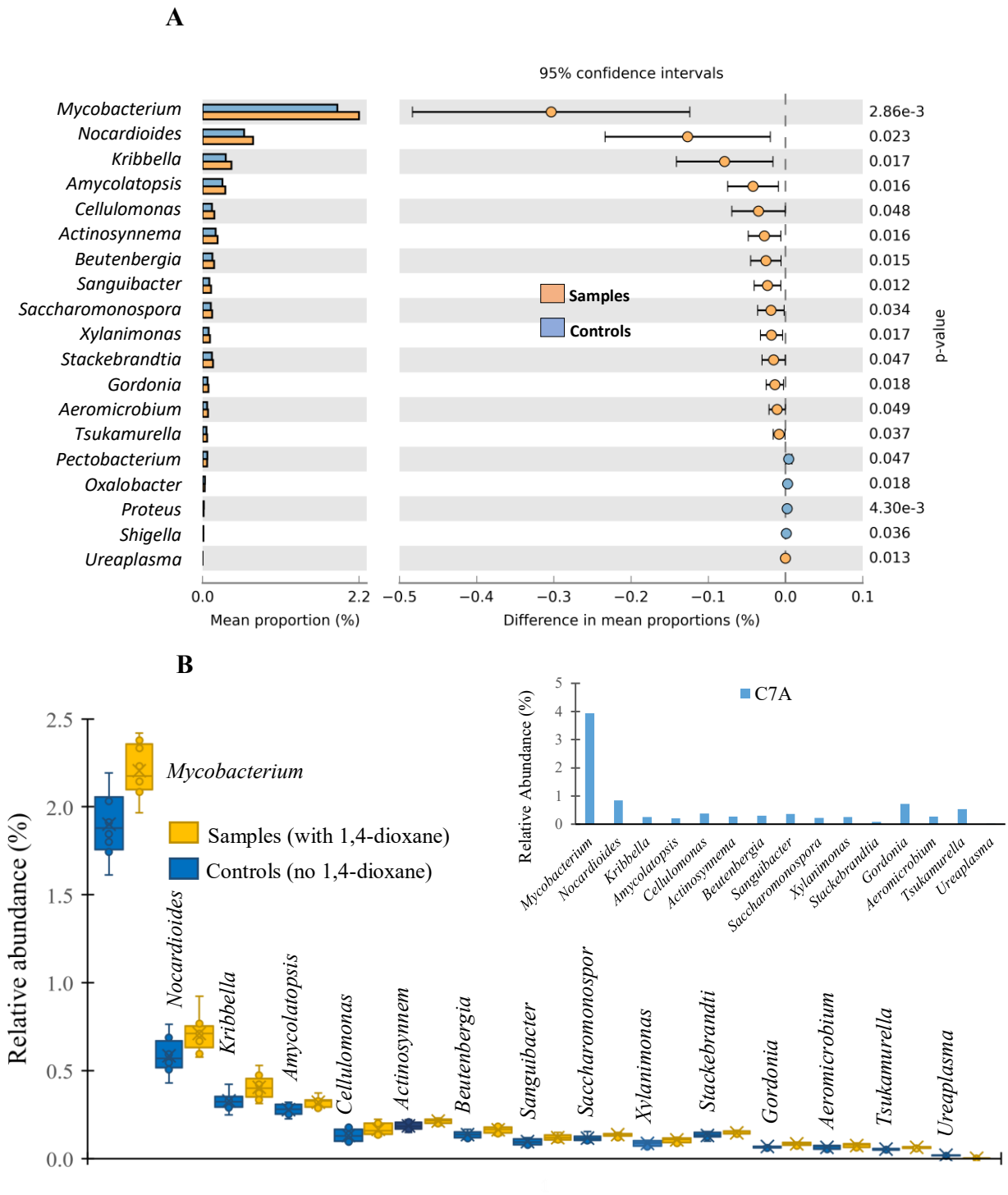

A
95% confidence intervals
Mycobacterium
Nocardioides
Kribbella
Amycolatopsis
Cellulomonas
Actinosynnema
Beutenbergia
Sanguibacter
Saccharomonospora
Xylanimonas
Stackebrandtia
Gordonia
Aeromicrobium
Tsukamurella
Pectobacterium
Oxalobacter
Proteus
Shigella
Ureaplasma
2.86e-3
0.023
0.017
0.016
0.048
0.016
0.015
0.012
0.034
0.017
0.047
0.018
0.049
0.037
0.047
0.018
4.30e-3
0.036
0.013
p-value
Samples
Controls
0.0
2.2
Mean proportion (%)
−0.5
−0.4
−0.3
−0.2
−0.1
0.0
0.1
Difference in mean proportions (%)
B
Relative abundance (%)
2.5
2.0
1.5
1.0
0.5
0.0
Mycobacterium
Samples (with 1,4-dioxane)
Controls (no 1,4-dioxane)
Nocardioides
Kribbella
Amycolatopsis
Cellulomonas
Actinosynnem
Beutenbergia
Sanguibacter
Saccharomonospor
Xylanimonas
Stackebrandti
Gordonia
Aeromicrobium
Tsukamurella
Ureaplasma
C7A
Relative Abundance (%)
5
4
3
2
1
0
Mycobacterium
Nocardioides
Kribbella
Amycolatopsis
Cellulomonas
Actinosynnema
Beutenbergia
Sanguibacter
Saccharomonospora
Xylanimonas
Stackebrandtia
Gordonia
Aeromicrobium
Tsukamurella
Ureaplasma

Figure 2

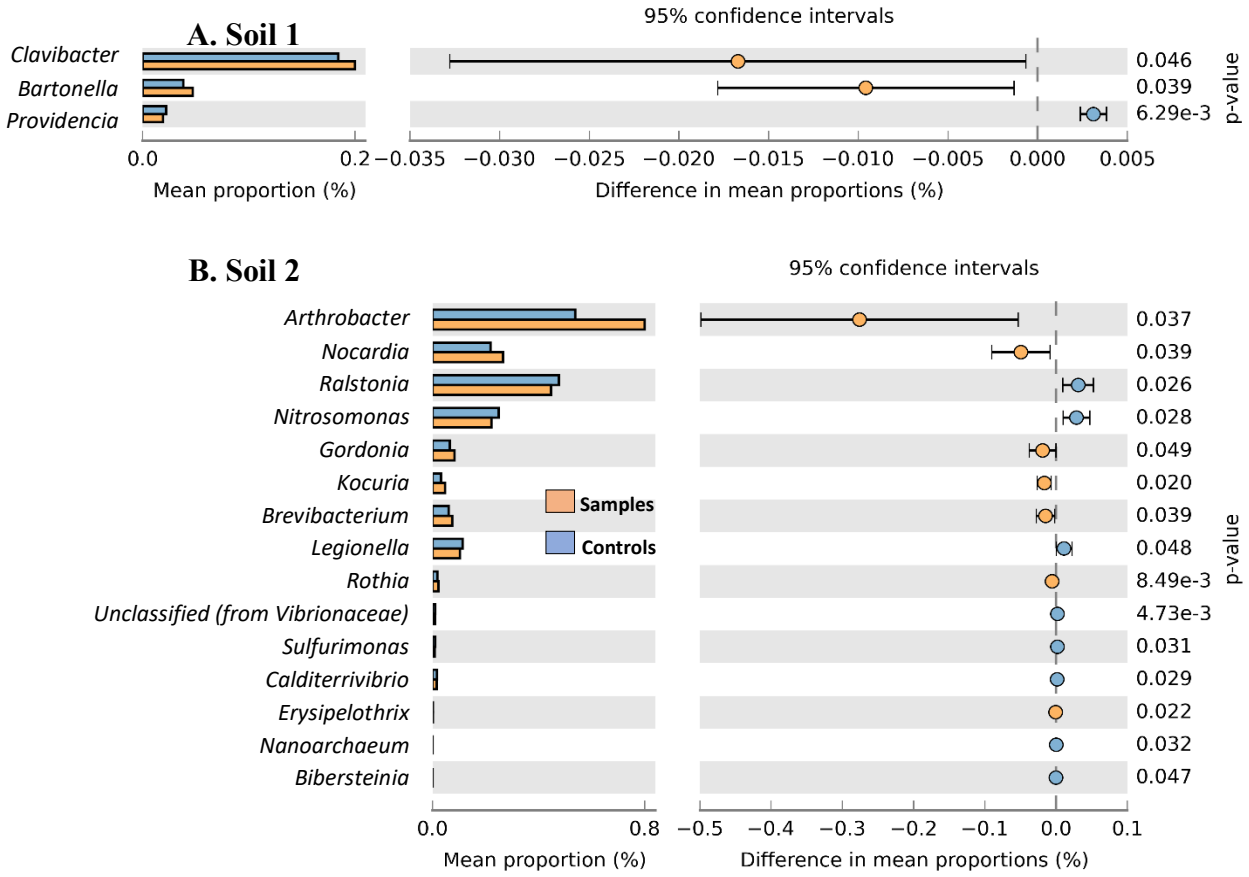

A. Soil 1
95% confidence intervals
Clavibacter
Bartonella
Providencia
0.046
0.039
6.29e-3
p-value
0.0
0.2
Mean proportion (%)
−0.035
−0.030
−0.025
−0.020
−0.015
−0.010
−0.005
0.000
0.005
Difference in mean proportions (%)
B. Soil 2
95% confidence intervals
Arthrobacter
Nocardia
Ralstonia
Nitrosomonas
Gordonia
Kocuria
Brevibacterium
Legionella
Rothia
Unclassified (from Vibrionaceae)
Sulfurimonas
Calditerrivibrio
Erysipelothrix
Nanoarchaeum
Bibersteinia
Samples
Controls
0.037
0.039
0.026
0.028
0.049
0.020
0.039
0.048
8.49e-3
4.73e-3
0.031
0.029
0.022
0.032
0.047
p-value
0.0
0.8
Mean proportion (%)
−0.5
−0.4
−0.3
−0.2
−0.1
0.0
0.1
Difference in mean proportions (%)

Figure 3

A. Soil F

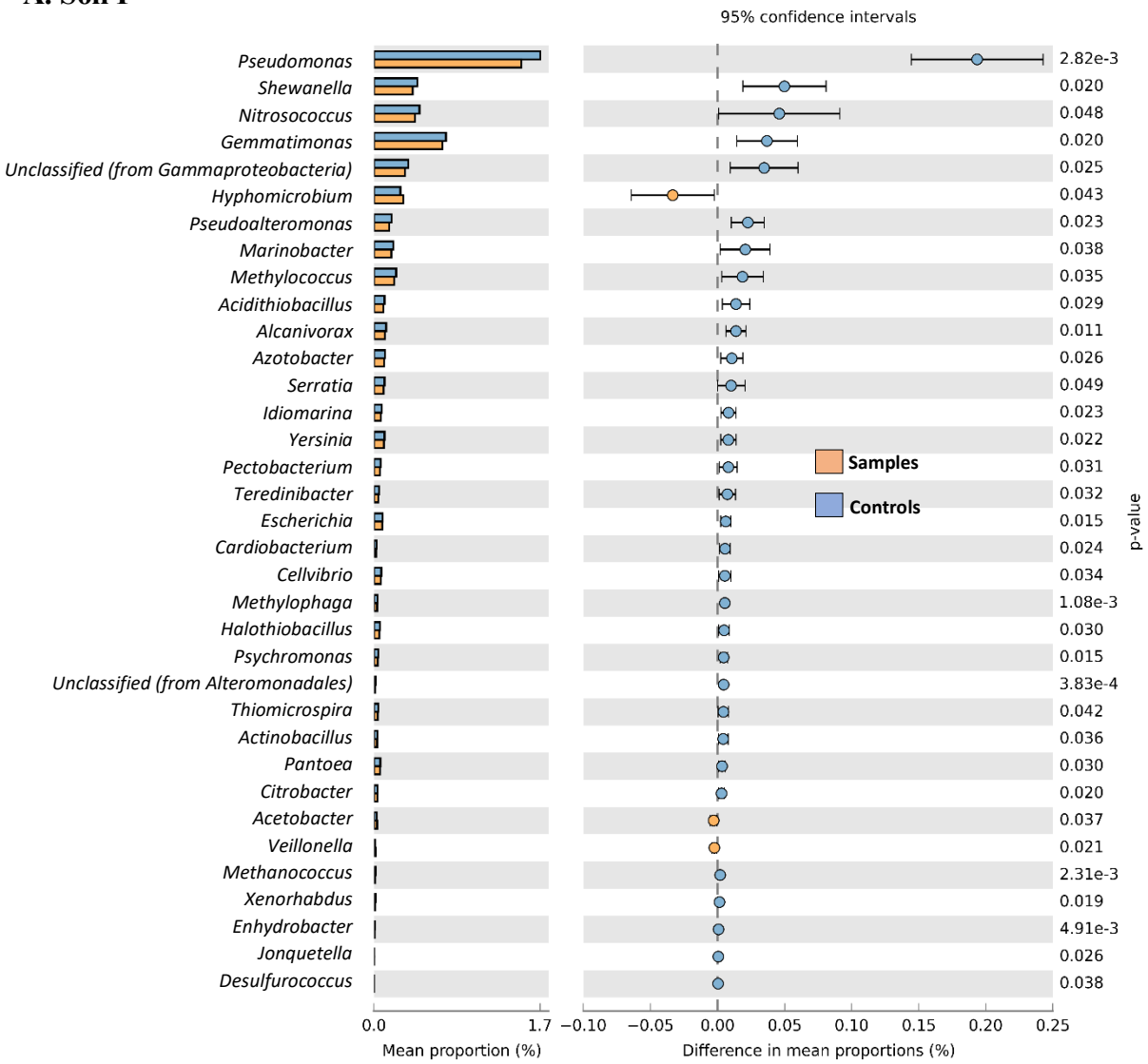
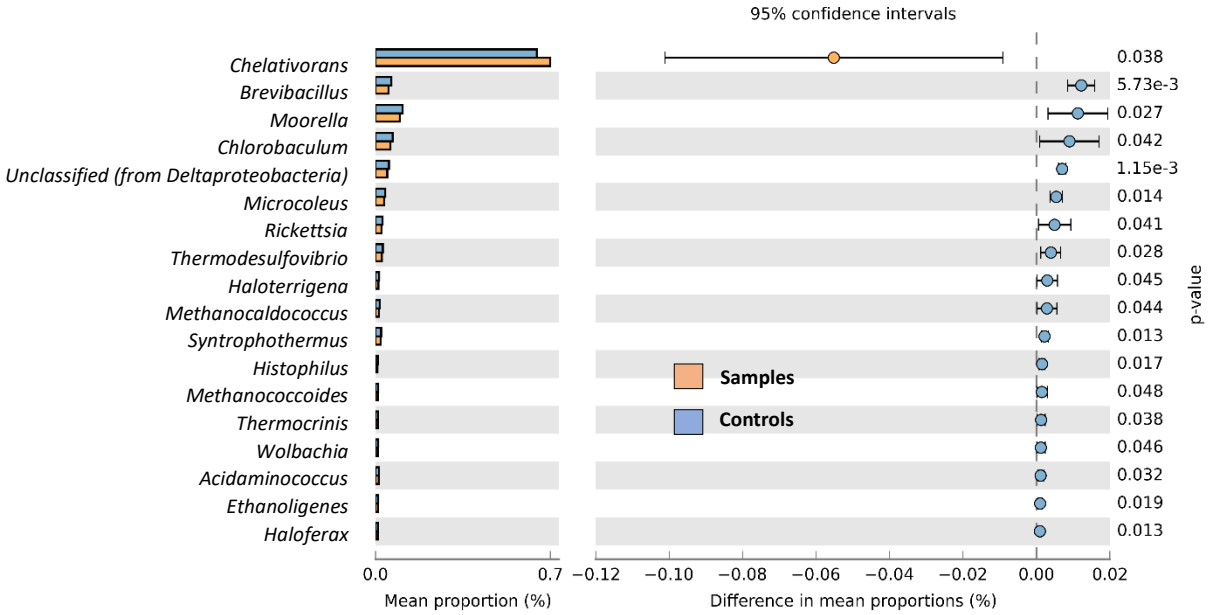


Figure 4

B. Soil G


95% confidence intervals
Chelativorans
Brevibacillus
Moorella
Chlorobaculum
Unclassified (from Deltaproteobacteria)
Microcoleus
Rickettsia
Thermodesulfovibrio
Haloterrigena
Methanocaldococcus
Syntrophothermus
Histophilus
Methanococcoides
Thermocrinis
Wolbachia
Acidaminococcus
Ethanoligenes
Haloferax
0.038
5.73e-3
0.027
0.042
1.15e-3
0.014
0.041
0.028
0.045
0.044
0.013
0.017
0.048
0.038
0.046
0.032
0.019
0.013
p-value
Samples
Controls
0.0
0.7
Mean proportion (%)
−0.12
−0.10
−0.08
−0.06
−0.04
−0.02
0.00
0.02
Difference in mean proportions (%)

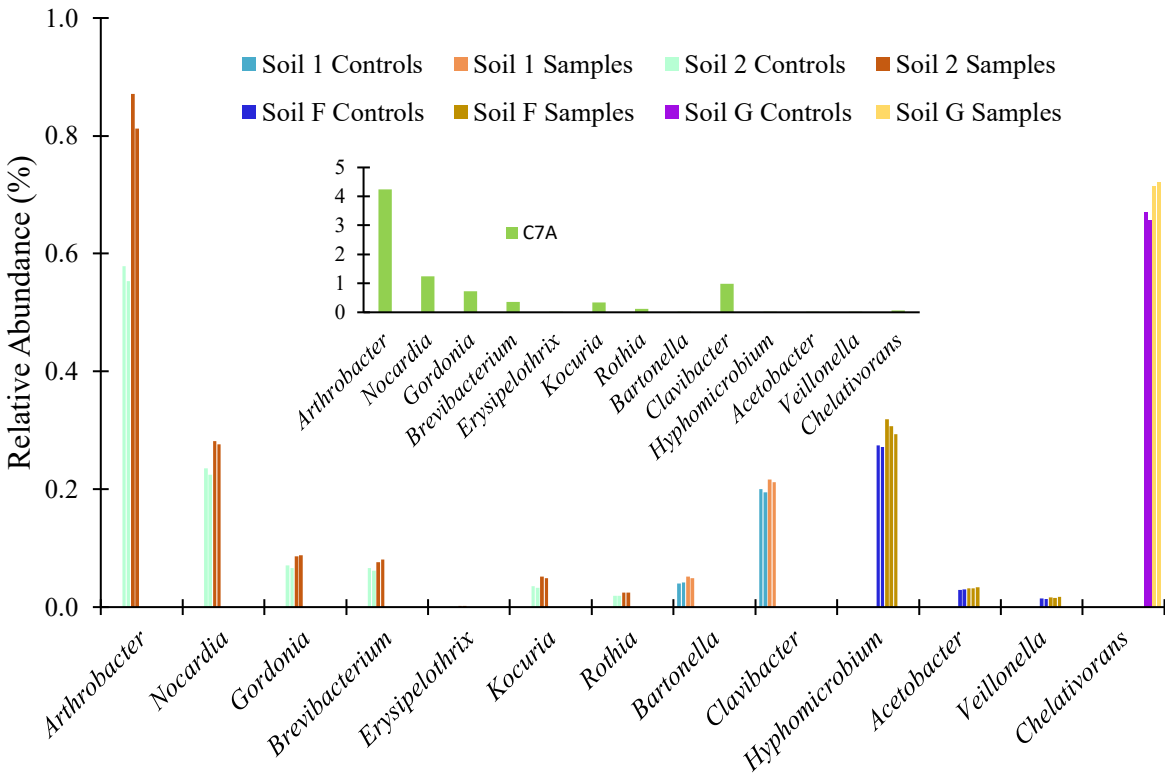

Soil 1 Controls
Soil 1 Samples
Soil 2 Controls
Soil 2 Samples
Soil F Controls
Soil F Samples
Soil G Controls
Soil G Samples
Relative Abundance (%)
1.0
0.8
0.6
0.4
0.2
0.0
C7A
5
4
3
2
1
0
Arthrobacter
Nocardia
Gordonia
Brevibacterium
Erysipelothrix
Kocuria
Rothia
Bartonella
Clavibacter
Hyphomicrobium
Acetobacter
Veillonella
Chelativorans

Figure 5

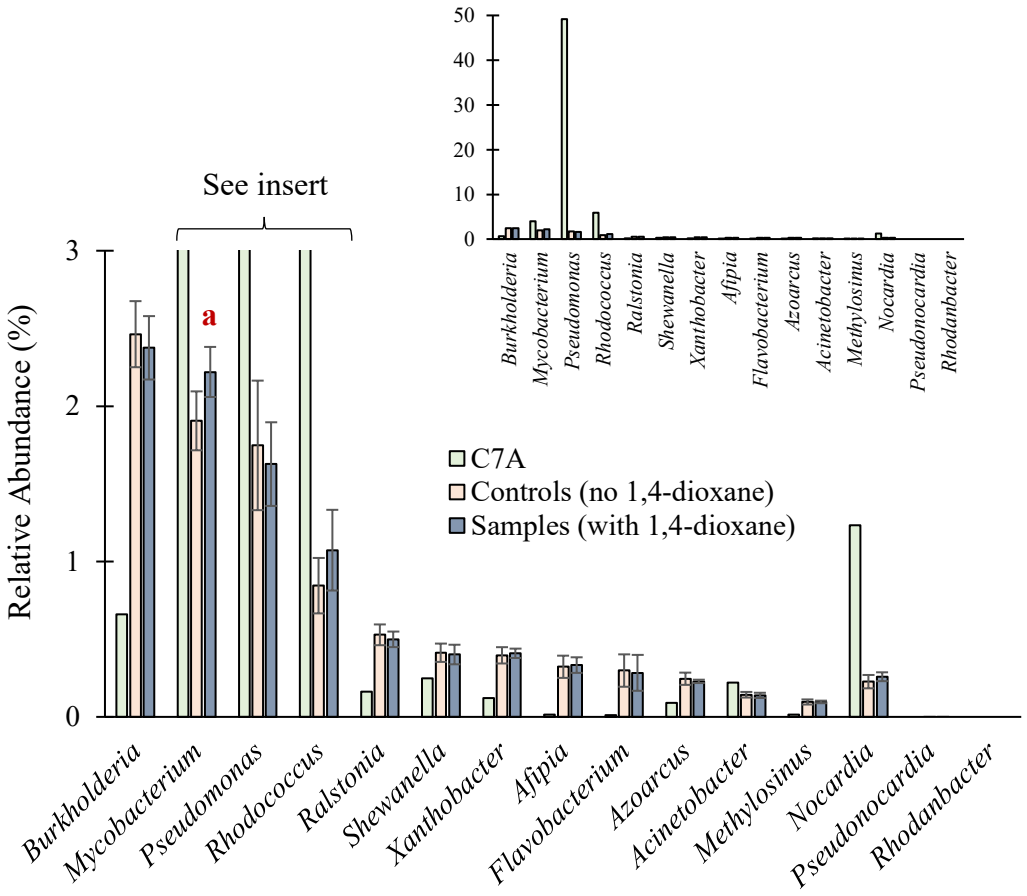

See insert
a
Relative Abundance (%)
C7A
Controls (no 1,4-dioxane)
Samples (with 1,4-dioxane)
Burkholderia
Mycobacterium
Pseudomonas
Rhodococcus
Ralstonia
Shewanella
Xanthobacter
Afipia
Flavobacterium
Azoarcus
Acinetobacter
Methylosinus
Nocardia
Pseudonocardia
Rhodanbacter

Figure 6

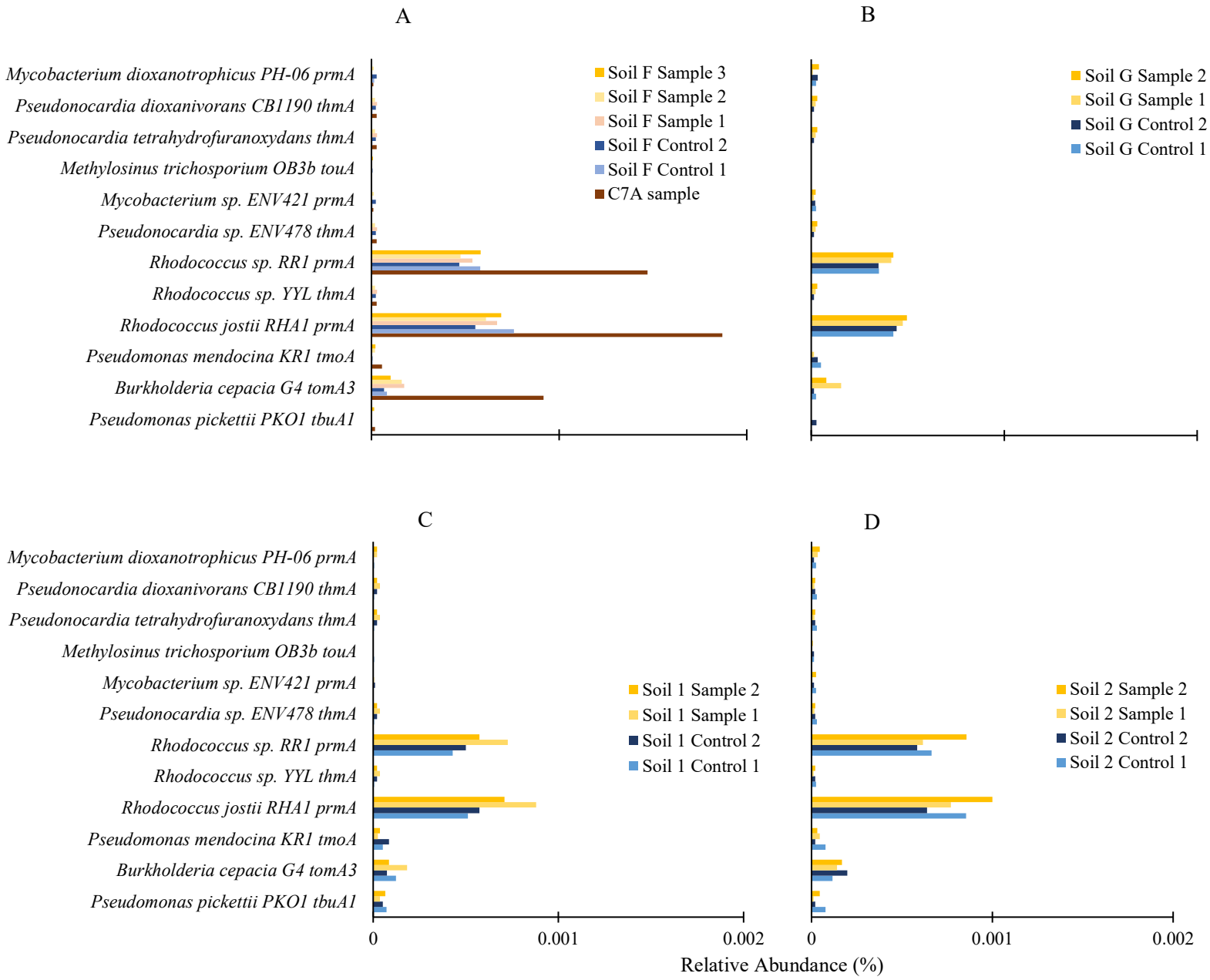

A
B
C
D
Mycobacterium dioxanotrophicus PH-06 prmA
Pseudonocardia dioxanivorans CB1190 thmA
Pseudonocardia tetrahydrofuranoxydans thmA
Methylosinus trichosporium OB3b touA
Mycobacterium sp. ENV421 prmA
Pseudonocardia sp. ENV478 thmA
Rhodococcus sp. RR1 prmA
Rhodococcus sp. YYL thmA
Rhodococcus jostii RHA1 prmA
Pseudomonas mendocina KR1 tmoA
Burkholderia cepacia G4 tomA3
Pseudomonas pickettii PKO1 tbuA1
Soil F Sample 3
Soil F Sample 2
Soil F Sample 1
Soil F Control 2
Soil F Control 1
C7A sample
Soil G Sample 2
Soil G Sample 1
Soil G Control 2
Soil G Control 1
Soil 1 Sample 2
Soil 1 Sample 1
Soil 1 Control 2
Soil 1 Control 1
Soil 2 Sample 2
Soil 2 Sample 1
Soil 2 Control 2
Soil 2 Control 1
0
0.001
0.002
Relative Abundance (%)

Figure 7

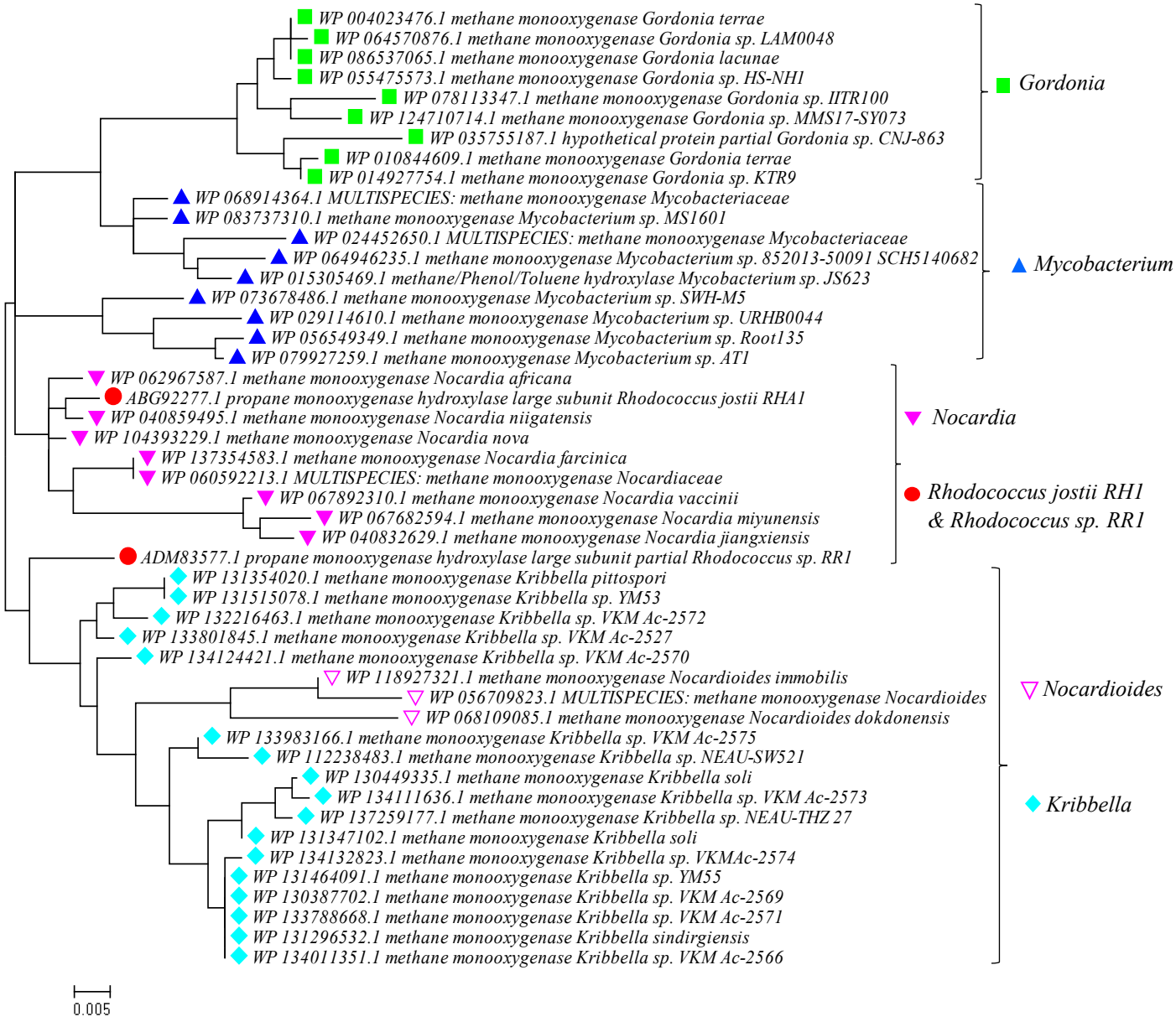

WP 004023476.1 methane monooxygenase Gordonia terrae
WP 064570876.1 methane monooxygenase Gordonia sp. LAM0048
WP 086537065.1 methane monooxygenase Gordonia lacunae
WP 055475573.1 methane monooxygenase Gordonia sp. HS-NH1
WP 078113347.1 methane monooxygenase Gordonia sp. IITR100
WP 124710714.1 methane monooxygenase Gordonia sp. MMS17-SY073
WP 035755187.1 hypothetical protein partial Gordonia sp. CNJ-863
WP 010844609.1 methane monooxygenase Gordonia terrae
WP 014927754.1 methane monooxygenase Gordonia sp. KTR9
Gordonia
WP 068914364.1 MULTISPECIES: methane monooxygenase Mycobacteriaceae
WP 083737310.1 methane monooxygenase Mycobacterium sp. MS1601
WP 024452650.1 MULTISPECIES: methane monooxygenase Mycobacteriaceae
WP 064946235.1 methane monooxygenase Mycobacterium sp. 852013-50091 SCH5140682
WP 015305469.1 methane/Phenol/Toluene hydroxylase Mycobacterium sp. JS623
WP 073678486.1 methane monooxygenase Mycobacterium sp. SWH-M5
WP 029114610.1 methane monooxygenase Mycobacterium sp. URHB0044
WP 056549349.1 methane monooxygenase Mycobacterium sp. Root135
WP 079927259.1 methane monooxygenase Mycobacterium sp. AT1
Mycobacterium
WP 062967587.1 methane monooxygenase Nocardia africana
ABG92277.1 propane monooxygenase hydroxylase large subunit Rhodococcus jostii RHA1
WP 040859495.1 methane monooxygenase Nocardia niigatensis
WP 104393229.1 methane monooxygenase Nocardia nova
WP 137354583.1 methane monooxygenase Nocardia farcinica
WP 060592213.1 MULTISPECIES: methane monooxygenase Nocardiaceae
WP 067892310.1 methane monooxygenase Nocardia vaccinii
WP 067682594.1 methane monooxygenase Nocardia miyunensis
WP 040832629.1 methane monooxygenase Nocardia jiangxiensis
Nocardia
ADM83577.1 propane monooxygenase hydroxylase large subunit partial Rhodococcus sp. RR1
Rhodococcus jostii RH1
& Rhodococcus sp. RR1
WP 131354020.1 methane monooxygenase Kribbella pittospori
WP 131515078.1 methane monooxygenase Kribbella sp. YM53
WP 132216463.1 methane monooxygenase Kribbella sp. VKM Ac-2572
WP 133801845.1 methane monooxygenase Kribbella sp. VKM Ac-2527
WP 134124421.1 methane monooxygenase Kribbella sp. VKM Ac-2570
WP 118927321.1 methane monooxygenase Nocardioides immobilis
WP 056709823.1 MULTISPECIES: methane monooxygenase Nocardioides
WP 068109085.1 methane monooxygenase Nocardioides dokdonensis
Nocardioides
WP 133983166.1 methane monooxygenase Kribbella sp. VKM Ac-2575
WP 112238483.1 methane monooxygenase Kribbella sp. NEAU-SW521
WP 130449335.1 methane monooxygenase Kribbella soli
WP 134111636.1 methane monooxygenase Kribbella sp. VKM Ac-2573
WP 137259177.1 methane monooxygenase Kribbella sp. NEAU-THZ 27
WP 131347102.1 methane monooxygenase Kribbella soli
WP 134132823.1 methane monooxygenase Kribbella sp. VKMAc-2574
WP 131464091.1 methane monooxygenase Kribbella sp. YM55
WP 130387702.1 methane monooxygenase Kribbella sp. VKM Ac-2569
WP 133788668.1 methane monooxygenase Kribbella sp. VKM Ac-2571
WP 131296532.1 methane monooxygenase Kribbella sindirgiensis
WP 134011351.1 methane monooxygenase Kribbella sp. VKM Ac-2566
Kribbella
0.005

Figure 8

Applied Microbiology and Biotechnology

Supplementary Section

Enrichment of Novel Actinomycetales and the Detection of Monooxygenases during Aerobic 1,4-Dioxane Biodegradation with Uncontaminated and Contaminated Inocula

Vidhya Ramalingam and A. M. Cupples

Department of Civil and Environmental Engineering, Michigan State University, East Lansing, Michigan, USA

*Corresponding Author:

Alison M. Cupples

A135, 1449 Engineering Research Court, Michigan State University, East Lansing, MI 48824,

cupplesa@egr.msu.edu

Table S1. Summary of the characteristics of the soils used to inoculate the sample and control microcosms.

Soil	Sand (%)	Silt (%)	Clay (%)	pH	Organic Matter %
E	70	19	11	7.3	1.9
F	64	25	11	6.6	1.5
G	68	20	12	5.5	1.6
T1	40	40	20	6.4	1.6
T2	36	42	22	6.1	1.9

Table S2. Summary of sequencing information processed by MG-RAST.

QC – Quality Control

ID	Name	Upload: bp Count	Upload: Sequences Count	Artificial Duplicate Reads: Sequence Count	Post QC: bp Count	Post QC: Sequences Count	Post QC: Mean Sequence Length bp
mgm4846244.3	C7A_6_S10_L001_R	928,461,523	3,828,893	730,886	745,308,161	3,067,261	243 ± 33
mgm4842040.3	SF_2_S3_L001_R	1,249,466,781	5,278,200	741,761	1,057,029,663	4,469,282	237 ± 36
mgm4846245.3	SF_3_S6_L001_R	1,497,947,613	6,267,333	1,045,636	1,227,596,223	5,142,267	239 ± 35
mgm4846246.3	SF_4_S8_L001_R	698,047,178	2,986,100	484,606	572,423,699	2,454,542	233 ± 36
mgm4846247.3	SF_5_S11_L001_R	1,455,181,423	6,244,496	875,408	1,226,035,914	5,268,011	233 ± 37
mgm4846248.3	SF_6_S13_L001_R	1,276,229,786	5,446,704	791,730	1,069,530,685	4,570,187	234 ± 36
mgm4846291.3	SG_1_S15_L001_R	362,019,482	1,517,474	232,203	300,578,338	1,261,641	238 ± 35
mgm4841972.3	SG_3_S17_L001_R	1,237,320,023	5,151,224	739,217	1,042,837,451	4,346,084	240 ± 34
mgm4841973.3	SG_4_S1_L001_R	1,183,825,731	4,906,054	779,630	975,862,599	4,048,734	241 ± 34
mgm4842102.3	SG_6_S4_L001_R	1,097,936,207	4,598,102	612,987	933,115,160	3,911,767	239 ± 35
mgm4841974.3	ST1_2_S7_L001_R	1,722,580,822	7,248,808	1,205,707	1,409,132,207	5,932,625	238 ± 35
mgm4846290.3	ST1_3_S9_L001_R	687,193,817	2,855,664	394,713	580,636,546	2,415,127	240 ± 35
mgm4842023.3	ST1_4_S12_L001_R	1,720,695,830	7,269,781	1,124,424	1,433,388,134	6,053,609	237 ± 35
mgm4842024.3	ST1_6_S14_L001_R	1,165,586,999	4,879,142	683,336	983,135,037	4,119,069	239 ± 35
mgm4842104.3	ST2_1_S16_L001_R	3,144,582,833	13,429,024	2,172,387	2,579,894,248	11,026,268	234 ± 36
mgm4842103.3	ST2_3_S18_L001_R	1,274,658,549	5,369,262	727,123	1,080,887,053	4,557,307	237 ± 35
mgm4842107.3	ST2_5_S2_L001_R	1,622,482,462	6,854,685	982,861	1,362,403,381	5,762,840	236 ± 35
mgm4842106.3	ST2_6_S5_L001_R	1,463,433,476	6,249,417	846,173	1,252,082,629	5,349,932	234 ± 36

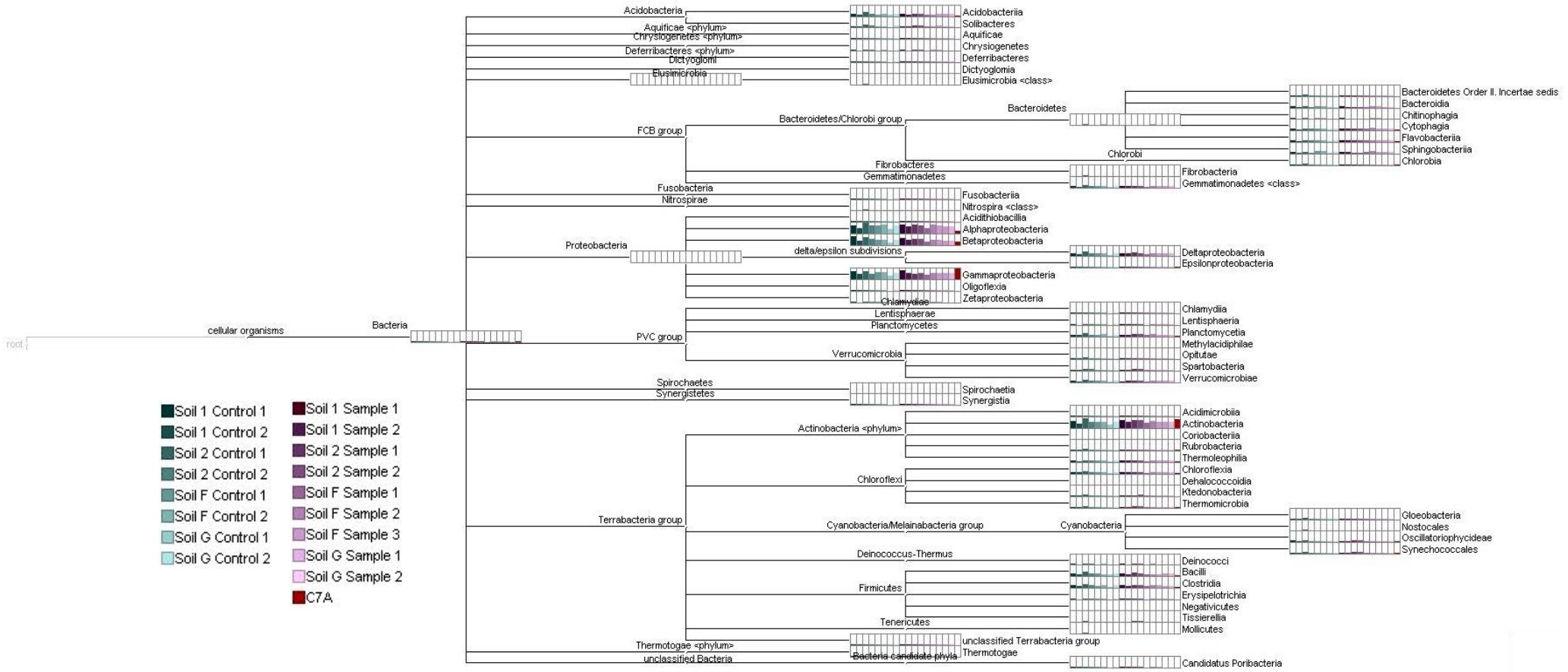


Figure S1. Phylogram (created with MEGAN6, version 6.11.7) illustrating the relative abundance and classification (Class Level) of all bacteria across all metagenomes (samples and controls).

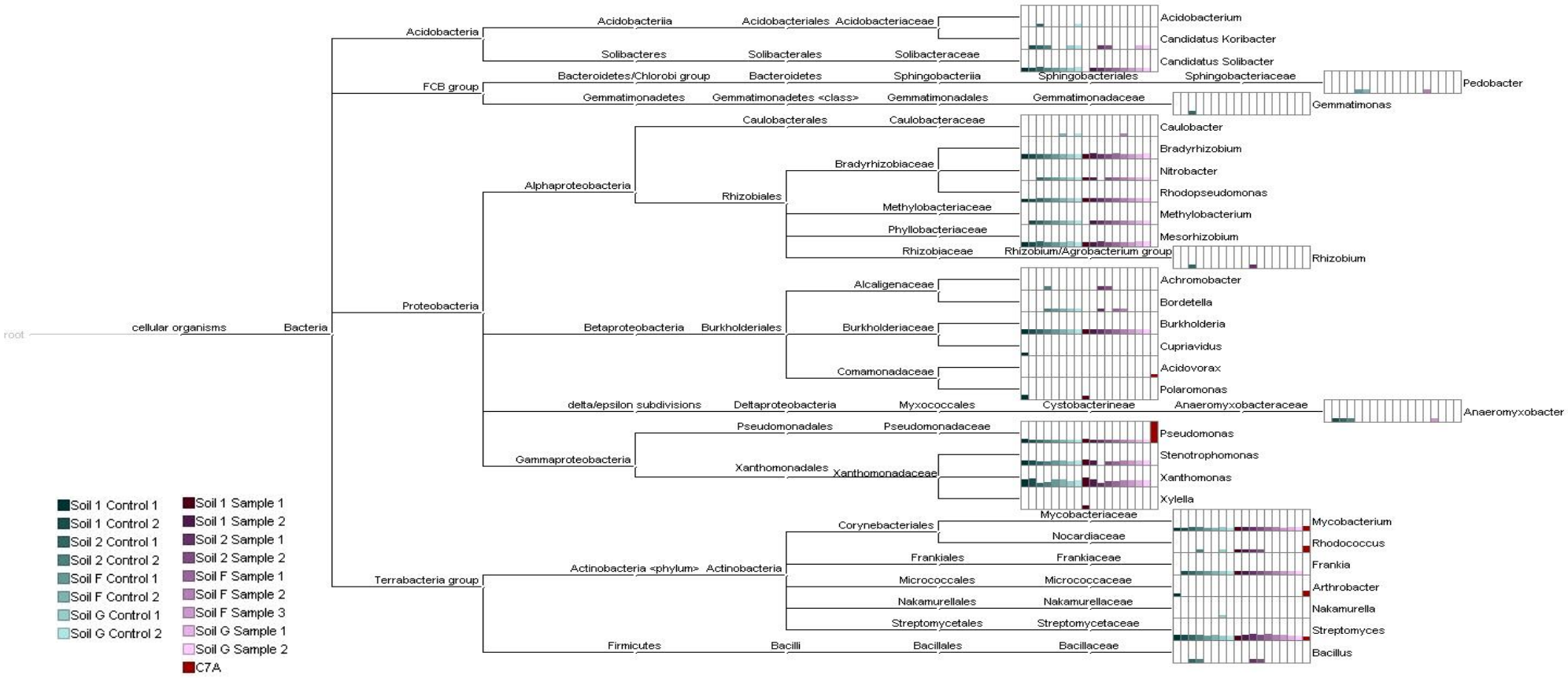


Figure S2. Phylogram (created with MEGAN6, version 6.11.7) illustrating the most abundant genera (ranked by average relative abundance, then selected if average relative abundance >0.5%) across all metagenomes (samples and controls).

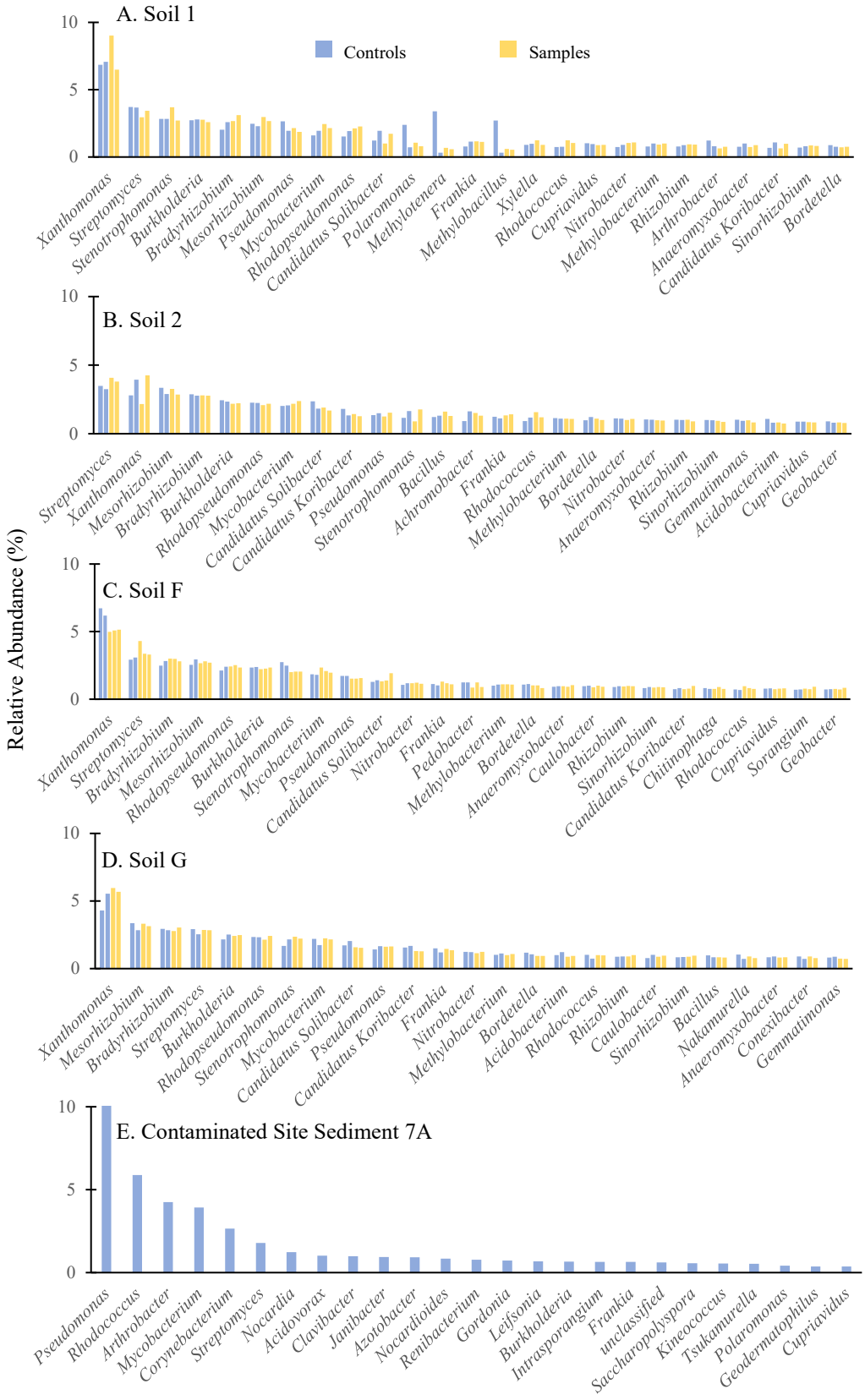


Figure S3. The twenty-five most common genera (by relative abundance, %), ranked by the averages of the samples and controls, in soil 1 (A), soil 2 (B), soil F (C), soil G (D) and the contaminated site sediment 7A (E).