Inhibition of Microbial Methylation via arsM in the Rhizosphere: Arsenic Speciation in the Soil to Plant Continuum

Characterization of Natural and Affected Environments

Inhibition of Microbial Methylation via arsM in the Rhizosphere: Arsenic Speciation in the Soil to Plant Continuum

Hasina Afroz, Shiming Su, Manus Patrick Carey, Andrew A. Meharg, and Caroline Meharg

Environ. Sci. Technol., **Just Accepted Manuscript** • DOI: 10.1021/acs.est.8b07008 • Publication Date (Web): 15 Mar 2019

Inhibition of Microbial Methylation via *arsM* in the Rhizosphere: Arsenic Speciation in the Soil to Plant Continuum

Hasina Afroz[1], Shiming Su[2], Manus Carey[1], Andy A. Meharg[*,1] and Caroline Meharg[*,1]

[1]*Institute for Global Food Security, Queen's University Belfast, David Keir Building, Malone Road, Belfast, BT9 5BN, UK.*

[2]*Institute of Environment and Sustainable Development in Agriculture, Chinese Academy of Agricultural Sciences/Key Laboratory of Agro-Environment, Ministry of Agriculture, Beijing, P.R. China.*

***Corresponding authors**

A.A. Meharg, aa.meharg@qub.ac.uk, +44 (0) 28 9097 5413

C. Meharg, caroline.meharg@qub.ac.uk, +44 (0) 28 9097 6880

ABSTRACT

The interplay between rice roots and manuring with respect to arsenic speciation, subsequent assimilation into roots, and translocation to shoots in paddy soil was investigated, alongside bacterial diversity characterization. Planting increased soil Eh and decreased soil solution arsenic species: inorganic arsenic (iAs), monomethylarsonic acid (MMA), trimethylarsenic oxide (TMAO) and dimethylarsinic acid (DMA). Presence of plant roots increased copy number of *Clostridium* and *Tumebacillus* 16S rRNA as well as *Streptomyces* arsenic methylating gene (*arsM*), but decreased *Acidobacteria_GP1* 16S rRNA and *R. palustris BisB5 arsM*. Sum of arsenic species decreased under root influence due to the interplay of inorganic arsenic mobilization in bulk soil under anaerobic and immobilization under oxygenated rhizospheric conditions. Manuring increased all soil solution arsenic species (>90%), shoot total arsenic (60%), copy number of *Geobacter* 16S rRNA and *R. palustris TIE-1 arsM*, indicative of a shift towards microbes with iron reduction and oxidation as well as arsenic methylation capabilities.

INTRODUCTION

Rice grain is *circa.* 10-fold elevated in inorganic arsenic, a class-one, non-threshold carcinogen, compared to virtually all other arable crops.[1] The levels of inorganic arsenic in rice grain is such that they are deemed a threat to human health and have led to legislation to regulate inorganic arsenic's concentrations in rice products by the European Union at 0.2 mg/kg As in white rice,[2] with the WHO setting advisory levels at this concentration.[3] Young children are considered to be at particular threat and stricter levels have been set at 0.1 mg/kg of arsenic for rice based baby foods by the European Union[2] and, at the time of writing, this threshold for rice based baby foods is also under consideration in the USA.[4] The introduction of food-standards will challenge the food processing industries to produce rice products lower in inorganic arsenic as baby foods currently exceed these standards in circa 50% of samples tested in the EU.[5]

The reason why rice is problematic with respect to inorganic arsenic is simply due to the soil reduced redox conditions that predominate under paddy rice cultivation.[6] Arsenic, as arsenate, from sources such as the weathering of rocks, builds up in iron (III) oxyhydroxide (iron plaque) that coats soil particles and sediments under aerobic conditions. Under anoxic conditions insoluble iron (III) is reduced to soluble iron (II) and relatively insoluble arsenate (As(V)) to the more soluble arsenite (As(III)). It is As(V) and As(III) together that are termed "inorganic arsenic", iAs, as they are readily interchangeable depending on prevailing abiotic and biotic conditions. As(V) and As(III) are analogues of the nutrients phosphate and silicic acid, respectively.[6] Rice has a high demand for both phosphate and is a silicic acid accumulator.[7] If iAs is available to roots, such as under paddy cultivation, rice will take up

high levels of iAs, to such an extent that rice shoot to soil arsenic concentrations are in the order of 1:1,[8] remarkable for a non-essential toxicant.

iAs species are not the only forms of arsenic found in rice grain. Dimethylarsinic acid (DMA) can contribute over 50% of the arsenic concentration in grain, with DMA concentrations in rice being highly variable around the globe.[1] DMA is not produced *in planta*[9] and derives from the methylation of iAs, metabolised by soil bacteria that express *arsM* genes,[10] with subsequent assimilation of DMA by roots.[11] Although roots are relatively inefficient at assimilating DMA, as compared to arsenate or arsenite, DMA is much more efficiently translocated to grain than iAs species, primarily as it is not as retarded in its progression through the plants vascular system by phytochelatin complexation that iAs species undergo in planta.[12] Under reducing conditions *arsM* catalyses a stepwise conversion of iAs to monomethylarsonic acid (MMA), DMA, trimethylarsenic oxide (TMAO) and finally tetramethylarsonium (TETRA).[10] All of these species, with the exception of TMAO, can be found in polished rice grain though DMA predominates.[13]

The reason why DMA is variable in rice throughout the globe is not known. Evidence suggests that soil organic matter (SOM), which is a key driver in soil redox has a role with methylated species in soil solution correlating with soluble SOM, and methylation increasing with different organic matter (OM) amendments to soil.[14,15] The interplay between the rhizosphere and arsenic speciation dynamics also may have a role. Rice aerates its roots through aerenchyma leading to their (partial) coating in iron plaque with subsequent sequestration of arsenic, as both arsenate and arsenite.[16-18] Bulk soil then acts as a source of labile As(III), which is moved due to advective transpiration-driven forces and via molecular diffusion along a concentration gradient to the root.[19-22] Microbes expressing *arsM* genes are

highly diverse in soil and root rhizosphere, both phylogenetically and ecologically,[20,23] which could have significant role in biogeochemical cycling of arsenic. However, it is not known how the rhizosphere and amendments with OM interplays with arsenic methylation and subsequent assimilation by the roots, and how this varies with soils of differing origin. Planted and unplanted microcosms, with soils sampled across one of the main EU rice growing regions, Valencia, Spain, were monitored through time to investigate methylation of arsenic to further understand the role of the rhizosphere in rice assimilation of arsenic. Here we investigated the arsenic speciation in the soil to shoot continuum with concurrent monitoring of soil bacterial genomic *arsM* and 16S rRNA copy number through qPCR. Following this, we chose 1 soil to investigate the interaction of planting and manuring, using the same methods plus paired amplicon sequencing[24,25] of *arsM* and 16S rRNA for in depth phylogenetic analysis of arsenic methylating soil microbial communities.

MATERIALS AND METHODS

Experimental design

Paddy soils, classified as Entisols, were collected from five different fields across the region of Valencia, Spain, as described and characterized by Signes-Pastor et al.,[26] were used in this study. The location and characteristics of soils are shown in Table S1. Field moist soils were sieved to < 2mm before the experiment. Rice seed (IRGA 425) was the plant cultivar used. Before planting, rice seeds were sterilized with H_2O_2 and then germinated at 25^0C for 5 days. Two different microcosm experiments were conducted.

Experiment 1 (5 soils, planted/unplanted): For the first microcosm experiment field moist soil (for details on soil 1-5 see Table S1) was added to 50 ml centrifuge tubes to an equivalent of 30 g dry weight. Pre-germinated seedlings were planted into half the microcosms resulting in unplanted and planted treatments. In total 90 microcosms (5 soils, 2 treatments, 9 time points) were set up to enable 9 destructive harvests at weekly spacing, with each soil represented by one planted and unplanted replicate at each harvest. On planting the microcosms had ultra-pure water added to maintain a 2cm layer of standing water above the soil surface. Microcosms were then placed into plant growth chambers (Fitotron) with day/night temperature 28^0C /25^0C, light period 16h per day. Rice seedlings were harvested weekly, starting at 2 weeks. Rice seedlings were taken out of tubes by shaking the tube and carefully removing the adhering soils. After repeated rinsing with ultra-pure water rice shoot and root were separated and stored at -20^0C until subsequent analysis. Once the plant was removed the centrifuge tubes were placed in a centrifuge (Sorvall Legend RT at 4600 rpm) for 20 minutes, soil porewater decanted and both soil and porewater immediately frozen and stored at -20^0C until subsequent analyses. Chemical and molecular analysis (qPCR) from soil, plant and pore water samples was conducted (described below).

Experiment 2 (1 soil, full factorial design, planted/unplanted and manured/non-manured): For the second microcosm experiment soil-5 (see Table S1) was incubated with four treatment combinations and five replicates of each treatment: non-manured unplanted (NMNP), non-manured planted (NMP), manured unplanted (MNP) and manured planted (MP), for molecular and chemical analysis. Bovine farm Yard Manure (FYM) was dried in an oven at 70°C to a constant weight and ground to fine powder. Following this, the total element content of FYM was analysed by X-ray florescence spectrophotometry (XRF). For manure treated microcosms the ground and dried FYM was applied to the soil @ 10% dry weight basis, and mixed very well before rice seedlings were planted into half of the manured microcosms. Ultra-pure water was added to maintain flooded condition during the incubation. Microcosms were then placed into a plant growth chamber (Fitotron) for 2 weeks using the same growth condition as in microcosm experiment 1. After 2 weeks, samples were collected as described above for experiment 1. Chemical and molecular analysis (qPCR) from soil, plant and pore water samples was conducted as described below. In addition to these, amplicon sequencing of soil DNA to assess the diversity of bacterial *arsM* and 16S rRNA was conducted (described below). The Eh and pH was measured in planted and non-planted microcosms by inserting an Eh/pH meter through the soil surface, with the live plant still in place in the planted microcosm, to a depth of 2 cm and then waiting for the reading to stabilize before recording that reading.

Arsenic speciation in soil pore water

Arsenic species in 0.4 micron Millipore membrane filtered soil porewaters (iAs, MMA, DMA, TMAO and TETRA) were determined using a Dionex IC chromatographic system interfaced with ICP-MS (ICS-5000 DC, Thermo Scientific) as outlined in detail in Signes-

Pastor et al. [5,26] Using an arsenic speciation specific IC-column, IC-ICP-MS, unlike HPLC – anion exchange – ICP-MS, separates TMAO, and we run authentic TMAO samples, and cross validate with cation exchange IC-ICP-MS[27]. We have added TMAO to the list of analytes detected in soil porewaters. Results were expressed on a per kg soil (μg/kg) and statistical analysis performed in Minitab as outlined below.

Plant analysis

Plant shoot and root material was digested using concentrated nitric acid microwave digestion, again following protocols outlined in Signes-Pastor et al. for total arsenic (experiment 2) and arsenic species (experiment 1).[5,26] Statistical analysis was performed in Minitab as outlined below. For further details see the supporting information 1.

DNA extraction

Soil DNA was extracted from 0.5g soil from each microcosm using the Powerlyzer® PowersSoil® DNA isolation kit (MOBIO Laboratories, Inc.) following the manufacturer's instructions. For further details see the supporting information 1.

Quantitative PCR

Relative bacterial 16S rRNA and *arsM* copy number were measured on the Eppendorf mastercycler (Realplex[4], Hamburg, Germany) using precisionPLUS MasterMix premixed with SYBR green (PrimerDesign, USA). Primer pairs of 16S_1369F/16S_1492R[21] and *arsM*F1/*arsM*R2[20] were used to amplify and quantify the relative copy number of bacterial 16S and *arsM* genes, respectively. For further details see the supporting information 1.

Amplicon sequencing and processing of the resultant data

Following quantification, 20 DNA samples extracted from soil of each microcosm of experiment 2, were submitted for 250bp paired end amplicon sequencing of the 16S rRNA (V4 region) and *arsM* gene on the Miseq. Initial processing of the data involved adapter trimming[28] and clipping[29] followed by quality control analysis using fastqc[30] and merging of forward and reverse paired reads with fastq-join[31]. For further details see the supporting information 1.

Bioinformatics and statistical analysis of the amplicon sequencing data

Both 16S rRNA and *arsM* merged sequences were processed using QIIME version 1.8[32] for generation of absolute and relative OTU count-tables, taxonomic assignment and alpha and beta diversity analysis. Quality filtering of both 16S rRNA and *arsM* OTUs involved removal of OTUs with lowest abundance. The *arsM* OTUs were also filtered to remove any OTUs with <70% identity and query coverage to *arsM* protein sequences (blastx against nr), followed by taxonomic annotation based on blastn against bacterial genomes and *arsM* related sequences downloaded from the fungene database[33]. For 16S rRNA, phyla (level 2) and genera (level 6) showing relative abundance >1% were identified, followed by statistical analysis (see section Minitab below). Statistical analysis on the overall *arsM* OTU and genus level 16S rRNA OTU count tables was performed in R with DESeq2 using the False Discovery Rate (FDR) <0.05 and an absolute log2 foldchange (Log2FC) >1 as cutoff for significance. Identified significant changes in copy number were assigned to 4 groups: Group I: OTUs with significant plant effect and increased copy number in unplanted soil, Group II: OTUs with significant plant effect and increased copy number in planted soil, Group III: OTUs with significant manure effect and increased copy number in manured soil, Group IV: OTUs with significant manure effect and increased copy number in non-manured soil. OTUs

within each group were visualised via a heatmap (R version 3.2.2, R package gplots and heatmap2), variance partition analysis (R version 3.2.2, R package variencePartition[34]), network analysis (Qiime version 1.8[32], Cytoskape version 3.7[35]) and for each *arsM* OTU in Group I-IV the most highly correlated Group I-IV genus level 16S rRNA OTU identified (Spearman's rank correlation, GraphPad Prism version 6.0 for MAC OS). For further details see the supplementary information 1.

General linear modelling

Minitab version 16 (Minitab, PA, USA) was used as the statistical platform for the analysis of chemical data, qPCR data and relative count data of abundant phyla and genera. For plant and soil solution speciation data, qPCR and relative count data of abundant phyla and genera (Level 2 and level 6 16S rRNA amplicon data), General Linear Modelling (GLM) was employed. For Eh and pH measurements the un-paired t-test was used for identification of significant differences. Results were plotted in GraphPad Prism (version 7). For further details see the supporting information.

RESULTS

Microcosm experiment 1 (5 soils, planted/unplanted)

The location and properties of the 5 soils used in experiment 1 are presented in Table S1. Soil of origin was significant for all arsenic species soil solution concentrations with the exception of TMAO ($P>0.05$), and both *arsM* and 16S relative copy number varied significantly between soils (Figure S1, Table S2). The iAs was the most variable between soils ranging, on average, 4-fold between soils (0.2-0.8 μg/kg). The soil solution concentrations of all arsenic species were significantly suppressed in the planted microcosms as compared to non-planted microcosms (Figure S1, Table S2). The As species were dominated by iAs and were

239 suppressed by >30% in planted microcosms in all 5 soils (Figure S1). Similar levels of
240 suppression were observed for DMA and TMAO. The greatest suppression of arsenic species
241 in soil solution when comparing unplanted versus planted soil was observed for MMA.
242 Further to that qPCR showed a significant decrease in relative *arsM* copy number in planted
243 compared to unplanted soils (Figure S1, Table S2).
244
245 Considering the interaction terms in the statistical model, for planted/unplanted*soil, there
246 was a significant interaction ($p<0.05$) for each arsenic species, and for 16S rRNA relative
247 copy number ($P<0.05$), but not for *arsM* relative copy number (Table S2). This showed that
248 the soils are inherently variable in the way they interact with planting with respect to arsenic
249 speciation, but that the *arsM* DNA copy number is relatively consistent across soils between
250 the two treatments, i.e. that the treatment was the dominant factor in *arsM* copy number as
251 this was highly significant ($P=0.002$). Treatment*time interaction for iAs and 16S relative
252 copy number were also significant ($P<0.05$). Soil*time interactions (Table S2) were
253 significant for iAs, MMA and DMA, indicating that the individual soil
254 microbiology/mineralogy varied with respect to temporal patterns of arsenic release.
255
256 Plant biomass production progressed in a linear manner throughout the experiment, positively
257 correlated with time, with root and shoot biomass being similar in quantity (Figure S2). iAs
258 concentrations were over an order of magnitude higher in roots compared to shoots, with
259 roots initially having 100 mg/kg iAs. The iAs content of both root and shoot declined over
260 time, in shoots more so than roots (Figure S2). Similar magnitudes of differences between
261 roots and shoots were also observed for MMA, though at week 1 MMA was quite high in
262 shoot compared to future time points. MMA concentrations were ~2-orders of magnitude
263 lower than iAs concentrations, though were similar to DMA concentrations in magnitude, at

least towards the end of the experiment. DMA in shoots was initially higher at week 1, but declined through the experiment, while root concentrations remained consistent (Figure S2). TMAO concentrations in both root and shoot converged at 0.01 mg/kg, 3-orders of magnitude lower than iAs in shoots (Figure S2). The sum of arsenic species in plant tissue (root and shoot) constituted 0.17 – 1.24 % of arsenic in the microcosms, from week 1 to week 9, respectively. Given that root biomass is intimately associated with soil, plant biomass arsenic is only a small sink relative to soil stores.

Microcosm experiment 2 (1 soil, planted/unplanted and manured/non-manured)

Soil solution arsenic species analysis and qPCR

Soil solution arsenic species concentrations in microcosm experiment 2 were suppressed significantly in planted soil compared to unplanted soil ($P>0.05$) (Figure 1, Table S3). The As species were dominated by iAs, which decreased in planted microcosm and increased in response to manure treatment (Figure 1). MMA, DMA and TMAO also decreased in planted microcosm and increased in response to manure treatment (Figure 1). Planting was shown to significantly increase Eh from -28 mV to -14 mV ($p<0.001$) and decrease pH from 7.5 to 7.25 ($p<0.001$) (Figure S3). As in experiment 1 there was a significant decrease in relative *arsM* copy number in the planted microcosms ($p=0.001$), as well as relative *arsM*/16S ($p=0.002$), while 16S relative copy number remained unchanged (Figure 1, Table S3).

FYM had significant influence on the soil solution As species concentration and 16S relative copy number (Figure 1, Table S3). In manured soil, the highest concentration was observed for iAs (0.90 μg/kg, mean of manured planted and unplanted) followed by TMAO (0.90 μg/kg, mean of manured planted and unplanted) and DMA (0.54 μg/kg, mean of manured planted and unplanted), with FYM significantly increasing all As species by 94% or more

289 when compared with non-manured soil. Highest percent increase due to FYM was observed
290 in MMA (from 0.02 mean of non-manure planted and non-manure unplanted) to 0.72 μg/kg
291 (mean of manure planted and manure unplanted). Soil *arsM* relative copy number was not
292 significantly increased by FYM (p=0.143) (Figure 1, Table S3). FYM significantly increased
293 the total As concentration in roots by 54% (from 39.67 mg/kg to 61.36 mg/kg) and in shoots
294 by 60% (from 0.54 mg/kg to 0.87 mg/kg) (Figure S4). Non-manured plants contained 0.82 %
295 of arsenic in the microcosms, and manured 0.54 % of this figure; again, as for experiment 1, a
296 small portion of the soil store.
297
298 The manure*plant interaction effect was only found to be significant for MMA ($p<0.001$),
299 with manure-planted soil showing a significantly higher concentration of MMA compared to
300 manured-unplanted soil ($p <0.05$, post-hoc multiple comparison test, Minitab) (Figure1,
301 Table S3). There was no significant manure*plant interaction observed for iAs, DMA,
302 TMAO and relative copy number of *arsM* and 16S rRNA ($p>0.05$), indicating that the
303 treatments individually influenced these parameters (Figure 1, Table S3).
304
305 Amplicon sequencing of the *arsM* and 16S rRNA
306 Amplicon sequencing of the *arsM* and 16S rRNA genes was used to investigate soil bacterial
307 community composition and diversity. After paired end joining of reads, a total of 6,733,356
308 sequences for the 16S rRNA gene were retained, ranging from 301,26 to 369,822 per sample
309 (Table S4). Rarefaction curves for alpha diversity (diversity within samples) measures
310 (Chao1, Shannon and Simpson) indicated that sequencing depth was sufficient to detect the
311 majority of the 16S rRNA sequence types in all samples (Figure S5A). Significant differences
312 in species richness (Chao1) and diversity (Shannon and Simpson) were observed. Manure
313 addition significantly increased species diversity (Shannon and Simpson, $P<0.001$, Figure

S6A) and species richness (Chao1, P<0.01, Figure S6A). Principle coordinate (PCoA) plots for overall bacterial community diversity between the samples is shown in Figure S7A. The principle component 1 and 2 explained 47.9 % and 15.5 % of the total variation, respectively (Figure S7A) with samples amended with FYM clearly separated from those without FYM and with a less defined level of separation also observed between the unplanted and planted treatments (Figure S7A).

With respect to 16S rRNA analysis in all treatments, Proteobacteria, Acidobacteria, Actinobacteria, Planctomycetes, Chloroflexi, Firmicutes and Bacteriodetes were the dominant phyla constituting >95% of total bacterial community and showing relative abundance >1% in at least 1 sample (Figure S8A). Among these phyla, Proteobacteria was the most abundant phylum in all treatments. Application of manure significantly increased (GLM, p<0.05, Table S5, Figure S8A) the relative abundance of Proteobacteria, Bacteroides and Firmicutes in both planted and unplanted non-manured treatments, while the Acidobacteria and Chloroflexi significantly decreased (GLM, p<0.05, Table S5, Figure S8A). At genus level, there were 17 dominant genera accounting for >1% in relative abundance this including OTUs significantly increased in manured (*Flavisolibacter, Sphaerobacter, Singulisphaera, Geobacter, Desulfocapsa*) and in planted (*Nocardiodes, Singulisphaera, Clostridium)* soil as well as OTUs significantly increased in non-manured (*Caldilinea, Gemmatimona, Defluviicoccus, Thiobacillus, Syntrophobacter, Steroidobacter*) and unplanted *(Caldilinea, Gemmatimona, Thiobacillus, Syntrophobacter, Steriodibacter)* soil (Table S6, Figure S8B). However, with respect to abundant genera, only for *Clostridium*, *Geobacter* and *Desulfocapsa* were the foldchanges shown to be large enough for these to be deemed as significant changes in copy number in response to manure treatment, and/or for

338 planting during DESeq2 analysis, as the significance threshold for Deseq2 was FDR<0.05
339 and log2FC>1 (Table1, Figure 2A, Table S6).
340
341 For all genus level 16S rRNA OTUs with at least 5 counts in at least 3 samples, the unique
342 and overlapping DESeq2 results for pairwise comparisons are illustrated in the Venn diagram
343 (Figure S9A). Manure effects (130 in total, Group III and Group IV, Figure 2A) showed 115
344 significant OTUs in planted, 83 significant OTUs in unplanted, with 68 overlapping OTUs.
345 Compared to this, plant effects (42 in total) showed 1 significant OTU in non-manured soil
346 and 41 in manured soil with 1 overlapping OTU. With respect to the 42 significant genus
347 level 16S rRNA OTUs identified for plant effect, 35 of these also showed a significant
348 manure effect (Figure S8A). Hence, only 7 genera level 16S rRNA OTUs responded solely to
349 planting (Group I and Group II, Figure 2A). DESeq2 analysis of genus level 16S rRNA
350 OTUs, furthermore, showed that the manure treatment resulted in significant >2-fold increase
351 (FDR<0.05) in copy number of 117 OTUs of specific genera within the phylum
352 Acidobacteria, Actinobacteria, Bacteroidetes, Cyanobacteria, Firmicutes, Proteobacteria and
353 Verrumicrobia (Group III in Figure 2A and Table 1); and 2-fold decrease (FDR<0.05) in 14
354 OTUs of another set of genera within the Phylum Acidobacteria, Bacteroidetes,
355 Cyanobacteria, Firmicutes, Proteobacteria (Group IV in Figure 2A and Table1). Of these only
356 *Desulfocapsa, and Geobacter* from Group III were also within the list of most abundant
357 (>1%) genera (Table1, Table S6). Planting resulted in significant >2-fold increase
358 (FDR<0.05) in copy number of 3 OTUs of genera (*Tumebacillus*, *Clostridium*,
359 *Desulfotomaculum*) within the phylum Firmicutes (Group II in Figure 3A and Table1); and a
360 decrease in 4 OTUs of genera (*Acidobacteria_GP1*, *Dehalogenimonas*, *Cryptomonadaceae*
361 and *Alterococcus*) within the phylum Acidobacteria, Chloroflexi, Cyanobacteria,
362 Verrumicrobia (Group 1 in Figure 2A and Table1). Of all these, only *Clostridium* (Group II)

was also within the list of most abundant (>1%) genera (Table1, Table S6). Hence 16S rRNA copy number changes of less abundant genera appear to contribute significantly to the overall effect of manure treatment and planting in this study.

The diversity of the *arsM* gene was also assessed using amplicon sequencing and after paired end joining of reads, a total of 5,690,027 sequences for the *arsM* rRNA gene were retained, ranging from 210,483 to 609,670 per sample and a median of 274,036 (Table S4). Rarefaction curves for alpha diversity (diversity within samples) measures (Chao1, Shannon and Simpson) indicated that sequencing depth was sufficient to detect the majority of the sequence types in all samples (Figure S5B). Both FYM and planting significantly decreased the *arsM* diversity (Shannon and Simpson) (Figure S6B). A significant manure*plant interaction effect was observed for *arsM* species richness (Chao1) indicating that planting decreased the overall number of *arsM* OTUs, in particular in non-manured soil (Figure S6B). PCoA plots of the overall diversity of *arsM* in different treatments (Figure S6B), shows a clear separation due to manuring and planting, very similar to that observed for 16S rRNA diversity. Annotation (blastn against bacterial genomes) of the 2680 representative *arsM* OTUs that were retained in the filtered OTU table, returned matches for 2104 OTUs against 97 different bacterial genera. Most *arsM* OTUs showed highest sequence similarity to the genus *Rhodopseudomonas* (25%), followed by *Roseiarcus* (6%), *Gemmatirosa* (5%), *Streptomyces* (4%), *Thiobacillus* (3%) and *Rhodoplanes* (2%); with the rest of the genera matching less that 2% of these OTUs (Table 2). The most numerous were matches of *arsM* OTUs to genomes of the species *Rhodopseudomonas palustris* (638 OTUs) with best matches obtained to 9 different strains within this species: *TIE-1* (212 OTUs), *BisB5* (180 OTUs), *YSC3* (94 OTUs), *XCP* (69 OTUs), *HaA2* (41 OTUs), *CGA009* (19 OTUs), *DX-1* (8 OTUs), *BAL398*, (6 OTUs), *R1* (9 OTUs). With regards to the genus *Rhodopseudomonas*, there were

also 33 OTUs with best matches to *R. pentothenatexigens strain JA575*. Of the 165 *arsM* OTUs with best match against *Roseiarcus ferementans,* 124 returned *R. palustris TIE-1* as second-best match (blastn against bacterial genomes) and 120 as best match when using blastn against the fungene_*arsM* database[33].

For all significant *arsM* OTUs, the unique and overlapping DESeq2 results for pairwise comparisons are illustrated in the Venn diagram (Figure S9B). Manure effects (319 in total) showed 258 significant OTUs in planted soil and 222 significant OTUs in unplanted soil, with 162 overlapping OTUs. Compared to this, plant effects (73 in total) showed 46 significant OTUs in non-manured soil and 34 in manured soil with 7 overlapping OTUs. With respect to the 73 significant OTUs identified for the plant effect, 12 of these also showed a significant manure effect (Figure S9B).

Corresponding *arsM* and genus level 16S rRNA OTUs for 4 identified significant copy number patterns were identified. These are visualized in a heatmap in Figure 2A (genus level 16S rRNA) and Figure 2B (*arsM* OTUs). For *arsM* OTUs 29 were assigned to Group I, 32 to Group II, 256 to Group III and 57 to Group IV. For genus level 16S rRNA, OTUs belonging to Group I, II, III and IV are summarized in Table 1. While overall species richness (number) of *arsM* OTUs was decreased in response to planting and manuring, the largest number of *arsM* OTUs with significantly changed copy number were observed in Group III (Figure 2B). With respect to annotation of significant *arsM* OTUs, blastn against bacterial genomes showed the following: Most Group 1 *arsM* OTUs showed closest sequence similarity to *R. palustris BisB5* (41%) followed by *R. palustris YSC3* (10%). Group 2 *arsM* OTUs showed closest sequence similarity to *Streptomyces aurantiacus* (12%) followed by *R. palustris TIE-1* (9%). Group 3 *arsM* OTUs showed closest sequence similarity to *R. palustris TIE-1* (29%)

and *Roseiarcus fermentans* (46%) (88% of the latter returned as second-best genome match also *R. palustris TIE-1*), followed by *Thioalbus denitrificans* (3%). Group 4 *arsM* OTUs showed closest sequence similarity to *R. palustris TIE-1* (16%) and *Gemmatirosa kalamazoonesis* (11%) followed by *R. palustris BisB5* (7%). Hence with respect to predominant *arsM* copy number changes, manuring resulted mainly in increase of *R. fermentans* (blast against the fungene-*arsM* database[33] returns for 85% of these *R. palustris TIE-1*) and *R. palustris TIE-1*, while planting resulted in decrease of *R. palustris BisB5* and increase in *S. aurantiacus* annotated OTUs. For further details see Table 2 & Supporting information File 2.

Variance partition analysis for Group I and II 16S rRNA OTUs verified that 60% of the variation in Acidobacteria_GP1 copy number was due to planting, while for the other 6 OTUs in Group I and II this was around 20-35% (Figure 3). For Group I and II *arsM* OTUs, this analysis verified that 45% of the variation in copy number was due to planting (Figure S10). Variance partition analysis for Group III and IV 16S rRNA OTUs showed that 28% and for Group III and IV *arsM* OTUs that 25% of the variation in copy number was due the manure effect (Figure S11 and Figure S12).

With regards to significant correlation (Spearman's rank correlation, $P<0.05$) of copy number patterns between genus level 16S rRNA and *arsM* OTUs the following results were obtained. Of the 29 *arsM* OTUs in Group I, 100% correlated with at least 1 Group I genus level 16S rRNA OTU. Of these 97% correlated most strongly with *Acidobacteria_GP1* and 3% with *Dehaligenimonas* (Table S7). Of the 32 *arsM* OTUs in Group II, 65% correlated with at least 1 Group II genus level 16S rRNA OTU. Of these 34% correlated most strongly with *Clostridium* and 31% with *Tumebacillus* (Table S8). All Group III *arsM* OTUs correlated

438 highly significantly with multiple Group III genus level 16S rRNA OTUs. Of these, highest
439 number of *arsM* OTUs (9%) correlated most significantly with *Luteolibacter* (phylum
440 Verrucomicrobia), followed by *Devosia* (7%), and *Herbaspirillum* (6%) (phylum
441 Proteobacteria). The rest of the Group III *arsM* OTUs correlated most significantly to a
442 highly diverse range of genus level 16S rRNA OTUs (Table S9) within the phyla
443 Proteobacteria (25 genera), Firmicutes (21 genera, this including *Pseudomonas*),
444 Bacteroidetes (5 genera), Acidobacteria (*Geothrix*) and Actinobacteria (*Thermobifida*). All
445 Group IV *arsM* OTUs correlated with at least one Group IV genus level 16S rRNA OTU. Of
446 these 23% correlated most significantly with *Acidobacteria_Gp25*, 14% with *Terrimonas,*
447 12% with *Acidobacteria_GP22,* 11% with *Streptophyta* and 11% with *Tepidibacter* (Table
448 S10).
449
450 Network analysis plots for genus level 16S rRNA OTUs that responded to planting (Group I
451 and II) are shown in Figure 4A-G. The network for *Acidobacteria_GP1* (Figure 4A) shows
452 the strongest connection to all unplanted samples (Group I), due to higher copy number in
453 unplanted soil and was also most highly correlated to Group I *arsM* OTUs (Table S7). In a
454 similar manner *Tumebacillus* and *Clostridium* (Figure 4E and 4F), show the strongest
455 connections to all planted samples (Group II), due to higher copy number in planted soil and
456 these were also most highly correlated to Group II *arsM* OTUs (Table S8). Network analysis
457 plots for genus level 16S rRNA OTUs that responded to manuring are shown in Figure S13.
458 Individual networks are shown for those Group IV genus level 16S rRNA OTUs (Figure
459 S13A-E) that showed highest levels of correlation with Group IV *arsM* OTUs (Table S10),
460 with *Acidobacteria_GP25* (Figure S13A), *Terrimonas* (Figure S13B), *Acidobacteria_GP22*
461 (Figure S13C), *Streptophyta* (Figure S13D), *Tepidobacter* (Figure S13E) all showing
462 strongest connections to non-manured samples, due to higher copy number in non-manured

soil, but with this effect clearly most pronounced for *Acidobacteria_GP22* (Figure S13A). *Acidobacteria_GP22* were also shown to be the most highly correlated to Group IV *arsM* OTUs (Table S10). For Group III genus level 16S rRNA OTUs the overall network shows a clear separation of manure and non-manure samples (Figure S13F), the same applies to networks for Group III and Group IV *arsM* OTUs (Figure 14A and 14B). Contrary to that, network analysis plots for *arsM* OTUs that responded to planting (Group I and Group II) show clear separation of planted versus unplanted samples (Figure S15A and S15B). Hence these results provide further evidence of treatment effects and present potential targets for the study of *arsM* and microbial arsenic methylation efficiency in paddy soil.

DISCUSSION

Microbial driven methylation, and potential subsequent volatilization, mediated by *arsM* plays a crucial role in biogeochemical cycling of arsenic.[36] Previous studies on *arsM* diversity in natural rice paddy microbial communities have used low throughput cloning and Sanger sequencing,[20,23,36] or have limited Illumina sequencing of *arsM* amplicons to the study of enrichment cultures from paddy soil[37], or metagenomics analysis to characterize arsenic metabolizing genes in 5 soils[38]. Our study provides a step-change as it employs arsenic speciation characterization combined with qPCR and high throughput Illumina amplicon sequencing (16S rRNA and *arsM* gene) to study arsenic mobilization and transformation in response to planting and manuring, alongside detailed analysis of the diversity of complex natural bacterial arsenic transforming communities in rice paddy soil.

Phylogenetic annotation based on the protein coding functional gene *arsM*, may provide additional information as well as a higher taxonomic resolution compared to the well-established 16S rRNA gene-based analysis, but is limited by the number of curated *arsM*

protein sequences in the database[33]. Correlated 16S rRNA and *arsM* gene copy numbers provide complementary information to highlight organisms that may promote arsenic methylation in paddy soil. The *arsM* based amplicon sequencing analysis revealed a great level of *arsM* sequence diversity with highest number of top matches against phototrophic purple non-sulfur bacteria of the genus *Rhodopseudomonas*. These were dominated by strain *R. palustris TIE-1* with smaller numbers of *arsM* OTUs showing matches to *R. palustris* strains *BisB5*, *YSC3*, *XCP*, *HaA2*, *CGA009* and *DX1*. High frequency of *R. palustris arsM* sequences in rice paddy soil has previously been reported[38]. The *R. palustris arsM* gene has been proven to methylate As(III) to the relatively nontoxic pentavalent species DMA(V) and TMAO and confer arsenic resistance to an arsenic sensitive *Escherichia coli* strain[10]. Further to that, members of this species have been proposed, based on their As detoxification mechanisms, for bioremediation of arsenic contaminated areas surrounding mines as have been shown to be able to oxidise As(III) to As (V)[39]. Members of this species can grow with or without light or oxygen, fix nitrogen and degrade a wide range of organic compounds, with different strains shown to harbour strain specific genes and exhibit distinctive physiological characteristics with respect to anaerobic fermentation, expanded biodegradation, or expanded light-harvesting capabilities[40,41] leading to the proposal that they should be reclassified into different species[42]. Hence, differences in copy number of these strains due to plant and manure induced changes in redox potential, pH and available carbon sources is of considerable interest in the context of arsenic in rice grain.

FYM increased all present arsenic species including methylated species in paddy soil solution and with respect to *arsM* gene OTUs led to significant increase in copy number of large numbers of *R. fermentans* and *R. palustris TIE-1* annotated OTUs, with *R. fermentans* reported to grow best under micro-oxic conditions by means of fermentation of sugars and

organic acids[43] and *R. palustris TIE-1* shown to be able to grow on a wide variety of carbon sources as well as perform phototrophic Fe(II) oxidation[44,45,46]. It is important to note in this context that *R. fermentans* annotated *arsM* OTU sequences (genome match), also showed high similarity to *R. plaustris TIE-1*. Further to that, 16S rRNA amplicon sequencing showed manure induced increase in *Geobacter*, an anaerobe shown to perform anaerobic oxidation of aromatic hydrocarbons to carbon dioxide via reduction of Fe(III)[47,48] as well as thought to be capable of nitrate dependent Fe(II) oxidation[49]. Hence, manuring appears to stimulate microbes involved in arsenic methylation as well as in both iron plaque formation (Fe(II) oxidation) and degradation (Fe(III) reduction), leading to binding and release of arsenic from Fe-plaque around roots.[22,50]

Higher redox potential of rhizosphere due to oxygenation from rice roots led to suppression of all present arsenic species, including methylated species. This corresponded to decreased *arsM* copy number (qPCR), species richness (amplicon sequencing, Chao1) and diversity (amplicon sequencing, Simpson, Shannon). This implies that in our study, the number of *arsM* organisms decreased in planted soil. With respect to *Rhodopseudomonas* annotated *arsM* OTUs, some with highest sequence similarity to *R. palustris BisB5* were suppressed by planting. *R. palustris* strain *BisB5* was shown to contain a gene cluster for anaerobic phenylacetate degradation and was proposed to be best adapted to degradation of plant material in oxygen depleted environments that are exposed to light[40]. In contrast to this *arsM* OTUs with closest match against *S. aurantiacus* were increased in planted soil. *Streptomyces* are generally aerobic bacteria and isolates from rice paddies have been shown to contribute to better growth of rice plants as well as increased levels of arsenic methylation and uptake of DMA into rice shoots[51]. Change in redox potential may therefore play a role in the observed copy number changes of these organisms in response to planting. In our study, the fact that

538 rhizosphere soils have lower methylated arsenic species than bulk soil coupled with decrease
539 and shift in *arsM* organisms could be attributable to oxygenation of the rhizosphere by the
540 root, as radial oxygen loss (ROL), regulates rice rhizosphere Eh,[52] and subsequent
541 assimilation of arsenic[18]. However, total arsenic species, inorganic as well as methylated,
542 decreased in the rhizosphere suggesting that redox, while part of the answer, does not fully
543 explain the reduction in methylated species.
544
545 The root itself is a sink for all arsenic species, but rates of MMA and DMA accumulation by
546 rice roots are low with much higher rates of arsenate and arsenite assimilation.[11] Thus, if the
547 root was acting as a major sink, to the extent that they affect soil solution concentrations, then
548 it may be expected that iAs species are preferentially depleted. This was not the case here.
549 Rather than the root tissue, *per se*, being the sink for arsenic species that leads to the lowering
550 of these species in porewater, it may be associated iron plaque in the rhizosphere. Iron plaque
551 preferentially binds arsenate to arsenite, but has a high capacity to sorb both species, and
552 methylated species.[17] It is interesting that our *arsM* amplicon data indicates increase in
553 bacteria with arsenic methylation and Fe(II) oxidation and Fe(III) reduction capacity in
554 manured soil, as this could lead to formation of iron plaques and subsequent release of
555 arsenic from iron plaque.[44-50,18] As well as the root potentially creating new sinks for arsenic
556 species, iron plaque and root tissue, the enhanced oxygenation of the planted microcosms
557 here may have led to lower mobilization of iAs from sediment stores as it is reducing
558 conditions that drive this mobilization.[19,22] Lower mobilization of iAs would mean lower
559 substrate for methylation and, hence, lower production of methylated arsenic species.
560
561 A number of studies have considered arsenic species mobility in the rhizosphere.[20,21] The
562 current study differs from these in that it compares planted versus non-planted microcosms,

563 while the other studies sampled a rhizosphere continuum from bulk soil to root surface. The
564 studies of Jia et al.[20,21] did not show the inhibition of in culture medium solution of arsenic
565 species concentration in the rhizosphere. In essence these studies have a relatively small and
566 constrained rhizosphere, to enable spatial sampling, with the root-soil interface surrounded by
567 a large volume of bulk soil to resupply arsenic species if they become depleted in the
568 rhizosphere. In the field, particularly in densely planted agronomic systems, the rhizosphere
569 can be considered as a continuum, with dense and overlapping roots.[53] In our experiments the
570 soil:fresh root biomass ratio was considerable, being 150:1 at week 2 when Eh measurements
571 were made. Even at these low biomass ratios the presence of the root had a profound effect
572 on both Eh and arsenic speciation.
573
574 Here, the detailed reporting of arsenic species in both roots and shoots over a time course
575 illustrates, again, that while iAs concentrations *in planta* are higher than methylated species,
576 it is the methylated species that are more readily translocated to the shoot.[12] The root
577 measurements do not distinguish between surface plaque bound and species actually
578 assimilated into plant tissues and this must be born in mind when interpreting the data as
579 plaque in roots can have high arsenic concentrations.[6] Here, the shoot measurements are free
580 from such qualifications and only represent arsenic species assimilated by root and then
581 translocated to shoot. Hydroponic experiments have shown that although iAs species are
582 more efficiently assimilated by the root,[11] their translocation to shoot are retarded as
583 compared to methylated species, with MMA less effectively translocated to the shoot from
584 roots than DMA.[12] Arsenate is reduced to arsenite in plant cells and then phytochelatin (PC)
585 complexed, with arsenite-PC complexes then transported into vacuoles via ABC
586 transporters.[6] MMA can be thiol complexed, though less efficiently than arsenite, while
587 DMA has lower affinity for thiol complexation,[12,54] with the affinity to PCs explaining their

588 relative efficiency in translocation to the shoot.[12] This pattern can also be seen here, with the
589 addition of TMAO which shows equivalency between root and shoot in concentration,
590 indicating further that increasing methylation leads to enhanced translocation to the shoot.
591 Although not tested, to date, it is likely that TMAO has poor affinity to plant thiols.
592
593 Most studies use anion exchange HPLC-ICP-MS to speciate arsenic, where TMAO elutes on
594 the solvent front.[13] As we used the more discriminatory IC-ICP-MS, TMAO was readily
595 detected and quantified. Even though iAs's movement is retarded through the plant, because
596 initial starting concentrations in soil solution and root tissues are much higher, shoot
597 concentrations were *circa.* 100-fold higher than for MMA, DMA and TMAO. However, in
598 rice grain only iAs and DMA are found in significant quantities, with DMA in many regions
599 of the world reaching equivalency with iAs or exceeding in cases.[1] Traces of MMA can be
600 found routinely, and TETRA in very elevated grain samples, but interestingly not TMAO.[13]
601 The differences in grain unloading of DMA and iAs have been studied for rice,[55] but MMA,
602 TMAO and TETRA not. It is clear that we still have a way to go in understanding the
603 physiology of all arsenic species found in rice with respect to uptake and translocation.
604
605 As in previous studies in paddy soil, [20,23,36-38,56] high diversity of bacterial 16S rRNA and
606 *arsM* were observed, and shown to be affected by soil amendments. Recent studies on *arsM*
607 have identified high levels of diversity with Proteobacteria, Gemmatmonadales and
608 Firmicutes[36] as the most common bacterial communities in paddy soil, this including high
609 frequency of *R. palustris*[38] and many phylogenetically divergent microbes have been shown
610 to harbor the *arsM* gene and to be able to convert aqueous iAs to methylated species[37]. Pure
611 culture studies have, furthermore, shown that microbes vary with respect to their methylation
612 efficiency with *Methanosarcina acetivorans* C2A[57] (phylum Euryarchaeota), *Clostridium* sp.

BXM[58] (phylum Firmicutes) shown to transform 10%, while *Streptomyces* sp.[59] (phylum Actinomycete), *Pseudomonas alcaligenes*[60] (phylum Proteobacteria) and *Arsenicibacter rosenii*[61] (phylum Bacteriodetes) shown to transform more than 50% of aqueous iAs to methylated species. In addition our observed correlated *arsM* and 16S rRNA copy numbers are worth further investigation. The copy number of many Group II *arsM* OTUs was for example correlated to 16S rRNA *Clostridium* OTUs, which is interesting as co-culture of *R. palustris* and *Clostridium* has previously been shown to lead to increased hydrogen production and *R. palustris* genes involved in organic compound catabolism and nitrogen fixation[62]. The copy number of Group III *arsM* OTUs on the other hand showed correlation to a very diverse range of genus level 16S rRNA OTUs, with highest matches to *Luteolibacter* and *Devosia. Luteolibacter* have previously reported as one of the 5 most abundant genera in arsenic contaminated soil[63] and *Devosia* were shown to exhibit high sequence similarity to arsenic tolerant organisms isolated from arsenic rich environments[64]. Therefore, both the genomic matches of *arsM* OTUs and the correlated copy number patterns in Group I, II, III and IV *arsM* and genera level 16S rRNA OTUs present potential targets for identification of *arsM* organisms and the study of arsenic transformation in paddy soil.

The relative *arsM* copy number is known to be variable between soils,[20,23,36-38,56] and we have shown this again in the current findings, with highly significant differences also observed between planted and unplanted microcosms. Similarly, arsenic speciation was profoundly influenced by soil origin, planting and manuring. Inherent variation in DMA in soil solution has been shown in other studies,[15] and the results shown here illustrate that this chemical is spatially and temporally variable in soils, interacting with environmental factors. Environmental factors such as soil solution soluble OM was found to correlate well with DMA in soil solution.[15] As plants obtain all their DMA from soil solution,[9] DMA variability

638 in soil solution will be reflected in root uptake, and ultimately in grain DMA concentrations.[1]
639 As DMA is less toxic than iAs to humans, altering grain arsenic through paddy soil
640 management can lead to higher DMA to iAs, but some caution is required before promoting
641 soils microbes to convert iAs to DMA. For example, SOM, while promoting arsenic
642 methylation in paddy soils also liberates more arsenic in general leading to enhanced plant
643 uptake to both inorganic and organic species.[14] Also, while less toxic to humans, DMA is
644 strongly implicated in "straight-head" disease of rice where yield is greatly lowered by DMA
645 (and MMA) exposure due to introducing sterility.[23,65] It is clear if rice is to be managed to
646 lower grain iAs content then caution needs to be heeded to ensure that adverse consequences
647 are negated. Understanding that the rhizosphere environment greatly alters arsenic
648 concentrations and speciation is a step forward in developing strategies for minimizing grain
649 iAs content without impacting yield. Further studies need to be conducted in a wide range of
650 paddy soils to validate the role of rhizosphere interactions, including microbial mediated
651 arsenic methylation as well as Fe(II) oxidation and Fe(III) reduction and its effect on arsenic
652 bound to iron-plaque, and how these can be manipulated to minimize iAs in rice grain.
653

654 **ACKNOWLEDGEMENTS**

655 We thank the Commonwealth Scholarships for funding Hasina Afroz' Ph.D. studentship.
656 Illumina Sequencing was performed by the Centre Of Genomics Research, Institute of
657 Integrative Biology University of Liverpool, Crown Street Liverpool L69 7ZB, UK.
658

659 **SUPPORTING INFORMATION.**

660 SupportingInformation1.docx (SI1.docx): additional information on materials and methods,
661 additional results in form of 10 tables and 15 figures, relevant references.
662 SupportingInformation2.xlsx (SI2.xlsx): table with information on significant Group1,

663 Group2, Group3 and Group4 *arsM* OTUs including, sequence, blast results and most highly
664 correlating 16S rRNA OTU.
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687

688 **REFERENCES**

689 (1) Meharg, A. A.; Williams, P. N.; Adomako, E.; Lawgali, Y. Y.; Deacon, C.; Villada, A.;
690 Cambell, R. C. J.; Sun, G.; Zhu, Y. G.; Feldmann, J.; Raab, A.; Zhao, F. J.; Islam, R.;
691 Hossain, S.; Yanai, J. Geographical variation in total and inorganic arsenic content of
692 polished (white) rice. *Environ. Sci. Technol*. **2009**, *43* (5), 1612−1617.
693
694 (2) EU Commission Regulation 2015/1006 of 25 June 2015 amending Regulation (EC) No
695 1881/2006 as regards maximum levels of inorganic arsenic in foodstuffs. **2015**. http://eur-
696 lex.europa.eu/legal-content/EN/TXT/?uri=uriserv%3AOJ.L_.2015.161.01.0014.01.ENG.
697
698 (3) WHO, 2014. Codex alimentarius commission. Project document proposal for new work
699 on a maximum level for arsenic in rice. **2014**
700 ftp://ftp.fao.org/codex/meetings/CCCF/cccf5/CRDs/CRD%2024%20ARSENIC%20IN%20R
701 ICE.pdf
702
703 (4) FDA, **2016**. FDA proposes limit for inorganic arsenic in infant rice cereal.
704 http://www.fda.gov/NewsEvents/Newsroom/PressAnnouncements/ucm493740.htm
705
706 (5) Signes-Pastor, A. J.; Carey, M.; Meharg, A. A. Inorganic arsenic in rice-based products
707 for infants and young children. *Food Chem*. **2016a**, *191*, 128−134.
708
709 (6) Meharg, A. A.; Zhao, F. J. *Arsenic & Rice*; Springer, The Netherlands, **2012**.
710
711 (7) Meharg, C.; Meharg, A. A. Silicon, the silver-bullet for mitigating biotic and abiotic
712 stress, and improving grain quality, in rice. *Environ. Exp. Bot*. **2015**, *120*, 8−17.

8. Williams, P. N.; Villada, A.; Deacon, C.; Raab, A.; Figuerola, J.; Green, A. J.; Feldmann, J.; Meharg, A. A. Greatly enhanced arsenic shoot assimilation in rice leads to elevated grain levels compared to wheat and barley. *Environ. Sci. Technol.* **2007**, *41* (19), 6854−6859.

(9) Lomax, C.; Liu, W. J.; Wu, L.; Xue, K.; Xiong, J.; Zhou, J.; McGrath, S. P.; Meharg, A. A.; Miller, A. J.; Zhao, F. J. Methylated arsenic species in plants originate from soil microorganisms. *New. Phytol*. **2012**, *193* (3), 665−672.

(10) Qin, J.; Rosen, B. P.; Zhang, Y.; Wang, G. J.; Franke, S.; Rensing, C. Arsenic detoxification and evolution of trimethylarsine gas by a microbial arsenite S-adenosylmethionine methyltransferase. *Proc. Natl. Acad. Sci.*, U. S. A. **2006**, *103* (7), 2075−2080.

(11) Abedin, M. J.; Cresser, M. S.; Meharg, A. A.; Feldmann, J.; Cotter-Howells, J. Arsenic accumulation and metabolism in rice (Oryza sativa L.). *Environ. Sci. Technol*. **2002**, *36* (5), 962−968.

(12) Raab, A.; Williams, P. N.; Meharg, A.; Feldmann, J. Uptake and translocation of inorganic and methylated arsenic species by plants. *Environ. Chem*. **2007**, *4* (3), 197−203.

(13) Hansen, H. R.; Raab, A.; Price, A. H.; Duan, G. L.; Zhu, Y. G.; Norton, G. J.; Feldmann, J.; Meharg, A. A. Identification of tetramethylarsonium in rice grains with elevated arsenic content. *J. Environ. Monit*. **2011**, 13 (1), 32−34.

738 (14) Norton, G. J.; Adamoko, E.; Deacon, C.; Carey, A. M.; Price, A. H.; Meharg, A. A..
739 Effect of organic matter amendment, arsenic amendment and water management regime on
740 rice grain arsenic species. *Environ. Poll*. **2013**, *177*, 38−47.
741
742 (15) Williams, P. N.; Zhang, H.; Davison, W.; Meharg, A. A.; Hossain, M.; Norton, G. J.;
743 Brammer, H.; Islam, M. R. Organic Matter-Solid Phase Interactions Are Critical for
744 Predicting Arsenic Release and Plant Uptake in Bangladesh Paddy Soils. *Environ. Sci.*
745 *Technol*. **2011**, *45* (14), 6080−6087.
746
747 (16) Liu, W. J.; Zhu, Y. G.; Smith, F. A.; Smith, S.E. Do iron plaque and genotypes affect
748 arsenate uptake and translocation by rice seedlings (*Oryza sativa* L.) grown in solution
749 culture? *J. Exp. Bot*. **2004**, *55* (403), 1707−1713.
750
751 (17) Liu, W. J.; Zhu, Y. G.; Hu, Y.; Williams, P. N.; Gault, A. G.; Meharg, A. A.; Charnock,
752 J. M.; Smith, F. A. Arsenic sequestration in iron plaque, its accumulation and speciation in
753 mature rice plants (*Oryza sativa* L.). *Environ. Sci. Technol.* **2006**, *40* (18), 5730−5736.
754
755 (18) Mei, X. Q.; Wong, M. H.; Yang, Y.; Dong, J. Y.; Qiu, R. L.; Ye, Z. H. The effects of
756 radial oxygen loss on arsenic tolerance and uptake in rice and on its rhizosphere. *Environ.*
757 *Poll*. **2012,** 165, 109−117.
758
759 (19) Bravin, M. N.; Travassac, F.; Le Floch, M.; Hinsinger, P.; Garnier, J.M. Oxygen input
760 controls the spatial and temporal dynamics of arsenic at the surface of a flooded paddy soil
761 and in the rhizosphere of lowland rice (*Oryza sativa* L.): a microcosm study. *Plant Soil* **2008,**
762 *312* (1-2), 207−218.

(20) Jia, Y.; Huang, H.; Zhong, M.; Wang, F. H.; Zhang, L. M.; Zhu, Y. G. Microbial arsenic methylation in soil and rice rhizosphere. *Environ. Sci. Technol*. **2013**, *47* (7), 3141−3148.

(21) Jia, Y.; Huang, H.; Chen, Z.; Zhu, Y. G. Arsenic uptake by rice is influenced by microbe-mediated arsenic redox changes in the rhizosphere. *Environ. Sci. Technol*. **2014**, *48* (2), 1001−1007.

(22) Williams, P. N.; Santner, J.; Larsen, M.; Lehto, N. J.; Oburger, E.; Wenzel, W.; Glud, R. N.; Davison, W.; Zhang, H. Localized Flux Maxima of Arsenic, Lead, and Iron around Root Apices in Flooded Lowland Rice. *Environ. Sci. Technol.* **2014,** *48* (15), 8498−8506.

(23) Zhao, F. J.; Harris, E.; Yan, J.; Ma, J.; Wu, L.; Liu, W. J.; McGrath, S. P.; Zhou, J. Z.; Zhu, Y. G. Arsenic methylation in soils and its relationship with microbial *arsM* abundance and diversity, and As speciation in rice. *Environ. Sci. Technol*. **2013**, *47* (13), 7147−7154.

(24) Caporaso, J. G.; Lauber, C. L.; Walters, W. A.; Berg-Lyons, D.; Huntley, J.; Fierer, N.; Owens, S.M.; Betley, J.; Fraser, L.; Bauer, M.; Gormley, N.; Gilbert, J. A.; Smith, G.; Knight, R. Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *Isme Journal* **2012,** *6* (8), 1621−1624.

(25) Degnan, P. H.; Ochman, H. Illumina-based analysis of microbial community diversity. *Isme Journal* **2012,** *6* (1), 183−194.

(26) Signes-Pastor, A.J.; Carey, M.; Carbonell-Barrachina, A.A.; Moreno-Jimenez, E.; Green,

A.J.; Meharg, A.A. Geographical variation in inorganic arsenic in paddy soils and rice from the Iberian Peninsula. *Food Chem*. **2016b**, *202*, 356−363.

(27) Savage, L.S. Carey, M. Hossain, M.; Islam, M.R.; de Silva P.MC.S.; Williams, P.N.; Meharg, A.A. Elevated trimethylarsine oxide and inorganic arsenic in northern hemisphere summer monsoonal wet deposition. *Environ. Sci. Technol.* **2017**, 51, 12210-12218 .

(28) Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet. J. **2011**, *17*, 10−12.

(29) Joshi, N. A.; Fass, J. N. Sickle: A sliding-window, adaptive, quality-based trimming tool for FastQ files (Version 1.33). **2011**. Available at https://github.com/najoshi/sickle.

(30) Andrews, S. FastQC: a quality control tool for high throughput sequence data. **2010**. Available online at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc

(31) Aronesty, E. ea-utils: "Command-line tools for processing biological sequencing data. **2011**. Available online at: http://code.google.com/p/ea-utils

(32) Caporaso, J. G.; Kuczynski, J.; Stombaugh, J.; Bittinger, K.; Bushman, F.D.; Costello, E. K.; Fierer, N.; Pena, A.G.; Goodrich, J. K.; Gordon, J. I.; Huttley, G. A.; Kelley, S. T.; Knights, D.; Koenig, J. E.; Ley, R. E.; Lozupone, C. A.; McDonald, D.; Muegge, B. D.; Pirrung, M.; Reeder, J.; Sevinsky, J. R.; Tumbaugh, P. J.; Walters, W. A.; Widmann, J.; Yatsunenko, T.; Zaneveld, J.; Knight, R. QIIME allows analysis of high-throughput community sequencing data. *Nature Methods* **2010,** *7* (5), 335−336.

(33) Fish, J. A.; Chai B.; Wang, Q.; Sun, Y.; Brown, C.T.; Tiedje, J.M.; Cole, J.R. FunGene: the Functional Gene Pipeline and Repository. *Front. Microbiol.* 2013, *4*:291.

(34) Hoffman, GE, Schadt, EE. variancePartition: interpreting drivers of variation in complex gene expression studies. BMC Bioinformatics. **2016**; 17(1):483

(35) Shannon, P, Markiel, A, Ozier, O, Baliga, NS, Wang, JT, Ramage, D, Amin, N, Schwikowski, B, Ideker, T. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. **2003**, 13(11):2498-504.

(36) Zhang, S.-Y.; Zhao, F.-J.; Sun, G.-X.; Su, J.-Q.; Yang, X.-R.; Li, H.; Zhu, Y.-G. Diversity and abundance of arsenic biotransformation genes in paddy soils from southern China. *Environ. Sci. Technol*. **2015**, *49* (7), 4138−4146.

(37) Reid, M. C.; Maillard, J.; Bagnoud, A.; Falquet, L.; Phu Le Vo; Bernier-Latmani, R. Arsenic Methylation Dynamics in a Rice Paddy Soil Anaerobic Enrichment Culture. *Environ. Sci. Technol.* **2017,** *51* (18), 10546−10554.

(38) Xiao, K.-Q.; Li, L.-G.; Ma, L.-P.; Zhang, S.-Y.; Bao, P.; Zhang, T.; Zhu, Y.-G. Metagenomic analysis revealed highly diverse microbial arsenic metabolism genes in paddy soils with low-arsenic contents. *Environ. Pollut*. **2016**, 211, 1−8.

(39) Nookongbut, P.; Kantachote, D.; Megharaj, M. Arsenic contamination in areas surrounding minesand selection of potential As-resistant purple nonsulfur bacteria for use in bioremediation based on their detoxification mechanism. *Ann Microbiol*. **2016**, *66*, 1419.

(40) Oda, Y.; Larimer, F.W.; Chain, P.S.; Malfatti, S.; Shin, M.V.; Vergez, L.M.; Hauser, L.; Land, M.L.; Braatsch, S.; Beatty, JT.; Pelletier, D.A.; Schaefer, A.L.; Harwood, C.S.; Multiple genome sequences reveal adaptations of a phototrophic bacterium to sediment microenvironments. *Proc Natl Acad Sci U S A*. **2008**. 105(47), 18543-18548.

(41) Lo, K.J.; Lin, S.S.; Lu, CW, Kuo, CH, Liu, CT. Whole-genome sequencing and comparative analysis of two plant-associated strains of *Rhodopseudomonas palustris* (*PS3* and *YSC3*). *Sci Rep.* **2018**. *8*(1),12769.

(42) Rayyan, A.; Meyer, T.; Kyndt, J.; Draft Whole-Genome Sequence of the Purple Photosynthetic Bacterium Rhodopseudomonas palustris XCP. *Microbiol Resour Announc.* **2018**. *7*(4). pii: e00855-18.

(43) Kulichevskaya, IS.; Danilova, O.V.; Tereshina, V.M.; Kevbrin, V.V.; Dedysh, S.N.; Descriptions of *Roseiarcus fermentans* gen. nov., sp. nov., a bacteriochlorophyll a-containing fermentative bacterium related phylogenetically to alphaproteobacterial methanotrophs, and of the family Roseiarcaceae fam. nov. *Int J Syst Evol Microbiol.* **2014.** *64*(Pt 8), 2558-2565.

(44) Jiao, Y.; Kappler, A.; Croal, L.R.; Newman, D.K.; Isolation and characterization of a genetically tractable photoautotrophic Fe(II)-oxidizing bacterium, *Rhodopseudomonas palustris* strain *TIE-1*. *Appl Environ Microbiol*. **2005.** *71*(8), 4487-96.

(45) Jiao, Y.; Newman, D.K.; The pio operon is essential for phototrophic Fe(II) oxidation in *Rhodopseudomonas palustris TIE-1*. *J Bacteriol*. **2007.** *189*(5), 1765-1773.

(46) Hedrich, S.; Schlömann, M.; Johnson, D.B.; The iron-oxidizing proteobacteria. Microbiology. **2011.** *157*(Pt 6), 1551-64.

(47) Lovley, D.R.; Giovannoni, S.J.; White, D.C.; Champine, J.E.; Phillips, E.J.; Gorby, Y.A.; Goodwin, S. *Geobacter metallireducens* gen. nov. sp. nov., a microorganism capable of coupling the complete oxidation of organic compounds to the reduction of iron and other metals. *Arch Microbiol*. **1993**.*159*(4):336-444.

(48) Mahadevan, R.; Bond, D.R.; Butler, J.E.; Esteve-Nuñez, A.; Coppi, M.V.; Palsson, B.O.; Schilling, C.H.; Lovley, D.R.; Characterization of metabolism in the Fe(III)-reducing organism *Geobacter sulfurreducens* by constraint-based modeling. *Appl Environ Microbiol.* **2006**. *72*(2), 1558-68.

(49) Coby, A.J.; Picardal, F.; Shelobolina, E.; Xu, H.; Roden, E.E. Repeated anaerobic microbial redox cycling of iron. *Appl Environ Microbiol*. **2011**. *77*(17), 6036-6042.

(50) Emerson, D.; Weiss, J.V.; Megonigal, J.P.; Iron-oxidizing bacteria are associated with ferric hydroxide precipitates (Fe-plaque) on the roots of wetland plants . *Appl Environ Microbiol.* **1999**. *65*(6), 2758-2761.

(51) Kuramata, M.; Sakakibara, F.; Kataoka, R.; Abe, T.; Asano, M.; Baba, K.; Takagi, K.; Ishikawa, S. Arsenic biotransformation by *Streptomyces* sp. isolated from rice rhizosphere. *Environ. Microbiol*. **2015**, *17* (6), 1897−1909.

(52) Kludze, H.; Delaune, R.; Patrick, W. Aerenchyma Formation and Methane and Oxygen-Exchange in Rice. *Soil Sci. Soc. Am. J.* **1993,** *57* (2), 386−391.

(53) York, L.M.; Carminati, A.; Mooney, S.J.; Ritz, K.; Bennett, M.J. The holistic rhizosphere: integrating zones, processes, and semantics in the soil influenced by roots. *J. Exp. Bot.* **2016,** *67* (12), 3629−3643.

(54) Park, S. G.; Butcher, D. J. Investigation of the primary interactions between arsenic species and thiols via electrospray ionization tandem mass spectrometry. *Microchem. J*. **2010**, *95*, 57−66.

(55) Carey, A. M.; Norton, G. J.; Deacon, C.; Scheckel, K. G.; Lombi, E.; Punshon, T.; Guerinot, M. L.; Lanzirotti, A.; Newville, M.; Choi, Y.; Price, A. H.; Meharg, A. A. Phloem transport of arsenic species from flag leaf to grain during grain filling. *New Phytol*. **2011**, *192*, 87−98.

(56) Gu, Y.; Van Nostrand, J. D.; Wu, L.; He, Z.; Qin, Y.; Zhao, F.-J.; Zhou, J. Bacterial

community and arsenic functional genes diversity in arsenic contaminated soils from different geographic locations. *PLoS One* **2017**, *12* (5), e0176696.

(57) Wang, P.-P.; Sun, G.-X.; Zhu, Y.-G. Identification and characterization of arsenite methyltransferase from an archaeon, *Methanosarcina acetivorans* C2A. *Environ. Sci. Technol*. **2014**, 48 (21), 12706-12713.

(58) Wang, P.-P.; Bao, P.; Sun, G.-X. Identification and catalytic residues of the arsenite methyltransferase from a sulfate-reducing bacterium, *Clostridium* sp. BXM. FEMS *Microbiology Letters* **2015**, *362* (1), 1-8.

(59) Kuramata, M.; Sakakibara, F.; Kataoka, R.; Abe, T.; Asano, M.; Baba, K.; Takagi, K.; Ishikawa, S. Arsenic biotransformation by *Streptomyces* sp. isolated from rice rhizosphere. *Environ Microbiol*. **2015.** *17*(6), 1897-1909.

(60) Zhang, J.; Cao, T.; Tang, Z.; Shen, Q.; Rosen, B. P.; Zhao, F.-J. Arsenic methylation and volatilization by arsenite S-adenosylmethionine methyltransferase in *Pseudomonas alcaligenes* NBRC14159. Appl. *Environ. Microbiol*. **2015**, *81* (8), 2852−2860.

(61) Huang, K.; Chen, C.; Zhang, J.; Tang, Z.; Shen, Q.; Rosen, B. P.; Zhao, F.-J. Efficient Arsenic Methylation and Volatilization Mediated by a Novel Bacterium from an Arsenic-Contaminated Paddy Soil. *Environ. Sci. Technol*. **2016**, *50* (12), 6389−6396.

(62) Lu, H.; Chen, J.; Jia, Y.; Cai, M.; Lee, PK.H. Transcriptomic Responses of the Interactions between *Clostridium cellulovorans 743B* and *Rhodopseudomonas palustris*

CGA009 in a Cellulose-Grown Coculture for Enhanced Hydrogen Production. *Appl Environ Microbiol.* **2016.** *82*(15), 4546-4559.

(63) Guan, X.; Yan, X.; Li, Y.; Jiang, B.; Luo, X.; Chi, X. Diversity and arsenic-tolerance potential of bacterial communities from soil and sediments along a gold tailing contamination gradient. Can J Microbiol. **2017**. *63*(9), 788-805.

(64) Mu, Y, Zhou, L, Zeng, XC, Liu, L, Pan, Y, Chen X, Wang J, Li S, Li WJ, Wang Y. *Arsenicitalea aurantiaca* gen. nov., sp. nov., a new member of the family Hyphomicrobiaceae, isolated from high-arsenic sediment. Int J Syst Evol Microbiol. **2016** 66(12), 5478-5484.

(65) Limmer, M.A.; Wise, P.; Dykes, G.E.; Seyfferth, A.L. Silicon decreases dimethylarsinic acid concentration in rice grain and mitigates straighthead disorder. *Environmental science & technology* **2018**. 52 (8), 4809-4816.

Table 1: Corresponding differential copy number patterns (see heatmap Figure 4a & 4b) identified for 16S rRNA OTUs and *arsM* gene OTUs in planted versus unplanted and manured versus non-manured pairwise comparisons. **Group-I = s**ignificant increase in unplanted soil, **Group-II** = significant increase in planted soil**, Group-III** = significant increase in manured soil**, Group-IV** = significant increase in non-manured soil, ** highest, * second highest (if applicable) correlation to *arsM* OTUs. In bold >1% abundant genera, increase= increase in copy number

Group	**Bacterial 16S rRNA OTUs**	
	Phylum level	**Genus level**
I	Acidobacteria	*Acidobacteria_GP1 sp.***
	Chloroflexi	*Dehalogenimonas sp.*
	Cyanobacteria	*Cryptomonadaceae sp.*
	Verrumicrobia	*Alterococcus*
II	Firmicutes	*Tumebacillus**, ***Clostridium***[32], *Desulfotomaculum*
III	Acidobacteria	*Geothrix*
	Actinobacteria	*Aeromirobium, Georgenia, Thermobifida, Patulibacter, Cellulomonadaceae sp.*
	Bacteroidetes	*Paludibacter, Flavobacterim, Muricauda, Chitonophaga, Parasegetibacter, Algoriphagus, Dyadobacter*
	Cyanobacteria	*Cyanobacteria_GpIV sp. Bacillariophyta sp.*
	Firmicutes	*Geobacillus, Ammoniphilus, Brevibacillus, Cohnella, Paenibacillus, Thermobacillus, Paenisporosarcina, Sporosarcina, Planifilum, Thermoactinomyces, Dehalobacter, Thermoflavimicrobium, Oxobacter, Alkalibacter, Desulfitobacterium, Papillibacter, Halocella, Lutispora, Anoxybacillus, Ocenobacillus, Ureibacillus, Desmospora, Garciella, Parasporobacterium, Desulfosporosinus, Sporobacter, Dethibacter, Tepianaerobacter*
	Proteobacteria	*Aquabacterium, Cellvibrio, Thermomonas, Brevundimonas, Beijerinckia, Bosea, Devosia*, Kaistia, Rhizobium, Xanthobacter, Paracoccus, Azospirillum, Magnetospirillum, Rickettsia, Novosphingobium, Sphingobium, Sphingopyxis, Sphingosinicella, Achromobacter, Ralstonia, Hydrogenophaga, Herbaspirillum, Herminiimonas, Methylophilus, Methylovorus, Aquaspirillum, Azoarcus, Shinella, **Desulfocapsa**, **Geobacter**, Rheinheimera, Halomonas, Pseudomonas*[33], Pseudoxanthomonas*
	Verrumicrobia	*Luteolibacter**, *Verrucomicrobiaceae sp., Optutaceae sp*
IV	Acidobacteria	*Acidobacteria_Gp25**, *Gp22, Gp11, Gp5 sp.*
	Bacteroidetes	*Bacteroides, Terrimonas*, Prevotellaceae sp.*
	Firmicutes	*Tepidibacter*
	Cyanobacteria	*Cyanobacteria_GpXI sp., Streptophyta sp.*
	Proteobacteria	*Phaselicystis, Methylosarcina, Elioraea sp., Sinobacteraceae sp.*

964 **Table2:** Annotation of representative *arsM* OTUs via Blastn against bacterial genomes using
965 1e-06 as cutoff for significance. The table shows the percentage of all sequences in each
966 group with best match against each genome: 1) all 2680 *arsM* OTUs in the final filtered OTU
967 table, 2) the 347 *arsM* OTU's with identified significant changes in copy number, G1= 29
968 OTUs increased in unplanted soil, G2= 32 OTUs increased in planted soil**,** G3= 256 OTUs
969 increased in manured soil**,** G4= 57 OTUs significant increase in non-manured soil.

blastn, genomes, top hit	All *arsM*	Sig. *arsM*	G1	G2	G3	G4
Rhodopseudomonas palustris TIE-1	7.9	23.3	3.4	9.4	28.9	15.8
Rhodopseudomonas palustris BisB5	6.7	5.6	41.4	0.0	2.0	7.0
*Roseiarcus fermentans strain DSM 24875**	6.2	31.8	0.0	3.1	45.7	1.8
Gemmatirosa kalamazoonesis strain KBS708	5.0	2.4	3.4	3.1	0.4	10.5
Rhodopseudomonas palustris strain YSC3	3.5	1.6	10.3	0.0	0.4	3.5
Streptomyces zinciresistens K42	3.3	1.1	0.0	3.1	0.8	1.8
Thiobacillus denitrificans DSM 12475	3.2	0.5	0.0	0.0	0.4	1.8
Rhodopseudomonas palustris strain XCP	2.6	1.6	3.4	0.0	1.6	1.8
Blastopirellula marina DSM 3645	1.6	0.8	0.0	6.3	0.4	0.0
Rhodoplanes roseus strain DSM 5909	1.6	1.6	0.0	0.0	1.6	3.5
Rhodopseudomonas palustris HaA2	1.5	1.6	0.0	0.0	1.2	5.3
Rhodopseudomonas pentothenatexigens	1.2	1.1	0.0	3.1	0.0	5.3
Thioalbus denitrificans strain DSM 26407	1.1	1.9	0.0	0.0	2.7	0.0
Intrasporangium chromatireducens Q5-1	1.0	0.5	0.0	3.1	0.4	0.0
Rhodomicrobium udaipurense JA643	0.9	0.8	0.0	0.0	1.2	0.0
Rhodoplanes elegans strain DSM 11907	0.8	0.8	3.4	0.0	0.0	3.5
Streptomyces aurantiacus JA 4570	0.7	1.6	6.9	12.5	0.0	0.0
Rhodopseudomonas palustris CGA009	0.7	0.8	0.0	3.1	0.8	0.0
Nocardioides sp. JS614	0.7	0.5	0.0	0.0	0.0	3.5
Acidobacteriaceae bacterium SbA1	0.6	0.5	0.0	0.0	0.4	1.8
Stackebrandtia nassauensis DSM 44728	0.4	0.5	0.0	0.0	0.0	3.5
Halobacterium sp. DL1	0.4	0.5	0.0	0.0	0.4	1.8
Streptomyces sp. Root1310	0.4	0.5	0.0	0.0	0.4	1.8
Methanoculleus taiwanensis strain CYW4	0.4	0.5	0.0	0.0	0.0	3.5
Rhodopseudomonas palustris DX-1	0.3	0.5	0.0	3.1	0.4	0.0
Thiocapsa sp. KS1	0.3	0.5	6.9	0.0	0.0	0.0
Planctomyces sp. SH-PL62	0.2	0.5	0.0	6.3	0.0	0.0
Microlunatus phosphovorus NM-1	0.2	0.5	6.9	0.0	0.0	0.0
Telmatospirillum siberiense strain 26-4b1	0.1	0.5	0.0	6.3	0.0	0.0
matches to other genomes	25.0	7.5	10.3	21.9	3.5	15.8
no genome match	21.5	7.2	3.4	15.6	6.6	7.0

970
971 *most of these also highly similar to *Rhodopseudomonas palustris TIE-1* (second best genome match)

Legends to figures

Figure 1. Box plot of soil solution arsenic speciation and relative soil *arsM* and 16S rRNA copy number in replicated microcosms (experiment2). The error bars are standard error of means as calculated in the GLM analysis. Paired lines above bars show identified significant manure (2 larger lines) and plant effect (2 smaller lines), with '*', '**', '***' above indicating $p<0.05$, $p<0.01$, ***$p<0.001$, respectively and '+' indicating significant interaction effect of manure and plants (GLM, Minitab). The associated probabilities are reported in Table S3.

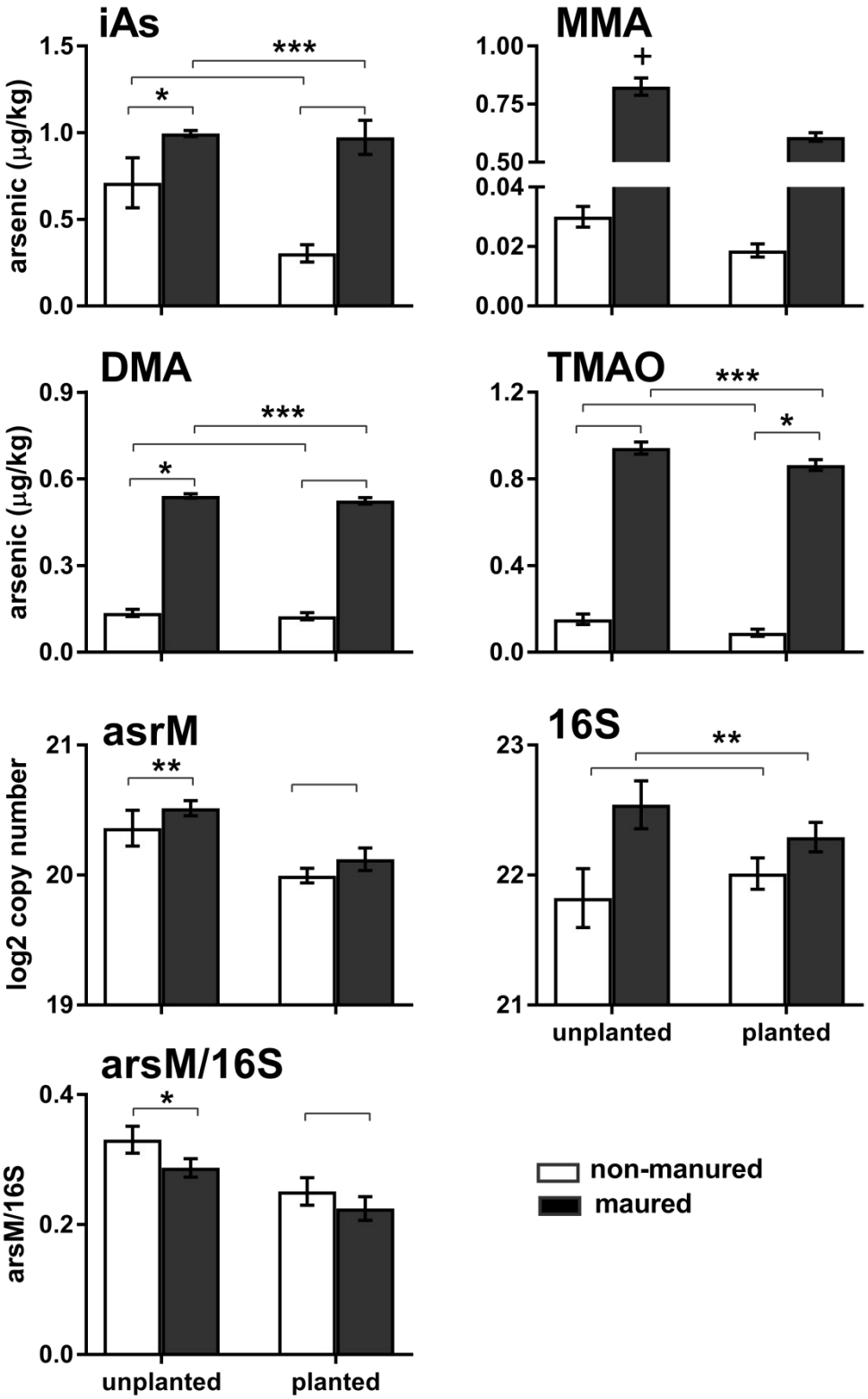
Figure 2. Heatmap showing normalized counts of OTUs with significant differences in copy number ($p<0.05$ and absolute log2FC >1, experiment 2). **(A)** level 6 16S rRNA; **(B)** *arsM* gene OTUs. Darker blue indicates greater normalized count. **Group-I = s**ignificantly higher copy number in unplanted soil, **Group-II = s**ignificantly higher copy number in planted soil**,** **Group-III = s**ignificantly higher copy number in manured soil**, Group-IV = s**ignificantly higher copy number in non-manured soil. **A:** Acidobacteria, **Ac:** Actinobacteria, **B:** Bacteroidetes, **Ch:** Chloroflexi, **Cy:** Cyanobacteria, **F:** Firmicutes, **P:** Proteobacteria and **V:** Verrumicrobia.

Figure 3. Variance partition analysis plot for genus level 16S rRNA OTUs with significant differences in copy number in response to planting (heatmap, Group I and II). A: Violin plot, showing the median and variance of the contribution of manure treatment and planting across all OTUs, B: Corresponding bar chart for each OTU.

997 **Figure 4.** Network analysis of genus level 16S rRNA OTUs that show significant differences
998 in copy number in response to planting. A, B, C, D= selected level 6 OTUs (genera) with
999 significantly higher copy number in non-planted soil (heatmap, Group I), A=
1000 *AcidobacteriaGP1*, B=*Dehalogenimonas*, C=*Cryptomonadaceae*, D= *Alterococcus*. E, F, G=
1001 selected level 6 OTUs with significantly higher copy number in planted soil (heatmap,
1002 GroupII), E= *Tumebacillus*, F= *Clostridium*, G= *Desulfomaculum*. Missing samples in the
1003 network for *Cryptomonadaceae* (C) are samples with a copy number count of 0.
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021

1022

1023 **Figure1.**



1024

1025

Figure 2.

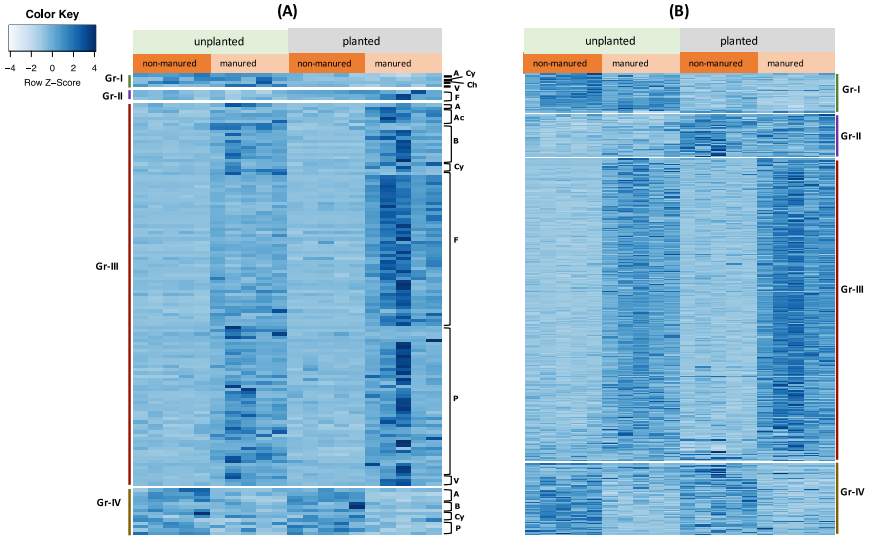
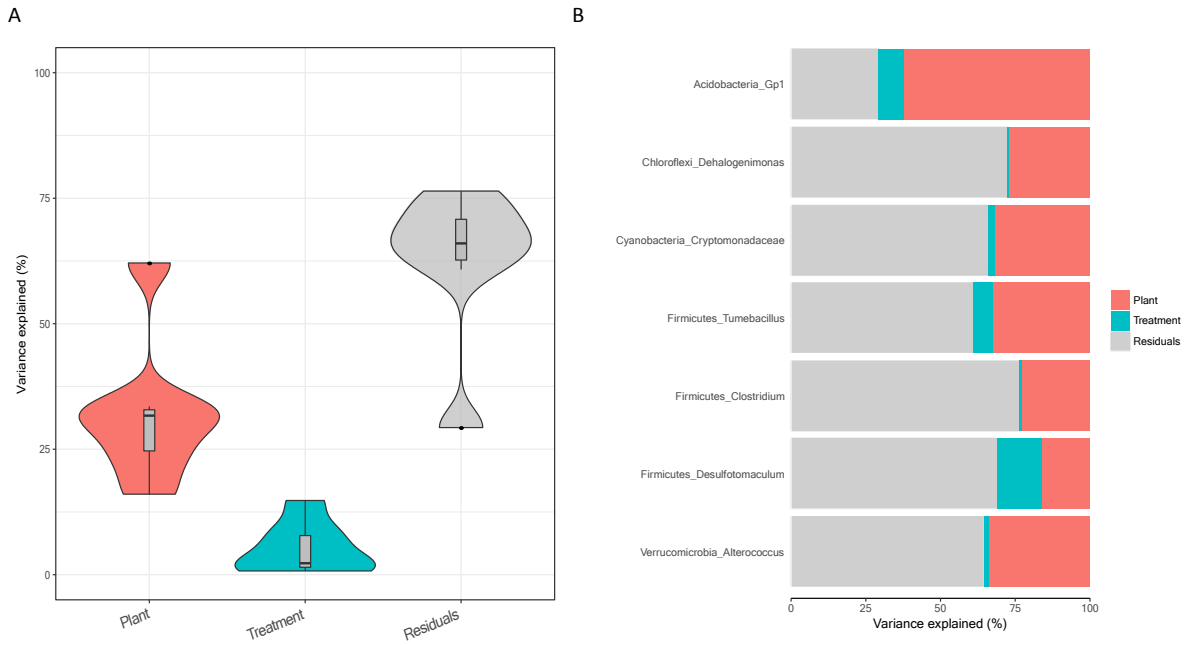
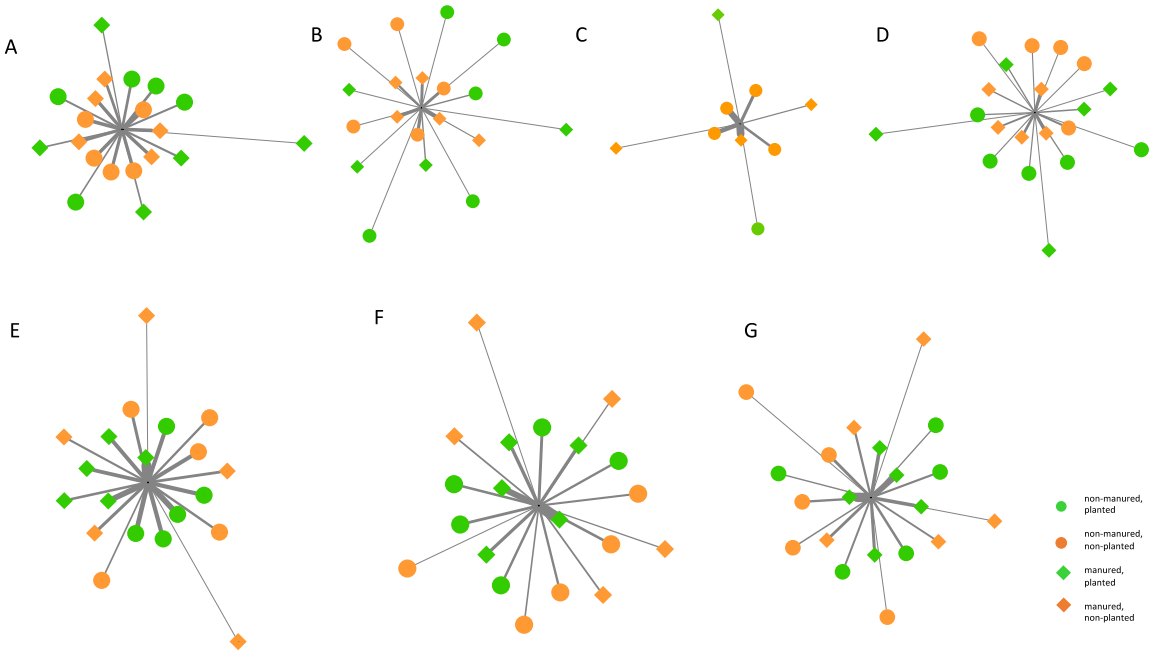

Color Key
−4 −2 0 2 4
Row Z−Score
(A)
unplanted
planted
non-manured
manured
non-manured
manured
Gr-I
Gr-II
Gr-III
Gr-IV
A
Cy
V
Ch
F
A
Ac
B
Cy
F
P
V
A
B
Cy
P
(B)
unplanted
planted
non-manured
manured
non-manured
manured
Gr-I
Gr-II
Gr-III
Gr-IV

Figure 3.


A
B
Variance explained (%)
100
75
50
25
0
Plant
Treatment
Residuals
Acidobacteria_Gp1
Chloroflexi_Dehalogenimonas
Cyanobacteria_Cryptomonadaceae
Firmicutes_Tumebacillus
Firmicutes_Clostridium
Firmicutes_Desulfotomaculum
Verrucomicrobia_Alterococcus
Variance explained (%)

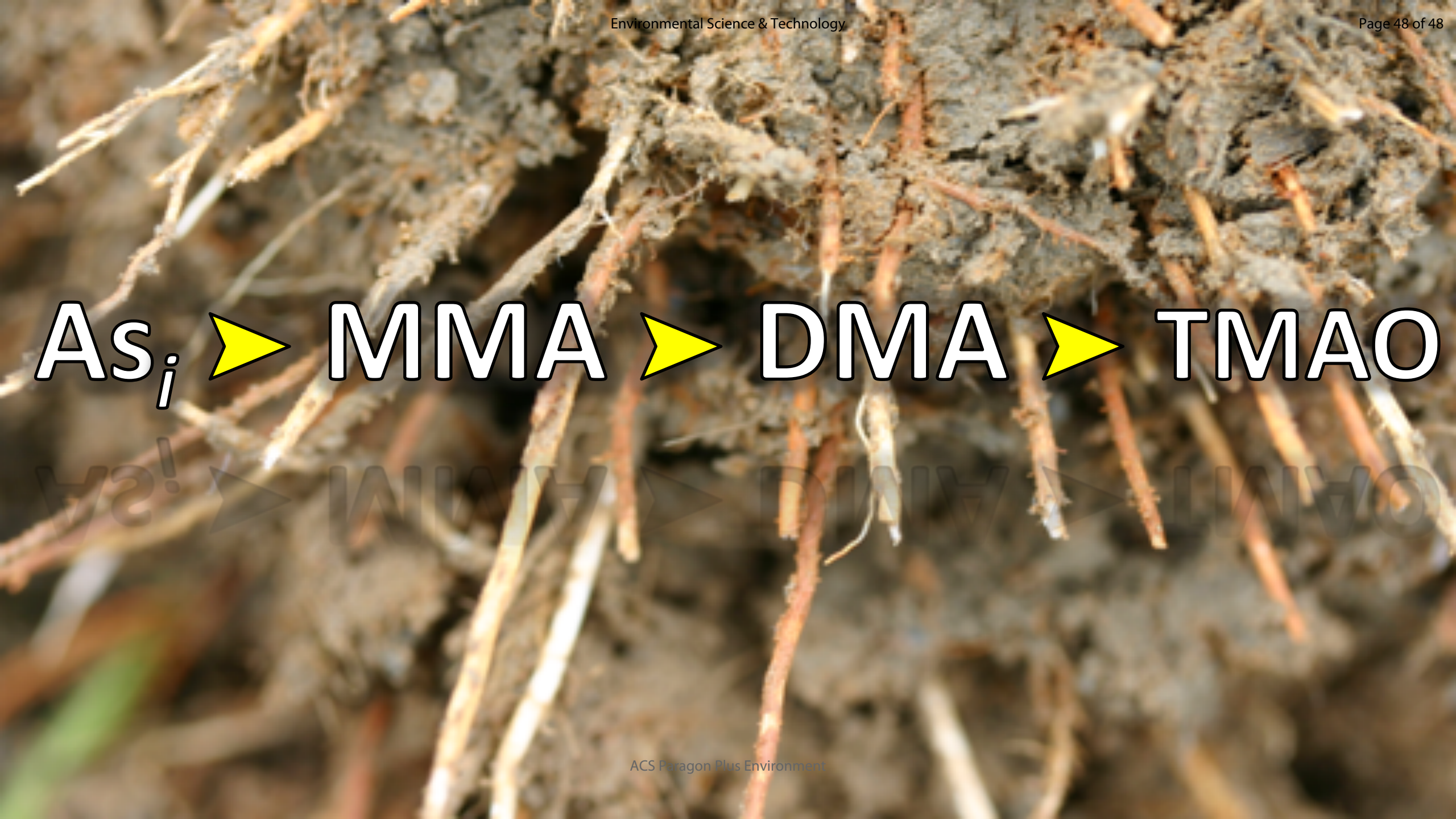
1060

1061 **Figure 4.**



1062

1063


As_i ➤ MMA ➤ DMA ➤ TMAO