

1 **Title:**
2 **Comparative genomics analyses indicate differential methylated amine**
3 **utilisation trait within the member of the genus *Gemmobacter***

4
5 Running title:
6 Methylated amine utilisation trait in *Gemmobacter*

7
8 Kröber, E.¹, Spurgin L.², Wischer, D., Kumaresan, D.^{3#}

9
10
11
12 ¹ Microbial Biogeochemistry, RA Landscape Functioning, ZALF Leibniz Centre for Landscape
13 Research, Müncheberg, Germany

14 ² School of Environmental Sciences, University of East Anglia, Norwich, UK

15 ⁴ School of Biological Sciences & Institute for Global Food Security, Queen's University
16 Belfast, Belfast, Northern Ireland, UK

17
18
19
20 [#]Correspondence to
21 Dr D Kumaresan, School of Biological Sciences, Queen's University Belfast, Belfast, Northern
22 Ireland, United Kingdom, d.kumaresan@qub.ac.uk

Abstract

Methylated amines are ubiquitous in the environment and play a role in regulating the earth's climate via a set of complex biological and chemical reactions. Microbial degradation of these compounds is thought to be a major sink. Recently we isolated a facultative methylotroph, *Gemmobacter* sp. LW-1, an isolate from the unique environment Movile Cave, Romania, which is capable of methylated amine utilisation as a carbon source. Here, using a comparative genomics approach, we investigate how widespread methylated amine utilisation trait is within the member of the bacterial genus *Gemmobacter*. Five genomes of different *Gemmobacter* species isolated from diverse environments, such as activated sludge, fresh water, sulphuric cave waters (Movile Cave) and the marine environment were available from the public repositories and used for the analysis. Our results indicate that some members of the genus *Gemmobacter*, namely *G. aquatilis*, *G. caeni* and *G. sp. LW-1* have the genetic potential of methylated amine utilisation while others (*G. megaterium* and *G. nectarophilus*) have not. Ancestral state reconstruction analysis also suggested that methylated amine utilisation trait might not be ancestral to members of *Gemmobacter* and has been gained. Based on our analysis, we suggest that the trait of methylated amine utilisation within the members of the genus *Gemmobacter* might be independent of their habitat and more randomly distributed.

Introduction

Methylated amines (MAs) are ubiquitous in the environment with a variety of natural and anthropogenic sources including the oceans, vegetation, sediments and organic-rich soils, animal husbandry, food industry, pesticides, sewage, and automobiles, to mention only a few (Schade and Crutzen, 1995; Latypova et al., 2010; Ge et al., 2011).

Methylated amines are also known to influence earth's climate, via a series of complex biological and chemical interactions (Carpenter et al., 2012). Some of the most abundant methylated amines found in the atmosphere are trimethylamine (TMA), dimethylamine (DMA) and monomethylamine (MMA) (Ge et al., 2011). Microbial metabolism of methylated amines involves both aerobic and anaerobic microorganisms, e.g. some methanogenic archaea such as *Methanosarcina* and *Methanomicrobium* can use MAs to produce methane (Burke et al., 1998;Liu and Whitman, 2008;Lyimo et al., 2009) while Gram-positive and Gram-negative methylotrophic bacteria can use MAs as carbon and nitrogen source (Chen et al., 2010b). Previously, MAs were typically associated with marine ecosystems as they are by-products of degradation of osmolytic chemicals such as glycine betaine, carnitine, choline and trimethylamine N-oxide (Chen et al., 2010b). However, recent studies have reported the detection and activity of aerobic methylotrophic bacteria that utilise MAs in a variety of natural and engineered environments (Chen et al., 2009;Chistoserdova et al., 2009;Chistoserdova, 2011;Ge et al., 2011;Wischer et al., 2015) and could play a major role in global C and N budgets.

Aerobic methylotrophs are a polyphyletic group of microorganisms capable of utilising one-carbon (C₁) compounds such as methane, methanol or methylated amines as their sole source of carbon and energy (Anthony, 1982;Lidstrom, 2006;Chistoserdova et al., 2009). Methylotrophs can degrade TMA to DMA by using the enzymes TMA dehydrogenase, TMA monooxygenase or TMA methyltransferase, encoded by the genes *tdm*, *tmm* and *mtt*, respectively (Paul et al., 2000;Chen, 2012;Lidbury et al., 2014). The enzymes DMA dehydrogenase (*dmd*) or DMA monooxygenase (*dmmDABC*) modulate the conversion of DMA to MMA (Lidstrom,

2006;Chen, 2012). Two distinct pathways have been characterised for the oxidation of MMA (Chistoserdova, 2011). The direct MMA-oxidation pathway mediated by a single enzyme (MMA dehydrogenase in Gram-negative bacteria and MMA oxidase in Gram-positive bacteria) converts MMA to formaldehyde and releases ammonium (McIntire et al., 1991;Chistoserdova et al., 1994). The alternate pathway, referred to as the *N*-methylglutamate (NMG) pathway or indirect MMA-oxidation pathway, is mediated by three individual enzymes via the oxidation of MMA to gamma-glutamylmethanamide (GMA) and its further degradation to *N*-methylglutamate (NMG) and 5,10-methylenetetrahydrofolate ($\text{CH}_2 = \text{H}_4\text{F}$) (Latypova et al., 2010;Chistoserdova, 2011). A stepwise conversion of MMA in the NMG pathway is modulated by the enzymes GMA synthetase (*gmaS*), NMG synthase (*mgsABC*) and NMG dehydrogenase (*mgdABCD*) (Chen et al., 2010b;Latypova et al., 2010). The capability to use MMA not only as a source for carbon but also for nitrogen is widespread in bacteria. Notably, the NMG pathway is not only restricted to methylotrophs but also present in non-methylotrophic bacteria that use MMA as a nitrogen but not as a carbon source (Chen et al., 2010a;Chen, 2012;Taubert et al., 2017).

In a recent study, we isolated an alphaproteobacterial facultative methylotrophic bacterium, *Gemmobacter* sp. LW-1 (recently renamed from *Catellibacterium* (Chen et al., 2013)) from the Movile Cave ecosystem (Kumaresan et al., 2014) (Mangalia, Romania) that can use methylated amines as both carbon and nitrogen source (Wischer et al., 2015) and subsequently obtained its genome sequence (Kumaresan et al., 2015). Using a ^{13}C -MMA DNA based stable-isotope probing (SIP) experiment we also showed that *Gemmobacter* sp. LW-1 was indeed an active MMA utiliser in

microbial mats from this environment (Wischer et al., 2015). This was the first report of methylated amine utilisation in a member of the bacterial genus *Gemmobacter*. However, growth on C₁ compounds (methanol and formate) has been reported for the genus *Gemmobacter*, e.g. in *G. caeni* (Zheng et al., 2011).

The genus *Gemmobacter* (family *Rhodobacteraceae*) currently comprises ten validated species: *Gemmobacter megaterium* (Liu et al., 2014), *G. nectarophilum* (Tanaka et al., 2004;Chen et al., 2013), *G. aquatilis* (Rothe et al., 1987), *G. caeni* (Zheng et al., 2011;Chen et al., 2013), *G. aquaticus* (Liu et al., 2010;Chen et al., 2013), *G. nanjingense* (Zhang et al., 2012;Chen et al., 2013), *G. intermedius* (Kämpfer et al., 2015), *G. lanyuensis* (Sheu et al., 2013b), *G. tilapiae* (Sheu et al., 2013a) and *G. fontiphilus* (Chen et al., 2013). These species were isolated from a wide range of environments including fresh water environments (freshwater pond (Rothe et al., 1987;Sheu et al., 2013a), freshwater spring (Chen et al., 2013;Sheu et al., 2013b)), coastal planktonic seaweed (Liu et al., 2014), white stork nestling (Kämpfer et al., 2015), waste water and activated sludge (Tanaka et al., 2004;Zheng et al., 2011;Zhang et al., 2012), suggesting that members of the genus *Gemmobacter* are widely distributed in engineered and natural environments.

Here, using a comparative genomics approach we study how widespread methylated amine utilisation trait is within the members of the genus *Gemmobacter*. We used five isolate genomes for members within the genus *Gemmobacter* (*G. sp.* LW-1, *G. caeni*, *G. aquatilis*, *G. nectariphilus* and *G. megaterium*) available from public repositories (Accessed June 2018). These genomes were used for comparative genomics and

ancestral state reconstruction analyses to understand patterns of gain/loss of methylated amine utilisation across the *Gemmobacter* phylogeny.

Materials and Methods

Genome data acquisition

Five *Gemmobacter* genomes (*G. caeni*, *G. aquatilis*, *G. nectarophilus*, *G. megaterium*, *Gemmobacter* sp. LW-1) available through the Integrated Microbial Genomes (IMG) database (<https://img.jgi.doe.gov/>) were used for comparative genome analysis (Markowitz et al., 2013). Accession numbers and genome characteristics are listed in Supplementary Table S1.

Phylogenetic analysis

Phylogenetic relatedness between the different members of the genus *Gemmobacter* was determined using phylogenetic trees constructed from 16S rRNA gene sequences (nucleotide) and metabolic gene sequences (*gmaS* and *mauA*; amino acids) involved in MMA utilisation. RNAmmer (Lagesen et al., 2007) was used to retrieve 16S rRNA gene sequences from the genome sequences. Multiple sequence alignment of 16S rRNA gene sequences was performed using the SINA alignment service via SILVA (Pruesse et al., 2007; Pruesse et al., 2012) and subsequently imported into MEGA7 (Kumar et al., 2016) to construct a maximum-likelihood nucleotide-based phylogenetic tree (Saitou and Nei, 1987). Bootstrap analysis was performed with 1000 replicates to provide confidence estimates for phylogenetic tree topologies (Felsenstein, 1985).

To determine phylogenetic affiliations for the protein encoding genes *gmaS* and *mauA*, gene sequences retrieved from the genome sequences were aligned to homologous sequences retrieved from the NCBI Genbank database using Basic Local Alignment Search Tool (BLAST, blastx) (Altschul et al., 1990) and curated *gmaS* sequences used for primer design in our previous study (Wischer et al., 2015). Amino acid sequences were aligned in MEGA7 (Kumar et al., 2016) using ClustalW (Thompson et al., 1994) and the alignment was subsequently used to construct maximum likelihood phylogenetic trees based on the JTT matrix-based model (Jones et al., 1992). Bootstrap analysis was performed with 1000 replicates to provide confidence estimates for phylogenetic tree topologies (Felsenstein, 1985).

Comparative genomic analyses

CGView Comparison Tool (CCT) was used to visually compare the genomes within the genus *Gemmobacter* (Grant et al., 2012). CCT utilises BLAST to compare the genomes and the BLAST results are presented in a DNA-based graphical map (Grant et al., 2012). Average Nucleotide Identity (ANI) (Rodriguez-R and Konstantinidis, 2016) between different genomes was estimated using one-way ANI (best hit) and two-way ANI (reciprocal best hit) based on Goris *et al.* (Goris et al., 2007). In addition the whole-genome based average nucleotide identity (gANI) and the $p_r^{\text{intra-species}}$ value were determined for *G. sp. LW-1* and *G. caeni* (these two genomes revealed the closest ANI) based on Konstantinidis and Tiedje (Konstantinidis and Tiedje, 2005) via the Joint Genome Institute (JGI) platform (<https://ani.jgi-psf.org/html/home.php>; Version 0.3, April 2014). In order to determine if two genomes belong to the same species, the computation of empirical probabilities ($p_r^{\text{intra-species}}$) can be calculated as follows,

$$p_r^{intra-species}[AF = a, ANI = b] = p_r^{intra-species}[AF = a] * p_r^{intra-species}[ANI = b|AF = a]$$

AF represents alignment fraction.

Pan-genome analysis including average amino acid identity (AAI) analysis, pan-genome tree construction and determination of core and dispensable genes and singletons (unique genes) was carried out using the Efficient Database framework for comparative Genome Analyses using BLAST score Ratios (EDGAR) platform (Blom et al., 2016).

In order to compare the genetic potential for methylated amine utilisation within the available *Gemmobacter* genomes, known protein sequences involved in methylated amine utilisation pathways (Latypova et al., 2010; Chen, 2012) were used as query sequences through the BLAST (blastp) program (Altschul et al., 1990) available within the Rapid Annotation using Subsystem Technology (RAST) server (Aziz et al., 2008). The list of protein queries used is given in Supplementary Table S2.

We examined patterns of gain/loss of methylated amine utilisation along the above 16S rRNA gene phylogeny, by performing ancestral state reconstruction analysis using the phytools package in R (Revell, 2012). We used stochastic character mapping (Nielsen, 2002) to map presence/absence of methylated amine utilisation, assigning a prior probability of one to species known to utilise methylated amines, zero to those known not to do so, and 0.5 to those where the trait value was unknown. We used MCMC (with the function *make.simmap*) to simulate 1000 stochastic maps, and then we used the function *densityMap* (Revell, 2013) to visualise the aggregate result from the stochastic mapping analysis.

209

210 **Results and discussion**

211 *Phylogenetic relatedness based on the 16S rRNA and metabolic gene*

212 The phylogenetic relatedness of the five members within the genus *Gemmobacter* (*G.*
 213 sp. LW-1, *G. caeni*, *G. aquatilis*, *G. nectariphilus* and *G. megaterium*) was resolved
 214 based on 16S rRNA gene sequences (Figure 1). Three members of the genus
 215 *Gemmobacter* (*G.* sp. LW-1, *G. caeni*, and *G. aquatilis*) clustered together with
 216 several other related *Gemmobacter* and *Rhodobacter* 16S rRNA gene sequences
 217 retrieved from fresh water, soil and sediment and activated sludge environments
 218 (Figure 1). *G. nectariphilus* and *G. megaterium* sequences clustered together with
 219 *Paracoccus kawasakiensis* and other related *Gemmobacter* sequences from marine,
 220 fresh water and activated sludge environments (Figure 1). Based on the 16S rRNA
 221 gene sequences retrieved from public database, we observed that the members of the
 222 genus *Gemmobacter* are widely distributed in engineered (such as activated sludge
 223 and clinical environments) and natural environments i.e. fresh water, soil and
 224 sediment, and marine environments (Figure 1). GMA synthetase, a key enzyme in the
 225 NMG pathway, is encoded by the gene *gmaS*. *gmaS* sequences retrieved from the
 226 isolate genomes along with other ratified *gmaS* sequences were used to construct an
 227 amino acid-based phylogenetic tree (Figure 2). *gmaS* gene sequences retrieved from
 228 genomes of *G.* sp. LW-1, *G. caeni* and *G. aquatilis* clustered within Group I of
 229 alphaproteobacterial *gmaS* sequences containing sequences from marine and non-
 230 marine bacteria within the orders *Rhodobacterales* and *Rhizobiales* as described in
 231 Wischer et al. (Wischer et al., 2015) and were closely related to *Paracoccus yeei*, *P.*
 232 sp. 1W-5 and *Rhodobacter* sp. 1W-5 (Figure 2).

233

Whilst *gmaS* gene sequences were detected in three of the five investigated *Gemmobacter* genomes, *mauA* gene sequences were identified only in the genomes of *G. caeni* and *G. sp. LW-1* (Supplementary Figure S1). It has been suggested that the NMG pathway for MMA utilisation is more universally distributed and more abundant across proteobacterial methylotrophs than the direct MMA oxidation pathway (Nayak and Marx, 2015). However, it should be noted that genes encoding for the enzymes within the NMG pathway (*gmaS*) can not only be detected in methylotrophs but also in non-methylotrophic bacteria that use MMA as a nitrogen source, but not as a carbon source (Chen, 2012; Wischer et al., 2015).

A comparative genome analysis of members within the genus Gemmobacter

At the time of the analysis, five *Gemmobacter* genomes obtained from isolates from different environments were available (Figure 1 and Table 2). *Gemmobacter* genome sizes range from ~3.96 Mb to ~5.14 Mb with GC contents between 64.71% to 66.19% (Table 2). Analysis of sequence annotations revealed that on average 91.13% of the genomes consist of coding sequences.

The genomes were compared using the CGView comparison tool (Grant et al., 2012) (Figure 3). *Gemmobacter sp. LW-1*, isolated from the Movile Cave ecosystem was used as the reference genome and the results of the BLAST comparison with other *Gemmobacter* genomes are represented as a BLAST ring for each genome (Figure 3). Similarities between segments of the reference genome sequence and the other genome sequences are shown by a coloured arc beneath the region of similarity indicating the percentage of similarity as a colour code. Our analysis (Figure 3) revealed low identity levels (mostly <88%) between *Gemmobacter sp. LW-1* and *G.*

aquaticus, *G. nectarophilus* and *G. megaterium* across the genomes. Moreover, the analysis suggested several sites of potential insertion/deletion events in the genome of *Gemmobacter* sp. LW-1. Possible insertion/deletion regions can be identified as those gaps in the map where no homology is detected. For example the region between 2200-2300 kbp (Figure 3) where a gap can be found in the otherwise contiguous homologous regions between the reference genome *G. sp. LW-1* and the first of the query genomes (*G. caeni*). This might likely be due to a lack of hits or hits with low identity that can be spurious matches. Since it covers a large region we could possibly rule out that it is not an artefact arising from a lack of sensitivity in the BLAST analysis. Even though the genomes of *G. sp. LW-1* and *G. caeni* are closely related, our analysis demonstrates that their genomes are not completely identical. Despite the fact that the majority of their genomes indicate very high identity levels (mostly >96-98% as shown by the dominance of dark red colours of the circle representing the BLAST hit identity between *G. sp. LW-1* and *G. caeni*), many segments appear to be exclusive to *G. sp. LW-1*.

In order to further resolve the similarity between these genomes we calculated the average nucleotide identity (ANI) (Rodriguez-R and Konstantinidis, 2016) (Supplementary Table S3 and Supplementary Figure S2A-D) and the average amino acid identity (AAI; Supplementary Figure S2E). It is generally accepted that an ANI value of >95-96% can be used for species delineation (Richter and Rossello-Mora, 2009; Kim et al., 2014). Our analysis revealed that *Gemmobacter* sp. LW-1 and *Gemmobacter caeni* share an ANI value of 98.62 (Supplementary Table S3) implying that both are in fact the same species. The genome-based average nucleotide identity (gANI) between *G. sp. LW-1* and *G. caeni* was calculated as 98.70. The AF was

calculated to be 0.91, which would result in a computed probability of 0.98 suggesting that both genomes might belong to the same species. However, it should be noted that these are draft genomes and a more in depth characterization of their physiology and phenotype is required to delineate these organisms at the level of strain.

Pan-genome analysis, carried out using the EDGAR platform (Blom et al., 2016), identified metabolic genes present in all *Gemmobacter* species (core genes), two or more *Gemmobacter* species (accessory or dispensable genes), and unique *Gemmobacter* species (singleton genes). A pan-genome tree was constructed (Figure 4A) based on the pan-genome dataset and neighbor-joining method (Saitou and Nei, 1987). As with the 16S-rRNA gene based phylogenetic tree (Figure 1), the five *Gemmobacter* species formed two main clusters in the pan-genome tree analysis (Figure 4A). The pan-genome tree also confirmed the phylogenetic closeness between *Gemmobacter caeni* and *Gemmobacter* sp. LW-1 (Figure 4A). According to pan-genome analysis of the five *Gemmobacter* genomes, a total of 9,286 genes were identified, consisting of 1,806 core genes, 3085 dispensable genes and 305, 1,072, 896, 1,165 and 957 singletons for *G. sp. LW-1*, *G. caeni*, *G. aquatilis*, *G. nectarophilus* and *G. megaterium*, respectively (Figure 4B).

Methylated amine utilisation, N assimilation and C₁ oxidation

Investigation of the methylated amine utilisation pathways in five *Gemmobacter* species revealed the presence of the genes encoding enzymes TMA dehydrogenase (*tmd*), TMA monooxygenase (*tmm*), TMAO demethylase (*tdm*) and DMA monooxygenase in genomes of *G. sp. LW-1*, *G. caeni* and *G. aquatilis* while none of

these genes were detected in *G. nectariphilus* or *G. megaterium* (Figure 5). These findings are supported by results from a previous study which showed growth of *G. sp.* LW-1 on TMA as a carbon and nitrogen source (Wischer et al., 2015). *G. sp.* LW-1, *G. caeni* and *G. aquatilis* could potentially use the TMA oxidation pathway to convert TMA to DMA. Based on the genome sequences, it can be suggested that these three species could use the enzyme DMA monooxygenase (*dmmDABC*) to oxidize DMA to MMA but not the DMA dehydrogenase since the corresponding protein encoding gene (*dmd*) was not found (Figure 5).

We also compared the distribution of the direct MMA-oxidation and the NMG pathways in the genomes of five *Gemmobacter* species (Figure 5). The direct MMA-oxidation pathway (*mauA*-dependent) is so far only known to be present in methylotrophic bacteria that can use MMA as a carbon source. Whereas the NMG pathway (*gmaS*-dependent) has been shown to be present in non-methylotrophic bacteria that can use MMA as a nitrogen source (Chen et al., 2010a; Nayak and Marx, 2015; Wischer et al., 2015; Nayak et al., 2016). Analysis of the genome sequences revealed that both *G. sp.* LW-1 and *G. caeni* possess genes for both MMA oxidation pathways (Figure 5). We have previously shown that *Gemmobacter sp.* LW-1 can use MMA and TMA as both a carbon and nitrogen source (Wischer et al., 2015). Genome sequence of *G. aquatilis* indicated the presence of genes involved only in the NMG pathway. In the facultative methylotroph *Methylobacterium extorquens* AM1 it has been shown that the NMG pathway is advantageous compared to the direct MMA-oxidation pathway (Nayak et al., 2016). NMG pathway enables facultative methylotrophic bacteria to switch between using MMA as a nitrogen source or as a carbon and energy source whereas the direct MMA oxidation pathway allows for

rapid growth on MMA only as the primary energy and carbon source (Nayak et al., 2016). This could suggest that *G. aquatilis* might use the NMG pathway for utilising MMA as both nitrogen and carbon source. However, growth assays are required to confirm whether *G. aquatilis* can use MMA as a carbon source. We did not detect genes for either MMA oxidation pathways in the genome sequences of *G. nectarophilus* and *G. megaterium* suggesting the lack of genetic potential of these organisms to use MMA as either C or N source.

The C₁ units derived from methylated amines need to be further oxidized when the nitrogen is sequestered without assimilation of the carbon from the methylated amines. Genome analysis confirmed that all five *Gemmobacter* species possess the genetic capability for C₁ oxidation and also indicate that tetrahydrofolate (H₄F) is the C₁ carrier (Figure 5). The bifunctional enzyme 5,10-methylene-tetrahydrofolate dehydrogenase/ cyclohydrolase, encoded by the gene *folD*, was detected in all the *Gemmobacter* genomes (Figure 5/ Table 1). Genes encoding key enzymes in the C₁ oxidation pathway via tetrahydromethanopterin (H₄MPT) were not detected. (Chistoserdova, 2011). The formate-tetrahydrofolate ligase, encoded by the gene *fhs* (Figure 5), provides C₁ units for biosynthetic pathways (Chen, 2012). However, the oxidation of formyl-H₄F (CHO-H₄F) can also be facilitated by *purU*, the gene encoding for the formyl-H₄F deformylase. The formate dehydrogenase (*fdh*) mediates the last step of the C₁ oxidation pathway, the oxidation of formate to CO₂. The genes for the C₁ oxidation pathway via H₄F were detected in all five *Gemmobacter* genomes.

The *fae* gene, encoding the formaldehyde-activating enzyme that catalyses the reduction of formaldehyde with H₄MPT was not detected in any of the five *Gemmobacter* genomes confirming that these members of the genus *Gemmobacter* lack the H₄MPT pathway for formaldehyde oxidation (Figure 5/ Table 1). Investigation of the nitrogen assimilation pathway revealed the presence of the genes encoding glutamine synthetase (GS; *gluL*) and glutamine synthase (GOGAT; *glxB*) in all five *Gemmobacter* genomes. In bacteria this pathway is essential for glutamate synthesis at low ammonium concentrations (Chen, 2012).

Using comparative genome analysis we provide genome-based evidence that the two *Gemmobacter* isolates *G. sp. LW-1* and *G. caeni* are capable of generating energy from complete oxidation of methylated amines via the H₄F-dependent pathway using either the NMG pathway or the direct MMA oxidation pathway. *Gemmobacter aquatilis* is genetically capable of methylated amine degradation to yield formaldehyde and only encodes the genes for the NMG pathway, which indicates that *G. aquatilis* could use this pathway to use MMA as a nitrogen source. Both *G. nectarophilus* and *G. megaterium* genomes indicate the lack of potential to use methylated amines (Figure 5/Table 1). Therefore, the question arises if the genes for methylated amine utilisation have been acquired or lost.

Stochastic character mapping along the 16S rRNA gene phylogeny suggested that the ability to use methylated amines has either been gained or lost multiple times (Supplementary Figure S3). Most likely is that methylated amine utilisation is not ancestral in *Gemmobacter*, and has evolved three times, once in the clade containing *G. sp. LW-1*, *G. caeni* and *G. aquatilis*, once in *Haematobacter* and once in

Paracoccus (Supplementary Figure S3). An alternative, but less plausible, scenario is that methylated amine utilisation is an ancestral trait in *Gemmobacter*, and has been lost and regained multiple times across the phylogeny.

Gemmobacter sp. LW-1 was isolated from the Movile Cave ecosystem (Wischer et al., 2015). Microbial mats and lake water within the cave have been shown to harbor a wide diversity of methylated amine-utilising bacteria (Wischer et al., 2015; Kumaresan et al., 2018). Whilst the mechanism of MAs production within the system has to be elucidated, it can be speculated that degradation of floating microbial mats (i.e. organic matter) could result in MAs (Wischer et al., 2015). Similarly, *G. caeni* isolated from activated sludge (Zheng et al., 2011) could possibly use the MAs generated from organic matter degradation. Interestingly, whilst *G. megaterium* was isolated from a marine environment (seaweed (Liu et al., 2014)) possibly encountering MAs from the degradation of osmolytes such as glycine betaine (*N,N,N*-trimethylglycine) we did not detect metabolic genes involved in methylated amine utilisation. Our analyses suggest that the trait for methylated amine utilisation could be independent of the habitat.

Conclusions

In summary, three of the five investigated *Gemmobacter* genomes (*G. sp.* LW-1, *G. caeni* and *G. aquatilis*) indicated metabolic potential to utilise methylated amines, of which only two (*G. sp.* LW-1 and *G. caeni*) possess the genes for both MMA oxidation pathways, the NMG pathway and the direct MMA oxidation pathway. *G. sp.* LW-1 and *G. caeni* are facultative methylotrophs which could potentially use these pathways to utilise MMA as both a carbon and nitrogen source, while

potentially *G. aquatilis* could only use the NMG pathway as a nitrogen source. Furthermore, the genomes of *G. sp.* LW-1 and *G. caeni* showed a high similarity to each other (>98%) suggesting that both belong to the same species. *G. megaterium* and *G. nectariphilus* genomes indicated no metabolic potential to utilise MAs. Phylogenetic, pan-genome and ANI analyses revealed that *G. sp.* LW-1 and *G. caeni* are closely related, although they were isolated from different environments. Whilst *G. caeni* and *G. nectariphilus* were isolated from a similar environment (activated sludge) it revealed a high amount of evolutionary change from the common ancestor. Overall, these results suggest that the trait for methylated amine utilisation could be independent from the habitat and localised factors or selection pressures could influence the ability of these organisms to use methylated amines. Access to *Gemmobacter* isolates with or without the genetic potential for methylated amine utilisation trait will allow us to perform physiological experiments in future to test how this trait can affect fitness of closely related organisms. Ancestral state reconstruction analysis confirms that across *Gemmobacter* and related genera, methylated amine utilisation has either evolved or been lost multiple times over the evolutionary history of this group. The adaptive or non-adaptive processes behind this pattern remain to be investigated.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgements

We thank Dr Felipe Hernandes Coutinho (Federal University of Rio de Janeiro) for a fruitful discussion on genomic regions of insertion and deletion in *Gemmobacter sp.*

433 LW-1. The EDGAR platform is financially supported by the BMBF grant FKZ
434 031A533 within the de.NBI network.

435

436

437

438

439

440

441

442

443

444

445

446

447

448

449

450

451

452

453

454

455

456

457

REFERENCES

- Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990). Basic local alignment search tool. *Journal of Molecular Biology* 215, 403-410.
- Anthony, C. (1982). *The biochemistry of methylotrophs*. London: Academic Press.
- Aziz, R.K., Bartels, D., Best, A.A., Dejongh, M., Disz, T., Edwards, R.A., Formsma, K., Gerdes, S., Glass, E.M., and Kubal, M. (2008). The RAST Server: rapid annotations using subsystems technology. *BMC Genomics* 9, 75.
- Blom, J., Kreis, J., Spanig, S., Juhre, T., Bertelli, C., Ernst, C., and Goesmann, A. (2016). EDGAR 2.0: an enhanced software platform for comparative gene content analyses. *Nucleic Acids Research* 44, W22-W28.
- Burke, S.A., Lo, S.L., and Krzycki, J.A. (1998). Clustered genes encoding the methyltransferases of methanogenesis from monomethylamine. *Journal of Bacteriology* 180, 3432-3440.
- Carpenter, L.J., Archer, S.D., and Beale, R. (2012). Ocean-atmosphere trace gas exchange. *Chemical Society Reviews* 41, 6473-6506.
- Chen, W.M., Cho, N.T., Huang, W.C., Young, C.C., and Sheu, S.Y. (2013). Description of *Gemmobacter fontiphilus* sp. nov., isolated from a freshwater spring, reclassification of *Catellibacterium nectariphilum* as *Gemmobacter nectariphilus* comb. nov., *Catellibacterium changlense* as *Gemmobacter changlensis* comb. nov., *Catellibacterium aquatile* as *Gemmobacter aquaticus* nom. nov., *Catellibacterium caeni* as *Gemmobacter caeni* comb. nov., *Catellibacterium nanjingense* as *Gemmobacter nanjingensis* comb. nov., and emended description of the genus *Gemmobacter* and of *Gemmobacter aquatilis*. *International Journal of Systematic and Evolutionary Microbiology* 63, 470-478.
- Chen, Y. (2012). Comparative genomics of methylated amine utilization by marine *Roseobacter* clade bacteria and development of functional gene markers (*tmm*, *gmaS*). *Environmental Microbiology* 14, 2308-2322.
- Chen, Y., Mcaleer, K.L., and Murrell, J.C. (2010a). Monomethylamine as a nitrogen source for a nonmethylotrophic bacterium, *Agrobacterium tumefaciens*. *Applied and Environmental Microbiology* 76, 4102-4104.
- Chen, Y., Scanlan, J., Song, L., Crombie, A., Rahman, M.T., Schäfer, H., and Murrell, J.C. (2010b). γ -Glutamylmethylamide is an essential intermediate in the metabolism of methylamine by *Methylocella silvestris*. *Applied and Environmental Microbiology* 76, 4530-4537.
- Chen, Y., Wu, L., Boden, R., Hillebrand, A., Kumaresan, D., Moussard, H., Baci, M., Lu, Y., and Colin Murrell, J. (2009). Life without light: microbial diversity and evidence of sulfur- and ammonium-based chemolithotrophy in Movile Cave. *ISME Journal* 3, 1093-1104.
- Chistoserdova, A.Y., Chistoserdova, L.V., McIntire, W.S., and Lidstrom, M.E. (1994). Genetic organization of the mau gene cluster in *Methylobacterium extorquens* AM1: complete nucleotide sequence and generation and characteristics of mau mutants. *Journal of Bacteriology* 176, 4052-4065.
- Chistoserdova, L. (2011). Modularity of methylotrophy, revisited. *Environmental Microbiology* 13, 2603-2622.
- Chistoserdova, L., Kalyuzhnaya, M.G., and Lidstrom, M.E. (2009). The expanding world of methylotrophic metabolism. *Annual Review of Microbiology* 63, 477-499.

507 Felsenstein, J. (1985). Confidence limits on phylogenies: an approach using the
508 bootstrap. *Evolution* 39, 783-791.

509 Ge, X., Wexler, A.S., and Clegg, S.L. (2011). Atmospheric amines – Part I. A review.
510 *Atmospheric Environment* 45, 524-546.

511 Goris, J., Konstantinidis, K.T., Klappenbach, J.A., Coenye, T., Vandamme, P., and
512 Tiedje, J.M. (2007). DNA–DNA hybridization values and their relationship
513 to whole-genome sequence similarities. *International Journal of*
514 *Systematic and Evolutionary Microbiology* 57, 81-91.

515 Grant, J.R., Arantes, A.S., and Stothard, P. (2012). Comparing thousands of
516 circular genomes using the CGView Comparison Tool. *BMC genomics* 13,
517 202.

518 Jones, D.T., Taylor, W.R., and Thornton, J.M. (1992). The rapid generation of
519 mutation data matrices from protein sequences. *Bioinformatics* 8, 275-
520 282.

521 Kämpfer, P., Jerzak, L., Wilharm, G., Golke, J., Busse, H.-J., and Glaeser, S.P. (2015).
522 *Gemmobacter intermedius* sp. nov., isolated from a white stork (*Ciconia*
523 *ciconia*). *International Journal of Systematic and Evolutionary Microbiology*
524 65, 778-783.

525 Kim, M., Oh, H.-S., Park, S.-C., and Chun, J. (2014). Towards a taxonomic
526 coherence between average nucleotide identity and 16S gene sequence
527 similarity for species demarcation of prokaryotes. *International Journal of*
528 *Systematic and Evolutionary Microbiology* 64, 346-351.

529 Konstantinidis, K.T., and Tiedje, J.M. (2005). Genomic insights that advance the
530 species definition for prokaryotes. *Proceedings of the National Academy of*
531 *Sciences of the United States of America* 102, 2567-2572.

532 Kumar, S., Stecher, G., and Tamura, K. (2016). MEGA7: Molecular Evolutionary
533 Genetics Analysis version 7.0 for bigger datasets. *Molecular Biology and*
534 *Evolution* 33, 1870-1874.

535 Kumaresan, D., Stephenson, J., Doxey, A.C., Bandukwala, H., Brooks, E.,
536 Hillebrand-Voiculescu, A., Whiteley, A.S., and Murrell, J.C. (2018). Aerobic
537 proteobacterial methylotrophs in Movile Cave: genomic and metagenomic
538 analyses. *Microbiome* 6.

539 Kumaresan, D., Wischer, D., Hillebrand-Voiculescu, A.M., and Murrell, J.C. (2015).
540 Draft genome sequences of facultative methylotrophs, *Gemmobacter* sp.
541 strain LW1 and *Mesorhizobium* sp. strain 1M-11, isolated from Movile
542 Cave, Romania. *Genome Announcements* 3.

543 Kumaresan, D., Wischer, D., Stephenson, J., Hillebrand-Voiculescu, A., and
544 Murrell, J.C. (2014). Microbiology of Movile Cave - A
545 chemolithoautotrophic ecosystem. *Geomicrobiology Journal* 31, 186-193.

546 Lagesen, K., Hallin, P., Rødland, E.A., Stærfeldt, H.-H., Rognes, T., and Ussery, D.W.
547 (2007). RNAmmer: consistent and rapid annotation of ribosomal RNA
548 genes. *Nucleic acids research* 35, 3100-3108.

549 Latypova, E., Yang, S., Wang, Y.S., Wang, T., Chavkin, T.A., Hackett, M., Schäfer, H.,
550 and Kalyuzhnaya, M.G. (2010). Genetics of the glutamate - mediated
551 methylamine utilization pathway in the facultative methylotrophic beta -
552 proteobacterium *Methyloversatilis universalis* FAM5. *Molecular*
553 *Microbiology* 75, 426-439.

554 Lidbury, I., Murrell, J.C., and Chen, Y. (2014). Trimethylamine N-oxide
555 metabolism by abundant marine heterotrophic bacteria. *Proceedings of*
556 *the National Academy of Sciences* 111, 2710-2715.

557 Lidstrom, M.E. (2006). "Aerobic Methylophilic Prokaryotes," in *The*
558 *Prokaryotes: Volume 2: Ecophysiology and Biochemistry*, eds. M. Dworkin,
559 S. Falkow, E. Rosenberg, K.-H. Schleifer & E. Stackebrandt. (New York, NY:
560 Springer New York), 618-634.

561 Liu, J.-J., Zhang, X.-Q., Chi, F.-T., Pan, J., Sun, C., and Wu, M. (2014). *Gemmobacter*
562 *megaterium* sp. nov., isolated from coastal planktonic seaweeds.
563 *International Journal of Systematic and Evolutionary Microbiology* 64, 66-
564 71.

565 Liu, Y., and Whitman, W.B. (2008). Metabolic, phylogenetic, and ecological
566 diversity of the methanogenic archaea. *Annals of the New York Academy of*
567 *Sciences* 1125, 171-189.

568 Liu, Y., Xu, C.-J., Jiang, J.-T., Liu, Y.-H., Song, X.-F., Li, H., and Liu, Z.-P. (2010).
569 *Catellibacterium aquatile* sp. nov., isolated from fresh water, and emended
570 description of the genus *Catellibacterium* Tanaka et al. 2004. *International*
571 *Journal of Systematic and Evolutionary Microbiology* 60, 2027-2031.

572 Lyimo, T.J., Pol, A., Jetten, M.S.M., and Op Den Camp, H.J.M. (2009). Diversity of
573 methanogenic archaea in a mangrove sediment and isolation of a new
574 *Methanococcoides* strain. *FEMS Microbiology Letters* 291, 247-253.

575 Markowitz, V.M., Chen, I.-M.A., Palaniappan, K., Chu, K., Szeto, E., Pillay, M.,
576 Ratner, A., Huang, J., Woyke, T., and Huntemann, M. (2013). IMG 4 version
577 of the integrated microbial genomes comparative analysis system. *Nucleic*
578 *acids research* 42, D560-D567.

579 McIntire, W.S., Wemmer, D.E., Chistoserdova, A., and Lidstrom, M.E. (1991). A
580 new cofactor in a prokaryotic enzyme: tryptophan tryptophylquinone as
581 the redox prosthetic group in methylamine dehydrogenase. *Science* 252,
582 817-824.

583 Nayak, D.D., Agashe, D., Lee, M.-C., and Marx, C.J. (2016). Selection maintains
584 apparently degenerate metabolic pathways due to tradeoffs in using
585 methylamine for carbon versus nitrogen. *Current Biology* 26, 1-11.

586 Nayak, D.D., and Marx, C.J. (2015). Experimental horizontal gene transfer of
587 methylamine dehydrogenase mimics prevalent exchange in nature and
588 overcomes the methylamine growth constraints posed by the sub-optimal
589 N-Methylglutamate pathway. *Microorganisms* 3, 60-79.

590 Nielsen, R. (2002). Mapping mutations on phylogenies. *Systematic Biology* 51,
591 729-739.

592 Paul, L., Ferguson, D.J., and Krzycki, J.A. (2000). The trimethylamine
593 methyltransferase gene and multiple dimethylamine methyltransferase
594 genes of *Methanosarcina barkeri* contain in-frame and read-through
595 amber codons. *Journal of Bacteriology* 182, 2520-2529.

596 Pruesse, E., Peplies, J., and Glöckner, F.O. (2012). SINA: accurate high-throughput
597 multiple sequence alignment of ribosomal RNA genes. *Bioinformatics* 28,
598 1823-1829.

599 Pruesse, E., Quast, C., Knittel, K., Fuchs, B.M., Ludwig, W., Peplies, J., and Glöckner,
600 F.O. (2007). SILVA: a comprehensive online resource for quality checked
601 and aligned ribosomal RNA sequence data compatible with ARB. *Nucleic*
602 *Acids Research* 35, 7188-7196.

603 Revell, L.J. (2012). phytools: an R package for phylogenetic comparative biology
604 (and other things). *Methods in Ecology and Evolution* 3, 217-223.

605 Revell, L.J. (2013). Two new graphical methods for mapping trait evolution on
606 phylogenies. *Methods in Ecology and Evolution* 4, 754-759.

607 Richter, M., and Rossello-Mora, R. (2009). Shifting the genomic gold standard for
608 the prokaryotic species definition. *Proceedings of the National Academy of*
609 *Sciences* 106, 19126-19131.

610 Rodriguez-R, L.M., and Konstantinidis, K.T. (2016). The enveomics collection: a
611 toolbox for specialized analyses of microbial genomes and metagenomes.
612 *PeerJ Preprints*.

613 Rothe, B., Fischer, A., Hirsch, P., Sittig, M., and Stackebrandt, E. (1987). The
614 phylogenetic position of the budding bacteria *Blastobacter aggregatus*
615 and *Gemmobacter aquatilis* gen., nov. sp. nov. *Archives of Microbiology*
616 147, 92-99.

617 Saitou, N., and Nei, M. (1987). The neighbor-joining method: a new method for
618 reconstructing phylogenetic trees. *Molecular Biology and Evolution* 4, 406-
619 425.

620 Schade, G.W., and Crutzen, P.J. (1995). Emission of aliphatic-amines from animal
621 husbandry and their reactions - potential source of N₂O and HC_n. *Journal*
622 *of Atmospheric Chemistry* 22, 319-346.

623 Sheu, S.-Y., Sheu, D.-S., Sheu, F.-S., and Chen, W.-M. (2013a). *Gemmobacter tilapiae*
624 sp. nov., a poly-β-hydroxybutyrate-accumulating bacterium isolated from
625 a freshwater pond. *International Journal of Systematic and Evolutionary*
626 *Microbiology* 63, 1550-1556.

627 Sheu, S.-Y., Shiau, Y.-W., Wei, Y.-T., and Chen, W.-M. (2013b). *Gemmobacter*
628 *lanyuensis* sp. nov., isolated from a freshwater spring. *International*
629 *Journal of Systematic and Evolutionary Microbiology* 63, 4039-4045.

630 Tanaka, Y., Hanada, S., Manome, A., Tsuchida, T., Kurane, R., Nakamura, K., and
631 Kamagata, Y. (2004). *Catellibacterium nectariphilum* gen. nov., sp. nov.,
632 which requires a diffusible compound from a strain related to the genus
633 *Sphingomonas* for vigorous growth. *International Journal of Systematic*
634 *and Evolutionary Microbiology* 54, 955-959.

635 Taubert, M., Grob, C., Howat, A.M., Burns, O.J., Pratscher, J., Jehmlich, N., Von
636 Bergen, M., Richnow, H.H., Chen, Y., and Murrell, J.C. (2017). Methylamine
637 as a nitrogen source for microorganisms from a coastal marine
638 environment. *Environmental Microbiology* 19, 2246-2257.

639 Thompson, J.D., Higgins, D.G., and Gibson, T.J. (1994). CLUSTAL W: improving the
640 sensitivity of progressive multiple sequence alignment through sequence
641 weighting, position-specific gap penalties and weight matrix choice.
642 *Nucleic Acids Research* 22, 4673-4680.

643 Wischer, D., Kumaresan, D., Johnston, A., El Khawand, M., Stephenson, J.,
644 Hillebrand-Voiculescu, A.M., Chen, Y., and Colin Murrell, J. (2015).
645 Bacterial metabolism of methylated amines and identification of novel
646 methylotrophs in Movile Cave. *ISME Journal* 9, 195-206.

647 Zhang, J., Chen, S.-A., Zheng, J.-W., Cai, S., Hang, B.-J., He, J., and Li, S.-P. (2012).
648 *Catellibacterium nanjingense* sp. nov., a propanil-degrading bacterium
649 isolated from activated sludge, and emended description of the genus
650 *Catellibacterium*. *International Journal of Systematic and Evolutionary*
651 *Microbiology* 62, 495-499.

Zheng, J.-W., Chen, Y.-G., Zhang, J., Ni, Y.-Y., Li, W.-J., He, J., and Li, S.-P. (2011).
Description of *Catellibacterium caeni* sp. nov., reclassification of
Rhodobacter changlensis Anil Kumar et al. 2007 as *Catellibacterium*
changelense comb. nov. and emended description of the genus
Catellibacterium. *International Journal of Systematic and Evolutionary*
Microbiology 61, 1921-1926.

Figures and Tables

Figure 1. Phylogenetic tree based on 16S rRNA gene sequences. The tree was constructed using the maximum likelihood method for clustering and the Tamura-Nei model for computing evolutionary distances. Numbers at branches are bootstrap percentages >50% of 1000 replicates. Star represents the *Gemmobacter* species used for comparative genome analysis. Coloured boxes represent the habitat where the sequence was retrieved: blue (fresh water), orange (soil and sediment), green (activated sludge), grey (marine), purple (clinical source). Triangles represent sequences that are listed as *Catellibacterium* in the NCBI database, which have been recently reclassified to *Gemmobacter* (Chen et al., 2013). Scale bar: 0.02 substitutions per nucleotide position.

Figure 2. Maximum-likelihood phylogenetic tree based on *gmaS* sequences. The tree was constructed using amino acid sequences (GmaS) using the maximum-likelihood method based on the JTT matrix-based model. Members of the genus *Gemmobacter* used for genome comparison are represented with a star. Numbers at branches are bootstrap percentages >50% of 1000 replicates. Amino acid sequences of the glutamine synthetase type III (GlnA) were used as out-group. Scale bar: 0.1 substitutions per amino acid position. MRC, marine *Roseobacter* clade.

Figure 3. DNA BLAST map of *Gemmobacter* genomes. *Gemmobacter* sp. LW-1 was used as a reference genome against *Gemmobacter megaterium* (inner ring), *Gemmobacter aquatilis* (second inner ring), *Gemmobacter nectarophilus* (third ring), and *Gemmobacter caeni* (fourth ring). The fifth and sixth ring (outer rings) represent the CDS (blue), tRNA (maroon), and rRNA (purple) on the reverse and forward

strand, respectively. The colour scale (inset) shows the level of sequence identity with the respective sequences from *G. megaterium*, *G. aquatilis*, *G. nectariphilus* and *G. caeni*. The locations of genes involved in methylotrophy are indicated at the outside of the map.

Figure 4. Pan-genome analysis. (A) Pan-genome tree consisting of five *Gemmobacter* species was constructed using the neighbour-joining method within the EDGAR platform. (B) Number of core, dispensable, and specific genes (singletons) of each *Gemmobacter* species.

Figure 5. Metabolic pathways involved in methylated amine utilisation and one-carbon utilisation annotated with presence/absence of specific genes in the genomes of *Gemmobacter*. The analysis was based on a five-way comparison among *Gemmobacter* sp. LW-1 (L), *Gemmobacter caeni* (C), *Gemmobacter aquatilis* (A), *Gemmobacter nectariphilus* (N) and *Gemmobacter megaterium* (M). The color-coded boxes next to the genes indicate the presence (green) or absence (orange) of a gene in each genome.

Table 1. Comparative genomic analysis of methylated amine-utilising genes in genomes-sequenced *Gemmobacters* in comparison to selected marine *Roseobacter* clade bacteria. Shown is the presence (+) or absence (-) of specific genes in the genome sequences.

Supplementary Figure S1. Maximum-likelihood phylogenetic tree (JTT matrix-based) of *mauA* sequences. Sequences from the genus *Gemmobacter* are marked

with a star. Amino acid sequences (MauA) were aligned using the ClustalW algorithm. Numbers at branches are bootstrap percentages >50% of 1000 replicates. Scale bar: 0.1 substitutions per amino acid. Coloured boxes indicate *Alphaproteobacteria* (yellow), *Gammaproteobacteria* (red) and *Betaproteobacteria* (black).

Supplementary Figure S2. (A-D) Average nucleotide identity (ANI) analysis of *Gemmobacter* sp. LW-1 and *Gemmobacter caeni*, *Gemmobacter aquatilis*, *Gemmobacter nectariphilus* and *Gemmobacter megaterium* and (E) AAI analysis between those species

Supplementary Figure S3. Ancestral state reconstruction of methylated amine utilisation along the 16S rRNA gene phylogeny, using stochastic mapping. Branch colour represents the posterior probability (computed as the relative frequency across stochastic maps) of methylated amine utilisation through the phylogeny. Red indicates a high posterior probability of methylated amine utilisation.

Supplementary Table S1. Genome characteristics of the five *Gemmobacter* isolate genomes used in this study.

Supplementary Table S2. List of protein queries used for the genome comparison with their accession number.

Supplementary Table S3. Average nucleotide identity (ANI) values between *Gemmobacter* sp. LW-1 and *Gemmobacter caeni*, *Gemmobacter aquatilis*,

769 *Gemmobacter nectarophilus*, *Gemmobacter megaterium*, *Rhodobacter sphaeroides*

770 and *Paracoccus denitrificans*

771

772

Figure 1

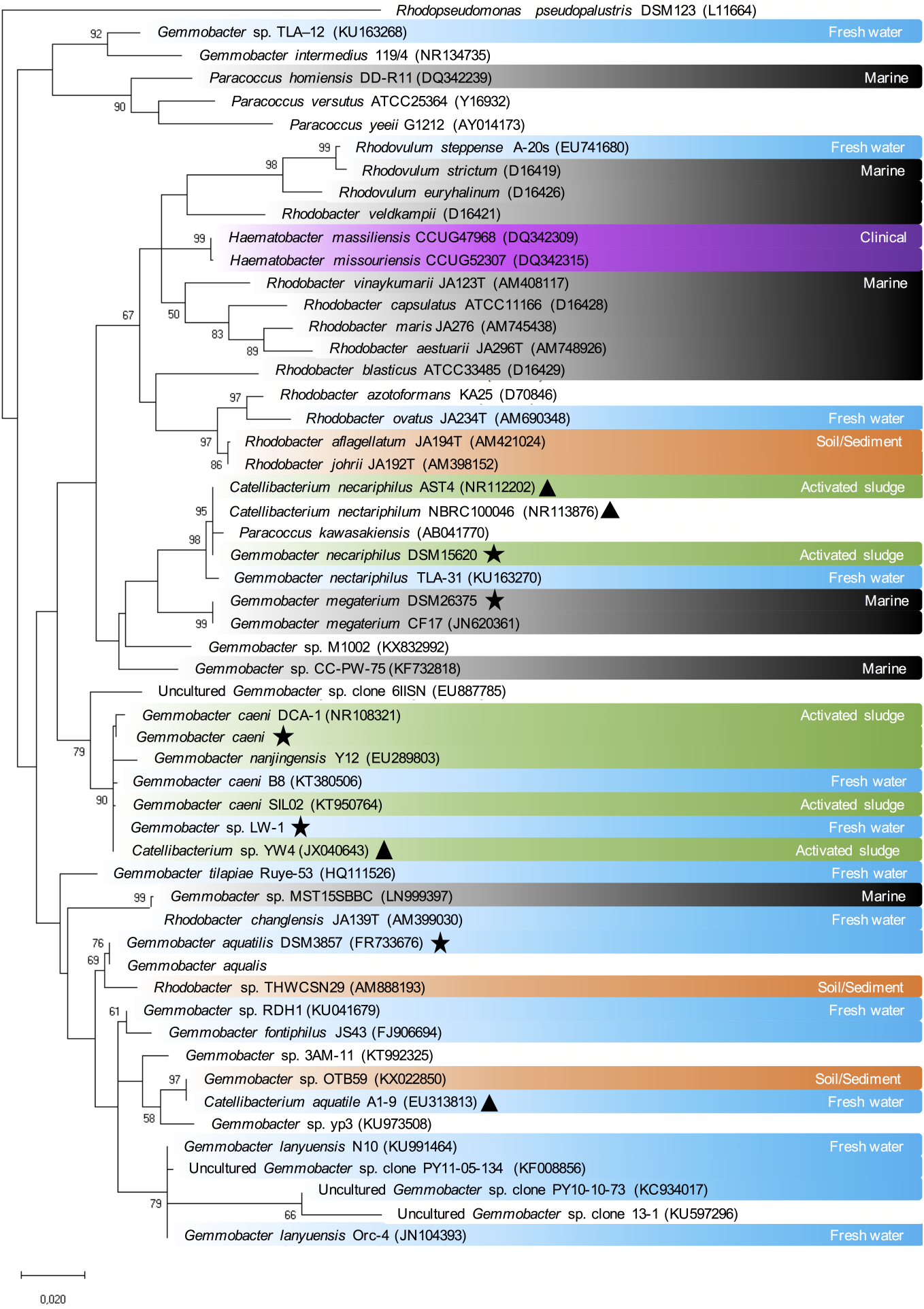
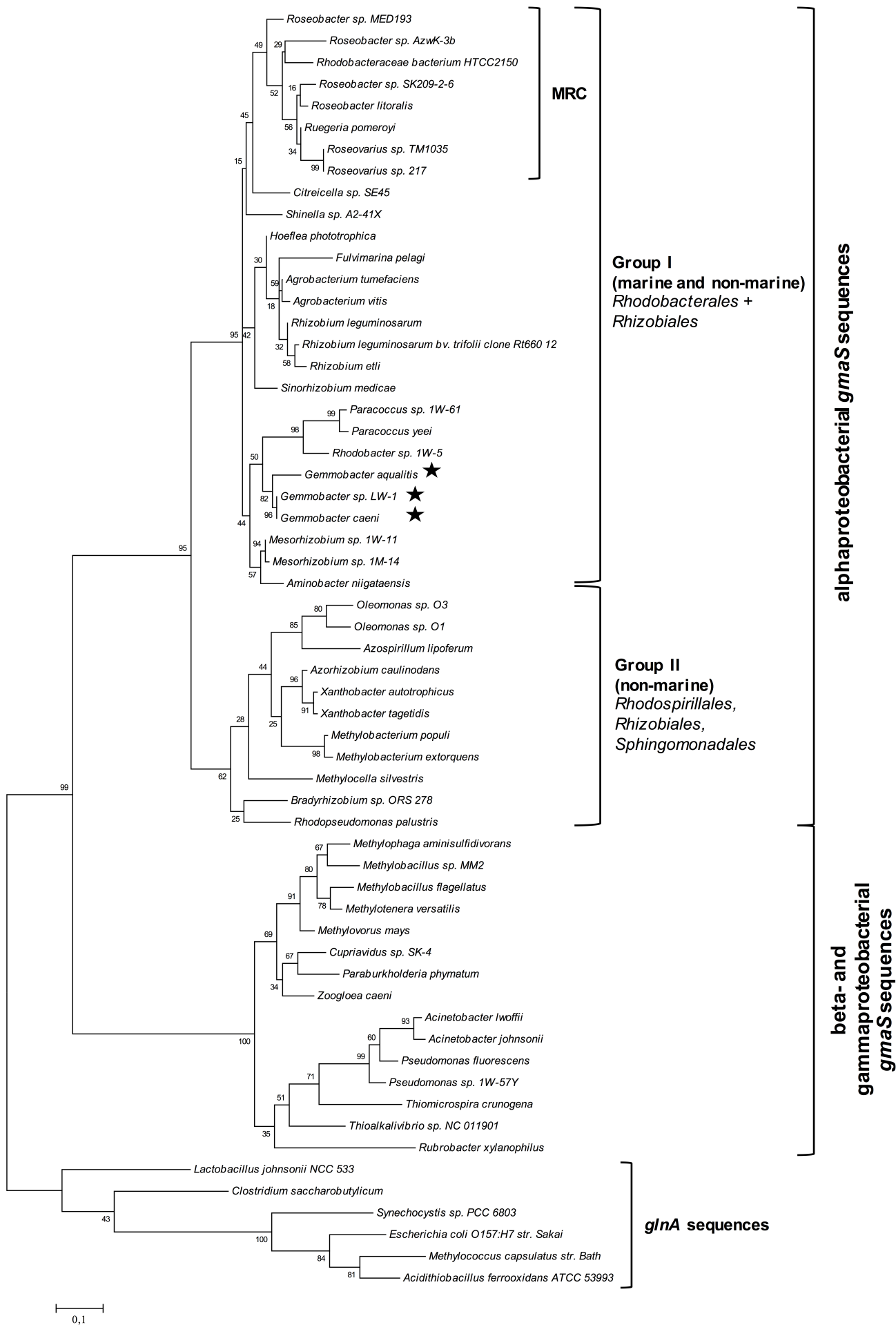
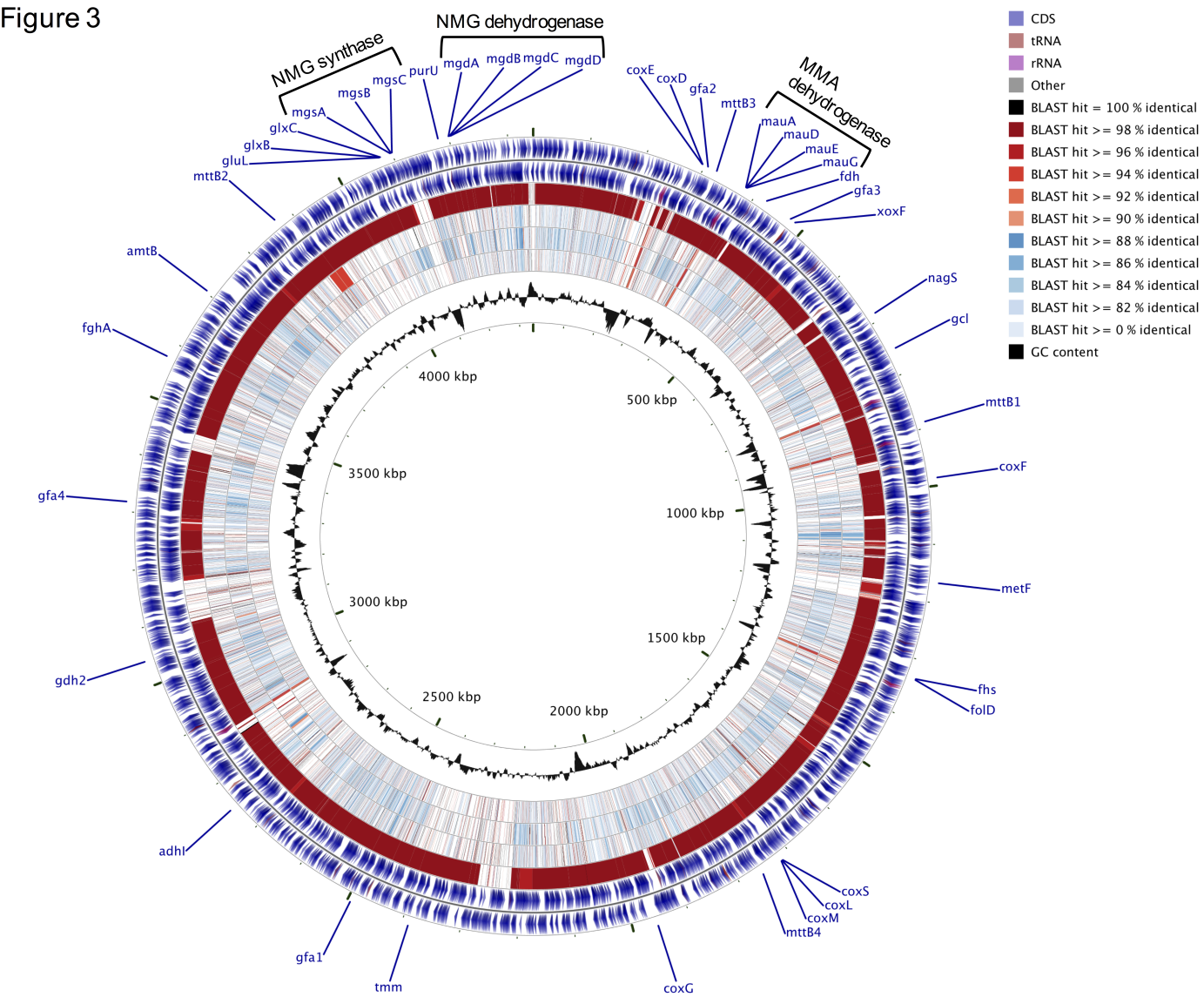


Figure 2





Protein	Corresponding gene(s)
TMA monooxygenase	tmm
TMA methyltransferase	mttB1, mttB2, mttB3, mttB4
MMA dehydrogenase	mauA, mauD, mauE, mauG
NMG synthase	mgsA, mgsB, mgsC
NMG dehydrogenase	mgdA, mgdB, mgdC, mgdD
HCHO-activating enzyme	gfa1, gfa2, gfa3, gfa4
GSH-dependent HCHO dehydrogenase	adhI
S-formyl-GSH hydrolase	fghA
Glutamine synthetase (GS)	gluL
Glutamate synthase (GOGAT)	glxB, glxC
Methylene-H ₄ F-dehydro-genase/cyclohydrolase	fold
Formate-H ₄ F ligase	fhs
Formyl-H ₄ F deformylase	purU
Formate dehydrogenase	fdh
Carbon monoxide dehydrogenase	coxD, coxE, coxF, coxG, coxL, coxM, coxS
Methanol dehydrogenase	xoxF
N-acetylglutamate synthase	nagS
Gamma-glutamylcysteine synthetase	gcl
Methylenetetrahydrofolate reductase	metF
Glutamate dehydrogenase	gdh2
Ammonium transporter	amtB

Figure 4A

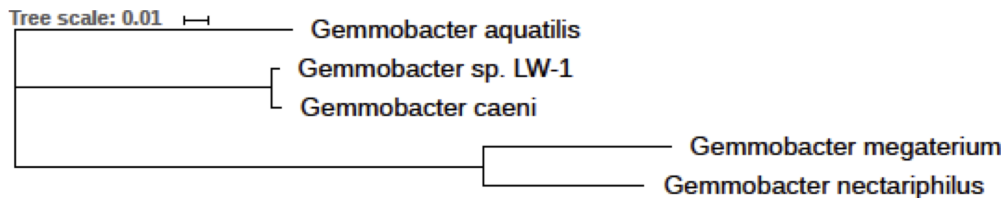


Figure 4B

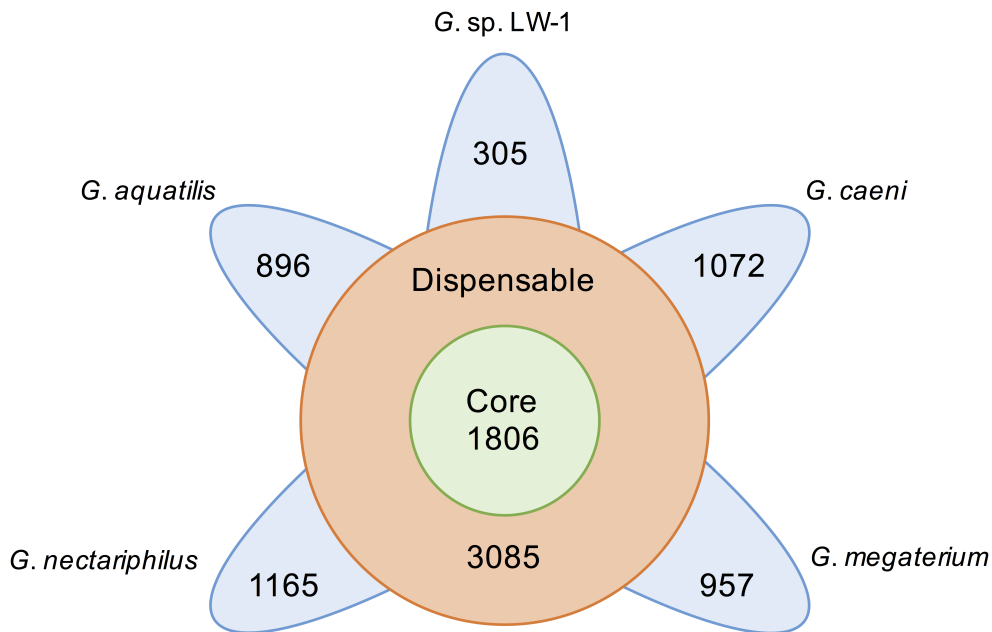


Figure 5

