Genome-scale Co-occurrence and Co-transcription Networks Reveal Key Microbial Taxa in Mangrove Sediments

Yang Liu
Shenzhen University
Zhichao Zhou
Shenzhen University
Yuchun Yang
Sun Yat-Sen University
Meng Li (limeng848@szu.edu.cn)
Shenzhen University https://orcid.org/0000-0001-8675-0758

Research

Keywords: Metagenomics, metatranscriptomics, network analysis, keystone taxon, genome-scale, Candidatus Mangrovidesulfobacterales

Posted Date: July 13th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-41034/v1

Abstract

Background

Mangroves are highly productive ecosystems, with one of the highest microbial diversities among all ecosystems. The concerted activity of microbial community in mangrove sediment mediates element cycling, but the underpinning mechanism of microbial synergy remains unknown.

Results

Here, we reconstructed 671 strain-resolved metagenome-assembled genomes (MAGs) from three mangrove and two mudflat sediments in Mai Po Nature Reserve. We then inferred the genome-scale co-occurrence and co-transcription networks based on metabolic capacity and transcriptional activity of the carbon, nitrogen, and sulfur cycles. We observed that the centrality was significantly higher in co-transcription networks than in co-occurrence networks, indicating that MAGs had stronger interrelationships when transcriptionally active. Further, we classified 57 microbes with low relative abundance (0.01–0.79%) as keystone taxa, which play key roles in the maintenance of co-transcription network structure, and participate in carbon transformations, denitrification, and sulfate reduction processes. One of the keystone taxa is a newly proposed deltaproteobacterial order, *Candidatus* Mangrovidesulfobacterales, capable of dissimilatory sulfate reduction and an anaerobic mixotrophic lifestyle. These findings highlight the ecological importance of rare species.

Conclusions

Collectively, this first screening of the potential keystone taxa in mangrove ecosystem based on genome-scale transcriptomic analysis revealed unique microbial functional assemblages, shedding light on microbial synergism in this ecosystem.

Background

Mangrove wetlands are distinct ecosystems. While they cover only 0.1% of the Earth's continental surface (approximately 81,485 km^2)[1], they store large amounts of carbon (C; 5–10.4 Pg globally)[2] andaccount for 14% of C sequestration by the global ocean[3]. Consequently, they are one of the most C-rich wetlands on the planet[4,5]. Mangrove sediments act as a large sink for recycled C, and play an essential role in reducing or preventing greenhouse gas emissions from the loss of C stocks, mitigating climate change[6].

Mangrove sediments are inhabited by a variety of unique microorganisms that play important roles in nutrient biogeochemical cycles[7–11]. According to a recent comparison of the microbial diversity in mangroves with those in other biomes in the Earth Microbiome Project, the total number of prokaryotic operational taxonomic units (OTUs) in mangroves exceeds 150,000, and the prokaryotic alpha-diversity in mangroves is significantly higher than that in other biomes[12]. Such high diversity reflects a great complexity of interconnected microbial communities that engage in synchronized activities (communicate, cross-feed, recombine, and coevolve) of a group. However, the scope and characteristics of microbial synergism are still largely unexplored[13].

Keystone taxon is a group of microorganisms that acts as a 'hub' or 'connector' [14], and is critical for the maintenance of the community structure and function[15], providing systematic insight into the microbial community structure and assembly patterns[16,17]. Computational removal of a keystone taxon from a network of co-occurrence of community members causes a dramatic shift in the composition of human and plant microbiome[18,19], highlighting its importance to the ecosystem function. While keystone taxa have been characterized in diverse terrestrial, aquatic, and human microbiomes[14,20,21], they have not yet been identified in mangroves.

The keystone taxa can be identified by OTU-based network analysis[14]; however, the resolution of the current OTU-based microbial co-occurrence network analysis is limited[22]. Consequently, it is not possible to distinguish strains with nearly identical 16S rRNA sequences but different functions in an ecosystem[22]. Genome-scale resolution of the metabolic potential of microorganisms in an ecosystem, not requiring laboratory isolation, could help to understand their relationships better and establish accurate strain-function linkages[23–25]. Further, the integration of a genome-scale network analysis based on metagenomics with metatranscriptomics could yield detailed insights into microbial synergism. To the best of our knowledge, no comprehensive reports on genome-resolved metatranscriptomics for microbial community in mangrove wetlands have been published to date, and the transcriptionally active interrelationships between different microbial functional groups remain elusive.

To address the above gaps, we performed in-depth shotgun DNA and RNA sequencing of five sediment samples from two sites of mangrove wetlands sites (vegetated mangrove and unvegetated mudflat) in the Mai Po Nature Reserve (Hong Kong SAR, China). We then used genome-scale metagenomics and metatranscriptomics analyses to recover high-quality genomes, reconstruct metabolic activities of major microbes in the mangrove sediments, and establish genome-scale co-occurrence and co-transcription networks to capture the keystone taxa. The presented investigation probes the microbial synergy mechanism from a genome-scale community-level view, and supports the notion that inter-microorganism interactions are vital modulators of the biogeochemical cycles of C, nitrogen (N), and sulfur (S) in mangrove ecosystems.

Results And Discussion

Reconstruction of metagenome-assembled genomes (MAGs) and key active metabolic modules in the mangrove sediments. To enable a genome-scale co-occurrence and co-transcription network analysis, we sampled the mangrove and intertidal mudflat sediments in Mai Po Nature Reserve (Hong Kong) at depth intervals of 0–2 cm, 10–15 cm, and 20–25 cm for the former (metagenomes MP7, MP8, and MP9), and at 0–5 cm and 13–16 cm for the latter (metagenomes MP10 and MP11). We reconstructed 671 curated, non-redundant MAGs, 67 archaeal and 604 bacterial MAGs, representing 7 archaeal and 43 bacterial phylum-level lineages from these five metagenomes. Overall, we mapped 0.24 billion reads (14.77% of the five metagenomes) to the MAGs; 338MAGs were

estimated to be ≥ 70% complete, 303 were estimated to be ≥ 80% complete, and 119were estimated to be ≥ 90% complete, with minimal domain-specific single-copy marker gene duplications (SupplementaryTable 1).We also assembled the reads mapped to the SILVA132 database and calculated the abundance of these assembled 16S rRNA gene sequences. We compared the normalized genome abundance (genome copies per million reads) with normalized abundances of the assembled 16S rRNA gene sequences at the phylum level. The R^2 (linear regression) for the correlation in each sample was higher than 0.84 (SupplementaryFig. 1), showing that the two parameters were highly correlated. This indicated that the reconstructed MAGs could represent the microbial community in these samples.

The majority of bacterial MAGs belonged to the lineages Gammaproteobacteria (n = 97), Chloroflexota (n = 78), Desulfobacterota (n = 77), Acidobacteriota (n = 63), and Bacteroidota (n = 55). The archaeal phyla of Asgard archaea (n = 18), Crenarchaeota (n = 15), Nanoarchaeota (n = 13), Euryarchaeota (n = 12), and Micrarchaeota (n = 7) were also identified. We profiled all Prodigal-predicted gene sequences from the MAGs using various databases (KEGG, TIGRfam, Pfam, and custom HMM databases) to determine their genome-scale metabolic capacities and transcript activities associated with C, N, and S metabolic pathways (Fig. 1, Supplementary Fig. 2). We considered the genome-resolved phylum/class level to establish accurate linkages between the microbial lineages and their biogeochemical activities (Supplementary Fig. 3). Gammaproteobacteria and Desulfobacterota were the two dominant lineages with the most transcripts in mangrove and mudflat sediments. Gammaproteobacteria were substantially active in various cycles of C, N (e.g. denitrification), and S (e.g. sulfate reduction) in the surface sediment. They were also prominently active in C fixation, denitrification, and methane oxidation in the subsurface layer. Desulfobacterota were mainly responsible for sulfite reduction to sulfide, both in the surface and subsurface sediments. Similarly, a recent functional gene-based metagenomics study revealed Gammaproteobacteria and Desulfobacterota as predominant microbial groups, implicated in the coupling of C, N, and S cycles[26].

Genome-scale co-occurrence and co-transcription networks of the C, N, and S metabolic modules in the mangrove sediments. We inferred genome-scale co-transcription networks based on significant correlations (Spearman's rank correlation, $P < 0.001$ and $R^2 > 0.8$) between the abundances of C, N, and S functional transcripts in pairwise MAGs. We then compared them with genome-scale co-occurrence networks inthe five sediment samples (Fig. 2a, 2b). The undirected networks captured 32,305 co-occurrence and 23,888 co-transcription associations among 664 microbial MAGs (Fig. 2, Table 1). We then merged and collapsed the networks at phylum/class level to determine whether these microbial lineages were interconnected (Supplementary Fig. 4). Consequently, 55.02% of network edges were identified as two microbial groups with significant associations of both metabolic capacity and transcriptional activity, while 31.39% and 13.59% of edges indicated paired microorganisms with significant connections of either type, accordingly. This indicated that more than half of the microbial lineages in the sampled ecosystem are metabolically and transcriptionally interconnected.

All generated co-occurrence networks exhibited scale-free characteristics, as indicated by the R^2 of power-law distribution of degree centrality, which ranged from 0.67 to 0.75 (Table 1); by contrast, the co-transcription networks were not scale-free (R^2 of power-law from 0.06 to 0.26, Table 1). These metrics indicated that the co-occurrence networks were nonrandomly structured, implying that the C, N, and S metabolic capacities of the microbiome responded differentially to the changing environment[21]. However, all the co-transcription network structures appeared to be random. It is reasonable to expect that the co-transcription network is a snapshot of microbiome activity, illustrating a response to changes in the environment. The shape of the degree distribution may affect the stability of architecture among different types of networks (e.g. biological, physical, and social networks)[27,28]; however, the universality of the scale-free network remains controversial[29,30]. Implementation of this analytical framework to time-series data could provide additional insight for understanding the unique characteristics, origins, evolutionary mechanisms, and dynamics of genome-scale co-occurrence and co-transcription networks.

We then examined whether the co-occurrence and co-transcription networks had unique topological features at MAG level. We focused on the betweenness centrality, which determines the number of shortest paths that pass through a given MAG, as a proxy for the location of the MAG in relation to others[28,31,32]. High betweenness centrality values indicate a core location of the MAG in the network, whereas low betweenness centrality values indicate a peripheral location. We observed significantly higher betweenness centrality values in the co-transcription networks than in the co-occurrence networks ($P = 4.7e\text{-}5$, Wilcox rank-sum test, Fig. 3, Supplementary Fig. 5a). The degree centrality value, an additional topological feature that takes into account only the immediate MAG neighbourhood, exhibited a similar significant tendency (Fig. 3, Supplementary Fig. 5b). Partitioning the MAGs using different grouping strategies did not result in significant differences in the betweenness centrality values of the co-occurrence networks in mangrove and mudflat sediments ($P = 0.3$, Wilcox rank-sum test, Supplementary Fig. 5c, 5d); however, the betweenness centrality value of the co-transcription network in mangrove sediments was significantly higher than that in mudflat ($P = 0.0087$, Wilcox rank-sum test, Supplementary Fig. 5e). Within mangrove sediments, the betweenness centrality value of co-transcription networks was significantly higher in the subsurface layer than in surface sediments ($P = 1.1e\text{-}13$, Wilcox rank-sum test, Supplementary Fig. 5i), with no significant difference observed in the co-occurrence networks ($P = 0.93$, Wilcox rank-sum test, Supplementary Fig. 5g). On the other hand, there was no significant difference between surface and subsurface mudflat sediments in terms of betweenness centrality values of the co-occurrence ($P = 0.054$, Wilcox rank-sum test, Supplementary Fig. 5k) and co-transcription networks ($P = 0.036$, Wilcox rank-sum test, Supplementary Fig. 5m). These findings indicated that microorganisms in the subsurface mangrove sediments had stronger interrelationships and were more actively interconnected than those in the mangrove surface and mudflat sediments.

Putative keystone taxa and their transcriptional abundances of C, N, and S metabolic modules in the mangrove sediments. To identify putative keystone taxa in the networks, we first separated the co-occurrence and co-transcription networks into modules, i.e. groups of MAGs that are highly connected within a group, with few connections outside the group[33], and are usually considered functional units in a biological system[34]. The five co-occurrence networks and five co-transcription networks inspected were modular, with a significantly higher modularity (M) than those of the corresponding randomly rewired networks (Table 1). We then classified MAGs into four categories based on their within-module connectivity (Zi) and among-module connectivity (Pi) values[28,31,32], as peripherals, connectors, module hubs, and network hubs (Supplementary Fig. 6). We considered both hubs and connectors as the putative keystone taxa in both co-occurrence and co-transcription networks. Overall, 443 MAGs were peripherals in the networks (Fig. 2). Myxobacterium MP_1.1789 was the only identified network hub in the co-occurrence network of surface mudflat sediment (MP10); it acted as a connector in the co-occurrence network of surface mangrove sediment (MP7). We identified more module hubs in the co-occurrence networks of subsurface sediments (12 ± 1) than in those of surface

sediments (4 ± 1). Further, except for six MAGs (five Desulfobacterota and one Methylomirabilota) in a surface mangrove sediment (MP7), we did not identify any module hubs in the co-transcription networks. These six module hubs were likely anaerobic autotrophs, based on the highly transcribed pathways of C fixation and sulfate reduction to sulfide (Fig. 2c).

We identified 57 MAGs as active keystone taxa across the five co-transcription networks, among which Desulfobacterota (n = 17), Chloroflexota (n = 9), and Acidobacteriota (n = 7) were the dominant microbial phyla. To the best of our knowledge, to date, Desulfobacteria had not been identified as a keystone taxon, either by computational inference or based on empirical evidence[14]. Our finding may be explained by the uniqueness of the mangrove ecosystem, since no other studies have focused on the keystone taxa in this biome. By contrast, Chloroflexota was reported as a keystone taxon in a groundwater hyporheic zone[35], whereas Acidobacteriota is commonly identified as a keystone taxon in various terrestrial biomes, e.g. woodlands[36,37], agricultural lands[37], grasslands[31,38], and plant-associated microbiota[39,40], but not in aquatic environments[14]. We also identified other bacterial groups, e.g. Burkholderiales and Rhizobiales, all reportedly prevalent in various terrestrial and aquatic ecosystems, as active keystone taxa in the analyzed samples (three Burkholderiales MAGs and one Rhizobiales MAG). Interestingly, we identified two Asgard archaeal MAGs (Thorarchaeotal MP_init.2226 and MP_5.633 in MP8) and one Euryarchaeotal MAG (Thermoplasmata, MP_5.1569 in MP9) as connectors in the co-transcription networks of subsurface mangrove sediments. The two Asgard archaea were transcriptionally active with respect to C fixation and fermentation, and Thermoplasmata were active in C fixation and sulfate reduction to sulfite. This was in agreement with our earlier studies where we reinforced the importance of these archaeal lineages in mangrove sediments[7,9,41]. It also highlighted the ecological roles of archaeal groups in biogeochemical cycles, which are frequently underestimated, in comparison with bacteria, in meta-community level investigations.

The relative genome abundances of the identified active keystone taxa ranged from less than 0.01% to 0.79% (Supplementary Table 2). Considering such low abundance, they thus could represent the 'rare biosphere'[42], which has a disproportionally high role in biogeochemical cycles[43]. The fundamental premise of keystone taxa is congruent with the rare taxa concept in that the abundance of a species is not the best determinant of its contribution to the community[42]. The roles of rare microorganisms in biogeochemical cycles have been investigated in the past[36]. As an example, simultaneous phylogenetic identification and quantitation of metabolic activities of single microbial cells in an oligotrophic lake revealed that *Chromatium okenii*, representing 0.3% of the total cell number, contributes more than 70% of the total uptake of C in the system[44]. Therefore, the relatively less abundant taxa can be as important as or more important than the abundant taxa in maintaining the microbial networks[32].

The analysis presented above adds another dimension to the commonly used gene-based co-occurrence network analysis and can be integrated with empirical evidence to identify microbial keystone taxa. Nevertheless, ecophysiological data, e.g. generated by culturomics[45] or microbiome-on-a-chip[46], are needed to verify the roles of the predicted keystone taxa in these ecosystems.

We identified the keystone MAGs that were most active in the co-transcription networks involved in sequential redox transformations, e.g. sulfate reduction to sulfite and sulfite reduction to sulfide, in surface sediments (both mangrove and mudflat sediments, Fig. 2c). On the other hand, in the subsurface sediments, keystone MAGs tended to cooperate by fulfilling individual steps of redox transformations of sulfate reduction to sulfide, although, based on the annotated gene sets, many MAGs were capable of performing sequential reactions. Meanwhile, no single keystone MAG was capable of performing the entire denitrification processes. These observations indicate that microorganisms have developed vertically stratified strategies of cooperation to fulfil sequential redox processes, in agreement with the notion that a highly interconnected co-transcription network in the mangrove subsurface sediments follows a scenario of metabolic handoffs[23].

***Candidatus* Mangrovidesulfobacterales, a new order in the class Deltaproteobacteria.** Among the putative keystone taxa in the co-transcription networks, one Desulfobacterota MAG (MP_5.128) formed a distinct clade with two peripheral MAGs (MP_5.1535 and MP_5.2156) within the class Deltaproteobacteria in a phylogenetic tree based on 122 concatenated bacterial marker genes (Fig. 4). The three MAGs branched with a recently proposed order *Ca.* Acidulodesulfobacterales (formerly Sva0485)[47] in the 16S rRNA gene phylogenetic tree. The averaged 16S rRNA gene sequence identity value between the two close groups was 85.19%, and the averaged nucleotide identity (OrthoANI) values between the three MAGs and four *Ca.* Acidulodesulfobacterales ranged from 59.64 to 61.27. We therefore propose that the three MAGs represent a new order in class Deltaproteobacteria and designate it as *Ca.* Mangrovidesulfobacterales (Supplementary Fig. 7).

The relative abundance of *Ca.* Mangrovidesulfobacterales was highest in subsurface mangrove and mudflat sediments, where pH value exceeds 7.31 (Supplementary Table 3), ranging from 0.44% to 0.81%. It was lowest in surface mangrove sediments (pH = 6.6), at 0.09%. Therefore, *Ca.* Mangrovidesulfobacterales are most likely anaerobes that do not preferentially dwell in the acidic environment, unlike the close relatives *Ca.* Acidulodesulfobacterales, which are mainly found in artificial acid mine drainage[47]. We investigated the global distribution pattern of the proposed new order by searching NCBI SRA databases for 16S rRNA sequences that were 97% identical with that of *Ca.* Mangrovidesulfobacterales MP_5.128. The analysis revealed that *Ca.* Mangrovidesulfobacterales are widely distributed in various habitats, e.g. terrestrial, freshwater, coastal, estuary, and deep-sea sediment ecosystems (Supplementary Fig. 8).

We also reconstructed the metabolic pathways of the three MAGs representing *Ca.* Mangrovidesulfobacterales. The analysis revealed the genetic potentials of C fixation via the reverse tricarboxylic acid cycle (rTCA) for autotrophs, glycolysis pathway, and non-oxidative pentose phosphate pathway for heterotrophs, indicating that they could be mixotrophic deltaproteobacteria. The transcriptomic analysis revealed that in the surface mudflat sediment, *Ca.* Mangrovidesulfobacterales actively fix C and subsequently produce acetate for ethanol fermentation. However, we did not detect transcripts for C fixation and glycolysis pathways in *Ca.* Mangrovidesulfobacterales genomes from subsurface mangrove sediments (Fig. 5). We also identified a complete dissimilatory sulfate reduction pathway in *Ca.* Mangrovidesulfobacterales genomes, suggesting that they are potential sulfate-reducing bacteria. Nevertheless, the *dsr*AB genes were not transcribed, and hence, *Ca.* Mangrovidesulfobacterales are only able to transform sulfate to sulfite (Fig. 5). This indicated that *Ca.* Mangrovidesulfobacterales might pass the end product, sulfite, to other sulfate-reducing bacteria, e.g. Desulfobacterota MP_init.1551 and Krumholzibacteriota

MP_5.1718, which were not transcriptionally active in the transformation of sulfate to sulfite but were active in the sulfite to sulfide transformation. Collectively, these observations demonstrated the metabolic versatility and transcriptional activity of *Ca.* Mangrovidesulfobacterales in the mangrove and mudflat sediments, revealing that these microorganisms are important participants in the biogeochemical cycles of C and S.

Conclusions

In the current study, in an effort to understand the mechanism of microbial synergy, we reconstructed hundreds of MAGs representing 50 distinct microbial lineages, from a typical coastal wetland ecosystem, and defined their metabolic capacities in the biogeochemical cycles of C, N, and S. We built co-occurrence and co-transcription networks of key metabolic processes of C, N, and S cycles for significantly correlated MAGs. Subsequently, we demonstrated that genome-scale network analysis can be used to describe the interrelationships between the microorganisms, and capture putative keystone taxa that actively participate in the biogeochemical cycles and maintain the topological structures of networks. We propose that one such keystone taxon represents a new deltaproteobacterial order, *Ca.* Mangrovidesulfobacterales. It is widely distributed in various habitats, and is likely anaerobic, mixotrophic, and capable of dissimilatory sulfate reduction. As a limitation, the keystone taxa inferred from network topology have not been confirmed by classical experimental validation; however, the study is the first initial screening of the potential keystone taxa in the enormously diverse and complex microbial community in the mangrove ecosystem. We are only beginning to fill a gap in knowledge on the unique microbial key functional assemblages of mangrove ecosystems. Understandings of their interrelationships will allow comprehensive documentation of microbial synergisms to underpin the roles of potentially underestimated microorganisms in biogeochemical cycles and mitigating climate change.

"Ca. Mangrovidesulfobacterium sediminis" (gen. nov., sp. nov.)

We propose the taxonomic epithets "*Ca.* Mangrovidesulfobacterium sediminis" (MP_5.128). The etymology and description are as follows: "Mangrovidesulfobacterium" (Man.gro.vi.de.sul.fo.bac.te'ri.um) N.L. neut. n. *Mangrovum*, mangrove; L. pref. *de*, from; L. n. *sulfur*, sulfur; N.L. pref. *desulfo-* (a dissimilatory sulfate-reducing bacterium); L. neut. n. *bacterium*, a rod; N.L. neut. n. *Mangrovidesulfobacterium*, a rod-shaped sulfate reducer in the mangrove ecosystem. The type species is "*Ca.* Mangrovidesulfobacterium sediminis". "*Ca.* Mangrovidesulfobacterium sediminis" (se.di'mi.nis) L. gen. n. *sediminis*, of sediment. The type material is the metagenomic bin MP_5.128.

Descriptions of Mangrovidesulfobacteraceae (fam. nov.) and Mangrovidesulfobacterales (ord. nov.)

In addition, phylogenomic analysis supports the proposal for a candidate family and order of "*Ca.* Mangrovidesulfobacterium" and "*Ca.* Mangrovidesulfobacteraceae" (Man.gro.vi.de.sul.fo.bac.ter.a'ce.ae) N.L. n. Mangrovidesulfobacterium type genus of the family; L. suff. -aceae ending to denote a family; N.L. neut. pl. n. Mangrovidesulfobacteraceae the family of the genus Mangrovidesulfobacterium. Similarly, a candidate order is proposed, "Mangrovidesulfobacterales" (Man.gro.vi.de.sul.fo.bac.ter.a'les) N.L. n. Mangrovidesulfobacterium type genus of the order; L. suff. -ales ending to denote an order; N.L. neut. pl. n. Mangrovidesulfobacterales the order of the genus Mangrovidesulfobacterium.

Methods

Sample collection and processing. Sediment samples were taken from a site covering mangrove forest (22°29.875'N, 114°01.767'E) and an intertidal mudflat (22°29.949'N, 114°01.656'E) from Mai Po Nature Reserve on September 12, 2014. The sediments were collected at depth intervals of 0-2, 10-15, 20-25 cm for mangrove (MP7, MP8, and MP9) and 0-5, 13-16 cm for intertidal mudflat (MP10 and MP11), respectively. The sampling strategy was described in detail elsewhere [7] and is stated here briefly. RNALater (Ambion, Life Technologies) was added immediately after collection and stored in a pre-cooled sampling box with ice cubes, then transported to the laboratory within < 4 h. We took two portions of 5 g of wet sediments stored at –20 °C for DNA and RNA extraction with the PowerSoil DNA Isolation Kit (MO BIO) and RNA Powersoil Total RNA Isolation Kit (QIAGEN), respectively, following the manufacturer's instructions. The ribosomal RNA genes were removed from the total RNA using the Ribo-Zero rRNA removal kit (Illumina, Inc., San Diego, CA, USA), and the remaining mRNA was reverse-transcribed. DNA and cDNA were sequenced using an Illumina HiSeq sequencer (Illumina) with 150-bp paired-end reads at BerryGenomics (Beijing, China). For each sample, the rest wet sediments were taken for physicochemical parameters measurement, including redox potential, pH, organic matter, and the concentrations of ammonium, nitrate, and nitrite in pore water via centrifugation from the sediments.

Metagenomic assembly and binning. We used "read_qc" module from metaWRAP (v.1.1)[48] to trim the raw shotgun sequencing metagenomic reads. All clean reads from five samples were pooled prior to *de novo* assemble to obtain one co-assembly. A total of 3,265,768,294 paired-end reads (150bp) were sent out to MEGAHIT (v1.1.2)[49] with flag "–presets meta-large" for co-assembling job. Sequencing coverage was determined for each assembled scaffold by mapping reads from each sample to the co-assembly using Bowtie2[50]. The binning analysis was carried out eight times by eight different combinations of specificity and sensitivity parameters using MetaBAT2[51] ("–maxP 60 or 95" AND "–minS 60 or 95" AND "–maxEdges 200 or 500") on the assembly with a minimum length of 2000 bp. DAS Tool[52] was used as a dereplication and aggregation strategy on those eight binning results to construct accurate bins. Manual curation was adapted for reducing the genome contamination based on differential coverages, GC contents, and the presence of duplicate genes. The completeness, contamination, and strain heterogeneity of the genomes were estimated by using CheckM (v.1.0.12)[53] and DAS Tool under the taxonomic scope of domain (i.e., Bacteria and Archaea).

Gene calling and metabolic annotation. We performed open reading frames (ORFs) prediction on genomic scaffolds using the single genome mode of Prodigal[54]. The Pfam[55], TIGRfam[56], KOfam[57], and a published in-house HMM database[23] were used to assign all predicted ORFs protein functions by using hmmsearch (sequence e-value < $1e^{-5}$ and domain e-value < $1e^{-4}$)[58]. The identifiers for genes involving specific metabolic pathways are listed in Supplementary Table 4. All the genes from 671 MAGs with such annotations are listed in Supplementary Table 5.

Genome coverage and functional gene abundance. Metatranscriptomic reads were quality-trimmed using Sickle (v1.33, https://github.com/najoshi/sickle), and the potential rRNA reads were removed using SortMeRNA (v2.0)[59] against both the SILVA132 database and the default databases (e-value <1e^{-5}). Metagenomic and metatranscriptomic reads were mapped against scaffolds from all MAGs using "quant_bins" module from metaWRAP, which uses Salmon[60] to estimate the abundance of each scaffold in each sample. The genome abundances are expressed as "genome copies per million reads" normalizing for differences in read counts between samples (Supplementary Table 2). The genomic and transcriptomic abundances of genes were expressed as "Transcript Per Million (TPM)" extracted from Salmon output. The abundances were averaged if a number of genes in one microorganism were annotated to multiple HMMs belonging to one metabolic module (Supplementary Table 6).

16S ribosomal RNA assembly. We used MATAM to reconstruct the full-length 16S rRNA sequences from reads and provide a taxonomic classification by RDP classifier against the SILVA132 database. The script "matam_compare_samples.py" was used to generate an abundance table of the assembled 16S rRNA sequences. The correlation coefficient (R^2) was calculated as a measure of linear association between the relative abundance of assembled 16S rRNA sequences and genome copies per million reads. The bin MP_5.128 as type material of *Ca*. Mangrovidesulfobacterium sediminis initially contained a partial 16S rRNA sequence with length of 532 bp, which had been extended by MATAM assembling the mapped reads to a 1272 bp length. The global alignment showed their similarity is 99.52% wit 2 bp mismatches.

Phylogenetic analyses. Phylogenetic analysis of the 671 MAGs was performed using a syntenic block of 15 single copy universal ribosomal proteins (L2, L3, L5, L6, L10, L11, L14b_L23e, L16_L10E, S2, S7, S10, S12_S23, S15P_S13e, S19). Each ribosomal protein was identified by using GraftM[61] and aligned using mafft-linsi[62] with default parameters. Individual ribosomal protein alignments were trimmed by BMGE (v.1.12) with "-m BLOSUM30 -g 0.2 -b 3" and concatenated before further analyses. In total, the concatenated alignment spanned 2,110 columns. Phylogenetic analysis of ribosomal protein was inferred by IQ-TREE[63] with the best-fit model of "LG4X" followed by 1000 times ultra-fast bootstrap approximation, and 1000 replicated SH-like approximate likelihood ratio tests. The phylogenetic tree was rooted with archaea, and visualized with figtree v.1.4.4 (http://tree.bio.ed.ac.uk/software/figtree/). Taxonomic assignments for bacterial and archaeal MAGs were given by running the genomes against GTDB-Tk[64] (GTDB r89) using FastANI[65]. The assignments were manually verified by phylogenetic trees of 120 bacterial/122 archaeal marker genes inferred by FastTree[66] with "-wag -gamma" flag on the alignments spanned 5040 columns and 5124 columns respectively. The 16S rRNA sequences were aligned using mafft-linsi and trimmed by trimAl[67] with "-gt 0.95 -cons 50", and associated phylogenetic analysis of *Ca*. Mangrovidesulfobacterales was inferred by IQ-TREE with best-fit model of "TIM3e+I+G4" followed by 1000 times ultra-fast bootstrap approximation, and 1000 replicated SH-like approximate likelihood ratio tests. The concatenation of 120 bacterial marker gene alignments was used to infer a phylogenetic tree of these deltaproteobacterial genomes by IQ-TREE with the best-fit model of "LG4X" followed by 1000 times ultra-fast bootstrap approximation, and 1000 replicated SH-like approximate likelihood ratio tests. The above two trees were rerooted by Nesterenkonia.

Network construction, module detection and identification of node roles. Networks were constructed for co-occurrence and co-transcription patterns of averaged TPM values for each of 22 metabolic modules of carbon, nitrogen and sulfur cycles for paired MAGs, yielding a total of ten networks. Random networks were generated using the Maslov-Sneppen procedure[68], which is to keep the numbers of nodes and edges unchanged but rewire all of the edges' positions in the networks, so the sizes of networks are the same and the random rewired networks are comparable to the original ones. A total of 100 permutations were generated for each empirical network. Network properties analysis and visualization was performed using {igraph} R packages[69]. We characterized network modularity for each network using seven cluster methods based on edge betweenness, fast greedy optimization, propagating labels, leading eigenvector, multi-level optimization, short random walks, and infomap algorithms from {igraph} R package, selecting the method with highest modularity value for final clustering. Infomap algorithm was used for MP7 co-occurrence network and MP8 co-transcription network, whereas fast greedy optimization algorithm was used for MP9 co-occurrence and co-transcription networks. Multi-level optimization was used for all the rest networks. The connectivity of each node was determined based on its within-module connectivity (*Zi*) and among-module connectivity (*Pi*)[34], which were then used to classify the nodes based on the topological roles they play in the network (Supplementary Table 7). Node topologies are organized into four categories: module hubs (highly connected nodes within modules, Zi > 2.5), network hubs (highly connected nodes within entire network, Zi > 2.5 and Pi > 0.62), connectors (nodes that connect modules, Pi > 0.62) and peripherals (nodes connected in modules with few outside connections, Zi < 2.5 and Pi < 0.62) [28,31,34].

Statistical analyses. All statistical analyses were performed by {stats} R packages.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and materials

The five metagenomic and five metatranscriptomic raw sequence reads for (MP7, MP8, MP9, MP10 and MP11), and metagenome-assembled genomes have been deposited in NCBI database under the project PRJNA360036.

Competing interests

The authors declare that they have no competing interests.

Funding

This research was financed by the National Natural Science Foundation of China (No. 91851105, 31622002, 31970105, 31600093, and 31700430), the Shenzhen Science and Technology Program (grant no. JCYJ20170818091727570 and KQTD20180412181334790), the Key Project of Department of Education of Guangdong Province (No. 2017KZDXM071).

Authors' contributions

Y.L., Z.Z., and M.L. conceived this study. Y.L., Z.Z., and Y.Y. collected samples and generated the metagenomics and metatranscriptomics data. Y.L. performed co-assembling, binning, genome reconstruction, phylogenetic, and network analyses. Y.L. and Z.Z. deposited the raw sequence data. Y.L., Z.Z., and M.L. wrote the first draft, and all authors wrote on the basis of the draft and approved the manuscript.

Acknowledgments

We thank Prof. Ji-Dong Gu of University of Hong Kong for his support to this project.

References

1. Hamilton SE, Casey D. Creation of a high spatio-temporal resolution global database of continuous mangrove forest cover for the 21st century (CGMFC-21). Global Ecology and Biogeography. 25:729–38.

2. Duarte CM, Losada IJ, Hendriks IE, Mazarrasa I, Marbà N. The role of coastal plant communities for climate change mitigation and adaptation. Nature Climate Change. 2013;3:961–8.

3. Alongi DM. Carbon sequestration in mangrove forests. Carbon Management. 2012;3:313–22.

4. Donato DC, Kauffman JB, Murdiyarso D, Kurnianto S, Stidham M, Kanninen M. Mangroves among the most carbon-rich forests in the tropics. Nature Geoscience. 2011;4:293–7.

5. Atwood TB, Connolly RM, Almahasheer H, Carnell PE, Duarte CM, Lewis CJE, et al. Global patterns in mangrove soil carbon stocks and losses. Nature Climate Change. 2017;7:523–8.

6. Siikamäki J, Sanchirico JN, Jardine SL. Global economic potential for reducing carbon dioxide emissions from mangrove loss. PNAS. 2012;109:14369–74.

7. Liu Y, Zhou Z, Pan J, Baker BJ, Gu J-D, Li M. Comparative genomic inference suggests mixotrophic lifestyle for Thorarchaeota. The ISME Journal. 2018;541:353.

8. Zhou Z, Meng H, Liu Y, Gu J-D, Li M. Stratified Bacterial and Archaeal Community in Mangrove and Intertidal Wetland Mudflats Revealed by High Throughput 16S rRNA Gene Sequencing. Frontiers in Microbiology. 2017;8:195.

9. Zhou Z, Liu Y, Lloyd KG, Pan J, Yang Y, Gu J-D, et al. Genomic and transcriptomic insights into the ecology and metabolism of benthic archaeal cosmopolitan, Thermoprofundales (MBG-D archaea). The ISME Journal. 2019;13:885.

10. Pan J, Chen Y, Wang Y, Zhou Z, Li M. Vertical Distribution of Bathyarchaeotal Communities in Mangrove Wetlands Suggests Distinct Niche Preference of Bathyarchaeota Subgroup 6. Microb Ecol. 2019;77:417–28.

11. Pan J, Zhou Z, Béjà O, Cai M, Yang Y, Liu Y, et al. Genomic and transcriptomic evidence of light-sensing, porphyrin biosynthesis, Calvin-Benson-Bassham cycle, and urea production in Bathyarchaeota. Microbiome. 2020;8:43.

12. Zhang C-J, Pan J, Duan C-H, Wang Y-M, Liu Y, Sun J, et al. Prokaryotic Diversity in Mangrove Sediments across Southeastern China Fundamentally Differs from That in Other Biomes. mSystems [Internet]. American Society for Microbiology Journals; 2019 [cited 2020 Mar 19];4. Available from: https://msystems.asm.org/content/4/5/e00442-19

13. Layeghifard M, Hwang DM, Guttman DS. Disentangling Interactions in the Microbiome: A Network Perspective. Trends in Microbiology. 2017;25:217–28.

14. Banerjee S, Schlaeppi K, van der Heijden MGA. Keystone taxa as drivers of microbiome structure and functioning. Nat Rev Microbiol. 2018;16:567–76.

15. Power ME, Tilman D, Estes JA, Menge BA, Bond WJ, Mills LS, et al. Challenges in the Quest for Keystones. BioScience. Oxford Academic; 1996;46:609–20.

16. Fuhrman JA. Microbial community structure and its functional implications. Nature. Nature Publishing Group; 2009;459:193–9.

17. Faust K, Raes J. Microbial interactions: from networks to models. Nature Reviews Microbiology. 2012;10:538–550.

18. van der Heijden MGA, Hartmann M. Networking in the Plant Microbiome. PLoS Biol. 2016;14:e1002378.

19. Barabási A-L, Gulbahce N, Loscalzo J. Network medicine: a network-based approach to human disease. Nat Rev Genet. Nature Publishing Group; 2011;12:56–68.

20. Berry D, Widder S. Deciphering microbial interactions and detecting keystone species with co-occurrence networks. Front Microbiol. 2014;5:219.

21. Vick-Majors TJ, Priscu JC, Amaral-Zettler LA. Modular community structure suggests metabolic plasticity during the transition to polar night in ice-covered Antarctic lakes. ISME J. 2014;8:778–89.

22. Röttjers L, Faust K. From hairballs to hypotheses–biological insights from microbial networks. FEMS Microbiol Rev. 2018;42:761–80.

23. Anantharaman K, Brown CT, Hug LA, Sharon I, Castelle CJ, Probst AJ, et al. Thousands of microbial genomes shed light on interconnected biogeochemical processes in an aquifer system. Nature Communications. 2016;7:13219.

24. Baker BJ, Lazar CS, Teske AP, Dick GJ. Genomic resolution of linkages in carbon, nitrogen, and sulfur cycling among widespread estuary sediment bacteria. Microbiome. 2015;3:305.

25. Dombrowski N, Seitz KW, Teske AP, Baker BJ. Genomic insights into potential interdependencies in microbial hydrocarbon and nutrient cycling in hydrothermal sediments. Microbiome. 2017;5:106.

26. Lin X, Hetharua B, Lin L, Xu H, Zheng T, He Z, et al. Mangrove Sediment Microbiome: Adaptive Microbial Assemblages and Their Routed Biogeochemical Processes in Yunxiao Mangrove National Nature Reserve, China. Microb Ecol. 2019;78:57–69.

27. Barabási A-L, Oltvai ZN. Network biology: understanding the cell's functional organization. Nat Rev Genet. Nature Publishing Group; 2004;5:101–13.

28. Zhou J, Deng Y, Luo F, He Z, Tu Q, Zhi X. Functional Molecular Ecological Networks. mBio [Internet]. American Society for Microbiology; 2010 [cited 2020 Mar 6];1. Available from: https://mbio.asm.org/content/1/4/e00169-10

29. Broido AD, Clauset A. Scale-free networks are rare. Nat Commun. Nature Publishing Group; 2019;10:1–10.

30. Holme P. Rare and everywhere: Perspectives on scale-free networks. Nat Commun. Nature Publishing Group; 2019;10:1–3.

31. Deng Y, Jiang Y-H, Yang Y, He Z, Luo F, Zhou J. Molecular ecological network analyses. BMC Bioinformatics. 2012;13:113.

32. Shi S, Nuccio EE, Shi ZJ, He Z, Zhou J, Firestone MK. The interconnected rhizosphere: High network complexity dominates rhizosphere assemblages. Ecology Letters. 2016;19:926–36.

33. Newman MEJ. Modularity and community structure in networks. Proceedings of the National Academy of Sciences. 2006;103:8577–82.

34. Olesen JM, Bascompte J, Dupont YL, Jordano P. The modularity of pollination networks. Proc Natl Acad Sci USA. 2007;104:19891–6.

35. Graham EB, Crump AR, Resch CT, Fansler S, Arntzen E, Kennedy DW, et al. Deterministic influences exceed dispersal effects on hydrologically-connected microbiomes. Environ Microbiol. 2017;19:1552–67.

36. Lupatini M, Suleiman AKA, Jacques RJS, Antoniolli ZI, de Siqueira Ferreira A, Kuramae EE, et al. Network topology reveals high connectance levels and few key microbial genera within soils. Front Environ Sci [Internet]. Frontiers; 2014 [cited 2020 Feb 25];2. Available from: https://www.frontiersin.org/articles/10.3389/fenvs.2014.00010/full

37. Banerjee S, Kirkby CA, Schmutter D, Bissett A, Kirkegaard JA, Richardson AE. Network analysis reveals functional redundancy and keystone taxa amongst bacterial and fungal communities during organic matter decomposition in an arable soil. Soil Biology and Biochemistry. 2016;97:188–98.

38. Zhou J, Deng Y, Luo F, He Z, Yang Y. Phylogenetic molecular ecological network of soil microbial communities in response to elevated CO2. mBio. 2011;2.

39. Jiang Y, Li S, Li R, Zhang J, Liu Y, Lv L, et al. Plant cultivars imprint the rhizosphere bacterial community composition and association networks. Soil Biology and Biochemistry. 2017;109:145–55.

40. Yan Y, Kuramae EE, de Hollander M, Klinkhamer PGL, van Veen JA. Functional traits dominate the diversity-related selection of bacterial communities in the rhizosphere. ISME J. 2017;11:56–66.

41. Cai M, Liu Y, Yin X, Zhou Z, Friedrich MW, Richter-Heitmann T, et al. Diverse Asgard archaea including the novel phylum Gerdarchaeota participate in organic matter degradation. Sci China Life Sci [Internet]. Science China Press; 2020 [cited 2020 Mar 17]; Available from: /doi/10.1007/s11427-020-1679-1

42. Lynch MDJ, Neufeld JD. Ecology and exploration of the rare biosphere. Nature Reviews Microbiology. 2015;13:217–29.

43. Jousset A, Bienhold C, Chatzinotas A, Gallien L, Gobet A, Kurm V, et al. Where less may be more: how the rare biosphere pulls ecosystems strings. ISME J. 2017;11:853–62.

44. Musat N, Halm H, Winterholler B, Hoppe P, Peduzzi S, Hillion F, et al. A single-cell view on the ecophysiology of anaerobic phototrophic bacteria. PNAS. National Academy of Sciences; 2008;105:17861–6.

45. Lagier J-C, Khelaifia S, Alou MT, Ndongo S, Dione N, Hugon P, et al. Culture of previously uncultured members of the human gut microbiota by culturomics. Nature Microbiology. 2016;1:16203.

46. Stanley CE, Heijden MGA van der. Microbiome-on-a-Chip: New Frontiers in Plant–Microbiota Research. Trends in Microbiology. Elsevier; 2017;25:610–3.

47. Tan S, Liu J, Fang Y, Hedlund BP, Lian Z-H, Huang L-Y, et al. Insights into ecological role of a new deltaproteobacterial order Candidatus Acidulodesulfobacterales by metagenomics and metatranscriptomics. ISME J. Nature Publishing Group; 2019;13:2044–57.

48. Uritskiy GV, DiRuggiero J, Taylor J. MetaWRAP—a flexible pipeline for genome-resolved metagenomic data analysis. Microbiome. 2018;6:158.

49. Li D, Liu CM, Luo R, Sadakane K, Lam TW. MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. Bioinformatics. 2015;31:1674–1676.

50. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods. Nature Publishing Group; 2012;9:357–9.

51. Kang DD, Li F, Kirton E, Thomas A, Egan R, An H, et al. MetaBAT 2: an adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies. PeerJ. 2019;7:e7359.

52. Sieber CMK, Probst AJ, Sharrar A, Thomas BC, Hess M, Tringe SG, et al. Recovery of genomes from metagenomes via a dereplication, aggregation and scoring strategy. Nat Microbiol. Nature Publishing Group; 2018;3:836–43.

53. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. Genome Research. 2015;25:1043–1055.

54. Hyatt D, Chen G-L, LoCascio PF, Land ML, Larimer FW, Hauser LJ. Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC Bioinformatics. 2010;11:119.

55. Finn RD, Bateman A, Clements J, Coggill P, Eberhardt RY, Eddy SR, et al. Pfam: the protein families database. Nucleic Acids Res. 2014;42:D222–30.

56. Haft DH, Selengut JD, White O. The TIGRFAMs database of protein families. Nucleic Acids Res. 2003;31:371–3.

57. Aramaki T, Blanc-Mathieu R, Endo H, Ohkubo K, Kanehisa M, Goto S, et al. KofamKOALA: KEGG ortholog assignment based on profile HMM and adaptive score threshold. bioRxiv. 2019;602110.

58. Eddy SR. Accelerated Profile HMM Searches. PLOS Computational Biology. Public Library of Science; 2011;7:e1002195.

59. Kopylova E, Noé L, Touzet H. SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. Bioinformatics. 2012;28:3211–7.

60. Patro R, Duggal G, Love MI, Irizarry RA, Kingsford C. Salmon: fast and bias-aware quantification of transcript expression using dual-phase inference. Nat Methods. 2017;14:417–9.

61. Boyd JA, Woodcroft BJ, Tyson GW. GraftM: a tool for scalable, phylogenetically informed classification of genes within metagenomes. Nucleic Acids Res. Oxford Academic; 2018;46:e59–e59.

62. Katoh K, Standley DM. MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. Molecular Biology and Evolution. 2013;30:772–780.

63. Nguyen L-T, Schmidt HA, von Haeseler A, Minh BQ. IQ-TREE: A Fast and Effective Stochastic Algorithm for Estimating Maximum-Likelihood Phylogenies. Molecular Biology and Evolution. 2014;32:268–274.

64. Chaumeil P-A, Mussig AJ, Hugenholtz P, Parks DH. GTDB-Tk: a toolkit to classify genomes with the Genome Taxonomy Database. Bioinformatics [Internet]. [cited 2020 Feb 21]; Available from: https://academic.oup.com/bioinformatics/advance-article/doi/10.1093/bioinformatics/btz848/5626182

65. Jain C, Rodriguez-R LM, Phillippy AM, Konstantinidis KT, Aluru S. High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. Nat Commun. Nature Publishing Group; 2018;9:1–8.

66. Price MN, Dehal PS, Arkin AP. FastTree 2 – Approximately Maximum-Likelihood Trees for Large Alignments. PLOS ONE. 2010;5:e9490.

67. Capella-Gutierrez S, Silla-Martinez JM, Gabaldon T. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics. 2009;25:1972–1973.

68. Clauset A, Newman MEJ, Moore C. Finding community structure in very large networks. Phys Rev E. American Physical Society; 2004;70:066111.

69. Csardi G, Nepusz T. The igraph software package for complex network research. InterJournal. 2006;Complex Systems:1695.

Tables

Table

Table 1 Major topological properties of the co-occurrence and co-transcription networks and their associated random networks

		Empirical networks										Random ne permutatio
Sample	Network type	MAG No.	Edge No.	R^2 of power law	Average connectivity	Average path length	Modules No.	Modularity	Average clustering coefficient	Average path length	Average clustering coefficient	Mo No.
MP7	Co-occurrence	594	5155	0.748	17.36	3.39	45	0.46	0.38	2.552±0.002	0.05±0.002	1.0
MP7	Co-transcription	389	2150	0.257	11.05	4.68	28	0.78	0.74	2.754±0.006	0.05±0.004	9.4
MP8	Co-occurrence	616	7357	0.67	23.89	3.21	17	0.48	0.41	2.353±0.002	0.061±0.001	9.4
MP8	Co-transcription	410	5711	0.078	27.86	3.91	28	0.61	0.66	2.111±0.003	0.093±0.002	1±0
MP9	Co-occurrence	611	7618	0.669	24.94	3.12	17	0.47	0.41	2.321±0.002	0.063±0.001	5.4
MP9	Co-transcription	346	5259	0.049	30.40	3.58	15	0.62	0.75	2.007±0.002	0.111±0.002	4.5
MP10	Co-occurrence	596	5860	0.746	19.66	3.33	21	0.48	0.40	2.471±0.002	0.054±0.001	9.7
MP10	Co-transcription	445	5272	0.234	23.69	4.02	23	0.65	0.68	2.245±0.003	0.079±0.002	7.9
MP11	Co-occurrence	594	6315	0.732	21.26	3.33	14	0.48	0.41	2.418±0.002	0.057±0.002	9.5
MP11	Co-transcription	370	5496	0.061	29.71	3.47	18	0.62	0.68	2.041±0.003	0.106±0.002	7.8

Figures

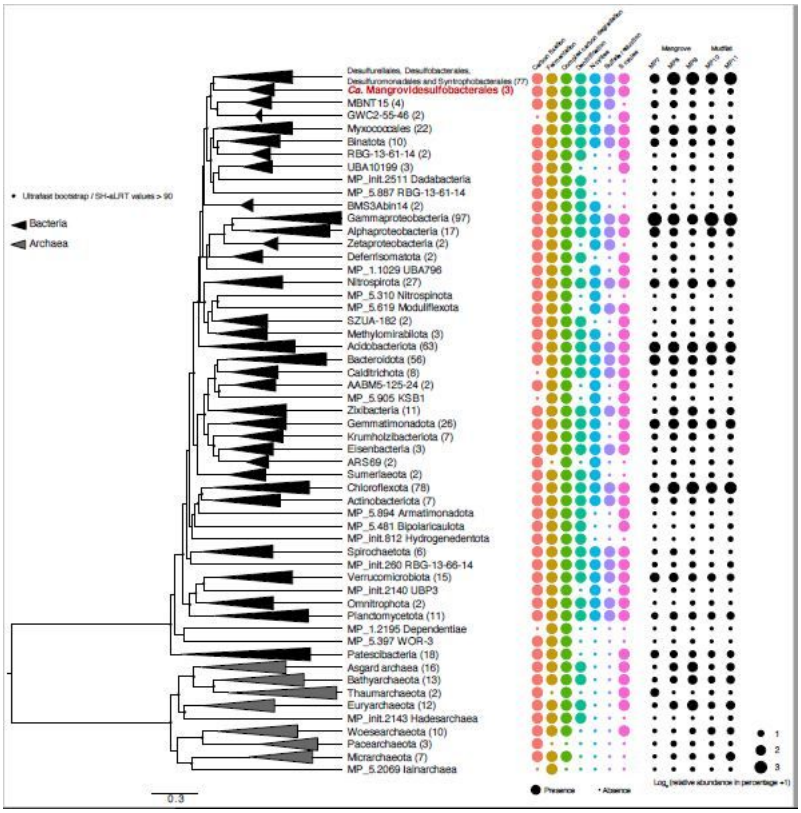

Figure 1

Phylogeny of bacterial and archaeal genomes inferred by maximum likelihood, and their metabolic modules and relative abundance across five metagenomes. The phylogenetic tree is based on 15 concatenated ribosomal proteins and was collapsed at the phylum/class level according to GTDB taxonomy with numbers of the genome in the bracket. Only genomes from this study are shown on this tree. The metabolic modules are shown in rainbow colors, and circles/dot represent presence/absence, respectively. The relative abundance of the sum of genomes from a branch is shown in different sizes of black circles.

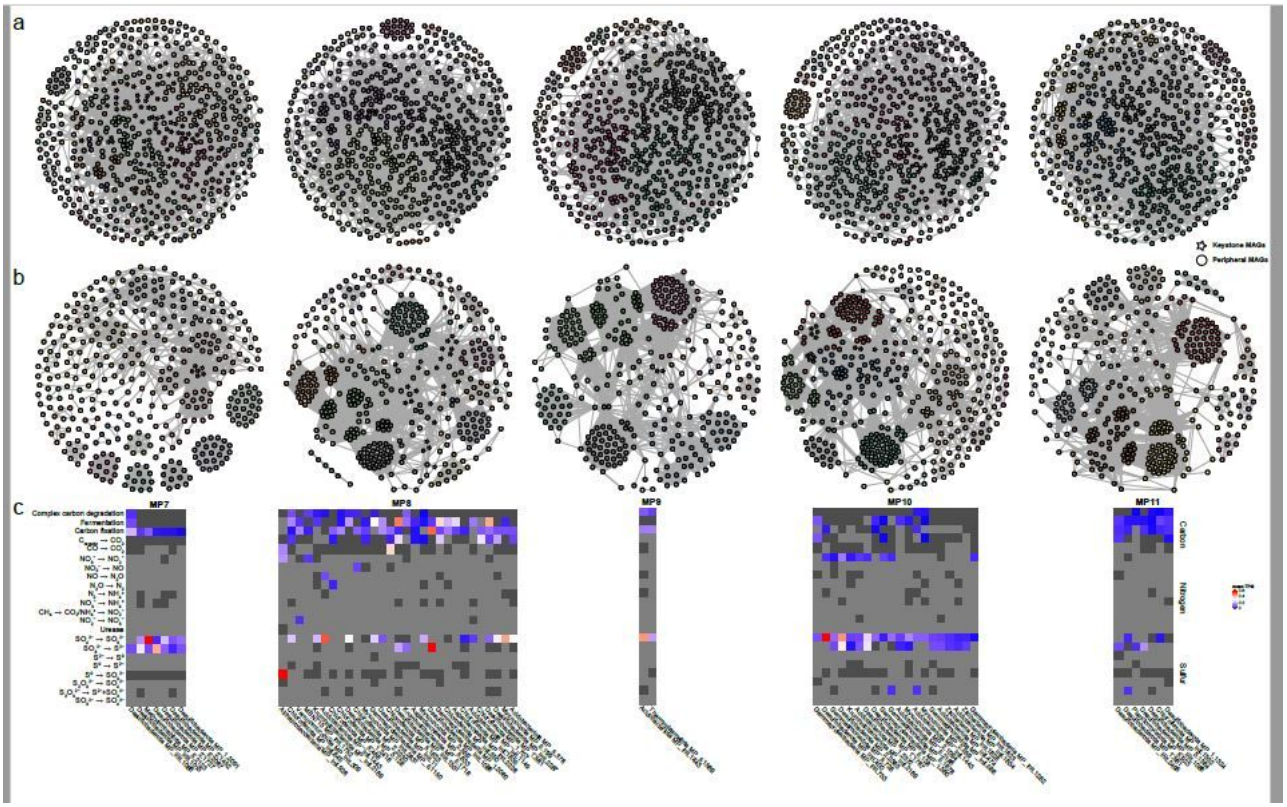

Figure 2

Genome-scale co-occurrence and co-transcription networks of the five sediment samples and the transcript abundances of 22 metabolic modules of carbon, nitrogen and sulfur cycles of active keystone taxa. Edges on the Co-occurrence (a) and co-transcription (b) networks represent significant correlations between paired MAGs (Spearman correlation test, R2 > 0.8). Nodes represent MAG. Modules are randomly colored. Circle-shaped nodes are peripheral MAGs, and star-shaped nodes are keystone MAGs. The transcript abundances are expressed as averaged TPM values.

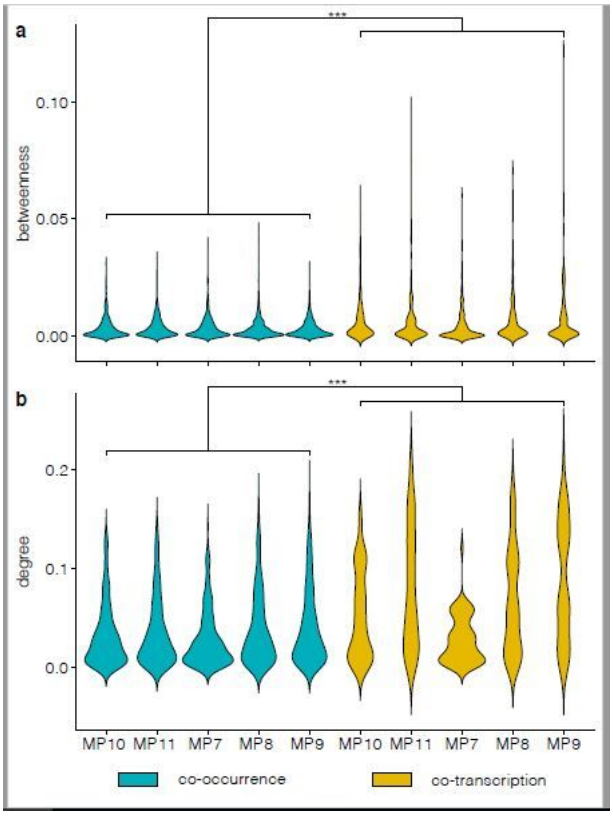

Figure 3

Betweenness (a) and degree (b) centralization associated with different co-occurrence and co-transcription networks. Asterisk (***) indicates significant difference P < 0.001 from Wilcox rank-sum test.

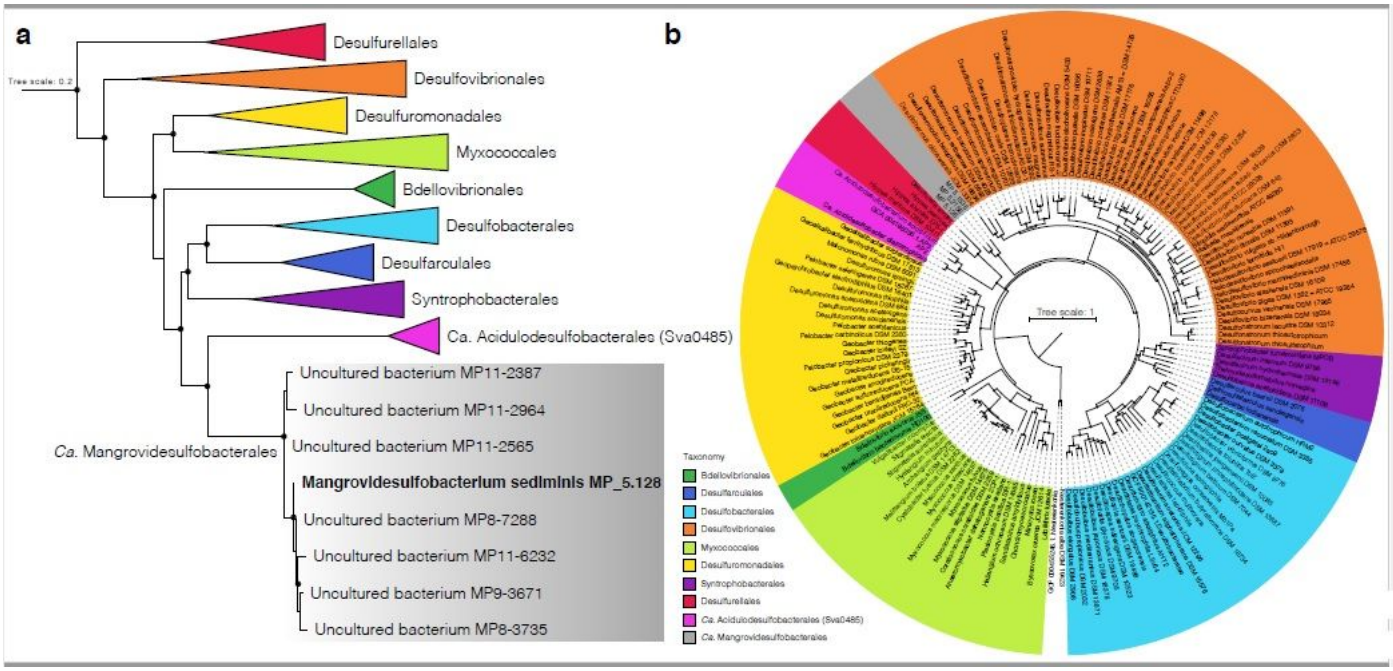

Figure 4

Phylogenetic placement of representatives of the class Ca. Mangrovidesulfobacterales. (a) a maximum-likelihood phylogenetic tree of 16S rRNA genes. (b) A maximum-likelihood phylogenetic tree was constructed using the concatenated alignment of 120 bacterial marker genes. Bootstrap values were based on, and percentages >80% are shown with black circles.

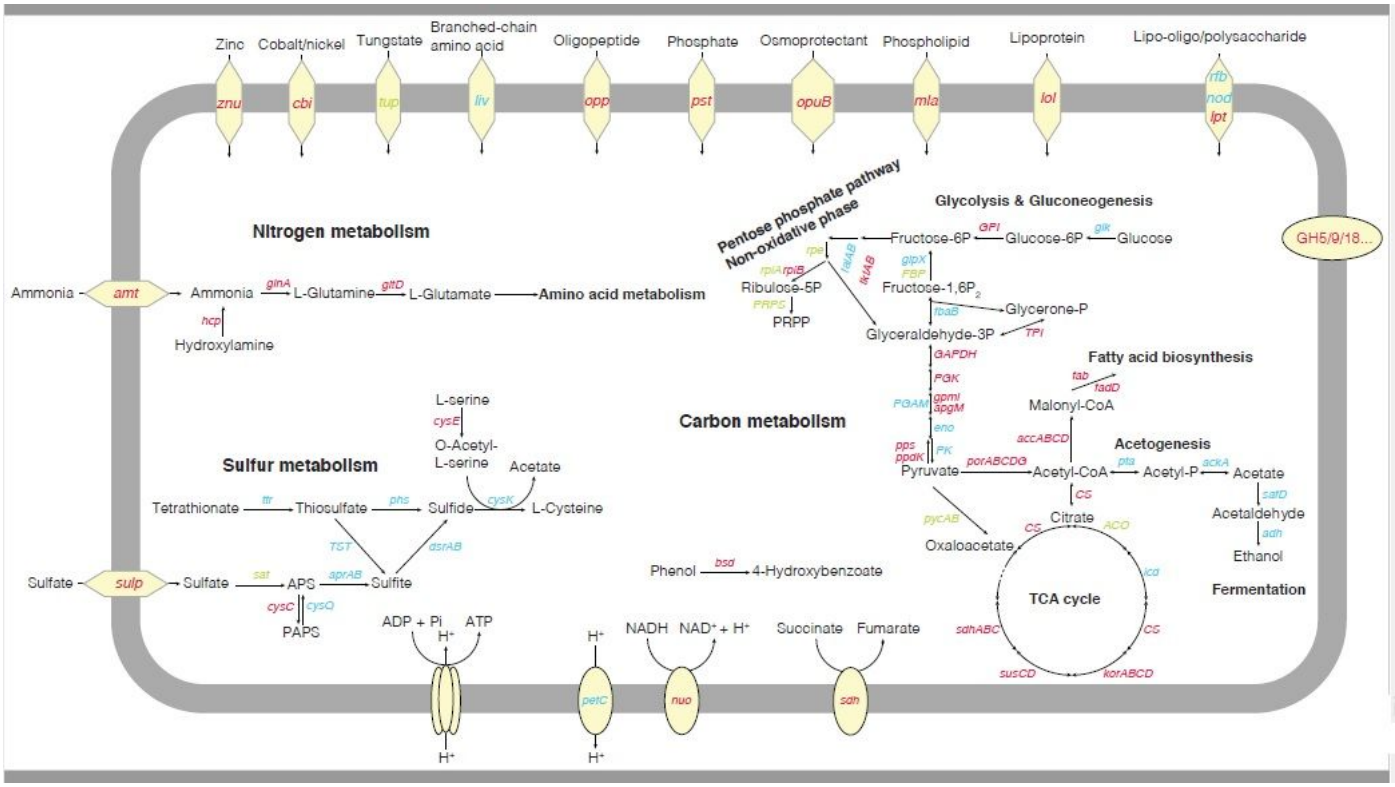

Figure 5

Metabolic capabilities of Ca. Mangrovidesulfobacterales. Pathways associated with carbon, nitrogen and sulfur metabolism are shown. Genes appearing in all three genomes, two genomes and one genome are shown in red, cyan and green, respectively.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryData.xlsx
- SupplementaryInformation.docx