MISSIONE 4
ISTRUZIONE
RICERCA

Hydrothermal versus microbial MEthane release From very shallow coastal systems: can differently sourced emISsions direcTly escape intO the atmosphere? (MEFISTO)

Deliverable 5. Methanotrophic communities - *Final report on the activities related to WP4*

Fonti V., Iacuzzo F., Caruso C., De Rosa G., De Vittor C., Esposito. V., Graziano M., Lazzaro G., Longo M., Bazzaro M.

Finanziato dall'Unione europea NextGenerationEU
Ministero dell'Università e della Ricerca
Italiadomani PIANO NAZIONALE DI RIPRESA E RESILIENZA
OGS Istituto Nazionale di Oceanografia e di Geofisica Sperimentale

MISSIONE 4
ISTRUZIONE
RICERCA

Table of Content

1. SUMMARY

Within the framework of the MEFISTO project, microbial communities associated with methane (CH_4) emissions were characterized at the Panarea hydrothermal vent area and the Northern Adriatic "Trezze". A total of 140 samples, encompassing seawater, hydrothermal fluids, interstitial water, sediments, and various biofilm structures, were collected and processed to define the taxonomic composition and community structure of the microbial community with a special focus on methanotrophic microorganisms. The methodological approach was based on a high-throughput metabarcoding strategy targeting the V4-V5 region of the 16S rRNA gene, which showed an optimal coverage for methanotrophs in both the Bacteria and Archaea domains. This report summarizes the sample management procedures, laboratory workflows, and bioinformatic strategies implemented to characterize the methanotrophic guilds across the investigated marine areas.

2. SCIENTIFIC FRAMEWORK

2.1 Biological Methane Oxidation and Marine Methanotrophs

In marine environments, methane-oxidizing microorganisms (a.k.a. methanotrophs) serve as biological sink for methane, acting like "biological filters" that prevent methane degassing from the seafloor to the atmosphere (Heintz et al., 2012). In marine sediments, these microorganisms can consume more than 90% of the methane produced locally before it can reach the water column (Knittel and Boetius, 2009; Nazaries et al., 2011). However, estimates support that methanotrophic bacteria may contribute to the oxidation of up to 50% of the dissolved CH_4 in the water column (Mao et al., 2022), playing a relevant role in regulating the Earth's atmospheric composition.
Methanotrophic microorganisms can grow on CH_4 as a sole carbon and energy source. Microbial methane consumption is an ubiquitous ecosystemic function (Guerrero-Cruz et al., 2021; Reis et al., 2022). Methanotrophs represent a widespread and polyphyletic functional guild that encompasses the two domains of prokaryotes: aerobic methanotrophic bacteria, which utilize oxygen to oxidize methane via methane monooxygenase (MMO) enzyme, and anaerobic methanotrophic archaea, which typically drive methane oxidation through a syntrophic partnership with sulfate-reducing bacteria in anoxic sediments.

2.1.1 Aerobic methane oxidation

Aerobic methane oxidation is confined to the Bacteria domain. Methane-oxidizing bacteria (MOB) constitute a functional rather than a phylogenetic group, as the ability to oxidize methane has evolved independently in multiple, phylogenetically distant bacterial lineages. Some MOB are found to be and able to grow on a multitude of other carbon compounds (facultative methanotrophs).
The aerobic oxidation of CH_4 involves several enzymatic steps. In modern methanotrophic bacteria, CH_4 is converted to methanol (CH_3OH) through a hydroxylation reaction catalyzed by the oxygen-dependent enzyme MMO. Methanol is subsequently oxidized by the methanol dehydrogenase to formaldehyde (HCHO), which is either assimilated into biomass or enters the catabolic pathway for energy conservation, with the consequent liberation of CO_2.
The MMO enzyme exists in two genetically unrelated forms:

- Particulate MMO (pMMO): Cu-bearing and membrane bound, present in most known methanotrophs;
- Soluble MMO (sMMO): Fe-containing and cytosolic, found only in a limited number of strains.

Methanotrophs can express i) both forms of MMOs, ii) only pMMO or iii) only sMMO. In methanotrophs possessing both forms, expression is regulated by copper availability. sMMO is expressed under low copper conditions, while increased Cu concentrations induce the production of pMMO, which is also associated to higher growth yields (Rani et al., 2024). As a consequence, the environmental dynamics of metal species significantly influence the CH_4 metabolism and, ultimately, the community structure (Fenibo et al., 2023; Steinle et al., 2016).
pMMO and sMMO are encoded by two distinct operons (i.e., pmoCAB and mmoXYBZDC, respectively) which carry the genes encoding for the protein subunits building the two enzymes. Molecular analysis showed that genes *pmoA* and *mmoX* (encoding for α- subunits of pMMO and sMMO, respectively) can

be used as reliable phylogenetic markers for MOB. Moreover, the phylogenetic profiles produced by the analyses of these two genes are largely congruent with those based on the 16S rRNA gene, allowing for simultaneous functional and taxonomic identification (Ahmadi and Lackner, 2024; Saunois et al., 2020). While the enzymatic machinery for initial steps of CH_4 oxidation is strongly conserved, methanotrophic bacteria exhibit significant diversity in their energy conservation and carbon assimilation strategies. These metabolic differences are reflected in their phylogenetic distribution across three main phyla:

- Proteobacteria: further divided into Gammaproteobacteria (Type I and Type X methanotrophs) and Alphaproteobacteria (Type II methanotrophs)
- Verrucomicrobia: Type III methanotrophs
- NC10: a candidate phylum, discovered about 20 years ago

These groups differ by cell morphology, physiology, and specific environmental preferences (Kalyuzhnaya et al., 2019). Three distinct metabolic pathways for the incorporation of C1 compounds into biomass are, currently, recognized (Fenibo et al., 2023):

1. The Ribulose Monophosphate (RuMP) pathway: typical of Type I methanotrophs (Gammaproteobacteria) and the NC10 phylum;
2. The Serine pathway: characteristic of Type II methanotrophs (Alphaproteobacteria);
3. The Calvin-Benson-Bassham (CBB) cycle: utilized by Verrucomicrobia and some Type X strains.

From an ecological point of view, the niche differentiation in methanotrophs seems to be driven by the variability of several environmental factors, such as methane concentration, oxygen, pH and temperature. Although MOB typically oxidize CH_4 in the presence of O_2, some groups can thrive under oxygen limitation due to an array of high-affinity O_2-binding proteins, such as cytochromes and hemerythrins (Guerrero-Cruz et al., 2021; Smith and Wrighton, 2019).
MOB distributed among three main phyla often show distinct physiological traits and environmental preferences. Type I and X methanotrophs (Gammaproteobacteria) generally dominate in CH_4-rich, hypoxic aquatic regions and exhibit high stress tolerance. Type II methanotrophs (Alphaproteobacteria) often prevail with low CH_4 concentrations and oxic surface waters (Kalyuzhnaya et al., 2019). However, the differential environmental regulation of the abundance and activity of type I and type II methanotrophs is still under debate, since experimental and environmental studies have yielded ambiguous and sometimes contrasting results (Crevecoeur et al., 2019).
The discovery of methanotrophs in the Verrucomicrobia phylum (Type III) in 2005 challenged the previous assumption that all aerobic methanotrophs were restricted to Proteobacteria. These bacteria are often extremophiles, capable of metabolizing H_2 and CH_4 in volcanic and highly acidic geothermal ecosystems. Their oxygen tolerance varies, with some strains preferring microoxic conditions (Schmitz et al., 2021).
Methanotrophs belonging to the NC10 phylum produce oxygen required for pMMO activity on their own (intra-aerobic metabolism) by dismutation of nitric oxide. This unique capability allows them to perform aerobic methane oxidation even in strictly anoxic environments (Ettwig et al., 2010; Fenibo et al., 2023).

2.1.2 Anaerobic methane oxidation

In marine environments, the anaerobic oxidation of CH_4 occurs mainly in the sediment and it is performed by microorganisms belonging to the Archaea domain, known as anaerobic methanotrophs (ANME). These organisms often operate through a syntrophic partnership with sulfate-reducing bacteria, utilizing the reverse-methanogenesis pathway.
Recent phylogenomic reclassifications have relocated ANME lineages into the Halobacteriota phylum, which also encompasses most methanogens. ANME sharing a homologous enzymatic machinery with methanogenic Halobacteriota, that they use for methane oxidation in the reversed methanogenesis pathway (Guerrero-Cruz et al., 2021). This methanotrophy pathway is thermodynamically unfavorable and yields very little energy, leading to the need of a strict association with other anaerobic microorganisms (syntrophy) for the efficient removal of metabolic intermediates (e.g. H_2), thus maintaining a negative ΔG – the thermodynamic condition necessary for methane oxidation to proceed. The *mcr*A gene, that encode for the methyl coenzyme reductase (MCR) is used to separate ANME from related methanogens (Smith and Wrighton, 2019).
ANME mainly occurs at the Sulfate Methane Transition Zone, where CH_4 and SO_4^{2-}overlap. According to Knittel and Boetius, 2009, in marine environments ANME are responsible for the oxidation of 90% of CH_4 produced by methanogens. In marine environments, the anaerobic oxidation of methane is mediated by three distinct groups of ANME, which belong to different orders or families with similar physiological features:

- ANME 1 (divided into the subclusters a and b)
- ANME 2 (divided into the subclusters a, b, and c)
- ANME-3

As for the aerobic methanotrophs, ANME exhibit a clear niche partitioning in the sediment, with: ANME-2a and 2b dominating upper layers, ANME-2c and/or ANME-1 increasing in deeper zones, and ANME-3 mainly found in mud volcanoes and in some seep sediments (Timmers et al., 2017). Interestingly, marine hydrothermal vent ecosystems do not offer many niches for ANME communities, due to lack of a real sediment cover, thus ANME are confined to small anoxic zones within vent chimneys (Knittel and Boetius, 2009).

2.2 Target Selection and Methodological Rationale

Functional markers such as *pmoA* and *mmoX* are well-established tools for high-resolution profiling of most methanotrophic guilds, as they target the key enzymes directly responsible for methane oxidation. However, they do not capture some lineages (e.g. ANME) and primers targeting them often suffer from coverage gaps due to divergency in the gene sequences. A multitarget approach is, thus, needed for a full description of the methanotrophic community. At the same time, targeting function genes does not allow for the identification of methanotrophs within their ecological context.
The 16S rRNA gene provides a robust phylogeny for methanotrophs, as it largely mirrors the taxonomic divergence found targeting methane-specific functional genes (Ahmadi and Lackner, 2024; Horz et al., 2001; Nagler et al., 2021; Saunois et al., 2020). In particular, the Next Generation Sequencing (NGS) of the 16S rRNA gene allows to target the of the microbial community as whole, providing both taxonomical information and insights in the community structure. While this enables capturing the microbial diversity

in high-throughput, making rare taxa visible, the choice of primers can introduce amplification bias, potentially underrepresenting certain methanotrophic groups.

In the MEFISTO project, we deployed a metabarcoding strategy targeting the V4-V5 region of the 16S rRNA gene. In spite of a small loss in taxonomic resolution, this approach allows for the simultaneous identification of aerobic methanotrophs and anaerobic methanotrophic archaea (ANME), providing a comprehensive snapshot of the 'biological methane filter' across the sampling sites, and placing the methanotrophic groups within the context of their local microbial community. The 16S gene target region was selected based on *in silico* coverage simulations testing the main universal primer pairs (Caporaso et al., 2012; Parada et al., 2016; Walters et al., 2016). We tested the primer set coverage for methanotrophs against the SILVA reference database (SSU r138.2 RefNR, (Quast et al., 2013). One mismatch was allowed, excluding the last 3 nucleotides at the 3' end, to ensure high taxonomic coverage while accounting for PCR stringency. The universal primers targeting the V4-V5 region (Parada et al., 2016) outperformed in detecting the ANME taxa and provided the best compromise, as shown in Table 2.1.

Table 2.1 – Parada et al. (2016) primer coverage analysis for methanotrophs. The performance of taxonomic assignments is based on total accessions in the SILVA reference database (SSU 138.2 nr). Coverage (%) is weighted to the total accessions per taxon in order to account for the uneven distribution of sequences within the database. Data are aggregated by methanotrophic functional guild.

Methanotroph functional group	*Total accessions*	*Mismatched*	*Not matchable*	***Weighted Coverage (%)***
Archaea				
ANME	499	40	5	**91.9**
Bacteria				
Aerobic Methanotrophs (Type I)	647	27	0	**95.8**
Aerobic Methanotrophs (Type II)	360	19	0	**94.7**
Anaerobic MOB (former Candidate Phylum NC10)	108	1	0	**99.1**
Verrucomicrobiota extremophilic methanotrophs	101	24	0	**76.2**

3. Sample management and processing

Within the MEFISTO project, the microbial communities associated with methane emissions were characterized across multiple environmental matrices. Collected samples included aliquots of seawater, hydrothermal fluid, interstitial water, sediment (including sandy substrates), and various types of biofilm structures (e.g. flocculant material, thin mats). Details about the sampling design and the sampling procedures are given in Deliverable 3.

A total of 140 samples were collected for the WP4 activities and processed. Table 3.1 shows the sample counts by matrix type and study area, while a description of sample pre-processing procedures is given in this section by sample type.

To avoid intrinsic biases, samples of the same type were processed using procedures as similar as possible, regardless of the collection site, sampling season, or study area.

Table 3.1. Detailed summary of samples collected for microbial community profiling, organized by sample type and study area.

Sample Type	*Trezze*	*Panarea*	*Total*
Biofilm			
Flocs	4	11	15
Flocs with sediment	3	1	4
Rock scrapings		2	2
Stone swab samples		9	9
Hydrothermal fluid		8	8
Interstitial water	6	11	17
Sediment			
Sediment	8	11	19
Sediment&Biofilm		6	6
Water column			
Surface	8	16	24
Intermediate	8	4	12
Bottom	8	16	24
Total	45	95	140

3.1 Water samples

In the MEFISTO project three types of water samples were collected: i) seawater samples (collected at two or three depth levels, depending on the water column height), ii) interstitial water samples, and iii) – exclusively for the Panarea's study area – samples of hydrothermal fluids, taken directly from scuba divers at the venting points. Water sample aliquots underwent filtration onto sterile 0.20 µm PES membrane filters that were then immediately frozen, until further processing. For long-term storage, the membrane were stored at -80°C.

For the Panarea's sampling campaigns, it was not possible to filter water samples on board. Thus, water samples were stored in the dark and cool soon after collection, and then taken to the ECCSEL Natlab facilities by 2 hours after collection. The filtration of the water column samples (water sample type i) was performed using reusable filtration devices mounting sterile 47 mm diameter PES membranes. The filtration apparatus were autoclaved prior each campaign, while between sample filtrations they were incubated with 1N HCl to prevent DNA cross-contamination. Instead, samples of hydrothermal fluid and interstitial water (water sample type ii and iii) were filtered onto single-use sterile filtration systems (ThermoScientific Nalgene Rapid-flow Filter Units) in order to prevent any possible form of contamination, given the low-DNA-input nature of these samples. Filtered volumes are shown in Table 3.1.

In the two sampling campaigns at the Trezze of the North Adriatic site, filtration was performed on-board, right after the sample collection. All samples were filtered using Sterivex filters mounting sterile 0.2 µm PES membranes, by drawing directly from the Niskin bottles through sterile silicone tubings and with the

aid of a portable peristaltic pump. On-board, the Sterivex filters were stored on dry-ice, then moved at -80°C until processing. Filtered volumes are shown in Table 3.2.

Table 3.1. Summary of volumes processed for the DNA isolation from samples collected at the Panarea's study area and the relative DNA extraction yields.

	Filtered volume (ml)		*Extracted DNA (ng/µl)*	
Water sample type	**Range**	**Median**	**Range**	**Average**
Water column	2500 – 9500	5000	9.07 – 39.30	21.64
Interstitial water	67 – 577	500	0.44 – 21.20	10.22
Hydrothermal fluid	270 – 600	500	0.03 – 14.10	3.51

Due to the fine grain size and the geochemical nature of the sediment at the N-Adriatic's sampling stations, the interstitial water samples were highly contaminated with fine sediment particle. Prefiltration to remove particles was not possible due to the contextual removal of prokaryotic cells.

Table 3.2. Summary of volumes processed for the DNA isolation from water samples collected at the N-Adriatic's study area and the relative DNA extraction yields.

	Filtered volume (ml)		*Extracted DNA (ng/µl)*	
Water sample type	**Range**	**Median**	**Range**	**Average**
Water column	1050 – 2650	2250	6.07 – 43.70	27.81
Interstitial water	170 – 310	270	4.03 – 52.00	29.05

Genomic DNA was isolated from filters using the DNeasy PowerWater Kit (Qiagen) with a few modifications to the manufacturer's protocol aimed at increasing the extraction yield (Fonti et al., 2021): after adding the lysing buffer provided by the manufacturer, samples underwent three cycles of incubation at 70°C for 5 min followed by 2 min vortexing at the maximum speed. We then proceeded as described in the kit handbook. Elution was performed twice to maximize DNA recovery from the silica buffer. Extraction blanks were included to asses potential contamination. DNA concentration was measured using a Qubit Fluorometer (ThermoFisher). Quality was inspected both visually by gel agarose electrophoresis and by Nanodrop spectroscopy. DNA extracts were stored at -80°C until further processing. DNA extraction yields are shown in Tables 3.1 and 3.2.

3.2 Sediment samples

Surface sediment samples, when present, were collected at each site by scuba divers using 50 ml polycarbonate tubes. On board, samples were either stored in the dark and cool and taken to the ECCSEL Natlab facilities within 2 hours after collection (for the sampling activities in Panarea), or placed on dry-ice in the dark (in the North-Adriatic sampling campaigns). The grain size of most sediment samples collected in the Panarea required the removal of excess water by centrifugation at 4.500 × g for 10 minutes. In some cases, the sediment was coarse sandy material with precipitates and featuring an incoherent flocculant biofilm that could not be separated (see Table 3.1). In the case of the two reference stations

(i.e. Zimmari and Blank OutCrop) sediment was collected by a 6L van Veen grab. Aliquots of surface layers were then collected by using 50 ml sterile polycarbonate tubes. All samples were stored at -80°C until processing.

For the DNA extraction, sediment samples were thawed, homogenized and sub-aliquoted. Genomic DNA was isolated using the DNeasy PowerSoil Pro Kit (Qiagen) following the manufacturer's protocol with a few modifications aimed at increasing the extraction yield: each sediment aliquot was split into three subaliquots to maximize the bead beating efficiency and the contaminant removal (e.g. exopolysaccharides, humic and fulvic acids, iron precipitates). After adding the lysis buffer, samples underwent three cycles of incubation at 70°C for 5 min followed by 3 min vortexing at the maximum speed. An addition 1-minute vortexing step was performed at the end. We then continued as described in the kit handbook. Elution was performed twice to maximize DNA recovery from the silica buffer. Extraction blanks were included to asses potential contamination. DNA quantification and quality checks were performed as described in Section 3.1. DNA extracts were stored at -80°C until further processing. The amount of sediment material used for DNA isolation and the corresponding extraction yields are summarized in Table 3.3.

Table 3.3. Summary of sediment amount processed for the DNA isolation and the relative DNA extraction yields. Sediment weight data refer to dry mass (calculated after full evaporation of the humidity at 60°C until weight stabilization).

	Amount of processed sediment (g)		*Extracted DNA (ng/µl)*	
Sampling site	**Range**	**Median**	**Range**	**Average**
Panarea	0.65 – 1.77	1.14	0.17 – 22.30	8.30
Trezze	0.55 – 3.42	2.13	10.50 – 58.00	38.39

3.3 Biofilm samples

In the MEFISTO project, additional seafloor samples (extended beyond standard sediment) were also collected by the scuba divers near the emissions, with the aim of analyzing the taxonomic composition of microbial benthic communities occurring as "biofilms" not directly associated with the sediment. These samples included: i) macroscopic mucilaginous flocs of varying consistency, sometimes containing inseparable sediment grains, ii) 10-15 cm size cobbles (10-15 cm) covered by a thin layer of polysaccharidic matrix, and iii) smaller encrusting clasts. Mucilaginous flocs were collected by using 200 ml long nosed syringe with wide bore nozzle, while the other sample type were collected using sterile plastic sample bags. On board, samples were stored in the dark and cool and the taken to the laboratory facilities by 6 hours after collection. Flocs-like samples were transferred into 50 ml sterile polycarbonate tubes and centrifuged at 4.500 × g for 5 minutes to remove the excess of water. The pelleted samples, gravels, and cobbles frozen were stored at -80°C until further processing.

DNA extraction tests were performed on a small subset of samples in order to define the best processing strategy for each biofilm sample type. The choice was made comparing the DNA extraction yields, quantified by Qubit. For the scope of the present report we will describe only the selected procedures.

Genomic DNA was isolated from flocs-like samples and the encrusting clasts (biofilm sample type i and iii, respectively) using the DNeasy PowerBiofilm Kit (Qiagen), following the manufacturer's protocol, with a double elution to maximize DNA recovery. In preparation to DNA extraction procedures, encrusting clasts were pretreated (after thawing) by scraping the surface to recover potentially containing microbial mat material. Scarping was performed with the aid of autoclaved and disinfected stainless steel spatulas. The scraped material was collected on sterile Petri dishes and then transferred into sterile tubes. Sub-aliquots of about 1 g were transferred into the bead-tubes and immediately processed for DNA extraction. The left-overs were stored at -80°C for biobanking.
The isolation of DNA from the cobbles (biofilm sample type ii) was performed adopting a different strategy. Cobbles were thawed at room temperature; the film covering the cobbles was collected by using sterile swabs that we, then, processed using the DNeasy PowerSoil Pro Kit (Qiagen) with the same modifications to the manufacturer's protocol described for the sediment samples (see Section 3.2). Three unused swabs were also extracted to assess potential contamination.
DNA quality checks and quantification were performed as described in Section 3.1. DNA extracts were stored at -80°C until further processing. The DNA extraction yields are shown in Table 3.4.

Table 3.4. DNA extraction yields from biofilm samples.

	Extracted DNA (ng/µl)	
Biofilm sample type	**Range**	**Average**
Panarea		
Flocs-like	4.53 – 37.30	20.80
Clast scrapings	5.70 – 16.30	11.00
Cobble swabs	0.87 – 8.20	3.43
Trezze		
Flocs-like	3.22 – 68.20	41.39

3.4 Metagenomic library and sequencing

The DNA extracts were diluted with nuclease-free water to a concentration of 5ng/µl and amplified with KAPA HiFi HotStart ReadyMix (Roche) using the universal primer pair 515F-Y/926R (Parada et al., 2016) following the Illumina protocol (available online: https://emea.illumina.com/content/dam/illumina-support/documents/documentation/chemistry_documentation/16s/16s-metagenomic-library-prep-guide-15044223-b.pdf). Some samples suffered amplification inhibition, likely due to residual polysaccharides or refractory organic compounds. For these samples, we applied an additional dilution factor and, in the worst cases, added bovine serum albumin (NEB, New England Labs) to the reaction mix to a final concentration of 0.2 µg/µl. None of these strategies resolved the PCR inhibition in the interstitial water sample collected in July 2024 at the San Pietro outcrop station. For this sample, we performed the amplification PCR using the AccuStart II PCR ToughMix (QuantaBio).
The PCR products were then sent to the sequencing facilities for the subsequent steps of the library preparation protocol, pooled in equal molar ratios, and sequenced on an Illumina NovaSeq 6000 platform (Illumina, San Diego, CA, USA) with a 2×250 bp strategy. Samples collected in the Panarea Island area

were sequenced at the ARGO Open Lab Platform (Area Science Park, Trieste, Italy), where the metagenetic libraries were prepared by enzymatic fragmentation followed by ligation of TruSeq adapters (Bentley et al., 2008). Samples collected in the Trezze of the North Adriatic Sea were sequenced by IGA Technology Services srl, where the metagenetic libraries were prepared by transposase-based tagmentation, with simultaneous ligation of Nextera adapters and DNA fragmentation (Adey et al., 2010).

4. BIOINFORMATIC ELABORATIONS

Quality inspection, sequence trimming, and filtering were performed locally. The remainder of the bioinformatic workflow was performed on the GALILEO100 HPC infrastructure of CINECA (Italian National Supercomputing Consortium), as part of the Partnership for Advanced Computing in Europe (PRACE).

Visual inspection of the read quality profiles was performed using FastQC and MultiQC. Primer removal was carried out with Cutadapt (v. 4.9, with Python v. 3.12.4) in a miniconda environment, using anchored primer trimming at both read ends, with the following settings: 15% maximum error rate, 10 bp minimum overlap and allowance for degenerate bases. Reads shorter than 50 bp, reads with terminal ambiguous bases (N), and reads not matching the primer sequence were discarded. The subsequent filtration and trimming steps were performed locally in a R environment (v. 4.5.1) using the DADA2 package (v. 1.36), based on quality profiles. The forward reads were trimmed at position 220, and the reverse at position 215; shorter reads were discarded. Reads contining N bases were discarded. Reads were also truncated at the first instance of a quality score less than or equal to 2. After truncation, reads with sum of expected errors greater than 2 were filtered out.

The remainder of the bioinformatic workflow was executed on the Galileo100 cluster, using resources allocated under the OGS23_PRACE_IT project. The analysis environment was managed via Conda (miniforge3) running R version 4.2. A customized local environment was built specifically for this workflow: all R packages and other dependencies were installed manually and managed within the user profile. Sequence data were processed using the DADA2 pipeline (v1.26), optimized for the HPC architecture. Multi-core parallelization was managed through RcppParallel, while job scheduling and resource allocation were handled by the SLURM workload manager.

Error estimation and sequence denoising accounted for differences in error profiles between samples sequenced in different runs. The distribution of sequencing errors was modeled using LOESS regression, with weights proportional to data abundance. We forced monotonicity in the estimated error rates, based on the biological assumption that the probability of error should not increase with higher Q-scores. The error-learning function used in this study is an adaptation of the solution proposed by Jona Lin (*loessErrfun*_mod4, https://github.com/jonalim/loessErrfun) for binned Q-score data, typical of Illumina NovaSeq sequencing. The 16S rRNA gene amplicon sequence variants (ASVs) were inferred by denoising in *pseudopooling* mode after a dereplication step. Subsequently, the sequence tables from all the sequencing runs were unified into a single dataset and prepared for chimera removal, which was performed with the *removeBimeraDenovo* function of the DADA2 package in *consensus* mode. Sequences outside the expected length range (360–390 bp) and singletons (i.e. sequences with one read detected across all samples) were removed prior to downstream analyses.

For taxonomy assignment, we used the naive Bayesian classifier provided within the DADA2 package against the SILVA SSU r138.2 RefNR reference database (released on July 11, 2024), setting the minimum bootstrap confidence level at 60.
Downstream analyses and data visualization were performed locally in the R environment (v. 4.5.1) using the phyloseq package (v. 1.52.0) for a unified sequence management, and other R packages, such as Biostrings (v. 2.76.0), microbiome (v. 1.30.0), stringr (v. 1.5.2), and tidyverse (v. 2.0.0). Sequences assigned to eukaryotes, mitochondria, chloroplasts, or to no domain were removed. ASVs with fewer than 5 reads were also removed.

5. PRELIMINARY RESULTS

Sample sequencing generated a total of 21.4 GB rawdata (fastq format). About the 25 % raw reads on a median were discharged over the bioinformatic analyses and post-processing filtration of unwanted sequences. A total number of 273,018 ASVs were resolved within the dataset, with 22,014 shared between the Panarea and the Trezze sampling sites. Other general statistics about the microbial community dataset are given in Tables 5.1.

Table 5.1. Statistics on the microbial community dataset. ASVs: amplicon sequence variants.

	Panarea	Trezze
Total reads	51453866	19162074
ASVs	184498	107769
Min reads	185068	270178
Max reads	3843019	889642
Sparsity index	0.95	0.89

Methanotrophic bacteria and archaea were extracted from the general dataset on the base of their taxonomic affiliation. A total number of 288 methanotrophic ASVs were found, representing up to the 2% of prokaryotic biodiversity within samples and belonging to the following families. ANME-2a-2b ANME-2c, Beijerinckiaceae, Hyphomicrobiaceae, Methanoperedenaceae, Methylacidiphilaceae, Methylococcaceae, Methylohalobiaceae, Methylomirabilaceae, Methylomonadaceae, Methylophagaceae and Methylophilaceae.

A preliminary exploration of the dataset showed a remarkable difference between the two sampling areas but an increased similarity between the seawater samples (Figure 5.1).

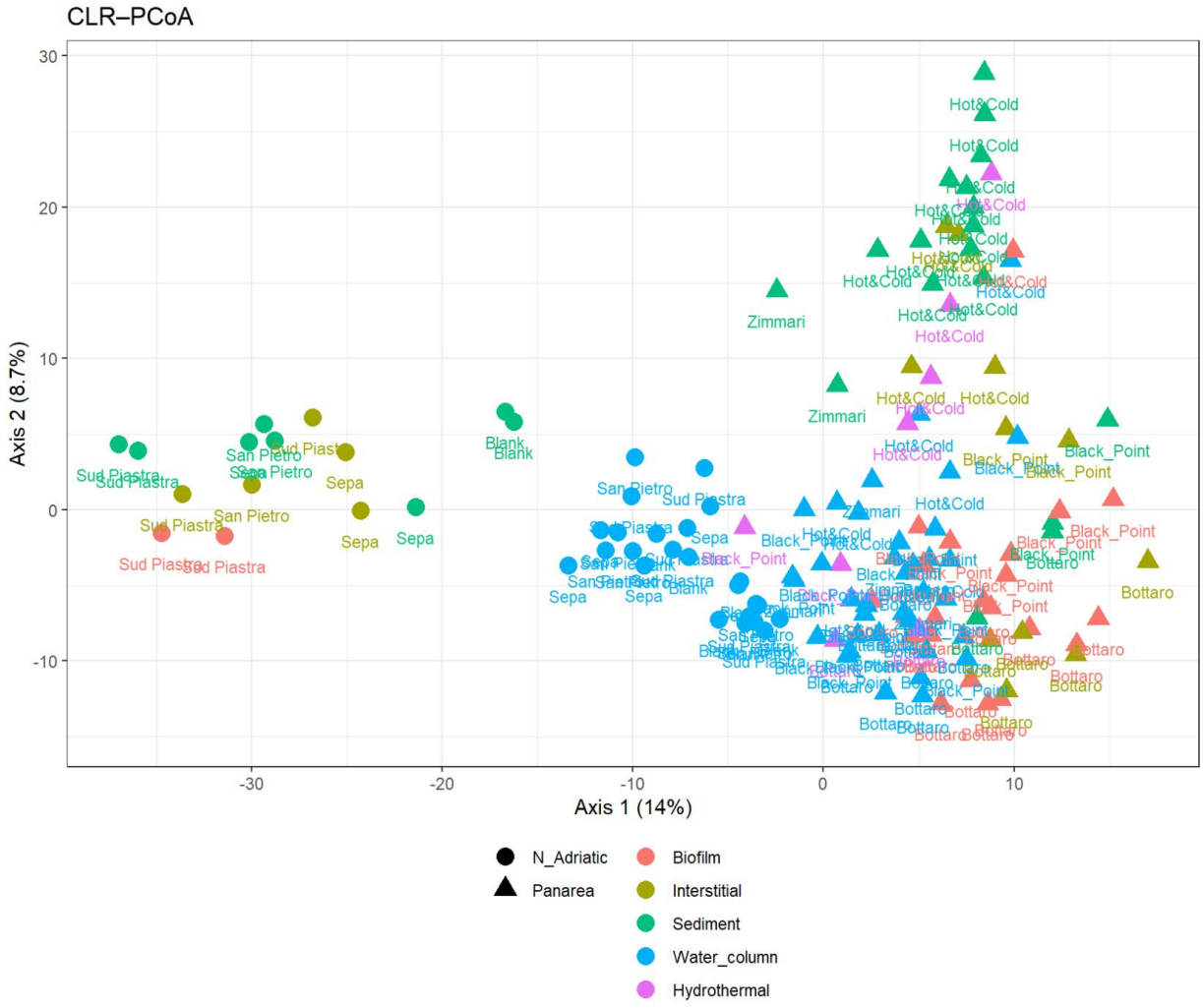

Figure 5.1. Sample ordination by PCoA of the methanotrophic taxonomic groups. Count data were transformed into centered log-ratios (CLR). Shapes represent the two sampling areas, while colors differentiate the environmental matrices.

6. ACKNOWLEDGEMENTS

The MEFISTO research team would like to acknowledge the European Union that financially supported the project through the "Fund for the National Research Programme and Research Projects of Significant National Interest (PRIN)" in the framework of the NRRP - Next Generation EU Mission 4 "Education and Research". Thanks go to Captain Angelo Portelli and Captain Stefano Caressa for the logistical support provided by the vessels M/N Corvo and Castorino 2 during the activities conducted at sea. The MEFISTO research team would also like to thank Associate Professor Francesca Malfatti and Dr. Elena Peresani from the University of Trieste, who were in Panarea to take samples as part of the FEMTO project, and who was glad to make an intellectual and practical contribution to the MEFISTO activities. Special thanks goes to Dr. Valentina Esposito, Dr. Marco Graziano and Dr. Giuseppe De Rosa from OGS ECCSEL NatLab-Italy, who granted the access to the facility and enthusiastically supported the onshore laboratory activities, and to Shoreline cooperative company and Grado port authority for the logistical support provided during the campaigns in the Northern Adriatic Sea. Thanks go to Dr. Stefano Salon from OGS for being manager and promoter or the OGS23_PRACE_IT partnership between OGS and CINECA, under which the HPC resources for the bioinformatic elaborations were provided at the benefit of the MEFISTO project.

7. REFERENCES

Adey A, Morrison HG, Asan, Xun X, Kitzman JO, Turner EH, et al. Rapid, low-input, low-bias construction of shotgun fragment libraries by high-density in vitro transposition. Genome Biol 2010;11:R119. https://doi.org/10.1186/gb-2010-11-12-r119.

Ahmadi F, Lackner M. Recent findings in methanotrophs: genetics, molecular ecology, and biopotential. Appl Microbiol Biotechnol 2024;108:60. https://doi.org/10.1007/s00253-023-12978-3.

Bentley DR, Balasubramanian S, Swerdlow HP, Smith GP, Milton J, Brown CG, et al. Accurate whole human genome sequencing using reversible terminator chemistry. Nature 2008;456:53–9. https://doi.org/10.1038/nature07517.

Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Huntley J, Fierer N, et al. Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. ISME J 2012;6:1621–4. https://doi.org/10.1038/ismej.2012.8.

Crevecoeur S, Ruiz-González C, Prairie YT, Del Giorgio PA. Large-scale biogeography and environmental regulation of methanotrophic bacteria across boreal inland waters. Mol Ecol 2019;28:4181–96. https://doi.org/10.1111/mec.15223.

Ettwig KF, Butler MK, Le Paslier D, Pelletier E, Mangenot S, Kuypers MMM, et al. Nitrite-driven anaerobic methane oxidation by oxygenic bacteria. Nature 2010;464:543–8. https://doi.org/10.1038/nature08883.

Fenibo EO, Selvarajan R, Wang H, Wang Y, Abia ALK. Untapped talents: insight into the ecological significance of methanotrophs and its prospects. Sci Total Environ 2023;903:166145. https://doi.org/10.1016/j.scitotenv.2023.166145.

Fonti V, Di Cesare A, Šangulin J, Del Negro P. Antibiotic resistance genes and potentially pathogenic bacteria in the central Adriatic Sea: are they connected to urban wastewater inputs? Water 2021.

Guerrero-Cruz S, Vaksmaa A, Horn MA, Niemann H, Pijuan M, Ho A. Methanotrophs: Discoveries, environmental relevance, and a perspective on current and future applications. Front Microbiol 2021;12:678057. https://doi.org/10.3389/fmicb.2021.678057.

Heintz MB, Mau S, Valentine DL. Physical control on methanotrophic potential in waters of the Santa Monica Basin, Southern California. Limnol Oceanogr 2012;57:420–32. https://doi.org/10.4319/lo.2012.57.2.0420.

Horz HP, Yimga MT, Liesack W. Detection of methanotroph diversity on roots of submerged rice plants by molecular retrieval of pmoA, mmoX, mxaF, and 16S rRNA and ribosomal DNA, including pmoA-based terminal restriction fragment length polymorphism profiling. Appl Environ Microbiol 2001;67:4177–85. https://doi.org/10.1128/AEM.67.9.4177-4185.2001.

Kalyuzhnaya MG, Gomez OA, Murrell JC. The methane-oxidizing bacteria (methanotrophs). Taxonomy, Genomics and Ecophysiology of Hydrocarbon-Degrading Microbes, Cham: Springer International Publishing; 2019, p. 245–78. https://doi.org/10.1007/978-3-030-14796-9_10.

Knittel K, Boetius A. Anaerobic oxidation of methane: progress with an unknown process. Annu Rev Microbiol 2009;63:311–34. https://doi.org/10.1146/annurev.micro.61.080706.093130.

Mao S-H, Zhang H-H, Zhuang G-C, Li X-J, Liu Q, Zhou Z, et al. Aerobic oxidation of methane significantly reduces global diffusive methane emissions from shallow marine waters. Nat Commun 2022;13:7309. https://doi.org/10.1038/s41467-022-35082-y.

Nagler M, Praeg N, Niedrist GH, Attermeyer K, Catalán N, Pilotto F, et al. Abundance and biogeography of methanogenic and methanotrophic microorganisms across European streams. J Biogeogr 2021;48:947–60. https://doi.org/10.1111/jbi.14052.

Nazaries L, Tate KR, Ross DJ, Singh J, Dando J, Saggar S, et al. Response of methanotrophic communities to afforestation and reforestation in New Zealand. ISME J 2011;5:1832–6. https://doi.org/10.1038/ismej.2011.62.

Parada AE, Needham DM, Fuhrman JA. Every base matters: assessing small subunit rRNA primers for marine microbiomes with mock communities, time series and global field samples. Environ Microbiol 2016;18:1403–14. https://doi.org/10.1111/1462-2920.13023.

Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, et al. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Res 2013;41:D590-6. https://doi.org/10.1093/nar/gks1219.

Rani A, Pundir A, Verma M, Joshi S, Verma G, Andjelković S, et al. Methanotrophy: A biological method to mitigate global methane emission. Microbiol Res (Pavia) 2024;15:634–54. https://doi.org/10.3390/microbiolres15020042.

Reis PCJ, Thottathil SD, Prairie YT. The role of methanotrophy in the microbial carbon metabolism of temperate lakes. Nat Commun 2022;13:43. https://doi.org/10.1038/s41467-021-27718-2.

Saunois M, Stavert AR, Poulter B, Bousquet P, Canadell JG, Jackson RB, et al. The global methane budget 2000–2017. Earth Syst Sci Data 2020;12:1561–623. https://doi.org/10.5194/essd-12-1561-2020.

Schmitz RA, Peeters SH, Versantvoort W, Picone N, Pol A, Jetten MSM, et al. Verrucomicrobial methanotrophs: ecophysiology of metabolically versatile acidophiles. FEMS Microbiol Rev 2021;45. https://doi.org/10.1093/femsre/fuab007.

Smith GJ, Wrighton KC. Metagenomic Approaches Unearth Methanotroph Phylogenetic and Metabolic Diversity. Curr Issues Mol Biol 2019;33:57–84. https://doi.org/10.21775/cimb.033.057.

Steinle L, Schmidt M, Bryant L, Haeckel M, Linke P, Sommer S, et al. Linked sediment and water-column methanotrophy at a man-made gas blowout in the North Sea: Implications for methane budgeting in seasonally stratified shallow seas. Limnol Oceanogr 2016;61:S367–86. https://doi.org/10.1002/lno.10388.

Timmers PHA, Welte CU, Koehorst JJ, Plugge CM, Jetten MSM, Stams AJM. Reverse methanogenesis and respiration in methanotrophic Archaea. Archaea 2017;2017:1–22. https://doi.org/10.1155/2017/1654237.

Walters W, Hyde ER, Berg-Lyons D, Ackermann G, Humphrey G, Parada A, et al. Improved bacterial 16S rRNA gene (V4 and V4-5) and fungal internal transcribed spacer marker gene primers for microbial community surveys. MSystems 2016;1. https://doi.org/10.1128/mSystems.00009-15.