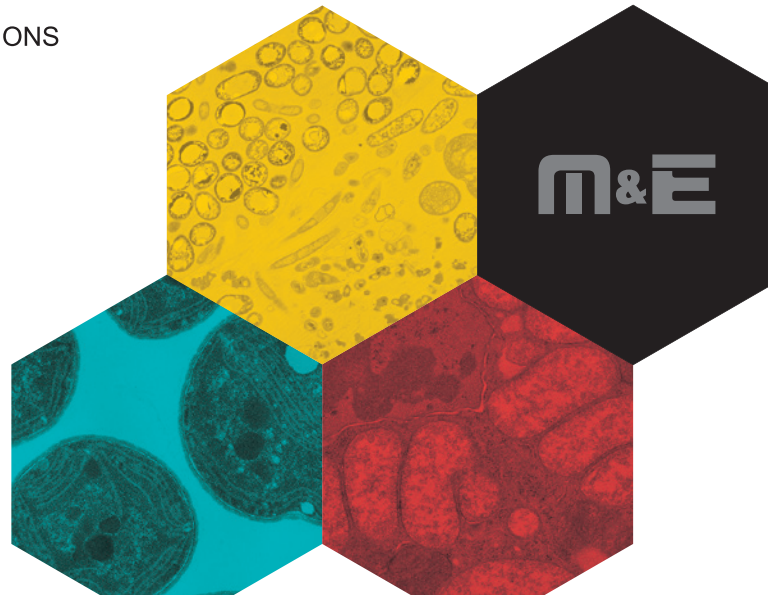
2024 Impact Factor
2.0
5 year IF 2.3

Microbes and Environments

DIGEST

JAPANESE SOCIETY OF MICROBIAL ECOLOGY
JAPANESE SOCIETY OF SOIL MICROBIOLOGY
TAIWAN SOCIETY OF MICROBIAL ECOLOGY
JAPANESE SOCIETY OF PLANT MICROBE INTERACTIONS
JAPANESE SOCIETY FOR EXTREMOPHILES



September 2025

Microbes and Environments
DIGEST 2025

CONTENTS

*Times cited at 2025/6/17

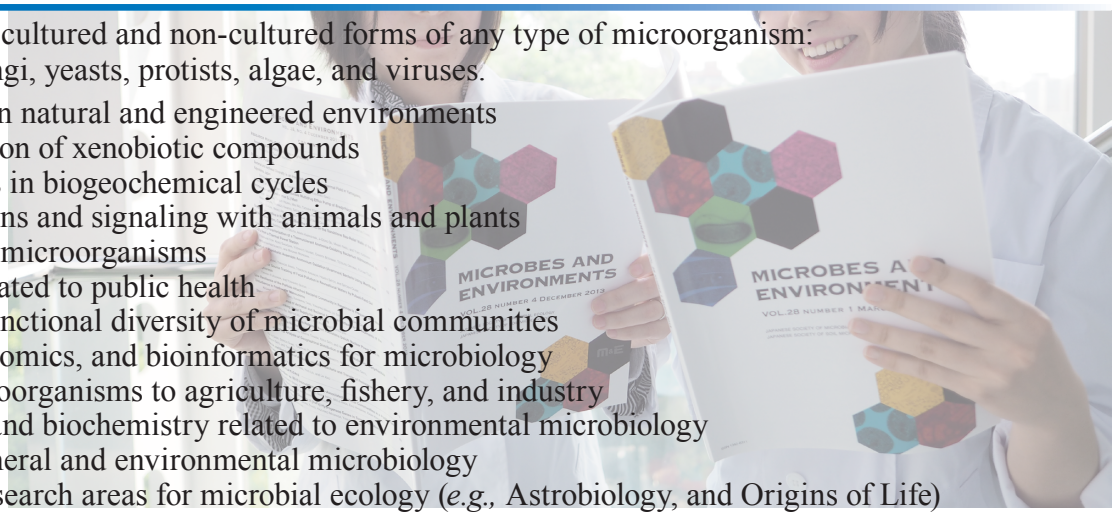
About Microbes and Environments

Microbes and Environments (M&E) is an international, peer-reviewed, and fully open-access journal devoted to the publication of original and review articles regarding all aspects of research on microbial ecology and environmental microbiology. It is published quarterly in an electronic version by Japanese Society of Microbial Ecology (JSME), Japanese Society of Soil Microbiology (JSSM), Taiwan Society of Microbial Ecology (TSME), Japanese Society of Plant Microbe Interactions (JSPMI), and Japanese Society for Extremophiles (JSE).

Benefits for Authors

- Open access policy for electric publication in PubMed, PubMed Central, and J-Stage
- Included in Web of Science® by Clarivate Analytics
- Page charges: ¥60,000/article for JSME/JSSM/TSME/JSPMI/JSE members
  ¥140,000/article for non-members

Interest of the Journal



Articles can deal with cultured and non-cultured forms of any type of microorganism: Bacteria, Archaea, fungi, yeasts, protists, algae, and viruses.

· Microbial ecology in natural and engineered environments
· Microbial degradation of xenobiotic compounds
· Microbial processes in biogeochemical cycles
· Microbial interactions and signaling with animals and plants
· Interactions among microorganisms
· Microorganisms related to public health
· Phylogenetic and functional diversity of microbial communities
· Genomics, metagenomics, and bioinformatics for microbiology
· Application of microorganisms to agriculture, fishery, and industry
· Molecular biology and biochemistry related to environmental microbiology
· Methodology in general and environmental microbiology
· Interdisciplinary research areas for microbial ecology (*e.g.,* Astrobiology, and Origins of Life)
· Taxonomic description of novel microorganisms with ecological perspective
· Physiology and metabolisms of microorganisms
· Evolution of genes and microorganisms
· Genome report of microorganisms with ecological perspective
· Life in extreme environments

Message from Editor-in-Chief

Microbes and Environments is now becoming one of the important journals for microbial ecology. The journal is being edited by five academic societies, but the submission is open to all researchers including non-members. Microbes and Environments is an open access and freely available online journal. Manuscripts are rigorously reviewed by credible peer reviewers to ensure excellent scientific quality. Once a manuscript is accepted, it will be published immediately after proof corrections so that readers can see and read the paper in a timely fashion. Different from commercial journals, we are providing the publication opportunity with an affordable price. We are looking forward to receiving your high-quality manuscripts.

Shigeto Otsuka
The University of Tokyo, Japan

Submit Your Manuscript

http://www.microbes-and-environments.jp/

Microbes Environ. 39(1), 2024
https://www.jstage.jst.go.jp/browse/jsme2 doi:10.1264/jsme2.ME23106

MOST VALUABLE PAPER OF THE YEAR 2024

Quest for Nitrous Oxide-reducing Bacteria Present in an Anammox Biofilm Fed with Nitrous Oxide

Kohei Oba[1], Toshikazu Suenaga[2,3], Shohei Yasuda[3,4], Megumi Kuroiwa[1], Tomoyuki Hori[5], Susanne Lackner[3,6], and Akihiko Terada[1,3]*

[1]*Department of Chemical Engineering, Tokyo University of Agriculture and Technology, 2–24–16 Naka-cho, Koganei, Tokyo, 184–8588, Japan;* [2]*Department of Chemical Engineering, Hiroshima University, 1–4–1 Kagamiyama, Higashi-Hiroshima, Hiroshima, 739–8527, Japan;* [3]*Global Innovation Research Institute, Tokyo University of Agriculture and Technology, 3–8–1 Harumi-cho, Fuchu, Tokyo, 185–8538, Japan;* [4]*Civil Engineering, School of Engineering, College of Science and Engineering, University of Galway, Galway H91 TK33, Ireland;* [5]*Environmental Management Research Institute, National Institute of Advanced Industrial Science and Technology, 16–1 Onogawa, Tsukuba, Ibaraki, 305–8569, Japan; and* [6]*Department of Civil and Environmental Engineering Science, Institute IWAR, Chair of Water and Environmental Biotechnology Technical University of Darmstadt, Franziska-Braun-Straße 7, 64287, Darmstadt, Germany*

(Received December 28, 2023—Accepted February 6, 2024—Published online March 28, 2024)

N_2O-reducing bacteria have been examined and harnessed to develop technologies that reduce the emission of N_2O, a greenhouse gas produced by biological nitrogen removal. Recent investigations using omics and physiological activity approaches have revealed the ecophysiologies of these bacteria during nitrogen removal. Nevertheless, their involvement in anammox processes remain unclear. Therefore, the present study investigated the identity, genetic potential, and activity of N_2O reducers in an anammox reactor. We hypothesized that N_2O is limiting for N_2O-reducing bacteria and an exogeneous N_2O supply enriches as-yet-uncultured N_2O-reducing bacteria. We conducted a 1200-day incubation of N_2O-reducing bacteria in an anammox consortium using gas-permeable membrane biofilm reactors (MBfRs), which efficiently supply N_2O in a bubbleless form directly to a biofilm grown on a gas-permeable membrane. A ^{15}N tracer test indicated that the supply of N_2O resulted in an enriched biomass with a higher N_2O sink potential. Quantitative PCR and 16S rRNA amplicon sequencing revealed Clade II *nosZ* type-carrying N_2O-reducing bacteria as protagonists of N_2O sinks. Shotgun metagenomics showed the genetic potentials of the predominant Clade II *nosZ*-carrying bacteria, *Anaerolineae* and *Ignavibacteria* in MBfRs. *Gemmatimonadota* and non-anammox *Planctomycetota* increased their abundance in MBfRs despite their overall lower abundance. The implication of N_2O as an inhibitory compound scavenging vitamin B12, which is essential for the synthesis of methionine, suggested its limited suppressive effect on the growth of B12-dependent bacteria, including N_2O reducers. We identified *Dehalococcoidia* and *Clostridia* as predominant N_2O sinks in an anammox consortium fed exogenous N_2O because of the higher metabolic potential of vitamin B12-dependent biosynthesis.

Key words: nitrous oxide reduction, nitrous oxide reductase gene *nosZ*, anammox, denitrifier

Increasing concerns for global environmental protection against nitrogen pollution has led to stricter emission standards and the consideration of biological nitrogen removal processes (McCarty, 2018). Anammox-based systems, *e.g.*, partial nitritation-anammox (PNA), are energy-saving and low-cost nitrogen removal processes that represent an alternative to conventional nitrification/denitrification. Since PNA reduces electricity consumption by up to 60% and lowers operation costs due to fewer requirements for aeration and external organic carbon (Gilbert *et al.*, 2015), more than 100 full-scale plants have hitherto implemented PNA in their operation since 2014 (Lackner *et al.*, 2014). The number of anammox-based systems, including PNA, for municipal and industrial wastewater treatment is increasing (Li *et al.*, 2018; Al-Hazmi *et al.*, 2023).

One of the challenges associated with anammox-based systems is the emission of nitrous oxide (N_2O). N_2O has a 273-fold higher global warming potential than CO_2 and is also known as an ozone-depleting substance (Smith *et al.*, 2021). Partial nitritation, the first step of PNA, is a hotspot for N_2O exhaustion triggered by low oxygen, but high nitrite concentrations (Domingo-Felez and Smets, 2019). Intensive N_2O emissions potentially offset the reduction in energy consumption and, in some cases, contrarily exceed the carbon footprint required for conventional nitrification/denitrification (Fenu *et al.*, 2019).

Aeration control is a common strategy to deter N_2O production in conventional wastewater treatment processes (Domingo-Felez and Smets, 2019; Duan *et al.*, 2021). This precautionary strategy is based on preventing the accumulation of nitrite, the primary source of N_2O production by ammonia-oxidizing and denitrifying bacteria (Peng *et al.*, 2017). Nevertheless, this control measure is not a panacea, resulting in N_2O emissions in some cases (Han *et al.*, 2023). In contrast, strategies that reduce N_2O produced via biotic and abiotic pathways have been proposed to mitigate N_2O

* Corresponding author. E-mail: akte@cc.tuat.ac.jp;
Tel: +81–042–388–7069; Fax: +81–042–388–7731.

Citation: Oba, K., Suenaga, T., Yasuda, S., Kuroiwa, M., Hori, T., Lackner, S., and Terada, A. (2024) Quest for Nitrous Oxide-reducing Bacteria Present in an Anammox Biofilm Fed with Nitrous Oxide. *Microbes Environ* **39**: ME23106.
https://doi.org/10.1264/jsme2.ME23106

Copyright: © Japanese Society of Microbial Ecology / Japanese Society of Soil Microbiology / Taiwan Society of Microbial Ecology / Japanese Society of Plant Microbe Interactions / Japanese Society for Extremophiles
https://creativecommons.org/licenses/by/4.0/

emissions. In one countermeasure strategy, gaseous N_2O in an off-gas line was fed to a bio-scrubber or biofilter, in which N_2O was converted into harmless nitrogen gas by N_2O-reducing bacteria (Frutos *et al.*, 2016; Yoon *et al.*, 2017; Han *et al.*, 2023). This concept has been successfully demonstrated in lab- (Yoon *et al.*, 2017) and pilot-scale studies (Han *et al.*, 2023), and is a promising option to reduce N_2O emissions from wastewater treatment plants (WWTPs) (Frutos *et al.*, 2016; Yoon *et al.*, 2017). In both strategies, harnessing N_2O-reducing bacteria is vital to accomplish the reduction of N_2O.

Regardless of the broad range of nitrogen loads, a consistent line-up of bacteria, *a.k.a.* the core microbiome, has been detected (Lawson *et al.*, 2017; Keren *et al.*, 2020; Xiao *et al.*, 2021). Furthermore, non-denitrifying N_2O-reducing bacteria, which are promising candidates as N_2O sinks and do not possess either or both nitrite reductase and nitric oxide reductase (Sanford *et al.*, 2012; Shan *et al.*, 2021), have been identified (Lawson *et al.*, 2017). In the core microbiome, some N_2O-reducing bacteria, consisting of Clade I and Clade II types, in an anammox biomass exhibited high activities devoid of external organic carbon sources as broadly available electron donors (Suenaga *et al.*, 2021). Based on their transcriptional activities, *Anaerolineaceae* (Clade II) and *Burkholderiaceae* (Clade I), counted as part of the core microbiome (Xiao *et al.*, 2021), potentially play an essential role in N_2O consumption in anammox reactors (Suenaga *et al.*, 2021). A metagenomic approach provides a more detailed understanding of the ecological and physiological functions of the core microbiome. Previous studies elucidated metabolic potentials (Speth *et al.*, 2016; Lawson *et al.*, 2017; Oshiki *et al.*, 2022a) by tracking the uptake of carbon labeled with radioactive and stable isotopes (^{14}C [Kindaichi *et al.*, 2012] and ^{13}C [Lawson *et al.*, 2021]), indicating the interdependence between anammox and heterotrophic bacteria in carbon metabolism, *e.g.*, vitamins, in anammox reactors (Lawson *et al.*, 2017; Keren *et al.*, 2020; Xiao *et al.*, 2021). However, this physiological interaction of N_2O-reducing bacteria with carbon and nitrogen compounds remains unclear and, thus, warrants further study.

Since N_2O-reducing bacteria use N_2O as an electron acceptor, an external supply of N_2O may promote the activity of N_2O-reducing bacteria when N_2O is a limiting factor. On the other hand, N_2O potentially inhibits bacterial growth, *e.g.*, *Paracoccus denitrificans*, because it reacts with vitamin B12 and deters methionine biosynthesis initiated from vitamin B12 at extracellular N_2O concentrations >2.8 mg N L^{-1} (Sullivan *et al.*, 2013). This concentration range was observed in an up-flow column reactor (Suenaga *et al.*, 2021) and the anaerobic regions of anammox granules (Okabe *et al.*, 2011). This suppression may be crucial in an anammox community in which most bacteria are interdependent on amino acids and vitamins provided by anammox bacteria (Keren *et al.*, 2020), particularly vitamin B12 (Lawson *et al.*, 2017; Oshiki *et al.*, 2022a). Under these conditions, the supply and retention of additional exogenous N_2O is a potential inhibitor. These controversial effects of N_2O on N_2O-reducing bacteria in an anammox community warrant thorough investigation with the goal of mitigating N_2O emissions from anammox-based systems.

Therefore, the present study attempted to identify the phylogeny of N_2O-reducing bacteria in an anammox reactor and characterize their metabolic functions based on their genotypes. Since the growth of N_2O-reducing bacteria in an anammox community is limited by the supply of N_2O (Suenaga *et al.*, 2021), we hypothesize that an external N_2O supply leads to the dominance of fast-growing N_2O-reducing bacteria under autotrophic conditions. To verify this hypothesis, we operated bioreactors designed to supply sufficient N_2O without bubble formation via a gas-permeable membrane (Kinh *et al.*, 2017; Suenaga *et al.*, 2019). We examined the effects of the exogenous N_2O supply on microbial community compositions and functions by the side-by-side operation of bioreactors with or without a N_2O supply with synthetic media containing ammonia and nitrite. 16S ribosomal RNA (rRNA) gene amplicon sequencing and shotgun metagenomic sequencing were performed for this evaluation.

Materials and Methods

Reactor set-up and operation

Two membrane biofilm reactors (MBfRs) (Suenaga *et al.*, 2019) were developed and applied for the enrichment of N_2O-reducing bacteria (Fig. S1). Each MBfR had liquid and gas compartments between which a flat-sheet silicon gas-permeable membrane was inserted. Their volume (0.26 L) and dimensions are referred to in a previous study (Kinh *et al.*, 2017). One MBfR, Reactor 1 (w/N_2O), was supplied 5% (v/v) N_2O (base gas: N_2) at 5 kPa as a feeding gas from the gas compartment (on the bottom) to the biomass grown on a flat-sheet gas-permeable silicone membrane (L×W of 170×30 mm with a wall thickness of 1 mm; Rubber) (Fig. S1). The other MBfR, Reactor 2 (w/o N_2O), had the same configuration and dimensions, but was not supplied with N_2O. A biomass from an up-flow column-bed anammox reactor (Suenaga *et al.*, 2021) was inoculated into Reactors 1 and 2.

The MBfRs were operated in a thermostatic chamber at 30°C for 1237 days. The medium was continuously supplied, and the liquid was recirculated through ports on the side walls of the MBfRs (Fig. S1). The medium used contained (L^{-1} of distilled water) 100 mg N of NH_4^+, 100 mg N of NO_2^-, 540 mg of $NaHCO_3$, 27 mg of KH_2PO_4, 300 mg of $MgSO_4{\cdot}7H_2O$, and 180 mg of $CaCl_2{\cdot}2H_2O$. One milliliter of trace element solutions with compositions described elsewhere (deGraaf *et al.*, 1996) was added to per liter of the medium. The mixed medium was continuously purged with N_2 gas to eliminate dissolved oxygen, and oxygen in the medium influent tank was maintained at a low concentration. The hydraulic retention time (HRT) was consistently set at 1 day.

A sample was taken from the influent and effluent ports (Fig. S1) and stored after filtration through a 0.45-μm membrane filter (A045A025A; Advantec). NH_4^+, NO_2^-, and NO_3^- concentrations were measured by ion chromatography (ICS1000 and ICS90; Thermo Fisher Scientific). In addition, 1 mL of the mixed medium was sampled into a gas vial (13 mL) sealed with a butyl rubber stopper filled with 12 mL of pure nitrogen gas to measure the gaseous concentration of N_2O. The dissolved concentration of N_2O was assessed by the liquid-gas equilibrium method, as previously reported (Isobe *et al.*, 2011; Riya *et al.*, 2012; Suenaga *et al.*, 2021), collecting a gaseous sample for the measurement. The concentration of N_2O in the headspace was measured by gas chromatography-quadrupole mass spectrometry (GCMS-QP2010 Ultra; Shimadzu).

Batch test using a ^{15}N-labeled tracer

A ^{15}N tracer test was performed using the biomass collected from Reactors 1 and 2 on day 776 to evaluate gross anammox, N_2O

production, and N_2O consumption activities. In the evaluation, 40 mg N L^{-1} of ^{15}N-labeled $^{15}NO_2^-$ (98% labeled; Shoko Science) and 40 mg N L^{-1} of unlabeled NH_4^+ were used. Fifteen milliliters of the mixed medium, the same volume as that supplied to the MBfRs, was poured into a 30-mL vial. Approximately 0.3 g of the dewatered biomass was inoculated, and $^{44}N_2O$ was added to evaluate the N_2O consumption potential. N_2O and N_2 concentrations in the headspace were measured by GCMS. Experiments were conducted in triplicate. The biomass pretreatment and measurement conditions are described in our previous study (Suenaga *et al.*, 2021).

Since $^{15}NO_2^-$ contained 98% of labeled ^{15}N (^{15}N fraction: F=0.98), it was assumed that $^{15}NO_2^-$ was converted to $^{30}N_2$ via $^{46}N_2O$ mediated by a heterotrophic denitrifying pathway and $^{15}NO_2^-$ and non-labeled NH_4^+ were converted into $^{29}N_2$ by the anammox reaction. N_2O production, N_2O consumption, and anammox activities were obtained using Eqs. 2–4. In these equations, $V_{obs,\ 44N2O}$, $V_{obs,\ 46N2O}$, $V_{obs,\ 29N2}$, and $V_{obs,\ 30N2}$ (note that $V_{obs,\ 44N2O}$ is a negative value) were obtained by the linear approximation of changes in the concentrations of $^{44}N_2O$, $^{46}N_2O$, $^{29}N_2$, and $^{30}N_2$ from 4 to 15 h. Detailed calculations are described in our previous study (Suenaga *et al.*, 2021).

$$N_2O \text{ production activity} = \left(V_{obs,30N2} + V_{obs,46N2O}\right) \bullet F^{-2} \quad \text{(Eq. 1)}$$

$$N_2O \text{ consumption activity} = V_{obs,30N2} \bullet F^{-2} - V_{obs,44N2O} \quad \text{(Eq. 2)}$$

$$\text{Anammox activity} = \left(V_{obs,29N2} - 2 \bullet V_{obs,30N2} \bullet F^{-1}(1-F)\right) F^{-1} \quad \text{(Eq. 3)}$$

Spatial localization of targeted bacteria by fluorescence in situ *hybridization (FISH)*

Biomass samples on day 1044 were subjected to FISH. The floc biomass was collected from Reactor 1 (w/N_2O) and Reactor 2 (w/o N_2O) (Fig. S2), followed by immediate fixation with 4% paraformaldehyde. Fixation and subsequent hybridization procedures were performed as previously described (Terada *et al.*, 2013). The oligonucleotide probes applied and formamide percentages are shown in Table S1. A microscopic analysis was performed using a confocal laser scanning microscope (LSM 900; Carl Zeiss) with a Diode Laser (488, 561, and 640 nm) and Airyscan 2. Image processing was performed using ZEN 3.0 (blue edition) (Carl Zeiss) and Imaris 10.0 (Oxford Instruments).

Taxonomy compositions and functional gene abundance

The biomass was routinely collected for a phylogenetic analysis targeting the 16S rRNA gene and the quantification of functional genes. The targeted genes were the *nirK* and *nirS* genes that encode nitrite reductase, the *cnorB* and *qnorB* genes that encode nitric oxide reductase, and the Clade I *nosZ* and Clade II *nosZ* genes that encode N_2O reductase (NOS). DNA extraction was conducted using the FastDNA Spin Kit for Soil (MP Biomedicals) according to the manufacturer's instructions. Functional gene quantification was performed using the CFX96 Real-Time PCR Detection System (BioRad Laboratories). The corresponding primers and sequences are summarized in Table S2, and PCR conditions are described in the Supplementary Information (SI).

PCR amplification of the V4 hypervariable region of the 16S rRNA gene was conducted using the primer set 515f-806r (Table S2). PCR conditions and library preparation procedures are described in SI. The DNA libraries generated and the initial control (bacteriophage PhiX; Illumina) were sequenced with a 300-cycle MiSeq Reagent kit (version 2, Illumina) using a MiSeq DNA sequencer (Illumina) in the paired-end sequencing mode.

Sequence data were analyzed using the following bioinformatics tools: removal of the adapter and low-quality reads were conducted using BBMap/BBduk (v. 38.84) (Bushnell, 2014), and trimmed reads were merged using FLASh (version: 2.2.00) (Magoc and Salzberg, 2011). Merged reads were imported into Dada2 (Prodan *et al.*, 2020) (version: 1.26.0). PhiX sequences were removed from imported reads with Dada2's FilterAndTrim command, and the reminders were then clustered into amplicon sequence variant (ASV) inference (default parameters), followed by the elimination of chimeras (default parameters). Samples with yields >10000 reads in total after the elimination of chimeras were used for a downstream analysis. 16S rRNA ASVs imported into Qiime2 (Bolyen *et al.*, 2019) (q2cli: 2023.5.1) were assigned by lineages using the QIIME 2 q2-feature-classifier plugin (Bokulich *et al.*, 2018) with the pre-trained Silva (138.1) database (Quast *et al.*, 2013; Robeson *et al.*, 2021). Diversity and statistical analyses were conducted using R version 4.2.2 (2022-10-31) and the phyloseq package (McMurdie and Holmes, 2013). The parameters used in the analysis are listed in SI and Table S3-S6.

Reconstruction and analyses of the metagenome-assembled genome

Three biomass samples, two from Reactor 1 (w/N_2O) and one from Reactor 2 (w/o N_2O) were collected on day 981 for shotgun metagenomic sequencing. DNA was extracted using a phenol-chloroform method (Butler, 2012; Yasuda *et al.*, 2020). Biomass samples were centrifuged (10000 rpm), and TE buffer (10 mM Tris/HCl and 10 mM EDTA, pH=8) and 10% SDS were added to the pelleted biomass. Genomic DNA was purified by repeating DNA and protein separation using phenol, chloroform, and CTAB/NaCl solution. RNA as a contaminant in genomic DNA was decomposed by RNaseA (TaKaRa Bio). After ethanol precipitation, genomic DNA was suspended in TE buffer and stored in a freezer until used. Library preparation and sequencing were performed at Azenta Life Science. Sequencing was performed on a Novaseq (Illumina) with a 150-bp paired-end sequencing protocol, and 10 Gb of data was obtained per sample.

The metagenomic pipeline employed in the present study is shown in Fig. S3. A quality check of raw sequencing data and low-quality read trimming were performed using fastp v0.22.0 (Chen *et al.*, 2018). Trimmed reads were assembled by Megahit (v1.2.9) (Li *et al.*, 2015) and metaSPAdes (v3.13.1) (Nurk *et al.*, 2017) in parallel with default parameters. Contigs and scaffolds were filtered using SeqKit (v2.2.0) (Shen *et al.*, 2016), and those longer than 500 bp were used in subsequent analyses.

To analyze the genomic profile of the *nosZ* gene in metagenome samples, gene predictions in filtered contigs and scaffolds were performed by Prodigal (v. 2.6.3) (Hyatt *et al.*, 2010). Predicted genes from all biomass samples were integrated to be non-redundant by Cd-hit (v. 4.8.1) (Li and Godzik, 2006) with '-c 0.95 -aS 0.9 -g 1'. Function assignments were conducted using EggNOG mapper (v2.1.9) (Cantalapiedra *et al.*, 2021) with the diamond mode and InterProScan (v. 5.65–97.0) (Jones *et al.*, 2014; Blum *et al.*, 2021). Quality-trimmed reads were mapped onto the assembled contigs using BWA-MEM (v. 2.2.1) (Vasimuddin *et al.*, 2019) with default parameters. The resulting sequence alignment mapping files were sorted using SAMtools (v. 1.13) (Danecek *et al.*, 2021), and the read coverage of each contig was calculated using featureCounts (v. 2.0.3) (Liao *et al.*, 2013). Read coverage was further normalized by the total number of mapped reads in each sample, yielding reads per kilobase per million mapped reads (RPKM) values.

To reconstruct the metagenome-assembled genome (MAG), filtered contigs and scaffolds were grouped into primary-bins with MaxBin (v2.2.6) (Wu *et al.*, 2016), MetaBAT (v2.12.1) (Kang *et al.*, 2019), and CONCOCT (v 1.0.0) (Wu *et al.*, 2016) with default parameters. The bins were consolidated into final bins by DASTools (v1.1.4) (Sieber *et al.*, 2018) using default parameters.

Redundant bins generated in parallel were de-replicated by dRep (v3.3.1) (Olm *et al.*, 2017) with '-l 50000 -pa 0.90 -sa 0.99 -comp 50 -con 25 -nc 0.1'. CheckM2 (v1.0.1) (Chklovski *et al.*, 2022 CheckM2: a rapid, scalable and accurate tool for assessing micro-

bial genome quality using machine learning. *bioRxiv*: https://doi.org/10.1101/2022.07.11.499243) was used to check the quality of bins, and bins with completeness <70% and contamination >5% were excluded. The presence of 16S/23S/5S rRNA was checked with barrnap (v0.9) (Seemann, 2018), and the numbers and types of tRNAs were counted using tRNAscan-SE (v 2.0.9) (Chan *et al.*, 2021). Taxonomy assignment was conducted using GTDB-Tk (v2.1.1) (Chaumeil *et al.*, 2022) with GTDB r207 (Parks *et al.*, 2022). DFAST (v1.2.14) (Tanizawa *et al.*, 2018) was used for gene predictions and functional annotations. Additional annotations for function assignments were conducted using EggNOG mapper (v2.1.9) (Cantalapiedra *et al.*, 2021) with a diamond mode and KofamScan (v 1.3.0) (Aramaki *et al.*, 2020). Annotation results were merged based on the Kyoto Encyclopedia of Genes and Genomes (KEGG) Orthology (KO) and mapped to the KEGG database (Kanehisa and Goto, 2000) using KEGG Decoder (v1.3) (Graham *et al.*, 2018). The relative abundance of bins in each sample was calculated using CoverM (v0.6.1). Since two biomass samples were collected from Reactor 1 (w/N_2O), the average relative abundance of the two samples was used.

Data availability

Raw 16S rRNA gene amplicon and metagenomic sequencing data are available in the DNA Data Bank of Japan (DDBJ) nucleotide sequence database under accession numbers DRA016674 and DRA016675, respectively. Assembled and annotated MAGs were deposited in the DDBJ nucleotide sequence database with the accession numbers shown in Table S7.

Results

Reactor operation and nitrogen removal performance

Similar changes were observed in effluent NH_4^+, NO_2^-, and NO_3^- concentrations in Reactors 1 (w/N_2O) and 2 (w/o N_2O) (Fig. S4 and S5). In the first 90 days of the incubation, approximately 50 mg N L^{-1} of NH_4^+ and NO_2^- remained in the effluent of both reactors. NH_4^+ and NO_2^- concentrations in the effluent then decreased in both reactors, reaching average NO_2^- and NH_4^+ concentrations of 18.2 and 20.7 mg N L^{-1}, respectively, from day 100 to 260 in Reactor 1 and 9.76 and 25.1 mg N L^{-1}, respectively, in Reactor 2. After day 295, NH_4^+ and NO_2^- concentrations in the effluent further decreased in Reactors 1 and 2. However, NH_4^+ and NO_2^- accumulated in the effluent from day 310 to 405 (Fig. S4b and S5b), possibly due to the lack of maintenance during the lockdown period of COVID-19. After day 426, NO_2^- and NH_4^+ concentrations again decreased (16.5 and 12.6 mg N L^{-1}, respectively, in Reactor 1 and 18.3 and 13.8 mg N L^{-1}, respectively, in Reactor 2 on average). Influent and effluent pH were 7.90±0.21 and 8.17±0.41, respectively, in Reactor 1 and 7.92±0.31 and 8.13±0.30, respectively, in Reactor 2. The exogenous supply of N_2O markedly affected dissolved N_2O concentrations in the bulk liquid. The dissolved concentration of N_2O in the bulk liquid was 3.96±2.67 mg N L^{-1} on average in Reactor 1 with a maximum of 8.95 mg N L^{-1} (day 805) and a minimum of 1.31 mg N L^{-1} (day 864) (Fig. S6). The dissolved concentration of N_2O was markedly lower in Reactor 2 (Fig. S7), with an average of 0.67±0.38 mg N L^{-1}, maximum of 1.72 mg N L^{-1} (day 776), and minimum of 0.34 mg N L^{-1} (day 1091). Since the total concentration of nitrogen in the influent was 200 mg N L^{-1}, the N_2O conversion ratio over total nitrogen in Reactor 2 was 0.34% under the assumption of marginal N_2O exhaustion to the gaseous layer due to the absence of the headspace, which was similar to that in an anammox column reactor (Suenaga *et al.*, 2021). The stoichiometric ratios of NO_2^- consumption and NO_3^- production over NH_4^+ consumption in both reactors were similar to those of enriched anammox bacteria (Strous *et al.*, 1998) (Fig. S4c and S5c), suggesting that an anammox reaction was responsible for biological nitrogen removal.

Regardless of the presence or absence of a N_2O supply to the reactors, their biomasses were mostly aggregated, consisting of granules or flocs. Although both reactors were constantly stirred by liquid recirculation, some of the biomass sedimented onto the gas-permeable membrane, while the remainder continued to be suspended in the bulk liquid or formed a biofilm on the reactor sidewall or gas-permeable membrane surface (Fig. S2a). However, the biofilm mass of these parts was insufficient to extract DNA (Fig. S2b) for routine functional gene quantification or a phylogenetic analysis. Therefore, samples for DNA extraction were collected from the aggregated biomass in a suspension.

Intrinsic activity test

Biomass-specific anammox rates were 0.90±0.22 mg N [g MLVSS]$^{-1}$ h^{-1} for Reactor 1 (w/N_2O) and 0.97±0.28 mg N [g MLVSS]$^{-1}$ h^{-1} for Reactor 2 (w/o N_2O), with no significant difference (P=0.735) (Fig. 1). Similar N_2O production rates were attained: 0.037±0.014 mg N [g MLVSS]$^{-1}$ h^{-1} for Reactor 1 (w/N_2O) and 0.039±0.0028 mg N [g MLVSS]$^{-1}$ h^{-1} for Reactor 2 (w/o N_2O). N_2O consumption rates were 0.21±0.064 mg N [g MLVSS]$^{-1}$ h^{-1} for Reactor 1 (w/o N_2O) and 0.14±0.014 mg N [g MLVSS]$^{-1}$ h^{-1} for Reactor 2 (w/o N_2O). Although no significant differences were observed in N_2O production (P=0.930) or consumption (P=0.452), an increase was noted in the N_2O consumption activity of the biomass fed N_2O.

Quantification of functional gene dynamics

The dynamics of functional gene abundance are shown in Fig. 2. The continuous external supply of N_2O allowed for the distinct dynamics of denitrifying genes. Gene copies of *nirK*, *nirS*, *cnorB*, and *qnorB* increased to day 280. Gene copies of *nirK* and *nirS* remained constant in Reactor 1 (w/o N_2O), while those in Reactor 2 (w/o N_2O) decreased from those in the inoculum. *nirK* gene abundance was similar in both reactors, whereas that of the *nirS* counterpart was significantly higher in Reactor 1 than in Reactor 2 after day 280, except on days 718 and 810 (P<0.05). The continuous external supply of N_2O increased *nosZ* gene abundance. Despite the presence or absence of the external N_2O supply, the Clade II *nosZ* gene was one order of magnitude higher than the Clade I *nosZ* gene. In contrast, Clade II *nosZ* gene copies in Reactor 2 (w/o N_2O) were lower on most sampling dates than those on day 0, which was opposite to that attained in Reactor 1 (w/N_2O). Clade II *nosZ* gene copies were significantly higher in Reactor 1 than in Reactor 2 after day 280 (P<0.05).

Taxonomy compositions

The top 6 taxa at the phylum level during the entire incu-

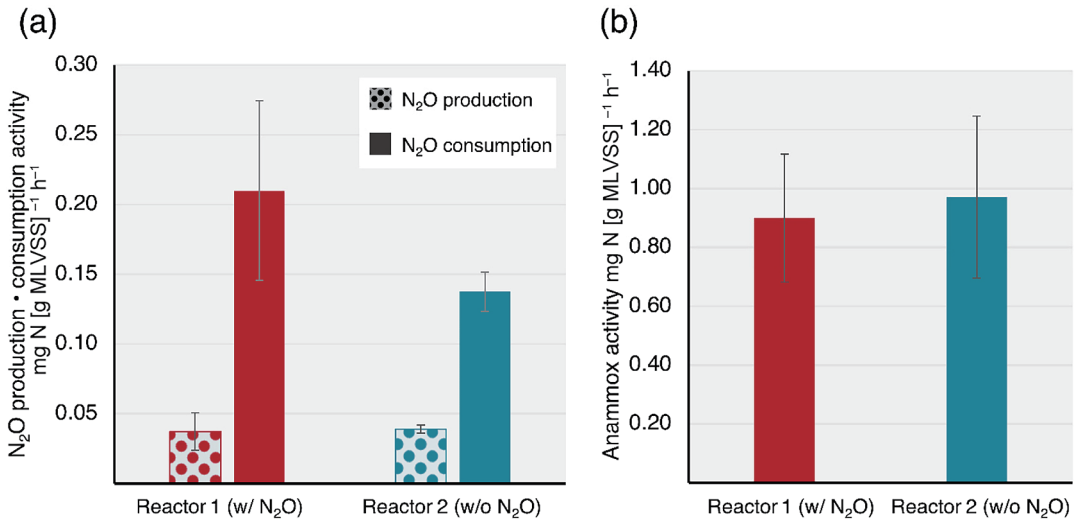


Fig. 1. Biomass-specific (a) N_2O production and N_2O consumption, (b) anammox activities calculated from the ^{15}N tracer test. Experiments were conducted in triplicate for reproducibility. Bars and error bars represent mean values and standard deviations, respectively. In the statistical analysis, Welch's *t*-test was performed (P>0.05).

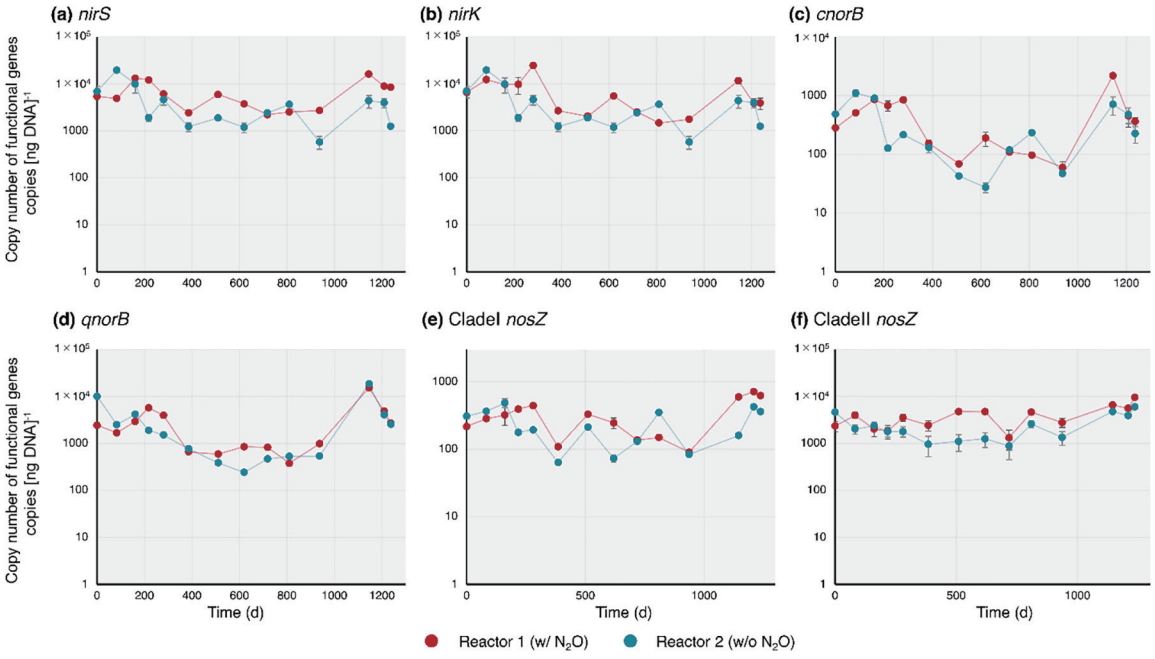


Fig. 2. Time courses of gene densities by quantitative PCR (n=3): (a) *nirS*, (b) *nirK*, (c) *cnorB*, (d) *qnorB*, (e) Clade I *nosZ*, and (f) Clade II *nosZ* genes normalized by DNA weight.

bation period are shown in Fig. 3. In both reactors, the phyla *Planctomycetota* and *Chloroflexota* accounted for more than 50% at all periods, except on day 246. The relative abundance of anammox bacteria (class *Brocadiales*) was higher in Reactor 2 than in Reactor 1. The relative abundance of anammox bacteria in Reactor 2 (w/o N_2O) was on average 24.8% for the period before day 385 and 51.2% after. Within the class *Brocadiales*, *Candidatus* Jettenia was the dominant bacterial species, regardless of the supply of N_2O.

Chloroflexota was the dominant phylum among non-anammox bacteria, with *Anaerolineae* (Reactor 1: 6.3–48.1% of the total, Reactor 2: 5.6–45.6%) and *Ardenticatenaceae* (Reactor 1: 1.9–48.1%, Reactor 2: 5.6–45.6%) as the predominant classes. Other members were *Ignavibacteria*, *Betaproteobacteria*/*Burkholderiales*, *Fimbriimonadia*, *Gammaproteobacteria*/*Xanthomonadales*, and *Gammaproteobacteria*/*Phycisphaerae*. The relative abundance of these non-anammox bacteria, potentially regarded as denitrifiers, was higher in Reactor 1 than in Reactor 2. The relative abundance of non-anammox bacteria changed over time, accounting for an average of approximately 60% and reaching a maximum of 77% in Reactor 1 after day 426, when nitrogen removal performance stabilized.

The bacterial community compositions of the biofilm adhering to the membrane and the aggregate biomass deposited onto the membrane-bound biofilm taken from Reactor 1 on day 199 (Fig. S2) are shown in Fig. S8. The major taxa

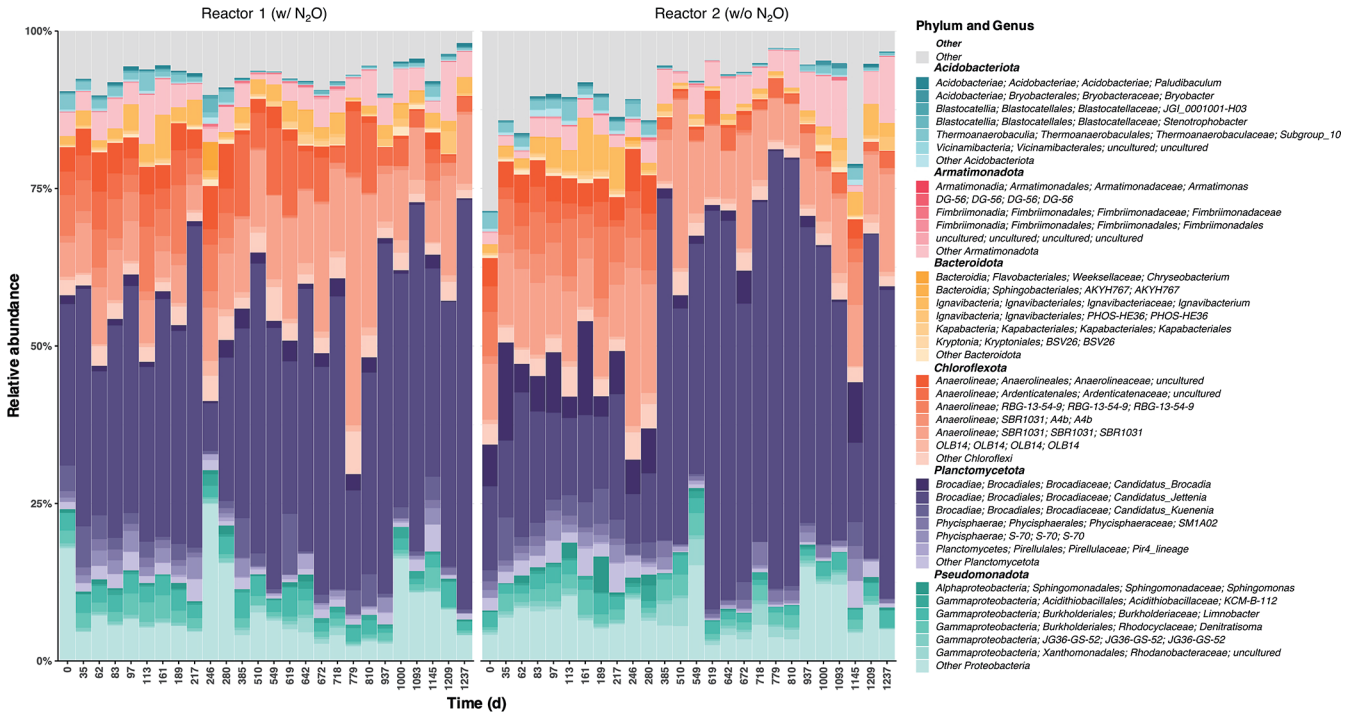


Fig. 3. Evolution of microbial community compositions by 16S rRNA gene amplicons in Reactors 1 (w/N_2O supply) and 2 (w/o N_2O supply). The top six phyla and genera during the overall incubation are listed.

constituting the community were analogous; however, their compositions differed. *Anaerolineae* was the most dominant biofilm component, of which the ASV assigned to the genus *SBR1031* had a higher relative abundance in the biofilm (32.6%>9.8% in the aggregate biomass). On the other hand, the relative abundance of the family *Brocadiaceae* was lower in the biofilm (3.5%) than in the aggregate biomass (15.7%).

Activity analysis with FISH imaging

FISH analyses showed sufficient fluorescence in samples taken from Reactor 1 (w/N_2O) (Fig. 4a, b, c, d, e, f, and g). Consistent with amplicon sequencing results, anammox and *Chloroflexota* were abundant. The spatial distribution of these bacteria in the floc did not markedly differ between the two reactors (data not shown). The mesh-like structure of the phylum *Chloroflexota*, possessing a filamentous cell morphology, was distributed over the floc, particularly in the floc exterior (Fig. 4a, b, c, d, and e). In contrast, anammox bacteria were present inside the floc. Bacteria other than *Chloroflexota* and anammox bacteria were detected near filamentous cells.

Analyses of MAGs

Of the reconstructed bins, after quality screening (completeness ≥70%, contamination <5%), 58 MAGs were acquired (Fig. 5 and Table S8). MAGs detected with higher relative abundance (>1.5-fold) in Reactor 1 than in Reactor 2 were assigned to *Vicinamibacteria* (bin 57), *Chthonomonadetes* (bin 19), *Anaerolineae* (bin 60), *Dehalococcoidia* (bin 23), *Clostridia* (bin 30), *Gemmatimonadetes* (bins 55 and 53), *Phycisphaerae* (bin 22), *Burkholderiales* (bin 52), *Rhodocyclales* (bin 45), *Nevskiales* (bin 37), and *Xanthomonadales* (bin 59). Abundant MAGs in Reactor 1 did not necessarily harbor either of the *nosZ* genes. It is important to note that all of the retrieved MAGs harboring the *nosZ* gene did not possess a complete set of denitrifying genes, suggesting that they were non-denitrifying genotypes.

MAGs (bins 48, 49, 50, and 51) were affiliated within *Ignavibacteriaceae*, a taxon containing bacteria frequently detected in anammox reactors that has been suggested to function as a N_2O sink (Lawson *et al.*, 2017) harboring the Clade II *nosZ* gene and dissimilatory nitrate reduction to ammonium (DNRA) gene (*nirBD* and/or *nrfAH*). In addition, bins 48 and 51 carried the *narGHI* genes encoding nitrate reductase. The MAG assigned to *Bacteroidia* (bin 79), which has potential as a N_2O sink, harbored the Clade II *nosZ* gene and possessed the *norBC* and *narGHI* genes.

Anaerolineae, one of the most dominant non-anammox bacterial genera, was annotated with the *narGHI* gene, *nirS* and *nirK* genes, and a group of genes related to DNRA. *Dehalococcoidia* harbored the nitrate reductase *nar* gene, the *nir* gene, and the Clade II *nosZ* gene. *Gemmatimonadetes* (bins 55 and 53), showing a highly transcribed Clade II *nosZ* gene in anammox reactors as previously reported (Suenaga *et al.*, 2021), contained the *nirK* gene encoding nitrite reductase, the *narGHI* gene, and genes involved in DNRA. The MAGs of *Planctomycetota* (bins 77, 58, and 47) were also annotated with the Clade II *nosZ* gene. The MAGs abundantly detected in Reactor 1 (w/N_2O) did not necessarily harbor any *nosZ* (*i.e.*, bins 19, 60, 6, 45, 37, and 59). The abundance of the *nosZ*-coding sequences collected from the assembled contigs revealed that the higher RPKM values derived from Clade II *nosZ* (174.2 RPKM in Reactor 1 and 164.0 RPKM in Reactor 2)

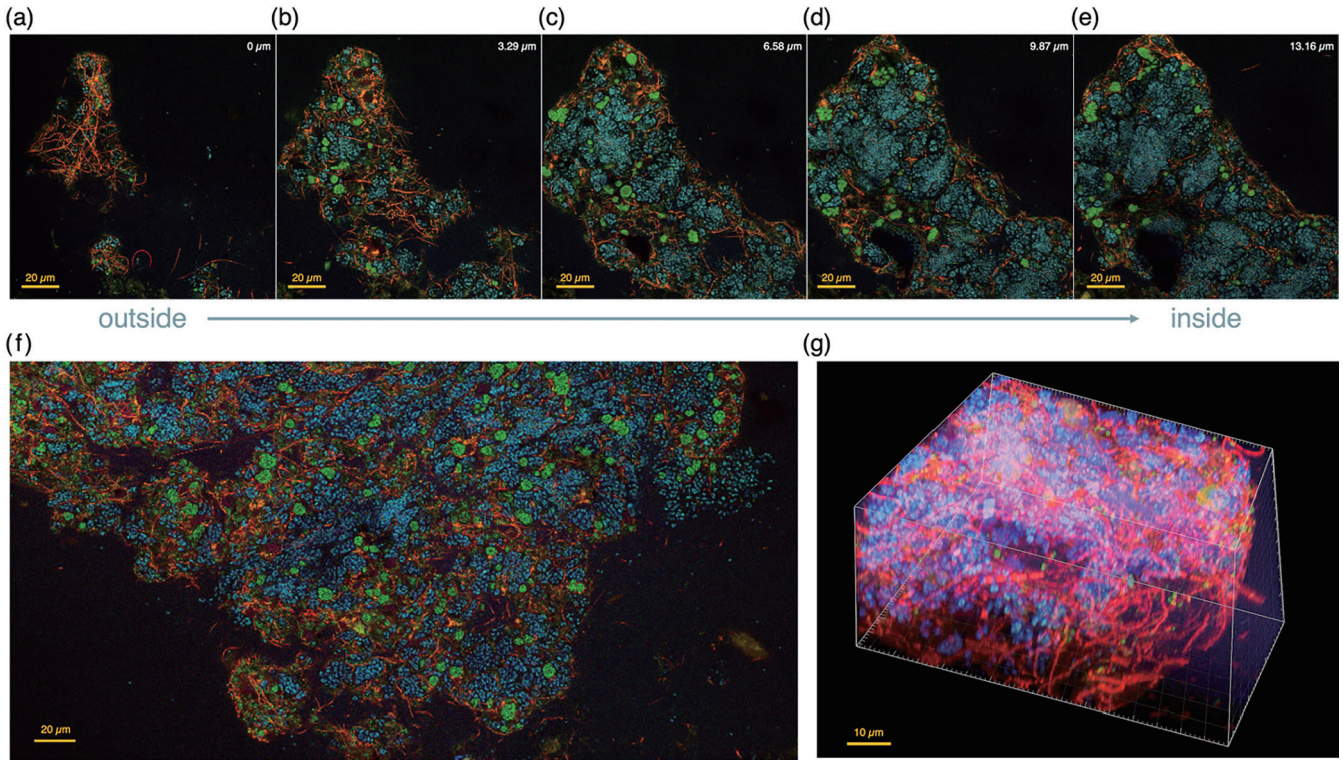


Fig. 4. Spatial distributions of anammox and *Chloroflexi* by fluorescence *in situ* hybridization, taken by confocal laser scanning microscopy under 630× magnification. Anammox and *Chloroflexi* were hybridized by probes of Amx 368 and CFX1228+GNSB941 with fluorochromes of Cy5 and Cy3, counterstained by the EUB338 mix with the fluorochrome of FITC. Cy5 is shown in blue, Cy3 in red, and FITC in green. Panels a, b, c, d, e, and f: Reactor 1 (with N_2O supply). Panel g: Reactor 2 (without N_2O supply). In Panels a, b, c, d, and e, the number in the upper right in each image indicates the distance from the aggregate surface.

than from Clade I *nosZ* (56.5 RPKM in Reactor 1 and 44.5 RPKM in Reactor 2) were retrieved (Fig. S9).

Genes involved in the metabolism and recycling of peptidoglycans involved in carbon degradation and alpha-amylase and beta-glucosidase, which play a role in the metabolism of polysaccharides, a constitute of extracellular polymeric substances (EPS), were annotated in *Armatimonadota* (bins 61, 19, and 81), *Ignavibacteria*, *Gemmatimonadetes*, *Myxococcota*, and *Planctomycetota*. *Chloroflexota* MAGs differed in their genotypic patterns associated with carbon degradation at the phylum level. For example, the *Anaerolineae* MAG (bin 60) possessed the genes encoding enzymes for peptidoglycan degradation, alpha-amylase and beta-glucosidase, whereas the *Dehalococcoidia* MAG (bin 80) carried an incomplete set of functional genes encoding enzymes for peptidoglycan degradation (Fig. 5).

The genes involved in cobalamin biosynthesis (Corrin ring biosynthesis and adenylation and nucleotide loop assembly) were well annotated in *Dehalococcoidia*, *Clostridia* (bin 30), and *Brocadiae* MAGs, suggesting the presence of important vitamin B12 producers in the anammox community. The coverage of these MAGs was *Dehalococcoidia* (bin 85: 0.08%, bin 23: 1.73%) and *Clostridia* (bin 30: 0.23%). *Burkholderiales* (bins 66, 72, and 78), *Anaerolineae*, and *Nitrospiraceae* (bin 67) MAGs harbored the genes responsible for a vitamin B12 transporter and nucleotide loop assembly for Cobalamin biosynthesis (3) with high annotation (>60%).

Discussion

Our reactor design with a flat-sheet gas-permeable membrane allowed a bubbleless N_2O supply to the biomasses in the MBfRs. Continuous incubations by the MBfRs successfully provided environments with high (Reactor 1) and low (Reactor 2) concentrations of N_2O (Fig. S6 and S7). The concentration of N_2O in Reactor 1 (3.96 mg N L^{-1} on average, *i.e.*, 141 μM) was more than one order of magnitude higher than the apparent half-saturation constant for N_2O ($K_{m,N2O}$) of 9.50±3.0 μM (0.266±0.083 mg N L^{-1}) of an enriched anammox biomass (Suenaga *et al.*, 2021). Therefore, this ensures an environment in which fast-growing N_2O reducers with a low affinity for N_2O, but a high N_2O consumption rate (Andrews and Harris, 1986; Yin *et al.*, 2022) preferentially grow. This study demonstrated that the application of an MBfR concept (Nerenberg, 2016) to enrich highly efficient N_2O-reducing bacteria in an anammox biomass secures a stable environment with high or low concentrations of N_2O with the biomass. The maintenance of a different N_2O level for a long period was essential in the present study, while we previously conducted a short-term (within 1 day) biomass exposure to a high N_2O concentration (Suenaga *et al.*, 2021).

By using the MBfR concept, the present study investigated whether an external N_2O supply led to the dominance of fast-growing N_2O-reducing bacteria. According to the microbial community compositions in the two MBfRs with and without the external N_2O supply (Fig. 3), slight changes

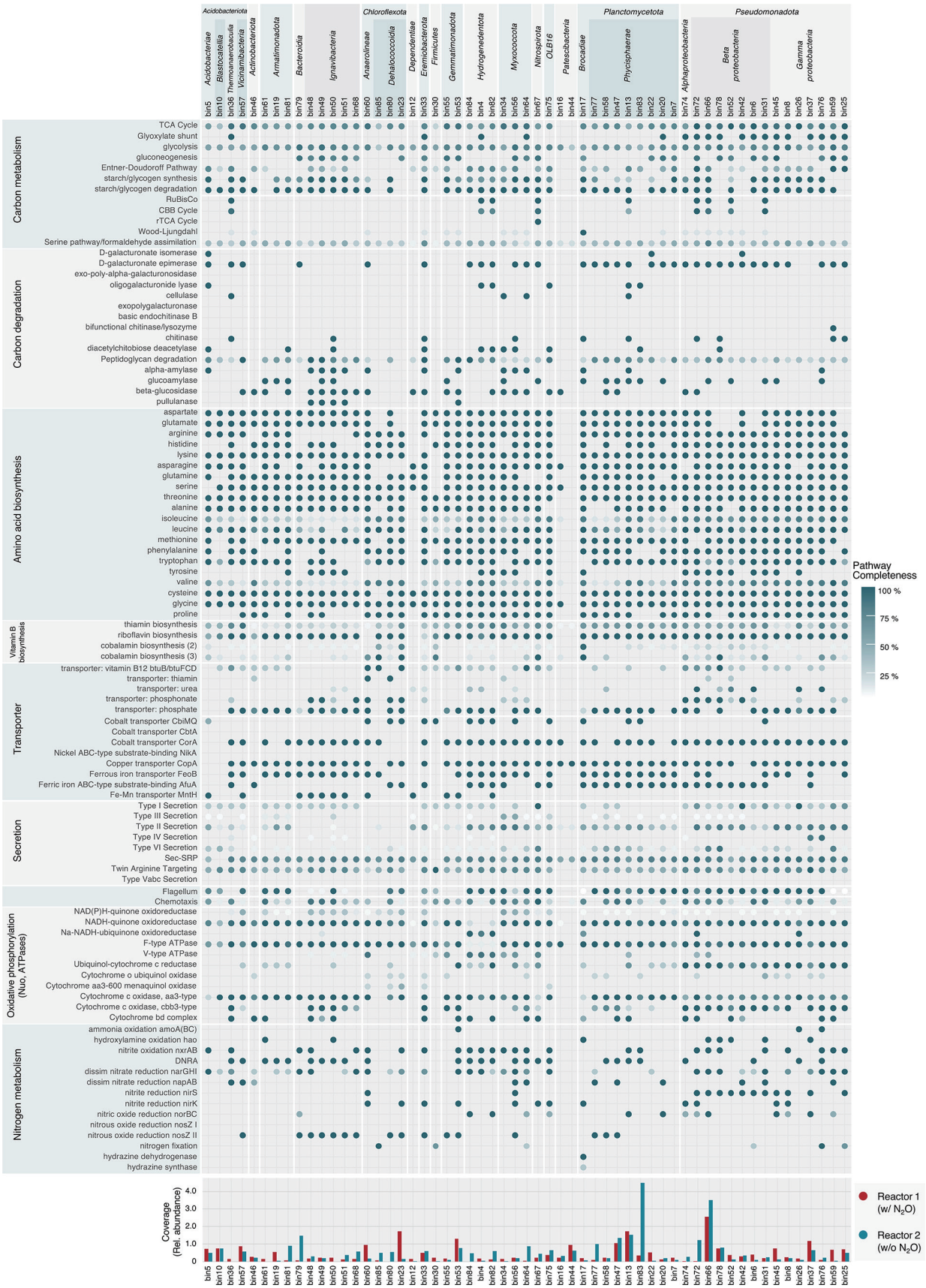


Fig. 5. Pathway coverage of high quality (completeness ≥70%, contamination <5%) metagenome-assembled genomes. The dot indicates metabolic pathway completeness calculated using the KEGG Decoder based on KEGG Orthology (KO) results assigned by DFAST, EggNOG mapper, and KofamKOALA.

in microbial community compositions were attained. This result refutes our hypothesis that an external N_2O supply leads to the dominance of fast-growing N_2O-reducing bacteria. Nevertheless, the biomass from Reactor 1 showed a higher N_2O consumption rate than that from Reactor 2 (Fig. 1). The quantitative values of Clade I and Clade II *nosZ* genes indicated that the supply of N_2O as an external electron acceptor was advantageous to the bacterial community harboring *nosZ*. This effect was significantly pronounced for Clade II *nosZ* (Fig. 2f). Although the contribution of the external N_2O supply to microbial community changes may be limited (Fig. 3), the N_2O supply was instrumental in the increases observed in Clade II *nosZ* gene abundance (Fig. 2f) and intrinsic N_2O consumption activities (Fig. 1).

An autotrophic environment with a high nitrogen concentration likely favors Clade II *nosZ* N_2O-reducing bacteria. This notion was supported by the present results showing that Clade II *nosZ* N_2O-reducing bacteria were consistently one order of magnitude and three-fold more abundant than the Clade I type in MBfRs regardless of the supply of N_2O by qPCR (Fig. 2e and 2f) and the metagenomic analysis (Fig. S9), respectively. This result may be attributed to an inoculum being highly enriched under autotrophic conditions (Suenaga *et al.*, 2021). Previous studies on N_2O-reducing bacteria harboring either Clade I or Clade II *nosZ* genes indicated that Clade II *nosZ* bacteria exhibited higher affinities (Yoon *et al.*, 2016; Suenaga *et al.*, 2019), which was not consistent with our results, where even a high concentration of N_2O allowed a higher abundance of Clade II *nosZ* bacteria. Clade II *nosZ* bacteria were found to be dominant during an incubation with a N_2O supply (Suenaga *et al.*, 2019). These findings and the present results indicate that the concentration of N_2O during the enrichment of N_2O-reducing bacteria is not the sole factor affecting the predominance of the clade of the *nosZ* gene.

Anaerolineae and *Ignavibacteria*, which were dominant among non-anammox bacteria in the present study (Fig. 3), may be critical N_2O reducers in anammox systems. They reportedly utilize endogenous organic matter when fed an organic-free medium, and are also the most prevalent when fermentable organic matter is supplied at low C/N ratios (Xiao *et al.*, 2021). *Ignavibacteria* MAGs (bins 48, 49, 50, and 51) and *Anaerolineae* MAG (bin 80) were annotated with metabolic pathways regarding soluble microbial products (SMP) by KEGG Decoder (Ni *et al.*, 2012), including EPS composed of polysaccharides and proteins secreted by anammox bacteria (Hou *et al.*, 2015; Ali *et al.*, 2018). The annotation count by KEGG Decoder accounted for over 65% of the total metabolic pathways. In addition, these associated genes were functionally assigned to carbon degradation to utilize cell debris (Fig. 5). Therefore, these taxa were likely to be N_2O consumers using SMP produced by anammox and others in the anammox reactor.

SBR1031 belonging to the class *Anaerolineae* showed a higher relative abundance in the biofilm grown on membrane surfaces (Fig. S8) with presumably high N_2O concentrations. The results of the metagenomic analysis suggest that this genus contains bacteria harboring the Clade II *nosZ* gene (Bovio-Winkler *et al.*, 2023), which may have a promising N_2O consumption capacity. Therefore, the class *Anaerolineae* has potential as a N_2O sink in the anammox system that is not susceptible to high N_2O concentrations. In contrast, a lower abundance of the family *Brocadiaceae* in the biofilm than the deposited aggregates on the biofilm outer surface (Fig. S8) suggested that a high N_2O concentration condition was unfavorable for anammox growth.

The phylum *Chloroflexota* coexisting with anammox bacteria was detected at the periphery of the biomass, surrounding anammox cell aggregates (Fig. 4). This spatial arrangement did not appear to change with or without the N_2O supply, which is consistent with previous findings (Wong *et al.*, 2023). In addition, filamentous *Chloroflexota* has been suggested to function as a junction for biomass aggregates and degrade EPS secreted by anammox bacteria, which are responsible for decomposing EPS to SMP available to other heterotrophic bacteria (Wong *et al.*, 2023). Our *Anaerolineae* MAG (bin 60) was also well annotated with pathways (alpha-amylase and beta-glucosidase) involved in the degradation of polysaccharide chains (Oshiki *et al.*, 2022a) (Fig. 5). Our FISH analysis also showed fluorescence derived from neither *Planctomycetota* nor *Chloroflexota* in the vicinity of *Chloroflexota* (Fig. 4), and its spatial coordination was similar to that observed in a previous study (Wong *et al.*, 2023). Therefore, the preference and availability of recalcitrant organic matter were deemed necessary for elucidating the composition of the non-anammox bacterial population. This unexplored population may be one of the reasons why N_2O-reducing bacteria with faster N_2O consumption rates were more dominant than N_2O-producing bacteria.

Although the phyla *Gemmatimonadota* and non-anammox *Planctomycetota* were not predominant in the microbial community, they were frequently detected taxa in Reactor 1 (Fig. 5), suggesting their involvement in another N_2O sink. These phyla exhibit high transcription activities of the *nosZ* gene in anammox processes (Park *et al.*, 2017; Suenaga *et al.*, 2021). The higher coverage of MAGs belonging to these strains in Reactor 1 suggested that the physiological traits matched the environment in the MBfR with external N_2O supply, alluding to N_2O reducers favorable in environments with high N_2O concentrations. The production of N_2O during denitrification, often found in anammox processes as a minor microbial reaction among nitrogen transformation, may be attributed to a N_2O production rate via the reduction of NO_3^- and NO_2^- to N_2O, overwhelming the N_2O consumption rate. A *Gemmatimonadetes* MAG (bin 53), harboring *narGHI* and Clade II *nosZ* genes, possessed the metabolic potentials of NO_3^- and N_2O reduction (Fig. 5), potentially consuming N_2O depending on N_2O concentrations according to the biokinetics in previous study (Oshiki *et al.*, 2022b). Moreover, non-anammox *Planctomycetes* (bins 77 and 47) possessed two denitrification-related genes, the *napAB* gene encoding nitrate reductase and the Clade II *nosZ* gene (Fig. 5). When bacteria carry the *nap* and *nos* genes, they preferentially use N_2O as an electron acceptor more than those with *nar* and *nos* genes (Gao *et al.*, 2021; Oba *et al.*, 2022), suggesting that non-anammox planctomycetes harboring *nosZ* genes are preferential N_2O sinks.

Theoretically, there has been a concern about the loss of nitrogen removal performance due to reduced microbial

activity caused by N_2O, which has yet to be examined in detail. Excessive N_2O in Reactor 1 may react with cob(I)alamin (vitamin B12) as previously reported (Drummond and Matthews, 1994), thereby inhibiting the function of the cobalamin-dependent enzyme (Shelton *et al.*, 2019). The N_2O level in Reactor 1 (3.96 mg N L^{-1}) exceeded the concentration (>2.8 mg N L^{-1}) at which N_2O reacts with MetH methionine synthase, the most widely used cobalamin-dependent enzyme, thereby retarding the cell growth of *P. denitrificans* (Sullivan *et al.*, 2013). To ameliorate the inhibition of N_2O, *P. denitrificans* and *Dehalococcoides mccartyi* either activate vitamin B12-independent MetE methionine synthase or maintain its activity by an endogenous and exogenous vitamin B12 supply (Sullivan *et al.*, 2013; Yin *et al.*, 2019). *Bacteroidia* and *Anaerolineae* (bins 79 and 60), predominant taxa based on 16S rRNA gene amplicon sequencing (Fig. 3), did not possess a *metE* gene encoding vitamin B12-independent MetE methionine synthase (Fig. 5). Bacteria possessing only vitamin B12-dependent enzymes may be disadvantageous for survival under excessive N_2O conditions. Nevertheless, their abundance did not decrease in the presence of excessive N_2O in Reactor 1 (Fig. 3). The results of the FISH analysis also indicated that the dominance of *Chloroflexota* was sustained (Fig. 4). Therefore, the anammox community in Reactor 1 appeared to have a bypassing mechanism by which vitamin B12 was relayed from its producers to acceptors. Further studies are needed to elucidate the underlying mechanisms.

Previous studies reported that *de novo* cobamide biosynthesis consists of approximately 30 steps, which may be broadly divided into the following steps: tetrapyrrole precursor biosynthesis, (aerobic/anaerobic) corrin ring biosynthesis and adenylation, and nucleotide loop assembly (Shelton *et al.*, 2019; Lu *et al.*, 2020; Balabanova *et al.*, 2021). Furthermore, among an anammox community, only *Brocadia* sp. synthesizes cobalamin (Lawson *et al.*, 2017; Keren *et al.*, 2020). In the present study, only anammox *Planctomycetes* (bin 17) had complete *de novo* biosynthetic pathways, which was consistent with previous findings.

The results obtained herein demonstrated that *Dehalococcoidia* (bins 23 and 85) and *Clostridia* (bin 30) carried a nucleotide loop assembly pathway following the cobalamin precursor synthesis pathway. Some bacteria have been shown to partially possess *de novo* cobalamin biosynthesis pathways (Shelton *et al.*, 2019; Lu *et al.*, 2020). As shown in physiological studies (Schipp *et al.*, 2013; Shelton *et al.*, 2019), some *Dehalococcoidia* and *Clostridia* synthesize cobamides from salvaged precursors or change the cobamide structure for their growth. The results of our metagenomic analysis suggest that Clade II *nosZ* N_2O-reducing bacteria, such as *Dehalococcoidia* and *Clostridia*, retain many genes associated with the cobalamin synthesis pathway. Putatively, they are prone to endogenously synthesize cobalamin and/or exogenously receive cobalamin supplied by *Brocadia* as a survival strategy. The present results imply the strong contribution of anammox bacteria as a *de novo* cobalamin producer to relay cobalamin to coexisting bacteria missing its production functions, thereby restricting bacteria from cobalamin inhibition caused by N_2O in the anammox community.

Conclusion

In the present study, we operated MBfRs that were externally supplied with N_2O for 1200 days for the continuous incubation of bacteria consuming N_2O. We demonstrated the long-term maintenance of a high abundance of Clade II *nosZ* and an increased capacity for N_2O consumption in the community enriched by an exogenous N_2O supply. On the other hand, slight differences were attained in the dominant taxa in the MBfRs with or without an external N_2O supply, which indicated that supplying N_2O as an additional electron acceptor did not necessarily result in the dominance of fast-growing N_2O-reducing bacteria. Therefore, other factors, particularly electron donors, limited the growth rate of heterotrophic N_2O-reducing bacteria. Despite growth-limiting conditions for N_2O-reducing bacteria, the metagenomic analysis revealed significant functions to acquire organic matter and vitamins that serve as electron donors. *Anaerolineae* and *Ignavibacteria*, regarded as parts of the core microbiome and dominant in the MBfR supplied with N_2O, have a broad availability of organic matter and survive at high N_2O concentrations. *Dehalococcoidia* and *Clostridia* have partial vitamin B12 production pathways, likely carrying a survival strategy in excessive N_2O environments. Moreover, we found that non-anammox *Planctomycetes* appeared to contribute to N_2O consumption despite their low abundance in the anammox community. Since the concentration of N_2O in the MBfR system supplied with N_2O was markedly higher than that often observed in anammox reactors, there have been concerns that the anammox community may lose its activity due to the inhibition of N_2O. Nevertheless, the N_2O-fed reactor operation underpinned the anammox community, exerting a stable N_2O sink at excessively high N_2O concentrations. Our in-depth analysis of microbial community compositions and functions provide insights on the as-yet-unknown functions of non-denitrifying Clade II *nosZ* bacteria. These ecophysiological descriptions will facilitate the utilization of N_2O-reducing bacteria inhabiting engineered systems that emit N_2O, such as anammox-based processes, and may be expanded to agricultural fields requiring N_2O mitigation. The may lead to the development of countermeasures against N_2O emissions, *e.g.*, bioaugmentation by inoculating biomasses enriching N_2O reducers.

Acknowledgements

We thank the late Ms. Kanako Mori for her experimental support. This research was funded by Grants-in-Aid for Scientific Research (Grant no. 20H04362 and 23H03565), Fostering Joint International Research (20KK0243) from the Japan Society for the Promotion of Science (JSPS), and the Kurita Water and Environment Foundation (22T012).

References

Al-Hazmi, H.E., Lu, X., Grubba, D., Majtacz, J., Badawi, M., and Mąkinia, J. (2023) Sustainable nitrogen removal in anammox-mediated systems: Microbial metabolic pathways, operational conditions and mathematical modelling. *Sci Total Environ* **868**: 161633.

Ali, M., Shaw, D.R., Zhang, L., Haroon, M.F., Narita, Y., Emwas, A.H., *et al.* (2018) Aggregation ability of three phylogenetically distant anammox bacterial species. *Water Res* **143**: 10–18.

Andrews JH, Harris RF. r- and K-Selection and Microbial Ecology. In: Marshall KC, editor. Advances in Microbial Ecology. Boston, MA: Springer US; 1986. p. 99-147.

Aramaki, T., Blanc-Mathieu, R., Endo, H., Ohkubo, K., Kanehisa, M., Goto, S., and Ogata, H. (2020) KofamKOALA: KEGG Ortholog assignment based on profile HMM and adaptive score threshold. *Bioinformatics* **36**: 2251–2252.

Balabanova, L., Averianova, L., Marchenok, M., Son, O., and Tekutyeva, L. (2021) Microbial and genetic resources for cobalamin (vitamin B12) biosynthesis: From ecosystems to industrial biotechnology. *Int J Mol Sci* **22**: 4522.

Blum, M., Chang, H.Y., Chuguransky, S., Grego, T., Kandasaamy, S., Mitchell, A., *et al.* (2021) The InterPro protein families and domains database: 20 years on. *Nucleic Acids Res* **49**: D344–D354.

Bokulich, N.A., Kaehler, B.D., Rideout, J.R., Dillon, M., Bolyen, E., Knight, R., *et al.* (2018) Optimizing taxonomic classification of marker-gene amplicon sequences with QIIME 2's q2-feature-classifier plugin. *Microbiome* **6**: 90.

Bolyen, E., Rideout, J.R., Dillon, M.R., Bokulich, N.A., Abnet, C.C., Al-Ghalith, G.A., *et al.* (2019) Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat Biotechnol* **37**: 852–857.

Bovio-Winkler, P., Guerrero, L.D., Erijman, L., Oyarzua, P., Suarez-Ojeda, M.E., Cabezas, A., and Etchebehere, C. (2023) Genome-centric metagenomic insights into the role of Chloroflexi in anammox, activated sludge and methanogenic reactors. *BMC Microbiol* **23**: 45.

Bushnell, B. (2014) BBTools software package. URL http://sourceforge.net/projects/bbmap

Butler JM. Chapter 2 - DNA Extraction Methods. In: Butler JM, editor. Advanced Topics in Forensic DNA Typing: Methodology. San Diego: Academic Press; 2012. p. 29–47.

Cantalapiedra, C.P., Hernández-Plaza, A., Letunic, I., Bork, P., and Huerta-Cepas, J. (2021) eggNOG-mapper v2: Functional annotation, orthology assignments, and domain prediction at the metagenomic scale. *Mol Biol Evol* **38**: 5825–5829.

Chan, P.P., Lin, B.Y., Mak, A.J., and Lowe, T.M. (2021) tRNAscan-SE 2.0: improved detection and functional classification of transfer RNA genes. *Nucleic Acids Res* **49**: 9077–9096.

Chaumeil, P.-A., Mussig, A.J., Hugenholtz, P., and Parks, D.H. (2022) GTDB-Tk v2: memory friendly classification with the genome taxonomy database. *Bioinformatics* **38**: 5315–5316.

Chen, S.F., Zhou, Y.Q., Chen, Y.R., and Gu, J. (2018) fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* **34**: 884–890.

Danecek, P., Bonfield, J.K., Liddle, J., Marshall, J., Ohan, V., Pollard, M.O., *et al.* (2021) Twelve years of SAMtools and BCFtools. *GigaScience* **10**: giab008.

deGraaf, A.A.V., deBruijn, P., Robertson, L.A., Jetten, M.S.M., and Kuenen, J.G. (1996) Autotrophic growth of anaerobic ammonium-oxidizing micro-organisms in a fluidized bed reactor. *Microbiology (Reading)* **142**: 2187–2196.

Domingo-Felez, C., and Smets, B.F. (2019) Regulation of key N(2)O production mechanisms during biological water treatment. *Curr Opin Biotechnol* **57**: 119–126.

Drummond, J.T., and Matthews, R.G. (1994) Nitrous oxide degradation by cobalamin-dependent methionine synthase: characterization of the reactants and products in the inactivation reaction. *Biochemistry* **33**: 3732–3741.

Duan, H., Zhao, Y., Koch, K., Wells, G.F., Zheng, M., Yuan, Z., and Ye, L. (2021) Insights into nitrous oxide mitigation strategies in wastewater treatment and challenges for wider implementation. *Environ Sci Technol* **55**: 7208–7224.

Fenu, A., Smolders, S., De Gussem, K., and Weemaes, M. (2019) Conflicting carbon footprint and energy saving in a side-stream Anammox Process. *Biochem Eng J* **151**: 107336.

Frutos, O.D., Quijano, G., Pérez, R., and Muñoz, R. (2016) Simultaneous biological nitrous oxide abatement and wastewater treatment in a denitrifying off-gas bioscrubber. *Chem Eng J* **288**: 28–37.

Gao, Y., Mania, D., Mousavi, S.A., Lycus, P., Arntzen, M.O., Woliy, K., *et al.* (2021) Competition for electrons favours N2 O reduction in denitrifying Bradyrhizobium isolates. *Environ Microbiol* **23**: 2244–2259.

Gilbert, E.M., Agrawal, S., Schwartz, T., Horn, H., and Lackner, S. (2015) Comparing different reactor configurations for Partial Nitritation/Anammox at low temperatures. *Water Res* **81**: 92–100.

Graham, E.D., Heidelberg, J.F., and Tully, B.J. (2018) Potential for primary productivity in a globally-distributed bacterial phototroph. *ISME J* **12**: 1861–1866.

Han, H., Kim, D.D., Song, M.J., Yun, T., Yoon, H., Lee, H.W., *et al.* (2023) Biotrickling filtration for the reduction of N(2)O emitted during wastewater treatment: results from a long-term in situ pilot-scale testing. *Environ Sci Technol* **57**: 3883–3892.

Hou, X., Liu, S., and Zhang, Z. (2015) Role of extracellular polymeric substance in determining the high aggregation ability of anammox sludge. *Water Res* **75**: 51–62.

Hyatt, D., Chen, G.L., Locascio, P.F., Land, M.L., Larimer, F.W., and Hauser, L.J. (2010) Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinf* **11**: 119.

Isobe, K., Suwa, Y., Ikutani, J., Kuroiwa, M., Makita, T., Takebayashi, Y., *et al.* (2011) Analytical techniques for quantifying (15)N/(14)N of nitrate, nitrite, total dissolved nitrogen and ammonium in environmental samples using a gas chromatograph equipped with a quadrupole mass spectrometer. *Microbes Environ* **26**: 46–53.

Jones, P., Binns, D., Chang, H.Y., Fraser, M., Li, W., McAnulla, C., *et al.* (2014) InterProScan 5: genome-scale protein function classification. *Bioinformatics* **30**: 1236–1240.

Kanehisa, M., and Goto, S. (2000) KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res* **28**: 27–30.

Kang, D.D., Li, F., Kirton, E., Thomas, A., Egan, R., An, H., and Wang, Z. (2019) MetaBAT 2: an adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies. *PeerJ* **7**: e7359.

Keren, R., Lawrence, J.E., Zhuang, W., Jenkins, D., Banfield, J.F., Alvarez-Cohen, L., *et al.* (2020) Increased replication of dissimilatory nitrate-reducing bacteria leads to decreased anammox bioreactor performance. *Microbiome* **8**: 7.

Kindaichi, T., Yuri, S., Ozaki, N., and Ohashi, A. (2012) Ecophysiological role and function of uncultured Chloroflexi in an anammox reactor. *Water Sci Technol* **66**: 2556–2561.

Kinh, C.T., Suenaga, T., Hori, T., Riya, S., Hosomi, M., Smets, B.F., and Terada, A. (2017) Counter-diffusion biofilms have lower N(2)O emissions than co-diffusion biofilms during simultaneous nitrification and denitrification: Insights from depth-profile analysis. *Water Res* **124**: 363–371.

Lackner, S., Gilbert, E.M., Vlaeminck, S.E., Joss, A., Horn, H., and van Loosdrecht, M.C. (2014) Full-scale partial nitritation/anammox experiences--an application survey. *Water Res* **55**: 292–303.

Lawson, C.E., Wu, S., Bhattacharjee, A.S., Hamilton, J.J., McMahon, K.D., Goel, R., and Noguera, D.R. (2017) Metabolic network analysis reveals microbial community interactions in anammox granules. *Nat Commun* **8**: 12.

Lawson, C.E., Nuijten, G.H.L., de Graaf, R.M., Jacobson, T.B., Pabst, M., Stevenson, D.M., *et al.* (2021) Autotrophic and mixotrophic metabolism of an anammox bacterium revealed by in vivo (13)C and (2)H metabolic network mapping. *ISME J* **15**: 673–687.

Li, D., Liu, C.M., Luo, R., Sadakane, K., and Lam, T.W. (2015) MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics* **31**: 1674–1676.

Li, J., Li, J., Gao, R., Wang, M., Yang, L., Wang, X., *et al.* (2018) A critical review of one-stage anammox processes for treating industrial wastewater: Optimization strategies based on key functional microorganisms. *Bioresour Technol* **265**: 498–505.

Li, W., and Godzik, A. (2006) Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* **22**: 1658–1659.

Liao, Y., Smyth, G.K., and Shi, W. (2013) featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**: 923–930.

Lu, X., Heal, K.R., Ingalls, A.E., Doxey, A.C., and Neufeld, J.D. (2020) Metagenomic and chemical characterization of soil cobalamin production. *ISME J* **14**: 53–66.

Magoc, T., and Salzberg, S.L. (2011) FLASH: fast length adjustment of short reads to improve genome assemblies. *Bioinformatics* **27**: 2957–2963.

McCarty, P.L. (2018) What is the best biological process for nitrogen removal: When and why? *Environ Sci Technol* **52**: 3835–3841.

McMurdie, P.J., and Holmes, S. (2013) phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS One* **8**: e61217.

Nerenberg, R. (2016) The membrane-biofilm reactor (MBfR) as a counter-diffusional biofilm process. *Curr Opin Biotechnol* **38**: 131–136.

Ni, B.J., Ruscalleda, M., and Smets, B.F. (2012) Evaluation on the microbial interactions of anaerobic ammonium oxidizers and heterotrophs in Anammox biofilm. *Water Res* **46**: 4645–4652.

Nurk, S., Meleshko, D., Korobeynikov, A., and Pevzner, P.A. (2017) metaSPAdes: a new versatile metagenomic assembler. *Genome Res* **27**: 824–834.

Oba, K., Suenaga, T., Kuroiwa, M., Riya, S., and Terada, A. (2022) Exploring the functions of efficient canonical denitrifying bacteria as N_2O sinks: Implications from (15)N tracer and transcriptome analyses. *Environ Sci Technol* **56**: 11694–11706.

Okabe, S., Oshiki, M., Takahashi, Y., and Satoh, H. (2011) N_2O emission from a partial nitrification-anammox process and identification of a key biological process of N_2O emission from anammox granules. *Water Res* **45**: 6461–6470.

Olm, M.R., Brown, C.T., Brooks, B., and Banfield, J.F. (2017) dRep: a tool for fast and accurate genomic comparisons that enables improved genome recovery from metagenomes through de-replication. *ISME J* **11**: 2864–2868.

Oshiki, M., Takaki, Y., Hirai, M., Nunoura, T., Kamigaito, A., and Okabe, S. (2022a) Metagenomic analysis of five phylogenetically distant anammox bacterial enrichment cultures. *Microbes Environ* **37**: ME22017.

Oshiki, M., Toyama, Y., Suenaga, T., Terada, A., Kasahara, Y., Yamaguchi, T., and Araki, N. (2022b) N(2)O reduction by gemmatimonas aurantiaca and potential involvement of gemmatimonadetes bacteria in N(2)O reduction in agricultural soils. *Microbes Environ* **37**: ME21090.

Park, D., Kim, H., and Yoon, S. (2017) Nitrous oxide reduction by an obligate aerobic bacterium, Gemmatimonas aurantiaca strain T-27. *Appl Environ Microbiol* **83**: e00502-17.

Parks, D.H., Chuvochina, M., Rinke, C., Mussig, A.J., Chaumeil, P.A., and Hugenholtz, P. (2022) GTDB: an ongoing census of bacterial and archaeal diversity through a phylogenetically consistent, rank normalized and complete genome-based taxonomy. *Nucleic Acids Res* **50**: D785–D794.

Peng, L., Carvajal-Arroyo, J.M., Seuntjens, D., Prat, D., Colica, G., Pintucci, C., and Vlaeminck, S.E. (2017) Smart operation of nitritation/denitritation virtually abolishes nitrous oxide emission during treatment of co-digested pig slurry centrate. *Water Res* **127**: 1–10.

Prodan, A., Tremaroli, V., Brolin, H., Zwinderman, A.H., Nieuwdorp, M., and Levin, E. (2020) Comparing bioinformatic pipelines for microbial 16S rRNA amplicon sequencing. *PLoS One* **15**: e0227434.

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., *et al.* (2013) The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res* **41**: D590–596.

Riya, S., Zhou, S., Watanabe, Y., Sagehashi, M., Terada, A., and Hosomi, M. (2012) CH4 and N_2O emissions from different varieties of forage rice (Oryza sativa L.) treating liquid cattle waste. *Sci Total Environ* **419**: 178–186.

Robeson, M.S., 2nd, O'Rourke, D.R., Kaehler, B.D., Ziemski, M., Dillon, M.R., Foster, J.T., and Bokulich, N.A. (2021) RESCRIPt: Reproducible sequence taxonomy reference database management. *PLoS Comput Biol* **17**: e1009581.

Sanford, R.A., Wagner, D.D., Wu, Q., Chee-Sanford, J.C., Thomas, S.H., Cruz-García, C., *et al.* (2012) Unexpected nondenitrifier nitrous oxide reductase gene diversity and abundance in soils. *Proc Natl Acad Sci U S A* **109**: 19709–19714.

Schipp, C.J., Marco-Urrea, E., Kublik, A., Seifert, J., and Adrian, L. (2013) Organic cofactors in the metabolism of Dehalococcoides mccartyi strains. *Philos Trans R Soc Lond B Biol Sci* **368**: 20120321.

Seemann, T. (2018) barrnap 0.9: rapid ribosomal RNA prediction. URL: https://github.com/tseemann/barrnap

Shan, J., Sanford, R.A., Chee-Sanford, J., Ooi, S.K., Loffler, F.E., Konstantinidis, K.T., and Yang, W.H. (2021) Beyond denitrification: The role of microbial diversity in controlling nitrous oxide reduction and soil nitrous oxide emissions. *Glob Chang Biol* **27**: 2669–2683.

Shelton, A.N., Seth, E.C., Mok, K.C., Han, A.W., Jackson, S.N., Haft, D.R., and Taga, M.E. (2019) Uneven distribution of cobamide biosynthesis and dependence in bacteria predicted by comparative genomics. *ISME J* **13**: 789–804.

Shen, W., Le, S., Li, Y., and Hu, F. (2016) SeqKit: A cross-platform and ultrafast toolkit for FASTA/Q file manipulation. *PLoS One* **11**: e0163962.

Sieber, C.M.K., Probst, A.J., Sharrar, A., Thomas, B.C., Hess, M., Tringe, S.G., and Banfield, J.F. (2018) Recovery of genomes from metagenomes via a dereplication, aggregation and scoring strategy. *Nat Microbiol* **3**: 836–843.

Smith, C., Nicholls, Z.R.J., Armour, K., Collins, W., Forster, P., Meinshausen, M., *et al.* (2021) The Earth's Energy Budget, Climate Feedbacks, and Climate Sensitivity Supplementary Material. *In Climate Change 2021. The Physical Science Basis*. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. Masson-Delmotte, V., Zhai, P., Pirani, A., Connors, S.L., Péan, C., Berger, S., *et al.*, (eds). URL: https://www.ipcc.ch/

Speth, D.R., In't Zandt, M.H., Guerrero-Cruz, S., Dutilh, B.E., and Jetten, M.S. (2016) Genome-based microbial ecology of anammox granules in a full-scale wastewater treatment system. *Nat Commun* **7**: 11172.

Strous, M., Heijnen, J.J., Kuenen, J.G., and Jetten, M.S.M. (1998) The sequencing batch reactor as a powerful tool for the study of slowly growing anaerobic ammonium-oxidizing microorganisms. *Appl Microbiol Biotechnol* **50**: 589–596.

Suenaga, T., Hori, T., Riya, S., Hosomi, M., Smets, B.F., and Terada, A. (2019) Enrichment, isolation, and characterization of high-affinity N_2O-reducing bacteria in a gas-permeable membrane reactor. *Environ Sci Technol* **53**: 12101–12112.

Suenaga, T., Ota, T., Oba, K., Usui, K., Sako, T., Hori, T., *et al.* (2021) Combination of (15)N tracer and microbial analyses discloses N_2O sink potential of the anammox community. *Environ Sci Technol* **55**: 9231–9242.

Sullivan, M.J., Gates, A.J., Appia-Ayme, C., Rowley, G., and Richardson, D.J. (2013) Copper control of bacterial nitrous oxide emission and its impact on vitamin B12-dependent metabolism. *Proc Natl Acad Sci U S A* **110**: 19926–19931.

Tanizawa, Y., Fujisawa, T., and Nakamura, Y. (2018) DFAST: a flexible prokaryotic genome annotation pipeline for faster genome publication. *Bioinformatics* **34**: 1037–1039.

Terada, A., Sugawara, S., Yamamoto, T., Zhou, S., Koba, K., and Hosomi, M. (2013) Physiological characteristics of predominant ammonia-oxidizing bacteria enriched from bioreactors with different influent supply regimes. *Biochem Eng J* **79**: 153–161.

Vasimuddin M, Misra S, Li H, Aluru S, editors. Efficient Architecture-Aware Acceleration of BWA-MEM for Multicore Systems. 2019 IEEE International Parallel and Distributed Processing Symposium (IPDPS); 2019 20-24 May 2019.

Wong, L.L., Lu, Y., Ho, J.C.S., Mugunthan, S., Law, Y., Conway, P., *et al.* (2023) Surface-layer protein is a public-good matrix exopolymer for microbial community organisation in environmental anammox biofilms. *ISME J* **17**: 803–812.

Wu, Y.W., Simmons, B.A., and Singer, S.W. (2016) MaxBin 2.0: an automated binning algorithm to recover genomes from multiple metagenomic datasets. *Bioinformatics* **32**: 605–607.

Xiao, R., Ni, B.J., Liu, S., and Lu, H. (2021) Impacts of organics on the microbial ecology of wastewater anammox processes: Recent advances and meta-analysis. *Water Res* **191**: 116817.

Yasuda, S., Suenaga, T., and Terada, A. (2020) Complete genome sequence of methylosinus sp. strain C49, a methane-oxidizing bacterium harboring phaABC genes for polyhydroxyalkanoate synthesis. *Microbiol Resour Announc* **9**: e00113-20.

Yin, Q., Sun, Y., Li, B., Feng, Z., and Wu, G. (2022) The r/K selection theory and its application in biological wastewater treatment processes. *Sci Total Environ* **824**: 153836.

Yin, Y., Yan, J., Chen, G., Murdoch, F.K., Pfisterer, N., and Loffler, F.E. (2019) Nitrous oxide is a potent inhibitor of bacterial reductive dechlorination. *Environ Sci Technol* **53**: 692–701.

Yoon, H., Song, M.J., and Yoon, S. (2017) Design and feasibility analysis of a self-sustaining biofiltration system for removal of low concentration N(2)O emitted from wastewater treatment plants. *Environ Sci Technol* **51**: 10736–10745.

Yoon, S., Nissen, S., Park, D., Sanford, R.A., and Loffler, F.E. (2016) Nitrous oxide reduction kinetics distinguish bacteria harboring clade I NosZ from those harboring clade II NosZ. *Appl Environ Microbiol* **82**: 3793–3800.

Microbes Environ. 38(2), 2023
https://www.jstage.jst.go.jp/browse/jsme2 doi:10.1264/jsme2.ME23015

MOST VALUABLE PAPER OF THE YEAR 2023

Light-driven Proton Pumps as a Potential Regulator for Carbon Fixation in Marine Diatoms

Susumu Yoshizawa[1,2,3]*, Tomonori Azuma[4], Keiichi Kojima[5,6], Keisuke Inomura[7], Masumi Hasegawa[1,2], Yosuke Nishimura[1], Masuzu Kikuchi[5], Gabrielle Armin[7], Yuya Tsukamoto[1], Hideaki Miyashita[4], Kentaro Ifuku[8], Takashi Yamano[9], Adrian Marchetti[10], Hideya Fukuzawa[9], Yuki Sudo[5,6], and Ryoma Kamikawa[8]

[1]*Atmosphere and Ocean Research Institute, The University of Tokyo, Chiba 277–8564, Japan;* [2]*Graduate School of Frontier Sciences, The University of Tokyo, Chiba 277–8563, Japan;* [3]*Collaborative Research Institute for Innovative Microbiology, The University of Tokyo, Tokyo 113–8657, Japan;* [4]*Graduate School of Human and Environmental Studies, Kyoto University, Kyoto 606–8501, Japan;* [5]*School of Pharmaceutical Sciences, Okayama University, Okayama 700–8530, Japan;* [6]*Faculty of Medicine, Dentistry and Pharmaceutical Sciences, Okayama University, Okayama 700–8530, Japan;* [7]*Graduate School of Oceanography, University of Rhode Island, Narragansett, RI, USA;* [8]*Graduate School of Agriculture, Kyoto University, Kyoto 606–8502, Japan;* [9]*Graduate School of Biostudies, Kyoto University, Kyoto, 606–8502, Japan; and* [10]*Department of Earth, Marine and Environmental Sciences, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA*

(Received February 10, 2023—Accepted May 12, 2023—Published online June 20, 2023)

Diatoms are a major phytoplankton group responsible for approximately 20% of carbon fixation on Earth. They perform photosynthesis using light-harvesting chlorophylls located in plastids, an organelle obtained through eukaryote-eukaryote endosymbiosis. Microbial rhodopsin, a photoreceptor distinct from chlorophyll-based photosystems, was recently identified in some diatoms. However, the physiological function of diatom rhodopsin remains unclear. Heterologous expression techniques were herein used to investigate the protein function and subcellular localization of diatom rhodopsin. We demonstrated that diatom rhodopsin acts as a light-driven proton pump and localizes primarily to the outermost membrane of four membrane-bound complex plastids. Using model simulations, we also examined the effects of pH changes inside the plastid due to rhodopsin-mediated proton transport on photosynthesis. The results obtained suggested the involvement of rhodopsin-mediated local pH changes in a photosynthetic CO_2-concentrating mechanism in rhodopsin-possessing diatoms.

Key words: microbial rhodopsins, diatom, marine microbiology, CO_2-concentrating mechanism

Diatoms are unicellular, photosynthetic algae found throughout aquatic environments and are responsible for up to 20% of annual net global carbon fixation (Nelson *et al.*, 1995; Field *et al.*, 1998). Since their contribution to primary production in the ocean is significant, their light utilization mechanisms are essential to correctly understand marine ecosystems. Diatoms contain chlorophylls *a* and *c* and carotenoids, such as fucoxanthin, as photosynthetic pigments in the plastids acquired by eukaryote-eukaryote endosymbiosis (Keeling, 2004). Some diatoms have recently been shown to contain microbial rhodopsin (henceforth rhodopsins), a light-harvesting antenna distinct from the chlorophyll-containing antenna for photosynthesis (Marchetti *et al.*, 2015). Although rhodopsin-mediated light-harvesting may support the survival of these diatoms in marine environments, the physiological role of rhodopsin in diatom cells remains unclear.

Microbial rhodopsins are a large family of seven transmembrane photoreceptor proteins (Spudich *et al.*, 2000). Rhodopsin has an all-*trans* retinal as the light-absorbing chromophore, and its protein function is triggered by the light-induced isomerization of the retinal. The first microbial rhodopsin, the light-driven proton pump bacteriorhodopsin (BR), was discovered in halophilic archaea (Oesterhelt and Stoeckenius, 1971). Although rhodopsin was initially considered to only occur in halophilic archaea inhabiting hypersaline environments, subsequent studies showed that the rhodopsin gene is widely distributed in all three domains of life (Beja *et al.*, 2000; Sineshchekov *et al.*, 2002). Rhodopsins are classified based on their functions into light-driven ion pumps, light-activated signal transducers, and light-gated ion channels. The two former functional types of rhodopsin have so far been identified in prokaryotes (Ernst *et al.*, 2014). Rhodopsins in prokaryotes, regardless of function, localize to the cell membrane in which they operate. For example, proton-pumping rhodopsins export protons from the cytosol across the cell membrane to convert light energy into a proton motive force (PMF) (Yoshizawa *et al.*, 2012). The PMF induced by rhodopsin ion transport is utilized by various physiological functions, such as ATP synthesis, substrate uptake, and flagellar movement.

Rhodopsins functioning as light-driven ion pumps and light-gated ion channels have been reported in eukaryotic microorganisms (Ernst *et al.*, 2014; Kikuchi *et al.*, 2021). A light-gated ion channel called channelrhodopsin, which localizes to the plasma membrane over the eyespot within the chloroplast of green algae, has been extensively exam-

* Corresponding author. E-mail: yoshizawa@aori.u-tokyo.ac.jp; Tel: +81–4–7136–6419; Fax: +81–4–7136–6419.

Citation: Yoshizawa, S., Azuma, T., Kojima, K., Inomura, K., Hasegawa, M., Nishimura, Y., et al. (2023) Light-driven Proton Pumps as a Potential Regulator for Carbon Fixation in Marine Diatoms. *Microbes Environ* **38**: ME23015.
https://doi.org/10.1264/jsme2.ME23015

ined for its role in phototaxis (Nagel *et al.*, 2002). The other type of rhodopsin in eukaryotes, light-driven ion-pumping rhodopsins has been detected in a number of organisms belonging to both photoautotrophic and heterotrophic protists (Slamovits *et al.*, 2011; Marchetti *et al.*, 2015). Since the intracellular membrane structure of eukaryotic cells is more complex than that of prokaryotes, containing various organelles, even light-driven ion-pumping rhodopsins may have distinct physiological roles depending on their subcellular localization (Slamovits *et al.*, 2011). However, due to the difficulties associated with identifying the exact localization of rhodopsins in eukaryotic cells, their subcellular localization remains unknown.

In the present study, to clarify the physiological function of rhodopsin in a marine pennate diatom, we investigated the phylogeny, protein function, spectroscopic characteristics, and subcellular localization of rhodopsin from a member of the genus *Pseudo-nitzschia*. Heterologous expression techniques were used to analyze protein functions and spectroscopic features. The expression of rhodopsin fused with a green fluorescent protein, eGFP revealed its subcellular localization in a model diatom (*Phaeodactylum tricornutum*). Furthermore, a model-based analysis was performed to evaluate the impact of the potential roles of rhodopsin in cellular biology.

Materials and Methods

Rhodopsin sequences and phylogenetic analysis

The rhodopsin sequence of the diatom *Pseudo-nitzschia granii* was previously reported (Marchetti *et al.*, 2015). All other rhodopsin sequences used in the phylogenetic analysis were collected from the National Center for Biotechnology Information. Detailed information on the strains used in this analysis is given in Extended Data Fig. 1. Sequences were aligned using MAFFT version 7.453 with the options '—genafpair' and '—maxiterate 1000' (Katoh and Standley, 2013). The phylogenetic tree was inferred using RAxML (v.8.2.12) with the 'PROTGAMMALGF' model using 1,000 rapid bootstrap searches (Stamatakis, 2014). Model selection was performed with the ProteinModelSelection.pl script in the RAxML package.

The search for eukaryotic rhodopsins belonging to the Xanthorhodopsin (XR)-like rhodopsin (XLR) clade was performed among the protein sequences of the Marine Microbial Eukaryote Transcriptome Sequencing Project (MMETSP) (Keeling *et al.*, 2014). The phylogenetic placement of rhodopsin proteins from MMETSP using pplacer (v1.1.alpha19) (Matsen *et al.*, 2010) was conducted on a prebuilt large-scale phylogenetic tree of rhodopsins and extracted placements on the XLR clade using gappa (v0.6.0) (Czech *et al.*, 2020).

Gene preparation, protein expression, and ion transport measurements of Escherichia coli *cells*

In the present study, a functional analysis of the rhodopsin possessed by the diatom *P. granii* (named PngR, accession no. AJA37445.1) was performed using a heterologous expression system. The full-length cDNA for PngR, the codons of which were optimized for *E. coli*, were chemically synthesized by Eurofins Genomics and inserted into the *Nde*I-*Xho*I site of the pET21a(+) vector as previously described (Hasegawa *et al.*, 2020). A hexahistidine-tag was fused at the C terminus of PngR, which was utilized for the purification of the expressed protein. The heterologous protein expression method is the same as that previously reported (Inoue *et al.*, 2018). *E. coli* BL21(DE3) cells harboring the cognate plasmid were grown at 37°C in LB medium supplemented with ampicillin (final concentration of 50 μg mL^{-1}). Protein expression was induced at an optical density at 600 nm of 0.7–1.2 with 1 mM isopropyl β-d-1-thiogalactopyranoside (IPTG) and 10 μM all-trans retinal, after which cells were incubated at 37°C for 3 h. The proton transport activity of PngR was measured as light-induced pH changes in suspensions of *E. coli* cells as previously described (Inoue *et al.*, 2018). Briefly, cells expressing PngR were washed more than three times in 150 mM NaCl and then resuspended in the same solution for measurements. Each cell suspension was placed in the dark for several min and then illuminated using a 300 W Xenon lamp (*ca.* 30 mW cm^{-2}, MAX-303; Asahi Spectra) through a >460 nm long-pass filter (Y48; HOYA) for 3 min. Measurements were repeated under the same conditions after the addition of the protonophore carbonyl cyanide m-chlorophenylhydrazone (CCCP) (final concentration=10 μM). Light-induced pH changes were monitored using a Horiba F-72 pH meter. All measurements were conducted at 25°C using a thermostat (Eyela NCB-1200; Tokyo Rikakikai).

Purification of PngR from E. coli *cells and spectroscopic measurements of the purified protein*

E. coli cells expressing PngR were disrupted by sonication for 30 min in ice-cold buffer containing 50 mM Tris–HCl (pH 7.0) and 300 mM NaCl. The crude membrane fraction was collected by ultracentrifugation (130,000×*g* at 4°C for 60 min) and solubilized with 1.0% (w/v) n-dodecyl-β-D-maltoside (DDM; DOJINDO Laboratories). The solubilized fraction was purified by Ni^{2+} affinity column chromatography with a linear gradient of imidazole as previously described (Kojima *et al.*, 2020b). The purified protein was concentrated by centrifugation using an Amicon Ultra filter (30,000 M_w cut-off; Millipore). The sample medium was then replaced with Buffer A (50 mM Tris–HCl, pH 7.0, 1 M NaCl, and 0.05% [w/v] DDM) by ultrafiltration 3 times.

The absorption spectra of purified proteins were recorded using a UV-2450 spectrophotometer (Shimadzu) at room temperature in Buffer A. The retinal composition in PngR was analyzed by high-performance liquid chromatography (HPLC) as previously described (Kojima *et al.*, 2020a). Regarding dark adaptation, samples were kept under dark conditions at 4°C for more than 72 h, whereas those for light adaptation were illuminated for 3 min at 520±10 nm, with light power being adjusted to approximately 10 mW cm^{-2}. The molar compositions of the retinal isomers were calculated from the areas of the peaks in HPLC patterns monitored at 360 nm using the extinction coefficients of retinal oxime isomers as previously described (Kojima *et al.*, 2020a). In pH titration experiments, samples were suspended in Buffer A. The pH values of the samples were adjusted to the desired acidic values by the addition of HCl, after which absorption spectra were measured at each pH value. All measurements were conducted at room temperature (approximately 25°C) under room light. After these measurements, the reversibility of spectral changes was examined to confirm that the sample was not denatured during measurements. Absorption changes at specific wavelengths were plotted against pH values and plots were fit to the Henderson–Hasselbalch equation assuming a single pK_a value as previously described (Inoue *et al.*, 2018).

The transient time-resolved absorption spectra of the purified proteins from 380 to 700 nm at 5-nm intervals were obtained using a homemade computer-controlled flash photolysis system equipped with an Nd: YAG laser as an actinic light source. Using an optical parametric oscillator, the wavelength of the actinic pulse was tuned at 510 nm for PngR. Pulse intensity was adjusted to 2 mJ per pulse. All data were averaged to improve the signal-to-noise ratio (n=30). All measurements were conducted at 25°C. In these experiments, samples were suspended in Buffer A. After measurements, the reproducibility of data was checked to confirm that the sample was not denatured during measurements. To investigate proton uptake and release during the photocycle, we used the pH indicator pyra-

nine (final concentration=100 μM; Tokyo Chemical Industry), which has been extensively used to monitor light-induced pH changes in various rhodopsins. pH changes in the bulk environment were measured as the absorption changes of pyranine at 450 nm. The absorption changes of pyranine were obtained by subtracting the absorption changes of samples without pyranine from those of samples with pyranine. Experiments using pyranine were performed in an unbuffered solution containing 1 M NaCl and 0.05% (w/v) DDM (pH 7.0) to enhance signals. The results of 1,000 traces were averaged to improve the signal-to-noise ratio.

Subcellular localization of PngR in the model diatom

The PngR:eGFP recombinant gene, coding the full length of PngR C-terminally tagged with eGFP, was cloned into the expression vector for the model diatom *P. tricornutum*, pPha-NR (Stork *et al.*, 2012), by CloneEZ (GenScript) according to the manufacturer's instructions. The plasmid was electroporated into cells of *P. tricornutum* UTEX642 with the NEPA21 Super Electroporator (NEPAGENE), and transformed cells were selected with a Zeocin-based antibiotic treatment as previously described (Miyahara *et al.*, 2013; Dorrell *et al.*, 2019). Selected clones were observed under an Olympus BX51 fluorescent microscope (Olympus) equipped with an Olympus DP72 CCD color camera (Olympus). The nucleus stained with DAPI and chlorophyll autofluorescence from the plastid were observed with a 420-nm filter at 330 to 385 nm excitation. GFP fluorescence was detected with a 510- to 550-nm filter at 470 to 495 nm excitation.

Quantitative model of C concentrations in a diatom: Cell Flux Model of C Concentration (CFM-CC)

Membrane transport model. We combined the membrane transport of CO_2 and C fixation. Parameter definitions, units, and values are provided in Supporting Information Table S1 and S2, respectively. The key model equation is the balance of the concentration of CO_2 in the cytosol, $[CO_2]_p$:

$$\frac{d[CO_2]_p}{dt} = D\left([CO_2]_m - [CO_2]_p\right) - V_{Cfix} \quad \text{[eq. 1]}$$

where t is time, D is the diffusion coefficient, and $[CO_2]_m$ is the concentration of CO_2 in the inner side of the outermost membrane of the plastid (hereafter "the middle space"). The first term represents the diffusion of CO_2 from the middle space to the cytosol, while the second term V_{Cfix} represents the C fixation rate following Michaelis–Menten kinetics (Berg *et al.*, 2010; Hopkinson, 2014):

$$V_{Cfix} = V_{max}\frac{[CO_2]_p}{[CO_2]_p + K} \quad \text{[eq. 2]}$$

where V_{max} is the maximum CO_2 fixation rate and K is the half saturation constant. $[CO_2]_m$ is obtained based on the carbonate chemistry in the middle space (see below). Under the steady state, [eq. 1] with [eq. 2] becomes the following quadratic relationship for $[CO_2]_p$:

$$[CO_2]_p^2 + \left(\frac{V_{max}}{D} + K - [CO_2]_m\right)[CO_2]_p - K[CO_2]_m = 0 \quad \text{[eq. 3]}$$

Solving this equation for $[CO_2]_p^2$ leads to:

$$[CO_2]_p = \frac{-\left(\frac{V_{max}}{D} + K - [CO_2]_m\right) + \sqrt{\left(\frac{V_{max}}{D} + K - [CO_2]_m\right)^2 + 4K}}{2} \quad \text{[eq. 4]}$$

Note that the other solution for the negative route is unrealistic because it may lead to the overall negative value of $[CO_2]_p$. Once we obtain $[CO_2]_p$, we may then calculate the rate of C fixation V_{Cfix} with [eq. 2].

Furthermore, from [eq. 4], we obtain two extreme solutions. In the case of $V_{max} \ll D$ (*i.e.*, when the CO_2 uptake capacity is small relative to the speed of CO_2 diffusion), [eq. 4] leads to

$$[CO_2]_p \sim [CO_2]_m \quad \text{[eq. 5]}$$

With this relationship and [eq. 2], V_{Cfix} is computed as follows:

$$V_{Cfix} \sim V_{max}\frac{[CO_2]_m}{[CO_2]_m + K} \quad \text{[eq. 6]}$$

In contrast, when $V_{max} \gg D$ (*i.e.*, when the CO_2 uptake capacity is high relative to the CO_2 diffusion across the membrane), [eq. 4] becomes

$$[CO_2]_p \sim 0 \quad \text{[eq. 7]}$$

Under the steady state, [eq. 1] becomes

$$V_{Cfix} = D\left([CO_2]_m - [CO_2]_p\right) \quad \text{[eq. 8]}$$

and plugging [eq. 7] into [eq. 8] leads to

$$V_{Cfix} \sim D[CO_2]_m \quad \text{[eq. 9]}$$

and V_{Cfix} is calculated. We note that $[CO_2]_p > [CO_2]_m$ may occur when there are membrane-bound transporters for HCO_3^- located on each membrane between the middle space and plastid (Hopkinson, 2014). However, such a set of transporters has not yet been discovered (Matsuda *et al.*, 2017). Therefore, our model conforms with the current state of knowledge. Even if $[CO_2]_p > [CO_2]_m$, moderately decreased pH_m and, thus, increased $[CO_2]_m$ may be useful since they may reduce the gradient of CO_2 across membranes (*i.e.*, $[CO_2]_p$ vs $[CO_2]_m$), thereby mitigating the diffusive loss of CO_2 from the plastid.

Carbonate chemistry in the middle space. The above equations may be solved once we obtain $[CO_2]_m$. The model uses a given DIC (dissolved inorganic C) concentration in the middle space $[DIC]_m$ to calculate $[CO_2]_m$ following the established equations for carbon chemistry (Emerson and Hedges, 2008).

$$[CO_2]_m = \frac{[DIC]_m}{1 + \frac{K_1}{[H^+]_m} + \frac{K_1K_2}{[H^+]_m^2}} \quad \text{[eq. 10]}$$

where $[H^+]_m$ is the concentration of H^+ (10^{-pH} mol L^{-1}) in the middle space and K_1 and K_2 are temperature- and salinity-dependent parameters (Lueker *et al.*, 2000; Emerson and Hedges, 2008):

$$K_1 = 10^{-pK_1} \quad \text{[eq. 11]}$$

$$K_2 = 10^{-pK_2} \quad \text{[eq. 12]}$$

where

$$pK_1 = \frac{3633.86}{T} - 61.2172 + 9.6777\ln(T) - 0.011555S + 0.0001152S^2 \quad \text{[eq. 13]}$$

$$pK_2 = \frac{471.78}{T} + 25.9290 - 3.16967\ln(T) - 0.01781S + 0.0001122S^2 \quad \text{[eq. 14]}$$

Code availability

The code for CFM-CC is freely available from GitHub/Zenodo at https://zenodo.org/record/5182712 (DOI: 10.5281/zenodo.5182712).

Results and Discussion

Rhodopsin sequences and phylogenetic analysis

We performed a phylogenetic analysis using the rhodopsin (named PngR, accession no. AJA37445.1) of the diatom *P. granii* and microbial rhodopsin sequences reported to date (Marchetti *et al.*, 2015). This phylogenetic tree revealed that PngR is not included in the proteorhodopsin (PR) clade

commonly found in oceanic organisms, but belongs to the Xanthorhodopsin (XR)-like rhodopsin (XLR) clade, which is presumed to have an outward proton transporting function (Fig. 1 and Extended Data Fig. 1). A comparison of the motif sequences necessary for ion transport showed that the amino acids in the putative proton donor and acceptor sites of XR and PR were conserved in PngR, suggesting that PngR functions as an outward proton pump (Extended Data Fig. 2). Furthermore, the homology search for rhodopsin sequences in the XLR clade from Marine Microbial Eukaryote Transcriptome Sequencing Project (MMETSP) revealed that not only diatoms (Ochrophyta, Stramenopiles), but also dinoflagellates (Dinophyceae, Alveolata) and haptophytes have rhodopsin genes in the same XLR clade (Supporting Information Table S3). These results indicate that rhodopsins of the XLR clade are widely distributed among the major phytoplankton groups, which are important primary producers in the ocean.

Function and spectroscopic features of diatom rhodopsin

To characterize the function of PngR, we heterologously expressed the synthesized rhodopsin gene in *E. coli* cells. A light-induced decrease in pH was observed in the suspension of *E. coli* cells expressing PngR, and this reduction was almost completely abolished in the presence of the protonophore carbonyl cyanide m-chlorophenylhydrazone (CCCP) (Fig. 2A). The pH changes observed clearly showed that PngR exported protons from the cytoplasmic side across the cell membrane.

We then examined the spectroscopic characteristics of PngR using the recombinant protein purified from *E. coli*. The absorption maximum of PngR was located at 511 nm (Fig. 2B), which was markedly shorter than those of XR (565 nm) and GR (*Gloeobacter* rhodopsin 541 nm) in the XLR clade (Balashov *et al.*, 2005). It is important to note that while *P. granii* is a marine species, XR and GR are both distributed in terrestrial organisms. Therefore, the present results are consistent with the shorter wavelength of the absorption maximum of rhodopsin in marine environments than in the terrestrial environment (Man *et al.*, 2003), indicating that PngR is well adapted to light conditions in the ocean, particularly the open ocean.

We then examined the retinal configuration in PngR by HPLC. In light- and dark-adapted samples, the isomeric state of retinal was predominantly all-*trans* (Extended Data Fig. 3), which was similar to the isomeric state of retinal in prokaryotic GR in the XLR clade, but different from that in BR (Miranda *et al.*, 2009). Since the pK_a value of the proton acceptor residue (Asp85 in BR) is an indicator of the efficiency of proton transport by rhodopsin, we estimated the

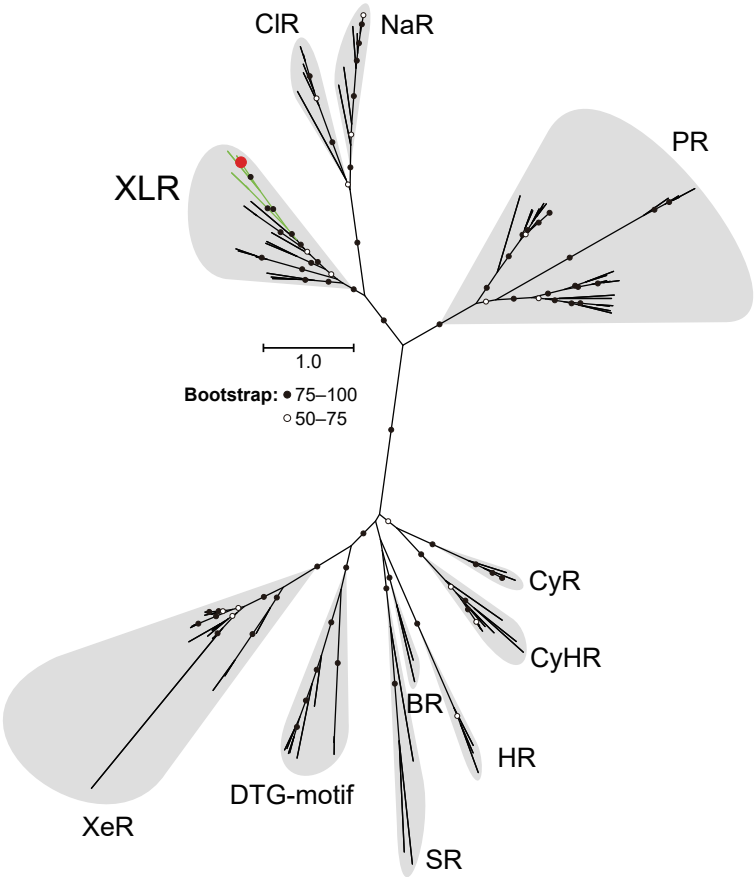


Fig. 1. Phylogenetic position of diatom rhodopsin. A maximum likelihood tree of the amino acid sequences of microbial rhodopsins. Diatom rhodopsin (PngR) is indicated by a red circle and bootstrap probabilities (≥50%) by black and white circles. Green branches indicate eukaryotic rhodopsins used in this analysis, while black branches indicate others. Rhodopsin clades are as follows: Xanthorhodopsin-like rhodopsin (XLR), Cl^--pumping rhodopsin (ClR), Na^+-pumping rhodopsin (NaR), proteorhodopsin (PR), xenorhodopsin (XeR), DTG-motif rhodopsin, sensory rhodopsin-I and sensory rhodopsin-II (SR), bacteriorhodopsin (BR), halorhodopsin (HR), cyanobacterial halorhodopsin (CyHR), and cyanorhodopsin (CyR).

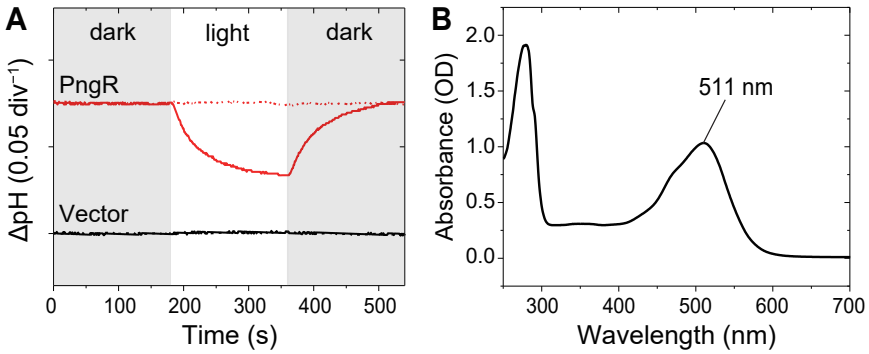


Fig. 2. Light-induced pH changes and absorption spectrum of PngR. (A) Outward proton pump activity of PngR in *E. coli* cells. Light-induced pH changes in solutions containing *E. coli* cells with the expression plasmid for PngR (upper panel) and the empty vector pET21a (lower panel) in the presence (red dashed line) or absence (red solid line) of CCCP. The white-filled region indicates the period of illumination. (B) Absorption spectrum of purified PngR in Buffer A (50 mM Tris–HCl, pH 7.0, 1 M NaCl, and 0.05% [w/v] DDM).

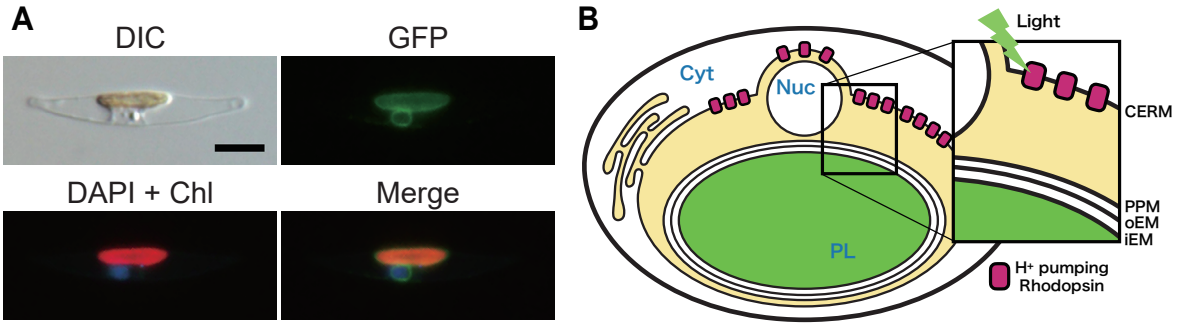


Fig. 3. Subcellular localization of rhodopsins in diatom cells. (A) A transformed diatom cell was observed with differential interface contrast (DIC) (Upper left). Green fluorescence from recombinant PngR (GFP) (Upper right). Nuclear DNA stained with DAPI and chlorophyll autofluorescence (DAPI + Chl) and a merged image (Merge) are shown in the bottom left and bottom right, respectively. The scale bar indicates 5 μm. (B) A mode for the subcellular localization of PngR. The proton transport of PngR acidifies or alkalizes the region (the middle space) surrounded by the membrane of CERM and PPM. Abbreviations are as follows: cytosol (Cyt), nucleus (Nuc), plastid (PL), chloroplast endoplasmic reticulum membrane (CERM), periplastidial membrane (PPM), outer plastid envelope membrane (oEM), and internal plastid envelope membrane (iEM).

pK_a values of the putative proton acceptor in PngR (Asp91) by a pH titration experiment (Extended Data Fig. 4). This experiment estimated that the p*K*a of this residue acceptor was approximately 5.0, indicating that the proton acceptor of PngR works well in marine and intracellular environments. Furthermore, the photochemical reactions that proceed behind the ion-transportation mechanism of PngR were examined by a flash photolysis analysis (Extended Data Fig. 5). All photocycles required for ion transportation in PngR were completed in approximately 300 ms, suggesting that the cycle was sufficiently fast to pump protons in a physiologically significant time scale. The results and a discussion of the flash photolysis analysis are described in the supplementary information.

Subcellular localization of PngR in a model diatom

The PngR sequence bears neither an apparent N-terminal extension nor a detectable N-terminal signal peptide, and, thus, *in silico* analyses are unable to predict the subcellular localization of PngR. To identify the subcellular localization of PngR, a C-terminal eGFP-fusion PngR was expressed in the model diatom *P. tricornutum*, which may be transformed by electroporation and is often used in a heterologous expression analysis (Nakajima *et al.*, 2013; Dorrell *et al.*, 2019). The transformed *P. tricornutum* cell was examined under differential interface contrast and epifluorescent microscopes (Fig. 3A). We observed the fluorescence of GFP, DAPI, and chlorophylls to establish the localization of recombinant PngR:eGFP, the nucleus, and chloroplast, respectively, in multiple cells (Extended Data Fig. 6 and 7). The fluorescence signal of the PngR:eGFP transformant appeared to localize at the periphery of chlorophyll fluorescence and DAPI signals, corresponding to the outermost plastid membrane, called the chloroplast endoplasmic reticulum membrane (CERM), which is physically connected to the nuclear membrane (Fig. 3A). A few cells also exhibited GFP signals within vacuolar membranes in addition to CERM (Extended Data Fig. 7). The insertion of the complete sequence of the PngR:eGFP gene in transformant DNA was confirmed by PCR followed by Sanger sequencing.

Based on the results of the heterologous expression experiment and microscopic observations, we concluded that PngRs primarily localized to the outermost membrane of the plastid. However, fluorescence signals were also observed to a lesser extent in the vacuolar membrane, suggesting the involvement of other factors, such as cell growth conditions, in their localization. These results imply that light-driven proton transport by PngR acidifies or alkalizes the inner region of CERM (Fig. 3B). Therefore, the physiological role of pH changes in this region in diatoms warrants further study. The electrochemical gradient formed by rhodopsin may be a driving force for various secondary

transport processes. Alternatively, based on the primary purpose of plastids, local pH changes may be related to photosynthesis. The pH in this region is considered to be important for the transport of inorganic carbon (Ci) to ribulose-1,5-bisphosphate carboxylase/oxygenase (RuBisCO) (Gee and Niyogi, 2017). This is because in the carbonate system, pH affects the proportion of carbonate species (CO_2, HCO_3^-, and CO_3^{2-}) in water.

Under weakly alkaline conditions in the ocean, the majority of dissolved inorganic carbon (DIC) is generally present in the form of HCO_3^-, with only approximately 1% being present in the form of CO_2. However, RuBisCO localized in the stroma only reacts with Ci in the form of CO_2, not HCO_3^-. The RuBisCO enzyme in diatoms exhibits low affinity even for CO_2 (*K*m of 25~68 µM, while CO_2aq in the ocean is approximately 10 µM at 25°C) and, thus, requires concentrated CO_2 for efficient fixation at the site of RuBisCO. In other words, the ocean is always a CO_2-limited environment for most phytoplankton (Riebesell *et al.*, 1993). Consequently, due to the membrane impermeability of HCO_3^-, phytoplankton have developed a number of CO_2-concentrating mechanisms (CCM) to efficiently transport Ci to the site of RuBisCO by placing HCO_3^- transporters in appropriate membranes and carbonic anhydrase (CA) in these compartments, the latter of which catalyzes the rapid interconversion between HCO_3^- and CO_2. However, since difficulties are associated with directly examining pH changes and the forms of Ci of the small compartment in eukaryotic microbial organelles, a model simulation is a powerful alternative approach (Hopkinson *et al.*, 2011). In the present study, we used a model simulation to investigate whether rhodopsin-mediated pH changes in this region were involved in CCM.

A quantitative model of carbon concentrations in diatoms: CFM-CC

Our subcellular localization analysis suggested that proton transport by rhodopsin acidified or alkalized the inner side of the outermost membrane of the plastid (the middle space). To quantitatively examine the effects of pH in the middle space on C fixation, we developed a simple quantitative model of carbonate chemistry combined with membrane transport and C fixation (CFM-CC: Cell Flux Model of C Concentration) (Fig. 4 upper panel). A comprehensive model of the concentration of CO_2 within diatoms was developed (Hopkinson *et al.*, 2011; Hopkinson, 2014). CFM-CC uses a conceptually similar structure to this model, focusing on more specific membrane layers, designed to test the effects of pH changes in the middle space.

Our model results showed that the concentrations of CO_2 in the middle space ($[CO_2]_m$) were strongly dependent on pH (pH_m), suggesting that proton pumping by rhodopsin affected C fixation (Fig. 4 bottom panel). The calculation of C chemistry in the middle space revealed that a decrease in

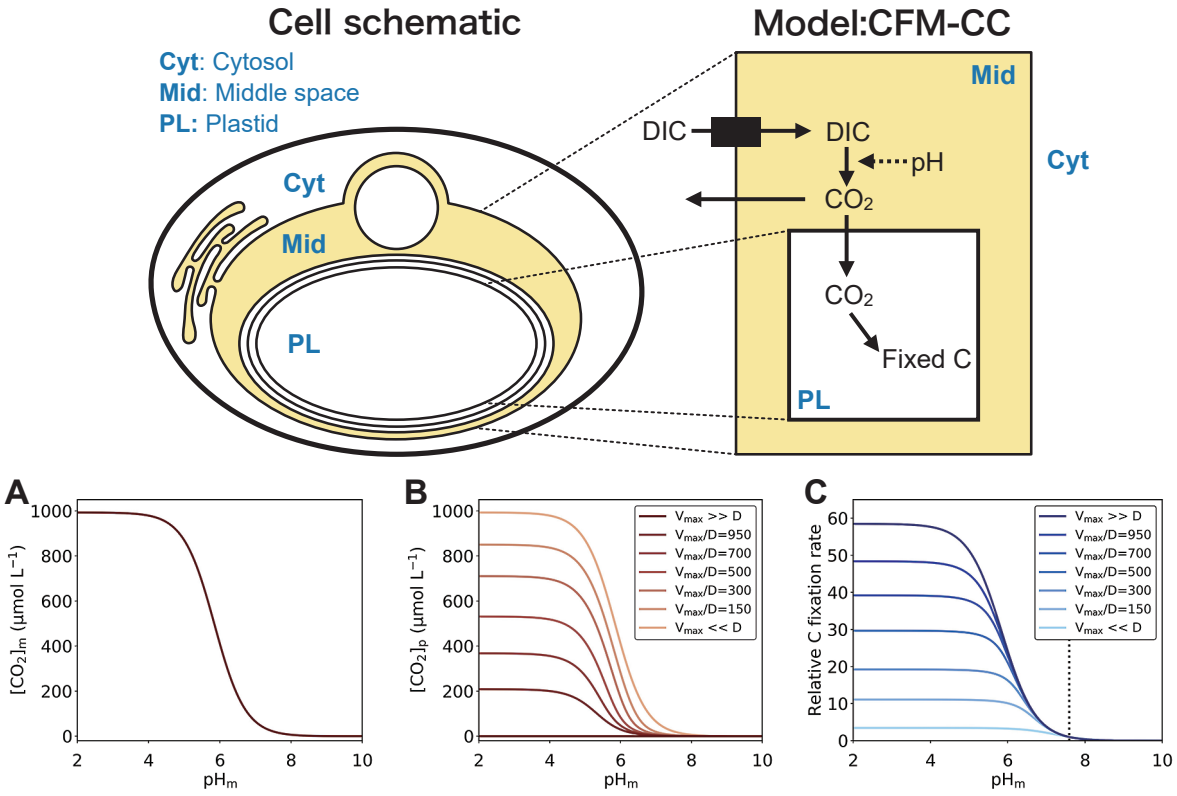


Fig. 4. A quantitative model of the concentration of carbon in diatoms. (Upper panel) Schematic of a Cell Flux Model of C Concentration (CFM-CC). The left panel represents the actual cell, while the right panel represents the model. Solid arrows show the net flux of C and the dashed arrow indicates the effects of pH. (Bottom panel) The effects of pH in the middle space on CO_2 concentrations and the photosynthesis rate. (A) CO_2 concentrations in the middle space $[CO_2]_m$. (B) CO_2 concentrations in the plastid $[CO_2]_p$. (C) The C fixation rate relative to that with pH in the middle space of 7.59, the only mean value we found for intracellular pH in a diatom (Burns and Beardall, 1987). (B) and (C) are plotted for various V_{max}/D. The solution for $[CO_2]_m$ in (A) is independent of V_{max}/D.

pH_m favored higher $[CO_2]_m$ at a given DIC concentration (Fig. 4A) (we used 993 μmol L^{-1} [Burns and Beardall, 1987]). We noted that the potential leaking of CO_2 into the cytosol may change DIC in the middle space, but used a constant DIC value because this effect has not been experimentally demonstrated and is difficult to quantify due to unknown factors (*e.g.*, the balance of DIC uptake and CO_2 leaking). At the reference point (we used pH_m=7.59 [Burns and Beardall, 1987]), $[CO_2]_m$ was 17 μmol L^{-1}, but increased to 64, 410, and 870 μmol L^{-1} for pH_m values of 7, 6, and 5, respectively (Fig. 4A).

Due to increased $[CO_2]_m$, the concentration of CO_2 in the plastid ($[CO_2]p$) also increased with lower pH_m; however, the level of this increase was dependent on V_{max}/D (the ratio of the maximum C fixation rate to the diffusion constant) (Fig. 4B). When V_{max}/D was small, the diffusion of CO_2 from the middle space to the plastid dominated the change, resulting in $[CO_2]_p$ similar to $[CO_2]_m$. In contrast, when V_{max}/D was large, CO_2 uptake dominated, and the effect of $[CO_2]_m$ on $[CO_2]_p$ was small.

The rate of C fixation increased with lower pH_m because increased $[CO_2]_m$ accelerated the transport of CO_2 into the plastid (Fig. 4C). However, the magnitude of this increase depended on V_{max}/D. The model showed that the effects of pH_m on C fixation were greater when V_{max}/D was large because the rate of C fixation was changed more by $[CO_2]_m$ [eq. 6], which is directly affected by pH_m. However, when V_{max}/D was small, the C fixation rate was changed more by C uptake kinetics [eq. 4], which were saturated at a relatively low $[CO_2]_p$ (K value of 44 μmol L^{-1} [Jensen *et al.*, 2020]). Based on the possible range of V_{max}/D, C fixation showed 2.1- to 3.8-, 3.2- to 24.2-, and 3.4- to 51.2-fold increases when pH_m decreased from 7.59 to 7, 6, 5 respectively. These results suggested that pH_m markedly affected C uptake at any V_{max}/D as well as also the benefit for cells to have high V_{max} relative to the diffusivity of CO_2 across the membrane. This was most likely the case because the D value was shown to be reduced when there were multiple membranes (Eichner *et al.*, 2019; Inomura *et al.*, 2019). Therefore, this simple yet elegant system with rhodopsin to manipulate pH_m provides a powerful mechanism in C concentrations and, thus, adjusts the C fixation rate to given physiological conditions in some rhodopsin-containing diatoms, enabling them to be more successful primary producers in the ocean.

In our model simulation, we examined the effects of rhodopsin-mediated pH changes in the middle space on CCM efficiency. The results obtained suggested that C fixation was enhanced when the pH of the middle space was acidified by a light-driven proton pump. CCM based on CO_2 diffusion (termed the pump-leak type) has been proposed as a possible mechanism by placing CAs in appropriate locations. For example, *Nannochloropsis oceanica* (*Ochrophyta*), possessing the same four membrane-bound complex plastids as those found in diatoms, is considered to generate CO_2 by placing CA in the middle space (Gee and Niyogi, 2017). Furthermore, the centric diatom *Chaetoceros gracilis* is considered to generate CO_2 by placing CAs outside the cell and allowing CO_2 to flow into the cell (Tsuji *et al.*, 2021). In contrast to the CA-based model, the acidification-based model was formerly proposed to facilitate the CO_2 fixation of RuBisCO in the thylakoid lumen of plastids; HCO_3^- is converted into CO_2 through acidification by photosynthetic proton pumping into the thylakoid lumen (Raven, 1997). Our rhodopsin-mediated Ci transform model, CFM-CC proposes that pH changes in the middle space by proton-pumping rhodopsin also plays the role of a CO_2 regulator. This proposed mechanism may be useful in most parts of the ocean where CO_2 chronically limits photosynthesis, but may be even more valuable in specific environments. For example, since CA, which plays a central role in CCM, requires cobalt or zinc ions as the reaction center, and photosynthetic proton-pumping systems need iron, rhodopsin-derived acidic pools may be useful for Ci uptake in oceans where these metal ions are depleted (such as the HNLC region of the North Pacific Ocean) (Moore *et al.*, 2013). In the HNLC region of the North Pacific, where *P. granii* rhodopsin-containing cells were initially identified (Marchetti *et al.*, 2012), primary production may be limited by iron and affected by other trace metals (Saito *et al.*, 2008). In other words, our proposed mechanism appears to be particularly effective in the ocean where trace metals involved in CCM are depleted.

In the present study, we clarified the function and subcellular localization of PngR in a photosynthetic diatom. The results obtained suggest that proton transport by rhodopsin changes pH inside the outermost membrane of the plastid (CERM). A quantitative simulation indicated that the creation of an acidic pool by light provides positive feedback on C fixation efficiency, while alkalization of the middle space may restrict C fixation. If PngR acidifies the middle space, diatom rhodopsin may contribute to CCM (Extended Data Fig. 8). Future analyses of cultured rhodopsin-bearing microbial eukaryotes will corroborate the present results and promote further research on the mechanisms by which rhodopsin-mediated proton transport promotes their growth in the ocean.

Acknowledgement

We thank Dr. Yusuke Matsuda for his useful discussions.

Funding

This work was supported by JSPS KAKENHI Grant Numbers 18K19224, 18H04136, and 22H00557 to S.Y., 19H04727, 21H00404, and 21H02446 to Y.S., and 19H03274 to R.K., and NSF grant OPP1745036 to A.M. This research was partially supported by the Interdisciplinary Collaborative Research Program of the Atmosphere and Ocean Research Institute, the University of Tokyo.

References

Balashov, S.P., Imasheva, E.S., Boichenko, V.A., Anton, J., Wang, J.M., and Lanyi, J.K. (2005) Xanthorhodopsin: a proton pump with a light-harvesting carotenoid antenna. *Science* **309**: 2061–2064.

Beja, O., Aravind, L., Koonin, E.V., Suzuki, M.T., Hadd, A., Nguyen, L.P., *et al.* (2000) Bacterial rhodopsin: evidence for a new type of phototrophy in the sea. *Science* **289**: 1902–1906.

Berg, J., Tymoczko, J., and Stryer, L. (2010) *Biochemistry, 7th edn*. New York, NY: WH Freeman.

Burns, B.D., and Beardall, J. (1987) Utilization of inorganic carbon by marine microalgae. *J Exp Mar Biol Ecol* **107**: 75–86.

Czech, L., Barbera, P., and Stamatakis, A. (2020) Genesis and Gappa: processing, analyzing and visualizing phylogenetic (placement) data. *Bioinformatics* **36**: 3263–3265.

Dorrell, R.G., Azuma, T., Nomura, M., de Kerdrel, G.A., Paoli, L., Yang, S.S., *et al.* (2019) Principles of plastid reductive evolution illuminated by nonphotosynthetic chrysophytes. *Proc Natl Acad Sci U S A* **116**: 6914–6923.

Eichner, M., Thoms, S., Rost, B., Mohr, W., Ahmerkamp, S., Ploug, H., *et al.* (2019) N_2 fixation in free-floating filaments of *Trichodesmium* is higher than in transiently suboxic colony microenvironments. *New Phytol* **222**: 852–863.

Emerson, S., and Hedges, J. (2008) *Chemical Oceanography and the Marine Carbon Cycle*. Cambridge, UK: Cambridge University Press.

Ernst, O.P., Lodowski, D.T., Elstner, M., Hegemann, P., Brown, L.S., and Kandori, H. (2014) Microbial and animal rhodopsins: structures, functions, and molecular mechanisms. *Chem Rev* **114**: 126–163.

Field, C.B., Behrenfeld, M.J., Randerson, J.T., and Falkowski, P. (1998) Primary production of the biosphere: Integrating terrestrial and oceanic components. *Science* **281**: 237–240.

Gee, C.W., and Niyogi, K.K. (2017) The carbonic anhydrase CAH1 is an essential component of the carbon-concentrating mechanism in Nannochloropsis oceanica. *Proc Natl Acad Sci U S A* **114**: 4537–4542.

Hasegawa, M., Hosaka, T., Kojima, K., Nishimura, Y., Nakajima, Y., Kimura-Someya, T., *et al.* (2020) A unique clade of light-driven proton-pumping rhodopsins evolved in the cyanobacterial lineage. *Sci Rep* **10**: 16752.

Hopkinson, B.M., Dupont, C.L., Allen, A.E., and Morel, F.M.M. (2011) Efficiency of the CO2-concentrating mechanism of diatoms. *Proc Natl Acad Sci U S A* **108**: 3830–3837.

Hopkinson, B.M. (2014) A chloroplast pump model for the CO2 concentrating mechanism in the diatom Phaeodactylum tricornutum. *Photosynth Res* **121**: 223–233.

Inomura, K., Wilson, S.T., and Deutsch, C. (2019) Mechanistic model for the coexistence of nitrogen fixation and photosynthesis in marine *Trichodesmium*. *mSystems* **4**: e00210-19.

Inoue, S., Yoshizawa, S., Nakajima, Y., Kojima, K., Tsukamoto, T., Kikukawa, T., and Sudo, Y. (2018) Spectroscopic characteristics of Rubricoccus marinus xenorhodopsin (RmXeR) and a putative model for its inward H+ transport mechanism. *Phys Chem Chem Phys* **20**: 3172–3183.

Jensen, E.L., Maberly, S.C., and Gontero, B. (2020) Insights on the functions and ecophysiological relevance of the diverse carbonic anhydrases in microalgae. *Int J Mol Sci* **21**: 2922.

Katoh, K., and Standley, D.M. (2013) MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Mol Biol Evol* **30**: 772–780.

Keeling, P.J. (2004) Diversity and evolutionary history of plastids and their hosts. *Am J Bot* **91**: 1481–1493.

Keeling, P.J., Burki, F., Wilcox, H.M., Allam, B., Allen, E.E., Amaral-Zettler, L.A., *et al.* (2014) The Marine Microbial Eukaryote Transcriptome Sequencing Project (MMETSP): Illuminating the functional diversity of eukaryotic life in the oceans through transcriptome sequencing. *PLoS Biol* **12**: e1001889.

Kikuchi, M., Kojima, K., Nakao, S., Yoshizawa, S., Kawanishi, S., Shibukawa, A., *et al.* (2021) Functional expression of the eukaryotic proton pump rhodopsin OmR2 in Escherichia coli and its photochemical characterization. *Sci Rep* **11**: 14765.

Kojima, K., Miyoshi, N., Shibukawa, A., Chowdhury, S., Tsujimura, M., Noji, T., *et al.* (2020a) Green-sensitive, long-lived, step-functional anion channelrhodopsin-2 variant as a high-potential neural silencing tool. *J Phys Chem* **11**: 6214–6218.

Kojima, K., Ueta, T., Noji, T., Saito, K., Kanehara, K., Yoshizawa, S., *et al*. (2020b) Vectorial proton transport mechanism of RxR, a phylogenetically distinct and thermally stable microbial rhodopsin. *Sci Rep* **10**: 282.

Lueker, T.J., Dickson, A.G., and Keeling, C.D. (2000) Ocean pCO2 calculated from dissolved inorganic carbon, alkalinity, and equations for K1 and K2: validation based on laboratory measurements of CO2 in gas and seawater at equilibrium. *Mar Chem* **70**: 105–119.

Man, D.L., Wang, W.W., Sabehi, G., Aravind, L., Post, A.F., Massana, R., *et al.* (2003) Diversification and spectral tuning in marine proteorhodopsins. *EMBO J* **22**: 1725–1731.

Marchetti, A., Schruth, D.M., Durkin, C.A., Parker, M.S., Kodner, R.B., Berthiaume, C.T., *et al.* (2012) Comparative metatranscriptomics identifies molecular bases for the physiological responses of phytoplankton to varying iron availability. *Proc Natl Acad Sci U S A* **109**: E317–E325.

Marchetti, A., Catlett, D., Hopkinson, B.M., Ellis, K., and Cassar, N. (2015) Marine diatom proteorhodopsins and their potential role in coping with low iron availability. *ISME J* **9**: 2745–2748.

Matsen, F.A., Kodner, R.B., and Armbrust, E.V. (2010) pplacer: linear time maximum-likelihood and Bayesian phylogenetic placement of sequences onto a fixed reference tree. *BMC Bioinf* **11**: 538.

Matsuda, Y., Hopkinson, B.M., Nakajima, K., Dupont, C.L., and Tsuji, Y. (2017) Mechanisms of carbon dioxide acquisition and CO2 sensing in marine diatoms: a gateway to carbon metabolism. *Philos Trans R Soc B* **372**: 20160403.

Miranda, M.R.M., Choi, A.R., Shi, L.C., Bezerra, A.G., Jung, K.H., and Brown, L.S. (2009) The photocycle and proton translocation pathway in a cyanobacterial ion-pumping rhodopsin. *Biophys J* **96**: 1471–1481.

Miyahara, M., Aoi, M., Inoue-Kashino, N., Kashino, Y., and Ifuku, K. (2013) Highly efficient transformation of the diatom Phaeodactylum tricornutum by multi-pulse electroporation. *Biosci Biotechnol Biochem* **77**: 874–876.

Moore, C.M., Mills, M.M., Arrigo, K.R., Berman-Frank, I., Bopp, L., Boyd, P.W., *et al*. (2013) Processes and patterns of oceanic nutrient limitation. *Nat Geosci* **6**: 701–710.

Nagel, G., Ollig, D., Fuhrmann, M., Kateriya, S., Mustl, A.M., Bamberg, E., and Hegemann, P. (2002) Channelrhodopsin-1: A light-gated proton channel in green algae. *Science* **296**: 2395–2398.

Nakajima, K., Tanaka, A., and Matsuda, Y. (2013) SLC4 family transporters in a marine diatom directly pump bicarbonate from seawater. *Proc Natl Acad Sci U S A* **110**: 1767–1772.

Nelson, D.M., Treguer, P., Brzezinski, M.A., Leynaert, A., and Queguiner, B. (1995) Production and dissolution of biogenic silica in the ocean—Revised global estimates, comparison with regional data and relationship to biogenic sedimentation. *Global Biogeochem Cycles* **9**: 359–372.

Oesterhelt, D., and Stoeckenius, W. (1971) Rhodopsin-like protein from the purple membrane of *Halobacterium halobium*. *Nature (London), New Biol* **233**: 149–152.

Raven, J.A. (1997) CO2-concentrating mechanisms: A direct role for thylakoid lumen acidification? *Plant, Cell Environ* **20**: 147–154.

Riebesell, U., Wolfgladrow, D.A., and Smetacek, V. (1993) Carbon-dioxide limitation of marine-phytoplankton growth-rates. *Nature* **361**: 249–251.

Saito, M.A., Goepfert, T.J., and Ritt, J.T. (2008) Some thoughts on the concept of colimitation: Three definitions and the importance of bioavailability. *Limnol Oceanogr* **53**: 276–290.

Sineshchekov, O.A., Jung, K.H., and Spudich, J.L. (2002) Two rhodopsins mediate phototaxis to low- and high-intensity light in Chlamydomonas reinhardtii. *Proc Natl Acad Sci U S A* **99**: 8689–8694.

Slamovits, C.H., Okamoto, N., Burri, L., James, E.R., and Keeling, P.J. (2011) A bacterial proteorhodopsin proton pump in marine eukaryotes. *Nat Commun* **2**: 183.

Spudich, J.L., Yang, C.S., Jung, K.H., and Spudich, E.N. (2000) Retinylidene proteins: Structures and functions from archaea to humans. *Annu Rev Cell Dev Biol* **16**: 365–392.

Stamatakis, A. (2014) RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**: 1312–1313.

Stork, S., Moog, D., Przyborski, J.M., Wilhelmi, I., Zauner, S., and Maier, U.G. (2012) Distribution of the SELMA translocon in secondary plastids of red algal origin and predicted uncoupling of ubiquitin-dependent translocation from degradation. *Eukaryot Cell* **11**: 1472–1481.

Tsuji, Y., Kusi-Appiah, G., Kozai, N., Fukuda, Y., Yamano, T., and Fukuzawa, H. (2021) Characterization of a CO2-concentrating mechanism with low sodium dependency in the centric diatom Chaetoceros gracilis. *Mar Biotechnol* **23**: 456–462.

Yoshizawa, S., Kawanabe, A., Ito, H., Kandori, H., and Kogure, K. (2012) Diversity and functional analysis of proteorhodopsin in marine *Flavobacteria*. *Environ Microbiol* **14**: 1240–1248.

Microbes Environ. 39(5), 2024
https://www.jstage.jst.go.jp/browse/jsme2 doi:10.1264/jsme2.ME23089

HONORABLE MENTION PAPER OF THE YEAR 2024

Special issue: *Molecular Biology, Biochemistry, and Ecology of Extremophiles*

Accelerated Iron Corrosion by Microbial Consortia Enriched from Slime-like Precipitates from a Corroded Metal Apparatus Deployed in a Deep-sea Hydrothermal System

Satoshi Wakai[1,2]*, Sanae Sakai[1], Tatsuo Nozaki[3,4,5], Masayuki Watanabe[3], and Ken Takai[1]

[1]*Institute for Extra-cutting-edge Science and Technology Avant-garde Research (X-star), Japan Agency for Marine-Earth Science and Technology (JAMSTEC), Yokosuka, Japan;* [2]*PRESTO, Japan Science and Technology Agency (JST), Tokyo, Japan;* [3]*Submarine Resources Research Center, Research Institute for Marine Resources Utilization, Japan Agency for Marine-Earth Science and Technology (JAMSTEC), Yokosuka, Japan;* [4]*Frontier Research Center for Energy and Resources, School of Engineering, The University of Tokyo, Japan; and* [5]*Department of Planetology, Graduate School of Science, Kobe University, Kobe, Japan*

(Received October 31, 2023—Accepted March 14, 2024—Published online June 6, 2024)

Microbiologically influenced corrosion refers to the corrosion of metal materials caused or promoted by microorganisms. Although some novel iron-corrosive microorganisms have been discovered in various manmade and natural freshwater and seawater environments, microbiologically influenced corrosion in the deep sea has not been investigated in detail. In the present study, we collected slime-like precipitates composed of corrosion products and microbial communities from a geochemical reactor set on an artificial hydrothermal vent for 14.5 months, and conducted culture-dependent and -independent microbial community analyses with corrosive activity measurements. After enrichment cultivation at 37, 50, and 70°C with zero-valent iron particles, some of the microbial consortia showed accelerated iron dissolution, which was approximately 10- to 50-fold higher than that of the abiotic control. In a comparative analysis based on the corrosion acceleration ratio and amplicon sequencing of the 16S rRNA gene, three types of corrosion were estimated: the methanogen-induced type, methanogen–sulfate-reducing bacteria cooperative type, and sulfate-reducing *Firmicutes*-induced type. The methanogen-induced and methanogen–sulfate-reducing bacteria cooperative types were observed at 50°C, while the sulfate-reducing *Firmicutes*-induced type was noted at 37°C. The present results suggest the microbial components associated with microbiologically influenced corrosion in deep-sea hydrothermal systems, providing important insights for the development of future deep-sea resources with metal infrastructures.

Key words: microbiologically influenced corrosion, deep sea, microbial community, methanogen

Metal corrosion phenomena cause significant economic losses, including the deterioration of existing infrastructures (Cámara *et al.*, 2022; Xu *et al.*, 2023). Microorganisms are involved in these phenomena, which are referred to as microbiologically influenced corrosion (MIC). Characteristic forms of corrosion have been observed in MIC, and localized corrosion significantly affects the service lifespan of metal materials. In addition to the corrosion form, abnormally fast corrosion rates and sudden occurrence are also problematic. When rapid and sudden localized corrosion occurs in pipelines, secondary environmental pollution results from leaked oil (Jacobson, 2007), leading to significant social issues.

Although research on MIC has a long history (Garett, 1891; Gaines, 1910), limited information is available on the causal microorganisms and their mechanisms. Various corrosive microorganisms with high corrosion activity have been identified in recent years, and in addition to sulfate-reducing bacteria (SRB) that have attracted attention for a long time, iron-corrosive methanogens, iron-corrosive nitrate-reducing bacteria, and iron-corrosive acetic acid-producing bacteria have been isolated (Dinh *et al.*, 2004; Mori *et al.*, 2010; Uchiyama *et al.*, 2010; McBeth *et al.*, 2011; Wakai *et al.*, 2014; Kato *et al.*, 2015; Iino *et al.*, 2015; Hirano *et al.*, 2022). The corrosion mechanism of these microorganisms comprises two types: chemical MIC (CMIC), which is indirectly affected by microbial metabolites, such as acids, corrosive gases, and oxidants, and electrical MIC (EMIC), which is affected by the electrochemical activity of microorganisms (Enning and Garrelfs, 2014). The corrosion mechanism of SRB and methanogens with high corrosion activity corresponds to EMIC, which is enhanced by extracellular electron transfer (EET) ability (Deng *et al.*, 2015; Tsurumaru *et al.*, 2018; Liang *et al.*, 2021). Although the corrosion mechanism of these microorganisms has gradually been revealed, the detection of corrosive microorganisms in suspected cases of MIC in the field is challenging.

MIC occurs in various settings, including soil, freshwater,

* Corresponding author. E-mail: wakais@jamstec.go.jp;
Tel: +81–46–867–9693.

Citation: Wakai, S., Sakai, S., Nozaki, T., Watanabe, M., and Takai, K. (2024) Accelerated Iron Corrosion by Microbial Consortia Enriched from Slime-like Precipitates from a Corroded Metal Apparatus Deployed in a Deep-sea Hydrothermal System. *Microbes Environ* **39**: ME23089.
https://doi.org/10.1264/jsme2.ME23089

Copyright: © Japanese Society of Microbial Ecology / Japanese Society of Soil Microbiology / Taiwan Society of Microbial Ecology / Japanese Society of Plant Microbe Interactions / Japanese Society for Extremophiles
https://creativecommons.org/licenses/by/4.0/

and seawater environments, and the degree of corrosiveness of MIC depends on a combination of metal materials, environmental factors, and microorganisms. In recent years, we have conducted long-term immersion tests in a freshwater environment and demonstrated that the microbial community structure markedly changes during the corrosion process of carbon steel and other steels, resulting in the accumulation of SRB (Wakai *et al.*, 2022a). Similarly, the eventual accumulation of SRB with the progression of corrosion has been reported not only in freshwater, but also in seawater environments (Ramírez *et al.*, 2016; Vigneron *et al.*, 2016; Li *et al.*, 2017). On the other hand, in the corrosion products of chromium content and martensitic stainless steel corroded in the same freshwater environment, the presence of SRB was negligible, while sulfur-oxidizing bacteria and electrochemically active microbes were clearly detected (Wakai *et al.*, 2022a, 2022b). Therefore, the composition and transition of microbial components associated with MIC are affected not only by environmental conditions and microbial sources, but also by the target metal materials; therefore, corrosion reproduction tests are important for understanding MIC in practical environments.

Although previous studies examined MIC in various terrestrial and marine environments, limited information is available on corrosion in deep-sea environments (Ma *et al.*, 2020; Melchers, 2021; Rajala *et al.*, 2022). Although deep-sea environments are far from human activities, a large amount of metal equipment has been installed for purposes such as submarine cable internet connections, seafloor resource development, and various forms of monitoring. Current knowledge of MIC may not be fully applicable to deep-sea environments with low or high temperatures under high hydrostatic pressure. A microbiome analysis of the corrosion of deep-sea carbon steel over the course of 10 years indicated that sulfur-metabolizing microorganisms contribute to corrosion (Rajala *et al.*, 2022). Although these findings form the basis for understanding MIC in cold deep-sea environments, the microbial populations and mechanisms of MIC in deep-sea hydrothermal systems remain unknown.

Therefore, we herein examined high-temperature MIC in a deep-sea hydrothermal system. We collected slime-like precipitates from a corroded metal apparatus deployed in an artificial deep-sea hydrothermal vent, which were then cultivated with zero-valent iron medium under moderately thermal conditions. Using enriched microorganisms, we conducted comparative analyses of corrosion abilities and microbial communities. The results obtained revealed accelerated corrosion by moderately thermophilic microorganisms and possible corrosion-associated components.

Materials and Methods

Operation of the Kuroko-ore cultivation apparatus

A geochemical reactor, called the Kuroko-ore (also called black ore, containing hydrothermal sulfide minerals enriched in sphalerite and galena) cultivation apparatus, was installed on the artificial hydrothermal vent at Hole C9017A during the cruise CK16-05. It consisted of an inflow pipe, a cultivation cell, four outlet pipes, and two P/T sensors (Fig. 1A), as previously reported (Nozaki *et al.*, 2021; Kinoshita *et al.*, 2022). Two P/T sensors with short and long probes were inserted into the inflow pipe (basal part of the cell) and the cultivation cell near its top, respectively. Hydrothermal fluid was supplied to the sulfide mineral cultivation cell through the inflow pipe and then flowed out from four outlet pipes on the top of the cultivation cell. After 14.5 months of cultivation, the Kuroko-ore cultivation apparatus was recovered during the cruise KR18-02C, along with the in-side precipitates.

Sampling

A seal of the Kuroko-ore cultivation apparatus was opened, and six slime-like precipitates were recovered from different positions of the cell (Supplementary Fig. S1). Sample numbers (R01 to R07) corresponded to the order from the top of the cell to the bottom, and sample R06 was only used in another Kuroko-ore cultivation study because of the small amount obtained. Sample R02 was collected from near the center of the cell only, while the other samples consisted of mixtures from near the center and inner walls. These precipitates were recovered into 100-mL medium bottles, and the gas phase of each bottle was replaced with 100% N_2 gas. These bottles were sealed with a butyl rubber stopper and stored at 4°C until used.

Microbial cultivation

To enrich iron-corroding microorganisms, 20 mL of modified artificial seawater medium supplied with 2.0 g of iron granules (Fe >99.98%, 1–2 mm in diameter; Alfa Aesar) was prepared using 70-mL serum bottles (Uchiyama *et al.*, 2010). Before the inoculation, the gas phase in the medium bottle was replaced by purging with N_2/CO_2 (80:20 ratio) gas. Approximately 1 mL of the precipitates recovered from the Kuroko ore cultivation cell was inoculated into the medium, which was then purged again with N_2/CO_2 gas. The bottles were sealed with butyl-rubber stoppers and aluminum caps. Bottles containing each precipitate sample were incubated at 37, 50, and 70°C in duplicate.

After the enrichment cultivation, 1 mL of the culture was inoculated into 20 mL of a newly prepared modified artificial seawater medium containing iron foil (Fe >99.98%, 10×10×0.1 mm; Sigma-Aldrich) to conduct the iron corrosion test.

Analytical method

To evaluate corrosion activity, a 0.5-mL aliquot of the solution containing insoluble matter in the medium was sampled every 7 days after vigorous mixing with a vortex mixer. This was then mixed with 0.5 mL of 6 M HCl and 1 mL of 1 M L-ascorbic acid. The concentration of iron in the mixture was assessed colorimetrically with orthophenanthroline (Sandell, 1959).

Molecular hydrogen (H_2) and methane (CH_4) that accumulated in the headspace of each bottle were measured using a gas chromatograph (GC-3200; GL Sciences) equipped with a thermal conductivity detector and a Molecular sieve 5A column (GL Sciences) at 100°C using N_2 as a carrier gas.

DNA extraction

After the enrichment cultivation, 10 mL of each culture with insoluble particles of corrosion products and iron particles was sampled, and microbial cells were precipitated by centrifuging at 10,000×*g* for 10 min. DNA from the pellet was extracted using a Fast DNA Spin kit for Soil (MP Biomedicals) according to the manufacturer's protocol with slight modifications. The DNA concentration was measured using a Qubit dsDNA HS assay kit (Thermo Fisher Scientific) and a Qubit 4 Fluorometer (Thermo Fisher Scientific).

Amplicon sequencing of 16S rRNA gene fragments

Partial 16S rRNA genes (V4–V5 regions) were amplified by PCR with the primer sets 530F and 907R (Caporaso *et al.*, 2012), which contain overhung adapters at the 5' ends. We then performed PCR amplification, enzymatic purification, the addition of multiplexing indices and Illumina sequencing adapters, and purification

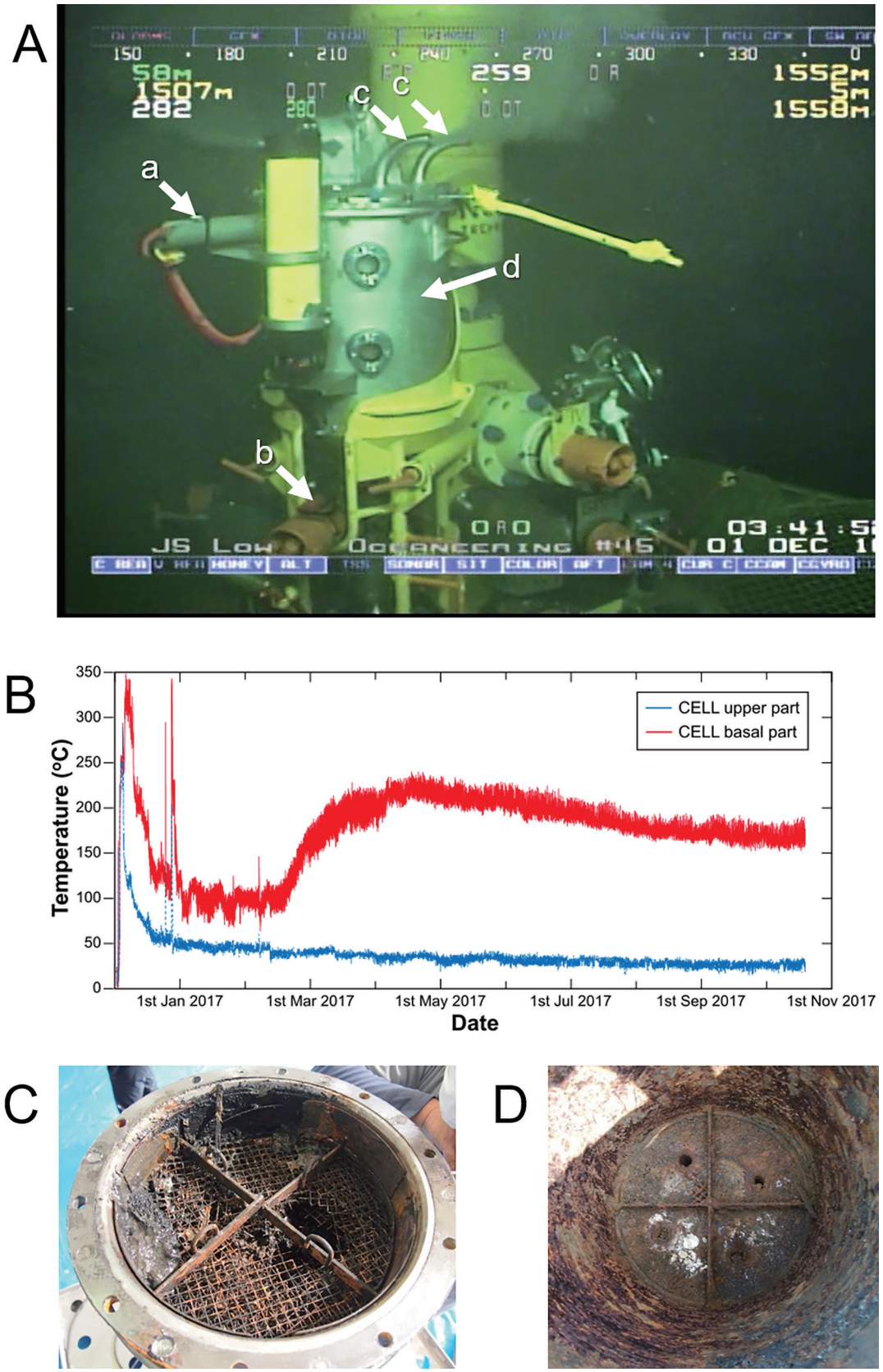


Fig. 1. Kuroko-ore cultivation apparatus. (A) Overview of the Kuroko-ore cultivation apparatus. Arrows show two long (a) and short (b) P/T sensors, outlet pipes (c), and the cultivation cell (d). (B) Secular changes in temperature in the upper and basal parts of the cell were monitored during the initial 11 months. (C and D) Pictures of the corroded stainless-steel mesh and the sediment/precipitate that formed in the cell of the Kuroko-ore cultivation apparatus and interior wall.

with magnetic beads as previously reported (Hirai *et al.*, 2017). Amplicon sequencing was conducted using an Illumina MiSeq platform and MiSeq v3 reagent (Illumina) for 300-bp paired-end reads, according to Illumina's standard protocol.

Raw FASTQ files generated by MiSeq were analyzed using the QIIME2 pipeline (Bolyen *et al.*, 2019). Paired-end FASTQ files were demultiplexed using a demux plugin based on their unique barcodes (Caporaso *et al.*, 2010). The demultiplexed sequences from each sample were treated using the dada2 plugin to obtain the feature table (Callahan *et al.*, 2016). A feature-classifier plugin was then used to align the feature sequences to a pretrained SILVA-138 99% database in order to generate the taxonomy table (Bokulich *et al.*, 2018). Data were rarefied prior to alpha- and beta-diversity analyses using a depth of 40,565 reads. Diversity metrics were calculated and plotted using the core-diversity and emperor plugins, respectively (Vázquez-Baeza *et al.*, 2013). To assess differences among the microbiomes of cultures at each temperature, a permutational multivariate analysis of variance (PERMANOVA) was conducted, which utilized weighted UniFrac distances and was performed with the qiime diversity beta-group significance tool. Each processed file was visualized via QIIME2 View (https://view.qiime2.org), and principal coordinate analysis plot data were exported with the qiime tools.

The datasets generated for this study may be obtained from NCBI using accession codes DRR426559–DRR426587.

Results

Corrosion behavior in the Kuroko-ore cultivation apparatus

During cultivation, fluid temperature in the Kuroko-ore cultivation apparatus was monitored in the upper and basal parts of the cultivation cell. The initial temperatures in the upper and basal parts were approximately 250 and 300°C, respectively (Fig. 1B). Fluid temperatures then suddenly decreased to approximately 50 and 100°C, respectively. The temperature in the upper part remained stable, whereas that in the basal part increased again by approximately 200°C, which was maintained for more than 6 months. Unfortunately, temperature measurements ceased in October 2017 because the memory of the P/T sensors reached their capacity limits. After the 14.5-month cultivation, large amounts of precipitates were observed within the cultivation cell, and stainless-steel meshes to trap generated minerals were severely corroded (Fig. 1C). The central part of each stainless-steel mesh had completely dissolved and disappeared. The remaining parts were also corroded, and their color had changed to black or reddish brown. After the removal of the stainless-steel meshes, large amounts of precipitates were observed at the basal part, and the interior wall of the cell had completely corroded (Fig. 1D). Since the dissolved oxygen concentration within the reactor was presumed to be nearly zero, mirroring that of the hydrothermal fluid, corrosion was expected to proceed under anaerobic conditions. Therefore, reddish brown-corrosion products may be formed by oxidization after recovery onboard.

Enrichment using zero-valent iron particle medium

Six slime-like precipitates containing microbial communities, corrosion products, and generated ore were collected and inoculated into modified artificial seawater medium with zero-valent iron particles. Cultivation was performed at 37, 50, and 70°C, and the dissolved iron concentration of each vial was periodically measured.

In the cultivation at 70°C, the dissolved iron concentration of the abiotic control was 4.3 mM, while those of the enriched cultures reached 12.1–33.5 mM (Fig. 2A). These concentrations increased linearly, and the acceleration ratio against the abiotic control was 2.2- to 8.6-fold higher.

In the cultivation at 50°C, the accelerated dissolution of iron was observed in all cultures (Fig. 2B). The dissolved iron concentration of the abiotic control was 3.3 mM, while those of the enriched cultures were 15.4–79.2 mM. Therefore, the degree of acceleration was divided into two levels: 1) a 30-fold increase for 50R01-1, 50R01-2, 50R03-2, 50R04-2, 50R05-1, and 50R05-2, and 2) a 4.3–13.7-fold increase for the remaining cultures. In the cultures with higher corrosion activity, the dissolved iron concentration was 2- to 4-fold higher than that in the cultures at 70°C. Dissolution in the 50R01-1, 50R04-2, 50R05-1, and 50R05-2 cultures showed a sigmoid curve, and extremely rapid dissolution was observed (35.6–47.8 mM 7 days^{-1}).

Since many mesophilic iron-corrosive microorganisms have been reported, we also conducted a cultivation at 37°C. The results obtained showed that the dissolved iron concentration of the abiotic control was 3.0 mM, while those of the enriched cultures were 10.0–67.6 mM (Fig. 2C). Culture 37R03-1 showed accelerated dissolution during the early cultivation phase, and the dissolution rate reached 45.8 mM 7 days^{-1}. In contrast, 37R03-2, 37R05-1, and 37R05-2

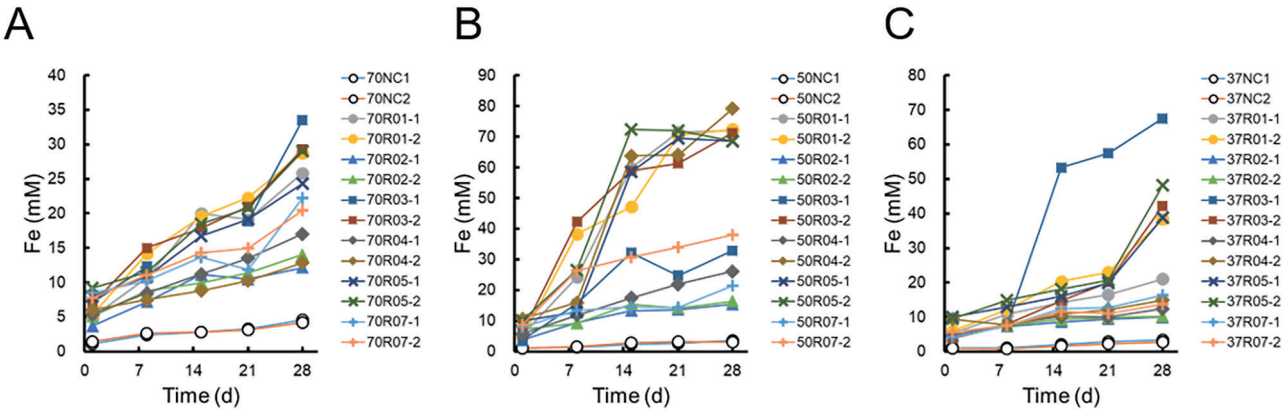


Fig. 2. Changes in iron concentrations during the 1st enrichment cultivation at 70°C (A), 50°C (B), and 37°C (C). Open circles represent abiotic controls: 70NC-1, 70NC-2, 50NC-1, 50NC-2, 37NC-1, and 37NC-2.

showed slow dissolution until 21 days and accelerated dissolution (approximately 20 mM 7 days^{-1}) in the late cultivation phase.

Iron-corroding test using enriched cultures

To investigate the relationship between corrosion behavior and gas production, a corrosion test was conducted using iron foil and enriched cultures. In the culture at 70°C, the amount of iron dissolution from the iron foil was slightly higher (<2-fold) than that of the abiotic control (Fig. 3A and B). The generation of similar amounts of H_2 was observed in all cultures containing the abiotic control (Fig. 3C). In contrast, methane was not detected in any of the samples (Fig. 3D).

In the cultures at 50°C, the amount of dissolved iron in all microbial cultures was higher than that in the abiotic control (Fig. 3E), and the ratio of accelerated dissolution was 4- to 10-fold higher (Fig. 3F). Culture 50R04-2 showed the highest ratio, reaching a 10.3-fold increase from the abiotic control. Approximately 20 μmol of H_2 was detected in the abiotic control, whereas only trace amounts were present in the other cultures (Fig. 3G). On the other hand, similar amounts of methane were detected in cultures 50R02-1, 50R02-2, 50R04-1, and 50R04-2 (Fig. 3H).

In the cultures at 37°C, the amount of iron dissolution was clearly divided into two levels: iron dissolution in cultures 37R02-1, 37R02-2, 37R04-1, 37R04-2, 37R07-1, and 37R07-2 was approximately 20 μmol, while that in cultures 37R01-1, 37R01-2, 37R03-1, 37R03-2, 37R05-1, and 37R05-2 ranged from 131 to 313 μmol (Fig. 3I). The cultures with higher iron dissolution activity showed acceleration ratios that were more than 20-fold higher than that of the abiotic control (Fig. 3J). Notably, culture 37R01-2 had an acceleration ratio that was approximately 50-fold higher. In contrast, cultures with low activity showed only 2- to 5-fold higher acceleration ratios than the abiotic control. Approximately 10 μmol of H_2 was detected in the abiotic control, 37R02-1, 37R04-1, and 37R04-2, while only trace amounts of H_2 were detected in the other cultures (Fig. 3K). Methane was detected in the cultures with trace amounts of

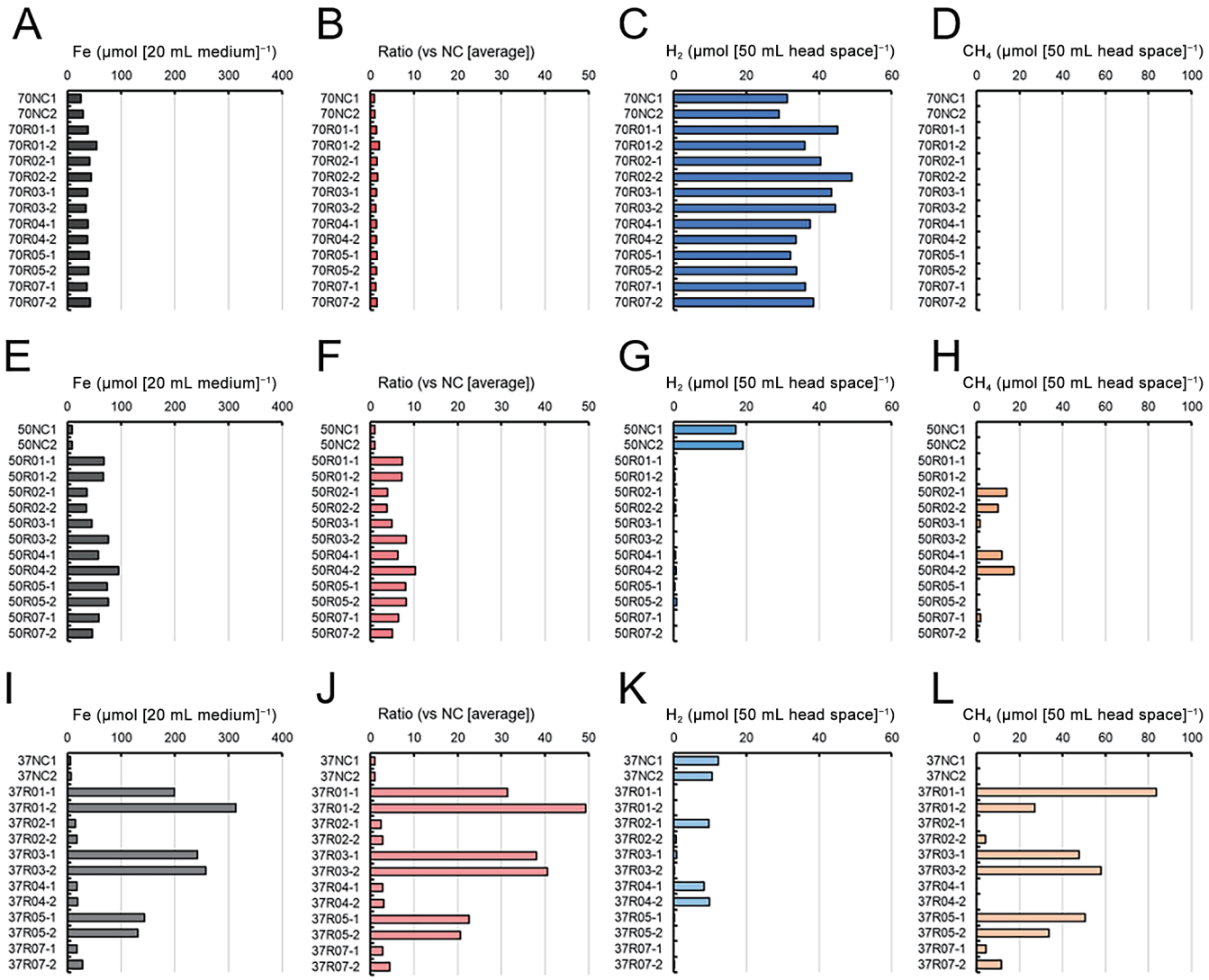


Fig. 3. Corrosion test using enriched cultures. Cultivation was conducted at 70°C (A, B, C, and D), 50°C (E, F, G, and H), and 37°C (I, J, K, and L). After cultivation, iron concentrations (A, E, and I) and the amounts of H_2 (C, G, and K) and CH_4 (D, H, and L) were measured. Each corrosion acceleration ratio (B, F, and J) was calculated from iron concentrations. NC represents the abiotic control.

H_2; notably, culture 37R01-1 had 83.7 μmol of CH_4 (Fig. 3H).

Relationship between iron dissolution and gas production

The electron equivalent for each reaction was calculated according to the following equations:

$Fe^0 \rightarrow Fe^{2+} + 2e^-$ (Eq. 1)
$2H^+ + 2e^- \rightarrow H_2$ (Eq. 2)
$CO_2 + 8H^+ + 8e^- \rightarrow CH_4 + 2H_2O$ (Eq. 3).

In the cultures at 70°C, each electron equivalent calculated from the amount of iron dissolution and H_2 generation was similar across all cultures (Fig. 4A). The amount of electrons released from iron foil through iron dissolution corresponded to that consumed during H_2 generation, according to the following equation:

$Fe^0 + 2H^+ \rightarrow Fe^{2+} + H_2$ (Eq. 4).

Therefore, these results indicate that corrosion at 70°C was induced by an abiotic chemical reaction.

In the cultures at 50°C, the electron equivalents consumed for methane production were higher than those for abiotic H_2 generation. Notably, the electron equivalents consumed for methane production in cultures 50R02-1, 50R02-2, 50R04-1, and 50R04-2 were similar to those of iron dissolution (Fig. 4B). In Eq. 1 and Eq. 3, this corrosion reaction may be represented as follows:

$4Fe^0 + CO_2 + 8H^+ \rightarrow 4Fe^{2+} + CH_4 + 2H_2O$ (Eq. 5).

In contrast, other cultures showed accelerated dissolution without methane production. These results indicated two types of corrosion, namely, methanogenesis-dependent and -independent types.

In the cultures at 37°C, those with higher acceleration ratios showed similar levels of electron equivalents between iron dissolution and methane production, according to Eq. 5; however, these levels in cultures 37R01-1 and 37R01-2 were significantly different (Fig. 4C). The electron equivalent of iron dissolution in culture 37R01-2 was approximately 3-fold higher than that of methane production, indicating a mixture of methanogenesis-dependent and -independent types. In contrast, in cultures 37R02-2, 37R07-1, and 37R07-2, which had low acceleration ratios and weak methane production, the electron equivalents of iron dissolution corresponded to those of methane production. Moreover, in cultures 37R02-1, 37R04-1, and 37R04-2, which had low acceleration ratios and weak H_2 generation, the electron equivalents of H_2 generation were similar to those of the abiotic control, while those of iron dissolution were slightly higher.

Alpha- and beta-diversities of microbial communities in enriched cultures

Since differences were observed in accelerated dissolution among the enriched cultures, amplicon sequencing was conducted. Out of 36 enriched cultures, the microbial communities in 29 cultures were analyzed because PCR products were not obtained from the other 7 samples. In the amplicon sequencing analysis, 1,905,643 reads were obtained with a range of 52,187–84,068 reads after denoising and removing chimeric reads (Supplementary Table S1). Using these reads, we calculated Chao1 and Shannon indices as a measure of alpha-diversity. Chao1 and Shannon indices in the enriched cultures at 70°C were 136–869 and 3.58–6.69, respectively, and were higher than those at 50 and 37°C (Fig. 5A and B, and Supplementary Table S1). The principal coordinate analysis plot based on the weighted UniFrac distance showed two clusters of cultures: one at 70°C and another at 37 and 50°C (Fig. 5C). This separation was verified by PERMANOVA (70°C–50°C: P=0.001, 70°C–37°C: P=0.001, and 50°C–37°C: P=0.07). These results suggest that unique microbial communities were formed by enrichment in the cultures at 37 and 50°C.

Microbial community structures at the phylum level

Firmicutes and *Euryarchaeota* sequences were observed

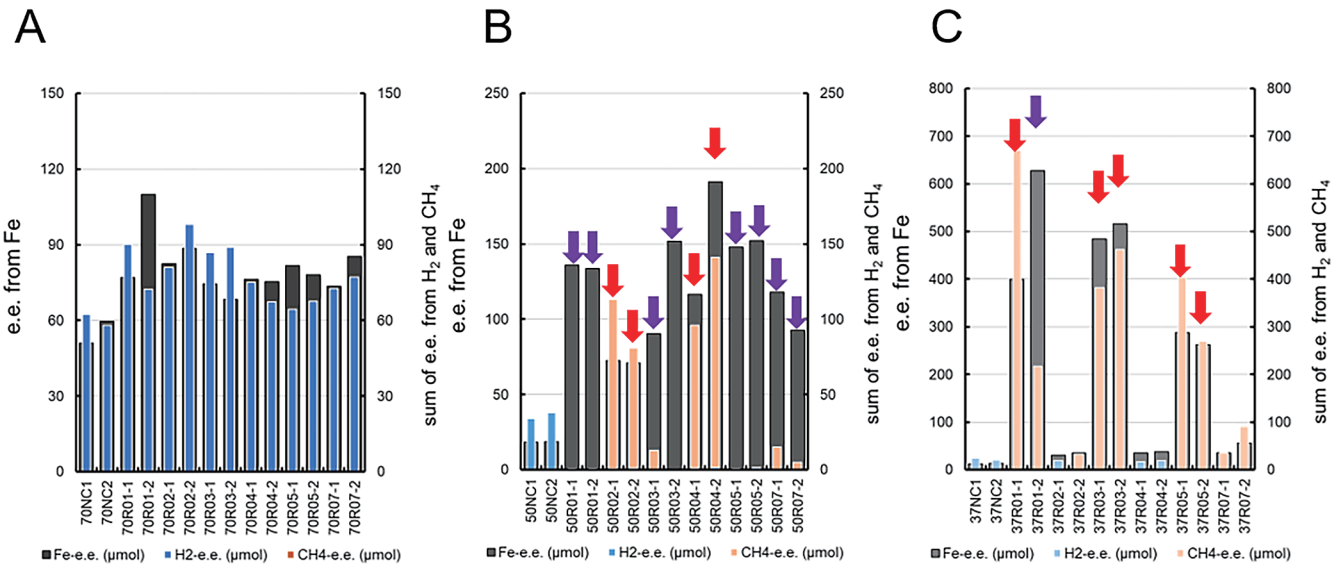


Fig. 4. Electron equivalents for iron dissolution and H_2 and CH_4 generation at 70°C (A), 50°C (B), and 37°C (C). Purple and red arrows represent H_2-consumed-type and methanogenesis-type corrosion, respectively. NC represents the abiotic control.

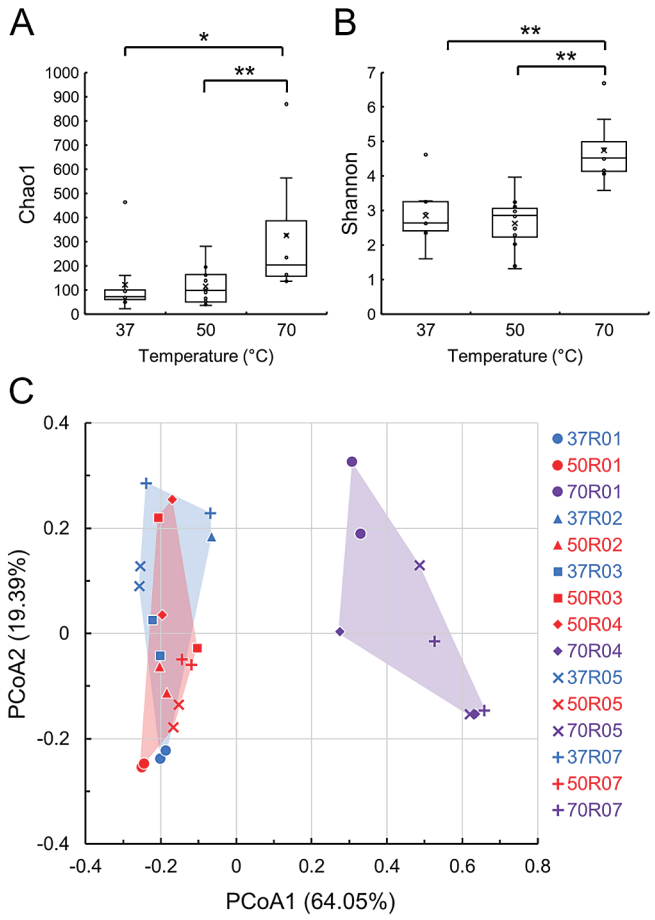


Fig. 5. Alpha- and beta-diversities of microbial communities. Chao1 (A) and Shannon (B) indices are represented as box-and-whisker plots. (A) and (B) were calculated based on each cultivation temperature (n=9, 12, and 8 for 37, 50, and 70°C, respectively). The line in the middle of the box, top and bottom of the box, whiskers, and cross symbols represent the median, 25 and 75 percentiles, min-to-max values, and average, respectively. Principal coordinate analysis (PCoA) plots based on unweighted UniFrac distances (C). Blue, red, and purple symbols indicate cultures at 37, 50, and 70°C, respectively.

in the majority of cultures (Fig. 6). In addition, the abundance of *Proteobacteria* and *Planctomycetota* sequences was higher in the culture at 70°C. On the other hand, *Halanaerobiaeota* members were observed in the cultures at 37 and 50°C, while *Spirochaeotota* sequences were only detected in cultures at 37°C. In addition, *Desulfobacterota* members, which are known to contain many SRB and are associated with MIC, showed a minor population of only 0.02–1.32% in amplicon sequence compositions at all temperatures. These results suggest that four major taxa (*Firmicutes*, *Euryarchaeota*, *Proteobacteria*, and *Planctomycetota*) contributed to the accelerated dissolution of iron observed in enrichment cultivation.

Microbial community structures at the amplicon sequence variant (ASV) level

Fig. 7 shows the relative abundance of the top 20 representatives and others at the ASV level. In all cultures, except for 70R04-2 (58.0%), 70R07-1 (79.5%), 70R07-2 (71.7%), and 37R02-2 (71.1%), the sum of these representatives was higher than 80%. Of the top 20 representatives, three ASVs were assigned as "genus *Desulfallas-Sporotomaculum*"-related bacteria. As a result of the BLAST search using each representative sequence, <97% homology was revealed in the following: "genus *Desulfallas-Sporotomaculum*", *Desulfoscipio geothermicus* (NR_119245, 96.05%); "genus *Desulfallas-Sporotomaculum* (BRH-c8a)", *D. geothermicus* (NR_119245, 96.02%); and "genus *Desulfallas-Sporotomaculum* (marine)", and *Desulfallas thermosapovorans* (NR_119247, 94.96%). Therefore, "genus *Desulfallas-Sporotomaculum*" and "genus *Desulfallas-Sporotomaculum* (BRH-c8a)" were denoted as *Desulfoscipio*-related bacterium 1 and 2, respectively, and "genus *Desulfallas-Sporotomaculum* (marine)" as *Desulfallas*-related bacterium.

In the cultures at 70°C, the sequence of *Lactobacillus* sp. was detected with the highest relative abundance (20.8–49.6%) (Fig. 7). In addition, the sequence of another *Lactobacillus* sp. was observed with 5.5–11.7% relative abundance, and the sum of *Lactobacillus* sp. reached 24.7–61.2%. On the other hand, the sequences of *Thermococcus* sp. (11.0–49.4%), family *Methanococcaceae* (0.6–7.7%), and an unassigned bacterium (5.7–12.3%) were also observed.

In the cultures at 50°C, the sequences of *Brassicibacter* sp., family *Methanococcaceae*, *Desulfohalotomaculum* sp., *Desulfotomaculum* sp., *Desulfallas*-related bacterium, *Anoxybacter* sp., *Halocella* sp., and *Thioreducter* sp. were found in high abundance (Fig. 7). In cultures 50R01-1 and 50R01-2, the percentages of *Brassicibacter* sp. sequences were 74.6 and 83.3%, respectively, while the sequences of *Desulfallas*-related bacterium and *Halocella* sp. were also detected, ranging from 4.3–13.2%. In 50R02-1, 50R02-2, 50R05-1, and 50R05-2, the sequence of *Brassicibacter* sp. had the highest relative abundance (40.2–48.7%). In 50R02-1 and 50R02-2, the *Desulfotomaculum* sp. sequence was the second most abundant (21.1 and 36.8%, respectively), followed by the sequence from the family *Methanococcaceae*, which accounted for 14.8 and 14.6%, respectively. However, in 50R05-1 and 50R05-2, the sequence of *Desulfallas*-related bacterium was the second highest (31.0–25.2%), and the sequences of family *Methanococcaceae* and *Thioreducter* sp. were observed at an abundance of 5.0–11.1%. In 50R03-2 and 50R04-2, the abundance of the family *Methanococcaceae* sequence was the highest (52.0 and 51.7%, respectively). In 50R07-1 and 50R07-2, the sequence of *Desulfohalotomaculum* sp. was observed at the highest percentage (44.9%). Representative microorganisms with the highest abundance in 50R03-1 and 50R04-1 were *Anoxybacter* sp. (32.2%) and the genus *Desulfallas*-related bacterium (36.2%), respectively.

In the cultures at 37°C, the sequences of *Desulfoscipio*-related bacterium 1 and 2, family *Methanococcaceae*, *Desulfallas*-related bacterium, *Brassicibacter* sp., and *Sphaerochaeta* sp. were detected as major representatives (Fig. 7). In 37R01-1 and 37R01-2, *Desulfoscipio*-related bacterium 2 and 1 showed the highest percentages (51.0 and 70.9%, respectively). In other cultures, the sequence of family *Methanococcaceae* was observed at a higher percentage (22.9–54.0%). In addition, in 37R03-1, 37R03-2, 37R05-1, and 37R05-2, the sequence of *Desulfallas*-related bacterium was present at a higher percentage (20.6–47.0%). Moreover,

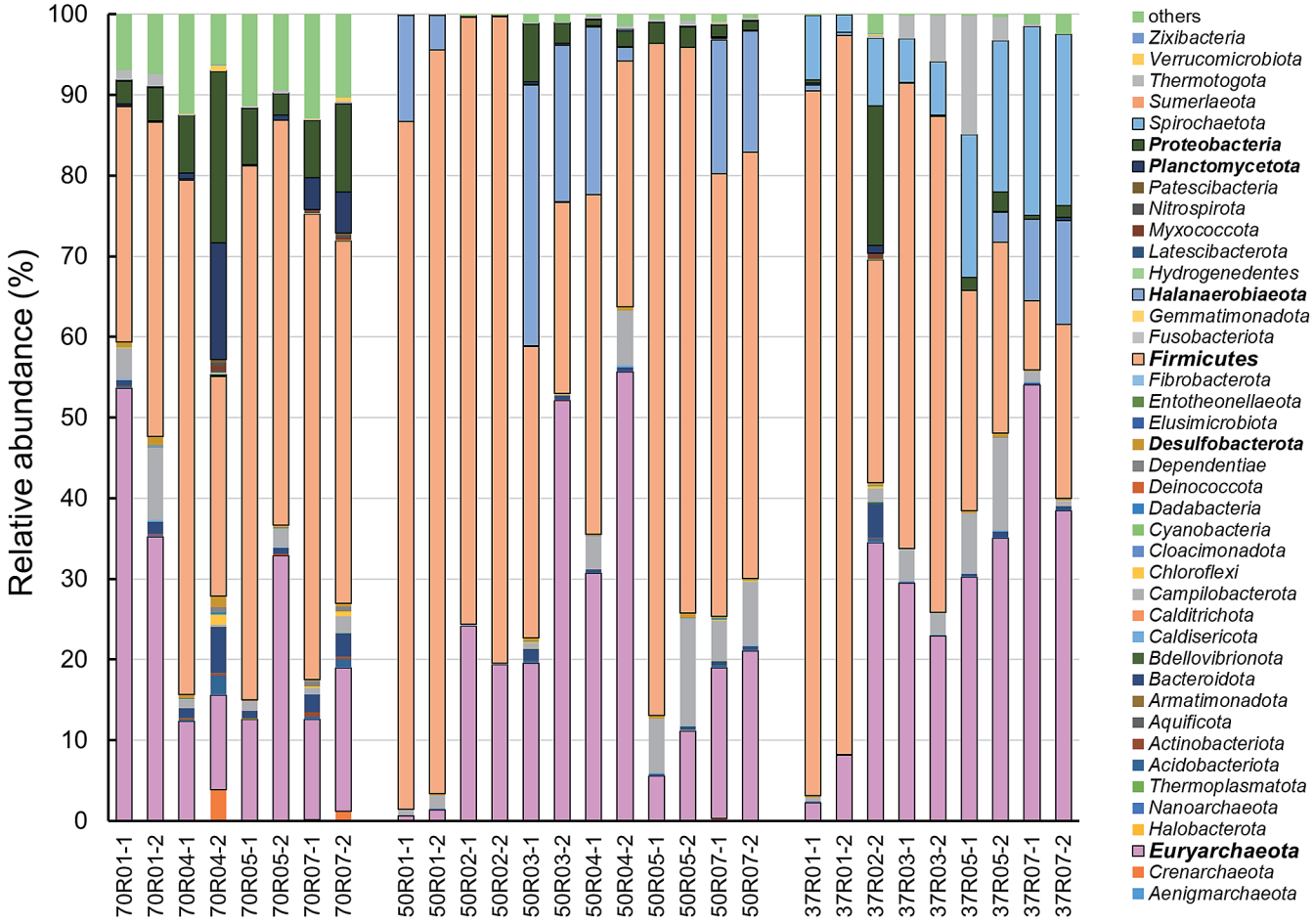


Fig. 6. Relative abundance (phylum level) of microorganisms in each sample.

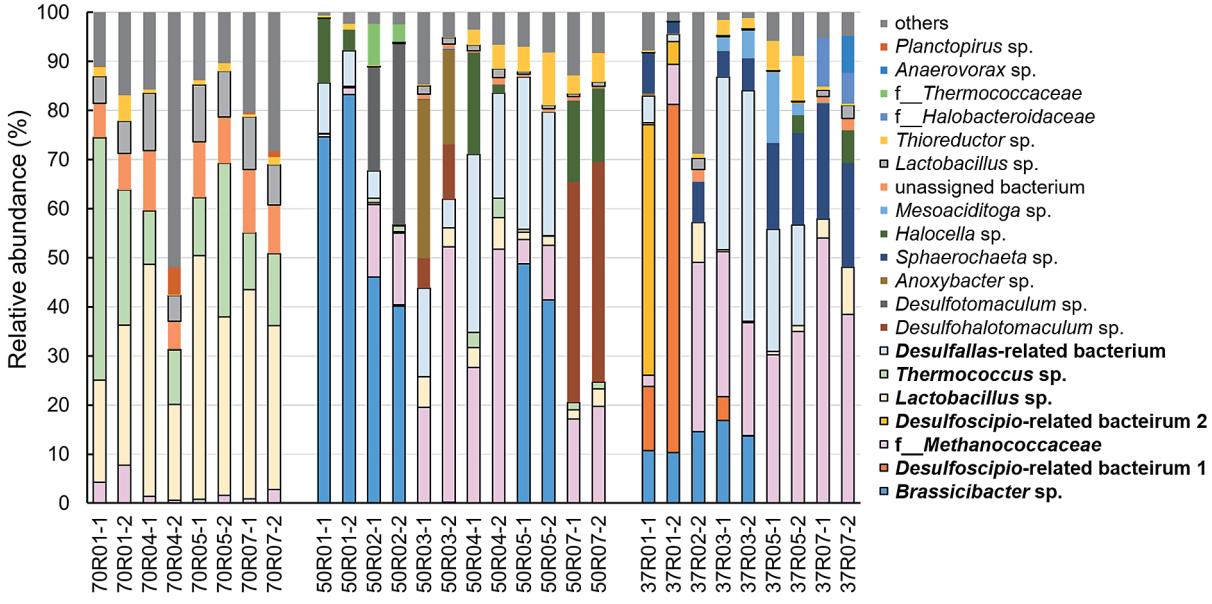


Fig. 7. Relative abundance (ASV level) of microorganisms in each sample.

the sequence of *Brassicibacter* sp. was only observed in 37R01-1, 37R01-2, 37R02-2, 37R03-1, and 37R03-2. In contrast, the relative abundance of the *Sphaerochaeta* sp. sequence was lower in these cultures than in the remaining cultures.

Discussion

In the present study, we investigated microbial populations and their role in MIC by examining corrosion precipitates in a deep-sea hydrothermal environment. Our culture-

based corrosion tests confirmed the occurrence of MIC at 50 and 37°C, while no significant MIC was evident at 70°C.

Microorganisms associated with MIC at 50°C

In the culture-based corrosion test at 50°C, two types of MIC were observed: a methanogenic type and non-methanogenic type. The methanogenic type was detected in cultures 50R02-1, 50R02-2, 50R04-1, and 50R04-2. The electron equivalents consumed for methane production in these cultures were similar to those used for iron dissolution, as indicated by Eq. 5. However, these equivalents were 2- to 4-fold higher than those consumed for H_2 generation in the abiotic control (Fig. 4). Accelerated corrosion cannot be explained by only methane production based on the consumption of H_2 resulting from chemical corrosion, as denoted in Eq. 4. Therefore, we classified accelerated corrosion as methanogenic-type MIC. Methanogenic metabolism appears to directly correlate with the accelerated dissolution of iron. It is important to note that MIC by some methanogens has already been reported (Dinh *et al.*, 2004; Mori *et al.*, 2010; Uchiyama *et al.*, 2010; Hirano *et al.*, 2022).

Although family *Methanococcaceae* sequences were observed in almost all cultures at 50°C, the acceleration ratios in methanogenesis-dependent corrosion positively correlated with the relative abundance of the archaeal sequences (Fig. 8A). A BLAST search using the representative sequence of the family *Methanococcaceae* observed in these cultures revealed the highest identity with the thermophilic, hydrogenotrophic methanogen, *Methanothermococcus okinawensis* strain IH-1 (accession number NR_102915, 100% identity). Since a culture strain of *M. okinawensis* was isolated from the hydrothermal chimney in the same deep-sea hydrothermal field (Takai *et al.*, 2002), the detection of this archaeon appears to be reasonable. Although MIC by the genera *Methanobacterium* and *Methanococcus* has been reported (Dinh *et al.*, 2004; Mori *et al.*, 2010, Uchiyama *et al.*, 2010, Hirano *et al.*, 2022), there is currently no information on MIC by the genus *Methanothermococcus*.

In addition to the family *Methanococcaceae*-related archaeon, a positive correlation was observed between the percentage of *Thioreductor* sp. sequences and acceleration ratios (Fig. 8B). The BLAST search, using the representative sequence of *Thioreductor* sp., showed the highest identity with the mesophilic, sulfate-reducing bacterium, *Thioreductor micantisoli* BKB25Ts-Y (accession number NR_041022, 98.11% identity). Although *T. micantisoli* BKB25Ts-Y was isolated from sediment near a hydrothermal vent in the same deep-sea field (Nakagawa *et al.*, 2005), this strain cannot grow at 50°C, which was the temperature in the enrichment and corrosion test. Therefore, the *Thioreductor* sequence observed in the culture may be a moderately thermophilic bacterium related to *T. micantisoli*. Similar to the family *Methanococcaceae*-related bacterium, there is currently no information on MIC by the genus *Thioreductor*.

The sequences of family *Methanococcaceae* and *Thioreductor* sp. were detected in the majority of cultures at 50°C. Additionally, a positive correlation was observed between the percentages of these sequences and acceleration ratios in methane-producing cultures. Therefore, methanogenic-type corrosion may result from the cooperative activity of methanogens and SRB, classifying them to methanogen–sulfate-reducing bacteria cooperative-type corrosion. In support of this result, corrosion by *Methanococcus maripaludis* KA1 was enhanced by a co-cultivation with SRB (Wakai, 2020). According to the proposed model, coexisting SRB have the capacity to generate hydrogen sulfide, which subsequently reacts with iron carbonate, a corrosion byproduct in MIC facilitated by iron-corrosive methanogens. This reaction generates electroconductive iron sulfide and releases carbonate ions. Iron-corrosive methanogens then utilize dissolved inorganic carbon released from the inert corrosion product as their methanogenic substrate. Iron-corrosive *M. maripaludis* KA1 and OS7 have an extracellular hydrogenase, which is homologous to the hydrogenase from the genus *Methanobacterium* and, thus, accelerated corrosion by these methanogens is considered to be the EMIC type based on the extracellular hydrogenase (Tsurumaru *et al.*, 2018). Therefore, the family *Methanococcaceae* member enriched in the cultivation test may also possess a similar corrosion mechanism. Alternatively, corrosion through electrosyntrophic metabolism has been proposed. Based on previous findings showing electrosyntrophic growth in microorganisms with electrochemical activity (Kato *et al.*, 2012), the co-existence of the family *Methanococcaceae* and

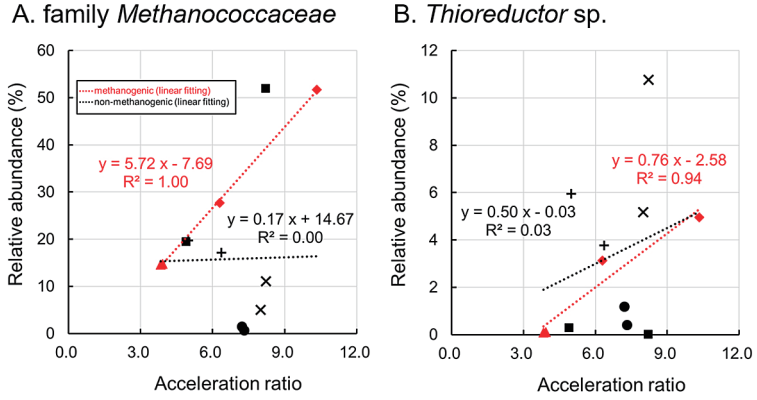


Fig. 8. Relationship between the acceleration ratio and relative abundance of family *Methanococcaceae* (A) and *Thioreductor* sp. (B) in the cultivation at 50°C. Red and black symbols represent methanogenic and non-methanogenic types, respectively. Circle, triangle, square, rhombus, cross, and plus represent samples R01, R02, R03, R04, R05, and R07, respectively.

Thioreductor sp. observed in the enriched cultures may be mutually beneficial via electrosyntrophic energy metabolism during the corrosion process.

Non-methanogenic corrosion was also observed. In these cultures, the percentage of specific representative microorganisms did not correlate with acceleration ratios (Supplementary Fig. S2A). In cultures 50R01-1 and 50R01-2, *Brassicibacter* sp. was significantly enriched (Fig. 7). However, the presence of this bacterium may not directly correlate with the corrosion process based on our BLAST search results. This search revealed that *Brassicibacter* sp. is closely related to the mesophilic heterotroph, *Brassicibacter mesophilus* (accession number NR_10884, 99.73%), which does not possess the ability to utilize reduced sulfur compounds or known EET systems (Fang *et al.*, 2012). In addition, the sequence of *Brassicibacter* sp. was detected at a relatively high abundance in cultures 50R02-1, 50R02-2, 50R05-1, and 50R05-2, among which 50R02-1 and 50R02-2 are methanogenic-type MIC. These results indicate that *Brassicibacter* sp. is irrelevant to the acceleration of corrosion. On the other hand, SRB, such as *Desulfallas*-related bacterium, *Desulfohalotomaculum* sp., and *Thioreductor* sp., were detected in non-methanogenic-type corrosion (Fig. 7). Therefore, these SRB may accelerate corrosion through the CMIC or EMIC system.

Microorganisms associated with MIC at 37°C

Six cultures showed highly accelerated methanogenic-type corrosion: 37R01-1, 37R01-2, 37R03-1, 37R03-2, 37R05-1, and 37R05-2 (Fig. 3 and 4). However, the abundance of family *Methanococcaceae* sequences negatively correlated with the acceleration ratio (Fig. 9A). Therefore, methanogenic energy metabolism may have contributed to some of the corrosion process, while other forms of microbial metabolism may have enhanced this process. In culture 37R01-2, which exhibited low methane production relative to the acceleration ratio, the sequence of *Desulfoscipio*-related bacterium 1 accounted for 70.9% of the microbial community's composition (Fig. 7). In addition, the sequences of other SRB, such as *Desulfoscipio*-related bacterium 2 and *Desulfallas*-related bacterium, were observed at higher percentages in the cultures with high acceleration ratios (Fig. 7). Although these SRB and other abundant microorganisms did not show a positive correlation with the acceleration ratio (Supplementary Fig. S2B), the sum of the percentages of these three SRB positively correlated with the acceleration ratio (Fig. 9B). Therefore, these SRB and members of family *Methanococcaceae* may be key players and electron consumers, respectively, in accelerated corrosion at 37°C.

The aforementioned SRB belong to the phylum *Firmicutes*, whereas the representative SRB related to MIC in many previous studies are members of the phylum *Desulfobacterota*, including the genus *Desulfovibrio* (Enning *et al.*, 2012; Waite *et al.*, 2020). Of these, *Desulfovibrio ferrophilus* and *Desulfopila corrodens* were shown to significantly accelerate corrosion through EMIC via EET (Dinh *et al.*, 2004; Enning *et al.*, 2012; Enning and Garrelfs, 2014; Kato *et al.*, 2016), and multi-heme-type outer membrane cytochromes played an important role. In addition, MIC by *Desulfovibrio vulgaris* has been demonstrated, and this microorganism has also been shown to induce CMIC-type corrosion because it lacks the outer membrane cytochrome required for EET. The relative species of *Firmicutes*-SRB enriched in our cultures included *D. geothermicus* (accession number NR_119245, 96.0% identity) and *D. thermosapovorans* (accession number NR_119247, 94.96% identity). Although their genome sequences have been analyzed, homologues of the outer membrane cytochrome have yet to be annotated, indicating that *Firmicutes*-SRB may not participate in EMIC-type corrosion. However, the acceleration ratios in these cultures were 20- to 49.6-fold higher than those of the abiotic control, and this significantly higher corrosive activity is difficult to explain through the corrosion mechanism of CMIC. Since biogenic iron sulfide nanoparticles have been shown to enhance EET (Deng *et al.*, 2020), in *Firmicutes*-SRB enriched cultures, these microorganisms may acquire EET through a similar mechanism, subsequently accelerating corrosion and, thus, we classified them to sulfate-reducing *Firmicutes*-induced-type corrosion.

Conclusion

In the present study, we provided evidence for MIC caused by moderately thermophilic and mesophilic microorganisms from deep-sea environments. In addition, we found that methanogenic-type corrosion predominated at 50°C,

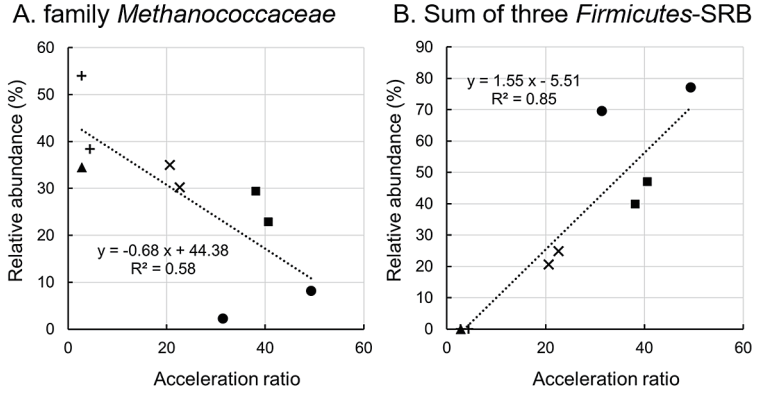


Fig. 9. Relationship between the acceleration ratio and relative abundance of family *Methanococcaceae* (A) and the sum of three *Firmicutes*-SRB (B) in the cultivation at 37°C. Circle, triangle, square, cross, and plus represent samples R01, R02, R03, R05, and R07, respectively.

while *Firmicutes*-SRB significantly accelerated corrosion at 37°C. While a quantitative assessment of the contribution of these microorganisms to corrosion is required, these microorganisms appear to play an important role in MIC in natural environments. However, previous studies on MIC primarily focused on sulfur-metabolizing bacteria in cold deep-sea environments. The present results provide insights into MIC in warm and hot deep-sea environments, bridging the gap in the literature and providing valuable insights for future research in this field.

Acknowledgements

We thank Miho HIRAI for her technical support as well as the captain, crew members, and ROV team of cruises CK16-05 and KR18-02C for the installation and recovery of the Kuroko-ore cultivation apparatus. This work was supported by a Japan Society for the Promotion of Science KAKENHI Grant (Grant Number: 17H04719), a JST PRESTO Grant (Grant Number: JPMJPR21NA), a Research Grant of ISIJ, and an MIC Research Grant of NACE-TJS.

References

Bokulich, N.A., Kaehler, B.D., Rideout, J.R., Dillon, M., Bolyen, E., Knight, R., *et al.* (2018) Optimizing taxonomic classification of marker-gene amplicon sequences with QIIME 2's q2-feature-classifier plugin. *Microbiome* **6**: 1–17.

Bolyen, E., Rideout, J.R., Dillon, M.R., Bokulich, N.A., Abnet, C.C., Al-Ghalith, G.A., *et al.* (2019) Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat Biotechnol* **37**: 852–857.

Callahan, B.J., McMurdie, P.J., Rosen, M.J., Han, A.W., Johnson, A.J.A., and Holmes, S.P. (2016) DADA2: High-resolution sample inference from Illumina amplicon data. *Nat Methods* **13**: 581–583.

Cámara, M., Green, W., MacPhee, C.E., Rakowska, P.D., Raval, R., Richardson, M.C., *et al.* (2022) Economic significance of biofilms: a multidisciplinary and cross-sectoral challenge. *npj Biofilms Microbiomes* **8**: 42.

Caporaso, J.G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F.D., Costello, E.K., *et al.* (2010) QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* **7**: 335–336.

Caporaso, J.G., Lauber, C.L., Walters, W.A., Berg-Lyons, D., Huntley, J., Fierer, N., *et al.* (2012) Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *ISME J* **6**: 1621–1624.

Deng, X., Nakamura, R., Hashimoto, K., and Okamoto, A. (2015) Electron extraction from an extracellular electrode by *Desulfovibrio ferrophilus* strain IS5 without using hydrogen as an electron carrier. *Electrochemistry* **83**: 529–531.

Deng, X., Dohmae, N., Kaksonen, A.H., and Okamoto, A. (2020) Biogenic iron sulfide nanoparticles to enable extracellular electron uptake in sulfate-reducing bacteria. *Angew Chem Int Ed Engl* **59**: 5995–5999.

Dinh, H.T., Kuever, J., Mußmann, M., Hassel, A.W., Stratmann, M., and Widdel, F. (2004) Iron corrosion by novel anaerobic microorganisms. *Nature* **427**: 829–832.

Enning, D., Venzlaff, H., Garrelfs, J., Dinh, H.T., Meyer, V., Mayrhofer, K., *et al.* (2012) Marine sulfate-reducing bacteria cause serious corrosion of iron under electroconductive biogenic mineral crust. *Environ Microbiol* **14**: 1772–1787.

Enning, D., and Garrelfs, J. (2014) Corrosion of iron by sulfate-reducing bacteria: new views of an old problem. *Appl Environ Microbiol* **80**: 1226–1236.

Fang, M.X., Zhang, W.W., Zhang, Y.Z., Tan, H.Q., Zhang, X.Q., Wu, M., and Zhu, X.F. (2012) *Brassicibacter mesophilus* gen. nov., sp. nov., a strictly anaerobic bacterium isolated from food industry wastewater. *Int J Syst Evol Microbiol* **62**: 3018–3023.

Gaines, R.H. (1910) Bacterial activity as a corrosive influence in the soil. *Ind Eng Chem* **2**: 128–130.

Garett, J.H. (1891) *The Action of Water on Lead*. London: H.K. Lewis.

Hirai, M., Nishi, S., Tsuda, M., Sunamura, M., Takaki, Y., and Nunoura, T. (2017) Library construction from subnanogram DNA for pelagic sea water and deep-sea sediments. *Microbes Environ* **32**: 336–343.

Hirano, S.I., Ihara, S., Wakai, S., Dotsuta, Y., Otani, K., Kitagaki, T., *et al.* (2022) Novel *Methanobacterium* strain induces severe corrosion by retrieving electrons from Fe^0 under a freshwater environment. *Microorganisms* **10**: 270.

Iino, T., Ito, K., Wakai, S., Tsurumaru, H., Ohkuma, M., and Harayama, S. (2015) Iron corrosion induced by non-hydrogenotrophic nitrate-reducing *Prolixibacter* sp. MIC1-1. *Appl Environ Microbiol* **81**: 1839–1846.

Jacobson, G.A. (2007) Corrosion at Prudhoe bay—a lesson on the line. *Mater Perform* **46**: 26–34.

Kato, S., Hashimoto, K., and Watanabe, K. (2012) Microbial interspecies electron transfer via electric currents through conductive minerals. *Proc Natl Acad Sci U S A* **109**: 10042–10046.

Kato, S., Yumoto, I., and Kamagata, Y. (2015) Isolation of acetogenic bacteria that induce biocorrosion by utilizing metallic iron as the sole electron donor. *Appl Environ Microbiol* **81**: 67–73.

Kinoshita, M., Kitada, K., and Nozaki, T. (2022) Tidally modulated temperature observed atop a drillsite at the Noho hydrothermal site, mid-Okinawa Trough. *J Geophys Res: Solid Earth* **127**: e2021JB023923.

Li, X., Duan, J., Xiao, H., Li, Y., Liu, H., Guan, F., and Zhai, X. (2017) Analysis of bacterial community composition of corroded steel immersed in Sanya and Xiamen seawaters in China via method of Illumina MiSeq sequencing. *Front Microbiol* **8**: 1737.

Liang, D., Liu, X., Woodard, T.L., Holmes, D.E., Smith, J.A., Nevin, K.P., *et al.* (2021) Extracellular electron exchange capabilities of *Desulfovibrio ferrophilus* and *Desulfopila corrodens*. *Environ Sci Technol* **55**: 16195–16203.

Ma, Y., Zhang, Y., Zhang, R., Guan, F., Hou, B., and Duan, J. (2020) Microbiologically influenced corrosion of marine steels within the interaction between steel and biofilms: a brief view. *Appl Microbiol Biotechnol* **104**: 515–525.

McBeth, J.M., Little, B.J., Ray, R.I., Farrar, K.M., and Emerson, D. (2011) Neutrophilic iron-oxidizing "Zetaproteobacteria" and mild steel corrosion in nearshore marine environments. *Appl Environ Microbiol* **77**: 1405–1412.

Melchers, R.E. (2021) Long-term corrosion of steels in deep, cold, low oxygen sea waters. *Corros Eng Sci Technol* **56**: 736–741.

Mori, K., Tsurumaru, H., and Harayama, S. (2010) Iron corrosion activity of anaerobic hydrogen-consuming microorganisms isolated from oil facilities. *J Biosci Bioeng* **110**: 426–430.

Nakagawa, S., Inagaki, F., Takai, K., Horikoshi, K., and Sako, Y. (2005) *Thioreductor micantisoli* gen. nov., sp. nov., a novel mesophilic, sulfur-reducing chemolithoautotroph within the epsilon-Proteobacteria isolated from hydrothermal sediments in the Mid-Okinawa Trough. *Int J Syst Evol Microbiol* **55**: 599–605.

Nozaki, T., Nagase, T., Torimoto, J., Takaya, Y., Ishibashi, J., Shimada, K., *et al.* (2021) Formation of highly Zn-enriched sulfide scale at a deep-sea artificial hydrothermal vent, Iheya-North Knoll, Okinawa Trough. *Miner Deposita* **56**: 975–990.

Rajala, P., Cheng, D.Q., Rice, S.A., and Lauro, F.M. (2022) Sulfate-dependant microbially induced corrosion of mild steel in the deep sea: a 10-year microbiome study. *Microbiome* **10**: 4.

Ramírez, G.A., Hoffman, C.L., Lee, M.D., Lesniewski, R.A., Barco, R.A., Garber, A., *et al.* (2016) Assessing marine microbial induced corrosion at Santa Catalina Island, California. *Front Microbiol* **7**: 1679.

Sandell, E.B. (1959) *Colorimetric Determination of Trace Metals. The Chemical Analysis Monograph Series*. *vol 3*. New York, NY: Interscience Publishers.

Takai, K., Inoue, A., and Horikoshi, K. (2002) *Methanothermococcus okinawensis* sp. nov., a thermophilic, methane-producing archaeon isolated from a Western Pacific deep sea hydrothermal vent system. *Int J Syst Evol Microbiol* **52**: 1089–1095.

Tsurumaru, H., Ito, N., Mori, K., Wakai, S., Uchiyama, T., Iino, T., *et al.* (2018) An extracellular [NiFe] hydrogenase mediating iron corrosion is encoded in a genetically unstable genomic island in *Methanococcus maripaludis*. *Sci Rep* **8**: 15149.

Uchiyama, T., Ito, K., Mori, K., Tsurumaru, H., and Harayama, S. (2010) Iron-corroding methanogen isolated from a crude-oil storage tank. *Appl Environ Microbiol* **76**: 1783–1788.

Vázquez-Baeza, Y., Pirrung, M., Gonzalez, A., and Knight, R. (2013) EMPeror: a tool for visualizing high-throughput microbial community data. *GigaScience* **2**: 16.

Vigneron, A., Alsop, E.B., Chambers, B., Lomans, B.P., Head, I.M., and Tsesmetzis, N. (2016) Complementary microorganisms in highly corrosive biofilms from an offshore oil production facility. *Appl Environ Microbiol* **82**: 2545–2554.

Waite, D.W., Chuvochina, M., Pelikan, C., Parks, D.H., Yilmaz, P., Wagner, M., *et al.* (2020) Proposal to reclassify the proteobacterial classes Deltaproteobacteria and Oligoflexia, and the phylum Thermodesulfobacteria into four phyla reflecting major functional capabilities. *Int J Syst Evol Microbiol* **70**: 5972–6016.

Wakai, S., Ito, K., Iino, T., Tomoe, Y., Mori, K., and Harayama, S. (2014) Corrosion of iron by iodide-oxidizing bacteria isolated from brine in an iodine production facility. *Microb Ecol* **68**: 519–527.

Wakai, S. (2020) Electron flow rate in microbiologically influenced corrosion and its applications. In *Electron Based Bioscience and Biotechnology (eBioX)*. Ishii, M., Wakai, S. (eds). Singapore: Springer, pp. 193–206.

Wakai, S., Eno, N., Miyanaga, K., Mizukami, H., Sunaba, T., and Miyano, Y. (2022a) Dynamics of microbial communities on the corrosion behavior of steel in freshwater environment. *npj Mater Degrad* **6**: 45.

Wakai, S., Eno, N., Mizukami, H., Sunaba, T., Miyanaga, K., and Miyano, Y. (2022b) Microbiologically influenced corrosion of stainless steel independent of sulfate-reducing bacteria. *Front Microbiol* **13**: 982047.

Xu, D., Gu, T., and Lovley, D.R. (2023) Microbially mediated metal corrosion. *Nat Rev Microbiol* **21**: 705–718.

Microbes Environ. 37(3), 2022
https://www.jstage.jst.go.jp/browse/jsme2 doi:10.1264/jsme2.ME22012

MOST CITED ORIGINAL PAPER PUBLISHED
IN THE LAST THREE YEARS [2022-2024] (30 citations)

Metabolic Potential of the Superphylum *Patescibacteria* Reconstructed from Activated Sludge Samples from a Municipal Wastewater Treatment Plant

Naoki Fujii[1], Kyohei Kuroda[2], Takashi Narihiro[2], Yoshiteru Aoi[3], Noriatsu Ozaki[1], Akiyoshi Ohashi[1], and Tomonori Kindaichi[1]*

[1]*Department of Civil and Environmental Engineering, Graduate School of Advanced Science and Engineering, Hiroshima University, 1–4–1, Kagamiyama, Higashihiroshima, Hiroshima 739–8527, Japan;* [2]*Bioproduction Research Institute, National Institute of Advanced Industrial Science and Technology (AIST), 2–17–2–1 Tsukisamu-Higashi, Toyohira-ku, Sapporo, Hokkaido 062–8517, Japan; and* [3]*Program of Biotechnology, Graduate School of Integrated Sciences for Life, Hiroshima University, 1–3–1, Kagamiyama, Higashihiroshima, Hiroshima 739–8530, Japan*

(Received February 25, 2022—Accepted May 14, 2022—Published online June 28, 2022)

Patescibacteria are widely distributed in various environments and often detected in activated sludge. However, limited information is currently available on their phylogeny, morphology, and ecophysiological role in activated sludge or interactions with other microorganisms. In the present study, we identified microorganisms that interacted with *Patescibacteria* in activated sludge via a correlation analysis using the 16S rRNA gene, and predicted the metabolic potential of *Patescibacteria* using a metagenomic analysis. The metagenome-assembled genomes of *Patescibacteria* consisted of three *Saccharimonadia*, three *Parcubacteria*, and one *Gracilibacteria*, and showed a strong positive correlation of relative abundance with *Chitinophagales*. Metabolic predictions from ten recovered patescibacterial and five *Chitinophagales* metagenome-assembled genomes supported mutualistic interactions between a member of *Saccharimonadia* and *Chitinophagales* via N-acetylglucosamine, between a member of *Parcubacteria* and *Chitinophagales* via nitrogen compounds related to denitrification, and between *Gracilibacteria* and *Chitinophagales* via phospholipids in activated sludge. The present results indicate that various interactions between *Patescibacteria* and *Chitinophagales* are important for the survival of *Patescibacteria* in activated sludge ecosystems.

Key words: *Patescibacteria*, candidate phyla radiation (CPR), activated sludge, metagenomic analysis, *Chitinophagales*

The phylogeny and physiology of microorganisms in activated sludge for wastewater treatment remain unclear due to their complexity and variations even though the activated sludge process has been used in wastewater treatment globally for more than 100 years. The structure of a microbial community depends on the climate, location, environment, and process configuration of the wastewater treatment plant (Zhang *et al.*, 2012). Recent studies reported that communities appeared to be stable at the genus and substrate specificity levels (Nielsen *et al.*, 2010; Kindaichi *et al.*, 2013). Most microorganisms, including those in nature and in engineered systems, cannot be cultured in laboratories and are called "microbial dark matter" (Rinke *et al.*, 2013). In this context, molecular biological methods, such as a 16S rRNA gene analysis and metagenomic analysis, have been widely used to predict the metabolic functions of uncultured bacteria. Large metagenomic data obtained from samples in various environments revealed the existence of a large group of bacteria, called *Patescibacteria* or candidate phyla radiation (CPR) (Rinke *et al.*, 2013; Brown *et al.*, 2015). The *Patescibacteria* or CPR (hereafter called *Patescibacteria*) group includes 35 phyla, accounts for 15–50% of all bacterial phyla, and has been reported to exist in various environments (Brown *et al.*, 2015; Hug *et al.*, 2016; Takebe *et al.*, 2020). To date, only a few of its members (*i.e.*, phyla *Saccharimonadia* and *Gracilibacteria*) have been cultured (Soro *et al.*, 2014; He *et al.*, 2015; Ibrahim *et al.*, 2021; Yakimov *et al.*, 2021). These bacteria are parasitic on other bacteria for sustenance; however, since most of them cannot be cultured, the mechanisms underlying their existence are unclear. *Patescibacteria* are commonly characterized by a small genome size (approximately 1.0 Mbp) (Lemos *et al.*, 2020; Nakai, 2020), limited metabolic potential, and fermentation-based metabolism (Wrighton *et al.*, 2012, 2014; Albertsen *et al.*, 2013; Lemos *et al.*, 2020). However, the physiology and phylogeny of *Patescibacteria* have not yet been elucidated in detail, except for some cultures in the phyla *Saccharimonadia* and *Gracilibacteria* (Soro *et al.*, 2014; He *et al.*, 2015; Moreira *et al.*, 2021; Yakimov *et al.*, 2021).

Activated sludge is an environment in which *Patescibacteria* are frequently detected. Among them, *Saccharimonadia*, *Parcubacteria*, and *Gracilibacteria* are the major phyla (Albertsen *et al.*, 2013; Kindaichi *et al.*, 2016; Singleton *et al.*, 2021). A moderate constituent of activated sludge is *Saccharimonadia*, a well-described *Patescibacteria* (Mielczarek *et al.*, 2012; Albertsen *et al.*, 2013; Kindaichi *et al.*, 2016). Based on the 16S rRNA gene classification, *Saccharimonadia* are primarily classified into

* Corresponding author. E-mail: tomokin@hiroshima-u.ac.jp;
Tel: +81–82–424–5718.

Citation: Fujii, N., Kuroda, K., Narihiro, T., Aoi, Y., Ozaki, N., Ohashi, A., and Kindaichi, T. (2022) Metabolic Potential of the Superphylum *Patescibacteria* Reconstructed from Activated Sludge Samples from a Municipal Wastewater Treatment Plant. *Microbes Environ* **37**: ME22012.
https://doi.org/10.1264/jsme2.ME22012

three subdivisions, with members having a filamentous morphology belonging to subdivision 1, and members with a coccus or rod morphology belonging to subdivisions 2 and 3 (Hugenholtz *et al.*, 2001). A complete genome belonging to subdivision 3 was reconstructed from activated sludge samples and the data obtained showed that *Saccharimonadia* are obligate fermentative metabolic bacteria that use heterolactic fermentation pathways (Albertsen *et al.*, 2013). In addition, filamentous *Saccharimonadia* were detected in activated sludge from wastewater treatment plants, and the characteristics of substrate utilization elucidated by microautoradiography combined with fluorescence *in situ* hybridization (FISH) revealed more diverse carbon metabolism, including the utilization of oleic acid and amino acids, which was not predicted from the available genome (Kindaichi *et al.*, 2016). *Parcubacteria* also belong to *Patescibacteria* and are found in activated sludge (Zhang *et al.*, 2012). *Parcubacteria* have a diverse distribution within the phylum, with most members being frequently found in anaerobic environments, such as groundwater. *Parcubacteria* are considered to be involved in hydrogen production, sulfur reduction, and nitrite reduction (Wrighton *et al.*, 2012; 2014; Rinke *et al.*, 2013; Danczak *et al.*, 2017). Additionally, a syntrophic relationship with other bacteria has been suggested as a putative benzene degrader in anaerobic environments (Phan *et al.*, 2021). However, some members of *Parcubacteria* harbor genes that are capable of using O_2 as a terminal electron acceptor (Nelson and Stegen, 2015). *Gracilibacteria* include three lineages and belong to *Patescibacteria*. Hanke *et al.* predicted that the terminal codon UGA encodes glycine in *Gracilibacteria* (Hanke *et al.*, 2014). These bacteria have poor metabolic potential (Sieber *et al.*, 2019), and some strains were reported to be parasitic on their hosts (Moreira *et al.*, 2021; Yakimov *et al.*, 2021).

Although many high-quality genomes related to *Patescibacteria* have been obtained from various environments, their detailed phylogeny, morphology, and ecophysiological role in activated sludge remain largely unknown. To clarify the phylogenetic and physiological diversities of *Patescibacteria* in activated sludge, obtaining high-quality genomes of *Patescibacteria* is necessary for further investigations in terms of visualization, *in situ* substrate utilization, and isolation. The purpose of the present study was to predict the metabolic potential of *Patescibacteria* in activated sludge and estimate their physiological role in activated sludge. A metagenomic approach using three activated sludge samples from a municipal wastewater treatment plant recovered 10 metagenome-assembled genomes (MAGs) related to *Saccharimonadia*, *Parcubacteria*, *and Gracilibacteria* within the superphylum *Patescibacteria*.

Materials and Methods

Sample collection

Four activated sludge samples were collected from aeration tanks in a wastewater treatment plant in Higashihiroshima city, which had previously been sampled (Kindaichi *et al.*, 2016; Table S1) in February 2019 (designated as AS201902), April 2020 (designated as AS202004), October 2020 (designated as AS202010R), and November 2020 (designated as AS202011). The collected sludge samples were immediately incubated to change the relative abundance of *Patescibacteria*. The AS202004 sample was anaerobically incubated for 3 d and was then designated as AA202004. The AS202010R sample was aerobically incubated for 3 d with washing and designated as AS202010A and without washing as AS202010B. In detail, 100 mL of the AS202004 sample was transferred into a 120-mL sterilized vial, which was sealed with a butyl rubber stopper. The gas phase was replaced with nitrogen gas, and the vial was then incubated anaerobically at 20°C for 3 d. One hundred milliliters of activated sludge from sample AS202010R was washed with Elix water (Merck) and then incubated at 20°C for 3 d. In the present study, the samples AS201902, AS202004, and AA202004 were used in a metagenomic analysis, while all seven samples were subjected to an amplicon analysis. Fresh and incubated sludge samples were stored at –18°C for further analyses.

Amplicon analysis of the 16S rRNA gene

DNA was extracted from activated sludge samples (0.5 g wet weight) (AS201902, AS202004, AA202004, AS202010A, AS202010B, AS202010, and AS202011) using a FastDNA SPIN kit for soil (MP Biomedicals). PCR amplification was performed using a primer set for the V3–V4 region of the 16S rRNA genes (341F and 805R). The primer sequences, detailed PCR conditions, and purification procedures used are as previously described (Dinh *et al.*, 2021). Purified DNA was sequenced using a MiSeq platform with paired-end sequencing (2×300 bp) and a MiSeq Reagent kit (v.3; Illumina). The obtained sequences were trimmed, merged, clustered, and analyzed using QIIME 2 core 2021.11, as previously described (Bolyen *et al.*, 2019; Awata *et al.*, 2021; Kambara *et al.*, 2022). The SILVA 138 database (Quast *et al.*, 2013) was used for the assignment. To elucidate the relationship between *Patescibacteria* and other co-existing bacteria, operational taxonomic units (OTUs) that showed a relative abundance of >0.1% were extracted, and Spearman's rank-order correlation coefficient was calculated for each OTU using Past 4.10 (Hammer *et al.*, 2001). OTUs that met the 5% significance level and correlated with *Patescibacteria* were investigated.

Metagenomic analysis

DNA was extracted from activated sludge samples (0.5 g wet weight) (AS201902, AS202004, and AA202004) using a FastDNA SPIN kit for soil (MP Biomedicals). Extracted DNA was purified using Agencourt AMPure XP magnetic beads (Beckman Coulter Life Sciences). Illumina sequencing libraries were prepared for the three samples using a TruSeq DNA PCR Free (350) kit (Illumina) and paired-end sequenced (2×151 bp) using shotgun sequencing on a HiSeq X system (Illumina). PacBio sequencing libraries were prepared for three samples using a 20 kb SMRTbell Express Template Prep kit (Pacific Biosciences of California) and sequenced on a PacBio Sequel II System (Pacific Biosciences of California). Circular consensus sequence (CCS) reads were generated from Sequel data with a Phred quality score above 20 (Q20, 99%).

A metagenomic analysis was conducted as previously described (Hosokawa *et al.*, 2021). Raw paired-end reads from HiSeq X were trimmed using Trimmomatic v.0.39 (Bolger *et al.*, 2014). The trimmed reads from HiSeq X and CCS reads from PacBio Sequel II were co-assembled using SPAdes v.3.13.1 (Bankevich *et al.*, 2012). BBtools v38.84 was used to obtain mapping information. Contigs from the assembly were binned using MetaBAT2.0 (Kang *et al.*, 2019). The relative abundance of the bins (multi-contigs classified into a taxonomic microorganism) were calculated based on information from the mapping file (*i.e.*, coverage) generated in MetaBAT2.0. The completeness and contamination of the bins were assessed using CheckM v1.1.2 (Parks *et al.*, 2015). The 43 marker genes proposed by Brown *et al.* (2015) likely provide improved estimates of CPR genome quality. Contamination in the obtained bins was manually removed. The bins with contamination removed were annotated using Prokka v1.13 (Seemann, 2014) and DRAM v1.2.2 (Shaffer *et al.*, 2020). Predicted amino acid

sequences were annotated using the KEGG BlastKEGG Orthology And Links Annotation (BlastKOALA) (Kanehisa *et al.*, 2016) and KEGG Automatic Annotation Server (KAAS) (Moriya *et al.*, 2007). BlastKOALA was used to visualize this pathway. A heatmap was created using KEGG-Decoder (Graham *et al.*, 2018) to visualize the percentage of gene possession related to each gene set. A phylogenetic tree of *Patescibacteria*, based on 400 marker protein sequences, was constructed using Phylophlan 3.0 (Asnicar *et al.*, 2020). The reference genome was selected from the genome registered in GenBank, and the complete genome was derived from activated sludge (Singleton *et al.*, 2021). Polyhydroxybutyrate (PHB) depolymerase-related sequences were aligned using mafft-linsi v7.480 (default parameters) (Katoh and Standley, 2013). Reference protein sequences were obtained from the top 500 hits for the identified PHB depolymerase-related protein (FNKGEGDK_00198) and known patescibacterial PHB depolymerase (OWK27304.1) using the NCBI-nr database. Protein sequences were clustered based on ≥70% similarity using CD-HIT version 4.8.1. (Fu *et al.*, 2012). A phylogenetic tree of PHB depolymerase-related proteins was constructed using iqtree2 version 2.1.2, with an automatically optimized substitution model of WAG+R10 (Minh *et al.*, 2020).

Nucleotide sequence accession number

The sequence data of the partial 16S rRNA gene sequence were deposited in the GenBank/EMBL/DDBJ databases under the accession number DRA013509. Metagenomic sequence data were deposited in the DDBJ database under the DDBJ/EMBL/GenBank accession number DRA013531.

Results and Discussion

Amplicon analysis of 16S rRNA genes

Amplicon sequencing of the 16S rRNA genes was performed to investigate the relative abundance of *Patescibacteria* in the seven activated sludge samples. On average, 34,573 reads and 549 OTUs were obtained from the seven samples (Table S3). In all activated sludge samples, except AS202004, *Patescibacteria* were predominant after *Proteobacteria* and *Bacteroidota*, with an average abundance of 12.1% (Fig. 1A). The most dominant group within *Patescibacteria* was *Saccharimonadia* in all samples, with the highest abundance of 13.7% in AS202010R. *Parcubacteria* and *Gracilibacteria* were the second and third most abundant groups, respectively (Fig. 1B). In addition to the above-mentioned groups, *Microgenomatia* (former candidate division OP11), ABY1, *Dojkabacteria* (former candidate division WS6), and *Berkelbacteria* were detected; however, their relative abundance was less than 0.2%. The ranges of the relative abundance of *Saccharimonadia*, *Parcubacteria*, and *Gracilibacteria* in untreated activated sludge samples (AS201902, AS202004, AS202010R, and AS202011) were 4.5–13.7%, 2.4–5.0%, and 0.4–1.8%, respectively. The relative abundance of *Patescibacteria* in all three treated samples (*i.e.*, aerobic and anaerobic incubations) decreased (Fig. S1). The relative abundance of *Saccharimonadia*, *Parcubacteria*,

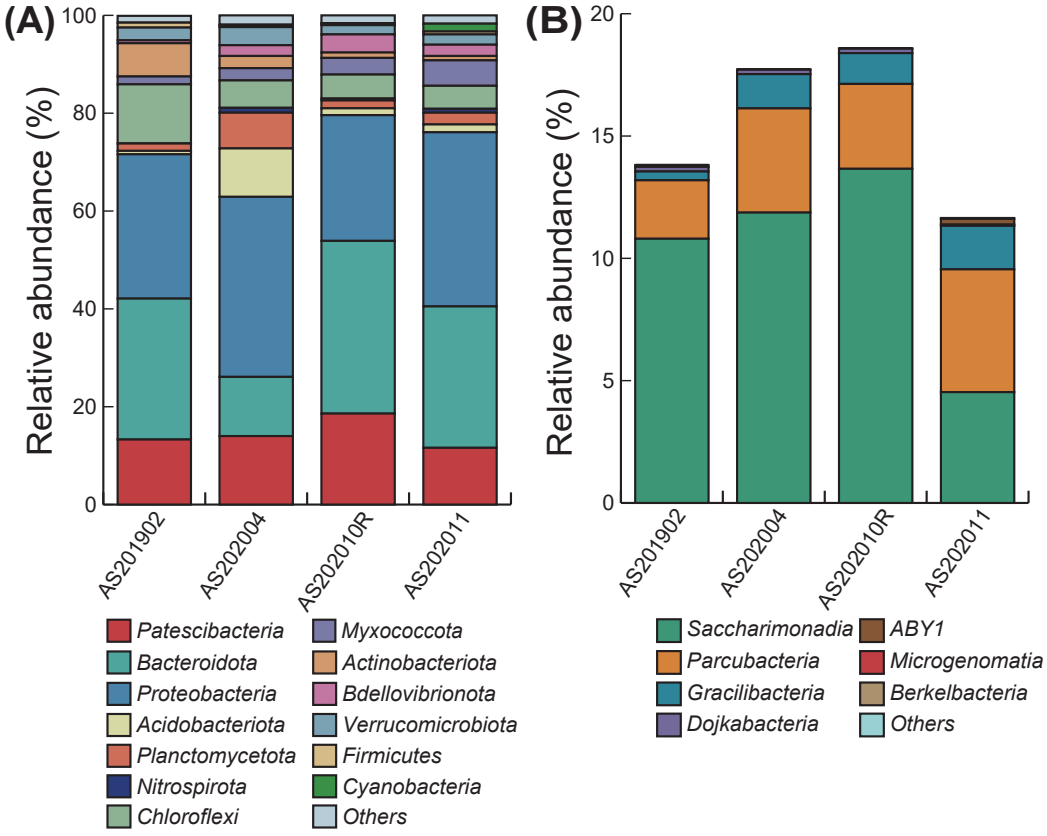


Fig. 1. Microbial community composition at the phylum level in four activated sludge samples examined in the present study (A), and the detailed composition of *Patescibacteria* in four activated sludge samples (B) based on 16S rRNA gene amplicon sequencing. The total relative abundance of each sample in panel (B) corresponds to the relative abundance of *Patescibacteria* (red) in each sample in panel (A).

and *Gracilibacteria* in AS202004 were 9.5, 3.5, and 1.1%, respectively, whereas those in AA202004 were 5.5, 2.8, and 0.6%, respectively. The relative abundance of *Saccharimonadia*, *Parcubacteria*, and *Gracilibacteria* in AS202010R, AS202010A, and AS202010B were 9.5, 3.5, and 1.1%, 5.5, 2.8, and 0.6%, and 5.5, 2.8, and 0.6%, respectively. This decrease may be associated with the oxygen level or abundance of coexisting bacteria. The relative abundance of *Patescibacteria* in groundwater samples ranged between 2.1 and 20.7%; however, it was not possible to compare these samples directly because they were enriched using a filter-based sampling method (Danczak *et al.*, 2017).

Correlation analysis

Pearson's correlation coefficients were calculated between the patescibacterial OTUs obtained from amplicon sequencing and other bacterial OTUs with more that 0.1% relative abundance. The OTUs of *Proteobacteria*, *Chloroflexi*, and *Planctomycetota* showed a positive correlation with the OTUs of *Patescibacteria*; however, in some cases, correlations were negative (data not shown). In contrast, the correlation between *Patescibacteria* and *Chitinophagales* belonging to the phylum *Bacteroidota* was positive (Fig. S2). Among the OTUs shown in Fig. S2, we extracted OTUs with a sequence that matched the reconstructed bin with 100% sequence identity (Fig. 2). Most *Chitinophagales* OTUs correlated with several patescibacterial OTUs (HS1, HS2, and HP2). In addition, two *Chitinophagales* OTUs correlated with the three lineages of *Patescibacteria*. A similar positive correlation between *Saccharimonadia* and *Chitinophagaceae* was found in acid mine drainage samples (Lemos *et al.*, 2019). Metabolic interactions between *Patescibacteria* and *Chitinophagales* are discussed in the following section.

Genome reconstruction and basic information on bins

In total, 0.77 billion reads and 0.06 million reads were obtained from HiSeq X and PacBio CCS sequencing of the three activated sludge samples, respectively (Table S2). The hybrid assembly using HiSeq X and PacBio CCS reads generated 12,097 contigs with an N50 value of 148,787 bp. A total of 8,211 contigs >1,500 bp were extracted and classified into 320 bins. Ten patescibacterial bins were reconstructed, which consisted of *Saccharimonadia* (five bins, HHAS1–HHAS5), *Parcubacteria* (four bins, HHAS6–HHAS9), and *Gracilibacteria* (one bin, HHAS10) (Table 1). The completeness of *Sacchrimonadia* and *Parcubacteria* ranged between 88.4 and 97.7 and between 62.8 and 90.7%, respectively, while that of *Gracilibacteria* was 97.7%.

A phylogenetic tree of the ten bins based on the protein sequence is shown in Fig. 3. Bins belonging to *Saccharimonadia* were classified into three groups. The group including HHAS3 and HHAS4 was related to the well-described saccharimonadial species *Candidatus* Saccharimonas aalborgensis (CP005957), reconstructed from a Danish activated sludge sample (Albertsen *et al.*, 2013). Since this species shows a small coccus morphology (Albertsen *et al.*, 2013), HHAS3 and HHAS4 were also considered to be small cocci. The HHAS1 and HHAS5 groups were related to the genomes of activated sludge samples. According to sequence similarities based on 16S rRNA genes, the morphology of this group was primarily filamentous (Kindaichi *et al.*, 2016). The morphology of filamentous *Saccharimonadia* needs to be confirmed using FISH in the future. The HHAS2 bin formed a different clade from other saccharimonadial genomes with a genome from activated sludge (Singleton *et al.*, 2021). However, the details of this group remain largely unknown. The parcubacterial bins were classified into three groups. The HHAS7 bin was classified as *Nomurabacteria* and was related to genomes from groundwater samples (Brown *et al.*, 2015). Groups HHAS6 and HHAS9 belonged to *Moranbacteria*. In addition, clades HHAS6 and HHAS9 consisted of genomes from activated sludge samples (Singleton *et al.*, 2021). The details of *Moranbacteria* in activated sludge samples are also unclear because the majority of information on *Moranbacteria* was obtained from groundwater samples (Anantharaman *et al.*, 2016). The HHAS8 bin did not belong to any parcubacterial subgroup. HHAS10 was classified as belonging to *Gracilibacteria*. Although some of the gracilibacterial genomes were also reconstructed from acti-

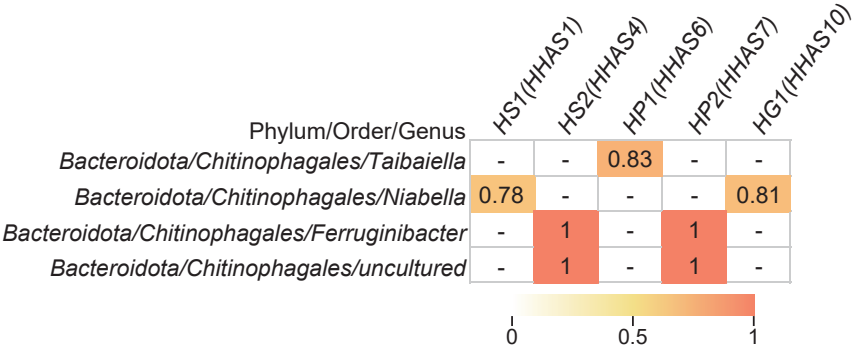


Fig. 2. Spearman's correlation between *Patescibacteria* and *Chitinophagales*. Correlation coefficients that met the 5% significance level are shown, and hyphens indicate that correlation coefficients did not meet the significance level. HS1 and HS2, HP1 and HP2, and HG1 indicate the OTUs of *Saccharimonadia*, *Parcubacteria*, and *Gracilibacteria*, respectively. The only OTUs belonging to *Patescibacteria* and *Bacteroidota* that matched with the reconstructed bins in the metagenomic analysis with 100% sequence identity are shown. Correlations for all OTUs are shown in Fig. S1. Correlation coefficients that met the 5% significance level are presented as a heatmap. The parentheses indicate the bin ID shown in Table 1 with 100% sequence identity to the OTU.

Table 1. Characteristics of patescibacterial bins obtained in the present study

Bin ID	Taxonomy	Bin size (Mbp)	Completeness (%)	Contamination (%)	Number of contigs	Number of CDSs	Relative abundance (%)†		
							AS201902	AS202004	AA202004
HHAS1	*Saccharimonadia*	0.91	88.37*	0*	3	946	1.79	0.04	0.03
HHAS2	*Saccharimonadia*	0.83	97.67*	0*	3	847	0.20	0	0
HHAS3	*Saccharimonadia*	1.00	93.02*	0*	2	1027	1.47	0	0
HHAS4	*Saccharimonadia*	0.73	90.70*	0*	3	759	2.26	0	0
HHAS5	*Saccharimonadia*	0.71	88.37*	0*	6	746	0.29	0	0
HHAS6	*Parcubacteria*	0.53	90.70*	0*	8	536	0.15	0	0
HHAS7	*Parcubacteria*	0.60	62.79*	0*	2	627	0.48	0	0
HHAS8	*Parcubacteria*	0.60	90.70*	0*	5	619	0.02	0.26	0.19
HHAS9	*Parcubacteria*	0.96	79.07*	0*	6	942	0.29	0.27	0.44
HHAS10	*Gracilibacteria*	1.30	97.67*	0*	3	1175	0.02	0.22	0.10
HHAS11	*Chitinophagales*	2.37	75.24	0	17	2051	0	0.17	0.23
HHAS12	*Chitinophagales*	2.79	75.2	3.96	21	2435	0.97	0	0.02
HHAS13	*Chitinophagales*	3.13	94.77	3.45	5	2669	1.85	0	0
HHAS14	*Chitinophagales*	2.99	91.21	2.72	14	2354	0.78	0	0

* Calculated using the CPR marker set.
† Calculated based on the mapping file generated in MetaBAT2.0

vated sludge samples (Singleton *et al.*, 2021), the HHAS10 bin formed a clade that included genomes from human oral samples (Dudek *et al.*, 2017).

Metabolic analysis

The predicted metabolic potential of *Gracilibacteria*, *Parcubacteria*, and *Saccharimonadia*, and the putative metabolic interactions between *Patescibacteria* and *Chitinophagales* based on the metagenomic analysis in this study are shown in Fig. 4. Patescibacterial bins revealed that *Patescibacteria* did not possess *de novo* nucleotide synthesis, amino acid synthesis, phospholipid synthesis, or a full TCA cycle. In addition, *Patescibacteria* possessed ABC transporters with unknown functions, the peptidoglycan biosynthesis pathway, and type IV pili (Fig. S3). The lack of *de novo* amino acid synthesis suggests that peptidases acquire amino acids. The presence of peptidases was also confirmed (Table S4). Several patescibacterial bins converted glycine to serine and harbored serine peptidases. These common features are consistent with the genomes of activated sludge samples as well as other natural samples (Wrighton *et al.*, 2012, 2014; Albertsen *et al.*, 2013; Danczak *et al.*, 2017; Starr *et al.*, 2018; Lemos *et al.*, 2019; 2020; Sieber *et al.*, 2019; Chaudhari *et al.*, 2021; Moreira *et al.*, 2021; Yakimov *et al.*, 2021). The incomplete nucleotide synthesis pathway and the presence of the *com*E gene and type IV pili support the acquisition of DNA from outside cells (Chen and Gotschlich, 2001; Starr *et al.*, 2018).

Saccharimonadial bins possessed glycolysis and the pentose phosphate pathway, with possession patterns depending on the subgroup (Fig. 4). The members of subdivision 1 (HHAS1 and HHAS5), which are putative filamentous *Saccharimonadia*, partially possessed glycolysis, whereas members of subdivision 3 (HHAS3 and HHAS4) possessed both glycolysis and the pentose phosphate pathway. The possession of genes to convert pyruvate to lactate, acetate, and malate and the lack of a TCA cycle supports fermentative metabolism. These results are consistent with previous findings (Wrighton *et al.*, 2012, 2014; Albertsen *et al.*, 2013; Danczak *et al.*, 2017; Starr *et al.*, 2018; Lemos *et al.*, 2019; 2020; Sieber *et al.*, 2019; Moreira *et al.*, 2021; Yakimov *et al.*, 2021). The fermentative pathway from pyruvate to lactate or malate may facilitate the production of NAD^+ (Starr *et al.*, 2018; Lemos *et al.*, 2019). The pentose phosphate pathway in subdivision 3 may be involved in the conversion of glucose-6P to glyceradehyde-3P and in energy conversion (NADPH to NADP production) (Albertsen *et al.*, 2013). However, genes involved in the synthesis of nucleic acids, such as ribose-phosphate pyrophosphokinase, were absent. Therefore, they are not expected to contribute to anabolism (the production of deoxyribonucleotides) (Castelle *et al.*, 2018). All reconstructed saccharimonadial bins in the present study possessed the NADH dehydrogenase-like protein and complete cytochrome o ubiquinol oxidase, which is related to the oxygen scavenging system, despite the absence of the TCA cycle (Kantor *et al.*, 2013; Starr *et al.*, 2018; Lemos *et al.*, 2019). In addition, Lemos *et al.* (2019) suggested that *Saccharimonadia* follow non-obligatory fermentative metabolism with occasional aerobic respiration. As previously reported by Lemos *et al.* (2019), *Saccharimonadia* have membrane-bound NADH dehydrogenase to supply NAD^+ and pass the electron to ubiquinone, which transfers it to cytochrome O ubiquinol oxidase. Cytochrome then reduces O_2 to H_2O as the final receptor, delivering protons through the plasma membrane to generate the proton electromotive force used for ATP synthesis by ATP synthase. Filamentous *Saccharimonadia* in activated sludge took up N-acetylglucosamine under aerobic conditions, as demonstrated by microautoradiography combined with FISH (Kindaichi *et al.*, 2016). *Chitinophagales* bins (HHAS12, HHAS13, and HHAS14) possessed chitinase (MHHEHLFG_00073, MHHEHLFG_01825, AECFEMFL_01184, and KHEBLPDM_00108) and all *Chitinophagales* bins harbored beta-acetylhexosamidase (MHHEHLFG_00810, CLEEFKKN_01010, AECFEMFL_02009, and KHEBLPDM_01401). *Chitinophagales* have the potential to convert chitin to N-acetylglucosamine via chitobiose (Fig. 4). All *Chitinophagales* bins encoded poly-beta-1,6 N-acetyl-d-glucosamine synthase (PgaC). This enzyme

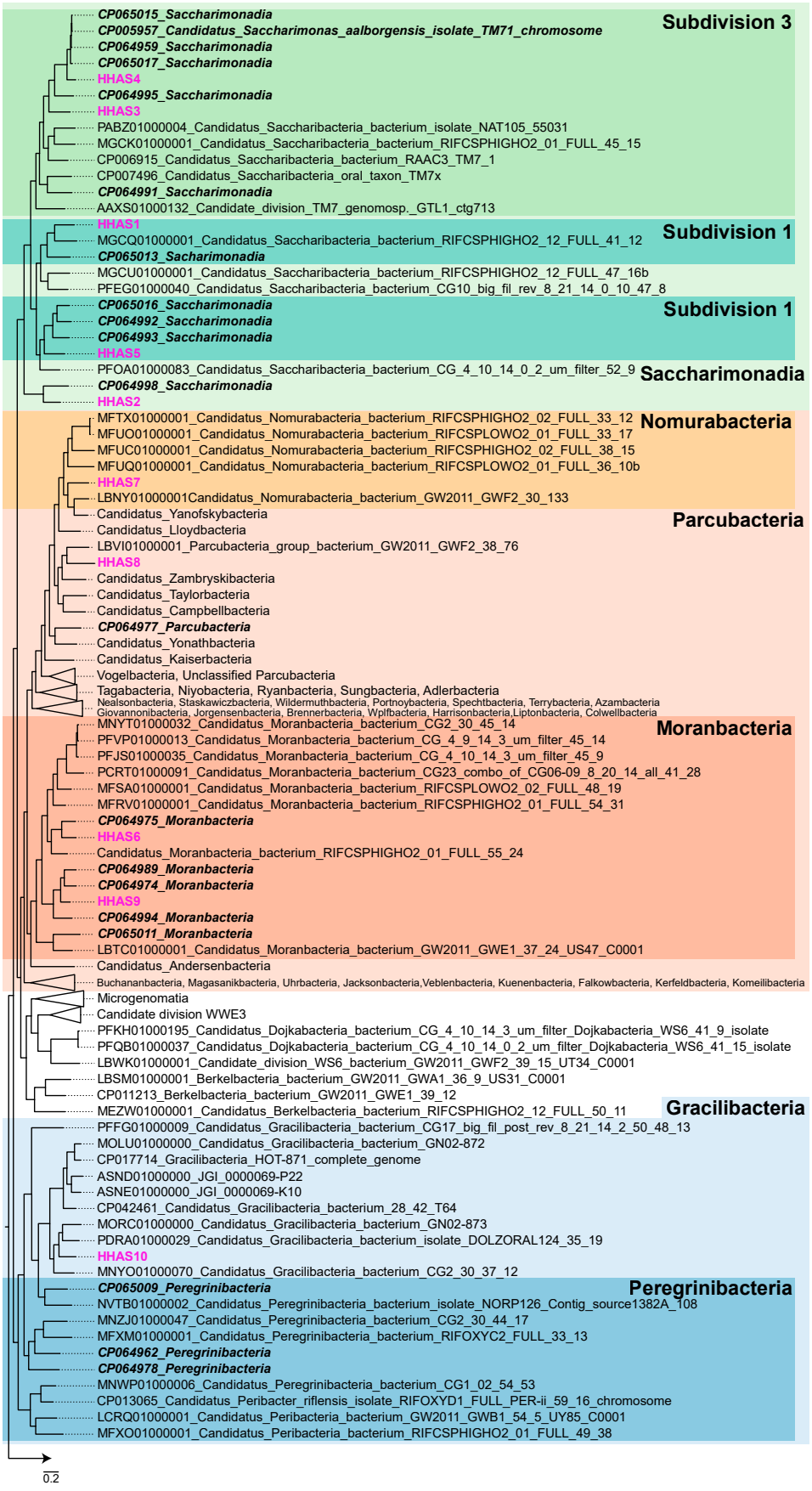


Fig. 3. Genome-based phylogenetic tree of reconstructed patescibacterial bins in activated sludge and related genomes. The bins found in the present study are shown in pink. Phylum-level designations (dotted line) are highlighted on the right. Bold italics indicate genomes obtained from activated sludge samples.

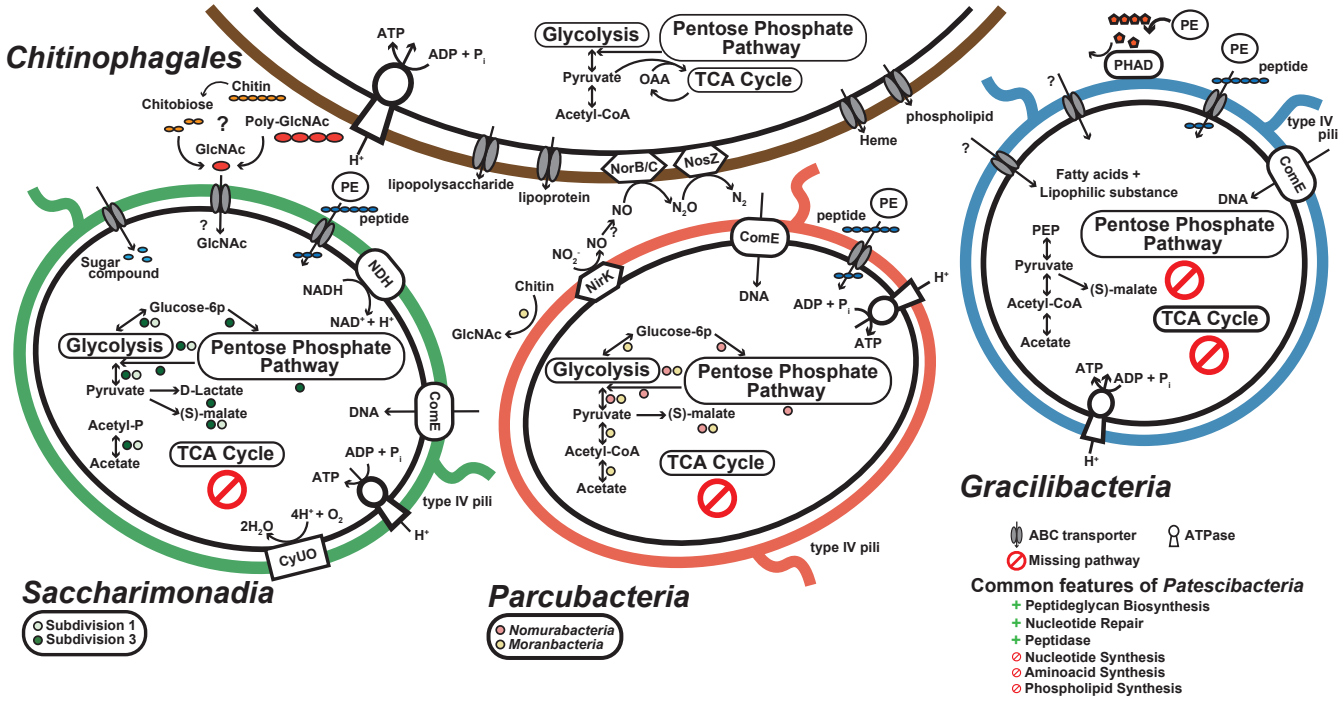


Fig. 4. Predicted metabolic potential of *Gracilibacteria*, *Parcubacteria*, and *Saccharimonadia*, and putative metabolic interactions between *Patescibacteria* and *Chitinophagales* based on the genetic information obtained from the metagenomic analysis. The colored circles below *Saccharimonadia* (light green and green) and *Parcubacteria* (pink and yellow) indicate the presence of genes only found in the subdivision/subgroup members. Semicircles indicate that one of the two possessed the genes. The common features of *Patescibacteria* are also shown. Abbreviations: CyUO, cytochrome O ubiquinol oxidase; NDH, NADH dehydrogenase; PHAD, polyhydroxyalkanoate depolymerase; ComEC, competence protein ComEC; PE, peptidases; OAA, oxaloacetate; GlcNAc, N-acetylglucosamine; PEP, phosphoenolpyruvate.

catalyzes the polymerization of uridine diphosphate-N-acetylglucosamine to produce poly-N-acetylglucosamine (PGA). There were other bins in active sludge belonging to *Ignavibacteria*, *Acidobacteria*, *Actinobacteria*, *Bacteroidota*, *Chloroflexi*, *Nitrospira*, *Proteobacteria*, and *Verrucomicrobia*. These bacteria possessed PgaC and were present in approximately 38–41% of samples from AS201902, AS202004, and AA202004. Filamentous *Saccharimonadia* in activated sludge took up N-acetylglucosamine, which strongly supports the metabolic interaction between *Chitinophagales* and *Saccharimonadia* via N-acetylglucosamine in activated sludge. In addition, *Saccharimonadia* have been suggested to use some of the PGA produced by bacteria (Hosokawa *et al.*, 2021). The mechanisms by which N-acetylglucosamine is assimilated or catabolized by *Saccharimonadia* currently remain unclear. Investigations on the metabolism of incorporated N-acetylglucosamine are highly challenging, but are warranted.

Parcubacterial bins possessed glycolysis and/or the pentose phosphate pathway. Parcubacterial bins also encoded genes involved in the conversion of pyruvate to acetate and malate (Fig. 4). The possession of these pathways and the lack of a TCA cycle are similar features to those of saccharimonadial bins and support fermentative metabolism, as reported in previous studies (Wrighton *et al.*, 2012; 2014; Albertsen *et al.*, 2013; Danczak *et al.*, 2017; Lemos *et al.*, 2019; 2020; Starr *et al.*, 2018; Sieber *et al.*, 2019; Moreira *et al.*, 2021; Yakimov *et al.*, 2021). A moranbacterial bin (HHAS9) possessed chitinase (DEAMOMGP_00832 and DEAMOMGP_00955) (Fig. 4), but not the genes to convert N-acetylglucosamine to other compounds. The nomurabacterial bin (HHAS7) possessed the copper-containing nitrite reductase gene (JCNBLHJH_00228, *nirK*) (Fig. 4). Some members of *Parcubacteria* are known to be involved in nitrite reduction (Castelle *et al.*, 2017; Danczak *et al.*, 2017; He *et al.*, 2021). In addition, several *Chitinophagales* bins (HHAS12 and HHAS13) possessed the nitric oxide reductase subunit B/C (*norB/C*) (CLEEFKKN_01663, CLEEFKKN_01664, AECFEMFL_01795, and AECFEMFL_01795) and nitrous-oxide reductase (*nosZ*) (CLEEFKKN_01654 and AECFEMFL_02605). Therefore, *Nomurabacteria* were partially responsible for denitrification along with *Chitinophagaceae* in the activated sludge process.

The gracilibacterial bin (HHAS10) had negligible central carbon metabolism. It possessed only pyruvate kinase (FNKGEGDK_00842), malate dehydrogenase (FNKGEGDK_00357), and 2-oxoglutarate/2-oxoacid ferredoxin oxidoreductase (FNKGEGDK_00164 and FNKGEGDK_00166). Although the poor metabolic potential of *Gracilibacteria* has also been reported (Sieber *et al.*, 2019), the gracilibacterial genomes in previous studies were mainly reconstructed from other habitats, such as oral and ground water samples. The genome size of the HHAS10 bin was 1.3 Mbp, which is similar to that of other *Gracilibacteria* (Sieber *et al.*, 2019), and completeness was relatively high (Table 1). Nevertheless, it was not possible to predict the metabolic potential of *Gracilibacteria* reconstructed in the present study using the current databases. The accumulation

Table 2. Summary of genes related to polyhydroxyalkanoate and polyhydroxybutyrate depolymerases in the *Gracilibacteria* bin HHAS10

Locus tags	Amino acid identity to known proteins of representative taxonomies based on the NCBI-nr database			
	Description	Accession no.	Identities	e-value
FNKGEGDK_00792	polyhydroxyalkanoate depolymerase [*Alteromonas oceani*]	WP_123327345.1	138/317 (44%)	6e-68
	polyhydroxyalkanoate depolymerase [*Polynucleobacter paneuropaeus*]	WP_215313699.1	144/340 (42%)	7e-68
	polyhydroxyalkanoate depolymerase [*Alteromonas lipolytica*]	WP_070177576.1	137/317 (43%)	2e-67
	polyhydroxyalkanoate depolymerase [*Polynucleobacter wuianus*]	WP_216235107.1	144/340 (42%)	2e-67
	polyhydroxyalkanoate depolymerase [*Marisediminitalea aggregata*]	WP_073324272.1	137/317 (43%)	4e-66
FNKGEGDK_00198	hypothetical protein US76_00085 [*Parcubacteria* GW2011_GWA2_38_13b]	OWK27304.1	41/145 (28%)	3e-15
	S-layer homology domain-containing protein [*Candidatus* Gracilibacteria bacterium]	MBP9812078.1	305/443 (69%)	0
	S-layer homology domain-containing protein [*Candidatus* Gracilibacteria bacterium]	MBC7498061.1	267/427 (63%)	0
	Ricin and poly(3-hydroxybutyrate) depolymerase fusion [*Myxococcales* bacterium]	MCA9656682.1	64/240 (27%)	9e-22
	Ricin and poly(3-hydroxybutyrate) depolymerase fusion [*Sorangium cellulosum*]	KYF79897.1	65/225 (29%)	4e-19
	poly(3-hydroxybutyrate) depolymerase [*Streptomyces* sp. TLI_55]	SNX55894.1	87/292 (30%)	4e-17

of genomic information on *Gracilibacteria* in activated sludge is necessary to construct substantial databases. In contrast, the gracilibacterial bin (HHAS10) possessed four copies of peptidase belonging to the M23 family (Table S4), which lyses the cell walls of other microorganisms. The HHAS10 bin also possessed a phospholipase gene (FNKGEGDK_00603, FNKGEGDK_00841), whereas *Chitinophagales* possessed an ABC transporter (MHHEHLFG_00269, MHHEHLFG_00294, MHHEHLFG_00428, MHHEHLFG_01772, MHHEHLFG_01773, MHHEHLFG_01879, KHEBLPDM_00308, KHEBLPDM_00309, and KHEBLPDM_02272), which releases phospholipids (Fig. 4). Therefore, the metabolic flow of phospholipids between *Chitinophagales* and *Gracilibacteria* in activated sludge was considerable. The HHAS10 bin possessed a homolog of polyhydroxyalkanoate (PHA) depolymerase (FNKGEGDK_00792), which showed >40% homology (<1e-65) to PHA depolymerases of known species (Table 2). This feature may help to obtain an energy source, even though *Gracilibacteria* have negligible central carbon metabolism. In general, PHA is degraded by PHA depolymerase to monomers, such as 3HB, which are then oxidized to acetoacetyl-CoA in a reaction catalyzed by 3HB dehydrogenase. This is then converted to acetyl-CoA by β-ketothiolase (Ong *et al.*, 2017). Although *Candidatus* Parcunitrobacter nitroensis belonging to *Parcubacteria* also possessed PHB, which is a PHA, depolymerase, and peptidase that acts extracellularly and converts PHB to acetate, suggesting that PHB may be used as a carbon source (Castelle *et al.*, 2017), no genes related to the reaction pathway of hydroxybutyrate in the HHAS10 bin were identified. Based on an amino acid sequence homology search using the NCBI-nr database, we found that the genome of HHAS10 bin had a surface layer protein (FNKGEGDK_00198) that was widely conserved in gracilibacterial genomes with high similarity (Table 2 and Fig. S4). Besides, the proteins showed 28% (41/145 bp, 3e-15) and 27–30% (<1e-16) homology with the PHB depolymerases of *Candidatus* Parcunitrobacter nitroensis (OWK27304.1) and other taxa (*Myxococcales*, *Sorangium cellulosum*, and *Streptomyces* sp.), respectively. Further studies are needed on the generality and roles of PHA/PHB depolymerases in *Gracilibacteria*.

Conclusions

In the present study, the metabolic potential of *Patescibacteria* was predicted from the MAGs of activated sludge samples, and the physiological role of *Patescibacteria* in activated sludge was estimated. The genomes of three *Saccharimonadia*, three *Parcubacteria*, and one *Gracilibacteria* species revealed a lack of *de novo* nucleotide synthesis, amino acid synthesis, phospholipid synthesis, and a full TCA cycle. Ten reconstructed genomes showed a strong positive correlation of relative abundance with *Chitinophagales* based on 16S rRNA genes. Metabolic interactions between a member of *Saccharimonadia* and *Chitinophagales* via N-acetylglucosamine, between a member of *Parcubacteria* and *Chitinophagales* via nitrogen compounds related to denitrification, and between *Gracilibacteria* and *Chitinophagales* via phospholipids in activated sludge were supported by metabolic predictions from 10 recovered *Patescibacteria* MAGs and five *Chitinophagales* MAGs. The high abundance of peptidases in *Gracilibacteria* suggests their role in cell lysis in activated sludge. Further studies related to visualization with FISH and the enrichment of *Patescibacteria* are necessary to elucidate the *in situ* physiological roles of *Patescibacteria* in the activated sludge process.

Acknowledgements

This work was supported by JSPS KAKENHI, Grant Numbers JP16H04833, and JP20H02287.

References

Albertsen, M., Hugenholtz, P., Skarshewski, A., Nielsen, K.L., Tyson, G.W., and Nielsen, P.H. (2013) Genome sequences of rare, uncultured bacteria obtained by differential coverage binning of multiple metagenomes. *Nat Biotechnol* **31**: 533–538.

Anantharaman, K., Brown, C.T., Hug, L.A., Sharon, I., Castelle, C.J., Probst, A.J., *et al.* (2016) Thousands of microbial genomes shed light on interconnected biogeochemical processes in an aquifer system. *Nat Commun* **7**: 13219.

Asnicar, F., Thomas, A.M., Beghini, F., Mengoni, C., Manara, S., Manghi, P., *et al.* (2020) Precise phylogenetic analysis of microbial isolates and genomes from metagenomes using PhyloPhlAn 3.0. *Nat Commun* **11**: 2500.

Awata, T., Goto, Y., Kuratsuka, H., Aoi, Y., Ozaki, N., Ohashi, A., *et al.* (2021) Reactor performance and microbial community structure of single-stage partial nitritation anammox membrane bioreactors inoculated with *Brocadia* and *Scalindua* enrichment cultures. *Biochem Eng J* **170**: 107991.

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A.A., Dvorkin, M., Kulikov, A.S., *et al.* (2012) SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* **19**: 455–477.

Bolger, A.M., Lohse, M., and Usadel, B. (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**: 2114–2120.

Bolyen, E., Rideout, J.R., Dillon, M.R., Bokulich, N.A., Abnet, C.C., Al-Ghalith, G.A., *et al.* (2019) Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat Biotechnol* **37**: 852–857.

Brown, C.T., Hug, L.A., Thomas, B.C., Sharon, I., Castelle, C.J., Singh, A., *et al.* (2015) Unusual biology across a group comprising more than 15% of domain Bacteria. *Nature* **523**: 208–211.

Castelle, C.J., Brown, C.T., Thomas, B.C., Williams, K.H., and Banfield, J.F. (2017) Unusual respiratory capacity and nitrogen metabolism in a Parcubacterium (OD1) of the Candidate Phyla Radiation. *Sci Rep* **7**: 40101.

Castelle, C.J., Brown, C.T., Anantharaman, K., Probst, A.J., Huang, R.H., and Banfield, J.F. (2018) Biosynthetic capacity, metabolic variety and unusual biology in the CPR and DPANN radiations. *Nat Rev Microbiol* **16**: 629–645.

Chaudhari, N.M., Overholt, W.A., Figueroa-Gonzalez, P.A., Taubert, M., Bornemann, T.L.V., Probst, A.J., *et al.* (2021) The economical lifestyle of CPR bacteria in groundwater allows little preference for environmental drivers. *Environ Microbiome* **16**: 24.

Chen, I., and Gotschlich, E.C. (2001) ComE, a competence protein from *Neisseria gonorrhoeae* with DNA-Binding Activity. *J Bacteriol* **183**: 3160–3168.

Danczak, R.E., Johnston, M.D., Kenah, C., Slattery, M., Wrighton, K.C., and Wilkins, M.J. (2017) Members of the candidate phyla radiation are functionally differentiated by carbon and nitrogen-cycling capabilities. *Microbiome* **5**: 112.

Dinh, H.T.T., Kambara, H., Harada, Y., Matsushita, S., Aoi, Y., Kindaichi, T., *et al.* (2021) Bioelectrical methane production with an ammonium oxidative reaction under the no organic substance condition. *Microbes Environ* **36**: ME21007.

Dudek, N.K., Sun, C.L., Burstein, D., Kantor, R.S., Goltsman, D.S.A., Bik, E.M., *et al.* (2017) Novel microbial diversity and functional potential in the marine mammal oral microbiome. *Curr Biol* **27**: 3752–3762.

Fu, L., Niu, B., Zhu, Z., Wu, S., and Li, W. (2012) CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics* **28**: 3150–3152.

Graham, E.D., Heidelberg, J.F., and Tully, B.J. (2018) Potential for primary productivity in a globally-distributed bacterial phototroph. *ISME J* **12**: 1861–1866.

Hammer, Ø., Harper, D.A.T. and Ryan, P.D. (2001) PAST: Paleontological statistics software package for education and data analysis. *Palaeontol Electronica* **4**: 9pp.

Hanke, A., Hamann, E., Sharma, R., Geelhoed, J.S., Hargesheimer, T., Kraft, B., *et al.* (2014) Recoding of the stop codon UGA to glycine by a BD1-5/SN-2 bacterium and niche partitioning between Alpha- and Gammaproteobacteria in a tidal sediment microbial community naturally selected in a laboratory chemostat. *Front Microbiol* **5**: 231.

He, C., Keren, R., Whittaker, M.L., Farag, I.F., Doudna, J.A., Cate, J.H.D., *et al.* (2021) Genome-resolved metagenomics reveals site-specific diversity of episymbiotic CPR bacteria and DPANN archaea in groundwater ecosystems. *Nat Microbiol* **6**: 354–365.

He, X.S., McLean, J.S., Edlund, A., Yooseph, S., Hall, A.P., and Liu, S.Y. (2015) Cultivation of a human-associated TM7 phylotype reveals a reduced genome and epibiotic parasitic lifestyle. *Proc Natl Acad Sci U S A* **112**: 244–249.

Hosokawa, S., Kuroda, K., Narihiro, T., Aoi, Y., Ozaki, N., Ohashi, A., *et al.* (2021) Cometabolism of the superphylum Patescibacteria with anammox bacteria in a long-term freshwater anammox column reactor. *Water (Basel, Switz)* **13**: 208.

Hug, L.A., Baker, B.J., Anantharaman, K., Brown, C.T., Probst, A.J., Castelle, C.J., *et al.* (2016) A new view of the tree of life. *Nat Microbiol* **1**: 16048.

Hugenholtz, P., Tyson, G.W., Webb, R.I., Wagner, A.M., and Blackall, L.L. (2001) Investigation of candidate division TM7, a recently recognized major lineage of the domain bacteria with no known pure-culture representatives. *Appl Environ Microbiol* **67**: 411–419.

Ibrahim, A., Maatouk, M., Rajaonison, A., Zgheib, R., Haddad, G., Khalil, J.B., *et al.* (2021) Adapted protocol for Saccharibacteria cocultivation: two new members join the club of candidate phyla radiation. *Microbiol Spectrum* **9**: e01069-21.

Kambara, H., Shinno, T., Matsuura, N., Matsushita, S., Aoi, Y., Kindaichi, T., *et al.* (2022) Environmental factors affecting the community of methane-oxidizing bacteria. *Microbes Environ* **37**: ME21074.

Kanehisa, M., Sato, Y., and Morishima, K. (2016) BlastKOALA and GhostKOALA: KEGG tools for functional characterization of genome and metagenome sequences. *J Mol Biol* **428**: 726–731.

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

Kantor, R.S., Wrighton, K.C., Handley, K.M., Sharon, I., Hug, L.A., Castelle, C.J., *et al.* (2013) Small genomes and sparse metabolisms of sediment-associated bacteria from four candidate phyla. *mBio* **4**: e00708-13.

Katoh, K., and Standley, D.M. (2013) MAFFT Multiple Sequence Alignment Software Version 7: improvements in performance and usability. *Mol Biol Evol* **30**: 772–780.

Kindaichi, T., Nierychlo, M., Kragelund., C., Nielsen, J.L., and Nielsen, P.H. (2013) High and stable substrate specificities of microorganisms in enhanced biological phosphorus removal plants. *Environ Microbiol* **15**: 1821–1831.

Kindaichi, T., Yamaoka, S., Uehara, R., Ozaki, N., Ohashi, A., Albertsen, M., *et al.* (2016) Phylogenetic diversity and ecophysiology of Candidate phylum Saccharibacteria in activated sludge. *FEMS Microbiol Ecol* **92**: fiw078.

Lemos, L.N., Medeiros, J.D., Dini-Andreote, F., Fernandes, G.R., Varani, A.M., Oliveira, G., *et al.* (2019) Genomic signatures and co-occurrence patterns of the ultra-small Saccharimonadia (phylum CPR/Patescibacteria) suggest a symbiotic lifestyle. *Mol Ecol* **28**: 4259–4271.

Lemos, L.N., Manoharan, L., Mendes, L.W., Venturini, A.M., Pylro, V.S., and Tsai, S.M. (2020) Metagenome assembled-genomes reveal similar functional profiles of CPR/Patescibacteria phyla in soils. *Environ Microbiol Rep* **12**: 651–655.

Mielczarek, A.T., Kragelund, C., Eriksen, P.S., and Nielsen, P.H. (2012) Population dynamics of filamentous bacteria in Danish wastewater treatment plants with nutrient removal. *Water Res* **46**: 3781–3795.

Minh, B.Q., Schmidt, H.A., Chernomor, O., Schrempf, D., Woodhams, M.D., Von H.A., *et al.* (2020) IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Mol Biol Evol* **37**: 1530–1534.

Moreira, D., Zivanovic, Y., López-Archilla, A.I., Iniesto, M., and López-Garcia, P. (2021) Reductive evolution and unique predatory mode in the CPR bacterium *Vampirococcus lugosii*. *Nat Commun* **12**: 2454.

Moriya, Y., Itoh, M., Okuda, S., Yoshizawa, A.C., and Kanehisa, M. (2007) KAAS: an automatic genome annotation and pathway reconstruction server. *Nucleic Acids Res* **35**: 182–185.

Nakai, R. (2020) Size matters: Ultra-small and filterable microorganisms in the environment. *Microbes Environ* **35**: ME20025.

Nelson, W.C., and Stegen, J.C. (2015) The reduced genomes of Parcubacteria (OD1) contain signatures of a symbiotic lifestyle. *Front Microbiol* **6**: 713.

Nielsen, P.H., Mielczarek, A.T., Kragelund, C., Nielsen, J.L., Saunders, A.M., and Kong, Y. (2010) A conceptual ecosystem model of microbial communities in enhanced biological phosphorus removal plants. *Water Res* **44**: 5070–5088.

Ong, S.Y., Chee, J.Y., and Sudesh, K. (2017) Degradation of polyhydroxyalkanoate (PHA): a review. *J Sib Fed Univ Biol* **10**: 211–225.

Parks, D.H., Imelfort, M., Skennerton, C.T., Hugenholtz, P., and Tyson, G.W. (2015) CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res* **25**: 1043–1055.

Phan, H.V., Kurisu, F., Kiba, K., and Furumai, H. (2021) Optimized cultivation and syntrophic relationship of anaerobic benzene-degrading enrichment cultures under methanogenic conditions. *Microbes Environ* **36**: ME21028.

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., *et al.* (2013) The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res* **41**: D590–D596.

Rinke, C., Schwientek, P., Sczyrba, A., Ivanova, N.N., Anderson, I.J., Cheng, J.F., *et al.* (2013) Insights into the phylogeny and coding potential of microbial dark matter. *Nature* **499**: 431–437.

Seemann, T. (2014) Prokka: rapid prokaryotic genome annotation. *Bioinformatics* **30**: 2068–2069.

Shaffer, M., Borton, M.A., McGivern, B.B., Zayed, A.A. La Rosa, S.L., Solden, L.M., *et al.* (2020) DRAM for distilling microbial metabolism to automate the curation of microbiome function. *Nucleic Acids Res* **48**: 8883–8900.

Sieber, C.M.K., Paul, B.G., Castelle, C.J., Hu, P., Tringe, S.G., Valentine, D.L., *et al.* (2019) Unusual metabolism and hypervariation in the genome of a Gracilibacterium (BD1-5) from an oil-degrading community. *mBio* **10**: e02128-19.

Singleton, C.M., Petriglieri, F., Kristensen, J.M., Kirkegaard, R.H., Michaelsen, T.Y., Andersen, M.H., *et al.* (2021) Connecting structure to function with the recovery of over 1000 high-quality metagenome-assembled genomes from activated sludge using long-read sequencing. *Nat Commun* **12**: 2009.

Soro, V., Dutton, L.C., Sprague, S.V., Nobbs, A.H., Ireland A.J., Sandy, J.R., *et al.* (2014) Axenic culture of a Candidate division TM7 bacterium from the human oral cavity and biofilm interactions with other oral bacteria. *Appl Environ Microbiol* **80**: 6480–6489.

Starr, E.P., Shi, S., Blazewicz, S.J., Probst, A.J., Herman, D.J., Firestone, M.K., *et al.* (2018) Stable isotope informed genome-resolved metagenomics reveals that Saccharibacteria utilize microbially-processed plant-derived carbon. *Microbiome* **6**: 122.

Takebe, H., Tominaga, K., Fujiwara, K., Yamamoto, K., and Yoshida, T. (2020) Differential responses of a coastal prokaryotic community to phytoplanktonic organic matter derived from cellular components and exudates. *Microbes Environ* **35**: ME20033.

Wrighton, K.C., Thomas, B.C., Sharon, I., Miller, C.S., Castelle, C.J., VerBerkmoes, N.C., *et al.* (2012) Fermentation, hydrogen, and sulfur metabolism in multiple uncultivated bacterial phyla. *Science* **337**: 1661–1665.

Wrighton, K.C., Castelle, C.J., Wilkins, M.J., Hug, L.A., Sharon, I., Thomas, B.C., *et al.* (2014) Metabolic interdependencies between phylogenetically novel fermenters and respiratory organisms in an unconfined aquifer. *ISME J* **8**: 1452–1463.

Yakimov, M.M., Merkel, A.Y., Gaisin, V.A., Pilhofer, M., Messina, E., Hallsworth, J.E., *et al.* (2021) Cultivation of a vampire: '*Candidatus* Absconditicoccus praedator'. *Environ Microbiol* **24**: 30–49.

Zhang, T., Shao, M.F., and Ye, L. (2012) 454 Pyrosequencing reveals bacterial diversity of activated sludge from 14 sewage treatment plants. *ISME J* **6**: 1137–1147.

Microbes Environ. Vol. 30, No. 2, 145-150, 2015
https://www.jstage.jst.go.jp/browse/jsme2 doi:10.1264/jsme2.ME14121

MOST CITED ORIGINAL PAPER PUBLISHED
IN THE LAST TEN YEARS [2015-2024] (198 citations)

A Comprehensive, Automatically Updated Fungal ITS Sequence Dataset for Reference-Based Chimera Control in Environmental Sequencing Efforts

R. Henrik Nilsson[1]*, Leho Tedersoo[2], Martin Ryberg[3], Erik Kristiansson[4], Martin Hartmann[5,6], Martin Unterseher[7], Teresita M. Porter[8], Johan Bengtsson-Palme[9], Donald M. Walker[10], Filipe de Sousa[1], Hannes Andres Gamper[11], Ellen Larsson[1], Karl-Henrik Larsson[12], Urmas Kõljalg[2,13], Robert C. Edgar[14], and Kessy Abarenkov[13]

[1]*Department of Biological and Environmental Sciences, University of Gothenburg, Box 461, 405 30 Gothenburg, Sweden;* [2]*Institute of Ecology and Earth Sciences, University of Tartu, Lai 40, Tartu 51005, Estonia;* [3]*Department of Organismal Biology, Uppsala University, Norbyvägen 18D, 75236 Uppsala, Sweden;* [4]*Department of Mathematical Statistics, Chalmers University of Technology, 412 96 Göteborg, Sweden;* [5]*Forest Soils and Biogeochemistry, Swiss Federal Research Institute WSL, Zuercherstrasse 111, 8903 Birmensdorf, Switzerland;* [6]*Molecular Ecology, Institute for Sustainability Sciences, Agroscope, Reckenholzstrasse 191, 8046 Zurich, Switzerland;* [7]*Ernst-Moritz-Arndt University, Institute of Botany and Landscape Ecology, Soldmannstr. 15, D-17487 Greifswald, Germany;* [8]*Department of Biology, McMaster University, Hamilton, Ontario, L8S 4K1, Canada;* [9]*Department of Infectious Diseases, Institute of Biomedicine, Sahlgrenska Academy, University of Gothenburg, Guldhedsgatan 10, SE-413 46 Gothenburg, Sweden;* [10]*Department of Natural Sciences, The University of Findlay, Findlay, OH, USA;* [11]*Group of Plant Nutrition, Institute of Agricultural Sciences, Department of Environmental Systems Science, ETH Zurich, Eschikon 33, 8315 Lindau (ZH), Switzerland;* [12]*Natural History Museum, P.O. Box 1172 Blindern, 0318 Oslo, Norway;* [13]*Natural History Museum, University of Tartu, Vanemuise 46, Tartu 51014, Estonia; and* [14]*Tiburon, CA, USA*

(Received August 25, 2014—Accepted January 14, 2015—Published online March 19, 2015)

The nuclear ribosomal internal transcribed spacer (ITS) region is the most commonly chosen genetic marker for the molecular identification of fungi in environmental sequencing and molecular ecology studies. Several analytical issues complicate such efforts, one of which is the formation of chimeric—artificially joined—DNA sequences during PCR amplification or sequence assembly. Several software tools are currently available for chimera detection, but rely to various degrees on the presence of a chimera-free reference dataset for optimal performance. However, no such dataset is available for use with the fungal ITS region. This study introduces a comprehensive, automatically updated reference dataset for fungal ITS sequences based on the UNITE database for the molecular identification of fungi. This dataset supports chimera detection throughout the fungal kingdom and for full-length ITS sequences as well as partial (ITS1 or ITS2 only) datasets. The performance of the dataset on a large set of artificial chimeras was above 99.5%, and we subsequently used the dataset to remove nearly 1,000 compromised fungal ITS sequences from public circulation. The dataset is available at http://unite.ut.ee/repository.php and is subject to web-based third-party curation.

Key words: fungi, chimera detection, reference dataset, molecular ecology, PCR artifacts

Fungi form a large and diverse group of heterotrophic organisms. Molecular (DNA sequence) data have gradually become a critical research tool in mycology, owing largely to the subterranean or otherwise inconspicuous nature of much of fungal life coupled with a general lack of tangible, discriminatory morphological characteristics in many fungi (12, 33). In many cases, DNA sequences represent the only means for high-precision species identification and delimitation (15). However, molecular mycology is not devoid of complications, and many technical issues and potential pitfalls need to be considered before DNA sequences can be applied for scientific purposes (13, 18). One of these complications is the unintentional generation of chimeric sequences during either PCR amplification or the assembly of individual sequence reads. Chimeras are artificial DNA sequences that are composed of two (or sometimes more) sequence fragments that do not naturally belong together (35). Most chimeras are produced during PCR when DNA templates of more than one sequence type are co-amplified and incomplete amplicons act as primers on not fully matching templates. These template switches are more likely to occur if the targeted gene/marker features a highly conserved segment that is very similar among different taxa in the mixed DNA template pool (9). The resulting chimeric sequence consists of two (or more) parts that originate from different parent sequence types. Chimeras of this kind lack a biological interpretation and need to be removed from any dataset in which they exist. Failure to do so will compromise any analyses the dataset is used in, including species identification and delimitation, richness estimation, and multiple sequence alignment/phylogenetic inference (23).

The nuclear ribosomal internal transcribed spacer (ITS) region is the formal fungal barcode and the most commonly sequenced genetic marker in mycology (2, 28). The average length of the ITS region is 550 base-pairs (bp) in the fungal kingdom, but varies markedly among lineages (8, 29). It is composed of the two variable spacers, ITS1 and ITS2, and the intercalary, highly conserved 5.8S ribosomal gene. The latter readily acts as a bridge-point for chimeric extension in mixed-template PCR, making chimera control an essential

* Corresponding author. E-mail: henrik.nilsson@bioenv.gu.se;
Tel: +46–31–7862623; Fax: +46-31-786 2560.

part of ITS-based mycological research (31, 32). Studies employing cloning of PCR amplicons need to be particularly vigilant against chimeras. In cloning, a single PCR fragment is selected, multiplied, and sequenced; therefore, any polymerase-generated artifact will penetrate to an extent not observed in direct Sanger sequencing, in which sequence chromatograms represent the averaged signal from numerous original templates rather than a single PCR fragment. Chimera control also has to be exercised when working with individual specimens such as fruiting bodies. Contamination in any of the laboratory steps or the presence of intra-sporocarp parasites or commensals such as *Hyphomycetes* in boletes (4) and lichen-inhabiting lineages in *Tremellales* and *Filboasidiales* (22) may produce chimeras in these cases. Chimera control is also essential in next-generation sequencing of fungal communities in environmental samples, in which the multi-species nature of the samples, sometimes coupled with the intrinsic properties of the sequencing platform, provide ample opportunities for chimera formation (27). Prior to this study, a total of 1,825 chimeras involving public sequences of the fungal ITS region were recorded in the UNITE database for the molecular identification of fungi (1), the largest tailored and actively curated public database for fungal ITS sequences. UNITE mirrors the (Sanger-derived) fungal ITS sequences in the International Nucleotide Sequence Database Collaboration (INSDC: GenBank, EMBL, and DDBJ; 20) such that the chimera count in UNITE is essentially that of the public fungal ITS sequence corpus generated to date by the scientific community.

The detection of chimeras is challenging. Obvious cases of chimeras can be identified in smaller, homogeneous datasets by simply examining the corresponding multiple sequence alignment or using the sequences for a BLAST search in INSDC (Supplemental Fig. S1). Manual approaches become problematic in larger datasets. Nilsson *et al.* (23) released a Perl-based semi-automated chimera finder for fungal ITS sequences, the advantage of which is its ability to detect Sanger-length chimeras occurring at and above the ordinal level; however, it is ineffective against chimeras occurring within the same order. Edgar *et al.* (6) introduced UCHIME, a powerful and feature-rich chimera checker for all major computer platforms and read lengths. In its "reference database mode", it cleaves all query sequences into four (default) segments in order to determine whether the constituent parts are best matched by different sequences in the reference database; on the UCHIME chimera scale, more obvious mismatches have higher chimera scores. The user is presented with a list of sequences that exceed the chimera score cut-off threshold, and, ideally, need to be examined by hand. UCHIME also offers a "*de novo* mode" of chimera detection for newly generated next-generation sequencing datasets, in which putative chimeras are deduced based on the abundances of the estimated amplicon sequences (denoising used) or unique reads (no denoising used).

A rich and reliable reference database lies at the core of the reference database mode of UCHIME and similar programs. However, such a database is not readily available for fungal ITS sequences. Difficulties have been associated with using the corpus of fungal ITS sequences downloaded from INSDC because this dataset contains a non-trivial number of chimeras and sequences of other technical or annotation-related problems. In the present study, we introduced an incremental, taxonomically inclusive, and high-quality set of fungal ITS sequences (http://unite.ut.ee/repository.php) derived from the INSDC as mirrored in UNITE for use in chimera detection pursuits.

Materials and Methods

Compilation of the ITS reference dataset

UNITE downloads all fungal ITS sequences from INSDC twice a year, and subjects them to a series of semi-automated quality control measures. All sequences are then clustered at 80% similarity in USEARCH 7 (5) to produce clusters at roughly the genus/subgenus level. A multiple sequence alignment is computed for each such genus-level cluster for graphical display, and the sequences in each cluster are subjected to a second round of clustering, at roughly the species level (97%–100% similarity in 0.5% steps). The resulting operational taxonomic units—called *species hypotheses* (SHs)—are given unique names of the accession number type to enable unambiguous communication across studies and datasets, and are reachable through URIs such as http://unite.ut.ee/sh/SH158651.06FU. Although a species hypothesis may be composed of a single sequence if sanctioned manually, the present study focused on species hypotheses consisting of two or more sequences. (As discussed below, many singleton sequences do not meet quality requirements.) When logged into UNITE, the user can view genus-level alignments with the species hypotheses indicated (Fig. 1). Based on the most frequent sequence type in each species hypothesis, a sequence is automatically chosen as a representative sequence (at the 98.5% similarity threshold) for that species hypothesis. These representative sequences serve as the basis for the chimera reference database.

However, additional control may be desired over the sequence chosen to represent a species hypothesis in some cases. Sequences stemming from type specimens, for example, form particularly good candidates to represent a species hypothesis in so far as they are of sufficient length and read quality (17). UNITE offers web-based third-party designation of representative sequences to its users; a user can log in and easily change the choice of representative sequences or the similarity level at which they should be applied. These manually chosen representative sequences are referred to as *reference* sequences. During a recent workshop and subsequent annotation effort (17, 25), the participants, primarily fungal taxonomists, re-selected 2,936 reference sequences and verified another several thousand representative sequences for reliability and representativeness. In addition, 97 sequences that came out as singletons in the 97% clustering step and, hence, did not qualify as species hypotheses were sanctioned as formal species hypotheses by the participants. Schoch *et al.* (29) similarly presented a large number of ITS sequences from type material; these were designated as reference sequences for species hypotheses in UNITE where applicable. The end product was a set of 3,973 reference and 17,086 representative sequences spanning all species hypotheses across the entire fungal tree of life.

Evaluation of the ITS reference dataset

To estimate the power of the reference dataset to detect chimeras under ideal (artificial) conditions, we manipulated the 21,059-sequence reference dataset to contain only chimeric sequences; all sequences were bisected in the middle of the 5.8S gene, and the fragments were reshuffled randomly to produce a total of 21,059 chimeras. Each of these consisted of ITS1 + half of the 5.8S gene from one sequence, and the remainder of the 5.8S gene + ITS2 from another. Fragments were not allowed to graft back onto their parent sequence. The procedure was repeated ten times to produce ten different (21,059-sequence) datasets of chimeras. We ran these new chimeras through UCHIME using the present reference sequence file as the reference corpus.

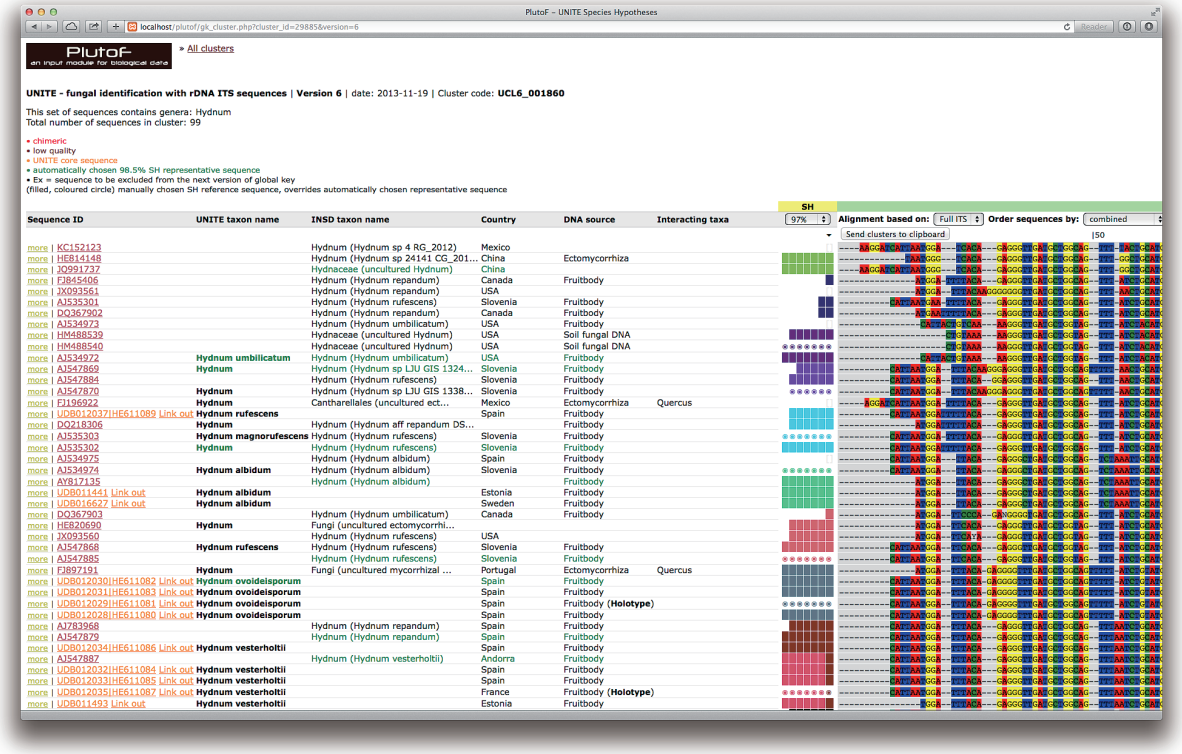


Fig. 1. A genus-level alignment in UNITE of the ectomycorrhizal genus *Hydnum*, with the individual species hypotheses (SHs) indicated by the colored boxes at different similarity levels (97–100%). One sequence (shown here in green) from each such species hypothesis was used to build the chimera reference dataset. Manually chosen reference sequences are indicated by filled circles in the SH column; these superseded the automatic choice of representative sequences for species hypotheses and are particularly suited for sequences from type (or otherwise authenticated) material. Two sequences from type specimens are indicated in the figure.

Chimeras among public ITS sequences

In an effort to detect and flag the worst cases of chimeric fungal ITS sequences in INSDC/UNITE, we ran the combined INSDC/UNITE dataset (376,840 more or less full-length fungal ITS sequences, excluding the 1,825 known chimeras) as a query in UCHIME 7.0.1090 using the default chimera score cut-off value (0.28) and the new 21,059-sequence dataset as the chimera-free reference dataset. We examined the 5,414 resulting putative chimeras manually for a chimeric nature following the procedure of Nilsson *et al.* (24). In order to evaluate the performance of UCHIME on these authentic sequences, we compared UCHIME scores for the sequences we deemed to be chimeric with the scores for those that were not considered to be chimeric by us. The Wilcoxon-Mann-Whitney test implemented in R 2.15.3 (http://cran.r-project.org) was used to statistically analyze differences in scores between the chimeric and non-chimeric sequences.

Results

Reference dataset

The reference dataset comprised 21,059 sequences (August 2014) and is available for download at http://unite.ut.ee/repository.php. It was designed for chimera control of more or less full-length fungal ITS sequences, and may, thus, not work well for sequences that contain non-trivial parts (200+ bp.) of the neighboring small-subunit (SSU/18S) and/or large subunit (LSU/28S) genes because the SSU and LSU are absent from the reference dataset. We also provide standalone ITS1-only and ITS2-only files extracted using ITSx for the same 21,059 sequences (3). This supports chimera detection in datasets containing only (full-length or partial) ITS1 or ITS2 sequences, notably those stemming from amplicon-based next-generation sequencing efforts.

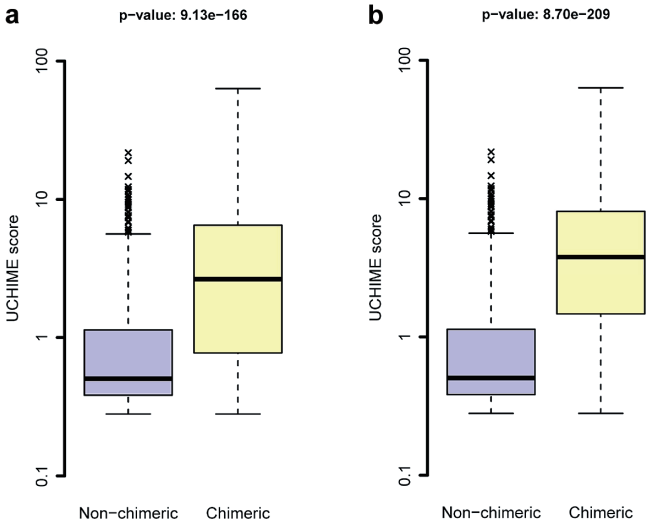


Fig. 2. Boxplots of UCHIME scores for non-chimeric and chimeric sequences. All sequences were included in panel (a), while sequences with low read quality were removed in panel (b). The score difference between the non-chimeric and chimeric sequences was statistically assessed through the Wilcoxon-Mann-Whitney test.

Evaluation of the ITS reference dataset

UCHIME identified an average of 99.82% (SD 0.0421) of the sequences in the ten 21,059-chimera datasets as chimeric by using default UCHIME settings.

Chimeras among public ITS sequences

A total of 5,414 (1.4%) out of the 376,840 UNITE/INSDC sequences were identified as putatively chimeric by UCHIME, ranging from cases with very high chimera scores (50+) to those barely exceeding the threshold value (0.28; Fig. 2). One hundred and eighty-seven sequences had a chimera score above 10; 768 had a chimera score between 1 and 10; and 4,459 sequences had a chimera score above the 0.28 threshold, but below 1. All these sequences were subjected to a manual examination in the INSDC and UNITE. We identified 724 (13.3%) as clear cases of chimeras and 239 (4.4%) that represented sequences of low read quality; however, we could not reach an unequivocal decision on the chimeric nature of the remaining 4,551 entries or a chimeric nature appeared to be unlikely. Screenshots of the BLAST results of these entries are shown in Supplemental Fig. S1. Approximately 25% of these sequences appeared to stem from fungi with conserved ITS2 regions, but variable ITS1 regions (*e.g.* lineages in *Aspergillus* and *Colletotrichum*; see 17), and approximately the same percentage of sequences appeared to be natural hybrids (cf. 10, 26), as evidenced by the multiple independent recoveries of many of these species. Although 10% of the sequences appeared to be of low read quality, explicit evidence was lacking. A decision on whether the remaining sequences were chimeric was dependent on one to five base-pairs. We concluded that unequivocal decisions were impossible in these cases, and the sequences were left in the database. Ten cases of incorrectly assembled sequences ("assembly chimeras") were also detected. Clear cases of chimeras, the assembly chimeras, and low-read quality entries were marked as compromised in UNITE and reported to the INSDC.

The Wilcoxon-Mann-Whitney test revealed significant differences in scores between sequences considered chimeric and those that were not found to be chimeric, with median UCHIME scores of 3.78 and 0.51 for the chimeric and non-chimeric groups, respectively ($p<10^{-16}$; Fig. 2).

Database implementation

We modified UNITE to hold and display UCHIME results in order to facilitate chimera detection for users with limited experience in command-line programs. Sequences that exceeded the default score at which UCHIME considered a sequence chimeric have now been marked as putatively chimeric in the database, with the score and two putative parent sequences indicated and hyperlinked (Supplemental Fig. S2). These data are re-generated after each re-computation of the species hypotheses in UNITE. The 5,414 sequences indicated as putatively chimeric in UNITE/INSDC in the present study were specified accordingly.

Discussion

We here present a 21,059-sequence fungal ITS dataset for use in chimera control in UCHIME or any other chimera detection program that relies on a chimera-free reference dataset. Amplicon-based next-generation sequencing studies are strongly advised to consider the reference-free *de novo* chimera detection mode of UCHIME/UPARSE (7); however, even in these studies, an established reference dataset may need to be referred to, at least for particularly problematic cases. Our sequence dataset, which is available at http://unite.ut.ee/repository.php, is updated automatically as the number of fungal ITS sequences in INSDC increases. It is furthermore subject to third-party annotation, such that those who feel they are in a position to improve the data (and particularly the choice or similarity levels of applications of reference sequences for species hypotheses) can do so. Nevertheless, it is not a dataset devoid of potential problems, with the most obvious one being its limited taxonomic depth. Although the ITS region is the formal fungal barcode, ITS sequences are only available for ~17,000 fully identified species out of the estimated one to several million extant species of fungi (11, 17). Chimera detection programs, in contrast, work better when both parent sequences of a chimera are present in the reference database. This will, in practice, not always be the case for those processing newly generated ITS sequences from environmental or taxonomic studies, suggesting that a certain proportion of false-positive (and false-negative) identifications will have to be tolerated by the mycological community for now. This testifies to the importance of manual verification of sequences indicated to be putatively chimeric (and possibly also sequences that are close to, but do exceed the threshold for what is regarded a chimeric sequence).

A second shortcoming lies in the requirement that a species hypothesis in UNITE must be composed of two or more sequences (although singleton sequences can be sanctioned manually as species hypotheses). This requirement is important from a sequence quality point of view because sequences that form singletons in clustering approaches of large datasets are often associated with quality problems (13, 24, 30). To automatically endow all singleton sequences with the status of species hypotheses may give rise to a large number of phantom species hypotheses. Such species hypotheses may not correspond to any biological reality because the sequences are in some way technically compromised. The inclusion of these sequences in the chimera reference dataset may reduce the power of the dataset for chimera detection. Nevertheless, a certain proportion of singleton sequences do correspond to high-quality DNA sequences that, although not finding any close match in the sequence databases, do correspond to actual species. These have, by default, not been included in the present reference database. UNITE provides the opportunity to manually sanction such singleton sequences as reference sequences for species hypotheses. Manual examinations and designations form a bottleneck here, and the absolute majority of reference sequences that have been designated in UNITE are found in species hypotheses composed of two or more sequences. This number is expected to rise because the number of third-party annotators is increasing (25). The requirement that a species hypothesis must be composed of

two or more sequences excludes most spurious sequences from being used as representative sequences, but does not offer full protection against unwanted representative sequences. For example, the exact same chimeric sequence may be formed twice or even more in the same study or across studies. Therefore, our "chimera-free" reference dataset may still contain a few chimeras. Smaller chimeric insertions (or untrimmed vectors) in sequences, in which the insertions are small enough not to penetrate to give rise to different species hypotheses, may also be present in the dataset. Although we are not currently aware of such a case, we ask the users to be aware of their potential existence in the dataset and to take action when they find any such entries. A final shortcoming lies in the fact that ITS chimeras spanning kingdoms, such as Fungi and Plantae, are, although unlikely, at least conceivable (21). The extent to which the present reference dataset, which covers fungi only, can be used to find cross-kingdom chimeras has yet to be established.

The intrinsic properties of the ITS region may also make successful chimera detection difficult. The highly variable ITS1 and ITS2 subregions may differ slightly in variability, with the length and sequence content of ITS2 being more conserved, at least in *Basidiomycota* (14, 34). An ITS2 sequence type may be conserved across several species whereas the corresponding ITS1 sequences differ in length/sequence content by a few base-pairs or more. These cases may give rise to false-positive chimera detections in which ITS1, due to its being unique, will be assigned to the correct parent sequence A, whereas ITS2 will be assigned by chance to parent sequence B. This designation occurs due to sequences A and B having identical ITS2 sequences. The multicopy nature of the ITS region, with the potential for several different allelic variants of the marker, may occasionally give rise to chimeras that may be both difficult to find and complicated to verify (19). A few cases of naturally occurring "chimeras" have been reported previously (16, 36). To summarize, the user needs to be aware that it is not always possible to unequivocally prove that a sequence is chimeric. This is particularly true for sequences downloaded from public sequence databases, in which the context of the sequence, as well as additional, explanatory data, will not always be available for examination.

We demonstrated that the UCHIME/reference dataset combination was useful in the pursuit of chimeric fungal ITS sequences in the public corpus; nearly 1,000 substandard publicly deposited sequences were identified and removed. Nevertheless, we do not want to imply that fungal ITS chimera detection will be trivial from this point on. When we manipulated the reference dataset to contain only chimeric sequences, the UCHIME/reference dataset combination detected an average of 99.82% out of the 21,059 chimeras as chimeric, which was considered to be a very satisfactory performance. However, this approach relied on the random regrafting of sequence halves, which, in reality, may be more common among closely related species (similar sequences). We similarly assumed chimeric breakpoints to occur in the very conserved 5.8S gene, which may be an oversimplification. Thus, the user should not expect a chimera detection efficiency of 99.82% in real-life datasets. Nevertheless, we believe that the present dataset will lead to noise reduction in fungal ITS datasets in medical, taxonomic, and environmental sequencing of fungi and fungal communities. We also hope that users will report any chimeric sequences detected in INSDC/UNITE to the database to prevent bad data from propagating through the literature. We invite the community to improve the present dataset through UNITE, particularly by designating reference sequences for species hypotheses and excluding substandard entries. Other improvements, such as taxonomic re-annotations, are also valuable to the mycological community. The time dedicated to third-party annotation is arguably a small price to pay for the knowledge that chimera detection and molecular identification will be performed in a richer and more informed way by the user and the remainder of the scientific community.

Acknowledgements

RHN acknowledges financial support from the Swedish Research Council of Environment, Agricultural Sciences, and Spatial Planning (FORMAS, 215-2011-498). TMP was funded by the Government of Canada through Genome Canada and the Ontario Genomics Institute through the Biomonitoring 2.0 project (OGI-050). The NEFOM network is acknowledged for infrastructural support. The authors have no conflict of interests to report.

References

1. Abarenkov, K., R.H. Nilsson, K.-H. Larsson, *et al.* 2010. The UNITE database for molecular identification of fungi—recent updates and future perspectives. New Phytol. 186:281–285.
2. Begerow, D., R.H. Nilsson, M. Unterseher, and W. Maier. 2010. Current state and perspectives of fungal DNA barcoding and rapid identification procedures. Appl. Microbiol. Biotechnol. 87:99–108.
3. Bengtsson-Palme, J., M. Ryberg, M. Hartmann, *et al.* 2013. Improved software detection and extraction of ITS1 and ITS2 from ribosomal ITS sequences of fungi and other eukaryotes for analysis of environmental sequencing data. Methods Ecol. Evol. 4:914–919.
4. Douhan, G.W., and D.M. Rizzo. 2003. Host-parasite relationships among bolete infecting *Hypomyces* species. Mycol. Res. 107:1342–1349.
5. Edgar, R.C. 2010. Search and clustering orders of magnitude faster than BLAST. Bioinformatics 26:2460–2461.
6. Edgar R.C., B.J. Haas, J.C. Clemente, C. Quince, and R. Knight. 2011. UCHIME improves sensitivity and speed of chimera detection. Bioinformatics 27:2194–2200.
7. Edgar R.C. 2013. UPARSE: highly accurate OUT sequences from microbial amplicon reads. Nature Methods 10:996–998.
8. Feibelman, T., P. Bayman, and W.G. Cibula. 1994. Length variation in the internal transcribed spacer of ribosomal DNA in chanterelles. Mycol. Res. 98:614–618.
9. Fonseca, V.G., B. Nichols, D. Lallias, C. Quince, G.R. Carvalho, D.M. Power, and S. Creer. 2012. Sample richness and genetic diversity as drivers of chimera formation in nSSU metagenetic analyses. Nucleic Acids Res. 40:e66.
10. González, S.S., E. Barrio, J. Gafner, and A. Querol. 2006. Natural hybrids from *Saccharomyces cerevisiae*, *Saccharomyces bayanus* and *Saccharomyces kudriavzevii* in wine fermentations. FEMS Yeast Res. 6:1221–1234.
11. Hawksworth, D.L. 2001. The magnitude of fungal diversity: the 1.5 million species estimate revisited. Mycol. Res. 105:1422–1432.
12. Hibbett, D.S., and J.W. Taylor. 2013. Fungal systematics: is a new age of enlightenment at hand? Nat. Rev. Microbiol. 11:129–133.
13. Hyde, K.D., D. Udayanga, D.S. Manamgoda, *et al.* 2013. Incorporating molecular data in fungal systematics: a guide for aspiring researchers. Curr. Res. Environ. App. Mycol. 3:1–32.
14. Ihrmark, K., I.T.M. Bödeker, K. Cruz-Martinez, *et al.* 2012. New primers to amplify the fungal ITS2 region—evaluation by 454-sequencing of artificial and natural communities. FEMS Microbiol. Ecol. 82:666–677.

15. Kang, S., M.A.M. Mansfield, B. Park, D.M. Geiser, K.L. Ivors, M.D. Coffey, N.J. Grünwald, F.N. Martin, C.A. Lévesque, and J.E. Blair. 2010. The promise and pitfalls of sequence-based identification of plant pathogenic fungi and oomycetes. Phytopathology 100:732–737.
16. Kauserud, H., I.B. Vegården, C. Decock, and N. Hallenberg. 2007. Hybridization among cryptic species of the cellar fungus *Coniophora puteana* (Basidiomycota). Mol. Ecol. 16:389–399.
17. Kõljalg, U., R.H. Nilsson, K. Abarenkov, *et al.* 2013. Towards a unified paradigm for sequence-based identification of fungi. Mol. Ecol. 22:5271–5277.
18. Lindahl, B.D., R.H. Nilsson, L. Tedersoo, *et al.* 2013. Fungal community analysis by high-throughput sequencing of amplified markers—a user's guide. New Phytol. 199:288–299.
19. Lindner, D.L., T. Carlsen, R.H. Nilsson, M. Davey, T. Schumacher, and H. Kauserud. 2013. Employing 454 amplicon pyrosequencing to reveal intragenomic divergence in the internal transcribed spacer (ITS) rDNA region in fungi. Ecol. Evol. 3:1751–1764.
20. Nakamura, Y., G. Cochrane, and I. Karsch-Mizrachi. 2013. The International nucleotide sequence database collaboration. Nucleic Acids Res. 41(D1):D21–D24.
21. Nguyen, N.H., D. Smith, K. Peay, and P. Kennedy. 2014. Parsing ecological signal from noise in next generation amplicon sequencing. New Phytol. doi:10.1111/nph.12923.
22. Millanes, A., P. Diederich, M. Westberg, T. Knutsson, and M. Wedin. 2014. *Tremella rhizocarpicola* sp. nov. and other interesting lichenicolous *Tremellales* and *Filobasidiales* in the Nordic countries. MycoKeys 8:31–41.
23. Nilsson, R.H., K. Abarenkov, V. Veldre, S. Nylinder, P. De Wit, S. Brosche, J.F. Alfredsson, M. Ryberg, and E. Kristiansson. 2010. An open source chimera checker for the fungal ITS region. Mol. Ecol. Res. 10:1076–1081.
24. Nilsson, R.H., L. Tedersoo, K. Abarenkov, *et al.* 2012. Five simple guidelines for establishing basic authenticity and reliability of newly generated fungal ITS sequences. MycoKeys 4:37–63.
25. Nilsson, R.H., K.D. Hyde, J. Pawlowska, *et al.* 2014. Improving ITS sequence data for identification of plant pathogenic fungi. Fungal Divers. 67:11–19.
26. Pryszcz, L.P., T. Németh, A. Gácser, and T. Gabaldón. 2014. Genome comparison of *Candida orthopsilosis* clinical strains reveals the existence of hybrids between two distinct subspecies. Genome Biol. Evol. 6:1069–1078.
27. Quince, C., A. Lanzen, R.J. Davenport, and P.J. Turnbaugh. 2011. Removing noise from pyrosequenced amplicons. BMC Bioinf. 12:38.
28. Schoch, C.L., K.A. Seifert, S. Huhndorf, *et al.* 2012. Nuclear ribosomal internal transcribed spacer (ITS) region as a universal DNA barcode marker for Fungi. Proc. Natl. Acad. Sci. U.S.A. 109:6241–6246.
29. Schoch, C.L., B. Robbertse, C. Robert, *et al.* 2014. Finding needles in haystacks: linking scientific names, reference specimens and molecular data for Fungi. Database 2014:bau061.
30. Tedersoo, L., R.H. Nilsson, K. Abarenkov, T. Jairus, A. Sadam, I. Saar, M. Bahram, E. Bechem, G. Chuyong, and U. Kõljalg. 2010. 454 Pyrosequencing and Sanger sequencing of tropical mycorrhizal fungi provide similar results but reveal substantial methodological biases. New Phytol. 188:291–301.
31. Tedersoo, L., K. Abarenkov, R.H. Nilsson, *et al.* 2011. Tidying up International Nucleotide Sequence Databases: ecological, geographical, and sequence quality annotation of ITS sequences of mycorrhizal fungi. PLoS ONE 6:e24940.
32. Tedersoo, L., M. Bahram, and I.A. Dickie. 2014. Does host plant richness explain diversity of ectomycorrhizal fungi? Re-evaluation of Gao *et al.* (2013) data sets reveals sampling effects. New Phytol. 23:992–995.
33. Tedersoo, L., M. Bahram, S. Põlme, *et al.* 2014. Global diversity and geography of soil fungi. Science 346:1256688.
34. Toju, H., A.S. Tanabe, S. Yamamoto, and H. Sato. 2012. High-coverage ITS primers for the DNA-based identification of ascomycetes and basidiomycetes in environmental samples. PLoS ONE 7:e40863.
35. Wang, G.C.Y., and Y. Wang. 1997. The frequency of chimeric molecules as a consequence of PCR co-amplification of 16S rRNA genes from different bacterial species. Microbiology 142:1107–1114.
36. Xie, J., Y. Fu, D. Jiang, G. Li, J. Huang, B. Li, T. Hsiang, and Y. Peng. 2008. Intergeneric transfer of ribosomal genes between two fungi. BMC Evol. Biol. 8:87.

Microbes Environ. 37(5), 2022
https://www.jstage.jst.go.jp/browse/jsme2 doi:10.1264/jsme2.ME22034

MOST CITED REVIEW PAPER PUBLISHED
IN THE LAST THREE YEARS [2022-2024] (8 citations)

Special issue: *Viruses in the Aquatic Realm*

Minireview

Eukaryotic Microbial RNA Viruses—Acute or Persistent? Insights into Their Function in the Aquatic Ecosystem

Syun-ichi Urayama[1,2*], Yoshihiro Takaki[3], Yuto Chiba[1], Yanjie Zhao[1], Misa Kuroki[1], Daisuke Hagiwara[1,2], and Takuro Nunoura[4]

[1]*Department of Life and Environmental Sciences, Laboratory of Fungal Interaction and Molecular Biology (donated by IFO), University of Tsukuba, 1–1–1 Tennodai, Tsukuba, Ibaraki 305–8577, Japan;* [2]*Microbiology Research Center for Sustainability (MiCS), University of Tsukuba, 1–1–1 Tennodai, Tsukuba, Ibaraki 305–8577, Japan;* [3]*Super-cutting-edge Grand and Advanced Research (SUGAR) Program, Japan Agency for Marine Science and Technology (JAMSTEC), 2–15 Natsushima-cho, Yokosuka, Kanagawa 237–0061, Japan; and* [4]*Research Center for Bioscience and Nanoscience (CeBN), JAMSTEC, 2–15 Natsushima-cho, Yokosuka, Kanagawa 237–0061, Japan*

(Received May 6, 2022—Accepted June 26, 2022—Published online August 3, 2022)

Isolated RNA viruses mainly parasitize eukaryotes. RNA viruses either expand horizontally by infecting hosts (acute type) or coexist with the host and are vertically inherited (persistent type). The significance of persistent-type RNA viruses in environmental viromes (the main hosts are expected to be microbes) was only recently reported because they had previously been overlooked in virology. In this review, we summarize the host-virus relationships of eukaryotic microbial RNA viruses. *Picornavirales* and *Reoviridae* are recognized as representative acute-type virus families, and most of the microbial viruses in *Narnaviridae*, *Totiviridae*, and *Partitiviridae* are categorized as representative persistent-type viruses. Acute-type viruses have only been found in aquatic environments, while persistent-type viruses are present in various environments, including aquatic environments. Moreover, persistent-type viruses are potentially widely spread in the RNA viral sequence space. This emerging evidence provides novel insights into RNA viral diversity, host-virus relationships, and their history of co-evolution.

Key words: aquatic, RNA virus, eukaryote

An RNA virus has single- or double-stranded RNA as its genome. The genome sizes of RNA viruses range from several kb to several tens of kb, and some harbor segmented genomes depending on the virus group/species (King *et al.*, 2012; Koonin *et al.*, 2020). To date, most RNA viruses have been isolated from eukaryotes, and more than 50% of isolated viruses from eukaryotes are RNA viruses (Nasir *et al.*, 2014). Since the majority of eukaryotes are expected to be infected by RNA viruses, RNA viruses and eukaryotes likely have a long history of co-evolution (Koonin *et al.*, 2015). In contrast, no or very few RNA virus families have been isolated from archaea and bacteria, respectively; however, the diversity of prokaryotic RNA viruses was very recently expanded by metatranscriptomics and subsequent *in silico* analyses (Callanan *et al.*, 2020; Neri, U., *et al.* 2022. A five-fold expansion of the global RNA virome reveals multiple new clades of RNA bacteriophages. bioRxiv https://doi.org/10.1101/2022.02.15.480533).

Historically, RNA viruses have been detected as infectious causative agents in humans and economically important plants and animals (Beijerinck, 1898; Loeffler and Frosch, 1898; Reed and Carroll, 1901). However, in the last several decades, fungal RNA viruses that persistently infect their hosts without causing clear phenotypic changes have been discovered, although a few have been found in edible mushrooms and pathogenic fungi that cause phenotypic symptoms (Nuss, 2005; Roossinck, 2014). Moreover, similar "cryptic" RNA viruses are widely present in plants (Roossinck *et al.*, 2010). To obtain a more detailed understanding of RNA viruses, Roossinck categorized the lifestyles of plant/fungal viruses into two types, persistent and acute (Roossinck, 2010, 2012, 2015). Briefly, viruses with a persistent lifestyle are transmitted vertically via cell division, while viruses with an acute lifestyle are most frequently transmitted horizontally (Fig. 1). This classification is completely different from "acute infection", "persistent (latent) infection", and "chronic infection" in human infections. These clinical categories indicate the course of viral dose/activity or the state of disease in humans. For example, hepatitis C virus (HCV), a positive-sense single-stranded RNA virus, sometimes establishes a chronic infection after an acute infection. However, this does not represent a

* Corresponding author. E-mail: urayama.shunichi.gn@u.tsukuba.ac.jp; Tel: +81–29–853–6636; Fax: +81–29–853–4605.

Citation: Urayama, S., Takaki, Y., Chiba, Y., Zhao, Y., Kuroki, M., Hagiwara, D., and Nunoura, T. (2022) Eukaryotic Microbial RNA Viruses—Acute or Persistent? Insights into Their Function in the Aquatic Ecosystem. *Microbes Environ* **37**: ME22034.
https://doi.org/10.1264/jsme2.ME22034

change in the viral lifestyle. HCV is categorized as an acute-type RNA virus by Roossinck's criteria because it mainly spreads by horizontal transmission.

The acute and persistent types have both been identified in eukaryotic microbial lineages of fungi and protists. Furthermore, viral metagenomics enables us to identify diverse RNA viruses and infer their lifestyles. To date, a large diversity of RNA viruses has been revealed in soil and aquatic microbial ecosystems (Culley *et al.*, 2006; Suttle, 2016; Urayama *et al.*, 2018a; Starr *et al.*, 2019; Wolf *et al.*, 2020; Edgar *et al.*, 2022; Zayed *et al.*, 2022), in addition to biological samples from organisms, such as invertebrates, deep-sea animals, and microbial consortia associated with sponges, macroalgae, and lichens (Shi *et al.*, 2016; Urayama *et al.*, 2018b, 2020a, 2020b; Chiba *et al.*, 2020). In this review, we summarize eukaryotic microbial RNA viruses based on their lifestyles and discuss their significance in microbial ecosystems. This review focuses on RNA viruses excluding retroviruses. Retroviruses are distinct from other RNA viruses in their life cycle because they are embedded into the host genome, whereas other RNA viruses are not. Retroviruses have a mixed strategy of vertical and horizontal transmission and change their strategy depending on the life stage.

Microbial RNA viruses have two life cycle strategies

We define acute- and persistent-type life cycles in microbial RNA viruses based on Roossinck's definition (Roossinck, 2010) (Fig. 1). Acute-type viruses are mainly transmitted horizontally via extracellular viral particles. These viruses have the ability to enter host cells from the outside and exit host cells to the outside. During the viral life cycle, acute-type viruses often cause "disease" as represented by cell lysis (Tomaru *et al.*, 2004). In contrast, persistent-type viruses are mainly transmitted vertically via the cell division of their host without cell lysis (Nuss, 2005). These viruses remain associated with their hosts for many generations with nearly 100% vertical transmission, and sometimes lack the ability to enter and exit host cells. In addition, some persistent-type RNA viruses lack capsid proteins (Dolja and Koonin, 2012). No gene is shared only in persistent-type RNA viruses. Persistence and vertical transmission in viruses has been suggested to correlate with commensal or mutualistic lifestyles (Roossinck, 2011; Marquez and Roossinck, 2012).

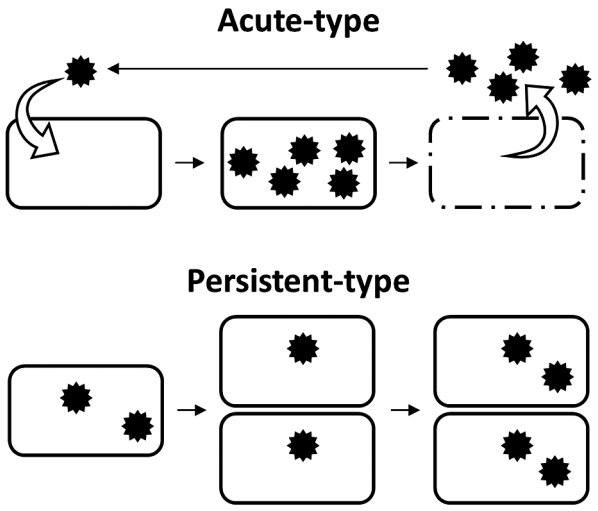


Fig. 1. Schematic of transmission routes of microbial RNA viruses.

The host cell phenotype is an attractive subject in virology, and persistent- and acute-type lifestyles overlap with symbiosis and antagonism, respectively. However, difficulties are associated with the classification of viruses according to the phenotype of the host organism because the host-virus relationship changes depending on the surrounding conditions (Xu *et al.*, 2008). Furthermore, we cannot test all environmental conditions that may affect the host phenotype. Therefore, we defined the two types of viruses based on the viral transmission route in this review.

Heterosigma akashiwo RNA virus (HaRNAV) (Tai *et al.*, 2003; Lang *et al.*, 2004) is a well-known acute-type RNA virus associated with microbial eukaryotes in aquatic ecosystems. HaRNAV was isolated from a *Heterosigma akashiwo* culture that showed the lysis of host cells after the inoculation of a virus particle fraction from seawater. To maintain HaRNAV in lab conditions, a 0.22-μm filtered cell lysate including HaRNAV is inoculated into *H. akashiwo* cultures, and after cell lysis, the lysate is used as an inoculum. In this case, since the host is re-infected with the daughter virus obtained by filter filtration from the lytic solution, we defined the RNA virus as an acute-type virus. Based on the genome sequence of HaRNAV, it belongs to *Marnaviridae* (Lang *et al.*, 2021), which includes several acute-type microbial RNA viral genera, such as *Marnavirus* and *Labyrnavirus* (Lang *et al.*, 2004; Takao *et al.*, 2006).

Among persistent-type RNA viruses, Saccharomyces cerevisiae virus L-A (ScV-L-A) (Wickner, 1996) is a representative strain. The presence of RNA viruses may be confirmed by long cellular dsRNA because it only consists of dsRNA viral genomes and/or replicative intermediates of non-retro ssRNA viruses (Morris and Dodds, 1979). ScV-L-A does not confer a detectable phenotype upon its host (yeast) cells (Schmitt and Breinig, 2002), and viral dsRNA is maintained in the host cell and transmitted to daughter cells via cell division. ScV-L-A may be maintained and amplified in *S. cerevisiae* cells under laboratory conditions. In this case, since the RNA virus is maintained by a host sub-culture, we define the RNA virus as a persistent-type virus. ScV-L-A belongs to the family *Totiviridae*, which includes other known totiviruses isolated from protists (Goodman *et al.*, 2011).

The definitions of acute- and persistent-type viruses are not applicable in some exceptional cases. For example, fungal persistent-type RNA viruses are mainly transmitted to daughter cells, but may also be transmitted between closely related fungal strains through anastomosis (Liu *et al.*, 2003; Vainio *et al.*, 2011). On a geological time scale, the transmission of RNA viruses between plants and fungi may also have occurred (Nibert *et al.*, 2014; Roossinck, 2019; Bian *et al.*, 2020).

Isolated microbial RNA viruses

We herein summarized and grouped isolated microbial

RNA viruses based on the infection type (Table S1). Microbial RNA viruses with acute-type lifestyles have been obtained from aquatic host strains (Sadeghi *et al.*, 2021). Among them, ssRNA viruses that lyse host cells have been reported, such as Heterosigma akashiwo RNA virus (HaRNAV) (*Picornavirales*), which infects the toxic bloom-forming raphidophyte *Heterosigma akashiwo* (Raphidophyceae) (Tai *et al.*, 2003); Heterocapsa circularisquama RNA virus 01 (*Alvernaviridae*), which infects the bivalve-killing dinoflagellate *Heterocapsa circularisquama* (Dinophyceae) (Tomaru *et al.*, 2004); and Aurantiochytrium single-stranded RNA virus 01 (*Picornavirales*), which infects the marine fungoide *Schizochytrium* sp. (Thraustochytriaceae) (Takao *et al.*, 2005). The dsRNA virus, Micromonas pusilla reovirus (*Reoviridae*), which infects and lyses the marine photosynthetic protist *Micromonas pusilla* (Mamiellophyceae) (Brussaard *et al.*, 2004), has also been isolated.

In contrast, microbial RNA viruses with persistent-type lifestyles have mainly been identified using intracellular dsRNA (a molecular biomarker) or RNA-seq technology. This type of RNA virus has been exclusively studied and identified in fungi (Ghabrial *et al.*, 2015; Kotta-Loizou and Coutts, 2017). To date, identified dsRNA viruses include unclassified RNA viruses that infect the cultivated mushroom *Agaricus bisporus* (van der Lende *et al.*, 1994); Penicillium chrysogenum virus (*Chrysoviridae*), which infects the penicillin-producing strains of *Penicillium chrysogenum* (Banks *et al.*, 1969); and Saccharomyces cerevisiae virus L-A (*Totiviridae*), which infects the budding yeast *S. cerevisiae* (Wickner, 1996). Oomycetes (Oomycota) have also been subject to dsRNA-based surveillance, and the plant pathogens *Phytophthora infestans* and *Asparagus officinalis* are known to harbor Phytophthora infestans RNA virus 3 (unclassified) and Phytophthora endornavirus 2 (and 3) (*Endornaviridae*), respectively (Cai *et al.*, 2013; Uchida *et al.*, 2021). In addition, protozoans, such as *Trichomonas*, *Leishmania*, and *Giardia*, harbor persistent-type RNA viruses (Wang and Wang, 1986; Stuart *et al.*, 1992; Khoshnan and Alderete, 1994); however, the presence of extracellular infection routes has been suggested (Wang and Wang, 1986; Torrecilhas *et al.*, 2020).

RNA viruses with acute-type lifestyles have been detected in aquatic environments, while those with persistent-type lifestyles have been found in terrestrial environments. However, ssRNA and dsRNA viruses with persistent-type lifestyles have both been recently identified from a marine oomycete strain in the genus *Halophytophthora* (Botella *et al.*, 2020) and a marine fungal strain isolated from the seagrass *Posidonia oceanica* (Nerva *et al.*, 2016). The identification of these strains was based on high-throughput sequencing methods, while culture-dependent isolation has typically been used to identify viruses in aquatic research. However, in terrestrial environments, few attempts have been made to obtain acute-type RNA viruses. We cannot exclude the possibility that the different distribution pattern observed between the two lifestyle types is a result of methodological bias in virus detection and isolation; however, acute-type RNA viruses may have advantages in dispersal via viral particles to access new host cells in aquatic environments. In the case of DNA phages, a number of theories have been proposed, including Kill-the-Winner and Piggyback-the-Winner (Obeng *et al.*, 2016; Pratama and van Elsas, 2018). Theoretical modeling in addition to further studies will provide more detailed insights into the distribution of acute- and persistent-type RNA viruses under diverse environmental conditions.

Persistent and acute RNA viruses in the RNA viral sequence space

To visualize the richness of persistent-type RNA viruses in the sequence space, RNA viruses isolated from eukaryotic microorganisms were classified into acute- or persistent-types based on culture-dependent laboratory observations (Table S1), as described above. In total, 96% (304/314) of eukaryotic microbial RNA viruses were recognized as persistent types (Fig. 2), while we were unable to exclude the possibility of experimental biases in culture-dependent analyses to obtain acute-type RNA viruses in eukaryotic microorganisms. To predict the distribution of persistent-type RNA viruses in the total RNA virome sequence space, we used this list as an operational reference virus list because of the lack of knowledge on persistent-type RNA viruses in other host organisms and limited metagenomic data. Based on the operational reference virus list, RNA viral clusters including all known viral RdRp sequences were analyzed, and clusters including RNA viruses related to microbial persistent-type viruses (>50% amino acid similarity) were distinguished from those of acute-type viruses (Fig. 3). In our current analysis, more than 1/3 of

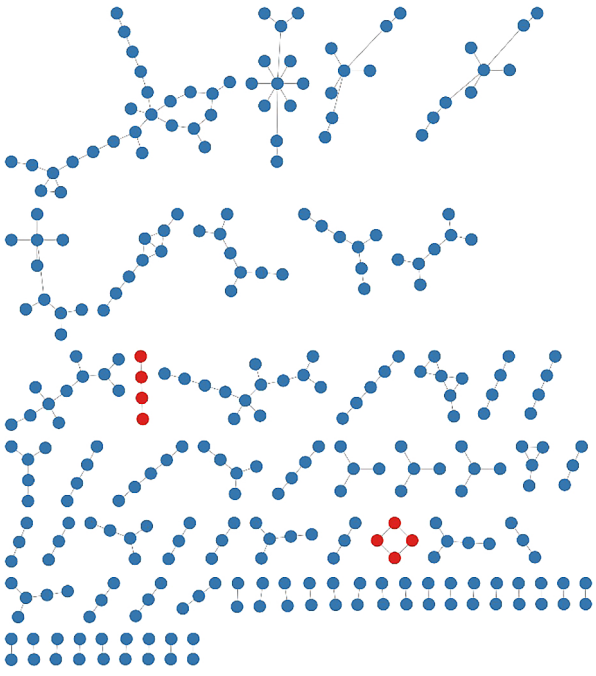


Fig. 2. A total of 293 isolated acute- and persistent-type microbial RNA viruses in sequence-based clusters. Colored circles indicate the lifestyle of each RNA virus: red, acute; blue, persistent. In total, 315 isolated microbial eukaryotic RNA viruses were collected from the manually curated Virus-Host database (Mihara *et al.*, 2016) downloaded on 2021.12.09. Sequences were clustered at 70% amino acid identity. Representative sequences were applied to a network analysis with MOCASSIN-prot (Keel *et al.*, 2018).

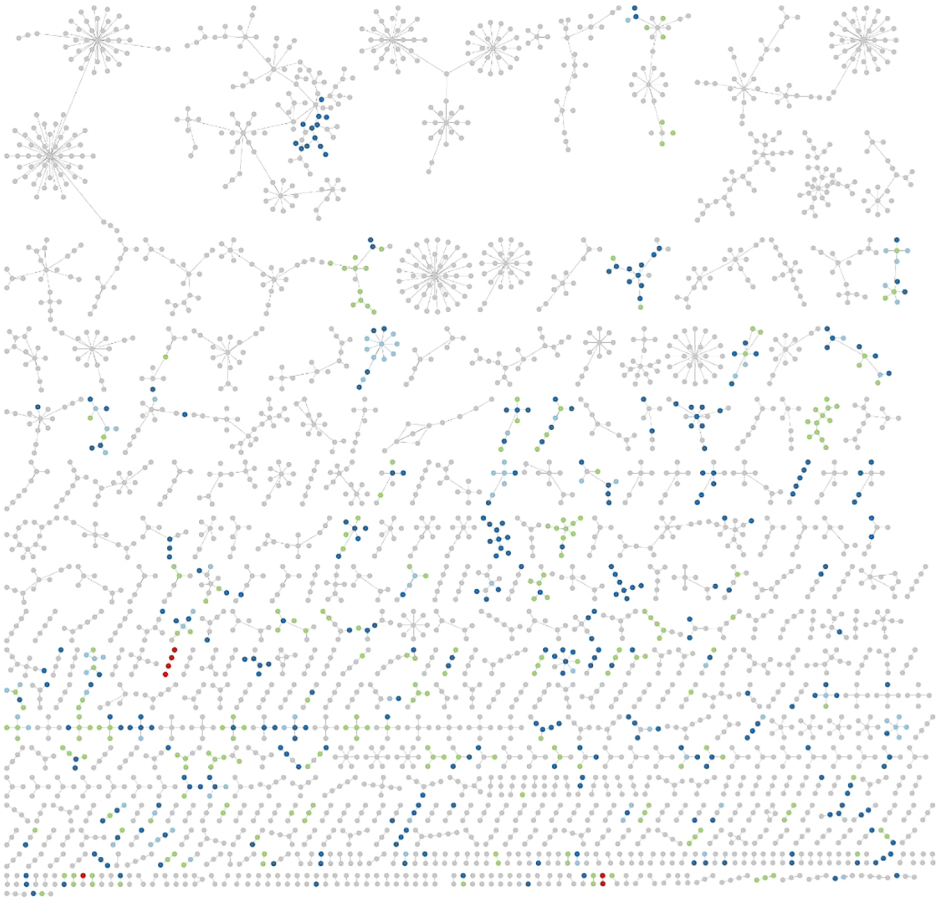


Fig. 3. RdRp sequences of known RNA viruses obtained from the Identical Protein Groups resource (https://www.ncbi.nlm.nih.gov/ipg) with keywords "rna dependent rna polymerase" and "viruses". After the removal of short (<200 aa) sequences, RdRp sequences were clustered at 70% using CD-HIT (Huang *et al.*, 2010). Representative sequences were applied to a network analysis with MOCASSIN-prot (Keel *et al.*, 2018). Colored circles indicate percent identity to persistent- or acute-type microbial RNA viruses: blue 100% to persistent; sky blue >70% to persistent; green >50% to persistent; red 100% to acute (>70 and 50% to acute were not identified).

the clusters (including five or more sequences) were predicted to include persistent-type viruses. Since the number of isolated persistent-type RNA viruses is limited, the true richness of persistent-type RNA viruses needs to be higher than that estimated.

Microbial RNA viruses in metagenomic data

Metagenomics is a powerful tool for identifying DNA viromes in nature (Breitbart *et al.*, 2002; Edwards and Rohwer, 2005; Suttle, 2016), and its use for RNA viromes has markedly increased (Culley *et al.*, 2006, 2010; Steward *et al.*, 2013; Wolf *et al.*, 2020; Edgar *et al.*, 2022; Zayed *et al.*, 2022). RNA virome research is conducted as either metagenomics (metatranscriptomics) of RNA virus particles or as cellular transcriptomics. RNA virus particles in extracellular environments are expected to predominantly be acute-type viruses, and cellular RNA may harbor viral RNA from both acute- and persistent-type viruses (Fig. 1). Accordingly, cellular RNA-specific viruses are expected to be of the persistent type. In addition, dsRNA-seq for cellular RNA has also been used to detect cellular viral RNA sequences in some studies in order to more efficiently retrieve RNA viral sequences (Decker and Parker, 2014; Decker *et al.*, 2019; Izumi *et al.*, 2021).

In aquatic environments, most of the acute-type RNA viruses identified from extracellular viral particles belong to specific RNA viral lineages, such as *Picornavirales* (ssRNA) and *Reoviridae* (dsRNA) (Culley *et al.*, 2006, 2014; Djikeng *et al.*, 2009; Steward *et al.*, 2013; Culley, 2017; Urayama *et al.*, 2018a; Wolf *et al.*, 2020). In addition, *Picobirnaviridae* (dsRNA) viruses that may infect bacterial hosts were identified as a dominant RNA virus population (Ghosh and Malik, 2021; Neri, U., *et al.* 2022. bioRxiv https://doi.org/10.1101/2022.02.15.480533.). The contribution of *Picornavirales* to viral lysis via their acute lifestyle and subsequent release of organic matter (previously defined as the virus shunt in DNA virus studies [Wilhelm and Suttle, 1999; Zimmerman *et al.*, 2020]) was suggested based on the correlation between the relative abundance of transcripts related to *Picornavirales* and the amount of particulate organic carbon in pelagic ecosystems (Kaneko *et al.*, 2021). These findings are also consistent with the results of culture-dependent isolation experiments as described above.

Based on pelagic microbial cellular RNA, in addition to *Picornavirales*, *Reoviridae*, and *Picobirnaviridae*, a wide range of RNA virus groups has been identified (Urayama *et al.*, 2016, 2018a). Among them, the dominant members belong to *Narnaviridae* (ssRNA) and *Partitiviridae*

(dsRNA), which have been recognized as persistent-type RNA viruses associated with fungi and plants/eukaryotic microorganisms, respectively (Hillman and Cai, 2013; Nibert *et al.*, 2014). Moreover, RNA virome studies have been conducted on macrocolonies of eukaryotic microorganisms, such as *Delisea pulchra* (red algae) and *Scytosiphon lomentaria* (brown algae). An association with *Totiviridae* (dsRNA) viral genomes, which are a persistent type, is generally observed; however, diverse RNA virus groups constitute the viromes (Lachnit *et al.*, 2016; Urayama *et al.*, 2016; Chiba *et al.*, 2020). In contrast, *Leviviridae* (ssRNA bacteriophage) and *Narnaviridae* virus operational taxonomic units (OTUs) have been identified as the predominant populations in soil RNA viromes, although the physical separation of cellular and viral particles has not been examined (Starr *et al.*, 2019).

We previously compared viral dsRNA in extracellular virus particle fractions and cellular dsRNA from surface seawater (Urayama *et al.*, 2018a). We showed 3.7- to 14.9-fold more dsRNA viral OTUs in cellular dsRNA metatranscriptomes than in extracellular dsRNA metatranscriptomes. Although we cannot rule out the possibility that some dsRNA virus particles were lost during sample processing, cellular dsRNA viruses, presumed to be of the persistent type, were abundant in the aquatic environment. In the present and related studies, we used a combination of fragmented and primer-ligated dsRNA sequencing (FLDS) and genome reconstruction *in silico* to identify viral sequences that are distinct from those in public databases. FLDS is a sequencing method that is applicable to long dsRNAs and enables the retrieval of the complete genomic sequences of both dsRNA and ssRNA viruses (Urayama *et al.*, 2016; Fukasawa *et al.*, 2020; Hirai *et al.*, 2021; Uehara-Ichiki *et al.*, 2021). In this technique, the terminal sequences of an RNA virus genome or genome segment may both be identified by sequence read mapping (Fig. 4). Therefore, it may be used to reconstruct multi-segmented RNA viral genomes based on terminally conserved sequences among segments (Fig. 5). In other words, a group of contigs sharing either/both terminal end sequences may represent a multi-segmented RNA virus genome. Moreover, if an RdRp gene or other virus proteins in any one of the segments in a reconstructed genome is identified based on sequence-dependent methods, all of the segments in the reconstructed genome may subsequently be identified as

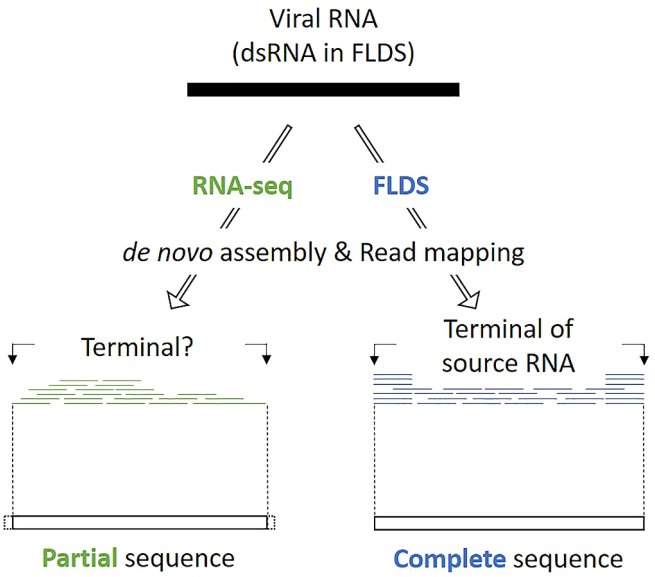


Fig. 4. Differences in conventional RNA-seq and FLDS for resultant contigs. In a *de novo* analysis, terminal sequence positions were not defined by RNA-seq data. However, FLDS data enabled us to identify terminal sequence positions because FLDS sequence reads included RACE (Rapid Amplification of cDNA Ends), which is widely used to assess the terminal sequences of RNA molecules (Urayama *et al.*, 2016).

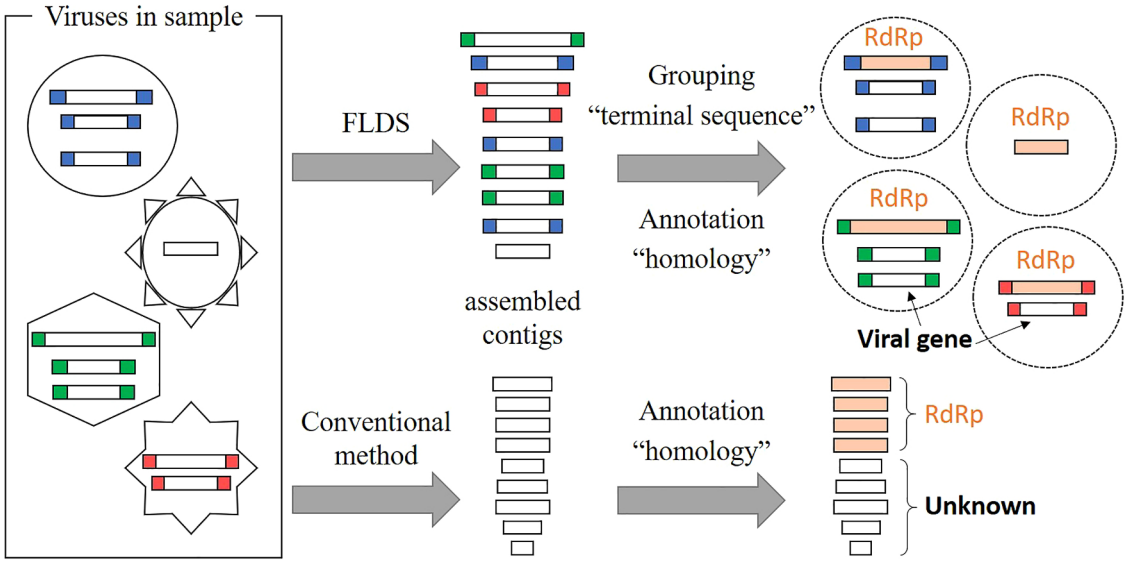


Fig. 5. Concept of RNA viral genome reconstruction based on conserved terminal sequences in segmented RNA virus genomes. Many segmented RNA viruses have conserved 5′- and 3′-terminal sequences (colored boxes). FLDS enabled us to obtain full-length RNA sequences, which are difficult to obtain with conventional RNA-seq technologies (Fig. 2). Based on terminal sequences, we reconstructed RNA viral genomes. If RdRp is identified in a potential RNA viral genome, we predict that other RNA sequences, which do not show significant similarity to known RNA viruses, will be segments of the RNA virus.

viral genome segments. We previously identified approximately 800 novel RNA viral genes that did not show significant (e-value less than 1×10^{-5}) similarities to known protein sequences (Urayama *et al.*, 2018a). Therefore, FLDS provides opportunities for deep surveys of RdRp and the identification of novel genes that are distinct from the viral genes in public databases.

Function of microbial RNA viruses in aquatic ecosystems

Knowledge on the ecological functions of acute-type microbial RNA viruses is based on the effects of host cell lysis. For example, previous studies suggested the contribution of RNA viruses to the virus shunt in pelagic zones (Kaneko *et al.*, 2021) and soil environments (Starr *et al.*, 2019). In addition, a sequence match between a CRISPR spacer and a bacterial RNA virus was reported (Wolf *et al.*, 2020). In contrast, their impacts on the evolution of microbial hosts have not yet been revealed. In the case of animals, RNA viral sequences incorporated into the host genome protect the host organism from related RNA viral infections (Suzuki *et al.*, 2020). However, a similar system has not been reported for microbes infected with RNA viruses. The ability to mediate horizontal gene transfer, which is often reported for DNA viruses, has also not been reported.

In contrast, persistent-type microbial RNA viruses have a wide range of effects through the manipulation of their host organisms. For example, fungi with an RNA virus may induce thermal tolerance to plants, while those without the RNA virus do not (Marquez *et al.*, 2007). Persistent-type RNA viruses also influence biological interactions between hosts and other organisms by changing secondary metabolites and pathogenicity (Chiba *et al.*, 2009; Zhang *et al.*, 2014; Aihara *et al.*, 2018; Okada *et al.*, 2018; Takahashi-Nakaguchi *et al.*, 2019; Ninomiya *et al.*, 2020). In some cases, persistent-type RNA viruses provide advantages to their hosts under specific conditions (Okada *et al.*, 2018). Therefore, these viruses may affect the adaptation of microbes to environmental challenges (Bao and Roossinck, 2013). Many of these are examples in fungi, which serve as model systems for persistent-type RNA viral research; however, their functions in aquatic environments remain largely unknown. Nevertheless, the host-virus relationship and the impact of persistent-type RNA viral infection on the physiology of their hosts in terrestrial ecosystems imply that persistent-type RNA viruses in aquatic ecosystems also influence host physiology and subsequent ecological functions, including the niche adaptation of their host organisms.

Climate change due to global warming and ocean acidification, is an important issue for our planet. These environmental changes are expected to affect pelagic microbial ecosystems, and other impacts on marine microbial ecosystems and subsequent biogeochemical cycles may also accelerate environmental changes. Therefore, observations of the marine microbial community are essential for understanding microbial responses to a changing ocean. Furthermore, current updates in our understanding of the RNA virome suggest that acute- and persistent-type microbial RNA viruses both play a significant role in biogeochemical cycles via the viral shunt and the regulation of host physiology. Accordingly, we need to pay close attention to marine RNA viromes in the changing ocean even though pelagic RNA viromes have so far been overlooked in marine microbial ecology.

Acknowledgements

This study was supported by a grant from the Institute for Fermentation, Osaka, by JSPS KAKENHI (Grant No. 18K19358, 20K20377, 20H05579, 21K18217, 21K20567, 22H04879, and 22J40078), and by Grants-in-Aid for Scientific Research on Innovative Areas from the Ministry of Education, Culture, Science, Sports, and Technology (MEXT) of Japan (Grant Nos. 16H06429, 16K21723, and 16H06437).

References

Aihara, M., Urayama, S., Le, M.T., Katoh, Y., Higashiura, T., Fukuhara, T., *et al.* (2018) Infection by Magnaporthe oryzae chrysovirus 1 strain A triggers reduced virulence and pathogenic race conversion of its host fungus, Magnaporthe oryzae. *J Gen Plant Pathol* **84**: 92–103.

Banks, G.T., Buck, K.W., Chain, E.B., Darbyshire, J.E., and Himmelweit, F. (1969) Virus-like particles in penicillin producing strains of Penicillium chrysogenum. *Nature* **222**: 89–90.

Bao, X., and Roossinck, M.J. (2013) Multiplexed interactions: viruses of endophytic fungi. In *Advances in Virus Research*. Amsterdam: Elsevier, pp. 37–58.

Beijerinck, M.W. (1898) Concerning a contagium vivum fluidum as cause of the spot disease of tobacco leaves. *Phytopathol Classics* **7**: 33–52.

Bian, R., Andika, I.B., Pang, T., Lian, Z., Wei, S., Niu, E., *et al.* (2020) Facilitative and synergistic interactions between fungal and plant viruses. *Proc Natl Acad Sci U S A* **117**: 3779–3788.

Botella, L., Janoušek, J., Maia, C., Jung, M.H., Raco, M., and Jung, T. (2020) Marine oomycetes of the genus Halophytophthora harbor viruses related to bunyaviruses. *Front Microbiol* **11**: 1467.

Breitbart, M., Salamon, P., Andresen, B., Mahaffy, J.M., Segall, A.M., Mead, D., *et al.* (2002) Genomic analysis of uncultured marine viral communities. *Proc Natl Acad Sci U S A* **99**: 14250–14255.

Brussaard, C.P., Noordeloos, A.A., Sandaa, R.A., Heldal, M., and Bratbak, G. (2004) Discovery of a dsRNA virus infecting the marine photosynthetic protist Micromonas pusilla. *Virology* **319**: 280–291.

Cai, G., Krychiw, J.F., Myers, K., Fry, W.E., and Hillman, B.I. (2013) A new virus from the plant pathogenic oomycete Phytophthora infestans with an 8 kb dsRNA genome: the sixth member of a proposed new virus genus. *Virology* **435**: 341–349.

Callanan, J., Stockdale, S.R., Shkoporov, A., Draper, L.A., Ross, R.P., and Hill, C. (2020) Expansion of known ssRNA phage genomes: from tens to over a thousand. *Sci Adv* **6**: eaay5981.

Chiba, S., Salaipeth, L., Lin, Y.H., Sasaki, A., Kanematsu, S., and Suzuki, N. (2009) A novel bipartite double-stranded RNA Mycovirus from the white root rot Fungus Rosellinia necatrix: molecular and biological characterization, taxonomic considerations, and potential for biological control. *J Virol* **83**: 12801–12812.

Chiba, Y., Tomaru, Y., Shimabukuro, H., Kimura, K., Hirai, M., Takaki, Y., *et al.* (2020) Viral RNA genomes identified from marine macroalgae and a diatom. *Microbes Environ* **35**: ME20016.

Culley, A. (2017) New insight into the RNA aquatic virosphere via viromics. *Virus Res* **244**: 84–89.

Culley, A.I., Lang, A.S., and Suttle, C.A. (2006) Metagenomic analysis of coastal RNA virus communities. *Science* **312**: 1795–1798.

Culley, A.I., Suttle, C.A., and Steward, G.F. (2010) Characterization of the diversity of marine RNA viruses. In *Manual of Aquatic Viral Ecology*. Wilheim, S., Weinbauer, M., and Suttle, C. (eds). Waco, TX: American Society of Limnology and Oceanography (ASLO), pp. 193–201.

Culley, A.I., Mueller, J.A., Belcaid, M., Wood-Charlson, E.M., Poisson, G., and Steward, G.F. (2014) The characterization of RNA viruses in tropical seawater using targeted PCR and metagenomics. *mBio* **5**: e01210–01214.

Decker, C.J., and Parker, R. (2014) Analysis of double-stranded RNA from microbial communities identifies double-stranded RNA virus-like elements. *Cell Rep* **7**: 898–906.

Decker, C.J., Steiner, H.R., Hoon-Hanks, L.L., Morrison, J.H., Haist, K.C., Stabell, A.C., *et al.* (2019) dsRNA-Seq: identification of viral infection by purifying and sequencing dsRNA. *Viruses* **11**: 943.

Djikeng, A., Kuzmickas, R., Anderson, N.G., and Spiro, D.J. (2009) Metagenomic analysis of RNA viruses in a fresh water lake. *PLoS One* **4**: e7264.

Dolja, V.V., and Koonin, E.V. (2012) Capsid-less RNA viruses. eLS. URL https://doi.org/10.1002/9780470015902.a0023269

Edgar, R.C., Taylor, J., Lin, V., Altman, T., Barbera, P., Meleshko, D., *et al.* (2022) Petabase-scale sequence alignment catalyses viral discovery. *Nature* **602**: 142–147.

Edwards, R.A., and Rohwer, F. (2005) Viral metagenomics. *Nat Rev Microbiol* **3**: 504–510.

Fukasawa, F., Hirai, M., Takaki, Y., Shimane, Y., Thomas, C.E., Urayama, S., *et al.* (2020) A new polycipivirus identified in Colobopsis shohki. *Arch Virol* **165**: 761–763.

Ghabrial, S.A., Castón, J.R., Jiang, D., Nibert, M.L., and Suzuki, N. (2015) 50-plus years of fungal viruses. *Virology* **479**: 356–368.

Ghosh, S., and Malik, Y.S. (2021) The true host/s of picobirnaviruses. *Front Vet Sci* **7**: 615293.

Goodman, R.P., Ghabrial, S.A., Fichorova, R.N., and Nibert, M.L. (2011) Trichomonasvirus: a new genus of protozoan viruses in the family Totiviridae. *Arch Virol* **156**: 171–179.

Hillman, B.I., and Cai, G. (2013) The family Narnaviridae: simplest of RNA viruses. *Adv Virus Res* **86**: 149–176.

Hirai, M., Takaki, Y., Kondo, F., Horie, M., Urayama, S., and Nunoura, T. (2021) RNA viral metagenome analysis of subnanogram dsRNA using fragmented and primer ligated dsRNA sequencing (FLDS). *Microbes Environ* **36**: ME20152.

Huang, Y., Niu, B., Gao, Y., Fu, L., and Li, W. (2010) CD-HIT Suite: a web server for clustering and comparing biological sequences. *Bioinformatics* **26**: 680–682.

Izumi, T., Morioka, Y., Urayama, S., Motooka, D., Tamura, T., Kawagishi, T., *et al.* (2021) DsRNA sequencing for RNA virus surveillance using human clinical samples. *Viruses* **13**: 1310.

Kaneko, H., Blanc-Mathieu, R., Endo, H., Chaffron, S., Delmont, T.O., Gaia, M., *et al.* (2021) Eukaryotic virus composition can predict the efficiency of carbon export in the global ocean. *iScience* **24**: 102002.

Keel, B.N., Deng, B., and Moriyama, E.N. (2018) MOCASSIN-prot: a multi-objective clustering approach for protein similarity networks. *Bioinformatics* **34**: 1270–1277.

Khoshnan, A., and Alderete, J. (1994) Trichomonas vaginalis with a double-stranded RNA virus has upregulated levels of phenotypically variable immunogen mRNA. *J Virol* **68**: 4035–4038.

King, A.M.Q., Adams, M.J., Carstens, E.B., and Lefkowitz, E.J. (2012) *Virus Taxonomy: Classification and Nomenclature of Viruses: Ninth Report of the International Committee on Taxonomy of Viruses*. Amsterdam: Elsevier.

Koonin, E.V., Dolja, V.V., and Krupovic, M. (2015) Origins and evolution of viruses of eukaryotes: The ultimate modularity. *Virology* **479–480**: 2–25.

Koonin, E.V., Dolja, V.V., Krupovic, M., Varsani, A., Wolf, Y.I., Yutin, N., *et al.* (2020) Global organization and proposed megataxonomy of the virus world. *Microbiol Mol Biol Rev* **84**: e00061-00019.

Kotta-Loizou, I., and Coutts, R.H. (2017) Mycoviruses in Aspergilli: a comprehensive review. *Front Microbiol* **8**: 1699.

Lachnit, T., Thomas, T., and Steinberg, P. (2016) Expanding our understanding of the seaweed holobiont: RNA viruses of the red alga Delisea pulchra. *Front Microbiol* **6**: 1489.

Lang, A.S., Culley, A.I., and Suttle, C.A. (2004) Genome sequence and characterization of a virus (HaRNAV) related to picorna-like viruses that infects the marine toxic bloom-forming alga Heterosigma akashiwo. *Virology* **320**: 206–217.

Lang, A.S., Vlok, M., Culley, A.I., Suttle, C.A., Takao, Y., Tomaru, Y., and ICTV Report Consortium (2021) ICTV virus taxonomy profile: Marnaviridae 2021. *J Gen Virol* **102**: 001633.

Liu, Y.C., Linder-Basso, D., Hillman, B.I., Kaneko, S., and Milgroom, M.G. (2003) Evidence for interspecies transmission of viruses in natural populations of filamentous fungi in the genus Cryphonectria. *Mol Ecol* **12**: 1619–1628.

Loeffler, F., and Frosch, P. (1898) Report of the commission for research on foot-and-mouth disease. *Zent Bakt Parasitkde Abt I* **23**: 371–391.

Marquez, L.M., Redman, R.S., Rodriguez, R.J., and Roossinck, M.J. (2007) A virus in a fungus in a plant: three-way symbiosis required for thermal tolerance. *Science* **315**: 513–515.

Márquez, L.M., and Roossinck, M.J. (2012) Do persistent RNA viruses fit the trade-off hypothesis of virulence evolution? *Curr Opin Virol* **2**: 556–560.

Mihara, T., Nishimura, Y., Shimizu, Y., Nishiyama, H., Yoshikawa, G., Uehara, H., *et al.* (2016) Linking virus genomes with host taxonomy. *Viruses* **8**: 66.

Morris, T., and Dodds, J. (1979) Isolation and analysis of double-stranded RNA from virus-infected plant and fungal tissue. *Phytopathology* **69**: 854–858.

Nasir, A., Forterre, P., Kim, K.M., and Caetano-Anollés, G. (2014) The distribution and impact of viral lineages in domains of life. *Front Microbiol* **5**: 194.

Nerva, L., Ciuffo, M., Vallino, M., Margaria, P., Varese, G.C., Gnavi, G., and Turina, M. (2016) Multiple approaches for the detection and characterization of viral and plasmid symbionts from a collection of marine fungi. *Virus Res* **219**: 22–38.

Nibert, M.L., Ghabrial, S.A., Maiss, E., Lesker, T., Vainio, E.J., Jiang, D., and Suzuki, N. (2014) Taxonomic reorganization of family Partitiviridae and other recent progress in partitivirus research. *Virus Res* **188**: 128–141.

Ninomiya, A., Urayama, S., Suo, R., Itoi, S., Fuji, S.I., Moriyama, H., and Hagiwara, D. (2020) Mycovirus-induced tenuazonic acid production in a rice blast fungus Magnaporthe oryzae. *Front Microbiol* **11**: 1641.

Nuss, D.L. (2005) Hypovirulence: mycoviruses at the fungal-plant interface. *Nat Rev Microbiol* **3**: 632–642.

Obeng, N., Pratama, A.A., and Elsas, J.D.V. (2016) The significance of mutualistic phages for bacterial ecology and evolution. *Trends Microbiol* **24**: 440–449.

Okada, R., Ichinose, S., Takeshita, K., Urayama, S., Fukuhara, T., Komatsu, K., *et al.* (2018) Molecular characterization of a novel mycovirus in Alternaria alternata manifesting two-sided effects: Down-regulation of host growth and up-regulation of host plant pathogenicity. *Virology* **519**: 23–32.

Pratama, A.A., and van Elsas, J.D. (2018) The 'neglected' soil virome—Potential role and impact. *Trends Microbiol* **26**: 649–662.

Reed, W., and Carroll, J. (1901) The prevention of yellow fever. *Public Health Pap Rep* **27**: 113–129.

Roossinck, M.J. (2010) Lifestyles of plant viruses. *Philos Trans R Soc Lond B Biol Sci* **365**: 1899–1905.

Roossinck, M.J., Saha, P., Wiley, G.B., Quan, J., White, J.D., Lai, H., *et al.* (2010) Ecogenomics: using massively parallel pyrosequencing to understand virus ecology. *Mol Ecol* **19** Suppl 1: 81–88.

Roossinck, M.J. (2011) The good viruses: viral mutualistic symbioses. *Nat Rev Microbiol* **9**: 99–108.

Roossinck, M.J. (2012) Persistent plant viruses: Molecular hitchhikers or epigenetic elements? In *Viruses: Essential Agents of Life*. Witzany, G. (ed.) Dordrecht: Springer Netherlands, pp. 177–186.

Roossinck, M.J. (2015) Metagenomics of plant and fungal viruses reveals an abundance of persistent lifestyles. *Front Microbiol* **5**: 767.

Roossinck, M.J. (2019) Evolutionary and ecological links between plant and fungal viruses. *New Phytol* **221**: 86–92.

Sadeghi, M., Tomaru, Y., and Ahola, T. (2021) RNA viruses in aquatic unicellular eukaryotes. *Viruses* **13**: 362.

Schmitt, M.J., and Breinig, F. (2002) The viral killer system in yeast: from molecular biology to application. *FEMS Microbiol Rev* **26**: 257–276.

Shi, M., Lin, X.D., Tian, J.H., Chen, L.J., Chen, X., Li, C.X., *et al.* (2016) Redefining the invertebrate RNA virosphere. *Nature* **540**: 539–543.

Starr, E.P., Nuccio, E.E., Pett-Ridge, J., Banfield, J.F., and Firestone, M.K. (2019) Metatranscriptomic reconstruction reveals RNA viruses with the potential to shape carbon cycling in soil. *Proc Natl Acad Sci U S A* **116**: 25900–25908.

Steward, G.F., Culley, A.I., Mueller, J.A., Wood-Charlson, E.M., Belcaid, M., and Poisson, G. (2013) Are we missing half of the viruses in the ocean? *ISME J* **7**: 672–679.

Stuart, K.D., Weeks, R., Guilbride, L., and Myler, P.J. (1992) Molecular organization of Leishmania RNA virus 1. *Proc Natl Acad Sci U S A* **89**: 8596–8600.

Suttle, C.A. (2016) Environmental microbiology: Viral diversity on the global stage. *Nat Microbiol* **1**: 16205.

Suzuki, Y., Baidaliuk, A., Miesen, P., Frangeul, L., Crist, A.B., Merkling, S.H., *et al.* (2020) Non-retroviral endogenous viral element limits cognate virus replication in Aedes aegypti ovaries. *Curr Biol* **30**: 3495–3506.e6.

Tai, V., Lawrence, J.E., Lang, A.S., Chan, A.M., Culley, A.I., and Suttle, C.A. (2003) Characterization of HaRNAV, a single-stranded RNA virus causing lysis of Heterosigma akashiwo (raphidophyceae) 1. *J Phycol* **39**: 343–352.

Takahashi-Nakaguchi, A., Shishido, E., Yahara, M., Urayama, S., Sakai, K., Chibana, H., *et al.* (2019) Analysis of an intrinsic mycovirus associated with reduced virulence of the human pathogenic fungus Aspergillus fumigatus. *Front Microbiol* **10**: 3045.

Takao, Y., Nagasaki, K., Mise, K., Okuno, T., and Honda, D. (2005) Isolation and characterization of a novel single-stranded RNA virus infectious to a marine fungoid protist, Schizochytrium sp. (Thraustochytriaceae, Labyrinthulea). *Appl Environ Microbiol* **71**: 4516–4522.

Takao, Y., Mise, K., Nagasaki, K., Okuno, T., and Honda, D. (2006) Complete nucleotide sequence and genome organization of a single-stranded RNA virus infecting the marine fungoid protist Schizochytrium sp. *J Gen Virol* **87**: 723–733.

Tomaru, Y., Katanozaka, N., Nishida, K., Shirai, Y., Tarutani, K., Yamaguchi, M., and Nagasaki, K. (2004) Isolation and characterization of two distinct types of HcRNAV, a single-stranded RNA virus infecting the bivalve-killing microalga Heterocapsa circularisquama. *Aquat Microb Ecol* **34**: 207–218.

Torrecilhas, A.C., Soares, R.P., Schenkman, S., Fernandez-Prada, C., and Olivier, M. (2020) Extracellular vesicles in trypanosomatids: host cell communication. *Front Cell Infect Microbiol* **10**: 602502.

Uchida, K., Sakuta, K., Ito, A., Takahashi, Y., Katayama, Y., Omatsu, T., *et al.* (2021) Two novel endornaviruses co-infecting a phytophthora pathogen of Asparagus officinalis modulate the developmental stages and fungicide sensitivities of the host oomycete. *Front Microbiol* **12**: 122.

Uehara-Ichiki, T., Urayama, S., Hirai, M., Takaki, Y., Nunoura, T., Fujinaga, M., and Hanada, K. (2021) Complete genome sequence of Sikte (Sitke) waterborne virus, a member of the genus Tombusvirus. *Arch Virol* **166**: 991–994.

Urayama, S., Takaki, Y., and Nunoura, T. (2016) FLDS: A comprehensive dsRNA sequencing method for intracellular RNA virus surveillance. *Microbes Environ* **31**: 33–40.

Urayama, S., Takaki, Y., Nishi, S., Yoshida-Takashima, Y., Deguchi, S., Takai, K., and Nunoura, T. (2018a) Unveiling the RNA virosphere associated with marine microorganisms. *Mol Ecol Resour* **18**: 1444–1455.

Urayama, S., Takaki, Y., Nunoura, T., and Miyamoto, N. (2018b) Complete genome sequence of a novel RNA virus identified from a deep-sea animal, Osedax japonicus. *Microbes Environ* **33**: 446–449.

Urayama, S., Doi, N., Kondo, F., Chiba, Y., Takaki, Y., Hirai, M., *et al.* (2020a) Diverged and active partitiviruses in lichen. *Front Microbiol* **11**: 561344.

Urayama, S., Takaki, Y., Hagiwara, D., and Nunoura, T. (2020b) dsRNA-seq reveals novel RNA virus and virus-like putative complete genome sequences from Hymeniacidon sp. sponge. *Microbes Environ* **35**: ME19132.

Vainio, E.J., Hakanpää, J., Dai, Y.-C., Hansen, E., Korhonen, K., and Hantula, J. (2011) Species of Heterobasidion host a diverse pool of partitiviruses with global distribution and interspecies transmission. *Fungal Biol* **115**: 1234–1243.

van der Lende, T.R., Harmsen, M.C., and Wessels, J.G. (1994) Double-stranded RNAs and proteins associated with the 34 nm virus particles of the cultivated mushroom Agaricus bisporus. *J Gen Virol* **75**: 2533–2536.

Wang, A.L., and Wang, C.C. (1986) Discovery of a specific double-stranded RNA virus in Giardia lamblia. *Mol Biochem Parasitol* **21**: 269–276.

Wickner, R.B. (1996) Double-stranded RNA viruses of Saccharomyces cerevisiae. *Microbiol Rev* **60**: 250–265.

Wilhelm, S.W., and Suttle, C.A. (1999) Viruses and nutrient cycles in the sea: viruses play critical roles in the structure and function of aquatic food webs. *Bioscience* **49**: 781–788.

Wolf, Y.I., Silas, S., Wang, Y., Wu, S., Bocek, M., Kazlauskas, D., *et al.* (2020) Doubling of the known set of RNA viruses by metagenomic analysis of an aquatic virome. *Nat Microbiol* **5**: 1262–1270.

Xu, P., Chen, F., Mannas, J.P., Feldman, T., Sumner, L.W., and Roossinck, M.J. (2008) Virus infection improves drought tolerance. *New Phytol* **180**: 911–921.

Zayed, A.A., Wainaina, J.M., Dominguez-Huerta, G., Pelletier, E., Guo, J., Mohssen, M., *et al.* (2022) Cryptic and abundant marine viruses at the evolutionary origins of Earth's RNA virome. *Science* **376**: 156–162.

Zhang, R., Liu, S., Chiba, S., Kondo, H., Kanematsu, S., and Suzuki, N. (2014) A novel single-stranded RNA virus isolated from a phytopathogenic filamentous fungus, Rosellinia necatrix, with similarity to hypo-like viruses. *Front Microbiol* **5**: 360.

Zimmerman, A.E., Howard-Varona, C., Needham, D.M., John, S.G., Worden, A.Z., Sullivan, M.B., *et al.* (2020) Metabolic and biogeochemical consequences of viral infection in aquatic ecosystems. *Nat Rev Microbiol* **18**: 21–34.

Microbes Environ. Vol. 32, No. 4, 300-313, 2017
https://www.jstage.jst.go.jp/browse/jsme2 doi:10.1264/jsme2.ME17017

MOST CITED REVIEW PAPER PUBLISHED
IN THE LAST TEN YEARS [2015-2024] (382 citations)

Minireview

Microbial Ecology along the Gastrointestinal Tract

Ethan T. Hillman[1], Hang Lu[2], Tianming Yao[3], and Cindy H. Nakatsu[4]*

[1]*Department of Agricultural and Biological Engineering, Purdue University, West Lafayette, Indiana 47907, USA;* [2]*Department of Animal Science, Purdue University, West Lafayette, Indiana 47907, USA;* [3]*Department of Food Science, Purdue University, West Lafayette, Indiana 47907, USA; and* [4]*Department of Agronomy, Purdue University, West Lafayette, Indiana 47907, USA*

(Received January 31, 2017—Accepted August 19, 2017—Published online November 10, 2017)

The ecosystem of the human gastrointestinal (GI) tract traverses a number of environmental, chemical, and physical conditions because it runs from the oral cavity to the anus. These differences in conditions along with food or other ingested substrates affect the composition and density of the microbiota as well as their functional roles by selecting those that are the most suitable for that environment. Previous studies have mostly focused on *Bacteria*, with the number of studies conducted on *Archaea*, *Eukarya*, and *Viruses* being limited despite their important roles in this ecosystem. Furthermore, due to the challenges associated with collecting samples directly from the inside of humans, many studies are still exploratory, with a primary focus on the composition of microbiomes. Thus, mechanistic studies to investigate functions are conducted using animal models. However, differences in physiology and microbiomes need to be clarified in order to aid in the translation of animal model findings into the context of humans. This review will highlight *Bacteria*, *Archaea*, *Fungi*, and *Viruses*, discuss differences along the GI tract of healthy humans, and perform comparisons with three common animal models: rats, mice, and pigs.

Key words: Microbiome, mycobiome, virome, human gastrointestinal (GI) tract, animal models, diet

Researchers have been investigating the ecology of the intestinal microbiota for decades (120, 165) in order to identify, characterize, and count their numbers. These extensive efforts are due to the important roles the intestinal microbiota play in digestion, the production of essential vitamins, and protection of the gastrointestinal (GI) tract from pathogen colonization (141). In the past few decades, molecular techniques targeting the 16S rRNA gene and other genetic markers have been developed to characterize and analyze bacterial communities. These methods have been used to reveal the important roles played by microbes in the GI tract (23, 180, 183, 184, 189, 212). In healthy individuals, the microbiome (microbial community) and host have a mutualistic relationship in which both partners benefit; however, pathogens may invade and cause disease under certain conditions. The initial aim of most studies was to elucidate the role of the microbiome in disease. More recently, surveys have been performed on healthy individuals in order to assess the contribution of the microbiota to health, particularly in response to dietary changes/supplementation with probiotics and/or prebiotics.

The human GI tract is a complex system that starts from the oral cavity, continues through the stomach and intestines, and finally ends at the anus (Fig. 1). The density and composition of the microbiome change along the GI tract, with major populations being selected by the functions performed at the various locations. Bacteria along the GI tract have several possible functions, many of which are beneficial for health including vitamin production, the absorption of ions (Ca, Mg, and Fe), protection against pathogens, histological development, enhancement of the immune system, and the fermentation of "non-digestible foods" to short chain fatty acids (SCFA) and other metabolites (19, 58, 63, 77, 138). The roles of fungi and viruses have not been examined in as much detail; however, they are known to play important roles in microbiota dynamics and host physiology/immunity related to health and disease (45, 94, 133).

Food passes through the GI tract and the absorption rate of nutrients is largely dependent on the activities of various enzymes in the digestive system, such as amylase in saliva, pepsin in the stomach, and pancreatic enzymes in the small intestine. These mechanisms have been extensively examined (61, 62), particularly in the stomach. However, many food components cannot be digested in the upper GI tract and are passed into the lower intestinal tract, in which they are fermented by microbes. Functional studies commonly use animal models in order to obtain a better understanding of the processes in the GI tract that may lead to better health or decrease disease. However, information from animal models may not be directly translatable to humans. Therefore, researchers need to consider the limitations of the selected animal model when extrapolating findings to humans.

Although microbiome studies often include an ecological component, most of the research performed to date has focused on *Bacteria* and not all of the biota. This represents a logical approach because *Bacteria* comprise most of the microbiome. However, even biota representing a small proportion of the microbiome may play important roles in the ecosystem (133). Therefore, researchers need to start shifting their approach to include eukaryotic, prokaryotic, and viral (33, 133) interactions in efforts to elucidate the roles of all components of the microbiome.

In recent years, a number of reviews have summarized findings from the increasing number of studies being performed in this field (36, 73, 176, 188). While most studies have focused on disease, the microbiome is also important for maintaining

* Corresponding author. E-mail: cnakatsu@purdue.edu;
Tel: +1 (765) 496–2997; Fax: +1 (765) 496–2926.

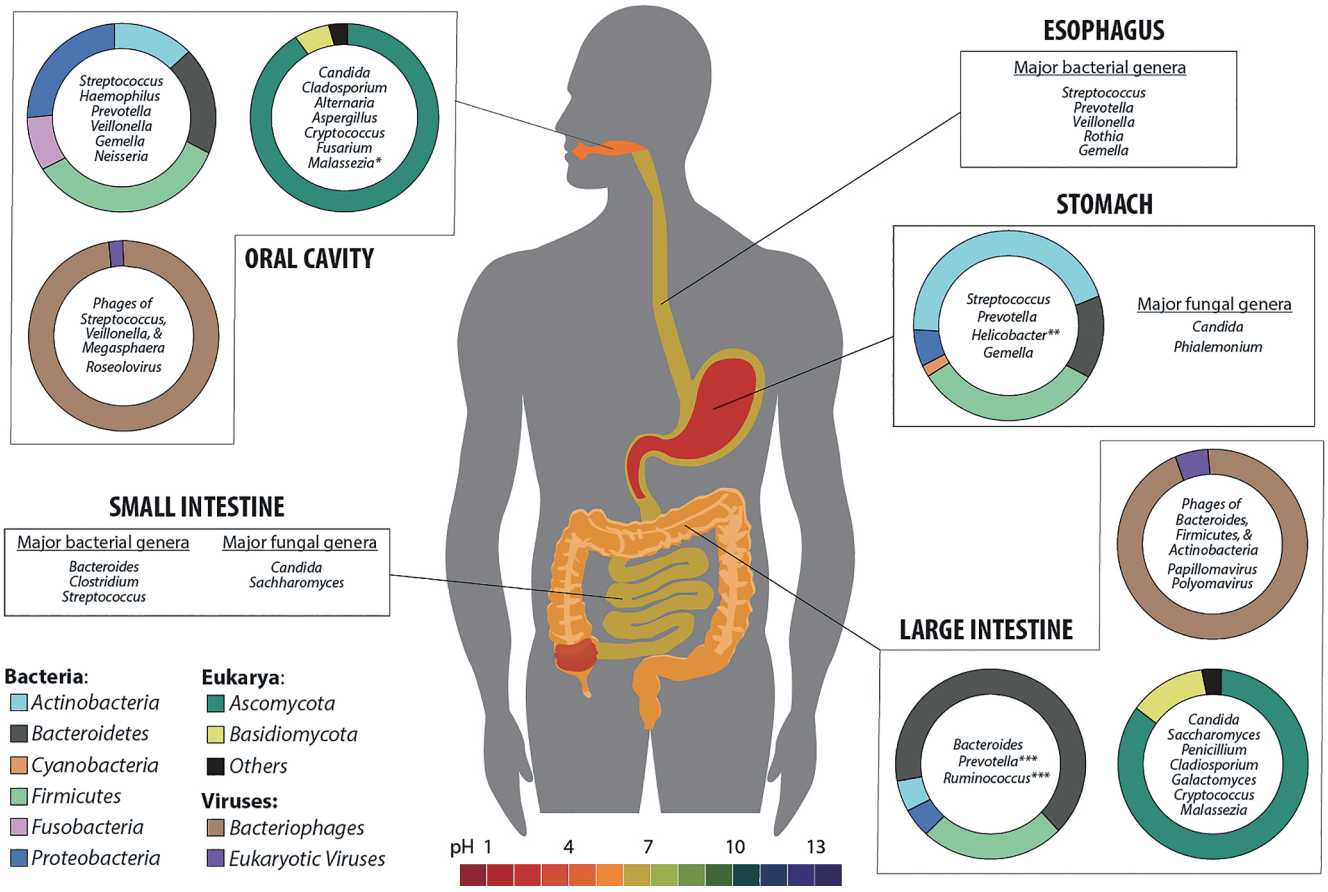


Fig. 1. Microbiome composition of *Bacteria* (1, 5, 20, 21, 43, 147, 156, 223), *Eukarya* (52, 85, 114, 126, 182, 197), and *Viruses* (45, 134, 151, 215) among the physiological niches of the human gastrointestinal (GI) tract. Phylum level compositional data are presented where available along with the most common genera in each GI tract location. The colors on the doughnut plots correspond to the legend in the lower left corner; the GI tract is colored according to the pH scale shown at the bottom of Fig. 1. (* *Malassezia* was very abundant in one study and was not detected in another study. ** The abundance of *Helicobacter* may vary greatly between individuals. *** Proportions of these and other colon genera vary with age, diet, & geographical location.)

health. We herein highlight differences in the microbiome (*Bacteria*, *Archaea*, *Fungi*, and *Viruses*) along the GI tract of healthy humans, and how it compares to those of typical animal models used in research. One finding that is consistent to most studies is that the microbiome of healthy individuals is unique; however, there are still some generalities that will be discussed in this review.

Microbiome diversity

Many factors contribute to the diversity of microbiomes, and most studies have demonstrated the individuality of microbiomes among subjects. Previous findings support microbial communities being more similar in subjects that are genetically related (191), of a similar age (135, 213), or with common diets (including the influences of ethnicity and geography) (63). Diseases will also have an impact on microbiome diversity, including autoimmune and neoplastic diseases, such as inflammatory bowel disease, diabetes, obesity, cardiovascular diseases, allergies, and cancer (37, 121). Treatments for diseases may also affect a patient's gut microbiota, and the consequences of antibiotic use have been intensively investigated (22, 95).

The host genotype has been shown to influence the development of the gut microbiota, and the immune system has been identified as a contributing factor (188). Crosstalk between the microbiome and human immune system occurs in response to a number of environmental factors, such as diet, xenobiotics, and pathogens. Microbial host interactions occur in the gut, mainly in the epithelial cell layer, myeloid cells, and innate lymphoid cells, in which crosstalk and feedback loops contribute to the microbiome composition, host physiology, and disease susceptibility. These interactions contribute not only to the bacterial community along the GI tract, but also to the other microbiota (*Fungi*, *Archaea*, and *Viruses*). Our understanding of the immunology associated with *Fungi* (150) and *Archaea* is currently limited. Transkingdom commensal relationships among microbiota (including *Viruses*) are considered to form from infancy (29, 30, 106, 200) and several co-occurring relationships have been identified (35, 75, 76, 85, 214).

Bacteria. A more complete picture of human-associated bacterial communities obtained using molecular techniques has revealed that their diversity is greater than initially considered through cultivation (9, 20, 56, 90, 113). Using almost full-length 16S rRNA gene sequences, predicted taxa numbers range from 100–300 (20, 56), while pyrosequencing suggests there are 1000s of phylotypes (38, 49). Most of the

gut bacteria identified by 16S rRNA gene sequencing belong to the five phyla originally identified by cultivation, namely, *Bacteroidetes*, *Firmicutes*, *Actinobacteria*, *Proteobacteria*, and *Verrucomicrobia* (90), and, at lower proportions, *Fusobacteria*, *Tenericutes*, *Spirochaetes*, *Cyanobacteria*, and TM7 (189). At lower levels of the taxonomic classification, microbiome compositions vary with each individual. Attempts have been made to identify a single core microbiome of *Bacteria* in the GI tract. Although this has not been possible in the lower GI tract (mainly using fecal samples) based on taxonomy, it appears there are core microbial functions (152, 189, 191). It is possible to identify some core microbiota in the oral cavity, esophagus, and stomach (148). Although extensive efforts have been made to cultivate representative gut microbiota in an attempt to gain a better understanding of the relationship between taxa and function (156), there are still many undescribed taxa with unknown functional roles in the gut.

As the price of sequencing decreases, it is becoming more common to use a metagenomic approach that provides information on all microbiota and potential functions (3, 70, 167, 189). This provides a means to go beyond *Bacteria* and obtain information on eukaryotic microbes (mainly fungi) and viruses. Although *Fungi*, *Archaea*, and *Viruses* in the microbiome are a part of the 'rare biosphere' (organisms that comprise <0.1% of the microbiome) (173), they still have a significant impact on host health.

Fungi

Fungi are considered to comprise approximately 0.03% of the fecal microbiome (143); making them approximately 3,300-fold less abundant than *Bacteria*. Fungal diversity in the human gut is also lower than that of *Bacteria* (143, 166), although more taxa are being found as the number of individuals being studied using next generation sequencing is increasing (44, 126, 166, 182). In 2015, a review of 36 fungal gut microbiome studies revealed that there have been at least 267 distinct fungi identified in the human gut (181), while another study reported 221 (72). Despite the number of taxa that have been reported, most fungi are highly variable among individuals, with few appearing to be common to all.

Cultivation-based analyses have typically identified *Candida* as the most common fungal genus (166), and it is also frequently identified using non-cultivation-based methods, whereas the other taxa identified have been variable, which may be because of the analytical method used and/or subject variability. For example, 66 genera of fungi were found using pyrosequencing when 98 individuals were examined, with the genera *Saccharomyces*, *Candida*, and *Cladosporium* being the most prevalent (85). *Mucor* was common in Spanish individuals (126) and the most common fungi in 16 vegetarians were *Fusarium*, *Malassezia*, *Penicillium*, and *Aspergillus* (182). These studies suggested that some taxa, *e.g.*, *Penicillium* and *Aspergillus*, are not resident in the gut and enter through environmental sources, such as food and water, in which they are commonly found. This may account for some of the variability in taxa reported in various studies and for the increasing number of fungi being identified as more studies are being performed, even those based on cultivation (71). Under certain conditions, some fungi may flourish and become pathogenic including *Candida*, *Aspergillus*, *Fusarium*, and *Cryptococcus* (44, 84, 140, 143). More information on fungal interactions and diseases is available in a review by Wang *et al.* (204).

Despite their low abundance, fungi appear to have developed in mammalian guts along with the rest of the body from infancy (106, 169). Although there is no consensus of a core mycobiome, *Candida*, *Saccharomyces*, and *Malassezia* have been commonly reported (72). Most of the fungal species detected appear to be either transient or environmental fungi that cannot colonize the gut and are often found in a single study and/or one host only. A previous study indicated that the fungal community is unstable; only 20% of the initially identified fungi were detected again 4 months later (78). More studies on the stability of the mycobiome are needed in order to establish the ecological roles of the components of the mycobiome. Many non-bacterial organisms have been found in numerous mammalian systems, which indicates that they play an important role that has been largely overlooked and may lead to important discoveries and understanding in the coming years.

Archaea

The most commonly reported genus of *Archaea* that has been found in the GI tract is *Methanobrevibacter* (51, 55, 66, 85, 109). Other genera that have also been detected are *Methanosphaera* (51), *Nitrososphaera*, *Thermogynomonas*, and *Thermoplasma* (85) and the new candidate species, *Methanomethylophilus alvus* (27, 131). Although *Archaea* comprise a very small proportion of the microbiota, *Methanobrevibacter* species are important contributors to methanogenesis (66). Differences in *Archaea* in microbiome samples may be due to the method used (51) and/or complex relationships with other microbiota. For example, *Methanobrevibacter* and *Nitrososphaera* were previously shown to be mutually exclusive and potentially related to carbohydrate intake (85). More studies are needed in order to clarify the interaction between *Archaea* and other microbiota groups, which may contribute to our understanding of their fitness and function (beyond methanogenesis) in the microbiome.

Viruses

Viruses in the human microbiome have also been understudied and available information is limited (161); the majority of data are related primarily to disease and do not address the commensal virome (34, 40). The majority of viral reads in studies that have been performed cannot be assigned to a known group; this has contributed to the difficulties associated with assessing their roles in the GI tract (124, 160). A number of teams have made extensive efforts in order to advance human virome studies (157, 161). In the last ten years, the number of identified polyomaviruses has increased from 4 to 13 species (some that cause disease and some that do not) (47), and the accuracy of identification techniques has been improved to identify taxa at the genus level (199) and use metagenomic information for viral taxonomy (172). Viral communities are mainly comprised of bacteria-infecting phage families (~90%), while eukaryotic viruses (~10%) are

in lower abundance (157, 161). Metagenomic analyses have suggested that the new bacteriophage, crAssphage associated with *Bacteroides*, is potentially common in humans (53). The greatest diversity of phages is considered to occur in infants and decreases with age, in contrast to increases in bacterial diversity (116, 117, 162). With the availability of methods to enrich viruses in samples (41), and with more metagenomic sequences and bioinformatics tools to identify viral sequences (53, 139), more information will be obtained on viral diversity and associated physiological factors in humans.

Similar to the microbiota, considerable variability appears to exist in the viral taxa found among subjects (133). Limited information is currently available on the functional roles of most viruses in the human GI tract. However, some possible functions are: to increase bacterial fitness as sources of genetic information (*e.g.*, the source of antibiotic resistance genes), to increase the immunity of bacteria or the human host, and to protect against pathogens (40, 64, 157). The general consensus is that the presence of bacteria is beneficial for viruses that are increasingly trying to evade the immune system. This relationship may also be beneficial to bacteria as viruses may be sources of potentially advantageous genes (resistance or tolerance to stress environments). Researchers are now examining the ecological and evolutionary influences of phages on bacterial ecosystems (102), and the findings obtained may provide insights into the important roles played by phages in the gut microbiome.

The GI tract

Many challenges are associated with studying the microbial ecology of the GI tract because it is composed of chemically and physically diverse microhabitats stretching from the esophagus to the rectum, providing a surface area of 150–200 m^2 for colonization or transient occupation by microbes (16). The adult GI tract was initially estimated to harbor 10^{14} bacteria, 10 times more cells than the human body (16, 120); however, a more recent calculation estimates there to be 10^{13} bacteria, which is equivalent to the number of human cells (170). Lower bacterial numbers (10^3 to 10^4 bacteria mL^{-1} of intestinal content) are found in the upper end of the GI tract, stomach, and small intestine, in which pH is low and the transit time is short (16). The highest biodiversity (richness and evenness) of bacteria (10^{10}–10^{11} bacteria g^{-1} of intestinal content) is in the colon, in which cell turnover rate is low, redox potential is low, and the transit time is long. This section highlights the different functions and associated microbiota along the human GI tract starting from the oral cavity, then the esophagus, stomach, and intestines (Fig. 1).

The oral cavity. Activity in the mouth may have a large impact on the further digestion of food in the lower GI tract. Food is mechanically ground into small particles, typically 0.1 mm, which increases the surface area. The oral microbiome is composed of transient and commensal populations that often form biofilms on soft and hard surfaces in the mouth (8). The most up-to-date information on taxa of the oral microbiome may be found in the Human Oral Microbiome Database (HOMD, http://www.homd.org/) (50). Information in this database is limited to *Bacteria* and one *Archaea*. Cultivation-independent analyses indicate that the most common genus is *Streptococcus*, while other genera include *Neisseria*, *Gemella*, *Granulicatella*, and *Veillonella*, but not in all individuals examined (1, 91, 92, 107). The taxa present appear to be dependent on interactions between microbes within the community. For example, using a graph theory-based algorithm of an organism's nutritional profile, the species *Streptococcus oralis* and *S. gordonii* have low metabolic complementarity and high metabolic competition, indicating they are antagonistic to each other (110). In contrast, *Porphyromonas gingivalis* was shown to have high metabolic complementarity, indicating its ability to grow symbiotically with diverse oral microbiota taxa. This computational method was tested and confirmed with growth assays, making it a viable means to assess the ability of species to inhabit the same environment. This has also been shown using an *in situ* spectral analysis of microbiota in biofilm plaques. Biofilms were shown to be composed of a number of taxa with *Corynebacterium* at the foundation (209). The other taxa are considered to play complementary roles driven by the environmental and chemical gradients formed in biofilms that control nutrient availability. These findings indicate that, despite the large number of taxa identified in oral microbiome studies, the core taxa of all microbiota may be identified in the future based on spatial locations and functional roles (10).

Similar to *Bacteria*, large variations have been noted in viruses found in the oral cavity among subjects (151). Most viruses are bacteriophages (approx. 99% of known sequences). Viral communities are reproducible across time points within a subject, suggesting that they are stable; however, the human and bacterial host significantly influence compositions (2, 151, 163). In addition to interactions among oral bacteria, many may associate with phages (57). Depending on the host range of the oral virome, this may make phages very common inhabitants of the oral cavity. Furthermore, in addition to survival within bacterial hosts, phages may also survive in the oral mucosa and contribute to host immunity (11). These are all new avenues of oral virome research that will likely be investigated in greater depth in the future.

In addition to the bacterial microbiome, two cultivation-independent studies have been conducted on oral fungi. Approximately 100 fungal species (20 genera) were detected in one study of the oral mycobiome of healthy individuals (68). Among the fungi detected, *Candida* species were the most common and abundant, while the other genera consisted of *Cladosporium*, *Aureobasidium*, *Saccharomycetales*, *Aspergillus*, *Fusarium*, and *Cryptococcus*. Most of these genera were also detected in a recent study on three subjects; however, *Malassezia*, a skin pathogen, accounted for the most sequence reads (52). Most of the other studies conducted on the oral mycobiome have focused on the role of fungi in disease (69, 136). Since the oral microbial community is directly exposed to the environment, the presence of a dynamic and transient community is expected, but warrants further study.

Esophagus. After swallowing, food is transported down the esophagus by peristalsis to the stomach. Limited information is available on microbes inhabiting the esophagus (5, 91, 147), and this may be due to the difficulties associated with obtaining samples because biopsies have typically been used. However, a less invasive method using an esophageal string has recently been demonstrated to be a feasible alternative

and yields similar findings to non-cultivation-based analyses of biopsies (60). Similar to the oral cavity, the most common genus found in the esophagus is *Streptococcus*; however, an overall comparison of the two communities has indicated that the number of taxa significantly differ between the two locations (15, 60). Among the few studies conducted on the viral and fungal microbiota of the esophagus, the focus has been on association with disease (204) and none of the pathogenic taxa inhabit healthy individuals.

Stomach. The stomach is the first digestive organ in the body (89). It holds food and mechanically mixes it with proteolytic enzymes and gastric acids that aid in the breakdown and subsequent absorption of nutrients. The growth of many common bacteria is inhibited by these acidic conditions (pH<4), making this a unique community with the lowest number of microbes, ranging between 10^1 and 10^3 CFU g^{-1}. In addition to digestion, the acidic conditions of the stomach are considered to have evolved as a means of protection from pathogens. This hypothesis is supported by the recent finding of a lower pH in the stomachs of scavengers and higher pH in herbivores, which are less likely to encounter pathogens in their food (13). Caution is needed when comparing the findings of various studies throughout the GI tract because gastric juice has a lower pH than the mucosal layer, resulting in differences in the microbiota present (89).

Despite the low pH, non-cultivation-based analyses on stomach biopsies revealed a more diverse microbiota than expected (5, 20, 115). Regardless of variations among subjects, there appears to be two major groups of individuals: those with and without *Helicobacter pylori* (20). There is a third subset in which *H. pylori* is present in lower proportions in some individuals that were negative using conventional testing. Microbiomes dominated by *H. pylori* had significantly greater proportions of the phylum *Proteobacteria*, of which it is a member, and lower alpha diversity (5, 20). Other common genera are *Streptococcus* and *Prevotella*, both of which are also found in the oral and esophageal communities; however, the communities at these locations appear to differ (5). Limited information is available on fungi analyzed in biopsy samples; although a cultivation study detected *Candida* species, this appeared to be associated more with disease (224). The major interaction currently studied in the stomach microbiota is with *Helicobacter* because of its association with gastritis, peptic ulcers, and gastric cancer. However, this taxon has been suggested to be beneficial for health, leading some to question whether the complete eradication of this microbe is the best option (67, 89).

In contrast, less information is available on the microbiome of stomach fluids; it appears to harbor fewer *Helicobacter* and an analysis of transcripts indicated that *Actinobacteria* are the most active phylum; however, the other major phyla, *Firmicutes*, *Bacteroidetes*, and *Proteobacteria*, are also present (197). In the same study, it also appeared to harbor novel fungi; 77.5% of the ITS reads were not identified at the phylum level or lower. *Candida* and *Phialemonium* were the only two identifiable fungal genera in all subjects tested, whereas an additional 66 genera were present in at least one of the nine subjects examined. Based on the infrequency and number of reads in this analysis, most of the taxa identified in stomach fluids appear to be transient, and those playing an active role are limited in this location.

Intestines. After mixing in the stomach, chime slowly passes through the pyloric sphincter and enters the intestines, in which the major digestion and absorption of nutrients begin (12). Humans have a small and large intestine. The small intestine, the main location in which food digestion and absorption occurs, is further divided into three parts, the duodenum, jejunum, and ileum. The duodenum, in which food chime enters from the stomach, is directly associated with digestion and is linked to the pancreas and gallbladder. Bile salts from the gallbladder and enzymes from the pancreas enter the duodenum and mix with stomach chime in order to start the digestion process. The epithelium in the jejunum and ileum is responsible for glucose absorption into the bloodstream via glucose transporters and sodium ions. The small intestine is followed by the large intestine (colon), which has a larger diameter, but shorter length and is divided into four sections: the ascending colon (cecum), transverse colon, descending colon, and sigmoid colon (123). Water and minerals are continuously absorbed along the colon before excretion. Furthermore, complex foods that cannot be digested by the host are used as growth substrates for the colonic microbiota (25, 178).

Spatial and temporal variabilities have been noted in the microbial composition among the different intestinal structures based on their functional roles and timing of food intake (18, 129, 186). Although spatial variability exists along the intestinal tract, the bacterial microbiome at the phylum level is considered to remain fairly stable over time (43, 155); however, many factors may affect its stability (119). Undigested food and most of the microbiota are found in the lumen, the central space surrounded by the mucosal layer of the tubular intestinal structure. The main absorption of growth substrates occurs through the epithelial cells of the mucosa, which also prevents the entry of the microbiota into host cells (174). A number of important host-microbe interactions occur within the mucosa. Energy from microbially produced metabolites, such as butyrate, contributes to epithelial metabolism (97). Most of the gut is anaerobic, but there is an oxygen gradient in the mucosa that provides a competitive advantage for facultative anaerobes (174). Recent studies have also shown the importance of metabolites produced by transkingdom microbiota to host physiology (185, 187, 188). Microbiota, such as *Akkermansia mucinophila*, are commonly found residing in the mucus layer and feed on mucin (39, 48). Therefore, the effects of host interactions with the gut microbiota, particularly those in the large intestine, have a prominent impact on overall human health, including energy reabsorption and immune system development.

Due to the difficulties associated with collecting multiple samples along a healthy human GI tract in order to capture the spatial heterogeneity of microbes in this environment, most studies use fecal samples as a surrogate. However, this limits the availability of regio-specific community information on the GI tract, resulting in portions, such as the small intestine, remaining poorly characterized. The few studies conducted on the small intestine have limited subject numbers because they used biopsy samples (4, 201, 203) or ileotomy patients (108, 195, 222). The bacterial genera most commonly found among these studies were *Clostridium*, *Streptococcus*, and *Bacteroides*. The number of studies that include fungi are

Table 1. Comparison of the anatomy of the intestinal tract in humans and animal models

	Human	Mouse	Rat	Pig
Stomach	Four regions: cardia, fundus, body, and pylorus	Three regions: forestomach, body, and pylorus	Three regions: forestomach, body, and pylorus	Four regions: esophagus, cardia, fundus, and pylorus
	pH 1.5 to 3.5	pH 3.0 to 4.0	pH 3.0 to 4.0	pH 1.5 to 2.5
Small intestine	5.5–6.4 m in length	350 mm in length	1,485 mm in length	1.2–2.1 m in length
	pH 6.4 to 7.3	pH 4.7 to 5.2	pH 5.0 to 6.1	pH 6.1 to 6.7
Cecum	Smaller than the colon	Larger than the colon	Larger than the colon	Smaller than the colon
	No fermentation	Main fermentation	Main fermentation	Some fermentation
	pH 5.7	pH 4.4 to 4.6	pH 5.9 to 6.6	pH 6.0 to 6.4
Appendix	Present	Absent	Absent	Absent
Colon	Divided into the ascending, transcending, and descending colon	Not divided	Not divided	Divided into the ascending, transcending, and descending colon
	Main fermentation	No fermentation	No fermentation	Main fermentation
	Thick mucosa	Thinner mucosa	Thinner mucosa	Thick mucosa
	pH 6.7	pH 4.4 to 5.0	pH 5.5 to 6.2	pH 6.1 to 6.6

Adapted from (59, 96, 128, 130, 137, 196)

even more limited, with the genera *Candida* and *Saccharomyces* being the most frequently detected (108, 114). Caution is also needed when extrapolating these findings to all individuals because the health of some subjects was compromised when samples were obtained.

Bacteria in the colon account for approximately 70% of all bacteria in the human body because it is the main site for the bacterial fermentation of non-digestible food components such as soluble fiber. The small number of studies that have examined microbial communities directly in the colon suggests that the bacterial composition is similar to that found in feces (86). However, fecal communities do not represent a single colonic environment, such as the mucosa (223), but a mixture of indigenous and transient microbes from the entire GI tract. In studies in which a global view of the GI tract microbial community is of interest, fecal material represents a good surrogate and is easily obtained, allowing for multiple samples to be obtained over short and long time periods from healthy individuals. The majority of microbiome reviews have extensively covered colonic communities using feces (74, 92, 121, 148, 189); therefore, we will not describe its composition in detail. However, later in this review, the impact of diet on the microbiome composition will be discussed. Furthermore, the above sections on fungi and viruses provide information on the taxa of these groups in the intestines.

Summary of the GI tract. The use of non-cultivation-based methods to investigate the microbiota in the GI tract has increased our knowledge of their diversity. One group that we neglected to mention in this review was *Protozoans*/*Protists*; however, recent reviews are available (79, 145). Despite representing a smaller biomass than fungi, they also appear to be important to the ecological structure of the gut microbiome. The predator-prey relationship they have with other microbiota (145) may, in some cases, lead to disease prevention (7). Difficulties are associated with elucidating the functional roles played by these various taxa at different points along the GI tract. Therefore, it is still important to obtain cultivated representatives to investigate their role and ecological significance along the GI tract. This consideration is important for all microbiota; however, it represents a larger issue for low diversity groups, such as fungi, which may not be numerically abundant, but still play a significant role (17).

Use of animal models

Animal models have been widely adopted in human gut microbiome research (28, 98, 220) to reduce confounding experimental factors such as genetics, age, and diet, which may be more easily controlled in laboratory animals. Additionally, animal models with modified genetic backgrounds are available for investigating potential mechanisms (137). Ideally, animal models with relatively similar genetic information (217), gut structures, metabolism (142), and diets and behavior patterns (202) to humans need to be selected. Comprehensive comparisons of mice (137) and pigs (217) to humans were recently conducted in order to aid in translating information from animal models to humans. In this section, we will highlight some of their findings and compare GI tract structures and microbial community compositions. Furthermore, some advantages and limitations associated with the use of animal models in human microbiome research will be discussed.

Similarities exist in the anatomy of the GI tract between humans and most animal models (Table 1). However, differences in anatomical structures and pH at different locations along the GI tract may contribute to differences in the microbiota found in humans versus animal models (26). The human colon also has a thicker mucosal layer than those of mice and rats (137), which may have an effect on the diversity of the microbiota colonizing the colon. Human gut bacteria are dominated by two phyla: *Firmicutes* and *Bacteroidetes* (189), which also dominate the GI tract of commonly used model animals (112). However, at lower taxonomic levels, some differences have been reported in microbiome compositions in the gut between humans and animal models (Table 2). The dominant taxa reported have varied as the number of comparisons performed has increased (137, 152); therefore, the findings shown in Table 2 need to be used cautiously.

A pig gut gene catalogue of metabolic function was recently developed and compared to catalogues available for humans and mice (217). They found that 96% of the KEGG orthologs in humans were also present in pigs, whereas the overlap at the gene level was markedly lower (9.46%). However, there was a greater overlap between humans and pigs than between humans and mice. Microbial activity also differs along the GI tract, with the most relevant being fermentation occurring in the ceca of most animal models, but

Table 2. Major taxa of the gut microbiota in humans and animal models

	Human	Mouse	Rat	Pig
Bacteria	*Firmicutes* *Bacteroidetes* *Actinobacteria* *Proteobacteria*	*Firmicutes* *Bacteroidetes*	*Firmicutes* *Bacteroidetes*	*Firmicutes* *Bacteroidetes*
Archaea	*Methanobrevibacter* *Nitrososphaera*	*Methanobrevibacter*	*Methanobrevibacter*	*Methanomicrobia*, *Methanosphaera*
Viruses	*Herpesviridae* *Papillomaviridae* *Polyomaviridae* *Adenoviridae*	Variable	Variable	*Picornaviridae* *Astroviridae* *Coronaviridae* *Caliciviridae*
Eukarya	*Candida* *Malassezia* *Saccharomyces* *Cladosporium*	*Ascomycota* *Basidiomycota* *Chytridiomycota* *Zygomycota*	*Ascomycota* *Basidiomycota* *Chytridiomycota* *Zygomycota*	*Kazachstania* *Candida* *Galactomyces* *Issatchenkia*

Adapted from (85, 103, 105, 112, 125, 137, 153, 154, 171, 179, 193, 194, 215, 216, 221)

not in humans (137). Strengths and weaknesses are associated with the major animal models being used, and these need to be taken into consideration when conducting translational research.

Rats. The use of rats as lab animals dates back to the 1850s. They were considered to be a good candidate for human microbiome research because the rat contains the same four dominant bacteria phyla in the GI tract (31), with *Firmicutes* (74%) and *Bacteroidetes* (23%) representing the largest proportions (221). The advantages of using rats in human microbiome research include quick reproduction, a fully sequenced genome, and easy handling and maintenance due to their relatively small size. The limitation of this model is that the diet used in rats differs from that for humans, and their behavior and living environment are also different, which will affect the gut microbiota. The diet used in rat studies is normal chow that is rich in fiber (205), and diet may rapidly alter gut microbiota diversity (46). Although most studies emphasize the impact of diet on the microbiota in the cecum and/or colon (feces), the oral cavity of rats has been used to clarify the impact of diet on the microbiome (93).

Mice. Many of the strengths and weaknesses associated with using rats are also applicable to mice. Similar to humans, the microbiota in the GI tract of mice is dominated by *Firmicutes* (74%) and *Bacteroidetes* (23%) at the phylum level (217). However, there are differences at the genus level, and this has led to the use of "humanized" mice. This is achieved by inoculating human gut microbiota into germ-free (GF) mice (192) or mice treated with antibiotics to eliminate their gut microbiome (83). The microbiome of these mice after fecal transplants may have a composition at the phyla level that is 100% similar to humans and 88% at the genus level (137). A recent study (175) used humanized mice to test microbiome diversity after feeding with poorly accessible carbohydrates, and found a similar reduction in OTU numbers to a human study (219). However, there are also some limitations to using these animals, including the diet and environmental living conditions. Furthermore, gnotobiotic mice may not reflect the human-microbe relationship due to their weaker immune system (6).

Approximately 10 years ago, Scupham (168) showed that all four major fungal phyla, *Ascomycota*, *Basidiomycota*, *Chytridiomycota*, and *Zygomycota*, were present in the murine gut. Additionally, many genera were identified, including *Acremonium*, *Monilinia*, *Fusarium*, *Cryptococcus*, *Filobasidium*, *Scleroderma*, *Catenomyces*, *Spizellomyces*, *Neocallimastix*, *Powellomyces*, *Entophlyctis*, *Mortierella*, and *Smittium*. When comparing these studies to the human gut, it is important to note that this study indicated a more diverse fungal community than those found in humans; the eukaryotic diversity of the human gut is low (143).

Pigs. Pigs have been used as surrogates for human microbiome research due to their highly similar genetics, physiological structures, behavior, metabolism, and immune functions to those of humans (81, 202). The greater similarities in the omnivorous diet and GI tract structure between pigs and humans are more advantageous than the murine model. The microbiome of pigs is dominated by two phyla: *Firmicutes* and *Bacteroidetes* (104); however, there are some notable differences at the genus level. The genus *Prevotella* was found to be common in two pig metagenomic studies (104, 118). Since the number of pigs used in most studies is less than humans, the pig core microbiome at the genus level may change as more pigs are studied. Another contributing factor to shaping the microbiome composition is diet. Most studies have found that the number of *Bifidobacteria* in pigs, even those on high fiber diets, is lower than that in humans (132, 218), while that of *Lactobacillus* is higher (149). In nutrition studies, humans and pigs are both dependent on the quality of the nutrient load; however, the pig cecum has a larger capacity to ferment indigestible compounds than the human cecum (54). The microbiota composition in pigs may differ from that in humans due in part to differences in diet (81). Similar to mice, humanized GF pigs have been developed and the microbiome after human fecal transplantation more closely resembles that of the donor than conventional pigs (144). However, the same disadvantages associated with using GF mice are also true for GF pigs.

The genome of pigs may be mutated to study human diseases; this is typically performed using miniature pigs such as those from the Ossabaw and Gottingen islands (146). Genetic mutations for metabolic syndrome and insulin resistance have successfully been performed using Ossabaw pigs to study human diseases such as type 2 diabetes (14, 177) and obesity (101). The ratio of *Firmicutes* to *Bacteroidetes* is higher in obese Ossabaw pigs than in lean pigs (146), similar to some

obese humans (111, 190). This finding suggests that Ossabaw pigs are a good model for researching the role of the microbiota in human obesity. However, disadvantages are associated with using miniature pigs, mainly the higher cost for maintenance and longer reproductive period than rodents (146).

Although more extensive efforts have been made to investigate fungi in pigs than in other animal models, many of these studies were cultivation-based or for use as probiotics. Fungi in pigs have been recently studied using a non-cultivation approach and up to 17 species of yeast (belonging to the genera *Kazachstania*, *Galactomyces*, *Candida*, *Issatchenkia*, *Pichia*, *Rhodotorula*, and *Trichosporon*) were common in the gut (194). The number of studies on viruses is limited, but the composition appears to be highly variable among samples (164, 171) and affected by disease (24). These groups need to be examined in more detail in order to establish whether pigs are good models for use in understanding fungi and viruses in humans.

Animal model summary. The convenience and cost of using animal models for human research are appealing. However, researchers need be very careful when selecting animal models appropriate for their objectives, particularly when the objective is to directly extrapolate findings from animals to humans, due to the significant differences in GI tract physiology and microbiome composition (65, 137, 217).

Diet in health

Many studies have found that diet is one of the main factors shaping the composition of gut microbial populations. Dietary approaches, such as the ingestion of non-digestible carbohydrates (prebiotics) and fermented food products containing live cultures (probiotics), have been suggested to confer health benefits by enhancing the growth of beneficial intestinal bacteria (100, 158). As described earlier, the microbiota may break down food components, such as non-digestible carbohydrates, which are indigestible by the host in order to aid in maximizing available nutrients (9) and produce metabolites that contribute to host health. Probiotics have been used as a means to replenish the gut with "beneficial" microbiota after antibiotic treatments or to treat diseases (82, 159). This section will highlight some studies that demonstrated the health benefits of prebiotics and probiotics and possible roles played by the microbiota.

Dietary prebiotics and probiotics. Non-digestible and fermentable food components are often consumed as prebiotics to selectively stimulate the growth and/or activity of endogenous colonic bacteria that may be beneficial to host health. The increased consumption of prebiotics often correlates with enhancements in certain bacterial genera (a common example is *Bifidobacterium* sp.); however, the reason they are beneficial remains unclear (208). Challenges are associated with elucidating the role being played by specific bacterial phylotypes because many of their processes are interactive (207). For example, SCFA produced by bacterial fermentation may lower intestinal pH, thereby increasing the solubility of essential minerals, such as calcium, iron, and magnesium, and consequently enhancing their absorption and improving health. Metabolites produced by microbes may also play an important role in cellular differentiation and proliferation in the colonic mucosa by inducing apoptosis and may confer protection against colitis and colorectal cancer by modulating oncogene expression. These functions do not appear to be performed by a single species; a number of different species may be acting independently or in combination. Research is leading to an understanding of microbial community structure and composition dynamics with respect to diet aids in establishing testable hypotheses for future research in health and beneficial microbes (32). Most research has been performed on the influence of beneficial intestinal bacteria such as *Bifidobacterium* spp. and *Lactobacillus* spp. on host health monitored using a cultivation approach. Cultivation-independent approaches have now become more popular, leading to the identification of new beneficial microbiota taxa and their potential functional roles in the gut as they relate to diet.

Dietary fibers and oligosaccharides are carbohydrate ingredients that vary in composition and structure, but are considered to be non-digestible because of the lack of appropriate intestinal enzymes to hydrolyze them or structural hindrances that prevent enzyme access in the gut. Although bacteria in the lower gut may ferment these carbohydrates, the rate and degree of fermentation vary with the polysaccharide (80). The range of fermentation in the colon for various fibers is broad, from approximately 5% for cellulose to nearly 100% for pectin (42). The resulting SCFA, including butyrate and propionate, are considered to reduce pH and solubilize minerals, thereby improving their absorption and subsequent utilization. Inulin, a long chain fructooligosaccharide (FOS) often obtained from chicory root, and FOS from other sources are the fibers that have been studied in the most detail (206). Several novel fibers have been tested in an *in vitro* large intestine model for their effects on the microbial stimulation and production of SCFA (122). All these novel fibers stimulated the growth of beneficial *Bifidobacteria* and some *Lactobacillus* species along with increases in SCFA production. Only a few studies have examined the effects of fibers and resistant starches on the human microbiome (87, 127, 198, 210, 211). A soluble corn fiber product has been demonstrated to increase Ca absorption in a number of different studies (210, 211). More benefits to human health may be attributed to the consumption of prebiotics and fermentation by the gut microbiome.

The number of studies that include diet effects on *Archaea*, *Fungi*, and/or *Viruses* are limited; however, some examples are included herein. Examinations of *Archaea*, *Fungi*, and *Bacteria* correlations in response to diet revealed a syntrophic model involving *Candida*, *Prevotella*, *Ruminococcus*, and *Methanobrevibacter* (85). *Candida* was considered to break down carbohydrates into metabolites used by *Prevotella* and *Ruminococcus* that produce CO_2 for *Methanobrevibacter* (85). However, shifts in carbon sources or breaking down starches via amylases from the human mouth may alter this relationship because *Prevotella* may no longer be dependent on *Candida*. This is a good example of how *Archaea*, which represent a very small portion of the microbiome, are a key contributor to methanogenesis and waste decomposition. The absence of *Archaea* may have severe effects on the surrounding community as hydrogen, glucose metabolites, and other carbon sources accumulate. Other organisms will eventually fill this niche, but may diminish or accumulate new metabolites that

ultimately shift the surrounding community based on their fitness for using these substrates.

A recent study investigated rapid changes in the microbiome composition when diets were either high in animal-based or plant-based fat and protein (46). The fungus *Candida* was found to increase in subjects placed on a plant-based diet, whereas *Penicillium* increased on animal-based diets. The most commonly found fungi in vegetarians were *Fusarium*, *Malassezia*, *Penicillium*, *Aspergillus*, and *Candida* (182). Caution is needed when interpreting findings because some of these fungi may be found on food prior to ingestion (46, 78, 182)

Phages assembled in the gut may also be modified by diet. A recent study examined changes in the fecal viral community over an 8-d period in six subjects supplied different diets (134). Shotgun sequencing of virus-like particles revealed that interpersonal differences in the virome were the largest source of variations in this study. However, the virome of subjects whose diets were changed differed more than in those who maintained their normal diet. Although this is only one study with a few human subjects, studies using a mouse model and different dietary fats support these findings (88, 99). Collectively, these findings indicate that diet plays a key role in shaping the gut virome, and further research is needed in order to investigate interactions between diet and the virome.

Summary

Advances have been made in the last decade in our understanding of the role of the GI tract microbiome in human health. This review has highlighted changes and differences in the microbiome along the GI tract that are due to changes in physical, chemical, and biological interactions. Although extensive research has been conducted on *Bacteria* in fecal samples, the main kingdom inhabiting the gut, our knowledge is still insufficient, particularly in other regions of the GI tract. Furthermore, other groups (*Archaea*, *Fungi*, and *Viruses*) have not yet been investigated in adequate detail, demonstrating a real void in knowledge. This highlights that the basic ecology of microbiomes is important for gaining a greater understanding to improve human health and decrease disease. In order to achieve this goal, it is important to include all microbiota in studies and remain cognizant of the limitations associated with understanding the entire GI tract of humans despite challenges in sampling and cultivation. Furthermore, the use of appropriate animal models in mechanistic studies requires careful consideration.

References

1. Aas, J.A., B.J. Paster, L.N. Stokes, I. Olsen, and F.E. Dewhirst. 2005. Defining the normal bacterial flora of the oral cavity. J. Clin. Microbiol. 43:5721–5732.
2. Abeles, S.R., R. Robles-Sikisaka, M. Ly, A.G. Lum, J. Salzman, T.K. Boehm, and D.T. Pride. 2014. Human oral viruses are personal, persistent and gender-consistent. ISME J. 8:1753–1767.
3. Abubucker, S., N. Segata, J. Goll, *et al.* 2012. Metabolic reconstruction for metagenomic data and Its application to the human microbiome. PLoS Comput. Biol. 8:e1002358.
4. Ahmed, S., G.T. Macfarlane, A. Fite, A.J. McBain, P. Gilbert, and S. Macfarlane. 2007. Mucosa-associated bacterial diversity in relation to human terminal ileum and colonic biopsy samples. Appl. Environ. Microbiol. 73:7435–7442.
5. Andersson, A.F., M. Lindberg, H. Jakobsson, F. Bäckhed, P. Nyrén, and L. Engstrand. 2008. Comparative analysis of human gut microbiota by barcoded pyrosequencing. PLoS One 3:e2836.
6. Atarashi, K., T. Tanoue, K. Oshima, *et al.* 2013. T-reg induction by a rationally selected mixture of Clostridia strains from the human microbiota. Nature 500:232–236.
7. Audebert, C., G. Even, A. Cian, A. Loywick, S. Merlin, E. Viscogliosi, and M. Chabe. 2016. Colonization with the enteric protozoa *Blastocystis* is associated with increased diversity of human gut bacterial microbiota. Sci. Rep. 6:25255.
8. Avila, M., D.M. Ojcius, and Ö. Yilmaz. 2009. The oral microbiota: Living with a permanent guest. DNA Cell Biol. 28:405–411.
9. Backhed, F., R.E. Ley, J.L. Sonnenburg, D.A. Peterson, and J.I. Gordon. 2005. Host-bacterial mutualism in the human intestine. Science 307:1915–1920.
10. Baker, J.L., B. Bor, M. Agnello, W.Y. Shi, and X.S. He. 2017. Ecology of the oral microbiome: Beyond bacteria. Trends Microbiol. 25:362–374.
11. Barr, J.J., R. Auro, M. Furlan, *et al.* 2013. Bacteriophage adhering to mucus provide a non-host-derived immunity. Proc. Natl. Acad. Sci. U.S.A. 110:10771–10776.
12. Barrett, K.E. 2014. Gastrointestinal Physiology, 2nd ed. Lange Medical Books/McGraw-Hill.
13. Beasley, D.E., A.M. Koltz, J.E. Lambert, N. Fierer, and R.R. Dunn. 2015. The evolution of stomach acidity and is relevance to the human microbiome. PLoS One 10:e.0134116.
14. Bellinger, D.A., E.P. Merricks, and T.C. Nichols. 2006. Swine models of type 2 diabetes mellitus: Insulin resistance, glucose tolerance, and cardiovascular complications. ILAR J. 47:243–258.
15. Benitez, A.J., C. Hoffmann, A.B. Muir, K.K. Dods, J.M. Spergel, F.D. Bushman, and M.-L. Wang. 2015. Inflammation-associated microbiota in pediatric eosinophilic esophagitis. Microbiome 3:23.
16. Berg, R.D. 1996. The indigenous gastrointestinal microflora. Trends Microbiol. 4:430–435.
17. Berrilli, F., D. Di Cave, S. Cavallero, and S. D'Amelio. 2012. Interactions between parasites and microbial communities in the human gut. Front. Cell. Infect. Microbiol. 2:141.
18. Biggs, M.B., G.L. Medlock, T.J. Moutinho, H.J. Lees, J.R. Swann, G.L. Kolling, and J.A. Papin. 2017. Systems-level metabolism of the altered Schaedler flora, a complete gut microbiota. ISME J. 11:426–438.
19. Bik, E.M. 2009. Composition and function of the human-associated microbiota. Nutr. Rev. 67:S164–S171.
20. Bik, E.M., P.B. Eckburg, S.R. Gill, K.E. Nelson, E.A. Purdom, F. Francois, G. Perez-Perez, M.J. Blaser, and D.A. Relman. 2006. Molecular analysis of the bacterial microbiota in the human stomach. Proc. Natl. Acad. Sci. U.S.A. 103:732–737.
21. Bik, E.M., C.D. Long, G.C. Armitage, *et al.* 2010. Bacterial diversity in the oral cavity of 10 healthy individuals. ISME J. 4:962–974.
22. Blaser, M.J. 2016. Antibiotic use and its consequences for the normal microbiome. Science 352:544–545.
23. Blaut, M., M.D. Collins, G.W. Welling, J. Dore, J. van Loo, and W. de Vos. 2002. Molecular biological methods for studying the gut microbiota: the EU human gut flora project. Brit. J. Nutr. 87:S203–S211.
24. Blomstrom, A.L., C. Fossum, P. Wallgren, and M. Berg. 2016. Viral metagenomic analysis displays the co-infection situation in healthy and PMWS affected pigs. PLoS One 11:e0166863.
25. Bolam, D.N., and J.L. Sonnenburg. 2011. Mechanistic insight into polysaccharide use within the intestinal microbiota. Gut microbes 2:86–90.
26. Booijink, C., E.G. Zoetendal, M. Kleerebezem, and W.M. de Vos. 2007. Microbial communities in the human small intestine: coupling diversity to metagenomics. Fut. Microbiol. 2:285–295.
27. Borrel, G., H.M. Harris, W. Tottey, *et al.* 2012. Genome sequence of "Candidatus *Methanomethylophilus alvus*" Mx1201, a methanogenic archaeon from the human gut belonging to a seventh order of methanogens. J. Bacteriol. 194:6944–6945.
28. Bowey, E., H. Adlercreutz, and I. Rowland. 2003. Metabolism of isoflavones and lignans by the gut microflora: a study in germ-free and human flora associated rats. Food Chem. Toxicol. 41:631–636.
29. Breitbart, M., I. Hewson, B. Felts, J.M. Mahaffy, J. Nulton, P. Salamon, and F. Rohwer. 2003. Metagenomic analyses of an uncultured viral community from human feces. J. Bacteriol. 185:6220–6223.

30. Breitbart, M., M. Haynes, S. Kelley, *et al.* 2008. Viral diversity and dynamics in an infant gut. Res. Microbiol. 159:367–373.
31. Brooks, S.P.J., M. McAllister, M. Sandoz, and M.L. Kalmokoff. 2003. Culture-independent phylogenetic analysis of the faecal flora of the rat. Can. J. Microbiol. 49:589–601.
32. Brownawell, A.M., W. Caers, G.R. Gibson, C.W.C. Kendall, K.D. Lewis, Y. Ringel, and J.L. Slavin. 2012. Prebiotics and the health benefits of fiber: Current regulatory status, future research, and goals. J. Nutr. 142:962–974.
33. Cadwell, K. 2015. Expanding the role of the virome: Commensalism in the gut. J. Virol. 89:1951–1953.
34. Cadwell, K. 2015. The virome in host health and disease. Immunity 42:805–813.
35. Cavalcanti, I.M.G., A.H. Nobbs, A.P. Ricomini, H.F. Jenkinson, and A.A.D. Cury. 2016. Interkingdom cooperation between *Candida albicans*, *Streptococcus oralis* and *Actinomyces oris* modulates early biofilm development on denture material. Pathog. Dis. 74:ftw002.
36. Charbonneau, M.R., L.V. Blanton, D.B. DiGiulio, D.A. Relman, C.B. Lebrilla, D.A. Mills, and J.I. Gordon. 2016. A microbial perspective of human developmental biology. Nature 535:48–55.
37. Cho, I., and M.J. Blaser. 2012. The human microbiome: at the interface of health and disease. Nat. Rev. Genet. 13.
38. Claesson, M.J., O. O'Sullivan, Q. Wang, J. Nikkilä, J.R. Marchesi, H. Smidt, W.M. de Vos, R.P. Ross, and P.W. O'Toole. 2009. Comparative analysis of pyrosequencing and a phylogenetic microarray for exploring microbial community structures in the human distal intestine. PLoS ONE 4:e6669.
39. Collado, M.C., M. Derrien, E. Isolauri, W.M. de Vos, and S. Salminen. 2007. Intestinal Integrity and *Akkermansia muciniphila*, a mucin-degrading member of the intestinal microbiota present in infants, adults, and the elderly. Appl. Environ. Microbiol. 73:7767–7770.
40. Columpsi, P., P. Sacchi, V. Zuccaro, S. Cima, C. Sarda, M. Mariani, A. Gori, and R. Bruno. 2016. Beyond the gut bacterial microbiota: The gut virome. J. Med. Virol. 88:1467–1472.
41. Conceição-Neto, N., M. Zeller, H. Lefrère, *et al.* 2015. Modular approach to customise sample preparation procedures for viral metagenomics: a reproducible protocol for virome analysis. Sci. Rep. 5:16532.
42. Cook, S.I., and J.H. Sellin. 1998. Review article: short chain fatty acids in health and disease. Ailment Pharmacol. Ther. 12:499–507.
43. Costello, E.K., C.L. Lauber, M. Hamady, N. Fierer, J.I. Gordon, and R. Knight. 2009. Bacterial community variation in human body habitats across space and time. Science 326:1694–1697.
44. Cui, L., A. Morris, and E. Ghedin. 2013. The human mycobiome in health and disease. Genome Med. 5:63.
45. Dalmasso, M., C. Hill, and R.P. Ross. 2014. Exploiting gut bacteriophages for human health. Trends Microbiol. 22:399–405.
46. David, L.A., C.F. Maurice, R.N. Carmody, *et al.* 2014. Diet rapidly and reproducibly alters the human gut microbiome. Nature 505:559–563.
47. DeCaprio, J.A., and R.L. Garcea. 2013. A cornucopia of human polyomaviruses. Nat. Rev. Microbiol. 11:264–276.
48. Derrien, M., E.E. Vaughan, C.M. Plugge, and W.M. de Vos. 2004. *Akkermansia muciniphila* gen. nov., sp nov., a human intestinal mucin-degrading bacterium. Internat. J. Syst. Evol. Microbiol. 54:1469–1476.
49. Dethlefsen, L., S. Huse, M.L. Sogin, and D.A. Relman. 2008. The pervasive effects of an antibiotic on the human gut microbiota, as revealed by deep 16S rRNA sequencing. PLoS Biol. 6:e280.
50. Dewhirst, F.E., T. Chen, J. Izard, B.J. Paster, A.C.R. Tanner, W.-H. Yu, A. Lakshmanan, and W.G. Wade. 2010. The human oral microbiome. J. Bacteriol. 192:5002–5017.
51. Dridi, B., D. Raoult, and M. Drancourt. 2011. Archaea as emerging organisms in complex human microbiomes. Anaerobe 17:56–63.
52. Dupuy, A.K., M.S. David, L. Li, T.N. Heider, J.D. Peterson, E.A. Montano, A. Dongari-Bagtzoglou, P.I. Diaz, and L.D. Strausbaugh. 2014. Redefining the human oral mycobiome with improved practices in amplicon-based taxonomy: Discovery of *Malassezia* as a prominent commensal. PLoS One 9:e90899.
53. Dutilh, B.E., C.W. Noriko, K. McNair, *et al.* 2014. A highly abundant bacteriophage discovered in the unknown sequences of human faecal metagenomes. Nat. Commun. 5:4498.
54. Eberhard, M., U. Hennig, S. Kuhla, R.M. Brunner, B. Kleessen, and C.C. Metges. 2007. Effect of inulin supplementation on selected gastric, duodenal, and caecal microbiota and short chain fatty acid pattern in growing piglets. Arch. Animal Nutr. 61:235–246.
55. Eckburg, P.B., P.W. Lepp, and D.A. Relman. 2003. Archaea and their potential role in human disease. Infect. Immun. 71:591–596.
56. Eckburg, P.B., E.M. Bik, C.N. Bernstein, E. Purdom, L. Dethlefsen, M. Sargent, S.R. Gill, K.E. Nelson, and D.A. Relman. 2005. Diversity of the human intestinal microbial flora. Science 308:1635–1638.
57. Edlund, A., T.M. Santiago-Rodriguez, T.K. Boehm, and D.T. Pride. 2015. Bacteriophage and their potential roles in the human oral cavity. J. Oral Microbiol. 7:27423.
58. Egert, M., A.A. de Graaf, H. Smidt, W.M. de Vos, and K. Venema. 2006. Beyond diversity: functional microbiomics of the human colon. Trends Microbiol. 14:86.
59. Fallingborg, J., L.A. Christensen, M. Ingeman-Nielsen, B.A. Jacobsen, K. Abildgaard, and H.H. Rasmussen. 1989. pH-profile and regional transit times of the normal gut measured by a radiotelemetry device. Aliment Pharmacol. Ther. 3:605–613.
60. Fillon, S.A., J.K. Harris, B.D. Wagner, *et al.* 2012. Novel device to sample the esophageal microbiome—The esophageal string test. PLoS One 7:e42938.
61. Fleet, J.C. 2006. Molecular regulation of calcium metabolism, p. 163–189. *In* C.M. Weaver and R.P. Heaney (ed.), Calcium in Human Health. Humana Press Inc., Totowa, NJ.
62. Fleet, J.C., and R.D. Schoch. 2010. Molecular mechanisms for regulation of intestinal calcium absorption by vitamin D and other factors. Crit. Rev. Clin.l Lab. Sci. 47:181–195.
63. Flint, H.J., S.H. Duncan, K.P. Scott, and P. Louis. 2007. Interactions and competition within the microbial community of the human colon: links between diet and health. Environ. Microbiol. 9:1101–1111.
64. Foca, A., M.C. Liberto, A. Quirino, N. Marascio, E. Zicca, and G. Pavia. 2015. Gut inflammation and immunity: What is the role of the human gut virome? Mediat. Inflamm. Article ID 326032.
65. Fritz, J.V., M.S. Desai, P. Shah, J.G. Schneider, and P. Wilmes. 2013. From meta-omics to causality: experimental models for human microbiome research. Microbiome 1:14.
66. Gaci, N., G. Borrel, W. Tottey, P.W. O'Toole, and J.-F. Brugère. 2014. Archaea and the human gut: New beginning of an old story. World J. Gastroenterol. 20:16062–16078.
67. Gagnaire, A., B. Nadel, D. Raoult, J. Neefjes, and J.-P. Gorvel. 2017. Collateral damage: insights into bacterial mechanisms that predispose host cells to cancer. Nat. Rev. Microbiol. 15:109–128.
68. Ghannoum, M.A., R.J. Jurevic, P.K. Mukherjee, F. Cui, M. Sikaroodi, A. Naqvi, and P.M. Gillevet. 2010. Characterization of the oral fungal microbiome (mycobiome) in healthy individuals. PLoS Pathog 6:e1000713.
69. Gholizadeh, P., H. Eslami, M. Yousefi, M. Asgharzadeh, M. Aghazadeh, and H.S. Kafil. 2016. Role of oral microbiome on oral cancers, a review. Biomed. Pharmacother. 84:552–558.
70. Gill, S.R., M. Pop, R.T. DeBoy, *et al.* 2006. Metagenomic analysis of the human distal gut microbiome. Science 312:1355–1359.
71. Gouba, N., D. Raoult, and M. Drancourt. 2014. Eukaryote culturomics of the gut reveals new species. PLoS One 9:e106994.
72. Gouba, N., and M. Drancourt. 2015. Digestive tract mycobiota: A source of infection. Med. Mal. Infect. 45:9–16.
73. Greenhalgh, K., K.M. Meyer, K.M. Aagaard, and P. Wilmes. 2016. The human gut microbiome in health: establishment and resilience of microbiota over a lifetime. Environ. Microbiol. 18:2103–2116.
74. Grice, E.A., and J.A. Segre. 2012. The human microbiome: Our second genome, p. 151–170. *In* A. Chakravarti and E. Green (ed.), Ann. Rev. Genom. Hum. Genet., vol. 13.
75. Grimaudo, N.J., W.E. Nesbitt, and W.B. Clark. 1996. Coaggregation of *Candida albicans* with oral Actinomyces species. Oral Microbiol. Immunol. 11:59–61.
76. Grimaudo, N.J., and W.E. Nesbitt. 1997. Coaggregation of *Candida albicans* with oral Fusobacterium species. Oral Microbiol. Immunol. 12:168–173.
77. Guarner, F., and J.-R. Malagelada. 2003. Gut flora in health and disease. The Lancet 361:512–519.
78. Hallen-Adams, H.E., S.D. Kachman, J. Kim, R.M. Legge, and I. Martínez. 2015. Fungi inhabiting the healthy human gastrointestinal tract: A diverse and dynamic community. Fung. Ecol. 15:9–17.

79. Hamad, I., D. Raoult, and F. Bittar. 2016. Repertory of eukaryotes (eukaryome) in the human gastrointestinal tract: Taxonomy and detection methods. Parasite Immunol. 38:12–36.
80. Hamaker, B.R., and Y.E. Tuncil. 2014. A perspective on the complexity of dietary fiber structures and their potential effect on the gut microbiota. J. Mol. Biol. 426:3838–3850.
81. Heinritz, S.N., R. Mosenthin, and E. Weiss. 2013. Use of pigs as a potential model for research into dietary modulation of the human gut microbiota. Nutr. Res. Rev. 26:191–209.
82. Hemarajata, P., and J. Versalovic. 2013. Effects of probiotics on gut microbiota: mechanisms of intestinal immunomodulation and neuromodulation. Therapeu. Adv. Gastroenterol. 6:39–51.
83. Hintze, K.J., J.E. Cox, G. Rompato, A.D. Benninghoff, R.E. Ward, J. Broadbent, and M. Lefevre. 2014. Broad scope method for creating humanized animal models for animal health and disease research through antibiotic treatment and human fecal transfer. Gut Microbes 5:183–191.
84. Hoarau, G., P.K. Mukherjee, C. Gower-Rousseau, *et al.* 2016. Bacteriome and mycobiome interactions underscore microbial dysbiosis in familial Crohn's disease. mBio 7:e01250-16.
85. Hoffmann, C., S. Dollive, S. Grunberg, J. Chen, H. Li, G.D. Wu, J.D. Lewis, and F.D. Bushman. 2013. Archaea and fungi of the human gut microbiome: Correlations with diet and bacterial residents. PLoS One 8:e66019.
86. Hold, G.L., S.E. Pryde, V.J. Russell, E. Furrie, and H.J. Flint. 2002. Assessment of microbial diversity in human colonic samples by 16S rDNA sequence analysis. FEMS Microbiol. Ecol. 39:33–39.
87. Hooda, S., B.M.V. Boler, M.C.R. Serao, J.M. Brulc, M.A. Staeger, T.W. Boileau, S.E. Dowd, G.C. Fahey, and K.S. Swanson. 2012. 454 Pyrosequencing reveals a shift in fecal microbiota of healthy adult men consuming polydextrose or soluble corn fiber. The J. Nutr. 142:1259–1265.
88. Howe, A., D.L. Ringus, R.J. Williams, Z.N. Choo, S.M. Greenwald, S.M. Owens, M.L. Coleman, F. Meyer, and E.B. Chang. 2016. Divergent responses of viral and bacterial communities in the gut microbiome to dietary disturbances in mice. ISME J. 10:1217–1227.
89. Hunt, R.H., M. Camilleri, S.E. Crowe, *et al.* 2015. The stomach in health and disease. Gut 64:1650–1668.
90. Huse, S.M., L. Dethlefsen, J.A. Huber, D. Mark Welch, D.A. Relman, and M.L. Sogin. 2008. Exploring microbial diversity and taxonomy using SSU rRNA hypervariable tag sequencing. PLoS Genet. 4:e1000255.
91. Huse, S.M., Y. Ye, Y. Zhou, and A.A. Fodor. 2012. A core human microbiome as viewed through 16S rRNA sequence clusters. PLoS ONE 7:e34242.
92. Huttenhower, C., D. Gevers, R. Knight, *et al.* 2012. Structure, function and diversity of the healthy human microbiome. Nature 486:207–214.
93. Hyde, E.R., B. Luk, S. Cron, *et al.* 2014. Characterization of the rat oral microbiome and the effects of dietary nitrate. Free Radic. Biol. Med. 77:249–257.
94. Iliev, I.D., V.A. Funari, K.D. Taylor, *et al.* 2012. Interactions between commensal fungi and the C-type lectin receptor dectin-1 Influence colitis. Science 336:1314–1317.
95. Jernberg, C., S. Lofmark, C. Edlund, and J.K. Jansson. 2010. Long-term impacts of antibiotic exposure on the human intestinal microbiota. Microbiol. 156:3216–3223.
96. Kararli, T.T. 1995. Comparison of the gastrointestinal anatomy, physiology, and biochemistry of humans and commonly used laboratory animals. Biopharm. Drug Dispos. 16:351–380.
97. Kelly, C.J., L. Zheng, E.L. Campbell, *et al.* 2015. Crosstalk between microbiota-derived short-chain fatty acids and intestinal epithelial HIF augments tissue barrier function. Cell Host Microbe 17:662–671.
98. Kibe, R., M. Sakamoto, H. Yokota, H. Ishikawa, Y. Aiba, Y. Koga, and Y. Benno. 2005. Movement and fixation of intestinal microbiota after administration of human feces to germfree mice. Appl. Environ. Microbiol. 71:3171–3178.
99. Kim, M.-S., and J.-W. Bae. 2016. Spatial disturbances in altered mucosal and luminal gut viromes of diet-induced obese mice. Environ. Microbiol. 18:1498–1510.
100. Kleerebezem, M., and E.E. Vaughan. 2009. Probiotic and gut Lactobacilli and Bifidobacteria: Molecular approaches to study diversity and activity. Ann. Rev. Microbiol. 63:269–290.
101. Koopmans, S.J., and T. Schuurman. 2015. Considerations on pig models for appetite, metabolic syndrome and obese type 2 diabetes: From food intake to metabolic disease. Eur. J. Pharmacol. 759:231–239.
102. Koskella, B., and M.A. Brockhurst. 2014. Bacteria–phage coevolution as a driver of ecological and evolutionary processes in microbial communities. FEMS Microbiol. Rev. 38:916–931.
103. Lam, V., J.D. Su, S. Koprowski, A.N. Hsu, J.S. Tweddell, P. Rafiee, G.J. Gross, N.H. Salzman, and J.E. Baker. 2012. Intestinal microbiota determine severity of myocardial infarction in rats. FASEB J. 26:1727–1735.
104. Lamendella, R., J.W.S. Domingo, S. Ghosh, J. Martinson, and D.B. Oerther. 2011. Comparative fecal metagenomics unveils unique functional capacity of the swine gut. BMC Microbiol. 11:103.
105. Lamendella, R., J.W. Santo Domingo, S. Ghosh, J. Martinson, and D.B. Oerther. 2011. Comparative fecal metagenomics unveils unique functional capacity of the swine gut. BMC Microbiol 11:103.
106. LaTuga, M.S., J.C. Ellis, C.M. Cotton, R.N. Goldberg, J.L. Wynn, R.B. Jackson, and P.C. Seed. 2011. Beyond bacteria: A study of the enteric microbial consortium in extremely low birth weight infants. PLoS ONE 6:e27858.
107. Lazarevic, V., K. Whiteson, S. Huse, D. Hernandez, L. Farinelli, M. Osteras, J. Schrenzel, and P. Francois. 2009. Metagenomic study of the oral microbiota by Illumina high-throughput sequencing. J. Microbiol. Meth. 79:266–271.
108. Leimena, M.M., J. Ramiro-Garcia, M. Davids, *et al.* 2013. A comprehensive metatranscriptome analysis pipeline and its validation using human small intestine microbiota datasets. BMC Genom. 14:530.
109. Lepp, P.W., M.M. Brinig, C.C. Ouverney, K. Palm, G.C. Armitage, and D.A. Relman. 2004. Methanogenic Archaea and human periodontal disease. Proc. Natl. Acad. Sci. U.S.A. 101:6176–6181.
110. Levy, R., and E. Borenstein. 2013. Metabolic modeling of species interaction in the human microbiome elucidates community-level assembly rules. Proc. Natl. Acad. Sci. U.S.A. 110:12804–12809.
111. Ley, R.E., P.J. Turnbaugh, S. Klein, and J.I. Gordon. 2006. Microbial ecology—Human gut microbes associated with obesity. Nature 444:1022–1023.
112. Ley, R.E., M. Hamady, C. Lozupone, *et al.* 2008. Evolution of mammals and their gut microbes. Science 320:1647–1651.
113. Li, M., B. Wang, M. Zhang, *et al.* 2008. Symbiotic gut microbes modulate human metabolic phenotypes. Proc. Natl. Acad. Sci. U.S.A. 105:2117–2122.
114. Li, Q., C. Wang, Q. Zhang, C. Tang, N. Li, B. Ruan, and J. Li. 2012. Use of 18S ribosomal DNA polymerase chain reaction–denaturing gradient gel electrophoresis to study composition of fungal community in 2 patients with intestinal transplants. Hum. Pathol. 43:1273–1281.
115. Li, X.-X., G.L.-H. Wong, K.-F. To, V.W.-S. Wong, L.H. Lai, D.K.-L. Chow, J.Y.-W. Lau, J.J.-Y. Sung, and C. Ding. 2009. Bacterial microbiota profiling in gastritis without *Helicobacter pylori* infection or non-steroidal anti-inflammatory drug use. PLoS One 4:e7985.
116. Lim, E.S., Y.J. Zhou, G.Y. Zhao, *et al.* 2015. Early life dynamics of the human gut virome and bacterial microbiome in infants. Nat. Med. 21:1228–1234.
117. Lim, E.S., D. Wang, and L.R. Holtz. 2016. The bacterial microbiome and virome milestones of infant development. Trends Microbiol. 24:801–810.
118. Looft, T., T.A. Johnson, H.K. Allen, *et al.* 2012. In-feed antibiotic effects on the swine intestinal microbiome. Proc. Natl. Acad. Sci. U.S.A. 109:1691–1696.
119. Lozupone, C.A., J.I. Stombaugh, J.I. Gordon, J.K. Jansson, and R. Knight. 2012. Diversity, stability and resilience of the human gut microbiota. Nature 489:220–230.
120. Luckey, T.D. 1972. Introduction to intestinal microecology Amer. J. Clin. Nutr. 25:1292–1294.
121. Lynch, S.V., and O. Pedersen. 2016. The human intestinal microbiome in health and disease. New Engl. J. Med. 375:2369–2379.
122. Maathuis, A., A. Hoffman, A. Evans, L. Sanders, and K. Venema. 2009. The effect of the undigested fraction of maize products on the activity and composition of the microbiota determined in a dynamic *in vitro* model of the human proximal large intestine. J. Am. Coll. Nutr. 28:657–666.

123. Macfarlane, G.T., and J.H. Cummings. 1991. The colonic flora, fermentation and large bowel digestive function. Raven Press, New York, N.Y.
124. Manrique, P., B. Bolduc, S.T. Walk, J. van der Oost, W.M. de Vos, and M.J. Young. 2016. Healthy human gut phageome. Proc. Natl. Acad. Sci. U.S.A. 113:10400–10405.
125. Mao, S.Y., C.F. Yang, and W.Y. Zhu. 2011. Phylogenetic analysis of methanogens in the pig feces. Curr. Microbiol. 62:1386–1389.
126. Mar Rodríguez, M., D. Pérez, F. Javier Chaves, *et al.* 2015. Obesity changes the human gut mycobiome. Sci. Rep. 5:14600.
127. Martinez, I., J. Kim, P.R. Duffy, V.L. Schlegel, and J. Walter. 2010. Resistant starches types 2 and 4 have differential effects on the composition of the fecal microbiota in human subjects. PLoS One 5:e15046.
128. McConnell, E.L., A.W. Basit, and S. Murdan. 2008. Measurements of rat and mouse gastrointestinal pH, fluid and lymphoid tissue, and implications for in-vivo experiments. J. Pharm. Pharmacol. 60:63–70.
129. McHardy, I.H., M. Goudarzi, M. Tong, *et al.* 2013. Integrative analysis of the microbiome and metabolome of the human intestinal mucosal surface reveals exquisite inter-relationships. Microbiome 1:17.
130. Merchant, H.A., E.L. McConnell, F. Liu, C. Ramaswamy, R.P. Kulkarni, A.W. Basit, and S. Murdan. 2011. Assessment of gastrointestinal pH, fluid and lymphoid tissue in the guinea pig, rabbit and pig, and implications for their use in drug development. Eur. J. Pharm. Sci. 42:3–10.
131. Mihajlovski, A., M. Alric, and J.-F. BrugËre. 2008. A putative new order of methanogenic Archaea inhabiting the human gut, as revealed by molecular analyses of the mcrA gene. Res. Microbiol. 159:516–521.
132. Mikkelsen, L.L., C. Bendixen, M. Jakobsen, and B.B. Jensen. 2003. Enumeration of bifidobacteria in gastrointestinal samples from piglets. Appl. Environ. Microbiol. 69:654–658.
133. Mills, S., F. Shanahan, C. Stanton, C. Hill, A. Coffey, and R.P. Ross. 2013. Movers and shakers: Influence of bacteriophages in shaping the mammalian gut microbiota. Gut Microbes 4:4–16.
134. Minot, S., R. Sinha, J. Chen, H.Z. Li, S.A. Keilbaugh, G.D. Wu, J.D. Lewis, and F.D. Bushman. 2011. The human gut virome: Inter-individual variation and dynamic response to diet. Genome Res. 21:1616–1625.
135. Mitsuoka, T. 1992. Intestinal flora and aging. Nutr. Rev. 50:438–446.
136. Mukherjee, P.K., J. Chandra, M. Retuerto, *et al.* 2014. Oral mycobiome analysis of HIV-infected patients: Identification of *Pichia* as an antagonist of opportunistic fungi. PLoS Pathog. 10:e1003996.
137. Nguyen, T.L.A., S. Vieira-Silva, A. Liston, and J. Raes. 2015. How informative is the mouse for human gut microbiota research? Dis. Mod. Mech. 8:1–16.
138. Nicholson, J.K., E. Holmes, and I.D. Wilson. 2005. Gut microorganisms, mammalian metabolism and personalized health care. Nat. Rev. Microbiol. 3:431–438.
139. Nielsen, H.B., M. Almeida, A.S. Juncker, *et al.* 2014. Identification and assembly of genomes and genetic elements in complex metagenomic samples without using reference genomes. Nat. Biotechnol. 32:822–828.
140. Noble, S.M., B.A. Gianetti, and J.N. Witchley. 2017. Candida albicans cell-type switching and functional plasticity in the mammalian host. Nat. Rev. Microbiol. 15:96–108.
141. O'Hara, A.M., and F. Shanahan. 2006. The gut flora as a forgotten organ. EMBO Rep. 7:688–693.
142. Odle, J., X. Lin, S.K. Jacobi, S.W. Kim, and C.H. Stahl. 2014. The suckling piglet as an agrimedical model for the study of pediatric nutrition and metabolism. Ann. Rev. Anim. Biosci. 2:419–444.
143. Ott, S.J., T. Kuhbacher, M. Musfeldt, *et al.* 2008. Fungi and inflammatory bowel diseases: Alterations of composition and diversity. Scand. J. Gastroenterol. 43:831–841.
144. Pang, X.Y., X.G. Hua, Q. Yang, D.H. Ding, C.Y. Che, L. Cui, W. Jia, P. Bucheli, and L.P. Zhao. 2007. Inter-species transplantation of gut microbiota from human to pigs. ISME J. 1:156–162.
145. Parfrey, L.W., W.A. Walters, and R. Knight. 2011. Microbial eukaryotes in the human microbiome: ecology, evolution, and future directions. Front. Microbiol. 2:153.
146. Pedersen, R., H.C. Ingerslev, M. Sturek, M. Alloosh, S. Cirera, B.O. Christoffersen, S.G. Moesgaard, N. Larsen, and M. Boye. 2013. Characterisation of gut microbiota in Ossabaw and Gottingen minipigs as models of obesity and metabolic syndrome. PLoS One 8:e 0056612.
147. Pei, Z.H., E.J. Bini, L.Y. Yang, M.S. Zhou, F. Francois, and M.J. Blaser. 2004. Bacterial biota in the human distal esophagus. Proc. Natl. Acad. Sci. U.S.A. 101:4250–4255.
148. Pflughoeft, K.J., and J. Versalovic. 2012. Human Microbiome in Health and Disease. Ann. Revi. Pathol. 7:99–122.
149. Pieper, R., P. Janczyk, A. Zeyner, H. Smidt, V. Guiard, and W.B. Souffrant. 2008. Ecophysiology of the developing total bacterial and *Lactobacillus* communities in the terminal small intestine of weaning piglets. Microb. Ecol. 56:474–483.
150. Pothoulakis, C. 2009. Review article: anti-inflammatory mechanisms of action of *Saccharomyces boulardii*. Alimen. Pharmacol. Therapeut. 30:826–833.
151. Pride, D.T., J. Salzman, M. Haynes, F. Rohwer, C. Davis-Long, R.A. White, III, P. Loomer, G.C. Armitage, and D.A. Relman. 2012. Evidence of a robust resident bacteriophage population revealed through analysis of the human salivary virome. ISME J. 6:915–926.
152. Qin, J., R. Li, J. Raes, *et al.* 2010. A human gut microbial gene catalogue established by metagenomic sequencing. Nature 464:59–65.
153. Qiu, X.Y., F. Zhang, X. Yang, N. Wu, W.W. Jiang, X. Li, X.X. Li, and Y.L. Liu. 2015. Changes in the composition of intestinal fungi and their role in mice with dextran sulfate sodium-induced colitis. Sci. Rep. 5:10416.
154. Queipo-Ortuno, M.I., L.M. Seoane, M. Murri, M. Pardo, J.M. Gomez-Zumaquero, F. Cardona, F. Casanueva, and F.J. Tinahones. 2013. Gut microbiota composition in male rat models under different nutritional status and physical activity and its association with serum leptin and ghrelin levels. PLoS One 8:e65465.
155. Rajilic-Stojanovic, M., H.G.H.J. Heilig, S. Tims, E.G. Zoetendal, and W.M. de Vos. 2013. Long-term monitoring of the human intestinal microbiota composition. Environ. Microbiol. 15:1146–1159.
156. Rajilic-Stojanovic, M., and W.M. de Vos. 2014. The first 1000 cultured species of the human gastrointestinal microbiota. FEMS Microbiol. Rev. 38:996–1047.
157. Rascovan, N., R. Duraisamy, and C. Desnues. 2016. Metagenomics and the human virome in asymptomatic individuals, p. 125–141. *In* S. Gottesman (ed.), Ann. Rev. Microbiol. 70:125–141.
158. Rastall, R.A., G.R. Gibson, H.S. Gill, F. Guarner, T.R. Klaenhammer, B. Pot, G. Reid, I.R. Rowland, and M.E. Sanders. 2005. Modulation of the microbial ecology of the human colon by probiotics, prebiotics and synbiotics to enhance human health: An overview of enabling science and potential applications. FEMS Microbiol. Ecol. 52:145.
159. Rauch, M., and S.V. Lynch. 2012. The potential for probiotic manipulation of the gastrointestinal microbiome. Curr. Opin. Biotechnol. 23:192–201.
160. Reyes, A., M. Haynes, N. Hanson, F.E. Angly, A.C. Heath, F. Rohwer, and J.I. Gordon. 2010. Viruses in the faecal microbiota of monozygotic twins and their mothers. Nature 466:334–338.
161. Reyes, A., N.P. Semenkovich, K. Whiteson, F. Rohwer, and J.I. Gordon. 2012. Going viral: next-generation sequencing applied to phage populations in the human gut. Nat. Rev. Microbiol. 10:607–617.
162. Reyes, A., L.V. Blanton, S. Cao, *et al.* 2015. Gut DNA viromes of Malawian twins discordant for severe acute malnutrition. Proc. Natl. Acad. Sci. U.S.A. 112:11941–11946.
163. Robles-Sikisaka, R., M. Ly, T. Boehm, M. Naidu, J. Salzman, and D.T. Pride. 2013. Association between living environment and human oral viral ecology. ISME J. 7:1710–1724.
164. Sachsenroder, J., S.O. Twardziok, M. Scheuch, and R. Johne. 2014. The general composition of the faecal virome of pigs depends on age, but not on feeding with a probiotic bacterium. PLoS One 9:e88888.
165. Savage, D.C. 1977. Microbial ecology of gastrointestinal-tract. Ann. Revi. Microbiol. 31:107–133.
166. Scanlan, P.D., F. Shanahan, C. O'Mahony, and J.R. Marchesi. 2006. Culture-independent analyses of temporal variation of the dominant fecal microbiota and targeted bacterial subgroups in Crohn's Disease. J. Clin. Microbiol. 44:3980–3988.
167. Schwartz, S., I. Friedberg, I.V. Ivanov, *et al.* 2012. A metagenomic study of diet-dependent interaction between gut microbiota and host in infants reveals differences in immune response. Genome Biol. 13:r32.
168. Scupham, A.J., L.L. Presley, B. Wei, *et al.* 2006. Abundant and diverse fungal microbiota in the murine intestine. Appl. Environ. Microbiol. 72:793–801.
169. Seed, P.C. 2015. The Human Mycobiome. Cold Spring Harbor Perspectives in Medicine 5.

170. Sender, R., S. Fuchs, and R. Milo. 2016. Revised estimates for the number of human and bacteria cells in the body. PLoS Biol. 14:e1002533.
171. Shan, T.L., L.L. Li, P. Simmonds, C.L. Wang, A. Moeser, and E. Delwart. 2011. The fecal virome of pigs on a high-density farm. J. Virol. 85:11697–11708.
172. Simmonds, P., M.J. Adams, M. Benko, *et al.* 2017. Virus taxonomy in the age of metagenomics. Nat. Rev. Microbiol. 15:161–168.
173. Sogin, M.L., H.G. Morrison, J.A. Huber, D.M. Welch, S.M. Huse, P.R. Neal, J.M. Arrieta, and G.J. Herndl. 2006. Microbial diversity in the deep sea and the underexplored "rare biosphere". Proc. Natl. Acad. Sci. U.S.A. 103:12115–12120.
174. Sommer, F., and F. Bäckhed. 2016. Know your neighbor: Microbiota and host epithelial cells interact locally to control intestinal function and physiology. BioEssays 38:455–464.
175. Sonnenburg, E.D., S.A. Smits, M. Tikhonov, S.K. Higginbottom, N.S. Wingreen, and J.L. Sonnenburg. 2016. Diet-induced extinctions in the gut microbiota compound over generations. Nature 529:212–215.
176. Sonnenburg, J.L., and F. Backhed. 2016. Diet-microbiota interactions as moderators of human metabolism. Nature 535:56–64.
177. Spurlock, M.E., and N.K. Gabler. 2008. The development of porcine models of obesity and the metabolic syndrome. J. Nutr. 138:397–402.
178. Stephen, A.M., and J.H. Cummings. 1980. Mechanism of action of dietary fibre in the human colon. Nature 284:283–284.
179. Su, Y., G.R. Bian, Z.G. Zhu, H. Smidt, and W.Y. Zhu. 2014. Early methanogenic colonisation in the faeces of Meishan and Yorkshire piglets as determined by pyrosequencing analysis. Archaea 2014:547908.
180. Suau, A., R. Bonnet, M. Sutren, J.J. Godon, G.R. Gibson, M.D. Collins, and J. Dore. 1999. Direct analysis of genes encoding 16S rRNA from complex communities reveals many novel molecular species within the human gut. Appl. Environ. Microbiol. 65:4799–4807.
181. Suhr, M.J., and H.E. Hallen-Adams. 2015. The human gut mycobiome: pitfalls and potentials—a mycologist's perspective. Mycologia 107:1057–1073.
182. Suhr, M.J., N. Banjara, and H.E. Hallen-Adams. 2016. Sequence-based methods for detecting and evaluating the human gut mycobiome. Lett. Appl. Microbiol. 62:209–215.
183. Tannock, G.W. 2002. Exploring the relationships between intestinal microflora and inflammatory conditions of the human bowel and spine. Anton. Leeuw. Int. J. G. 81:529–535.
184. Tannock, G.W. 2008. Molecular analysis of the intestinal microflora in IBD. Mucosal Immunol. 1:S15–S18.
185. Thaiss, C.A., D. Zeevi, M. Levy, *et al.* 2014. Transkingdom control of microbiota diurnal oscillations promotes metabolic homeostasis. Cell 159:514–529.
186. Thaiss, C.A., D. Zeevi, M. Levy, E. Segal, and E. Elinav. 2015. A day in the life of the meta-organism: diurnal rhythms of the intestinal microbiome and its host. Gut Microbes 6:137–142.
187. Thaiss, C.A., M. Levy, T. Korem, *et al.* 2016. Microbiota diurnal rhythmicity programs host transcriptome oscillations. Cell 167: 1495–1510.e12.
188. Thaiss, C.A., N. Zmora, M. Levy, and E. Elinav. 2016. The microbiome and innate immunity. Nature 535:65–74.
189. The Human Microbiome Project Consortium. 2012. Structure, function and diversity of the healthy human microbiome. Nature 486:207–214.
190. Turnbaugh, P., R. Ley, M.A. Mahowald, V. Magrini, E.R. Mardis, and J. Gordon. 2006. An obesity-associated gut microbiome with increased capacity for energy harvest. Nature 444:1027–1131.
191. Turnbaugh, P.J., and J.I. Gordon. 2009. The core gut microbiome, energy balance and obesity. J. Physiol. 587:4153–4158.
192. Turnbaugh, P.J., V.K. Ridaura, J.J. Faith, F.E. Rey, R. Knight, and J.I. Gordon. 2009. The effect of diet on the human gut microbiome: A metagenomic analysis in humanized gnotobiotic mice. Sci. Trans. Med. 1:6ra14–6ra14.
193. Underhill, D.M., and L.D. Lliev. 2014. The mycobiota: interactions between commensal fungi and the host immune system. Nat. Rev. Immunol. 14:405–416.
194. Urubschurov, V., P. Janczyk, W.B. Souffrant, G. Freyer, and A. Zeyner. 2011. Establishment of intestinal microbiota with focus on yeasts of unweaned and weaned piglets kept under different farm conditions. FEMS Microbiol. Ecol. 77:493–502.
195. van den Bogert, B., O. Erkus, J. Boekhorst, M. de Goffau, E.J. Smid, E.G. Zoetendal, and M. Kleerebezem. 2013. Diversity of human small intestinal *Streptococcus* and *Veillonella* populations. FEMS Microbiol. Ecol. 85:376–388.
196. Vdoviaková, K., E. Petrovová, M. Maloveská, L. Krešáková, J. Teleky, M.Z.J. Elias, and D. Petrášová. 2016. Surgical anatomy of the gastrointestinal tract and its vasculature in the laboratory rat. Gastroenterol. Res. Pract. 2016, Article ID 2632368.
197. von Rosenvinge, E.C., Y. Song, J.R. White, C. Maddox, T. Blanchard, and W.F. Fricke. 2013. Immune status, antibiotic medication and pH are associated with changes in the stomach fluid microbiota. ISME J. 7:1354–1366.
198. Walker, A.W., J. Ince, S.H. Duncan, *et al.* 2011. Dominant and diet-responsive groups of bacteria within the human colonic microbiota. ISME J. 5:220–230.
199. Waller, A.S., T. Yamada, D.M. Kristensen, J.R. Kultima, S. Sunagawa, E.V. Koonin, and P. Bork. 2014. Classification and quantification of bacteriophage taxa in human gut metagenomes. ISME J. 8:1391–1402.
200. Wampach, L., A. Heintz-Buschart, A. Hogan, *et al.* 2017. Colonization and succession within the human gut microbiome by Archaea, Bacteria, and Microeukaryotes during the first year of life. Front. Microbiol. 8:738.
201. Wang, M., S. Ahrn, B. Jeppsson, and G. Molin. 2005. Comparison of bacterial diversity along the human intestinal tract by direct cloning and sequencing of 16S rRNA genes. FEMS Microbiol. Ecol. 54:219.
202. Wang, M., and S.M. Donovan. 2015. Human microbiota-associated swine: Current progress and future opportunities. ILAR J. 56:63–73.
203. Wang, X., S.P. Heazlewood, D.O. Krause, and T.H.J. Florin. 2003. Molecular characterization of the microbial species that colonize human ileal and colonic mucosa by using 16S rDNA sequence analysis. J. Appl. Microbiol. 95:508–520.
204. Wang, Z.K., Y.S. Yang, A.T. Stefka, G. Sun, and L.H. Peng. 2014. Review article: fungal microbiota and digestive diseases. Alimen. Pharmacol. Therapeu. 39:751–766.
205. Warden, C.H., and J.S. Fisler. 2008. Comparisons of diets used in animal models of high fat feeding. Cell Metabol. 7:277.
206. Weaver, C.M., L.D. McCabe, G.M. McCabe, R. Novotny, M.D. Van Loan, S.B. Going, C. Boushey, D.A. Savaiano, and V. Matkovic. 2005. Bone mineral and predictors of whole body, total hip, and lumbar spine for 740 early pubertal white, Hispanic, and Asian girls. J. Bone Miner. Res. 20:S314–S314.
207. Weaver, C.M., B.R. Martin, J.A. Story, I. Hutchinson, and L. Sanders. 2010. Novel fibers Increase bone calcium content and strength beyond efficiency of large intestine fermentation. J. Ag. Food Chem. 58:8952–8957.
208. Weaver, C.M., B.R. Martin, C.H. Nakatsu, *et al.* 2011. Galactooligosaccharides improve mineral absorption and bone properties in growing rats through gut fermentation. J. Ag. Food Chem. 59:6501–6510.
209. Welch, J.L.M., B.J. Rossetti, C.W. Rieken, F.E. Dewhirst, and G.G. Borisy. 2016. Biogeography of a human oral microbiome at the micron scale. Proc. Natl. Acad. Sci. U.S.A. 113:E791–E800.
210. Whisner, C.M., B.R. Martin, C.H. Nakatsu, G.P. McCabe, L.D. McCabe, M. Peacock, and C.M. Weaver. 2014. Soluble maize fibre affects short-term calcium absorption in adolescent boys and girls: a randomised controlled trial using dual stable isotopic tracers. Brit. J. Nutr. 112:446–456.
211. Whisner, C.M., B.R. Martin, C.H. Nakatsu, J.A. Story, C.J. MacDonald-Clarke, L.D. McCabe, G.P. McCabe, and C.M. Weaver. 2016. Soluble corn fiber increases calcium absorption associated with shifts in the gut microbiome: A randomized dose-response trial in free-living pubertal girls. Nutr. J. 146:1298–1306
212. Wilson, K., and R. Blitchington. 1996. Human colonic biota studied by ribosomal DNA sequence analysis. Appl. Environ. Microbiol. 62:2273–2278.
213. Woodmansey, E.J. 2007. Intestinal bacteria and ageing. J. Appl. Microbiol. 102:1178–1186.
214. Wright, C.J., L.H. Burns, A.A. Jack, C.R. Back, L.C. Dutton, A.H. Nobbs, R.J. Lamont, and H.F. Jenkinson. 2013. Microbial interactions in building of communities. Mol. Oral Microbiol. 28:83–101.
215. Wylie, K.M., K.A. Mihindukulasuriya, Y. Zhou, E. Sodergren, G.A. Storch, and G.M. Weinstock. 2014. Metagenomic analysis of double-stranded DNA viruses in healthy adults. BMC Biol. 12:71.
216. Xiao, L., Q. Feng, S. Liang, *et al.* 2015. A catalog of the mouse gut metagenome. Nat Biotechnol. 33:1103–1108.

217. Xiao, L., J. Estellé, P. Kiilerich, *et al.* 2016. A reference gene catalogue of the pig gut microbiome. Nat. Microbiol. 1:16161.
218. Yan, H., R. Potu, H. Lu, *et al.* 2013. Dietary fat content and fiber type modulate hind gut microbial community and metabolic markers in the pig. PLoS One 8:e59581.
219. Yatsunenko, T., F.E. Rey, M.J. Manary, *et al.* 2012. Human gut microbiome viewed across age and geography. Nature 486:222–227.
220. Zhang, Q., G. Widmer, and S. Tzipori. 2013. A pig model of the human gastrointestinal tract. Gut Microbes 4:193–200.
221. Zhu, Y.Y., H. Li, X.L. Xu, C.B. Li, and G.H. Zhou. 2016. The gut microbiota in young and middle-aged rats showed different responses to chicken protein in their diet. BMC Microbiol. 16:281.
222. Zoetendal, E.G., A. von Wright, T. Vilpponen-Salmela, K. Ben-Amor, A.D.L. Akkermans, and W.M. de Vos. 2002. Mucosa-associated bacteria in the human gastrointestinal tract are uniformly distributed along the colon and differ from the community recovered from feces. Appl. Environ. Microbiol. 68:3401–3407.
223. Zoetendal, E.G., J. Raes, B. van den Bogert, *et al.* 2012. The human small intestinal microbiota is driven by rapid uptake and conversion of simple carbohydrates. ISME J. 6:1415–1426.
224. Zwolinska-Wcislo, M., A. Budak, D. Trojanowska, J. Bogdal, and J. Stachura. 1998. Fungal colonization of the stomach and its clinical relevance. Mycoses 41:327–334.

Microbes and Environments

EDITORIAL BOARD (2025)

Editor-in-Chief

Shigeto Otsuka, *The University of Tokyo, Tokyo, Japan*

Senior Editors

Makoto Ikenaga, *Kagoshima University, Kagoshima, Japan*
Satoshi Ishii, *University of Minnesota, St. Paul, MN, USA*
Yoichi Kamagata, *National Institute of Advanced Industrial Science and Technology (AIST)/Hokkaido University, Sapporo, Japan*
Kiwamu Minamisawa, *Tohoku University, Sendai, Japan*
Yuji Nagata, *Tohoku University, Sendai, Japan*
Kazuhiko Narisawa, *Ibaraki University, Ami, Japan*
Nobuhiko Nomura, *University of Tsukuba, Tsukuba, Japan*
Takuro Nunoura, *Japan Agency for Marine-Earth Science and Technology (JAMSTEC), Yokosuka, Japan*
Shino Suzuki, *Institute of Physical and Chemical Research (RIKEN), Wako, Japan*
Sen-Lin Tang, *Academia Sinica, Taipei, Taiwan*
Naoko Yoshida, *Nagoya University, Nagoya, Japan*

Managing Editors

Michinari Sunamura (General affairs & accounting), *The University of Tokyo, Tokyo, Japan*
Shino Suzuki (General affairs/Manuscript Reception), *Institute of Physical and Chemical Research (RIKEN), Wako, Japan*
Shunichi Ishii (Online Submission System), *Japan Agency for Marine-Earth Science and Technology (JAMSTEC), Yokosuka, Japan*
Satoshi Nakagawa (Online Submission System), *Kyoto University, Kyoto, Japan*
Takashi Narihiro (Information Analysis), *National Institute of Advanced Industrial Science and Technology (AIST), Sapporo, Japan*
Satoko Shinozaki-Yabana, (Manuscript Edition/Manuscript Reception), *Japan Science and Technology Agency (JST), Tokyo, Japan*
Eri Hara (Information analysis), *National Institute of Advanced Industrial Science and Technology (AIST), Tsukuba, Japan*

Associate Editors

Yoshiteru Aoi, *Hiroshima University, Higashi-Hiroshima, Japan*
Peter Dunfield, *University of Calgary, Calgary, Canada*
Manabu Fukui, *Hokkaido University, Sapporo, Japan*
Natsuko Hamamura, *Kyushu University, Fukuoka, Japan*
Yasufumi Hikichi, *Kochi University, Nankoku, Japan*
Yuichi Hongoh, *Tokyo Institute of Technology, Tokyo, Japan*
Shir-Ly Huang, *National Yang Ming Chiao Tung University, Taipei, Taiwan*
Shun'ichi Ishii, *Japan Agency for Marine-Earth Science and Technology (JAMSTEC), Yokosuka, Japan*
Tsukasa Ito, *Gunma University, Kiryu, Japan*
Wataru Iwasaki, *The University of Tokyo, Kashiwa, Japan*
Yasuhiro Kasahara, *Hokkaido University, Sapporo, Japan*
Takafumi Kataoka, *Fukui Prefectural University, Obama, Japan*
Yoshitomo Kikuchi, *National Institute of Advanced Industrial Science and Technology (AIST), Sapporo, Japan*
Hiroyuki Kimura, *Shizuoka University, Shizuoka, Japan*
Tomonori Kindaichi, *Hiroshima University, Higashi-Hiroshima, Japan*
Masaaki Konishi, *Kitami Institute of Technology, Kitami, Japan*
Chi-Te Liu, *National Taiwan University, Taipei, Taiwan*
Fumito Maruyama, *Hiroshima University, Higashi-Hiroshima, Japan*
Shawn McGlynn, *Tokyo Institute of Technology, Tokyo, Japan*
Keisuke Miyauchi, *Tohoku Gakuin University, Tagajo, Japan*
Ryo Miyazaki, *National Institute of Advanced Industrial Science and Technology (AIST), Tsukuba, Japan*
Atsushi Nakabachi, *Toyohashi University of Technology, Toyohashi, Japan*
Satoshi Nakagawa, *Kyoto University, Kyoto, Japan*
Tatsunori Nakagawa, *Nihon University, Fujisawa, Japan*
Shin-ichi Nakano, *Kyoto University, Otsu, Japan*
Tomoyasu Nishizawa, *Ibaraki University, Ami, Japan*
Naohiro Noda, *National Institute of Advanced Industrial Science and Technology (AIST), Tsukuba, Japan*
Satoko Noda, *Ibaraki University, Mito, Japan*
Ryo Ohtomo, *Institute for Agro-Environmental Sciences, NARO, Tsukuba, Japan*
Shin Okazaki, *Tokyo University of Agriculture and Technology, Fuchu, Japan*
Torahiko Okubo, *Hokkaido University, Sapporo, Japan*
Katsuharu Saito, *Shinshu University, Kamiina, Japan*
Shusei Sato, *Tohoku University, Sendai, Japan*
Toru Shigematsu, *Niigata University of Pharmacy and Medical and Life Sciences, Niigata, Japan*
Masafumi Shimizu, *Gifu University, Gifu, Japan*
Masaki Shintani, *Shizuoka University, Hamamatsu, Japan*
Mio Takeuchi, *National Institute of Advanced Industrial Science and Technology (AIST), Osaka, Japan*
Hideyuki Tamaki, *National Institute of Advanced Industrial Science and Technology (AIST), Tsukuba, Japan*
Masaru Usui, *Rakuno Gakuen University, Ebetsu, Japan*
Takayuki Wada, *Osaka Metropolitan University, Osaka, Japan*
Satoshi Wakai, *Japan Agency for Marine-Earth Science and Technology (JAMSTEC), Yokosuka, Japan*
Akinori Yabuki, *Japan Agency for Marine-Earth Science and Technology (JAMSTEC), Yokosuka, Japan*
Rie Yatsunami, *Tokyo Institute of Technology, Yokohama, Japan*
Takashi Yoshida, *Kyoto University, Kyoto, Japan*
Susumu Yoshizawa, *The University of Tokyo, Kashiwa, Japan*

JAPANESE SOCIETY OF MICROBIAL ECOLOGY

President (2025-2027): Hiroyuki FUTAMATA, Shizuoka University
Secretary/Treasurer (2025-2027): Hiroyuki KIMURA, Shizuoka University
Business affairs (2025-2027): Masaki SHINTANI, Shizuoka University
Business affairs (2025-2027): Natsuko HAMAMURA, Kyushu University
Business affairs (2025-2027): Takashi NARIHIRO, National Institute of Advanced Industrial Science and Technology (AIST)
Treasurer (2025-2027): Yosuke TASHIRO, Shizuoka University
Treasurer (2025-2027): Michihiro ITO, University of The Ryukyus
Treasurer (2025-2027): Tomo AOYAGI, National Institute of Advanced Industrial Science and Technology (AIST)

JAPANESE SOCIETY OF SOIL MICROBIOLOGY

President (2023–2025): Masahiro SHISHIDO, Chiba University
Vice President (2023–2025): Koki TOYOTA, Tokyo University of Agriculture and Technology
Secretary (2024–2026): Ryo OHTOMO, Institute for Agro-Environmental Sciences, NARO
Treasurer (2024–2026): Hiroyuki SEKIGUCHI, Institute for Plant Protection, NARO

TAIWAN SOCIETY OF MICROBIAL ECOLOGY

President (2024-): Jyh-Yih Leu, Fu Jen Catholic University
Vice President (2024-): Chi-Te Liu, National Taiwan University
Secretary-General (2024-): Hui-Ping Chuang, National Cheng Kung University
International Business Affairs (2024-):Chang-Ping Yu, National Taiwan University
International Business Affairs (2024-): Shir-Ly HUANG, National Yang Ming Chiao Tung University

JAPANESE SOCIETY OF PLANT MICROBE INTERACTIONS

Chair of Society: Toshiki UCHIUMI, Kagoshima University
Secretary-General: Shusei SATO, Tohoku University
Treasurer: Hisayuki MITSUI, Tohoku University
General affairs: Masayuki SUGAWARA, Obihiro University of Agriculture and Veterinary Medicine

JAPANESE SOCIETY FOR EXTREMOPHILES

President (2024–2025): Yoshizumi ISHINO, Kyushu University
Vice President (2024–2025): Satoshi NAKAMURA, Tokyo Institute of Technology
Vice President (2024–2025): Masahiro ITO, Toyo University

This journal is partly subsidized by the Grant-in-Aid for Publication of Scientific Research Results from the Japan Society for the Promotion of Science (JSPS).

c/o Nakanishi Printing Co.
Shimodachiuri Ogawa Higashi Kamikyoku, Kyoto 602-8048, Japan TEL: +81–75–441–3155, FAX: +81–75–417–2050