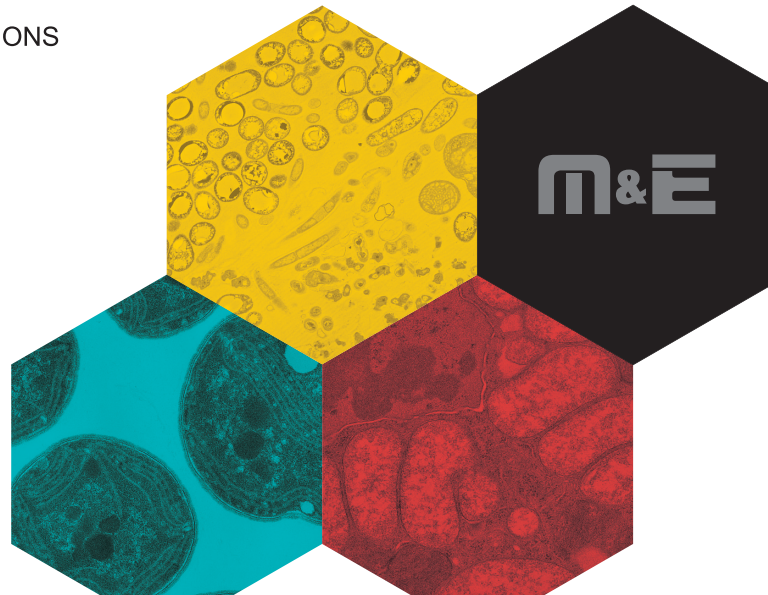
2023 Impact Factor
2.1
5 year IF 2.6

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



October 2024

Microbes and Environments
DIGEST 2024

CONTENTS

*Times cited at 2024/6/27

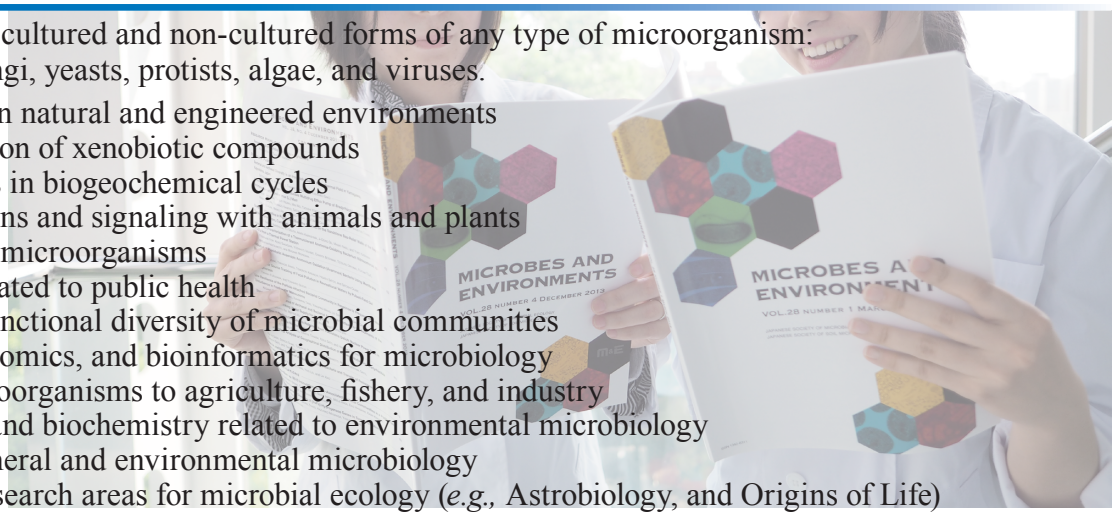
About Microbes and Environments

Microbes and Environments (M&E) is an international, peer-reviewed, 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 a printed and 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. 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

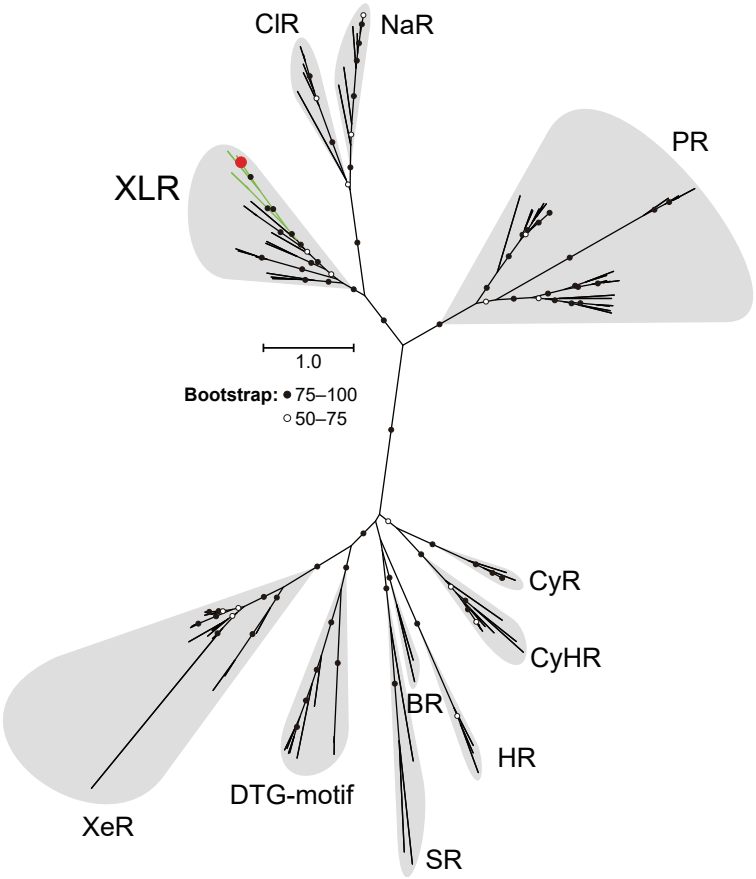


Fig. 1. Phylogenetic position of diatom rhodopsin. A maximum likelihood tree of the amino acid sequences of microbial rhodopsins. Diatom rhodopsin (PngR) is indicated by a red circle and bootstrap probabilities (≥50%) by black and white circles. Green branches indicate eukaryotic rhodopsins used in this 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).

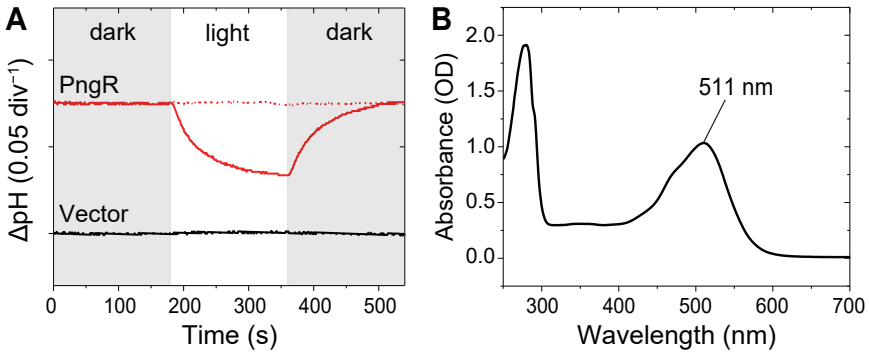


Fig. 2. Light-induced pH changes and absorption spectrum of PngR. (A) Outward proton pump activity of PngR in *E. coli* cells. Light-induced pH changes in solutions containing *E. coli* cells with the expression plasmid for PngR (upper panel) and the empty vector pET21a (lower panel) in the presence (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).

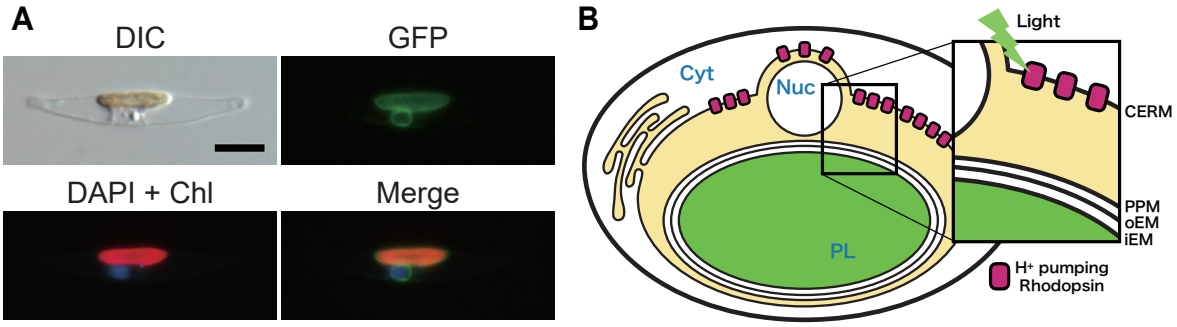


Fig. 3. Subcellular localization of rhodopsins in diatom cells. (A) A transformed diatom cell was observed with differential interface contrast (DIC) (Upper left). Green fluorescence from recombinant PngR (GFP) (Upper right). Nuclear DNA stained with DAPI and chlorophyll autofluorescence (DAPI + Chl) 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

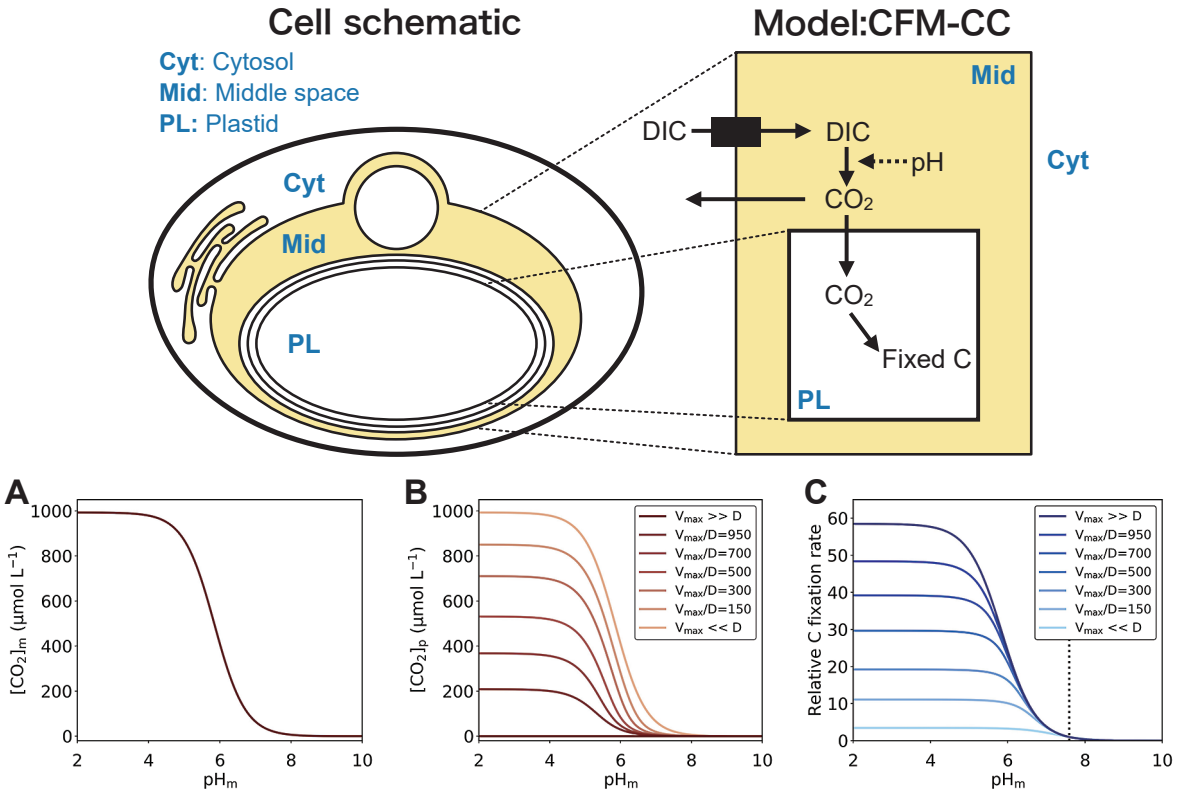


Fig. 4. A quantitative model of the concentration of carbon in diatoms. (Upper panel) Schematic of a Cell Flux Model of C Concentration (CFM-CC). The left panel represents the actual cell, while the right panel represents the model. Solid arrows show the net flux of C and the dashed arrow indicates the 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. 37(1), 2022
https://www.jstage.jst.go.jp/browse/jsme2 doi:10.1264/jsme2.ME21094

MOST VALUABLE PAPER OF THE YEAR 2022

Phenolic Acids Induce Nod Factor Production in *Lotus japonicus–Mesorhizobium* Symbiosis

Masayuki Shimamura[1†], Takashi Kumaki[2†], Shun Hashimoto[2], Kazuhiko Saeki[3], Shin-ichi Ayabe[1], Atsushi Higashitani[2], Tomoyoshi Akashi[1*], Shusei Sato[2*], and Toshio Aoki[1]

[1]*Department of Applied Biological Sciences, Nihon University, Fujisawa, Kanagawa 252–0880, Japan;* [2]*Graduate School of Life Sciences, Tohoku University, Sendai, Miyagi 980–8577, Japan; and* [3]*Department of Biological Sciences and Kyousei Science Center for Life and Nature, Nara Women's University, Nara 630–8506, Japan*

(Received December 28, 2021—Accepted January 22, 2022—Published online March 12, 2022)

In legume–rhizobia symbiosis, partner recognition and the initiation of symbiosis processes require the mutual exchange of chemical signals. Chemicals, generally (iso)flavonoids, in the root exudates of the host plant induce the expression of *nod* genes in rhizobia, and, thus, are called *nod* gene inducers. The expression of *nod* genes leads to the production of lipochitooligosaccharides (LCOs) called Nod factors. Natural *nod* gene inducer(s) in *Lotus japonicus–Mesorhizobium* symbiosis remain unknown. Therefore, we developed an LCO detection method based on ultra-high-performance liquid chromatography–tandem-quadrupole mass spectrometry (UPLC-TQMS) to identify these inducers and used it herein to screen 40 phenolic compounds and aldonic acids for their ability to induce LCOs in *Mesorhizobium japonicum* MAFF303099. We identified five phenolic acids with LCO-inducing activities, including *p*-coumaric, caffeic, and ferulic acids. The induced LCOs caused root hair deformation, and nodule numbers in *L. japonicus* inoculated with *M. japonicum* were increased by these phenolic acids. The three phenolic acids listed above induced the expression of the *nodA*, *nodB*, and *ttsI* genes in a strain harboring a multicopy plasmid encoding NodD1, but not that encoding NodD2. The presence of *p*-coumaric and ferulic acids in the root exudates of *L. japonicus* was confirmed by UPLC-TQMS, and the induction of *ttsI*::*lacZ* in the strain harboring the *nodD1* plasmid was detected in the rhizosphere of *L. japonicus*. Based on these results, we propose that phenolic acids are a novel type of *nod* gene inducer in *L. japonicus–Mesorhizobium* symbiosis.

Key words: *nod* gene inducer, *Lotus japonicus–Mesorhizobium* symbiosis, phenolic acids, lipochitooligosaccharides, ultra-high-performance liquid chromatography–tandem-quadrupole mass spectrometry

Leguminous plants are characterized by their ability for symbiosis with a number of Gram-negative bacteria, collectively known as rhizobia. Rhizobia are free-living in soil, but change into bacteroids in the cells of specific host plants, in which they produce ammonium from atmospheric nitrogen and provide it to the host. Host–symbiont recognition and the initiation of symbiosis require the mutual exchange of chemical signals between leguminous plants and rhizobia. In host plants, the processes leading to root nodulation are triggered by rhizobial signal molecules called nod factors (NFs). NFs are lipochitooligosaccharides (LCOs) consisting of the oligomeric backbone of β-1,4-linked N-acetyl-D-glucosamine residues N-acylated at the non-reducing end. Chemical groups such as sulfate, fucose, and acetate, which vary according to the rhizobial strain, may substitute the oligosaccharide backbone (Liang *et al.*, 2014). The transmembrane receptor kinases of the host plant recognize specific NF structures and transmit a signal that triggers a series of symbiotic events, including root hair deformation, the formation and elongation of infection threads, and the induction of nodule primordia (Suzaki *et al.*, 2015). The induction of NF biosynthesis requires specific low-molecular-weight compounds exuded from the roots of the host plant, which are recognized by the rhizobial receptor NodD (Liu and Murray, 2016). When bound to the host-derived ligand, NodD serves as a transcription factor; it binds to *cis* elements called *nod* boxes and induces the transcription of a series of flanking genes, including *nod* genes, which encode enzymes involved in NF biosynthesis (Recourt *et al.*, 1989; Begume *et al.*, 2001; Liu and Murray, 2016). Therefore, these plant factors are called *nod* gene inducers. They include flavonoids and related compounds, such as flavanones, flavones, isoflavones, and chalcones, and function at very low concentrations of 10–100 μM (Peters *et al.*, 1986; Kosslak *et al.*, 1987; Hungria *et al.*, 1991). A *nod* box is also located upstream of the rhizobial type III secretion system (T3SS) cluster containing the *ttsI* gene (Okazaki *et al.*, 2010). T3SS secretes effector proteins that affect symbiosis in host plant cells (Okazaki *et al.*, 2013; Sugawara *et al.*, 2018; Kusakabe *et al.*, 2020).

Lotus japonicus, together with barrel medic (*Medicago truncatula*) and soybean (*Glycine max*), is a leguminous model system for molecular genetics and genomics

* Corresponding authors.
Tomoyoshi Akashi: E-mail: akashi.tomoyoshi@nihon-u.ac.jp;
Tel: +81–466–84–3353; Fax: +81–466–84–3353.
Shusei Sato: E-mail: shuseis@ige.tohoku.ac.jp;
Tel: +81–22–217–5688; Fax: +81–22–217–5691.
† These authors contributed equally to this work.
Toshio Aoki: deceased.

Citation: Shimamura, M., Kumaki, T., Hashimoto, S., Saeki, K., Ayabe, S., Higashitani, A., et al. (2022) Phenolic Acids Induce Nod Factor Production in *Lotus japonicus–Mesorhizobium* Symbiosis. *Microbes Environ* **37**: ME21094.
https://doi.org/10.1264/jsme2.ME21094

(Handberg and Stougaard, 1992; Udvardi *et al.*, 2005; Sato and Tabata, 2006). Data from whole-genome analyses of these species have provided insights into the origin and adaptive evolution of diverse leguminous plants (Cannon *et al.*, 2006; Hougaard *et al.*, 2008; Sato *et al.*, 2008; Young and Bharti, 2012; O'Rourke *et al.*, 2014). A more detailed understanding of symbiotic nitrogen fixation may be obtained by comparing the underlying molecular mechanisms among the three models. The reported *nod* gene inducers that act on *Mesorhizobium* strains, the symbionts of *L. japonicus*, are aldonic acids and related compounds, such as erythronic acid, tetronic acid, and succinic anhydride (Gagnon and Ibrahim, 1998). However, the induction of NF production in *Mesorhizobium* strains requires 10 mM tetronic acid or even higher concentrations of erythronic acid and succinic anhydride, more than a thousand-fold higher than those required for NF production in the symbionts of other legumes (Gagnon and Ibrahim, 1998). Furthermore, aldonic acids have not been detected in the root exudates or root tissues of *L. japonicus*. Therefore, the natural *nod* gene inducers of *L. japonicus–Mesorhizobium* symbiosis have not yet been identified.

Structural analyses using ultra-high-performance liquid chromatography coupled to quadrupole-time-of-flight (UPLC-QTOF) mass spectrometry (MS) and high-magnetic-field nuclear magnetic resonance spectroscopy revealed that *Mesorhizobium* NFs are a mixture of four major and four minor LCOs that vary in the fatty acid type, the number of carbamoyl groups at the non-reducing end, and the number of acetyl groups attached to fucose at the reducing end (Bek *et al.*, 2010). In the present study, we aimed to identify natural substances that are exuded from *L. japonicus* roots and act on *Mesorhizobium* strains as *nod* gene inducers by directly assaying NF production. We developed an assay method in which LCOs were extracted from small-scale cultures of *Mesorhizobium* strains in the presence of the candidate chemical compound and specifically detected by UPLC–tandem-quadrupole MS (TQMS). UPLC-TQMS is a convenient method for routine assays; however, its resolution is inferior to QTOF MS. We used this method to screen authentic samples of aldonic acids, flavonoids, and related phenolic compounds (instead of the fractionation of plant extracts or root exudates), and then examined the presence of the identified target compounds in the root exudates of *L. japonicus*. We identified five phenolic acids that induce NF production in *Mesorhizobium* strains.

NodD in *Rhizobium leguminosarum* is activated by naringenin, a *nod* gene inducer from *Medicago sativa* (Firmin *et al.*, 1986). When the heterologous *nodD* of *R. leguminosarum* was introduced into *Mesorhizobium japonicum* MAFF303099 (reclassified from *M. loti* based on genome sequence information; Martínez-Hidalgo *et al.*, 2016) via a multicopy plasmid, the application of naringenin induced NF production (López-Lara *et al.*, 1995; Niwa *et al.*, 2001; Bek *et al.*, 2010) and the expression of the *ttsI* gene, the regulator of T3SS (Okazaki *et al.*, 2010). Using a similar approach, namely, the introduction of endogenous *nodD* genes encoded by multicopy plasmids, we herein established which of the two NodD receptors, NodD1 or NodD2, of *M. japonicum* MAFF303099 interacted with the identified phenolic acids.

Materials and Methods

Chemicals

The following materials were purchased from the suppliers indicated in parentheses: chlorogenic acid (MP Biomedicals); gossypetin (Indofine Chemical Company); butein, eriodictyol, formononetin, herbacetin, isorhamnetin, kaempferol, myricetin, and quercetin (Extrasynthèse); daidzein and genistein (LC Laboratories); vestitol (Plantech); apigenin, biochanin A, coniferyl alcohol, *o*-coumaric acid, *m*-coumaric acid, coumestrol, 5-hydroxyferulic acid, luteolin, naringenin, phloretic acid, sinapic acid, and umbelliferone (Merck); *p*-coumaric acid, erythronic acid, isoferulic acid, and succinic anhydride (TCI); caffeic acid, *trans*-cinnamic acid, 3,4-dimethoxycinnamic acid, ferulic acid, *p*-methoxycinnamic acid, phenylalanine, tetronic acid, and L-tyrosine (FUJIFILM Wako Pure Chemical). Umbellic acid was prepared from umbelliferone by a treatment with 1 M NaOH at 90°C for 1 h. Isoliquiritigenin was obtained from our laboratory stock (Shimamura *et al.*, 2007).

NFs (LCOs) and Mesorhizobium *strains*

The plasmid pMP2112 encoding *R. leguminosarum* bv. *trifolii* nodD (López-Lara *et al.*, 1995) and a sample of the NF derived from *M. japonicum* MAFF303099 harboring pMP2112 were provided by H. Kouchi of the International Christian University, Japan. pMP2112 was transferred into *M. japonicum* MAFF303099 (Kaneko *et al.*, 2000; Saeki and Kouchi, 2000).

The bacterial strains and plasmids used in the present study are summarized in Supplementary Table S1. The *nodD1* and *nodD2* deletion (Δ*nodD1-nodD2*) variant of *M. japonicum* MAFF303099 was generated by homologous recombination as described by Hattori *et al.* (2002). The cosmid c243 was digested with *Bam*HI and ligated with a 1.9-kbp *Bam*HI fragment of the kanamycin resistance gene *neo* from pUCKM1 (Saeki *et al.*, 1991). In the resultant knockout plasmid pEMA49, the *nodD1* (*mll6179*)–*nolL* (*mlr8757*)–*nodD2* (*mlr6182*)–*mll6183–mlr6185* region was replaced with the *neo* gene. The generated allele was homogenotized with the endogenous genomic locus in *M. japonicum* to produce the Δ*nodD* variant. The construct was verified by Southern hybridization using c243 as a probe.

The *nodA* deletion (Δ*nodA*) variant was constructed essentially as described above, but with the precise in-frame deletion of the NodA coding region. An allele in which *nodA* was replaced with the spectinomycin resistance gene *aadA* was constructed by PCR amplification from pKST001R (Hanyu *et al.*, 2009) with primers to add overhangs (wan_mlr8755_upper and wan_mlr8755_lower). The amplified allele was then exchanged in *Escherichia coli* with the endogenous *mlr8755* allele in the cosmid c242.1 (Hattori *et al.*, 2002) in the presence of the phage lambda Red recombinase (Datsenko and Wanner, 2000) to generate the knockout plasmid pML8755DA. The correct construction of the Δ*nodA* variant was verified by PCR with the primers KS_nodSJ_F01S and KSnodC_Rev01 as well as by Southern hybridization using the wild-type PCR product as a probe.

In the functional analysis of *nodD*, the *nodD1* and *nodD2* genes were introduced separately into *M. japonicum* MAFF303099 and ML033 (Okazaki *et al.*, 2010) as follows. The 1,272-bp fragment containing the coding and promoter regions of *nodD1* (*mlr6182*) and the 1,365-bp fragment containing those of *nodD2* (*mll6179*) were amplified by PCR with the primer pairs pBBR1_nodD1_Fw and pBBR1_nodD1_Rv, and pBBR1_nodD2_Fw and pBBR1_nodD2_Rv, respectively (Supplementary Table S2). PCR products were cloned into pBBR1MCS-2 (Kovach *et al.*, 1995) by In-Fusion HD cloning (Clontech). The plasmids obtained (p*Mj*-NodD1 and p*Mj*-NodD2, respectively) were introduced separately into *E. coli* DH5α and mobilized into *M. japonicum* MAFF303099 using the previously described bacterial conjugation system (Kusakabe *et al.*, 2020). One day after

conjugation, transformants containing p*Mj*-NodD1 or p*Mj*-NodD2 were selected on tryptone–yeast-extract plates containing 100 μg mL^{-1} phosphomycin and 50 μg mL^{-1} kanamycin. Plasmid transfer was confirmed by PCR.

Culture conditions for M. japonicum *MAFF303099 and extraction of LCOs*

M. japonicum MAFF303099 was pre-cultured in TY medium at 28°C overnight. An aliquot of the culture was diluted with fresh TY medium (10 mL, OD_{660}=0.001) and supplemented with the antibiotics and phenolics shown in Table 1. Regarding *M. japonicum* MAFF303099 carrying pMP2112, antibiotics and naringenin (final concentration, 1 μM) were added to the culture medium. Diluted cultures were grown at 28°C for 24 h, centrifuged (8,000×*g*, room temperature, 2 min), and LCOs were then extracted from the supernatants with *n*-butanol.

Extraction of root exudates from legume seedlings

The seeds of alfalfa (*M. sativa*), red clover (*Trifolium pratense*), and *L. japonicus* B-129 Gifu were sterilized with solution containing 2% (v/v) sodium hypochlorite and 0.02% (v/v) Tween-20 for 10 min, rinsed five times with sterilized distilled water, and immersed in sterilized distilled water at room temperature overnight. They were then sown in a plastic container containing B&D liquid medium (pH 6.8) and cultivated at 25°C (16 h light/8 h dark) for 7 days. Media containing seedling exudates were collected and loaded onto Oasis HLB cartridges (Waters), and exudate components were eluted with ethanol.

UPLC-TQMS analysis

The butanol extracts of bacterial cultures and the ethanol extracts of root exudates were concentrated by evaporation, dissolved in 50% acetonitrile, and filtered through polytetrafluoroethylene membrane filters (Merck). UPLC analyses were conducted on Quattro Premier XE (Waters). Separation was performed on an Acquity UPLC BEH C18 column (2.1×100 mm, Waters) at 40°C and a flow rate of 0.38 mL min^{-1}. Gradient elution was performed with 0.1% formic acid in water (A) and 0.1% formic acid in acetonitrile (B) as follows. LCOs: 44% B (0–4 min), 44–85% B (4–7 min), 85–99.5% B (7–7.1 min), 99.5% B (7.1–7.2 min); root exudates: 5–15% B (0–7 min), 15–99% B (7–10 min).

Table 1. NF (LCO)-inducing activities of compounds tested in the present study.

Class	Compound	PubChem CID	Activity*
Phenylpropanoids	L-Phenylalanine (**1**)	6140	–
	L-Tyrosine (**2**)	6057	–
	trans-Cinnamic acid (**3**)	444539	–
	p-Coumaric acid (**4**)	637542	+
	Caffeic acid (**5**)	689043	+
	Ferulic acid (**6**)	445858	++
	5-Hydroxyferulic acid (**7**)	446834	+
	Sinapic acid (**8**)	637775	–
	Coniferyl alcohol (**9**)	1549095	–
	Umbellic acid (**10**)	446611	–
	Umbelliferone (**11**)	5281426	–
	Chlorogenic acid (**12**)	1794427	–
	Phloretic acid (**13**)	10394	+
	o-Coumaric acid (**14**)	637540	–
	m-Coumaric acid (**15**)	637541	–
	3,4-Dimethoxycinnamic acid (**16**)	717531	–
	Isoferulic acid (**17**)	736186	–
	p-Methoxycinnamic acid (**18**)	699414	–
Chalcones	Butein	5281222	–
	Isoliquiritigenin	638278	–
Flavanones	(*2S*)-Eriodictyol	440735	–
	(*2S*)-Naringenin	439246	–
Flavones	Apigenin	5280443	–
	Luteolin	5280445	–
Flavonols	Gossypetin	5280647	–
	Herbacetin	5280544	–
	Isorhamnetin	5281654	–
	Kaempferol	5280863	–
	Myricetin	5281672	–
	Quercetin	5280343	–
Isoflavones	Biochanin A	5280373	–
	Daidzein	5281708	–
	Formononetin	5280378	–
	Genistein	5280961	–
	Pseudobaptigenin	5281805	–
Isoflavans	(*3R*)-Vestitol	182259	–
Coumestans	Coumestrol	5281707	–
Aldonic acids	Erythronic acid (lactonized)	5325915	–
	Succinic anhydride	7922	–
	Tetronic acid (lactonized)	521261	–

*: ++ and + indicate LCO-inducing activity at 0.3 mg L^{-1} (++) and 15 mg L^{-1} (+); – no activity.

MS spectra were acquired with Quattro Premier v. 4.1 software (Waters) under the following conditions. Qualitative analyses: the selected ion recording (SIR) mode, electrospray ionization (ESI) positive mode, capillary voltage 3.0 kV, cone voltage 30 V, desolvation gas flow rate 800 L h^{-1} at 400°C, cone gas flow rate 50 L h^{-1}, and source temperature 120°C. Quantitative analyses: the selected reaction monitoring (SRM) mode; conditions described above except that the capillary voltage was 3.5 kV and the cone voltage was 50 V. SRM conditions for phenolic acids in root exudates are shown in Supplementary Table S3.

Isolation of rhizobia from root nodules of L. japonicus

Mature *L. japonicus* plants collected near the coast of Kanagawa, Japan, were dubbed "Bishamon" (34°54'10.3"N 139°53'15.7"E) and "Nojimazaki" (35°08'26.5"N 139°39'36.0"E). Nodules were harvested, and their surfaces were sterilized with solution containing 2% (v/v) sodium hypochlorite and 0.02% (v/v) Tween-20 for 10 min and then rinsed five times with sterilized distilled water. Nodules were crushed individually with a pestle in the presence of 40% glycerol, and the homogenates were spread on TY agar medium at 28°C for 10–14 days. A rhizobial colony was isolated from each nodule and named Bishamon1-c2 or Nojimazaki1-a1. Nodule formation by the isolated rhizobia in *L. japonicus* B-129 Gifu was examined according to Aoki *et al.* (2021). The procedure for LCO extraction from rhizobia, namely, Bishamon1-c2, Nojimazaki1-a1, and Tono (Kawaguchi *et al.*, 2002), was the same as that for *M. japonicum* MAFF303099 described above.

Root hair deformation assay

The root hair deformation assay was performed with the *n*-butanol extract of *M. japonicum* MAFF303099 according to previously described methods, except that B&D liquid and agar media adjusted to pH 6.8 (Broughton and Dilworth, 1971) were used instead of half-strength nitrogen-free HM nutrients (Imaizumi-Anraku *et al.*, 1997). LCOs were extracted with *n*-butanol from cultures of *M. japonicum* MAFF303099 treated with caffeic acid and purified on a solid-phase extraction column. An *n*-butanol extract prepared from mock-treated *M. japonicum* was used as a control. Purified LCOs and the control sample were used to treat 7-day-old *L. japonicus* B-129 Gifu for 24 h, and the roots were stained with toluidine blue and observed by light microscopy.

Expression analysis of ttsI*,* nodA*, and* nodB *genes*

To analyze the transcriptional regulation of the *ttsI* gene by phenolic acids, we used a chromosomally integrated translational *lacZ* fusion with the ML033 *ttsI* promoter (ML033) as previously described (Okazaki *et al.*, 2010). In the β-galactosidase assay, approximately 50 µL of pre-cultured (stationary phase) *Mesorhizobium* strains were inoculated into 5 mL of TY liquid medium (OD_{660} 0.01) in 50-mL tubes and grown for 21 h with or without phenolic acids at a final concentration of 10 µM.

β-Galactosidase activity was assessed in a microplate assay as previously described (Griffith and Wolf, 2002). The expression of the *nodA* and *nodB* genes was analyzed by qRT-PCR. Pre-cultured (mid-log phase) *Mesorhizobium* strains (1.5 mL) were inoculated into 1.5 mL of TY liquid medium in 15-mL tubes with or without phenolic acids at a final concentration of 10 µM and grown for 4 h. Bacterial RNA was stabilized by adding RNAprotect Bacteria Reagent (Qiagen) and extracted with an RNeasy Mini kit (Qiagen), and purified total RNA was then treated with Recombinant DNase I (Takara Bio). cDNA was synthesized from total RNA using an ExScript RT Reagent Kit (Takara Bio). The primer pairs used in the qRT-PCR analysis are listed in Supplementary Table S2. All qRT-PCR measurements were performed in a C1000 Thermal Cycler (Bio-Rad) with a Kapa SYBR Fast qPCR Kit (Kapa Biosystems). The relative expression of the selected genes was calculated as $2^{-\Delta\Delta Ct}$ using the *16S rRNA* gene as a reference. All experiments were performed for three technical and three biological replicates. The primers used to assess the expression levels of the various target genes are listed in Supplementary Table S2.

Root X-Gal staining assay

L. japonicus MG-20 plants were grown for 10 days on 1/2 MS medium agar plates, and ML033, ML033/p*Mj*-nodD1, or ML033/p*Mj*-nodD2 suspensions containing approximately 10^8 cells were inoculated onto the root surface with low melting point agar containing 0.02% X-Gal. Low melting point agar containing 0.02% X-Gal without rhizobia cells was used as a mock control. Blue staining was observed after a 2-days incubation at 25°C.

Results

Screening of compounds for LCO-inducing activity in rhizobia based on direct LCO detection by UPLC-TQMS

To detect NF production with high sensitivity and reproducibility, we established an analytical method to detect LCOs in rhizobial culture media using UPLC-TQMS. When an authentic LCO sample was analyzed in the SIR mode by UPLC-TQMS, a major peak appeared with a retention time of 5–5.5 min and a minor peak appeared after 6 min (Fig. 1a). To monitor an authentic LCO, we set the *m*/*z* value to 1502.7 based on the consecutive mass spectra of the major peak (Fig. 1b), which we presumed to correspond to previously reported NodMl-V (C18:1, Me, Cb, AcFuc) (López-Lara *et al.*, 1995; Niwa *et al.*, 2001; Bek *et al.*, 2010). The analysis of LCOs extracted from the culture medium of *M. japonicum* MAFF303099 carrying the pMP2112 plasmid, which harbored *nodD* from *R. leguminosarum*, in the presence of 1 µM naringenin also revealed a major peak at 5–5.5 min (Fig. 1a).

The established analytical method was used to test the LCO-inducing activities of 40 phenolic compounds, including phenylpropanoids, chalcones, flavanones, flavones, flavonols, isoflavones, an isoflavan, and a coumestan, as well as aldonic acids that were previously reported to exhibit *nod* gene-inducing activity (Table 1). LCO induction was analyzed by culturing *M. japonicum* MAFF303099 with 0.3 or 15 mg L^{-1} of the tested compounds, followed by *n*-butanol extraction of the culture medium and UPLC-TQMS analyses using the SIR mode at *m*/*z* 1502.7 (Fig. 1b). LCO was produced in the presence of five phenolic acids at 15 mg L^{-1}: *p*-coumaric acid (**4**), caffeic acid (**5**), ferulic acid (**6**), 5-hydroxyferulic acid (**7**), and phloretic acid (**13**) (Fig. 1c and Table 1). Only ferulic acid (**6**) promoted LCO production at 0.3 mg L^{-1} (Table 1). Among the 40 compounds tested, the remaining 35 compounds, including cinnamic acid, produced few or no LCOs (Table 1). We evaluated the reproducibility of LCO-inducing activities at various concentrations of these five phenolic acids employing the SRM mode of UPLC-TQMS and found that their activity increased in different concentration-dependent manners up to 100 µM (Fig. 1d). At 1 µM, only ferulic acid (**6**) was active, and its activity was the highest among the five inducers at 10 µM. In contrast, caffeic acid (**5**) showed weak activity up to 10 µM, but was the most active inducer at 100 µM. *p*-Coumaric acid (**4**), 5-hydroxyferulic acid (**7**), and phloretic acid (**13**) were only active at 100 µM (Fig. 1d).

Using the deletion variants of *M. japonicum* MAFF303099, we confirmed that the LCO-inducing activi-

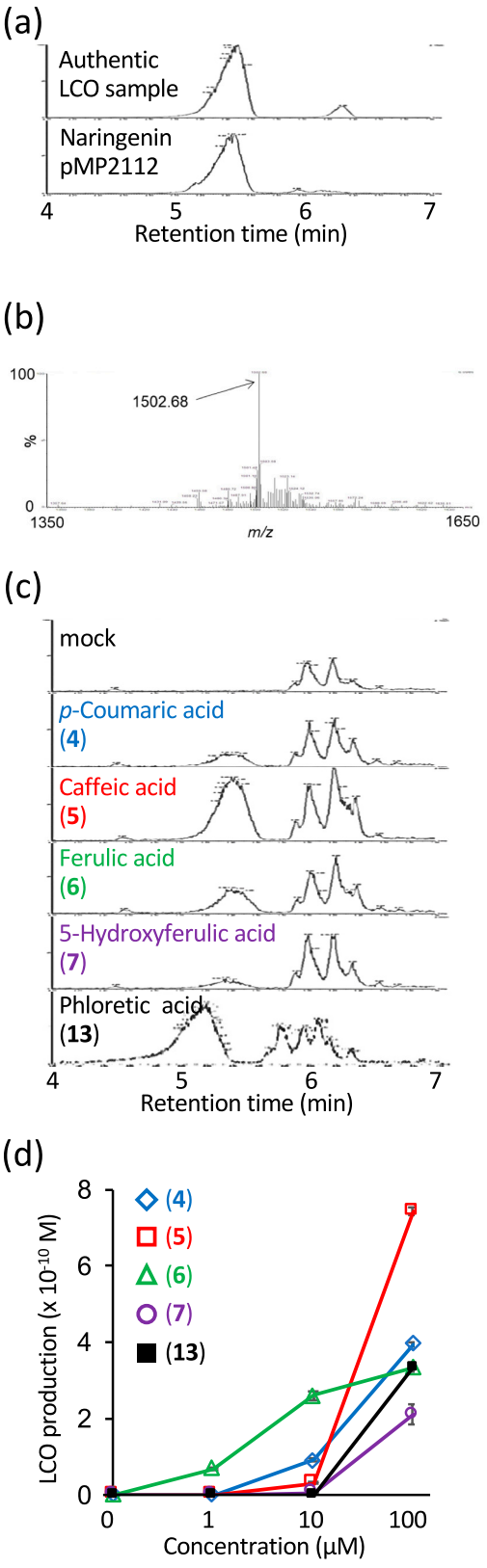


Fig. 1. Phenolic acids induce the production of NFs (LCOs) in *Mesorhizobium japonicum* MAFF303099. UPLC-TQMS analysis of an authentic LCO sample and naringenin-induced LCOs in *M. japonicum* harboring pMP2112. Chromatograms recorded by the SIR mode at m/z 1502.7 (a) and the mass spectrum at a retention time of 5.41 min (b). (c) UPLC-TQMS analysis of LCOs produced in *M. japonicum* MAFF303099 after the application of the indicated compounds at 15 mg L^{-1}. (d) LCO production induced by different concentrations of phenolic acids listed in (c). Error bars show S.E. (n=4).

ties of *p*-coumaric acid (**4**), caffeic acid (**5**), and ferulic acid (**6**) depended on the *nodA* and *nodD* genes (Supplementary Fig. S1).

Phenolic acids induce LCO production in native Mesorhizobium *strains*

To investigate whether the phenolic acids identified by screening using *M. japonicum* MAFF303099 generally induce LCOs in native *L. japonicus* rhizobia, we tested three additional *Mesorhizobium* strains isolated from native *L. japonicus*: Tono (Kawaguchi *et al.*, 2002), Nojimazaki 1-a1, and Bishamon 1-c2 (the present study). The phenolic acids tested, namely, *p*-coumaric acid (**4**), caffeic acid (**5**), ferulic acid (**6**), and 5-hydroxyferulic acid (**7**), produced LCOs with good reproducibility and at levels that were significantly higher (25- to 60-fold) in all native *Mesorhizobium* strains than in *M. japonicum* MAFF303099 (Fig. 2a). No adverse effects of these phenolic acids on the growth of rhizobia were observed at the concentrations used (Fig. 2b).

Biological activities of phenolic acids toward the host plant L. japonicus

To examine the abilities of the LCOs produced to function as NFs for *L. japonicus*, we assessed their root hair deformation activities. The root hairs of *L. japonicus* treated with 1 nM LCOs, which were induced by 100 μM caffeic acid, showed bending deformation, and some root hair tips were curled (Supplementary Fig. S2a, b, and c). The control sample had no effect on root hairs (Supplementary Fig. S2d).

Exogenous *nod* gene inducers have been shown to increase the nodule number in pea and soybean (Novák *et al.*, 2002; Pan *et al.*, 2008); therefore, we tested the effects of phenolic acids on *L. japonicus* inoculated with *M. japonicum*. The number of mature nodules significantly increased after 4 weeks in the presence of 10 μM ferulic acid (Fig. 3a and b) or 10 μM caffeic acid (Fig. 3c).

NodD1 acts as a receptor for phenolic acids

Rhizobia perceive *nod* gene inducers, such as flavonoids, through their binding to the transcriptional activator NodD, which up-regulates the expression of *nod* genes and the *ttsI* gene, a regulator of T3SS (Firmin *et al.*, 1986; Okazaki *et al.*, 2010). To examine transcriptional responses to the identified phenolic acids, we used *M. japonicum* ML033, in which the translational fusion of *lacZ* with *ttsI* was integrated into the chromosome of *M. japonicum* MAFF303099 (Okazaki *et al.*, 2010). No significant increases in β-galactosidase activity were detected in the presence of 1 μM *p*-coumaric acid (**4**), caffeic acid (**5**), or ferulic acid (**6**) (Fig. 4a). We then introduced one of the endogenous *nodD* genes via a multicopy plasmid, followed by an analysis using pMP2112 harboring *nodD* from *R. leguminosarum*. We constructed separate multicopy plasmids harboring *nodD1* (p*Mj*-NodD1) or *nodD2* (p*Mj*-NodD2), and transferred them into *M. japonicum* ML033.

After a 21-h culture with 1 μM *p*-coumaric acid (**4**), caffeic acid (**5**), or ferulic acid (**6**), β-galactosidase activity significantly increased in ML033/p*Mj*-NodD1, but not in ML033/p*Mj*-NodD2 (Fig. 4a). Induction levels were similar

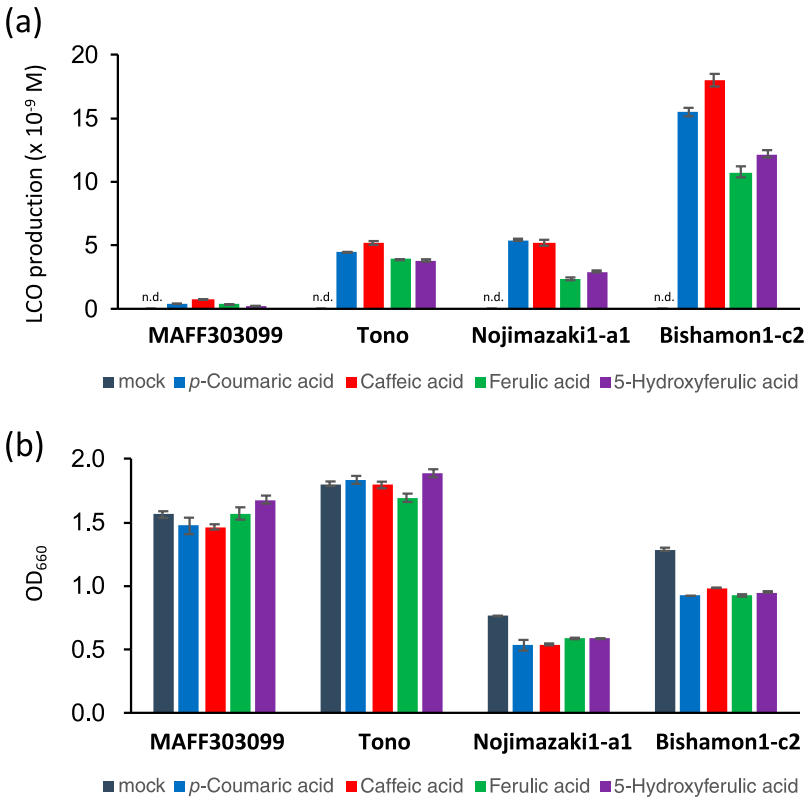


Fig. 2. Production of LCOs and growth of *Mesorhizobium* strains isolated from different locations in the presence of phenolic acids. (a) LCO production in *M. japonicum* MAFF303099, Tono, Nojimazaki 1-a1, and Bishamon 1-c2. (b) Growth of the strains tested in panel (a). Error bars show S.E. (n=3 or 4).

to that with 1 μM naringenin in ML033/pMP2112 (Fig. 4a), which carries *nodD* from *R. leguminosarum* (López-Lara *et al.*, 1995). In contrast, cinnamic acid (**3**), which did not induce LCO production in the screening described above, did not induce β-galactosidase activity in any of the strains (Fig. 4a).

The concentration dependence of β-galactosidase activity in ML033/p*Mj*-NodD1 induced by each phenolic acid (10 nM to 10 μM) is shown in Fig. 4b. *p*-Coumaric acid (**4**) and ferulic acid (**6**) both significantly induced activity at the lowest concentration tested (10 nM), whereas caffeic acid (**5**) only induced it at 1 and 100 μM. This concentration dependence was similar to that observed in the direct detection of LCO production (Fig. 1).

We then conducted qRT-PCR to investigate whether phenolic acids induce the transcriptional activation of the NF biosynthesis genes *nodA* and *nodB* using the ML033/pMj-NodD1 strain. The *nodA* gene was significantly induced by 1 μM ferulic acid (**6**) and caffeic acid (**5**), while the *nodB* gene was significantly induced at both 1 and 10 μM, similar to the *ttsI* gene (Fig. 5). Caffeic acid (**5**) exerted similar effects to ferulic acid (**6**) (Fig. 5). As in the β-galactosidase assay, cinnamic acid (**3**) did not induce *nodA*, *nodB*, or *ttsI* expression (data not shown). The above results revealed that the *M. japonicum* NodD1 receptor recognizes phenolic acids, such as ferulic and caffeic acids, and activates the transcription of the *nod* genes.

Phenolic acids exuded from legume roots

To elucidate whether phenolic compounds were exuded from the roots of *L. japonicus*, we used the SRM mode of UPLC-TQMS to analyze the components of culture media 7 days after hydroponic cultures of *L. japonicus*, red clover, and alfalfa. The levels of five phenolic acids that induce the production of NFs (*p*-coumaric acid [**4**], caffeic acid [**5**], ferulic acid [**6**], 5-hydroxyferulic acid [**7**], and phloretic acid [**13**]) and those of phenolic acids that do not (cinnamic acid [**3**] and sinapic acid [**8**]) were quantified. *p*-Coumaric acid (**4**) and ferulic acid (**6**) were detected in hydroponic media from all three plant species, with the levels of *p*-coumaric acid (**4**) being higher in exudates from *L. japonicus* (Table 2). Cinnamic acid (**3**) was only secreted by *L. japonicus* (Table 2). Trace levels (0.2 nmol g^{-1} FW plants or 0.6 nmol mg^{-1} root exudate) of caffeic acid (**5**), sinapic acid (**8**), and phloretic acid (**13**) were detected, whereas 5-hydroxyferulic acid (**7**) was undetectable (less than 0.2 nmol g^{-1} FW plant or 0.6 nmol mg^{-1} root exudate) in either legume.

To monitor the induction of the *ttsI*::*lacZ* fusion in ML033 series strains in the rhizosphere of *L. japonicus*, approximately 10^8 rhizobial cells were spread on the roots of 10-day-old seedlings grown on agar medium with X-Gal. The inoculation with the ML033/p*Mj*-NodD1 strain resulted in X-gal blue staining on and around the root surface (Supplementary Fig. S3). The inoculation of the ML033/p*Mj*-NodD2 strain resulted in X-gal staining only on the root surface. These results strongly suggest that phenolic acid(s) activating NodD1 were exuded not only on the root surface, but also into the rhizosphere of *L. japonicus*.

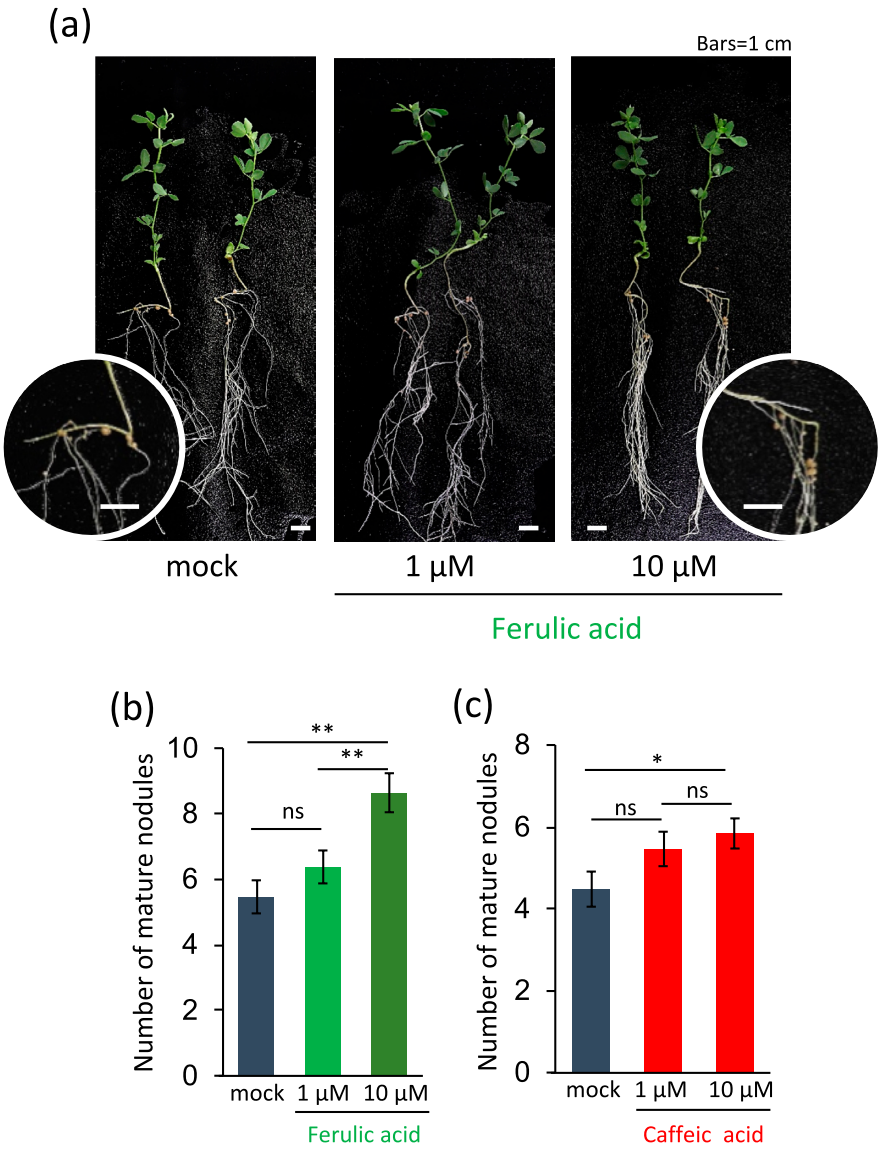


Fig. 3. Promotion of nodulation by phenolic acids. (a) Typical images of 4-week-old *Lotus japonicus* MG-20 inoculated with *Mesorhizobium japonicum* MAFF303099 in the absence or presence of ferulic acid. Scale bars: 1 cm. (b, c) Number of mature nodules in the absence or presence of (b) ferulic acid or (c) caffeic acid. Error bars show S.E. (n=22–27). Significant differences between the absence (mock) and presence of phenolic acids were assessed by the Student's t-test (** P<0.01, * P<0.05, *ns*, not significant).

Discussion

The *nod* gene inducer in *L. japonicus–Mesorhizobium* symbiosis has not been identified despite decades of research. Therefore, we developed a highly sensitive method to analyze LCOs using UPLC-TQMS. MS was previously applied to the study of LCOs, mainly for a structural analysis (Niwa *et al.*, 2001; Bek *et al.*, 2010). To use MS for screening, we simplified purification and reduced the required culture volume, which allowed us to evaluate 40 phenolic compounds. Among them, we identified five phenolic acids that had the potential to induce LCO production in *M. japonicum* MAFF303099. The production of LCOs was enhanced by increases in the concentrations of each of these phenolic acids. LCOs produced by *M. japonicum* MAFF303099 in the presence of caffeic acid (**5**) induced root hair deformation, and nodule numbers in *L. japonicus* inoculated with *M. japonicum* were increased by the addition of ferulic acid (**6**) and caffeic acid (**5**). These results clearly identified phenolic compounds, but not flavonoids, as *nod* gene inducers. Phenolic acids are produced via shikimic acid through the phenylpropanoid pathway, and also as intermediates of the monolignol pathway in vascular plants. A previous study reported that rhizobia utilized phenolic acids as carbon sources (Blum *et al.*, 2000). A number of *nod* gene inducers have been identified in legumes, and most of them are (iso)flavonoids (Liu and Murray, 2016); to the best of our knowledge, this is the first study to demonstrate that phenolic acids function as *nod* gene inducers.

The identified candidate *nod* gene inducers of *L. japonicus* are phenylpropanoids with a carboxylic acid group, in contrast to coniferyl alcohol (**9**) and chlorogenic acid (**12**) (Fig. 6). A comparison with phenolic acids that did not induce LCOs (Table 2), *i.e.* L-phenylalanine (**1**),

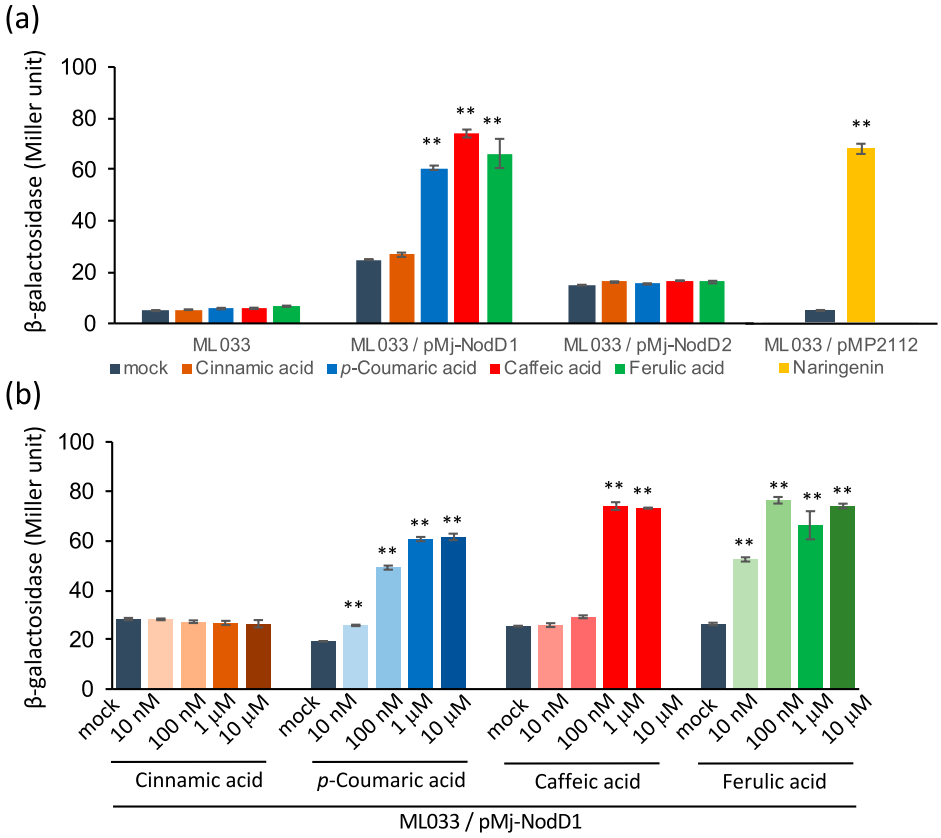


Fig. 4. Effects of phenolic acids on the expression of *lacZ*-fused *ttsI*. (a) Effects of harboring p*Mj*-NodD1, the p*Mj*-NodD2 plasmid, or pMP2112 on β-galactosidase induction by the indicated phenolic acids (1 μM). (b) β-Galactosidase activity in *Mesorhizobium japonicum* MAFF303099 harboring p*Mj*-NodD1 in the presence of different concentrations of the phenolic acids tested in the panel (a). Error bars show S.E. (n=3 or 4). Significant differences between the absence (mock) and presence of phenolic acids were assessed by the Student's t-test (** P<0.01).

trans-cinnamic acid (**3**), umbellic acid (**10**), *o*-coumaric acid (**14**), *m*-coumaric acid (**15**), 3,4-dimethoxycinnamic acid (**16**), isoferulic acid (**17**), and *p*-methoxycinnamic acid (**18**), suggested that a hydroxyl at C-4 and a hydrogen at C-2 or C-6 are important for the ability to induce NFs in *M. japonicum* (Fig. 6). A single methoxy group did not preclude LCO-inducing activity, whereas two methoxy groups, as in sinapic acid (**8**), did. A double bond between α and β carbons appears to be important for the induction of LCO production, but is not essential because weak induction was detected with phloretic acid (**13**). Therefore, the basic structure of a *nod* gene inducer appears to be that of *p*-coumaric acid (**4**) with at most a single methoxy group at C-3 or C-5 and hydrogens at C-2 and C-6. We suggest that the carboxylic acid group and the C-3, C-4, and C-5 positions of the phenyl ring were recognized by NodD.

Aldonic acids have been shown to promote LCO biosynthesis in *Mesorhizobium* strains (Gagnon and Ibrahim, 1998). In our experiments, none of the three aldonic acid compounds tested induced LCO production at the concentrations at which the five phenolic acids induced it. Gagnon and Ibrahim (1998) identified aldonic acids in the root exudates of *Lupinus albus* by screening based on measurements of the β-galactosidase activities of *Rhizobium lupini* strains harboring *nodC*::*lacZ* fusions, and 10 mM tetronic acid was required to induce detectable LCO production in *M. japonicum* R7A. We only tested concentrations up to 100 μM, which may explain why the aldonic acids tested did not induce LCOs. Since tetronic acid was not detected in the root exudates or seed metabolites of *L. japonicus* (Hashiguchi *et al.*, 2018), aldonic acids cannot be endogenous *nod* gene inducers of *L. japonicus*.

In contrast to aldonic acids, the presence of *p*-coumaric acid (**4**), ferulic acid (**6**), and *trans*-cinnamic acid (**3**) was confirmed in the root exudates of *L japonicus* (Table 2). In addition, *p*-coumaric acid (**4**) and ferulic acid (**6**) are listed as metabolites in the seeds of experimental and wild accessions of *L. japonicus* in LegumeBase, the resource database of National BioResource Project *Lotus*/*Glycine* (Hashiguchi *et al.*, 2018). Phenolic acids have also been reported in the root exudates of other legume and non-legume plants (Mandal *et al.*, 2010). We identified *p*-coumaric acid (**4**) and ferulic acid (**6**) in the root exudates of *T. pratense* and *M. sativa* (Table 2), and, thus, these phenolic acids do not appear to contribute to host specificity. Since phenolic acids are generally present in the rhizosphere, the responses of *Mesorhizobium* strains to them may contribute to their associations with a broad range of plants, including non-host plants. *M. japonicum* MAFF303099 associates with non-host plants, such as *Arabidopsis thaliana* (Poitout *et al.*, 2017), as a root epiphyte. T3SS may play a role in this relationship, as reported in *Bradyrhizobium* strains (Piromyou *et al.*, 2015). In the present study, the expression of *ttsI*, a regulator of the T3SS gene cluster, was induced by phenolic

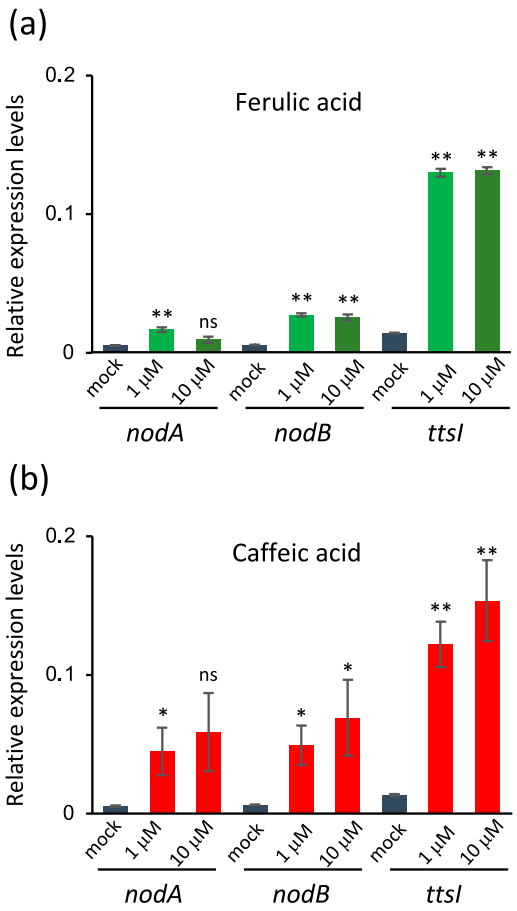


Fig. 5. qRT-PCR analysis of the expression of *nodA*, *nodB*, and *ttsI* genes in *Mesorhizobium japonicum* MAFF303099 harboring p*Mj*-NodD1 in the absence or presence of (a) ferulic acid or (b) caffeic acid. Error bars show S.E. (n=3 or 4). Significant differences between the absence (mock) and presence of phenolic acids were assessed by the Student's t-test (** P<0.01, * P<0.05, *ns*, not significant).

acids at higher levels than that of *nodA* in *M. japonicum* MAFF303099 carrying p*Mj*-NodD1 (Fig. 5). Therefore, phenolic acid recognition by *Mesorhizobium* strains may have a function against non-host plants by inducing T3SS. Regarding host specificity, differences in the concentrations of phenolic acids in root exudates may affect host specificity in *L. japonicus–Mesorhizobium* symbiosis because the concentration of *p*-coumaric acid (**4**) in the root exudates of *L. japonicus* was more than ten-fold higher than those in the root exudates of red clover and alfalfa (Table 2). The NF receptors of host plants are important for recognition that affects the host range in plant–rhizobia symbiotic interactions (Radutoiu *et al.*, 2007; Bek *et al.*, 2010). We confirmed the function of LCOs induced by caffeic acid (**5**) as endogenous NFs for *L. japonicus* by demonstrating their ability to induce root hair deformation (Supplementary Fig. S2). The number of mature nodules was increased by the addition of phenolic acid-type *nod* gene inducers together with an inoculation with *M. japonicum* (Fig. 3), as previously reported in pea and soybean (Novák *et al.*, 2002; Pan *et al.*, 2008). Therefore, the recognition of phenolic acids by *Mesorhizobium* strains may function in two ways: the production of T3SS may contribute to associations with a wide range of plants, and, at the same time, the production of LCOs may function in host recognition in symbiotic interactions.

Although we confirmed LCO production after treatments with five phenolic acids, we failed to detect the induction of *nod* genes by RT-PCR or *ttsI* expression using its promoter fused to *lacZ* in the genome (Fig. 4a). The transcript levels of genes regulated by NodD may be below the detection level of normal RT-PCR and a single copy of *lacZ* in the genome of *M. japonicum* MAFF303099; this may explain our failure to identify *nod* gene inducers in *L. japonicus–Mesorhizobium* symbiosis, even though it has been widely used as a model of symbiosis in legume plants (Liu and Murray, 2016). The *nodA* promoter fused to *lacZ* in a multicopy plasmid (pMP220) has been used to detect *lacZ* expression in *M. japonicum* MAFF303099 (Kojima *et al.*, 2012). We adopted this approach and attempted to increase the copy number of *nodD* genes by introducing *nodD1* or *nodD2* into a multicopy plasmid. This analysis revealed that NodD1 was more sensitive to phenolic acids than NodD2, indicating a functional differentiation between NodD1 and NodD2 with regards to the perception of phenolic acid signals. In *M. japonicum* R7A, NodD1 and NodD2 are functionally redundant for nodulation, with *nodD1* mutants exhibiting only a slight delay in nodulation (Rodpothong *et al.*, 2009). Kelly *et al.* (2018) showed the preferential activation of NodD1 and NodD2 by different compounds produced at defined stages of symbiotic infection. NodD1 is primarily involved in the induction of downstream genes within root hair infection threads. Since phenolic acids are intermediates in the biosynthesis of a number of phenolic compounds, such as flavonoids and monolignols, it is reasonable that they act as *nod* gene inducers in root hair infection threads. However, we detected *p*-coumaric acid (**4**) and ferulic acid (**6**) in the root exudates of *L. japonicus* (Table 2) and *lacZ* gene induction in the rhizosphere using an *M. japonicum* strain with the *nodD1* plasmid (Supplementary Fig. S3). Therefore, the *nod* gene–inducing activities of phenolic acids are not restricted to root hair infection threads,

Table 2. Characteristics of hydroponic cultures and phenolic acid contents in root exudates of three leguminous plants. Values are mean total contents±S.E. (n=4).

	Seed germinated (g DW)	Plant FW (g)	Root exudates (mg)	Contents (nmol g^{-1} FW plant)			Contents (nmol mg^{-1} root exudate)		
				3	**4**	**6**	**3**	**4**	**6**
Lotus japonicus	0.82	1.44	2.7	4.0±0.1	11.9±0.3	2.0±0.0	21.1±0.6	63.2±1.3	10.5±0.1
red clover	0.69	3.38	2.7	n.d.	0.3±0.0	1.7±0.0	n.d.	4.0±0.0	21.6±0.1
alfalfa	2.07	10.21	8.2	n.d.	0.4±0.0	0.6±0.0	n.d.	5.0±0.0	8.0±0.1

FW, fresh weight. **3**, *trans*-cinnamic acid; **4**, *p*-coumaric acid; **6**, ferulic acid; n.d., not detected.

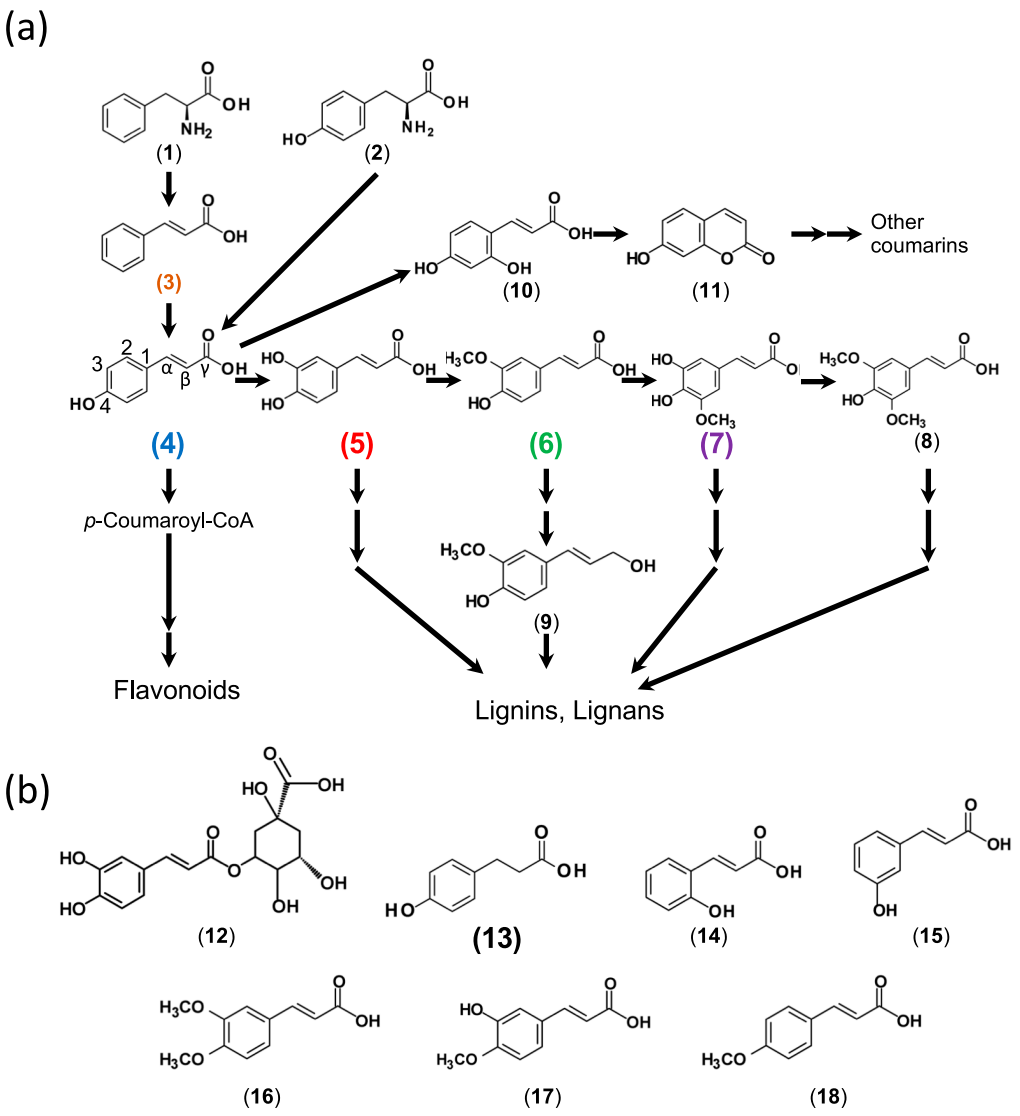


Fig. 6. Chemical structures of phenolic acids used in the present study. (a) Phenolic acids in the major phenylpropanoid pathway towards coumarins, lignins, and lignans. (b) Other phenolic acids. See Table 1 for compound names.

they may also be involved in a wide range of associations in the rhizosphere.

In the present study, we used the direct detection of LCOs to screen for *nod* gene inducers in *L. japonicus–Mesorhizobium* symbiosis. We identified five candidate compounds in the group of phenolic acids, and detected two in the root exudates of *L. japonicus*. By increasing the copy number of one of the two *nodD* genes in *M. japonicum*, we revealed that phenolic acids as *nod* gene inducers were mainly recognized by NodD1. Overall, we propose that phenolic acids are a novel type of *nod* gene inducer in the *L. japonicus–Mesorhizobium* symbiosis system. Therefore, substances that act as mutual symbiotic signals from both sides, *L. japonicus* and *M. japonicum*, are now revealed. The present results will accelerate the elucidation of the regulatory mechanisms in this symbiotic system.

Acknowledgements

We dedicate this report to the memory of the late Professor Toshio Aoki (1961–2019), the last author and principle investigator of this work. The late Professor Aoki designed and conducted the central part of this research, including the establishment of the LCO detection system using UPLC-TQMS, the screening of phenolic compounds, the confirmation of the NF activity of LCO induced by phenolic acids, and the detection of phenolic acids in root exudates.

This work was supported by JSPS KAKENHI Grant Number JP20250915, JST-Mirai Program Grant Number JPMJMI20E4, and JST CREST Grant Number JPMJCR16O1 Japan. We would like to thank Ms. Chikako Mistuoka for her excellent technical assistance. We would also like to express our gratitude to Prof. Hiroshi Kouchi for NF from *M. japonicum* and the plasmid pMP2112, Prof. Hisayuki Mitsui for the plasmid pBBR1MCS-2, and Prof. Shin Okazaki for the *M. japonicum* ML033 strain. The accessions of the *L. japonicus* and *Mesorhizobium* strains were provided by the National BioResource Project '*Lotus*/*Glycine*'.

References

Aoki, T., Kawaguchi, M., Imaizumi-Anraku, H., Akao, S., Ayabe, S., and Akashi, T. (2021) Mutants of *Lotus japonicus* deficient in flavonoid biosynthesis. *J Plant Res* **134**: 341–352.

Begume, A.A., Leibovitch, S., Migner, P., and Zhang, F. (2001) Specific flavonoids induced *nod* gene expression and pre-activated *nod* genes of *Rhizobium leguminosarum* increased pea (*Pisum sativum* L.) and lentil (*Lens culinaris* L.) nodulation in controlled growth chamber environments. *J Exp Bot* **52**: 1537–1543.

Bek, A.S., Sauer, J., Thygesen, M.B., Duus, J.Ø., Petersen, B.O., Thirup, S., *et al.* (2010) Improved characterization of nod factors and genetically based variation in LysM receptor domains identify amino acids expendable for nod factor recognition in *Lotus* spp. *Mol Plant-Microbe Interact* **23**: 58–66.

Blum, U., Staman, K.L., Flint, L.J., and Shafer, S.R. (2000) Induction and/or selection of phenolic acid-utilizing bulk-soil and rhizosphere bacteria and their influence on phenolic acid phytotoxicity. *J Chem Ecol* **26**: 2059–2078.

Broughton, W.J., and Dilworth, M.J. (1971) Control of leghaemoglobin synthesis in snake beans. *Biochem J* **125**: 1075–1080.

Cannon, S.B., Sterck, L., Rombauts, S., Sato, S., Cheung, F., Gouzy, J., *et al.* (2006) Legume genome evolution viewed through the *Medicago truncatula* and *Lotus japonicus* genomes. *Proc Natl Acad Sci U S A* **103**: 14959–14964.

Datsenko, K.A., and Wanner, B.L. (2000) One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc Natl Acad Sci U S A* **97**: 6640–6645.

Firmin, J.L., Wilson, K.E., Rossen, L., and Johnston, W.B. (1986) Flavonoid activation of nodulation genes in Rhizobium reversed by other compounds present in plants. *Nature* **324**: 90–92.

Gagnon, H., and Ibrahim, R.K. (1998) Aldonic acids: A novel family of *nod* gene inducers of *Mesorhizobium loti*, *Rhizobium lupini*, and *Sinorhizobium meliloti*. *Mol Plant-Microbe Interact* **11**: 988–998.

Griffith, K.L., and Wolf, R.E. (2002) Measuring β-galactosidase activity in bacteria: Cell growth, permeabilization, and enzyme assays in 96-well arrays. *Biochem Biophys Res Commun* **290**: 397–402.

Handberg, K., and Stougaard, J. (1992) *Lotus japonicus*, an autogamous, diploid legume species for classical and molecular genetics. *Plant J* **2**: 487–496.

Hanyu, M., Fujimoto, H., Tejima, K., and Saeki, K. (2009) Functional differences of two distinct catalases in *Mesorhizobium loti* MAFF303099 under free-living and symbiotic conditions. *J Bacteriol* **191**: 1463–1471.

Hashiguchi, M., Tanaka, H., Muguerza, M., Akashi, R., Sandal, N.N., Andersen, S.U., *et al.* (2018) *Lotus japonicus* genetic, mutant, and germplasm resources. *Curr Protoc Plant Biol* **3**: e20070.

Hattori, Y., Omori, H., Hanyu, M., Kaseda, N., Mishima, E., Kaneko, T., *et al.* (2002) Ordered cosmid library of the *Mesorhizobium loti* MAFF303099 genome for systematic gene disruption and complementation analysis. *Plant Cell Physiol* **43**: 1542–1557.

Hougaard, B.K., Madsen, L.H., Sandal, N., Moretzsohn, M.C. Fredslund, J., Schauser, L., *et al.* (2008) Legume anchor markers link syntenic regions between *Phaseolus vulgaris*, *Lotus japonicus*, *Medicago truncatula* and Arachis. *Genetics* **179**: 2299–2312.

Hungria, M., Joseph, C.M., and Phillips, D.A. (1991) Anthocyanidins and flavanols, major *nod* gene inducers from seeds of black-seeded common bean (*Phaseolus vulgaris* L.). *Plant Physiol* **97**: 751–775.

Imaizumi-Anraku, H., Kawaguchi, M., Koiwa, H., Akao, S., and Syōno, K. (1997) Two ineffective-nodulating mutants of *Lotus japonicus*—Different phenotypes caused by the blockage of endocytotic bacterial release and nodule maturation. *Plant Cell Physiol* **38**: 871–881.

Kaneko, T., Nakamura, Y., Sato, S., Asamizu, E., Kato, T., Sasamoto, S., *et al.* (2000) Complete genome structure of the nitrogen-fixing symbiotic bacterium *Mesorhizobium loti*. *DNA Res* **7**: 331–338.

Kawaguchi, M., Imaizumi-Anraku, H., Koiwa, H., Niwa, S., Ikuta, A., Syono, K., and Akao, S. (2002) Root, root hair, and symbiotic mutants of the model legume *Lotus japonicus*. *Mol Plant-Microbe Interact* **15**: 17–26.

Kelly, S., Sullivan, T.J.S., Kawaharada, Y., Radutoiu, S., Ronson, W.C., and Stougaard, J. (2018) Regulation of Nod factor biosynthesis by alternative NodD proteins at distinct stages of symbiosis provides additional compatibility scrutiny. *Environ Microbiol* **20**: 97–110.

Kojima, K., Yokoyama, T., Ohkama-Ohtsu, N., Saengkerdsub, S., Itakura, M., Mitsui, H., *et al.*, (2012) Exploration of natural nod gene inducers for Mesorhizobium loti in seed and root exudates of *Lotus corniculatus*. *Soil Microorganisms* **66**: 12–21.

Kosslak, R.M., Bookland, R., Barkei, J., Paaren, H.E., and Appelbaum, E.R. (1987) Induction of *Bradyrhizobium japonicum* common *nod* genes by isoflavones isolated from *Glycine max*. *Proc Natl Acad Sci U S A* **84**: 7428–7432.

Kovach, M.E., Elzer, P.E., Hill D.S., Robertson G.T., Farris, M.A., Roop II, R.M., *et al.* (1995) Four new derivatives of the broad-host-range cloning vector pBBR1MCS, carrying different antibiotic-resistance cassettes. *Gene* **166**: 175–176.

Kusakabe, S., Higasitani, N., Kaneko, T., Yasuda, M., Miwa, H., Okazaki, S., *et al.* (2020) *Lotus* accessions possess multiple checkpoints triggered by different type III secretion system effectors of the wide-host-range symbiont *Bradyrhizobium elkanii* USDA61. *Microbes Environ* **35**: ME19141.

Liang, Y., Tóth, K., Cao, Y., Tanaka, K., Espinoza, C., and Stacey, G. (2014) Lipochitooligosaccharide recognition: an ancient story. *New Phytol* **204**: 289–296.

Liu, C.W., and Murray, J.D. (2016) The role of flavonoids in nodulation host-range specificity: An update. *Plants* **5**: 33.

López-Lara, I.M., van den Berg, J.D.J., Thomas-Oates, J.E., Glushka, J., Lugtenberg, B.J.J., and Spaink, H.P. (1995) Structural identification of the lipo-chitin oligosaccharide nodulation signals of *Rhizobium loti*. *Mol Microbiol* **15**: 627–638.

Mandal, S.M., Chakraborty, D., and Dey, S. (2010) Phenolic acids act as signaling molecules in plant-microbe symbioses. *Plant Signal Behav* **5**: 359–368.

Martínez-Hidalgo, P., Ramírez-Bahena, M.H., Flores-Félix, J.D., Igual, J.M., Sanjuán J., León-Barrios, M., *et al.* (2016) Reclassification of strains MAFF 303099^{T} and R7A into *Mesorhizobium japonicum* sp. nov. *Int J Syst Bacteriol* **66**: 4936–4941.

Niwa, S., Kawaguchi, M., Imaizumi-Anraku, H., Chechetka, S.A., Ishizaka, M., Ikuta, A., *et al.* (2001) Responses of a model legume *Lotus japonicus* to lipochitin oligosaccharide nodulation factors purified from *Mesorhizobium loti* JRL501. *Mol Plant-Microbe Interact* **14**: 848–856.

Novák, K., Chovanec, P., Škrdleta, V., Kropáčová, M., Lisá, L., and Němcová, M. (2002) Effect of exogenous flavonoids on nodulation of pea (*Pisum sativum* L.). *J Exp Bot* **53**: 1735–1745.

Okazaki, S., Okabe, S., Higashi, M., Shimoda, Y., Sato, S., Tabata, S., *et al.* (2010) Identification and functional analysis of type 3 effector proteins in *Mesorhizobium loti*. *Mol Plant-Microbe Interact* **23**: 223–234.

Okazaki, S., Kaneko, T., Sato, S., and Saeki, K. (2013) Hijacking of leguminous nodulation signaling by the rhizobial type III secretion system. *Proc Natl Acad Sci U S A* **110**: 17131–17136.

O'Rourke, J.A., Bolon, Y.T., Bucciarelli, B., and Vance, C.P. (2014) Legume genomics: understanding biology through DNA and RNA sequencing. *Ann Bot* **113**: 1107–1120.

Pan, B., Zhang, F., and Smith, D.L. (2008) Genistein addition to the rooting medium of soybean at the onset of nitrogen fixation increases nodulation. *J Plant Nutr* **21**: 1631–1639.

Peters, N.K., Frost, J.W., and Long, S.R. (1986) A plant flavone, luteolin, induces expression of *Rhizobium meliloti* nodulation genes. *Science* **233**: 977–980.

Piromyou, P., Songwattana, P., Greetatorn, T., Okubo, T., Chiba-Kakizaki, K., Prakamhang, J., *et al.* (2015) The Type III Secretion System (T3SS) is a determinant for rice-endophyte colonization by non-photosynthetic *Bradyrhizobium*. *Microbes Environ* **30**: 291–300.

Poitout, A., Martinière, A., Kucharczyk, B., Queruel, N., Silva-Andia, J., Mashkoor, S., *et al.* (2017) Local signalling pathways regulate the *Arabidopsis* root developmental response to *Mesorhizobium loti* inoculation. *J Exp Bot* **68**: 1199–1211.

Radutoiu, S., Madsen, L.H., Madsen, E.B., Jurkiewicz, A., Fukai, E., Quistgaard, E.M.H., *et al.* (2007) LysM domains mediate lipochitin–oligosaccharide recognition and *Nfr* genes extend the symbiotic host range. *EMBO J* **26**: 3923–3935.

Recourt, K., Brussel, A.A.N., Driessen, A.J.M., and Lugtenberg, B.J.J. (1989) Accumulation of a *nod* gene inducer, the flavonoid naringenin, in the cytoplasmic membrane of *Rhizobium leguminosarum* biovar *viciae* is caused by the pH-dependent hydrophobicity of naringenin. *J Bacteriol* **171**: 4370–4377.

Rodpothong, P., Sullivan, J.T., Songsrirote, K., Sumpton, D., Cheung, K.W., Thomas-Oates, J., *et al.* (2009) Nodulation gene mutants of *Mesorhizobium loti* R7A-nodZ and nolL mutants have host-specific phenotypes on *Lotus* spp. *Mol Plant-Microbe Interact* **22**: 1546–1554.

Saeki, K., Suetsugu, Y., Tokuda, K., Miyatake, Y., Young, D.A., Marrs, B.L., *et al.* (1991) Genetic analysis of functional differences among distinct ferredoxins in *Rhodobacter capsulatus*. *J Biol Chem* **266**: 12889–12895.

Saeki, K., and Kouchi, H. (2000) The *Lotus* symbiont, *Mesorhizobium loti*: molecular genetic techniques and application. *J Plant Res* **113**: 457–465.

Sato, S., and Tabata, S. (2006) *Lotus japonicus* as a platform for legume research. *Curr Opin Plant Biol* **9**: 128–132.

Sato, S., Nakamura, Y., Kaneko, T., Asamizu, E., Kato, T., Nakao, M., *et al*. (2008) Genome structure of the legume, *Lotus japonicus*. *DNA Res* **15**: 227–239.

Shimamura, M., Akashi, T., Sakurai, N., Suzuki, H., Saito, K., Shibata, D., *et al.* (2007) 2-Hydroxyisoflavanone dehydratase is a critical determinant of isoflavone productivity in hairy root cultures of *Lotus japonicus*. *Plant Cell Physiol* **48**: 1652–1657.

Sugawara, M., Takahashi, S., Umehara, Y., Iwano, H., Tsurumaru, H., Odake, H., *et al.* (2018) Variation in bradyrhizobial NopP effector determines symbiotic incompatibility with Rj2-soybeans via effector-triggered immunity. *Nat Commun* **9**: 3139.

Suzaki, T., Yoro, E., and Kawaguchi, M. (2015) Leguminous plants: inventors of root nodules to accommodate symbiotic bacteria. *Int Rev Cell Mol Biol* **316**: 111–158.

Udvardi, M.K., Tabata, S., Parniske, M., and Stougaard, J. (2005) *Lotus japonicus*: legume research in the fast lane. *Trends Plant Sci* **10**: 222–228.

Young, N.D., and Bharti, A.K. (2012) Genome-enabled insights into legume biology. *Annu Rev Plant Biol* **63**: 283–305.

Microbes Environ. 36(2), 2021
https://www.jstage.jst.go.jp/browse/jsme2 doi:10.1264/jsme2.ME20137

Community Analysis-based Screening of Plant Growth-promoting Bacteria for Sugar Beet

Kazuyuki Okazaki[1], Hirohito Tsurumaru[2], Megumi Hashimoto[2], Hiroyuki Takahashi[1], Takashi Okubo[2], Takuji Ohwada[3], Kiwamu Minamisawa[2], and Seishi Ikeda[1]*

[1]*Memuro Research Station, Hokkaido Agricultural Research Center, National Agriculture and Food Research Organization, 9–4 Shinsei-minami, Memuro, Kasai-gun, Hokkaido 082–0081, Japan;* [2]*Graduate School of Life Science, Tohoku University, 2–1–1 Katahira, Aoba-ku, Sendai, Miyagi 980–8577, Japan; and* [3]*Department of Agricultural and Life Sciences, Obihiro University of Agriculture and Veterinary Medicine, Obihiro, Hokkaido 080–8555, Japan*

(Received October 20, 2020—Accepted March 9, 2021—Published online April 27, 2021)

Clone libraries of bacterial 16S rRNA genes (a total of 1,980 clones) were constructed from the leaf blades, petioles, taproots, and lateral roots of sugar beet (*Beta vulgaris* L.) grown under different fertilization conditions. A principal coordinate analysis revealed that the structures of bacterial communities in above- and underground tissues were largely separated by PC1 (44.5%). The bacterial communities of above-ground tissues (leaf blades and petioles) were more tightly clustered regardless of differences in the tissue types and fertilization conditions than those of below-ground tissues (taproots and lateral roots). The bacterial communities of below-ground tissues were largely separated by PC2 (26.0%). To survey plant growth-promoting bacteria (PGPBs), isolate collections (a total of 665 isolates) were constructed from the lateral roots. As candidate PGPBs, 44 isolates were selected via clustering analyses with the combined 16S rRNA gene sequence data of clone libraries and isolate collections. The results of inoculation tests using sugar beet seedlings showed that eight isolates exhibited growth-promoting effects on the seedlings. Among them, seven isolates belonging to seven genera (*Asticcacaulis*, *Mesorhizobium*, *Nocardioides*, *Sphingobium*, *Sphingomonas*, *Sphingopyxis*, and *Polaromonas*) were newly identified as PGPBs for sugar beet at the genus level, and two isolates belonging to two genera (*Asticcacaulis* and *Polaromonas*) were revealed to exert growth-promoting effects on the plant at the genus level for the first time. These results suggest that a community analysis-based selection strategy will facilitate the isolation of novel PGPBs and extend the potential for the development of novel biofertilizers.

Key words: 16S rRNA gene, biofertilizer, community analysis, plant growth-promoting bacteria, sugar beet

Approximately 20% of the world's sucrose production is derived from sugar beet (*Beta vulgaris* L.), the most important crop in temperate regions for sugar production (Godshall, 2012). Although the initial growth of sugar beet seedlings is often inhibited by environmental stress, such as nutrient deficiency and frost damage, sugar beet grows well under harsh environmental conditions once initial growth is established (Steingrobe, 2001; 2005). Sugar beet has recently been attracting attention as a source of bioenergy (Koga, 2008) because of its higher biomass production than other temperate crops (de Vries *et al.*, 2010). However, the mechanisms underlying the high productivity and stress tolerance of sugar beet have not yet been elucidated in detail. One possible explanation for these features may be the colonization of plant growth-promoting bacteria (PGPBs), which confer stress tolerance and growth-promoting effects in the seedling stage (Steingrobe, 2005; Toyota and Watanabe, 2013). To date, a number of bacterial species have been reported as PGPBs for sugar beet, including *Acinetobacter* (Shi *et al.*, 2009, 2010), *Bacillus* (Çakmakçi *et al.*, 1999, 2001; Shi *et al.*, 2009, 2010), *Burkholderia* (Çakmakçi *et al.*, 2001), *Chryseobacterium* (Shi *et al.*, 2009, 2010), *Pseudomonas* (Kloepper *et al.*, 1980; Dunne *et al.*, 1998; Çakmakçi *et al.*, 2001, 2006), *Paenibacillus* (Çakmakçi *et al.*, 2006), *Rhodobacter* (Çakmakçi *et al.*, 2006), and *Stenotrophomonas* spp. (Dunne *et al.*, 1998). However, the phylogenetic diversity of these PGPBs is limited to a certain range of taxonomic groups when considering the entire phylogenetic diversity of sugar beet-associated bacteria, as revealed in our previous studies (Okazaki *et al.*, 2014; Tsurumaru *et al.*, 2015). Surveys of PGPBs have frequently been conducted on the basis of plant growth-promoting traits (PGPTs), including 1-aminocyclopropane-1-carboxylic acid deaminase production, indole acetic acid (IAA) production, N_2 fixation, phosphate solubilization, pyrroloquinoline quinone production, siderophore production, and plant disease suppression (Dunne *et al.*, 1998; Çakmakçi *et al.*, 2001; Ahmad *et al.*, 2008; Shi *et al.*, 2009; Kumar *et al.*, 2012; Tani *et al.*, 2012; Bal *et al.*, 2013; Gupta *et al.*, 2014). However, assays for these traits are both time- and labor-intensive, which limits large-scale surveys on PGPBs. Recent studies revealed that a single PGPT is not a fully reliable marker for selecting PGPBs; multiple PGPTs are considered to contribute to the beneficial effects of a PGPB (Ahmad *et al.*, 2008; Kumar *et al.*, 2012; Quecine *et al.*, 2012). In addition, bacterial iso-

* Corresponding author. E-mail: sikeda67@affrc.go.jp;
Tel: +81–155–62–9276; Fax: +81–155–61–2127.

Citation: Okazaki, K., Tsurumaru, H., Hashimoto, M., Takahashi, H., Okubo, T., Ohwada, T., et al. (2021) Community Analysis-based Screening of Plant Growth-promoting Bacteria for Sugar Beet. *Microbes Environ* **36**: ME20137.
https://doi.org/10.1264/jsme2.ME20137

lates selected using known PGPTs often fail to produce the desired growth-promoting effects when they are inoculated onto plants even under experimental conditions, such as a growth chamber or greenhouse. Therefore, screening using known PGPTs is not currently considered to be an efficient strategy for identifying PGPBs. One possible explanation for this failure is the insufficient consideration of the colonization ability of PGPBs on plant tissue (Lugtenberg and Kamilova, 2009; Compant *et al.*, 2010).

It is conceivable that the colonization ability of a plant-associated microorganism is reflected in its abundance in the tissue of a plant species. Therefore, a community analysis-based screening of PGPBs may be an ecologically reasonable and powerful strategy for identifying and selecting novel PGPBs that are highly compatible with a plant (tissue). The diversity of plant-associated bacteria is also markedly affected by fertilization conditions in several major crops (Ikeda *et al.*, 2014; Unno *et al.*, 2015; Masuda *et al.*, 2016). Stable colonization on and in plant tissues under diverse environmental conditions is regarded as a favorite characteristic of PGPBs for their use in agricultural practice under field conditions. A metagenomic analysis of the phylogenetic diversity and functionality of taproot-associated bacterial community in sugar beet (Tsurumaru *et al.*, 2015) revealed the dominance of *Alphaproteobacteria* in taproot tissue, which is consistent with the findings of Shi *et al.* (2014), revealing the potential importance of the functionality of this bacterial group for high biomass production by sugar beet. Collectively, these findings suggest that *Alphaproteobacteria* is a promising candidate group for screening PGPBs for sugar beet.

To obtain a more detailed understanding of the role of plant-associated bacteria in the growth of sugar beet, the present study aimed to (i) reveal dominant bacterial groups in the above- and underground tissues of sugar beet under different fertilization conditions via a clone library analysis, (ii) build bacterial isolate collections as a resource for surveying PGPBs, and (iii) conduct a large-scale screening of PGPBs for sugar beet by employing the combined data of 16S rRNA gene sequences derived from clone libraries and isolate collections. As a result, a subset of novel PGPBs were efficiently selected and identified for sugar beet at the genus and species levels, indicating that a community analysis-based screening strategy is a powerful tool for surveying and selecting novel PGPBs for practical agricultural use.

Materials and Methods

Plant materials and sampling

Seeds of the sugar beet cultivar "Amahomare" were sown in pots (paper pot no. 1; Nippon Beet Sugar Manufacturing) under greenhouse conditions on March 16, 2010 and grown for 41 days. They were planted in a plot with standard fertilization (NPK plot), only P and K fertilization (PK plot, no N fertilization), or only K fertilization (K plot, no N or P fertilization). All plots were 31.2 m^2 in size, and planting was performed on April 26, 2010 in a long-term experimental field in Japan (42°89′20″N, 143°07′70″E, 94 m a.s.l.) that had been maintained under the rotation of upland crops with potato, maize, sugar beet, or soybean grown during the summer and no cultivation during the winter since 1994 at the Memuro Research Station of the Hokkaido Agricultural Research Center (Memuro, Hokkaido, Japan). Ammonium sulfate (150 kg of N $hectare^{-1}$ for the NPK plot), calcium superphosphate (250 kg of P_2O_5 $hectare^{-1}$ for the NPK and PK plots), and potassium sulfate (160 kg of K_2O kg $hectare^{-1}$ for the NPK, PK, and K plots) were applied as basal fertilizers. On July 12, 2010, based on visual inspections, nine healthy plants were randomly sampled from the NPK, PK, and K plots. They were carefully washed with tap water to remove loosely adhering soil and organic debris, rinsed with sterilized water, and individually separated into taproots, lateral roots, leaf blades, and petioles. Lateral roots on a taproot were collected using forceps. These tissues were stored at –30°C until used for the construction of a clone library or bacterial isolate collection. Soil samples from the NPK, PK, and K plots were also taken from between plants using an auger (between 5 and 15 cm in depth) at the time of sampling, and the chemical characteristics of these soils were elucidated by the Tokachi Nokyoren Agricultural Research Institute (Obihiro, Hokkaido, Japan). The analysis of each chemical characteristic was conducted using the following methods: pH (H_2O); pH meter, P_2O_5; Truog's method, K_2O, MgO, CaO, and CEC; Schollenberger's method, total nitrogen; dry combustion method, NO_3-N; hydrazine reduction method, NH_4-N; indophenol method, Phosphate absorption coefficient; SPAD simple method, according to Nakatsu *et al.* (2012).

Construction of 16S rRNA gene clone libraries from sugar beet-associated bacteria

The 16S rRNA gene clone libraries of sugar beet-associated bacteria were constructed using the leaf blades, petioles, taproots, and lateral roots of sugar beet grown in the NPK, PK, and K plots. Bacterial cells were individually extracted from the leaf blades, petioles, taproots, and lateral roots of nine plants, as previously reported (Ikeda *et al.*, 2009). Briefly, 25 g of leaf blades or petioles were collected from a plant and homogenized in a blender with 250 mL of bacterial cell extraction buffer. Taproots were cut into several pieces, and 100 g of tissue was homogenized in a blender with 500 mL of bacterial cell extraction buffer. Lateral roots (approximately 1 g) were ground in liquid nitrogen with a mortar and pestle and homogenized in a blender with 250 mL of bacterial cell extraction buffer. The metagenomic DNA of extracted bacterial cells was prepared according to the protocol of Ikeda *et al.* (2004). DNA samples were used as the template for the PCR amplification of the 16S rRNA gene. PCR amplification was performed using the universal primers 27F (5′-AGAGTTTGATCMTG GCTCAG-3′) and 1525R (5′-AAGGAGGTGWTCCARCC-3′), as previously reported (Someya *et al.*, 2013). After electrophoresis of the PCR products on a 1% agarose gel, amplicons of the expected size (approximately 1,500 bp) were purified using NucleoSpin Extract II (Macherey-Nagel). Equal amounts of amplicons derived from each of the nine plants were combined for each tissue and then cloned using the pGEM-T Easy TA cloning vector (Promega). A partial sequence of the 16S rRNA gene (corresponding to bases 109–665 of the gene in *Escherichia coli*) for each clone of the library was elucidated using the 27F primer by Takara Bio.

Construction of bacterial isolate collections of lateral root-associated bacteria

Lateral root-associated bacteria were isolated from the lateral roots of sugar beet grown in the PK or K plots based on the presumption of the proliferation of beneficial bacteria for plant growth as a compensatory effect under nutrient-limited conditions. Lateral roots derived from three plants grown in a plot (1 g from each plant, approximately 3 g in total) were homogenized in a mortar and pestle with 30 mL of 67 mM phosphate buffer (pH 7.0). The homogenate was filtered through a layer of Miracloth (Calbiochem), and an aliquot (100 μL) of the filtrate was spread onto a R2A (Becton, Dickinson, and Company) or tryptic soy agar (TSA; Becton, Dickinson, and Company) plate containing cycloheximide (50 mg mL^{-1}; Wako Pure Chemical Industries) as the antifungal

agent. After 1 week of cultivation at 24°C in the dark, colonies were randomly selected and subjected to single colony isolation twice. These isolates were suspended into R2A or tryptic soy broth liquid medium containing 15% glycerol and stored at –80°C until later use for DNA extraction and inoculation tests.

The sequencing of 16S rRNA genes for isolate collections was also conducted. Isolates were cultivated on R2A or TSA plates, and a portion of each colony was used for genomic DNA extraction as previously described by Okazaki *et al.* (2014). In addition, a partial sequence of the 16S rRNA gene for each isolate of the collections was elucidated using the 27F primer as previously described for the sequencing of clone libraries. Nearly the full length of the 16S rRNA gene sequence was elucidated for growth-promoting or growth-inhibiting bacteria using the r1L (5′-GTATTACCGCGG CTGCTGG-3′), 926f (5′-AAACTCAAAGGAATTGACGG-3′), and f2L (5′-CCAGCAGCCGCGGTAATAG-3′) primers by Takara Bio.

Clustering analysis of 16S rRNA gene sequences

Clustering analyses of 16S rRNA gene sequences were performed with the combined data of clone libraries and isolate collections. The sequence orientation and presence of non-16S rRNA gene sequences in the libraries were examined using Orientation-Checker (Ashelford *et al.*, 2006). Chimeric sequences in the libraries were removed using MALLARD (Ashelford *et al.*, 2006). Any sequence identified at the 99.9% threshold was discarded as a chimera. The remaining sequences were aligned using CLUSTAL W (Thompson *et al.*, 1994). Based on the alignment, a distance matrix was constructed using the DNADIST program from PHYLIP ver. 3.66 (http://evolution.genetics.washington.edu/phylip.html) with the default parameters. The resulting matrices were used as the input for the Mothur program (Schloss *et al.*, 2009) to create operational taxonomic units (OTUs) with a threshold value of 97% sequence identity and calculate the diversity indices of clone libraries and isolate collections. Library coverage was calculated as described by Kemp and Aller (2004).

Principal coordinate analysis (PCoA) of 16S rRNA gene clone libraries

The UniFrac program (Lozupone *et al.*, 2006) was applied to examine the similarities in community structures between clone libraries. A tree file generated by CLUSTAL W and an environment file that links a file to the library were used as the input for UniFrac to conduct PCoA with the abundance-weighted option.

Phylogenetic analysis of 16S rRNA gene sequences

The phylogenetic compositions (the relative abundance of taxa) in clone libraries and isolate collections were analyzed using the Classifier program in the Ribosomal Database Project (RDP; Wang *et al.*, 2007) with a confidence threshold of 80%. Statistical comparisons among clone libraries were conducted using the Library Compare program in RDP (Cole *et al.*, 2014).

Selection of sugar beet growth-promoting bacteria

To select sugar beet growth-promoting bacteria, OTUs in isolate collections were selected based on their phylogenetic novelty (less than 97% identity with the closest known species), tissue specificity (restricted presence in taproots or lateral roots), persistence among libraries or collections (stable presence regardless of fertilizer application conditions), or high abundance in a tissue (more than 1% relative abundance in a clone library or isolate collection) in taproots or lateral roots. In addition, OTUs exhibiting high identity to the 16S rRNA gene sequences of known PGPBs reported in previous studies were selected. From a practical viewpoint, OTUs displaying high identity to a plant, animal, or human pathogen were eliminated in the selection process of the present study.

One isolate was selected from each of the OTUs matching the criteria described above as an inoculum and cultivated on an R2A or TSA plate at 24°C for 3 days in the dark, and colonies were suspended in sterilized water. The bacterial cell suspension was washed and adjusted to an optical density at 660 nm of 0.1 with sterilized water as an inoculant.

Sugar beet seedlings were prepared as follows. Sugar beet seeds (cultivar "Rycka") were sterilized via soaking in 70% ethanol for 1 min followed by 1% sodium hypochlorite (containing 0.01% Tween 20) for 15 min. After rinsing with sterilized water, surface-sterilized seeds were covered with wet filter paper. They were placed in a sterile Petri dish and germinated at 25°C for 1 day in the dark. Commercial soil (Pot-ace N; Katakura & Co-op Agri Corporation, 200 mg N L^{-1}, 800 mg P L^{-1}, 200 mg K L^{-1}, 60 mg Mg L^{-1}) for nursing seedlings was sterilized via autoclaving at 121°C for 5 min, and 80 mL of soil was added to a pot (41×41×43.5 mm^3; Cell box, Meiwa). Two germinated seeds were planted in a pot and covered with 20 mL of soil. One milliliter of the bacterial inoculant was applied to a pot. Control seeds were inoculated with sterilized water. Seedlings were then grown in a plant growth chamber (16 h of light at 25°C and 8 h of darkness at 20°C) (NK system E5ZS-34; Nippon Medical & Chemical Instruments), and distilled water was supplied as needed to maintain the moisture content. After 1 week of cultivation, seedlings were thinned to one plant per pot. After 4 weeks of cultivation, a whole seedling was sampled and separated into shoots and roots. These tissues were air-dried at 80°C for 3 days and dry matter weight was measured. Twelve seedlings in a tray were used in an inoculation test with each isolate, and this test was repeated three or four times at different dates to ensure the reproducibility of PGP effects. In order to correct data variations among repeated tests at different dates, data for PGP effects were evaluated by Welch's *t*-tests (two-tailed) using the ratio data of dry weight relative to a control.

Phylogenetic tree analysis

In the phylogenetic tree analysis, sequences were aligned using the CLUSTAL W program. The neighbor-joining method was used to build the trees (Saitou and Nei, 1987). The PHYLIP format tree output was applied using the bootstrapping procedure with 1,000 replicates (Felsenstein, 1985). Trees were constructed with TreeView software (Page, 1996).

Statistical analysis

Welch's *t*-tests (two-tailed) were performed using JMP software version 12 (SAS Institute). A P value <0.05 was considered to be significant.

Accession numbers of nucleotide sequences

Nucleotide sequences were deposited into the DDBJ/EMBL/GenBank database. The sequence data for clone libraries and isolate collections were deposited under the accession nos. LC038237–LC040216 and LC040217–LC040864, respectively (Table S1). The nearly full length of the 16S rRNA gene sequences (approximately 1,400 bp) for sugar beet growth-promoting and growth-inhibiting bacteria were deposited under the accession nos. LC040865–LC040881 and LC602158–LC602165.

Results and Discussion

Clone library analyses of sugar beet-associated bacteria under different fertilization conditions

The effects of fertilization conditions on the chemical characteristics of soil in the NPK, PK, and K plots are summarized in Table S2. Although a statistical analysis was not performed due to the lack of replications at the plot level in field experiments, the following changes were generally observed. N levels in soils were similar among the plots. However, the shoot length and number of leaves in the PK and K plots was markedly shorter and smaller, respectively, than those in the NPK plot (Fig. 1).

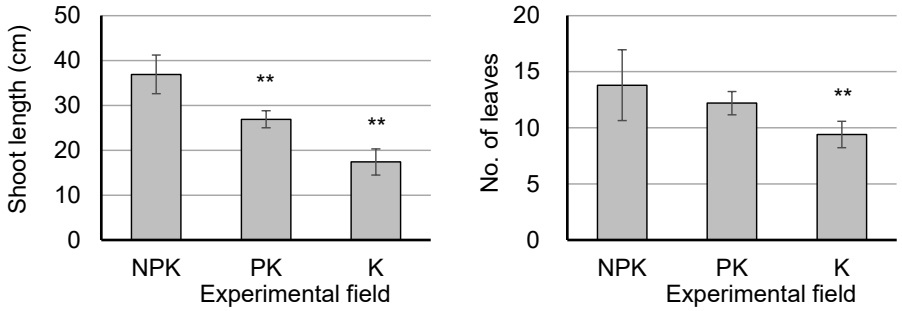


Fig. 1. Shoot length and number of leaves on sugar beet plants at the time of sampling. NPK, PK, and K denote plots with standard fertilization, only P and K fertilization (no N fertilization), and only K fertilization (no fertilization with N or P), respectively. Each value indicates the mean ±standard deviation of 10 individual plants. Double asterisks indicate a significant difference from NPK by Welch's *t*-test (two-tailed) at $P<0.01$.

Table 1. Alpha diversity indices of 16S rRNA gene sequences for clone libraries and isolate collections derived from sugar beet-associated bacteria

Libraries/Collections	Clone libraries												Isolate collections			
Tissues	Leaf blade (LB)			Petiole (PE)			Tap root (TR)			Lateral root (LR)			Lateral root (LR)			
Experimental fields	NPK	PK	K	NPK	PK	K	NPK	PK	K	NPK	PK	K	PK		K	
Isolation media	—	—	—	—	—	—	—	—	—	—	—	—	R2A	TSA	R2A	TSA
Library/collection names	LB-NPK	LB-PK	LB-K	PE-NPK	PE-PK	PE-K	TR-NPK	TR-PK	TR-K	LR-NPK	LR-PK	LR-K	LR-PK-R	LR-PK-T	LR-K-R	LR-K-T
Statistics																
No. of sequences	175	177	167	145	178	178	152	164	178	180	135	151	171	148	180	166
No. of OTUs[a]	42	37	35	40	35	31	70	72	71	97	75	85	52	48	54	48
No. of singletons	20	14	16	17	18	12	47	49	42	61	52	58	28	24	30	28
Library coverage (%)[b]	88.6	92.1	90.4	88.3	89.9	93.3	69.1	70.1	76.4	66.1	61.5	61.6	83.6	83.8	83.3	83.1
Diversity indexes																
Chao1	63	48	55	59	73	44	224	170	133	199	186	212	115	73	163	86
ACE	105	49	76	59	90	52	164	291	197	318	300	378	124	102	148	97
Shannon index (*H'*)	3.1	2.9	2.9	3.2	2.9	2.6	3.7	3.6	3.7	4.3	4.0	4.2	3.4	3.3	3.5	3.1
Simpson index (1/*D*)	15.3	12.5	11.9	17.9	13.6	8.0	25.8	16.1	25.1	79.0	54.8	67.0	21.3	15.5	26.1	11.8

[a] OTUs were defined at 97% sequence identity.
[b] Library coverage Cx was calculated as follows: Cx=1–(*n*/N), where *n* is the number of singletons that are encountered only once in a library or collection, and N is the total number of clones or isolates.

Library coverages were lower in the libraries for below-ground tissues (61.5–76.4%) than in those for above-ground tissues (88.3–93.3%; Table 1). As expected, the Chao1, ACE, and Shannon indices were higher in the libraries for below-ground tissues than in those for above-ground tissues (Table 1). The numbers of OTUs and the Shannon and Simpson indices were similar between leaf blades and petioles under the same fertilization conditions. Although these diversity indices for leaf blades and petioles were lower in response to the degree of deterioration in fertilization conditions in the PK and K plots than in the NPK plot, the same indices for taproots and lateral roots were both stable under all field conditions. These results suggest that the bacterial diversity of above-ground tissues is more sensitive to fertilization management than that of below-ground tissues. PCoA revealed that the diversity of sugar beet-associated bacteria was mainly clustered into three groups (above-ground tissues, taproots, and lateral roots; Fig. 2). The size of these clusters also suggested that the bacterial diversity of above-ground tissues was more sensitive to fertilization management than that of below-ground tissues in terms of the phylogenetic composition. Fluctuations in the relative abundance of many taxa strongly depended on fertilization management (Tables 2 and 3). As a result of the deterioration of plant nutrition in the PK and K plots, bacterial diversity in the above-ground tissues decreased. The relative abundance and number of OTUs belonging to *Firmicutes* decreased in the PK and K plots (Tables 2 and S3). In isolate collections, *Firmicutes* bacteria were rarely isolated on R2A medium. A high nutritional condition medium may be preferred for the efficient isolation of this bacterial group, or some pretreatment, such as heat shock, to break their dormancy may be required under nutrient-deficient condition. These results most likely reflect the differences in nutrient conditions for bacterial communities between above- and below-ground plant tissues. Therefore, above-ground plant tissue-associated bacteria almost totally depended on most of their nutrients through plant metabolism, while below-ground tissue-associated bacteria, such as some bacterial groups living on roots, depended on their nutrients through not only plant metabolism, but also soil.

Clusters of the bacterial communities of above-ground tissues were separated from those of below-ground tissues along PC1 (44.5%), whereas those of taproots and lateral roots were separated more clearly along PC2 (26.0%) than those for leaf blades and petioles, indicating that taproots and lateral roots harbor unique and distinct bacterial diversity.

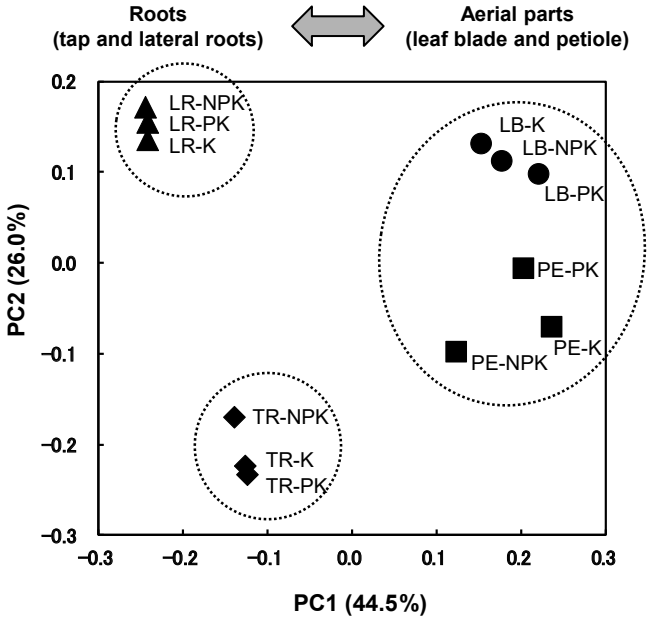


Fig. 2. Principal coordinate analysis (PCoA) of 16S rRNA gene sequences of sugar beet-associated bacteria. The library name is indicated on the right side of each symbol. Circles, squares, diamonds, and triangles denote leaf blade (LB)-, petiole (PE)-, taproot (TR)-, and lateral root (LR)-derived libraries, respectively. NPK, PK, and K denote plots with standard fertilization, only P and K fertilization (no N fertilization), and only K fertilization (no fertilization with N or P), respectively.

Phylogenetic composition analyses revealed that *Proteobacteria*, particularly *Alpha-* and *Gammaproteobacteria*, largely dominated the entire phytosphere of sugar beet (Table 2). In *Alphaproteobacteria*, *Rhizobiales* was exclusively found in the leaf blades, petioles, and taproots, followed by *Sphingomonadales* with less abundance. Two genera, namely, *Methylobacterium* and *Phyllobacterium*, were mainly responsible for the high abundance of *Rhizobiales* in above-ground tissues (Table 3). The high abundance of *Methylobacterium* in above-ground tissues (approximately 15–46%) was also observed in other plant species (Delmotte *et al.*, 2009; Someya *et al.*, 2013; Okubo *et al.*, 2014; Minami *et al.*, 2016; Hara *et al.*, 2019). *Devosia*, *Mesorhizobium*, and unclassified *Bradyrhizobiaceae* bacteria were uniquely found in below-ground tissues. Including these genera, the dominance of *Rhizobiales* in taproots (28–39%; Table 2) was demonstrated in our previous metagenome analysis of taproots (Tsurumaru *et al.*, 2015). At the genus level, *Rhizobium* was a common taxonomic group throughout the entire phytosphere of sugar beet, as previously reported for other plant species, such as *Arabidopsis* (Bodenhausen *et al.*, 2013) and potato (Someya *et al.*, 2013; Unno *et al.*, 2015). The high persistency of this genus in various tissues of diverse plant species suggests an unknown ecological role for *Rhizobium* in the phytosphere.

In *Sphingomonadales*, *Sphingomonas* and *Novosphingobium* were dominantly present in sugar beet tissues (Table 3). *Sphingomonas* was mainly found in above-ground tissues. In contrast, *Novosphingobium* was exclusively present in below-ground tissues, particularly in taproots as the most dominant genus. The distribution patterns of these two genera in the sugar beet phytosphere suggested an equivalent ecological role for these genera in the above- and below-ground tissues of sugar beet. The presence of *Novosphingobium* in taproots was reported in our previous metagenome analysis (Tsurumaru *et al.*, 2015); however, the relative abundance of *Novosphingobium* was small (approximately 3%) and the most dominant taxon in taproots was *Mesorhizobium* (14%). These differences between previous findings and the present results may be attributed to the lack of sufficient genomic data for the genus *Novosphingobium*. Alternately, the abundance of dominant taxa may be markedly affected by differences in the methodologies employed, climate conditions, growth stages (a single time point for sampling), and cultivars of sugar beet.

Gammaproteobacteria in sugar beet tissues mainly consisted of three taxonomic groups (*Enterobacteriaceae*, *Xanthomonadaceae*, and *Pseudomonadaceae*). *Enterobacteriaceae* and *Xanthomonadaceae* were exclusively found in above-ground tissues and lateral roots, respectively, whereas *Pseudomonadaceae* was detected in all tissues as a dominant genus (Table 2). *Verrucomicrobia* and *Planctomycetes* were mainly observed in below-ground tissues under all fertilization conditions. *Acidobacteria* and *Bacteroidetes* were only detected in lateral roots. The high abundance of *Niastella* in lateral roots is one of the unique characteristics of the sugar beet phytosphere (Table 3). *Niastella* was recently reported to be more abundant in sugarcane roots than in bulk soils (Yeoh *et al.*, 2016). *Niastella* sp. may be aggressively attracted to the rhizospheres of high sugar-accumulating crops.

Isolate collections of lateral root-associated bacteria

Bacteria were isolated from the lateral roots of sugar beet grown in the PK or K plot using R2A or TSA medium, and four isolate collections were constructed with 665 isolates (Table 1). Alpha diversity indices for the isolate collections revealed that all indices were higher in the isolate collections derived from R2A medium (LR-PK-R and LR-K-R) than in those derived from TSA medium (LR-PK-T and LR-K-T). Marked differences were observed between the LR-K-R and LR-K-T isolate collections. These results suggest that R2A medium is more suitable for isolating phylogenetically diverse bacteria than TSA medium, as reported in our previous studies (Okubo *et al.*, 2009; Okazaki *et al.*, 2014). For example, *Mesorhizobium*, *Neorhizobium*, *Nocardioides*, *Polaromonas*, and *Sphingomonas* were exclusively isolated with R2A medium, but not with TSA medium, whereas *Pantoea* was isolated with TSA medium only (Table 3). Low concentrations of carbon sources or NaCl in R2A medium may contribute to enhancing the propagation of oligotrophic bacteria. Nutrient-rich media, such as TSA and NA media, most of which were developed in medical science, appear to preferentially enhance the propagation of copiotrophic bacteria, such as *Pantoea*.

The results of the phylogenetic composition analysis (Table 3) suggested that the greatest difference between clone libraries and isolate collections was the relative abundance of *Niastella* bacteria. All currently known species in this genus have been isolated from soil, and may be grown on R2A agar plates (Weon *et al.*, 2006; Zhang *et al.*, 2010;

Table 2. Relative abundance of major taxa for clone libraries or isolate collections derived from sugar beet-associated bacteria[a]

Libraries/Collections	Clone libraries (%)												Isolate collections (%)			
Tissues	Leaf blade			Petiole			Tap root			Lateral root			Lateral root			
Experimental fields	NPK	PK	K	NPK	PK	K	NPK	PK	K	NPK	PK	K	PK		K	
Isolation medium	—	—	—	—	—	—	—	—	—	—	—	—	R2A	TSA	R2A	TSA
Library/collection name	LB-NPK	LB-PK	LB-K	PE-NPK	PE-PK	PE-K	TR-NPK	TR-PK	TR-K	LR-NPK	LR-PK	LR-K	LR-PK-R	LR-PK-T	LR-K-R	LR-K-T
Phylum																
Acidobacteria	—	—	—	—	—	—	—	—	—	3.9	2.2	4.0	—	—	—	—
Actinobacteria	5.1	7.9	5.4	9.7	2.2	2.8	7.2	6.1	3.9	6.7	1.5	4.6	34.5	52.0	14.4	42.2
Bacteroidetes	—	—	—	—	—	—	—	—	—	26.1	19.3	19.2	1.2	—	3.9	—
Firmicutes	6.9	4.0	1.2	8.3	2.2	1.1	2.6	0.6	1.1	0.6	—	1.3	—	3.4	0.6	3.6
Planctomycetes	—	0.6	0.6	—	—	—	4.6	4.3	6.7	2.8	0.7	1.3	—	—	—	—
Proteobacteria	88.0	87.6	92.2	81.4	95.5	94.4	79.6	86.0	83.1	56.1	69.6	61.6	64.3	44.6	81.1	54.2
Verrucomicrobia	—	—	—	—	—	—	4.6	2.4	2.2	1.7	1.5	4.0	—	—	—	—
Others	—	—	0.6	0.7	—	1.7	1.3	0.6	2.8	2.2	5.2	4.0	—	—	—	—
Class																
Actinobacteria	5.1	7.9	5.4	9.7	2.2	2.8	7.2	6.1	3.9	6.7	1.5	4.6	34.5	52.0	14.4	42.2
Alphaproteobacteria	29.1	41.2	22.8	57.2	58.4	69.7	46.7	59.8	59.6	15.6	25.9	23.2	52.0	29.7	61.1	27.7
Bacilli	6.9	2.8	1.2	8.3	2.2	1.1	2.6	0.6	1.1	—	—	0.7	—	3.4	0.6	3.6
Betaproteobacteria	2.9	1.1	10.2	9.0	2.8	2.2	13.2	6.7	9.6	13.3	9.6	14.6	1.8	2.7	9.4	—
Deltaproteobacteria	—	—	—	—	—	—	—	1.2	0.6	3.9	1.5	2.0	—	—	—	—
Gammaproteobacteria	56.0	45.2	59.3	15.2	34.3	22.5	19.7	17.7	11.8	22.8	31.9	21.2	10.5	12.2	10.6	26.5
Planctomycetia	—	0.6	0.6	—	—	—	4.6	4.3	6.7	2.8	0.7	1.3	—	—	—	—
Spartobacteria	—	—	—	—	—	—	3.9	1.8	1.1	—	—	0.7	—	—	—	—
Sphingobacteriia	—	—	—	—	—	—	—	—	—	23.3	17.8	16.6	1.2	—	1.1	—
Others	—	1.1	0.6	0.7	—	1.7	2.0	1.8	5.6	11.7	11.1	15.2	—	—	2.8	—
Order																
Actinomycetales	5.1	7.9	5.4	9.7	1.1	2.8	7.2	6.1	3.9	6.7	0.7	4.6	34.5	52.0	14.4	42.2
Bacillales	4.6	0.6	—	7.6	0.6	1.1	2.6	0.6	1.1	—	—	0.7	—	3.4	0.6	3.6
Burkholderiales	2.9	1.1	10.2	9.0	2.8	2.2	13.2	6.7	9.6	12.2	8.1	12.6	1.8	2.7	9.4	—
Caulobacterales	—	—	—	—	—	—	0.7	0.6	0.6	2.2	2.2	1.3	4.7	1.4	1.1	—
Enterobacteriales	33.1	28.8	41.3	9.0	19.1	12.9	0.7	—	1.1	1.1	—	1.3	0.6	1.4	2.2	19.9
Legionellales	—	—	—	—	—	0.6	3.9	3.0	0.6	—	—	—	—	—	—	—
Planctomycetales	—	0.6	0.6	—	—	—	4.6	4.3	6.7	2.8	0.7	1.3	—	—	—	—
Pseudomonadales	22.3	15.8	15.0	4.8	10.1	6.7	2.6	2.4	2.8	8.9	11.1	7.9	2.9	4.7	4.4	3.6
Rhizobiales	25.1	36.7	21.0	44.8	48.9	63.5	28.3	35.4	38.8	8.3	11.9	11.3	28.1	25.7	33.3	26.5
Sphingobacteriales	—	—	—	—	—	—	—	—	—	23.3	17.8	16.6	1.2	0.0	1.1	—
Sphingomonadales	4.0	2.8	1.8	11.7	9.6	4.5	16.4	23.2	18.5	3.9	11.1	9.9	19.3	2.7	25.0	1.2
Xanthomonadales	0.6	0.6	2.4	1.4	4.5	1.7	2.0	0.6	1.1	4.4	14.8	3.3	7.0	6.1	3.9	3.0
Unclassified *Gammaproteobacteria*	—	—	0.6	—	0.6	0.6	10.5	11.6	6.2	5.6	5.9	7.9	—	—	—	—
Others	2.3	5.1	1.8	2.1	2.8	3.4	7.2	5.5	9.0	20.6	15.6	21.2	—	—	4.4	—
Family																
Bacillaceae 1	3.4	—	—	2.8	—	1.1	1.3	—	0.6	—	—	0.7	—	2.7	0.6	3.0
Bradyrhizobiaceae	—	—	0.6	—	0.6	—	7.2	7.9	15.7	1.7	3.0	1.3	1.8	1.4	2.2	3.0
Burkholderiaceae	—	—	4.8	2.1	—	—	4.6	2.4	1.7	2.2	0.7	0.0	1.2	—	—	—
Caulobacteraceae	—	—	—	—	—	—	0.7	0.6	0.6	2.2	2.2	1.3	4.7	1.4	1.1	—
Chitinophagaceae	—	—	—	—	—	—	—	—	—	21.1	17.0	15.2	—	—	0.6	—
Comamonadaceae	1.1	1.1	4.8	6.2	2.8	2.2	3.3	1.8	1.7	4.4	3.0	8.6	—	0.7	8.9	—
Enterobacteriaceae	33.1	28.8	41.3	9.0	19.1	12.9	0.7	—	1.1	1.1	—	1.3	0.6	1.4	2.2	19.9
Hyphomicrobiaceae	—	—	—	0.7	—	—	1.3	1.2	4.5	1.7	0.7	2.0	4.1	8.1	5.0	7.8
Methylobacteriaceae	15.4	29.9	14.4	24.1	32.6	45.5	—	—	0.6	—	—	—	—	—	—	—
Microbacteriaceae	1.1	2.8	0.6	0.7	—	—	0.7	0.6	0.6	—	—	—	8.8	18.9	1.1	3.0
Micrococcaceae	3.4	2.8	4.2	4.8	0.6	2.2	—	—	—	—	—	—	1.2	0.7	1.1	3.0
Mycobacteriaceae	0.6	—	0.6	2.8	0.6	0.0	2.6	1.8	1.7	—	—	—	1.2	—	—	1.2
Nocardioidaceae	—	1.1	—	—	—	0.6	—	—	—	—	—	—	4.1	0.7	1.1	—
Oxalobacteraceae	1.7	—	0.6	0.7	—	—	3.3	1.8	5.6	5.6	4.4	4.0	—	—	0.6	—
Phyllobacteriaceae	5.1	2.8	0.6	13.8	10.1	9.6	3.3	6.7	7.3	0.6	0.0	2.0	10.5	6.1	6.7	1.2
Planctomycetaceae	—	0.6	0.6	—	—	—	4.6	4.3	6.7	2.8	0.7	1.3	—	—	—	—
Pseudomonadaceae	21.7	15.8	15.0	4.8	9.6	6.7	2.6	2.4	2.8	8.9	11.1	7.9	2.9	4.7	4.4	3.6
Rhizobiaceae	2.3	3.4	5.4	5.5	3.4	6.7	15.8	17.1	9.6	3.9	6.7	4.6	11.1	9.5	18.3	13.9
Sphingomonadaceae	4.0	2.8	1.8	11.7	9.6	4.5	16.4	23.2	18.5	3.9	11.1	9.9	19.3	2.7	24.4	1.2
Staphylococcaceae	—	—	—	4.1	0.6	—	—	—	—	—	—	—	—	—	—	—
Streptomycetaceae	—	—	—	—	—	—	2.6	2.4	1.1	5.0	0.7	4.6	16.4	25.7	9.4	33.1
Xanthomonadaceae	0.6	0.6	2.4	1.4	4.5	1.7	2.0	0.6	1.1	4.4	14.1	3.3	7.0	6.1	3.9	3.0
Others	6.3	7.3	2.4	4.8	6.2	6.2	27.0	25.0	18.5	30.6	24.4	31.8	5.3	9.5	8.3	3.0

[a] Gray indicates the taxa with a significant difference (P<0.05) to the NPK field library in each tissue.

Table 3. Relative abundance of major genera for clone libraries or isolate collections derived from sugar beet-associated bacteria[a]

Libraries/Collections	Clone libraries (%)												Isolate collections (%)			
Tissues	Leaf blade			Petiole			Tap root			Lateral root			Lateral root			
Experimental fields	NPK	PK	K	NPK	PK	K	NPK	PK	K	NPK	PK	K	PK		K	
Isolation media	—	—	—	—	—	—	—	—	—	—	—	—	R2A	TSA	R2A	TSA
Library/collection names	LB-NPK	LB-PK	LB-K	PE-NPK	PE-PK	PE-K	TR-NPK	TR-PK	TR-K	LR-NPK	LR-PK	LR-K	LR-PK-R	LR-PK-T	LR-K-R	LR-K-T
Genus																
Arthrobacter	2.9	2.8	4.2	4.8	0.6	2.2	—	—	—	—	—	—	0.6	—	0.6	3.0
Bacillus	2.9	—	—	2.8	—	1.1	1.3	—	0.6	—	—	0.7	—	2.7	0.6	3.0
Curtobacterium	—	2.8	—	—	—	—	—	—	—	—	—	—	—	—	—	—
Devosia	—	—	—	—	—	—	1.3	1.2	4.5	1.1	0.7	2.0	4.1	8.1	5.0	7.8
Enterobacter	—	—	3.0	—	—	—	—	—	—	—	—	—	—	—	—	—
Mesorhizobium	—	—	—	—	—	—	2.6	5.5	6.7	—	—	2.0	6.4	—	6.1	—
Methylobacterium	15.4	29.9	14.4	24.1	32.6	45.5	—	—	0.6	—	—	—	—	—	—	—
Microbacterium	0.6	—	—	—	—	—	—	0.6	—	—	—	—	5.3	16.9	1.1	1.8
Mycobacterium	0.6	—	0.6	2.8	0.6	—	2.6	1.8	1.7	—	—	—	1.2	—	—	1.2
Neorhizobium	—	—	—	—	—	—	—	1.2	1.7	2.2	3.7	2.0	2.9	—	7.8	—
Niastella	—	—	—	—	—	—	—	—	—	11.7	11.1	7.9	—	—	—	—
Nocardioides	—	—	—	—	—	0.6	—	—	—	—	—	—	4.1	—	1.1	—
Novosphingobium	—	—	—	—	0.6	—	15.8	22.0	14.0	3.3	8.1	4.6	1.8	2.0	10.6	1.2
Pantoea	5.1	2.3	6.6	2.8	—	5.6	—	—	1.1	1.1	—	0.7	—	1.4	—	4.8
Phyllobacterium	5.1	2.8	0.6	13.8	10.1	9.6	—	1.2	0.6	—	—	—	0.6	2.7	—	0.6
Polaromonas	—	—	—	—	—	—	2.6	1.2	0.6	2.2	1.5	3.3	—	—	7.2	—
Pseudomonas	21.7	15.8	15.0	4.8	9.6	6.7	2.6	2.4	2.8	7.2	7.4	6.0	2.9	4.7	3.9	3.0
Ralstonia	—	—	3.6	1.4	—	—	—	—	—	—	—	—	—	—	—	—
Rhizobium	2.3	3.4	4.8	5.5	3.4	6.7	15.1	15.2	7.9	1.7	3.0	2.6	7.0	9.5	9.4	13.3
Sphingomonas	4.0	2.8	1.8	11.7	9.0	4.5	—	—	—	0.6	1.5	1.3	13.5	—	7.2	—
Streptomyces	—	—	—	—	—	—	2.6	2.4	1.1	5.0	0.7	4.6	16.4	24.3	9.4	33.1
Variovorax	1.1	1.1	4.8	6.2	2.8	1.7	0.7	0.6	0.6	1.7	—	—	—	0.7	1.1	—
Yersinia	2.9	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
Unclassified *Bradyrhizobiaceae*	—	—	—	—	—	—	0.7	1.2	5.1	—	1.5	0.7	—	—	0.6	—
Unclassified *Chitinophagaceae*	—	—	—	—	—	—	—	—	—	7.2	2.2	6.0	—	—	—	—
Unclassified *Enterobacteriaceae*	22.3	20.9	28.1	4.1	16.9	5.1	0.7	—	—	—	—	0.7	—	—	1.1	10.8
Others	13.1	15.3	12.6	15.2	14.0	10.7	51.3	43.3	50.6	55.0	58.5	55.0	33.3	27.0	27.2	16.3

[a] Gray highlight indicates taxa with a significant difference ($P<0.05$) from the NPK field library in each tissue.

Kim *et al.*, 2015). The genus *Niastella* comprises six species, two of which (*Niastella koreensis* and *Niastella yeongjuensis*) were isolated from soil cultivated with ginseng (Weon *et al.*, 2006). However, Vendan *et al.* (2010) failed to isolate a *Niastella* bacterium from ginseng tissue. These findings and the present results suggest that plant-associated *Niastella* is recalcitrant to isolation with standard R2A medium.

Clustering analysis of 16S rRNA gene sequences of clones and isolates and selection of sugar beet growth-promoting bacteria

A total of 2,645 sequences of the 16S rRNA gene from culture-independent clones (1,980 clones) and culture-dependent isolates (665 isolates) were clustered into 456 OTUs at a sequence identity of 97% (Table S3). Among 279 lateral root-relating OTUs, 93 consisted of only isolates, while 30 contained both clone and isolate sequences. Among the remaining 156 OTUs, no isolates were obtained in the present study (Table S3). Based on the criteria for the selection of OTUs, such as the specificity and stability of colonization to a tissue and phylogenetic novelty, the representative isolates of 44 OTUs were selected (Table S4) as candidate PGPBs and inoculated onto the seedlings of sugar beet. The results obtained revealed eight and six isolates as plant growth-promoting and growth-inhibiting bacteria, respectively (Fig. 3 and S1). BLAST search analyses using 16S rRNA gene sequences indicated that two isolates exerting plant growth-promoting effects (*Asticcacaulis* sp. RK043 and *Mesorhizobium* sp. TP027) are potential novel species based on their identity to known species (Table 4). Furthermore, phylogenetic tree analyses suggested that three isolates (*Asticcacaulis* sp. RK043, *Mesorhizobium* sp. TP027, and *Rhizobacter* sp. RK021) are novel species based on their phylogenetic positions (Fig. S2, S3, and S4).

Variovorax are typical phylogenetic groups of PGPBs for a wide range of plant species, including sugar beet (Zhou *et al.*, 2017). Seven genera (*Asticcacaulis*, *Mesorhizobium*, *Nocardioides*, *Sphingobium*, *Sphingomonas*, *Sphingopyxis*, and *Polaromonas*) were newly identified as PGPBs for sugar beet, and two (*Asticcacaulis* and *Polaromonas*) were demonstrated to exert growth-promoting effects on a plant for the first time.

The selection of PGPBs has often been conducted based on known PGPTs, such as nitrogen fixation and IAA production. However, the examination of known PGPTs is a time-consuming and laborious task that has been the bottleneck for the large-scale selection of PGPBs. More impor-

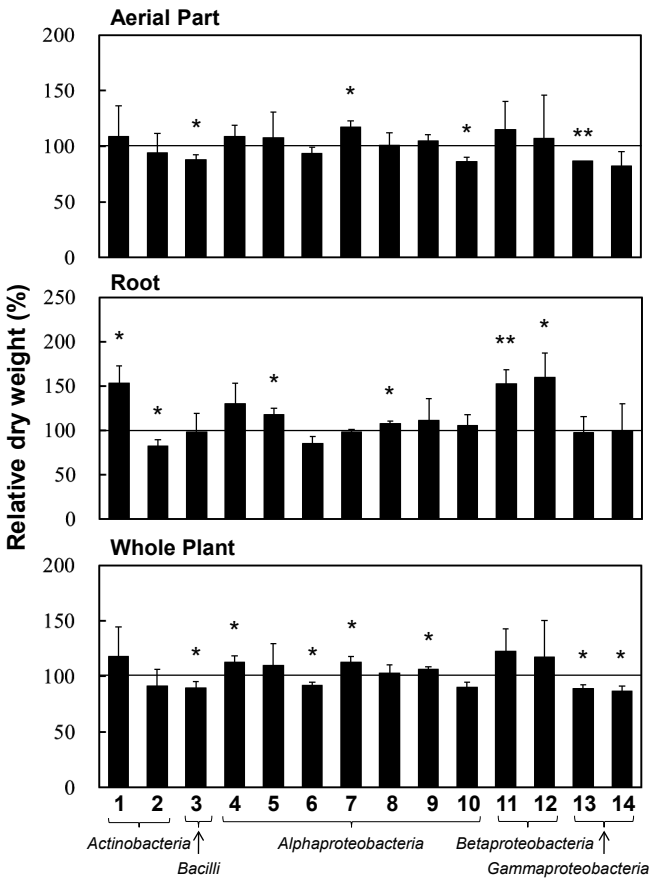


Fig. 3. Inoculation effects of sugar beet lateral root-associated bacteria on sugar beet seedling growth. Lateral root-associated bacteria were inoculated onto sugar beet seedlings grown in a pot, and the dry weights of seedlings were measured after 4 weeks of cultivation. Control plants were inoculated with sterilized water. Twelve seedlings in a tray were used in an inoculation test with each isolate, and this test was repeated three or four times at different dates to ensure the reproducibility of PGP effects. The dry weights of inoculated seedlings were compared to those of control seedlings (non-inoculated seedlings) by Welch's *t*-test (two-tailed). Error bars indicate the standard deviation. Single and double asterisks indicate a significant difference at $P<0.05$ and $P<0.01$, respectively. Bacterial isolates: 1, *Nocardioides* sp. RP110; 2, *Streptomyces* sp. TP071; 3, *Bacillus* sp. TP182; 4, *Asticcacaulis* sp. RK043; 5, *Mesorhizobium* sp. TP027; 6, *Neorhizobium* sp. RK064; 7, *Sphingobium* sp. RK166; 8, *Sphingomonas* sp. RP195; 9, *Sphingopyxis* sp. RK106; 10, *Tardiphaga* sp. RK140; 11, *Polaromonas* sp. RK103; 12, *Variovorax* sp. RK170; 13, *Pantoea* sp. RK126; 14, *Rhizobacter* sp. RK021.

tantly, possession of the activity of a known PGPT does not always guarantee that a PGPB will exert growth-promoting effects in inoculation tests (Barazani and Friedman, 1999; Cardinale *et al.*, 2015). In contrast, in the present study, large-scale isolate collections were initially constructed from target tissues via random isolation, and the candidate isolates of PGPBs were then selected based on 16S rRNA gene sequence data. Clustering analyses using the combined sequence data of clone libraries and isolate collections proposed four criteria (the novelty of the sequence, relative abundance and persistence in a target tissue, and its identity to known beneficial or nonpathogenic bacteria) as an indication for selecting candidate isolates of PGPB (Table S4). The ability to colonize a plant tissue is an important trait for PGPBs (Lugtenberg and Kamilova, 2009; Quecine *et al.*, 2012), and the relative abundance of an OTU in a tissue is considered to reflect its compatibility and persistence in a plant tissue.

In addition, deleterious bacteria for sugar beet seedlings were identified as growth-inhibiting bacteria in the present study (*Bacillus* sp. TP182, *Neorhizobium* sp. RK064, *Pantoea* sp. RK126. *Rhizobacter* sp. RK021, *Tardiphaga* sp. RK140, and *Streptomyces* sp. TP071 in Table 4). *Bacillus* spp. are generally considered to be beneficial bacteria because of their growth-promoting effects on many plants, including sugar beet (Çakmakçi *et al.*, 2006; Shi *et al.*, 2010; Park *et al.*, 2017). The present results and previous findings revealed the difficulty of selecting PGPB based solely on phylogenetic information, and indicate the importance of an inoculation test to screen PGPBs on plants. Based on the relative abundance of OTUs in root tissues, deleterious bacteria are considered to exhibit a competitive colonization ability relative to PGPBs (Table 5). The characterization and ecological control of deleterious bacteria are also important for maximizing the effects of PGPBs. Deleterious bacteria may have the ability to interfere with the beneficial effects of PGPBs on plant tissue, and in addition to several physiochemical factors and the genetic background of crops, this interference may partially explain why the effects of PGPBs are often unstable, even under practical field conditions (Timmusk *et al.*, 2005).

Conclusion

The present study revealed some of the characteristics of the phylogenetic composition of sugar beet-associated bacteria and identified eight isolates of novel PGPBs and six isolates of deleterious bacteria for sugar beet at the species level. The majority of these PGPBs belonging to seven genera (*Nocardioides*, *Asticcacaulis*, *Mesorhizobium*, *Sphingobium*, *Sphingomonas*, *Sphingopyxis*, and *Polaromonas*) were newly identified as PGPBs for sugar beet at the genus level, and two isolates belonging to two genera (*Asticcacaulis* and *Polaromonas*) were identified as PGPBs on a plant at the genus level for the first time. These results demonstrated that a community analysis-based selection is a highly efficient strategy for the initial selection of PGPBs in combination with large-scale isolate collections that increases the likelihood of identifying novel PGPBs. Further analyses of the biochemical and ecological characteristics of beneficial/deleterious bacteria isolated in the present study will provide a more detailed understanding of plant-microbe interactions under field conditions and possibly facilitate the utilization of these beneficial bacteria in agricultural practice for reducing chemical use in sugar beet production.

Acknowledgements

The present study was supported by the Ministry of Agriculture, Forestry and Fisheries of Japan through projects entitled "Development of technologies for mitigation and adaptation to climate change in agriculture, forestry and fisheries". A part of this study was also supported by JST CREST Grant Number JPMJCR1512 and JSPS KAKENHI Grant Number 19K05759 from the Ministry of Education, Science, Sports and Culture of Japan.

Table 4. BlastN search results with nearly the full sequence of the 16S rRNA gene of growth-promoting and -inhibiting bacteria for sugar beet seedlings

OTU No.	Isolate name	Inoculation effect[a]	BlastN search results[b]	
			Closest known species	Identity (%)
	Actinobacteria			
OTU-301	*Nocardioides* sp. RP110 (LC040866)[c]	+	*Nocardioides cavernae* (NR_156135)	100
OTU-272	*Streptomyces* sp. TP071 (LC602158)	–	*Streptomyces mirabilis* (EF371431)	100
	Bacilli			
OTU-329	*Bacillus* sp. TP182 (LC602159)	–	*Bacillus gibsonii* (FJ937920)	100
	Alphaproteobacteria			
OTU-226	*Asticcacaulis* sp. RK043 (LC040869)	+	*Asticcacaulis benevestitus* (NR_042433)	98
OTU-170	*Mesorhizobium* sp. TP027 (LC040873)	+	*Mesorhizobium chacoense* (NR_025411)	98
OTU-166	*Neorhizobium* sp. RK064 (LC602160)	–	*Neorhizobium galegae* (HG938355)	100
OTU-189	*Sphingobium* sp. RK166 (LC602162)	+	*Sphingobium aromaticiconvertens* (MF101093)	100
OTU-191	*Sphingomonas* sp. RP195 (LC602164)	+	*Sphingomonas asaccharolytica* (NR_029327)	100
OTU-199	*Sphingopyxis* sp. RK106 (LC602163)	+	*Sphingopyxis taejonensis* (NR_024999)	100
OTU-218	*Tardiphaga* sp. RK140 (LC602161)	–	*Tardiphaga robiniae* (CP050292)	99
	Betaproteobacteria			
OTU-87	*Polaromonas* sp. RK103 (LC040879)	+	*Polaromonas ginsengisoli* (AB245355)	100
OTU-86	*Variovorax* sp. RK170 (LC040880)	+	*Variovorax paradoxus* (CP002417)	100
	Gammaproteobacteria			
OTU-69	*Pantoea* sp. RK126 (LC602165)	–	*Pantoea ananatis* (CP020943)	100
OTU-96	*Rhizobacter* sp. RK021 (LC040878)	–	*Methylibium petroleiphilum* (CP000555)	99

[a] "+" and "–" indicate growth-promoting and -inhibiting bacteria, respectively.
[b] Results with approximately 1,400 bp using the NCBI database (https://blast.ncbi.nlm.nih.gov/Blast.cgi) are shown.
[c] The numbers in parentheses indicate the accession number.

Table 5. Relative abundance of OTUs in clone libraries of 16S rRNA genes derived from sugar beet-associated bacteria

Tissues	Leaf blade			Petiole			Tap root			Lateral root			Isolation
Experimental fields	NPK	PK	K	NPK	PK	K	NPK	PK	K	NPK	PK	K	
Growth-inhibiting bacteria													
OTU-69	—	—	—	—	—	—	—	—	—	—	—	—	*Pantoea* sp. RK126 (LC602165)
OTU-96	—	—	—	—	—	—	—	—	—	0.6	4.4	2.0	*Rhizobacter* sp. RK021 (LC040878)
OTU-166	—	—	—	—	—	—	—	1.2	1.7	2.2	3.7	2.0	*Neorhizobium* sp. RK064 (LC602160)
OTU-218	—	—	—	—	—	—	2.6	1.8	4.5	0.6	0.7	—	*Tardiphaga* sp. RK140 (LC602161)
OTU-272	—	—	—	—	—	—	—	—	—	1.7	0.7	0.7	*Streptomyces* sp. TP071 (LC602158)
OTU-329	1.7	—	—	2.8	—	—	1.3	—	0.6	—	—	—	*Bacillus* sp. TP182 (LC602159)
Growth-promoting bacteria													
OTU-86	1.1	1.1	4.8	6.2	2.8	1.7	0.7	0.6	0.6	1.7	—	—	*Variovorax* sp. RK170 (LC040880)
OTU-87	—	—	—	—	—	—	2.6	1.2	0.6	2.2	1.5	3.3	*Polaromonas* sp. RK103 (LC040879)
OTU-170	—	—	—	—	—	—	—	—	—	—	—	—	*Mesorhizobium* sp. TP027 (LC040873)
OTU-189	—	—	—	—	—	—	0.7	—	2.8	—	1.5	2.0	*Sphingobium* sp. RK166 (LC602162)
OTU-191	—	—	—	—	—	—	—	—	—	0.6	1.5	1.3	*Sphingomonas* sp. RP195 (LC602164)
OTU-199	—	—	—	—	—	—	—	—	—	—	—	—	*Sphingopyxis* sp. RK106 (LC602163)
OTU-226	—	—	—	—	—	—	—	—	—	—	0.7	—	*Asticcacaulis* sp. RK043 (LC040869)
OTU-301	—	—	—	—	—	—	—	—	—	—	—	—	*Nocardioides* sp. RP110 (LC040866)

References

Ahmad, F., Ahmad, I., and Khan, M.S. (2008) Screening of free-living rhizospheric bacteria for their multiple plant growth promoting activities. *Microbiol Res* **163**: 173–181.

Ashelford, K.E., Chuzhanova, N.A., Fry, J.C., Jones, A.J., and Weightman, A.J. (2006) New screening software shows that most recent large 16S rRNA gene clone libraries contain chimeras. *Appl Environ Microbiol* **72**: 5734–5741.

Bal, H.B., Nayak, L., Das, S., and Adhya, T.K. (2013) Isolation of ACC deaminase producing PGPR from rice rhizosphere and evaluating their plant growth promoting activity under salt stress. *Plant Soil* **366**: 93–105.

Barazani, O.Z., and Friedman, J. (1999) Is IAA the major root growth factor secreted from plant-growth-mediating bacteria? *J Chem Ecol* **25**: 2397–2406.

Bodenhausen, N., Horton, M.W., and Bergelson, J. (2013) Bacterial communities associated with the leaves and the roots of *Arabidopsis thaliana*. *PLoS One* **8**: e56329.

Çakmakçi, R., Kantar, F., and Algur, Ö.F. (1999) Sugar beet and barley yields in relation to *Bacillus polymyxa* and *Bacillus megaterium var. phosphaticum* inoculation. *J Plant Nutr Soil Sci* **162**: 437–442.

Çakmakçi, R., Kantar, F., and Sahin, F. (2001) Effect of N2-fixing bacterial inoculations on yield of sugar beet and barley. *J Plant Nutr Soil Sci* **164**: 527–531.

Çakmakçi, R., Dönmez, F., Aydın, A., and Şahin, F. (2006) Growth promotion of plants by plant growth-promoting rhizobacteria under greenhouse and two different field soil conditions. *Soil Biol Biochem* **38**: 1482–1487.

Cardinale, M., Ratering, S., Suarez, C., Zapata Montoya, A.M., Geissler-Plaum, R., and Schnell, S. (2015) Paradox of plant growth promotion potential of rhizobacteria and their actual promotion effect on growth of barley (*Hordeum vulgare* L.) under salt stress. *Microbiol Res* **181**: 22–32.

Cole, J.R., Wang, Q., Fish, J.A., Chai, B., McGarrell, D.M., Sun, Y., *et al.* (2014) Ribosomal Database Project: data and tools for high throughput rRNA analysis. *Nucleic Acids Res* **42**: D633–D642.

Compant, S., Clément, C., and Sessitsch, A. (2010) Plant growth-promoting bacteria in the rhizo-and endosphere of plants: their role, colonization, mechanisms involved and prospects for utilization. *Soil Biol Biochem* **42**: 669–678.

de Vries, S.C., van de Ven, G.W., van Ittersum, M.K., and Giller, K.E. (2010) Resource use efficiency and environmental performance of nine major biofuel crops, processed by first-generation conversion techniques. *Biomass Bioenergy* **34**: 588–601.

Delmotte, N., Knief, C., Chaffron, S., Innerebner, G., Roschitzki, B., Schlapbach, R., *et al.* (2009) Community proteogenomics reveals insights into the physiology of phyllosphere bacteria. *Proc Natl Acad Sci U S A* **106**: 16428–16433.

Dunne, C., Moënne-Loccoz, Y., McCarthy, J., Higgins, P., Powell, J., Dowling, D.N., and O'gara, F. (1998) Combining proteolytic and phloroglucinol-producing bacteria for improved biocontrol of *Pythium*-mediated damping-off of sugar beet. *Plant Pathol* **47**: 299–307.

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

Godshall, M.A. (2012) Sugar and other sweeteners. In *Handbook of Industrial Chemistry and Biotechnology*. Kent, J.A. (ed.) New York, NY: Springer Science+Business Media, pp. 1403–1430.

Gupta, S., Meena, M.K., and Datta, S. (2014) Isolation, characterization of plant growth promoting bacteria from the plant *Chlorophytum borivilianum* and *in-vitro* screening for activity of nitrogen fixation, phosphate solubilization and IAA production. *Int J Curr Microbiol Appl Sci* **3**: 1082–1090.

Hara, S., Matsuda, M., and Minamisawa, K. (2019) Growth stage-dependent bacterial communities in soybean plant tissues: *Methylorubrum* transiently dominated in the flowering stage of the soybean shoot. *Microbes Environ* **34**: 446–450.

Ikeda, S., Watanabe, K.N., Minamisawa, K., and Ytow, N. (2004) Evaluation of soil DNA from arable land in Japan using a modified direct-extraction method. *Microbes Environ* **19**: 301–309.

Ikeda, S., Kaneko, T., Okubo, T., Rallos, L.E., Eda, S., Mitsui, H., *et al.* (2009) Development of a bacterial cell enrichment method and its application to the community analysis in soybean stems. *Microb Ecol* **58**: 703–714.

Ikeda, S., Suzuki, K., Kawahara, M., Noshiro, M., and Takahashi, N. (2014) An assessment of urea-formaldehyde fertilizer on the diversity of bacterial communities in onion and sugar beet. *Microbes Environ* **29**: 231–234.

Kemp, P.F., and Aller, J.Y. (2004) Bacterial diversity in aquatic and other environments: what 16S rDNA libraries can tell us. *FEMS Microbiol Ecol* **47**: 161–177.

Kim, S.J., Ahn, J.H., Weon, H.Y., Hong, S.B., Seok, S.J., Kim, J.S., and Kwon, S.W. (2015) *Niastella gongjuensis* sp. nov., isolated from greenhouse soil. *Int J Syst Evol Microbiol* **65**: 3115–3118.

Kloepper, J.W., Leong, J., Teintze, M., and Schroth, M.N. (1980) Enhanced plant growth by siderophores produced by plant growth-promoting rhizobacteria. *Nature* **286**: 885–886.

Koga, N. (2008) An energy balance under a conventional crop rotation system in northern Japan: Perspectives on fuel ethanol production from sugar beet. *Agric, Ecosyst Environ* **125**: 101–110.

Kumar, P., Dubey, R.C., and Maheshwari, D.K. (2012) *Bacillus* strains isolated from rhizosphere showed plant growth promoting and antagonistic activity against phytopathogens. *Microbiol Res* **167**: 493–499.

Lozupone, C., Hamady, M., and Knight, R. (2006) UniFrac—an online tool for comparing microbial community diversity in a phylogenetic context. *BMC Bioinf* **7**: 371.

Lugtenberg, B., and Kamilova, F. (2009) Plant-growth-promoting rhizobacteria. *Annu Rev Microbiol* **63**: 541–556.

Masuda, S., Bao, Z., Okubo, T., Sasaki, K., Ikeda, S., Shinoda, R., *et al.* (2016) Sulfur fertilization changes the community structure of rice root-, and soil-associated bacteria. *Microbes Environ* **31**: 70–75.

Minami, T., Anda, M., Mitsui, H., Sugawara, M., Kaneko, T., Sato, S., *et al.* (2016) Metagenomic analysis revealed methylamine and ureide utilization of soybean-associated *Methylobacterium*. *Microbes Environ* **31**: 268–278.

Nakatsu, S., Ueno, T., Sakurai, M., Nakamoto, H., Takeuchi, H., and Sakaguchi, M. (2012) Soil chemical properties. In *Diagnostic Criteria for Soil and Crop Nutrition-Analysis Method 2012*. Shiga, H. (ed.) Sapporo, Japan: Hokkaido Research Organization, pp. 57–93 (in Japanese).

Okazaki, K., Iino, T., Kuroda, Y., Taguchi, K., Takahashi, H., Ohwada, T., *et al.* (2014) An assessment of the diversity of culturable bacteria from main root of sugar beet. *Microbes Environ* **29**: 220–223.

Okubo, T., Ikeda, S., Kaneko, T., Eda, S., Mitsui, H., Sato, S., *et al.* (2009) Nodulation-dependent communities of culturable bacterial endophytes from stems of field-grown soybeans. *Microbes Environ* **24**: 253–258.

Okubo, T., Ikeda, S., Sasaki, K., Ohshima, K., Hattori, M., Sato, T., *et al.* (2014) Phylogeny and functions of bacterial communities associated with field-grown rice shoots. *Microbes Environ* **29**: 329–332.

Page, R.D. (1996) Tree View: An application to display phylogenetic trees on personal computers. *Bioinformatics* **12**: 357–358.

Park, Y.G., Mun, B.G., Kang, S.M., Hussain, A., Shahzad, R., Seo, C.W., *et al.* (2017) *Bacillus aryabhattai* SRB02 tolerates oxidative and nitrosative stress and promotes the growth of soybean by modulating the production of phytohormones. *PLoS One* **12**: e0173203.

Quecine, M.C., Araújo, W.L., Rossetto, P.B., Ferreira, A., Tsui, S., Lacava, P.T., *et al.* (2012) Sugarcane growth promotion by the endophytic bacterium *Pantoea agglomerans* 33.1. *Appl Environ Microbiol* **78**: 7511–7518.

Saitou, N., and Nei, M. (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* **4**: 406–425.

Schloss, P.D., Westcott, S.L., Ryabin, T., Hall, J.R., Hartmann, M., Hollister, E.B., *et al.* (2009) Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl Environ Microbiol* **75**: 7537–7541.

Shi, Y., Lou, K., and Li, C. (2009) Promotion of plant growth by phytohormone-producing endophytic microbes of sugar beet. *Biol Fertil Soils* **45**: 645–653.

Shi, Y., Lou, K., and Li, C. (2010) Growth and photosynthetic efficiency promotion of sugar beet (*Beta vulgaris* L.) by endophytic bacteria. *Photosynth Res* **105**: 5–13.

Shi, Y., Yang, H., Zhang, T., Sun, J., and Lou, K. (2014) Illumina-based analysis of endophytic bacterial diversity and space-time dynamics in sugar beet on the north slope of Tianshan mountain. *Appl Microbiol Biotechnol* **98**: 6375–6385.

Someya, N., Kobayashi, Y.O., Tsuda, S., and Ikeda, S. (2013) Molecular characterization of the bacterial community in a potato phytosphere. *Microbes Environ* **28**: 295–305.

Steingrobe, B. (2001) Root renewal of sugar beet as a mechanism of P uptake efficiency. *J Plant Nutr Soil Sci* **164**: 533–539.

Steingrobe, B. (2005) A sensitivity analysis for assessing the relevance of fine-root turnover for P and K uptake. *J Plant Nutr Soil Sci* **168**: 496–502.

Tani, A., Sahin, N., Matsuyama, Y., Enomoto, T., Nishimura, N., Yokota, A., and Kimbara, K. (2012) High-throughput identification and screening of novel *Methylobacterium* species using whole-cell MALDI-TOF/MS analysis. *PLoS One* **7**: e40784.

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

Timmusk, S., Grantcharova, N., and Wagner, E.G.H. (2005) *Paenibacillus polymyxa* invades plant roots and forms biofilms. *Appl Environ Microbiol* **71**: 7292–7300.

Toyota, K., and Watanabe, T. (2013) Recent trends in microbial inoculants in agriculture. *Microbes Environ* **28**: 403–404.

Tsurumaru, H., Okubo, T., Okazaki, K., Hashimoto, M., Kakizaki, K., Hanzawa, E., *et al.* (2015) Metagenomic analysis of the bacterial community associated with the taproot of sugar beet. *Microbes Environ* **30**: 63–69.

Unno, Y., Shinano, T., Minamisawa, K., and Ikeda, S. (2015) Bacterial community shifts associated with high abundance of Rhizobium spp. in potato roots under macronutrient-deficient conditions. *Soil Biol Biochem* **80**: 232–236.

Vendan, R.T., Yu, Y.J., Lee, S.H., and Rhee, Y.H. (2010) Diversity of endophytic bacteria in ginseng and their potential for plant growth promotion. *J Microbiol* **48**: 559–565.

Wang, Q., Garrity, G.M., Tiedje, J.M., and Cole, J.R. (2007) Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl Environ Microbiol* **73**: 5261–5267.

Weon, H.Y., Kim, B.Y., Yoo, S.H., Lee, S.Y., Kwon, S.W., Go, S.J., and Stackebrandt, E. (2006) *Niastella koreensis* gen. nov., sp. nov. and *Niastella yeongjuensis* sp. nov., novel members of the phylum *Bacteroidetes*, isolated from soil cultivated with Korean ginseng. *Int J Syst Evol Microbiol* **56**: 1777–1782.

Yeoh, Y.K., Paungfoo-Lonhienne, C., Dennis, P.G., Robinson, N., Ragan, M.A., Schmidt, S., and Hugenholtz, P. (2016) The core root microbiome of sugarcanes cultivated under varying nitrogen fertilizer application. *Environ Microbiol* **18**: 1338–1351.

Zhang, K., Wang, Y., Tang, Y., Dai, J., Zhang, L., An, H., *et al.* (2010) *Niastella populi* sp. nov., isolated from soil of Euphrates poplar (*Populus euphratica*) forest, and emended description of the genus *Niastella*. *Int J Syst Evol Microbiol* **60**: 542–545.

Zhou, N., Zhao, S., and Tian, C.Y. (2017) Effect of halotolerant rhizobacteria isolated from halophytes on the growth of sugar beet (*Beta vulgaris* L.) under salt stress. *FEMS Microbiol Lett* **364**: fnx091.

Microbes Environ. Vol. 29, No. 4, 427-430, 2014
https://www.jstage.jst.go.jp/browse/jsme2 doi:10.1264/jsme2.ME14054

Short Communication

Intestinal Colonization by a *Lachnospiraceae* Bacterium Contributes to the Development of Diabetes in Obese Mice

Keishi Kameyama[1]*, and Kikuji Itoh[2]

[1]*Institute for Innovation, Ajinomoto Co., Inc., Kawasaki 210–8681, Japan; and* [2]*Department of Veterinary Public Health, Graduate School of Agricultural and Life Sciences, The University of Tokyo, Tokyo 113–8657, Japan*

(Received April 4, 2014—Accepted July 17, 2014—Published online October 4, 2014)

The aim of the present study was to identify bacteria that may contribute to the onset of metabolic dysfunctions. We isolated and identified a candidate bacterium belonging to *Lachnospiraceae* (strain AJ110941) in the feces of hyperglycemic obese mice. The colonization of germ-free *ob/ob* mice by AJ110941 induced significant increases in fasting blood glucose levels as well as liver and mesenteric adipose tissue weights, and decreases in plasma insulin levels and HOMA-β values. These results indicated that the specific gut commensal bacterium AJ110941 influenced the development of obesity and diabetes in *ob/ob* mice with genetic susceptibility for obesity.

Key words: gut microbiota, diabetes, T-RFLP, gnotobiotic mouse, *Lachnospiraceae*

The gut microbiota has been shown to influence host energy homeostasis, metabolism, and inflammation; thus, microbial communities within the gut are now recognized as an important environmental factor in the onset of obesity and type 2 diabetes (T2D) (5, 7, 12, 14, 23). As a consequence, interest in the development of methods to target gut microbiota as a therapy for T2D is increasing (2, 12, 18, 21).

Several studies have conducted metagenomic analyses in an attempt to identify members of the gut microbial community that contribute most significantly to the progression of T2D. Qin *et al.* performed a metagenomic analysis on fecal samples obtained from 345 Chinese men and women with or without T2D (19). Karlsson *et al.* also conducted a metagenomic analysis to compare microbial communities in fecal samples collected from 145 European women who had either T2D, impaired glucose metabolism, or were healthy (13). Both of these studies reported that butyrate-producing *Roseburia* species and *Faecalibacterium prauznitzii* were rarer in subjects with T2D; however, the identified bacteria that existed in high population in T2D subjects differed between these studies (8). Cani *et al.* proposed that lipopolysaccharide (LPS), a cell wall component of almost all Gram-negative bacteria, may be a key substance responsible for metabolic endotoxemia, low-grade systemic inflammation, and insulin resistance in mouse models (1, 3). In support of this, a previous study reported that experimental endotoxemia induced adipose inflammation and insulin resistance in human subjects (17).

We herein searched for the specific gut commensal bacterium related to metabolic syndrome using the terminal restriction fragment length polymorphism (T-RFLP) analysis of fecal samples from a mouse model of T2D. In this study, a comparison of homozygous *db/db* (diabetic) with heterozygous *db/+* (non-diabetic) mice revealed that a specific fragment of the T-RFLP analysis was present at significantly higher levels in *db/db* mice than in *db/+* mice. Even though *ob/ob* obese mice gain weight rapidly and show the symptoms of insulin resistance, most of the obese mice do not develop severe hyperglycemia compared with *db/db* mice (22). However, one of the *ob/ob* mice in the present study showed an abnormally high fasting blood glucose (FBG) level similar to *db/db* mice. The specific fragment observed in *db/db* mice was also the most prominent in the abnormally hyperglycemic *ob/ob* mouse. We hypothesized that the specific bacterium harboring the fragment may have contributed to the progression of T2D; therefore, we isolated and cultured the bacterium (strain AJ110941). We then determined that colonization by the isolate with the LPS producer *E. coli* induced hyperglycemia and the accumulation of adipose tissue in the gnotobiotic *ob/ob* mouse model.

All experimental procedures were reviewed and approved by the Animal Care Committee of Ajinomoto Co., Inc. Male 5-week-old homozygous BKS.Cg-Dock7^{m} +/+Lepr^{db}/J (*db/db*, diabetic) mice, heterozygous control (*db/+*, non-diabetic) mice, and homozygous B6.V-Lep^{ob}/J (*ob/ob*, obese) mice were obtained from Charles River Japan (Yokohama, Japan). Male 5-week-old germ-free *ob/ob* mice were obtained from Sankyo Lab Service (Tokyo, Japan). These mice were housed in a controlled environment (on a 12-h light/dark cycle with lights turning off at 19:00) with free access to standard chow CRF-1 (Oriental Yeast, Tokyo, Japan) and water and kept in specific pathogen-free (SPF) or germ-free/gnotobiotic conditions throughout the experimental period. Fresh feces and blood were collected from 5- and 11-week-old mice after fasting for 16 h.

Blood glucose (FBG) levels were determined using DRI-CHEM 7000V (Fujifilm, Tokyo, Japan). Plasma insulin levels were determined using an ultrasensitive mouse insulin kit (Morinaga Institute of Biological Science, Yokohama, Japan). Plasma glucagon levels were determined using the

* Corresponding author. E-mail: keishi_kameyama@ajinomoto.com

Glucagon ELISA Kit Wako (Wako Pure Chemical Industries, Osaka, Japan). Insulin resistance (HOMA-IR) and β-cell function (HOMA-β) were both calculated on the basis of the fasting levels of plasma glucose and insulin according to the homeostasis model assessment (HOMA) method (16).

Fresh fecal samples were collected from mice, added to 99% ethanol, and stored at −30°C. Bacterial DNA was extracted from the fecal samples using the FastDNA spin kit for soil (MP Biomedicals, Santa Ana, CA) using the FastPrep instrument (MP Biomedicals). T-RFLP analyses of the mouse gut microbiota were performed as previously described (10). Two universal primers, 27F labeled with 6-carboxyfluorescein (FAM): 5'-FAM-AGAGTTTGATCCTGGCTCAG-3' and 1492R: 5'-GGTTACCTTGTTACGACTT-3' were used in PCR to amplify the 16S rRNA gene coding region. Purified PCR products of the 16S rRNA gene were digested with *Msp*I (Takara Bio, Otsu, Japan).

To isolate and culture the bacterium, fresh feces were collected and immediately weighed and transferred to an anaerobic chamber. The feces were homogenized with a 50-fold volume (v/w) of an anaerobic diluent, serially diluted, and then plated on Eggerth-Gagnon (EG) agar plates (11). The agar plates were incubated at 37°C for 4 d in the anaerobic chamber. All of the colonies were identified by their fragment size, and a target colony with a fragment size of 282 bp was then selected. The colony was passaged onto new EG agar plates, and the bacterial purity of the isolate was ensured by re-streaking and microscopic observations. The 16S rRNA gene sequence of the isolated bacterium was determined as previously described (15). The sequence was assembled using GENETYX version 7 (GENETYX, Tokyo, Japan). The identification and phylogenic tree analysis of the 16S rRNA gene sequence were carried out using the Ribosomal Database Project (RDP) (6).

When the mice were 8 weeks old, germ-free *ob/ob* mice were separated into 3 groups for the germ-free/gnotobiotic experiment: Group-1 (n=4), germ-free; Group-2 (n=4), colonization by the non-pathogenic *Escherichia coli* strain E-17, which had been isolated from SPF mice; Group-3 (n=4), colonization by AJ110941 with *E. coli* E-17. Groups-2 and -3 were then orally inoculated with *E. coli* E-17 or AJ110941 with *E. coli* E-17 (approximately 1×10^8 cells suspended in anaerobic PBS, respectively). Group-1 was orally administered PBS only. These mice were maintained under the germ-free or gnotobiotic conditions for 8 weeks.

We compared differences in gut microbiota patterns between *db/+* and *db/db* mice. Hyperglycemia was observed in 11-week-old, but not in 5-week-old mice (Fig. S1). The 282 bp fragment was significantly higher in *db/db* than *db/+* mice at both 5 and 11 weeks old (Fig. S2). We then evaluated the presence of the 282 bp fragment in another metabolic syndrome model, the *ob/ob* mouse. The highest FBG level observed among all *ob/ob* mice was in Mouse Number 4 (Fig. S3). The 282 bp fragment was also more prominent in hyperglycemic Mouse Number 4 than in the normal glycemic *ob/ob* mouse at both 5 and 11 weeks old (Fig. S4). Based on these *db/db* and *ob/ob* mice results, we focused on this 282 bp fragment bacteria.

We isolated a colony harboring the 282 bp fragment and established an axenic strain (strain AJ110941). The isolate was determined to be closely related to the genus *Anaerostipes* in the family *Lachnospiraceae* (Fig. 1).

We then generated gnotobiotic *ob/ob* mice colonized by AJ110941 with *E. coli*. Eight weeks after the inoculation, the cecum, liver, adipose tissue, and blood were collected after a 16-h fast. The gut microbiota of the cecal contents from the three groups was determined by T-RFLP analysis. No PCR amplicon was detected in Group-1. Only a single fragment (488 bp) that corresponded to *E. coli* was detected in Group-2. Only two fragments (282 bp and 488 bp) were detected in Group-3. These results clearly indicated that germ-free or gnotobiotic conditions were maintained during the experimental period. The weights of the liver and mesenteric adipose tissue significantly increased in Group-3, whereas no significant difference was observed in body weights between the three groups. FBG and plasma glucagon levels were significantly higher in Group-3, while plasma insulin levels were significantly lower. On the other hand, no significant differences were observed in these parameters between Groups-1 and -2. The homeostasis model assessment was calculated from FBG and plasma insulin levels as an index of insulin resistance (HOMA-IR) and pancreatic β cell function (HOMA-β). HOMA-β was significantly lower in Group-3, whereas no significant differences were noted in HOMA-IR between the three groups (Table 1). These results suggested that colonization by AJ110941 may have promoted the dysfunction of pancreatic β-cells.

Therefore, AJ110941 should be regarded as one of the important causative gut bacteria for the induction of T2D. A previous study reported that the relative abundance of the

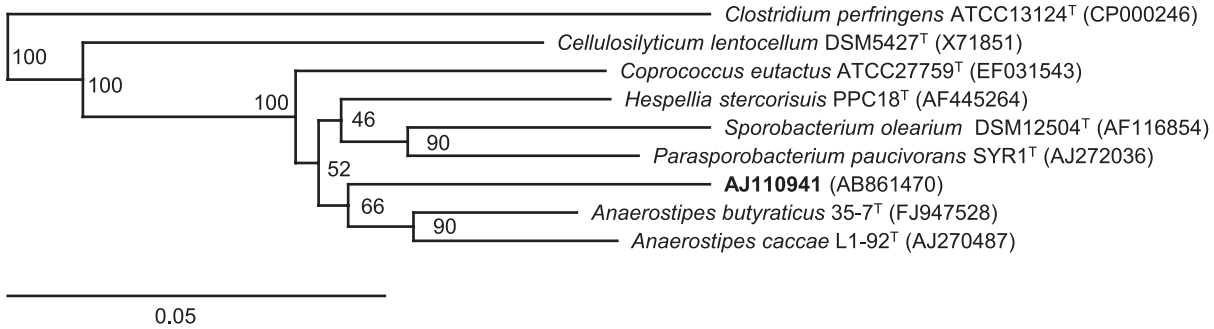


Fig. 1. Phylogenetic tree based on the 16S rRNA gene constructed by Weighbor, the weighted neighbor-joining tree building algorithm, showing AJ110941 and other strains of the family *Lachnospiraceae*. *Clostridium perfringens* ATCC13124^T was used as an outgroup. The numbers close to the nodes represent bootstrap values (n=100 replicates). The scale bar represents 0.05 substitutions per nucleotide position.

Table 1. Phenotypes of germ-free and gnotobiotic *ob/ob* mice

	Group-1	Group-2	Group-3
Body weight without the caecum (g)	58.10 ± 5.98^{a}	61.75 ± 4.29^{a}	63.48 ± 10.37^{a}
Cecum (% of body weight)	11.63 ± 4.89^{a}	12.88 ± 1.94^{a}	4.47 ± 0.93^{b}
Liver (% of body weight without the cecum)	6.91 ± 0.66^{a}	7.29 ± 1.83^{a}	9.70 ± 0.46^{b}
Mesenteric adipose tissue (% of body weight without the cecum)	1.64 ± 0.46^{a}	1.81 ± 0.39^{a}	2.63 ± 0.30^{b}
Fasting blood glucose (mg dL^{-1})	216.5 ± 113.6^{a}	286.5 ± 54.1^{ab}	433.25 ± 65.4^{b}
Fasting plasma insulin (μU mL^{-1})	339.7 ± 46.44^{a}	299.8 ± 105.0^{a}	159.4 ± 50.9^{b}
Fasting plasma glucagon (pg mL^{-1})	179.9 ± 43.0^{a}	324.2 ± 42.4^{a}	529.0 ± 18.3^{b}
HOMA-IR	132.5 ± 23.3^{a}	187.5 ± 70.6^{a}	184.4 ± 64.1^{a}
HOMA-β	1350.4 ± 468.9^{a}	638.5 ± 411.9^{ab}	141.7 ± 40.7^{b}

Group-1: germ-free, Group-2: colonization by *E. coli*, Group-3: colonization by AJ110941 with *E. coli*. Data are expressed as means±SD. Values not sharing a common letter are significantly different at $p<0.05$ by the Tukey–Kramer multiple comparisons test ($n=4$ per group) using the JMP 10.0.0 statistical software package (SAS Institute, Cary, NC).

taxonomic family *Lachnospiraceae* was increased by early-life subtherapeutic antibiotic treatments in an obese mouse model (4). Additionally, a metagenomic study indicated that the taxonomic family *Lachnospiraceae* may be associated with T2D (19). However, it remains unclear whether bacteria belonging to the family *Lachnospiraceae* actually affect obesity and FBG levels *in vivo*. Therefore, we generated gnotobiotic *ob/ob* mice colonized by AJ110941 with *E. coli* to reveal a possible causal relationship. In a preliminary examination, we microscopically observed that AJ110941 did not singly colonize the intestinal tract of *ob/ob* germ-free mice. *E. coli*, which is a facultative anaerobe, may be needed to maintain an oxygen-free environment in the intestinal tract because AJ110941 needs strict anaerobic conditions for growth. In our study, a mono-association with the LPS producer *E. coli* induced neither hyperglycemia nor the accumulation of adipose tissues. In contrast, LPS derived from Gram-negative bacteria in the gastrointestinal tract was previously identified as one of the most important factors inducing the development of T2D (1, 3). Our results indicated that the presence of LPS in the intestinal tract was necessary, but not sufficient for the pathogenesis of diabetes. We speculated that AJ110941 may have assisted with the translocation of LPS into the blood from the intestinal tract. We are currently investigating the effects of AJ110941 on LPS translocation in *in vitro*/*vivo* models.

Regarding the particular gut bacterium involved in metabolic syndrome, *Methanobrevibacter smithii* and *Bacteroides thetaiotaomicron* were previously shown to enhance host energy storage in di-associated mice (20). Furthermore, Fei and Zhao reported that the strain *Enterobacter cloacae* B29, which was isolated from an obese human subject, induced obesity and insulin resistance accompanied by serum endotoxemia in mono-associated mice (9).

This is the first study to have successfully identified a specific *Lachnospiraceae* bacterium involved in metabolic disorders. Future studies are needed to elucidate the molecular mechanisms underlying the adverse effects of AJ110941 on glucose and lipid metabolism in mouse models. It remains unclear whether AJ110941 or its closely-related species inhabits the human intestinal tract. We intend to perform a preliminary epidemiological study on obese and diabetic subjects to address this question.

The 16S rRNA gene sequence of the isolated bacterium (strain AJ110941) is available in the DDBJ/EMBL/GenBank databases under the accession number AB861470.

Acknowledgements

We thank Ranko Nishi and Mizuki Kobayashi for their research assistance.

References

1. Cani, P.D., J. Amar, M.A. Iglesias, *et al.* 2007. Metabolic endotoxemia initiates obesity and insulin resistance. Diabetes 56:1761–1772.
2. Cani, P.D., and N.M. Delzenne. 2012. The gut microbiome as therapeutic target. Pharmacol. Ther. 130:202–212.
3. Cani, P.D., S. Possemiers, T. Van de Wiele, *et al.* 2009. Changes in gut microbiota control inflammation in obese mice through a mechanism involving GLP-2-driven improvement of gut permeability. Gut 58:1091–1103.
4. Cho, I., S. Yamanishi, L. Cox, *et al.* 2012. Antibiotics in early life alter the murine colonic microbiome and adiposity. Nature 488:621–626.
5. Clemente, J.C., L.K. Ursell, L.W. Parfrey, and R. Knight. 2012. The impact of the gut microbiota on human health: an integrative view. Cell 148:1258–1270.
6. Cole, J.R., Q. Wang, J.A. Fish, *et al.* 2014. Ribosomal Database Project: data and tools for high throughput rRNA analysis. Nucleic Acids Res. 42:D633–642.
7. Cox, L.M., and M.J. Blaser. 2013. Pathways in microbe-induced obesity. Cell Metab. 17:883–894.
8. de Vos, W.M., and M. Nieuwdorp. 2013. Genomics: A gut prediction. Nature 498:48–49.
9. Fei, N., and L. Zhao. 2013. An opportunistic pathogen isolated from the gut of an obese human causes obesity in germfree mice. ISME J. 7:880–884.
10. Hayashi, H., M. Sakamoto, M. Kitahara, and Y. Benno. 2003. Molecular analysis of fecal microbiota in elderly individuals using 16S rDNA library and T-RFLP. Microbiol. Immunol. 47:557–570.
11. Itoh, K., and T. Mitsuoka. 1985. Characterization of clostridia isolated from faeces of limited flora mice and their effect on caecal size when associated with germ-free mice. Lab. Anim. 19:111–118.
12. Johnson, A.M., and J.M. Olefsky. 2013. The origins and drivers of insulin resistance. Cell 152:673–684.
13. Karlsson, F.H., V. Tremaroli, I. Nookaew, G. Bergström, C.J. Behre, B. Fagerberg, J. Nielsen, and F. Bäckhed. 2013. Gut metagenome in European women with normal, impaired and diabetic glucose control. Nature 498:99–103.
14. Le Chatelier, E., T. Nielsen, J. Qin, *et al.* 2013. Richness of human gut microbiome correlates with metabolic markers. Nature 500:541–546.
15. Liu, W.T., T.L. Marsh, H. Cheng, and L.J. Forney. 1997. Characterization of microbial diversity by determining terminal restriction fragment length polymorphisms of genes encoding 16S rRNA. Appl. Environ. Microbiol. 63:4516–4522.
16. Matthews, D.R., J.P. Hosker, A.S. Rudenski, B.A. Naylor, D.F. Treacher, and R.C. Turner. 1985. Homeostasis model assessment: insulin resistance and beta-cell function from fasting plasma glucose and insulin concentrations in man. Diabetologia 28:412–419.
17. Mehta, N.N., F.C. McGillicuddy, P.D. Anderson, *et al.* 2010. Experimental endotoxemia induces adipose inflammation and insulin resistance in humans. Diabetes 59:172–181.

18. Panwar, H., H.M. Rashmi, V.K. Batish, and S. Grover. 2013. Probiotics as potential biotherapeutics in the management of type 2 diabetes—prospects and perspectives. Diabetes Metab. Res. Rev. 29:103–112.
19. Qin, J., Y. Li, Z. Cai, *et al.* 2012. A metagenome-wide association study of gut microbiota in type 2 diabetes. Nature 490:55–60.
20. Samuel, B.S., and J.I. Gordon. 2006. A humanized gnotobiotic mouse model of host-archaeal-bacterial mutualism. Proc. Natl. Acad. Sci. U.S.A. 103:10011–10016.
21. Sanz, Y., A. Santacruz, and P. Gauffin. 2010. Gut microbiota in obesity and metabolic disorders. Proc. Nutr. Soc. 69:434–441.
22. Shafrir, E., E. Ziv, and L. Mosthaf. 1999. Nutritionally induced insulin resistance and receptor defect leading to beta-cell failure in animal models. Ann. N. Y. Acad. Sci. 892:223–246.
23. Tremaroli, V., and F. Bäckhed. 2012. Functional interactions between the gut microbiota and host metabolism. Nature 489:242–249.

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

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).

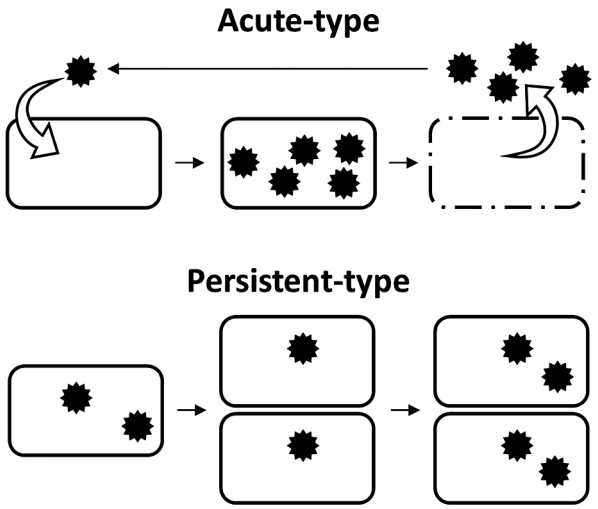


Fig. 1. Schematic of transmission routes of microbial RNA viruses.

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

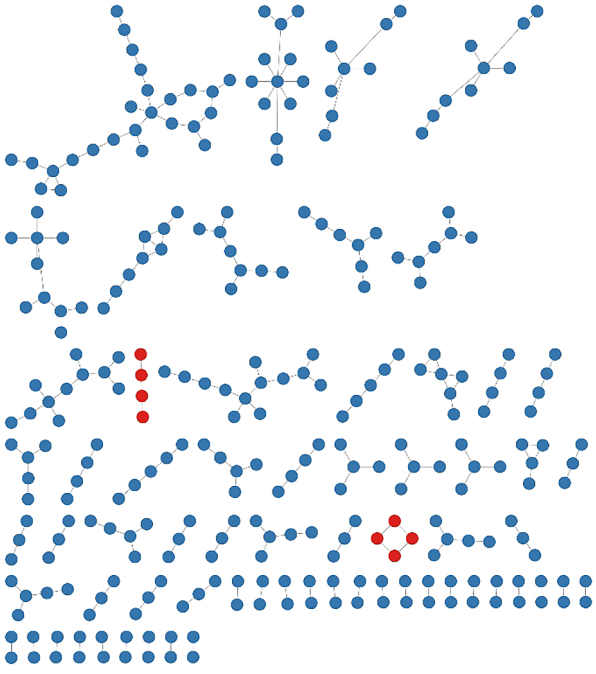


Fig. 2. A total of 293 isolated acute- and persistent-type microbial RNA viruses in sequence-based clusters. Colored circles indicate the lifestyle of each RNA virus: red, acute; blue, persistent. In total, 315 isolated microbial eukaryotic RNA viruses were collected from the manually curated Virus-Host 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).

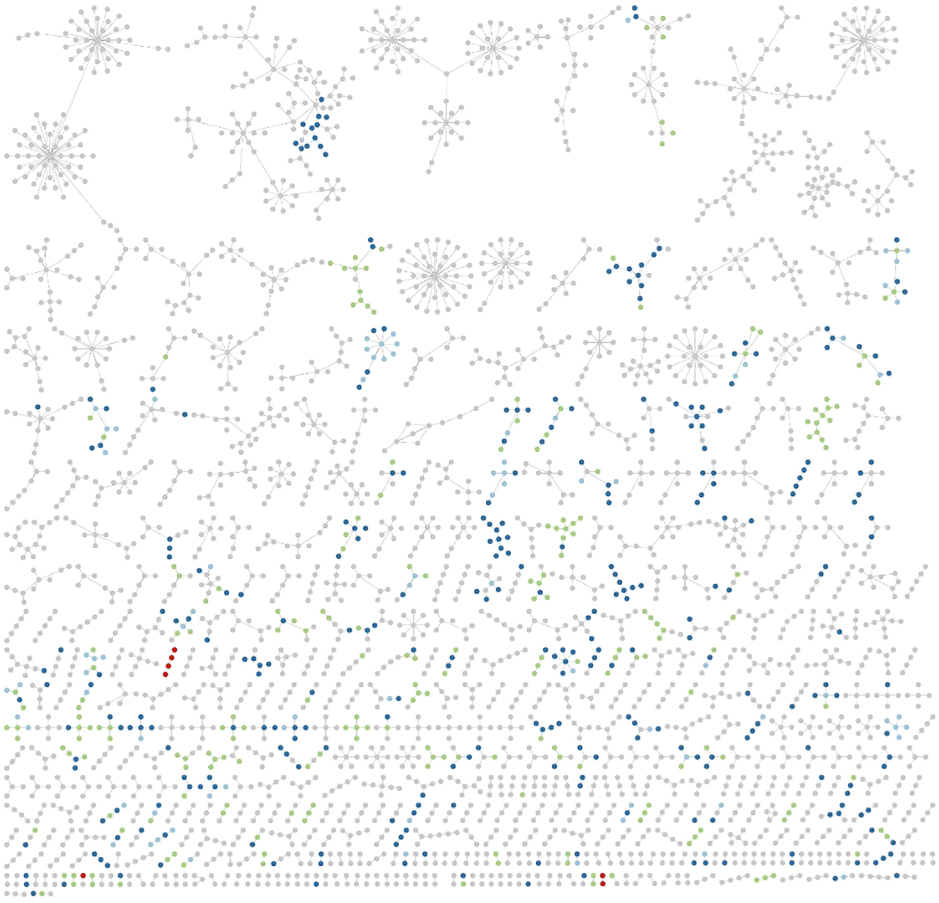


Fig. 3. RdRp sequences of known RNA viruses obtained from the Identical Protein Groups resource (https://www.ncbi.nlm.nih.gov/ipg) with keywords "rna dependent rna polymerase" and "viruses". After the removal of short (<200 aa) sequences, RdRp sequences were clustered at 70% using CD-HIT (Huang *et al.*, 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

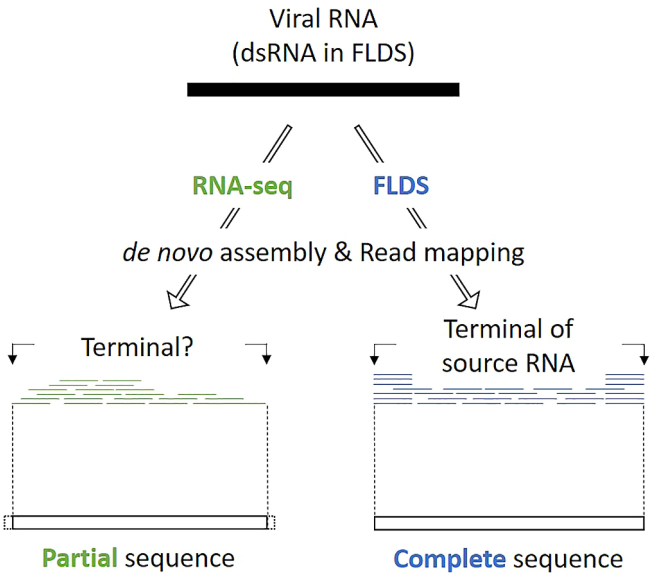


Fig. 4. Differences in conventional RNA-seq and FLDS for resultant contigs. In a *de novo* analysis, terminal sequence positions were not defined by RNA-seq data. However, FLDS data enabled us to identify terminal sequence positions because FLDS sequence reads included RACE (Rapid Amplification of cDNA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

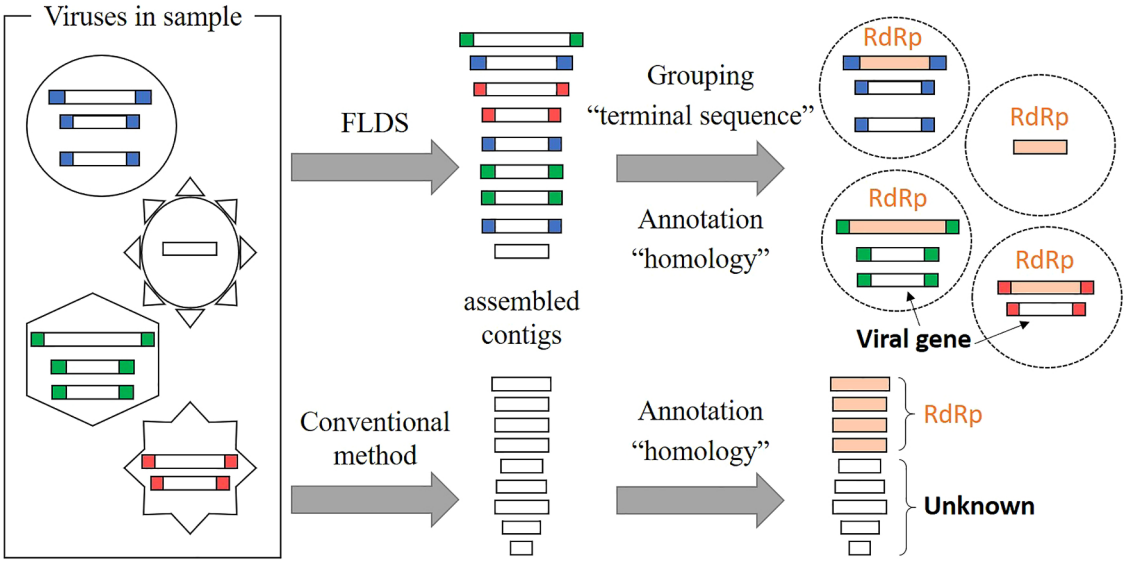


Fig. 5. Concept of RNA viral genome reconstruction based on conserved terminal sequences in segmented RNA virus genomes. Many segmented RNA viruses have conserved 5′- and 3′-terminal sequences (colored boxes). FLDS enabled us to obtain full-length RNA sequences, which are difficult to obtain with conventional 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

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.

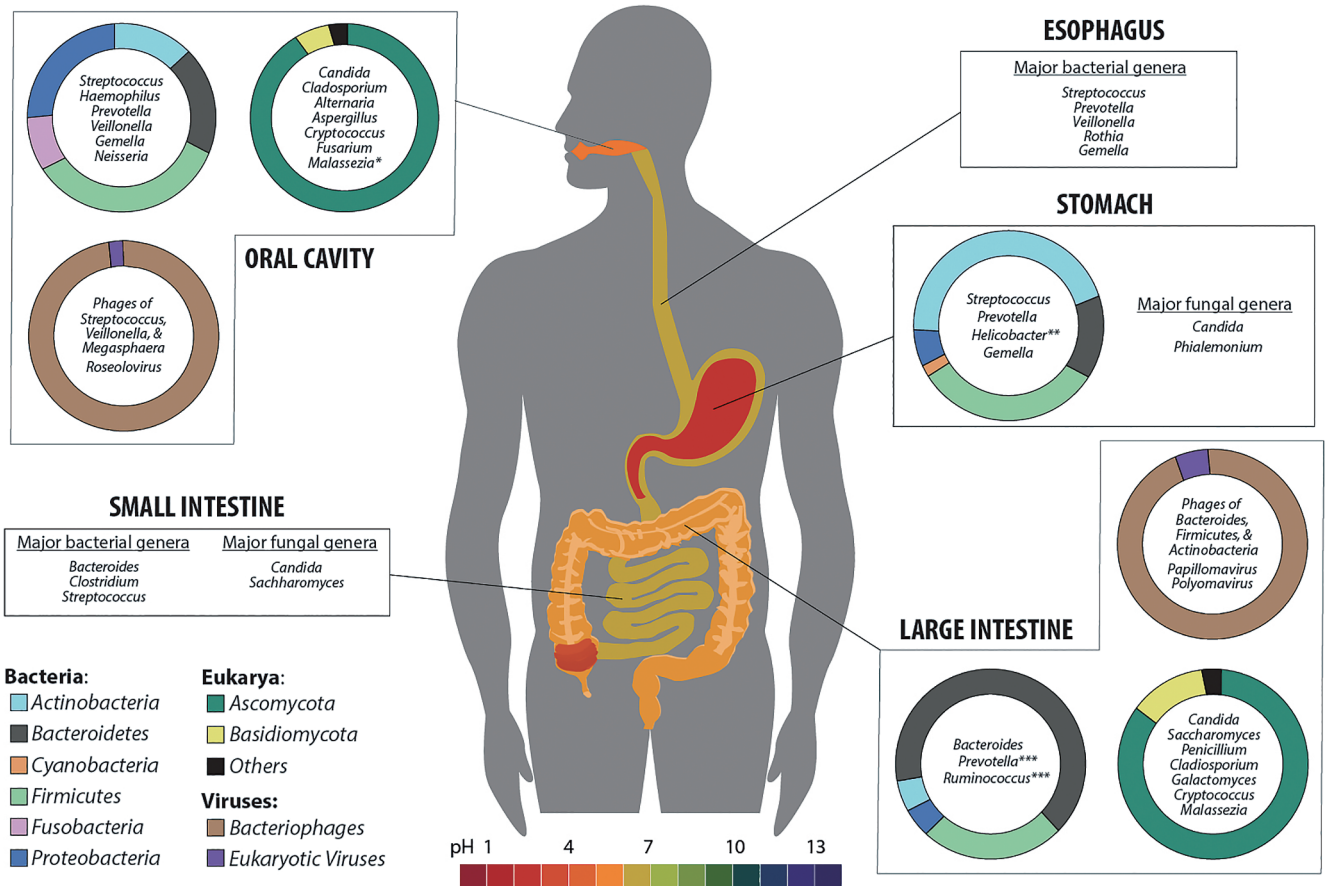


Fig. 1. Microbiome composition of *Bacteria* (1, 5, 20, 21, 43, 147, 156, 223), *Eukarya* (52, 85, 114, 126, 182, 197), and *Viruses* (45, 134, 151, 215) among the physiological niches of the human gastrointestinal (GI) tract. Phylum level compositional data are presented where available along with the 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 (2024)

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