



OPEN *Selenobaculum gbiensis* gen. nov. sp. nov., a new bacterium isolated from the gut microbiota of a patient with Crohn's disease

Soyoung Yeo^{1,2,5}, Hyunjoon Park^{2,5}, Heebal Kim¹, Chang Beom Ryu³ & Chul Sung Huh^{2,4}✉

The human gut microbiota is a complex ecology comprising approximately 10 to 100 trillion microbial cells. Most of the bacteria detected by 16s rRNA sequencing have yet to be cultured, but intensive attempts to isolate the novel bacteria have improved our knowledge of the gut microbiome composition and its roles within human host. In our culturomics study, a novel gram-negative, motile, obligately anaerobic, rod-shaped bacteria, designated as strain ICN-92133^T, was isolated from a fecal sample of a 26-year-old patient with Crohn's disease. Based on the 16s rRNA sequence of strain ICN-92133^T, the phylogeny analysis placed the strain into the family *Selenomonadaceae*, showing 93.91% similarity with the closely related *Massilibacillus massiliensis* strain DSM 102838^T. Strain ICN-92133^T exhibited a genome size of 2,679,003 bp with a GC content of 35.5% which was predicted to contain 26 potential virulence factors and five antimicrobial resistance genes. In comparative genomic analysis, strain ICN-92133^T showed digital DNA–DNA Hybridization and OrthoANI values lower than 21.9% and 71.9% with the closest type strains, respectively. In addition, comparing phenotypic, biochemical, and cellular fatty acids with those of closely related strains revealed the distinctiveness of strain ICN-92133^T. Based on the taxonogenomic results, strain ICN-92133^T is proposed as a novel species belonging to a new genus. Therefore, we suggest the name of the new genus *Selenobaculum* gen. nov. within the family *Selenomonadaceae* and strain ICN-92133^T (= KCTC 25622^T = JCM 36070^T) as a type strain of new species *Selenobaculum gbiensis* sp. nov.

The gut microbiota is linked to human health and disease by maintaining intestinal homeostasis and regulating metabolism^{1–3}. Since the first large-scale 16s rRNA sequencing in 2005 to identify intestinal bacteria, information on the diversity of gut ecology has increased exponentially^{4,5}. However, the culture of bacteria had been very biased toward partial phylogenetic groups, and the quantitative gap between the number of novel cultured species and the results obtained by the culture-independent method had widened^{6,7}. Microbial culturomics, introduced in 2012, is a research concept that combines high-throughput cultivating microorganisms with various culture conditions, a matrix-assisted laser desorption-ionization time-of-flight mass spectrometry (MALDI-TOF MS), and 16s rRNA gene sequencing⁸. The culturomics approach has enabled high-throughput microbial identification and cultivation, resulting in many previously unexplored bacterial lineages reported^{9–11}.

Nevertheless, a significant fraction of the bacteria detected in the 16s rRNA sequencing has yet to be cultured. The number is expected to be higher when considering the bacteria not detected in metagenomics for reasons such as low abundances and sequencing bias^{12,13}. Since 2021, we have conducted a culturomics study focusing on the intestine of healthy individuals and inflammatory bowel disease (IBD) patients to contribute to discovering novel gut bacteria. As part of the results, we isolated a novel bacterial strain ICN-92133^T (= KCTC 25622^T = JCM 36070^T) from a stool sample in a patient with Crohn's disease. Using the taxonogenomics, a new research concept that proposes the novelty of new species based on whole genome sequence along with polyphasic properties¹⁴, we present distinct properties of strain ICN-92133^T from its phylogenetic neighbors, thereby proposing strain

¹Department of Agricultural Biotechnology, College of Agriculture and Life Sciences, Seoul National University, Seoul 08826, South Korea. ²Research Institute of Eco-Friendly Livestock Science, Institute of Green-Bio Science and Technology, Seoul National University, Pyeongchang 25354, South Korea. ³Department of Internal Medicine, Digestive Disease Center and Research Institute, Soon Chun Hyang University School of Medicine, Bucheon 14584, South Korea. ⁴Graduate School of International Agricultural Technology, Seoul National University, Pyeongchang 25354, South Korea. ⁵These authors contributed equally: Soyoung Yeo and Hyunjoon Park. ✉email: chulsunghuh@gmail.com

ICN-92133^T as a type species *Selenobaculum gbiensis* gen. nov., sp. nov. of the new genus within the family *Selenomonadaceae*.

Materials and methods

Sample collection under ethical approval. Before the sample collection, this study was approved by the Institutional Review Board of Soon Chun Hyang University Bucheon Hospital in South Korea under the number SCHBC 2021-01-028-002. The protocol was carried out in accordance with relevant guidelines and regulations approved by our ethics committee. The donor was a 26-year-old male diagnosed with Crohn's disease who gave his informed and signed consent. The fecal samples were kept below 4 °C under anaerobic conditions immediately after defecation until processed in the laboratory within 24 h.

Strain isolation and identification by MALDI-TOF MS. The sample was homogenized, resuspended to 0.25 g/L with 0.85% sodium chloride saline, and immobilized on polysaccharide gel beads for long-term pre-incubation¹⁵. The stool gel beads were inoculated in an anaerobic culture bottle (BioMérieux, Marcy l'Etoile, France) supplemented with 0.2 µm filtered rumen fluid¹⁶ and defibrinated sheep blood and then incubated at 37 °C for 30 days. The cultured broth was sampled regularly and plated onto modified Gifu Anaerobic Medium (mGAM; Nissui Pharmaceutical, Tokyo, Japan). Bacterial colonies were isolated and initially identified using MALDI-TOF MS on a Biotyper Sirius system (Bruker Daltonics, Bremen, Germany)¹⁷. The process was conducted in an anaerobic chamber consisting of 5% CO₂, 10% H₂, and 85% N₂.

Whole genome sequencing and analysis. Genomic DNA of strain ICN-92133^T was extracted using Quick-DNA HMW MagBead kit (Zymo Research Corporation, Irvine, CA, USA) with lysozyme chloride (Sigma-Aldrich, St. Louis, MO, USA) and quantified by QuantiFluor dsDNA System method (Promega, Madison, WI, USA) using Victor Nivo Multimode Microplate Reader (PerkinElmer, Waltham, MA, USA). Whole genome sequencing was performed using a PacBio Sequel (Pacific Biosciences, Menlo Park, CA, USA) and a NovaSeq 6000 system (Illumina, San Diego, CA, USA) by Macrogen Corporation (Seoul, South Korea). PacBio long-read sequences from approximately 12 kb SMRTbell templates were assembled de novo to construct genome sequences. They were then corrected using Illumina sequencing data from paired-end (2 × 150 bp) sequencing to obtain high quality. In this study, Microbial Assembly Application was applied for assembly¹⁸, and the genome was annotated using the PATRIC RAST tool kit (RASTtk)¹⁹ in the BV-BRC server²⁰. The circular genome was constructed in a CGView tool²¹. The virulence factor database (VFDB)²² and the Comprehensive Antibiotic Resistance Database (CARD)²³ were used to predict potential virulence factors and antimicrobial resistance genes.

Phylogenetic analysis based on 16s rRNA gene sequences. The NCBI BLASTn search²⁴ and Eztaxon database²⁵ were used to classify the strain based on 16s rRNA gene sequence similarities with phylogenetically closest strains. The threshold of 94.5% 16s rRNA sequence identity, previously established^{26,27}, was applied to determine a new genus classification. A phylogenetic tree was constructed using the Maximum-likelihood method²⁸ based on the Tamura-Nei model²⁹ after alignment of the 16s rRNA sequences using MUSCLE³⁰ in the MEGA11 program³¹.

Morphological, phenotypical, and biochemical assays. Gram-stained morphology was observed using a phase contrast microscope (Eclipse Ci-L, Nikon, Tokyo, Japan). The outer and inner structures of strain ICN-92133^T were observed using the scanning electron microscope (SEM) and energy-filtered transmission electron microscope (EFTEM) acquired on a Regulus 8100 (Hitachi, Tokyo, Japan) and LIBRA 120 (Carl Zeiss, Germany), respectively. The main characteristics, such as oxidase, catalase, nitrate reduction, and sporulation activities, of strain ICN-92133^T and phylogenetically closest type strains were compared. The optimal growth conditions were determined by culturing at different temperatures (20, 26, 30, 37, and 45 °C), pH (5, 5.5, 6, 6.5, 7, 7.5, 8.5), and osmotic conditions (0, 1, 3, 5, 10, 15 and 20 g of NaCl/L). According to the manufacturer's instructions, enzymatic activities and substrate utilization properties were investigated using API ZYM, API 20A, and API 50CH kits (BioMérieux) at 30 °C.

Cellular fatty acids methyl ester analysis. Cellular fatty acids methyl esters (FAMES) were extracted from colonies grown on mGAM agar and analyzed by Agilent 8890 gas chromatography system (Agilent Technologies, Inc., Palo Alto, CA, USA) at Korean Collection for Type Cultures (KCTC; Jeonbuk, South Korea). The spectra of cellular fatty acids were determined using the Sherlock MIDI software version 6.5 based on the ANAER6 method³².

Minimum inhibitory concentration assay. The minimum inhibitory concentrations (MICs) of strain ICN-92133^T were determined in mGAM agar at 30 °C using ETEST (BioMérieux) of nine antibiotics: ampicillin, gentamicin, kanamycin, streptomycin, erythromycin, clindamycin, vancomycin, tetracycline, and chloramphenicol.

Genomic comparison. Whole genome-based phylogeny was inferred by cross-genus protein families (PGfams) Codon Tree pipeline³³ and Type (Strain) Genome Server (TYGS)³⁴. The genomic features of the comparing strains were obtained from the BV-BRC server²⁰. To calculate in silico genome-to-genome similarity,

digital DNA-DNA hybridization (dDDH)³⁵ and Orthologous Average Nucleotide Identity (OrthoANI)³⁶ were applied.

Results

Strain isolation, morphology, and identification of strain ICN-92133^T. A colony of strain ICN-92133^T was first isolated on mGAM agar plates after 21 days of pre-incubation. The morphology of the colonies was creamy white, convex, and circular forms less than 1 mm in diameter. The spectrum of strain ICN-92133^T did not match the bacterial spectra in the MALDI Biotyper database showing scores lower than 1.5 (Fig. 1). The strain was gram-negative rod-shaped with 0.4–0.5 × 1.6–3.2 μm in size. However, some cells had blunt ends sporadically, which were observed in surface observations (Fig. 2a). In addition, structures presumed to be granules were observed inside the cells (Fig. 2b). The 16s rRNA gene sequence of strain ICN-92133^T was 1557 bp in length. Strain ICN-92133^T exhibited sequence similarity with *M. massiliensis* strain DSM 102838^T (93.91%)³⁷, *Propionispira raffinovorans* strain DSM 20765^T (93.08%)³⁸, *Propionispira arcuata* strain KCTC 15499^T (92.73%)³⁹, *Propionispira arboris* strain DSM 2179^T (92.51%)⁴⁰, *Anaerobaculum glycerini* strain DSM 5192^T (92.54%)^{41,42}, and *Propionispira paucivorans* strain DSM 20756^T (91.85%)³⁸. Phylogenetic analysis based on 16s rRNA placed strain ICN-92133^T into a member of the family *Selenomonadaceae* (Fig. 3). Since the highest identity value is lower than the threshold of 95% for determining a new genus, strain ICN-92133^T showed the possibility of being classified as a new genus of a new species and required further characterization.

Comparison of phenotypic and biochemical characteristics. The condition from which strain ICN-92133^T was isolated was mGAM medium at 37 °C in an anaerobic atmosphere. Strain ICN-92133^T is motile with peritrichous flagella, but oxidase, nitrate reduction, and sporulation activities were not observed under the growth conditions (Table 1). The optimal growth conditions for strain ICN-92133^T were 26 to 37 °C, pH 6.5 to 8.5, and less than 0.5% salinity. To compare the main characteristics with strain ICN-92133^T, four strains belonging to the genera *Massilibacillus*, *Propionispira*, and *Anaerobaculum* were selected according to phylogeny results. Strain ICN-92133^T had enzymatic activities regarding pyroglutamyl aminopeptidase, alkaline phosphatase, esterase C4, esterase lipase C8, leucine arylamidase, acid phosphatase, and naphthol-AS-BI-phosphohydrolase.

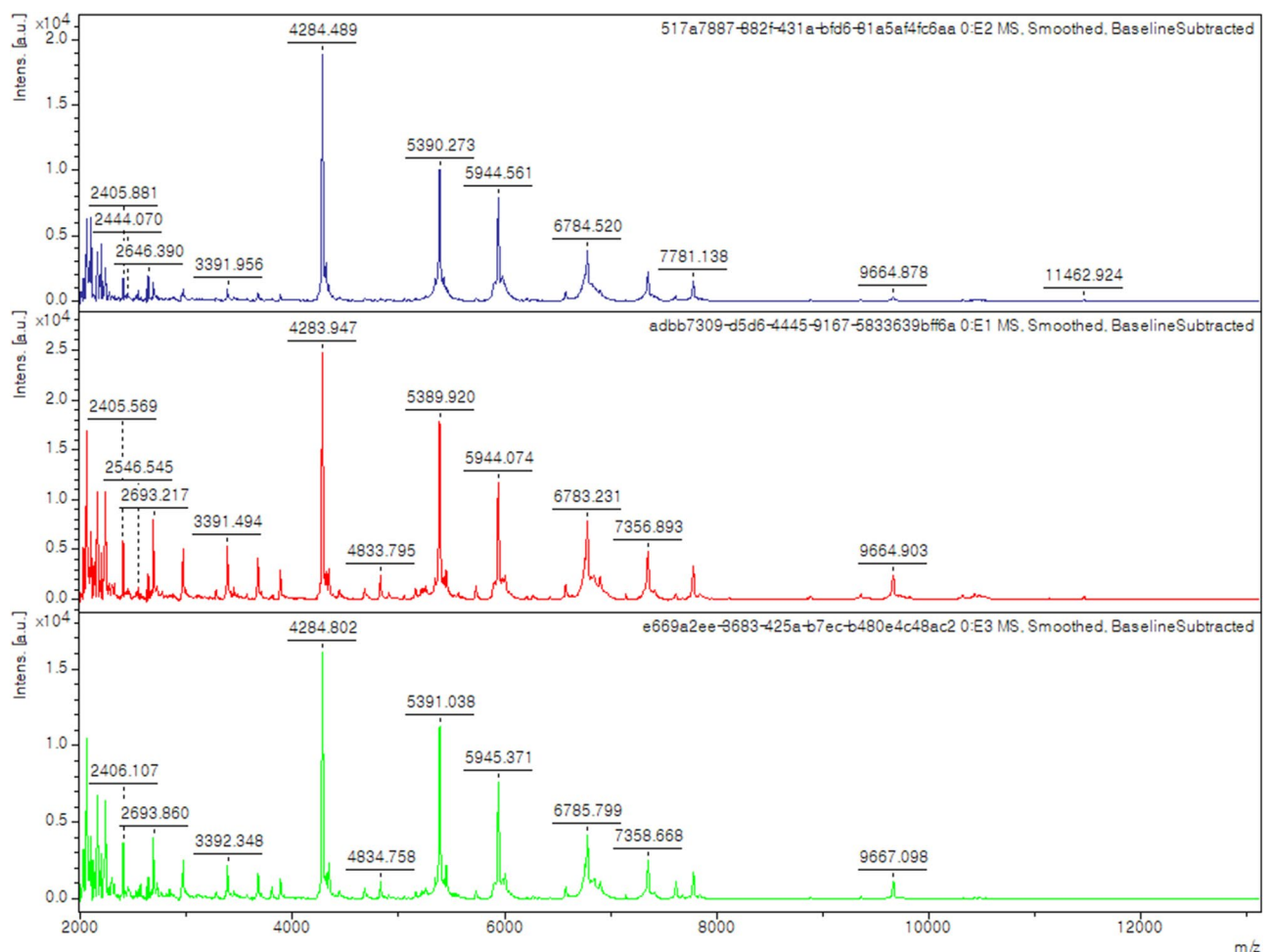


Figure 1. MALDI-TOF MS spectra of *Selenobaculum gbiensis* strain ICN-92133^T.

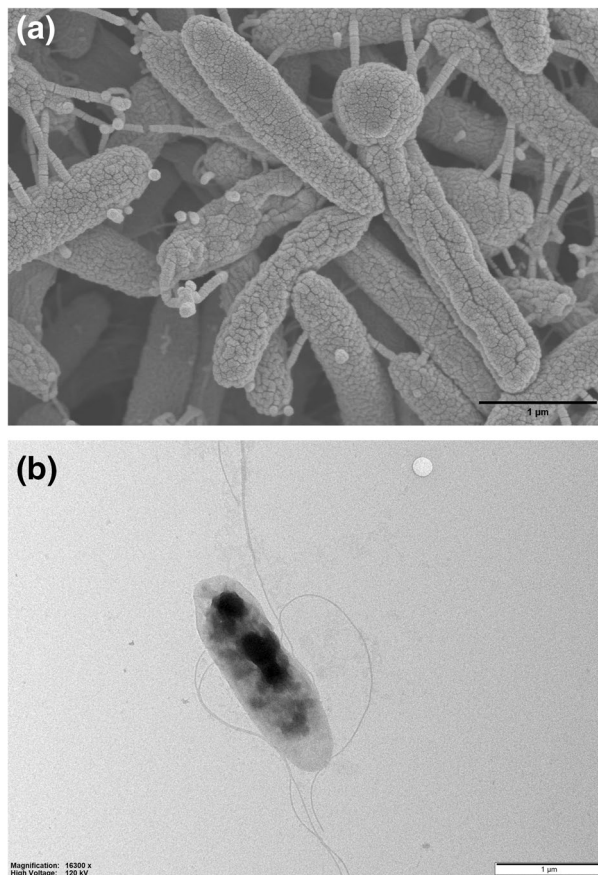


Figure 2. Morphology of *Selenobaculum gbiensis* strain ICN-92133^T. Images were obtained using (a) scanning electron microscope (SEM) and (b) energy-filtered transmission electron microscope (EFTEM). Scale bars of 1 µm were represented.

Moreover, the strain utilized glycerol, D-ribose, D-adonitol, D-glucose, D-fructose, D-mannose, and D-mannitol. Since *A. glycerini* strain DSM 5192^T utilizes glycerol as a primary carbon source, some reactions were not observed under the test conditions. In cellular fatty acids analysis, it is revealed that the primary cellular fatty acids of strain ICN-92133^T were C_{17:1,cis-9} (17.6%) and C_{18:1,cis-9} (19.5%) (Table 2).

Antibiotic susceptibility of strain ICN-92133^T. Strain ICN-92133^T showed MICs (µg/mL) of 0.094 for ampicillin, 4 for gentamicin, 4 for kanamycin, 6 for streptomycin, 6 for erythromycin, 0.032 for clindamycin, 0.19 for tetracycline, 0.5 for chloramphenicol, and more than 256 for vancomycin.

Genomic properties of strain ICN-92133^T. The complete genome of strain ICN-92133^T was a circular genome with a 2,679,003 bp size and a GC content of 35.5% (Fig. 4a). The genome contained 2526 protein-coding sequences (CDS), 86 tRNA genes, and 18 rRNA genes. The annotated proteins included 1016 hypothetical proteins and 1510 proteins assigned to 237 RAST subsystems (Fig. 4b). The clusters of orthologous group (COG) functional categories in which most of the genes involved were amino acids and derivatives (19.6%), protein metabolism (13.5%), cofactors, vitamins, prosthetic groups, pigments (13.4%), carbohydrates (8.3%) and DNA metabolism (6.4%). In addition, virulence factor candidates and antibiotic-resistance genes were predicted within the genome of strain ICN-92133^T to evaluate the potential risk to the host. A total of 38 sequences were matched to virulence genes in the VFDB database regarding adherence (*pilA*, *tcpl*, *groEL*, *tufA*, and *htpB*), exotoxin (*hlyB*), chemotaxis (*acfB*), immune modulation (*cpsB/cdsA*, *pseC*, *rfaD*, *gmd*, and *manC*), motility (*flhA*, *tlpB*, *gleQ*, and *pseB*), stress survival (*clpE* and *clpC*), and Type III secretion system (T3SS; *ysaV*, *escV*, and *invA*) (Table 3). In addition, five genes related to antibiotic efflux pump (*adeF* and *qacJ*), antibiotic target alteration (*vanT* and *vanW*), and inactivation of antibiotics (*fos^{XCC}*) were detected in CARD.

Comparative genomic characteristics. The whole genome of strain ICN-92133^T was compared with the available genomes of seven strains which were phylogenetic neighbors (Table 4). The genome of strain ICN-92133^T had the third shortest genome length and the lowest GC content and CDS numbers. In addition, strain

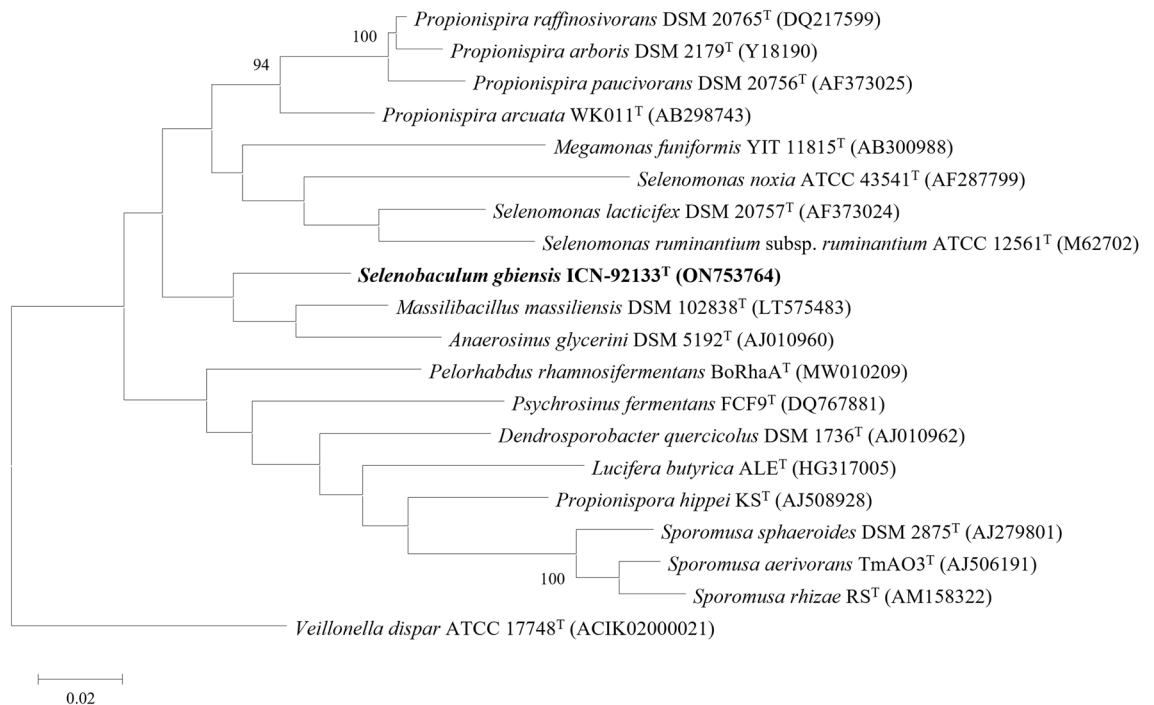


Figure 3. Phylogeny of *Selenobaculum gbiensis* strain ICN-92133^T and closely related species. The 16s rRNA sequences were aligned using MUSCLE with default parameters, and the phylogenetic tree was calculated using the maximum-likelihood method within MEGA software version 11.0. GenBank sequence accession numbers were described in parentheses. Among the bootstrap values obtained from 500 replications, values higher than 90% were indicated at branch points. *Veillonella dispar* was used as outgroup.

ICN-92133^T showed a G+C content difference of 2.23 to 17.21% from the type strains excluding *Pectinatus brassicae* strain DSM 24661^T. The subsystems annotated using the RASTtk are described in Table 5. The phylogenetic tree-based whole genomes constructed in PGfams Codon Tree pipeline showed that strain ICN-92133^T formed a clade with *M. massiliensis* strain DSM 102838^T and *Propionispira* type strains (Fig. 5a). Strain ICN-92133^T exhibited dDDH values lower than 21.9% and OrthoANI values ranging from 71.9% with *M. massiliensis* strain DSM 102838^T to 64.6% with *S. montiformis* strain DSM 106892^T (Fig. 5b). The dDDH and OrthoANI values were lower than the thresholds (70% and 95%, respectively), indicating that strain ICN-92133^T represented a novel species^{43,44}. Taken together, we propose strain ICN-92133^T as a type species of a new genus within the family *Selenomonadaceae*.

Discussion

Since gut microbiome and its association with host health and disease have been focused on, intensive efforts are being made to cultivate the gut microorganisms to elucidate microbe-microbe and microbe-host interactions^{9,45–47}. The culturomics approach can contribute to acquiring diverse gut microbes, including new species that have never been cultured^{8–12}. We obtained novel strains from the gut microbiota of healthy individuals and IBD patients during the culturomics study. Among them, the characteristics of strain ICN-92133^T are presented here with the taxonogenomic strategy.

The family *Selenomonadaceae*, which has the characteristics of strictly anaerobic and gram-negative bacilli, was first classified as a new family within the order *Selenomonadales* in the class *Negativicutes* in 2015⁴⁸. The family currently contains 43 species in 8 genera, including *Anaerovibrio*, *Centipeda*, *Megamonas*, *Mitsuokella*, *Pectinatus*, *Propionispira*, *Schwartzia*, and *Selenomonas*. The phylogenetic analysis based on 16s rRNA gene sequences placed strain ICN-92133^T with species belonging to the family *Selenomonadaceae*, *Veillonellaceae*, and *Sporomusaceae* (Fig. 3). Phylogeny based on both amino acids and nucleotide sequences in the Codon Tree pipeline was inferred with currently available genomes within the three families to confirm its taxonomic classification (see Supplementary Fig. S1). Although strain ICN-92133^T not only showed the highest sequences similarity of 93.91% in 16s rRNA with *M. massiliensis* strain Marseille-P2411^T (= DSM 102838^T), which belongs to family *Veillonellaceae*, but also constituted a clade in whole genome-based phylogeny, it shared a common ancestor with the other species of the family *Selenomonadaceae*. Therefore, we classified strain ICN-92133^T as a new genus belonging to the family *Selenomonadaceae*.

The complete genome of strain ICN-92133^T was smaller than those of *M. massiliensis* DSM 102838^T and *P. arboris* DSM 2179^T, which could be the basis for the relatively fewer ranges of enzymatic and fermentation activities (Tables 1, 4). In electron microscopic observations of strain ICN-92133^T, we found granular cytoplasmic

Properties	1	2	3	4	5
Isolation source	Stool from a patient with Crohn's disease	Stool from a healthy infant	Methanogen reactor	Alkaline wet woods of poplar trees	Black freshwater sediment
Cell morphology	Rods with a rounded end	Rods with a rounded end	Curved rods	Curved rods with a rounded end	Curved rods
Cell size (µm)	0.4–0.5 × 1.6–3.2	0.4–0.6 × 1.7–3.5	0.3–0.5 × 1.3–2.7	0.6–0.9 × 3.1–7.2	0.3–0.6 × 2.3–4.1
Gram-staining	–	– (+)	–	–	–
Catalase activity	–	–	–	–	–
Oxidase activity	–	–	–	–	–
Nitrate reduction activity	–	–	–	–	–
Sporulation ability	–	–	–	–	–
Optimal growth conditions					
Temperature (20–45 °C)	26–37	26–37	20–30	20–30	26–37
pH (5.0–8.5)	6.5–8.5	6.5–8.5	6.5–7.5	6.0–7.0	6.5–8.5
NaCl (0–20 g/L)	< 5 g/L	< 15 g/L	< 3 g/L	< 3 g/L	< 3 g/L
Enzymatic activity					
Pyroglutamyl aminopeptidase	+	+	+	+	+
Alkaline phosphatase	+	+	+	+	–
Esterase (C4)	+	+	+	+	+
Esterase lipase (C8)	+	–	+	–	–
Leucine arylamidase	+	+	+	+	–
Acid phosphatase	+	+	+	+	+
Naphtol-AS-BI-phosphohydrolase	+	+	+	+	+
α-Galactosidase	–	–	–	+	–
β-Galactosidase	–	–	–	+	–
α-Glucosidase	–	–	–	+	–
β-Glucosidase	–	+	–	+	–
N-Acetyl-β-glucosaminidase	–	–	+	+	–
Substrate utilization					
Glycerol	+	–	+	+	+
Erythritol	–	–	+	+	NA
D-Arabinose	–	+	–	+	NA (–)
L-Arabinose	–	+	–	+	NA (–)
D-Ribose	+	+	+	+	NA (–)
Properties	1	2	3	4	5
D-Xylose	–	+	–	+	NA (–)
D-Adonitol	+	+	+	–	NA
D-Galactose	–	–	+	+	NA
D-Glucose	+	+	+	+	NA (–)
D-Fructose	+	+	+	+	NA
D-Mannose	+	+	+	+	NA (–)
L-Sorbose	–	–	–	+	NA (–)
L-Rhamnose	–	–	–	+	NA
Dulcitol	–	+	–	+	NA
Inositol	–	+	–	+	NA
D-Mannitol	+	+	+	+	NA (–)
D-Sorbitol	–	–	–	+	NA
Methyl-α-D-Glucopyranoside	–	–	–	+	NA
N-Acetyl-glucosamine	–	+	–	+	NA
Amygdalin	–	+	–	+	NA
Arbutin	–	+	–	+	NA
Esculin, ferric citrate	–	+	–	+	NA
Salicin	–	+	–	+	NA (–)
D-Cellobiose	–	+	–	+	NA
D-Maltose	–	–	–	+	NA (–)
D-Lactose (bovine origin)	–	–	–	+	NA (–)
D-Melibiose	–	–	–	+	NA (–)
Continued					

Properties	1	2	3	4	5
D-Saccharose (sucrose)	–	–	+	+	NA (–)
D-Trehalose	–	–	–	+	NA (–)
D-Raffinose	–	–	–	+	NA (–)
Amidon (starch)	–	–	–	+	NA
Gentiobiose	–	+	–	+	NA
D-Turanose	–	–	–	+	NA
D-Tagatose	–	+	+	+	NA
L-fucose	–	–	–	+	NA
D-Arabitol	–	+	–	+	NA
L-Arabitol	–	+	+	–	NA
Potassium gluconate	–	+	–	+	NA
Potassium 5-ketogluconate	–	+	–	–	NA

Table 1. Biochemical characteristics of strain ICN-92133^T and its phylogenomically related type strains. Strains: 1, ICN-92133^T; 2, *Massilibacillus massiliensis* DSM 102838^T; 3, *Propionispira arcuate* KCTC 15499^T; 4, *Propionispira arboris* DSM 2179^T; 5, *Anaerosinus glycerini* DSM 5192^T. Data different from the original paper are indicated in parentheses. NA not available.

Cellular fatty acids	1	2	3	4	5
C9:0	2.0	–	0.7	–	0.8
C10:0 iso	0.6	–	–	–	2.1
C10:0	2.4	–	1.1	–	3.4
C11:0	6.1	7.9	7.5	8.6	1.9
C13:0	0.3	–	–	2.6	–
C12:0 3-OH	3.9	3.6	0.7	–	8.3
C14:0 DMA	9.4	13.2	10.0	9.4	3.0
C15:1 Cis 7	8.3	10.8	14.7	27.3	3.3
C15:0	8.6	8.8	12.5	12.6	1.8
C15:0 DMA	–	–	1.4	3.3	–
C16:1 Cis 7	1.6	–	2.8	2.6	2.1
C16:0	3.3	3.6	2.3	1.3	6.0
C17:1 Cis 8	–	–	–	3.3	–
C17:1 Cis 9	17.6	18.5	23.3	5.7	2.0
C17:0	8.1	3.1	5.1	1.7	0.5
UN 17.223	–	–	1.2	3.6	–
C18:1 AT 17.254 DMA	–	–	4.3	5.1	–
C18:2 Cis 9,12	0.7	–	–	–	2.3
C18:1 Cis 9	19.5	30.7	10.3	9.5	53.6
C18:1c11/t9/t6	1.0	–	–	–	2.7

Table 2. Cellular fatty acids compositions (%). Strains: 1, ICN-92133^T; 2, *Massilibacillus massiliensis* DSM 102838^T; 3, *Propionispira arcuate* KCTC 15499^T; 4, *Propionispira arboris* DSM 2179^T; 5, *Anaerosinus glycerini* DSM 5192^T. The values above 2% in at least one strain are indicated. –, not detected.

inclusions in strain ICN-92133^T (Fig. 2b). Although the contained substances have not been verified in this study, they can be assumed to be a kind of energy storage based on the results of studies that *P. arboris* DSM 2179^T has carbohydrate-containing granules⁴⁰. In addition, we investigated potential virulence and antibiotic

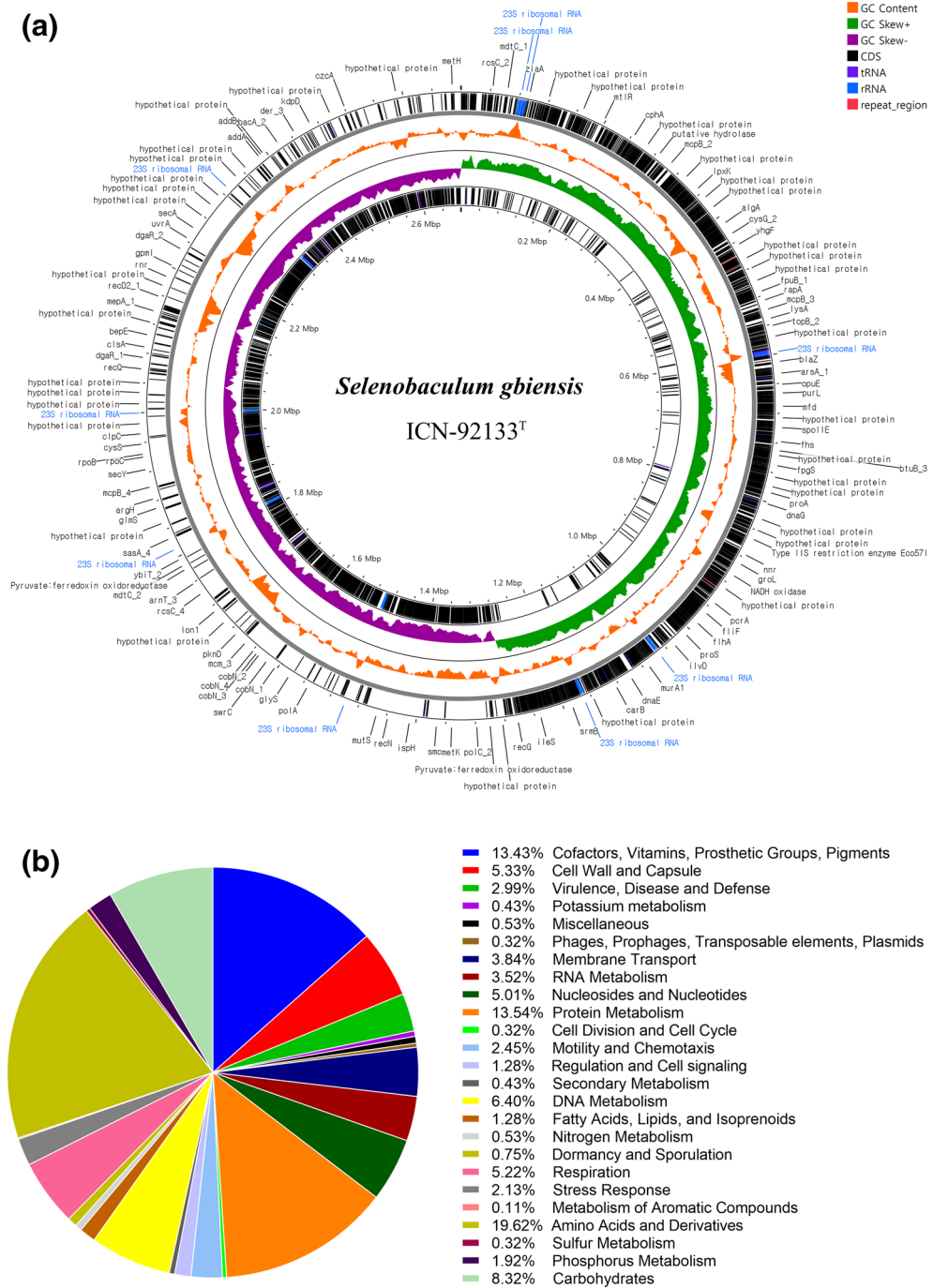


Figure 4. Complete genome and RAST subsystem category distributions of *Selenobaculum gbiensis* strain ICN-92133^T. **(a)** A circular chromosome map represents annotated proteins, GC contents, forward GC skew, reverse GC skew, CDS, tRNA, rRNA, and repeat regions from outside to the center. **(b)** COG categories distributions were analyzed in the RAST server.

Gene	Protein	Identity%	Database ID
Virulence factor database (VFDB)			
<i>tufA</i>	Elongation factor Tu	80.3/81.9/88.7 ^a	VFG046465
<i>htpB</i>	Hsp60, 60 K heat shock protein HtpB	82.2/90.6/92.3 ^a	VFG001855
<i>groEL</i>	Chaperonin GroEL	83.7/86.4 ^b	VFG012095
<i>clpC</i>	Endopeptidase Clp ATP-binding chain C	88.1/90.4/95.7 ^a	VFG000079
<i>escV</i>	Type III secretion system major export apparatus protein	84.9	VFG033545
<i>escV</i>	Type III secretion system major export apparatus protein	84.9	VFG000813
<i>clpE</i>	ATP-dependent protease	86.5/89.1/94.3 ^a	VFG000080
<i>pseC</i>	UDP-4-amino-4, 6-dideoxy-N-acetyl-beta-L-altrosamine transaminase	83.7	VFG037927
<i>invA</i>	Type III secretion system major export apparatus protein InvA	82.7	VFG000557
<i>hlyB</i>	Hemolysin B	88.3/89.1 ^b	VFG000841
<i>pseB</i>	UDP-GlcNAc-specific C4,6 dehydratase/C5 epimerase	84.5/88.5 ^b	VFG012002
<i>gmd</i>	GDP-mannose 4,6-dehydratase	83.9	VFG002225
<i>fleQ</i>	Transcriptional regulator FleQ	94.74	VFG043319
<i>gmd</i>	GDP-mannose 4,6-dehydratase	79.75	VFG048885
<i>tlpB</i>	Membrane-bound chemoreceptor sensing pH and autoinducer-2	83.95	VFG043367
<i>pseB</i>	UDP-N-acetylglucosamine 4,6-dehydratase	92.68	VFG037912
<i>pilA</i>	Type IV major pilin protein PilA	94.44	VFG006896
<i>acfB</i>	Accessory colonization factor AcfB	83.75	VFG000104
<i>gmd</i>	GDP-mannose 4,6-dehydratase	79.35	VFG002365
<i>flhA</i>	Flagellar biosynthesis protein	90.7	VFG006577
<i>rfaD</i>	ADP-L-glycero-D-mannoheptose-6-epimerase	89.36	VFG000332
<i>ysaV</i>	EscV/YscV/HrcV family type III secretion system export apparatus	81.37	VFG003525
<i>tcpI</i>	Negative regulator of the major pilin TcpA	85.48	VFG000088
<i>manC</i>	Mannose-1-phosphate guanylyltransferase	92.11	VFG047006
<i>hlyB</i>	Hemolysin B	87.0/92.1 ^b	VFG000907
<i>cpsB/cdsA</i>	Phosphatidate cytidylyltransferase	92.1	VFG002189
Comprehensive antibiotic resistance database (CARD)			
<i>adeF</i>	Resistance-nodulation-cell division (RND) antibiotic efflux pump	42.7	NA
<i>qacJ</i>	Small multidrug resistance (SMR) antibiotic efflux pump	47.6	NA
<i>vanT</i>	Glycopeptide resistance gene cluster	37.7	NA
<i>vanW</i>	Vancomycin B-type resistance protein VanW	28.1	NA
<i>fos^{XCC}</i>	Fosfomycin thiol transferase	62.4	NA

Table 3. Prediction of virulence factors and antimicrobial resistance genes in the genome of *Selenobaculum gbiensis* strain ICN-92133^T based on VFDB and CARD databases, respectively. NA not available. ^aThree sequences were detected in the VFDB database. Each percentage of identity is represented. ^bTwo sequences were detected in the VFDB database. Each percentage of identity is represented.

Strains	GenBank accession number	Contigs	Genome size (bp)	CDS	G+C (mol, %)	Diff. G+C (%) ^a
<i>Selenobaculum gbiensis</i> ICN-92133 ^T	CP120678	1	2,679,003	2526	35	0
<i>Massilibacillus massiliensis</i> DSM 102838 ^T	FLKF000000000	7	4,000,953	3876	38	2.23
<i>Propionispira raffinosivorans</i> DSM 20765 ^T	ARLE000000000	66	4,131,539	3945	38	2.65
<i>Propionispira arboris</i> DSM 2179 ^T	FNZK000000000	40	4,261,607	4164	38	2.53
<i>Schwartzia succinivorans</i> DSM 10502 ^T	FQUG000000000	21	2,659,575	2556	47	11.35
<i>Selenomonas ruminantium</i> subsp. <i>lactilytica</i> DSM 2872 ^T	FNQG000000000	22	3,039,002	2950	50	14.06
<i>Selenomonas montiformis</i> DSM 106892 ^T	VUNL000000000	26	2,640,693	2686	53	17.21
<i>Pectinatus brassicae</i> DSM 24661 ^T	JACHFH000000000	90	2,849,197	2824	36	0.46

Table 4. General genomic features of *Selenobaculum gbiensis* ICN-92133^T and related species. Strains were selected based on TYGS analysis, and the genomic data was retrieved from the BV-BRC server. ^aThe G+C content difference (%) from the sequence of *Selenobaculum gbiensis* strain ICN-92133^T.

Subsystems	ICN-92133 ^T	DSM 102838 ^T	DSM 20765 ^T	DSM 2179 ^T	DSM 10502 ^T	DSM 2872 ^T	DSM 106892 ^T	DSM 24661 ^T
Metabolism	378	515	1056	1103	414	786	419	407
Protein processing	149	210	414	420	199	402	200	195
Energy	79	164	328	329	62	186	90	94
Cellular processes	88	85	148	160	61	150	80	68
DNA processing	57	78	128	155	60	124	76	64
Stress response, defense, virulence	81	66	103	100	43	90	60	57
RNA processing	51	47	93	93	40	89	48	43
Membrane transport	18	26	77	89	21	86	25	22
Cell envelope	21	25	50	47	14	46	17	20
Miscellaneous	4	6	14	12	6	10	6	7
Regulation and cell signaling	4	4	12	12	3	8	3	4

Table 5. Comparison of subsystems based on PATRIC annotation results. Strains: ICN-92133^T, *Selenobaculum gbiensis*; DSM 102838^T, *Massilibacillus massiliensis*; DSM 20765^T, *Propionispira raffinovorans*; DSM 2179^T, *Propionispira arboris*; DSM 10502^T, *Schwartzia succinivorans*; DSM 2872^T, *Selenomonas ruminantium* subsp. *lactilytica*; DSM 106892^T, *Selenomonas montiformis*; DSM 24661^T, *Pectinatus brassicae*.

resistance genes within the complete genome of strain ICN-92133^T to predict its pathogenicity because the strain was isolated from fecal specimens of a patient with Crohn's disease (Table 4). Based on the strain ICN-92133^T results showing high MIC values for vancomycin in vitro, the resistance might be derived from efflux pumps and reduced binding affinity to vancomycin. Our results also included several housekeeping proteins in the VFDB results, such as EF-Tu and GLoEL, which may play a role in the moonlighting functions in certain species^{49,50}. However, it is necessary to confirm whether the detected genes in strain ICN-92133^T are expressed in the host and show pathogenicity through further research.

In conclusion, we propose strain ICN-92133^T obtained through culturomics as a type species *Selenobaculum gbiensis* sp. nov., of a new genus *Selenobaculum* within the family *Selenomonadaceae*, according to the taxonomic results, including MALDI-TOF MS spectra, morphology, phenotypic properties, phylogenetic analysis, FAME composition, ANI, and dDDH calculation.

Description of *Selenobaculum* gen. nov.. *Selenobaculum* (Se.le.no.ba.cu.lum. N.L. fem. pl. n. *Selenomonadaceae*, a bacterial family name; L. neut. n. *baculum*, a stick, staff, rod; N.L. neut. n. *Selenobaculum*, a rod-shaped bacterium that belongs to the family *Selenomonadaceae*). Cells are gram-negative rods with rounded ends, obligately anaerobic, motile with flagellum, oxidase, and catalase negative. The type species is *Selenobaculum gbiensis*.

Description of *Selenobaculum gbiensis* gen. nov., sp. nov.. *Selenobaculum gbiensis* (g.bi.en.sis. N.L. masc. adj. *gbiensis* for the acronym of Green-Bio Institute, where the type strain was isolated). Cells grow at 26–37 °C, pH 6.5–8.5, a salinity of less than 0.5%, and the size was 0.4–0.5 × 1.6–3.2 μm. After 2 days of incubation at 37 °C in anaerobic conditions, colonies that were creamy and circular forms less than 1 mm in diameter were observed on mGAM agar. Among the 19 enzymatic activities tested in the API ZYM system, a positive reaction was observed for pyroglutamyl aminopeptidase, alkaline phosphatase, esterase (C4), esterase lipase (C8), leucine arylamidase, acid phosphatase, and naphthol-AS-BI-phosphohydrolase. In API 20A and API 50CH tests, it was observed to be utilized glycerol, D-ribose, D-adonitol, D-glucose, D-fructose, D-mannose, and D-mannitol as carbon sources. In contrast, utilization for other substrates was not observed. The primary cellular fatty acids were C_{17:1,cis-9} and C_{18:1,cis-9}. The type strain ICN-92133^T was first isolated from a patient's stool with Crohn's disease, and its genome was found to be 2,679,003 base-pair lengths and 35.5 mol% of GC contents. Strain ICN-92133^T

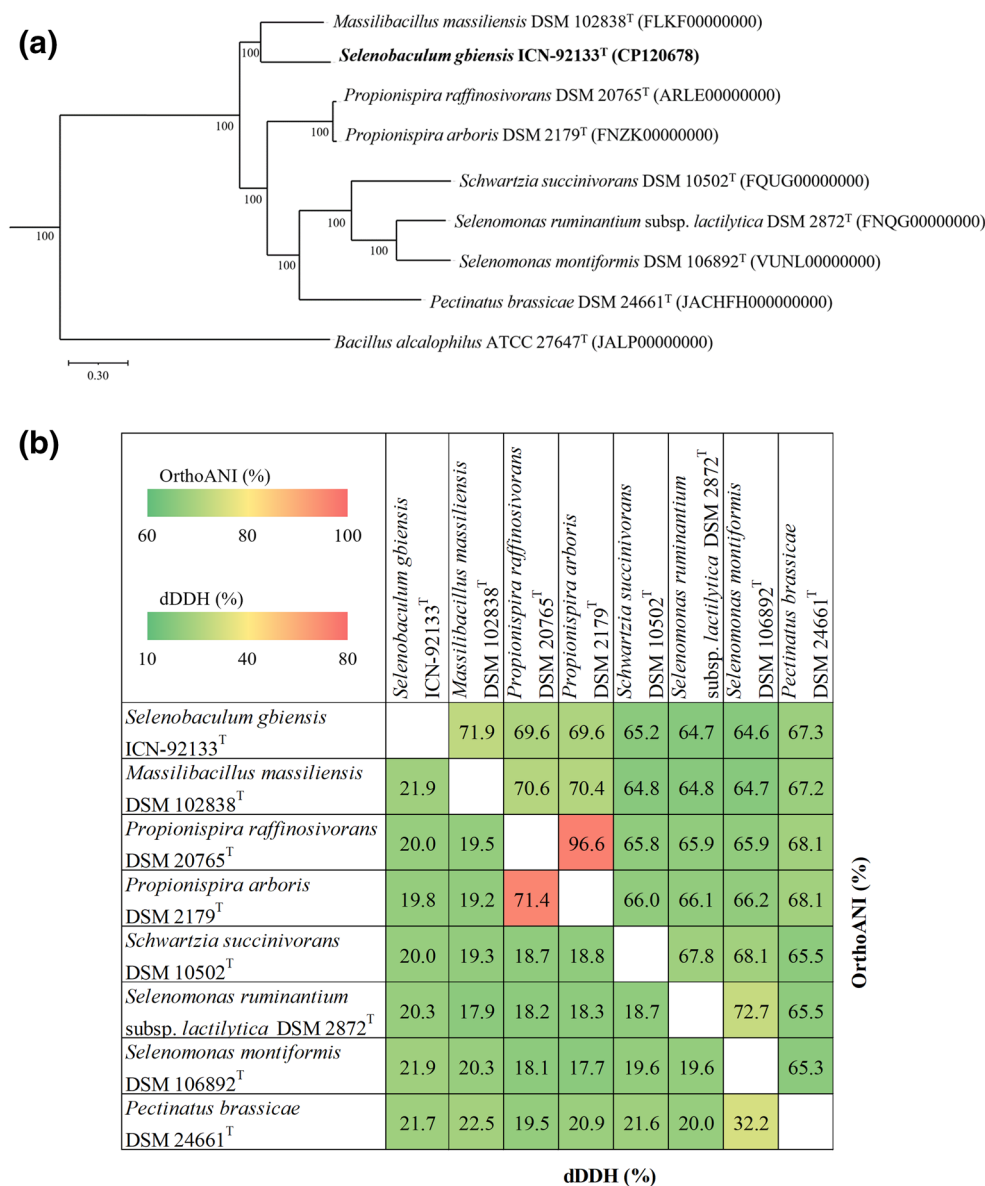


Figure 5. Whole genome-based phylogeny and pairwise comparison. **(a)** Phylogenetic tree based on Codon Tree pipeline in BV-BRC PGFams. GenBank sequence accession numbers were described in parentheses. *Bacillus alcalophilus* was used as outgroup. **(b)** Comparison of ANI and dDDH values between *Selenobaculum gbiensis* strain ICN-92133^T and closely related species. ANI values were obtained from OAT software, and dDDH was calculated in a formula 2 equation (DDH estimates based on identities/high-scoring segment pair length) in GGDC.

has been deposited at Korean Collection for Type Cultures (KCTC) and the Japan Collection of Microorganisms (JCM) under the numbers KCTC 25622^T and JCM 36070^T, respectively.

Data availability

The 16s rRNA gene sequence and complete genome sequence for strain ICN-92133^T have been deposited under GenBank Accession No. ON753764 and CP120678 (BioProject accession no. PRJNA944274, BioSample accession no. SAMN33746654, and SRA accession no. SRR23852393–SRR23852394), respectively.

Received: 8 June 2023; Accepted: 4 September 2023

Published online: 08 September 2023

References

1. Wu, H. J. & Wu, E. The role of gut microbiota in immune homeostasis and autoimmunity. *Gut Microbes* 3, 4–14. <https://doi.org/10.4161/gmic.19320> (2012).

2. Zheng, D., Liwinski, T. & Elinav, E. Interaction between microbiota and immunity in health and disease. *Cell Res.* **30**, 492–506. <https://doi.org/10.1038/s41422-020-0332-7> (2020).
3. Fan, Y. & Pedersen, O. Gut microbiota in human metabolic health and disease. *Nat. Rev. Microbiol.* **19**, 55–71. <https://doi.org/10.1038/s41579-020-0433-9> (2021).
4. Eckburg, P. B. *et al.* Diversity of the human intestinal microbial flora. *Science* **308**, 1635–1638. <https://doi.org/10.1126/science.1110591> (2005).
5. Oulas, A. *et al.* Metagenomics: Tools and insights for analyzing next-generation sequencing data derived from biodiversity studies. *Bioinform. Biol. Insights* **9**, 75–88. <https://doi.org/10.4137/BBI.S12462> (2015).
6. Overmann, J., Abt, B. & Sikorski, J. Present and future of culturing bacteria. *Annu. Rev. Microbiol.* **71**, 711–730 (2017).
7. Vrancken, G., Gregory, A. C., Huys, G. R. B., Faust, K. & Raes, J. Synthetic ecology of the human gut microbiota. *Nat. Rev. Microbiol.* **17**, 754–763. <https://doi.org/10.1038/s41579-019-0264-8> (2019).
8. Lagier, J. C. *et al.* Microbial culturomics: Paradigm shift in the human gut microbiome study. *Clin. Microbiol. Infect.* **18**, 1185–1193. <https://doi.org/10.1111/1469-0691.12023> (2012).
9. Lagier, J. C. *et al.* Culture of previously uncultured members of the human gut microbiota by culturomics. *Nat. Microbiol.* **1**, 16203. <https://doi.org/10.1038/nmicrobiol.2016.203> (2016).
10. Diakite, A., Dubourg, G. & Raoult, D. Updating the repertoire of cultured bacteria from the human being. *Microb. Pathog.* **150**, 104698. <https://doi.org/10.1016/j.micpath.2020.104698> (2021).
11. Bilen, M. *et al.* The contribution of culturomics to the repertoire of isolated human bacterial and archaeal species. *Microbiome* **6**, 94. <https://doi.org/10.1186/s40168-018-0485-5> (2018).
12. Lagier, J. C. *et al.* Culturing the human microbiota and culturomics. *Nat. Rev. Microbiol.* **16**, 540–550. <https://doi.org/10.1038/s41579-018-0041-0> (2018).
13. Bellali, S. *et al.* Running after ghosts: Are dead bacteria the dark matter of the human gut microbiota? *Gut Microbes* **13**, 1–12. <https://doi.org/10.1080/19490976.2021.1897208> (2021).
14. Lagier, J. C. *et al.* The rebirth of culture in microbiology through the example of culturomics to study human gut microbiota. *Clin. Microbiol. Rev.* **28**, 237–264. <https://doi.org/10.1128/CMR.00014-14> (2015).
15. Tanner, S. A. *et al.* In vitro continuous fermentation model (PolyFermS) of the swine proximal colon for simultaneous testing on the same gut microbiota. *PLoS ONE* **9**, e94123. <https://doi.org/10.1371/journal.pone.0094123> (2014).
16. Goodman, A. L. *et al.* Extensive personal human gut microbiota culture collections characterized and manipulated in gnotobiotic mice. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 6252–6257. <https://doi.org/10.1073/pnas.1102938108> (2011).
17. Singhal, N., Kumar, M., Kanaujia, P. K. & Virdi, J. S. MALDI-TOF mass spectrometry: An emerging technology for microbial identification and diagnosis. *Front. Microbiol.* **6**, 791. <https://doi.org/10.3389/fmicb.2015.00791> (2015).
18. *Pacific Biosciences, SMRT Tools.* <https://www.pacb.com/wp-content/uploads/SMRT-Link-User-Guide-v8.0.pdf>.
19. Brettin, T. *et al.* RASTtk: A modular and extensible implementation of the RAST algorithm for building custom annotation pipelines and annotating batches of genomes. *Sci. Rep.* **5**, 8365. <https://doi.org/10.1038/srep08365> (2015).
20. Olson, R. D. *et al.* Introducing the bacterial and viral bioinformatics resource center (BV-BRC): A resource combining PATRIC, IRD and ViPR. *Nucleic Acids Res.* **51**, D678–D689. <https://doi.org/10.1093/nar/gkac1003> (2023).
21. Stothard, P., Grant, J. R. & Van Domselaar, G. Visualizing and comparing circular genomes using the CGView family of tools. *Brief. Bioinform.* **20**, 1576–1582. <https://doi.org/10.1093/bib/bbx081> (2019).
22. Liu, B., Zheng, D., Zhou, S., Chen, L. & Yang, J. VFDB 2022: A general classification scheme for bacterial virulence factors. *Nucleic Acids Res.* **50**, D912–D917. <https://doi.org/10.1093/nar/gkab1107> (2022).
23. Alcock, B. P. *et al.* CARD 2023: Expanded curation, support for machine learning, and resistome prediction at the comprehensive antibiotic resistance database. *Nucleic Acids Res.* **51**, D690–D699. <https://doi.org/10.1093/nar/gkac920> (2023).
24. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410. [https://doi.org/10.1016/S0022-2836\(05\)80360-2](https://doi.org/10.1016/S0022-2836(05)80360-2) (1990).
25. Yoon, S. H. *et al.* Introducing EzBioCloud: A taxonomically united database of 16S rRNA gene sequences and whole-genome assemblies. *Int. J. Syst. Evol. Microbiol.* **67**, 1613–1617. <https://doi.org/10.1099/ijsem.0.001755> (2017).
26. Stackebrandt, E. & Ebers, J. Taxonomic parameters revisited: Tarnished gold standards. *Microbiol. Today* **8**, 6–9 (2006).
27. Yarza, P. *et al.* Uniting the classification of cultured and uncultured bacteria and archaea using 16S rRNA gene sequences. *Nat. Rev. Microbiol.* **12**, 635–645. <https://doi.org/10.1038/nrmicro3330> (2014).
28. Yang, Z. Phylogenetic analysis using parsimony and likelihood methods. *J. Mol. Evol.* **42**, 294–307. <https://doi.org/10.1007/BF02198856> (1996).
29. Tamura, K. & Nei, M. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Mol. Biol. Evol.* **10**, 512–526. <https://doi.org/10.1093/oxfordjournals.molbev.a040023> (1993).
30. Edgar, R. C. MUSCLE: A multiple sequence alignment method with reduced time and space complexity. *BMC Bioinform.* **5**, 113. <https://doi.org/10.1186/1471-2105-5-113> (2004).
31. Tamura, K., Stecher, G. & Kumar, S. MEGA11: Molecular evolutionary genetics analysis version 11. *Mol. Biol. Evol.* **38**, 3022–3027. <https://doi.org/10.1093/molbev/msab120> (2021).
32. Sasser, M. C. Identification of bacteria by gas chromatography of cellular fatty acids. *MIDI Tech. Note* **101**, 1–6 (1990).
33. Davis, J. J. *et al.* PATtyFams: Protein families for the microbial genomes in the PATRIC database. *Front. Microbiol.* **7**, 118. <https://doi.org/10.3389/fmicb.2016.00118> (2016).
34. Meier-Kolthoff, J. P. & Göker, M. TYGS is an automated high-throughput platform for state-of-the-art genome-based taxonomy. *Nat. Commun.* **10**, 2182. <https://doi.org/10.1038/s41467-019-10210-3> (2019).
35. Meier-Kolthoff, J. P., Carbasse, J. S., Peinado-Olarte, R. L. & Göker, M. TYGS and LPSN: A database tandem for fast and reliable genome-based classification and nomenclature of prokaryotes. *Nucleic Acids Res.* **50**, D801–D807. <https://doi.org/10.1093/nar/gkab902> (2022).
36. Lee, I., Ouk Kim, Y., Park, S. C. & Chun, J. OrthoANI: An improved algorithm and software for calculating average nucleotide identity. *Int. J. Syst. Evol. Microbiol.* **66**, 1100–1103. <https://doi.org/10.1099/ijsem.0.000760> (2016).
37. Tidjani Alou, M. *et al.* 'Khelaifiella massiliensis', 'Niameybacter massiliensis', 'Brachybacterium massiliense', 'Enterobacter timonensis', 'Massilibacillus massiliensis', new bacterial species and genera isolated from the gut microbiota of healthy infants. *New Microbes New Infect.* **19**, 1–7. <https://doi.org/10.1016/j.nmni.2017.02.002> (2017).
38. Schleifer, K. H. *et al.* Taxonomic study of anaerobic, gram-negative, rod-shaped bacteria from breweries: Emended description of *Pectinatus cerevisiiphilus* and description of *Pectinatus frisingensis* sp. nov., *Selenomonas lactificex* sp. nov., *Zymophilus raffinovorans* gen. nov., sp. nov., and *Zymophilus paucivorans* sp. nov. *Int. J. Syst. Bacteriol.* **40**, 19–27. <https://doi.org/10.1099/00207713-40-1-19> (1990).
39. Ueki, A., Watanabe, M., Ohtaki, Y., Kaku, N. & Ueki, K. Description of *Propionispira arcuata* sp. nov., isolated from a methanogenic reactor of cattle waste, reclassification of *Zymophilus raffinovorans* and *Zymophilus paucivorans* as *Propionispira raffinovorans* comb. nov. and *Propionispira paucivorans* comb. nov. and emended description of the genus *Propionispira*. *Int. J. Syst. Bacteriol.* **64**, 3571–3577. <https://doi.org/10.1099/ijls.0.063875-0> (2014).
40. Schink, B., Thompson, T. E. & Zeikus, J. G. Characterization of *Propionispira arboris* gen. nov. sp. nov., a nitrogen-fixing anaerobe common to wetwoods of living trees. *J. Gen. Microbiol.* **128**, 2771–2779 (1982).

41. Schauder, R. & Schink, B. *Anaerovibrio glycerini* sp. nov., an anaerobic bacterium fermenting glycerol to propionate, cell matter, and hydrogen. *Arch. Microbiol.* **152**, 473–478 (1989).
42. Strompl, C. *et al.* A re-evaluation of the taxonomy of the genus *Anaerovibrio*, with the reclassification of *Anaerovibrio glycerini* as *Anaerosinus glycerini* gen. nov., comb. nov., and *Anaerovibrio burkinabensis* as *Anaeroarcus burkinensis* [corrig.] gen. nov., comb. nov.. *Int. J. Syst. Bacteriol.* **49**, 1861–1872 (1999).
43. Auch, A. F., von Jan, M., Klenk, H. P. & Göker, M. Digital DNA–DNA hybridization for microbial species delineation by means of genome-to-genome sequence comparison. *Stand. Genom. Sci.* **2**, 117–134. <https://doi.org/10.4056/sigs.531120> (2010).
44. Jain, C., Rodriguez-R, L. M., Phillippy, A. M., Konstantinidis, K. T. & Aluru, S. High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nat. Commun.* **9**, 5114. <https://doi.org/10.1038/s41467-018-07641-9> (2018).
45. Forster, S. C. *et al.* A human gut bacterial genome and culture collection for improved metagenomic analyses. *Nat. Biotechnol.* **37**, 186–192. <https://doi.org/10.1038/s41587-018-0009-7> (2019).
46. Zou, Y. *et al.* 1,520 reference genomes from cultivated human gut bacteria enable functional microbiome analyses. *Nat. Biotechnol.* **37**, 179–185. <https://doi.org/10.1038/s41587-018-0008-8> (2019).
47. Poyet, M. *et al.* A library of human gut bacterial isolates paired with longitudinal multiomics data enables mechanistic microbiome research. *Nat. Med.* **25**, 1442–1452. <https://doi.org/10.1038/s41591-019-0559-3> (2019).
48. Campbell, C., Adeolu, M. & Gupta, R. S. Genome-based taxonomic framework for the class *Negativicutes*: division of the class *Negativicutes* into the orders *Selenomonadales* emend., *Acidaminococcales* ord. nov. and *Veillonellales* ord. nov.. *Int. J. Syst. Evol. Microbiol.* **65**, 3203–3215. <https://doi.org/10.1099/ijs.0.000347> (2015).
49. Harvey, K. L., Jarocki, V. M., Charles, I. G. & Djordjevic, S. P. The diverse functional roles of elongation factor Tu (EF-Tu) in microbial pathogenesis. *Front. Microbiol.* **10**, 2351. <https://doi.org/10.3389/fmicb.2019.02351> (2019).
50. Henderson, B., Fares, M. A. & Lund, P. A. Chaperonin 60: A paradoxical, evolutionarily conserved protein family with multiple moonlighting functions. *Biol. Rev. Camb. Philos. Soc.* **88**, 955–987. <https://doi.org/10.1111/brv.12037> (2013).

Acknowledgements

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF), funded by the Ministry of Education (NRF-2021R1A6A3A13038425 and 2021R1I1A1A01057496). We also thank Doctor Tae Kyu Lee, Doctor Jun Young Bae, and researcher Yu Min Han for their help.

Author contributions

S.Y. and H.P. isolated, identified, and cultured bacterial strains, designed the experiments, conceived the experiments, performed genomic analysis, acquired funding, and wrote the manuscript; S.Y. performed the phenotypic tests; C.B.R. collected the fecal samples from the patient; S.Y., H.P., H.K., and C.S.H. finalized the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-023-42017-0>.

Correspondence and requests for materials should be addressed to C.S.H.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023