Applied and Environmental Microbiology

Functional Screening of a Metagenomic Library Reveals Operons Responsible for Enhanced Intestinal Colonization by Gut Commensal Microbes

Mi Young Yoon, Kang-Mu Lee, Yujin Yoon, Junhyeok Go, Yongjin Park, Yong-Joon Cho, Gerald W. Tannock and Sang Sun Yoon

Appl. Environ. Microbiol. **2013, 79(12):3829. DOI: 10.1128/AEM.00581-13.**
Published Ahead of Print 12 April 2013.

AEM
Journals.ASM.org

Functional Screening of a Metagenomic Library Reveals Operons Responsible for Enhanced Intestinal Colonization by Gut Commensal Microbes

Mi Young Yoon,[a] Kang-Mu Lee,[a] Yujin Yoon,[a,b] Junhyeok Go,[a,b] Yongjin Park,[a] Yong-Joon Cho,[d] Gerald W. Tannock,[e] Sang Sun Yoon[a,b,c]

Department of Microbiology,[a] Brain Korea 21 Project for Medical Sciences,[b] Institute for Immunology and Immunological Diseases, Yonsei University College of Medicine,[c] Seoul, Republic of Korea; ChunLab, Inc., Seoul, Republic of Korea[d]; Department of Microbiology and Immunology, University of Otago, Dunedin, New Zealand[e]

Evidence suggests that gut microbes colonize the mammalian intestine through propagation as an adhesive microbial community. A bacterial artificial chromosome (BAC) library of murine bowel microbiota DNA in the surrogate host *Escherichia coli* DH10B was screened for enhanced adherence capability. Two out of 5,472 DH10B clones, 10G6 and 25G1, exhibited enhanced capabilities to adhere to inanimate surfaces in functional screens. DNA segments inserted into the 10G6 and 25G1 clones were 52 and 41 kb and included 47 and 41 protein-coding open reading frames (ORFs), respectively. DNA sequence alignments, tetranucleotide frequency, and codon usage analysis strongly suggest that these two DNA fragments are derived from species belonging to the genus *Bacteroides*. Consistent with this finding, a large portion of the predicted gene products were highly homologous to those of *Bacteroides* spp. Transposon mutagenesis and subsequent experiments that involved heterologous expression identified two operons associated with enhanced adherence. *E. coli* strains transformed with the 10a or 25b operon adhered to the surface of intestinal epithelium and colonized the mouse intestine more vigorously than did the control strain. This study has revealed the genetic determinants of unknown commensals (probably resembling *Bacteroides* species) that enhance the ability of the bacteria to colonize the murine bowel.

Metagenomics aims to characterize a collection of genetic materials as they exist in a microbial ecosystem (1). This method stands in contrast to characterization by isolation of individual colonies. Because metagenomics offers a unique opportunity to study organisms that are not cultured in a laboratory, it opens access to a reservoir of novel microbial genes.

The large bowels of mammalian species are colonized by microbial communities that are referred to as the gut microbiota. The communities, mostly bacterial in composition, have considerable biodiversity and gain much of their energy and carbon requirements from the hydrolysis of plant glycans and fermentation of the hydrolysis products. Additionally, some members of the community utilize mucins from mucus and the components of enterocytes sloughed from the intestinal mucosal surface (2, 3). Both the metabolic activities and antigenicity of the microbiota have important physiological and immunological repercussions for the host (4–7).

Although many of the bacterial commensals of the human intestine have now been cultured (8), most information with regard to the bacterial community has been derived from high-throughput sequencing studies (3, 9). This strategy has revealed the complexity and functional potential of the communities but relies on gene annotations in public databases. However, these annotations, confounded by the detection of numerous hypothetical proteins of unknown function, may not reveal the full potential of proteins encoded by genes detected in as-yet uncultivated bacteria.

Functional screens of metagenomic libraries of microbiota DNA can uncover important functional information about bacterial inhabitants. Of particular note was the discovery of proteorhodopsins in marine bacteria by functional metagenomics (10, 11). Therefore, we hypothesized that a bacterial artificial chromosome (BAC) metagenomic library of murine large-bowel microbiota encoded proteins with functions associated with intestinal colonization by commensal bacteria. In support of this hypothesis, bowel commensals were detected in association with mucosal biopsy specimens (12). In this work, we screened a library of *Escherichia coli* clones for enhanced adherence of the surrogate host to surfaces. Each clone harbored cloned DNA derived from the large-bowel microbiota of BALB/c mice (13). We further characterized two operons that were found to play key roles in adhesion when expressed heterologously in *E. coli*. We also tested the effects of these genetic elements on intestinal colonization *in vivo*. Detection of such genetic elements can increase our knowledge and understanding of how commensal microbes colonize the mammalian intestine.

MATERIALS AND METHODS

Bacterial strains and growth conditions. The entire procedure for the library construction has been described elsewhere (13). The metagenome library was prepared in the pIndigoBAC-5 vector (Epicentre, Madison, WI). Bacterial strains were cultured in Luria broth (LB; 10 g tryptone, 5 g NaCl, 5 g yeast extract per liter) supplemented with 34 μg/ml of chloramphenicol (Sigma-Aldrich Co., St. Louis, MO) at 37°C. For anaerobic growth, bacteria were grown in an anaerobic chamber (Coylab Inc., Grass

Received 21 February 2013 Accepted 6 April 2013
Published ahead of print 12 April 2013
Address correspondence to Sang Sun Yoon, sangsun_yoon@yuhs.ac.
Supplemental material for this article may be found at http://dx.doi.org/10.1128/AEM.00581-13.

doi:10.1128/AEM.00581-13

Lake, MI) that was filled with mixed gas (nitrogen, 90%; hydrogen, 5%; carbon dioxide, 5%). Oxbile (catalog no. 7216; Neogen Corp., Lansing, MI) was used to examine the effect of bile acids on biofilm formation.

Biofilm assay. The BAC clones were inoculated with a 96-well pin replicator (Boekel Scientific Corp., Feasterville, PA) into 200 μl of LB plus chloramphenicol placed in wells of 96-well plates (SPL Lifesciences Inc., Pocheon-Si, South Korea) and grown for 2 days at 37°C. To prevent evaporation-mediated volume loss during growth, the plates were situated in a polyvinyl bag together with damp tissues. After 48 h, before biofilm assays were performed, bacterial growth was assessed by measuring the optical density at 600 nm (OD_{600}) using a Thermomax microplate reader (Molecular Devices Inc., Sunnyvale, CA). Biofilm assays were performed by following the procedures previously reported (14). For biofilm three-dimensional (3D) architecture examination, circular dishes (model no. D27F; InfiniteBio Inc., San Jose, CA) that contained 2 ml of biofilm growth medium were inoculated with precultures of each strain that had been grown overnight. After biofilm growth for 2 days at 37°C, biofilm cells were washed with phosphate-buffered saline (PBS) and stained with 1 ml of 3.34 μM Syto 9 nucleic acid dye, which is freely penetrable to cells. Image acquisition and data processing were accomplished using a confocal laser scanning microscope (FV-1000; Olympus Optical Co. Ltd., Japan) and its operating software, FV10-ASW (ver. 02.01). After scanning with a 488-nm excitation laser at a sampling speed of 2 μs/pixel, a 512- by 512-pixel, 12 bits/pixel, 2D *x-y* image (126.728 by 126.728 μm) was acquired. For the *z* dimension, 80 slices that gave up to 40-μm depths (0.5 μm/slice) were obtained. SYTO 9 green fluorescence was detected through a 500- to 540-nm bandpass filter. A UPLSAPO 100XO (Olympus) objective lens was used for the bacterial cell image analysis. Images were saved as TIF files with embedded *x-y-z* scale lines.

***De novo* sequencing and annotation.** Sequencing of BAC plasmids was performed by Macrogen Inc. (Seoul, South Korea) using a shotgun sequencing method. Sequence data were assembled using software described elsewhere (15). Open reading frames (ORFs) were assigned by Glimmer 3.0 software (16). The start and stop codons of each ORF were manually confirmed, and the presence of a promoter sequence was identified in the upstream sequence of each gene. The search for homologous proteins was performed against the database in the National Center of Biotechnology Information using a BLASTp algorithm. Clustering of genes into an operon was performed using FGENESB, a program for the prediction of bacterial operons (SoftBerry, Mount Kisco, NY). The ORF map shown in Fig. 3A and B was constructed using CloneMap (ver. 2.11) software (CGC Scientific, Inc., Ballwin, MO).

Tetranucleotide frequencies and codon usage. The tetranucleotide frequency of the 10G6 or 25G1 sequence was compared to that of each of the 31 bacterial genomes used for the analysis. These genomes are of bacterial species that are mainly gut commensals and are listed in the supplemental material. The correlation coefficient between two query sequences was calculated using the TETRA stand-alone program (http://www.megx.net/tetra). The similarity of codon usage between two sequences was computed using a web-based algorithm (http://gcua.schoedl.de).

RNA extraction and RT-PCR analysis. To examine whether transcription occurred from the cloned genes in the *E. coli* host, reverse transcriptase PCR (RT-PCR) analyses were performed. DH10B strains harboring the 10G6 or 25G1 BAC were grown in LB with chloramphenicol to an OD_{600} of ~1.5. RNA extraction and RT-PCR were performed as described previously (17), and primers used for PCR are listed in the supplemental material. Transcript levels of *rpoA* and *rpoD* genes were used as internal controls.

***In vitro* transposon mutagenesis of 10G6 and 25G1 BACs and secondary screening for defective biofilm producers.** To further elucidate the genes that play crucial roles in biofilm formation, *in vitro* transposon (Tn) insertion mutagenesis was employed using the EZ-Tn5 <TET-1> insertion kit (Epicentre Biotechnologies, Madison, WI). An *in vitro* transposition reaction using the 10G6 or 25G1 BAC and the transposon was achieved by following the manufacturer's instructions and is outlined in Fig. S3 in the supplemental material. After random transposon insertion, *E. coli* DH10B was transformed with a pool of mutated BACs by electroporation. Electroporation was performed using a Gene Pulser (Bio-Rad Laboratories, Hercules, CA) with the pulse controller set at 200 Ω, 25 μF, and 1.5 kV. After 16 to 18 h of incubation, transposon insertion clones were selected on LB plates containing 10 μg/ml tetracycline (Tc). A total of ~120 10G6-derived and ~480 25G1-derived Tc-resistant colonies were picked, and their biofilm-forming capabilities were evaluated using the simple microtiter dish biofilm assay as described earlier. The transposon insertion sites of the mutants selected on the basis of their defective biofilm were precisely determined by sequencing the transposon-BAC junctions using the transposon-specific flanking primers provided in the kit.

Construction of *E. coli* DH10B strains that heterologously express the 10a or 25b operon. The DNA fragment of the 10a (~2.65 kbp) or 25b (~5.38 kbp) operon was PCR amplified to cover the entire operon sequence and its endogenous promoter region. PCR products were cloned into the multicloning site of the pUC19 plasmid, and the resultant plasmids, named pUC19::10a and pUC19::25b, were transformed into the DH10B *E. coli* host strain. PCR primers used for cloning are listed in the supplemental material.

Effects of heterologous expression of 10a or 25b operon on the cell adhesion capability and *in vivo* colonization. The bacterial adhesion assay was performed as described elsewhere (18). The adhesion capability of bacterial cells was assessed by counting the number of bacteria adherent to the HT29 cells. An *in vivo* colonization assay was performed using antibiotic-pretreated adult BALB/c mice ($n = 6$) as described in Fig. S4 in the supplemental material. Isolated intestinal tissue (from duodenum to rectum) was individually placed in sterile PBS and ground using a tissue homogenizer for bacterial counting. Ampicillin (100 μg/ml) was used for selective growth of *E. coli* strains. The animal study was approved by the Yonsei University Animal Research Ethics Committee, and antibiotic treatment before bacterial infection was performed as described previously (19).

Preparation of mouse gut microbiota metagenome and quantitative real-time PCR. Whole intestinal tissue was individually removed and immediately frozen in liquid nitrogen. The frozen gut tissues were ground by mortar and pestle. Approximately 200 mg of the homogenate was suspended in 1.4 ml of ASL buffer provided in the QIAamp DNA stool Mini kit (Qiagen Inc., Valencia, CA), and genomic DNA was extracted by following the manufacturer's instructions. The quantity of purified DNA was assessed spectrophotometrically, and 400 ng of DNA was used as a template to amplify the *E. coli* 16S rRNA gene, *Bacteroidetes* phylum 16S rRNA gene, ORF20 in the 10a operon, and ORF13 in the 25b operon. The real-time PCR was performed as described previously (20), and primers used for the PCR are listed in the supplemental material.

Construction of a DH10B reporter strain containing chromosomally encoded green fluorescent protein (GFP). A 500-bp noncoding region between the *lacY* and *lacI* genes in the genome of *E. coli* DH10B was PCR amplified and ligated with a *gfp* gene cassette containing its own promoter. The ligated fragment was cloned into pCVD442 vector containing both gentamicin and ampicillin markers (20), and the resultant plasmid (pCVD442-*gfp*) was conjugated into *E. coli* DH10B. DH10B transconjugants, in which pCVD442-*gfp* was integrated in the 500-bp noncoding region via homologous recombination, were selected and sequence verified.

Live animal imaging analysis. Infant mice were orogastrically inoculated with 50 μl of cultures (5×10^7 CFU). At 24 h postinfection, the mice were anesthetized with Zoletil and Rompun. Whole-body images were obtained with the IVIS spectrum system (Caliper Life Sciences, Alameda, CA), and the acquired fluorescent signals were processed using Living Image 4.1 software (Caliper Life Sciences).

Statistical analysis. Data are expressed as means ± standard deviations (SD). Unpaired Student's *t* test was used to analyze the data, and a *P*

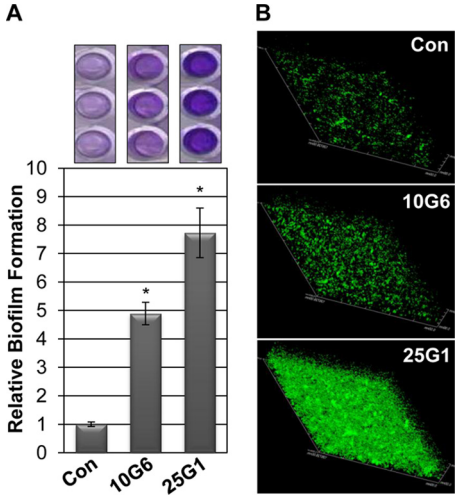


FIG 1 Biofilm formation of 10G6 and 25G1 clones. (A) Representative images of crystal violet staining of biofilms formed by the control (Con), 10G6, and 25G1 clones in the 96-well microtiter plates. Quantification of crystal violet staining (OD_{540}) normalized to cell growth (OD_{600}) was used to represent each strain's biofilm robustness. Means ± SD are shown by each bar; *, $P < 0.01$ versus control biofilm. (B) Three-dimensional confocal microscopic images of 2-day-old biofilms. Bacterial cells were stained with Syto 9, a green fluorescence dye, and 126.728- by 126.728- by 21.5-μm *x-y-z* images were obtained as described in Materials and Methods.

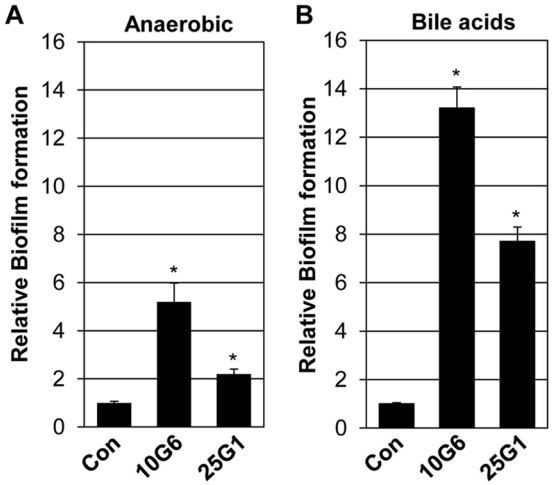


FIG 2 Effects of anaerobiosis and the presence of bile acids on biofilm formation. Biofilm formation of 10G6 and 25G1 clones under anaerobic growth conditions (A) and in the presence of 0.6% bile acids (B) was quantified. Experimental conditions were identical to those described for the experiment shown in Fig. 1A. Values displayed in each bar are the means ± SD. *, $P < 0.01$ versus control biofilm.

value of <0.05 was considered statistically significant. All experiments were repeated to verify reproducibility.

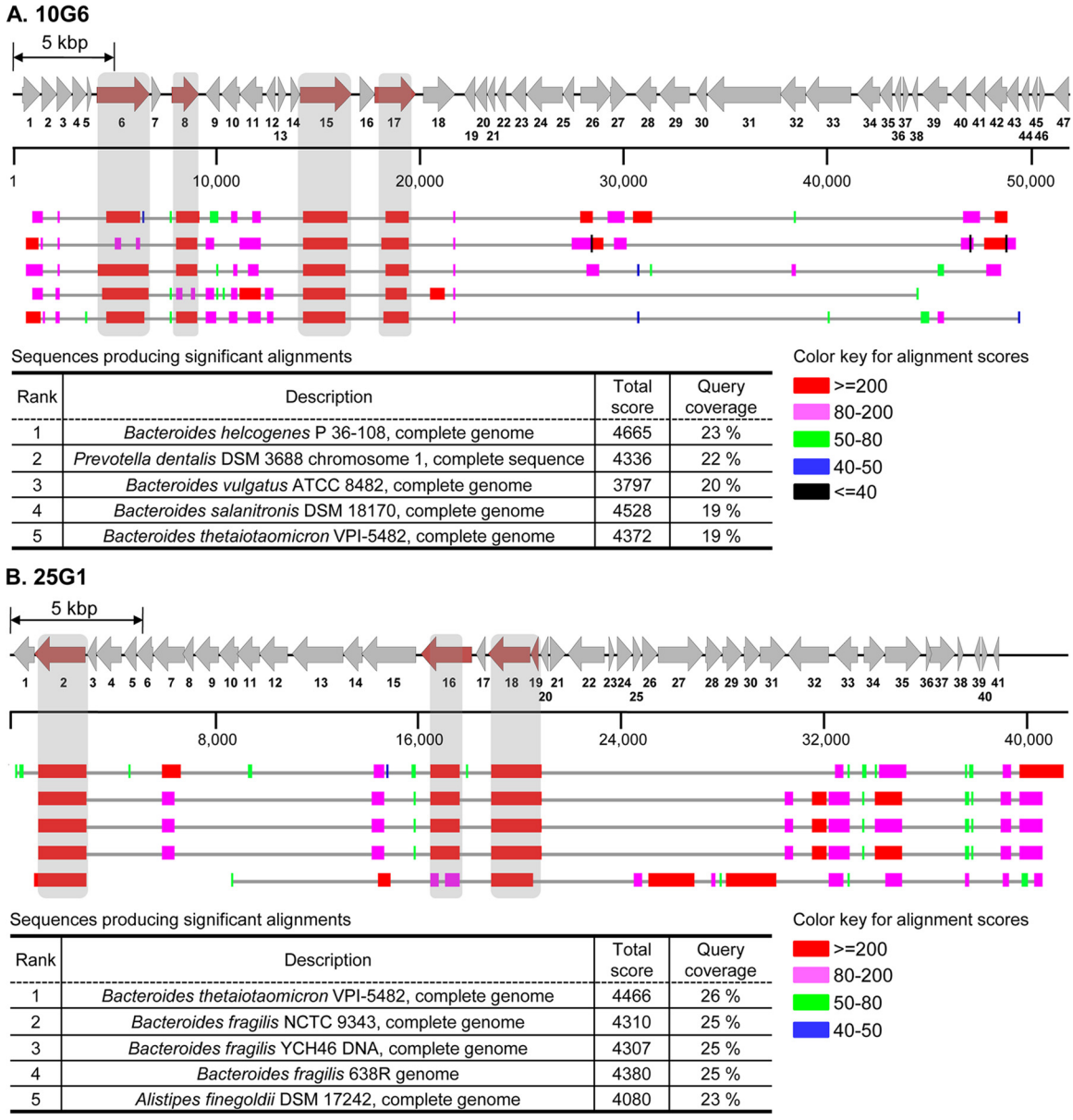
Nucleotide sequence accession numbers. Full-length sequences for clones 10G6 and 25G1 are available in the NCBI database (accession numbers KC595276 and KC595277).

RESULTS

Screening of a gut microbiota metagenome library for clones with adhesive capability. To reveal genetic determinants, expression of which results in enhanced adhesion of *E. coli* host cells to a surface, we screened a gut microbiota metagenome library for biofilm producers. A schematic diagram of the screening procedure is shown in Fig. S1A in the supplemental material. Using a simple, highly reproducible biofilm assay performed in 96-well plates, we screened 5,472 *E. coli* clones for their biofilm-forming capability (14). Since biofilm formation is proportional to cell growth, biofilm formation was normalized with bacterial cell growth, which was monitored by measuring the OD_{600} (see Fig. S1A). *E. coli* DH10B that harbored an empty pIndigoBAC-5 vector was used as a baseline control. Figure S1B shows the distribution of the biofilm-forming capability of all clones tested. Approximately 83% of clones formed biofilms that were either comparable to or slightly larger than that of the control strain. Notably, about 2% of the tested clones were determined to form biofilms that were >4 times more robust than the control in our initial primary screen (fold induction, >4) (see Fig. S1B). Each biofilm-forming BAC clone was retested, and two clones were validated (10G6 and 25G1). Based on our crystal violet staining assay, biofilms formed by clones 10G6 and 25G1 were ~4.8- and ~7.5-fold, respectively, more robust than the control biofilm (Fig. 1A). The biofilms were grown on coverslips and examined by confocal laser scanning microscopy for comparison with respect to spatial architecture. Figure 1B shows the three-dimensional images (126.852 by 126.728 by 21.5 μm) of the biofilms. Consistent with biofilm formation on the surfaces of 96-well plates (Fig. 1A), the two selected clones formed robust biofilms on glass surfaces, with the 25G1 clone forming the thicker and denser biofilm.

Effects of anaerobiosis and the presence of bile acids on biofilm formation by selected clones. Diverse obligate anaerobes colonize the human intestine (21), suggesting that the human intestine maintains an anaerobic environment. Bile acids present in the large intestine also may influence the growth of gut microbes (22, 23). Therefore, we tested the biofilm-forming capability of selected clones during growth under anaerobic conditions or in the presence of bile acids. Although anaerobic conditions retarded bacterial growth, clone 10G6 still formed a robust biofilm, suggesting that biofilm formation by this clone does not require aerobiosis. In contrast, the clone 25G1 formed weaker biofilms during anaerobic growth than when grown aerobically (Fig. 2A). During static growth in the presence of 0.6% (wt/vol) bile acids, the 10G1 strain formed biofilms that were ~13-fold denser than the control biofilm, although bile acids reduced bacterial growth (Fig. 2B). Under the same growth conditions, 25G1 formed an ~8-fold-denser biofilm than the control (Fig. 2B).

Sequence analysis of 10G6 and 25G1 BACs. To better understand the molecular basis for the formation of the adhesive bacterial community, clones 10G6 and 25G1 were sequenced. DNA inserts assembled after shotgun sequencing and subsequent gap filling were 51,956 and 41,474 bp for the clones 10G6 and 25G1, respectively. These insert sizes are close to the average insert length of ~55,000 bp for the library (13). Our bioinformatic analysis revealed that the cloned sequences of 10G6 and 25G1 contained 47 and 41 protein-coding open reading frames (ORFs), respectively (Fig. 3A and B). The BLASTn search using each full-length sequence as a query failed to retrieve a known sequence element, suggesting that these two inserted DNA sequences were derived from microbes with unknown genome sequences. BLASTn results presented in Fig. 3A and B, however, clearly demonstrate that both of the fragments are highly similar to homologue regions of *Bac-*



Rank	Description	Total score	Query coverage
1	*Bacteroides helcogenes* P 36-108, complete genome	4665	23 %
2	*Prevotella dentalis* DSM 3688 chromosome 1, complete sequence	4336	22 %
3	*Bacteroides vulgatus* ATCC 8482, complete genome	3797	20 %
4	*Bacteroides salanitronis* DSM 18170, complete genome	4528	19 %
5	*Bacteroides thetaiotaomicron* VPI-5482, complete genome	4372	19 %

Rank	Description	Total score	Query coverage
1	*Bacteroides thetaiotaomicron* VPI-5482, complete genome	4466	26 %
2	*Bacteroides fragilis* NCTC 9343, complete genome	4310	25 %
3	*Bacteroides fragilis* YCH46 DNA, complete genome	4307	25 %
4	*Bacteroides fragilis* 638R genome	4380	25 %
5	*Alistipes finegoldii* DSM 17242, complete genome	4080	23 %

FIG 3 Open reading frame (ORF) maps of the 10G6 and 25G1 clones and BLASTn search results. ORF maps of the 10G6 (A) and 25G1 (B) clones. The length and direction of the arrow indicate the relative size and the transcriptional direction of each ORF. Further details pertaining to the predicted proteins are provided in the supplemental material. The five sequences most similar to those of 10G6 (A) or 25G1 (B) are shown with color-coded BLASTn alignment scores. Bacterial genomes containing homologous regions are sorted by query coverage and are listed in the table.

teroides spp., a major group of the gut microbiota (24). Among the 5 highest-ranked sequences that produced significant alignments, 4 sequences were from genomes of the genus *Bacteroides* in each BLASTn search. Four genes in the 10G6 sequence (i.e., genes 6, 8, 15, and 17; gray-shaded regions in Fig. 3A) are highly conserved, and most of them are present in the genome of the top-ranked *Bacteroides* species. Likewise, genes 2, 16, 18, and 19 in the 25G1 sequence are also highly conserved, and all of them are located in the core genomes (Fig. 3B).

We then compared the tetranucleotide frequencies (TFs) of the 10G6 and 25G1 sequences of each of the 31 bacterial genomes used in the analysis. As shown in Table 1, the highest correlation coefficient was obtained with the *Bacteroides uniformis* genome sequence. Interestingly, 4 bacterial species (i.e., *Bacteroides uniformis*, *Tannerella* sp., *Bacteroides oleiciplenus*, and *Bacteroides vulgatus*) are commonly included in the top-ranked species in the analysis with 10G6 and 25G1 sequences (Table 1). The entire matrix of TF analysis is shown in the supplemental material. Codon usage in the 25G1 sequence was most similar to that of *Bacteroides uniformis*, and *Bacteroides* sp. strain D2 is among the species that use genetic codes similar to genes in the 10G6 sequence (Table 2).

We next sought to assign putative functions to each predicted protein by a BLASTp search. Among the 47 genes in the 10G6 clone, 27 genes produced homologous proteins of a known source. Importantly, 11 out of those 27 proteins were found to be derived from *Bacteroides* spp. (see the ORF analysis tab in the

TABLE 1 Bacterial species that exhibit significant similarities of tetranucleotide frequencies to those of the 10G6 and 25G1 sequences

	Species exhibiting similarities to:			
	10G6		25G1	
Rank	Correlation coefficient	Species	Correlation coefficient	Species
1	0.76516	*Bacteroides uniformis* ATCC 8492	0.62521	*Bacteroides uniformis* ATCC 8492
2	0.74825	*Tannerella* sp. strain 6_1_58FAA_CT1	0.60482	*Tannerella* sp. strain 6_1_58FAA_CT1
3	0.70313	*Bacteroides oleiciplenus* YIT 12058	0.59633	*Bacteroides oleiciplenus* YIT 12058
4	0.69374	*Bacteroides vulgatus* ATCC 8482	0.55767	*Bacteroides* sp. strain D2
5	0.68189	*Desulfovibrio vulgaris* Hildenborough	0.55723	*Bacteroides vulgatus* ATCC 8482

Excel file in the supplemental material). In clone 25G1, 29% of the predicted proteins of known sources showed homology to proteins of *Bacteroides* spp. (see the ORF analysis tab in the Excel file in the supplemental material). A smaller number of proteins were similar to proteins of *Barnesiella*, *Clostridium*, *Prevotella*, *Parabacteroides*, *Odoribacter*, or *Porphyromonas* spp. In general, proteins with homology to proteins of *Bacteroides* spp. were determined to be homologous with greater statistical certainty than were proteins similar to those of other bacterial groups. Taken together, our comprehensive sequence analyses suggest that 10G6 and 25G1 sequences are derived from species of *Bacteroides* lineage.

We then performed reverse transcriptase PCR analysis to determine whether the expression of genes present in the cloned insert indeed occurred in the surrogate *E. coli* host. As shown in Fig. S2 in the supplemental material, all selected genes, three from 10G6 and five from 25G1, were transcribed in *E. coli* transformed with the corresponding BAC clone, but they were not transcribed in the control strain. It is worth noting that transcript levels varied among the different genes. The *orf12* and *orf13* genes were transcribed to a higher level than the other genes in 25G1. These results strongly suggested that enhanced biofilm formation of strains containing clones 10G6 and 25G1 was due to the heterologous expression of the genes in *E. coli*.

Identification of operons crucial for biofilm formation. Sequence analysis provided a basis to explore genetic determinants for biofilm formation of the 10G6 and 25G1 clones. To precisely define the genes involved, we sought to identify mutants of the 10G6 or 25G1 clone that had lost their abilities to form biofilms on the same surface. Mutagenesis was achieved by *in vitro* transposon (Tn) insertion, and biofilm-defective mutants were isolated by screening a library of *E. coli* DH10B transformed with a collection of Tn-marked 10G6 or 25G1 BACs. The detailed procedure is described in Fig. S3 in the supplemental material. Among ~120 mutants derived from 10G6, 2 mutants (3C10 and 2F2) were determined to form defective biofilms compared to the biofilm of the parental strain (Fig. 4A). Two mutants (7B9 and 4B10) were recovered from screening ~480 mutants derived from 25G1 (Fig. 4B). The biofilm-forming capability of each mutant compared to that of the parental strain (i.e., 10G6 or 25G1 clone) is displayed in the far right column of each table presented above the ORF map (Fig. 4A and B). To map the Tn insertion sites, we directly sequenced the Tn-BAC junctions. Genes interrupted by the Tn insertion are indicated with dashed arrows in the middle columns of the tables (Fig. 4A and B). A 10G6-derived mutant (2F2) was found to possess a transposon insertion in ORF20, and the 3C10 mutant harbored an insertion in the intergenic region between ORF21 and ORF22 (Fig. 4A). The two 25G1-derived mutants (7B9 and 4B10) both harbored Tn insertions in ORF13 (Fig. 4B).

To further support functional roles of these two genetic elements in biofilm formation, we constructed *E. coli* DH10B strains that heterologously expressed the 10a or 25b operon (red lines in Fig. 4A and B) and tested their biofilm-forming capabilities. DNA fragments, including the entire operon sequences and corresponding upstream promoter regions, were cloned into the multicopy plasmid, pUC19. As shown in Fig. 5A, biofilms formed by *E. coli* strains harboring pUC19::10a (here called 10a) or pUC19::25b (here called 25b) were significantly more robust than the control biofilm. Consistent with the results presented in Fig. 1, the biofilm formed by the 25b strain was denser than that of the 10a strain. The depths of biofilms formed by the 10a and 25b strains were 15.12 and 33.82 μm, respectively, while that of the control was 9.79 μm (Fig. 5A). These results clearly support roles for these two genetic elements in the formation of the bacterial community with biofilm-forming capability.

Relative abundance of 10a or 25b sequence in the mouse gut microbiota metagenome. We next sought to gain an idea of how abundantly the 10a and 25b elements are present in the mouse gut microbiota metagenome. To address this issue, we performed quantitative real-time PCR (qRT-PCR) using primers that specifically amplify ORF20 and ORF13 of the 10a and 25b operons,

TABLE 2 Bacterial species that exhibit significant similarities of codon usage to those of the 10G6 and 25G1 sequences

	Species exhibiting similarities to:			
	10G6		25G1	
Rank	Similarity (%)	Species	Similarity (%)	Species
1	80.37	*Bifidobacterium longum* NCC 2705	83.08	*Bacteroides uniformis* ATCC 8492
2	80.20	*Bacteroides* sp. strain D2	82.09	*Porphyromonas gingivalis* W83
3	80.11	*Geobacter sulfurreducens* PCA	77.61	*Bdellovibrio bacteriovorus* HD100
4	76.23	*Bordetella bronchiseptica* RB50	75.92	*Bifidobacterium longum* NCC 2705
5	76.16	*Bacteroides uniformis* ATCC 8492	75.91	*Geobacter sulfurreducens* PCA

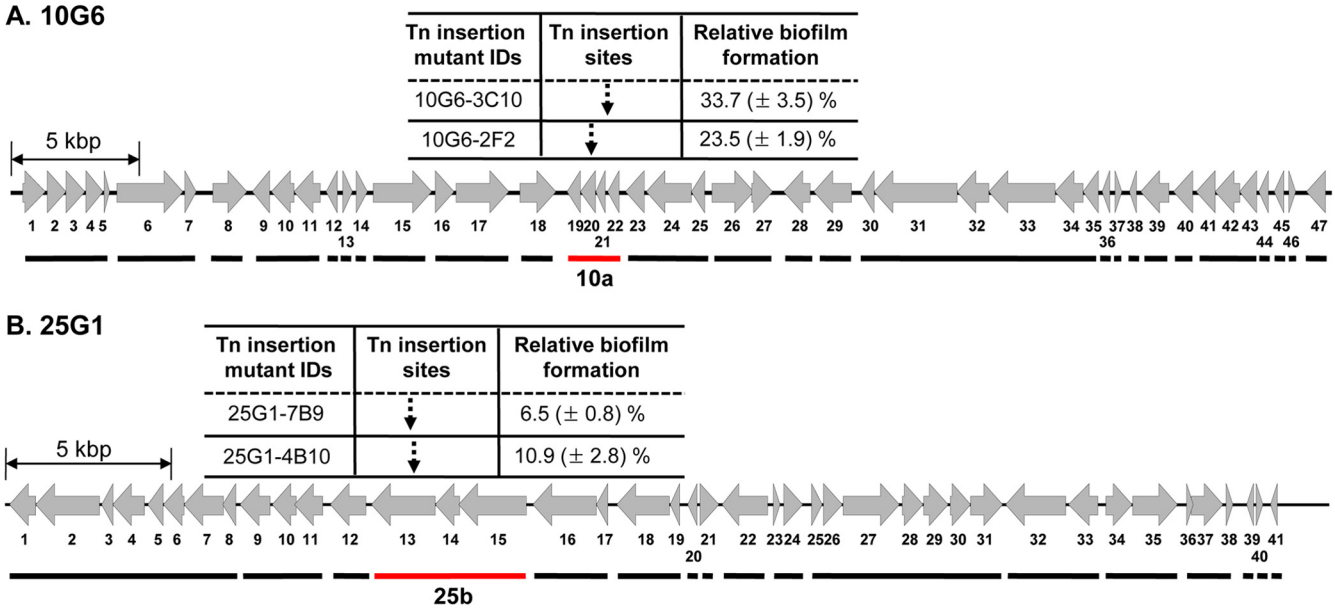


FIG 4 Identification of Tn-inserted mutant clones defective in biofilm formation. Tn insertion mutants that were identified as defective biofilm producers and genes disrupted in such mutants (dashed downward arrows) are shown in the table placed at the top of each ORF map. Relative biofilm formation of each mutant is displayed as the percentage of biofilm formed by the parental clone in the far right column of the table. Solid lines at the bottom of each map denote genes clustered in an operon.

respectively. These two genes were selected because they were originally identified in our *in vitro* transposon mutant library screen (Fig. 4). For quantitative analysis, we compared RQ (relative quantity) values of target genes to those of the *E. coli* 16S rRNA gene, which were normalized to 1.0. We also measured the RQ values of 16S rRNA genes present in species classified as *Bacteroidetes* at the phylum level (25). The levels of 16S rRNA genes present in the *Bacteroidetes* phylum were >2,000-fold higher than that of the *E. coli* 16S rRNA gene (Fig. 5B). This result verified that the qRT-PCR technique is indeed useful to measure the RQ of a particular DNA sequence in a mixed-DNA sample. The relative abundances of ORF20 and ORF13 in the mouse gut microbiota metagenome were ~24 and ~ 31%, respectively, of that of the *E. coli* 16S rRNA gene (Fig. 5B). Given that *E. coli* is one of the core members of the gut microbiota, these results suggest that bacterial species that contain either gene in their genome likely are present in sufficient quantity in the mouse intestine.

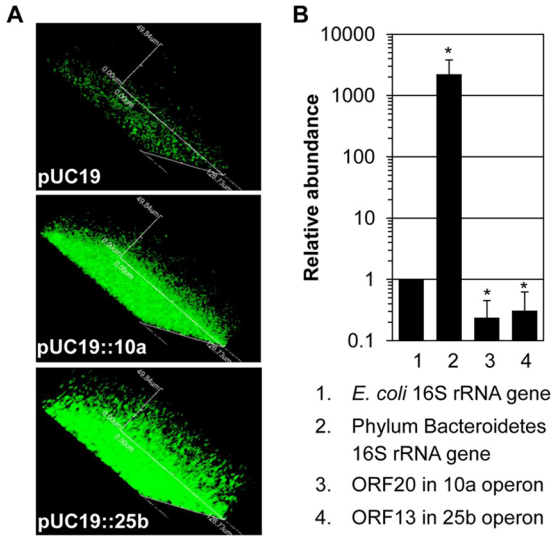


FIG 5 Biofilm formation of 10a and 25b strains and relative abundance of 10a and 25b sequences in the mouse gut microbiota metagenome. (A) Three-dimensional biofilm images of *E. coli* strains harboring pUC19, pUC19::10a, or pUC19::25b. Experimental conditions were identical to those for the experiment shown in Fig. 1B. (B) Quantitative RT-PCR was conducted on genomic DNA extracted from mouse gut microbiota. Relative abundance of the *E. coli* 16S rRNA gene was normalized to 1.0. Three independent experiments were performed, and values (means ± SD) are displayed in each bar. *, $P < 0.05$ versus the relative quantity of *E. coli* 16S rRNA gene.

10a and 25b strains exhibited enhanced *in vitro* adherence capability and *in vivo* colonization. We examined the capabilities of 10a and 25b strains to adhere to intestinal epithelium. After 2 h of incubation with confluent HT29 cells followed by rigorous washing, the number of 10a cells remaining adherent was ~17-fold higher than the number of control *E. coli* cells (Fig. 6A). Under the same experimental conditions, about 4 times more 25b cells than control cells remained adhered to the cultured epithelial tissue (Fig. 6A). During the incubation with bacteria, the HT29 cells maintained viability (data not shown).

We next investigated whether the increased adherence capability of the two strains to cultured epithelium was reflected in their abilities to colonize mouse intestine. It was reported that colonization resistance of adult mice was significantly decreased in antibiotic-pretreated versus conventionally reared mice (26). Prior to being infected, therefore, a group of BALB/c mice was pretreated with antibiotics for a total of 4 days, as outlined in Fig. S4 in the supplemental material. The bacterial content of feces of antibiotic-pretreated mice was significantly reduced (see Fig. S4). Similar to results for cell adhesion, inoculation with either strain resulted in enhanced bacterial colonization in the mouse intestine.

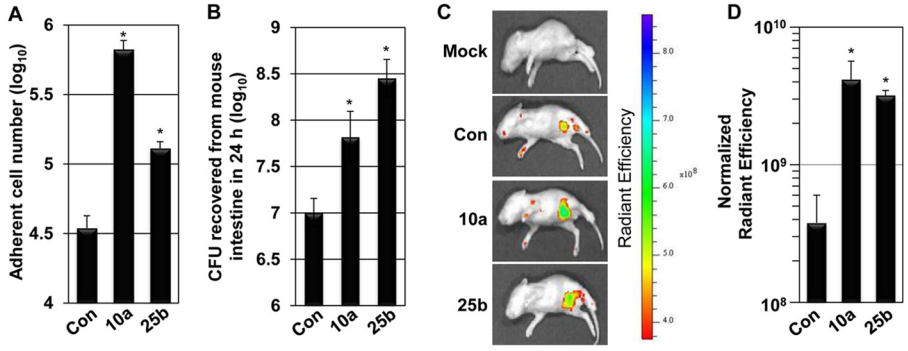
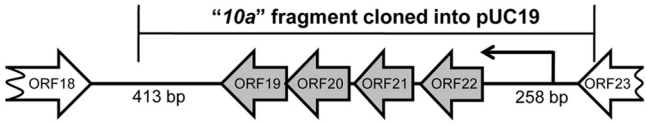


FIG 6 Capability of 10a and 25b strains to adhere to cultured intestinal epithelium and to colonize mouse intestine. (A) Three different *E. coli* DH10B strains harboring the control plasmid, pUC19::10a, or pUC19::25b (1×10^7 CFU) were incubated with confluent HT29 cells for 2 h. At the end of the incubation, unbound bacteria were removed by washing three times with PBS. HT29 cells and attached bacterial cells then were released from each well with a trypsin-EDTA mixture, and the remaining adherent bacteria were enumerated by plating on agar plates. Values (means ± SD) are displayed by each bar. *, $P < 0.01$ versus the number for the adherent control strain. (B) Antibiotic-treated BALB/c adult mice were infected with the same set of bacterial strains by oral gavage (5×10^9 CFU). After 24 h, animals were sacrificed, and total gut tissue samples, placed in PBS, were homogenized. Bacterial cells were enumerated by plating the serially diluted cells on agar plates. Values (means ± SD) are displayed by each bar ($n = 6$). *, $P < 0.01$ versus the number for the control strain. (C) Five- to 6-day-old BALB/c infant mice were orogastrically inoculated with bacterial strains (5×10^7 CFU). Mice were anesthetized at 24 h postinfection, and the whole body was scanned to acquire fluorescent signals using an IVIS imaging system. Fluorescent intensity was calculated as radiant efficiency (RE) using the accompanying software (Living Image 4.1). (D) The RE in each group of mice was subtracted from that of mock-infected control animals and quantitatively compared. Values (means ± SD) are displayed by each bar ($n = 4$). *, $P < 0.05$ versus the value for control mice.

Notably, the 25b strain colonized the mouse intestine to a higher level than the 10a strain. The quantities of 10a and 25b strains recovered from the mouse gut homogenates were ~6.6- and ~28-fold higher, respectively, than the quantity of control *E. coli* cells harboring the empty pUC19 vector (Fig. 6B). These results indicate that the 10a and 25b strains, which show enhanced *in vitro* adhesion capabilities, can also colonize the mouse intestine.

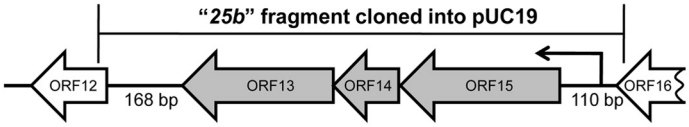
Finally, GFP-tagged strains were inoculated into the intestine of infant mice, and the extent of intestinal colonization was visualized using an IVIS imaging system. The large bowels of infant mice have a microbiota in which facultative anaerobes such as *E. coli* predominate. In addition, the thin, hairless abdominal skin of the infant mouse offers a technical advantage for successful live-imaging analysis. When mouse intestines were imaged at 24 h postinoculation, fluorescent signals with various degrees of intensity were detected in the lower ventral region of mice (Fig. 6C), suggesting that orogastrically inoculated bacterial cells did establish in the large intestine. The fluorescent signals from the mice infected with 10a and 25b were ~11.0- and ~8.4-fold more intense, respectively, than the signal from control infected mice (Fig. 6D). These results demonstrate that the expression of 10a or 25b genetic elements contributes to intestinal colonization.

BLASTp search of genes in operons 10a and 25b. Operon 10a contains four genes (Fig. 7). For ORF19, a BLASTp search retrieved a hypothetical protein of unknown source with very low statistical significance. ORF21 and ORF22 encode hypothetical proteins, and no distinct domain was identified in these two proteins by our bioinformatic search (Fig. 7). Hence, we could not



ORF #	Predicted Proteins (size) [Source species]	Characteristic Domains	E-value
ORF19	Hypothetical protein (174 aa) [null]	none	-1
ORF20	Conserved hypothetical protein (191 aa) [*Bacteroides* sp. 9_1_42FAA]	• DUF847 : Predicted lysozyme • PG_binding_3 : Peptidoglycan binding domain	3.42E-56
	Putative secretion-activator protein (191 aa) [*Prevotella multiformis* DSM 16608]		4.97E-47
	N-acetylmuramidase (191 aa) [*Teredinibacter turnerae* T7901]		3.12E-14
ORF21	Hypothetical protein (116 aa) [null]	none	-1
ORF22	Hypothetical protein (147 aa) [null]	none	-1

FIG 7 Genetic map of the 10a operon and the characteristics of predicted proteins encoded by each gene. The four genes included in the 10a operon and the DNA fragment amplified for cloning into pUC19 are drawn to scale. Predicted proteins with BLASTp similarity to proteins of other gut microbes are listed along with E values that measure the statistical significance. The presence of characteristic domains in each predicted protein is displayed in column 3. For the protein encoded by *orf20*, three different proteins were retrieved, each with strong statistical significance.



ORF #	Predicted Proteins (size) [Source species]	Characteristic Domains	E-value
ORF13	Competence protein (652 aa) [*Bacteroides sp.* D2]	• Competence superfamily	5.00E-36
ORF14	Putative shikimate 5-dehydrogenase (242 aa) [*Bacteroides vulgatus PC510*]	• Shikimate dehydrogenase superfamily	1.31E-69
ORF15	Hypothetical protein (672 aa) [*Barnesiella intestinihominis YIT 11860*]	• M13 Peptidase family	0

FIG 8 Genetic map of the 25b operon and the characteristics of predicted proteins encoded by each gene. The three genes that make up the 25b operon and the DNA fragment amplified for cloning into pUC19 are drawn to scale. Predicted proteins retrieved from the BLASTp search are listed with E values. The presence of characteristic domains in each predicted protein is displayed in column 3.

deduce the roles of ORF19, ORF21, and ORF22 in biofilm formation based on the available information. ORF20, the site of the original interruption in the mutant 2F2 (Fig. 4), displays significant homology to a conserved hypothetical protein of *Bacteroides* sp. strain 9_1_42FAA. This protein possesses two distinct, well-characterized domains, a predicted lysozyme domain and a peptidoglycan (PG)-binding domain, in its N and C termini, respectively (Fig. 7). This protein was also determined to be highly homologous to a putative secretion-activator protein of *Prevotella multiformis*, a Gram-negative rod isolated from human subgingival plaque (27), and an *N*-acetylmuramidase of *Teredinibacter turnerae*, a cellulolytic Gram-negative bacterium (28) (Fig. 7).

Three genes in operon 25b encode relatively well-characterized proteins, and these proteins are presented, with their characteristic domains, in Fig. 8. The genes grouped in this operon, however, appear to be functionally distinct, because their predicted functions most likely are associated with distinct cellular processes. ORF13, which was disrupted in two independent Tn-marked mutants (7B9 and 4B10) (Fig. 4), is a competence protein (29), whereas ORF14 and ORF15 encode proteins that contain a shikimate dehydrogenase superfamily domain (30) and an M13 peptidase domain (31), respectively (Fig. 8). The transmembrane protein encoded by ORF13 harbors a conserved competence superfamily domain that is essential for the uptake of extracellular DNA (32).

DISCUSSION

Collectively, the commensals that comprise the microbiota of the large bowel have a diverse metabolic capacity that enables mammalian hosts to extract energy from substrates that they would otherwise lack the ability to utilize (33–35). Indeed, most information about bowel commensals relates to their phylogeny and metabolic activities, while the lifestyle of these commensals, notably the mechanisms of intestinal colonization and biofilm formation, attracts less interest. The ability of commensals to maintain a population in a self-regulating community of complex composition, however, probably includes dynamic relationships with surfaces in order to obtain nutrients or to minimize the wash-out effect of peristalsis. Digesta-associated surfaces (i.e., other bacterial cells, plant cell structures, and partly degraded mucus) or mucosa-associated surfaces (i.e., mucus and apical surfaces of epithelial cells) could be important. Little is known of these relationships or the molecular or structural mechanisms that mediate them. Investigation of microbial adhesion in the gut is difficult not only because gut microbes are largely unculturable but also because of the lack of effective means to preserve the intestinal mucus layer, where microbial communities are formed (36, 37). Therefore, an alternative approach is required to explore the basis for the formation of the adhesive community by gut microbiota.

Our functional metagenomic screen, followed by mutational inactivation using a surrogate host, revealed two operons that were associated with enhanced adherence to surfaces. These operons probably originated in unknown *Bacteroides* species. Expression of the heterologous DNA resulted, in turn, in enhanced colonization of the surrogate hosts in the bowels of mice. While obligate anaerobes predominate in the large bowel of adults, facultative anaerobic bacteria such as *E. coli* are numerous in the bowels of infant mice (38). Hence, the attributes encoded by the cloned genes enhanced the colonization ability of the surrogate strain in competition with autochthonous strains of *E. coli*.

Although annotations of the cloned ORFs were obtained, they did not seem to have direct relevance to enhanced adherence to surfaces by bacterial cells. BLASTp alignments may be misleading, because there is a paucity of knowledge relating to mechanisms by which commensal bacteria associate with surfaces in the gut (39). The rationale behind homology-based annotation is that if two sequences have a high degree of similarity, then they have evolved from a common ancestor and should have similar, if not identical, functions. However, with increasing numbers of sequences in databases and considering the effects of gene duplications, which might be followed by divergence of function, the power of homology-based annotation has lessened. Adding to this issue is the problem of errors in annotation which spread misannotations when homology-based approaches are used. Additionally, most of the newly identified proteins do not show significant sequence similarity to experimentally characterized proteins (40). Some genes have multiple annotations reflecting functions of gene products in different experimental systems (41).

Among the genes present in the 10a operon, ORF20, which encodes a lysozyme-like protein, is presumed to play the most crucial role both in adhesion and in intestinal colonization. The biofilm-forming capability of *Lactococcus lactis* was substantially enhanced by PG breaks mediated by PG hydrolase, which also possesses both PG-binding and lysozyme domains (42). An *L. lactis* mutant strain defective in PG hydrolase failed to adhere to

solid surfaces and form biofilm. Moreover, the addition of lysozyme completely reversed these phenotypes (42). A hypothetical protein harboring a glycosyl hydrolase domain positively influences biofilm formation by *Streptococcus mutans* (43). Similarly, DNA release mediated by autolysin, a lysozyme-like protein, can participate in biofilm formation by *Staphylococcus epidermidis* (44). These results demonstrate that peptidoglycan autolysis, a natural process in the synthesis and degradation of the PG molecule (45), may contribute to the bacterial adhesion process.

Involvement of a competence-related protein, such as ORF13 of the 25b operon, in biofilm formation has been clearly established in other bacterial species (46–49). For example, a competence-defective mutant of *Streptococcus mutans* forms biofilms with decreased biomass (48). In addition, *S. mutans* cells growing in a biofilm are more competent than planktonic counterparts (47). Therefore, our findings further support a relationship between the ability to interact with external DNA and biofilm formation. Intestinal epithelium is a rapidly self-renewing tissue (50, 51). Cell turnover rate is suspected to be influenced by the presence of gut microflora, because the rate of *de novo* DNA synthesis as determined by [^{3}H]thymidine incorporation is lower in the colon of germfree mice than in conventional mice (52). Importantly, the exfoliation of epithelial cells into the intestinal lumen releases a steady supply of DNA (53, 54). Furthermore, extracellular DNA is suggested to be a major structural component of microbial biofilms (55). Although this DNA may be recycled or degraded, our results suggest that gut microbes take advantage of the availability of liberated DNA to promote the formation of an adhesive microbial community. Interestingly, DNA release is higher in patients with Crohn's disease, which is characterized by chronic hyperimmune activity (56). It would be of particular interest to investigate whether biofilm formation in the intestines of patients with Crohn's disease is increased in proportion to the level of DNA present.

Although we successfully unveiled two gut microbiota-derived genetic determinants for the colonization of the mouse intestine, our approach has the following limitations. (i) The BAC library used in the current study may represent only a portion of the mouse gut microbiota metagenome. (ii) The *E. coli* DH10B strain may not be the best host to express genes derived from Gram-positive organisms, including ones that belong to the dominant phylum, *Firmicutes*. Therefore, additional experiments using a broader collection of the gut microbiota metagenome and a new host that is better suited to express genes of diverse species will help provide more insights into how gut microbes establish a beneficial community in the host intestine.

We believe it would be beneficial to assemble a databank of gut colonization-associated operons that have demonstrated functions in the bowel. We propose that the proteins encoded by the operons that we have detected be termed *Bacteroides* niche factors (57) until they have been studied in more detail. These operons could be studied in terms of proteomics and molecular modeling so that the function of these relatively unknown genetic determinants can be defined. As shown by our results, BAC metagenomic libraries will clearly have an important role in building this databank.

ACKNOWLEDGMENTS

This work was supported by grants from the National Research Foundation (NRF) of Korea, funded by the Korean government (MEST), no. 2011-0016210 (to S.S.Y.). This work was also supported by a grant from the Korea Healthcare Technology R&D Project, Ministry for Health, Welfare and Family Affairs, A110096 (to S.S.Y.). This work was supported by grants from the NRF of Korea, funded by the Korean government (MEST), no. 2011-355-E00046 and no. 2012-3015650 (to M.Y.Y.).

REFERENCES

1. **Handelsman J.** 2004. Metagenomics: application of genomics to uncultured microorganisms. Microbiol. Mol. Biol. Rev. **68**:669–685.
2. **Flint HJ, Bayer EA, Rincon MT, Lamed R, White BA.** 2008. Polysaccharide utilization by gut bacteria: potential for new insights from genomic analysis. Nat. Rev. Microbiol. **6**:121–131.
3. **Ley RE, Lozupone CA, Hamady M, Knight R, Gordon JI.** 2008. Worlds within worlds: evolution of the vertebrate gut microbiota. Nat. Rev. Microbiol. **6**:776–788.
4. **Martin R, Nauta AJ, Ben Amor K, Knippels LM, Knol J, Garssen J.** 2010. Early life: gut microbiota and immune development in infancy. Benef. Microbes **1**:367–382.
5. **Sjogren YM, Tomicic S, Lundberg A, Bottcher MF, Bjorksten B, Sverremark-Ekstrom E, Jenmalm MC.** 2009. Influence of early gut microbiota on the maturation of childhood mucosal and systemic immune responses. Clin. Exp. Allergy **39**:1842–1851.
6. **Turnbaugh PJ, Ley RE, Mahowald MA, Magrini V, Mardis ER, Gordon JI.** 2006. An obesity-associated gut microbiome with increased capacity for energy harvest. Nature **444**:1027–1031.
7. **Vijay-Kumar M, Aitken JD, Carvalho FA, Cullender TC, Mwangi S, Srinivasan S, Sitaraman SV, Knight R, Ley RE, Gewirtz AT.** 2010. Metabolic syndrome and altered gut microbiota in mice lacking Toll-like receptor 5. Science **328**:228–231.
8. **Goodman AL, Kallstrom G, Faith JJ, Reyes A, Moore A, Dantas G, Gordon JI.** 2011. Extensive personal human gut microbiota culture collections characterized and manipulated in gnotobiotic mice. Proc. Natl. Acad. Sci. U. S. A. **108**:6252–6257.
9. **Claesson MJ, Wang Q, O'Sullivan O, Greene-Diniz R, Cole JR, Ross RP, O'Toole PW.** 2010. Comparison of two next-generation sequencing technologies for resolving highly complex microbiota composition using tandem variable 16S rRNA gene regions. Nucleic Acids Res. **38**:e200.
10. **Beja O, Aravind L, Koonin EV, Suzuki MT, Hadd A, Nguyen LP, Jovanovich SB, Gates CM, Feldman RA, Spudich JL, Spudich EN, DeLong EF.** 2000. Bacterial rhodopsin: evidence for a new type of phototrophy in the sea. Science **289**:1902–1906.
11. **DeLong EF, Beja O.** 2010. The light-driven proton pump proteorhodopsin enhances bacterial survival during tough times. PLoS Biol. **8**:e1000359. doi:10.1371/journal.pbio.1000359.
12. **Zoetendal EG, von Wright A, Vilpponen-Salmela T, Ben-Amor K, Akkermans AD, de Vos WM.** 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.
13. **Walter J, Mangold M, Tannock GW.** 2005. Construction, analysis, and beta-glucanase screening of a bacterial artificial chromosome library from the large-bowel microbiota of mice. Appl. Environ. Microbiol. **71**:2347–2354.
14. **Yoon SS, Hennigan RF, Hilliard GM, Ochsner UA, Parvatiyar K, Kamani MC, Allen HL, DeKievit TR, Gardner PR, Schwab U, Rowe JJ, Iglewski BH, McDermott TR, Mason RP, Wozniak DJ, Hancock RE, Parsek MR, Noah TL, Boucher RC, Hassett DJ.** 2002. Pseudomonas aeruginosa anaerobic respiration in biofilms: relationships to cystic fibrosis pathogenesis. Dev. Cell **3**:593–603.
15. **Ewing B, Green P.** 1998. Base-calling of automated sequencer traces using phred. II. Error probabilities. Genome Res. **8**:186–194.
16. **Delcher AL, Bratke KA, Powers EC, Salzberg SL.** 2007. Identifying bacterial genes and endosymbiont DNA with Glimmer. Bioinformatics **23**:673–679.
17. **Yoon MY, Lee KM, Park Y, Yoon SS.** 2011. Contribution of cell elongation to the biofilm formation of Pseudomonas aeruginosa during anaerobic respiration. PLoS One **6**:e16105. doi:10.1371/journal.pone.0016105.
18. **Crociani J, Grill JP, Huppert M, Ballongue J.** 1995. Adhesion of different bifidobacteria strains to human enterocyte-like Caco-2 cells and comparison with in vivo study. Lett. Appl. Microbiol. **21**:146–148.
19. **Reikvam DH, Erofeev A, Sandvik A, Grcic V, Jahnsen FL, Gaustad P,**

McCoy KD, Macpherson AJ, Meza-Zepeda LA, Johansen FE. 2011. Depletion of murine intestinal microbiota: effects on gut mucosa and epithelial gene expression. PLoS One **6**:e17996. doi:10.1371/journal.pone.0017996.
20. **Lee KM, Go J, Yoon MY, Park Y, Kim SC, Yong DE, Yoon SS.** 2012. Vitamin B12-mediated restoration of defective anaerobic growth leads to reduced biofilm formation in Pseudomonas aeruginosa. Infect. Immun. **80**:1639–1649.
21. **Lee A, Gordon J, Lee CJ, Dubos R.** 1971. The mouse intestinal microflora with emphasis on the strict anaerobes. J. Exp. Med. **133**:339–352.
22. **Binder HJ, Filburn B, Floch M.** 1975. Bile acid inhibition of intestinal anaerobic organisms. Am. J. Clin. Nutr. **28**:119–125.
23. **Floch MH, Binder HJ, Filburn B, Gershengoren W.** 1972. The effect of bile acids on intestinal microflora. Am. J. Clin. Nutr. **25**:1418–1426.
24. **Hoyles L, McCartney AL.** 2009. What do we mean when we refer to Bacteroidetes populations in the human gastrointestinal microbiota? FEMS Microbiol. Lett. **299**:175–183.
25. **Armougom F, Henry M, Vialettes B, Raccah D, Raoult D.** 2009. Monitoring bacterial community of human gut microbiota reveals an increase in Lactobacillus in obese patients and Methanogens in anorexic patients. PLoS One **4**:e7125. doi:10.1371/journal.pone.0007125.
26. **van der Waaij D, Berghuis-de Vries JM, Lekkerkerk L-V.** 1971. Colonization resistance of the digestive tract in conventional and antibiotic-treated mice. J. Hyg. (London) **69**:405–411.
27. **Sakamoto M, Huang Y, Umeda M, Ishikawa I, Benno Y.** 2005. Prevotella multiformis sp. nov., isolated from human subgingival plaque. Int. J. Syst. Evol. Microbiol. **55**:815–819.
28. **Ekborg NA, Morrill W, Burgoyne AM, Li L, Distel DL.** 2007. CelAB, a multifunctional cellulase encoded by Teredinibacter turnerae T7902T, a culturable symbiont isolated from the wood-boring marine bivalve Lyrodus pedicellatus. Appl. Environ. Microbiol. **73**:7785–7788.
29. **Facius D, Meyer TF.** 1993. A novel determinant (comA) essential for natural transformation competence in Neisseria gonorrhoeae and the effect of a comA defect on pilin variation. Mol. Microbiol. **10**:699–712.
30. **Hoch JA, Nester EW.** 1973. Gene-enzyme relationships of aromatic acid biosynthesis in Bacillus subtilis. J. Bacteriol. **116**:59–66.
31. **Bland ND, Pinney JW, Thomas JE, Turner AJ, Isaac RE.** 2008. Bioinformatic analysis of the neprilysin (M13) family of peptidases reveals complex evolutionary and functional relationships. BMC Evol. Biol. **8**:16. doi:10.1186/1471-2148-8-16.
32. **Dubnau D.** 1999. DNA uptake in bacteria. Annu. Rev. Microbiol. **53**:217–244.
33. **Cebra JJ.** 1999. Influences of microbiota on intestinal immune system development. Am. J. Clin. Nutr. **69**:1046S–1051S.
34. **Ley RE.** 2010. Obesity and the human microbiome. Curr. Opin. Gastroenterol. **26**:5–11.
35. **Sekirov I, Russell SL, Antunes LC, Finlay BB.** 2010. Gut microbiota in health and disease. Physiol. Rev. **90**:859–904.
36. **Bollinger RR, Barbas AS, Bush EL, Lin SS, Parker W.** 2007. Biofilms in the normal human large bowel: fact rather than fiction. Gut **56**:1481–1482.
37. **Strugala V, Allen A, Dettmar PW, Pearson JP.** 2003. Colonic mucin: methods of measuring mucus thickness. Proc. Nutr. Soc. **62**:237–243.
38. **Savage DC, Dubos R, Schaedler RW.** 1968. The gastrointestinal epithelium and its autochthonous bacterial flora. J. Exp. Med. **127**:67–76.
39. **Macfarlane S, Macfarlane GT.** 2006. Composition and metabolic activities of bacterial biofilms colonizing food residues in the human gut. Appl. Environ. Microbiol. **72**:6204–6211.
40. **Janga SC, Diaz-Mejia JJ, Moreno-Hagelsieb G.** 2011. Network-based function prediction and interactomics: the case for metabolic enzymes. Metab. Eng. **13**:1–10.
41. **Naumoff DG, Xu Y, Glansdorff N, Labedan B.** 2004. Retrieving sequences of enzymes experimentally characterized but erroneously annotated: the case of the putrescine carbamoyltransferase. BMC Genomics **5**:52. doi:10.1186/1471-2164-5-52.
42. **Mercier C, Durrieu C, Briandet R, Domakova E, Tremblay J, Buist G, Kulakauskas S.** 2002. Positive role of peptidoglycan breaks in lactococcal biofilm formation. Mol. Microbiol. **46**:235–243.
43. **Brown TA, Jr, Ahn SJ, Frank RN, Chen YY, Lemos JA, Burne RA.** 2005. A hypothetical protein of Streptococcus mutans is critical for biofilm formation. Infect. Immun. **73**:3147–3151.
44. **Qin Z, Ou Y, Yang L, Zhu Y, Tolker-Nielsen T, Molin S, Qu D.** 2007. Role of autolysin-mediated DNA release in biofilm formation of Staphylococcus epidermidis. Microbiology **153**:2083–2092.
45. **Shockman GD, Daneo-Moore L, Kariyama R, Massidda O.** 1996. Bacterial walls, peptidoglycan hydrolases, autolysins, and autolysis. Microb. Drug Resist. **2**:95–98.
46. **Lang E, Haugen K, Fleckenstein B, Homberset H, Frye SA, Ambur OH, Tonjum T.** 2009. Identification of neisserial DNA binding components. Microbiology **155**:852–862.
47. **Li YH, Lau PC, Lee JH, Ellen RP, Cvitkovitch DG.** 2001. Natural genetic transformation of Streptococcus mutans growing in biofilms. J. Bacteriol. **183**:897–908.
48. **Li YH, Tang N, Aspiras MB, Lau PC, Lee JH, Ellen RP, Cvitkovitch DG.** 2002. A quorum-sensing signaling system essential for genetic competence in Streptococcus mutans is involved in biofilm formation. J. Bacteriol. **184**:2699–2708.
49. **Trappetti C, Gualdi L, Di Meola L, Jain P, Korir CC, Edmonds P, Iannelli F, Ricci S, Pozzi G, Oggioni MR.** 2011. The impact of the competence quorum sensing system on Streptococcus pneumoniae biofilms varies depending on the experimental model. BMC Microbiol. **11**:75. doi:10.1186/1471-2180-11-75.
50. **Barker N, van Es JH, Kuipers J, Kujala P, van den Born M, Cozijnsen M, Haegebarth A, Korving J, Begthel H, Peters PJ, Clevers H.** 2007. Identification of stem cells in small intestine and colon by marker gene Lgr5. Nature **449**:1003–1007.
51. **van der Flier LG, Clevers H.** 2009. Stem cells, self-renewal, and differentiation in the intestinal epithelium. Annu. Rev. Physiol. **71**:241–260.
52. **Webb P, Chanana AD, Cronkite EP, Laissue JA, Joel DD.** 1980. Comparison of DNA renewal in germ-free and conventional mice using [125I]iododeoxyuridine and [3H]thymidine. Cell Tissue Kinet. **13**:227–237.
53. **Cline WS, Lorenzsonn V, Benz L, Bass P, Olsen WA.** 1976. The effects of sodium ricinoleate on small intestinal function and structure. J. Clin. Invest. **58**:380–390.
54. **Gullikson GW, Cline WS, Lorenzsonn V, Benz L, Olsen WA, Bass P.** 1977. Effects of anionic surfactants on hamster small intestinal membrane structure and function: relationship to surface activity. Gastroenterology **73**:501–511.
55. **Whitchurch CB, Tolker-Nielsen T, Ragas PC, Mattick JS.** 2002. Extracellular DNA required for bacterial biofilm formation. Science **295**:1487.
56. **Casellas F, Guarner F, Antolin M, Rodriguez R, Salas A, Malagelada JR.** 1994. Abnormal leukotriene C4 released by unaffected jejunal mucosa in patients with inactive Crohn's disease. Gut **35**:517–522.
57. **Hill C.** 2012. Virulence or niche factors; what's in a name? J. Bacteriol. **194**:5725–5727.