Complete genome sequence of acetate-producing *Klebsiella pneumoniae* L5-2 isolated from infant feces

Yong-Soo Park^{1,4} · Jisu Kang^{1,3} · Won-Hyong Chung¹ · Mi Young Lim¹ · Myung-Ji Seo² · Young-Do Nam^{1,3} · Jung-Hoon Yoon⁴ · Dong-Ho Seo¹

Received: 15 September 2018 / Accepted: 17 January 2019 / Published online: 15 February 2019
© King Abdulaziz City for Science and Technology 2019

Abstract

Acetate is an important metabolite in infants as it can affect metabolism as well as immune and inflammatory responses. However, there have been no studies on acetate production by *Klebsiella pneumoniae* isolated from infant feces. In this study, we isolated a *K. pneumoniae* strain, L5-2, from infant feces, and we found it produces acetate. The genome of L5-2 consisted of a 5,237,123-bp single chromosome and a 139,211-bp single plasmid. The G + C content was 57.27%. By whole-genome analysis of *K. pneumoniae* L5-2, we identified seven genes related to acetate production (*poxA*, *pta*, *eutD*, *ackA*, *eutP*, *eutQ*, and *adhE*). We confirmed acetate production by *K. pneumoniae* L5-2 by ion chromatography. The aldehyde/alcohol dehydrogenase (*adhE*) activity of *K. pneumoniae* L5-2 was significantly higher than that of the *K. pneumoniae* subsp. *ozaenae* ATCC 11296. Thus, the acetate-producing ability of *K. pneumoniae* L5-2 was influenced by the *adhE* gene. In addition, *K. pneumoniae* L5-2 had significantly less virulence factor-encoding genes than other *K. pneumoniae* strains isolated from humans. In conclusion, *K. pneumoniae* L5-2 isolated from infant feces has less virulence factors and higher *adhE* activity than other *K. pneumoniae* strains.

Keywords *Klebsiella pneumoniae* · Infant · Gut microbiota · Whole-genome sequencing · Short chain fatty acid · Acetate

Yong-Soo Park and Jisu Kang contributed equally to this work.
Jung-Hoon Yoon and Dong-Ho Seo contributed equally to this work.

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s13205-019-1578-y>) contains supplementary material, which is available to authorized users.

✉ Dong-Ho Seo
sdh83@kfri.re.kr
Jung-Hoon Yoon
jhyoon69@skku.edu

- ¹ Research Group of Healthcare, Korea Food Research Institute, 245, Nongsaengmyeong-ro, Iseo-myeon, Wanju-gun, Jeollabuk-do 55365, South Korea
- ² Division of Bioengineering, Incheon National University, Incheon 22012, South Korea
- ³ Department of Food Biotechnology, Korea University of Science and Technology, Daejeon 34113, South Korea
- ⁴ Department of Food Science and Biotechnology, Sungkyunkwan University, Suwon 16419, South Korea

Introduction

The gut microbiota comprises approximately 1000 species and 3.3 million genes, which is 150 times the number of genes in the human genome (Zhu et al. 2010). Although the importance of the gut microbiota and microbial metabolites in the maintenance of host health has been demonstrated (Ríos-Covián et al. 2016), the genetic and functional diversity of the gut microbiota is not fully elucidated (Li et al. 2014). Infant gut microbial communities are easily influenced by various factors, such as the environment, disease, and diet, and are thus more dynamic than those of adults (Matamoros et al. 2013). *Klebsiella* species are commonly found in soils, plants, insects, and human patients (Downie et al. 2013; Jiang et al. 2016). Recently, a study indicated that *Klebsiella* species are present in the gut of healthy infants, with *Klebsiella pneumoniae* being a dominant species (Gómez et al. 2016).

Short-chain fatty acids (SCFAs), also known as volatile fatty acids, are the major metabolites produced by gut microbiota (Ríos-Covián et al. 2016; Rooks and Garrett 2016) through anaerobic fermentation of dietary fibers or

complex polysaccharides (Flint et al. 2008; Trompette et al. 2014). SCFAs can alter the gastrointestinal environment and modulate inflammation and energy balance through interaction with host factors (Read and Holmes 2017). SCFAs produced by the decomposition of prebiotics have been shown to promote the growth of microorganisms beneficial to infant health, such as *Bifidobacteria* and *Lactobacillus* (Groer et al. 2014). Acetate, propionate, and butyrate are representative SCFAs (White 2017), and an increase in acetate levels affects appetite regulation through homeostatic mechanisms (Frost et al. 2014) along with metabolism, immunity, and inflammation through interaction with G-protein-coupled receptor 43 (Maslowski et al. 2009).

We previously isolated *Shigella flexneri*, *Citrobacter youngae*, *Enterococcus gallinarum*, *Enterococcus avium*, *Enterococcus raffinosus*, and *K. pneumoniae* from infant feces (unpublished). Although *K. pneumoniae* isolated from the natural environment is known to produce acetate as a metabolite (Biebl et al. 1998), it is unclear whether the same is true for *K. pneumoniae* isolated from the gut of infants. To address this question and to identify a number of genes involved in acetate production, we performed whole-genome sequencing of *K. pneumoniae* L5-2 from infant feces and comparative genome analysis using a *K. pneumoniae* subsp. *ozaenae* ATCC 11296.

Materials and methods

Isolation of *K. pneumoniae*

Klebsiella pneumoniae L5-2 isolated from Korean infant feces was cultured in 1 l of selective LC medium composed of 10.0 g peptone, 4.0 g meat extract, 3.0 g sodium acetate, 1.0 g KH_2PO_4 , 1.0 g yeast extract, 1.0 g ammonium citrate, 1.0 g skim milk, 0.1 g MgSO_4 , 0.05 g MnSO_4 , 1.0 ml of Tween-80 (Sigma-Aldrich, St. Louis, MO, USA), and 15 g bacteriological agar (Junsei Chemical Co., Tokyo, Japan) (Ravula and Shah 1998). The pH was adjusted to 7.0 ± 0.2 , and the bacteria were anaerobically incubated at 37°C for 48 h. *K. pneumoniae* subsp. *ozaenae* ATCC 11296 (Leibniz Institute DSMZ, Braunschweig, Germany) was used as a representative type strain for comparative analyses (Fierer et al. 2005). Strains L5-2 and ATCC 11296 were preserved in 20% glycerol at -80°C (Kitamoto et al. 2002).

Microbial growth measurement

Strains L5-2 and ATCC 11296 were cultured in 0.5% glucose-supplemented tryptic soy agar composed of 17.0 g/L pancreatic digest of casein, 3.0 g/L papaic digest of soybean, 2.5 g/L dextrose, 5.0 g/L sodium chloride, 2.5 g/L dipotassium phosphate (BD Biosciences, Franklin Lakes,

NJ, USA), and 15 g/L bacteriological agar. The two strains were incubated in tryptic soy broth with glucose at 37°C in a shaking incubator for 24 h. Microbial growth was determined by measuring the $\text{OD}_{600\text{nm}}$ in a 96-well plate using a SpectraMax i3 spectrophotometer (Molecular Devices, Sunnyvale, CA, USA). The experiment was repeated three times.

Analysis of acetate production

One milliliter of each culture was collected at 2-h intervals from 0 to 24 h. The samples were centrifuged at $15,500 \times g$ for 1 min using a Microfuge 20R (Beckman Coulter, Brea, CA, USA). One milliliter of the supernatant was filtered through a $0.45\text{-}\mu\text{m}$ syringe filter (GVS, Zola Predosa, Italy) and was transferred to a vial for analysis by ion chromatography (IC) (Metrohm, Herisau, Switzerland). The run time for each sample was set to 40 min (Lu et al. 2010). Integrations were automated. A Metrosep organic acid 250/7.8 column (Metrohm) was used. The eluent was made up to 1 l with sulfuric acid (0.5 mM) (99.999%; Sigma-Aldrich, St. Louis, MO, USA) and was used after filtering (500-ml bottle top filter, $0.22\text{-}\mu\text{m}$; Corning Inc., Corning, NY, USA) and degassing (Powersonic 510; Hwashin, Yeongcheon, Korea). The liquid flow rate was set to 0.5 ml/min. IC pressure during analysis was maintained between 540 and 550 psi, and the instrument temperature was set to 30°C . Sample analysis was started after 5 min of pump priming and 1 h of stabilization. The acetate concentration was calculated based on peak height because the unknown peak was not completely separated from acetate in IC analysis (Supplementary Fig. S1). The calibration curve for acetate concentration was of quadratic curve type, H (peak height) = $3.245 + 4.054 \times Q$ (acetate amount) $- 0.016 \times Q^2$, the relative standard deviation was 6.554%, and the correlation coefficient was 0.999.

Genome sequencing, assembly, and annotation

Genomic DNA was extracted from strain L5-2 using a QIAmp DNA Mini kit (Qiagen, Hilden, Germany), and DNA quantity and quality were determined with a NanoDrop 2000 spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA) and a Qubit 2.0 fluorometer (Life Technologies, Carlsbad, CA, USA).

Whole-genome sequencing was performed using the PacBio RS II platform. A 10-kb DNA library was constructed and sequenced by single-molecule real-time sequencing technology using P6-C4 chemistry (Pacific Biosciences, Menlo Park CA, USA). Sequence reads were assembled using HGAP 4.0 (Pacific Biosciences) with 5.3 Mb as the genome size option. Functional gene annotation was performed with the National Center for Biotechnology Information (NCBI) Prokaryotic Genome Annotation Pipeline (PGAAP) (Tatusova et al. 2016). Genes involved in

acetate production were identified based on PGAAP annotation, and acetate metabolism pathways were predicted by Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis (Kanehisa et al. 2017). DNA plotter (Carver et al. 2009) was used to generate a genome map and examine genome topology.

Phylogenetic and comparative genomic analyses

Acetate production-related genes identified in *K. pneumoniae* L5-2 were validated by comparisons to the *K. pneumoniae* subsp. *ozaenae* ATCC 11296 genome. Phylogenetic analysis was performed with MEGA 6, and a tree was constructed by the neighbor-joining method with 1000 bootstraps (Tamura et al. 2013). Average nucleotide identity (ANI) values were calculated to evaluate evolutionary distance using pyani version 0.2.7; option “-m ANIb” (Pritchard et al. 2016). The EzBioCloud, an online taxon search tool, was used to identify the 16S rRNA gene sequences of L5-2 and ATCC 11296 (Yoon et al. 2017). Complete or draft genome sequences of ATCC 11296, ATCC 13884, HS11286, TH1, and W14 were downloaded from the NCBI genome database [<http://www.ncbi.nlm.nih.gov/genome/>; accession nos. GCF_000826585 (ATCC 11,296); GCF_000163455 (ATCC13884); GCF_000240185 (HS11286); GCF_001676825 (TH1); GCF_001646625 (W14)].

Functional classification and identification of genes encoding virulence factors

Coding genes were aligned to the Clusters of Orthologous Groups (COG) database using BLASTP (E value cut-off: $1e-5$) and were classified into COG categories (Galperin et al. 2015). Genes encoding virulence factors were searched in the pathogenic bacteria database (Chen et al. 2005, 2016) with the following criteria: minimum protein identity, 80%; minimum coverage, 80%; and minimum alignment length, 50 bp.

Aldehyde dehydrogenase activity assay

Aldehyde dehydrogenase activity was analyzed using the Aldehyde Dehydrogenase Activity Colorimetric Assay kit (Sigma-Aldrich, St. Louis, MO, USA) according to manufacturer's instructions. Briefly, cell lysates were incubated with 50 μ l of lysis buffer, 5 mM NAD^+ , and 5 mM propionaldehyde as a substrate in a 96-well plate at 37 °C. The initial absorbance at 340 nm was measured immediately using a SpectraMax i3 plate reader (Molecular Devices, Sunnyvale, CA, USA). The rate of change in absorbance at 340 nm was measured in triplicates over 30 min. A control reaction in which the substrate was not added was included to monitor the endogenous rate of NAD^+ reduction. Aldehyde dehydrogenase activity was expressed in nmol/ 10^8 cells/min.

Results and discussion

Genome comparison

Sequence analysis of the 16S rRNA gene sequence of L5-2 was carried out using the NCBI and EzBioCloud databases (Kim et al. 2012). *K. pneumoniae* subsp. *ozaenae* showed the highest similarity to L5-2 (99.79%) in the EzBioCloud database (Table 1). Thus, we selected *K. pneumoniae* subsp. *ozaenae* ATCC 11296 as the type strain for comparative analyses. A phylogenetic tree constructed on the basis of 16S rRNA gene sequences confirmed the evolutionary relatedness of L5-2 and *K. pneumoniae* subsp. *ozaenae* ATCC 11296 (Fig. 1a). ANI values ranged from 98.8 to 99% (Fig. 1b).

The complete genome sequence of *K. pneumoniae* L5-2 was obtained with 170-fold coverage. The total genome size was 5,376,334 bp, and the GC content was 57.27%. There were two contigs: a single circular chromosome (5,237,123 bp) and a single circular plasmid (139,211 bp) (Table 2; Fig. 2). In total, 5341 genes were assigned, including 5067 protein-coding genes, 155 pseudogenes, and 124 RNA genes (25 rRNA, 87 tRNA, and 12 non-coding RNA).

Table 1 Similarities based on 16S rRNA gene sequence were calculated using EzBioCloud and NCBI database

Strain	EzBioCloud		NCBI	
	Similarity (%)	Completeness (%)	Identity (%)	Query cover (%)
<i>Klebsiella pneumoniae</i> subsp. <i>ozaenae</i> ATCC 11296	99.79	99.3	99	93
<i>Klebsiella pneumoniae</i> subsp. <i>rhinoscleromatis</i> ATCC 13884	99.59	100	99	92
<i>Klebsiella pneumoniae</i> W14	98.09	93.8	—	—

—, not detected

Fig. 1 Phylogenetic tree of *K. pneumoniae* strains based on 16S rRNA gene sequences (**a**) and average nucleotide identity (**b**)

Table 2 Genome summary of L5-2 and ATCC 11296

Characteristic	L5-2	ATCC 11296
Genome		
Accession	CP025683	GCA_000826585.2
Assembly level	Complete	Scaffold
Sequence category	1 Chromosome, 1 plasmid	23 Scaffolds
Total size	5,376,334	4,944,477
GC content	57.27	57.26
Annotation		
Genes	5341	5014
CDS	5217	4937
Coding genes	5062	4433
Pseudogenes	155	504
RNAs	124	77
rRNAs (5S, 16S, 23S)	25 (9, 8, 8)	3 (1, 1, 1)
tRNAs	87	64
ncRNAs	12	10
Repeat region	0	2
CRISPRs	0	2
misc_feature	0	0

CDS coding sequence, *CRISPR* clustered regularly interspaced short palindromic repeats, *ncRNA* non-coding RNA, *rRNA* ribosomal RNA, *tRNA* transfer RNA

The assembly level of ATCC 11296 was scaffold genome representation.

Comparison of genomic features and functional classification

Protein-coding genes were functionally classified based on COG definitions. L5-2 and ATCC 11286 were assigned 25

out of 26 COG functional codes, excluding “Nuclear structure (Y)” (Table 3). In total, 5091 and 4676 genes in L5-2 and ATCC 11296, respectively, were assigned COG codes, whereas 636 and 838 genes, respectively, were unassigned. Eighteen COG categories of L5-2, including “Signal transduction mechanisms (T)”, “Extracellular structures (W)”, “Intracellular trafficking, secretion, and vesicular transport (U)”, and “Secondary metabolites biosynthesis, transport,

Fig. 2 Complete circular genome map of *K. pneumoniae* L5-2. **a** Chromosome. **b** Plasmid. Track 1 (light blue), forward coding sequence (CDS); track 2 (blue), reverse CDS, track 3 (light purple), rRNA (5S, 16S, and 23S); track 4 (green), tRNA; track 5 (red), clus-

tered regularly interspaced short palindromic repeats (CRISPRs); track 6 (light green and purple), GC content; and track 7 (light green and purple), GC skew

Table 3 Cluster of orthologous group (GOG) functional categories

CODE	COG functional description	L5-2	ATCC 11,296
J	Translation, ribosomal structure, and biogenesis	272	272
A	RNA processing and modification	1	1
K	Transcription	488	421
L	Replication, recombination, and repair	150	175
B	Chromatin structure and dynamics	1	1
D	Cell cycle control, cell division, and chromosome partitioning	51	52
Y	Nuclear structure	0	0
V	Defense mechanisms	122	110
T	Signal transduction mechanisms	238	195
M	Cell wall/membrane/envelope biogenesis	278	256
N	Cell motility	78	50
Z	Cytoskeleton	1	1
W	Extracellular structures	49	27
U	Intracellular trafficking, secretion, and vesicular transport	77	63
O	Posttranslational modification, protein turnover, and chaperones	196	165
X	Mobilome: prophages and transposons	42	128
C	Energy production and conversion	292	262
G	Carbohydrate transport and metabolism	573	512
E	Amino acid transport and metabolism	508	468
F	Nucleotide transport and metabolism	100	96
H	Coenzyme transport and metabolism	242	235
I	Lipid transport and metabolism	160	127
P	Inorganic ion transport and metabolism	354	326
Q	Secondary metabolites biosynthesis, transport, and catabolism	136	111
R	General function prediction only	431	388
S	Function unknown	251	234
—	NA	636	838

Fig. 4 Growth curves and acetate production by *K. pneumoniae*. **a** Comparison of growth curves of *K. pneumoniae* L5-2 and *K. pneumoniae* subsp. *ozaenae* ATCC 11296 from 0 to 24 h of culture. **b**

Comparison of acetate production in *K. pneumoniae* L5-2 and *K. pneumoniae* subsp. *ozaenae* ATCC 11296 from 0 to 24 h of culture

Genes and pathways involved in acetate production

Genes related to acetate production and metabolism were searched based on previous studies (Lama et al. 2017; Garsin 2010) and using the KEGG database (Fig. 3) (Kanehisa and Goto 2000). We focused on genes encoding four enzymes involved in acetate production, i.e., pyruvate dehydrogenase (*poxA*), phosphate acetyltransferase (*pta* and *eutD*), acetate kinase (*ackA*, *eutP*, and *eutQ*), and aldehyde/alcohol dehydrogenase (*adhE*) (Table 4 and Supplementary Table S1). L5-2 and ATCC 11296 harbored one *poxA* gene each; the gene product regulates pyruvate oxidase (*poxB*), which directly converts pyruvate to acetate in glycolysis (Wolfe 2005). A comparison of the growth curves revealed that L5-2 consistently grew better than ATCC 11296 (Fig. 4a). Therefore, we predicted that acetate production by L5-2 would be higher than that by ATCC 11296. Acetate was confirmed to be the only SCFA produced by *K. pneumoniae* by IC analysis (Fig. S1). However, this was partially contradicted by the IC results: acetate production was higher in L5-2 than in ATCC 11296 up to 10 h, whereas the reverse was observed thereafter (Fig. 4b). This may be because L5-2 has an additional copy of *adhE*. Initially (< 10 h), L5-2 produces acetate using glucose, whereas later, it might produce ethanol from acetate by *adhE*. Therefore, it can be assumed that ATCC 11296 has higher acetate production than L5-2 after 10 h of culture. Previous study showed that increased ethanol production through metabolism reduces acetate production (Yazdani and Gonzalez 2008). Our results suggest that the copy number of *adhE* involved in the conversion of acetaldehyde to ethanol affects acetate production. Moreover, the intracellular *adhE* activity in L5-2 (103.33 ± 4.95 U) was significantly higher than that of ATCC 11296 (1.33 ± 0.19 U) (Fig. 5).

The *poxA* gene of ATCC 11296 is a confirmed pseudogene and may be affected by chromosome-level scaffolding (Table 4). Except for *adhE*, gene copy numbers were similar

Fig. 5 Aldehyde dehydrogenase activity of *K. pneumoniae* L5-2 and *K. pneumoniae* subsp. *ozaenae* ATCC 11296

between the two strains. *adhE* encodes a bifunctional alcohol and aldehyde dehydrogenase that is important for the generation of ethanol from acetyl-CoA (Lo et al. 2015); a previous study showed that acetate production of *adhE*⁺ mutants (*Caldicellulosiruptor bescii* and *Clostridium thermocellum*) was decreased (Scully and Orlygsson 2017; Biswas et al. 2014). These findings also explain why L5-2 produces less acetate than ATCC 11296. L5-2 has fewer *adhE* genes than *K. pneumoniae* strains isolated from humans. L5-2 might have the highest acetate production among *K. pneumoniae* from humans, suggesting that it might create a beneficial environment in the infant's intestines.

Comparison of virulence factors

Klebsiella pneumoniae causes hospital-acquired infections and accounts for a large number of cases of hepatitis, pneumonia, sepsis, and soft tissue infections (Podschun and Ullmann 1998). We found that L5-2 has fewer genes encoding virulence factors than did ATCC 11296; for instance,

Table 5 Number of virulence factor-encoding genes in various strains

Virulence gene ID	L5-2	ATCC 11296	HS11286	ATCC 13884	TH1	W14
Type 1 fimbriae (VF0221)	5	0	5	4	5	5
LOS (CVF494)	1	1	1	1	1	1
OmpA (VF0236)	1	1	1	1	1	1
Enterobactin (VF0228)	8	8	8	0	8	8
Enterobactin (IA019)	1	1	1	0	1	1
ECP (VF0404)	6	0	6	6	6	6
Aerobactin (VF0123)	0	2	0	2	0	0
Yersiniabactin (VF0136)	0	2	11	12	11	11
Aerobactin (VF0229)	0	3	0	3	0	0

L5-2, *K. pneumoniae* L5-2; HS 11286, *K. pneumoniae* subsp. *pneumoniae* HS11286; TH1, *K. pneumoniae* TH1; W14, *K. pneumoniae* W14; ATCC 11296, *K. pneumoniae* subsp. *ozaenae* ATCC 11296; ATCC 13884, *K. pneumoniae* subsp. *rhinoscleromatis* ATCC 13884

aerobactin (VF0123), yersiniabactin (VF0136), and aerobactin (VF0229) were found only in ATCC 11296 (Supplementary Table S2). However, five genes encoding type 1 fimbriae (VF0221) and six encoding *Escherichia coli* common pilus (ECP; VF0404) were found in the L5-2 genome. These differences can be explained by the source of each strain: ATCC 11296 was isolated from soil, whereas L5-2 was isolated from infant feces. Interestingly, L5-2 has very few virulence factors compared to other *K. pneumoniae* strains from humans (Table 5). This suggests that L5-2 does not act as a pathogen in infancy.

Type 1 fimbriae mediate and promote the adhesion of bacterial cells to the host (Khanal et al. 2015; Vizcarra et al. 2016). ECP also is an adherence factor (Rendón et al. 2007). The fact that genes encoding these factors are present in the *K. pneumoniae* L5-2 genome suggests an evolutionary adaptation that allows the microorganism to adhere to host intestinal epithelial cells (Table 5).

Acetate produced by gut microbiota has a variety of beneficial health effects. *K. pneumoniae* L5-2 isolated from infant's intestine was confirmed to produce acetate through fermentation. In addition, we identified several genes that may be involved in acetate production and metabolism. Our results provide a useful resource for future investigations into the role of the gut microbiota in infant health.

Nucleotide sequence accession number

The complete genome sequence has been deposited in GenBank under accession numbers CP025683-CP025684.

Acknowledgements This research was supported by Main Research Programs (E0170602-02) of the Korea Food Research Institute funded by the Ministry of Science and Information and Communications Technologies.

Author contributions D-HS designed and coordinated all the experiments; Y-SP performed the experiments, analyzed the genome data and wrote manuscript; SK performed the sequence assembly and genomic analysis and wrote manuscript; W-HC and MYL performed gene prediction and gene annotation; M-JS and Y-DN checked and edited the manuscript. J-HY measured aldehyde dehydrogenase activity. All authors have read and approved the final manuscript.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest in the publication.

References

- Biebl H, Zeng AP, Menzel K, Deckwer WD (1998) Fermentation of glycerol to 1,3-propanediol and 2,3-butanediol by *Klebsiella pneumoniae*. Appl Microbiol Biotechnol 50:24–29
- Biswas R, Prabhu S, Lynd LR, Guss AM (2014) Increase in ethanol yield via elimination of lactate production in an ethanol-tolerant mutant of *Clostridium thermocellum*. PLoS One 9:e86389
- Carver T, Thomson N, Bleasby A, Berriman M, Parkhill J (2009) DNAPlotter: circular and linear interactive genome visualization. Bioinformatics 25:119–120
- Chen L, Yang J, Yu J, Yao Z, Sun Y et al (2005) VFDB: a reference database for bacterial virulence factors. Nucleic Acids Res 33:D325–D328
- Chen L, Zheng D, Liu B, Yang J, Jin Q (2016) VFDB 2016: hierarchical and refined dataset for big data analysis-10 years on. Nucleic Acids Res 44:D694–D697
- Downie L, Armiento R, Subhi R, Kelly J, Clifford V, Duke T (2013) Community-acquired neonatal and infant sepsis in developing countries: efficacy of WHO's currently recommended antibiotics—systematic review and meta-analysis. Arch Dis Child 98:146–154
- Fierer N, Jackson JA, Vilgalys R, Jackson RB (2005) Assessment of soil microbial community structure by use of taxon-specific quantitative PCR assays. Appl Environ Microbiol 71:4117–4120
- 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
- Frost G, Sleeth ML, Sahuri-Arisoylu M, Lizarbe B, Cerdan S et al (2014) The short-chain fatty acid acetate reduces appetite via a central homeostatic mechanism. Nat Commun 5:3611

- Galperin MY, Makarova KS, Wolf YI, Koonin EV (2015) Expanded microbial genome coverage and improved protein family annotation in the COG database. *Nucleic Acids Res* 43:D261–D269
- Garsin DA (2010) Ethanolamine utilization in bacterial pathogens: roles and regulation. *Nat Rev Microbiol* 8:290–295
- Gómez M, Moles L, Melgar A, Ureta N, Bustos G et al (2016) Early gut colonization of preterm infants: effect of enteral feeding tubes. *J Pediatr Gastroenterol Nutr* 62:893–900
- Groer MW, Luciano AA, Dishaw LJ, Ashmeade TL, Miller E, Gilbert JA (2014) Development of the preterm infant gut microbiome: a research priority. *Microbiome* 2:38
- Jiang SF, Liu Y, Xiao MY, Ruan CJ, Lu ZJ (2016) Draft genome sequence of *Klebsiella variicola* strain KV321 isolated from rhizosphere soil of *Pisolithus tinctorius-Eucalyptus* Mycorrhiza. *Genome Announc* 4:e00676–e00616
- Kanehisa M, Goto S (2000) KEGG: Kyoto Encyclopedia of genes and genomes. *Nucleic Acids Res* 28:27–30
- Kanehisa M, Furumichi M, Tanabe M, Sato Y, Morishima K (2017) KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res* 45:D353–D361
- Khanal M, Larssonneur F, Raks V, Barras A, Baumann JS et al (2015) Inhibition of type 1 fimbriae-mediated *Escherichia coli* adhesion and biofilm formation by trimeric cluster thiomannosides conjugated to diamond nanoparticles. *Nanoscale* 7:2325–2335
- Kim OS, Cho YJ, Lee K, Yoon H, Kim M et al (2012) Introducing EzTaxon-e: a prokaryotic 16S rRNA gene sequence database with phylotypes that represent uncultured species. *Int J Syst Evol Microbiol* 62:716–721
- Kitamoto Y, Suzuki A, Shimada S, Yamanaka K (2002) A new method for the preservation of fungus stock cultures by deep-freezing. *Mycoscience* 43:0143–0149
- Lama S, Seol E, Park S (2017) Metabolic engineering of *Klebsiella pneumoniae* J2B for the production of 1,3-propanediol from glucose. *Bioresour Technol* 245:1542–1550
- Li J, Jia H, Cai X, Zhong H, Feng Q et al (2014) An integrated catalog of reference genes in the human gut microbiome. *Nat Biotechnol* 32:834–841
- Lo J, Zheng T, Hon S, Olson DG, Lynd LR (2015) The bifunctional alcohol and aldehyde dehydrogenase gene, *adhE*, is necessary for ethanol production in *Clostridium thermocellum* and *Thermoanaerobacterium saccharolyticum*. *J Bacteriol* 197:1386–1393
- Lu W, Clasquin MF, Melamud E, Amador-Noguez D, Caudy AA et al (2010) Metabolomic analysis via reversed-phase ion-pairing liquid chromatography coupled to a stand alone orbitrap mass spectrometer. *Anal Chem* 82:3212–3221
- Maslowski KM, Vieira AT, Ng A, Kranich J, Sierro F et al (2009) Regulation of inflammatory responses by gut microbiota and chemottractant receptor GPR43. *Nature* 461:1282–1286
- Matamoros S, Gras-Leguen C, Le Vacon F, Potel G, de La Cochetiere MF (2013) Development of intestinal microbiota in infants and its impact on health. *Trends Microbiol* 21:167–173
- Podschun R, Ullmann U (1998) *Klebsiella* spp. as nosocomial pathogens: epidemiology, taxonomy, typing methods, and pathogenicity factors. *Clin Microbiol Rev* 11:589–603
- Pritchard L, Glover RH, Humphris S, Elphinstone JG, Toth IK (2016) Genomics and taxonomy in diagnostics for food security: soft-rotting enterobacterial plant pathogens. *Anal Methods* 8:12–24
- Ravula RR, Shah NP (1998) Selective enumeration of *Lactobacillus casei* from yogurts and fermented milk drinks. *Biotechnol Tech* 12:819–822
- Read MN, Holmes AJ (2017) Towards an integrative understanding of diet–host–gut microbiome interactions. *Front Immunol* 8:538
- Rendón MA, Saldaña Z, Erdem AL, Monteiro-Neto V, Vázquez A et al (2007) Commensal and pathogenic *Escherichia coli* use a common pilus adherence factor for epithelial cell colonization. *Proc Natl Acad Sci USA* 104:10637–10642
- Ríos-Covián D, Ruas-Madiedo P, Margolles A, Gueimonde M, de los CG, Reyes-Gavilán CG et al (2016) Intestinal short chain fatty acids and their link with diet and human health. *Front Microbiol* 7:185
- Rooks MG, Garrett WS (2016) Gut microbiota, metabolites and host immunity. *Nat Rev Immunol* 16:341–352
- Scully SM, Orlygsson J (2017) Engineering of microorganisms for the production of chemicals and biofuels from renewable resources. In: Gosset G (ed) Recent advances in genetic engineering of thermophilic ethanol producing bacteria. Springer International Publishing, Berlin, pp 1–29
- Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013) MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol Biol Evol* 30:2725–2729
- Tatusova T, Di Cuccio M, Badretdin A, Chetvernin V, Nawrocki EP et al (2016) NCBI prokaryotic genome annotation pipeline. *Nucleic Acids Res* 44:6614–6624
- Trompette A, Gollwitzer ES, Yadava K, Sichelstiel AK, Sprenger N et al (2014) Gut microbiota metabolism of dietary fiber influences allergic airway disease and hematopoiesis. *Nat Med* 20:159–166
- Vizcarra A, Hosseini V, Kollmannsberger P, Meier S, Weber SS et al (2016) How type 1 fimbriae help *Escherichia coli* to evade extracellular antibiotics. *Sci Rep* 6:18109
- White K (2017) Effects of short-chain fatty acids propionate, acetate, butyrate on the growth of *Clostridium difficile* in co-culture with an *Escherichia coli* *atoE* mutant (Master's thesis)
- Wolfe AJ (2005) The acetate switch. *Microbiol Mol Biol Rev* 69:12–50
- Yazdani SS, Gonzalez R (2008) Engineering *Escherichia coli* for the efficient conversion of glycerol to ethanol and co-products. *Metab Eng* 10:340–351
- Yoon SH, Ha SM, Kwon S, Lim J, Kim Y et al (2017) Introducing EzBioCloud: a taxonomically united database of 16S rRNA gene sequences and whole-genome assemblies. *Int J Syst Evol Microbiol* 67:1613–1617
- Zhu B, Wang X, Li L (2010) Human gut microbiome: the second genome of human body. *Protein Cell* 1:1718–1725