AMERICAN SOCIETY FOR MICROBIOLOGY | genomeAnnouncements™

PROKARYOTES

Complete Genome Sequence of *Collinsella aerofaciens* Isolated from the Gut of a Healthy Indian Subject

Satyabrata Bag, Tarini Shankar Ghosh, Bhabatosh Das

Molecular Genetics Laboratory, Centre for Human Microbial Ecology, Translational Health Science and Technology Institute, NCR Biotech Science Cluster, Faridabad, India

ABSTRACT *Collinsella aerofaciens*, a rod-shaped nonmotile obligate anaerobe, is the most abundant actinobacterium in the gastrointestinal tract of healthy humans. An altered abundance of *C. aerofaciens* may be linked with several health disorders, including irritable bowel syndrome. In the present study, we report the complete genome sequence of *C. aerofaciens* strain indica.

Received 31 October 2017 **Accepted** 1 November 2017 **Published** 22 November 2017

Citation Bag S, Ghosh TS, Das B. 2017. Complete genome sequence of *Collinsella aerofaciens* isolated from the gut of a healthy Indian subject. Genome Announc 5:e01361-17. https://doi.org/10.1128/genomeA.01361-17.

Address correspondence to Bhabatosh Das, bhabatosh@thsti.res.in.

S.B. and T.S.G. contributed equally to this article.

The bacterium *Collinsella aerofaciens* is well known for its ability to ferment a range of plant and animal origin carbohydrates and for producing H_2, ethanol, short-chain fatty acids, and lactate in the human colon (1). *C. aerofaciens* is the major utilizer of lactose in the human colon. Several studies demonstrated that *Collinsella* and *Bifidobacterium* can modify the host bile acids to modulate the virulence and pathogenicity of enteric pathogens (2). Recently, it was reported that an altered abundance of *Collinsella* may also influence host plasma cholesterol levels (3). To understand the importance of *C. aerofaciens* in health and disease, it is important to explore its genomic repertoire and identify functions that potentially influence host physiology.

In the present study, *C. aerofaciens* strain indica was isolated from a fecal sample of a healthy adult Indian subject. An approximately 500-mg fecal sample was resuspended in 1 ml of phosphate-buffered saline (PBS), diluted serially in the same buffer, and plated on a prereduced Trypticase soy agar plate (pH 7.3) supplemented with 5% (vol/vol) defibrinated sheep blood. Bacterial cells were spread over the surface of the plates using four or five glass beads (3.00 mm). Plates were incubated for 48 h at 37°C in an anaerobic workstation (Whitley DG250) filled with 80% N_2, 10% CO_2, and 10% H_2. A single colony of *C. aerofaciens* was grown in 5 ml of tryptic soy broth (TSB) for 48 h. Growth of the cells in TSB was monitored by spectrophotometer. *C. aerofaciens* cells were harvested from 2 ml of culture by centrifugation (8,000 × *g* for 3 min), and the genomic DNA was extracted by the THSTI method (4).

The whole-genome sequencing of *C. aerofaciens* was carried out by utilizing 2 different DNA sequencing platforms, those from Illumina (HiSeq 2500 system) and Oxford Nanopore Technologies (MinION). The SPAdes tool was used to assemble error-corrected long Nanopore reads and Illumina reads, which generated a single contig. The assembled complete genome sequence of *C. aerofaciens* was evaluated by Sanger sequencing. The complete genome of *C. aerofaciens* strain indica is 2,306,349 bp in length, with 60.1% GC content.

The analysis of the 2.30-Mb genome sequence of *C. aerofaciens* identified 1,995 genes, including 75 RNA-encoding genes. The genome has 276 subsystems and is enriched with protein (231 open reading frames [ORFs]), carbohydrate (226 ORFs), amino acid (200 ORFs), cofactor, and vitamin (102 ORFs) metabolic functions. The *C. aerofaciens* genome is also enriched with fatty acid, lipid, and isoprenoid metabolic functions (56 ORFs). However, the *C. aerofaciens* genome has several antibiotic resistance genes, such as β-lactamase, tetracycline resistance and ribosome protection

functions, and multidrug resistance efflux pumps, but no gene was detected for the subcategory of virulence, pathogenicity, and disease development. The complete genome sequence of *C. aerofaciens* strain indica will contribute to a better understanding of the biology of the commensal and the molecular basis of its dominance in the gut of Indian subjects.

Accession number(s). The whole-genome shotgun project has been deposited at DDBJ/ENA/GenBank under the accession number CP024160. The version described in this paper is CP024160.1.

ACKNOWLEDGMENT

The present study is supported by the Department of Biotechnology, Government of India (grant BT/MB/THSTI/HMC-SFC/2011).

REFERENCES

1. Kageyama A, Benno Y, Nakase T. 1999. Phylogenetic and phenotypic evidence for the transfer of *Eubacterium aerofaciens* to the genus *Collinsella* as *Collinsella aerofaciens* gen. nov., comb. nov. Int J Syst Bacteriol 49:557–565. https://doi.org/10.1099/00207713-49-2-557.
2. Rajilić-Stojanović M, De Vos WM. 2014. The first 1000 cultured species of the human gastrointestinal microbiota. FEMS Microbiol Rev 38:996–1047. https://doi.org/10.1111/1574-6976.12075.
3. Lahti L, Salonen A, Kekkonen RA, Salojärvi J, Jalanka-Tuovinen J, Palva A, Orešič M, de Vos WM. 2013. Associations between the human intestinal microbiota, *Lactobacillus rhamnosus* GG and serum lipids indicated by integrated analysis of high-throughput profiling data. PeerJ 1:e32. https://doi.org/10.7717/peerj.32.
4. Bag S, Saha B, Mehta O, Anbumani D, Kumar N, Dayal M, Pant A, Kumar P, Saxena S, Allin KH, Hansen T, Arumugam M, Vestergaard H, Pedersen O, Pereira V, Abraham P, Tripathi R, Wadhwa N, Bhatnagar S, Prakash VG, Radha V, Anjana RM, Mohan V, Takeda K, Kurakawa T, Nair GB, Das B. 2016. An improved method for high quality metagenomics DNA extraction from human and environmental samples. Sci Rep 6:26775. https://doi.org/10.1038/srep26775.