Draft genome sequence of *Fusobacterium animalis* KCOM 1280 isolated from a human subgingival plaque of periodontitis lesion

Soon-Nang Park[1†], Yun Kyong Lim[1†], Ja Young Shin[2], Hanseong Roh[2], and Joong-Ki Kook[1,3*]

[1]*Korean Collection for Oral Microbiology and Department of Oral Biochemistry, School of Dentistry, Chosun University, Gwangju 61452, Republic of Korea*
[2]*Macrogen Inc., Seoul 08511, Republic of Korea*
[3]*Oral Biology Research Institute, Chosun University, Gwangju 61452, Republic of Korea*

사람 치주염 병소의 치은연하지면세균막에서 분리된 *Fusobacterium animalis* KCOM 1280의 유전체 염기서열 해독

박순낭[1†] · 임윤경[1†] · 신자영[2] · 노한성[2] · 국중기[1,3*]
[1]조선대학교 치과대학 구강생화학교실 및 한국구강미생물자원은행, [2]마크로젠, [3]조선대학교 구강생물학연구소

Fusobacterium animalis (formerly *Fusobacterium nucleatum* subsp. *animalis*) is a Gram-negative, anaerobic, and filament-shaped bacterium. *F. animalis* may be a part of normal flora and a periodontopathogen of human oral cavity. *F. animalis* KCOM 1280 (= ChDC F318) was isolated from a human periodontitis lesion. In this report, we present the draft genome sequence of *F. animalis* KCOM 1280.

Keywords: *Fusobacterium animalis*, human, periodontitis

Fusobacterium nucleatum subsp. *animalis* was classified as one of four or five subspecies of *Fusobacterium nucleatum* by DNA-DNA hybridization (DDH) and electrophoretic patterns of glutamate dehydrogenase and 2-oxoglutarate reductase (Gharbia and Shah, 1990, 1992). Recently, *F. nucleatum* subsp. *animalis* was reclassified as *Fusobacterium animalis* based on the average nucleotide identity and genome-to-genome distance analyses (Kook *et al.*, 2017). *F. animalis* is a Gram-negative,

anaerobic, and filament-shaped bacterium (Gharbia and Shah, 1990). *F. animalis* may be a commensal species of human oral cavity. *F. animalis* KCOM 1280 (= ChDC F318) was isolated from a human periodontitis lesion. In this report, we present the draft genome sequence of *F. animalis* KCOM 1280.

The *F. animalis* KCOM 1280 was grown on brain heart infusion (BHI, Difco Laboratories) medium supplemented with 0.5% yeast extract, 0.05% cysteine HCl-H_2O, 0.5 mg/ml of hemin, 2 µg/ml of vitamin K_1, and 5% sheep blood in an anaerobic chamber (Model Bactron I, Sheldon Manufacturing Inc.) maintaining with a gas mixture of 10% H_2, 5% CO_2, and 85% N_2 (Park *et al.*, 2013). The bacterial genomic DNA was prepared as described previously and DNA concentration was determined by the Epoch™ Microplate Spectrophotometer (BioTek Instruments Inc.) at wavelengths of 260 and 280 nm (Cho *et al.*, 2015).

The genomic DNA of *F. animalis* KCOM 1280 was sequenced using the Illumina Hiseq 2000 platform by Macrogen Inc. The library of 5 kb mate-pair was sequenced which reached coverage of 2,056×. The *de novo* assembly was performed by ALLPATHS-LG (Gnerre *et al.*, 2011) and SSPACE (Boetzer

---

[†]These authors contributed equally to this work.
*For correspondence. E-mail: jkkook@chosun.ac.kr;
Tel.: +82-62-230-6877; Fax: +82-62-224-3706

and Pirovano, 2014). All gaps among the scaffolds were filled by GapCloser (Luo *et al*., 2012; http://sourceforge.net/projects/soapdenovo2/files/GapCloser). And we confirmed the scaffolds were placed at gaps on the largest scaffold by dot plot analysis. Finally, the assembly was polished by iCORN2 (Otto *et al*., 2010). Genome annotation was conducted by the NCBI Prokaryotic Genome Annotation Pipeline (PGAP) (https://www.ncbi.nlm.nih.gov/ genome/annotation_prok/).

The complete genome of *F. animalis* KCOM 1280 was 2,674,036 bp in length and had a G + C content of 26.9% (Table 1). A total of 2,399 protein-coding sequences (CDSs), 14 rRNAs, and 53 tRNAs were annotated (Table 1). The genome sequence contained virulence factors such as haemagglutinin, hemolysin transporter protein ShlB, hemolysin A, protease 3 precursor, FtsH protease regulator HflK, metalloprotease LoiP/YpwA, ATP-dependent zinc metalloprotease FtsH, protease HtpX, protease 4, papain family cysteine protease, putative protease YhbU precursor, metalloprotease LoiP precursor, putative metalloprotease YpwA, sialic acid-binding periplasmic protein SiaP precursor, sialic acid TRAP transporter permease protein SiaT, outer membrane porin F, multidrug resistance protein MexA/MdtC/MdtK/NorM, multidrug export ATP-binding/permease protein, multidrug export protein MepA, multidrug-efflux transporter 1 regulator, putative multidrug export ATP-binding/permease protein, macrolide export protein MacA, macrolide export ATP-binding/permease protein MacB, antitoxin ChpS/YwqK/YefM/RelB/MazE, toxin YoeB, toxin zeta, plasmid encoded toxin Txe, outer membrane porin F precursor, and transport protein TonB. The genome contained phage genes such as phage portal protein, phage capsid family protein, phage tail sheath protein,

phage-like element PBSX protein XkdM, phage-related minor tail protein, putative prophage phiRv2 integrase, phage tail protein (Tail_P2_I), phage P2 GpU, phage late control gene D protein (GPD), phage tail tube protein FII, phage major capsid protein E, phage terminase large subunit (GpA), N-terminal phage replisome organiser (Phage_rep_org_N), and SPBc2 prophage-derived glycosyltransferase SunS. The complete genome encodes proteins for the biofilm formation, autoinducer-2 (AI-2) modifying protein LsrG, autoinducer 2 sensor kinase/phosphatase LuxQ, and glycosyltransferase family 28 C-terminal domain. It also contained type II secretion system protein D/E/F, type IV secretion system protein virB4/virB9/ virB10/virB11/PtlG, preprotein translocase subunit YajC, and protein translocase subunit SecA/SecD/SecE/ SecY.

The genome also contained the oxidative stress-response genes such as anaerobic nitric oxide reductase flavorubredoxin, anaerobic sulfite reductase subunit A/B/C, thioredoxin reductase, glutaredoxin, peptide methionine sulfoxide reductase MsrB, thiol-disulfide oxidoreductase ResA, nitroreductase A, pyruvate-flavodoxin oxidoreductase, rubrerythrin (conferring superoxide dismutase-like activity), and NADH oxidase. The genome contained the three two-component systems (PdtaS/PdtaR, YpdA/YpdB, and YehU/putative response regulatory protein).

The *F. animalis* KCOM 1280 strain was deposited in the Korean Collection for Oral Microbiology.

Nucleotide sequence accession number

This whole Genome sequence has been deposited at DDBJ/ENA/GenBank under the accession NJGJ00000000. The version described in this paper is version NJGJ01000000.

Table 1. Genome features of *Fusobacterium animalis* KCOM 1280

Attribute	Value
Genome size (bp)	2,674,036
GC content (%)	26.9
Number of contig	4
Number of total gene	2,610
Number of protein-coding gene	2,399
Number of tRNA	53
Number of complete rRNA (5S, 16S, 23S)	14 (4, 5, 5)
Number of partial rRNA (23S)	1
Number of ncRNA	3
Number of pseudogene	141

적 요

Fusobacterium animalis (예전에 *Fusobacterium nucleatum* subsp. *animalis*으로 알려짐)는 그람 음성이면서, 혐기성 및 선형의 세균이다. *F. animalis*는 사람 구강 내 정상 세균총의 하나이며 치주질환원인균이라 여겨지고 있다. *F. animalis* KCOM 1280 (= ChDC F318) 균주는 사람 치주질환 병소에서 분리되었다. 본 논문에서 *F. animalis* KCOM 1280 균주 유전체 염기서열을 해독하여 보고하고자 한다.

Acknowledgements

References

Boetzer, M. and Pirovano, W. 2014. SSPACE-LongRead: scaffolding bacterial draft genomes using long read sequence information. *BMC Bioinformatics* **15**, 211.

Cho, E., Park, S.N., Lim, Y.K., Shin, Y., Paek, J., Hwang, C.H., Chang, Y.H., and Kook, J.K. 2015. *Fusobacterium hwasookii* sp. nov., isolated from a human periodontitis lesion. *Curr. Microbiol.* **70**, 169–175.

Gharbia, S.E. and Shah, H.N. 1990. Heterogeneity within *Fusobacterium nucleatum*, proposal of four subspecies. *Lett. Appl. Microbiol.* **10**, 105–108.

Gharbia, S.E. and Shah, H.N. 1992. *Fusobacterium nucleatum* subsp. *fusiforme* subsp. nov. and *Fusobacterium nucleatum* subsp. *animalis* subsp. nov. as additional subspecies within *Fusobacterium nucleatum*. *Int. J. Syst. Bacteriol.* **42**, 296–298.

Gnerre, S., Maccallum, I., Przybylski, D., Ribeiro, F.J., Burton, J.N., Walker, B.J., Sharpe, T., Hall, G., Shea, T.P., Sykes, S., *et al.* 2011. High-quality draft assemblies of mammalian genomes from massively parallel sequence data. *Proc. Natl. Acad. Sci. USA* **108**, 1513–1518.

Kook, J.K., Park, S.N., Lim, Y.K., Cho, E., Jo, E., Roh, H., Shin, Y., Paek, J., Kim, H.S., Kim, H., *et al.* 2017. Genome-based reclassification of *Fusobacterium nucleatum* subspecies at the species level. *Curr. Microbiol.* **74**, 1137–1147.

Luo, R., Liu, B., Xie, Y., Li, Z., Huang, W., Yuan, J., He, G., Chen, Y., Pan, Q., Liu, Y., *et al.* 2012. SOAPdenovo2: an empirically improved memory-efficient short-read *de novo* assembler. *Gigascience.* **1**, 18. Erratum in: *Gigascience.* 2015. **4**, 30.

Otto, T.D., Sanders, M., Berriman, M., and Newbold, C. 2010. Iterative correction of reference nucleotides (iCORN) using second generation sequencing technology. *Bioinformatics* **26**, 1704–1707.

Park, S.N., Lim, Y.K., and Kook, J.K. 2013. Development of quantitative real-time PCR primers for detecting 42 oral bacterial species. *Arch. Microbiol.* **195**, 473–482.