

Complete genome sequence of *Neisseria* sp. KEM232 isolated from a human smooth surface caries

Eun Mi Kim¹ and Chi Nam Seong^{2*}

¹Department of Dental Hygiene, Gwangju Health University, Gwangju 62287, Republic of Korea

²Department of Biology, College of Life Science and Natural Resources, Suncheon National University, Suncheon 57922, Republic of Korea

사람 평활면 치아우식에서 분리한 *Neisseria* sp. KEM232 균주의 유전체 서열 분석

김은미¹ · 성치남^{2*}

¹광주보건대학교 치위생과, ²순천대학교 생명산업과학대학 생물학과

(Received January 31, 2018; Revised February 5, 2018; Accepted February 5, 2018)

We sequenced the genome of the *Neisseria* sp. KEM232 isolated from the smooth surface caries of human cavity of a 7-year old male in Republic of Korea by using the standard dilution plating technique. The genome comprises a single circular 2,371,912 bp chromosome with a G + C content of 58.5%, 2,210 protein-coding genes, 108 pseudo genes, 51 RNA genes, and one CRISPR array. Based on the 16S rRNA gene sequence similarity and average nucleotide identity, the strain KEM232 is most closely related to *Neisseria bacilliformis*.

Keywords: *Neisseria* sp. KEM232, genome sequence, smooth surface caries

The genus *Neisseria* was proposed in 1885 by Trevisan with type species *Neisseria gonorrhoeae* and approved in 1980 as the type genus of the family *Neisseriaceae* (Skerman *et al.*, 1980). The members of the genus *Neisseria* are Gram-stain-negative and non-spore forming bacteria and classified into a member of the order *Neisseriales*, class *Betaproteobacteria*, phylum *Proteobacteria*. To date, the genus comprises 29 species and 2

subspecies with validly published names (LPSN, <http://www.bacterio.net>). Although two species *N. gonorrhoeae* and *Neisseria meningitidis* play as human pathogens, most of *Neisseria* species are considered commensals of human and animals, particularly in oral cavity (Sneath and Barrett, 1996; Bennett *et al.*, 2014; Liu *et al.*, 2015). Here, we describe the complete genome sequence and annotation of *Neisseria* sp. KEM232 (= KCTC 22739) isolated from the smooth surface caries of a 7-year old male in Republic of Korea. The genomic DNA was extracted from the stationary phase cells using a QIAamp DNA mini kit (QIAGEN Bioinformatics). The whole genome of KEM232 was sequenced using an Illumina MiSeq system (Illumina Inc.) following a paired-end 2 × 300 bp protocol. Sequencing data were assembled with CLCGenomics Workbench 7.5.1 software (QIAGEN). The annotation of each CDS was made through National Center for Biotechnology Information prokaryotic genome Annotation Pipeline (Tatusova *et al.*, 2016). Resistance-related genes were analyzed using ResFinder 3.0 with a 90% threshold for identification of genes (Zankari *et al.*, 2012). Identification of clustered regularly interspaced short palindromic repeat (CRISPR) sequences was predicted by

*For correspondence. E-mail: scnu@scnu.ac.kr;
Tel.: +82-61-750-3613; Fax: +82-61-750-3208

application CRISPRFinder program online (<http://crispr.i2bc.paris-saclay.fr/Server/>) (Grissa *et al.*, 2007). Calculation of pairwise 16S rRNA gene sequence similarity and average nucleotide identity (ANI) was achieved using the EzBioCloud server (<http://www.ezbiocloud.net/>) and OrthoANlu tool (<https://www.ezbiocloud.net/tools/orthoaniu>; Yoon *et al.*, 2017), respectively.

The genome statistics are provided in Table 1. The complete

Table 1. Genome statistics of *Neisseria* sp. KEM232

Attribute	Value
Genome size (bp)	2,371,912
G + C content (%)	58.5
Number of contigs	1
Number of total genes	2,369
Number of protein-coding genes	2,210
Number of pseudogenes	108
Number of rRNA (5S, 16S, 23S) genes	3 (1, 1, 1)
Number of tRNA genes	44
Number of ncRNA genes	4
CRISPR array	1

genome of KEM232 was composed of a single circular chromosome and did not contain any plasmid DNA. The 2,371,912 bp genome with a G + C content of 58.5% contains 2,210 coding regions (CDS), 108 pseudogenes and 51 RNA genes (3 rRNA genes, 44 tRNA genes, and 4 non-coding RNA genes) based on NCBI's Prokaryotic Genome Annotation Pipeline (PGAP) (Fig. 1). The pseudogenes make a stop codon in the middle of nucleotides sequences that encode proteins. The genome revealed the presence of 23 phage related genes and phosphoglucomutase gene and absence of potent genes responsible for the antibiotic resistance except for tetracycline gene *tetR*. The genome sequence of this organism will allow for its further characterization and assessment of pathogenic potential.

Based on the 16S rRNA gene sequence similarity and average nucleotide identity, the strain KEM232 is most closely related to *Neisseria bacilliformis* ATCC BAA-1200^T (GenBank under accession number GL878494) with the values of 96.8% and 84%, respectively.

The *Neisseria* sp. KEM232 is available with deposit numbers of KCTC 22739 and JCM 16337.

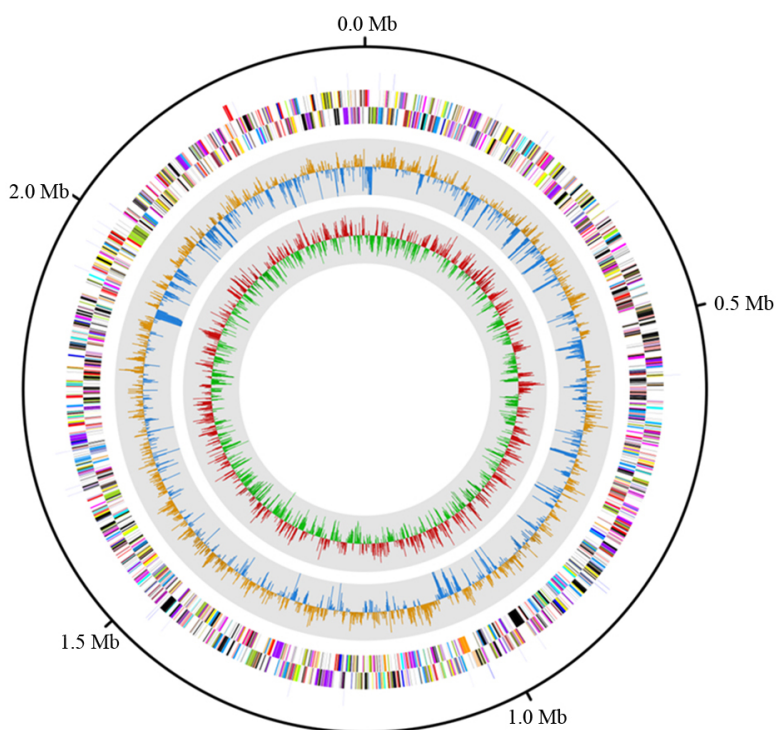


Fig. 1. Graphical circular map of *Neisseria* sp. KEM232. Marked characteristics are shown from outside to the center; CDS on forward strand, CDS on reverse strand, tRNA, rRNA, GC content, and GC skew.

Nucleotide sequence accession number

The genome sequence of *Neisseria* sp. KEM232 has been deposited in NCBI GenBank under accession number CP022527.

적 요

Neisseria 속 균주 KEM232는 사람 평활면 치아우식 부위로부터 분리하였다. 균주 KEM232의 유전체는 G + C 비율이 58.5%, 2,369개의 유전자와 2,210개의 단백질 코딩 유전자, 108개의 위유전자, 51개의 RNA 유전자 그리고 한 개의 CRISPR array를 포함한 단일 원형 염색체로 구성되었으며 그 크기는 2,371,912 bp였다. 균주 KEM232의 최 근연종은 *Neisseria baciliformis* 로서 두 균주 사이의 16S rRNA 유전자 염기서열의 유사도는 96.8% 그리고 유전체의 평균 염기 동일성은 84%였다.

Acknowledgements

The Research has been conducted by the Research Grant of Gwangju Health University in 2014 (3014009), Republic of Korea.

References

- Bennett, J.S., Bratcher, H.B., Brehony, C., Harrison, O.B., and Maiden, M.C.** 2014. The genus *Neisseria*, pp. 881–900. In Rosenberg, E., DeLong, E.F., Lory, S., Stackebrandt, E., and Thompson, F. (eds.), *The Prokaryotes -Alphaproteobacteria and Betaproteobacteria*, 4th ed. Verlag, Springer.
- Grissa, I., Vergnaud, G., and Pourcel, C.** 2007. CRISPRfinder: a web tool to identify clustered regularly interspaced short palindromic repeats. *Nucleic Acids Res.* **35**, W52–W57.
- Liu, G., Tang, C.M., and Exley, R.M.** 2015. Non-pathogenic *Neisseria*: members of an abundant, multi-habitat, diverse genus. *Microbiology* **161**, 1297–1312.
- Skerman, V.B.D., McGowan, V., and Sneath, P.H.A.** 1980. Approved lists of bacterial names. *Int. J. Syst. Bacteriol.* **30**, 225–420.
- Sneath, P.H.A. and Barrett, S.J.** 1996. A new species of *Neisseria* from the dental plaque of the domestic cow, *Neisseria dentiae* sp. nov. *Lett. Appl. Microbiol.* **23**, 355–358.
- Tatusova, T., DiCuccio, M., Badretdin, A., Chetvernin, V., Nawrocki, E.P., Zaslavsky, L., Lomsadze, A., Pruitt, K.D., Borodovsky, M., and Ostell, J.** 2016. NCBI prokaryotic genome annotation pipeline. *Nucleic Acids Res.* **44**, 6614–6624.
- Yoon, S.H., Ha, S.M., Lim, J.M., Kwon, S.J., and Chun, J.** 2017. A large-scale evaluation of algorithms to calculate average nucleotide identity. *Antonie van Leeuwenhoek* **110**, 1281–1286.
- Zankari, E., Hasman, H., Cosentino, S., Vestergaard, M., Rasmussen, S., Lund, O., Aarestrup, F.M., and Larsen, M.V.** 2012. Identification of acquired antimicrobial resistance genes. *J. Antimicrob. Chemother.* **67**, 2640–2644.