JOURNAL OF BACTERIOLOGY, Sept. 2008, p. 6035–6036 Vol. 190, No. 17
0021-9193/08/$08.00+0 doi:10.1128/JB.00566-08

Genome Announcement

Complete Genome Sequence of *Neisseria gonorrhoeae* NCCP11945▽

Gyung Tae Chung,[1] Jeong Sik Yoo,[1] Hee Bok Oh,[1] Yeong Seon Lee,[1] Sun Ho Cha,[2] Sang Jun Kim,[2] and Cheon Kwon Yoo[3]*

Centers for Infectious Diseases, National Institute of Health, Korea Centers for Disease Control and Prevention, Seoul 122-701, Republic of Korea[1]; GenoTech Corporation, 59-5 Jang-Dong, Yuseong-Gu, Daejeon 305-343, Republic of Korea[2]; and Division of Biosafety Evaluation and Control, National Institute of Health, Korea Centers for Disease Control and Prevention, Seoul 122-701, Republic of Korea[3]

Received 23 April 2008/Accepted 16 June 2008

***Neisseria gonorrhoeae* is an obligate human pathogen that is the etiological agent of gonorrhea. We explored variations in the genes of a multidrug-resistant *N. gonorrhoeae* isolate from a Korean patient in an effort to understand the prevalence, antibiotic resistance, and importance of horizontal gene transfer within this important, naturally competent organism. Here, we report the complete annotated genome sequence of *N. gonorrhoeae* strain NCCP11945.**

Neisseria gonorrhoeae NCCP11945 was isolated from a vaginal smear of a Korean patient. This strain fit the pattern of antimicrobial resistance that is prevalent in the Republic of Korea, namely, chromosome-mediated resistance to penicillin and tetracycline and a ciprofloxacin MIC of 16 mg/liter.

The complete genome sequence of NCCP11945 was determined by whole-genome shotgun sequencing. Two genome libraries were generated by random shearing of genomic DNA; one was a fosmid library with a mean insert size of ~32 kb (CopyControl fosmid library production kit; Epicentre, Madison, WI), and the other had a size of approximately 1 to 2 kb. Automated DNA sequencing chromatograms were analyzed by the Phred/Phrap/Consed software package (http://www.phrap.org). Gap closure and additional sequencing of low-coverage regions were accomplished via primer walking of gap-spanning clones and direct sequencing of PCR products. In particular, the order of 184 contigs was predicted by comparison with the genome sequence of *N. gonorrhoeae* FA1090 (strain AE004969) and then confirmed by PCR. To solve problems with misassembled regions caused by repetitive sequences and to close remaining sequence gaps, we used long PCR along with six fosmid clones. Their relationships were reassembled manually based on location information for paired-end reads using Consed. The completed genome sequence had eightfold sequence coverage, with an error rate of 0.15 error per 10,000 bases.

Open reading frame (ORF) prediction and annotation were performed using GLIMMER3 (3) and BLAST. The functional assignment of genes was performed by searching translated ORFs against sequences in the COG (9) and KEGG (5) databases.

The genome of NCCP11945 consists of one circular chromosome (2,232,025 bp) encoding 2,662 predicted ORFs and one plasmid (4,153 bp) encoding 12 predicted ORFs. The estimated coding density over the entire genome is 87%, and the average G+C content is 52.4%, values that are similar to those of strain FA1090. The strain NCCP11945 genome encodes 54 tRNAs and four copies of 16S-23S-5S rRNA operons.

Genome structure comparisons between NCCP11945 and FA1090 were performed using the programs ACT (2) and MUMMER (6). Genome colinearity between these strains is interrupted by NCCP11945-specific and FA1090-specific regions as well as by several inversions and translocations. The strain NCCP11945 genome is 82,256 bp larger than that of FA1090, and the overall genome sequence identity is 95.2%. This difference in genome size is caused by a gonococcal genetic island (GGI) in NCCP11945. The GGI is present in 80% of gonococcal isolates and encodes a type IV secretion system (4). The GGI of NCCP11945 encodes 61 predicted ORFs (GenBank accession number CP001050) and is similar to the GGI of *N. gonorrhoeae* strain MS11A (57 kb). The GGI sequence similarity between these two strains is 99.6%.

As with other *Neisseria* spp., the NCCP11945 genome contains hundreds of repetitive sequence elements. We analyzed these repetitive elements using the Emboss (8) and Nicolas (1) methods. The most abundant repeat type is the DNA uptake sequence (5′-GCCGTCTGAA-3′), comprising 1,966 copies throughout the genome. The next most abundant repeat types are neisseria intergenic mosaic elements (repetitive sequence, 123 copies; duplicate repetitive sequence 3, 215 copies) (7). The NCCP11945 genome also contains 79 copies of Correia elements (1).

Nucleotide sequence accession number. The complete genome sequence of *N. gonorrhoeae* NCCP11945 has been assigned GenBank accession numbers CP001050 and CP001051.

* Corresponding author. Mailing address: Division of Biosafety Evaluation and Control, National Institute of Health, Korea Centers for Disease Control and Prevention, Seoul 122-701, Republic of Korea. Phone: 82-2-380-2971. Fax: 82-2-380-2280. E-mail: ckyoo@nih.go.kr.

▽ Published ahead of print on 27 June 2008.

The funding for this sequencing project was provided by the National Institute of Health, Ministry of Health and Welfare, Republic of Korea.

REFERENCES

1. **Buisine, N., C. M. Tang, and R. Chalmers.** 2002. Transposon-like Correia elements: structure, distribution and genetic exchange between pathogenic *Neisseria* sp. FEBS Lett. **522:**52–58.
2. **Carver, T. J., K. M. Rutherford, M. Berriman, M. A. Rajandream, B. G. Barrell, and J. Parkhill.** 2005. ACT: the Artemis Comparison Tool. Bioinformatics **21:**3422–3423.
3. **Delcher, A. L., D. Harmon, S. Kasif, O. White, and S. L. Salzberg.** 1999. Improved microbial gene identification with GLIMMER. Nucleic Acids Res. **27:**4636–4641.
4. **Hamilton, H. L., N. M. Dominguez, K. J. Schwartz, K. T. Hackett, and J. P. Dillard.** 2005. *Neisseria gonorrhoeae* secretes chromosomal DNA via a novel type IV secretion system. Mol. Microbiol. **55:**1704–1721.
5. **Kanehisa, M., S. Goto, S. Kawashima, Y. Okuno, and M. Hattori.** 2004. The KEGG resource for deciphering the genome. Nucleic Acids Res. **32:**D277–D280.
6. **Kurtz, S., A. Phillippy, A. L. Delcher, M. Smoot, M. Shumway, C. Antonescu, and S. L. Salzberg.** 2004. Versatile and open software for comparing large genomes. Genome Biol. **5:**R12.
7. **Liu, S. V., N. J. Saunders, A. Jeffries, and R. F. Rest.** 2002. Genome analysis and strain comparison of Correia repeats and Correia repeat-enclosed elements in pathogenic *Neisseria*. J. Bacteriol. **184:**6163–6173.
8. **Rice, P., I. Longden, and A. Bleasby.** 2000. EMBOSS: the European molecular biology open software suite. Trends Genet. **16:**276–277.
9. **Tatusov, R. L., D. A. Natale, I. V. Garkavtsev, T. A. Tatusova, U. T. Shankavaram, B. S. Rao, B. Kiryutin, M. Y. Galperin, N. D. Fedorova, and E. V. Koonin.** 2001. The COG database: new developments in phylogenetic classification of proteins from complete genomes. Nucleic Acids Res. **29:**22–28.