JOURNAL OF BACTERIOLOGY, Aug. 2011, p. 4017–4018 Vol. 193, No. 15
0021-9193/11/$12.00 doi:10.1128/JB.05184-11

Complete Genome Sequence of *Bordetella pertussis* CS, a Chinese Pertussis Vaccine Strain▽

Shumin Zhang,[1] Yinghua Xu,[1,3] Zhemin Zhou,[2] Shaojing Wang,[2] Ruifu Yang,[3] Junzhi Wang,[1]* and Lei Wang[2]*

National Institute for the Control of Pharmaceutical and Biological Products, Temple of Heaven, Beijing 100050, People's Republic of China[1]; TEDA School of Biological Sciences and Biotechnology, Nankai University, Tianjin 300457, People's Republic of China[2]; and Laboratory of Analytical Microbiology, State Key Laboratory of Pathogen and Biosecurity, Beijing Institute of Microbiology and Epidemiology, Beijing 100071, People's Republic of China[3]

Received 28 April 2011/Accepted 16 May 2011

***Bordetella pertussis* is the causative agent of pertussis. Here, we report the genome sequence of *Bordetella pertussis* strain CS, isolated from an infant patient in Beijing and widely used as a vaccine strain for production of an acellular pertussis vaccine in China.**

Bordetella pertussis is a strict human pathogen that is the causative agent of pertussis (whooping cough). Despite widespread immunization with pertussis vaccines, many countries still report outbreaks (1, 12, 14). Resurgence of pertussis has been recently observed in some countries with high vaccination coverage (4, 8, 10). It is suggested that pathogen adaptation may play a role in the reemergence of pertussis (10, 15, 16). In this present work, the complete genome sequence of *B. pertussis* strain CS, which was isolated from an infant patient in 1951 in Beijing and widely used as a vaccine strain for production of an acellular pertussis vaccine in China, was determined and compared with the published genome of Tohama I (11).

Whole-genome sequencing of *B. pertussis* CS was performed with a combined strategy of the Sanger shotgun approach (6) and 454 fragment sequencing technology (9). A total of 329,480 reads, giving 28-fold coverage of the genome, were generated using the GS FLX system (454 Life Sciences Corporation) and assembled into 287 contigs with the 454 Newbler assembler. Artificial 1-kb reads representing the Roche/454 assembly were generated using mktrace and assembled with 11,444 paired-end ABI3730 reads (3.2-kb library) using the Phred/Phrap/Consed package (7). Sequence gaps were filled through sequencing of PCR products. Prediction and annotation of protein-encoding genes were performed as described previously (5) and verified manually using the annotation of Tohama I (11).

The complete genome of *B. pertussis* CS contains a circular 4,124,236-bp chromosome with an average G+C content of 67.3%. There are 3,456 protein-coding sequences with an average size of 327 amino acids, 51 tRNA genes, and three rRNA operons in the genome.

Compared with strain Tohama I, 2 large fragments (>10 kb) were exclusively present in strain CS, of which contained 18 (BPTD_2835 to BPTD_2852) and 21 (BPTD_0387 to BPTD_0407) genes. These differences were involved mainly in the transcriptional regulator system, the metabolism system, mobile elements, and the restriction modification system. The two regions can be found in both *Bordetella bronchiseptica* and *Bordetella parapertussis* (11), suggesting that these genes are lost in the Tohama I lineage and are not obtained in CS, which is consistent with the previously reported tendency (2). Both of these two regions were located adjacent to copies of the insertion element IS*481*, which is present in high numbers in the *B. pertussis* chromosome and has been reported to provide targets for homologous recombination, resulting in deletion of intervening sequences (3, 13).

Except for the two large insertions above, only 341 polymorphic sites were found between CS and Tohama I, consisting of 301 single nucleotide polymorphisms (SNPs) and 40 small insertion or deletion events (indels). Of the 301 SNPs, 204 were located in open reading frames and 97 were in the intergenic regions, indicating an estimated SNP density of 1 SNP per 13,701 bases. This places *B. pertussis* to be one of the known most monomorphic human pathogens.

In summary, the complete genome sequence of *B. pertussis* CS will facilitate additional bioinformatic and phylogenetic analyses and enable in-depth studies of *B. pertussis* sequence variations.

Nucleotide sequence accession number. The genome sequence of *B. pertussis* strain CS has been deposited in GenBank under the accession number CP002695.

This work was supported by the National Science & Technology Pillar Program from the Ministry of Science and Technology, People's Republic of China (grant no. 2008BAI54B03 and 2009ZX10608).

* Corresponding author. Mailing address for Junzhi Wang: National Institute for the Control of Pharmaceutical and Biological Products, Temple of Heaven, Beijing 100050, People's Republic of China. Phone: 86-10-67094993. Fax: 86-10-67051912. E-mail: wangjz@nicpbp.org.cn. Mailing address for Lei Wang: TEDA School of Biological Sciences and Biotechnology, Nankai University, 23 Hongda Street, TEDA, Tianjin 300457, People's Republic of China. Phone: 86-22-66229588. Fax: 86-22-66229596. E-mail: wanglei@nankai.edu.cn.

▽ Published ahead of print on 27 May 2011.

REFERENCES

1. **Berger, F., et al.** 2010. Investigation on a pertussis outbreak in a military school: risk factors and approach to vaccine efficacy. Vaccine **28:**5147–5152.
2. **Bouchez, V., V. Caro, E. Levillain, G. Guigon, and N. Guiso.** 2008. Genomic content of *Bordetella pertussis* clinical isolates circulating in areas of intensive children vaccination. PLoS One **3:**e2437.
3. **Brinig, M. M., et al.** 2006. Significant gene order and expression differences

in *Bordetella pertussis* despite limited gene content variation. J. Bacteriol. **188:**2375–2382.
4. **de Melker, H. E., et al.** 2000. Reemergence of pertussis in the highly vaccinated population of the Netherlands: observations on surveillance data. Emerg. Infect. Dis. **6:**348–357.
5. **Feng, L., et al.** 2008. A recalibrated molecular clock and independent origins for the cholera pandemic clones. PLoS One **3:**e4053.
6. **Fleischmann, R. D., et al.** 1995. Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. Science **269:**496–512.
7. **Gordon, D., C. Abajian, and P. Green.** 1998. Consed: a graphical tool for sequence finishing. Genome Res. **8:**195–202.
8. **Guris, D., et al.** 1999. Changing epidemiology of pertussis in the United States: increasing reported incidence among adolescents and adults, 1990-1996. Clin. Infect. Dis. **28:**1230–1237.
9. **Margulies, M., et al.** 2005. Genome sequencing in microfabricated high-density picolitre reactors. Nature **437:**376–380.
10. **Mooi, F. R., I. H. van Loo, and A. J. King.** 2001. Adaptation of *Bordetella pertussis* to vaccination: a cause for its reemergence? Emerg. Infect. Dis. **7:**526–528.
11. **Parkhill, J., et al.** 2003. Comparative analysis of the genome sequences of *Bordetella pertussis*, *Bordetella parapertussis* and *Bordetella bronchiseptica*. Nat. Genet. **35:**32–40.
12. **Sin, M. A., et al.** 2009. Pertussis outbreak in primary and secondary schools in Ludwigslust, Germany demonstrating the role of waning immunity. Pediatr. Infect. Dis. J. **28:**242–244.
13. **Stibitz, S., and M. S. Yang.** 1999. Genomic plasticity in natural populations of *Bordetella pertussis*. J. Bacteriol. **181:**5512–5515.
14. **Takum, T., D. Gara, H. Tagyung, and M. V. Murhekar.** 2009. An outbreak of pertussis in Sarli Circle of Kurung-kumey district, Arunachal Pradesh, India. Indian Pediatr. **46:**1017–1020.
15. **van Boven, M., F. R. Mooi, J. F. Schellekens, H. E. de Melker, and M. Kretzschmar.** 2005. Pathogen adaptation under imperfect vaccination: implications for pertussis. Proc. Biol. Sci. **272:**1617–1624.
16. **van der Zee, A., F. Mooi, J. Van Embden, and J. Musser.** 1997. Molecular evolution and host adaptation of *Bordetella* spp.: phylogenetic analysis using multilocus enzyme electrophoresis and typing with three insertion sequences. J. Bacteriol. **179:**6609–6617.