Draft Genome Sequence of *Porphyromonas gingivalis* Strain SJD2, Isolated from the Periodontal Pocket of a Patient with Periodontitis in China

Dali Liu,[a] Yanbin Zhou,[a] Mariko Naito,[b] Hiromichi Yumoto,[c] Qiang Li,[d] Yoichiro Miyake,[e] Jingping Liang,[f] Rong Shu[a]

Department of Periodontology, Ninth People's Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai Key Laboratory of Stomatology, Shanghai, China[a]; Division of Microbiology and Oral Infection, Department of Molecular Microbiology and Immunology, Nagasaki University Graduate School of Biomedical Sciences, Nagasaki, Japan[b]; Department of Conservative Dentistry, Institute of Health Biosciences, The Tokushima University Graduate School of Dentistry, Tokushima, Japan[c]; Shanghai Majorbio Bio-pharm Technology Company Limited, Shanghai, China[d]; Department of Oral Microbiology, Institute of Health Biosciences, The Tokushima University Graduate School of Dentistry, Tokushima, Japan[e]; Department of Endodontics, Ninth People's Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai Key Laboratory of Stomatology, Shanghai, China[f]

D.L. and Y.Z. contributed equally to this work.

Porphyromonas gingivalis strain SJD2 was isolated from subgingival plaque of a patient in China with chronic periodontitis. Here, we report the draft genome of this strain, with a size of 2,328,850 bp, average G+C content of 48.3%, and 2,020 predicted protein-coding sequences.

*P*orphyromonas gingivalis is a black-pigmented bacterium from the phylum *Bacteroidetes* and is a major pathogen in periodontal disease (1). The complete genomic sequences of *P. gingivalis* strains W83, ATCC 33277, and TDC60 have been available (2–4). *P. gingivalis* SJD2 was obtained from subgingival dental plaque of a patient in China with chronic periodontitis (5). The 16S rRNA sequence of this strain showed 99% identity with those of type strains of W83 and ATCC 33277, as well as with clinical strain TDC60 (2). In addition, this strain was shown to have high virulent properties comparable with those of the type strain of W83 in a mouse abscess model.

The whole-genome sequence of *P. gingivalis* SJD2 was generated with short-insert DNA (496 bp) libraries using the high-throughput sequencing Illumina HiSeq 2000 platform (BGI, Shenzhen, China), which generated 100-bp paired-end sequences at a coverage of 100×. The resulting 257.5-Mb raw data were cleaned by removing adapter, low-quality reads, poly-N, and error pair-end reads. Quality-trimmed reads (Sanger Q, ≥20, 91.8% of total reads) were then assembled with SOAP*denovo* version 1.04 (http://soap.genomics.org.cn/soapdenovo.html). The draft genome, as submitted to GenBank, is 2,328,850 bp in length and comprises 140 contigs (>300 bp in size), with a mean contig size of 16,635 bp, a median size of 32,167 bp, and a maximum length of 85,472 bp. The mean G+C content of the genome is 48.3%.

Gene prediction was performed using the Prokaryotic Genome Automatic Annotation Pipeline (PGAAP) provided by National Center for Biotechnology Information (NCBI). A total of 2,021 protein-coding sequences (CDSs) and 51 structural RNAs, together with 64 pseudogenes and 3 clustered regularly interspaced short palindromic repeats (CRISPRs), were predicted. Each CDS was then reviewed manually for the presence of a start codon and

a potential ribosome-binding Shine-Dalgarno (SD) sequence that should be related to a part of an AGAAAGGAGG(T) sequence, or an AANGA sequence with a sufficient size of 6 to 3 nucleotides (3). Three hundred four CDSs without SD sequences were performed with the homology search through the public protein database from NCBI using the BLASTp program (http://blast.ncbi.nlm.nih.gov/Blast.cgi?CMD=Web&PAGE_TYPE=BlastHome). One CDS that has a size of >180 bp showing an e-value of homology >e-10 was excluded from the 2,021 CDSs. Finally, 2,020 CDSs were predicted, and the details were given in the GenBank database.

Overall, BLASTp analysis with the complete genome sequences of W83, ATCC 33277, and TDC60 under 70% of identity covered 50 to 150% length revealed that 13.7% CDSs (278 of 2,020 CDSs) in SJD2 are represented as strain specific, which is higher than the percentages (6.5 to 9.6%) of strain-specific CDSs in other strain genomes. The higher number of SJD2-specific CDSs suggests that strains isolated from a periodontal pocket of Chinese patients with chronic periodontitis may have distinct genes. Further investigation of the genomes of this highly virulent clinical strain may yield more insights into the considerable pathogenicity of *P. gingivalis*.

Nucleotide sequence accession numbers. This whole-genome shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession no. ASYL00000000. The version described in this paper is version ASYL01000000.

ACKNOWLEDGMENT

REFERENCES

1. **Holt SC, Ebersole JL.** 2005. *Porphyromonas gingivalis, Treponema denticola,* and *Tannerella forsythia:* the "red complex", a prototype polybacterial pathogenic consortium in periodontitis. Periodontol. 2000 **38:**72–122. http://dx.doi.org/10.1111/j.1600-0757.2005.00113.x.

2. **Nelson KE, Fleischmann RD, DeBoy RT, Paulsen IT, Fouts DE, Eisen JA, Daugherty SC, Dodson RJ, Durkin AS, Gwinn M, Haft DH, Kolonay JF, Nelson WC, Mason T, Tallon L, Gray J, Granger D, Tettelin H, Dong H, Galvin JL, Duncan MJ, Dewhirst FE, Fraser CM.** 2003. Complete genome sequence of the oral pathogenic bacterium *Porphyromonas gingivalis* strain W83. J. Bacteriol. **185:**5591–5601. http://dx.doi.org/10.1128/JB.185.18.5591-5601.2003.

3. **Naito M, Hirakawa H, Yamashita A, Ohara N, Shoji M, Yukitake H, Nakayama K, Toh H, Yoshimura F, Kuhara S, Hattori M, Hayashi T, Nakayama K.** 2008. Determination of the genome sequence of *Porphyromonas gingivalis* strain ATCC 33277 and genomic comparison with strain W83 revealed extensive genome rearrangements in *P. gingivalis.* DNA Res. **15:**215–225. http://dx.doi.org/10.1093/dnares/dsn013.

4. **Watanabe T, Maruyama F, Nozawa T, Aoki A, Okano S, Shibata Y, Oshima K, Kurokawa K, Hattori M, Nakagawa I, Abiko Y.** 2011. Complete genome sequence of the bacterium *Porphyromonas gingivalis* TDC60, which causes periodontal disease. J. Bacteriol. **193:**4259–4260. http://dx.doi.org/10.1128/JB.05269-11.

5 **Li Y, Liu D, Shu R, Zhou Y.** 2012. Isolation, identification and growth characteristics of *Porphyromonas gingivalis* clinical strains. Shanghai Jiao Tong Da Xue Xue Bao Yi Xue Ban **32:**64–68.