Draft Genome Sequence of *Streptococcus suis* S10, a Virulent Strain Used in Experimental Pig Infections

Rogier A. Gaiser,[a]* Aldert L. Zomer,[b] Jerry M. Wells,[a] Peter van Baarlen[a]

aHost-Microbe Interactomics, Wageningen University & Research, Wageningen, The Netherlands
bVeterinary Medicine, Department of Infectious Diseases and Immunology, Utrecht University, Utrecht, The Netherlands

ABSTRACT Here, we report the draft whole-genome sequence of *Streptococcus suis* strain S10, isolated from the tonsils of a healthy pig. *S. suis* S10 belongs to the highly virulent serotype 2, which includes isolates that cause infectious diseases, including meningitis, in pigs and human. The genome contains a complete prophage that encodes a candidate virulence gene.

*S*treptococcus suis bacteria are commonly part of the porcine tonsillar microbiota (1), comprising carriage strains and strains that cause infectious disease in pigs and humans (2). Here, we report the genome sequence of *S. suis* strain S10, sampled in 1992 from porcine tonsils, which has been used for experimental infections of pigs (3).

S. suis S10 was cultured without agitation at 37°C with 5% CO_2 in Todd-Hewitt Broth (THB) (Oxoid, UK). Genomic DNA was extracted from 2 ml of exponentially growing bacteria that were lysed in Nuclei Lysis solution (Promega) with proteinase K and protein precipitation solution (Promega) following the manufacturer's protocol. DNA was precipitated in isopropanol and purified using phenol-chloroform-isoamyl alcohol following the protocols described by Barker (4). DNA purity and quality were assessed by gel electrophoresis and spectrometric analysis (ND-1000; NanoDrop Technologies).

DNA library preparation using the Illumina Nextera XT kit following the manufacturer's protocol and genomic DNA sequencing using an Illumina HiSeq 2500 instrument on a paired-end library (125 cycles) were carried out at BaseClear B.V. (Leiden, The Netherlands). FASTQ sequence files were generated using the Illumina CASAVA pipeline 1.8.3. A BaseClear in-house pipeline carried out quality assessment of paired-end reads by Illumina Chastity filtering, removal of reads containing adapters, PhiX control signal, and FASTQC Quality Control Tool 0.10.0. CLC Genomics Workbench 8 trimmed low-quality bases and assembled reads into contigs. KmerGenie (5) determined the optimal k-mer size. SSPACE Premium Scaffolder 2.3 (6) linked contigs and placed them into scaffolds. GapFiller 1.10 (7) closed gapped regions within scaffolds. For all *in silico* analyses, default parameters were used.

Sequencing generated 4,268,836 paired-end reads (average length, 126 bases) from which 2,011,091 bases were aligned (266.38-fold coverage). Filtered reads were assembled in 32 gap-closed scaffolds. Quast (8) showed that the GC content was 41.25%, the N_{50} and L_{50} values were 170,355 and 5, and the N_{75} and L_{75} values were 86,491 and 9. The number of uncalled bases per assembled 100 kbp was 0.10.

The S10 genome was annotated via Prokka (9). ANItools Web (10) showed that *S. suis* strain P1/7 (11) was most identical to S10 (99.97%; not shown). CONTIGuator (12) aligned S10 genomic contigs to the P1/7 genome; 23 S10 contigs with a combined base content of 1,997,858 base pairs were mapped to the total of 2,007,491 base pairs of P1/7.

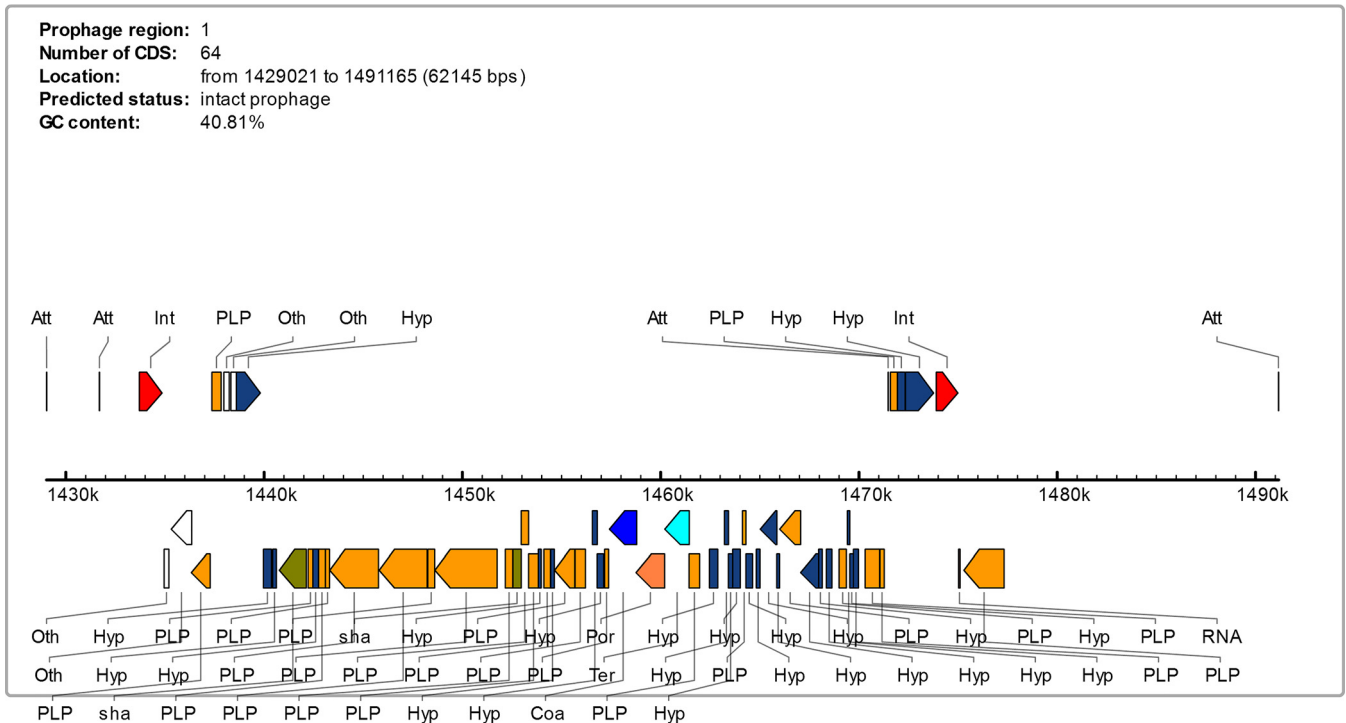In contig_00009, Phage Search Tool Enhanced Release (PHASTER) (13) predicted a complete prophage containing 64 coding sequences (CDS) spanning a region of

Address correspondence to Peter van Baarlen, peter.vanbaarlen@wur.nl.

* Present address: Rogier A. Gaiser, Karolinska Institutet, Department of Dental Medicine, Huddinge, Stockholm, Sweden.

FIG 1 Phage Search Tool Enhanced Release (PHASTER) analysis output displaying the 64 coding sequences (CDS) of phage origin and their functional annotations; the annotated region shown is part of the largest S10 region that did not align to *S. suis* strain P1/7. Att, attachment site; Hyp, hypothetical protein; Int, integrase; Oth, other protein; PLP, phage-like protein.

62,145 base pairs (Fig. 1). The predicted S10 phage contains a candidate virulence gene with high identity (E value, 7E−164) with the *S. mitis* (14) and *S. pneumoniae* (15) phage-encoded gene *pblB*, which encodes a platelet adhesin that plays a role during multiple steps in endocardial infection, including direct binding of bacteria to platelets.

The genome sequence of *S. suis* S10 will enable the research community to make targeted gene deletion mutants.

Data availability. The genome sequence of strain *S. suis* S10 has been deposited at the European Molecular Biology Laboratory, European Bioinformatics Institute (EMBL-EBI); the raw reads have been deposited at the European Nucleotide Archive (ENA) under the study identifier PRJEB30600.

REFERENCES

1. Arends JP, Hartwig N, Rudolphy M, Zanen HC. 1984. Carrier rate of *Streptococcus suis* capsular type 2 in palatine tonsils of slaughtered pigs. J Clin Microbiol 20:945–947.
2. Goyette-Desjardins G, Auger JP, Xu J, Segura M, Gottschalk M. 2014. *Streptococcus suis*, an important pig pathogen and emerging zoonotic agent—an update on the worldwide distribution based on serotyping and sequence typing. Emerg Microbes Infect 3:e45. https://doi.org/10.1038/emi.2014.45.
3. Vecht U, Wisselink HJ, van Dijk JE, Smith HE. 1992. Virulence of *Streptococcus suis* type 2 strains in newborn germfree pigs depends on phenotype. Infect Immun 60:550–556.
4. Barker K. 2005. At the bench: a laboratory navigator, updated edition. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
5. Chikhi R, Medvedev P. 2014. Informed and automated k-mer size selection for genome assembly. Bioinformatics 30:31–37. https://doi.org/10.1093/bioinformatics/btt310.
6. Boetzer M, Henkel CV, Jansen HJ, Butler D, Pirovano W. 2011. Scaffolding pre-assembled contigs using SSPACE. Bioinformatics 27:578–579. https://doi.org/10.1093/bioinformatics/btq683.
7. Boetzer M, Pirovano W. 2012. Toward almost closed genomes with GapFiller. Genome Biol 13:R56. https://doi.org/10.1186/gb-2012-13-6-r56.
8. Gurevich A, Saveliev V, Vyahhi N, Tesler G. 2013. QUAST: quality assessment tool for genome assemblies. Bioinformatics 29:1072–1075. https://doi.org/10.1093/bioinformatics/btt086.
9. Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. Bioinformatics 30:2068–2069. https://doi.org/10.1093/bioinformatics/btu153.
10. Han N, Qiang Y, Zhang W. 2016. ANItools Web: a Web tool for fast genome comparison within multiple bacterial strains. Database (Oxford) 2016:baw084. https://doi.org/10.1093/database/baw084.
11. Holden MT, Hauser H, Sanders M, Ngo TH, Cherevach I, Cronin A, Goodhead I, Mungall K, Quail MA, Price C, Rabbinowitsch E, Sharp S, Croucher NJ, Chieu TB, Mai NT, Diep TS, Chinh NT, Kehoe M, Leigh JA, Ward PN, Dowson CG, Whatmore AM, Chanter N, Iversen P, Gottschalk M, Slater JD, Smith HE, Spratt BG, Xu J, Ye C, Bentley S, Barrell BG, Schultsz C, Maskell DJ, Parkhill J. 2009. Rapid evolution of virulence and drug resistance in the emerging zoonotic pathogen *Streptococcus suis*. PLoS One 4:e6072. https://doi.org/10.1371/journal.pone.0006072.
12. Galardini M, Biondi EG, Bazzicalupo M, Mengoni A. 2011. CONTIGuator: a bacterial genomes finishing tool for structural insights on draft genomes. Source Code Biol Med 6:11. https://doi.org/10.1186/1751-0473-6-11.
13. Arndt D, Grant JR, Marcu A, Sajed T, Pon A, Liang Y, Wishart DS. 2016. PHASTER: a better, faster version of the PHAST phage search tool. Nucleic Acids Res 44:W16–W21. https://doi.org/10.1093/nar/gkw387.
14. Bensing BA, Siboo IR, Sullam PM. 2001. Proteins PblA and PblB of *Streptococcus mitis*, which promote binding to human platelets, are encoded within a lysogenic bacteriophage. Infect Immun 69:6186–6192. https://doi.org/10.1128/IAI.69.10.6186-6192.2001.
15. Tunjungputri RN, Mobegi FM, Cremers AJ, van der Gaast-de Jongh CE, Ferwerda G, Meis JF, Roeleveld N, Bentley SD, Pastura AS, van Hijum SA, van der Ven AJ, de Mast Q, Zomer A, de Jonge MI. 2017. Phage-derived protein induces increased platelet activation and is associated with mortality in patients with invasive pneumococcal disease. mBio 8:e01984-16. https://doi.org/10.1128/mBio.01984-16.