AMERICAN SOCIETY FOR MICROBIOLOGY | Microbiology Resource Announcements

GENOME SEQUENCES

Complete Genome Sequencing of the Attenuated Strain *Mycoplasma synoviae* 5-9

Xiaorong Zhang,[a,b] Yang Chen,[a] Zehua Wei,[a] Shuang Ma,[b] Mengjiao Guo,[a] Dianfeng Chu,[b] Chengcheng Zhang,[a] Yongzhong Cao,[c] Yantao Wu[a,c]

[a]Jiangsu Co-Innovation Center for the Prevention and Control of Animal Infectious Disease and Zoonoses, College of Veterinary Medicine, Yangzhou University, Yangzhou, Jiangsu, China

[b]State Key Laboratory of Genetically Engineered Veterinary Vaccines, Qingdao Yebio Biological Engineering Co., Ltd., Qingdao, Shandong, China

[c]The International Joint Research Laboratory of Agricultural and Agri-product Safety, Yangzhou University, Yangzhou, Jiangsu, China

Xiaorong Zhang, Yang Chen, and Zehua Wei contributed equally to this article. Order of authors was ranked by contributions.

ABSTRACT Here, we present the complete genome sequence of *Mycoplasma synoviae* strain 5-9. Strain 5-9 was attenuated by chemical mutagenesis from a field strain isolated from egg breeders in Ningxia, China. It was completely sequenced and its genome annotated; it is presented with the relevant data as a potential vaccine candidate.

Mycoplasma synoviae is an important pathogen infecting chickens, causing acute or chronic respiratory diseases, infectious synovitis, air sacculitis (1–3), and eggshell apex abnormalities (4). Ningxia/2017-1 is the virulent parent strain of 5-9, isolated in Ningxia, China. *M. synoviae* isolation and identification using *Mycoplasma* broth (MB) and *Mycoplasma* agar (MA) were conducted as previously described (5). A Ningxia/2017-1 culture was exposed to 100 μg/ml N-methyl-N′-nitro-N-nitrosoguanidine (NTG) for 15 min and continuously passaged 5 times in 10 μg/ml NTG at 33°C before being plated onto MA. Single colonies were selected to grow in MB at 33°C (6).

High-quality genomic DNA (optical density at 260/280 nm [$OD_{260/280}$] = 1.8 ~ 2.0) from 5-9 was extracted for Illumina and ONT sequencing using a bacterial genomic DNA kit (CWBIO Biotech, China) until the color changed from red to orange-yellow when cultured at 33°C and passaged 2 times in MB (pH 6.9). For next-generation sequencing (NGS), paired-end libraries with insert sizes of ~400 bp were prepared using the Illumina TruSeq Nano DNA high-throughput library prep kit. The purified DNA was sheared into smaller fragments of the desired size (Covaris). Finally, the qualified Illumina paired-end library was used for Illumina NovaSeq 6000 sequencing (2 × 150-bp format; Shanghai BIOZERON Co., Ltd., China). A total of 2,333 Mb raw data was produced, and we obtained 2,222 Mb clean data after filtering out the low-quality reads using Trimmomatic (7) (http://www.usadellab.org/cms/?page=trimmomatic) (Table 1). The raw paired-end reads were trimmed and quality controlled using Trimmomatic v0.36 (7) (http://www.usadellab.org/cms/?page=trimmomatic) (parameters, SLIDINGWINDOW:4:15 MINLEN:75).

A Nanopore library was prepared using the ONT 1D ligation sequencing kit (SQK-LSK108) with the native barcoding expansion kit (EXP-NBD103). At least 1 μg DNA was treated with the end-repair/dA-tailing module, but the DNA was eluted in 24 μl following AMPure XP bead cleanup. Following the barcode ligation reaction, the DNA was cleaned again using AMPure XP beads and elution in 10 μl. The run was performed on a MinION MK1b device using the NC_48 h_Sequencing_Run_FLOMIN106_SQK-LSK108 protocol.

The Illumina data were used to evaluate the complexity of the genome and correct the Nanopore long reads. First, we used ABySS v2.02 (8, 9) (http://www.bcgsc.ca/platform/bioinfo/software/abyss) to perform the genome assembly with multiple kmer parameters and obtain

Editor David Rasko, University of Maryland School of Medicine

Address correspondence to Yantao Wu, ytwu@yzu.edu.cn.

The authors declare no conflict of interest.

Received 2 October 2021
Accepted 17 November 2021
Published 9 December 2021

TABLE 1 Genome assembly, Nanopore sequencing, and Illumina sequencing data

					Nanopore sequencing			Illumina sequencing			
								Raw data		Clean data	
Sample	Total length (bp)	N_{50} length (bp)	G+C content (%)	Coverage (×)	No. of reads	No. of bases	Avg length (bp)	Total no. of reads	Total size (Mb)	Total no. of reads	Total size (Mb)
5-9	786,872	786,872	28.47	1,015	125,369	798,728,344	6,371	15,553,532	2,333	14,881,648	2,222.2

optimal assembly results. Second, Canu v2.0 (10) (https://github.com/marbl/canu) was used to assemble the Nanopore corrected long reads. Finally, GapCloser v1.12 (11) (https://sourceforge.net/projects/soapdenovo2/files/GapCloser/) was applied to fill the remaining local inner gaps and correct single-base polymorphisms for the final assembly results. The complete circularized genome was drawn using Circos v0.64 (12) (http://circos.ca/). Default parameters were used for all software unless otherwise specified. The consensus assembly generated one contig of 786,872 bp (1,015-fold coverage). The G+C content was 28.47% (Table 1).

Gene models were identified using GeneMark. Then, all gene models were searched using blastp against the NCBI nonredundant (NR), Swiss-Prot (http://uniprot.org), KEGG (http://www.genome.jp/kegg/), and COG (http://www.ncbi.nlm.nih.gov/COG) databases. Additionally, tRNAs were identified using tRNAscan-SE v1.23 (13) (http://lowelab.ucsc.edu/tRNAscan-SE), and rRNAs were determined using RNAmmer v1.2 (14) (https://services.healthtech.dtu.dk/service.php?RNAmmer-1.2). There were 652 protein-coding genes and 44 RNA genes (34 tRNAs, 3 5S rRNAs, 2 16S rRNAs, 2 23S rRNAs, and 3 noncoding RNA [ncRNA] genes).

Data availability. The complete genome sequence and annotation of *M. synoviae* strain 5-9 have been deposited at GenBank under accession number CP083748. The raw data were deposited in the Sequence Read Archive (SRA) database under the accession numbers SRR16005423 (Oxford Nanopore) and SRR16005422 (Illumina). The BioProject accession number is PRJNA763075. The BioSample accession number is SAMN21422653.

ACKNOWLEDGMENTS

This work was supported by the Key Special Project "Science and Technology Promote Economy of 2020" of the National Key Research and Development Program (SQ2020YFF0426460), China Agriculture Research System of MOF and MARA (CARS-40), the State Key Laboratory of Genetically Engineered Veterinary Vaccines (AGVSKL-ZD-202001), and a project funded by the Priority Academic Program Development of Jiangsu Higher Education Institutions (PAPD).

REFERENCES

1. Catania S, Gobbo F, Ramirez AS, Guadagnini D, Baldasso E, Moronato ML, Nicholas RAJ. 2016. Erratum for "Laboratory investigations into the origin of Mycoplasma synoviae isolated from a lesser flamingo (Phoeniconaias minor)." BMC Vet Res 12:75. https://doi.org/10.1186/s12917-016-0699-3.
2. Fletcher OJ, Anderson DP, Kleven SH. 1976. Histology of air sac lesions induced in chickens by contact exposure to Mycoplasma synoviae. Vet Pathol 13:303–314. https://doi.org/10.1177/030098587601300407.
3. Michiels T, Welby S, Vanrobaeys M, Quinet C, Rouffaer L, Lens L, Martel A, Butaye P. 2016. Prevalence of Mycoplasma gallisepticum and Mycoplasma synoviae in commercial poultry, racing pigeons and wild birds in Belgium. Avian Pathol 45:244–252. https://doi.org/10.1080/03079457.2016.1145354.
4. Feberwee A, de Wit JJ, Landman WJ. 2009. Induction of eggshell apex abnormalities by Mycoplasma synoviae: field and experimental studies. Avian Pathol 38:77–85. https://doi.org/10.1080/03079450802662772.
5. Zhang X, Chen Y, Xie D, Guo M, Ma S, Chen M, Chu D, Wu Y. 2021. Multilocus sequence typing analysis of Mycoplasma synoviae isolates reveals unique sequence types in China. Vet Microbiol 259:109101. https://doi.org/10.1016/j.vetmic.2021.109101.
6. Morrow CJ, Markham JF, Whithear KG. 1998. Production of temperature-sensitive clones of Mycoplasma synoviae for evaluation as live vaccines. Avian Dis 42:667–670. https://doi.org/10.2307/1592700.
7. Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30:2114–2120. https://doi.org/10.1093/bioinformatics/btu170.
8. Simpson JT, Wong K, Jackman SD, Schein JE, Jones SJ, Birol I. 2009. ABySS: a parallel assembler for short read sequence data. Genome Res 19:1117–1123. https://doi.org/10.1101/gr.089532.108.
9. Jackman SD, Vandervalk BP, Mohamadi H, Chu J, Yeo S, Hammond SA, Jahesh G, Khan H, Coombe L, Warren RL, Birol I. 2017. ABySS 2.0: resource-efficient assembly of large genomes using a Bloom filter. Genome Res 27:768–777. https://doi.org/10.1101/gr.214346.116.
10. Koren S, Walenz BP, Berlin K, Miller JR, Bergman NH, Phillippy AM. 2017. Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. Genome Res 27:722–736. https://doi.org/10.1101/gr.215087.116.
11. Luo R, Liu B, Xie Y, Li Z, Huang W, Yuan J, He G, Chen Y, Pan Q, Liu Y, Tang J, Wu G, Zhang H, Shi Y, Liu Y, Yu C, Wang B, Lu Y, Han C, Cheung

DW, Yiu SM, Peng S, Xiaoqian Z, Liu G, Liao X, Li Y, Yang H, Wang J, Lam TW, Wang J. 2012. SOAPdenovo2: an empirically improved memory-efficient short-read de novo assembler. Gigascience 1:18. https://doi.org/10.1186/2047-217X-1-18.

12. Krzywinski M, Schein J, Birol I, Connors J, Gascoyne R, Horsman D, Jones SJ, Marra MA. 2009. Circos: an information aesthetic for comparative genomics. Genome Res 19:1639–1645. https://doi.org/10.1101/gr.092759.109.
13. Lowe TM, Chan PP. 2016. tRNAscan-SE On-line: integrating search and context for analysis of transfer RNA genes. Nucleic Acids Res 44:W54–W57. https://doi.org/10.1093/nar/gkw413.
14. Lagesen K, Hallin P, Rodland EA, Staerfeldt HH, Rognes T, Ussery DW. 2007. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. Nucleic Acids Res 35:3100–3108. https://doi.org/10.1093/nar/gkm160.