JB Journals.ASM.org

GENOME ANNOUNCEMENT

Complete Genome Sequence of *Salmonella enterica* Serovar Pullorum RKS5078

Ye Feng,[a] Hua-Feng Xu,[a] Qing-Hai Li,[a] Si-Yao Zhang,[a] Chun-Xiao Wang,[a] Da-Ling Zhu,[b] Feng-Lin Cao,[c] Yong-Guo Li,[c] Randal N. Johnston,[d] Jin Zhou,[c] Gui-Rong Liu,[a,c] and Shu-Lin Liu[a,c,e,f]

Genomics Research Center,[a] Institute of Biopharmaceutical Sciences, College of Pharmacy,[b] and Genetic Detection Center, First Hospital,[c] Harbin Medical University, Harbin, China; Departments of Biochemistry and Molecular Biology[d] and Microbiology and Infectious Diseases,[e] University of Calgary, Calgary, Alberta, Canada; and Department of Microbiology, Peking University Health Science Center, Beijing, China[f]

***Salmonella enterica* serovar Pullorum is a chicken-adapted pathogen, causing pullorum disease. Its strict host adaptation has been suspected to result in gene decay. To validate this hypothesis and identify the decayed genes, we sequenced the complete genome of *S.* Pullorum RKS5078. We found 263 pseudogenes in this strain and conducted functional analyses of the decayed genes.**

Salmonellae are important pathogenic bacteria infecting humans or animals. *Salmonella enterica* serovar Pullorum, a strictly chicken-adapted pullorum agent, is one of the more than 2,500 documented *Salmonella* lineages (4). It is very closely related to another chicken pathogen, *Salmonella enterica* serovar Gallinarum, and, to a lesser degree, to a host generalist, *Salmonella enterica* serovar Enteritidis. Therefore, *S.* Pullorum has the special genetic and pathogenic features suitable for studies to elucidate the genetic basis and evolution of host adaptation and acquisition of novel pathogenicity in *Salmonella*. To this end, we sequenced the genome of *S.* Pullorum RKS5078 and carried out systematic analyses.

The sequenced *S.* Pullorum strain RKS5078 was obtained from R. K. Selander. The genomic DNA was sheared into 3-kb fragments by the Hydroshear instrument and then was sequenced on a SOLiD sequencer by the mate-pair strategy (2 × 50 bp), as detailed in the manual for the instrument (Applied Biosystems). The final coverage depth reached 325-fold for an estimated genome size of 4.6 Mb. The derived color space reads were assembled by SOLiD *de novo* accessory tools V2.0. The relationships between contigs were determined following the physical map published earlier (3), and the gaps between contigs were closed by PCR amplification of the genomic DNA and sequencing. To avoid potential misassembly caused by the short-reads sequencing platform, we compared the finished genome with that of *S.* Gallinarum 287/91 (accession no. AM933173.1): all nucleotide differences longer than 50 bp (e.g., insertions and deletions) were further validated by PCR.

The chromosome of *S.* Pullorum RKS5078 was 4,637,962 bp, and its average G+C content was 52.2%. The sequence was submitted to the web service RAST for automatic annotation (1) followed by manual checking. For identification of pseudogenes, *S.* Enteritidis P125109 (accession no. AM933172.1) was used as a reference for comparison by the Psi-Fi Perl script (2). Genes that contain frameshifts, nonsense mutations, truncations, or insertion/deletions (indels) that altered >20% of the amino acid sequence in comparison with the reference sequence were treated as pseudogenes.

On the chromosome of *S.* Pullorum RKS5078, we annotated 4,325 protein-coding genes and 263 pseudogenes. The genome was highly syntenic with those of *S.* Enteritidis P125109 and *S.* Gallinarum 287/91, with average nucleotide identities of 99.7% to the former and 99.8% to the latter. More detailed comparative analyses with additional *Salmonella* strains will provide further insights into the evolution of the genetically related but pathogenically distinct *Salmonella* lineages.

Nucleotide sequence accession number. The genome sequence of *S.* Pullorum RKS5078 has been deposited in GenBank under accession no. CP003047.

ACKNOWLEDGMENTS

We thank R. K. Selander for the kind gift of the bacterial strains used in this study.

This work was supported by a Canadian Institutes of Health Research grant to R.N.J., a grant from the National Natural Science Foundation of China (NSFC30970078) and a grant from the Natural Science Foundation of Heilongjiang Province of China to G.R.L.; a grant from Harbin Medical University, grants from the National Natural Science Foundation of China (NSFC30870098, -30970119, and -81030029), a 985 Project grant from the Beijing University Health Science Center, and a Specialized Research Fund for the Doctoral Program of Higher Education (no. 20092307110001) to S.L.L.; and a grant from the Harbin Medical University Innovation Program in Graduate Education (HCXB2009016) to H.-F.X.

REFERENCES

1. **Aziz RK, et al.** 2008. The RAST Server: rapid annotations using subsystems technology. BMC Genomics **9**:75.
2. **Lerat E, Ochman H.** 2004. Psi-Phi: exploring the outer limits of bacterial pseudogenes. Genome Res. **14**:2273–2278.
3. **Liu GR, et al.** 2002. The evolving genome of *Salmonella enterica* serovar Pullorum. J. Bacteriol. **184**:2626–2633.
4. **Shivaprasad HL.** 2000. Fowl typhoid and pullorum disease. Rev. Sci. Tech. **19**:405–424.

Received 10 November 2011 **Accepted** 18 November 2011

Address correspondence to Jin Zhou, jinzhouh85@163.com, or Shu-Lin Liu, slliu@ucalgary.ca.

Y. Feng and H.-F. Xu contributed equally to this work.

doi:10.1128/JB.06507-11