Draft Genome Sequence of *Yersinia pestis* Strain 2501, an Isolate from the Great Gerbil Plague Focus in Xinjiang, China

Surong Sun, a Xianwei Yang, b,c Yanting Yuan, c,d Xiang Dai, a Yanfeng Yan, b Hanli Cao, e Tao Luo, e Rong Guo, e Xinhui Wang, e Yajun Song, a Rufu Yang, b Yujian Zhang, a and Yujun Cui b

Xinjiang Key Laboratory of Biological Resources and Genetic Engineering, College of Life Science and Technology, Xinjiang University, Urumqi, People’s Republic of China; State Key Laboratory of Pathogen and Biosecurity, Beijing Institute of Microbiology and Epidemiology, Beijing, People’s Republic of China; BGI-Shenzhen, Shenzhen, People’s Republic of China; State Key Laboratory of Bioelectronics, School of Biological Science and Medical Engineering, Southeast University, Nanjing, People’s Republic of China; and The Center for Disease Control and Prevention of Xinjiang Uygur Autonomous Region, Urumqi, People’s Republic of China

We deciphered the genome of *Yersinia pestis* strain 2501, isolated from the Junggar Basin, a newly discovered great gerbil plague focus in Xinjiang, China. The total length of assembly was 4,597,322 bp, and 4,265 coding sequences were predicted within the genome. It is the first *Y. pestis* genome from this plague focus.

Plague has claimed millions of people in the three major historical pandemics and has remained endemic in many natural plague foci around the world (1,8). The causative agent of plague, *Yersinia pestis*, can be spread to humans from natural rodent reservoirs through the bite of an infected flea, contact with infected animals and/or persons, or respiratory droplets (9). As the human infection cases of *Y. pestis* increased in recent years, plague was listed as one of the 20 reemergent infectious diseases by the WHO (10).

The Junggar Basin is located in Xinjiang province, in the northwest of China, with an area of around 330,000 km². The North Tianshan Mountain plague focus neighbors the southern side of the Junggar Basin (2,6,7), and the Pre-Balkhash plague focus (belonging to the Central Asian desert region, Kazakhstan) is on the west of the Junggar Basin (1,4). Although encircled by two large, active natural plague foci and although the natural environment of the local region was adapted to the establishment of a plague focus (affluent species diversity of wild rodents and fleas present in the region), no plague epidemic case was reported and no *Y. pestis* strain was isolated from the wild animals/fleas during around 50 years of surveillance in this region (12). Since May 2005, great gerbils (*Rhombomys opimus*) infected with *Y. pestis* were observed in multiple sites in the Junggar Basin, and the following epidemic investigation provided evidence for the emergence of a new natural plague focus in this region (11,12).

Strain 2501 was isolated from a dead great gerbil in the south of the Junggar Basin in 2005 (12). It belongs to the biovar Medievalis and reveals high virulence to mouse (50% lethal dose [LD50] = 13 CFU). The strain was cultured on Luria-Bertani broth, and DNA was extracted using the conventional SDS lysis and phenol-chloroform extraction method. Whole-genome sequencing was performed using Illumina HiSeq 2000 (Illumina Inc.) by generating paired-end libraries with an insert size of 500 bp according to the manufacturer’s instructions. The read length is 90 bp, and 500 Mb of high-quality data were generated. The paired-end reads were de novo assembled using SOAPdenovo v1.06 (5), and gaps were filled using mapping information. The coding sequences (CDSs) were predicted by using Glimmer v3.02 (3), and homologous comparison to a nonredundant public database for function annotations was performed by BLAST. Finally, we obtained 185 scaffolds consisting of 201 contigs with a total length of 4,597,322 bp and predicted 4,265 CDSs within the assembly result. The G+C content of the genome is 47.53%, and the average G+C content for each gene is 48.79%.

The comparative genomic analysis between strain 2501 and the genomes of other available *Y. pestis* isolates will provide information on the evolution and spread process of *Y. pestis* from other foci to this region. A detailed report of a full comparative genomic analysis will be included in a future publication.

**Nucleotide sequence accession numbers.** This whole-genome shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession number AKVQ00000000. The version described in this paper is the first version, AKVQ01000000.

**ACKNOWLEDGMENTS**

This work was supported by the Industry Research Special Foundation of China Ministry of Health (201202021), the National Natural Science Foundation of China (30960348), and the Key Research Program Foundation of Xinjiang Province (200933120).

**REFERENCES**


