Journal of Bacteriology

Whole-Genome Sequences and Comparative Genomics of Salmonella enterica Serovar Typhi Isolates from Patients with Fatal and Nonfatal Typhoid Fever in Papua New Guinea

Ramani Baddam, Kwai-Lin Thong, Tiruvayipati Suma Avasthi, Sabiha Shaik, Kien-Pong Yap, Cindy Shuan Ju Teh, Lay-Ching Chai, Narender Kumar and Niyaz Ahmed

J. Bacteriol. **2012, 194(18):5122. DOI: 10.1128/JB.01051-12.**

Updated information and services can be found at:
http://jb.asm.org/content/194/18/5122

These include:

REFERENCES This article cites 23 articles, 15 of which can be accessed free at: http://jb.asm.org/content/194/18/5122#ref-list-1

CONTENT ALERTS Receive: RSS Feeds, eTOCs, free email alerts (when new articles cite this article), more»

Information about commercial reprint orders: **http://journals.asm.org/site/misc/reprints.xhtml**
To subscribe to to another ASM Journal go to: **http://journals.asm.org/site/subscriptions/**

Journals.ASM.org

Downloaded from http://jb.asm.org/ on March 26, 2013 by UNIVERSITY OF MALAYA

JB Journals.ASM.org

GENOME ANNOUNCEMENT

Whole-Genome Sequences and Comparative Genomics of *Salmonella enterica* Serovar Typhi Isolates from Patients with Fatal and Nonfatal Typhoid Fever in Papua New Guinea

Ramani Baddam,[a] Kwai-Lin Thong,[b,c] Tiruvayipati Suma Avasthi,[a,b] Sabiha Shaik,[a] Kien-Pong Yap,[b,c] Cindy Shuan Ju Teh,[c] Lay-Ching Chai,[b,c] Narender Kumar,[a] and Niyaz Ahmed[a,b,d]

Pathogen Biology Laboratory, Department of Biotechnology, School of Life Sciences (Centre for Advanced Studies—UGC-SAP-CAS-I), University of Hyderabad, Gachibowli, Hyderabad, India[a]; Institute of Biological Sciences, Faculty of Science, University of Malaya, Kuala Lumpur, Malaysia[b]; Laboratory of Biomedical Science and Molecular Microbiology, UMBIO Research Cluster, University of Malaya, Kuala Lumpur, Malaysia[c]; and Institute of Life Sciences, University of Hyderabad Campus, Gachibowli, Hyderabad, India[d]

Many of the developing countries of the Southeast Asian region are significantly affected by endemic typhoid fever, possibly as a result of marginal living standards. It is an important public health problem in countries such as Papua New Guinea, which is geographically close to some of the foci of endemicity in Asia. The severity of the disease varies in different regions, and this may be attributable to genetic diversity among the native strains. Genome sequence data on strains from different countries are needed to clearly understand their genetic makeup and virulence potential. We describe the genomes of two *Salmonella* Typhi isolates from patients with fatal and nonfatal cases of typhoid fever in Papua New Guinea. We discuss in brief the underlying sequencing methodology, assembly, genome statistics, and important features of the two draft genomes, which form an essential step in our functional molecular infection epidemiology program centering on typhoid fever. The comparative genomics of these and other isolates would enable us to identify genetic rearrangements and mechanisms responsible for endemicity and the differential severity of pathogenic salmonellae in Papua New Guinea and elsewhere.

Typhoid fever is a major pestilence in the developing world (20), and its prevalence is significant (1,000 cases per 100,000 individuals per year) in Papua New Guinea, which is geographically close to Southeast Asia (15). DNA profiling, which was used previously (6, 14, 20, 21, 22), could not fully explore genome diversity, as *Salmonella enterica* serovar Typhi isolates from Papua New Guinea showed limited heterogeneity, perhaps because of recent clonal expansion from a single endemic/ancestral strain (the disease was rarely seen before 1985) (17, 20). Minimal selection pressure and confinement to a specific geographical region might explain this limited genetic diversity (21, 22) despite horizontal gene transfer (13).

We hypothesized that the genome sequences of *Salmonella* Typhi isolates from patients with typhoid fever due a fatal strain (UJ308A) or a nonfatal strain (UJ816A) would provide significant insights into the association among disease phenotypes and strain characteristics. Two such strains were isolated from blood samples from patients and were found to be sensitive to common antibiotics. Strain UJ308A (phage type VS1) was obtained from a patient who died of typhoid, while UJ816A (phage type DI) was from a patient who recovered.

The 73-bp paired-end sequence data (insert size, 300 bp) were determined with an Illumina Genome Analyzer (GA2x, pipeline version 1.6). About 95× and 105× coverage was achieved for strains UJ308A and UJ816A, respectively, comprising 1.9 and 2.0 Gb of data, respectively. *De novo* assembly was done as described previously (1, 2, 4, 8, 19); initial assembly generated 416 and 335 contigs for UJ308A and UJ816A, respectively, using Velvet (23) with a hash length of 39. The scaffolds were generated from contigs by using SSPACE (5) and further assembled and curated to give a consensus draft. The following statistics were gleaned upon analysis at RAST (3). The sizes of the chromosomes for UJ308A and UJ816A were approximately 4,724,875 and 4,736,723 bp, respectively, with a G+C contents of 51.89 and 51.94%, respectively. The coding percentage for both strains was ~86.8%; UJ308A and UJ816A contained approximately 4,720 and 4,710 protein coding sequences with average lengths of 869 and 873 bp, respectively. The data were further validated by Glimmer (7) and EasyGene (12). RNAmmer (11) revealed that the genome of UJ308A has 78 tRNA and 21 rRNA genes and the genome of UJ816A contains 77 tRNA and 22 rRNA genes. All of the major virulence markers encoded by pathogenicity islands and the genes relevant to the assembly of a type III secretion system (16) were identified in both the strains. The Vi antigen (10, 18), which plays major role in immune evasion, was present in both strains, as in *Salmonella* Typhi CT18 (16). The homologues of *Campylobacter* toxin *cdtB* and *Bordetella* pertussis toxin (9) were also present.

In view of this, further efforts are needed to determine the true extent of strain diversity in terms of (i) gene gains and losses over an evolutionary time scale, (ii) geographic gene flow, (iii) core versus accessory genome dynamics, (iv) virulence acquisition and attenuation, and (v) the preponderance of highly virulent versus "docile" strains across the regions of Asia where typhoid fever is endemic.

Nucleotide sequence accession numbers. The GenBank accession numbers for the genomes reported here are AJTD00000000 (UJ308A) and AJTE00000000 (UJ816A).

Received 12 June 2012 **Accepted** 29 June 2012

Address correspondence to Niyaz Ahmed, niyazSL@uohyd.ernet.in.

Copyright © 2012, American Society for Microbiology. All Rights Reserved.

doi:10.1128/JB.01051-12

Downloaded from http://jb.asm.org/ on March 26, 2013 by UNIVERSITY OF MALAYA

ACKNOWLEDGMENTS

We thankfully acknowledge support received from the University of Malaya High Impact Research Grant (UM.C/625/1HIR/MOHE/02 [A000002-5000 1]) Molecular Genetics. This genome program was completed under the wider umbrella of the Indo-German International Research Training Group, Internationales Graduiertenkolleg (GRK1673), Functional Molecular Infection Epidemiology, an initiative of the German Research Foundation (DFG) and the University of Hyderabad (India), of which N.A. is a speaker. N.A. is an Adjunct Professor of Molecular Biosciences at the University of Malaya, Kuala Lumpur, Malaysia, and an Adjunct Professor of Chemical Biology at the Institute of Life Sciences, Hyderabad, India. We are also grateful to M/s Genotypic Technology Pvt. Ltd., Bengaluru, India, for their untiring efforts with Illumina sequencing. We acknowledge the Bioinformatics Facility (BIF) at the Department of Biotechnology, University of Hyderabad, for the use of its computational infrastructure. We are thankful to Akash Ranjan for enabling access to the high-speed computing infrastructure at CDFD, Hyderabad.

All of the members of the Ahmed and Thong labs are gratefully acknowledged for their help and support.

REFERENCES

1. **Avasthi TS, et al.** 2011. Genomes of two chronological isolates (*Helicobacter pylori* 2017 and 2018) of the West African *Helicobacter pylori* strain 908 obtained from a single patient. J. Bacteriol. **193**:3385–3386.
2. **Avasthi TS, et al.** 2011. Genome of multidrug-resistant uropathogenic *Escherichia coli* strain NA114 from India. J. Bacteriol. **193**:4272–4273.
3. **Aziz RK, et al.** 2008. The RAST server: rapid annotations using subsystems technology. BMC Genomics **9**:75. doi:10.1186/1471-2164-9-75.
4. **Baddam R, et al.** 2012. Genetic fine structure of a *Salmonella enterica* serovar Typhi strain associated with the 2005 outbreak of typhoid fever in Kelantan, Malaysia. J. Bacteriol. **194**:3565–3566.
5. **Boetzer M, et al.** 2011. Scaffolding pre-assembled contigs using SSPACE. Bioinformatics **27**:578–579.
6. **Combs BG, et al.** 2005. Ribotyping of *Salmonella enterica* serovar Typhi isolates from Papua New Guinea over the period 1977 to 1996. P. N. G. Med. J. **48**:158–167.
7. **Delcher AL, et al.** 1999. Improved microbial gene identification with GLIMMER. Nucleic Acids Res. **27**:4636–4641.
8. **Devi SH, et al.** 2010. Genome of *Helicobacter pylori* strain 908. J. Bacteriol. **192**:6488–6489.
9. **Haghjoo E, Galan JE.** 2004. Salmonella typhi encodes a functional cytolethal distending toxin that is delivered into host cells by a bacterial-internalization pathway. Proc. Natl. Acad. Sci. U. S. A. **101**:4614–4619.
10. **Hirose K, et al.** 1997. Survival of Vi-capsulated and Vi-deleted *Salmonella typhi* strains in cultured macrophage expressing different levels of CD14 antigen. FEMS Microbiol. Lett. **147**:259–265.
11. **Lagesen K, et al.** 2007. RNammer: consistent annotation of rRNA genes in genomic sequences. Nucleic Acids Res. **35**:3100–3108.
12. **Larsen TS, Krogh A.** 2003. EasyGene—a prokaryotic gene finder that ranks ORFs by statistical significance. BMC Bioinformatics **4**:21. doi: 10.1186/1471-2105-4-21.
13. **Liu SL, Sanderson KE.** 1995. Rearrangements in the genome of the bacterium Salmonella typhi. Proc. Natl. Acad. Sci. U. S. A. **92**:1018–1022.
14. **Nair S, Schreiber E, Thong KL, Pang T, Altwegg M.** 2000. Genotypic characterization of *Salmonella typhi* by amplified fragment length polymorphism fingerprinting provides increased discrimination as compared to pulsed-field gel electrophoresis and ribotyping. J. Microbiol. Methods **41**:35–43.
15. **Pang T, Bhutta ZA, Finlay BB, Altwegg M.** 1995. Typhoid fever and other salmonelloses: a continuing challenge. Trends Microbiol. **3**:253–255.
16. **Parkhill J, et al.** 2001. Complete genome sequence of a multiple drug resistant Salmonella enterica serovar Typhi CT18. Nature **413**:848–852.
17. **Passey M, et al.** 1995. Highly endemic typhoid fever in Papua New Guinea. Southeast Asian J. Trop. Med. Public Health **26**:83–84.
18. **Raffatellu M, et al.** 2006. Capsule-mediated immune evasion: a new hypothesis explaining aspects of typhoid fever pathogenesis. Infect. Immun. **74**:19–27.
19. **Siddavattam D, et al.** 2011. Genome of a novel isolate of *Paracoccus* denitrificans capable of degrading *N*,*N*-dimethylformamide. J. Bacteriol. **193**:5598–5599.
20. **Thong KL, et al.** 1995. Analysis of *Salmonella typhi* isolates from Southeast Asia by pulsed-field gel electrophoresis. J. Clin. Microbiol. **33**:1938–1941.
21. **Thong KL, et al.** 1996. Molecular analysis of isolates of *Salmonella typhi* obtained from patients with fatal and nonfatal typhoid fever. J. Clin. Microbiol. **34**:1029–1033.
22. **Thong KL, et al.** 2002. Increasing genetic diversity of *Salmonella enterica* serovar Typhi isolates from Papua New Guinea over the period from 1992 to 1999. J. Clin. Microbiol. **40**:4156–4160.
23. **Zerbino DR, Birney E.** 2008. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. Genome Res. **18**:821–829.

Downloaded from http://jb.asm.org/ on March 26, 2013 by UNIVERSITY OF MALAYA