JOURNAL OF BACTERIOLOGY, Oct. 2010, p. 5556–5557 Vol. 192, No. 20
0021-9193/10/$12.00 doi:10.1128/JB.00878-10

Complete Genome Sequence of *Staphylococcus aureus* Strain JKD6159, a Unique Australian Clone of ST93-IV Community Methicillin-Resistant *Staphylococcus aureus*▽

Kyra Chua,[1,2,4] Torsten Seemann,[3] Paul F. Harrison,[3] John K. Davies,[1] Scott J. Coutts,[1] Honglei Chen,[6] Volker Haring,[6] Robert Moore,[6,7] Benjamin P. Howden,[1,2,4,5]† and Timothy P. Stinear[1,2]*

Department of Microbiology, Monash University, Melbourne, Victoria, Australia[1]; Department of Microbiology and Immunology, University of Melbourne, Parkville, Victoria, Australia[2]; Victorian Bioinformatics Consortium, Monash University, Melbourne, Victoria, Australia[3]; Departments of Infectious Diseases[4] and Microbiology,[5] Austin Health, Melbourne, Victoria, Australia; CSIRO, Australian Animal Health Laboratory, Geelong, Victoria, Australia[6]; and ARC Centre of Excellence in Structural and Functional Microbial Genomics, Monash University, Melbourne, Victoria, Australia[7]

Received 26 July 2010/Accepted 9 August 2010

Community methicillin-resistant *Staphylococcus aureus* (cMRSA) is an emerging issue that has resulted in multiple worldwide epidemics. We report the first complete genome sequence of an ST93-MRSA-IV clinical isolate that caused severe invasive infection and a familial outbreak of skin infection. This isolate is a representative of the most common Australian clone of cMRSA that is more distantly related to the previously sequenced genomes of *S. aureus*.

Staphylococcus aureus is a major cause of both hospital- and community-acquired infections, with rapid emergence of antibiotic resistance, in particular methicillin resistance, adding complexity to the treatment of this organism (3). While previously a hospital problem, methicillin-resistant *S. aureus* (MRSA) is now being increasingly documented in healthy patients in the community, and these isolates are termed "community MRSA" (cMRSA). A number of cMRSA genomes have been sequenced; however, these are phylogenetically closely related to each other. In contrast, ST93-MRSA-IV, a unique Australian clone, is a singleton by multilocus sequence typing (MLST) eBURST analysis (4). It is now the dominant cMRSA clone in Australia and is associated with both skin infection and severe invasive infection, including necrotizing pneumonia, deep-seated abscesses, and septicemia (5, 10). JKD6159 is a representative ST93-MRSA-IV clinical isolate which caused septicemia and multifocal pulmonary and musculoskeletal abscesses in a previously well intravenous drug user and also resulted in a familial outbreak of skin infection.

The genome sequence of *S. aureus* strain JKD6159 was determined by high-throughput whole-genome shotgun sequencing, using both Illumina GAII (Illumina, CA) and Roche GS FLX Titanium (Roche Diagnostics, Basel, Switzerland) sequencing technologies, producing approximately 164× and 32× coverage of the genome, respectively. The GS FLX Titanium reads were assembled using Newbler 2.0.01.12, resulting in 56 contigs totaling 2.8 Mbp (9). The paired GAII reads were aligned to the contigs using SHRiMP 1.3.2 to identify and correct 74 homopolymeric sequencing errors (11). Optical mapping was used to produce a high-resolution XbaI chromosome restriction map, and the contigs were ordered and oriented against this map using MapSolver 2.1.1 (Opgen). Gap closures were performed by PCR followed by Sanger sequencing of amplification products (3730S genetic analyzer sequencer; Applied Biosystems, CA). The finished sequence was validated by reference to the XbaI optical map, Roche GS FLX Titanium mate pair analysis, and Illumina paired-end-read analysis.

Protein coding regions were predicted using GeneMarkS 4.6b, tRNA genes using tRNAscan-SE 1.23, and rRNA genes using RNAmmer 1.2 (2, 7, 8). Gene products were assigned using HMMER 3.0 against the Pfam database (release 23) and BLAST 2.2.23 against RefSeq Proteins (April 2010) and the Conserved Domain Database (v2.22) (1, 6). These automated analyses were followed by manual curation, including comparison with other completed *S. aureus* genomes.

The genome of *S. aureus* strain JKD6159 consists of a circular 2,811,435-bp chromosome with 33% G+C content—similar to those of other staphylococci—and one circular plasmid of 20,730 bp. A total of 2,605 coding regions, 57 tRNA genes, and 5 rRNA loci were detected. Over 67% of genes were assigned to specific Clusters of Orthologous Groups (COG) Database functional groups, and 40% were assigned an enzyme classification number (12).

Initial analysis of the whole-genome sequence of JKD6159 confirms that ST93-MRSA-IV is distantly related to other previously sequenced *S. aureus* genomes. ST93-MRSA-IV has a distinct accessory genome. There were a number of regions of difference in JKD6159 that contain coding sequences (CDS) not present in any other published *S. aureus* genomes. Additionally, the *ssl* gene cluster in JKD6159 appears distinct from other sequenced *S. aureus* isolates. Comparison with other *S. aureus* genomes also shows that although JKD6159 carries

* Corresponding author. Mailing address: Department of Microbiology and Immunology, University of Melbourne, Parkville 3010, Victoria, Australia. Phone: 61 3 8344 5711. Fax: 61 3 9347 1540. E-mail: tstinear@unimelb.edu.au.
† Joint senior author.
▽ Published ahead of print on 20 August 2010.

lukSF-PV (the genes encoding Panton-Valentine leukocidin), there is a relative paucity of virulence factors such as *tst-1*, genes encoding staphylococcal enterotoxins A to U, and the ACME locus. Further analysis of the genome is now under way to identify factors that might explain the emergence of this MRSA strain in the community.

Nucleotide sequence accession numbers. The complete genome sequence has been deposited in NCBI GenBank under accession no. CP002114 and CP002115.

This work was supported by the National Health and Medical Research Council of Australia, the CASS Foundation, and the Austin Hospital Medical Research Foundation.

REFERENCES

1. **Altschul, S. F., T. L. Madden, A. A. Schäffer, J. Zhang, Z. Zhang, W. Miller, and D. J. Lipman.** 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. **25:**3389–3402.
2. **Besemer, J., A. Lomsadze, and M. Borodovsky.** 2001. GeneMarkS: a self-training method for prediction of gene starts in microbial genomes. Implications for finding sequence motifs in regulatory regions. Nucleic Acids Res. **29:**2607–2618.
3. **Chambers, H. F., and F. Deleo.** 2009. Waves of resistance: Staphylococcus aureus in the antibiotic era. Nat. Rev. Microbiol. **7:**629–641.
4. **Coombs, G. W., G. R. Nimmo, J. M. Bell, F. Huygens, F. G. O'Brien, M. J. Malkowski, J. C. Pearson, A. J. Stephens, P. M. Giffard, and Australian Group for Antimicrobial Resistance.** 2004. Genetic diversity among community methicillin-resistant *Staphylococcus aureus* strains causing outpatient infections in Australia. J. Clin. Microbiol. **42:**4735–4743.
5. **Coombs, G. W., G. R. Nimmo, J. C. Pearson, K. J. Christiansen, J. M. Bell, P. J. Collignon, M. L. McLaws, and Australian Group for Antimicrobial Resistance.** 2009. Prevalence of MRSA strains among Staphylococcus aureus isolated from outpatients, 2006. Commun. Dis. Intell **33:**10–20.
6. **Eddy, S. R.** 2009. A new generation of homology search tools based on probabilistic inference. Genome Inform. **23:**205–211.
7. **Lagesen, K., P. Hallin, E. A. Rødland, H. H. Staerfeldt, T. Rognes, and D. W. Ussery.** 2007. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. Nucleic Acids Res. **35:**3100–3108.
8. **Lowe, T. M., and S. R. Eddy.** 1997. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res. **25:**955–964.
9. **Margulies, M., M. Egholm, W. E. Altman, S. Attiya, J. S. Bader, L. A. Bemben, J. Berka, M. S. Braverman, Y. J. Chen, Z. Chen, S. B. Dewell, L. Du, J. M. Fierro, X. V. Gomes, B. C. Godwin, W. He, S. Helgesen, C. H. Ho, C. H. Ho, G. P. Irzyk, S. C. Jando, M. L. Alenquer, T. P. Jarvie, K. B. Jirage, J. B. Kim, J. R. Knight, J. R. Lanza, J. H. Leamon, S. M. Lefkowitz, M. Lei, J. Li, K. L. Lohman, H. Lu, V. B. Makhijani, K. E. McDade, M. P. McKenna, E. W. Myers, E. Nickerson, J. R. Nobile, R. Plant, B. P. Puc, M. T. Ronan, G. T. Roth, G. J. Sarkis, J. F. Simons, J. W. Simpson, M. Srinivasan, K. R. Tartaro, A. Tomasz, K. A. Vogt, G. A. Volkmer, S. H. Wang, Y. Wang, M. P. Weiner, P. Yu, R. F. Begley, and J. M. Rothberg.** 2005. Genome sequencing in microfabricated high-density picolitre reactors. Nature **437:**376–380.
10. **Peleg, A. Y., W. J. Munckhof, S. L. Kleinschmidt, A. J. Stephens, and F. Huygens.** 2005. Life-threatening community-acquired methicillin-resistant Staphylococcus aureus infection in Australia. Eur. J. Clin. Microbiol. Infect. Dis. **24:**384–387.
11. **Rumble, S. M., P. Lacroute, A. V. Dalca, M. Fiume, A. Sidow, and M. Brudno.** 2009. SHRiMP: accurate mapping of short color-space reads. PLoS Comput. Biol. **5:**e1000386.
12. **Tatusov, R. L., N. D. Fedorova, J. D. Jackson, A. R. Jacobs, B. Kiryutin, E. V. Koonin, D. M. Krylov, R. Mazumder, S. L. Mekhedov, A. N. Nikolskaya, B. S. Rao, S. Smirnov, A. V. Sverdlov, S. Vasudevan, Y. I. Wolf, J. J. Yin, and D. A. Natale.** 2003. The COG database: an updated version includes eukaryotes. BMC Bioinformatics **4:**41.