Genome Sequence of *Klebsiella pneumoniae* Ecl8, a Reference Strain for Targeted Genetic Manipulation

Maria Fookes,[a] Jing Yu,[b] Shyamasree De Majumdar,[b] Nicholas Thomson,[a] Thamarai Schneiders[b]

Wellcome Trust, Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge, United Kingdom[a]; Center for Infection and Immunity, Queen's University Belfast, Belfast, United Kingdom[b]

We report the genome sequence of *Klebsiella pneumoniae* subsp. *pneumoniae* Ecl8, a spontaneous streptomycin-resistant mutant of strain ECL4, derived from NCIB 418. *K. pneumoniae* Ecl8 has been shown to be genetically tractable for targeted gene deletion strategies and so provides a platform for in-depth analyses of this species.

*K*lebsiella pneumoniae is an important nosocomial pathogen and ubiquitous colonizer of the environment (1, 2). Given the increasing antimicrobial resistance within this species (3), it is critical to establish phenotypic roles for gene loci linked both directly and indirectly to reduced antimicrobial susceptibility. *K. pneumoniae* strain Ecl8 is of human origin (4) and is amenable to targeted gene deletion (5, 6) using a modified strategy described by Merlin et al. (7). To ensure the accurate mapping of targeted gene deletions using *K. pneumoniae* Ecl8, whole-genome sequencing was undertaken. Our studies indicate that genetic deletions or mutations generated in Ecl8 are widely applicable to other *Klebsiella* strains derived from clinical, environmental, and veterinary niches.

Total bacterial DNA was extracted using the Purelink genomic DNA extraction kit (Life Technologies, United Kingdom), with which the whole genome sequence was determined using both the Illumina HiSeq 2000 and the Roche 454 GS FLX sequencers. The 454 data were produced from an 8-kb-insert library, generating 590,000 reads with an average length of 539 bp by using Ti$^+$ chemistry. The 1.2-Gb Illumina data were 100-bp paired reads. A merged assembly was generated by the Newbler program (v 2.6) and improved using IMAGE (8) and iCORN (9). Chromosomal and plasmid sequences were ordered against the genome sequences of *K. pneumoniae* KCTC 2242 (accession no. CP002910 and CP002911), as the KCTC 2242 and Ecl8 chromosomes were found to differ in only 258 single-nucleotide polymorphisms (SNPs), based on a reconstructed phylogenetic tree (10).

The complete genome assembly for *K. pneumoniae* Ecl8 is composed of a chromosome of 5,324,709 bp in 5 contigs, including the 454 scaffolds and the 206,102-bp-long plasmid in a single scaffold (3 gaps). The N_{50} values for the chromosome and plasmid were 279,996 and 52,921 bp, respectively, and the largest contig values were 449,727 and 77,767 bp. Published genome sequences of *K. pneumoniae* strains 342 and KCTC 2242 were used to inform annotation, prior to manual curation.

The chromosome of *K. pneumoniae* Ecl8 has a G+C content of 57.2%. Of its 5,006 predicted coding sequences (CDS), 44 are pseudogenes sustaining mutations likely to ablate function; 85 tRNAs, 25 rRNAs, and 88 sRNAs are also predicted. TBLASTx comparisons between the KCTC 2242 and Ecl8 chromosomes showed that the latter carries two insertions: the first (8.6 kb) carries the beta-lactamase resistance gene *ampH* and the ATP-binding cassette (ABC) transporter domain-containing gene *smbA*, among others; the second insertion is a 39.2-kb novel prophage inserted alongside tRNA^{Arg}. Database analysis showed that this prophage is similar in regions to the *Salmonella enterica* serovar Typhimurium ST64B prophage, particularly in relation to capsid-encoding genes; there are no obvious regions of low G+C content within this prophage or gene or protein motifs that indicate the carriage of cargo genes associated with virulence or lipopolysaccharide modification. The plasmid is identical to the previously reported *K. pneumoniae* plasmid pKTCT2242 except for two additional insertion sequence elements.

Nucleotide sequence accession numbers. Raw data were submitted to the European Nucleotide Archive (ENA) with sample accession no. ERS088743. Genome assemblies, qualifying as an improved high-quality draft standard (11) of the chromosome and plasmid of *Klebsiella pneumoniae* Ecl8, have EMBL accession numbers HF536482 and HF536483, respectively.

ACKNOWLEDGMENTS

REFERENCES

1. **Podschun R, Ullmann U.** 1998. *Klebsiella* spp. as nosocomial pathogens: epidemiology, taxonomy, typing methods, and pathogenicity factors. Clin. Microbiol. Rev. **11**:589–603.
2. **Podschun R, Pietsch S, Höller C, Ullmann U.** 2001. Incidence of *Klebsiella* species in surface waters and their expression of virulence factors. Appl. Environ. Microbiol. **67**:3325–3327.
3. **Ho J, Tambyah PA, Paterson DL.** 2010. Multiresistant gram-negative infections: a global perspective. Curr. Opin. Infect. Dis. **23**:546–553.
4. **Forage RG, Lin EC.** 1982. DHA system mediating aerobic and anaerobic dissimilation of glycerol in *Klebsiella pneumoniae* NCIB 418. J. Bacteriol. **151**:591–599.

5. **Rosenblum R, Khan E, Gonzalez G, Hasan R, Schneiders T.** 2011. Genetic regulation of the *ramA* locus and its expression in clinical isolates of *Klebsiella pneumoniae*. Int. J. Antimicrob. Agents **38**:39–45.

6. **Veleba M, Higgins PG, Gonzalez G, Seifert H, Schneiders T.** 2012. Characterization of RarA, a novel AraC family multidrug resistance regulator in *Klebsiella pneumoniae*. Antimicrob. Agents Chemother. **56**:4450–4458.

7. **Merlin C, McAteer S, Masters M.** 2002. Tools for characterization of *Escherichia coli* genes of unknown function. J. Bacteriol. **184**:4573–4581.

8. **Tsai IJ, Otto TD, Berriman M.** 2010. Improving draft assemblies by iterative mapping and assembly of short reads to eliminate gaps. Genome Biol. **11**:R41.

9. **Otto TD, Sanders M, Berriman M, Newbold C.** 2010. Iterative correction of reference nucleotides (iCORN) using second generation sequencing technology. Bioinformatics **26**:1704–1707.

10. **Harris SR, Feil EJ, Holden MT, Quail MA, Nickerson EK, Chantratita N, Gardete S, Tavares A, Day N, Lindsay JA, Edgeworth JD, de Lencastre H, Parkhill J, Peacock SJ, Bentley SD.** 2010. Evolution of MRSA during hospital transmission and intercontinental spread. Science **327**(5964):469–474.

11. **Chain PS, Grafham DV, Fulton RS, Fitzgerald MG, Hostetler J, Muzny D, Ali J, Birren B, Bruce DC, Buhay C, Cole JR, Ding Y, Dugan S, Field D, Garrity GM, Gibbs R, Graves T, Han CS, Harrison SH, highlander S, Hugenholtz P, Khouri HM, Kodira CD, Kolker E, Kyrpides NC, Lang D, Lapidus A, Malfatti SA, Markowitz V, Metha T, Nelson KE, Parkhill J, Pitluck S, Qin X, Read TD, Schmutz J, Sozhamannan S, Sterk P, Strausberg RL, Sutton G, Thomson NR, Tiedje JM, Weinstock G, Wollam A, Detter JC.** 2009. Genomics. Genome project standards in a new era of sequencing. Science **326**:236–237.