Complete Genome Sequencing of the Mouse Intestinal Isolate *Escherichia coli* Mt1B1

Debora Garzetti,[a,b] Claudia Eberl,[a] Bärbel Stecher[a,b]

[a]Max von Pettenkofer Institute of Hygiene and Medical Microbiology, Faculty of Medicine, LMU Munich, Munich, Germany

[b]German Center for Infection Research (DZIF), Partner Site Munich, Munich, Germany

ABSTRACT *Escherichia coli* Mt1B1, a mouse isolate, is a facultative anaerobic bacterium which was shown to counteract *Salmonella enterica* serovar Typhimurium infection in a mouse model. In the present study, we describe the complete genome sequence of *E. coli* Mt1B1, composed of a 5.1-Mb chromosome and a 62.6-kb plasmid.

Colonization resistance, defined as the ability of the gut microbiota to prevent infection by potential pathogens, is a key function fulfilled by the heterogeneous microbial community populating the mammalian gastrointestinal tract (1). To study this important feature, gnotobiotic mouse models can be employed for cause-effect analyses and to elucidate the role of individual bacteria (1). We have recently demonstrated that members of the Oligo-Mouse-Microbiota (Oligo-MM), a defined bacterial collection of 12 murine strains, confer partial colonization resistance against *Salmonella enterica* serovar Typhimurium (2). The addition of three facultative anaerobes, including *Escherichia coli* strain Mt1B1, to the Oligo-MM increased colonization resistance to the level of conventional mice. Importantly, *Escherichia coli* Mt1B1 showed high similarity to *S.* Typhimurium, both in terms of hierarchical KEGG module clustering and occupation of physical niches (2).

E. coli Mt1B1 (deposited at the German Culture Collection of Microorganisms and Cell Cultures [DSMZ] with number DSM-28618) has been isolated from the ileal mucosa of a conventional laboratory mouse, and its genome has been sequenced in frame of the mouse intestinal bacteria collection (miBC) (3). Analysis of the complete genome to discover specific metabolic functions and pathways is crucial for uncovering the interplay between *E. coli* Mt1B1 and the gut microbiota in the context of resistance against infection by enteric pathogens, like *S.* Typhimurium.

The complete genome of *E. coli* Mt1B1 was obtained by extracting the genomic DNA with a phenol-chloroform-based protocol from a culture grown overnight in LB medium, inoculated by a single bacterial colony. Sequences obtained on a PacBio RS II system with a 20-kb insert size (Eurofins Genomics, Ebersberg, Germany) were assembled using the Canu 1.7 software (4). The publicly available Illumina MiSeq reads (SRA accession number ERX1284231), after trimming and filtering with Trimmomatic 0.36 (5), were used for error correction by mapping them against the complete genome with BWA 7.10 (6), followed by extraction of the consensus sequence with CLC Genomics Workbench 7.5.1 (Qiagen Bioinformatics). The obtained genome sequence of *E. coli* Mt1B1 is composed of two contigs, with one chromosome of 5,115,630 bp and 50.6% G+C content, and one contig of 62,649 bp and 41.1% G+C content. The small sequence was found to be highly similar to plasmid pCROD2 (IncX4 type) from *Citrobacter rodentium* ICC168 and pCROD2-like plasmids from other *E. coli* mouse isolates by BLAST (7) and PlasmidFinder (8) searches. Gene prediction and subsystem analysis were performed using the Rapid Annotations using Subsystems Technology (RAST) server (9) for both the chromosome and plasmid sequences, resulting in 4,859

Address correspondence to Debora Garzetti, garzetti@mvp.uni-muenchen.de.

coding sequences, 22 rRNAs, and 86 tRNAs for the chromosome and 97 coding sequences for the plasmid. Genome-scale metabolic models were reconstructed with the model SEED (10), and the prediction of secondary metabolites and biosynthetic gene clusters was done using antiSMASH (11).

Accession number(s). The assembled complete genome sequence has been deposited in DDBJ/ENA/GenBank under the accession numbers CP028714 and CP028715.

ACKNOWLEDGMENTS

REFERENCES

1. Stecher B, Berry D, Loy A. 2013. Colonization resistance and microbial ecophysiology: using gnotobiotic mouse models and single-cell technology to explore the intestinal jungle. FEMS Microbiol Rev 37:793–829. https://doi.org/10.1111/1574-6976.12024.

2. Brugiroux S, Beutler M, Pfann C, Garzetti D, Ruscheweyh H-J, Ring D, Diehl M, Herp S, Lötscher Y, Hussain S, Bunk B, Pukall R, Huson DH, Münch PC, McHardy AC, McCoy KD, Macpherson AJ, Loy A, Clavel T, Berry D, Stecher B. 2016. Genome-guided design of a defined mouse microbiota that confers colonization resistance against *Salmonella enterica* serovar Typhimurium. Nat Microbiol 2:16215. https://doi.org/10.1038/nmicrobiol.2016.215.

3. Lagkouvardos I, Pukall R, Abt B, Foesel BU, Meier-Kolthoff JP, Kumar N, Bresciani A, Martínez I, Just S, Ziegler C, Brugiroux S, Garzetti D, Wenning M, Bui TPN, Wang J, Hugenholtz F, Plugge CM, Peterson DA, Hornef MW, Baines JF, Smidt H, Walter J, Kristiansen K, Nielsen HB, Haller D, Overmann J, Stecher B, Clavel T. 2016. The Mouse Intestinal Bacterial Collection (miBC) provides host-specific insight into cultured diversity and functional potential of the gut microbiota. Nat Microbiol 1:16131. https://doi.org/10.1038/nmicrobiol.2016.131.

4. Koren S, Walenz BP, Berlin K, Miller JR, Bergman NH, Phillippy AM. 2017. Canu: scalable and accurate long-read assembly via adaptive *k*-mer weighting and repeat separation. Genome Res 27:722–736. https://doi.org/10.1101/gr.215087.116.

5. Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30:2114–2120. https://doi.org/10.1093/bioinformatics/btu170.

6. Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25:1754–1760. https://doi.org/10.1093/bioinformatics/btp324.

7. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. J Mol Biol 215:403–410. https://doi.org/10.1016/S0022-2836(05)80360-2.

8. Carattoli A, Zankari E, García-Fernández A, Voldby Larsen M, Lund O, Villa L, Møller Aarestrup F, Hasman H. 2014. *In silico* detection and typing of plasmids using PlasmidFinder and plasmid multilocus sequence typing. Antimicrob Agents Chemother 58:3895–3903. https://doi.org/10.1128/AAC.02412-14.

9. Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formsma K, Gerdes S, Glass EM, Kubal M, Meyer F, Olsen GJ, Olson R, Osterman AL, Overbeek RA, McNeil LK, Paarmann D, Paczian T, Parrello B, Pusch GD, Reich C, Stevens R, Vassieva O, Vonstein V, Wilke A, Zagnitko O. 2008. The RAST server: Rapid Annotations using Subsystems Technology. BMC Genomics 9:75. https://doi.org/10.1186/1471-2164-9-75.

10. Henry CS, DeJongh M, Best AA, Frybarger PM, Linsay B, Stevens RL. 2010. High-throughput generation, optimization and analysis of genome-scale metabolic models. Nat Biotechnol 28:977–982. https://doi.org/10.1038/nbt.1672.

11. Weber T, Blin K, Duddela S, Krug D, Kim HU, Bruccoleri R, Lee SY, Fischbach MA, Müller R, Wohlleben W, Breitling R, Takano E, Medema MH. 2015. antiSMASH 3.0 – a comprehensive resource for the genome mining of biosynthetic gene clusters. Nucleic Acids Res 43:W237–W243. https://doi.org/10.1093/nar/gkv437.