



Complete Genome Sequences of the *Escherichia coli* Donor Strains ST18 and MFD π ir

 Simon A. Jackson,^{a,b} Bridget J. Fellows,^{a,b}  Peter C. Fineran^{a,b,c}

^aDepartment of Microbiology and Immunology, University of Otago, Dunedin, New Zealand

^bGenetics Otago, University of Otago, Dunedin, New Zealand

^cBio-Protection Research Centre, University of Otago, Dunedin, New Zealand

ABSTRACT *Escherichia coli* ST18 and MFD π ir are donors commonly used to transfer *oriT*_{RP4}-containing plasmids to diverse bacteria via conjugation. ST18 and MFD π ir were constructed via multiple genetic manipulations involving several *E. coli* strains. Here, we used Illumina and Nanopore sequencing to determine the complete genomes of these widely used strains.

A common laboratory method for transferring plasmids to diverse bacteria is via conjugation from an *Escherichia coli* donor strain. Several *E. coli* K-12 derivative strains have been engineered as donors, typically based on mobilization machinery derived from the broad-host-range plasmid RP4 (1–4). To facilitate efficient postmating donor counterselection, several auxotrophic *E. coli* donors, such as ST18, have been constructed (5). ST18 is a 5-aminolevulinic acid auxotroph derivative of S17-1 λ π ir that was derived by deletion of *hemA* (5–8). During mating with S17-1 λ π ir (by implication also with ST18), parts of the donor *E. coli* genome can inadvertently transfer to the recipient strain via Mu phage mobilization or high-frequency recombination (Hfr) transfer from an *oriT* located within the chromosome-integrated RP4 region (*oriT*_{chrRP4}) (2, 3). To overcome these issues, a Mu-free donor strain, MFD π ir, was constructed by removal of the Mu prophages and mutation of *oriT*_{chrRP4} to prevent Hfr transfer (2, 3). Additionally, MFD π ir is a diaminopimelic acid auxotroph and *recA* deletion mutant (2). The construction of ST18 and MFD π ir involved multiple genetic manipulations of several *E. coli* strains from the K-12 lineage (1–3, 5). To confirm the genotypes of these widely used donors, we determined the complete genome sequences of both strains. The ST18 strain reported here has undergone three transfers between research groups since the source laboratory (5). Therefore, there may be divergence between the isolate reported here and other laboratory strains in circulation. The MFD π ir isolate reported here was obtained directly from the laboratory that constructed the strain (2).

Cultures of ST18 and MFD π ir were grown in LB medium (10 g/liter Bacto tryptone, 5 g/liter yeast extract, and 5 g/liter NaCl) at 37°C for 20 h with shaking at 200 rpm and then pelleted by centrifugation. Genomic DNA was extracted using either a DNeasy blood and tissue kit (Qiagen) or a NucleoBond high-molecular-weight (HMW) DNA kit (Macherey-Nagel). Isolated genomic DNA was cleaned and concentrated using AMPure XP beads (Beckman Coulter). Short-read data (Illumina NextSeq custom library preparation, with 150-bp paired-end reads) for both strains and long-read data (Nanopore MinION platform, using the SQK-LSK109 library kit) for MFD π ir were obtained through the Microbial Genome Sequencing Center (MiGS) (Pittsburgh, PA) (9). Nanopore data for ST18 were obtained using a MinION R9.4.1 flow cell (rapid barcoding kit SQK-RBK004), base called using Guppy (Oxford Nanopore Technologies), and demultiplexed using Deepbiner v0.2.0 (10). Raw Illumina reads (ST18, 1,508,686 reads; MFD π ir, 2,736,618 reads) were trimmed and adapter sequences were removed using Trimmomatic v0.39

Citation Jackson SA, Fellows BJ, Fineran PC. 2020. Complete genome sequences of the *Escherichia coli* donor strains ST18 and MFD π ir. Microbiol Resour Announc 9:e01014-20. <https://doi.org/10.1128/MRA.01014-20>.

Editor Catherine Putonti, Loyola University Chicago

Copyright © 2020 Jackson et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Simon A. Jackson, simon.jackson@otago.ac.nz.

Received 31 August 2020

Accepted 13 October 2020

Published 5 November 2020

(11). The Nanopore Fast5 reads (ST18, 47,264 reads, with an N_{50} value of 10,262 bp; MFD*pir*, 275,897 reads, with an N_{50} value of 20,005 bp) were converted to Fastq files using Poretools v0.6.0 (12), and reads of <2,000 bp were discarded. The first 50 bp and last 20 bp of each Nanopore read were trimmed and reads with average quality scores of <9 were removed using NanoFilt v2.6.0 (13). The processed Illumina and Nanopore reads were used for hybrid genome assembly using Unicycler v0.4.9 with default parameters (14). Both ST18 and MFD*pir* genomes were resolved to single closed circular chromosomes (ST18, 4,825,686 bp, with 50.9% GC content; MFD*pir*, 4,686,434 bp, with 50.9% GC content). Coverage was determined by mapping the input reads to the final assemblies using Bowtie 2 (15) and minimap2 (16) for the Illumina and Nanopore reads, respectively. The coverages for ST18 were 74× (Illumina) and 46× (Nanopore) and for MFD*pir* were 143× (Illumina) and 507× (Nanopore).

Data availability. The complete genome sequences have been deposited in NCBI GenBank with accession numbers CP060709 (ST18) and CP060708 (MFD*pir*) and annotated with the NCBI Prokaryotic Genome Annotation Pipeline (PGAP) v4.12 (17, 18). Illumina and Nanopore sequence data have been deposited in the SRA under the accession number PRJNA658821.

ACKNOWLEDGMENTS

We thank Jean-Marc Ghigo (Institut Pasteur) for providing the MFD*pir* strain.

This work was supported by the Royal Society of New Zealand Te Apārangi (RSNZ) Marsden Fund and the Bio-Protection Research Centre (Tertiary Education Commission, New Zealand). B.J.F. was supported by a Duffus Lubecki Postgraduate Scholarship in Applied Science.

REFERENCES

- Simon R, Priefer U, Pühler A. 1983. A broad host range mobilization system for in vivo genetic engineering: transposon mutagenesis in Gram negative bacteria. *Nat Biotechnol* 1:784–791. <https://doi.org/10.1038/nbt1183-784>.
- Ferrieres L, Hemery G, Nham T, Guerout AM, Mazel D, Beloin C, Ghigo JM. 2010. Silent mischief: bacteriophage Mu insertions contaminate products of *Escherichia coli* random mutagenesis performed using suicidal transposon delivery plasmids mobilized by broad-host-range RP4 conjugative machinery. *J Bacteriol* 192:6418–6427. <https://doi.org/10.1128/JB.00621-10>.
- Babic A, Guerout AM, Mazel D. 2008. Construction of an improved RP4 (RK2)-based conjugative system. *Res Microbiol* 159:545–549. <https://doi.org/10.1016/j.resmic.2008.06.004>.
- Miller VL, Mekalanos JJ. 1988. A novel suicide vector and its use in construction of insertion mutations: osmoregulation of outer membrane proteins and virulence determinants in *Vibrio cholerae* requires *toxR*. *J Bacteriol* 170:2575–2583. <https://doi.org/10.1128/jb.170.6.2575-2583.1988>.
- Thoma S, Schobert M. 2009. An improved *Escherichia coli* donor strain for diparental mating. *FEMS Microbiol Lett* 294:127–132. <https://doi.org/10.1111/j.1574-6968.2009.01556.x>.
- Săsărman A, Surdeanu M, Horodniceanu T. 1968. Locus determining the synthesis of δ -aminolevulinic acid in *Escherichia coli* K-12. *J Bacteriol* 96:1882–1884. <https://doi.org/10.1128/JB.96.5.1882-1884.1968>.
- de Lorenzo V, Timmis KN. 1994. Analysis and construction of stable phenotypes in Gram-negative bacteria with Tn5- and Tn10-derived minitransposons. *Methods Enzymol* 235:386–405. [https://doi.org/10.1016/0076-6879\(94\)35157-0](https://doi.org/10.1016/0076-6879(94)35157-0).
- de Lorenzo V, Eltis L, Kessler B, Timmis KN. 1993. Analysis of *Pseudomonas* gene products using *lacIq/Ptrp-lac* plasmids and transposons that confer conditional phenotypes. *Gene* 123:17–24. [https://doi.org/10.1016/0378-1119\(93\)90533-9](https://doi.org/10.1016/0378-1119(93)90533-9).
- Baym M, Kryazhimskiy S, Lieberman TD, Chung H, Desai MM, Kishony R. 2015. Inexpensive multiplexed library preparation for megabase-sized genomes. *PLoS One* 10:e0128036. <https://doi.org/10.1371/journal.pone.0128036>.
- Wick RR, Judd LM, Holt KE. 2018. Deepbiner: demultiplexing barcoded Oxford Nanopore reads with deep convolutional neural networks. *PLoS Comput Biol* 14:e1006583. <https://doi.org/10.1371/journal.pcbi.1006583>.
- 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>.
- Loman NJ, Quinlan AR. 2014. Poretools: a toolkit for analyzing Nanopore sequence data. *Bioinformatics* 30:3399–3401. <https://doi.org/10.1093/bioinformatics/btu555>.
- De Coster W, D'Hert S, Schultz DT, Cruts M, Van Broeckhoven C. 2018. NanoPack: visualizing and processing long-read sequencing data. *Bioinformatics* 34:2666–2669. <https://doi.org/10.1093/bioinformatics/bty149>.
- Wick RR, Judd LM, Gorrie CL, Holt KE. 2017. Unicycler: resolving bacterial genome assemblies from short and long sequencing reads. *PLoS Comput Biol* 13:e1005595. <https://doi.org/10.1371/journal.pcbi.1005595>.
- Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nat Methods* 9:357–359. <https://doi.org/10.1038/nmeth.1923>.
- Li H. 2018. minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* 34:3094–3100. <https://doi.org/10.1093/bioinformatics/bty191>.
- Tatusova T, DiCuccio M, Badretdin A, Chetvernin V, Nawrocki EP, Zaslavsky L, Lomsadze A, Pruitt KD, Borodovsky M, Ostell J. 2016. NCBI Prokaryotic Genome Annotation Pipeline. *Nucleic Acids Res* 44:6614–6624. <https://doi.org/10.1093/nar/gkw569>.
- Haft DH, DiCuccio M, Badretdin A, Brover V, Chetvernin V, O'Neill K, Li W, Chitsaz F, Derbyshire MK, Gonzales NR, Gwazd M, Lu F, Marchler GH, Song JS, Thanki N, Yamashita RA, Zheng C, Thibaud-Nissen F, Geer LY, Marchler-Bauer A, Pruitt KD. 2018. RefSeq: an update on prokaryotic genome annotation and curation. *Nucleic Acids Res* 46:D851–D860. <https://doi.org/10.1093/nar/gkx1068>.