

AUTHOR CONTRIBUTIONS

Patrick J. Biggs, Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Resources, Software, Supervision, Validation, Visualization, Writing – review and editing | Marie Moinet, Data curation, Formal analysis, Investigation, Methodology, Project administration, Supervision, Writing – review and editing | Lynn E. Rogers, Data curation, Formal analysis, Methodology, Project administration, Resources, Writing – review and editing | Megan Devane, Conceptualization, Formal analysis, Funding acquisition, Investigation, Methodology, Writing – review and editing | Richard Muirhead, Conceptualization, Data curation, Methodology, Supervision, Writing – review and editing | Rebecca Stott, Conceptualization, Formal analysis, Investigation, Supervision, Writing – review and editing | Jonathann C. Marshall, Conceptualization, Formal analysis, Funding acquisition, Investigation, Methodology, Supervision, Writing – review and editing | Adrian L. Cookson, Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Writing – original draft, Writing – review and editing

DATA AVAILABILITY

The draft genome assemblies ([JAVGZV010000000–JAVHJK010000000](https://doi.org/10.6084/m9.figshare.24152268)) and short reads were deposited at DDBJ/ENA/GenBank under BioProject accession number [PRJNA1007421](https://doi.org/10.6084/m9.figshare.24152268). A full listing of the source and phylogroup information for the 500 genomes can be found at <https://doi.org/10.6084/m9.figshare.24152268> (17).

REFERENCES

- Till D, McBride G, Ball A, Taylor K, Pyle E. 2008. Large-scale freshwater microbiological study: rationale, results and risks. *J Water Health* 6:443–460. <https://doi.org/10.2166/wh.2008.071>
- Walk ST, Alm EW, Gordon DM, Ram JL, Toranzos GA, Tiedje JM, Whittam TS. 2009. Cryptic lineages of the genus *Escherichia*. *Appl Environ Microbiol* 75:6534–6544. <https://doi.org/10.1128/AEM.01262-09>
- Walk ST. 2015. The "Cryptic" *Escherichia*. *EcoSal Plus* 6. <https://doi.org/10.1128/ecosalplus.ESP-0002-2015>
- Devane ML, Moriarty E, Weaver L, Cookson A, Gilpin B. 2020. Fecal indicator bacteria from environmental sources; strategies for identification to improve water quality monitoring. *Water Res* 185:116204. <https://doi.org/10.1016/j.watres.2020.116204>
- Gilroy R, Ravi A, Getino M, Pursley I, Horton DL, Alikhan N-F, Baker D, Gharbi K, Hall N, Watson M, Adriaenssens EM, Foster-Nyarko E, Jarju S, Secka A, Antonio M, Oren A, Chaudhuri RR, La Ragione R, Hildebrand F, Pallen MJ. 2021. Extensive microbial diversity within the chicken gut microbiome revealed by metagenomics and culture. *PeerJ* 9:e10941. <https://doi.org/10.7717/peerj.10941>
- van der Putten BCL, Matamoros S, Mende DR, Scholl ER, Consortium C, Schultsz C. 2021. *Escherichia ruysiae* sp. nov., a novel gram-stain-negative bacterium, isolated from a faecal sample of an international traveller. *Int J Syst Evol Microbiol* 71:004609. <https://doi.org/10.1099/ijsem.0.004609>
- Liu S, Jin D, Lan R, Wang Y, Meng Q, Dai H, Lu S, Hu S, Xu J. 2015. *Escherichia marmotae* sp. nov., isolated from faeces of *Marmota himalayana*. *Int J Syst Evol Microbiol* 65:2130–2134. <https://doi.org/10.1099/ijms.0.000228>
- Cookson AL, Marshall JC, Biggs PJ, Rogers LE, Collis RM, Devane ML, Stott R, Wilkinson DA, Kamke J, Brightwell G. 2022. Whole-genome sequencing and virulome analysis of *Escherichia coli* isolated from New Zealand environments of contrasting observed land use. *Appl Environ Microbiol* 88:e0027722. <https://doi.org/10.1128/aem.00277-22>
- Andrews A. 2010. FastQC: A Quality Control Tool for High Throughput Sequence Data. Available from: <https://www.bioinformatics.babraham.ac.uk/projects/fastqc>
- Seemann T, Goncalves S, Silva A, Bulach DM, Schultz MB, Kwong JC, Howden BP. 2019. Nullarbor - pipeline to generate complete public health microbiology reports from sequenced isolates. Available from: <https://github.com/tseemann/nullarbor>
- 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>
- Souvorov A, Agarwala R, Lipman DJ. 2018. SKESA: strategic K-MER extension for scrupulous assemblies. *Genome Biol* 19:153. <https://doi.org/10.1186/s13059-018-1540-z>
- 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>
- Beghain J, Bridier-Nahmias A, Le Nagard H, Denamur E, Clermont O. 2018. Clermonttyping: an easy-to-use and accurate in silico method for *Escherichia* genus strain phylotyping. *Microb Genom* 4:e000192. <https://doi.org/10.1099/mgen.0.000192>
- Clermont O, Christenson JK, Denamur E, Gordon DM. 2013. The clermont *Escherichia coli* phylo-typing method revisited: improvement of specificity and detection of new phylo-groups. *Environ Microbiol Rep* 5:58–65. <https://doi.org/10.1111/1758-2229.12019>
- Clermont O, Dixit OVA, Vangchhia B, Condamine B, Dion S, Bridier-Nahmias A, Denamur E, Gordon D. 2019. Characterization and rapid identification of phylogroup G in *Escherichia coli*, a lineage with high virulence and antibiotic resistance potential. *Environ Microbiol* 21:3107–3117. <https://doi.org/10.1111/1462-2920.14713>
- Biggs PJ, Moinet M, Rogers LE, Devane M, Muirhead RW, Stott R, Marshall JC, Cookson AL. 2023. Supplemental Table 1. assembly and Strain Information for 500 genomes sourced from the New Zealand Environment. Available from: <https://doi.org/10.6084/m9.figshare.24152268>