



# Genome Sequence of *Pluralibacter gergoviae* ECO77, a Multireplicon Isolate of Industrial Origin

Edward Cunningham-Oakes,<sup>a</sup> Tom Pointon,<sup>b</sup> Barry Murphy,<sup>b</sup> Thomas R. Connor,<sup>a</sup> Eshwar Mahenthalingam<sup>a</sup>

<sup>a</sup>Microbiomes, Microbes, and Informatics Group, Organisms and Environment Division, Cardiff School of Biosciences, Cardiff University, Cardiff, United Kingdom

<sup>b</sup>Unilever Research and Development, Port Sunlight, Bebington, United Kingdom

**ABSTRACT** In order to expand the limited understanding of the genomics of antimicrobial-resistant industrial bacteria, we report the genome sequence of *Pluralibacter gergoviae* ECO77, a historical contaminant strain of industrial origin. The multireplicon 6.16-Mbp genome of ECO77 consists of a 5.37-Mbp main chromosome, a megaplasmid (605,666 bp), and a large plasmid (182,007 bp).

*Pluralibacter gergoviae* is a member of the *Enterobacteriaceae* family (1) that is occasionally associated with human infection but is commonly associated with cosmetic spoilage (2). In previous studies, we showed that industrial *Pseudomonas aeruginosa* strains uniquely possess large segments of DNA known as megaplasmids, which confer extra genomic capacity with  $\geq 500$  genes and thus may play a role in preservative resistance (3). Furthermore, studies with *Pluralibacter* strains and other Gram-negative bacteria have shown that there are clear links between preservative resistance and outer membrane modifications (4, 5). The genetic changes underpinning these mechanisms of resistance are not well understood but may be elucidated with advances in genomics (6).

Genomic DNA was extracted from pelleted 3-ml overnight cultures of strain ECO77 using a Maxwell 16 tissue DNA purification kit and instrument (Promega) (3). Short-read sequencing was performed with an Illumina NextSeq 500 system, using an NEBNext Ultra II DNA library preparation kit for Illumina (New England Biolabs). Long-read sequencing libraries were produced using a rapid barcoding kit (product number SQK-RBK004; the barcode for ECO77 was BC06) and were sequenced using a MinION R9.4.1 flow cell (Oxford Nanopore Technologies), and bases were called using MinKNOW base-calling software (Oxford Nanopore Technologies). Adaptors were trimmed from short reads using Trim Galore v0.4.4 (7) and from long reads using Porechop v0.2.4 (<https://github.com/rwick/Porechop>). Short-read quality was assessed using FastQC v0.11.7 (8), and long-read quality was assessed using Canu v1.8 (9). A hybrid assembly of 6.16 Mbp was generated using Unicycler v0.4.8 (10). Genome quality was assessed using QUAST v5.0.2 (11). All summary statistics are shown in Table 1.

A total of 5,632 coding DNA sequences were identified in the genome by Prokka v1.14.5 (12). The ECO77 strain also contained 22 rRNAs, 85 tRNAs, and 1 transfer-messenger RNA (tmRNA). The main chromosome was subjected to average nucleotide identity (ANI) analysis against all *P. gergoviae* NCBI RefSeq genomes using PyANI (<https://github.com/widdowquinn/pyani>), which confirmed the species-level identity as *P. gergoviae*.

The plasmid *repA* gene (13) (a highly conserved plasmid gene responsible for copy number), as derived from the Prokka gene annotation of each replicon, was used to identify related homologous plasmids by using BLASTn to align the sequences with the NCBI nucleotide collection. Top hits for the megaplasmid were plasmids from multiple genera within the family *Enterobacteriaceae*, and those for the large plasmid were plasmids from the genus *Klebsiella*.

**Citation** Cunningham-Oakes E, Pointon T, Murphy B, Connor TR, Mahenthalingam E. 2020. Genome sequence of *Pluralibacter gergoviae* ECO77, a multireplicon isolate of industrial origin. *Microbiol Resour Announc* 9:e01561-19. <https://doi.org/10.1128/MRA.01561-19>.

**Editor** Kenneth M. Stedman, Portland State University

**Copyright** © 2020 Cunningham-Oakes 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 Edward Cunningham-Oakes, [cunningham-oakes@cardiff.ac.uk](mailto:cunningham-oakes@cardiff.ac.uk).

**Received** 19 December 2019

**Accepted** 31 January 2020

**Published** 27 February 2020

**TABLE 1** Summary of ECO77 sequencing read and assembly quality metrics

Component	No. of reads ( $N_{50}$ [bp]) with:		Assembly size (bp)	GC content (%)	$N_{50}$ (bp)
	Illumina sequencing (accession no. <a href="#">ERS3910022</a> )	Nanopore sequencing (accession no. <a href="#">ERS3910023</a> )			
Whole genome (accession no. <a href="#">ERS4259183</a> )	2,351,986 (150)	96,499 (62,034)	6,155,741	57.84	5,368,409
Main chromosome			5,368,409	59.27	5,368,409
Megaplasmid			605,666	47.74	294,260
Large plasmid			182,007	53.17	182,007

The contribution of the multireplicon genome of *P. gergoviae* to preservative tolerance and fitness in industrial settings remains to be determined. Nonetheless, the contribution of the ECO77 genome to public databases provides an invaluable reference genome from an industrial background, where genomic characterization is limited.

**Data availability.** Illumina and Nanopore raw sequence reads and the genome have been deposited in the European Nucleotide Archive (ENA) under ENA project accession number [PRJEB34950](#) and BioSample accession numbers [ERS3910022](#), [ERS3910023](#), and [ERS4259183](#).

## ACKNOWLEDGMENTS

E.C.-O. was supported by the Biotechnology and Biological Sciences Research Council South West doctoral training partnership and Unilever R&D industrial CASE studentship awards (grant BB/M009122/1).

Sequencing was carried out at the Cardiff School of Biosciences Genomics Research Hub. All analysis was performed using the Cloud Infrastructure for Microbial Bioinformatics, funded by Medical Research Council grant MR/L015080/1.

## REFERENCES

- Brady C, Cleenwerck I, Venter S, Coutinho T, de Vos P. 2013. Taxonomic evaluation of the genus *Enterobacter* based on multilocus sequence analysis (MLSA). *Syst Appl Microbiol* 36:309–319. <https://doi.org/10.1016/j.syapm.2013.03.005>.
- Périmé M, Pagès JM, Davin-Regli A. 2014. *Enterobacter gergoviae* adaptation to preservatives commonly used in cosmetic industry. *Int J Cosmet Sci* 36:386–395. <https://doi.org/10.1111/ics.12140>.
- Weiser R, Green AE, Bull MJ, Cunningham-Oakes E, Jolley KA, Maiden MCJ, Hall AJ, Winstanley C, Weightman AJ, Donoghue D, Amezcua A, Connor TR, Mahenthalingam E. 2019. Not all *Pseudomonas aeruginosa* are equal: strains from industrial sources possess uniquely large multi-replicon genomes. *Microb Genomics*. <https://doi.org/10.1099/mgen.0.000276>.
- Périmé M, Philippe N, Condell O, Fanning S, Pagès JM, Davin-Regli A. 2015. Phenotypic changes contributing to *Enterobacter gergoviae* biocide resistance. *Lett Appl Microbiol* 61:121–129. <https://doi.org/10.1111/lam.12435>.
- Rushton L, Sass A, Baldwin A, Dowson CG, Donoghue D, Mahenthalingam E. 2013. Key role for efflux in the preservative susceptibility and adaptive resistance of *Burkholderia cepacia* complex bacteria. *Antimicrob Agents Chemother* 57:2972–2980. <https://doi.org/10.1128/AAC.00140-13>.
- Boolchandani M, D'Souza AW, Dantas G. 2019. Sequencing-based methods and resources to study antimicrobial resistance. *Nat Rev Genet* 20:356–370. <https://doi.org/10.1038/s41576-019-0108-4>.
- Krueger F. 2015. Trim Galore!: a wrapper tool around Cutadapt and FastQC to consistently apply quality and adapter trimming to FastQ files. [https://www.bioinformatics.babraham.ac.uk/projects/trim\\_galore](https://www.bioinformatics.babraham.ac.uk/projects/trim_galore).
- Andrews S, Krueger F, Semonds-Pichon A, Biggins F, Wingett S. 2015. FastQC: a quality control tool for high throughput sequence data. <https://www.bioinformatics.babraham.ac.uk/projects/fastqc>.
- 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>.
- 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>.
- Gurevich A, Saveliev V, Vyahhi N, Tesler G. 2013. QUILT: quality assessment tool for genome assemblies. *Bioinformatics* 29:1072–1075. <https://doi.org/10.1093/bioinformatics/btt086>.
- Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 30:2068–2069. <https://doi.org/10.1093/bioinformatics/btu153>.
- Petersen J. 2011. Phylogeny and compatibility: plasmid classification in the genomics era. *Arch Microbiol* 193:313–321. <https://doi.org/10.1007/s00203-011-0686-9>.