AMERICAN SOCIETY FOR MICROBIOLOGY | Microbiology Resource Announcements

GENOME SEQUENCES

Draft Genome Sequences of Antibiotic-Resistant *Escherichia coli* Isolates from U.S. Wastewater Treatment Plants

Vicente Gomez-Alvarez,[a] Jill Hoelle[a]

[a]U.S. Environmental Protection Agency, Office of Research and Development, Cincinnati, Ohio, USA

ABSTRACT The spread of antibiotic-resistant microorganisms is a major public health concern. Here, we report the draft genome sequences of three *Escherichia coli* isolates from primary effluent collected from geographically dispersed U.S. wastewater treatment plants (WWTPs). Genomic analysis confirms the presence of genes encoding resistance to a broad spectrum of antibiotics.

A survey for antibiotic-resistant (AR) *Escherichia coli* was undertaken by collecting samples of primary effluents from geographically dispersed U.S. wastewater treatment plants (WWTPs) (1). *E. coli*, a Gram-negative bacterium and member of the family *Enterobacteriaceae*, is a commensal inhabitant of the gastrointestinal tract (2) and is commonly used as an indicator of fecal pollution (1). AR *E. coli* in wastewater could provide information on the occurrence and dissemination of sequence types (ST) within a given community or population (1, 3).

Strains EPA165, EPA233, and EPA336 were isolated from effluent collected from primary clarifiers at WWTPs located in New Jersey, California, and Maryland, as previously described by Hoelle et al. (1). Briefly, samples were filtered and transferred to membrane fecal coliform (m-FC) agar (Becton, Dickinson, Franklin Lakes, NJ) supplemented with imipenem (1 mg/liter), ciprofloxacin (4 mg/liter), cefotaxime (4 mg/liter), or ceftazidime (16 mg/liter) and incubated at 44.5°C. Single colonies were transferred to m-FC plates supplemented with 4-methylumbelliferyl-β-D-glucuronide (MUG). The DNA from MUG-positive colonies was extracted using the UltraClean DNA microbial isolation kit following the manufacturer's instructions (Mo Bio Laboratories, Solana Beach, CA). Genomic libraries were prepared using the Nextera XT index kit v2 set A and sequenced on the HiSeq 4000 platform (Illumina, Inc., San Diego, CA) with a HiSeq 3000/4000 PE cluster kit (2 × 150 bp).

A total of 29,806,503 paired-end reads (EPA165, 9,001,282; EPA223, 11,064,325; EPA336, 9,740,896) were generated. Prior to assembly, libraries were cleaned from adapters and phiX artifacts, error corrected, normalized (≤100×), and filtered to a minimum length of 100 nucleotides (nt) using the software package BBMap v38.22 (ktrim=r k=23 mink=11 hdist=1 tbo tpe maxns=0 trimq=10 qtrim=r maq=12 minlength=100 ecco=t eccc=t ecct=t target=100) (4). A reference-assisted *de novo* assembly approach was used to assemble the processed reads using Unicycler v0.4.7 (5). Average nucleotide identity (ANI), a similarity index between two genomes (6), was determined using FastANI v1.1 (7). The *in silico* multilocus sequence type (MLST) based on seven alleles (*adk*, *fumC*, *gyrB*, *icd*, *mdh*, *purA*, and *recA*) was obtained using mlst v2.16.1 (8, 9) and phylotyping was performed with the EzClermont Web tool (10), serotyping (O-antigen and flagellin genes) with SerotypeFinder v2.0 (11), subtyping (*fumC* and *fimH* alleles) with CHTyper v1.0 (12), antibiotic resistance gene determination with ResFinder v3.1 (13), and chromosomal point mutation determination with PointFinder v3.1 (14). Default parameters were used for all software unless otherwise specified. Genome quality and statistics were estimated with BBMap and annotated with Prokka v1.13.1 (15) (Table 1).

Citation Gomez-Alvarez V, Hoelle J. 2019. Draft genome sequences of antibiotic-resistant *Escherichia coli* isolates from U.S. wastewater treatment plants. Microbiol Resour Announc 8:e00351-19. https://doi.org/10.1128/MRA.00351-19.

Editor David Rasko, University of Maryland School of Medicine

Address correspondence to Vicente Gomez-Alvarez, Gomez-Alvarez.Vicente@epa.gov.

Received 25 March 2019
Accepted 9 May 2019
Published 6 June 2019

TABLE 1 Summary statistics of whole-genome assemblies

Strain[a]	Coverage (×)	Genetic element[b]	No. of contigs	Assembly size (bp)	Contig N_{50} (bp)	G+C content (%)	Gene annotation data (no.)[c]				PT[d]	ST[e]	Serotype[f]	Subtype[g]	Reference genome	GenBank accession no.
							Genes	CDS	rRNAs	tRNAs						
EPA165	255	Chromosome	110	5,040,423	288,542	50.53	4,816	4,714	15	86	D	973	O21:H15	187:95	GCA_001542545.1	SJTD00000000
		Plasmid 1	1	4,936	4,936	47.81										
		Plasmid 2	1	4,074	4,074	49.93										
		Plasmid 3	4	72,427	58,169	50.21										
EPA233	328	Chromosome	59	4,669,471	271,035	50.79	4,559	4,460	14	84	B1	156	O9:H9	29:38	GCA_000010385.1	SJTC00000000
		Plasmid 1	1	36,144	36,144	42.66										
		Plasmid 2	10	126,415	22,867	50.95										
EPA336	293	Chromosome	66	4,712,839	477,434	50.83	4,599	4,501	10	87	B1	205	O100:H12	23:54	GCA_001007915.1	SJTB00000000
		Plasmid 1	1	8,379	8,379	55.48										
		Plasmid 2	1	113,422	113,422	46.17										

[a] The numbers of paired-end reads were 9,001,282 (EPA165), 11,064,325 (EPA223), and 9,740,896 (EPA336).

[b] Plasmid contig identifiers for EPA165 are contig000025 (plasmid 1), contig000027 (plasmid 2), and contig000022, contig000024, contig000015, and contig000034 (plasmid 3); for EPA223, contig000027 (plasmid 1) and contig000038, contig000036, contig000032, contig000033, contig000023, contig000053, contig000054, contig000037, contig000048, and contig000031 (plasmid 2); and for EPA336, contig000019 (plasmid 1) and contig000013 (plasmid 2).

[c] CDS, coding sequences.

[d] PT, phylotype.

[e] ST, sequence type (*in silico* MLST; *adk*, *fumC*, *gyrB*, *icd*, *mdh*, *purA*, *recA*).

[f] O, O-antigen; H, flagellin gene.

[g] *fumC*:*fimH* alleles.

ANI calculations revealed an average genome similarity of 99.10% between strains EPA233 and EPA336, which were both closely related to EPA165, with 97.30% similarity. Despite their high genome similarity, further analysis confirmed the clustering of the three strains to a different sequence type, serotype, and subtype complex (Table 1). EPA233 and EPA336 were identified as members of phylotype B1 and EPA165 was identified as phylotype D. Pangenome analysis confirmed the presence of genes encoding resistance to a broad spectrum of antibiotics, such as β-lactams (bla_{CMY-2} and bla_{TEM-1B}), the multiple drug resistance (MDR) family (*mdfA*), aminoglycoside [*aac(6′)-Ib3*, *aph(6)-Id*, *aadA1*, and *aph(3″)-Ib*], fluoroquinolone [*aac(6′)-Ib-cr*], phenicol (*catA1* and *catB3*), sulfonamide (*sul1* and *sul2*), tetracycline (*tetA* and *tetB*), and trimethoprim (*dfrA1*). Furthermore, chromosomal point mutations associated with antimicrobial resistance in *gyrA* and *parC* genes were detected. The plasmids carried antimicrobial resistance genes and contained putative conjugal transfer modules, including an *oriT*-like region, relaxase, and components of the type IV coupling protein (T4CP) and type IV secretion system (T4SS) (16).

Data availability. This whole-genome shotgun project has been deposited in DDBJ/ENA/GenBank under the accession numbers listed in Table 1. The raw sequence reads have been submitted to the NCBI SRA under the accession numbers SRR8648405, SRR8648406, and SRR8648407. The versions described in this paper are the first versions.

ACKNOWLEDGMENTS

We thank James R. Johnson, Brian D. Johnston, Laura Boczek, and Hodon Ryu for valuable assistance for this project.

The opinions expressed are those of the authors and do not necessarily reflect the official positions and policies of the U.S. EPA. Any mention of product or trade names does not constitute recommendation for use by the U.S. EPA.

REFERENCES

1. Hoelle J, Johnson JR, Johnston BD, Kinkle B, Boczek L, Ryu H, Hayes S. 2019. Survey of US wastewater for carbapenem-resistant *Enterobacteriaceae*. J Water Health 17:219–226. https://doi.org/10.2166/wh.2019.165.
2. Vila J, Sáez-López E, Johnson JR, Römling U, Dobrindt U, Cantón R, Giske CG, Naas T, Carattoli A, Martínez-Medina M, Bosch J, Retamar P, Rodríguez-Baño J, Baquero F, Soto SM. 2016. *Escherichia coli*: an old friend with new tidings. FEMS Microbiol Rev 40:437–463. https://doi.org/10.1093/femsre/fuw005.
3. Boczek L, Rice EW, Johnston B, Johnson JR. 2007. Occurrence of antibiotic-resistant uropathogenic *Escherichia coli* clonal group A in wastewater effluents. Appl Environ Microbiol 73:4180–4184. https://doi.org/10.1128/AEM.02225-06.
4. Bushnell B. 2016. BBMap short-read aligner, and other bioinformatics tools. http://sourceforge.net/projects/bbmap/.
5. Wick RR, Judd LM, Gorrie CL, Holt KE. 2017. Unicycler: resolving bacterial genome assemblies from short and long sequencing read. PLoS Comput Biol 13:e1005595. https://doi.org/10.1371/journal.pcbi.1005595.
6. Goris J, Konstantinidis KT, Klappenbach JA, Coenye T, Vandamme P, Tiedje JM. 2007. DNA-DNA hybridization values and their relationship to whole-genome sequence similarities. Int J Syst Evol Microbiol 57:81–91. https://doi.org/10.1099/ijs.0.64483-0.
7. Jain C, Rodriguez-R LM, Phillippy AM, Konstantinidis KT, Aluru S. 2018. High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. Nat Commun 9:5114. https://doi.org/10.1038/s41467-018-07641-9.
8. Seemann T. 2019. mlst. https://github.com/tseemann/mlst.
9. Jolley KA, Maiden MC. 2010. BIGSdb: scalable analysis of bacterial genome variation at the population level. BMC Bioinformatics 11:595. https://doi.org/10.1186/1471-2105-11-595.
10. Waters N, Brennan F, Holmes A, Abram F, Pritchard L. 2018. Easily phylotyping *E. coli* via the EzClermont Web app and command-line tool. bioRxiv. https://doi.org/10.1101/317610.
11. Joensen KG, Tetzschner AM, Iguchi A, Aarestrup FM, Scheutz F. 2015. Rapid and easy *in silico* serotyping of *Escherichia coli* using whole genome sequencing (WGS) data. J Clin Microbiol 53:2410–2426. https://doi.org/10.1128/JCM.00008-15.
12. Roer L, Johannesen TB, Hansen F, Stegger M, Tchesnokova V, Sokurenko E, Garibay N, Allesøe R, Thomsen MCF, Lund O, Hasman H, Hammerum AM. 2018. CHTyper, a Web tool for subtyping of extraintestinal pathogenic *Escherichia coli* based on the *fumC* and *fimH* alleles. J Clin Microbiol 56:e00063-18. https://doi.org/10.1128/JCM.00063-18.
13. Zankari E, Hasman H, Cosentino S, Vestergaard M, Rasmussen S, Lund O, Aarestrup FM, Larsen MV. 2012. Identification of acquired antimicrobial resistance genes. J Antimicrob Chemother 67:2640–2644. https://doi.org/10.1093/jac/dks261.
14. Zankari E, Allesøe R, Joensen KG, Cavaco LM, Lund O, Aarestrup FM. 2017. PointFinder: a novel Web tool for WGS-based detection of antimicrobial resistance associated with chromosomal point mutations in bacterial pathogens. J Antimicrob Chemother 72:2764–2768. https://doi.org/10.1093/jac/dkx217.
15. Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. Bioinformatics 30:2068–2069. https://doi.org/10.1093/bioinformatics/btu153.
16. Li X, Xie Y, Liu M, Tai C, Sun J, Deng Z, Ou HY. 2018. oriTfinder: a Web-based tool for the identification of origin of transfers in DNA sequences of bacterial mobile genetic elements. Nucleic Acids Res 46:W229–W234. https://doi.org/10.1093/nar/gky352.