AMERICAN SOCIETY FOR MICROBIOLOGY | Microbiology Resource Announcements

GENOME SEQUENCES

Complete Genome Sequences of Two Shiga Toxin-Producing *Escherichia coli* Strains Isolated from Crows

Michelle Q. Carter,[a] Antares Pham,[a] Diana K. Carychao,[a] Michael B. Cooley[a]

[a]Western Regional Research Center, Produce Safety, and Microbiology Research Unit, U.S. Department of Agriculture, Agricultural Research Service, Albany, California, USA

ABSTRACT *Escherichia coli* strains RM9088 and RM10410 were isolated from crows near a leafy greens-growing region in California in April and July 2009, respectively. Both strains carry genes encoding Shiga toxins and other virulence factors in enteric pathogens. Here, we report the complete genome sequences of RM9088 and RM10410.

Shiga toxin-producing *Escherichia coli* (STEC) is one of the main bacterial causal agents of foodborne illness outbreaks associated with fresh produce (1). Although STEC O157:H7 has been considered the most frequent cause of STEC-associated outbreaks, recent studies suggest that non-O157 STEC strains are causing a large number of human infections worldwide (2–4). STEC naturally resides in ruminant animals, primarily cattle; however, diverse STEC strains have been isolated from birds (5–7), implying that birds might be an environmental source of STEC transmission. To better understand the pathogenicity of avian STEC, we sequenced the genomes of the two isolates obtained from crows.

Strains RM9088 and RM10410 were isolated by cloacal swab, as described previously (8). Genomic DNA was extracted from the mid-exponential-phase cultures grown in LB broth, as described previously (9). Genomic libraries were prepared according to the PacBio 20-kb library standard protocol (10) using the SMRTbell DNA template prep kit 3.0, followed by size selection with the BluePippin size selection system (Sage Science, Inc.) and then template binding with the P6v2 kit. DNA sequencing was performed on an RS II instrument (Pacific Biosciences) with P6-C4 sequencing chemistry and a 360-min data collection protocol. The sequence reads were filtered with PreAssembler filter prior to *de novo* assembly with RS_HGAP_Assembly v.3. Detailed sequencing metrics and filter parameters for each strain are listed in Table 1. The closed genomes were annotated using the NCBI Prokaryotic Genome Annotation Pipeline (PGAP) (11).

The RM9088 genome is composed of a 5,270,611-bp chromosome and two plasmids, encoding a total of 5,500 coding DNA sequences (CDSs). The RM10410 genome is composed of only a 5,227,472-bp chromosome, encoding 5,114 CDSs (Table 1). The serotypes of RM9088 and RM10410 were determined to be O109:H48 and O113:H4, respectively, using SerotypeFinder 2.0 (12), with the default settings (thresholds for percent identity [%ID], 85%, and minimum length, 60%). The sequence types (STs) of RM9088 and RM10410 are ST339 and ST10, respectively, using the Warwick scheme (13). *In silico* phylo-typing using the Clermont method (14) placed both strains in phylogroup A.

The *stx* genes in strain RM9088 encode Stx1a, located on a 43,766-bp prophage (chromosome positions, base pairs 4689207 to 4732972) that was identified using PHASTER (15, 16). The p1RM9088 plasmid (167,256 bp) is a typical pEHEC (the large virulence plasmid of enterohemorrhagic *E. coli* [EHEC]) (17) containing genes (*hlyCABD*) encoding enterohemolysin. Interestingly, this plasmid also carries genes such as EAST1 (annotated as *astA* in RM9088) (GenBank accession number AB042002) and *sta1* (GenBank accession number AJ555214) encoding heat-stable toxins, a common viru-

Citation Carter MQ, Pham A, Carychao DK, Cooley MB. 2019. Complete genome sequences of two Shiga toxin-producing *Escherichia coli* strains isolated from crows. Microbiol Resour Announc 8:e01082-19. https://doi.org/10.1128/MRA.01082-19.

Editor David Rasko, University of Maryland School of Medicine

Address correspondence to Michelle Q. Carter, michelle.carter@usda.gov.

Received 30 August 2019
Accepted 14 October 2019
Published 7 November 2019

TABLE 1 Genome attributes of two avian STEC strains

Feature	Value for strain: RM9088	RM10410
Serotype[a]	O109:H48	O113:H4
PacBio sequencing metrics		
No. of reads	81,165	67,867
Avg read length (bp)	6,931	10,205
No. of bases	1,351,682,302	1,354,076,155
Sequence coverage (×)	152	200
PreAssembler Filter v.1		
Minimum subread length	1,000	500
Minimum polymerase read quality	0.80	0.80
Minimum polymerase read length	100	100
Chromosome		
Size (bp)	5,270,611	5,227,472
G+C content (%)	50.8	50.5
No. of CDSs	5,188	5,114
No. of tRNAs	94	91
No. of rRNAs	22	22
No. of prophages	14	11
stx gene(s)	stx_{1a}	stx_{1a}, stx_{2d}
Plasmids		
pEHEC	p1RM9088	None
Size (bp)	167,256	None
G+C content (%)	45.9	None
No. of CDSs	198	None
Second plasmid	p2RM9088	None
Size (bp)	86,529	None
G+C content (%)	49.2	None
No. of CDSs	114	None

[a] The serotype of each strain was determined *in silico*.

lence factor of enterotoxigenic *E. coli*. No known STEC virulence genes were identified on the second plasmid, p2RM9088 (86,529 bp). The two sets of *stx* genes in strain RM10410 encode Stx1a and Stx2d, respectively. The stx_{1a} gene is located on a 94,727-bp prophage (chromosome positions, base pairs 1786478 to 1881204); the stx_{2d} gene is located on a 59,845-bp prophage (chromosome positions, base pairs 2420694 to 2480538). A search of additional virulence factors using VirulenceFinder 2.0 (18), with the default settings (threshold for %ID, 90%; minimum length, 60%), failed to identify the pathogenicity island locus of enterocyte effacement (LEE) or genes encoding any type III secretion effectors in either of the two genomes. Since strain RM10410 has no plasmids, it lacks genes encoding enterohemolysin.

Data availability. The sequences described in this study are available under BioProject accession number PRJNA557687. The GenBank accession numbers are CP042298, CP042296, CP042297, and CP042350 for the RM9088 chromosome, plasmid p1RM9088, plasmid p2RM9088, and RM10410 chromosome, respectively. The raw reads are available under Sequence Read Archive (SRA) accession numbers SRR9953605 and SRR9953217 for RM9088 and RM10410, respectively.

ACKNOWLEDGMENT

This research work was supported by the U.S. Department of Agriculture Agricultural Research Service under grant CRIS 2030-42000-050-00D.

REFERENCES

1. Rangel JM, Sparling PH, Crowe C, Griffin PM, Swerdlow DL. 2005. Epidemiology of *Escherichia coli* O157:H7 outbreaks, United States, 1982–2002. Emerg Infect Dis 11:603–609. https://doi.org/10.3201/eid1104.040739.
2. Scallan E, Hoekstra RM, Angulo FJ, Tauxe RV, Widdowson MA, Roy SL, Jones JL, Griffin PM. 2011. Foodborne illness acquired in the United States—major pathogens. Emerg Infect Dis 17:7–15. https://doi.org/10.3201/eid1701.p11101.
3. Gould LH, Mody RK, Ong KL, Clogher P, Cronquist AB, Garman KN,

Lathrop S, Medus C, Spina NL, Webb TH, White PL, Wymore K, Gierke RE, Mahon BE, Griffin PM, Emerging Infections Program FoodNet Working Group. 2013. Increased recognition of non-O157 Shiga toxin-producing *Escherichia coli* infections in the United States during 2000–2010: epidemiologic features and comparison with *E. coli* O157 infections. Foodborne Pathog Dis 10:453–460. https://doi.org/10.1089/fpd.2012.1401.

4. Luna-Gierke RE, Griffin PM, Gould LH, Herman K, Bopp CA, Strockbine N, Mody RK. 2014. Outbreaks of non-O157 Shiga toxin-producing *Escherichia coli* infection: USA. Epidemiol Infect 142:2270–2280. https://doi.org/10.1017/S0950268813003233.
5. Fadel HM, Afifi R, Al-Qabili DM. 2017. Characterization and zoonotic impact of Shiga toxin producing *Escherichia coli* in some wild bird species. Vet World 10:1118–1128. https://doi.org/10.14202/vetworld.2017.1118-1128.
6. Gioia-Di Chiacchio RM, Cunha MPV, de Sá LRM, Davies YM, Pereira CBP, Martins FH, Munhoz DD, Abe CM, Franzolin MR, Dos Santos LF, Guth BEC, Elias WP, Piazza RMF, Knöbl T. 2018. Novel hybrid of typical enteropathogenic Escherichia coli and Shiga-toxin-producing *E. coli* (tEPEC/STEC) emerging from pet birds. Front Microbiol 9:2975. https://doi.org/10.3389/fmicb.2018.02975.
7. van Hoek A, van Veldhuizen JNJ, Friesema I, Coipan C, Rossen JWA, Bergval IL, Franz E. 2019. Comparative genomics reveals a lack of evidence for pigeons as a main source of stx_{2f}-carrying *Escherichia coli* causing disease in humans and the common existence of hybrid Shiga toxin-producing and enteropathogenic *E. coli* pathotypes. BMC Genomics 20:271. https://doi.org/10.1186/s12864-019-5635-z.
8. Cooley MB, Jay-Russell M, Atwill ER, Carychao D, Nguyen K, Quinones B, Patel R, Walker S, Swimley M, Pierre-Jerome E, Gordus AG, Mandrell RE. 2013. Development of a robust method for isolation of Shiga toxin-positive *Escherichia coli* (STEC) from fecal, plant, soil and water samples from a leafy greens production region in California. PLoS One 8:e65716. https://doi.org/10.1371/journal.pone.0065716.
9. Miller WG, Pearson BM, Wells JM, Parker CT, Kapitonov VV, Mandrell RE. 2005. Diversity within the *Campylobacter jejuni* type I restriction-modification loci. Microbiology 151:337–351. https://doi.org/10.1099/mic.0.27327-0.
10. Pacific Biosciences. 2015. Procedure and checklist: 20 kb template preparation using BluePippin size-selection system. Pacific Biosciences, Menlo Park, CA. https://www.pacb.com/wp-content/uploads/2015/09/Procedure-Checklist-20-kb-Template-Preparation-Using-BluePippin-Size-Selection.pdf.
11. 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.
12. Joensen KG, Tetzschner AM, Iguchi A, Aarestrup FM, Scheutz F. 2015. Rapid and easy *in silico* serotyping of *Escherichia coli* isolates by use of whole-genome sequencing data. J Clin Microbiol 53:2410–2426. https://doi.org/10.1128/JCM.00008-15.
13. Clermont O, Gordon D, Denamur E. 2015. Guide to the various phylogenetic classification schemes for *Escherichia coli* and the correspondence among schemes. Microbiology 161:980–988. https://doi.org/10.1099/mic.0.000063.
14. 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.
15. Arndt D, Grant JR, Marcu A, Sajed T, Pon A, Liang Y, Wishart DS. 2016. PHASTER: a better, faster version of the PHAST phage search tool. Nucleic Acids Res 44:W16–W21. https://doi.org/10.1093/nar/gkw387.
16. Zhou Y, Liang Y, Lynch KH, Dennis JJ, Wishart DS. 2011. PHAST: a fast phage search tool. Nucleic Acids Res 39:W347–W352. https://doi.org/10.1093/nar/gkr485.
17. Cooper KK, Mandrell RE, Louie JW, Korlach J, Clark TA, Parker CT, Huynh S, Chain PS, Ahmed S, Carter MQ. 2014. Comparative genomics of enterohemorrhagic *Escherichia coli* O145:H28 demonstrates a common evolutionary lineage with *Escherichia coli* O157:H7. BMC Genomics 15:17. https://doi.org/10.1186/1471-2164-15-17.
18. Joensen KG, Scheutz F, Lund O, Hasman H, Kaas RS, Nielsen EM, Aarestrup FM. 2014. Real-time whole-genome sequencing for routine typing, surveillance, and outbreak detection of verotoxigenic *Escherichia coli*. J Clin Microbiol 52:1501–1510. https://doi.org/10.1128/JCM.03617-13.