

Complete genome sequence of *Escherichia coli* strain GCU-H1 isolated from a captive-bred common bottlenose dolphin (*Tursiops truncatus*) in Korea[§]

Eun Jeong Park^{1†} , Jong Beom Na^{1†} , Seon Young Park^{2†} , Seungbin Song³ , Young Min Lee⁴ ,
Soojin Lim¹ , Keeman Lee¹ , Ye Bin Kim¹ , Tae Seon Cha¹ , Seunghui Lee¹ , Jee Eun Han⁵ ,
Kyunglee Lee^{4*} , and Ji Hyung Kim^{1*} 

¹Department of Food Science and Bioengineering, College of Bionano Technology, Gachon University, Seongnam 13120, Republic of Korea

²Veterinary Drugs and Biologics Division, Animal and Plant Quarantine Agency, Gimcheon 39660, Republic of Korea

³i-SENS, Inc., Seoul 06646, Republic of Korea

⁴Cetacean Research Institute, National Institute of Fisheries Science, Ulsan 44780, Republic of Korea

⁵Institute for Veterinary Biomedical Science, Department of Veterinary Medicine, Kyungpook National University, Daegu 41566, Republic of Korea

(Received July 11, 2025; Revised August 5, 2025; Accepted August 8, 2025)

Although *Escherichia coli* is considered a commensal bacterium in humans and mammals, including cetaceans, some specific toxin-producing or sequence type (ST) strains are known to cause clinical infections. Notably, the ST131 *E. coli* is a globally prevalent, multidrug-resistant clone responsible for potential risks to public and animal health. Here, we report the complete genome of ST131 *E. coli* strain GCU-H1, isolated from the feces of a captive common bottlenose dolphin (*Tursiops truncatus*) bred in Korea. The strain GCU-H1 contained a 5,095,519 bp genome with a G + C content of 50.8%, which comprises a single chromosome and three plasmids. Multilocus sequence typing analysis confirmed the strain as ST131/PST43, specifically belonging to the C1/H30R subclade. Moreover, several genes coding for virulence, antimicrobial resistance, and prophages were also detected. This study will provide fundamental information on the possible emergence of potentially pathogenic *E. coli* in captive marine mammals, thus contributing to our understanding of its dissemination in dolphinarium and public health.

Keywords: *Escherichia coli*, *Tursiops truncatus*, ST131, marine mammal

The global dissemination of zoonotic antimicrobial-resistant (AMR) bacteria is a serious threat to human and animal health, and their emergence in marine mammals, including cetaceans, is concerning (Dold, 2014). Furthermore, the emergence of AMR bacteria in marine mammals is more plausible in captive than in free-ranging individuals because of their frequent contact with humans and the antimicrobials used for prophylactic or curative treatment. Several cases of fatal infections caused by zoonotic AMR bacteria have been reported in captive dolphins bred in a dolphinarium (Park et al., 2020); however, the mechanisms of AMR in these bacterial isolates have not been adequately investigated owing to limited genetic information. Therefore, there is an urgent need for genomic information on AMR bacteria in captive cetaceans to clarify their potential health risks to marine mammals and the humans they encounter.

As in humans and other animals, *Escherichia coli* has been isolated from various aquatic mammals and is one of the most prevalent groups of microorganisms in the gastrointestinal

[†]These authors contributed equally to this work.

*For correspondence. (K. Lee) E-mail: moby19@korea.kr;
Tel.: +82-52-270-0970; Fax: +82-52-270-0913 /

(J.H. Kim) E-mail: kzh81@gachon.ac.kr;

Tel.: +82-31-750-5383; Fax: +82-31-750-5389

[§]Supplemental material for this article may be found at
<http://www.kjom.org/main.html>

tracts of both free-ranging and captive cetaceans (Bai et al., 2024; Fulham et al., 2020). Considering that certain toxin-producing or sequence type (ST) strains of *E. coli* have virulence traits that can cause disease in various host species, the bacteria have a strong potential to cause opportunistic or fatal infections in marine mammals. However, little is known about the diversity and prevalence of *E. coli* in those animals (Gross et al., 2022; Melendez et al., 2019; Sellera et al., 2022).

Among the various STs of potentially pathogenic *E. coli*, ST131 is a globally prevalent AMR clone responsible for human and animal infections and represents a common extraintestinal pathogenic *E. coli* (ExPEC) associated with urinary tract and bloodstream infections (Hassuna et al., 2024; Pitout and DeVinney, 2017). Although the geographical distribution of ST131 is not fully understood, it is commonly found in clinical infections in Europe, North America, Canada, Japan, and Korea (Bevan et al., 2017). Moreover, the potential presence of ST131 *E. coli* has been reported in marine sediments and wild marine mammals (Manageiro et al., 2015; Mora et al., 2018; Vignaroli et al., 2013). However, its presence in captive-bred marine mammals has not yet been reported.

The strain GCU-H1 was isolated from the feces of a captive-bred *T. truncatus* in a dolphinarium in Korea in 2024. The bacterial isolate was cultured overnight on tryptic soy

agar (TSA; Difco) at 37°C, and preliminarily identified as *E. coli* based on the 16S rRNA sequence analysis. Then, the bacterial genomic DNA was extracted using the QIAamp DNA Mini Kit (Qiagen) and sequenced using a hybrid approach combining the PacBio RSII System (Pacific Biosciences) and Illumina HiSeq X-10 platform (Illumina), as previously described (Lim et al., 2024). The Microbial Genome Assembly application (<http://www.pacb.com/products-and-services/pacbio-systems/sequel/sequel-software/>) was used to assemble PacBio long reads (599,706,396 bp; 69,247 reads) and Illumina paired-end reads (905,820,202 bp; 6,006,018 reads) for error correction using Pilon v1.21 (Walker et al., 2014). Genome annotation was performed using the NCBI Prokaryotic Genome Annotation Pipeline (https://www.ncbi.nlm.nih.gov/genome/annotation_prok/).

The sequenced GCU-H1 genome comprised 5,095,519 bp (50.8% G + C content, 103.7× coverage), consisting of a single chromosome (5,052,029 bp, 50.8% G + C content) along with three plasmids, designated pGCU-H1-1 (33,826 bp, 41.6% G + C content), pGCU-H1-2 (4,946 bp, 48.1% G + C content), and pGCU-H1-3 (4,718 bp, 46.9% G + C content), comprising 5,061 genes, including 4,728 protein-coding sequences, 22 rRNAs, 90 tRNAs, 5 noncoding RNAs, and 216 pseudogenes (Fig. 1, Table 1). Comparative genome analysis

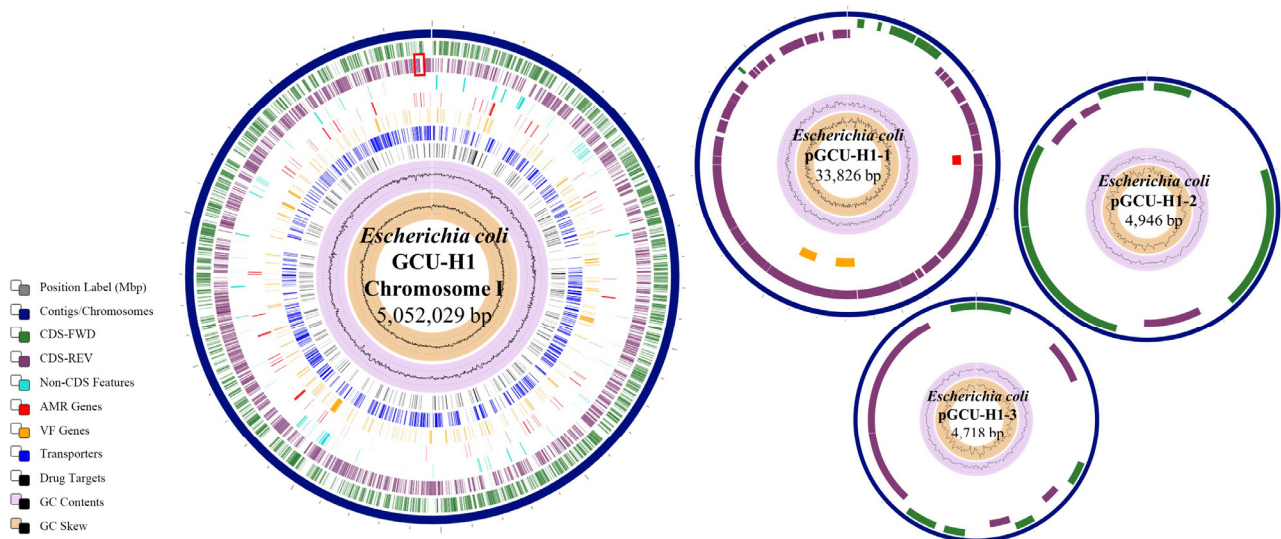


Fig. 1. Circular genome maps of the ST131 *E. coli* strain GCU-H1. The genome was visualized and classified into subsystems using the Comprehensive Genome Analysis of Pathosystems Resource Integration Center (PATRIC) server v.3.6.12 (Wattam et al., 2017). CDS-FWD, coding sequences on the forward strand; CDS-REV, coding sequences on the reverse strand; AMR, antimicrobial resistance genes; VF, virulence factors. The putative origin of replication (*oriC*) was marked with a red box.

Table 1. General genomic features of the ST131 *E. coli* strain GCU-H1

Features	Value
Genome size (bp)	5,095,519
Chromosome I	5,052,029
Plasmid (pGCU-H1-1)	33,826
Plasmid (pGCU-H1-2)	4,946
Plasmid (pGCU-H1-3)	4,718
G + C content (%)	50.8
Coverage (total)	103.73 ×
No. of contigs	4
Chromosomes	1
Plasmids	3
Total genes	5,061
Protein coding genes	4,728
tRNAs	90
rRNAs	22
ncRNAs	5
Pseudogenes	216

revealed that strain GCU-H1 exhibited the highest genomic relatedness to the available *E. coli* genomes in the GenBank database; thus, the isolate was classified as *E. coli*.

Multilocus sequence typing (MLST) was performed to determine the ST of *E. coli* strain GCU-H1, using the Achtman scheme (seven housekeeping genes: *adhA*, *fimC*, *gyrB*, *icd*, *mdh*, *purA*, and *recA*) and the Pasteur scheme (eight housekeeping genes: *dinB*, *icdA*, *pabB*, *polB*, *putP*, *trpA*, *trpB*, and *uidA*) was performed through the PubMLST database (<https://pubmlst.org/organisms/escherichia-spp>), and their respective allele numbers and ST were generated, respectively. The allele profiles of strain GCU-H1 were 53-40-47-13-36-28-29 (Achtman scheme) and 9-1-15-7-4-9-6-9 (Pasteur scheme), and their STs were assigned as ST131 (Achtman scheme) and PST43 (Pasteur scheme).

Potential virulence- and AMR-related genes in strain GCU-H1 were screened as previously described (Lim et al., 2024), and several genes (*fimH*, *sat*, *tsh*, *iucD*, and *usp*), commonly found in ST131 *E. coli* (Blanco et al., 2013), were detected in the genome (Table S1). Therefore, the *fimH* type and serotype of GCU-H1 were predicted in silico using the Center for Genomic Epidemiology (<https://www.genomicepidemiology.org/>), revealing that the isolate possessed *fimH*30 and the O25:H4 serotype, and is classified as Clade C

(Camacho et al., 2009), which could be considered to be a virulent type (Barrios-Villa et al., 2018). Among the virotype-determining genes (Blanco et al., 2013; Nicolas-Chanoine et al., 2014), only *sat* (secreted autotransporter toxin) gene was detected, thus indicating that the strain belonged to virotype C.

Amino acid substitutions in *gyrA* (D87N and S83L) and *parC* (S80I) within the quinolone resistance-determining regions, which confer resistance to quinolones/fluoroquinolones (FQ) (Nicolas-Chanoine et al., 2014), were detected among the screened AMR-related genes in the GCU-H1 genome (Table S2). These results were consistent with the in vitro phenotypic antibiotic susceptibility of the isolate, which was resistant to all tested FQ antimicrobials (ciprofloxacin, levofloxacin, nalidixic acid, norfloxacin, and ofloxacin; Table S3). Generally, ST131 is divided into subclades C1/H30R (associated with FQ-resistant strains) and C2/H30-Rx (associated with the ESBL encoding CTX-M-15) (Rodríguez-Martínez et al., 2021). Therefore, strain GCU-H1 was categorized as subclade C1/H30R of clade C in ST131.

Overall, seven prophage regions were identified in the GCU-H1 chromosome using the PHASTEST (PHAge Search Tool with Enhanced Sequence Translation, <https://phastest.ca>), all of which were intact (score > 90). The estimated sizes of the intact prophage regions were 57.9 (Region 1), 50.4 (Region 2), 51.5 (Region 3), 50.5 (Region 4), 54.2 (Region 5), 49.4 (Region 6), and 56 kb (Region 7; Table S4).

This study presents the genomic characteristics of ST131 *E. coli*, which first emerged in marine mammals in Korea. The findings will provide insights into the genomic features and potential zoonotic transmission risks and contribute to the development of surveillance strategies for monitoring and controlling ST131 *E. coli* in marine animals and environments in Korea.

Nucleotide sequence accession number

Escherichia coli GCU-H1 was deposited in the Korean Collection for Type Cultures (KCTC) under the deposition number KCTC 8663. The complete genome sequence of strain GCU-H1 has been deposited in the GenBank database under the accession numbers CP178558 (chromosome), CP178559 (pGCU-H1-1), CP178560 (pGCU-H1-2), and CP178561 (pGCU-H1-3).

Acknowledgments

This research was supported by the National Research Foundation of Korea (NRF) grant funded by the Korean government (MSIT) (RS-2024-00336046) and Development of Technology for the Biomaterialization of Marine Fishery Byproducts of the Korea Institute of Marine Science and Technology Promotion [KIMST], funded by the Ministry of Oceans and Fisheries [KIMST-20220128]. This research was also supported by the National Institute of Fisheries Science (R2025004) of the Ministry of Oceans and Fisheries.

Conflict of Interest

All authors disclose no conflicts of interest relevant to this research.

References

- Bai S, Zhang P, Qi X, Liu J, Liu D, et al. 2024. Similarities and differences in gastrointestinal microbiomes between wild and human-cared spotted dolphins (*Stenella attenuata*) in natural waters. *Water Biol Secur.* 3: 100288.
- Barrios-Villa E, Cortés-Cortés G, Lozano-Zaraín P, Arenas-Hernández MMP, Martínez de la Peña CF, et al. 2018. Adherent/invasive *Escherichia coli* (AIEC) isolates from asymptomatic people: new *E. coli* ST131 O25:H4/H30-Rx virotypes. *Ann Clin Microbiol Antimicrob.* 17: 42.
- Bevan ER, Jones AM, Hawkey PM. 2017. Global epidemiology of CTX-M β -lactamases: temporal and geographical shifts in genotype. *J Antimicrob Chemother.* 72: 2145–2155.
- Blanco J, Mora A, Mamani R, López C, Blanco M, et al. 2013. Four main virotypes among extended-spectrum- β -lactamase-producing isolates of *Escherichia coli* O25b:H4-B2-ST131: bacterial, epidemiological, and clinical characteristics. *J Clin Microbiol.* 51: 3358–3367.
- Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, et al. 2009. BLAST+: architecture and applications. *BMC Bioinformatics.* 10: 421.
- Dold C. 2014. Chapter 43 - Cetacea (whales, dolphins, porpoises). In Miller RE, Fowler ME. (eds.), *Fowler's zoo and wild animal medicine*, vol. 8th, pp. 422–436. Elsevier.
- Fulham M, Power M, Gray R. 2020. Diversity and distribution of *Escherichia coli* in three species of free-ranging Australian pinniped pups. *Front Mar Sci.* 7: 571171.
- Gross S, Müller A, Seinige D, Wohlsein P, Oliveira M, et al. 2022. Occurrence of antimicrobial-resistant *Escherichia coli* in marine mammals of the North and Baltic Seas: Sentinels for human health. *Antibiotics.* 11: 1248.
- Hassuna NA, Rabea EM, Mahdi WKM, Abdelraheem WM. 2024. Biofilm formation and antimicrobial resistance pattern of uropathogenic *E. coli* ST131 isolated from children with malignant tumors. *J Antibiot.* 77: 324–330.
- Lim S, Park SY, Kim YB, Han SY, Han JE, et al. 2024. Complete genome sequence of the biotype 1 *Vibrio vulnificus* isolated from diseased Pacific shortfin eel (*Anguilla bicolor pacifica*) cultured in Korea. *Korean J Microbiol.* 60: 115–119.
- Manageiro V, Clemente L, Jones-Dias D, Albuquerque T, Ferreira E, et al. 2015. CTX-M-15-producing *Escherichia coli* in dolphin, Portugal. *Emerg Infect Dis.* 21: 2249–2251.
- Melendez D, Roberts MC, Greninger AL, Weissman S, No D, et al. 2019. Whole-genome analysis of extraintestinal pathogenic *Escherichia coli* (ExPEC) MDR ST73 and ST127 isolated from endangered southern resident killer whales (*Orcinus orca*). *J Antimicrob Chemother.* 74: 2176–2180.
- Mora A, García-Peña FJ, Alonso MP, Pedraza-Díaz S, Ortega-Mora LM, et al. 2018. Impact of human-associated *Escherichia coli* clonal groups in Antarctic pinnipeds: presence of ST73, ST95, ST141, and ST131. *Sci Rep.* 8: 4678.
- Nicolas-Chanoine MH, Bertrand X, Madec JY. 2014. *Escherichia coli* ST131, an intriguing clonal group. *Clin Microbiol Rev.* 27: 543–574.
- Park SY, Lee K, Cho Y, Lim SR, Kwon H, et al. 2020. Emergence of third-generation cephalosporin-resistant *Morganella morganii* in a captive breeding dolphin in South Korea. *Animals.* 10: 2052.
- Pitout JDD, DeVinney R. 2017. *Escherichia coli* ST131: a multidrug-resistant clone primed for global domination. *F1000Res.* 6: 195.
- Rodríguez-Martínez JM, López-Cerero L, García-Duque A, Rodríguez-Baño J, Pascual A. 2021. Interplay between IncF plasmids and topoisomerase mutations conferring quinolone resistance in the *Escherichia coli* ST131 clone: stability and resistance evolution. *Eur J Clin Microbiol Infect Dis.* Online ahead of print. doi: 10.1007/s10096-021-04358-4.
- Sellera FP, Cardoso B, Fuentes-Castillo D, Esposito F, Sano E, et al. 2022. Genomic analysis of a highly virulent NDM-1-producing *Escherichia coli* ST162 infecting a pygmy sperm whale (*Kogia breviceps*) in South America. *Front Microbiol.* 13: 915375.
- Vignaroli C, Luna GM, Pasquaroli S, Di Cesare A, Petruzzella R, et al. 2013. Epidemic *Escherichia coli* ST131 and *Enterococcus faecium* ST17 in coastal marine sediments from an Italian beach. *Environ Sci Technol.* 47: 13772–13780.
- Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, et al. 2014. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS One.* 99: e112963.
- Wattam AR, Davis JJ, Assaf R, Boisvert S, Brettin T, et al. 2017. Improvements to PATRIC, the all-bacterial bioinformatics database and analysis resource center. *Nucl Acids Res.* 45: D535–D542.