




Whole-Genome Draft Sequences of Nine Asymptomatic *Escherichia coli* Bacteriuria Isolates from Diabetic Patients

Christoph Stork,^a Beáta Kovács,^{b,c} Eva Trost,^{a*} Tamás Kovács,^d György Schneider,^b Barnabás Rózsai,^e Monika Kerényi,^b Levente Emődy,^b  Ulrich Dobrindt^a

^aInstitute of Hygiene, University of Münster, Münster, Germany

^bDepartment of Medical Microbiology and Immunology, University of Pécs, Pécs, Hungary

^c1st Department of Medicine, University of Pécs, Pécs, Hungary

^dNanophagetherapy Center, Enviroinvest Corp., Pécs, Hungary

^eDepartment of Pediatrics, University of Pécs, Pécs, Hungary

ABSTRACT *Escherichia coli* can colonize the urinary bladder without causing a disease response in the host. This asymptomatic bacteriuria (ABU) can protect against recurrent symptomatic urinary tract infection by virulent bacteria. Here, we report the whole-genome sequences of nine *E. coli* ABU isolates from diabetic patients.

Urinary tract infection (UTI) constitutes a serious global health problem associated with considerable morbidity and death and can affect patient quality of life. Symptomatic UTI is caused most often by *Escherichia coli*. Besides provoking symptoms, *E. coli* can also asymptotically colonize the human bladder for extended time periods. Asymptomatic bacteriuria (ABU) does not, with some exceptions, require any treatment (1). Deliberate colonization of the urinary bladder thus represents a promising alternative to antibiotic treatment and prevents colonization by more virulent bacterial strains (2–4). To characterize bacterial traits that contribute to the fitness and competitiveness of *E. coli* ABU strains in urine, we sequenced nine ABU isolates from patients with diabetes mellitus in Hungary.

Total genomic DNA was isolated from the *E. coli* ABU isolates using the MagAttract high-molecular-weight (HMW) DNA kit (Qiagen, Hilden, Germany). To prepare 500-bp paired-end libraries of all ABU isolates, we used the Nextera XT DNA library preparation kit (Illumina, San Diego, CA, USA). Libraries were sequenced on the Illumina MiSeq sequencing platform using v2 sequencing chemistry. *E. coli* ABU isolate 65 was sequenced twice on the Illumina platform. In addition, this strain was subjected to whole-genome bidirectional sequencing using the Roche GS Junior sequencer with titanium chemistry (Roche Hungary Ltd., Budaörs, Hungary). *E. coli* ABU strain 65 reads from the Illumina and Roche sequencing platforms were included in a hybrid assembly. Prior to assembly, raw reads of all ABU isolates were quality checked with FastQC (v11.5) (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc>), and low-quality reads were trimmed using Sickle (v1.33) (<https://github.com/najoshi/sickle>). Subsequently, the quality-filtered reads were *de novo* assembled using SPAdes (v3.9.0) (5). Final genome assemblies were analyzed using QUAST (v4.3) (6), resulting in 43 to 197 contigs of >1 kb, and accounting for total genome sizes ranging from 4,578,015 to 5,285,335 bp (Table 1). Coding sequences (CDSs) and orthologous genes were determined using Prokka (v1.12) (7) and Proteinortho (v5.15) (8). Between 4,204 and 5,035 coding DNA sequences were identified in the corresponding genomes (Table 1). To predict and identify plasmids, serotypes, and acquired antibiotic resistance mechanisms, we performed a systematic analysis, querying the databases PlasmidFinder (v1.3) (9), SerotypeFinder (v1.1) (10), and ResFinder (v2.1) (11), respectively. For the determination of virulence factors (VFs), we extended the freely available *E. coli* VF collection

Received 3 November 2017 Accepted 20 November 2017 Published 11 January 2018

Citation Stork C, Kovács B, Trost E, Kovács T, Schneider G, Rózsai B, Kerényi M, Emődy L, Dobrindt U. 2018. Whole-genome draft sequences of nine asymptomatic *Escherichia coli* bacteriuria isolates from diabetic patients. *Genome Announc* 6:e01369-17. <https://doi.org/10.1128/genomeA.01369-17>.

Copyright © 2018 Stork 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 Ulrich Dobrindt, dobrindt@uni-muenster.de.

* Present address: Eva Trost, Division of Enteropathogenic Bacteria and *Legionella*, Robert Koch-Institute, Wernigerode Branch, Wernigerode, Germany.

TABLE 1 Features and assembly metrics of the genome sequences of nine *E. coli* ABU isolates from diabetic patients

<i>E. coli</i> strain	ECOR phylogroup (ST) ^a	No. of raw paired reads	No. of trimmed reads	No. of contigs	Genome size (bp)	N ₅₀ (bp)	L ₅₀	Coverage (×)	No. of CDSs ^b	GenBank accession no.
ABU 1	B1 (205)	11,310,408	8,774,314	43	4,578,015	236,117	8	285	4,204	PENN000000000
ABU 9	B2 (95)	10,142,410	7,982,006	106	5,212,357	220,362	7	228	4,862	PEN000000000
ABU 61	F (59)	6,927,866	3,800,294	197	5,183,452	55,491	28	158	4,888	PENP000000000
ABU 65	F (117)	5,847,714, ^c 230,941 ^d	4,315,886, ^c 217,712 ^d	124	5,184,119	150,223	12	179, ^c 16 ^d	5,035	PENQ000000000
ABU 84	B2 (73)	9,577,806	7,435,548	78	5,223,185	301,227	6	212	4,871	PENR000000000
ABU 91	B2 (95)	8,845,416	6,756,004	81	5,285,335	268,769	7	190	5,002	PENJ000000000
ABU 106	B2 (95)	1,220,066	774,552	108	4,939,518	129,842	10	33	4,601	PENK000000000
ABU 123	D (69)	6,759,080	2,783,684	96	4,917,555	115,220	13	116	4,547	PENL000000000
ABU 148	B2 (73)	9,488,588	7,435,092	87	5,154,177	248,638	7	214	4,795	PENM000000000

^aECOR, *E. coli* Reference Collection; ST, sequence type.^bCDS, coding sequence.^cCombined Illumina reads.^d454 reads.

(v0.1) (<https://doi.org/10.5281/zenodo.56686>). Core genome multilocus sequence typing (cgMLST) was performed with SeqSphere+ (v3.4.0) (Ridom GmbH, Münster, Germany).

The draft genomes of nine clinical ABU isolates from diabetic patients will serve as a useful source for upcoming comparative studies of *E. coli* isolate virulence- and fitness-related traits that promote efficient growth and competitiveness in urine. The identification of bacterial characteristics relevant for efficient and safe therapeutic bladder colonization will improve the application of bacterial interference.

Accession number(s). The assembled draft genome sequences obtained from this project were deposited into DDBJ/ENA/GenBank under the accession numbers listed in Table 1.

ACKNOWLEDGMENTS

This work was supported by the Deutsche Forschungsgemeinschaft (grants SFB1009 and TP B05) and ERA-NET PathoGenoMics II (OTKA grant NN78915).

We thank K. Tegelkamp (Münster) for excellent technical assistance.

REFERENCES

- Cai T, Mazzoli S, Mondaini N, Meacci F, Nesi G, D'Elia C, Malossini G, Boddi V, Bartoletti R. 2012. The role of asymptomatic bacteriuria in young women with recurrent urinary tract infections: to treat or not to treat? *Clin Infect Dis* 55:771–777. <https://doi.org/10.1093/cid/cis534>.
- Köves B, Salvador E, Grönberg-Hernández J, Zdziarski J, Wullt B, Svanborg C, Dobrindt U. 2014. Rare emergence of symptoms during long-term asymptomatic *Escherichia coli* 83972 carriage without an altered virulence factor repertoire. *J Urol* 191:519–528. <https://doi.org/10.1016/j.juro.2013.07.060>.
- Sundén F, Håkansson L, Ljunggren E, Wullt B. 2010. *Escherichia coli* 83972 bacteriuria protects against recurrent lower urinary tract infections in patients with incomplete bladder emptying. *J Urol* 184:179–185. <https://doi.org/10.1016/j.juro.2010.03.024>.
- Wullt B, Connell H, Röllano P, Månsson W, Colleen S, Svanborg C. 1998. Urodynamic factors influence the duration of *Escherichia coli* bacteriuria in deliberately colonized cases. *J Urol* 159:2057–2062. [https://doi.org/10.1016/S0022-5347\(01\)63246-4](https://doi.org/10.1016/S0022-5347(01)63246-4).
- Sundén F, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* 19:455–477. <https://doi.org/10.1089/cmb.2012.0021>.
- Gurevich A, Saveliev V, Vyahhi N, Tesler G. 2013. QUASt: 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>.
- Lechner M, Findeiss S, Steiner L, Marz M, Stadler PF, Prohaska SJ. 2011. Proteinortho: detection of (co-)orthologs in large-scale analysis. *BMC Bioinformatics* 12:124. <https://doi.org/10.1186/1471-2105-12-124>.
- Carattoli A, Zankari E, García-Fernández A, Voldby Larsen M, Lund O, Villa L, Möller Aarestrup F, Hasman H. 2014. *In silico* detection and typing of plasmids using PlasmidFinder and plasmid multilocus sequence typing. *Antimicrob Agents Chemother* 58:3895–3903. <https://doi.org/10.1128/AAC.02412-14>.
- Joensen KG, Tetzschner AM, Iguchi A, Aarestrup FM, Scheut 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>.
- 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>.