AMERICAN SOCIETY FOR MICROBIOLOGY

genomeAnnouncements™

Check for updates

Draft Genome Sequences of Semiconstitutive Red, Dry, and Rough Biofilm-Forming Commensal and Uropathogenic *Escherichia coli* Isolates

Annika Cimdins,[a] Petra Lüthje,[a]* Fengyang Li,[a] Irfan Ahmad,[a]* Annelie Brauner,[a,b] Ute Römling[a]

Department of Microbiology, Tumor and Cell Biology, Karolinska Institutet, Stockholm, Sweden[a]; Division of Clinical Microbiology, Karolinska University Hospital, Stockholm, Sweden[b]

ABSTRACT Strains of *Escherichia coli* exhibit diverse biofilm formation capabilities. *E. coli* K-12 expresses the red, dry, and rough (rdar) morphotype below 30°C, whereas clinical isolates frequently display the rdar morphotype semiconstitutively. We sequenced the genomes of eight *E. coli* strains to subsequently investigate the molecular basis of semiconstitutive rdar morphotype expression.

Biofilms are multicellular microbial communities surrounded by a self-produced extracellular matrix adherent to each other, to interfaces, and/or to biotic or abiotic surfaces (1). The red, dry, and rough (rdar) biofilm morphotype defined by the expression of amyloid curli fibers and the exopolysaccharide cellulose is expressed by several species of *Enterobacteriaceae*, among them *Escherichia coli* (2). Formation of the rdar morphotype is activated by the orphan transcriptional regulator CsgD and occurs predominantly at temperatures below 30°C in model strains of *E. coli* and *Salmonella enterica* serovar Typhimurium (3–5). Pathogenic, commensal, and probiotic strains of *E. coli*, however, frequently express the rdar morphotype semiconstitutively (5–9).

In this study, we sequenced the genomes of eight *E. coli* strains with rdar biofilm formation at 28°C and 37°C (Table 1). Recently isolated and minimally passaged strains include three commensal *E. coli* isolates from human feces (6) and four uropathogenic strains obtained from patients' samples (reference 7 and this study). As a historical strain, we included ECOR 31 isolated from leopard feces from the *E. coli* reference strain collection (10).

Genomic DNA from Tob1 and ECOR 31 was sequenced with the PacBio RS II system (Pacific Biosciences; NGI Uppsala, Science For Life Laboratory [SciLifeLab], Uppsala, Sweden). The assembly was done on SMRT portal version 2.3, using HGAP3 with default settings. The other strains were sequenced using an Illumina MiSeq version 3 platform with read length up to 2×300 bp (NGI Stockholm, SciLifeLab, Solna, Sweden). *De novo* assembly was performed using SPAdes (http://bioinf.spbau.ru/spades) (11). For calculations of coverage according to Lander and Waterman (12), see Table 1. Analysis by the Rapid Annotations using Subsystems Technology (RAST; version 2.0) server (http://rast.nmpdr.org/rast.cgi) (13–15) indicated a genome size between 4.97 and 5.65 Mbp and a G+C content between 50.5% and 50.9% (Table 1).

Contigs smaller than 500 bp were omitted, and the sequences were submitted to DDBJ/ENA/GenBank and annotated with the NCBI Prokaryotic Genome Annotation Pipeline. The number of genes is between 4,886 (*E. coli* 80//6) and 5,328 (ECOR 31) (Table 1).

Published genome sequences of *E. coli* isolates lack coupling to a distinct biofilm phenotype, but it is well known that even *E. coli* K-12 derivatives designated the same

TABLE 1 Strain information, sequencing, and genomic features of the eight *E. coli* draft genomes

Strain	Isolation source[a]	No. of sequence entities	Coverage (×)	N_{50} contig size (bp)[b]	Size (Mbp)[c]	G+C content (%)[c]	No. of genes (total/coding)[d]	No. of tRNAs[d]	Nucleotide accession no.
Tob1	Commensal human fecal isolate, Germany	4 polished contigs	68	NA	5.19	50.9	5,147/4,882	90	MIIH00000000
Fec67	Commensal human fecal isolate, Germany	97 scaffolds	173	411,704	5.20	50.6	5,115/4,918	79	MDRZ00000000
Fec101	Commensal human fecal isolate, Germany	100 scaffolds	192	418,989	4.97	50.7	4,942/4,751	79	MDYZ00000000
ECOR31	Commensal fecal isolate from leopard, USA	6 polished contigs	159	NA	5.44	50.6	5,328/5,095	90	MIIL00000000
No. 12	Human pyelonephritis isolate, Slovakia	112 scaffolds	125	349,216	5.10	50.5	5,009/4,807	78	MIIG00000000
B-11870	Human urosepsis isolate, Sweden	116 scaffolds	138	377,062	5.65	50.5	5,188/4,970	77	MIIJ00000000
80//6	Human UTI isolate, Estonia	90 scaffolds	172	367,346	4.98	50.6	4,886/4,697	79	MIII00000000
B-8638	Human urosepsis isolate, Sweden	140 scaffolds	170	190,633	5.21	50.7	5,143/4,917	77	MIIK00000000

[a]UTI, urinary tract infection.
[b]NA, not applicable.
[c]According to RAST server.
[d]According to NCBI Prokaryotic Genome Annotation Pipeline.

name derived from different laboratories can have different biofilm phenotypes (3, 16). The sequences of *E. coli* genomes with distinct biofilm phenotypes provide a profound basis to investigate the molecular mechanisms of semiconstitutive rdar morphotype expression.

Accession number(s). The whole-genome shotgun projects have been deposited at DDBJ/ENA/GenBank under the accession numbers listed in Table 1. The versions described in this paper are the first versions, respectively.

REFERENCES

1. Costerton JW, Lewandowski Z, Caldwell DE, Korber DR, Lappin-Scott HM. 1995. Microbial biofilms. Annu Rev Microbiol 49:711–745. https://doi.org/10.1146/annurev.mi.49.100195.003431.

2. Römling U. 2005. Characterization of the rdar morphotype, a multicellular behaviour in *Enterobacteriaceae*. Cell Mol Life Sci 62:1234–1246. https://doi.org/10.1007/s00018-005-4557-x.

3. Hammar M, Arnqvist A, Bian Z, Olsén A, Normark S. 1995. Expression of two *csg* operons is required for production of fibronectin- and Congo red-binding curli polymers in *Escherichia coli* K-12. Mol Microbiol 18:661–670. https://doi.org/10.1111/j.1365-2958.1995.mmi_18040661.x.

4. Römling U, Rohde M, Olsén A, Normark S, Reinköster J. 2000. AgfD, the checkpoint of multicellular and aggregative behaviour in *Salmonella typhimurium* regulates at least two independent pathways. Mol Microbiol 36:10–23. https://doi.org/10.1046/j.1365-2958.2000.01822.x.

5. Römling U, Sierralta WD, Eriksson K, Normark S. 1998. Multicellular and aggregative behaviour of *Salmonella* Typhimurium strains is controlled by mutations in the *agfD* promoter. Mol Microbiol 28:249–264. https://doi.org/10.1046/j.1365-2958.1998.00791.x.

6. Bokranz W, Wang X, Tschäpe H, Römling U. 2005. Expression of cellulose and curli fimbriae by *Escherichia coli* isolated from the gastrointestinal tract. J Med Microbiol 54:1171–1182. https://doi.org/10.1099/jmm.46064-0.

7. Kai-Larsen Y, Lüthje P, Chromek M, Peters V, Wang X, Holm A, Kádas L, Hedlund KO, Johansson J, Chapman MR, Jacobson SH, Römling U, Agerberth B, Brauner A. 2010. Uropathogenic *Escherichia coli* modulates immune responses and its curli fimbriae interact with the antimicrobial peptide LL-37. PLoS Pathog 6:e1001010. https://doi.org/10.1371/journal.ppat.1001010.

8. Monteiro C, Saxena I, Wang X, Kader A, Bokranz W, Simm R, Nobles D, Chromek M, Brauner A, Brown RM, Jr, Römling U. 2009. Characterization of cellulose production in *Escherichia coli* Nissle 1917 and its biological consequences. Environ Microbiol 11:1105–1116. https://doi.org/10.1111/j.1462-2920.2008.01840.x.

9. Bian Z, Brauner A, Li Y, Normark S. 2000. Expression of and cytokine activation by *Escherichia coli* curli fibers in human sepsis. J Infect Dis 181:602–612. https://doi.org/10.1086/315233.

10. Ochman H, Selander RK. 1984. Standard reference strains of *Escherichia coli* from natural populations. J Bacteriol 157:690–693.

11. Bankevich A, 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.

12. Lander ES, Waterman MS. 1988. Genomic mapping by fingerprinting random clones: a mathematical analysis. Genomics 2:231–239. https://doi.org/10.1016/0888-7543(88)90007-9.

13. Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formsma K, Gerdes S, Glass EM, Kubal M, Meyer F, Olsen GJ, Olson R, Osterman AL, Overbeek RA, McNeil LK, Paarmann D, Paczian T, Parrello B, Pusch GD, Reich C, Stevens R, Vassieva O, Vonstein V, Wilke A, Zagnitko O. 2008. The RAST server: rapid annotations using subsystems technology. BMC Genomics 9:75. https://doi.org/10.1186/1471-2164-9-75.

14. Overbeek R, Olson R, Pusch GD, Olsen GJ, Davis JJ, Disz T, Edwards RA, Gerdes S, Parrello B, Shukla M, Vonstein V, Wattam AR, Xia F, Stevens R. 2014. The SEED and the rapid annotation of microbial genomes using subsystems technology (RAST). Nucleic Acids Res 42:D206–D214. https://doi.org/10.1093/nar/gkt1226.

15. Brettin T, Davis JJ, Disz T, Edwards RA, Gerdes S, Olsen GJ, Olson R, Overbeek R, Parrello B, Pusch GD, Shukla M, Thomason JA III, Stevens R, Vonstein V, Wattam AR, Xia F. 2015. RASTtk: a modular and extensible implementation of the RAST algorithm for building custom annotation pipelines and annotating batches of genomes. Sci Rep 5:8365. https://doi.org/10.1038/srep08365.

16. Serra DO, Richter AM, Hengge R. 2013. Cellulose as an architectural element in spatially structured *Escherichia coli* biofilms. J Bacteriol 195:5540–5554. https://doi.org/10.1128/JB.00946-13.