



Complete Draft Genome Sequence of *Escherichia coli* K802

Gabriele R. M. Kleiner-Grote,^a Daniel Wibberg,^b Anika Winkler,^b Jörn Kalinowski,^b Karl Friehs^{a,b}

Fermentation Engineering, Bielefeld University, Bielefeld, Germany^a; Center for Biotechnology, Bielefeld University, Bielefeld, Germany^b

ABSTRACT *Escherichia coli* K802 is an old strain used for cloning experiments, as well as for the production of recombinant proteins. To understand the genomic background of *E. coli* K802 better, we present here its complete draft genome sequence.

Escherichia coli K802 (CGSC 5610), also designated WA802, is a K-12 derivative. Because of its impaired restriction abilities, it was often used as cloning host, especially for restriction modification (R-M) cloning experiments, as well as for the study of the bacterial R-M system against foreign DNA (1–3). Furthermore, *E. coli* K802 served as an expression strain for secretory protein production of, e.g., antibodies or single-chain fusion proteins (4, 5).

The strain was developed by multiple mutation steps based on *E. coli* K-12. Consecutive X-ray and UV light treatment, as well as phage infection steps, resulted in strain C600. This ancestor requires the supplementation of leucine, threonine, and thiamine; is unable to utilize lactose; and is resistant to coliphages T1 and T5 but sensitive toward phage λ (6–8). Subsequently, *met* and *gal* mutants were introduced by mating with *Hfr Calvalli* and by the use of strain 144 and P1 transduction, respectively. Finally, the nonrestricting mutant K802 was derived by spontaneous mutation (1). The strain was described as lacking *McrA* and *McrB* (9). *E. coli* K802 is available from the *E. coli* Genetic Stock Center (CGSC, New Haven, CT, USA). The currently published genotype by CGSC is as follows: [F⁻, *lacY1* or $\Delta(\text{cod-lacI})6$, *glnX44*(AS), *galk2*(Oc), *galT22*, λ^- , *e14* mutant, *mcrA0*, *rbcC1*, *metB1*, *mcrB1*, *hsdR2*], 2016-04-20.

To establish the draft genome sequence of isolate K802, sequencing of a paired library on an Illumina MiSeq system was performed. The sequencing run (2 × 300 bp) resulted in 1,084,357 reads yielding approximately 242 Mb of sequence information. A *de novo* assembly of the obtained reads by means of the *gsAssembler* software version 2.8 applying default settings generated 169 contigs, accounting for a total length of 4.5 Mb featuring a GC content of 50.76%. Assembled contigs were arranged in 66 scaffolds based on the paired-end sequence information.

Gene prediction and genome annotation were performed as recently described for *E. coli* JF733 (10). Gene prediction resulted in 4,176 predicted genes, 71 tRNAs, and 2 rRNAs, which were annotated using the GenDB platform (11).

The sequence of *E. coli* K802 was analyzed and compared to that of the *E. coli* K-12 strain W3110 (GenBank accession number AP009048). Most of the published mutations could be confirmed, except for *galk2*(Oc) and *rbcC1*. Instead of *rbcC1*, a 1-bp deletion within *rbcD* was detected, resulting in a frameshift. Furthermore, the published deletion $\Delta(\text{cod-lacI})6$ for K802 is different and larger than assumed. In *E. coli* K802, only the *lacY* and *lacA* genes of the *lac* operon, not *lacI* and *lacZ*, are affected by mutations. More precisely, a total region of 38,729 bp within and downstream of *lacY* is deleted in comparison to W3110, including not only the *cyn* and *cod* operon as already indicated but also the *bet* and *prp* operon, as well as several predicted and hypothetical proteins.

Received 26 July 2017 Accepted 31 July 2017 Published 28 September 2017

Citation Kleiner-Grote GRM, Wibberg D, Winkler A, Kalinowski J, Friehs K. 2017. Complete draft genome sequence of *Escherichia coli* K802. *Genome Announc* 5:e00934-17. <https://doi.org/10.1128/genomeA.00934-17>.

Copyright © 2017 Kleiner-Grote 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 Jörn Kalinowski, joern.kalinowski@cebitec.uni-bielefeld.de.

Accession number(s). The *E. coli* K802 draft genome sequence was deposited in the EMBL database under the accession numbers [FMKC01000001](#) to [FMKC01000088](#).

ACKNOWLEDGMENTS

The bioinformatics support of the BMBF-funded project Bielefeld-Gießen Center for Microbial Bioinformatics–BiGi (grant number 031A533) within the German Network for Bioinformatics Infrastructure (de.NBI) is gratefully acknowledged.

We acknowledge support for the article processing charge from the Deutsche Forschungsgemeinschaft and the Open Access Publication Fund of Bielefeld University. Further, we thank John E. Wertz (*E. coli* Genetic Stock Center, Yale University, USA) for providing additional information about the history and development of *E. coli* K802.

REFERENCES

1. Wood WB. 1966. Host specificity of DNA produced by *Escherichia coli*: bacterial mutations affecting the restriction and modification of DNA. *J Mol Biol* 16:118–133. [https://doi.org/10.1016/S0022-2836\(66\)80267-X](https://doi.org/10.1016/S0022-2836(66)80267-X).
2. Zakharova MV, Beletskaya IV, Kravetz AN, Pertzov AV, Mayorov SG, Shlyapnikov MG, Solonin AS. 1998. Cloning and sequence analysis of the plasmid-borne genes encoding the *Eco29kI* restriction and modification enzymes. *Gene* 208:177–182. [https://doi.org/10.1016/S0378-1119\(97\)00637-9](https://doi.org/10.1016/S0378-1119(97)00637-9).
3. Pingoud A, Wilson GG, Wende W. 2014. Type II restriction endonucleases—a historical perspective and more. *Nucleic Acids Res* 42:7489–7527. <https://doi.org/10.1093/nar/gku447>.
4. Wich G, Dassler T. September 2007. Process for the fermentative production of antibodies. US Patent 2008/0206818 A1.
5. Yang J, Moyana T, Mackenzie S, Xia Q, Xiang J. 1998. One hundred seventy-fold increase in excretion of an FV fragment-tumor necrosis factor alpha fusion protein (sFV/TNF- α) from *Escherichia coli* caused by the synergistic effects of glycine and triton X-100. *Appl Environ Microbiol* 64:2869–2874.
6. Tatum EL, Lederberg J. 1947. Gene recombination in the bacterium *Escherichia coli*. *J Bacteriol* 53:673–684.
7. Lederberg EM. 1952. Allelic relationships and reverse mutation in *Escherichia coli*. *Genetics* 37:469–483.
8. Appleyard RK. 1954. Segregation of new lysogenic types during growth of a doubly lysogenic strain derived from *Escherichia coli* K12. *Genetics* 39:440–452.
9. Raleigh EA, Murray NE, Revel H, Blumenthal RM, Westaway D, Reith AD, Rigby PWJ, Elhai J, Hanahan D. 1988. McrA and McrB restriction phenotypes of some *E. coli* strains and implications for gene cloning. *Nucleic Acids Res* 16:1563–1575. <https://doi.org/10.1093/nar/16.4.1563>.
10. Kleiner GRM, Wibberg D, Winkler A, Kalinowski J, Wertz JE, Friehs K. 2016. Complete draft genome sequence of *Escherichia coli* JF733. *Genome Announc* 4(2):e00298-16. <https://doi.org/10.1128/genomeA.00298-16>.
11. Meyer F, Goesmann A, McHardy AC, Bartels D, Bekel T, Clausen J, Kalinowski J, Linke B, Rupp O, Giegerich R, Pühler A. 2003. GenDB—an open source genome annotation system for prokaryote genomes. *Nucleic Acids Res* 31:2187–2195. <https://doi.org/10.1093/nar/gkg312>.