Escherichia coli str. K-12 substr. MG1655, complete genome

NCBI Reference Sequence: NC_000913.3

FASTA Graphics

LOCUS NC_000913 1482 bp DNA linear CON 09-DEC-2025

DEFINITION Escherichia coli str. K-12 substr. MG1655, complete genome.

ACCESSION NC_000913 REGION: 3178115..3179596

VERSION NC_000913.3

DBLINK BioProject: PRJNA57779

BioSample: SAMN02604091

KEYWORDS RefSeq.

SOURCE Escherichia coli str. K-12 substr. MG1655

ORGANISM Escherichia coli str. K-12 substr. MG1655

Bacteria; Pseudomonadati; Pseudomonadota; Gammaproteobacteria;

Enterobacterales; Enterobacteriaceae; Escherichia.

REFERENCE 1 (bases 1 to 1482)

AUTHORS Riley,M., Abe,T., Arnaud,M.B., Berlyn,M.K., Blattner,F.R.,

Chaudhuri,R.R., Glasner,J.D., Horiuchi,T., Keseler,I.M.,
Kosuge,T.,

Mori,H., Perna,N.T., Plunkett,G. III, Rudd,K.E., Serres,M.H.,

Thomas,G.H., Thomson,N.R., Wishart,D. and Wanner,B.L.

TITLE Escherichia coli K-12: a cooperatively developed annotation

snapshot--2005

JOURNAL Nucleic Acids Res. 34 (1), 1-9 (2006)

PUBMED 16397293

REMARK Publication Status: Online-Only

REFERENCE 2 (bases 1 to 1482)

AUTHORS Hayashi,K., Morooka,N., Yamamoto,Y., Fujita,K., Isono,K.,

Choi,S.,
Ohtsubo,E., Baba,T., Wanner,B.L., Mori,H. and Horiuchi,T.
TITLE Highly accurate genome sequences of Escherichia coli K-12 strains MG1655 and W3110
JOURNAL Mol. Syst. Biol. 2, 2006 (2006)
PUBMED 16738553
REFERENCE 3 (bases 1 to 1482)
AUTHORS Blattner,F.R., Plunkett,G. III, Bloch,C.A., Perna,N.T., Burland,V.,
Riley,M., Collado-Vides,J., Glasner,J.D., Rode,C.K., Mayhew,G.F., Gregor,J., Davis,N.W., Kirkpatrick,H.A., Goeden,M.A., Rose,D.J., Mau,B. and Shao,Y.
TITLE The complete genome sequence of Escherichia coli K-12
JOURNAL Science 277 (5331), 1453-1462 (1997)
PUBMED 9278503
REFERENCE 4 (bases 1 to 1482)
AUTHORS Arnaud,M., Berlyn,M.K.B., Blattner,F.R., Galperin,M.Y., Glasner,J.D., Horiuchi,T., Kosuge,T., Mori,H., Perna,N.T., Plunkett,G. III, Riley,M., Rudd,K.E., Serres,M.H., Thomas,G.H. and Wanner,B.L.
TITLE Workshop on Annotation of Escherichia coli K-12
JOURNAL Unpublished
REMARK Woods Hole, Mass., on 14-18 November 2003 (sequence corrections)
REFERENCE 5 (bases 1 to 1482)
AUTHORS Glasner,J.D., Perna,N.T., Plunkett,G. III, Anderson,B.D., Bockhorst,J., Hu,J.C., Riley,M., Rudd,K.E. and Serres,M.H.
TITLE ASAP: Escherichia coli K-12 strain MG1655 version m56

JOURNAL Unpublished

REMARK ASAP download 10 June 2004 (annotation updates)

REFERENCE 6 (bases 1 to 1482)

AUTHORS Hayashi,K., Morooka,N., Mori,H. and Horiuchi,T.

TITLE A more accurate sequence comparison between genomes of Escherichia coli K12 W3110 and MG1655 strains

JOURNAL Unpublished

REMARK GenBank accessions AG613214 to AG613378 (sequence corrections)

REFERENCE 7 (bases 1 to 1482)

AUTHORS Perna,N.T.

TITLE Escherichia coli K-12 MG1655 yqiK-rfaE intergenic region, genomic sequence correction

JOURNAL Unpublished

REMARK GenBank accession AY605712 (sequence corrections)

REFERENCE 8 (bases 1 to 1482)

AUTHORS Rudd,K.E.

TITLE A manual approach to accurate translation start site annotation: an E. coli K-12 case study

JOURNAL Unpublished

REFERENCE 9 (bases 1 to 1482)

CONSRTM NCBI Genome Project

TITLE Direct Submission

JOURNAL Submitted (11-JUL-2025) National Center for Biotechnology Information, NIH, Bethesda, MD 20894, USA

REFERENCE 10 (bases 1 to 1482)

AUTHORS Auchincloss,A.H., Kothari,A., Karp,P.D., Moore,L.R. and

Plunkett,G.

III.

TITLE Direct Submission

JOURNAL Submitted (06-NOV-2024) EcoCyc Project, SRI International, Menlo Park, CA 94025, USA

REMARK Protein update by submitter

REFERENCE 11 (bases 1 to 1482)

AUTHORS Blattner,F.R. and Plunkett,G. III.

TITLE Direct Submission

JOURNAL Submitted (30-JUL-2014) Laboratory of Genetics, University of Wisconsin, 425G Henry Mall, Madison, WI 53706-1580, USA

REMARK Protein update by submitter

REFERENCE 12 (bases 1 to 1482)

AUTHORS Blattner,F.R. and Plunkett,G. III.

TITLE Direct Submission

JOURNAL Submitted (15-NOV-2013) Laboratory of Genetics, University of Wisconsin, 425G Henry Mall, Madison, WI 53706-1580, USA

REMARK Protein update by submitter

REFERENCE 13 (bases 1 to 1482)

AUTHORS Blattner,F.R. and Plunkett,G. III.

TITLE Direct Submission

JOURNAL Submitted (26-SEP-2013) Laboratory of Genetics, University of Wisconsin, 425G Henry Mall, Madison, WI 53706-1580, USA

REMARK Sequence update by submitter

REFERENCE 14 (bases 1 to 1482)

AUTHORS Rudd,K.E.

TITLE Direct Submission

JOURNAL Submitted (06-FEB-2013) Department of Biochemistry and Molecular Biology, University of Miami Miller School of Medicine, 118 Gautier Bldg., Miami, FL 33136, USA

REMARK Sequence update by submitter

REFERENCE 15 (bases 1 to 1482)

AUTHORS Rudd,K.E.

TITLE Direct Submission

JOURNAL Submitted (24-APR-2007) Department of Biochemistry and Molecular Biology, University of Miami Miller School of Medicine, 118 Gautier Bldg., Miami, FL 33136, USA

REMARK Annotation update from ecogene.org as a multi-database collaboration

REFERENCE 16 (bases 1 to 1482)

AUTHORS Plunkett,G. III.

TITLE Direct Submission

JOURNAL Submitted (07-FEB-2006) Laboratory of Genetics, University of Wisconsin, 425G Henry Mall, Madison, WI 53706-1580, USA

REMARK Protein updates by submitter

REFERENCE 17 (bases 1 to 1482)

AUTHORS Plunkett,G. III.

TITLE Direct Submission

JOURNAL Submitted (10-JUN-2004) Laboratory of Genetics, University of Wisconsin, 425G Henry Mall, Madison, WI 53706-1580, USA

REMARK Sequence update by submitter

REFERENCE 18 (bases 1 to 1482)

AUTHORS Plunkett,G. III.

TITLE Direct Submission

JOURNAL Submitted (13-OCT-1998) Laboratory of Genetics, University of Wisconsin, 425G Henry Mall, Madison, WI 53706-1580, USA

REFERENCE 19 (bases 1 to 1482)

AUTHORS Blattner,F.R. and Plunkett,G. III.

TITLE Direct Submission

JOURNAL Submitted (02-SEP-1997) Laboratory of Genetics, University of Wisconsin, 425G Henry Mall, Madison, WI 53706-1580, USA

REFERENCE 20 (bases 1 to 1482)

AUTHORS Blattner,F.R. and Plunkett,G. III.

TITLE Direct Submission

JOURNAL Submitted (16-JAN-1997) Laboratory of Genetics, University of Wisconsin, 425G Henry Mall, Madison, WI 53706-1580, USA

COMMENT REVIEWED REFSEQ: This record has been curated by NCBI staff. The reference sequence is identical to U00096.

On Nov 3, 2013 this sequence version replaced NC_000913.2.
Protein update by submitter; annotation updates are derived from EcoCyc https://ecocyc.org/. Suggestions for updates can be sent to biocyc-support@ai.sri.com. This update was generated from a collaboration that includes EcoCyc, the University of Wisconsin, UniProtKB/Swiss-Prot, and the National Center for Biotechnology Information (NCBI).
COMPLETENESS: full length.

FEATURES Location/Qualifiers

source 1..1482

/organism="Escherichia coli str. K-12 substr. MG1655"

/mol_type="genomic DNA"

/strain="K-12"

/sub_strain="MG1655"

/db_xref="taxon:511145"

gene 1..1482

/gene="tolC"

/locus_tag="b3035"

/gene_synonym="colE1-i; ECK3026; mtcB; mukA; refI; toc;

weeA"

/db_xref="ASAP:ABE-0009965"

/db_xref="ECOCYC:EG11009"

/db_xref="GeneID:947521"

CDS 1..1482

/gene="tolC"

/locus_tag="b3035"

/gene_synonym="colE1-i; ECK3026; mtcB; mukA; refI; toc;

weeA"

/codon_start=1

/transl_table=11

/product="outer membrane channel TolC"

/protein_id="NP_417507.2"

/db_xref="UniProtKB/Swiss-Prot:P02930"

/db_xref="ASAP:ABE-0009965"

/db_xref="ECOCYC:EG11009"

/db_xref="GeneID:947521"

/translation="MKKLLPILIGLSLSGFSSLSQAENLMQVYQQARLSNPELRKSAA

DRDAAFEKINEARSPLLPQLGLGADYTYSNGYRDANGINSNATSASLQLTQSIFDMSK

WRALTLQEKAAGIQDVTYQTDQQTLILNTATAYFNVLNAIDVLSYTQAQKEAIYRQLD

QTTQRFNVGLVAITDVQNARAQYDTVLANEVTARNNLDNAVEQLRQITGNYYPELAAL

NVENFKTDKPQPVNALLKEAEKRNLSLLQARLSQDLAREQIRQAQDGHLPTLDLTAST

GISDTSYSGSKTRGAAGTQYDDSNMGQNKVGLSFSLPIYQGGMVNSQVKQAQYNFVGA

SEQLESAHRSVVQTVRSSFNNINASISSINAYKQAVVSAQSSLDAMEAGYSVGTRTIV

DVLDATTTLYNAKQELANARYNYLINQLNIKSALGTLNEQDLLALNNALSKPVSTNPE

NVAPQTPEQNAIADGYAPDSPAPVVQQTSARTTTSNGHNPFRN"

ORIGIN

```
        1 atgaagaaat tgctccccat tcttatcggc ctgagccttt ctgggttcag ttcgttgagc

       61 caggccgaga acctgatgca agtttatcag caagcacgcc ttagtaaccc ggaattgcgt

      121 aagtctgccg ccgatcgtga tgctgccttt gaaaaaatta atgaagcgcg cagtccatta

      181 ctgccacagc taggtttagg tgcagattac acctatagca acggctaccg cgacgcgaac

      241 ggcatcaact ctaacgcgac cagtgcgtcc ttgcagttaa ctcaatccat ttttgatatg

      301 tcgaaatggc gtgcgttaac gctgcaggaa aaagcagcag ggattcagga cgtcacgtat

      361 cagaccgatc agcaaacctt gatcctcaac accgcgaccg cttatttcaa cgtgttgaat

      421 gctattgacg ttctttccta tacacaggca caaaaagaag cgatctaccg tcaattagat

      481 caaaccaccc aacgttttaa cgtgggcctg gtagcgatca ccgacgtgca gaacgcccgc

      541 gcacagtacg ataccgtgct ggcgaacgaa gtgaccgcac gtaataacct tgataacgcg

      601 gtagagcagc tgcgccagat caccggtaac tactatccgg aactggctgc gctgaatgtc

      661 gaaaacttta aaaccgacaa accacagccg gttaacgcgc tgctgaaaga agccgaaaaa

      721 cgcaacctgt cgctgttaca ggcacgcttg agccaggacc tggcgcgcga gcaaattcgc

      781 caggcgcagg atggtcactt accgactctg gatttaacgg cttctaccgg gatttctgac

      841 acctcttata gcggttcgaa aacccgtggt gccgctggta cccagtatga cgatagcaat

      901 atgggccaga acaaagttgg cctgagcttc tcgctgccga tttatcaggg cggaatggtt

      961 aactcgcagg tgaaacaggc acagtacaac tttgtcggtg ccagcgagca actggaaagt

     1021 gcccatcgta gcgtcgtgca gaccgtgcgt tcctccttca acaacattaa tgcatctatc
```

1081 agtagcatta acgcctacaa acaagccgta gtttccgctc aaagctcatt agacgcgatg

1141 gaagcgggct actcggtcgg tacgcgtacc attgttgatg tgttggatgc gaccaccacg

1201 ttgtacaacg ccaagcaaga gctggcgaat gcgcgttata actacctgat taatcagctg

1261 aatattaagt cagctctggg tacgttgaac gagcaggatc tgctggcact gaacaatgcg

1321 ctgagcaaac cggtttccac taatccggaa aacgttgcac cgcaaacgcc ggaacagaat

1381 gctattgctg atggttatgc gcctgatagc ccggcaccag tcgttcagca aacatccgca

1441 cgcactacca ccagtaacgg tcataaccct ttccgtaact ga