

Salmonella sp. WGH-01 chromosome, complete genome

GenBank: CP096171.1

[FASTA Graphics](#)

[Go to:](#)

LOCUS CP096171 1001 bp DNA linear BCT 16-MAY-2022

DEFINITION Salmonella sp. WGH-01 chromosome, complete genome.

ACCESSION [CP096171](#) REGION: 3643480..3644480

VERSION CP096171.1

DBLINK BioProject: [PRJNA827233](#)

BioSample: [SAMN27606337](#)

KEYWORDS .

SOURCE Salmonella sp. WGH-01

ORGANISM [Salmonella sp. WGH-01](#)

Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacterales;
Enterobacteriaceae; Salmonella.

REFERENCE 1 (bases 1 to 1001)

AUTHORS Zhao,J., Lu,X., Tie,A., Ngegba,E., Wang,L., Sun,L., Liang,Y.,
Abdulai,M.K., Bah,S., Wang,G., Dong,X., Harding,D. and Kan,B.

TITLE Molecular diagnostics and next-generation sequencing reveal real
etioloical characteristics of invasive Salmonella infection in
febrile illness in Freetown, Sierra Leone

JOURNAL Emerg Microbes Infect, 1-20 (2022) In press

PUBMED [35537043](#)

REMARK Publication Status: Available-Online prior to print

REFERENCE 2 (bases 1 to 1001)

AUTHORS Liu,J.

TITLE Direct Submission

JOURNAL Submitted (19-APR-2022) Chinese Center for Disease Control and
Prevention, Chinese Center for Disease Control and Prevention, 155
Changbai Road, Changping District, Beijing 102200, China

COMMENT The annotation was added by the NCBI Prokaryotic Genome
Annotation

Pipeline (PGAP). Information about PGAP can be found here:

https://www.ncbi.nlm.nih.gov/genome/annotation_prok/

NCBI analysis of this genome has determined that >30% of the
predicted CDS features contain frameshifts, which is well above

the

usually observed value of <10%.

##Genome-Assembly-Data-START##

Assembly Method :: SPAdes v. V3

Genome Representation :: Full

Expected Final Version :: Yes

Genome Coverage :: 100.0x

Sequencing Technology :: Oxford Nanopore MinION

##Genome-Assembly-Data-END##

##Genome-Annotation-Data-START##

Annotation Provider :: NCBI
Annotation Date :: 04/21/2022 12:52:20
Annotation Pipeline :: NCBI Prokaryotic Genome
Annotation Pipeline (PGAP)
Annotation Method :: Best-placed reference protein
set; GeneMarkS-2+
Annotation Software revision :: 6.1
Features Annotated :: Gene; CDS; rRNA; tRNA; ncRNA;
repeat_region
Genes (total) :: 5,597
CDSs (total) :: 5,483
Genes (coding) :: 1,476
CDSs (with protein) :: 1,476
Genes (RNA) :: 114
rRNAs :: 8, 7, 7 (5S, 16S, 23S)
complete rRNAs :: 8, 7, 7 (5S, 16S, 23S)
tRNAs :: 79
ncRNAs :: 13
Pseudo Genes (total) :: 4,007
CDSs (without protein) :: 4,007
Pseudo Genes (ambiguous residues) :: 0 of 4,007
Pseudo Genes (frameshifted) :: 3,826 of 4,007
Pseudo Genes (incomplete) :: 402 of 4,007
Pseudo Genes (internal stop) :: 493 of 4,007
Pseudo Genes (multiple problems) :: 651 of 4,007
##Genome-Annotation-Data-END##

FEATURES Location/Qualifiers
 source 1..1001
 /organism="Salmonella sp. WGH-01"
 /mol_type="genomic DNA"
 /strain="WGH-01"
 /isolation_source="blood samples from outpatients"
 /host="Homo sapiens"
 /db_xref="taxon:[2935857](#)"
 /country="Sierra Leone"
 /collection_date="2019"
 gene <1..>1001
 /gene="invA"
 /locus_tag="MYA98_21175"
 /pseudo

CDS <1..>1001
 /gene="invA"
 /locus_tag="MYA98_21175"
 /inference="COORDINATES: similar to AA
 sequence:RefSeq:NP_461817.1"
 /note="frameshifted; Derived by automated computational
 analysis using gene prediction method: Protein Homology."
 /pseudo
 /codon_start=3
 /transl_table=[11](#)
 /product="type III secretion system export apparatus
 protein InvA"

ORIGIN

```

1  caacgtttcc tgcggtactg ttaattacca cgctctttcg tctggcatta tcgatcagta
61  ccagccgtct tatcttgatt gaagccgatg ccggtgaaat tatcgccacg ttcgggcaat
121  tcgttattgg cgatagcccg gcggtgggtt ttgttgtctt ctctattgtc accgtgggtc
181  agtttatcgt tattaccaa gggtcagaac gcgtcgcgga agtcgcggcc cgattttctc
241  tggatgggat gcccggtaaa cagatgagta ttgatgccga tttgaaggcc ggtattattg
301  atgcggatgc tgcgcgcgaa cggcgaagcg tactggaaag ggaaagccag ctttacggtt
361  cctttgacgg tgcgatgaag tttatcaaag gtgacgctat tgccggcatc attatcatct
421  ttgtgaactt tattggcggg atttcgggtg ggatgaccgg ccatgggatg gatttgtcct
481  ccgctctgtc tacttatacc atgctgacca ttggtgatgg tcttgtcgcc cagatccccg
541  cattgttgat tgcgattagt gccggtttta tcgtgactcg cgtaaatggc gatagcgata
601  atatggggcg gaatatcatg acgcagctgt tgaacaaccc atttgtattg gttgttacgg
661  ctattttgac catttcaatg ggaactctgc cgggattccc gctgccggta tttgttattt
721  tatcggtggt tttaagcgta ctcttctatt ttaaattccg tgaagcaaaa cgtagcgccg
781  ccaaacctaa aaccagcaaa ggcgagcagc cgctcagtat tgaggaaaag aagggtcgtc
841  gttaggactg attggcgatc tcgataaagt ctctacagag actgtaccgt tgatattact
901  tgtgccgaag agccggcgtg aagatctgga aaaagctcaa cttgcggagc gcctacgtag
961  tcagttcttt attgattatg gcgtgcmcct gccggaagta t

```