



**University of
Zurich**^{UZH}

**Zurich Open Repository and
Archive**

University of Zurich
Main Library
Strickhofstrasse 39
CH-8057 Zurich
www.zora.uzh.ch

Year: 2018

Complete and assembled genome sequence of *Salmonella enterica* subsp. *enterica* serotype Senftenberg N17-509, a strain lacking *Salmonella* pathogen island SPI-1

Stevens, Marc J A ; Zurfluh, Katrin ; Althaus, D ; Corti, S ; Lehner, Angelika ; Stephan, Roger

Abstract: The genome of *Salmonella enterica* subsp. *enterica* serotype Senftenberg N17-509, a strain isolated from desiccated coconut, was sequenced using single-molecule real-time sequencing. It consists of a 5.1-Mbp chromosome and a 29-kb linear plasmid.

DOI: <https://doi.org/10.1128/genomeA.00156-18>

Posted at the Zurich Open Repository and Archive, University of Zurich

ZORA URL: <https://doi.org/10.5167/uzh-167635>

Journal Article

Accepted Version

Originally published at:

Stevens, Marc J A; Zurfluh, Katrin; Althaus, D; Corti, S; Lehner, Angelika; Stephan, Roger (2018). Complete and assembled genome sequence of *Salmonella enterica* subsp. *enterica* serotype Senftenberg N17-509, a strain lacking *Salmonella* pathogen island SPI-1. *Genome Announcements*, 6(12):e00156-18.

DOI: <https://doi.org/10.1128/genomeA.00156-18>

1 **Complete and assembled genome sequence of *Salmonella enterica* subsp. *enterica***
2 **serotype Senftenberg N17-509, a strain lacking *Salmonella* pathogen island SPI-1**

3

4 Marc J. A. Stevens, Katrin Zurfluh, Denise Althaus, Sabrina Corti, Angelika Lehner,
5 Roger Stephan*

6

7 Institute for Food Safety and Hygiene, Vetsuisse Faculty, University of Zurich,
8 Switzerland.

9

10 *Corresponding author: roger.stephan@uhz.ch

11

12

13 **Abstract.**

14 The genome of *Salmonella enterica* subsp. *enterica* serotype Senftenberg N17-509, a
15 strain isolated from desiccated coconut, was sequenced using single molecule real time
16 sequencing. It consists of a 5.1-Mbp chromosome, and a 29-kb linear plasmid.

17

18 The *Salmonella enterica* subsp. *enterica* strain N17-509 was isolated from desiccated
19 coconut and provided different identification and typing challenges. The biochemical
20 tests performed to confirm the *Salmonella* genus were inconclusive. A commercial real-
21 time PCR identification system (Assurance GDS[®] *Salmonella* kit (BioControl Systems,
22 Bellevue, WA, USA)) based on the *invA* gene yielded a negative result. A further rRNA
23 based fast method (VIT *Salmonella* kit (vermicon, München, Germany) as well as the

24 MALDI TOF based identification was positive for *Salmonella*. According to the
25 KAUFFMANN-WHITE Scheme (Institut Pasteur, Paris), the strain was serotyped as
26 *Salmonella enterica* subsp. *enterica* serovar 1,3,19:nt.

27 Genomic DNA was extracted using Wizard[®] Genomic DNA Purification Kit according to
28 the manufacturer's protocol (Promega AG, Dübendorf, Switzerland). The genome was
29 sequenced using a two SMRT cells on a PacBio RS II (Pacific Biosciences, Menlo Park,
30 CA, USA) and was performed at the Functional Genomics Center Zurich (Zurich,
31 Switzerland). The raw reads were filtered using the RS_Filter_Only protocol in the
32 SMRT-portal (Pacific Biosciences) with 0.75 as polymerase read quality cut off and a
33 minimal length of 500 bp. A total of 91,709 reads with an average length of 11,001 bp
34 were selected, corresponding to 1,008,898,178 sequenced base pairs and a 197-fold
35 coverage. The reads were assembled using the Canu assembler 1.6 (1) with the option
36 "pacbio-raw" and an estimated genome size of 5.1 Mbp. The Canu output consisted of 3
37 contigs, which were further polished in CLC workbench 7 (CLC, Aarhus, Denmark).
38 Two contigs were combined and formed the chromosomal sequence. The third contig
39 contained a plasmid-like sequence. The start of the chromosome was set at the origin of
40 replication, determined using OriCfinder (2).

41 A blastN with the complete chromosome against all *Enterobacteriales* genomes in the
42 NCBI nucleotide database revealed highest identity of 99% and a coverage of 91% with
43 the genome of *Salmonella enterica* subsp. *enterica* serotype Senftenberg 775W (=ATCC
44 43845). Hence, strain N17-509 was classified as *Salmonella enterica* subsp. *enterica*
45 serotype Senftenberg. The genome of N17-509 consists of a 5,121,989-bp chromosome
46 and a 29,600-bp extra-chromosomal element designated pN17-509. The GC-contents of

47 the chromosome and pN17-509 are 52.2% and 36,5%, respectively. The genome was
48 annotated by the NCBI Prokaryotic Genomes Automatic Annotation Pipeline and
49 contains 5361 genes of which 5241 encoding proteins. Further, 84 tRNA genes and 7
50 rRNA operons are present. Remarkably, the *Salmonella* pathogen island SPI-1,
51 harbouring e.g. the *invA* gene, is absent in this strain. The plasmid pN17-509 has 97%
52 identity at 87% coverage with the linear plasmid pBSSB1 from *Salmonella enterica*
53 subsp. *enterica* serovar Typhi that mediates flagellar variation (3).

54

55 **Nucleotide sequence accession number.** Sequence and annotation data of the complete
56 genome of *Salmonella enterica* subsp. *enterica* serotype Senftenberg N17-509 are
57 deposited in the GenBank database with the accession number CP026379 for the
58 chromosome and CP026380 for the plasmid.

59

60 REFERENCES

61

62 **1) Koren S, Walenz BP, Berlin K, Miller JR, Bergman NH, Phillippy AM. 2017.**

63 Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and
64 repeat separation. *Genome Res* **27**:722-736.

65

66 **2) Gao F, Zhang CT. 2008. Ori-Finder: a web-based system for finding oriCs in**
67 **unannotated bacterial genomes. *BMC Bioinformatics* **9**:79.**

68

69 3) **Baker S, Hardy J, Sanderson KE, Quail M, Goodhead I, Kingsley RA,**
70 **Parkhill J, Stocker B, Dougan G.** 2007. A novel linear plasmid mediates
71 flagellar variation in *Salmonella* Typhi. PLoS Pathog **11**;3(5):e59.