

**Whole-Genome Sequences of Streptococcus alactolyticus Strain An1F4 and Escherichia coli Strains Ae3A3 and Ae3B3, Isolated from Feces of Domestic Pigs**

Sigurdarson, Jens Jakob; Kawaichi, Satoshi; Svane, Simon; Karring, Henrik

*Published in:*  
Microbiology Resource Announcements

*DOI:*  
10.1128/mra.00903-22

*Publication date:*  
2023

*Document version:*  
Final published version

*Document license:*  
CC BY

*Citation for published version (APA):*  
Sigurdarson, J. J., Kawaichi, S., Svane, S., & Karring, H. (2023). Whole-Genome Sequences of Streptococcus alactolyticus Strain An1F4 and Escherichia coli Strains Ae3A3 and Ae3B3, Isolated from Feces of Domestic Pigs. *Microbiology Resource Announcements*, 12(2). <https://doi.org/10.1128/mra.00903-22>

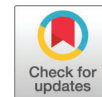
Go to publication entry in University of Southern Denmark's Research Portal

**Terms of use**

This work is brought to you by the University of Southern Denmark.  
Unless otherwise specified it has been shared according to the terms for self-archiving.  
If no other license is stated, these terms apply:

- You may download this work for personal use only.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying this open access version

If you believe that this document breaches copyright please contact us providing details and we will investigate your claim.  
Please direct all enquiries to [puresupport@bib.sdu.dk](mailto:puresupport@bib.sdu.dk)



# Whole-Genome Sequences of *Streptococcus alactolyticus* Strain An1F4 and *Escherichia coli* Strains Ae3A3 and Ae3B3, Isolated from Feces of Domestic Pigs

Jens Jakob Sigurdarson,<sup>a\*</sup> Satoshi Kawaichi,<sup>a</sup> Simon Svane,<sup>a</sup>  Henrik Karring<sup>a</sup>

<sup>a</sup>Department of Green Technology, University of Southern Denmark, Odense, Denmark

Jens Jakob Sigurdarson and Satoshi Kawaichi contributed equally to this work. Author order was determined on the basis of seniority and contribution to the work.

**ABSTRACT** Here, we present the genome sequences of a strain of *Streptococcus alactolyticus* and two strains of *Escherichia coli* that were isolated from feces samples from domestic pigs in Denmark. The genome sequences contribute to a better understanding of the microbiological processes in the feces and manure of domestic pigs.

The production of domestic pigs (*Sus domesticus*) is associated with different environmental and societal challenges (1). The microbiological communities in pig feces and manure contribute to the production of different biogenic pollutants, including ammonia, methane, sulfur compounds, and volatile organic compounds (2). Hence, sequencing the genomes of bacteria isolated from pig feces will provide the basis for a better understanding of the microbiological and biochemical processes in pig feces and manure.

*Streptococcus alactolyticus* strain An1F4 and *Escherichia coli* strains Ae3A3 and Ae3B3 were isolated from feces samples from domestic pigs. Fresh pig feces samples were collected at Grønhøj Research Station (Karup, Denmark) without touching the pigs, and the animals were never stimulated or forced to excrete feces. Because the animals experienced no pain, suffering, anxiety, or lasting harm, approval from the Danish Inspectorate for Animal Experiments was not necessary, according to the relevant Danish legislation (Bekendtgørelse af lov om dyreforsøg). The feces samples were suspended in phosphate-buffered saline and used as an inoculum on a modified yeast casitone fatty acid (YCFA) agar medium without cysteine, sodium bicarbonate, and resazurin (3) and supplemented with 2 g/L glucose, 2 g/L maltose, and 2 g/L cellobiose as carbon and energy sources (4). Agar plates incubated under anoxic and oxic conditions at 37°C were used to isolate *S. alactolyticus* strain An1F4 and *E. coli* strains Ae3A3 and Ae3B3, respectively. The strains were purified by streaking single colonies onto new agar plates at least three times before the isolates were identified to the species level (identification scores of >2.3) using a matrix-assisted laser desorption ionization–time of flight mass spectrometry (MALDI-TOF MS) Biotyper system (Bruker Daltonics, USA). Single colonies of the purified *S. alactolyticus* and *E. coli* strains were grown in modified YCFA liquid medium at 37°C under anoxic and oxic conditions, respectively.

Genomic DNA of each strain was extracted from cultured cells using the standard protocol for the Genra Puregene kit (Qiagen, Germany) before sequencing and processing by DNASense (Aalborg, Denmark). Paired-end sequence libraries (2 × 301 bp) were prepared using a NEBNext Ultra II DNA library preparation kit for Illumina (New England Biolabs, USA) and then were sequenced on a MiSeq instrument (Illumina, USA) using a MiSeq reagent kit v3 (600 cycles; Illumina, USA). Forward and reverse reads were trimmed based on adaptors and quality using Cutadapt v1.10 (5). The trimmed reads were assembled using SPAdes v3.7.1 (6) and mapped back to the respective assembled genomes (An1F4, Ae3A3, and Ae3B3) using BWA-MEM v0.7.17 (r1188) (7) to generate individual coverage profiles.

**Editor** Vanja Klepac-Ceraj, Wellesley College

**Copyright** © 2023 Sigurdarson 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 Henrik Karring, [hka@igt.sdu.dk](mailto:hka@igt.sdu.dk).

\*Present address: Jens Jakob Sigurdarson, BioPol, Skagaströnd, Iceland.

The authors declare no conflict of interest.

**Received** 21 September 2022

**Accepted** 16 December 2022

**TABLE 1** Assembly results and statistics, numbers of essential genes, and accession numbers for the sequenced genomes of *S. alactolyticus* and *E. coli* strains isolated from pig feces

Bacterial isolate	Total length (bp)	No. of scaffolds	Size of longest scaffold (bp)	$N_{50}$ (bp)	Mean scaffold length (bp)	Coverage (x)	Mean G+C content (%)	No. of predicted coding sequences	Total no. of essential genes	No. of unique essential genes	BioSample accession no.	GenBank genome accession no.
<i>S. alactolyticus</i> strain An1F4	1,699,682	75	128,731	32,139	22,662	157	40	1,677	107	106	SAMD00517187	BRXM01000000
<i>E. coli</i> strain Ae3A3	5,300,654	138	341,114	102,593	38,410	50	51	4,951	107	106	SAMD00517185	BRXL01000000
<i>E. coli</i> strain Ae3B3	4,874,471	159	193,214	70,132	30,657	54	51	4,524	107	106	SAMD00517186	BRXM01000000

Prodigal v2.6.3 (8), hmmsearch (HMMER v3.1b2) (9), and BLAST+ v2.7.1 (10) were used to find essential genes. MEGAN v4.70.4 (11) was subsequently used to extract consensus taxonomic assignments for the mmgenome coverage plots (12). Barnap v0.8 (13) was used to extract the rRNA genes from each of the three assemblies (An1F4, Ae3A3, and Ae3B3), and the extracted 16S rRNA genes were classified using SINA v1.2.11 (<https://www.arb-silva.de/aligner>) (14). The results were analyzed in R v3.4.3 (15) through the RStudio integrated development environment (IDE) v1.1.383 using the mmgenome package v0.7.1 (12). The PhiX control DNA contig was removed from the assembly manually, and PhiX-screened genomes were annotated using the DDBJ Fast Annotation and Submission Tool (DFAST) v1.2.15 (16).

The genome of *S. alactolyticus* strain An1F4 has a total length of 1,699,682 bp, harbors 1,677 predicted coding sequences, and has a G+C content of 40%. *E. coli* strains Ae3A3 and Ae3B3 have genome sizes of approximately 5 Mbp, with 4,951 and 4,524 predicted coding sequences, respectively, and overall G+C contents of 51% (Table 1).

**Data availability.** The assembly and annotation for the three genomes are available in DDBJ/ENA/GenBank. The genome accession numbers are BRXN01000000 for *S. alactolyticus* strain An1F4, BRXL01000000 for *E. coli* strain Ae3A3, and BRXM01000000 for *E. coli* strain Ae3B3.

## ACKNOWLEDGMENTS

This work was supported by a grant from the Green Development and Demonstration Program (ManUREA Technology, file number 34009-15-0934) and a grant from the Danish Agricultural Agency (NoGas, file number 33010-NIFA-19-731) under the Danish Ministry of Environment and Food. The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

We thank Ulrik S. Justesen for assistance with the species identifications using MALDI-TOF MS.

J.J.S. and S.K. performed the experimental work under the supervision of S.S. and H.K. All authors contributed to the conception and design of the study and the analysis and interpretation of the data. J.J.S. and S.K. drafted the manuscript, and all authors contributed critically to the revisions of the manuscript and gave final approval for publication.

## REFERENCES

- McAuliffe GA, Takahashi T, Mogensen L, Hermansen JE, Sage CL, Chapman DV, Lee MRF. 2017. Environmental trade-offs of pig production systems under varied operational efficiencies. *J Clean Prod* 165:1163–1173. <https://doi.org/10.1016/j.jclepro.2017.07.191>.
- Petersen SO, Dorno N, Lindholm S, Feilberg A, Eriksen J. 2013. Emissions of CH<sub>4</sub>, N<sub>2</sub>O, NH<sub>3</sub> and odorants from pig slurry during winter and summer storage. *Nutr Cycl Agroecosyst* 95:103–113. <https://doi.org/10.1007/s10705-013-9551-3>.
- Duncan SH. 2002. Growth requirements and fermentation products of *Fusobacterium prausnitzii*, and a proposal to reclassify it as *Faecalibacterium prausnitzii* gen. nov., comb. nov. *Int J Syst Evol Microbiol* 52:2141–2146. <https://doi.org/10.1099/ijs.0.02241-0>.
- Browne HP, Forster SC, Anonye BO, Kumar N, Neville BA, Stares MD, Goulding D, Lawley TD. 2016. Culturing of 'unculturable' human microbiota reveals novel taxa and extensive sporulation. *Nature* 533:543–546. <https://doi.org/10.1038/nature17645>.
- Martin M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J* 17:10. <https://doi.org/10.14806/ej.17.1.200>.
- Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Pribelski 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>.
- Li H. 2013. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv* 1303.3997. <https://doi.org/10.48550/arXiv.1303.3997>.
- Hyatt D, Chen G, LoCascio PF, Land ML, Larimer FW, Hauser LJ. 2010. Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* 11:119. <https://doi.org/10.1186/1471-2105-11-119>.
- Mistry J, Finn RD, Eddy SR, Bateman A, Punta M. 2013. Challenges in homology search: HMMER3 and convergent evolution of coiled-coil regions. *Nucleic Acids Res* 41:e121. <https://doi.org/10.1093/nar/gkt263>.
- Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL. 2009. BLAST+: architecture and applications. *BMC Bioinformatics* 10:421. <https://doi.org/10.1186/1471-2105-10-421>.
- Huson DH, Mitra S, Ruscheweyh H, Weber N, Schuster SC. 2011. Integrative analysis of environmental sequences using MEGAN4. *Genome Res* 21:1552–1560. <https://doi.org/10.1101/gr.120618.111>.
- Karst SM, Kirkegaard RH, Albertsen M. 2016. mmgenome: a toolbox for reproducible genome extraction from metagenomes. *bioRxiv*. 059121. <https://doi.org/10.1101/059121>.
- Seemann T. 2013. Barnap 0.7: rapid ribosomal RNA prediction. <https://github.com/tseemann/barnap>.
- Pruesse E, Peplies J, Glöckner FO. 2012. SINA: accurate high-throughput multiple sequence alignment of ribosomal RNA genes. *Bioinformatics* 28:1823–1829. <https://doi.org/10.1093/bioinformatics/bts252>.
- R Core Team. 2017. R: a language and environment for statistical computing. <https://www.R-project.org>.
- Tanizawa Y, Fujisawa T, Nakamura Y. 2018. DFAST: a flexible prokaryotic genome annotation pipeline for faster genome publication. *Bioinformatics* 34:1037–1039. <https://doi.org/10.1093/bioinformatics/btx713>.