

# Draft Genome Sequence of a Novel *Lactobacillus salivarius* Strain Isolated from Piglet

Donald A. MacKenzie,<sup>a</sup> Kirsten McLay,<sup>b</sup> Stefan Roos,<sup>c</sup> Jens Walter,<sup>d</sup> David Swarbreck,<sup>b</sup> Nizar Drou,<sup>b</sup> Lisa C. Crossman,<sup>b,e\*</sup> Nathalie Juge<sup>a</sup>

The Gut Health and Food Safety Programme, Institute of Food Research, Norwich Research Park, Norwich, United Kingdom<sup>a</sup>; The Genome Analysis Centre, Norwich Research Park, Norwich United Kingdom<sup>b</sup>; Department of Microbiology, Swedish University of Agricultural Sciences, Uppsala, Sweden<sup>c</sup>; Department of Food Science and Technology, University of Nebraska, Lincoln, Nebraska, USA<sup>d</sup>; School of Biological Sciences, University of East Anglia, Norwich, United Kingdom<sup>e</sup>

\* Present address: Lisa C. Crossman, SequenceAnalysis.co.uk, Norwich Research Park, Norwich, United Kingdom.

***Lactobacillus salivarius* is part of the vertebrate indigenous microbiota of the gastrointestinal tract, oral cavity, and milk. The properties associated with some *L. salivarius* strains have led to their use as probiotics. Here we describe the draft genome of the pig isolate *L. salivarius* cp400, providing insights into host-niche specialization.**

Received 17 December 2013 Accepted 23 January 2014 Published 13 February 2014

Citation MacKenzie DA, McLay K, Roos S, Walter J, Swarbreck D, Drou N, Crossman LC, Juge N. 2014. Draft genome sequence of a novel *Lactobacillus salivarius* strain isolated from piglet. *Genome Announc.* 2(1):e01231-13. doi:10.1128/genomeA.01231-13.

Copyright © 2014 MacKenzie et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 3.0 Unported license](https://creativecommons.org/licenses/by/3.0/).

Address correspondence to Nathalie Juge, [nathalie.juge@ifr.ac.uk](mailto:nathalie.juge@ifr.ac.uk).

The Gram-positive bacterium *Lactobacillus salivarius* is an important member of the pig microbiota (1) and a promising probiotic candidate frequently isolated from human, porcine, and avian gastrointestinal tracts, many of which are producers of unmodified bacteriocins (2). The apparent ability of *L. salivarius* to produce very different bacteriocins using the same cellular machinery suggests an unknown versatility mechanism with respect to the production of this dominant probiotic trait (3).

To gain novel insights into the molecular basis underlying *L. salivarius* adaptation to the intestinal ecological niche, we have determined the genome sequence of the pig isolate *L. salivarius* cp400, isolated from preweaned piglet feces (Stefan Roos, unpublished data). Genomic DNA was purified using a modified method from Oh and colleagues (4) and used to generate over 4.65 Gbp of sequence on the Illumina HiSeq2000 platform. Reads passing default filter settings were assembled using Abyss (5) and generated 89 scaffolds containing 87 large scaffolds (>500 bp) and spanning 2,155,512 bp of sequence. This genome assembly of *L. salivarius* cp400 is 2,156,840 bp in length, with an average G+C content of 33.2%. Automatic gene prediction was performed using Glimmer3 and GeneMark (6, 7). Annotation was transferred from the related strain, *L. salivarius* ATCC 11741. Unique regions were manually annotated using Artemis (8), augmented with InterPro (9), TMHMM (10), and SignalP (11). A total of 2,044 protein-coding sequences was predicted, with a coding percentage of 83.7%. Coding density was 0.947 genes per kb with an average gene length of 883 bp. Phylogenetic analysis of the whole-genome sequence of cp400 with the genome sequences available for *L. salivarius* strains ATCC 11741 (human), CECT 5713 (human), UCC118 (human), GJ-24 (human), NIAS840 (chicken), and SMXD51 (chicken) indicates that cp400 is genetically distinct, falling into a separate phylogenetic clade. Comparison of cp400 with the above strains revealed the presence of a megaplasmid of approximately 0.29 Mb and 31.9% G+C, harboring a bacteriocin

gene cluster with flanking transposase genes. The cp400 megaplasmid showed similarity with the plasmids pMP188 from the human strain UC118 and pLS51A from the chicken isolate SMXD51 at 91.9% and 91.3% DNA identity, respectively, for matches of >2,000 bp. Analysis of the bacteriocin gene cluster revealed high homology to the Abp118 gene cluster, encoding a class IIb two-peptide bacteriocin composed of Abp118alpha, exhibiting antimicrobial activity, and Abp118beta, which enhanced the antimicrobial activity in *L. salivarius* UCC118 (12). The cp400 bacteriocin displayed 100% identity to the mature alpha and beta peptides of salivaricin P produced by five other pig intestinal isolates of *L. salivarius* (13), and the cluster also contained genes encoding homologues of AbpIP (inducer), AbpK (sensory transduction histidine kinase), AbpR (response regulator), and AbpD (accessory export), suggesting the presence of a functional salivaricin P cluster in the cp400 megaplasmid. There was lower similarity with the salivaricin cluster from human strain CRL1328 (14) at 87.2% DNA identity for matches over 2,000 bp.

Detailed analysis of the assembled cp400 genome will help determine specific genetic features driving genome-wide specialization of *L. salivarius* in response to a particular niche.

**Nucleotide sequence accession numbers.** This genome sequencing project has been deposited at EMBL/DDBJ/GenBank via the European Nucleotide Archive (ENA). The 89 fully annotated scaffolds have been deposited under the accession numbers [CBVR010000001](https://www.ncbi.nlm.nih.gov/nuccore/CBVR010000001) through [CBVR010000089](https://www.ncbi.nlm.nih.gov/nuccore/CBVR010000089).

## ACKNOWLEDGMENTS

We gratefully acknowledge the support of the Biotechnology and Biological Sciences Research Council (BBSRC). This research was funded by the BBSRC Institute Strategic Programme for The Gut Health and Food Safety (BB/J004529/1).

## REFERENCES

1. Leser TD, Amenuvor JZ, Jensen TK, Lindecrona RH, Boye M, Møller K. 2002. Culture-independent analysis of gut bacteria: the pig gastrointestinal tract microbiota revisited. *Appl. Environ. Microbiol.* 68:673–690. <http://dx.doi.org/10.1128/AEM.68.2.673-690.2002>.
2. Messaoudi S, Manai M, Kergourlay G, Prévost H, Connil N, Chobert JM, Dousset X. 2013. *Lactobacillus salivarius*: bacteriocin and probiotic activity. *Food Microbiol.* 36:296–304. <http://dx.doi.org/10.1016/j.fm.2013.05.010>.
3. O’Shea EF, O’Connor PM, Raftis EJ, O’Toole PW, Stanton C, Cotter PD, Ross RP, Hill C. 2012. Subspecies diversity in bacteriocin production by intestinal *Lactobacillus salivarius* strains. *Gut Microbes* 3:468–473. <http://dx.doi.org/10.4161/gmic.21417>.
4. Oh PL, Benson AK, Peterson DA, Patil PB, Moriyama EN, Roos S, Walter J. 2010. Diversification of the gut symbiont *Lactobacillus reuteri* as a result of host-driven evolution. *ISME J* 4:377–387. <http://dx.doi.org/10.1038/ismej.2009.123>.
5. Simpson JT, Wong K, Jackman SD, Schein JE, Jones SJ, Birol I. 2009. ABySS: a parallel assembler for short read sequence data. *Genome Res.* 19:1117–1123. <http://dx.doi.org/10.1101/gr.089532.108>.
6. Delcher AL, Bratke KA, Powers EC, Salzberg SL. 2007. Identifying bacterial genes and endosymbiont DNA with Glimmer. *Bioinformatics* 23:673–679. <http://dx.doi.org/10.1093/bioinformatics/btm009>.
7. Besemer J, Lomsadze A, Borodovsky M. 2001. GeneMarkS: a self-training method for prediction of gene starts in microbial genomes. Implications for finding sequence motifs in regulatory regions. *Nucleic Acids Res.* 29:2607–2618. <http://dx.doi.org/10.1093/nar/29.12.2607>.
8. Rutherford K, Parkhill J, Crook J, Horsnell T, Rice P, Rajandream MA, Barrell B. 2000. Artemis: sequence visualization and annotation. *Bioinformatics* 16:944–945. <http://dx.doi.org/10.1093/bioinformatics/16.10.944>.
9. Hunter S, Apweiler R, Attwood TK, Bairoch A, Bateman A, Binns D, Bork P, Das U, Daugherty L, Duquenne L, Finn RD, Gough J, Haft D, Hulo N, Kahn D, Kelly E, Laugraud A, Letunic I, Lonsdale D, Lopez R, Madera M, Maslen J, McAnulla C, McDowall J, Mistry J, Mitchell A, Mulder N, Natale D, Orengo C, Quinn AF, Selengut JD, Sigrist CJ, Thimmma M, Thomas PD, Valentin F, Wilson D, Wu CH, Yeats C. 2009. Interpro: the integrative protein signature database. *Nucleic Acids Res.* 37:D211–D215. <http://dx.doi.org/10.1093/nar/gkn785>.
10. Krogh A, Larsson B, von Heijne G, Sonnhammer EL. 2001. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J. Mol. Biol.* 305:567–580. <http://dx.doi.org/10.1006/jmbi.2000.4315>.
11. Emanuelsson O, Brunak S, von Heijne G, Nielsen H. 2007. Locating proteins in the cell using TargetP, SignalP and related tools. *Nat. Protoc.* 2:953–971. <http://dx.doi.org/10.1038/nprot.2007.131>.
12. Flynn S, van Sinderen D, Thornton GM, Holo H, Nes IF, Collins JK. 2002. Characterization of the genetic locus responsible for the production of ABP-118, a novel bacteriocin produced by the probiotic bacterium *Lactobacillus salivarius* subsp. *salivarius* UCC118. *Microbiology* 148: 973–984.
13. Barrett E, Hayes M, O’Connor P, Gardiner G, Fitzgerald GF, Stanton C, Ross RP, Hill C. 2007. Salivaricin P, one of a family of two-component antilisterial bacteriocins produced by intestinal isolates of *Lactobacillus salivarius*. *Appl. Environ. Microbiol.* 73:3719–3723. <http://dx.doi.org/10.1128/AEM.00666-06>.
14. O’Shea EF, O’Connor PM, Raftis EJ, O’Toole PW, Stanton C, Cotter PD, Ross RP, Hill C. 2011. Production of multiple bacteriocins from a single locus by gastrointestinal strains of *Lactobacillus salivarius*. *J. Bacteriol.* 193:6973–6982. <http://dx.doi.org/10.1128/JB.06221-11>.