Data Article

Proteomic dataset comparing strains of *Leptospira borgpetersenii* serovar Hardjo cultured at different temperatures

Ellie J. Putz\(^a\), Luis G.V. Fernandes\(^{a,b}\), Darrell O. Bayles\(^a\), John D. Lippolis\(^a\), Jarlath E. Nally\(^{a,*}\)

\(^a\)Agriculture Research Service, National Animal Disease Center, 1920 Dayton Ave, Ames, IA 50010 USA

\(^b\)Laboratorio de Desenvolvimento de Vacinas, Instituto Butantan, Avenida Vital Brasil, 1500, Sao Paulo, SP 05503-000, Brasil

**ARTICLE INFO**

Article history:
Received 8 July 2022
Revised 21 October 2022
Accepted 25 October 2022
Available online 29 October 2022

Dataset link: Proteome comparison of *Leptospira borgpetersenii* serovar Hardjo strains cultured at different temperatures (Original data)

Keywords:
Leptospirosis
*Leptospira*
Proteomics
HaN media
Hardjo

**ABSTRACT**

Leptospirosis is a global zoonotic bacterial disease which is a threat for humans and most mammals. Bacterin vaccines for leptospirosis are available however they are severely limited in cross protection between serogroups. *Leptospira* typically colonize the kidneys of reservoir hosts where they are subsequently shed in the urine and persist in the environment and can thus be indirectly or directly transmitted to incidental hosts. *Leptospira borgpetersenii* serovar Hardjo is the primary cause of leptospirosis in cattle which can result in abortion, unhealthy calves, and rebreed problems. This dataset comprises proteomic profiles of four strains of *L. borgpetersenii* serovar Hardjo propagated at the routinely utilized culture temperature of 29 °C, and a newly achieved culture temperature of 37 °C, which more closely emulates the temperature of an infected host. The strains analyzed include JB197 (established strain that causes Hardjo atypical acute disease in the hamster model of leptospirosis), HB203 (established strain, causes typical chronic disease in hamsters), as well as TC129 and TC273 (recently isolated strains from the central United States). Differential expression profiles were detected not only between strains but also within strains between cul-

DOI of original article: 10.1016/j.jprot.2022.104602

* Corresponding author.
E-mail address: jarlath.nally@usda.gov (J.E. Nally).

https://doi.org/10.1016/j.dib.2022.108713
2352-3409/Published by Elsevier Inc. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/)
ture temperatures. Mass spectrometry data are available via ProteomeXchange with identifier PXD032831.

Published by Elsevier Inc.
This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/)

### Specifications Table

| Subject                          | Microbiology: Bacteriology                  |
|---------------------------------|---------------------------------------------|
| Specific subject area           | Proteomic profile comparisons of strains of Leptospira borgpetersenii |
| Type of data                    | Table                                       |
| How the data were acquired      | Mass spectrometry and peptide chromatography utilized the Thermo Dionex UltiMate 3000 RSLCnano system; Mass spectrometry raw data files were processed with the MaxQuant software, and additional analysis was conducted with the Perseus software. |
| Data format                     | Analyzed.                                  |
| Description of data collection  | JB197, HB203, TC129, and TC273 strains of L. borgpetersenii serovar Hardjo were cultured at 29 °C and 37 °C. Whole cell lysate proteomic profiles were evaluated by mass spectrometry and analyzed with MaxQuant and Perseus software. |
| Data source location            | USDA Agricultural Research Service, National Animal Disease Center Ames, IA, USA |
| Data accessibility              | Repository name: PRIDE Proteomics Identifications Database Data identification number: PXD032831. (https://www.ebi.ac.uk/pride/archive/projects/PXD032831) |
| Related research article        | E.J. Putz, L.G.V Fernandes, S. Sivasankaran, D.O. Bayles, D.P. Alt, J.D. Lippolis, J.E. Nally, Some like it hot, some like it cold: Proteome comparison of Leptospira borgpetersenii serovar Hardjo strains propagated at different temperatures, Journal of Proteomics 262 (2022) 104.602. (https://doi.org/10.1016/j.jprot.2022.104602) |

### Value of the Data

- These data illustrate distinct proteomic profiles associated with highly similar genomic strains of L. borgpetersenii serovar Hardjo and demonstrate the role of temperature, as encountered during host infection, in regulating protein expression.
- These data illustrate strain to strain variation in protein expression and provide reference comparisons for additional strains and serovars of L. borgpetersenii, as well as additional species of Leptospira.
- Understanding strain to strain variation is critical to bacterin vaccine design and efficacy.
- The proteomic expression information provided here serves as a reference for the behavior of L. borgpetersenii serovar Hardjo strains in response to change in temperatures encountered during host infection.

### 1. Data Description

Data describes whole cell proteomics of four strains of L. borgpetersenii serovar Hardjo (JB197, HB203, TC129, TC273) cultured at 29 °C and 37 °C. Raw mass spectrometry data was analyzed by MaxQuant and Perseus for differential protein expression of interest. The 50 most highly expressed proteins across all strains and conditions are shown in Table 1. Full data and specific
| Protein IDs | Fasta headers |
|-------------|---------------|
| Q6J0P4 | Immunodominant outer membrane protein LipL32 |
| P61438 | 60 kDa chaperonin |
| Q04PT6 | Elongation factor Tu |
| Q04UY1 | LigB lipoprotein |
| Q04QD6 | Electron transfer flavoprotein, alpha subunit |
| Q054E1 | DNA-directed RNA polymerase subunit beta |
| M6XN61 | Flagellin |
| P61443 | Chaperone protein DnaK |
| Q04PQ5 | Uncharacterized protein |
| Q8F0S2 | DNA-directed RNA polymerase subunit beta n |
| Q04RL2 | Uncharacterized protein |
| Q04TE2 | Cysteine synthase |
| Q04S18 | ATP synthase subunit beta |
| M6U9A6 | Uncharacterized protein |
| P61436 | 10 kDa chaperonin |
| A0A1B9F1H9 | ArsR family transcriptional regulator |
| A0A0E3AZB2 | LysM domain-containing protein |
| M6A6Z4 | 4Fe-4S cluster domain protein |
| Q04UP0 | Transcription elongation factor GreA |
| M3ERF2 | Glutamine synthetase, type I |
| Q72MM7 | Uncharacterized protein |
| Q04PW5 | DNA-directed RNA polymerase subunit alpha |
| M6A6K1 | Redoxin |
| Q04Y01 | Elongation factor G |
| Q04UP1 | Endoflagellar filament sheath protein |
| Q04NS0 | HSP90 |
| A0A1D7US12 | OmpA-like domain-containing protein |
| Q04P78 | ABC transp_aux domain-containing protein |
| Q04VX0 | Isocitrate dehydrogenase [NADP] |
| Q04QD7 | Electron transfer flavoprotein, beta subunit |
| A0A1D7V0C6 | TPR, REGION domain-containing protein |
| Q72SY1 | ATP synthase subunit alpha |
| M3GL54 | Methyl-accepting chemotaxis protein signaling domain protein |
| M6CD43 | Uncharacterized protein |
| A0A0C5 | Cytoplasmic membrane lipoprotein LipL31 |
| Q04TN1 | Methylesterase/methyltransferase |
| Q04S8 | Nucleoside S1 |
| Q04QS3 | Succinate-CoA ligase [ADP-forming] subunit beta |

Table 1
The fifty most highly expressed proteins across all strains and conditions based on abundance scores generated by MaxQuant and Perseus analysis. Proteins identified by UniProt ID and Fasta annotation. Shown are average abundance scores by strain and condition combination over four biological replicates.

(continued on next page)
Table 1 (continued)

| Protein IDs | Fasta headers                                                                 | JB197 29 °C | JB197 37 °C | HB203 29 °C | HB203 37 °C | TC129 29 °C | TC129 37 °C | TC273 29 °C | TC273 37 °C |
|-------------|-------------------------------------------------------------------------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|
| V6iAH4      | Acetyl-CoA C-acyltransferase                                                   | 29.63       | 29.23       | 29.80       | 29.54       | 30.19       | 30.26       | 29.69       | 29.70       |
| Q6GXD1      | Transmembrane outer membrane protein                                          | 30.87       | 30.50       | 30.06       | 28.45       | 30.24       | 29.31       | 29.54       | 28.84       |
| Q72VT0      | Phosphoenolpyruvate carboxykinase (ATP)                                        | 29.29       | 29.33       | 29.37       | 29.66       | 30.02       | 30.01       | 29.98       | 29.78       |
| Q04PG8      | Cytochrome c oxidase subunit 2                                                | 29.81       | 29.45       | 29.55       | 29.73       | 29.68       | 29.57       | 29.86       | 29.78       |
| Q04QF8      | Uncharacterized protein                                                        | 29.99       | 29.86       | 30.13       | 29.18       | 29.51       | 29.59       | 29.73       | 29.31       |
| Q04J9       | Polyribonucleotide nucleotidyldtransferase                                    | 29.54       | 29.21       | 29.81       | 29.50       | 29.81       | 29.69       | 30.00       | 29.52       |
| A0A540UB21  | Re/Si-specific NAD(P)\(^{+}\) transhydrogenase subunit alpha                 | 29.80       | 29.69       | 29.16       | 29.70       | 29.70       | 29.82       | 29.48       | 29.68       |
| Q04P68      | 3-hydroxyacyl-CoA dehydrogenase                                                | 29.29       | 29.22       | 29.03       | 29.62       | 30.23       | 30.45       | 29.44       | 29.65       |
| Q04W30      | Membrane protein insertase YidC                                                | 29.81       | 29.44       | 29.29       | 29.45       | 30.00       | 29.81       | 29.72       | 29.34       |
| Q04U49      | Chemotaxis protein histidine kinase                                            | 29.72       | 29.51       | 30.37       | 29.72       | 28.70       | 29.56       | 29.36       | 29.57       |
| A0A2H1XGQ9  | Fructose-bisphosphate aldolase class I                                         | 29.35       | 29.36       | 29.32       | 29.65       | 29.45       | 29.87       | 29.45       | 29.70       |
| Q04NN6      | Adenosylhomocysteinase                                                         | 29.04       | 29.27       | 29.07       | 29.68       | 29.88       | 29.72       | 29.60       | 29.87       |

contrasts are available in Supplementary Table 1 (includes all output data from Perseus DE analysis. Strain JB197 at 29 °C was used as the reference condition).

2. Experimental Design, Materials and Methods

2.1. Culture of Leptospira Strains

*Leptospira borgpetersenii* serovar Hardjo-bovis strains JB197 (accession: PRJNA16148), HB203 (accession: PRJNA384237), TC129, and TC273 (accession: PRJNA759631, [1]) were cultured in HAN media at 29 °C and 37 °C respectively [2]. At mid-late log phase, bacteria were centrifuged (4 °C, 10,000 x g, 30 min), and the pellet was washed with cold TE buffer twice, and stored at -80 °C. Four biological replicates per each condition were prepared for proteomics.

2.2. Trypsin Digestion

Samples were digested according to manufacturer’s directions with Trypsin/Lys-C Mix (Promega Corporation, Madison, WI) as previous reported [3]. After digestion, samples were run through a desalting Pierce C18 spin column (Thermo Fisher Scientific, West Palm Beach, FL) according to manufacturer’s instructions. After drying by vacuum (Speedvac), samples were resuspended in 95% H₂O 5% acetonitrile and 0.1% formic acid.

2.3. Mass Spectrometry

Mass spectrometry and peptide chromatography were completed as done previously [4]. The Thermo Dionex UltiMate 3000 RSLCnano system (Thermo Fisher Scientific, FL, USA) instrument
was utilized for peptide chromatography. Chromatography was accomplished with a gradient of buffer A (95% H2O: 5% acetonitrile and 0.1% formic acid) and buffer B (5% H2O: 95% acetonitrile and 0.1% formic acid) over an Acclaim PepMap 100 C18 column. The LC was connected to a Nanospray Ion Source (Thermo Fisher Scientific, FL, USA) on a LTQ Orbitrap Velos Pro (Thermo Fisher Scientific, FL, USA) mass spectrometer. The mass spectrometer used these settings: MSMS scanning resolution was 15,000, the 400 to 2000 m/z MS resolution was 30,000, the repeated mass duration exclusion was 1 min, and CID activation was utilized (normalized with a collision energy of 35). A signal minimum of 5000 was required for the MS Top 10. The L. borgpetersenii serovar Hardjo-Bovis database was downloaded from UniProt (August 2020). The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE [5] partner repository with the dataset identifier PXD032831.

2.4. Analysis of Proteomic Data

Raw mass spectrometry data went through initial processing by MaxQuant (Version 1.6.7.0) [6] followed by subsequent Perseus analysis [7] as reported previously [3]. MaxQuant fixed modifications consisted of Carbamidomethyl (C) and variable modifications included acetyl (Protein N-term), deamidation (NQ), oxidation (M), and Phosphorylation (STY). In Perseus, data was log2 transformed, single peptide identified proteins were removed, and differential expression contrasts were calculated and reported in completion in Supplementary Table 1. For summary of the fifty most highly expressed proteins across strain and temperature shown in Table 1, all abundance scores generated in MaxQuant and Perseus were summed and sorted to identify the fifty most abundant proteins. Table one shows the average abundance score per strain and temperature condition among four biological replicates.

2.5. Statistical Analysis

Perseus software was used to analyze proteomics data [7]. Differential expression utilized a P-value < 0.01 for significant differences.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data Availability

Proteome comparison of Leptospira borgpetersenii serovar Hardjo strains cultured at different temperatures (Original data) (PRIDE Proteomics Identification Database).

CRediT Author Statement

Ellie J. Putz: Conceptualization, Methodology, Formal analysis, Investigation, Visualization, Writing – original draft, Writing – review & editing; Luis G.V. Fernandes: Methodology, Formal analysis, Data curation, Writing – review & editing; Darrell O. Bayles: Methodology, Formal analysis, Validation, Writing – review & editing; John D. Lippolis: Conceptualization, Methodology, Formal analysis, Writing – review & editing; Jarlath E. Nally: Conceptualization, Validation, Formal analysis, Investigation, Data curation, Supervision, Writing – original draft, Writing – review & editing.
Acknowledgments

The authors extend their utmost appreciation to Duane Zimmerman for his excellent mass spectrometer sample processing. This work was solely funded by ARS, USDA. USDA is an equal opportunity provider and employer.

Supplementary Materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.dib.2022.108713.

References

[1] E.J. Putz, D.O. Bayles, D.P. Alt, J.E. Nally, Complete genome sequence of four strains of *Leptospira borgpetersenii* serovar Hardjo isolated from cattle in the Central United States, J. Genom. 10 (2022) 45–48.
[2] R.L. Hornsby, D.P. Alt, J.E. Nally, Isolation and propagation of leptospires at 37 °C directly from the mammalian host, Sci. Rep. 10 (1) (2020) 9620.
[3] L.G. Fernandes, E.J. Putz, J.A. Stasko, J.D. Lippolis, A.L.T.O. Nascimento, J.E. nally, Evaluation of LipL32 and LigA/LigB knockdown mutants in *Leptospira interrogans* serovar Copenhageni; impacts to proteome and virulence, Front. Microbiol. 12 (2022) 799012.
[4] T.A. Reinhardt, J.D. Lippolis, Characterization of bovine mammary gland dry secretions and their proteome from the end of lactation through day 21 of the dry period, J. Proteom. 223 (2020) 103831.
[5] Y. Perez-Riverol, J. Bai, C. Bandla, D. García-Seisdedos, S. Hewapathirana, S. Kamatchinathan, D.J. Kundu, A. Prakash, A. Frericks-Zipper, M. Eisenacher, M. Walzer, S. Wang, A. Brazma, J.A. Vizcaíno, The PRIDE database resources in 2022: a hub for mass spectrometry-based proteomics evidences, Nucleic. Acids Res. 50 (D1) (2022) D543–d552.
[6] J. Cox, M. Mann, MaxQuant enables high peptide identification rates, individualized P.P.B.-range mass accuracies and proteome-wide protein quantification, Nat. Biotechnol. 26 (12) (2008) 1367–1372.
[7] S. Tyanova, T. Temu, P. Sinticyn, A. Carlson, M.Y. Hein, T. Geiger, M. Mann, J. Cox, The Perseus computational platform for comprehensive analysis of (pro)teomics data, Nat. Methods 13 (9) (2016) 731–740.