



ELSEVIER

Contents lists available at ScienceDirect

## Data in Brief

journal homepage: [www.elsevier.com/locate/dib](http://www.elsevier.com/locate/dib)



### Data Article

# Mass spectrometry data from identification of host-defense related proteins using label-free quantitative proteomic analysis of milk whey from cows with *Staphylococcus aureus* subclinical mastitis



Shaimaa Abdelmegid<sup>a</sup>, Jayaseelan Murugaiyan<sup>b,\*</sup>,  
Mohamed Abo-Ismael<sup>c</sup>, Jeff Caswell<sup>d</sup>, David Kelton<sup>e</sup>,  
Gordon Kirby<sup>a,\*</sup>

<sup>a</sup> Department of Biomedical Sciences, Ontario Veterinary College, University of Guelph, Guelph, Canada

<sup>b</sup> Freie Universität Berlin, Institute of Animal Hygiene and Environmental Health, Centre for Infectious Medicine, Berlin, Germany

<sup>c</sup> Department of Animal Biosciences, Ontario Agriculture College, University of Guelph, Guelph, Canada

<sup>d</sup> Department of Pathobiology, Ontario Veterinary College, University of Guelph, Guelph, Canada

<sup>e</sup> Department of Population Medicine, Ontario Veterinary College, University of Guelph, Guelph, Canada

### ARTICLE INFO

#### Article history:

Received 12 February 2018

Received in revised form

2 January 2019

Accepted 4 January 2019

Available online 10 January 2019

### ABSTRACT

This dataset is associated with our research article 'Identification of host-defence related proteins using label-free quantitative proteomic analysis of milk whey from cows with *Staphylococcus aureus* Subclinical mastitis' published in International Journal of Molecular Sciences. Milk samples were collected from cows suffering from *S. aureus*-associated subclinical mastitis and the proteins abundance were identified in comparison with samples collected from the control cows using liquid chromatography-mass spectrometry (LC-MS)-based label free proteomics analysis. Following the MS measurements, the raw spectra were processed using MaxQuant-Andromeda software and the protein identification was carried out through a search against Uniprot FASTA files of the *Bos taurus* reference proteome. Perseus software analysis was applied for computation of protein abundance. The raw file

\* Corresponding authors.

E-mail addresses: [sabdelme@uoguelph.ca](mailto:sabdelme@uoguelph.ca) (S. Abdelmegid), [jayaseelan.murugaiyan@fu-berlin.de](mailto:jayaseelan.murugaiyan@fu-berlin.de) (J. Murugaiyan), [maboisma@uoguelph.ca](mailto:maboisma@uoguelph.ca) (M. Abo-Ismael), [jcaswell@uoguelph.ca](mailto:jcaswell@uoguelph.ca) (J. Caswell), [dkelton@ovc.uoguelph.ca](mailto:dkelton@ovc.uoguelph.ca) (D. Kelton), [gkirby@uoguelph.ca](mailto:gkirby@uoguelph.ca) (G. Kirby).

<https://doi.org/10.1016/j.dib.2019.01.013>

2352-3409/© 2019 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

presented in this DiB was deposition at PRIDE repository in the PRIDE repository with the dataset identifier PXD007516.

© 2019 The Authors. 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 area	Biology
More specific subject area	Label-free quantitative proteomics, <i>Staphylococcus aureus</i> , Bovine mastitis, cow milk, whey samples
Type of data	Raw data, table and Excel output files
How data was acquired	Mass spectrometry data acquired using Orbitrap Elite™ Hybrid Ion Trap-Orbitrap Mass Spectrometer coupled with an EASY-nLC™ 1200 system (ThermoFisher Scientific).
Data format	Raw and excel sheet
Experimental factors	Milk samples were collected from normal and cows suffering from sub-clinical associated with <i>Staphylococcus aureus</i> . The whey samples were separated, whey proteins subjected to trypsin digestion and label free quantitative analysis using liquid chromatography-mass spectrometry (LCMS).
Experimental features	Milk sample collection, assessment of sub-clinical mastitis status, whey separation, protein estimation, trypsin digestion and mass spectrometry analysis using LCMS method. MaxQuant software was used for protein identification through search against FASTA file of the <i>Bos taurus</i> reference proteome and identification of comparative protein abundance out using Perseus software.
Data source location	University of Guelph, Guelph, Ontario, Canada
Data accessibility	Data available at PRIDE: <a href="https://www.ebi.ac.uk/pride/archive/study/PXD007516">PXD007516</a>
Related research article	S. Abdelmegid, J. Murugaiyan, M. Abo-Ismael, J. L. Caswell, D. Kelton, G. M. Kirby. Identification of host defense-related proteins using label-free quantitative proteomic analysis of milk whey from cows with <i>Staphylococcus aureus</i> subclinical mastitis. <i>Int J Mol Sci.</i> (2017), 19(1): 78

## Value of the data

- Comparative listings of proteins associated with *S. aureus* associated subclinical mastitis are useful to those examining the relative abundance of host defense proteins in milk during intramammary infection
- Availability of raw LC-MS data of whey samples isolated from normal and *S. aureus* associated sub-clinical mastitis milk samples to other researchers
- This raw data is useful in direct comparison with milk samples collected with similar or different diseases.

## 1. Data

This mass spectrometry data-in-brief is associated with the research article aimed at identification of protein abundantly present in milk samples of cows suffering from sub-clinical mastitis due to *S. aureus* [1]. The raw MS data and associated files can freely downloadable from the PRIDE repository

with an identified PXD007516 [2]. The identified proteins are listed in a [Supplementary Table](#), while those proteins identified as differentially abundant proteins discussed in Abdelmegid et al. [1] are shown in Fig. 1. Log<sup>2</sup>-fold changes range from a 2.9-fold reduction (Q9XSG3: Isocitrate dehydrogenase) to a 7.9-fold increase (Q8SPP7: Peptidoglycan recognition protein 1). Fig. 2 illustrates the steps involved in preparation of the milk samples and analysis, identification and quantification of the whey proteins.

## 2. Experimental design, materials and methods

Fig. 2 depicts the experimental design and work flow of the data discussed in Abdelmegid et al. [1]. In brief, milk was collected from dairy cows that were classified into two groups: healthy cows to serve as controls and cows with subclinical mastitis associated with *S. aureus*. Label-free quantitative proteomics analysis was applied for identification of proteins and their comparative abundance among the control and mastitis groups.

## 3. Milk samples and classification

Milk samples were collected from dairy cows at the Elora Research Station- Dairy Facility at the University of Guelph in accordance to the guidelines set by the Animal Care and Use Committee of the University of Guelph. (AUP # 1424, Approved time line – Monday, April 30, 2012 to Monday, July 11, 2016). The cows were classified as being either healthy or having subclinical mastitis based on history, physical examination and testing of the milk samples by Somatic Cell Counting and bacterial culture.

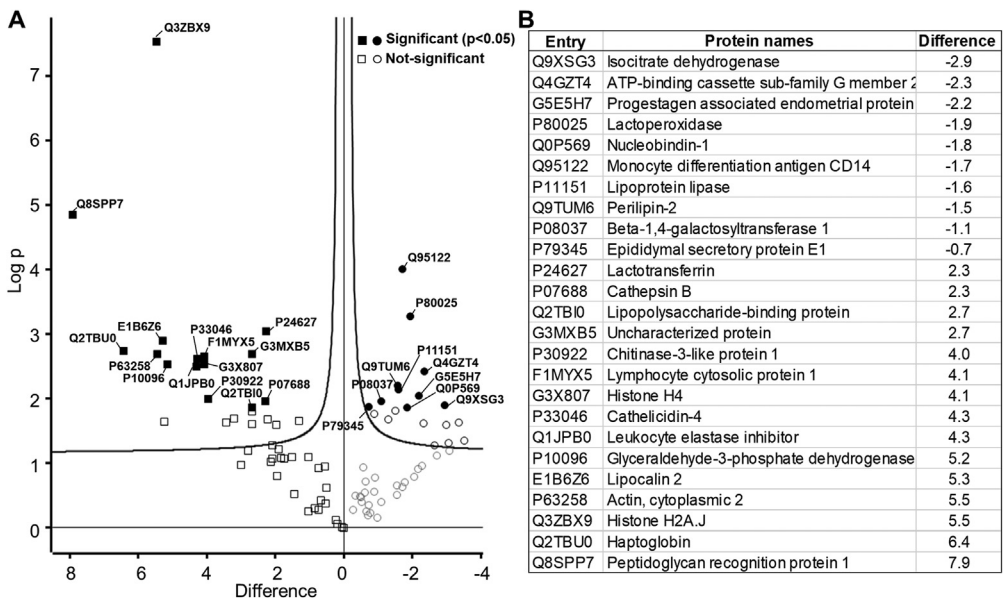


Fig. 1. (A) Volcano plot comparing the milk whey control and *S. aureus* associated sub-clinical mastitis sample groups. The *p*-values were calculated by Student *t*-test analysis using log<sub>2</sub> transformed peptide intensities. (B) List of the proteins identified as significantly differentially abundant.

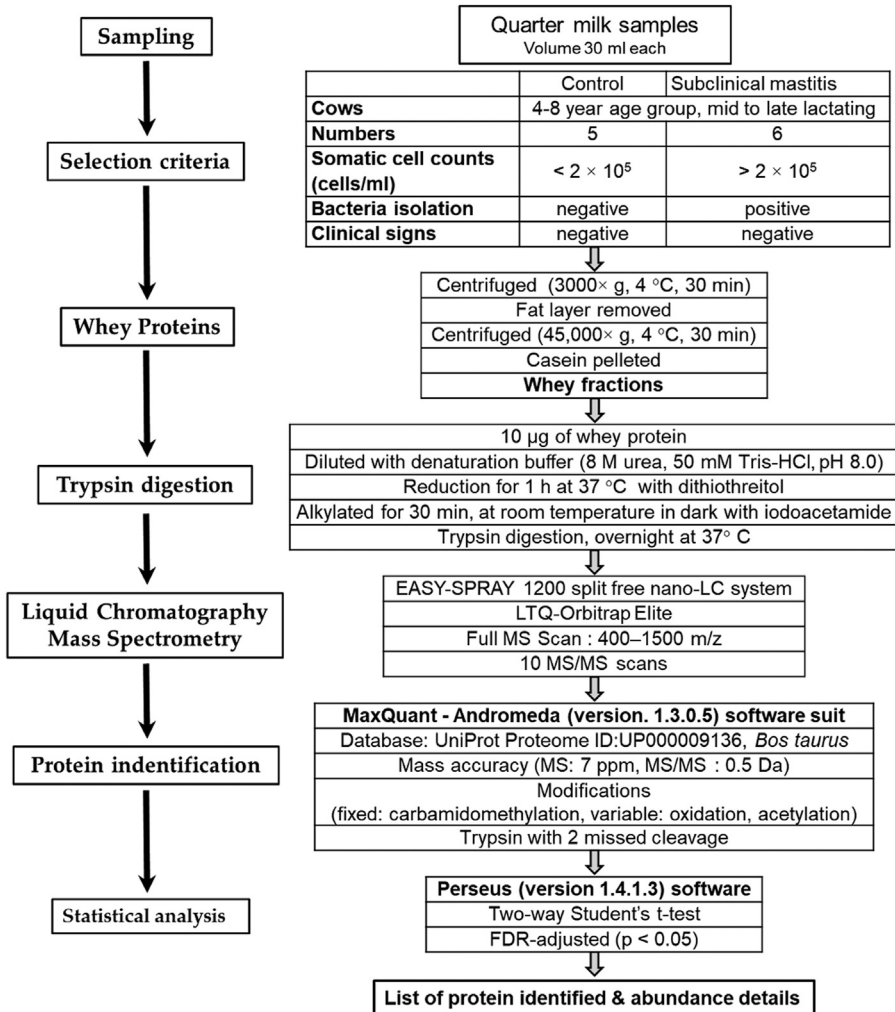


Fig. 2. Experimental design and work flow.

#### 4. Mass spectrometry analysis

Whey protein was prepared by removing fat and casein micelles by centrifugation followed by trypsin digestion. Separation of peptides was carried out by liquid chromatography and an Orbitrap-based mass spectrometry system (Orbitrap Elite™ Hybrid Ion Trap coupled with an EASY-nLC™ 1200 system; ThermoFischer Scientific) as described in [1].

#### 5. Data analysis

MaxQuant-Andromeda software suit was utilised to process the raw spectra and to search against the *Bos taurus* protein reference proteomes (UniProt Proteome ID: UP000009136). Label-free quantification of identified proteins was performed and statistically significant differences in protein abundance in the subclinical group in comparison to that of the control group was determined using Perseus 1.4.1.3 software ([http://141.61.102.17/perseus\\_doku/doku.php?id=start](http://141.61.102.17/perseus_doku/doku.php?id=start)) [1,3–5].

## 6. Mass spectrometry dataset deposit

The mass spectrometry data was deposited at the ProteomeXchange (PX) Consortium [2] via the PRIDE (PRoteomics IDentifications) partner repository and is accessible with the dataset identifier [PXD007516](https://doi.org/10.1016/j.dib.2019.01.013).

## Acknowledgements

We thank the PRIDE team for their assistance in the MS data deposition.

## Transparency document. Supplementary material

Transparency data associated with this article can be found in the online version at <https://doi.org/10.1016/j.dib.2019.01.013>.

## Appendix A. Supplementary material

Supplementary data associated with this article can be found in the online version at <https://doi.org/10.1016/j.dib.2019.01.013>.

## References

- [1] S. Abdelmegid, J. Murugaiyan, M. Abo-Ismael, J.L. Caswell, D. Kelton, G.M. Kirby, Identification of host defense-related proteins using label-free quantitative proteomic analysis of milk whey from cows with *Staphylococcus aureus* subclinical mastitis, *Int. J. Mol. Sci.* 19 (1) (2017).
- [2] E.W. Deutsch, A. Csordas, Z. Sun, A. Jarnuczak, Y. Perez-Riverol, T. Ternent, D.S. Campbell, M. Bernal-Llinares, S. Okuda, S. Kawano, The ProteomeXchange consortium in 2017: supporting the cultural change in proteomics public data deposition, *Nucleic Acids Res.* 45 (2016) D1100–D1106.
- [3] Y. Benjamini, Y. Hochberg, Controlling the false discovery rate—a practical and powerful approach to multiple testing, *J. R. Stat. Soc. Ser. B Methodol.* 57 (1995) 289–300.
- [4] J. Cox, N. Neuhauser, A. Michalski, R.A. Scheltema, J.V. Olsen, M. Mann, Andromeda: a peptide search engine integrated into the MaxQuant environment, *J. Proteome Res.* 10 (2011) 1794–1805.
- [5] J. Cox, M.Y. Hein, C.A. Lubner, I. Paron, N. Nagaraj, M. Mann, Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ, *Mol. Cell. Proteom.* 13 (2014) 2513–2526.