Data in Brief 30 (2020) 105665

Contents lists available at ScienceDirect

Data in Brief

journal homepage: www.elsevier.com/locate/dib

Data Article

Proteomic datasets of uninfected and *Staphylococcus aureus*-infected goat milk

Salvatore Pisanu[a], Carla Cacciotto[a], Daniela Pagnozzi[a], Sergio Uzzau[a,b,c], Claudia Pollera[d], Martina Penati[d], Valerio Bronzo[d], Maria Filippa Addis[a,c,d,*]

[a] *Porto Conte Ricerche, Alghero, Italy*
[b] *Dipartimento di Scienze Biomediche, Università degli Studi di Sassari, Sassari, Italy*
[c] *Mediterranean Center for Disease Control, Sassari, Italy*
[d] *Dipartimento di Medicina Veterinaria, Università degli Studi di Milano, Milan, Italy*

ARTICLE INFO

Article history:
Received 21 April 2020
Accepted 28 April 2020
Available online 6 May 2020

Keywords:
Goat milk
Proteomics
FASP
Mass spectrometry
NSAF

ABSTRACT

We present a proteomic dataset generated from half-udder Alpine goat milk. The milk samples belonged to 3 groups: i) mid-lactation, low somatic cell count, uninfected milk (MLU, n=3); ii) late lactation, high somatic cell count, uninfected milk (LHU, n=3); and late lactation, high somatic cell count, *Staphylococcus aureus* subclinically infected milk (LHS, n=3). The detailed description of results is reported in the research article entitled "Impact of *Staphylococcus aureus* infection on the late lactation goat milk proteome: new perspectives for monitoring and understanding mastitis in dairy goats". After milk defatting, high speed centrifugation and trypsin digestion of milk with the FASP protocol, peptide mixtures were analyzed by LC-MS/MS on a Q-Exactive. Peptide identification was carried out using Sequest-HT in Proteome Discoverer. Then, the Normalized Abundance Spectrum Factor (NSAF) value was calculated by label free quantitation using the spectral counting approach, and Gene Ontology (GO) annotation by Uniprot was carried out by reporting biological process, molecular function and cellular component. The MS data have been deposited to the ProteomeXchange via the PRIDE with the dataset identifier PXD017243.

* Corresponding author: Maria Filippa Addis, Department of Veterinary Medicine, University of Milan, Via dell'Università 6, 26900 Lodi (LO). Tel. ++39 02 503 34515
E-mail address: filippa.addis@unimi.it (M.F. Addis).

https://doi.org/10.1016/j.dib.2020.105665

Specifications Table

Subject	Biology
Specific subject area	Proteomics
Type of data	Tables Files raw Files .msf
How data was acquired	Sample preparation: filter-aided sample preparation (FASP). Instrument: Q Exactive mass spectrometer interfaced with an UltiMate 3000 RSLCnanoLC system (Thermo Fisher Scientific). Peptide identification: Proteome Discoverer (version 1.4; Thermo Scientific) SEQUEST-HT. Gene ontology annotation: UniProt Knowledgebase (UniProtKB).
Data format	RAW data (Thermo proprietary), msf files (Proteome Discoverer) and supplementary Excel files (1-3) with data analysis output.
Parameters for data collection	Half-udder goat milk collected along the course of an entire lactation (middle and late lactation) from Alpine goat half-udders.
Description of data collection	Half-udder goat milk was defatted, submitted to high-speed centrifugation, and proteins were digested by Filter-aided sample preparation (FASP). Peptide mixtures were analyzed by mass spectrometry with a Q-Exactive interfaced with an UltiMate 3000 RSLCnanoLC system (Thermo Scientific, San Jose, CA, USA). Identifications and Peptide spectrum matches (PSMs) data were obtained with Proteome Discoverer and used to calculate protein relative abundance and to perform Gene Ontology annotation with the UniprotKB database.
Data source location	Tramariglio, Alghero (Sassari), Italy
Data accessibility	Analyzed data is provided in supplementary files 1-3 within this article. Raw data repository name: PRIDE [1]. Dataset identifier: PXD017243. direct URL to data (raw and msf files): http://proteomecentral.proteomexchange.org/cgi/GetDataset?ID=PXD017243
Related research article	S. Pisanu, C. Cacciotto, P. Pagnozzi, S. Uzzau, C. Pollera, M. Penati, V. Bronzo, M.F. Addis. Impact of *Staphylococcus aureus* infection on the late lactation goat milk proteome: new perspectives for monitoring and understanding mastitis in dairy goats**.** Journal of Proteomics, submitted [2].

Value of the data

- This data helps to identify the proteomic changes specifically occurring in the milk of late lactation goats with a subclinical *Staphylococcus aureus* intramammary infection (IMI) when compared to the physiological changes occurring in the same period in uninfected goats.
- This data can serve to identify potential biomarkers for a more reliable monitoring of intramammary infections (IMIs) in late lactation goats than the somatic cell count.
- Protein identification, relative abundance and GO annotation data can provide an additional knowledge to better understand the mechanisms involved in infectious/inflammatory processes in the goat udder.

1. Data Description

The goat milk proteome was analyzed by a shotgun proteomics approach. Stage of lactation, somatic cell count (SCC) and bacteriological culture results were used to categorize milk samples and assemble the experimental groups, as reported in Table 1. The three proteomic profiles of goat milk are reported in Supplementary Tables 1, 2 and 3.

Table 1
Somatic cell count and bacteriological culture results of the half-udder goat milk samples used for generating the dataset.

Sample number	Stage of lactation[b]	SCC[a] (cells/mL x 10^3)	Bacteriology result	Acronym
1	Middle	13	Sterile	MLU[c]
2	Middle	16	Sterile	MLU
3	Middle	27	Sterile	MLU
4	Late	2488	Sterile	LHU[d]
5	Late	2942	Sterile	LHU
6	Late	3365	Sterile	LHU
7	Late	3936	*Staphylococcus aureus*	LHS[e]
8	Late	3938	*Staphylococcus aureus*	LHS
9	Late	4066	*Staphylococcus aureus*	LHS

[a] SCC: Somatic Cell Count
[b] Middle, 30-60 DIM; Late, > 250 DIM
[c] MLU: middle, low, uninfected
[d] LHU: late, high, uninfected
[e] LHS: late, high, *S. aureus*.

A total of 173, 290 and 446 non-redundant proteins were identified in the mid-lactation, uninfected (MLU, Supplementary Table 1), late lactation, uninfected (LHU, Supplementary Table 2) and late lactation, *Staphylococcus aureus*-infected (LHS, Supplementary Table 3) half-udder goat milk, respectively.

Moreover, for each milk proteome, protein abundances expressed as NSAF and GO analysis components are listed in the Supplementary Tables. GO analysis includes biological process, molecular function, and cellular components.

2. Experimental Design, Materials, and Methods

2.1. Goat milk samples

Goat milk samples were selected from a sample bank collected from Alpine goats, as described previously [2,3]. Bacteriological analysis and somatic cell count were carried out on all samples. Hereafter, milk samples were centrifuged to remove fat ring and underwent the filter-aided sample preparation (FASP) protocol to obtain peptide mixtures, as described by Pisanu et al. [2,4].

2.2. LC-MS/MS analysis

All peptide mixtures were carried out on a Q Exactive interfaced with an UltiMate 3000 RSLCnanoLC system (Thermo Fisher Scientific, San Jose, CA, USA), as previously described by Pisanu et al [2,4], using 4 µg of peptide mixture of each sample. Separation of peptide mixtures was performed using a linear gradient of 245 minutes from 5 to 37.5% eluent B (0.1% formic acid in 80% acetonitrile) in eluent A (0.1% formic acid) [2,4].

Peptide identification was performed using Proteome Discoverer (version 1.4; Thermo Scientific) SEQUEST-HT as the search engine with database custom composed by *Bos taurus* (Swiss-Prot), *Capra hircus* (TrEMBL) and *Staphylococcus* (TrEMBL) databases.

2.3. Data analysis and Gene Ontology

The spectral count approach was used to calculate the protein relative abundance that was expressed as normalized spectral abundance factor (NSAF) according to Old et al. [5]. All iden-

tified proteins were loaded in the UniProt Knowledgebase (UniProtKB) database to obtain the Gene Ontology (GO) annotations associated to biological process, molecular function, and cellular component.

Acknowledgments

This work was funded by Sardegna Ricerche with grant Art. 26 2013, and with funds from the University of Milan.

Conflict of 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.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.dib.2020.105665.

References

[1] J.A. Vizcaíno, A. Csordas, N. Del-Toro, J.A. Dianes, J. Griss, I. Lavidas, G. Mayer, Y. Perez-Riverol, F. Reisinger, T. Ternent, Q.W. Xu, R. Wang, H. Hermjakob, pdate of the PRIDE database and its related tools, Nucleic Acids Res 44 (2016 u2016) D447–D456 https://doi.org/, doi:10.1093/nar/gkv1145.

[2] S. Pisanu, C. Cacciotto, D. Pagnozzi, S. Uzzau, C. Pollera, M. Penati, V. Bronzo, M.F. Addis, Impact of Staphylococcus aureus infection on the late lactation goat milk proteome: new perspectives for monitoring and understanding mastitis in dairy goats, J. Proteomics. (n.d.).

[3] V. Tedde, V. Bronzo, G.M.G. Puggioni, C. Pollera, A. Casula, G. Curone, P. Moroni, S. Uzzau, M.F. Addis, Milk cathelicidin and somatic cell counts in dairy goats along the course of lactation, J, Dairy Res 86 (2019) 217–221.

[4] S. Pisanu, C. Cacciotto, D. Pagnozzi, S. Uzzau, P. Ciaramella, J. Guccione, M. Penati, C. Pollera, P. Moroni, V. Bronzo, M.F. Addis, Proteomic changes in the milk of water buffaloes (Bubalus bubalis) with subclinical mastitis due to intramammary infection by Staphylococcus aureus and by non-aureus staphylococci, Sci. Rep. (2019) in press.

[5] W.M. Old, K. Meyer-Arendt, L. Aveline-Wolf, K.G. Pierce, A. Mendoza, J.R. Sevinsky, K.A. Resing, N.G. Ahn, Comparison of label-free methods for quantifying human proteins by shotgun proteomics, Mol. Cell. Proteomics. 4 (2005) 1487–1502 https://doi.org/, doi:10.1074/mcp.M500084-MCP200.