

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

2 Salvatore Pisanu<sup>1</sup>, Carla Cacciotto<sup>1</sup>, Daniela Pagnozzi<sup>1</sup>, Sergio Uzzau<sup>1,2,3</sup>, Claudia Pollera<sup>4</sup>, Martina  
3 Penati<sup>4</sup>, Valerio Bronzo<sup>4</sup>, Maria Filippa Addis<sup>1,3,4\*</sup>

4

5 1 Porto Conte Ricerche, Alghero, Italy

6 2 Dipartimento di Scienze Biomediche, Università degli Studi di Sassari, Sassari, Italy

7 3 Mediterranean Center for Disease Control, Sassari, Italy

8 4 Dipartimento di Medicina Veterinaria, Università degli Studi di Milano, Milan, Italy

9

10 **Corresponding author:** Maria Filippa Addis, Department of Veterinary Medicine, University of  
11 Milan, Via dell'Università 6, 26900 Lodi (LO). Tel. ++39 02 503 34515; E-mail:  
12 [filippa.addis@unimi.it](mailto:filippa.addis@unimi.it).

13 **Abstract**

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

27

28 **Keywords**

29 Goat milk, Proteomics; FASP, Mass spectrometry, NSAF

30

32 **Specifications Table**

<b>Subject</b>	Biology
<b>Specific subject area</b>	Proteomics
<b>Type of data</b>	Tables Files .raw Files .msf
<b>How data was acquired</b>	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).
<b>Data format</b>	RAW data (Thermo proprietary), msf files (Proteome Discoverer) and supplementary Excel files (1-3) with data analysis output.
<b>Parameters for data collection</b>	Half-udder goat milk collected along the course of an entire lactation (middle and late lactation) from Alpine goat half-udders.
<b>Description of data collection</b>	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.
<b>Data source location</b>	Tramariglio, Alghero (Sassari), Italy
<b>Data accessibility</b>	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): <a href="http://proteomecentral.proteomexchange.org/cgi/GetDataset?ID=PX017243">http://proteomecentral.proteomexchange.org/cgi/GetDataset?ID=PX017243</a>

<b>Related research article</b>	S. Pisanu, C. Cacciotto, P. Pagnozzi, S. Uzzau, C. Pollera, M. Penati, V. Bronzo, M.F. Addis. Impact of <i>Staphylococcus aureus</i> infection on the late lactation goat milk proteome: new perspectives for monitoring and understanding mastitis in dairy goats. Journal of Proteomics, submitted [2].
---------------------------------	--

33

34

35 **Value of the Data**

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

45

46

47 **1. Data Description**

48 The goat milk proteome was analyzed by a shotgun proteomics approach. Stage of lactation,

49 somatic cell count (SCC) and bacteriological culture results were used to categorize milk samples

50 and assemble the experimental groups, as reported in Table 1. The three proteomic profiles of

51 goat milk are reported in Supplementary Tables 1, 2 and 3.

52 A total of 173, 290 and 446 non-redundant proteins were identified in the mid-lactation,

53 uninfected (MLU, Supplementary Table 1), late lactation, uninfected (LHU, Supplementary Table 2)

54 and late lactation, *Staphylococcus aureus*-infected (LHS, Supplementary Table 3) half-udder goat

55 milk, respectively.

56

57

58

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

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

61 <sup>a</sup>SCC: Somatic Cell Count. <sup>b</sup>Mid-lactation 30-60 DIM; Late lactation > 250 DIM. <sup>c</sup>MLU: middle, low,  
 62 uninfected. <sup>d</sup>LHU: late, high, uninfected. <sup>e</sup>LHS: late, high, *S. aureus*.

63

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

67

68

## 69 **2. Experimental Design, Materials, and Methods**

70

### 71 *2.1. Goat milk samples*

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

76

### 77 *2.2. LC-MS/MS analysis*

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

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

86

87 *2.3. Data analysis and Gene Ontology*

88 The spectral count approach was used to calculate the protein relative abundance that was  
89 expressed as normalized spectral abundance factor (NSAF) according to Old et al. [5]. All identified  
90 proteins were loaded in the UniProt Knowledgebase (UniProtKB) database to obtain the Gene  
91 Ontology (GO) annotations associated to biological process, molecular function, and cellular  
92 component.

93

94 **Acknowledgments**

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

97

98 **Conflict of interest**

99 The authors declare that they have no known competing financial interests or personal  
100 relationships which have, or could be perceived to have, influenced the work reported in this  
101 article.

102

103 **References**

- 104 [1] J.A. Vizcaíno, A. Csordas, N. Del-Toro, J.A. Dienes, J. Griss, I. Lavidas, G. Mayer, Y. Perez-  
105 Riverol, F. Reisinger, T. Ternent, Q.W. Xu, R. Wang, H. Hermjakob, 2016 update of the PRIDE  
106 database and its related tools, *Nucleic Acids Res.* 44 (2016) D447–D456.  
107 <https://doi.org/10.1093/nar/gkv1145>.
- 108 [2] S. Pisanu, C. Cacciotto, D. Pagnozzi, S. Uzzau, C. Pollera, M. Penati, V. Bronzo, M.F. Addis,  
109 Impact of *Staphylococcus aureus* infection on the late lactation goat milk proteome: new  
110 perspectives for monitoring and understanding mastitis in dairy goats, *J. Proteomics.* (n.d.).
- 111 [3] V. Tedde, V. Bronzo, G.M.G. Puggioni, C. Pollera, A. Casula, G. Curone, P. Moroni, S. Uzzau,  
112 M.F. Addis, Milk cathelicidin and somatic cell counts in dairy goats along the course of  
113 lactation, *J. Dairy Res.* 86 (2019) 217–221.
- 114 [4] S. Pisanu, C. Cacciotto, D. Pagnozzi, S. Uzzau, P. Ciaramella, J. Guccione, M. Penati, C.  
115 Pollera, P. Moroni, V. Bronzo, M.F. Addis, Proteomic changes in the milk of water buffaloes

116 (Bubalus bubalis) with subclinical mastitis due to intramammary infection by  
117 Staphylococcus aureus and by non-aureus staphylococci, Sci. Rep. (2019) in press.

118 [5] W.M. Old, K. Meyer-Arendt, L. Aveline-Wolf, K.G. Pierce, A. Mendoza, J.R. Sevinsky, K.A.  
119 Resing, N.G. Ahn, Comparison of label-free methods for quantifying human proteins by  
120 shotgun proteomics., Mol. Cell. Proteomics. 4 (2005) 1487–502.  
121 <https://doi.org/10.1074/mcp.M500084-MCP200>.

122