UNIVERSITÀ
DEGLI STUDI
DI PADOVA

Head Office: Università degli Studi di Padova

Department of Agronomy, Food, Natural resources, Animal and Environment

Ph.D. COURSE IN: ANIMAL AND FOOD SCIENCE

SERIES XXXVI

PHENOTYPIC AND MOLECULAR ASPECTS OF SUBCLINICAL MASTITIS IN DAIRY CATTLE

Coordinator: Prof.ssa Angela Trocino

Supervisor: Prof. Alessio Cecchinato

Co-Supervisor: Dr. Sara Pegolo

Ph.D. student: Vittoria Bisutti

PhD Course in Animal and Food Science, series XXXVI

University of Padova, Italy

A thesis for applying for the degree of Doctor of Philosophy, submitted in fulfilment of the requirements of University of Padova, by the authority of the Rector Magnificus, Prof. Daniela Mapelli.

PHENOTYPIC AND MOLECULAR ASPECTS OF SUBCLINICAL MASTITIS IN DAIRY CATTLE

Vittoria Bisutti

Thesis Committee

Supervisor

Prof. Alessio Cecchinato, Department of Agronomy, Food, Natural Resources, Animals and Environment (DAFNAE), University of Padova, Viale dell'Università 16, 35020, Legnaro (PD), Italy

Co-supervisor

Dr. Sara Pegolo, Department of Agronomy, Food, Natural Resources, Animals and Environment (DAFNAE), University of Padova, Viale dell'Università 16, 35020, Legnaro (PD), Italy

Legnaro (PD), Italy

September 2023

Declaration

I declare that the present thesis has not been previously submitted as an exercise for a degree at the University of Padova (Italy) or any other University.

I further declare this work embodied to be mine.

CONTENTS

ABSTRACT

Taking into consideration the implication that subclinical mastitis has for the dairy industry and the welfare of the animals, the focus of this PhD thesis was to better characterize the phenotypic and molecular aspects associated with udder inflammation and subclinical IMI in dairy cows, with the specific focus on the effects of subclinical mastitis on some milk important traits (i.e. protein fractions) and also for better explaining the biological complexity behind the development of mastitis.

To achieve this goal we have analyzed the topic from different angles and split it in three contributions (chapters). The dataset used consisted in individual milk information belonging to two different projects (BENELAT and LATSAN) which had the overall aim of developing short- and long- term interventions for the improvement of animal welfare, efficiency, and quality of dairy cattle production. For the first contribution (**Chapter I**), we collected pooled individual milk samples from a population of 1,482 clinically healthy cows reared in six different commercial farms with the aim of evaluating the association existing between somatic cell count (SCC) and differential somatic cell count (DSCC) on the detailed milk protein profile. The protein fractions were determined with a validated Reversed-Phase High Performance Liquid Cromatography (RP-HPLC) which allowed the identification of four caseins (i.e., α-$_{S1}$, α-$_{S2}$, β- and κ-CN) and three whey proteins (β- lactoglobulin, α-lactalbumin and lactoferrin) which were later expressed both quantitatively (g/L) and qualitatively (percentage of total milk nitrogen).

In the second contribution (**Chapter II**) we aimed at investigating the associations between naturally occurring subclinical intramammary infection (IMI) caused by different etiological agents (i.e., *Staphylococcus aureus*, *Streptococcus agalactiae*, *Streptococcus uberis* and *Prototheca* spp) in combination with SCC on the detailed protein profile at the individual mammary quarter level. For this study, an initial bacteriological screening (time 0, T0) on 450 cows was carried out to identify healthy animals and the ones having subclinical IMI. We selected 78 infected animals which were followed up at the quarter level at two different time points (time 1 and 2, T1 and T2) after two and

four weeks from T0. Finally, to get more insights on the molecular mechanism underling the development of mastitis, as well as to better understand the relationship between host-pathogen interaction, we used the RNA-Sequencing (RNA-Seq) technology with the final goal of evaluating the milk somatic cell (SC) transcriptomic signatures upon natural infection from *S. agalactiae* and *Prototheca* spp. In this third contribution (**Chapter III**) we also included the integration of transcriptomic and phenotypic information to better explain the complexity underlying the molecular mechanism of mastitis and to identify putative informative hub variables for early mastitis detection and prediction.

Results from the first contribution firstly confirmed the detrimental effects of the increase in SCC on the milk protein profile, and especially on the casein fractions. Then, when we focused on the contribution of the leucocyte's populations extracted from the DSCC parameter, which were expressed as combined count of polymorphonuclear cells and lymphocytes (PMN-LYM) and macrophages (MAC) count, we observed that the most detrimental effects towards the protein profile, and once again, especially on the casein fractions, was carried out by the MAC proportion. This finding is probably ascribable to the proteolytic activity derived from macrophages, as they were present in higher number given the non-pathological status of the animals.

The second contribution allowed us to dig deeper on the behavior of milk protein fractions in animals having subclinical IMI by exploring in an extensive way the pattern of alteration at the mammary gland at the quarter level. Indeed, we observed that both subclinical IMI, and especially increased SCC significantly reduced the proportion of two of the most abundant casein fractions (i.e., β-CN and α_{S1}-CN), which is ascribed to the increased activity of both milk endogenous and microbial proteases. In terms of whey proteins, the most significant alteration involved lactoferrin, which being a glycoprotein with direct and undirect antimicrobial activity, increased both with IMI and SCC. Overall, we observed that the inflammation status driven by the increase in SCC, was associated with the most significant alteration in the protein profile. This suggests that the increase in endogenous

proteolytic enzymes related to the onset of inflammation, and not the infection, might be the pivotal aspect driving the alteration of the milk protein profile.

Shifting from the phenotypic to the molecular level in the third contribution, allowed us to shed some light on the molecular mechanisms behind subclinical IMI and on the immune reaction of the animal in response to the infection from *S. agalactiae* and *Prototheca* spp. Results from the RNA-Seq allowed us to observe the changes in the gene expression from healthy animals to the ones infected by *S. agalactiae* or *Prototheca*. The evidence collected from the pathway analyses highlighted that, even though some differences in terms of immune-related pathways and gene expression between the two infections have been observed, there was a strong "core" immune response that was commonly shared by the two pathogens. Then, integrating the transcriptomic and phenotypic information highlighted a strong correlation between the transcriptome and the immune cell populations, and udder health traits (i.e., SCC, DSCC, lactose and conductivity). Also, we identified some putative hub genes that could be helpful for the animal's response to subclinical mastitis. In conclusion, the hub variables identified with the integrative approach, once validated in a larger population, could be included in screening and diagnostic tools for subclinical IMI detection at the herd level.

RIASSUNTO

Considerando l'impatto che la mastite subclinica ha sul settore lattiero-caseario e sul benessere degli animali, l'obiettivo principale di questa tesi di dottorato è stato quello di caratterizzare in modo più approfondito da un punto di vista sia fenotipico che molecolare, l'infezione intramammaria (IMI) subclinica e all'infiammazione della ghiandola mammaria nelle bovine da latte, prestando inoltre particolare attenzione sui loro effetti su alcune importanti componenti del latte (le frazioni proteiche).

Per raggiungere questo obiettivo abbiamo analizzato l'argomento da diverse prospettive che hanno portato alla realizzazione di tre contributi scientifici. Il dataset utilizzato consisteva in informazioni relative a campioni di latte individuali raccolti durante due diversi progetti (BENELAT e LATSAN), i quali avevano l'obiettivo generale di sviluppare interventi a breve e lungo termine per migliorare il benessere degli animali, l'efficienza e la qualità della produzione di latte bovino. Nel primo contributo (**Capitolo I**), abbiamo prelevato dei campioni di latte individuali da 1,482 bovine di razza Frisona, clinicamente sane, provenienti da sei diversi allevamenti, con l'obiettivo di andare a studiare in modo approfondito l'associazione esistente tra la conta delle cellule somatiche (SCC) e la conta differenziale delle cellule somatiche (DSCC) sul profilo proteico del latte. Esso è stato misurato tramite l'utilizzo di una metodica validata di Cromatografia Liquida ad Alta Prestazione in Fase Inversa (RP-HPLC) che ha permesso l'identificazione di quattro caseine (α_{S1}-, α_{S2}-, β- e κ-) e tre sieroproteine (β-lattoglobulina, α-lattoalbumina e lattoferrina), che sono state successivamente espresse sia in modo quantitativo (g/L) che qualitativo (% di azoto presente nel latte, %N). Con il secondo contributo (**Capitolo II**) siamo andati ad investigare in modo mirato l'associazione tra infezione intramammaria (IMI) subclinica causata da diversi agenti eziologici (*Staphylococcus aureus*, *Streptococcus agalactiae*, *Streptococcus uberis* e *Prototheca* spp), lo stato di infiammazione della ghiandola mammaria (SCC), e la loro interazione, nei confronti del profilo proteico del latte, a livello di singolo quarto mammario. Per raggiungere questo obiettivo è stato effettuato un primo screening batteriologico (tempo 0, T0) su 450 bovine per identificare gli animali sani e quelli affetti

da IMI subclinica. Sulla base di questo screening, abbiamo selezionato 78 animali infetti che sono stati seguiti a livello del quarto mammario in due tempi diversi (T1 e T2) dopo rispettivamente due e quattro settimane dal T0.

Infine, nel terzo studio (**Capitolo III**), per ottenere una migliore comprensione dei meccanismi molecolari alla base dello sviluppo della mastite subclinica, nonché sulla interazione tra animale e agente-patogeno, attraverso l'utilizzo della tecnologia di RNA-Sequencing (RNA-Seq) abbiamo valutato il trascrittoma delle cellule somatiche del latte (SC) in animali sani e aventi IMI subclinica spontanea causata da due diversi microorganismi: *S. agalactiae* e *Prototheca* spp. In aggiunta, siamo andati ad integrare le informazioni trascrittomiche con quelle fenotipiche a nostra disposizione (popolazioni leucocitarie ottenute tramite cito fluorimetria, caratteri di composizione del latte e salute della ghiandola mammaria) non solo per spiegare meglio la complessità dei meccanismi molecolari che stanno alla base dello sviluppo della mastite ma anche per identificare potenziali variabili informative (fenotipiche e molecolari) potenzialmente informative per l'identificazione e diagnosi precoce di questa patologia.

I risultati del primo studio hanno in primis confermato gli effetti negativi dell'incremento della SCC sul profilo proteico del latte, e in particolare sulle frazioni caseiniche. Focalizzandoci in seguito sul contributo delle popolazioni leucocitarie estratte dal parametro DSCC (che fornisce una proporzione combinata di linfociti [LYM] e neutrofili polimorfonucleati [PMN] , da cui è possibile calcolare come complemento a 100 la percentuale di macrofagi [MAC]), abbiamo osservato come i MAC fossero la popolazione leucocitaria che influenzava in modo peggiore le frazioni proteiche, e le caseine in particolare. Questa associazione sfavorevole dei MAC nei confronti delle caseine sembra essere ascrivibile alla loro azione proteolitica endogena. Essendo che gli animali coinvolti si trovavano in buone condizioni di salute si è potuto assumere che la popolazione immunitaria prevalente fosse appunto quella dei macrofagi.

Nel secondo contributo, focalizzarci a livello del singolo quarto mammario ci ha permesso di spiegare meglio il ruolo degli attori coinvolti nell'alterazione delle frazioni proteiche del latte. Infatti,

abbiamo osservato che sia l'IMI subclinica che l'incremento di SCC riducevano significativamente la proporzione delle due frazioni caseiniche più abbondanti (β-CN e α_{S1}-CN), attribuibile all'incremento dell'attività di proteasi endogene e microbiche del latte. Per quanto riguarda le proteine del siero, l'alterazione più significativa ha riguardato la lattoferrina, una glicoproteina con attività antimicrobica diretta e indiretta, che è aumentata sia in presenza di IMI che con l'incremento di SCC. Nel complesso, abbiamo però osservato che l'incremento di SCC è stato il fattore associato con le alterazioni più significative nel profilo proteico del latte. Questo suggerisce che l'incremento degli enzimi proteolitici endogeni legati all'insorgenza dell'infiammazione, e non l'infezione (cioè, lo stato batteriologico), potrebbe essere l'aspetto chiave che guida l'alterazione del profilo proteico del latte.

Nel terzo contributo, grazie all'utilizzo dell'RNA-Seq, è stato possibile investigare più a fondo i meccanismi molecolari legati alla mastite subclinica. I risultati ottenuti dall'analisi trascrittomica ci hanno permesso di osservare dei cambiamenti a livello dell'espressione genica tra animali sani a quelli infetti da *S. agalactiae* o *Prototheca,* legati per lo più a *pathways* coinvolti nella risposta immunitaria e nel metabolismo energetico. Nonostante alcune differenze riscontrate tra i due microorganismi, ciò che è emerso è che la presenza di una risposta "immuno-trascrittomica" condivisa dai due patogeni, che potrebbe risultare utile in futuro per identificare dei biomarcatori per l'identificazione di IMI, a prescindere dal patogeno. Infine, integrando le informazioni trascrittomiche e fenotipiche, abbiamo osservato una forte correlazione tra il trascrittoma e le popolazioni leucocitarie del latte, e i caratteri di salute della ghiandola mammaria (SCC, DSCC, lattosio e conducibilità). Questo approccio integrativo ha permesso inoltre di identificare alcune "*hub variables*" (sia molecolari che fenotipiche) che, una volta validate in una popolazione più ampia, potrebbero essere incluse come strumenti di screening e diagnostici per la rilevazione delle IMI subcliniche.

GENERAL INTRODUCTION

When it comes to the dairy production, Italy represents one of the most impactful countries in Europe both concerning the amount of milk produced and the quality of the derived dairy products. In fact, among the EU countries, Italy contributes for the 13% to the production of cheese, placing as the third cheese producer, after Germany (23%) and France (18%) (EUROSTAT 2021). Not only Italy is one of the top cheese producers in terms of quantity, but together with France, is responsible for the greatest amount of high-quality cheese labelled with the Protected Designation of Origin (PDO) certification. Indeed, in Italy more than 75% of the produced milk is destined to cheesemaking (CLAL, 2022), of which 50% is specifically directed to the production of PDO cheese like Grana Padano and Parmigiano Reggiano (ISMEA, 2023).

In this economic context, milk proteins are important traits related to milk quality and technological characteristics. Bovine milk has an average protein content of 35 g/L and from a nutritional point of view, it is an important source of high-quality proteins, as not only contains all of the nine essential amino acids (Davoodi et al., 2016) but also protein-derived bioactive compounds with several physiological activities, such as immunomodulatory and gastrointestinal properties (Korhonen and Pihlanto, 2006). Milk proteins are mainly composed by caseins (CNs), which account for almost 80% of the protein content, and whey proteins, which represent the remaining 20%. Milk caseins, which have the characteristic of precipitating at pH 4.6 at room temperature, are represented by α-$_{S1}$-, α-$_{S2}$-, β- and κ-CN. Whey proteins, on the other hand are found in solution in milk and are mainly composed by β-lactoglobulin (β-LG) and α-lactalbumin (α-LA), followed by lactoferrin (LF), and a small proportion of serum albumin and immunoglobulins. Several studies have evidenced that the quantity and proportion of milk protein fractions, especially caseins, can significantly influence the variation of coagulation and cheesemaking ability of milk (Caroli et al., 2009; Bittante et al., 2012), resulting in the assessment of the genetic variability of these traits for selective breeding (Heck et al., 2009).

Several animal individual factors can affect the protein profile in bovine milk such as parity, stage of lactation, season and feeding regime (Bobe et al., 1999). Together with these factors, also the health condition of the mammary gland concurs in influencing the quantity of milk protein fractions. In fact, several studies have reported that with the onset of inflammatory conditions (i.e., mastitis), milk undergoes through compositional changes which negatively affect its composition and quality (Bobbo et al., 2016; Pegolo et al., 2021). Concerning the protein fractions, the inflammation status in the udder results in an increased proteolytic activity derived by the release of several types of proteases which have a strong impact on milk proteins, and especially on caseins (Caggiano et al., 2019).

In this context, it is fundamental to maintain and improve herd management strategies focused on the care of udder health to keep a sustainable and efficient milk production. In fact, despite the continuous improvement in hygiene and management strategies, mastitis continues to be the most impactful disease which firstly hinders the welfare of dairy cows, but also causes significant economic losses, due to veterinary expenses, premature culling and antimicrobial treatments (Neculai-Valeanu and Ariton, 2022).

Bovine mastitis is a complex multifactorial disease of the mammary gland. Even though it can be caused by physical trauma, toxic chemicals (Reyher and Dohoo, 2011) and even poor nutrition of the animals (Ajose et al., 2022), mastitis typically occurs consequent to the intramammary infection (IMI) from a wide range of microorganisms. Bovine mastitis can occur in clinical or subclinical forms, depending on the presence (or absence) of observable manifestations of clinical signs. Clinical mastitis is characterized by visible signs of inflammation, such as red and swollen udder, and physical alteration of milk and of its components. In fact, it is common to observe changes in color, consistency and presence of clots and blood (Roberson, 2012). On the contrary, subclinical mastitis shows no signs of alteration and usually is associated only with an increase in the milk somatic cells. Being unnoticed, it represents an important source of infection for the animals within the farms leading to a

significant decrease in milk production (Ruegg, 2017) but also to a possible decrease in the bulk-tank milk quality, as this milk is mixed with the ones from the healthy cows (Haxhiaj et al., 2022). For these reasons, and since subclinical is estimated to be 15-40 times more frequent than its clinical counterpart (Martin et al., 2018), it is probably the most insidious and challenging form.

It has been reported that the most causative agents of IMI, accounting for over 80% of the infections, are *Escherichia coli* (*E. coli*), *Staphylococcus aureus* (*S. aureus*), *Streptococcus agalactiae* (*S. agalactiae*), *Streptococcus uberis* (*S. uberis*) and *Streptococcus dysgalactiae* (*S. dysgalactiae*) (Ranjan et al., 2006). *Escherichia coli* is a Gram-negative bacterium which mainly causes clinical mastitis (Ju et al., 2020), while *S. aureus*, *Strep. agalactiae* and *Strep. uberis* are all Gram-positive bacteria which are mostly responsible for inducing subclinical and/or chronic forms of the disease (Lewandowska-Sabat et al., 201; Chen et al., 2021). Aside from these microorganisms, in the last years a change in the pattern distribution of the pathogens has been observed, and other microorganisms which originally were labelled as uncommon mastitis agents, are now being considered as emerging pathogens, and therefore requiring additional attention. That is the case of *Prototheca spp*. *Prototheca* is a colorless non-photosynthetic microalgae that in the last years has rapidly emerged as a non-negligible causative agents of mastitis (Shave et al., 2021): for example, in Italy the prevalence of this pathogen is estimated to be around 11.2 % (Zecconi et al., 2020). Little is known about the mechanism of action by which this alga can infect hosts and cause disease. Moreover, up to date no effective treatments have been found towards this pathogen (Shave et al., 2021), which represents a serious problem for the dairy sector and calls for the need of finding alternative management strategies rather than premature culling.

Udder health is of pivotal importance for efficient and sustainable dairy production, since mastitis, as above stated, is not just a medical condition but it impacts milk quality, cattle performance, and farm antimicrobial use (Neculai-Valeanu and Ariton, 2022). In this context, milk

somatic cells can provide useful information for the assessment of the mammary gland health and therefore for the screening and detection of mastitis.

The inflammatory response that can be observed in a cow's udder can be a sign of the presence of mastitis. However, the identification of the causative agents causing the disease is the ultimate way of confirming IMI and, up to date, the gold standard analysis for achieving the confirmation of the infection status has remained the microbiological examination (NMC, 2017). However, its systemic application at a farm level is not always feasible as it is a time consuming and expensive technique, which requires also expert personnel. In this context, the somatic cell count (SCC) in milk provides an indication of the inflammatory response in the mammary gland and for this reason has become the most widespread proxy for identifying mastitis at the quarter, cow, herd and population level (Schukken et al., 2003). Milk somatic cells (SCs) are mainly composed by cells derived from the immune system (i.e., lymphocytes, polymorphonuclear cells and macrophages) and by a small proportion of epithelial cells derived from the shedding of the mammary gland. Each type of immune cells contributes in a different way to the inflammatory response and therefore their proportion may be subjected to variation according to the stage of infection. In healthy conditions the level of SCs is normally low, and usually lymphocytes and macrophages prevail, whereas polymorphonuclear cells (PMN) are present only in small proportion (Dosogne et al., 2003; Schwarz et al., 2011). In the presence of infection and/or inflammatory condition somatic cells tend to increase not only to help fighting the infection but also for repairing tissue damage (Alhussien and Dang, 2018). In fact, once the mammary gland are infected, a surge in PMN can be observed triggered by the resident cells while few days later PMN start to decrease leaving space to macrophages for the elimination of bacteria and debris (Rivas et al., 2001).

Conventionally, the cut-off point to distinguish healthy cows from the ones having subclinical mastitis, has been set to 200,000 cells/mL (Ruegg, 2011), even if scientific evidence has reported that SCC might have some limits in giving information of the cows' health condition (Bobbo et al., 2017).

For example, some studies observed animals with milk with low SCC but positive at the microbial examination and, conversely, animals with cultured-negative milk samples but with high SCC (Bobbo et al., 2017; Pegolo et al., 2022).

The somatic cell count has proved to be a robust indicator of the intensity of the mammary gland inflammation. However, SCC does not provide specific information on the distribution of each immune cell population and therefore cannot provide insights on the stage of inflammation or the progression of the disease, which could be pivotal for improving knowledge on this disease. Up until recently, the major obstacle for the large-scale evaluation of the main leucocytes populations was the unavailability of high throughput technologies, as well as the cost the analysis. Indeed microscopy, which is still the gold standard method for the evaluation of the immune cells distribution, it is extremely time consuming and it has poor repeatability (Zecconi et al., 2019). Flow cytometry, on the other hand, is more efficient, but very expensive and allows only the analysis of few sample per day, being therefore not suitable for large-scale analysis outside of the research field (Koess and Hamann, 2008). A step forward in this direction was recently made with the introduction of a new high-throughput flow cytometry-based analyzer which led to the development of a new parameter: the differential somatic cell count (DSCC). This trait represents the sum of PMN combined with lymphocytes as a percentage of total SCC (Damm et al., 2017). Moreover, proportion of macrophages can be calculated as 100 – DSCC. The use of DSCC, in combination with SCC, could provide large-scale additional insights on the mammary gland health status and informative screening tools for the detection of mastitis.

Aside from SCC and DSCC, there are other traits that could be helpful for the detection of suspected udder inflammation, such as milk lactose and electric conductivity. In fact, during the inflammation, the altered permeability of the mammary gland-blood barrier results in an increase in lactose and ions like in potassium, magnesium, sodium and calcium, which can together increase the electrical conductivity in milk (Viguier et al., 2009). Although lactose and conductivity might be

informative indicators, on their own they might be not reliable enough for conclusive diagnosis (Norberg, 2005), but as proposed by Ebrahimie et al. (2018), if considered together with SCC and DSCC, they could be systematically implemented in health prevention programs as biomarkers for the detection of subclinical mastitis.

In order to improve practical diagnostic and therapeutic tools for battling subclinical (but also clinical) mastitis it is fundamental also to further study the complex molecular processed that are behind this disease. Thanks to the advances achieved in high-throughput technologies and bioinformatic analyses, in the last decade we have observed a significant increase in the use of the so called "omic" technologies (i.e. genomic, transcriptomic, epigenomic and metabolomic) in the animal science field (Naserkheil et al., 2022), with the final goal of identifying putative reliable biomarkers for large-scale disease detection. Among these omic technologies, RNA sequencing (RNA-Seq) represents an useful tool for studying the complexity behind host-pathogen interaction, as well as for better elucidating the molecular mechanisms behind the disease (Wang et al., 2020a). Up to date several studies have investigated the changes in gene expression in the mammary gland after intramammary infection from different pathogens (Jensen et al., 2013; Wang et al., 2020b; Lewandowska-Sabat et al., 2019), however most of the studies so far focused on experimentally induced models of clinical mastitis which do not provide the exact picture of what happens in case of natural and subclinical conditions. Moreover, since there is still a difficulty in identifying reliable and reproducible molecular biomarkers for mastitis detection and prediction, the integration of multi-layer information (i.e., the system biology approach) might represent the ultimate step to strengthen the information on this complex molecular system, but also to create more reliable large scale prediction models.

REFERENCES

Ajose, D.J., B.O. Oluwarinde, T.O. Abolarinwa, J. Fri, K.P. Montso, O.E. Fayemi, A.O. Aremu, and C.N. Ateba. 2022. Combating Bovine Mastitis in the Dairy Sector in an Era of Antimicrobial Resistance: Ethno-veterinary Medicinal Option as a Viable Alternative Approach. Front. Vet. Sci. 9:1–15. doi:10.3389/fvets.2022.800322.

Alhussien, M.N., and A.K. Dang. 2018. Milk somatic cells, factors influencing their release, future prospects, and practical utility in dairy animals: An overview. Vet. World 11:562–577. doi:10.14202/vetworld.2018.562-577.

Bittante, G., M. Penasa, and A. Cecchinato. 2012. Invited review: Genetics and modeling of milk coagulation properties. J. Dairy Sci. 95:6843–70. doi:10.3168/jds.2012-5507.

Bobbo, T., C. Cipolat-Gotet, G. Bittante, and A. Cecchinato. 2016. The nonlinear effect of somatic cell count on milk composition, coagulation properties, curd firmness modeling, cheese yield, and curd nutrient recovery. J. Dairy Sci. 99:5104–5119. doi:10.3168/jds.2015-10512.

Bobbo, T., P.L. Ruegg, G. Stocco, E. Fiore, M. Gianesella, M. Morgante, D. Pasotto, G. Bittante, and A. Cecchinato. 2017. Associations between pathogen-specific cases of subclinical mastitis and milk yield, quality, protein composition, and cheese-making traits in dairy cows. J. Dairy Sci. 100:4868–4883. doi:10.3168/jds.2016-12353.

Bobe, G., D.C. Beitz, A.E. Freeman, and G.L. Lindberg. 1999. Effect of milk protein genotypes on milk protein composition and its genetic parameter estimates. J. Dairy Sci. 82:2797–2804. doi:10.3168/jds.S0022-0302(99)75537-2.

Caggiano, N., A. Lorenzo Smirnoff, J.M. Bottini, and E.A. De Simone. 2019. Protease activity and protein profile in milk from healthy dairy cows and cows with different types of mastitis. Int. Dairy J. 89:1–5. doi:10.1016/j.idairyj.2018.07.020.

Caroli, A.M., S. Chessa, and G.J. Erhardt. 2009. Invited review: Milk protein polymorphisms in cattle: Effect on animal breeding and human nutrition. J. Dairy Sci. 92:5335–5352. doi:10.3168/jds.2009-2461.

Chen, Y., H. Jing, M. Chen, W. Liang, J. Yang, G. Deng, and M. Guo. 2021. Transcriptional Profiling of Exosomes Derived from Staphylococcus aureus -Infected Bovine Mammary Epithelial Cell Line MAC-T by RNA-Seq Analysis. Oxid. Med. Cell. Longev. 1–18.

CLAL. 2022. Italy: Dairy sector. Accessed Sep. 25, 2023. https://www.clal.it/en/?section=quadro_italia.

Damm, M., C. Holm, M. Blaabjerg, M.N. Bro, and D. Schwarz. 2017. Differential somatic cell count—A novel method for routine mastitis screening in the frame of Dairy Herd Improvement testing programs. J. Dairy Sci. 100:4926–4940. doi:10.3168/jds.2016-12409.

Davoodi, S.H., R. Shahbazi, S. Esmaeili, S. Sohrabvandi, A.M. Mortazavian, S. Jazayeri, and A. Taslimi. 2016. Health-related aspects of milk proteins. Iran. J. Pharm. Res. 15:573–591.

Dosogne, H., F. Vangroenweghe, J. Mehrzad, A.M. Massart-Leën, and C. Burvenich. 2003. Differential leukocyte count method for bovine low somatic cell count milk. J. Dairy Sci. 86:828–834. doi:10.3168/jds.S0022-0302(03)73665-0.

Ebrahimie, E., F. Ebrahimi, M. Ebrahimi, S. Tomlinson, and K.R. Petrovski. 2018. A large-scale study of indicators of sub-clinical mastitis in dairy cattle by attribute weighting analysis of milk composition features : highlighting the predictive power of lactose and electrical conductivity. J. Dairy Res. 85:193–200. doi:10.1017/S0022029918000249.

EUROSTAT. 2021. Milk and milk products statistics. Accessed Sep. 25, 2023 https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Milk_and_milk_product_statistics

Haxhiaj, K., D.S. Wishart, and B.N. Ametaj. 2022. Mastitis: What It Is, Current Diagnostics, and

the Potential of Metabolomics to Identify New Predictive Biomarkers. Dairy 3:722–746. doi:10.3390/dairy3040050.

Heck, J.M.L.;, A.; Schennink, H.J.F.. van Valenberg, H.; Bovenhuis, M.H.P.W.. Visker, J.A.. ; van Arendok, and A.C.. van Hooijdonk. 2009. Effects of milk protein variants on the protein composition of bovine milk. J. Dairy Sci. 92:1192–1202. doi:10.3168/jds.2008-1208.

ISMEA. 2023. Scheda di settore latte e derivati bovini. Rome, Italy. Accessed Sep. 25, 2023. https://www.ismeamercati.it/lattiero-caseari/latte-derivati-bovini

Jensen, K., J. Günther, R. Talbot, W. Petzl, H. Zerbe, H.J. Schuberth, H.M. Seyfert, and E.J. Glass. 2013. Escherichia coli- and Staphylococcus aureus-induced mastitis differentially modulate transcriptional responses in neighbouring uninfected bovine mammary gland quarters. BMC Genomics 14. doi:10.1186/1471-2164-14-36.

Ju, Z., Q. Jiang, J. Wang, X. Wang, C. Yang, Y. Sun, Y. Zhang, C. Wang, Y. Gao, X. Wei, M. Hou, and J. Huang. 2020. Genome-wide methylation and transcriptome of blood neutrophils reveal the roles of DNA methylation in affecting transcription of protein-coding genes and miRNAs in E. coli-infected mastitis cows. BMC Genomics 21:1–14. doi:10.1186/s12864-020-6526-z.

Koess, C., and J. Hamann. 2008. Detection of mastitis in the bovine mammary gland by flow cytometry at early stages. J. Dairy Res. 75:225–232. doi:10.1017/S0022029908003245.

Korhonen, H., and A. Pihlanto. 2006. Bioactive peptides: Production and functionality. Int. Dairy J. 16:945–960. doi:10.1016/j.idairyj.2005.10.012.

Lewandowska-Sabat, A.M., E. Kirsanova, C. Klopp, T.R. Solberg, B. Heringstad, O. Østerås, P. Boysen, and I. Olsaker. 2019. Transcription profiling of monocyte-derived macrophages infected in vitro with two strains of Streptococcus agalactiae Reveals candidate pathways affecting subclinical mastitis in cattle. Front. Genet. 10:1–14. doi:10.3389/fgene.2019.00689.

Martin, P., H.W. Barkema, L.F. Brito, S.G. Narayana, and F. Miglior. 2018. Symposium review: Novel strategies to genetically improve mastitis resistance in dairy cattle. J. Dairy Sci. 101:2724–2736. doi:10.3168/jds.2017-13554.

Naserkheil, M., F. Ghafouri, S. Zakizadeh, N. Pirany, Z. Manzari, S. Ghorbani, M.H. Banabazi, M.R. Bakhtiarizadeh, M.A. Huq, M.N. Park, H.W. Barkema, D. Lee, and K.S. Min. 2022. Multi-Omics Integration and Network Analysis Reveal Potential Hub Genes and Genetic Mechanisms Regulating Bovine Mastitis. Curr. Issues Mol. Biol. 44:309–328. doi:10.3390/cimb44010023.

Neculai-Valeanu, A.S., and A.M. Ariton. 2022. Udder Health Monitoring for Prevention of Bovine Mastitis and Improvement of Milk Quality. Bioengineering 9. doi:10.3390/bioengineering9110608.

N.M.C. 2017. Laboratory Handbook on Bovine Mastitis. third. National Mastitis Council, Inc.

Norberg, E. 2005. Electrical conductivity of milk as a phenotypic and genetic indicator of bovine mastitis: A review. Livest. Prod. Sci. 96:129–139. doi:10.1016/j.livprodsci.2004.12.014.

Pegolo, S., D. Giannuzzi, V. Bisutti, R. Tessari, M.E. Gelain, L. Gallo, S. Schiavon, F. Tagliapietra, E. Trevisi, P. Ajmone Marsan, G. Bittante, and A. Cecchinato. 2021. Associations between differential somatic cell count and milk yield, quality, and technological characteristics in Holstein cows. J. Dairy Sci. 104:4822–4836. doi:10.3168/jds.2020-19084.

Pegolo, S., R. Tessari, V. Bisutti, A. Vanzin, D. Giannuzzi, M. Gianesella, A. Lisuzzo, E. Fiore, A. Barberio, E. Schiavon, E. Trevisi, F. Piccioli Cappelli, L. Gallo, P. Ruegg, R. Negrini, and A. Cecchinato. 2022. Quarter-level analyses of the associations among subclinical intramammary infection and milk quality, udder health, and cheesemaking traits in Holstein cows. J. Dairy Sci. 105. doi:10.3168/jds.2021-21267.

Ranjan, R., D. Swarup, R.C. Patra, and D. Nandi. 2006. Bovine protothecal mastitis: A review.

CAB Rev. Perspect. Agric. Vet. Sci. Nutr. Nat. Resour. 1:1–7. doi:10.1079/PAVSNNR20061017.

Reyher, K.K., and I.R. Dohoo. 2011. Diagnosing intramammary infections: Evaluation of composite milk samples to detect intramammary infections. J. Dairy Sci. 94:3387–3396. doi:10.3168/jds.2010-3907.

Rivas, A.L., F.W. Quimby, J. Blue, and O. Coksaygan. 2001. Longitudinal evaluation of bovine mammary gland health status by somatic cell counting, flow cytometry, and cytology. J. Vet. Diagnostic Investig. 13:399–407. doi:10.1177/104063870101300506.

Roberson, J.R. 2012. Treatment of Clinical Mastitis. Vet. Clin. North Am. - Food Anim. Pract. 28:271–288. doi:10.1016/j.cvfa.2012.03.011.

Ruegg, P.L. 2017. A 100-Year Review: Mastitis detection, management, and prevention. J. Dairy Sci. 100:10381–10397. doi:10.3168/jds.2017-13023.

Schukken, Y.H., D.J. Wilson, F. Welcome, Garrison-Tikofsky, and R.N. Gonzalez. 2003. Monitoring udder health and milk quality using somatic cell counts 35:467–483. doi:10.1051/vetres.

Schwarz, D., U.S. Diesterbeck, S. König, K. Brügemann, K. Schlez, M. Zschöck, W. Wolter, and C.P. Czerny. 2011. Flow cytometric differential cell counts in milk for the evaluation of inflammatory reactions in clinically healthy and subclinically infected bovine mammary glands. J. Dairy Sci. 94:5033–5044. doi:10.3168/jds.2011-4348.

Shave, C.D., L. Millyard, and R.C. May. 2021. Now for something completely different: Prototheca, pathogenic algae. PLoS Pathog. 17:1–7. doi:10.1371/JOURNAL.PPAT.1009362.

Viguier, C., S. Arora, N. Gilmartin, K. Welbeck, and R. O'Kennedy. 2009. Mastitis detection: current trends and future perspectives. Trends Biotechnol. 27:486–493.

doi:10.1016/j.tibtech.2009.05.004.

Wang, D., L. Liu, S.M.A. Augustino, T. Duan, T.J. Hall, D.E. MacHugh, J. Dou, Y. Zhang, Y. Wang, and Y. Yu. 2020a. Identification of novel molecular markers of mastitis caused by Staphylococcus aureus using gene expression profiling in two consecutive generations of Chinese Holstein dairy cattle. J. Anim. Sci. Biotechnol. 11:1–17. doi:10.1186/s40104-020-00494-7.

Wang, X., F. Su, X. Yu, N. Geng, L. Li, R. Wang, M. Zhang, J. Liu, Y. Liu, and B. Han. 2020b. RNA-Seq Whole Transcriptome Analysis of Bovine Mammary Epithelial Cells in Response to Intracellular Staphylococcus aureus. Front. Vet. Sci. 7:1–14. doi:10.3389/fvets.2020.00642.

Zecconi, A., F. dell'Orco, N. Rizzi, D. Vairani, M. Cipolla, P. Pozzi, and L. Zanini. 2020. Cross-sectional study on the prevalence of contagious pathogens in bulk tank milk and their effects on somatic cell counts and milk yield. Ital. J. Anim. Sci. 19:66–74. doi:10.1080/1828051X.2019.1693282.

Zecconi, A., D. Vairani, M. Cipolla, N. Rizzi, and L. Zanini. 2019. Assessment of subclinical mastitis diagnostic accuracy by differential cell count in individual cow milk. Ital. J. Anim. Sci. 18:460–465. doi:10.1080/1828051X.2018.1533391.

AIMS AND OUTLINES OF THE THESIS

Taking into consideration the implication that subclinical mastitis has for the dairy industry and the welfare of the animals, the focus of this PhD thesis was to better evaluate the phenotypic and molecular aspects associated with udder inflammation and subclinical IMI in dairy cows, with the specific focus on the effects of subclinical mastitis on some milk important traits (i.e. protein fractions) and also for better explaining the biological complexity behind the development of mastitis. This topic was dealt from different perspectives and with different approaches, therefore specific aims of the thesis were:

- To study the association between DSCC, firstly combined with SCC and then expressed in a quantitative way as combined count of PMN-LYM and MAC, and the detailed milk protein profile expressed both quantitatively (g/L) and qualitatively (% of milk nitrogen) in 1,482 Holstein cows (**Chapter I**);
- To deeper investigate the association between naturally occurring subclinical intramammary infection (IMI) caused by different pathogens (*Staphylococcus aureus*, *Streptococcus agalactiae*, *Streptococcus uberis* and *Prototheca* spp) in combination with the inflammatory status (i.e. SCC), as well as the specific interaction between IMI and SCC, on the detailed milk protein profile assessed at the quarter level in Holstein cows (**Chapter II**);
- To better assess the molecular mechanisms behind the pathogenesis of naturally occurring subclinical mastitis and the potential difference in the host immune response by: i) evaluating the milk SCs transcriptome in cows with natural and subclinical infection from *S. agalactiae* and *Prototheca* spp., and ii) integrate the transcriptomic and phenotypic information (i.e. milk immune cells population, milk composition traits, SCC and DSCC) to better explain the complexity behind the development of mastitis and the identification of putative informative variables for its detection and prediction (**Chapter III**).

CHAPTER I

DIFFERENTIAL SOMATIC CELL COUNT AND MILK PROTEINS

Impact of somatic cell count combined with differential somatic cell count on milk protein fractions in Holstein cattle

V. Bisutti[1], A. Vanzin[1], A. Toscano[1], S. Pegolo[1*], D. Giannuzzi[1], F. Tagliapietra[1], S. Schiavon[1], L. Gallo[1], E. Trevisi[2], R. Negrini[2], A. Cecchinato[1]

[1]Department of Agronomy, Food, Natural Resources, Animals and Environment (DAFNAE), University of Padova, Viale dell' Università 16, 35020, Legnaro PD, Italy.

[2]Department of Animal Science, Food and Nutrition (DIANA) and Research Center Romeo and Enrica Invernizzi for Sustainable Dairy Production (CREI), Faculty of Agricultural, Food and Environmental Sciences, Università Cattolica del Sacro Cuore, 29122 Piacenza, Italy

Published in: Journal of Dairy Science 105 (2022) 6447-6459

https://doi.org/10.3168/jds.2022-22071

INTERPRETATIVE SUMMARY

Impact of somatic cell count combined with differential somatic cell count on milk protein fractions in Holstein cattle

By Bisutti et al., page 6447. In this study, we present new insights into the role of different immune cells'population (i.e. lymphocytes, polymorphonuclear neutrophilis and macrophages) in explaining the variation in milk protein fractions in clinically healthy Holstein cows. Overall, the macrophage content, expressed quantitatively (i.e. as a percentage of the total somatic cell count) had the most significant effect on the protein profile, causing a reduction in the casein fractions, especially β-casein, and an increase in whey proteins. The effect of the combined proportions of polymorphonuclear cells and lymphocytes had a smaller impact and in the opposite direction.

ABSTRACT

Udder health in dairy herds has in recent years become a very important issue given its implications for animal welfare and the production of high-quality milk. Somatic cell count (SCC) is the most widely used means of assessing udder health status. However, differential somatic cell count (DSCC) has recently been proposed as a new and more effective means of evaluating intra-mammary infection dynamics. Differential somatic cell count represents the combined percentage of polymorphonuclear neutrophils and lymphocytes (PMN-LYM) in the total somatic cell count, with macrophages (MAC) making up for the remaining proportion. This study's aim was to evaluate the association between SCC and DSCC and the detailed milk protein profile in a population of 1,482 Holstein cows. A validated reversed-phase high-performance liquid chromatography (RP-HPLC) method was used to quantify four caseins (CN), namely αs1-, αs2-, κ-, and β-CN, and three whey protein fractions, namely β-LG, α-LA and LF, which were expressed both quantitatively (g/L) and qualitatively (as a percentage of the total milk nitrogen content, %N). A linear mixed model was fitted to explore the associations between SCS combined with DSCC and the protein fractions expressed as g/L and %N. We ran an additional model that included DSCC expressed as PMN-LYM and MAC counts, obtained by multiplying the percentages of PMN-LYM and MAC by SCC for each cow in the data set. When the protein fractions were expressed as g/L, SCS was significantly negatively associated with almost all the casein fractions, and positively associated with the whey protein α-LA, while DSCC was significantly associated with αs1-CN, β-CN and α-LA, but in the opposite direction to SCS. We observed the same pattern with the qualitative data (i.e. %N), confirming opposite effects of SCS and DSCC on milk protein fractions. The PMN-LYM count was only slightly associated with the traits of concern, although the pattern observed was the same as when both SCS and DSCC were included in the model. The MAC count, however, generally had a greater impact on many casein fractions, in particular decreasing both β-CN content (g/L) and proportion (%N), and exhibited the opposite pattern to the PMN-LYM count. Our results show that information obtained from both SCS and

DSCC may be useful to in assessing milk quality and protein fractions, and demonstrate the potential of MAC count as a novel udder health trait.

Keywords: differential somatic cell count, macrophages, milk quality, protein fractions, HPLC, dairy cattle

INTRODUCTION

The Italian dairy sector has a significant impact on the country's economy as most of the milk produced is processed into high quality cheese, including Protected Designation of Origin products (Bertoni et al., 2001; Bittante et al., 2012). It is widely acknowledged that the quantities and proportions of milk protein fractions are one of the most important factors involved in the coagulation and cheese-making process (Guinee, 2003; Caroli et al., 2009; Bittante et al., 2012). Caseins, in particular, which account for almost 80% of the total milk proteins, are directly involved in curd formation and cheese yield (Wedholm et al., 2006). For example, β-CN content is positively associated with cheese yield, while κ-CN is responsible for the stability of casein micelles and influences dry matter yield (Cipolat-Gotet et al., 2018). Whey proteins, on the other hand, despite their high nutritional value, have less economic importance since they are not directly involved in the milk coagulation process, although they influence nutrient recovery in the curd. For example, Cipolat-Gotet et al. (2018) observed that higher milk β-LG contents were associated with lower fat and protein recoveries.

When it comes to efficient dairy production, an important factor having a decisive influence on milk production, quality and, consequently, cheese production is udder health (Urech et al., 1999), which is the subject of continuously improving farm management strategies (Nørstebø et al., 2019). At present, the most widely used indicator able to rapidly provide information on both udder health and milk quality, is somatic cell count (SCC) (Guzzo et al., 2018). Somatic cells comprise a small, but constant proportion of mammary epithelial cells, and consist in different immune cell populations, mainly lymphocytes (LYM), polymorphonuclear neutrophils (PMN) and macrophages (MAC), each making different contributions to the inflammatory response, also according to the infection stage. Lymphocytes, for example, are responsible for the regulation and suppression of immune responses (Nickerson, 1989), while PMN are known to defend against invading bacteria during the early stage of an acute inflammatory process (Oviedo-Boyso et al., 2006). Finally, macrophages, having

phagocytic and pinocytotic activity, are appointed for the non-specific cell-induced immunity (Sordillo et al., 1997).

SCC is a robust indicator of the intensity of the mammary gland immune response, although it does not provide information on the proportions of the individual cell populations, and therefore provides no information on the stage of the inflammatory response and the progression of the disease. A step forward was recently made with the development of a new parameter to be coupled with SCC: the differential somatic cell count (DSCC). This trait represents the combined proportions of PMN and LYM in the milk expressed as a percentage, from which the MAC proportion can be calculated by subtracting the DSCC from 100% (Damm et al., 2017).

Milk with a high SCC undergoes changes to its composition due to increased proteolytic activity (Wickström et al., 2009), which has detrimental effects on its technological properties, such as reduced clotting ability and cheese yield (Bobbo et al., 2016; Pegolo et al., 2021). Proteolytic activity seems, in particular, to affect the casein fraction rather than the whey proteins, and has commonly been attributed to the PMN (Le Roux et al., 2003).

Recent studies have started to conduct phenotypic investigation of DSCC and how variations in it are associated with total SCC (Damm et al., 2017), udder health traits (Zecconi et al., 2019), milk yield and composition (Zecconi et al., 2020; Stocco et al., 2020), and technological properties (Pegolo et al., 2021a ; Pegolo et al., 2021b)

To date, no studies have yet investigated the possible associations between DSCC and the fine composition of milk, and, specifically, the detailed protein profile. Exploring these associations could shed light on some of the effects already found with respect to coagulation properties (Pegolo et al., 2021), as the latter are strictly connected with milk protein fractions (especially caseins). In addition, since the different leukocyte populations exhibit different levels of proteolytic activity, the results might be useful for identifying potential novel indicator traits of mammary gland inflammation. Therefore, the aim of this work was to evaluate the association between DSCC, either combined with

SCS or expressed as PMN-LYM and MAC counts, and the detailed milk protein profile, expressed both quantitatively (g/L) and qualitatively (% N), in 1,482 Holstein cows.

MATERIALS AND METHODS

Field Data and Sample Collection

The present study was carried out as part of the LATSAN and BENELAT projects, whose aim is to develop short- and long-term interventions for improving animal welfare and efficiency, and the quality of dairy cattle production.

Individual milk samples were collected from 1,482 Holstein-Friesian cows belonging to six different herds (herds size ranging from 27 to 951 lactating cows) located in northern Italy. Three of the herds were located in the production area of the Grana Padano Protected Designation of Origin (PDO) cheese. All animals were housed in free stalls and fed TMR formulated in accordance with the nutritional requirements set by the NRC (2001) and based on corn silage, sorghum silage and concentrates. Drinking water was distributed through automatic water bowls and the animals were milked twice per day. Cows with any clinical signs of disease or under medical treatment were excluded from the sampling. Milk samples were collected once during the evening milking from March 2019 to February 2020 (one or more samplings per herd, depending on the size of the herd, 21 herd-date combinations in total).

After collection, each sample was divided into two aliquots: bronopol preservative was added to one, which was then transferred to the laboratory of the Breeders Association of the Veneto where it was stored at 4 °C until analysis of milk composition and udder health traits (within 48 hours of collection); the other was taken to the laboratory of the Department of Agronomy, Food, Natural Resources, Animals and Environment (DAFNAE) of the University of Padova, Italy, within 3 hours of sampling, where it was stored at -80 °C until detailed analysis of the protein profile (within 4 months of collection).

Analysis of Milk Composition and Udder Health Traits

Samples were analyzed for milk composition traits (fat, protein, casein, lactose and urea contents) with an FT6000 Milkoscan infrared analyzer (Foss A/S, Hillerød, Denmark), while SCC (cells/mL) and DSCC (%) were determined with a Fossomatic TM 7DC analyzer (Foss A/S, Hillerød, Denmark). To obtain a normal distribution, the SCC were transformed to SCS according to the equation proposed by Ali and Shook (1980).

In accordance with Pegolo et al. (2021), the percentages of DSCC were converted to cell counts as follows:

PMN-LYM count, 10^3/mL = DSCC × SCC (10^3/mL);

MAC count, 10^3/mL = (1 – DSCC) × SCC (10^3/mL).

Also the PMN-LYM and MAC counts were log-transformed as log(PMN-LYM or MAC count/100,000) + 3, in the same way as SCS, to obtain a normal distribution of the data.

Milk Protein Profiling

The validated RP-HPLC method proposed by Maurmayr et al. (2013) and recently applied in Bisutti et al. (2022) was used to quantify the contents of κ-CN, αs1-CN, αs2-CN, β-CN, β-lactoglobulin (β-LG), α-lactalbumin (α-LA) and lactoferrin (LF) in 1,264 individual milk samples. We were also able to discriminate between the proportions of glycosylated and non-glycosylated κ-CN, thus the total κ-CN content was the sum of the two peaks. Briefly, an Agilent 1260 Series chromatograph (Agilent Technologies, Santa Clara, CA, USA) coupled to a quaternary pump (Agilent 1260 Series, G1311B) was used to separate the protein fractions using a C:8 reverse-phase analytical column (Aeris WIDEPORE XB-C8, Phenomenex, Torrance, CA) with large pore core-shell packaging (3.6 μm, 300 Å, 250 x 2.1 I.D). A gradient elution was carried out with a combination of two solvents as reported in Maurmayer et al. (2018): solvent A was 94.9% water, 5.0% acetonitrile

and 0,1% trifluoroacetic acid and Solvent B was 0,1% TFA in acetonitrile. The flow rate was 0.5 ml/min and the injection volume was 2 μL.

All protein fractions were expressed both quantitatively (g/L) and qualitatively (%N), as reported in detail in Amalfitano et al. (2020). The quantitative measures were calculated by multiplying the proportion of each protein fraction determined by HPLC by the proportion of casein determined by the FT6000 infrared analyzer, while the qualitative measures were calculated as their percentages of the total nitrogen content.

Statistical Analysis

The dataset underwent preliminary editing for all the investigated traits in order to exclude records falling outside the interval between ±3 standard deviations of the mean. Only milk samples with >50,000 SCC/ml and <1,500,000 SCC/ml were used in the analysis, as the Fossomatic 7DC analyzer has a reportedly poor level of accuracy for extremely low or high SCC values (Damm et al., 2017). Following Pegolo et al. (2021), we did not assume any linear relationship between predictors and response variables, which was why we discretized the explanatory variables (i.e. SCC, DSCC, PMN-LYM count, and MAC count) and created classes on the basis of the 25th, 50th and 75th percentiles. This approach ensures better assessment of the pattern of DSCC and SCS effects and ensures the classes are balanced. Moreover, SCC and DSCC were included in the same model to account for different degrees of mammary gland health as it was reported by Damm et al. (2017) and Zecconi et al. (2020) DSCC values may differ in animals with low SCC (healthy or at an early stage of the disease) and medium-high SCC (acute or chronic stage of the disease).

The associations between SCS and DSCC on the one hand and milk yield, composition and protein fractions, expressed both as g/L and %N, on the other, were estimated using the following linear mixed model implemented in the *lme4* package of the R software:

$$y_{ijklmn} = \mathrm{DIM}_i + \mathrm{Parity}_j + \mathrm{SCS}_k + \mathrm{DSCC}_l + \mathrm{Herd/date}_m + e_{ijklmn} \ [\mathbf{M1}]$$

where y_{ijklmn} is the measure of the trait, DIM_i is the fixed effect of i^{th} class of DIM (6 levels; i = <60; 60-120; 121-180; 181-240; 241-300; >300), $Parity_j$ is the fixed effect of the j^{th} parity (4 levels; j = 1, 2, 3, ≥4), SCS_k is the fixed effect of the k^{th} class of SCS defined by quartiles (4 levels; k = <2.63; 2.63-3.33; 3.34-4.29; >4.29), $DSCC_l$ is the fixed effect of the l^{th} class of DSCC defined by quartiles (4 levels; l = <59.9; 59.9-71.2; 72.3-79.7; >79.7, %), $Herd/date_m$ is the random effect of the m^{th} herd-date (21 levels), and e_{ijklmn} is the random residual assumed to have a normal distribution with $e_{ijklmn} \sim N(0, A\sigma_e^2)$.

In addition to **M1**, another linear mixed model (**M2**) was built to evaluate the specific effects of PMN-LYM and MAC counts in explaining the total variance for the investigated traits:

$$y_{ijklmn} = DIM_i + Parity_j + PMN\_LYM_k + MAC_l + Herd/date_m + e_{ijklmn} \quad [\mathbf{M2}]$$

In **M2**, SCS and DSCC were excluded from the vector of fixed effects**,** while the model included $PMN\text{-}LYM_k$ as the fixed effect of PMN-LYM count class defined by quartiles (k = 4 levels: <2.01; 2.01-2.68; 2.69-3.78; >3.78), and MAC_l as the fixed effect of MAC count class defined by quartiles (l = 4 levels: <1.03; 1.03-1.61; 1.62-2.36; >2.36). All the other terms were the same as in **M1**. Polynomial contrasts (P <0.05) were estimated to describe the pattern of the effects of SCS, DSCC (%), and PMN-LYM and MAC counts; first order comparisons measured the linear relationships, second order comparisons the quadratic relationships, and third order comparisons the cubic relationships. A model including the SCS × DSCC interaction was also tested, but it was significant for only a few traits and exhibited erratic patterns, so the results are not reported here.

RESULTS AND DISCUSSION

Previous studies have investigated the phenotypic associations between DSCC traits and milk composition (Stocco et al., 2020; Zecconi et al., 2020) or technological properties (Pegolo et al., 2021b). Furthermore, correlations previously obtained between SCS and DSCC (r=0.44), and

between PMN-LYM count and MAC (r=0.65) count appear to suggest that they capture different biological features (Pegolo et al., 2021). Given the essential role of protein fractions in determining the cheese-making ability of milk, and the effects of some leucocytes populations on the integrity and functionality of milk proteins, this study focused, for the first time, on evaluating the relationship between DSCC and the detailed milk protein profile.

Descriptive Statistics of Udder Health Traits and Protein Fractions

Descriptive statistics for the investigated traits are reported in Table 1. The average values of the main milk constituents are consistent with previous studies on Holstein cows (Maurmayr et al., 2018; Saha et al., 2020). Among the udder health traits, average SCS was 2.65 (± 1.94), in agreement with other studies on Holstein cattle (Malchiodi et al., 2014), showing that the farming systems and the overall hygiene conditions in which the sampled herds were kept were good. Mean DSCC was 69.09% (± 13.51%), while the mean MAC percentage, calculated as 100% minus DSCC, was 30.97% (± 13.62%). When DSCC was expressed quantitatively, i.e. as logPMN-LYM and logMAC counts, these averaged 3.09 ± 1.35 and 1.82 ± 1.04, respectively. Also in the case of DSCC, the mean value was found to be consistent with previous studies (Stocco et al., 2020).

Regarding the protein profile expressed both quantitatively and qualitatively, the predominant casein fractions were α_{s1}-CN and β-CN, with mean values of 9.06 g/L (26.47 %N) and 8.97 g/L (26.19 %N), respectively. The fractions α_{s2}-CN and κ-CN (calculated as the sum of the glycosylated and carbohydrate-free forms) were less represented and averaged 3.48 (10.13 %N) and 5.25 g/L (15.29 %N), respectively. The most abundant of the whey protein fractions was β-LG, with a mean value of 3.44 g/L (10.02 %N), followed by α-LA (0.99 g/L, 2.88 %N) and lactoferrin (LF) (0.16 g/L, 0.47 %N). Overall, the mean values for both the quantitative (g/L) and qualitative (%N) measures were similar to those found by Amalfitano et al. (2020) and Maurmayr et al. (2018), with the exception of total caseins and β-CN, which were lower than those reported by Amalfitano et al. (2020) (28.36 g/L and 78.49 %N, 10.69 g/L and 29.69 %N, respectively), and α_{s2}-CN, which was instead slightly higher

(2.78 g/L and 7.2 %N). These differences could be due to several factors, such as differences in population size and breed, but also in the analytical method used (Bonfatti et al., 2011; Gebreyesus et al., 2016).

Effects of Herd-Date

We observed that herd-date explained a relatively low proportion of the total variance (<25%) for all the investigated traits (Table 2), with the exception of urea (56%), which is well known to be influenced by diet (Roseler et al., 1993; Schiavon et al., 2015), as well as environmental and management conditions (Maurmayr et al., 2018). This result is in agreement with previous studies (Schopen et al., 2009; Bonfatti et al., 2011), confirming that variation in the milk protein composition is mainly related to genetic factors.

Association of SCS and DSCC with Milk Yield and Composition

The results of the ANOVA between SCS combined with DSCC and milk composition and the protein profile are reported in Table 2. In agreement with Amalfitano et al. (2020), DIM and Parity resulted important sources of variation for almost all the traits considered, confirming the importance of including these factors in the statistical models used to analyze these kind of traits.

The effects of SCS and DSCC on milk yield and composition have been extensively investigated by Stocco et al. (2020) and Pegolo et al. (2021). A significant decrease in milk yield and lactose was observed as SCS increased (data not shown) while increased DSCC was associated with increased milk yield and lactose (data not shown). Our results generally agree with previous studies, with the only exception of the fat proportion, which was not significantly affected by either SCS or DSCC in our study, likewise reported by Bansal et al. (2005).

Association of SCS and DSCC with Milk Detailed Protein Profile

Regarding protein composition measured quantitatively (g/L), Table 2 shows that SCS had a significant effect on true protein (TP) (P <0.05). Among the casein fractions, SCS appeared

significantly associated to κ-CN (P <0.01) and α_{s1}-CN (P <0.05), which surprisingly linearly increased at increasing SCS (Figure1). Moving on to the whey proteins, SCS had a significant positive effect not only on total whey protein content (P <0.01), but also on β-LG and α-LA (P <0.01). All whey proteins linearly increased at increasing SCS (Figure 1). Also DSCC evidenced a significant (P <0.05) and positive effect on β-CN, but a significant negative effect on the total whey protein content (P <0.01), β-LG (P <0.01) and α-LA (P <0.001), an opposite pattern to SCS (Table 2 and Figure1). When expressing the protein fractions as percentages of total milk nitrogen (%N), SCS negatively affected the total casein proportion (P <0.001) and β-CN (P <0.001), while DSCC showed the opposite pattern, and its increase resulted associated to a linear increase of both total casein and β-CN (P <0.001, Figure 2). Both SCS and DSCC had significant effects on the total proportion of whey protein (P <0.05 and P <0.01, respectively) and β-LG (P <0.05 and P <0.01, respectively), confirming once again an opposite pattern between them: a higher SCS was associated with higher proportions of total whey protein and β-LG, while higher DSCC was associated with lower proportions of those fractions. α-LA was significantly affected only by DSCC (P <0.01), the association being negative, as with the other whey proteins. β-LG is the most abundant whey protein in milk and it is synthetized in the epithelial cells of the mammary gland. The reduction of β-LG with the increase of DSCC needs to be further evaluated, as at the moment no clear information are known for its biological role. However, we might speculate that β-LG, belonging to lipocalin family, which includes, among others, fatty acid-binding proteins and bacterial metalloprotease inhibitors (Broersen, 2020), could have a role in the innate immune response of the mammary gland.

A significant reduction in β-LG and α-LA has been previously observed in mastitic animals with elevated SCC (Ishikawa et al., 1982; Hogarth et al., 2004; Ramos et al., 2015). As no cow enrolled in our study exhibited clinical signs of mastitis, we could assume that no inflammatory damage of mammary secretory tissue or increased leakage between tight junctions in the mammary gland could have affected the reduction of either β-LG and α-LA.

The reduction in the proportion of total casein, and more specifically the β-CN fraction, observed at the increasing of SCS could be ascribed either to reduced synthesis of caseins or to the proteolytic activity of plasmin, the principal indigenous proteinase in milk, which seems to break the β-casein protein chain, in particular (Le Roux et al., 2003; Crudden et al., 2005; Forsbäck et al., 2010). The opposite pattern shown by DSCC could, on the other hand, be ascribed to the fact that the present study was carried out on a population of clinically healthy animals, and it is therefore likely that the LYM predominated over the PMN and MAC, cells which are known for their proteolytic activity.

Different Leucocytes Populations Have Contrasting Behavior on Milk Protein Fractions

Since DSCC is a parameter that provides information on the proportions of the different leucocyte populations in milk, it would be interesting evaluate if the trends previously observed could either change or be confirmed when considering the quantitative variables, i.e. the PMN-LYM and MAC counts. Including both variables in the model did not induce any bias since they were only moderately correlated (Pegolo et al., 2021).

Table 3 reports the results of the ANOVA for the associations between PMN-LYM and MAC counts and milk composition and protein profile. Expressed quantitatively, β-CN was the only casein fraction significantly affected by both PMN-LYM (P <0.05) and MAC (P <0.001). In particular, an increase across PMN-LYM count classes was associated with an increase in β-CN contents (+5.63%), while an increase across MAC count classes was associated with lower β-CN contents (-6.75%) (Figure 3). Moreover, MAC count significantly affected total whey protein and β-LG (P <0.05), as well as α-LA (P <0.001), which was also significantly associated with PMN-LYM count (P <0.01). The effect of PMN-LYM and MAC counts on the whey protein fractions was the opposite to what we observed for β-CN, with MAC count associated with an increase in total whey proteins, (+9.98%), β-LG (+ 11.31%) and α-LA (+10.44%), while PMN-LYM was negatively associated with α-LA (-7.29%).

Regarding the protein fractions expressed as percentages of total nitrogen (%N), we confirmed the previously observed trend, with MAC count having a significant negative association with total

caseins and β-CN (P <0.001), and a positive association with total whey proteins (P <0.01), β-LG (P <0.05) and α-LA (P <0.001) (Figure 4). Once again, PMN-LYM count showed the opposite trend to MAC count, being associated with an increase in the β-CN proportion (P <0.05), and a decrease in the α-LA proportion (P <0.001).

In order to correctly interpret these results, it is important to remember that DSCC represents the combined proportions of PMN and LYM, which have very different distributions in healthy and diseased udders. It is in fact thought that LYM and MAC predominate among the immune cells in a healthy mammary gland, while PMN constitute only a small proportion (Wickström et al., 2009; Schwarz et al., 2011), but predominate once inflammation occurs (Kehrli and Shuster, 1994). There is still some uncertainty concerning the distribution of immune cells during inflammation. Damm et al. (2017) and Stocco et al. (2020) observed that PMN and MAC seem to change dramatically and in opposite directions across the whole SCC range, while LYM remained fairly constant and were able to sustain the immune response. In contrast, other studies reported that the most marked change related to inflammation process concerned the proportions of LYM and PMN, with MAC remaining relatively unchanged (Schwarz et al., 2011; Pilla et al., 2013).

The general low SCS values that we observed highlights the overall good conditions of the cows, so it was most likely that the predominant leucocyte populations were LYM and MAC, rather than PMN. The main function of macrophages, which contain lysosomes with proteolytic enzymes, such as collagenase, elastase and cathepsins, is to defend the udder from infection (Kelly and McSweeney, 2003). They have also been found to produce *in vitro* the urokinase-type plasminogen activator (u-PA) (Politis et al., 1991), which converts inactive plasminogen into plasmin, the main native proteolytic enzyme in milk. The contribution of MAC to proteolysis in milk has been investigated by Caroprese et al. (2007), who found appreciable casein hydrolysis in bulk ewes' milk, which regarded α-CN, in particular, while β-CN did not undergo any appreciable degradation. This difference might

be related to the slightly different structures and distributions of casein fractions in different animal species.

CONCLUSIONS

Our results confirm the unfavorable effect of SCC on milk caseins, and specifically on β-CN, which is probably related to the proteolytic activity of milk native enzymes, like plasmin, and shed new light on the association between DSCC and milk proteins. In particular, the unfavorable association we found between MAC count and casein fractions might be related to the proteolytic activity of MAC and the fact that they were present in larger amounts than PMN, given the non-pathological status of the animals. These results highlight the importance of this trait as a potential novel udder health indicator and could be of interest for the dairy industry, given the important role of caseins in milk coagulation properties and cheese production.

To validate these findings and assess the association between DSCC and milk protein fractions in greater detail, it will be necessary to be able to accurately differentiate the PMN and LYM populations – which are combined in the DSCC parameter – and investigate their individual effects. Furthermore, possible developments of this study will concern the study of genetic correlations between DSCC and protein fractions which could provide useful information from breeding perspectives.

ACKNOWLEDGEMENTS

This research was carried out as part of the BENELAT project, funded by the Regione Lombardia (Bando d.d.s. no. 4403, 28/03/2018) and the LATSAN Project, funded by the Ministero delle Politiche Agricole, Alimentari, Forestali (MIPAAF), Rome, Italy. The authors wish to thank Dr. Sofia Ton and Marco Franzoi of the Breeders Association of the Veneto Region (ARAV, Padua, Italy) for the milk composition analyses. Vittoria Bisutti holds a PhD grant from the CARIPARO foundation.

REFERENCES

Ali, A.K.A., and G.E. Shook. 1980. An Optimum Transformation for Somatic Cell Concentration in Milk. J. Dairy Sci. 63:487–490. doi:10.3168/jds.S0022-0302(80)82959-6.

Amalfitano, N., G. Stocco, A. Maurmayr, S. Pegolo, A. Cecchinato, and G. Bittante. 2020. Quantitative and qualitative detailed milk protein profiles of 6 cattle breeds: Sources of variation and contribution of protein genetic variants. J. Dairy Sci. 103:11190–11208. doi:10.3168/jds.2020-18497.

Bansal, B.K., J. Hamann, N.T. Grabowski, and K.B. Singh. 2005. Variation in the composition of selected milk fraction samples from healthy and mastitic quarters, and its significance for mastitis diagnosis. J. Dairy Res. 72:144–152. doi:10.1017/S0022029905000798.

Bertoni, G., L. Calamari, and M. Grazia Maianti. 2001. Producing specific milks for speciality cheeses. Proc. Nutr. Soc. 60:231–246. doi:10.1079/pns200080.

Bittante, G., M. Penasa, and A. Cecchinato. 2012. Invited review: Genetics and modeling of milk coagulation properties. J. Dairy Sci. 95:6843–70. doi:10.3168/jds.2012-5507.

Bobbo, T., S. Biffani, C. Taccioli, M. Penasa, and M. Cassandro. 2021. Comparison of machine learning methods to predict udder health status based on somatic cell counts in dairy cows. Sci. Rep. 11:1–10. doi:10.1038/s41598-021-93056-4.

Bobbo, T., M. Penasa, and M. Cassandro. 2019. Short communication: Genetic aspects of milk differential somatic cell count in Holstein cows: A preliminary analysis. J. Dairy Sci. 102:4275–4279. doi:10.3168/jds.2018-16092.

Bonfatti, V., A. Cecchinato, L. Gallo, A. Blasco, and P. Carnier. 2011. Genetic analysis of detailed milk protein composition and coagulation properties in Simmental cattle. J. Dairy Sci. 94:5183–5193. doi:10.3168/jds.2011-4297.

Broersen, K. 2020. Milk processing affects structure, bioavailability and immunogenicity of β-lactoglobulin. Foods 9. doi:10.3390/foods9070874.

Caroli, A.M., S. Chessa, and G.J. Erhardt. 2009. Invited review: Milk protein polymorphisms in cattle: Effect on animal breeding and human nutrition. J. Dairy Sci. 92:5335–5352. doi:10.3168/jds.2009-2461.

Cipolat-Gotet, C., A. Cecchinato, M. Malacarne, G. Bittante, and A. Summer. 2018. Variations in milk protein fractions affect the efficiency of the cheese-making process. J. Dairy Sci. 101:8788–8804. doi:10.3168/jds.2018-14503.

Crudden, A., F. Fox, and A.L.K. Ã. 2005. Factors affecting the hydrolytic action of plasmin in milk 15:305–313. doi:10.1016/j.idairyj.2004.08.008.

Damm, M., C. Holm, M. Blaabjerg, M.N. Bro, and D. Schwarz. 2017. Differential somatic cell count—A novel method for routine mastitis screening in the frame of Dairy Herd Improvement testing programs. J. Dairy Sci. 100:4926–4940. doi:10.3168/jds.2016-12409.

Forsbäck, L., H. Lindmark-Månsson, A. Andrén, and K. Svennersten-Sjaunja. 2010. Evaluation of quality changes in udder quarter milk from cows with low-to-moderate somatic cell counts. Animal 4:617–626. doi:10.1017/S1751731109991467.

Gebreyesus, G., M.S. Lund, L. Janss, N.A. Poulsen, L.B. Larsen, H. Bovenhuis, and A.J. Buitenhuis. 2016. Short communication: Multi-trait estimation of genetic parameters for milk protein composition in the Danish Holstein. J. Dairy Sci. 99:2863–2866. doi:10.3168/jds.2015-10501.

Guinee, T.. 2003. Role of protein in cheese and cheese products. P.F. Fox and P.L.. McSweeney, ed. Springer, New York, NY.

Guzzo, N., C. Sartori, and R. Mantovani. 2018. Genetic parameters of different measures of somatic cell counts in the Rendena breed. J. Dairy Sci. 101:8054–8062. doi:10.3168/jds.2017-14047.

Hogarth, C.J., J.L. Fitzpatrick, A.M. Nolan, F.J. Young, A. Pitt, and P.D. Eckersall. 2004. Differential protein composition of bovine whey: A comparison of whey from healthy animals and from those with clinical mastitis. Proteomics 4:2094–2100. doi:10.1002/pmic.200300723.

Ishikawa, H., T. Shimizu, H. Hirano, N. Saito, and T. Nakano. 1982. Protein Composition of Whey from Subclinical Mastitis and Effect of Treatment with Levamisole. J. Dairy Sci. 65:653–658. doi:10.3168/jds.S0022-0302(82)82244-3.

Kehrli, M.E., and D.E. Shuster. 1994. Factors Affecting Milk Somatic Cells and Their Role in Health of the Bovine Mammary Gland. J. Dairy Sci. 77:619–627. doi:10.3168/jds.S0022-0302(94)76992-7.

Maurmayr, A., A. Cecchinato, L. Grigoletto, and G. Bittante. 2013. Detection and quantification of α S1-, α-S2-, β-, κ-casein, α-lactalbumin, β-lactoglobulin and lactoferrin in bovine milk by reverse-phase high-performance liquid chromatography. Agriculturae Conspectus Scientificus 78:201–205.

Maurmayr, A., S. Pegolo, F. Malchiodi, G. Bittante, and A. Cecchinato. 2018a. Milk protein composition in purebred Holsteins and in first/second-generation crossbred cows from Swedish Red, Montbeliarde and Brown Swiss bulls. Animal 12:2214–2220. doi:10.1017/S1751731117003640.

Maurmayr, A., S. Pegolo, F. Malchiodi, G. Bittante, and A. Cecchinato. 2018b. Milk protein composition in purebred Holsteins and in first/second-generation crossbred cows from Swedish Red, Montbeliarde and Brown Swiss bulls. Animal 1–7. doi:10.1017/S1751731117003640.

Nickerson, S.C. 1989. Immunological Aspects of Mammary Involution. J. Dairy Sci. 72:1665–1678. doi:10.3168/jds.S0022-0302(89)79278-X.

Nørstebø, H., G. Dalen, A. Rachah, B. Heringstad, A.C. Whist, A. Nødtvedt, and O. Reksen. 2019. Factors associated with milking-to-milking variability in somatic cell counts from healthy cows

in an automatic milking system. Prev. Vet. Med. 172:104786. doi:10.1016/j.prevetmed.2019.104786.

Oviedo-Boyso, J., J.J. Valdez-Alarcón, M. Cajero-Juárez, A. Ochoa-Zarzosa, J.E. López-Meza, A. Bravo-Patiño, and V.M. Baizabal-Aguirre. 2007. Innate immune response of bovine mammary gland to pathogenic bacteria responsible for mastitis. J. Infect. 54:399–409. doi:10.1016/j.jinf.2006.06.010.

Pegolo, S., D. Giannuzzi, V. Bisutti, R. Tessari, M.E. Gelain, L. Gallo, S. Schiavon, F. Tagliapietra, E. Trevisi, P. Ajmone Marsan, G. Bittante, and A. Cecchinato. 2021. Associations between differential somatic cell count and milk yield, quality, and technological characteristics in Holstein cows. J. Dairy Sci. 104:4822–4836. doi:10.3168/jds.2020-19084.

Pilla, R., M. Malvisi, G.G.M. Snel, D. Schwarz, S. König, C.P. Czerny, and R. Piccinini. 2013. Differential cell count as an alternative method to diagnose dairy cow mastitis. J. Dairy Sci. 96:1653–1660. doi:10.3168/jds.2012-6298.

Ramos, T., F. Costa, I. Pinto, S. Pinto, and L. Abreu. 2015. Effect of Somatic Cell Count on Bovine Milk Protein Fractions. J. Anal. Bioanal. Tech. 6:1–7. doi:10.4172/2155-9872.1000269.

Roseler, D.K., J.D. Ferguson, C.J. Sniffen, and J. Herrema. 1993. Dietary Protein Degradability Effects on Plasma and Milk Urea Nitrogen and Milk Nonprotein Nitrogen In Holstein Cows. J. Dairy Sci. 76:525–534. doi:10.3168/jds.S0022-0302(93)77372-5.

Le Roux, Y., F. Laurent, and F. Moussaoui. 2003. Polymorphonuclear proteolytic activity and milk composition change. Vet. Res. 34:629–645. doi:10.1051/vetres:2003021.

Saha, S., N. Amalfitano, G. Bittante, and L. Gallo. 2020. Milk coagulation traits and cheese yields of purebred Holsteins and 4 generations of 3-breed rotational crossbred cows from Viking Red, Montbéliarde, and Holstein bulls. J. Dairy Sci. 103:3349–3362. doi:10.3168/jds.2019-17576.

Schiavon, S., G. Cesaro, F. Tagliapietra, L. Gallo, and G. Bittante. 2015. Influence of N shortage and conjugated linoleic acid supplementation on some productive , digestive , and metabolic parameters of lactating cows. Anim. Feed Sci. Technol. 208:86–97. doi:10.1016/j.anifeedsci.2015.07.016.

Schopen, G.C.B., J.M.L. Heck, H. Bovenhuis, M.H.P.W. Visker, H.J.F. Van Valenberg, and J.A.M. Van Arendonk. 2009. Genetic parameters for major milk proteins in Dutch Holstein-Friesians. J. Dairy Sci. 92:1182–1191. doi:10.3168/jds.2008-1281.

Schwarz, D., U.S. Diesterbeck, S. König, K. Brügemann, K. Schlez, M. Zschöck, W. Wolter, and C.P. Czerny. 2011. Flow cytometric differential cell counts in milk for the evaluation of inflammatory reactions in clinically healthy and subclinically infected bovine mammary glands. J. Dairy Sci. 94:5033–5044. doi:10.3168/jds.2011-4348.

Sordillo, L.M., K. Shafer-Weaver, and D. DeRosa. 1997. Immunobiology of the Mammary Gland. J. Dairy Sci. 80:1851–1865. doi:10.3168/jds.S0022-0302(97)76121-6.

Stocco, G., A. Summer, C. Cipolat-Gotet, L. Zanini, D. Vairani, C. Dadousis, and A. Zecconi. 2020. Differential somatic cell count as a novel indicator of milk quality in dairy cows. Animals 10:1–14. doi:10.3390/ani10050753.

TM, R., F.. Costa, I.S.. Pinto, and L.. Abreu. 2015. Effect of Somatic Cell Count on Bovine Milk Protein Fractions. J. Anal. Bioanal. Tech. 6:1–7. doi:10.4172/2155-9872.1000269.

Urech, E., Z. Puhan, and M. Schällibaum. 1999. Changes in milk protein fraction as affected by subclinical mastitis. J. Dairy Sci. 82:2402–2411. doi:10.3168/jds.S0022-0302(99)75491-3.

Wedholm, A., L.B. Larsen, H. Lindmark-Månsson, A.H. Karlsson, and A. Andrén. 2006. Effect of protein composition on the cheese-making properties of milk from individual dairy cows. J. Dairy Sci. 89:3296–3305. doi:10.3168/jds.S0022-0302(06)72366-9.

Wickström, E., K. Persson-Waller, H. Lindmark-Månsson, K. Östensson, and Å. Sternesjö. 2009. Relationship between somatic cell count, polymorphonuclear leucocyte count and quality parameters in bovine bulk tank milk. J. Dairy Res. 76:195–201. doi:10.1017/S0022029909003926.

Zecconi, A., F. Dell'orco, D. Vairani, N. Rizzi, M. Cipolla, and L. Zanini. 2020. Differential somatic cell count as a marker for changes of milk composition in cows with very low somatic cell count. Animals 10:1–14. doi:10.3390/ani10040604.

Zecconi, A., D. Vairani, M. Cipolla, N. Rizzi, and L. Zanini. 2019. Assessment of subclinical mastitis diagnostic accuracy by differential cell count in individual cow milk. Ital. J. Anim. Sci. 18:460–465. doi:10.1080/1828051X.2018.1533391.

TABLES AND FIGURES

Table 1. Descriptive statistics of single test-day milk yield, composition, udder health, and protein fractions expressed both as quantitative (g/L) and qualitative measures (%N)

Trait[1]	N	Mean	SD	P1[2]	P99[2]
Milk yield, kg/d	1,481	33.60	9.41	11.78	57.30
Milk composition					
Fat, %	1,472	3.66	0.83	1.47	5.71
Protein, %	1,479	3.43	0.34	2.72	4.32
Casein, %	1,480	2.68	0.28	2.09	3.43
Lactose, %	1,470	4.86	0.23	4.21	5.29
Urea, mg/100g	1,480	26.87	5.90	13.80	40.60
Udder health					
SCC, 10^3/ml	1,482	272.30	1192.41	7.00	2972.03
SCS, units	1,482	2.65	1.94	-0.84	7.89
DSCC, %	1,260	69.09	13.51	30.97	91.90
MAC, %	1,260	30.97	13.62	8.10	69.48
Log PMN-LYM count	1,260	3.09	1.35	0.81	6.43
Log MAC count	1,260	1.82	1.04	-0.14	4.68
Protein composition, g/L of milk					
True protein	1,264	31.38	3.30	24.39	39.96
Caseins	1,262	26.75	2.77	21.00	33.78
α_{s1}-CN	1,260	9.06	1.08	6.65	11.88
α_{s2}-CN	1,263	3.48	0.68	2.07	5.20
β-CN	1,264	8.97	1.34	5.80	12.24
κ-CN	1,262	5.25	1.05	3.16	8.14
Glycosylated κ-CN	1,243	1.60	0.58	0.66	3.33
Whey proteins	1,259	4.58	0.92	2.77	7.19
β-LG	1,259	3.44	0.84	1.76	5.94
α-LA	1,258	0.99	0.14	0.68	1.36
LF	1,256	0.16	0.06	0.05	0.33
NPN compounds	1,252	2.93	0.91	0.70	5.30
Urea	1,263	0.28	0.06	0.15	0.41
Minor N compounds	1,234	2.13	0.84	0.40	4.35
Protein composition, % of total milk N					
True protein	1,253	91.41	2.62	85.31	98.19
Caseins	1,259	78.07	1.20	75.08	80.63
α_{s1}-CN	1,261	26.47	2.05	22.07	31.83
α_{s2}-CN	1,262	10.13	1.64	6.25	14.13
β-CN	1,257	26.19	2.60	19.45	31.57
κ-CN	1,260	15.29	2.60	10.15	22.63
Glycosylated κ-CN	1,245	4.67	1.62	1.98	9.34
Whey proteins	1,257	13.38	2.35	8.30	19.88
β-LG	1,258	10.02	2.21	5.31	16.09
α-LA	1,258	2.88	0.39	2.03	3.99
LF	1,257	0.47	0.16	0.16	0.91
NPN compounds	1,252	8.59	2.62	1.81	14.69
Urea	1,236	2.45	0.54	1.31	3.78
Minor N compounds	1,235	6.23	2.39	1.11	12.17

[1]SCS = log2 (SCC/100,000) + 3; DSCC = differential SCC; log PMN-LYM count = polymorphonuclear neutrophils-lymphocytes count expressed as log2 (PMN-LYM count/100,000) + 3; log MAC count =macrophages count expressed as log2 (MAC count/100,000) + 3; caseins= sum of α_{s1}-CN + α_{s2}-CN + β-CN + κ-CN; Whey proteins= sum of β-LG + α-LA + LF; LF= lactoferrin; NPN compounds= non protein nitrogen compounds, sum of urea and minor non protein N compounds (e.g., small peptides, ammonia, creatine, creatinine).
[2]P1 = 1st percentile; P99 = 99th percentile

Table 2. Results from ANOVA (*F*-value and significance) for single test-day milk yield, composition traits and protein fractions, expressed both as quantitative (g/L) and qualitative measures (%N)

Trait[1]	Parity	DIM	SCS[2]	DSCC[3]	Herd-date
Milk yield, kg/d	9.52***	38.90***	5.10**	8.52***	10
Milk composition					
Fat, %	3.07*	2.73*	2.49	1.62	25
Protein, %	7.78***	30.16***	3.02*	1.05	11
Casein, %	9.65***	31.78***	1.83	0.91	9
Lactose, %	18.61***	4.96***	17.99***	14.74***	7
Urea, mg/100g	0.02	2.26*	0.86	0.50	56
Protein composition, g/L of milk					
True protein	7.13***	33.20***	3.43*	2.12	7
Caseins	10.27***	31.88***	1.42	0.82	8
α_{s1}-CN	2.17	12.39***	2.68*	2.10	20
α_{s2}-CN	3.27*	6.76***	0.74	1.94	25
β-CN	16.52***	15.09***	1.48	3.56*	17
κ-CN	1.05	14.80***	4.57**	0.91	5
Glycosylated κ-CN	1.33	25.30***	2.47	1.70	10
Whey proteins	1.15	8.22***	3.88**	4.27**	5
β-LG	1.82	6.34***	4.79**	4.73**	3
α-LA	2.08	1.63	5.11**	6.87***	16
LF	1.22	7.84***	0.08	0.39	16
N compound	0.03	0.56	0.13	1.60	14
Urea	0.03	2.40*	0.98	0.43	56
Minor N compound	0.09	0.67	0.11	1.50	13
Protein composition, % of total milk N					
True protein	0.26	3.08**	0.45	2.60	11
Caseins	14.85***	10.55***	7.55***	9.32***	24
α_{s1}-CN	1.23	2.43*	0.59	0.55	23
α_{s2}-CN	1.36	1.53	0.47	1.46	33
β-CN	13.33***	2.88*	11.50***	11.86***	18
κ-CN	1.72	2.70*	2.20	0.74	8
Glycosylated κ-CN	0.90	14.77***	1.82	1.34	11
Whey proteins	2.11	0.76	2.70*	4.85**	8
β-LG	2.71*	1.40	2.73*	3.85**	6
α-LA	3.37*	4.77***	2.21	3.87**	18
LF	0.36	3.14**	0.20	0.44	16
N compound	0.26	3.08**	0.45	2.60	11
Urea	1.61	10.66	2.96	1.45	50
Minor N compound	0.26	1.77	0.09	2.11	10

[1]caseins= sum of α_{s1}-CN + α_{s2}-CN + β-CN + κ-CN; Whey proteins= sum of β-LG + α-LA + LF; LF= lactoferrin; NPN compounds= non protein nitrogen compounds, sum of urea and minor non protein N compounds (e.g., small peptides, ammonia, creatine, creatinine).

[2]SCS = log2 (SCC/100,000) + 3. [3]DSCC= differential SCC. [4]L = linear; Q = quadratic, C = cubic.

[5]Herd/Date effect expressed as proportion of variance explained by herd/test date calculated by dividing the corresponding variance component by the total variance.

*$P < 0.05$; **$P < 0.01$; ***$P < 0.001$.

Table 3. Results from ANOVA (*F*-value and significance) for single test-day milk yield, composition traits and protein fractions expressed both as quantitative (g/L) and qualitative measures (%N) using log counts of PMN and lymphocytes (PMN-LYM) and macrophages (MAC).

Trait[1]	Parity	DIM	Log PMN-LYM[2]	Log MAC[3]	Herd-Date, %[4]
Milk yield, kg/d	9.61***	39.39***	1.88	7.71***	10
Milk composition					
Fat, %	3.59*	2.77*	0.33	3.06*	25
Protein, %	8.16***	30.67***	1.78	1.25	11
Casein, %	10.01***	32.27***	2.39	1.49	9
Lactose, %	19.29***	5.35***	2.82*	20.52***	7
Urea, mg/100g	0.04	2.27*	0.99	0.43	56
Protein composition, g/L of milk					
True protein	7.21***	33.90***	1.14	0.85	7
Caseins	10.69***	32.20***	2.09	1.73	8
α_{s1}-CN	2.42	12.68***	1.88	0.78	20
α_{s2}-CN	3.22*	6.99***	0.50	1.54	25
β-CN	17.54***	15.13***	2.97*	6.07***	16
κ-CN	1.14	14.81***	1.25	0.79	5
Glycosylated κ-CN	1.48	25.56***	0.43	1.71	10
Whey proteins	1.20	8.90***	0.87	4.44*	5
β-LG	1.84	7.11***	0.64	3.73*	4
α-LA	1.57	1.80	4.76**	8.50***	16
LF	1.26	8.38	0.09	0.32	16
N compound	0.05	0.50	1.47	0.71	14
Urea	0.04	2.41*	0.95	0.42	56
Minor N compound	0.11	0.62	1.46	0.67	13
Protein composition, % of total milk N					
True protein	0.20	3.30**	0.66	0.61	11
Caseins	15.08***	10.59***	2.37	13.85***	25
α_{s1}-CN	1.38	2.27*	1.35	0.10	23
α_{s2}-CN	1.31	1.78	2.40	1.67	33
β-CN	14.14***	3.05**	3.16*	13.16***	18
κ-CN	1.90	2.55*	0.81	0.77	8
Glycosylated κ-CN	0.98	14.99***	0.61	1.24	11
Whey proteins	2.35	1.09	2.51	5.40**	8
β-LG	2.87*	1.78	1.38	3.62*	6
α-LA	3.21*	4.53***	7.71***	8.15***	18
LF	0.37	3.69**	0.09	0.45	16
N compound	0.20	3.30**	0.66	0.61	11
Urea	1.60	10.87***	1.51	0.43	50
Minor N compound	0.20	1.90	1.04	0.58	10

[1]caseins= sum of α_{s1}-CN + α_{s2}-CN + β-CN + κ-CN; Whey proteins= sum of β-LG + α-LA + LF; LF= lactoferrin; NPN compounds= non protein nitrogen compounds, sum of urea and minor non protein N compounds (e.g., small peptides, ammonia, creatine, creatinine). [2]Log PMN-LYM count = polymorphonuclear neutrophils-lymphocytes count expressed as log2 (PMN-LYM count/100,000) + 3; [3]Log MAC count =macrophages count expressed as log2 (MAC count/100,000) + 3. [4]Herd/Date effect expressed as proportion of variance explained by herd/test date calculated by dividing the corresponding variance component by the total variance. *$P < 0.05$; **$P < 0.01$; ***$P < 0.001$.

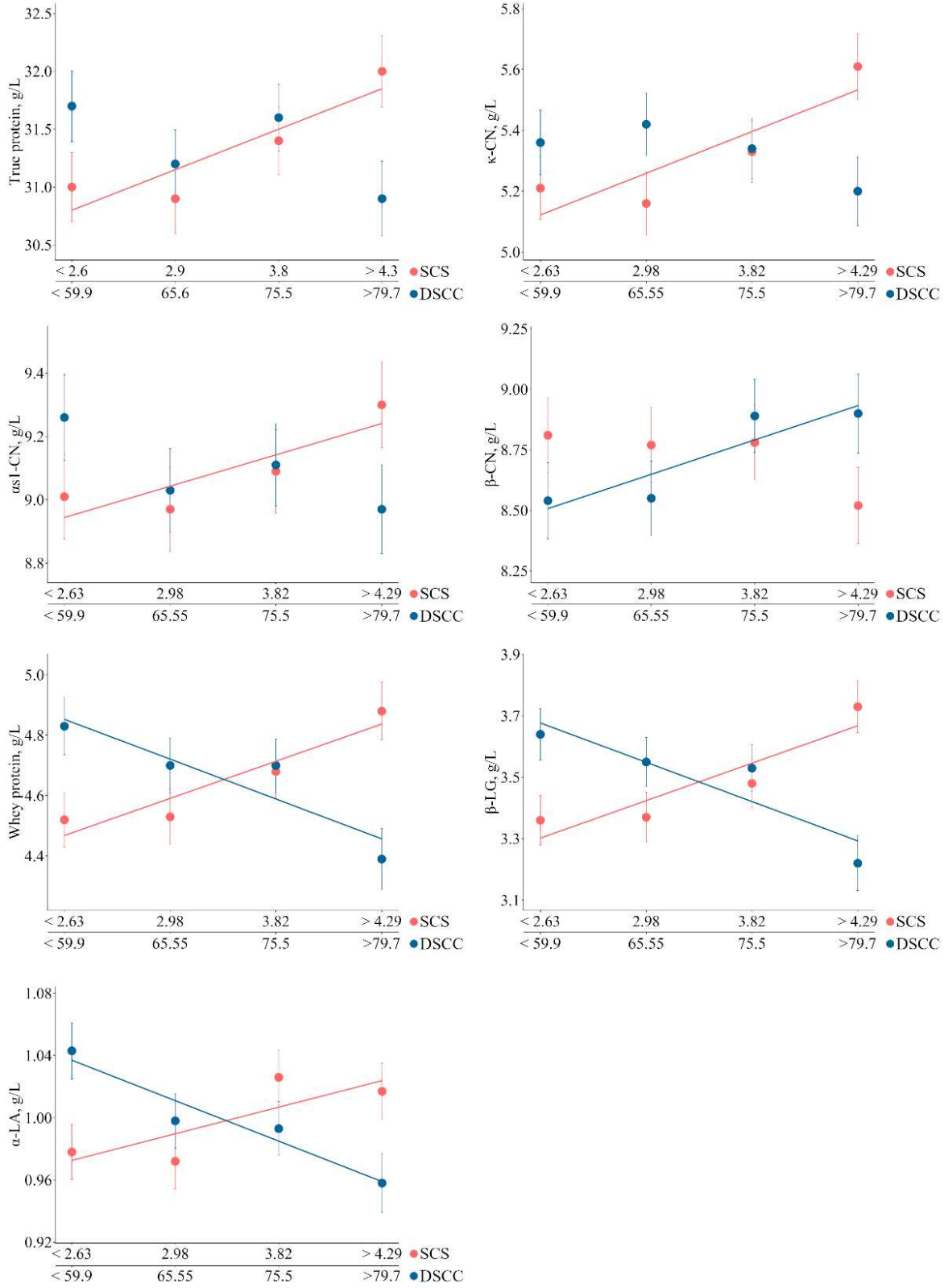


Figure 1. Least squares (LS) means and SE of milk protein fractions expressed as quantitative measures (g/L) across somatic cell score (SCS) and differential somatic cell count (DSCC) classes.

Only significant results ($P < 0.05$) are displayed. When the trend line across SCS or DSCC LS-means classes is presented, it means that the polynomial contrast was significant at $P < 0.05$. Only linear contrasts were significant. Values on the x-axis refer to the average value for each class of SCS [top row; expressed as log2 (SCC/100,000) + 3] and DSCC [bottom row; expressed as percentage].

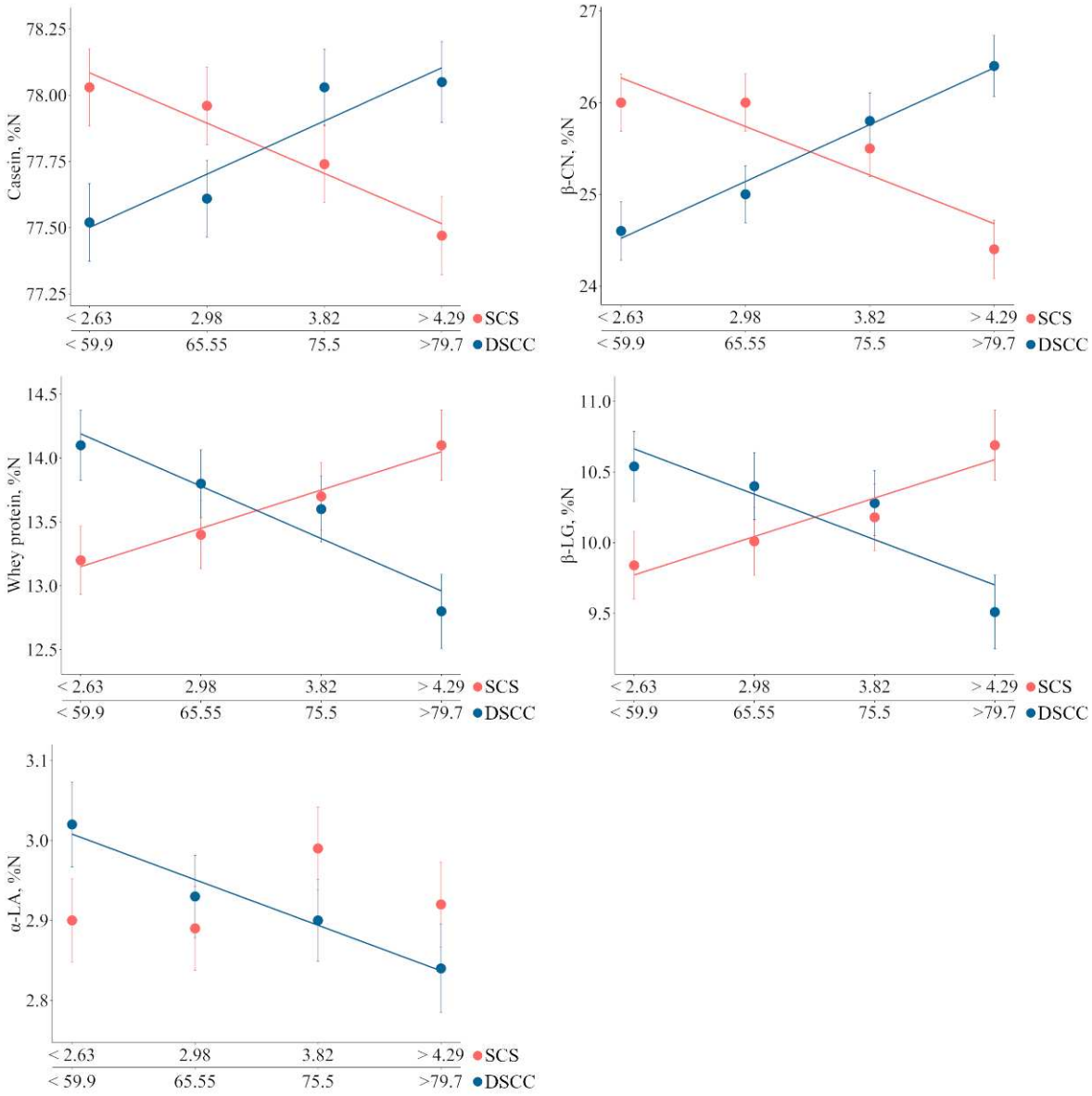


Figure 2. Least squares (LS) means and SE of milk protein fractions expressed as qualitative measures (%N) across somatic cell score (SCS) and differential somatic cell count (DSCC) classes. Only significant results ($P < 0.05$) are displayed. When the trend line across SCS or DSCC LS-means classes is presented, it means that the polynomial contrast was significant at $P < 0.05$. Only linear contrasts were significant. Values on the x-axis refer to the average value for each class of SCS [top row; expressed as log2 (SCC/100,000) + 3] and DSCC [bottom row; expressed as percentage].

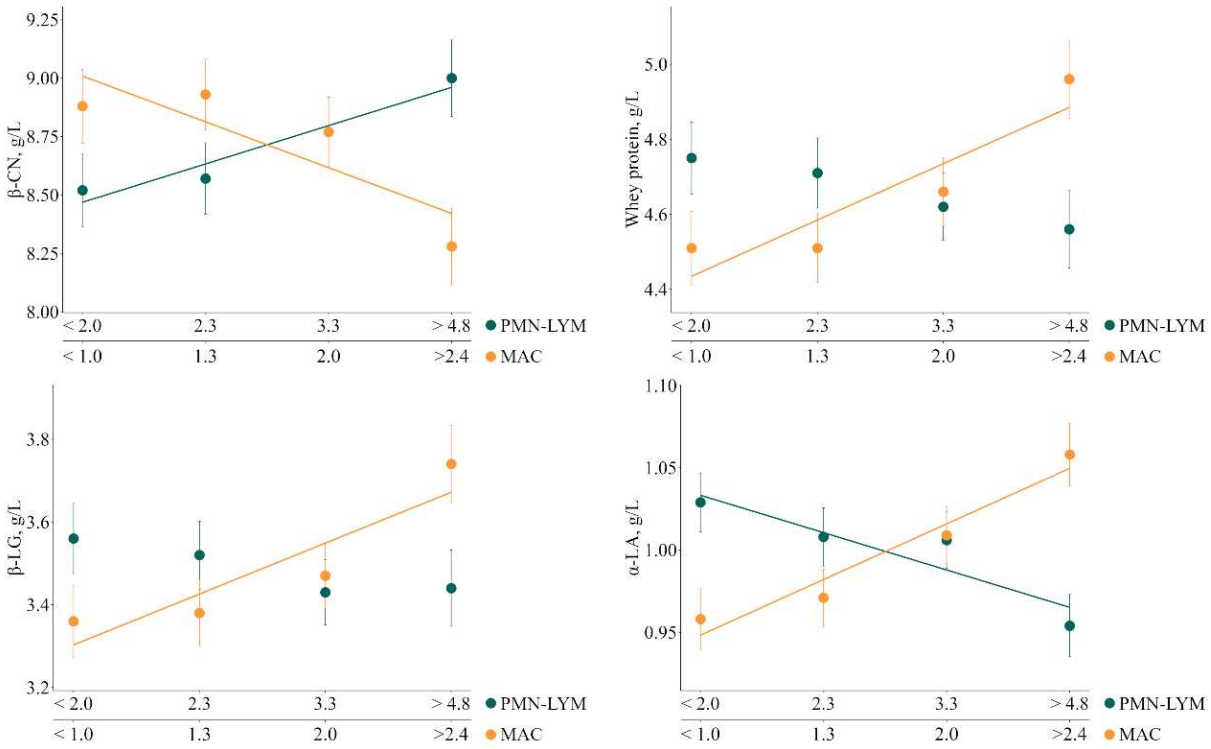


Figure 3. Least squares (LS) means and SE of milk protein fractions expressed as quantitative (g/L) measures across polymorphonuclear neutrophils and lymphocytes (PMN-LYM), and macrophages (MAC) classes. Only significant results ($P < 0.05$) are displayed. When the trend line across SCS or DSCC LS-means classes is presented, it means that the polynomial contrast was significant at $P < 0.05$. Only linear contrasts were significant. Values on the x-axis refer to the average value for each class of PMN-LYM [top row; expressed as log2 (PMN-LYM count/100,000) + 3] and MAC [bottom row; expressed as log2 (MAC count/100,000) + 3].

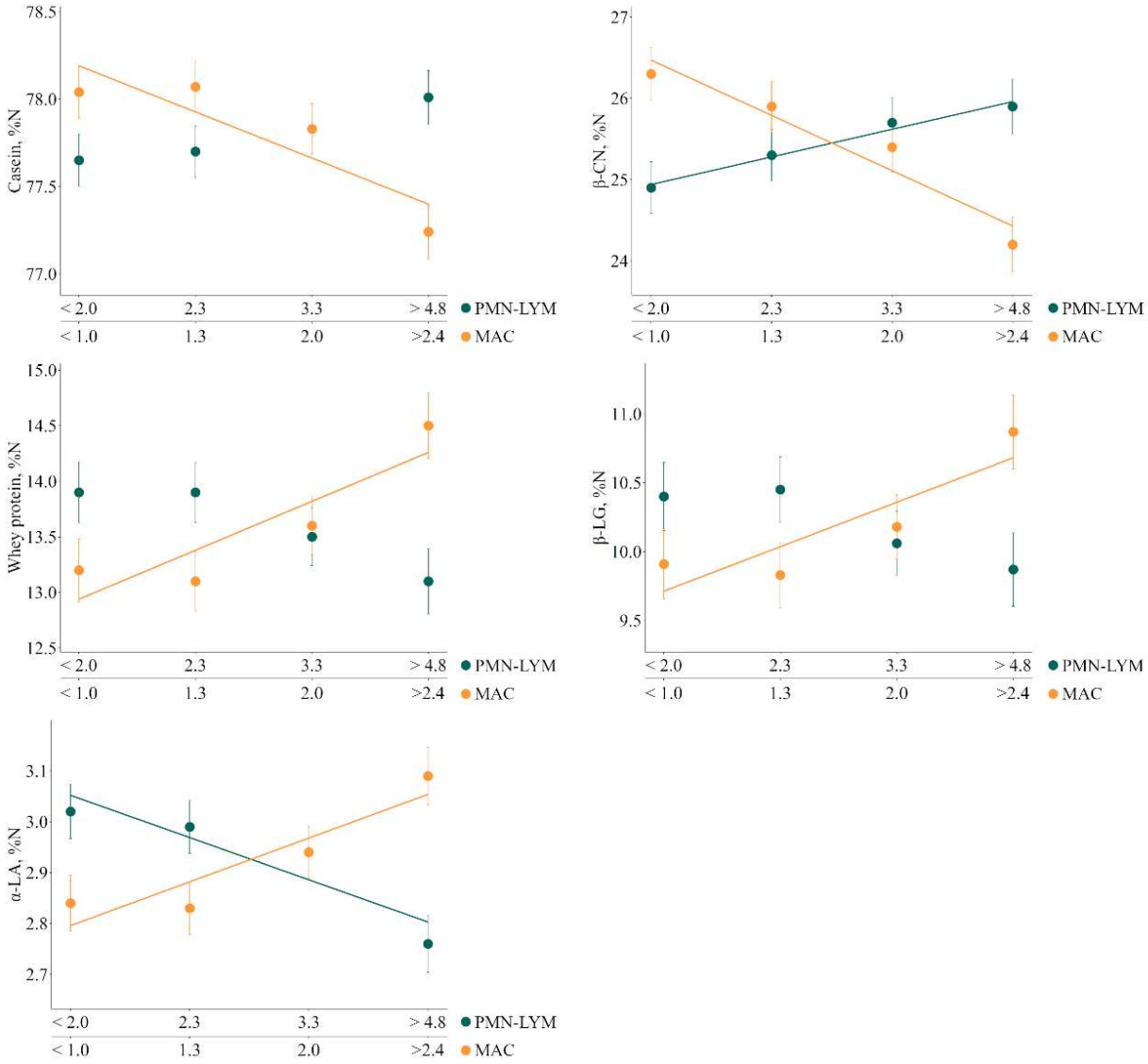


Figure 4. Least squares (LS) means and SE of milk protein fractions expressed as qualitative measures (%N) across polymorphonuclear neutrophils and lymphocytes (PMN-LYM), and macrophages (MAC) classes. Only significant results ($P < 0.05$) from model are displayed. When the trend line across SCS or DSCC LS-means classes is presented, it means that the polynomial contrast was significant at $P < 0.05$. Only linear contrasts were significant. Values on the x-axis refer to the average value for each class of PMN-LYM [top row; expressed as log2 (PMN-LYM count/100,000) + 3] and MAC [bottom row; expressed as log2 (MAC count/100,000) + 3].

CHAPTER II

SUBCLINICAL MASTITIS AND MILK PROTEIN FRACTIONS

Effect of intramammary infection and inflammation on milk protein profile assessed at the quarter level in Holstein cows

V.Bisutti[1], A. Vanzin[1], S. Pegolo[1*], A. Toscano[1], M. Gianesella[2], E. Sturaro[1], S. Schiavon[1], L. Gallo[1], F. Tagliapietra[1], D. Giannuzzi[1], A. Cecchinato[1]

[1]Department of Agronomy, Food, Natural Resources, Animals and Environment (DAFNAE), University of Padova, Viale dell' Università 16, 35020, Legnaro (PD), Italy

[2]Department of Animal Medicine, Productions and Health, University of Padua, Viale dell' Università, 16, 35020, Legnaro (PD), Italy

In press in: Journal of Dairy Science (Available online from the 19th of October)

https://doi.org/10.3168/jds.2023-23818

INTERPRETATIVE SUMMARY

Effect of intramammary infection and inflammation on milk protein profile assessed at the quarter level in Holstein cows

By Bisutti et al., page 000. In this study we evaluated the existing association between naturally occurring subclinical intramammary infection, in combination with somatic cell count, with the detailed protein profile in Holstein cows assessed at the individual quarter level. We observed that the infection, but mostly the inflammation status of the mammary gland, affected the milk protein profile, and in particular the casein fractions. These results helped elucidate the behavior of protein fractions during subclinical intramammary infection.

SUBCLINICAL MASTITIS AND MILK PROTEIN FRACTIONS

Effect of intramammary infection and inflammation on milk protein profile assessed at the quarter level in Holstein cows

V.Bisutti[1], A. Vanzin[1], S. Pegolo[1*], A. Toscano[1], M. Gianesella[2], E. Sturaro[1], S. Schiavon[1], L. Gallo[1], F. Tagliapietra[1], D. Giannuzzi[1], A. Cecchinato[1]

[1]*Department of Agronomy, Food, Natural Resources, Animals and Environment (DAFNAE), University of Padova, Viale dell' Università 16, 35020, Legnaro (PD), Italy*

[2]*Department of Animal Medicine, Productions and Health, University of Padua, Viale dell' Università, 16, 35020, Legnaro (PD), Italy*

*Corresponding author: Sara Pegolo, Department of Agronomy, Food, Natural resources, Animals and Environment (DAFNAE), University of Padova, Viale dell'Università 16, 35020 Legnaro (PD), Italy, Email: sara.pegolo@unipd.it

ABSTRACT

In this study we wanted to investigate the associations between naturally occurring subclinical IMI caused by different aetiological agents (i.e., *Staphylococcus aureus*, *Streptococcus agalactiae*, *Streptococcus uberis* and *Prototheca* spp.), in combination with SCC, on the detailed milk protein profile measured at the individual mammary gland quarter. An initial bacteriological screening (time 0; T0) conducted on individual composite milk from 450 Holstein cows reared in three herds, was performed to identify cows with subclinical IMI. We identified 78 infected animals which were followed up at the quarter level at two different sampling times: T1 and T2, two and six weeks after T0, respectively. A total of 529 quarter samples belonging to the previously selected animals were collected at the two sampling points and analyzed with a reversed phase high performance liquid chromatography (RP-HPLC) validated method. Specifically, we identified and quantified four caseins (CN), namely α_{S1}-CN, α_{S2}-CN, κ-CN, and β-CN, and three whey protein fractions, namely β-lactoglobulin, α-lactalbumin, and lactoferrin (LF), which were later expressed both quantitatively (g/L) and qualitatively (as a percentage of the total milk nitrogen content, %N). Data were analyzed with a hierarchical linear mixed model with the following fixed effects: days in milk (DIM), parity, herd, SCC, bacteriological status (BACT), and the SCC × BACT interaction. The random effect of individual cow, nested within herd, DIM and parity was used as the error term for the latter effects. Both IMI (i.e., BACT) and SCC significantly reduced the proportion of β-CN and α_{S1}-CN, ascribed to the increased activity of both milk endogenous and microbial proteases. Less evident alterations were found for whey proteins, except for LF, which being a glycoprotein with direct and undirect antimicrobial activity, increased both with IMI and SCC, suggesting its involvement in the modulation of both the innate and adaptive immune response. Finally, increasing SCC in the positive samples was associated with a more marked reduction of total caseins at T1, and α_{S1}-CN at T2, suggesting a synergic effect of infection and inflammation, more evident at high SCC. In conclusion, our work helps clarify the behavior of protein fractions at quarter level in animals having subclinical

IMI. The inflammation status driven by the increase in SCC, rather the infection, was associated with the most significant changes, suggesting that the activity of endogenous proteolytic enzymes related to the onset of inflammation might have a pivotal role in directing the alteration of the milk protein profile.

Keywords: Subclinical mastitis, somatic cell count (SCC), udder health, dairy cows

INTRODUCTION

Mastitis is a mammary gland inflammatory condition, which typically develops because of the penetration of a wide range of microorganisms (mostly bacteria) in the udder. In the context of dairy production, despite the ongoing effort towards the improvement of farm management strategies, bovine mastitis remains, especially in the subclinical form, one of the most common and impactful diseases, leading to the impairment of animal welfare and huge economic losses, mostly related to the decreased milk productivity and quality, veterinary costs, and premature culling (Antanaitis et al., 2021; Nayan et al., 2022).

To date, the gold standard for the detection of intramammary infection (**IMI**) and the identification of the respective causative pathogens remains the bacteriological examination (Nagasawa et al., 2020), even if its systemic application at a farm level is not always possible as it is a time consuming and expensive technique. Milk somatic cell count (**SCC**) gives instead information on the inflammatory status of the mammary gland, and at present is the most widespread tool for screening rapidly and on-site the udder condition (Ruegg, 2017). In this context, in dairy farm practice a threshold of 200,000 somatic cells/mL is generally considered appropriate for identifying animals having subclinical mastitis (NMC, 2017; Dal Prà et al., 2022). However, there have been recent studies that demonstrated that SCC may not always provide a complete picture of the animals' health condition (Bobbo et al., 2017). Cases of animals positive at the microbiological examination with low somatic cells have been reported as well as culture-negative samples with high SCC (Bobbo et al., 2017; Pegolo et al., 2022).

It has been highly demonstrated that milk having high SCC undergoes through compositional changes, that negatively affect its composition and quality (Bobbo et al., 2016; Pegolo et al., 2020) and, therefore, its economic value. In this context, it has been observed that the increased proteolytic activity associated with high SCC seems to have a particularly strong impact on milk protein fractions

(Zhang et al., 2015) and especially on caseins (Caggiano et al., 2019), which are important traits influencing milk quality and technological characteristics.

In the context of the Italian dairy sector, where most of the produced milk is processed into high quality cheese, milk proteins, and especially caseins, play a pivotal role as their quantity and proportion can directly affect the milk coagulation and cheesemaking ability (Cipolat-Gotet et al., 2018; Amalfitano et al., 2019). Several studies have investigated the effects of clinical and subclinical mastitis on milk composition and quality (Paixão et al., 2017; Stocco et al., 2020) and technological traits (Pegolo et al., 2021a). However, to the best of our knowledge, no extensive research focused on the specific interaction between subclinical IMI and SCC (i.e., the inflammatory status) and the detailed milk protein profile, when measured at the mammary gland quarter level.

Therefore, the aim of this study was to investigate the association between naturally occurring subclinical IMI caused by four different pathogens (*Staphylococcus aureus*, *Streptococcus agalactiae*, *Streptococcus uberis* and *Prototheca* spp.) in combination with SCC on the detailed milk protein profile assessed at the quarter level in Holstein cows.

MATERIALS AND METHODS

Study Design, Animals and Sampling

This study was conducted within the LATSAN project, which aimed at developing innovative tools for improving mammary gland health and milk quality in dairy cattle. The research was approved by the Ethical Animal Care and Use Committee (OPBA – Organismo Preposto al Benessere degli Animali) of the Università Cattolica del Sacro Cuore and by the Italian Ministry of Health (protocol number 510/2019-PR of 19/07/2019).

This research was a longitudinal observational study aimed at inferring the associations between milk protein profile, determined at the quarter level, and subclinical IMI. The null hypothesis was that no association existed between them. The experimental unit was the mammary quarter within

cow, therefore all comparisons were made between positive and negative samples as well as between different classes of SCC, within cow. The required sample size for this study was calculated assuming that β-CN, being together with αs1-CN the predominant milk casein fraction, could be the most representative milk protein fractions affected by udder health (Forsback et al., 2010). Therefore, the minimum sample size was calculated to achieve a 95% confidence and 80% power to detect differences between negative and positive milk samples at the bacteriological screening of 1.6-point percentage of total milk N, corresponding to roughly 38% of the standard deviation of the trait. Considering an expected prevalence of animals affected by subclinical IMI of 20% and the variability of milk protein profile, we considered 280 quarters enough to detect desired differences. Milk samples were collected during the evening milking from 450 lactating Holstein cows belonging to three different commercial farms in Veneto region (N = 144, 71 and 235 for herd A, B and C, respectively). Herd selection was based on a prevalence study conducted by the *Istituto Zooprofilattico Sperimentale delle Venezie* (IZSVe) that investigated some of the most common causative agents of mastitis in the Veneto region (*Streptococcus. agalactiae*, *Staphylococcus. aureus*, *Streptococcus. uberis* and *Prototheca* spp.). The detailed description of the structure of the experimental design, as well as the management and farming conditions, are reported in the work of Pegolo et al. (2022) and Pegolo et al. (2023a). In brief, the first sampling (time 0, T0) consisted in performing a bacteriological screening on all lactating cows within each herd, by collecting aseptically 50 mL of composite milk, to identify healthy subjects (Neg) from the ones having subclinical IMI. Among the ones positive to the bacteriological screening, we selected those with no clinical signs of mastitis during that lactation period, and no previous history of antibiotic and other drug treatments. In total 78 animals were selected. Two subsequent milk sampling were performed after two (T1) and six (T2) weeks, from T0, respectively, in order to monitor the progress of the disease. On such occasions, we collected two milk aliquots of 50 mL from each mammary quarter. One aliquot was sent to the regional breeders association (*Associazione Regionale degli Allevatori del Veneto*, ARAV) for milk composition analysis and SCC determination. The second aliquot was instead transferred to the Milk Laboratory

of the Department of Agronomy, Food, Natural Resources, Animals and Environment (DAFNAE) of the University of Padova for the evaluation of the detailed milk protein profile through high performance liquid chromatography (HPLC). A schematic summary of the structure of the experimental design is also reported in Supplementary figure S1.

All milk samplings were performed in accordance with the guidelines published by the National Mastitis Council (NMC, 2017). Briefly, before the aseptic milk collection, teats were cleaned by the veterinarian with a commercial pre-milking disinfectant, dried with individual towels then cleaned again with sterile gauzes and alcohol. Cows were fed TMR mostly based on corn silage, sorghum silage and concentrate, and drinking water was available in automatic water bowls. Detailed description of the feeding regime adopted in the three herds is available in the study of Pegolo et al. (2023b) and also reported in Supplementary table S1.

Microbiological Analysis

Milk samples microbiological examination was carried out by the IZSVe laboratory (Legnaro, PD, Italy). Samples, which were delivered within 4 hours from collection, were frozen and analysed within three days. Pegolo et al. (2022) reported all the detailed procedures of the microbiological analyses. Briefly, ten µL of milk from the composite samples was inoculated on each of the following selective media: Baird-Parker agar with rabbit plasma fibrinogen (BP-RPF, Biokar Diagnostic, Beauvais, France; identification of *S. aureus*), thallium sulfate-crystal violet-B toxin blood agar (TKT, IZSVe internal production; identification of *Strep. agalactiae*), and *Prototheca* isolation medium (PIM, IZSVe internal production). BP-RPF plates were prepared following the ISO 6888-2 (2021) standards, TKT according to the method described by Hauge and Ellingsen (1953), and PIM in accordance with the NMC guidelines (NMC, 2017). All inoculated media were incubated at 37 °C. For BP- RPF plates colonies growth was observed at 24 and 48 hours from the incubation, and suspected *S. aureus* colonies were confirmed through the tube coagulase test (NMC, 2017). Concerning the TKT plates, growth was observed at 24 h, and suspected positive colonies for *Strep. agalactiae* were confirmed by CAMP test (NMC, 2017). *Prototheca* isolation medium plates were

observed at 24, 48, and 72 h, and suspected colonies were confirmed by wet mount method (NMC, 2017).

Quarter milk samples were cultured following the NMC guidelines (NMC, 2017). Briefly, 10 µL of milk was streaked on blood esculin agar, incubated at 37 °C and, after 16-18 h, single isolated colonies were picked, subcultured and identified through both morphological examination and additional NMC tests (NMC, 2017). For mixed species colonies, it was possible to isolate and identify up to five different colonies, if more were present the sample was labelled as "contaminated".

Bacterial colonies with questionable test results were identified by matrix-assisted laser desorption ionization time of flight mass spectrometry (MALDI-TOF MS) performed by Microflex Biotyper LT (Bruker Daltonics GmbH, Bremen, Germany). A log (score) ≥ 2 represented the threshold for species level identification.

Analysis of milk composition and SCC

Milk composition analysis (fat, protein, casein, lactose, and urea contents) was carried out using the Milkoscan FT6000 infrared Analyser (Foss A/S, Hillerød, Denmark) while SCC were determined using the FossomaticTM 7 DC analyzer (Foss A/S, Hillerød, Denmark).

Milk Protein Profiling

The validated Reversed Phase High Performance Liquid Cromatography (RP-HPLC) technique (Maurmayr et al., 2013) was applied to identify and quantify the following protein fractions: κ-CN (κ-casein), α_{S1}-CN, α_{S2}-CN, β-CN, β-LG (β-lactoglobulin), α-LA (α-lactalbumin), and LF (lactoferrin). With this method we were also able to identify the portion of glycosylated and non-glycosylated (carbohydrate-free) κ-CN by separating the two chromatographic peaks. Total κ-CN was expressed as the sum of these two portions. The instrument used consisted in an Agilent 1260 Series chromatograph (Agilent Technologies) coupled to a quaternary pump (Agilent 1260 Series, G1311B). A C:8 reverse-phase analytical column (Aeris WIDEPORE XB-C8, Phenomenex) with large pore core-shell packaging (3.6 µm, 200 Å, 250 × 2.1 mm i.d.) was used to separate the protein fractions. The injection volume was set at 2 µL and the flow rate was 0.5 mL/min. The

cromatographic run was carried out in a gradient elution, whose specifics are reported in detail by Maurmayr et al. (2013), with the combination of two solutions: solvent A was 94.9 % water, 5.0 % acetonitrile, and 0.1 % trifluoroacetic acid, and solvent B was 0.1 % trifluoroacetic acid in acetonitrile. All protein fractions were later expressed both quantitatively (g/L) and qualitatively (%N) as previously reported by Amalfitano et al. (2020).

Statistical Analysis

The association between quarter-level protein fractions and subclinical mastitis was explored using a hierarchical linear mixed model and the PROC MIXED procedure (SAS institute Inc., Cary, North Carolina, USA). Two separate analyses were conducted for T1 and T2 due to heterogeneity of the variances and to the possible change in bacteriological status of the quarter between T1 and T2. The following model was run for T1 and T2, where the experimental unit was the individual mammary gland quarter:

$$y_{ijklmn} = \mu + \mathrm{DIM}_i + \mathrm{Parity}_j + \mathrm{Herd}_k + \mathrm{SCC}_l + \mathrm{BACT}_m + (\mathrm{SCC} \times \mathrm{BACT})_{lm} + Cow_{n:ijk} + e_{ijklmn}$$

Where: y_{ijklmn} is the investigated trait (milk protein fractions, expressed both quantitatively [g/L] and qualitatively as a percentage of the total milk nitrogen content, %N); μ is the overall mean; DIM_i is the fixed effect of the *i*th class of days in milk (*i* = 3 classes; class 1 ≤ 120; 120 < class 2 ≤ 240; class 3 > 240); Parity_j is the fixed effect of the *j*th parity (*j* = primiparous or multiparous); Herd_k is the fixed effect of the *k*th herd/date (*k* = 3 herds); SCC_l is the fixed effect of the *l*th class of SCC (*l* = 4 classes; class 1 < 50,000 cell/mL; 50,000 cell/mL ≤ class 2 < 200,000 cell/mL; 200,000 cell/mL ≤ class 3 < 400,000 cell/mL; class 4 ≥ 400,000 cell/mL); BACT_m is the fixed effect of the *m*th class of bacteriological status (*m* = positive or negative); $(\mathrm{SCC} \times \mathrm{BACT})_{lm}$ is the two-way interaction between SCC_l and BACT_m; $Cow_{n:ijk}$ is the random effect of the *n*th animal nested within DIM_i, parity_j, and herd_k; and e_{ijklmn} is the random residual. Cow and the residuals were assumed to be normally distributed with a mean of zero and variances of σ^2_{Cow} and σ^2_e, respectively. With this nested

design, the effects of DIM_i, $parity_j$, and $herd_k$ were tested on the error line of the nth animal nested within DIM_i, $parity_j$, and $herd_k$, while the SCC_l, $BACT_m$, $(SCC \times BACT)_{lm}$, and animal effects were tested on the residual term.

We also tested the position and number of affected quarters, but since no significant outcome was obtained, we excluded them for the final analyses.

For SCC classes the orthogonal contrasts were set as follows:

- class 1 + class 2 *vs* class 3 + class 4 (low SCC *vs* high SCC)
- class 1 *vs* class 2 (within low SCC)
- class 3 *vs* class 4 (within high SCC)

While for BACT x SCC interaction, orthogonal contrasts were built as following:

- negative (SCC class 1 + class 2 + class 3 +class 4) *vs* positive (SCC class 1 + class 2 + class 3 + class 4)
- negative (low SCC) *vs* positive (low SCC)
- negative (high SCC) *vs* positive (high SCC)
- negative *vs* positive (SCC class 1)
- negative *vs* positive (SCC class 2)
- negative *vs* positive (SCC class 3)
- negative *vs* positive (SCC class 4)

The significance was set at $P < 0.05$.

RESULTS AND DISCUSSION

This longitudinal study evaluated for the first time the effect of naturally occurring subclinical IMI, inflammation status determined by the SCC level, and their combination, on milk protein fractions in Holstein cows at the quarter level.

The detailed results of the herd prevalence and the pathogens distribution for each farm have been described in the study of Pegolo et. al (2022). Briefly, among the positive animals identified during the first bacteriological screening (T0), *Strep. agalactiae* was the most prevalent microorganism found (n = 51, two herds), followed by *Prototheca* spp. (n = 19, two herds), *S. aureus* (n = 6, two herds) and finally *S. uberis* (n = 5, one herd). Three animals from one farm were co-infected by both *Strep. agalactiae* and *Prototheca*, and therefore excluded from the trial.

As reported by Pegolo et al. (2022), no significant differences were found in terms of milk production and composition between negative and positive animals to the bacteriological examination, underlying that subclinical infection may lead to no detectable alterations. On the other hand, as expected, SCC significantly differed between negative (199,700 cells/mL) and positive animals (688,600 cells/mL) ($P < 0.001$).

Milk protein fractions were identified and quantified on 529 quarter milk samples from the 78 subclinically infected animals identified at T0, divided according to the sampling time as follows: 253 for T1 and 276 for T2, respectively. Nineteen samples (8 and 11 at T1 and T2, respectively) were found contaminated, thus they were excluded from the statistical analysis. In total, 63 quarter samples tested positive for the presence of microorganisms at T1 while the number of positive samples was reduced to 47 at T2. As previously stated, *Strep. agalactiae* was the prevalent pathogen, with 41 positive samples at T1 and 34 positive samples at T2 (Supplementary table S2). *Strep. agalactiae* is a reemerging pathogen in European countries, and is becoming one of the most important causative agents of subclinical mastitis, despite the implemented eradication programs (Barsi et al., 2022). In Italy, no systematic studies have been carried out to evaluate the prevalence of this microorganism.

Only two regions, Emilia Romagna and Lombardia have investigated it, with an estimated herd prevalence of 7-10% (Sora et al., 2022)

The descriptive statistics of the traits of concerns are reported in Table 1. As expected, the predominant casein fractions were β-CN and α_{S1}-CN, which averaged 9.83 g/L (27.35 %N) and 9.26 g/L (25.89 %N), respectively. Among the whey proteins, β-LG was the most abundant (4.47 g/L, 12.47 %N), followed by α-LA (1.04 g/L, 2.91 %N) and LF (0.16 g/L, 0.44 %N). These values were overall in accordance with previous works carried out using the same RP-HPLC method on composite milk samples (Amalfitano et al., 2020; Bisutti et al., 2022a).

The pattern of the protein fractions at the two sampling points was overall similar. Moreover, we did not find differences when comparing the effects of IMI and inflammation status between the two ways of expressing the protein fractions (quantitatively, g/L and qualitatively, %N). Therefore, to avoid redundancy we focused on the results expressed as percentage of the total milk nitrogen. The plot displaying the results of the protein fractions expressed as g/L are reported in the Supplementary materials (Supplementary figures S2, S3, S4 and S5). Moreover, as the behavior of the glycosylated and carbohydrate free κ-CN had the same pattern as the total κ-CN (calculated as the sum of these two fractions; data not shown), we focused only on the effects towards the latter.

Associations between subclinical intramammary infection and milk protein fractions

The results of the ANOVA for the IMI effect on the investigated traits are reported in Table 2. The evaluation of pathogen-specific effects towards the investigated traits was not possible due to an excessive fragmentation of the dataset which would have led to a significant reduction in the test power.

At T1, subclinical IMI was associated with most of the casein fractions. In particular, we observed that mammary quarters positive at the bacteriological examination had higher proportions of κ-CN ($P < 0.05$; + 5.8 %) and α_{s2}-CN ($P < 0.01$; + 5.7 %) but significantly lower β-CN ($P < 0.01$; -7.5 %) (Figure 1). The milk compositional changes in the infected mammary gland are related to the activity of immune cells like polymorphonuclear (**PMN**) cells and macrophages that release a great

amount of proteases and other cellular components in the milk (Kelly et al., 2006; Pisanu et al., 2015) while fighting invading pathogens. Moreover, infecting bacteria can release into the milk exogenous enzymes, like elastases (Guerrero et al., 2016), contributing to proteolysis. Furthermore, some bacteria can secrete activators of protease zymogens, which, synergically with native bovine enzymes lead to an increase in proteolytic potential of milk (Larsen et al., 2006). Enzymatic degradation seems to affect protein fractions in a different way. For example, Considine et al. (2000) observed a cleavage specificity of elastase on α_{s1}-CN and β-CN , suggesting that among the casein fractions those two might be the most suitable substrates for the proteolysis. Moreover, Hinz et al. (2012) assessed with two-dimensional gel electrophoresis (2-DE) the peptides produced by proteolysis in milk samples from cows infused with *Escherichia coli* LPS and observed that almost all the obtained peptides derived from α_{s1}-CN and β-CN, in accordance with what observed in our study.

At T2, subclinical IMI was associated with higher LF ($P < 0.05$; + 51%; Figure 1). Lactoferrin is an iron-binding glycoprotein synthesized by specific granules in PMN and glandular epithelial cells (Cheng et al., 2008), that is characterized by antimicrobial properties. The increased concentration of LF found in our study was in accordance to Chaneton et al. (2008) who also found a greater content of this protein in infected quarters of animal having clinical mastitis.

Apart from LF, we did not observe further significant effects of the bacteriological status towards the whey proteins. Different studies on the milk peptidome in cows with mastitis highlighted that peptides derived from the degradation of whey proteins were few or not even detected (Baum et al., 2013; Mansor et al., 2013). Hence, previous studies that observed a reduction of β-LG and α-LA in animals affected from clinical mastitis (Hogarth et al., 2004; Guerrero et al., 2016) attributed this alteration as a result of either the physical damage of the mammary epithelial cells from the pathogen invasion or the reduced synthesis and secretion of those proteins (McFadden et al., 1988), rather than from enzymatic degradation. Conversely, evaluation of whey from cows with subclinical mastitis through by two-dimensional gel electrophoresis (2-DE) did not observe a reduction in either one of those proteins (Baeker et al., 2002), in accordance with what we observed in our study.

We might speculate that this result might partly be related to the chronic nature of infection, and partially to the fact that animals involved in our study, even if all displayed a subclinical condition, were not in the same stage of infection, and therefore the changes in the whey population might have been less evident than the ones found for caseins.

Associations between milk somatic cell count and protein fractions

The results of the ANOVA for the SCC effect on the investigated traits are reported in Table 2. At T1, SCC was associated with all the casein fractions. Increases in SCC led to a reduction α_{s1}-CN (-8 %; $P < 0.001$) and β-CN (-8 %; $P < 0.01$). On the other hand, κ-CN and α_{s2}-CN showed an opposite trend, as we observed for both these casein fractions a higher proportion at increasing SCC (+17% and 4.9% for κ-CN and α_{s2}-CN, respectively; Figure 2). Indeed, the reduction in α_{s1}-CN and β-CN, which are the two most abundant protein fractions, is probably to be ascribed to the increase of milk endogenous proteases, which include plasmin, enzymes produced by leucocytes and those derived from the transition through the altered blood-milk barrier (Le Roux et al., 2003; L. Forsback et al., 2010). Plasmin, the main milk native enzyme (Crudden et al., 2005), exists as a component of a complex system, which includes its zymogen, plasminogen and plasminogen activators (PAs) (Kelly et al., 2006). Even though it is normally present in milk, both plasmin and plasminogen significantly increase with elevated SCC (Dallas et al., 2016), together with the urokinase type I plasminogen activator (u-PA), which activates plasminogen into plasmin (Guerrero et al., 2016). Along with plasmin, other proteases like cathepsin B and D, collagenases, and elastases, are released by lysosomes contained in macrophages and PMN cells with the onset of inflammation. Some of these enzymes (like elastase and cathepsin B) have the ability to kill different pathogens that are causative agents of mastitis (Ezzat Alnakip et al., 2014; Abdelmegid et al., 2020), but at the same time they significantly concur in the degradation of milk protein fractions (Le Roux et al., 2003; Wedholm et al., 2008).

It was observed that, among all caseins, β-CN, followed by α_{s1}-CN were the fractions most subjected to the enzymatic degradation (Urech et al., 1999; Ramos et al., 2015). In addition, studies

that evaluated changes in peptidome in animals affected by subclinical mastitis evidenced a more pronounced production of β-CN and α_{s1}-CN derived peptides (Guerrero et al., 2016; Addis et al., 2020), suggesting that κ-CN might be more resistant to the proteolysis by these enzymes than the other casein fractions. Therefore, we might speculate that the higher sensitivity to proteolysis of these two casein fractions might be attributable either to their aminoacidic structure or to their tertiary structure, which could make the proteases cleavage sites more accessible.

Among the whey proteins, SCC was associated with β-LG ($P < 0.05$) and LF ($P < 0.001$; + 95%) (Figure 3). However, in our study the pattern of variation of β-LG across the SCC classes was erratic, suggesting further evaluations. Aside from protein fractions, we observed a significant reduction in the urea proportion at the increase of SCC ($P < 0.001$; - 9.6%). Urea is the primary metabolite derived from dietary protein and tissue protein turnover and it is released from liver to the bloodstream and then to the milk in a stable form (Hayton et al., 2012). Urea is greatly influenced by management and feeding conditions, like the lack of rumen degradable proteins, alteration in dry matter intake, or the energy in the feed ration (Rezamand et al., 2007). There have been studies which reported a reduction in milk urea proportion in cows with IMI and increased SCC (Nyman et al., 2014; Timonen et al., 2017). The molecular mechanisms behind this association are not fully understood and therefore require further evaluations, however we might speculate being partially linked to the reduction of feed intake that is observed in animals having masitits .

Overall, the patterns of milk protein fractions according to SCC classes agreed at the two sampling points. In fact, at T2 SCC was associated with almost all the protein fractions, except for β-LG and α-LA. We observed a 9% reduction in β-CN ($P < 0.001$). On the flipping side, we confirmed the same behavior as at T1 for κ-CN (+ 14%; $P < 0.001$) and α_{s2}-CN (+ 5%; $P < 0.05$), whose proportion increased with higher SCC. The increase in κ-CN was found in accordance to a previous study which evaluated the impact of SCC and differential somatic cell count (DSCC) in clinically healthy animals (Bisutti et al., 2022b). We observed the same trend also for LF ($P < 0.001$; + 121%). Aside from its antimicrobial effect, it has been reported that LF can exert a wide range of actions

towards the immune system, from the inhibition of the inflammatory process to the modulation of both the innate and adaptive immune response (Drago-Serrano et al., 2017; Shimazaki and Kawai, 2017) . In our study we observed an increase of LF from T1 to T2, and this result might be ascribed to the fact that, with the progression of inflammation, LF might have accumulated due to both its induced production and additional degranulation of PMN (de la Rosa et al., 2008).

Association of the interaction between BACT and SCC with milk protein fractions

The results of the ANOVA for the IMI combined with SCC effect on the investigated traits are reported in Table 2. Evaluating the interaction between the infection status and inflammation process allowed us to focus directly on the effect of subclinical IMI within each SCC classes, to better understand which could be the major forces of alteration of milk protein fractions.

At T1, the BACT x SCC interaction was associated with the total casein proportion ($P < 0.01$, Figure 3). Specifically, both positive and negative samples showed the highest proportion of casein with very low SCC (< 50,000 cell/mL), and then decreased across higher SCC classes, The behavior of negative and positive samples differed for the SCC class ≥ 400,000, with these latter having the lowest casein content (Figure 4). At T2, we observed an association of the BACT x SCC with α_{s1}-CN. In positive samples we observed an increase in α_{s1}-CN from the first SCC class to the second one (≥ 50,000 and < 200,000 cells/mL) but then a reduction, reaching the lowest value with SCC > 400,000 cells/mL.

These results suggest that the effect of IMI on milk caseins seems to be more evident at very high SCC value, probably linked to a synergic effect of infection and inflammation. In fact, aside from the pathogen release of exogenous enzymes, some bacteria are able to induce the production of protease activators, which by enhancing the proteolysis in the mammary gland, release important substrates for the bacterial growth (Mehrzad et al., 2005; Kelly et al., 2006).

These results further support that the level of inflammation (i.e., SCC) strongly affects the milk protein profile, which is known to be a factor of interest for the dairy sector, given the crucial role that protein fractions play in determining the milk cheesemaking aptitude. It is however worth

remembering that IMI, especially when subclinical, is not always associated with increased SCC. Indeed, cases of animals positive at the microbiological examination but with low SCC (< 200,000 cells/mL) have been reported, as well as culture-negative samples with high SCC (Bobbo et al., 2017; Pegolo et al., 2022). This can be due to different reasons, including the stage of infection, the animals' immune response to infection and/or the pathogens' infection patterns. For instance, *Staphylococcus aureus* can induce chronic infections with intermittent and cyclical shedding (Sears et al., 1990).

CONCLUSIONS

In conclusion, this study added new insights on the behavior of milk protein fractions in animals having subclinical IMI. Moreover, it explored for the first time and in an extensive way the pattern of alteration of the protein fractions at the quarter level. The IMI had a significant impact on mostly casein fractions, negatively affecting β-CN and αs1-CN proportions. However, we have substantiated that the increase in SCC seemed to be the factor associated with the most significant changes in the protein fractions. This suggests that the increase in endogenous proteolytic enzymes related to the onset of inflammation, and not the infection (i.e., the bacteriological status), might be the pivotal aspect driving the alteration of the milk protein profile. Finally, our results further confirmed the importance of SCC as an indicator of both udder health and milk quality.

ORCIDs

Vittoria Bisutti https://orcid.org/0000-0001-5681-4001

Alice Vanzin https://orcid.org/0000-0001-8215-5265

Sara Pegolo https://orcid.org/0000-0001-6390-9826

Alessandro Toscano https://orcid.org/0000-0003-3190-2608

Matteo Gianesella https://orcid.org/0000-0001-8117-5969

Enrico Sturaro https://orcid.org/0000-0001-9508-5622

Stefano Schiavon https://orcid.org/0000-0002-5539-8947

Luigi Gallo https://orcid.org/0000-0002-8908-5105

Franco Tagliapietra https://orcid.org/0000-0002-0593-1600

Diana Giannuzzi https://orcid.org/0000-0003-2975-0385

Alessio Cecchinato https://orcid.org/0000-0003-3518-720X

ACKNOWLEDGMENTS

This study was part of the LATSAN project, funded by the Ministero dell'agricoltura, della sovranità alimentare e delle foreste (MASAF), Rome (Italy). The authors thank Sofia Ton of the Breeders Association of the Veneto Region (ARAV, Padova, Italy) for the analyses of milk composition. The PhD grant of Vittoria Bisutti is founded by CARIPARO foundation. The authors state no conflict of interest.

REFERENCES

Abdelmegid, S., D. Kelton, J. Caswell, and G. Kirby. 2020. Proteomic 2d-dige analysis of milk whey from dairy cows with staphylococcus aureus mastitis reveals overexpression of host defense proteins. Microorganisms 8:1–14. doi:10.3390/microorganisms8121883.

Addis, M.F., E.M. Maffioli, F. Ceciliani, G. Tedeschi, V. Zamarian, F. Tangorra, M. Albertini, R. Piccinini, and V. Bronzo. 2020. Influence of subclinical mastitis and intramammary infection by coagulase-negative staphylococci on the cow milk peptidome. J. Proteomics 226. doi:10.1016/j.jprot.2020.103885.

Amalfitano, N., C. Cipolat-Gotet, A. Cecchinato, M. Malacarne, A. Summer, and G. Bittante. 2019. Milk protein fractions strongly affect the patterns of coagulation, curd firming, and syneresis. J. Dairy Sci. 102:2903–2917. doi:10.3168/jds.2018-15524.

Amalfitano, N., G. Stocco, A. Maurmayr, S. Pegolo, A. Cecchinato, and G. Bittante. 2020. Quantitative and qualitative detailed milk protein profiles of 6 cattle breeds: Sources of variation and contribution of protein genetic variants. J. Dairy Sci. 103:11190–11208. doi:10.3168/jds.2020-18497.

Antanaitis, R., V. Juozaitienė, V. Jonike, W. Baumgartner, and A. Paulauskas. 2021. Milk lactose as a biomarker of subclinical mastitis in dairy cows. Animals 11:1–11. doi:10.3390/ani11061736.

Baeker, R., S. Haebel, K. Schlatterer, and B. Schlatterer. 2002. Lipo- calin-type prostaglandin D synthase in milk: a new biomarker for bovine mastitis. Prostaglandins Other Lipid Mediat. 67:75–88.

Barsi, F., E. Carra, M. Ricchi, G. Gnali, G. Pisoni, S. Russo, A. Filippi, N. Arrigoni, R.N. Zadoks, and C. Garbarino. 2022. Circulation of Streptococcus agalactiae ST103 in a Free Stall Italian Dairy Farm. Appl. Environ. Microbiol. 88. doi:10.1128/aem.00383-22.

Baum, F., M. Fedorova, J. Ebner, R. Hoffmann, and M. Pischetsrieder. 2013. Analysis of the endogenous peptide profile of milk: Identification of 248 mainly casein-derived peptides. J. Proteome Res. 12:5447–5462. doi:10.1021/pr4003273.

Bisutti, V., S. Pegolo, D. Giannuzzi, L.F.M. Mota, A. Vanzin, A. Toscano, E. Trevisi, P. Ajmone Marsan, M. Brasca, and A. Cecchinato. 2022a. The β-casein (CSN2) A2 allelic variant alters milk protein profile and slightly worsens coagulation properties in Holstein cows. J. Dairy Sci.. doi:10.3168/jds.2021-21537.

Bisutti, V., A. Vanzin, A. Toscano, S. Pegolo, D. Giannuzzi, L. Gallo, E. Trevisi, R. Negrini, A. Cecchinato, F. Tagliapietra, and S. Schiavon. 2022b. Impact of somatic cell count combined with differential somatic cell count on milk protein fractions in Holstein cattle. J. Dairy Sci.. doi:10.3168/jds.2022-22071.

Bobbo, T., C. Cipolat-Gotet, G. Bittante, and A. Cecchinato. 2016. The nonlinear effect of somatic cell count on milk composition, coagulation properties, curd firmness modeling, cheese yield, and curd nutrient recovery. J. Dairy Sci. 99:5104–5119. doi:10.3168/jds.2015-10512.

Bobbo, T., P.L. Ruegg, G. Stocco, E. Fiore, M. Gianesella, M. Morgante, D. Pasotto, G. Bittante, and A. Cecchinato. 2017. Associations between pathogen-specific cases of subclinical mastitis and milk yield, quality, protein composition, and cheese-making traits in dairy cows. J. Dairy Sci. 100:4868–4883. doi:10.3168/jds.2016-12353.

Caggiano, N., A. Lorenzo Smirnoff, J.M. Bottini, and E.A. De Simone. 2019. Protease activity and protein profile in milk from healthy dairy cows and cows with different types of mastitis. Int. Dairy J. 89:1–5. doi:10.1016/j.idairyj.2018.07.020.

Chaneton, L., L. Tirante, J. Maito, J. Chaves, and L.E. Bussmann. 2008. Relationship between milk lactoferrin and etiological agent in the mastitic bovine mammary gland. J. Dairy Sci. 91:1865–1873. doi:10.3168/jds.2007-0732.

Cheng, J.B., J.Q. Wang, D.P. Bu, G.L. Liu, C.G. Zhang, H.Y. Wei, L.Y. Zhou, and J.Z. Wang. 2008. Factors affecting the lactoferrin concentration in bovine milk. J. Dairy Sci. 91:970–976. doi:10.3168/jds.2007-0689.

Cipolat-Gotet, C., A. Cecchinato, G. Bittante, M. Malacarne, and A. Summer. 2018. Variations in milk protein fractions affect the efficiency of the cheese-making process. J. Dairy Sci. 101:8788–8804. doi:10.3168/jds.2018-14503.

Considine, T., A. Healy, A.L. Kelly, and P.L.H. McSweeney. 2000. Proteolytic specificity of elastase on bovine as1-casein. Food Chem. 66:463–470. doi:10.1016/S0308-8146(99)00065-5.

Crudden, A., F. Fox, and A.L.K. Ã. 2005. Factors affecting the hydrolytic action of plasmin in milk 15:305–313. doi:10.1016/j.idairyj.2004.08.008.

Dal Prà, A., F. Biscarini, G.L. Cavani, S. Bacchelli, A. Iotti, S. Borghi, M. Nocetti, and P. Moroni. 2022. Relationship between total and differential quarter somatic cell counts at dry-off and early lactation. PLoS One 17:1–14. doi:10.1371/journal.pone.0275755.

Dallas, D.C., N.M. Murray, and J. Gan. 2016. Function within the Mammary Gland and the Infant. J. Mammary Gland Biol. Neoplasia 20:133–147. doi:10.1007/s10911-015-9334-3.Proteolytic.

Drago-Serrano, M.E., R. Campos-Rodríguez, J.C. Carrero, and M. Delagarza. 2017. Lactoferrin: Balancing ups and downs of inflammation due to microbial infections. Int. J. Mol. Sci. 18. doi:10.3390/ijms18030501.

Ezzat Alnakip, M., M. Quintela-Baluja, K. Böhme, I. Fernández-No, S. Caamaño-Antelo, P. Calo-Mata, and J. Barros-Velázquez. 2014. The Immunology of Mammary Gland of Dairy Ruminants between Healthy and Inflammatory Conditions. J. Vet. Med. 2014:1–31. doi:10.1155/2014/659801.

Guerrero, A., D.C. Dallas, S. Contreras, A. Bhandari, A. Islas-trejo, J.F. Medrano, E.A. Parker, M.

Wang, S. Chee, J.B. German, D. Barile, C.B. Lebrilla, O.S. Avenue, U. States, S. Avenue, U. States, O.S. Avenue, U. States, O.S. Avenue, U. States, D. Science, T. Group, F. Quality, and D. Group. 2016. Peptidomic analysis of healthy and subclinically mastitic bovine milk. Int. Dairy J. 46:46–52. doi:10.1016/j.idairyj.2014.09.006.Peptidomic.

Hauge, S., and J. K. Ellingsen. 1953. Selective agar medium (TKT Medium) for demonstration of group B streptococci in samples of producer milk. Nord. Vet. Med. 5:539–547

Hayton, A., J. Husband, and R. Vecqueray. 2012. Nutritional manage- ment of herd health. Page 248 in Dairy Herd Health, 1st ed. M. Green, ed. CABI North American Office, Cambridge, MA.

Hinz, K., L.B. Larsen, O. Wellnitz, R.M. Bruckmaier, and A.L. Kelly. 2012. Proteolytic and proteomic changes in milk at quarter level following infusion with Escherichia coli lipopolysaccharide. J. Dairy Sci. 95:1655–1666. doi:10.3168/jds.2011-4813.

Hogarth, C.J., J.L. Fitzpatrick, A.M. Nolan, F.J. Young, A. Pitt, and P.D. Eckersall. 2004. Differential protein composition of bovine whey: A comparison of whey from healthy animals and from those with clinical mastitis. Proteomics 4:2094–2100. doi:10.1002/pmic.200300723.

International Organization for Standardization. (2005). ISO 6888-2:2021. Microbiology of the food chain — Horizontal method for the enumeration of coagulase-positive staphylococci (Staphylococcus aureus and other species) — Part 2: Method using rabbit plasma fibrinogen agar medium.

Kelly, A.L., F. O'Flaherty, and P.F. Fox. 2006. Indigenous proteolytic enzymes in milk: A brief overview of the present state of knowledge. Int. Dairy J. 16:563–572. doi:10.1016/j.idairyj.2005.10.019.

L. Forsback, H. Lindmark-Ma˚nsson, A. Andre´n, and K. Svennersten-Sjaunja. 2010. Evaluation of quality changes in udder quarter milk from cows with low-to-moderate somatic cell counts. Animal 4:617–626. doi:10.1017/S1751731109991467.

de la Rosa, G., D. Yang, P. Tewary, A. Varadhachary, and J.J. Oppenheim. 2008. Lactoferrin Acts as an Alarmin to Promote the Recruitment and Activation of APCs and Antigen-Specific Immune Responses. J. Immunol. 180:6868–6876. doi:10.4049/jimmunol.180.10.6868.

Larsen, L.B., P.L.H. McSweeney, M.G. Hayes, J.B. Andersen, K.L. Ingvartsen, and A.L. Kelly. 2006. Variation in activity and heterogeneity of bovine milk proteases with stage of lactation and somatic cell count. Int. Dairy J. 16:1–8. doi:10.1016/j.idairyj.2005.01.009.

Mansor, R., W. Mullen, A. Albalat, P. Zerefos, H. Mischak, D.C. Barrett, A. Biggs, and P.D. Eckersall. 2013. A peptidomic approach to biomarker discovery for bovine mastitis. J. Proteomics 85:89–98. doi:10.1016/j.jprot.2013.04.027.

Maurmayr, A., A. Cecchinato, L. Grigoletto, and G. Bittante. 2013. Detection and quantification of αS1-, αS2-, β-, κ-casein, α-lactalbumin, β-lactoglobulin and lactoferrin in bovine milk by reverse-phase high- performance liquid chromatography. Agric. Conspec. Sci. 78:201–205.

McFadden, T.B., R.M. Akers, and A. V. Capuco. 1988. Relationship of Milk Proteins in Blood with Somatic Cell Counts in Milk of Dairy Cows. J. Dairy Sci. 71:826–834. doi:10.3168/jds.S0022-0302(88)79623-X.

Mehrzad, J., C. Desrosiers, K. Lauzon, G. Robitaille, X. Zhao, and P. Lacasse. 2005. Proteases involved in mammary tissue damage during endotoxin-lnduced mastitis in dairy cows. J. Dairy Sci. 88:211–222. doi:10.3168/jds.S0022-0302(05)72679-5.

Nagasawa, Y., Y. Kiku, K. Sugawara, N. Yabusaki, and K. Oono. 2020. Detection From Clinical Mastitis Milk by Colloidal Gold Nanoparticle-Based Immunochromatographic Strips 6:1–12. doi:10.3389/fvets.2019.00504.

National Research Council 2001. 2001. Nutrient Requirements of Dairy Cattle. National Academies Press, Washington, D.C.

Nayan, V., K. Singh, M.A. Iquebal, S. Jaiswal, A. Bhardwaj, C. Singh, T. Bhatia, S. Kumar, R. Singh, M.N. Swaroop, R. Kumar, S.K. Phulia, A. Bharadwaj, T.K. Datta, A. Rai, and D. Kumar. 2022. Genome-Wide DNA Methylation and Its Effect on Gene Expression During Subclinical Mastitis in Water Buffalo. Front. Genet. 13:1–14. doi:10.3389/fgene.2022.828292.

N.M.C. 2017. Laboratory Handbook on Bovine Mastitis. third. National Mastitis Council, Inc.

Nyman, A., K.P. Waller, T.W. Bennedsgaard, T. Larsen, and U. Emanuelson. 2014. Associations of udder-health indicators with cow factors and with intramammary infection in dairy cows. J. Dairy Sci. 97:5459–5473. doi:10.3168/jds.2013-7885.

Paixão, M.G., L.R. Abreu, R. Richert, and P.L. Ruegg. 2017. Milk composition and health status from mammary gland quarters adjacent to glands affected with naturally occurring clinical mastitis. J. Dairy Sci. 100:7522–7533. doi:10.3168/jds.2017-12547.

Pegolo, S., M. Momen, G. Morota, G.J.M. Rosa, D. Gianola, G. Bittante, and A. Cecchinato. 2020. Structural equation modeling for investigating multi-trait genetic architecture of udder health in dairy cattle 1–15. doi:10.1038/s41598-020-64575-3.

Pegolo, S., D. Giannuzzi, V. Bisutti, R. Tessari, M.E. Gelain, L. Gallo, S. Schiavon, F. Tagliapietra, E. Trevisi, P. Ajmone Marsan, G. Bittante, and A. Cecchinato. 2021. Associations between differential somatic cell count and milk yield, quality, and technological characteristics in Holstein cows. J. Dairy Sci. 104:4822–4836. doi:10.3168/jds.2020-19084.

Pegolo, S., R. Tessari, V. Bisutti, A. Vanzin, D. Giannuzzi, M. Gianesella, A. Lisuzzo, E. Fiore, A. Barberio, E. Schiavon, E. Trevisi, F. Piccioli Cappelli, L. Gallo, P. Ruegg, R. Negrini, and A. Cecchinato. 2022. Quarter-level analyses of the associations among subclinical intramammary infection and milk quality, udder health, and cheesemaking traits in Holstein cows. J. Dairy Sci. 105. doi:10.3168/jds.2021-21267.

Pegolo, S., A. Toscano, V. Bisutti, A. Vanzin, D. Giannuzzi, L. Gallo, F. Tagliapietra, P. Ruegg, M. Gianesella, M. Mele, S. Schiavon, and A. Cecchinato. 2023a. Associations between subclinical intramammary infection and milk fatty acid profile at the quarter level in Holstein cattle. animal 100978. doi:10.1016/j.animal.2023.100978.

Pegolo, S., D. Giannuzzi, F. Piccioli-Cappelli, L. Cattaneo, M. Gianesella, P. Ruegg, E. Trevisi, and A. Cecchinato. 2023b. Blood biochemical changes upon subclinical intramammary infection and inflammation in Holstein cattle. J. Dairy Sci. 106:6539–6550. doi:10.3168/jds.2022-23155.

Pisanu, S., T. Cubeddu, D. Pagnozzi, S. Rocca, C. Cacciotto, A. Alberti, G. Marogna, S. Uzzau, and M.F. Addis. 2015. Neutrophil extracellular traps in sheep mastitis. Vet. Res. 46. doi:10.1186/s13567-015-0196-x.

Ramos, T., F. Costa, I. Pinto, S. Pinto, and L. Abreu. 2015. Effect of Somatic Cell Count on Bovine Milk Protein Fractions. J. Anal. Bioanal. Tech. 6:1–7. doi:10.4172/2155-9872.1000269.

Rezamand, P., T.A. Hoagland, K.M. Moyes, L.K. Silbart, and S.M. Andrew. 2007. Energy Status , Lipid-Soluble Vitamins , and Acute Phase Proteins in Periparturient Holstein and Jersey Dairy Cows With or Without Subclinical Mastitis. J. Dairy Sci. 90:5097–5107. doi:10.3168/jds.2007-0035.

Le Roux, Y., F. Laurent, and F. Moussaoui. 2003. Polymorphonuclear proteolytic activity and milk composition change. Vet. Res. 34:629–645. doi:10.1051/vetres:2003021.

Ruegg, P.L. 2017. A 100-Year Review: Mastitis detection, management, and prevention. J. Dairy Sci. 100:10381–10397. doi:10.3168/jds.2017-13023.

Sears, P.M., B.S. Smith, P.B. English, P.S. Herer, and R.N. Gonzalez. 1990. Shedding Pattern of Staphylococcus aureus from Bovine Intramammary Infections. J. Dairy Sci. 73:2785–2789. doi:10.3168/jds.S0022-0302(90)78964-3.

Shimazaki, K.I., and K. Kawai. 2017. Advances in lactoferrin research concerning bovine mastitis. Biochem. Cell Biol. 95:69–75. doi:10.1139/bcb-2016-0044.

Sora, V.M., S. Panseri, M. Nobile, F. Di Cesare, G. Meroni, L.M. Chiesa, and A. Zecconi. 2022. Milk Quality and Safety in a One Health Perspective: Results of a Prevalence Study on Dairy Herds in Lombardy (Italy). Life 12:1–10. doi:10.3390/life12060786.

Stocco, G., A. Summer, C. Cipolat-Gotet, L. Zanini, D. Vairani, C. Dadousis, and A. Zecconi. 2020. Differential somatic cell count as a novel indicator of milk quality in dairy cows. Animals 10:1–15. doi:10.3390/ani10050753.

Timonen, A.A.E., J. Katholm, A. Petersen, K. Mõtus, and P. Kalmus. 2017. Within-herd prevalence of intramammary infection caused by Mycoplasma bovis and associations between cow udder health, milk yield, and composition. J. Dairy Sci. 100:6554–6561. doi:10.3168/jds.2016-12267.

Urech, E., Z. Puhan, and M. Schällibaum. 1999. Changes in milk protein fraction as affected by subclinical mastitis. J. Dairy Sci. 82:2402–2411. doi:10.3168/jds.S0022-0302(99)75491-3.

Wedholm, A., H.S. Møller, H. Lindmark-Månsson, M.D. Rasmussen, A. Andrén, and L.B. Larsen. 2008. Identification of peptides in milk as a result of proteolysis at different levels of somatic cell counts using LC MALDI MS/MS detection. J. Dairy Res. 75:76–83. doi:10.1017/S0022029907002968.

Zhang, L., S. Boeren, A.C.M. Van Hooijdonk, J.M. Vervoort, and K.A. Hettinga. 2015. A proteomic perspective on the changes in milk proteins due to high somatic cell count. J. Dairy Sci. 98:5339–5351. doi:10.3168/jds.2014-9279.

TABLES AND FIGURES

Table 1. Descriptive statistics of the milk protein fractions expressed both quantitatively (g/L) and qualitatively (%N)

Item[1]	N	Mean	SD	P1[2]	P99[2]
Protein composition, g/L of milk					
True protein	528	33.60	4.04	24.43	43.09
Caseins	528	27.92	3.28	19.98	35.07
κ-CN	528	5.49	1.40	3.27	9.58
Glycosylated κ-CN	528	2.13	0.86	1.03	5.33
Non glycosylated κ-CN	528	3.36	1.09	1.56	6.87
αs2-CN	528	3.36	0.67	1.98	5.07
αs1-CN	528	9.26	1.42	6.31	13.00
β-CN	527	9.83	2.19	5.30	14.50
Whey proteins	528	5.68	1.33	3.41	10.20
β-LG	528	4.47	1.19	2.32	8.30
α-LA	528	1.04	0.25	0.61	2.08
LF	526	0.16	0.19	0.03	0.97
N compounds					
Urea	524	0.22	0.05	0.12	0.36
Protein composition, % of total milk N					
True protein	503	93.20	2.65	87.51	99.17
Caseins	528	77.81	1.13	74.05	79.73
κ-CN	528	15.24	3.07	10.18	24.78
Glycosylated κ-CN	528	5.87	1.98	2.95	12.65
Non glycosylated κ-CN	528	9.37	2.73	4.23	17.82
αs2-CN	528	9.36	1.53	5.63	13.50
αs1-CN	528	25.89	3.27	18.92	32.93
β-CN	527	27.35	5.17	16.51	36.91
Whey proteins	528	15.84	3.34	10.40	27.61
β-LG	527	12.47	3.06	7.01	21.18
α-LA	528	2.91	0.67	1.74	5.04
LF	526	0.44	0.49	0.10	2.52
N compounds					
Urea	483	1.55	0.63	0.47	3.12

[1] κ-CN: sum of glycosylated κ-CN and carbohydrate-free κ-CN; Caseins= sum of κ-CN + α_{s2}-CN + α_{s1}-CN + β-CN; Whey proteins= sum of β-LG + α-LA + LF; NPN compounds= sum of Urea and Minor NPN compounds.

[2] P1 = 1st percentile; P99 = 99th percentile.

Table 2. F-values and significances of the linear mixed model tested for quarter-level protein fractions at sampling time 1 (T1) and time 2 (T2).

Item[1]	T1			T2		
	BACT[2]	SCC[3]	SCC*BACT[4]	BACT[2]	SCC[3]	SCC*BACT[4]
Protein composition, g/L of milk						
True protein	0.66	0.73	0.33	1.39	0.48	0.24
Caseins	0.00	1.91	2.16	0.61	0.97	0.67
κ-CN	5.94*	18.67***	1.00	0.19	8.86***	0.27
Glycosylated κ-CN	2.24	3.52*	1.51	0.01	9.99***	0.33
Carbohydrate-free κ-CN	3.32	15.56***	2.85*	0.97	5.06**	0.23
αs2-CN	9.15**	3.96**	0.48	1.20	6.41***	1.19
αs1-CN	0.77	5.35**	0.22	0.55	3.13*	3.53*
β-CN	6.97**	4.05**	1.33	0.34	3.53*	0.31
Whey proteins	3.58	1.94	0.73	0.03	0.12	0.63
β-LG	2.92	2.69*	1.03	0.44	0.20	1.03
α-LA	1.49	1.97	0.30	0.35	0.51	2.29
LF	1.72	19.24***	2.11	3.92*	5.77**	1.58
N compounds						
Urea	0.96	10.98***	0.35	3.22	8.26***	0.44
Protein composition, % of total milk N						
True protein	1.54	1.86	0.93	0.27	5.24**	0.95
Caseins	4.81*	27.64***	4.58**	0.13	27.57***	0.72
κ-CN	5.95*	12.75***	0.87	0.03	8.68***	0.79
Glycosylated κ-CN	1.68	2.83*	1.84	0.03	9.13***	0.53
Carbohydrate-free κ-CN	3.59	11.04***	2.20	0.67	4.25**	0.62
αs2-CN	9.47**	2.35	0.43	0.72	2.78*	0.79
αs1-CN	0.11	9.76***	0.12	1.93	15.28***	3.74*
β-CN	10.36**	5.19**	1.26	0.02	9.51***	0.24
Whey proteins	1.54	1.87	1.22	0.18	0.49	0.59
β-LG	1.75	3.15*	1.64	0.70	0.67	0.81
α-LA	0.61	1.49	0.33	0.49	0.51	2.37
LF	0.55	16.90***	1.19	4.38*	5.91***	1.83
N compounds						
Urea	1.18	18.21***	1.10	1.40	5.98***	0.47

[1] κ-CN: sum of glycosylated κ-CN and carbohydrate-free κ-CN; Caseins= sum of κ-CN + α_{s2}-CN + α_{s1}-CN + β-CN; Whey proteins= sum of β-LG + α-LA + LF; NPN compounds= sum of Urea and Minor NPN compounds.

[2] BACT: Bacteriological status (negative, positive).

[3] SCC: Somatic cell count, expressed in classes (4 classes: class 1: SCC < 50,000 cells/mL, 50,000 cells/mL ≤ class 2 < 200,00 cells/mL; 200,00 cells/mL≤ class 3 < 400,000 cells/mL; class 4 ≥ 400,00 cells/mL.

[4] BACT x SCC: Interaction between the bacteriological status and the somatic cell count effects.

* $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$.

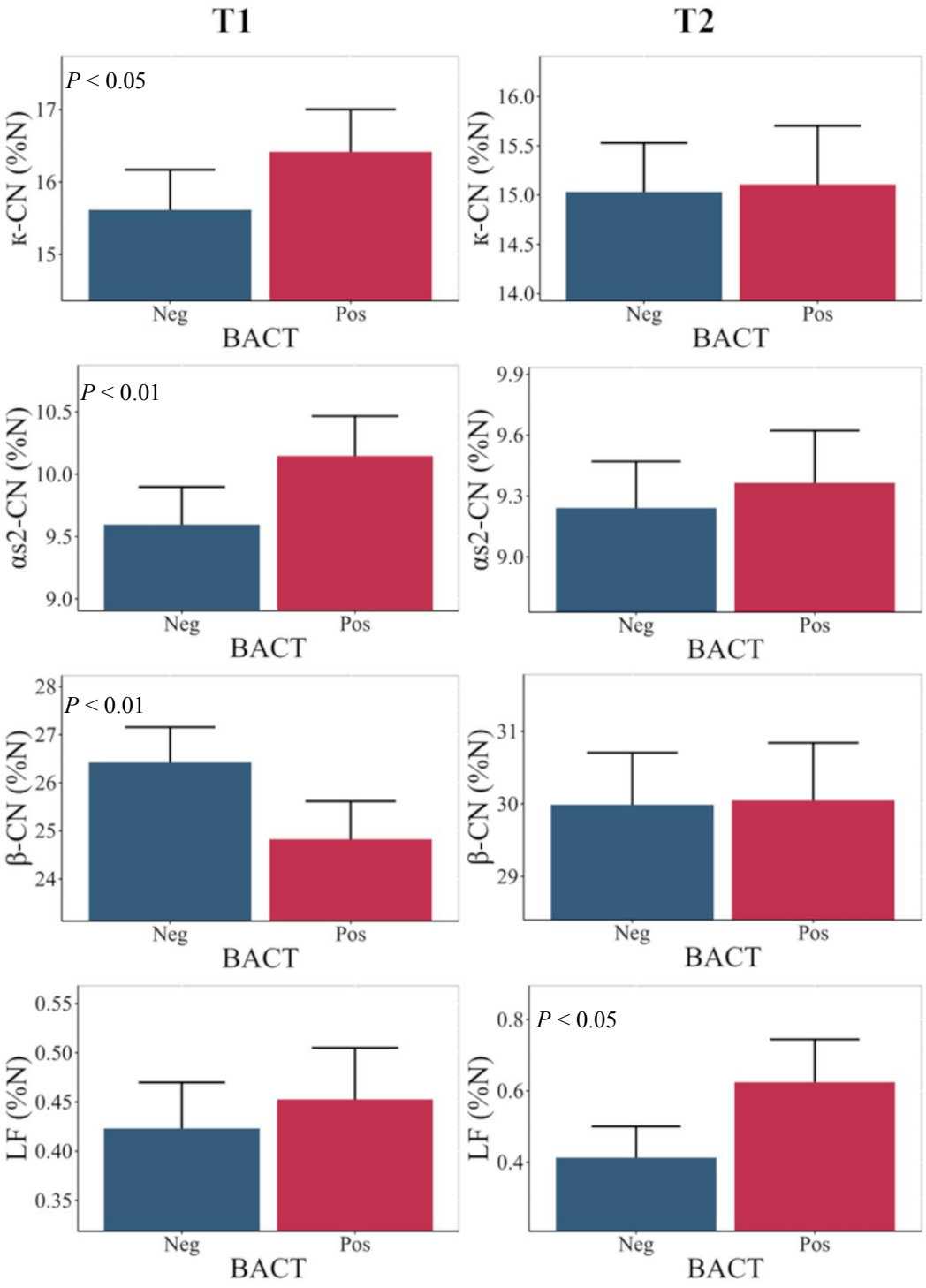


Figure 1. Least square means and standard error (SE) of the bacteriological status on the protein fractions expressed qualitatively (%N) and measured at the quarter level, at T1 and T2, respectively. BACT classes were defined as: negative at bacteriological examination (neg) and positive at bacteriological examination (pos). Only significant ($P < 0.05$) traits for at least one sampling time were displayed.

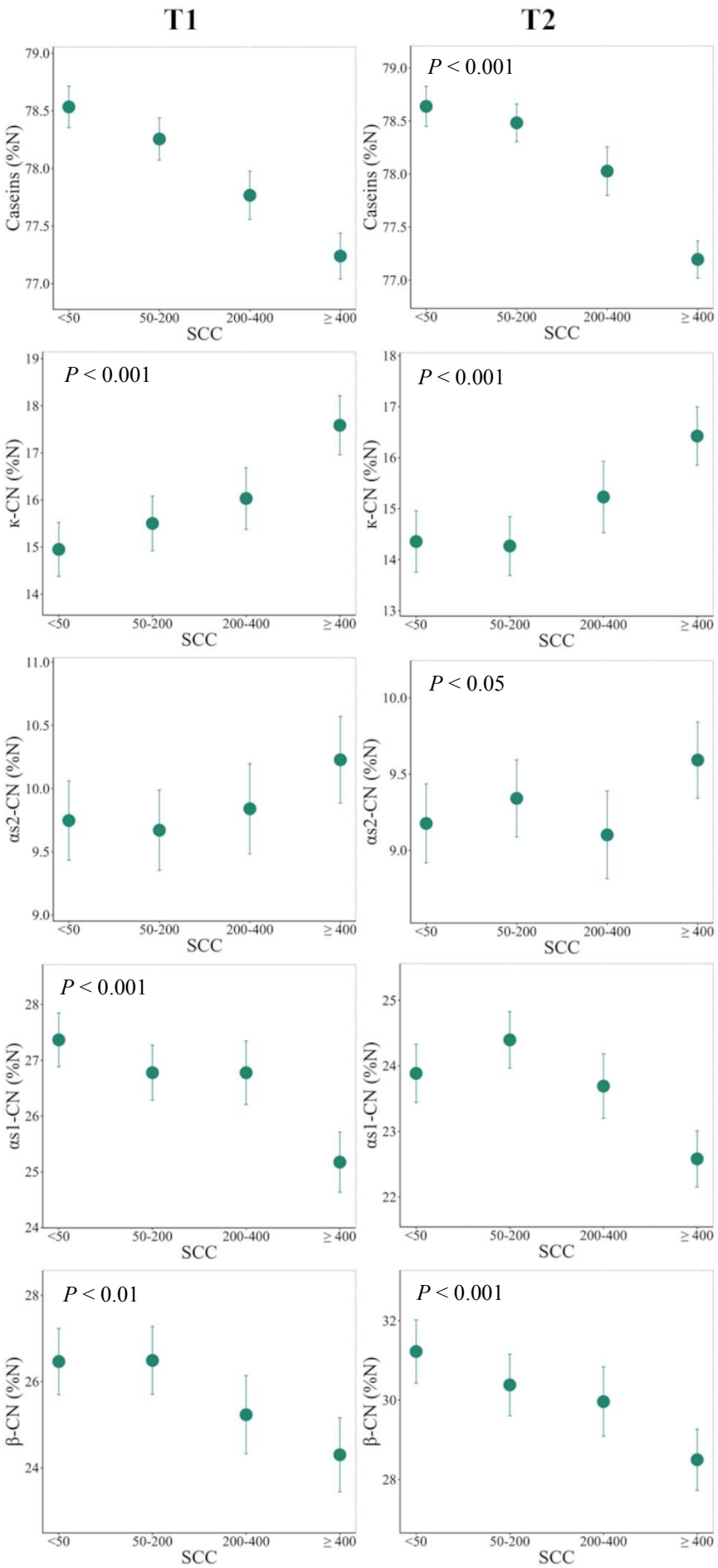

T1
T2
Caseins (%N)
κ-CN (%N)
αs2-CN (%N)
αs1-CN (%N)
β-CN (%N)
SCC
<50
50-200
200-400
≥ 400
P < 0.001
P < 0.001
P < 0.05
P < 0.001
P < 0.01
P < 0.001

Figure 2. Least square means and standard error (SE) of the SCC classes on the casein fractions expressed qualitatively (%N) and measured at the quarter level, at T1 and T2, respectively. Only significant traits ($P < 0.05$) for at least one sampling time were displayed.

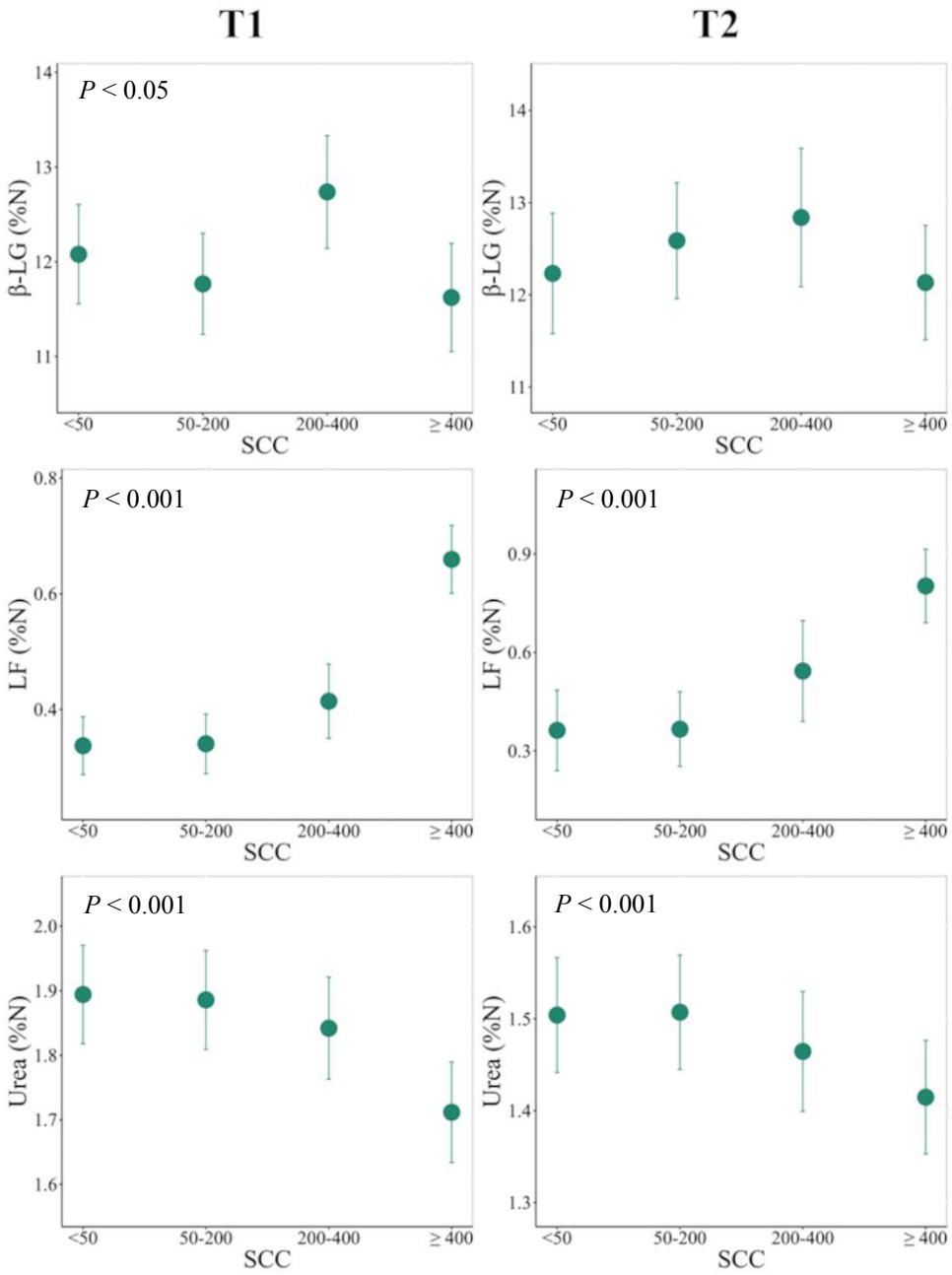


Figure 3. Least square means and standard error (SE) of the SCC classes on the whey protein fractions and milk urea expressed qualitatively (%N) and measured at the quarter level, at T1 and T2, respectively. Only significant traits ($P < 0.05$) for at least one sampling time were displayed.

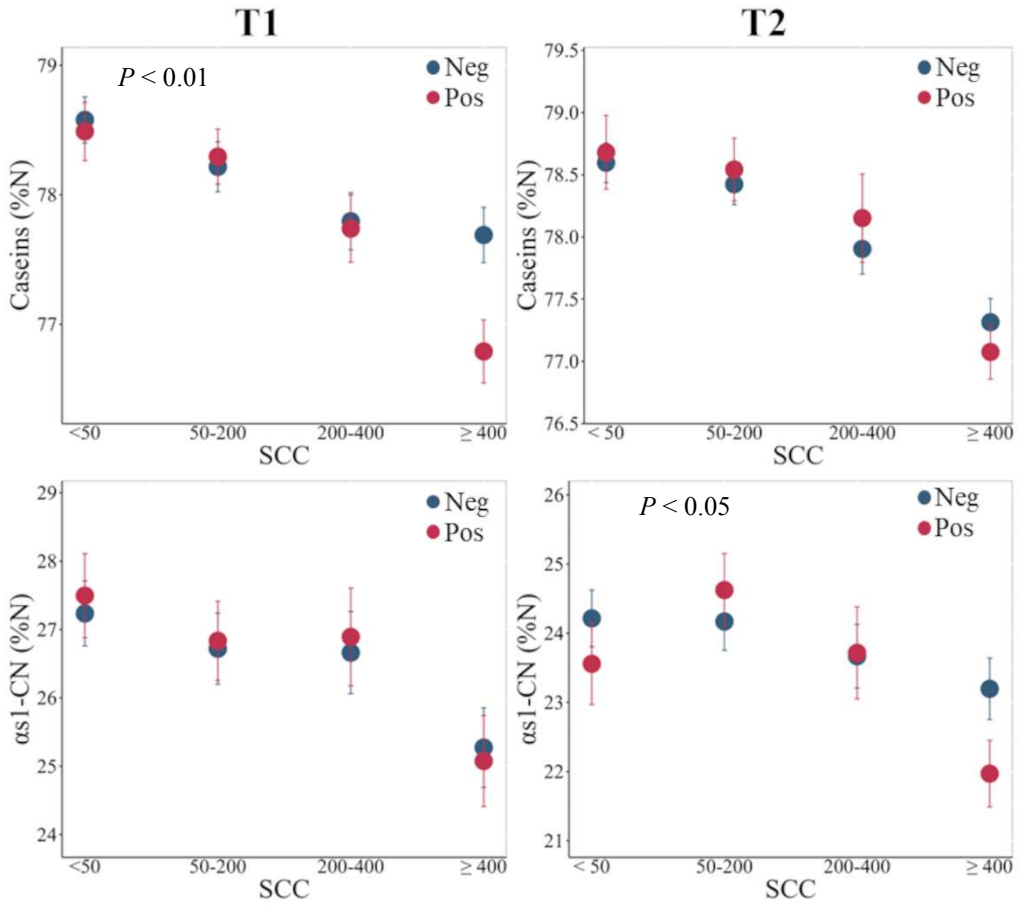


Figure 4. Least square means and standard error (SE) of the bacteriological status (BACT) × somatic cell count (SCC) interaction on the protein fractions expressed qualitatively (%N) and measured at the quarter level, at T1 and T2, respectively. Only significant traits ($P < 0.05$) for at least one sampling time were displayed.

SUPPLEMENTARY INFORMATION

Supplementary Table S1. Feed ingredients and main nutritional values of the rations used in the three herds.

	Herd A	Herd B	Herd C
Feed ingredients (kg DM/d)			
Corn silage	4.0	7.0	4
Sorghum silage	3.0		2
Meadow hay, mid maturity	3.0	3.0	4
Barley grain	3.0		
Corn grain	3.0	8.0	6.5
Soybean meal solvent (44% crude protein)	3.0	2.5	2.8
Mineral – vitamin mixture	1.0	0.8	1
Total	20.0	21.3	20.3
Nutritional values:			
Net energy, MJ/kg DM	6.48	6.92	6.59
Crude protein, g/kg	160	142	151
Ether extract g/kg	27	32	30
NDF, g/kg	337	282	313
ADF, g/kg	202	169	191
Ash, g/kg	50	40	48
Ca, g/kg	8.4	8.2	8.1
P, g/kg	4.6	4.4	4.5

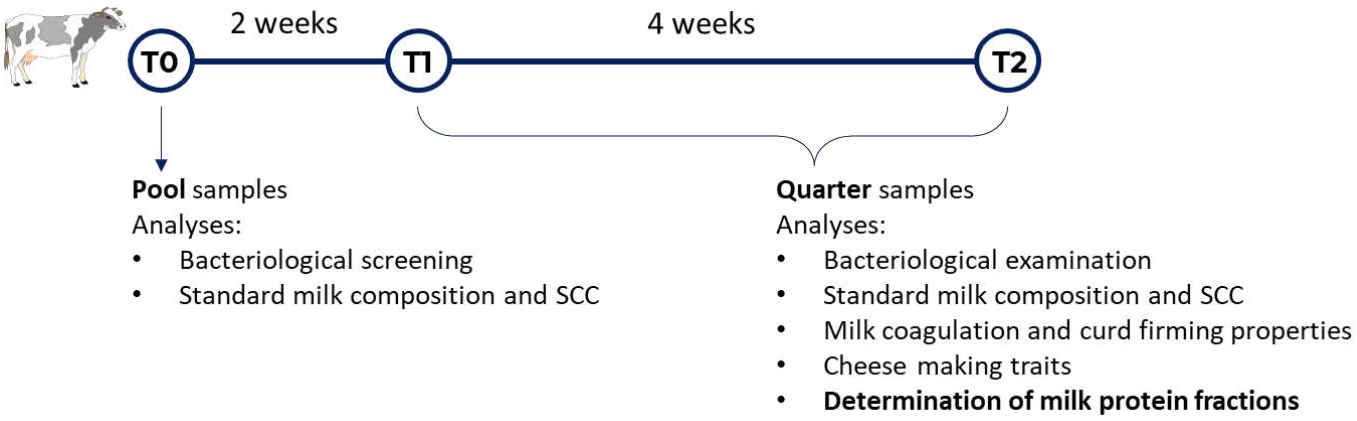
The rations characteristics were provided by the extension service of the farmers. Data were based on chemical analysis for the forages produced in the farms, and on tabular values (NRC, 2001) for the concentrates From: Pegolo et al. (2023b).

Supplementary Table S2. Microbiological results of quarter milk samples collected at time 1 (T1) and time 2 (T2).

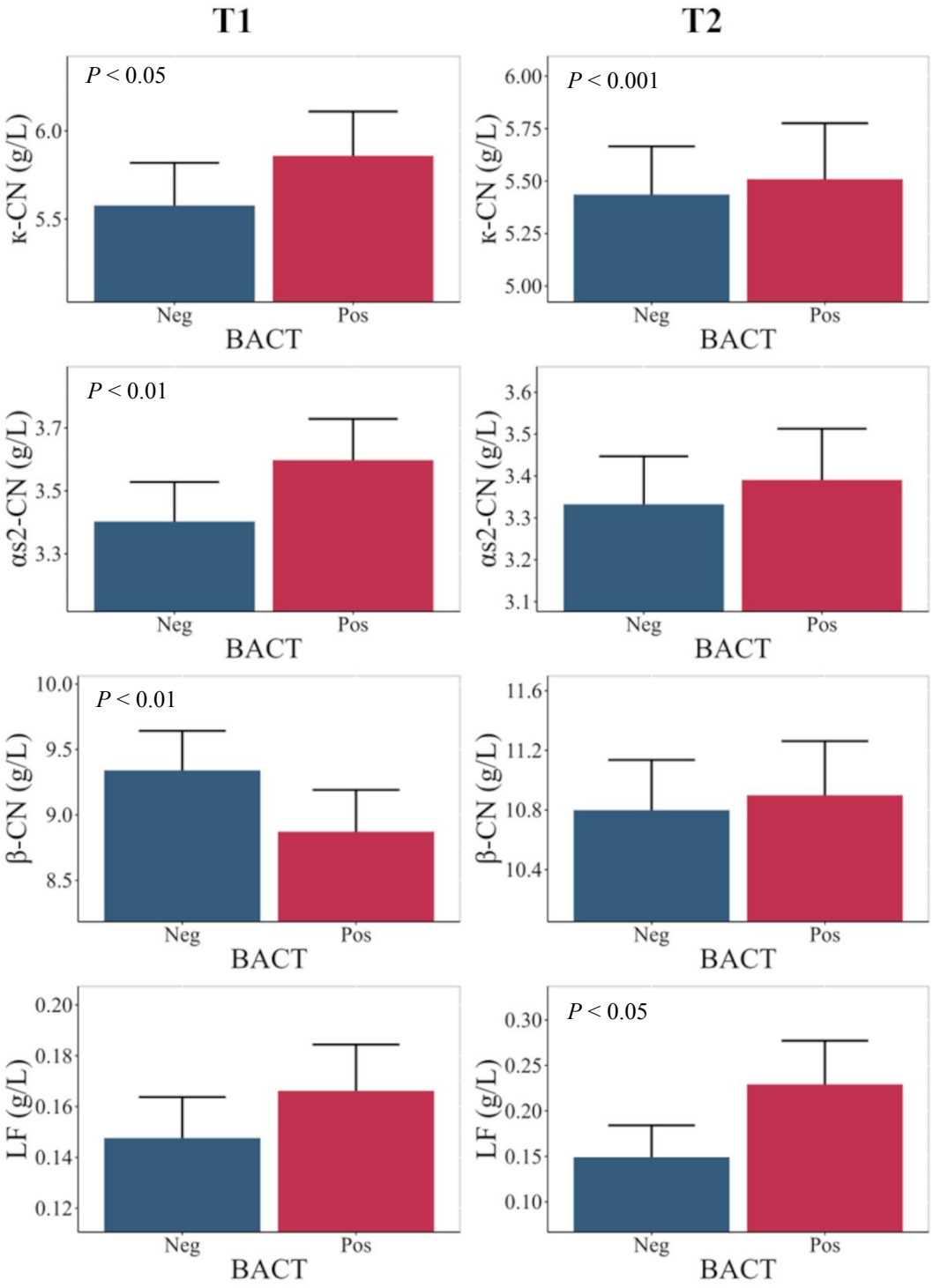
Microbiological results	T1		T2	
	N	%[1]	N	%[1]
S. aureus	4	1.6	3	1.1
Strep. agalactiae	41	16.2	34	12.3
Prototheca spp.	15	5.9	8	3.3
Strep. uberis	3	1.2	1	0.4
Total positive quarter samples	63	24.9	47	17.02
Contaminated samples[2]	8	3.2	11	4.0
Total quarter samples	253	-	276	-

[1]Percentage of samples on the total quarter samples analyzed according to the bacteriological results.

[2]Samples contaminated by environmental microorganisms (*Acinetobacter* spp.*, Bacillus* spp.*, Corynebacterium* spp.*, E. coli, Enterococcus* spp.*, Klebsiella* spp.*, Proteus* spp.*, Staphylococcus* spp. *coagulase negative*)

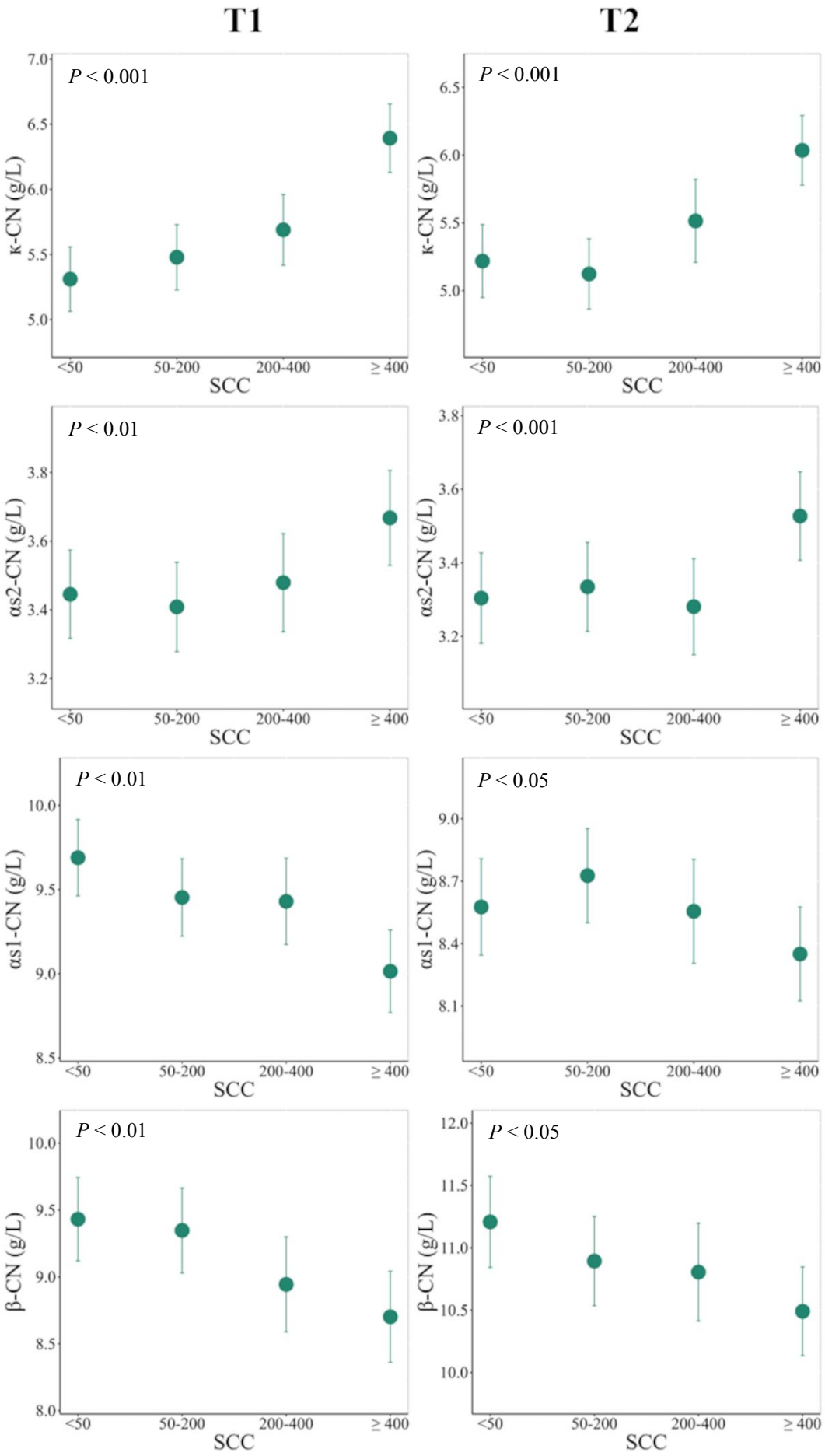


Supplementary figure S1. Schematic representation of the experimental design and sampling activity. T0: Time 0, T1: time 1, after two weeks from T0; T2: time 2, after four weeks from T1. SCC: somatic cell count. Modified from Pegolo et al., (2022).

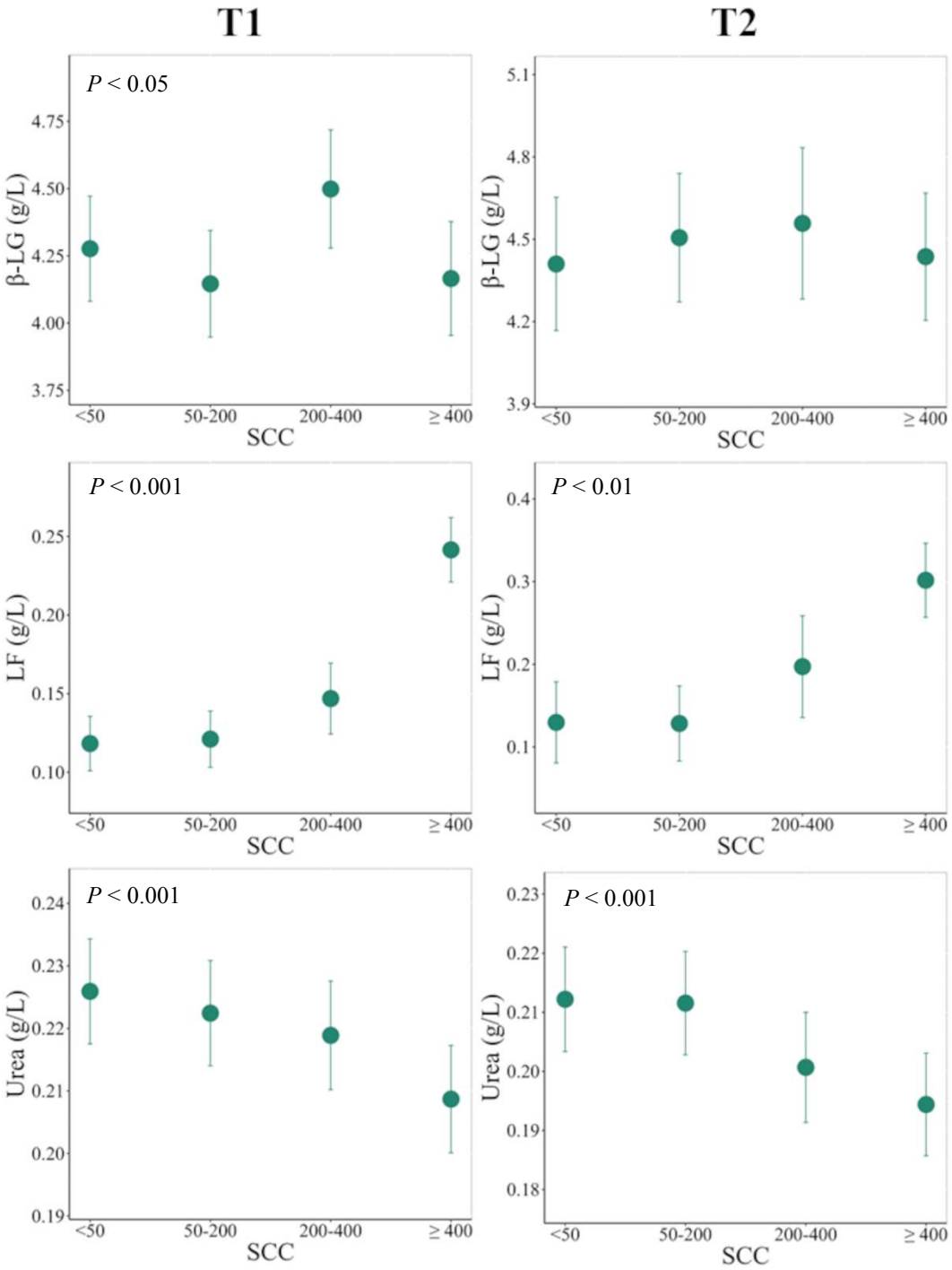


Supplementary Figure S2. Least square means and standard error (SE) of the bacteriological status on the protein fractions expressed quantitatively (g/L) and measured at the quarter level at T1 and T2, respectively. BACT classes were defined as: negative at bacteriological examination (neg) and positive at bacteriological examination (pos).

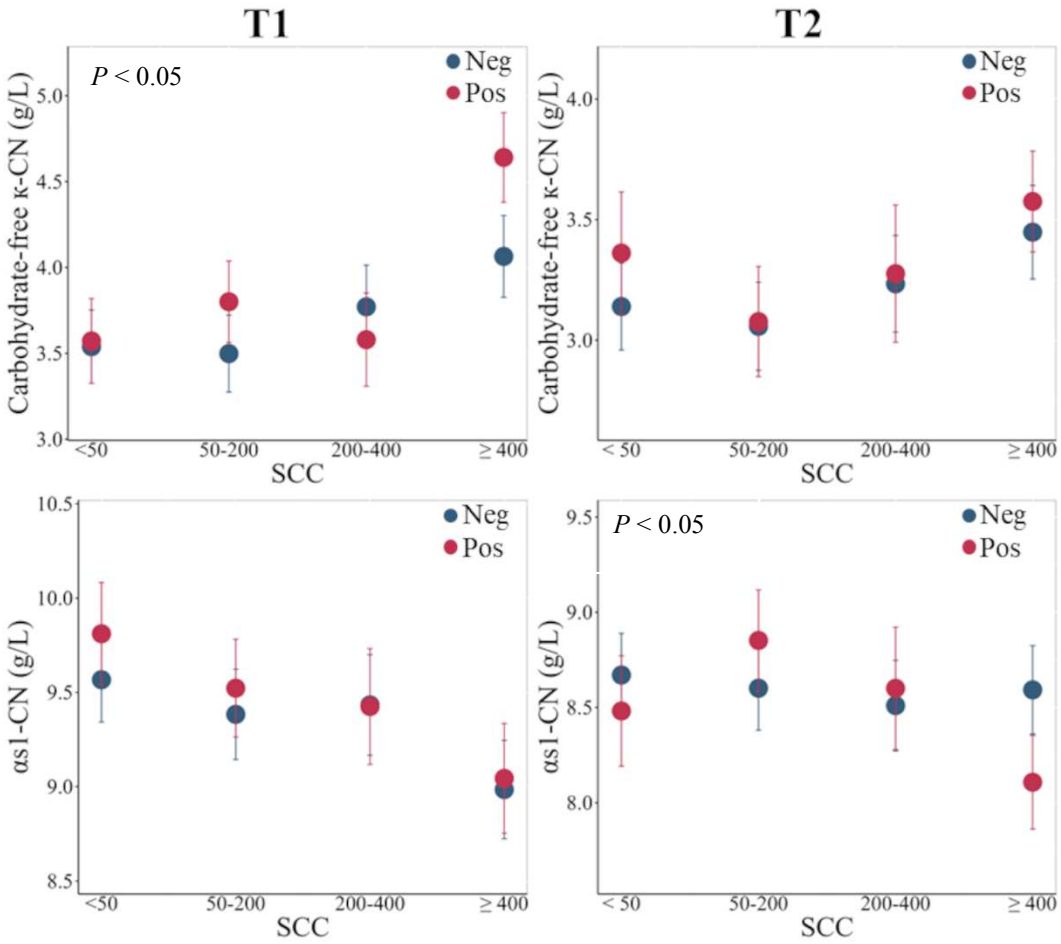
Only significant traits ($P < 0.05$) for at least one sampling time were displayed.


T1
T2
P < 0.001
P < 0.001
κ-CN (g/L)
7.0
6.5
6.0
5.5
5.0
6.5
6.0
5.5
5.0
<50
50-200
200-400
≥ 400
SCC
P < 0.01
P < 0.001
αs2-CN (g/L)
3.8
3.6
3.4
3.2
3.8
3.6
3.4
3.2
P < 0.01
P < 0.05
αs1-CN (g/L)
10.0
9.5
9.0
8.5
9.0
8.7
8.4
8.1
P < 0.01
P < 0.05
β-CN (g/L)
10.0
9.5
9.0
8.5
8.0
12.0
11.5
11.0
10.5
10.0

Supplementary Figure S3. Least square means and standard error (SE) of the SCC classes on the casein fractions expressed quantitatively (g/L) and measured at the quarter level at T1 and T2, respectively. Only significant traits ($P < 0.05$) for at least one sampling time were displayed.



Supplementary Figure S4. Least square means and standard error (SE) of the SCC classes on the whey protein fractions and milk urea expressed quantitatively (g/L) and measured at the quarter level at T1 and T2, respectively. Only significant traits ($P < 0.05$) for at least one sampling time were displayed.



Supplementary Figure S5. Least square means and standard error (SE) of the bacteriological status (BACT) × somatic cell count (SCC) interaction on the protein fractions expressed quantitatively (g/L) and measured at the quarter level at T1 and T2, respectively. Only significant traits ($P < 0.05$) for at least one sampling time were displayed.

CHAPTER III

Transcriptome-wide mapping of milk somatic cells upon subclinical mastitis infection in dairy cattle

Vittoria Bisutti[1†*], Núria Mach[2†], Diana Giannuzzi[1], Alice Vanzin[1], Emanuele Capra[3], Riccardo Negrini[4], Maria Elena Gelain[5], Alessio Cecchinato[1], Paolo Ajmone-Marsan[4], Sara Pegolo[1]

[1]*DAFNAE, University of Padova, viale dell'Università 16, Legnaro (PD), Italy*

[2]*IHAP, Université de Toulouse, INRAE, ENVT, 23 Chemin des Capelles, Toulouse, France*

[3]*IBBA, National Research Council, via Einstein, Lodi, Italy.*

[4]*DIANA, Università Cattolica del Sacro Cuore, via E. Parmense 84, Piacenza, Italy*

[5]*BCA, University of Padova, viale dell'Università 16, Legnaro (PD), Italy*

† Equally contributing authors

Published in: Journal of Animal Science and Biotechnology 93 (2023) 1-19

https://doi.org/10.1186/s40104-023-00890-9

ABSTRACT

Subclinical intramammary infection (IMI) represents a significant problem in maintaining dairy cows' health. Disease severity and extent depend on the interaction between the causative agent, environment, and host. To investigate the molecular mechanisms behind the host immune response, we used RNA-Seq for the milk somatic cells (SC) transcriptome profiling in healthy cows (n = 9), and cows naturally affected by subclinical IMI from *Prototheca* spp. (n = 11) and *Streptococcus agalactiae (S. agalactiae; n* = 11). Data Integration Analysis for Biomarker discovery using Latent cOmponents (DIABLO) was used to integrate transcriptomic data and host phenotypic traits related to milk composition, SC composition, and udder health to identify hub variables for subclinical IMI detection.

A total of 1,682 and 2,427 differentially expressed genes (DEGs) were identified when comparing *Prototheca* spp. and *S. agalactiae* to healthy animals, respectively. Pathogen-specific pathway analyses evidenced that *Prototheca*'s infection upregulated antigen processing and lymphocyte proliferation pathways while *S. agalactiae* induced a reduction of energy-related pathways like the tricarboxylic acid cycle, and carbohydrate and lipid metabolism. The integrative analysis of commonly shared DEGs between the two pathogens (n = 681) referred to the core-mastitis response genes, and phenotypic data evidenced a strong covariation between those genes and the flow cytometry immune cells (r^2 = 0.72), followed by the udder health (r^2 = 0.64) and milk quality parameters (r^2 = 0.64). Variables with r ≥ 0.90 were used to build a network in which the top 20 hub variables were identified with the Cytoscape cytohubba plug-in. The genes in common between DIABLO and cytohubba (n = 10) were submitted to a ROC analysis which showed they had excellent predictive performances in terms of discriminating healthy and mastitis-affected animals (sensitivity > 0.89, specificity > 0.81, accuracy > 0.87, and precision > 0.69). Among these genes, *CIITA* could play a key role in regulating the animals' response to subclinical IMI. Despite some differences in the enriched pathways, the two mastitis-causing pathogens seemed to

induce a shared host immune-transcriptomic response. The hub variables identified with the integrative approach might be included in screening and diagnostic tools for subclinical IMI detection.

Keywords: Data integration, Immune response, Milk somatic cells, RNA-sequencing, Subclinical mastitis

INTRODUCTION

Mastitis in dairy cattle is a well-established problem that firstly affects animal welfare but also hinders milk production and quality, leading to significant economic losses at the expense of farmers (Halasa et al., 2007). Mastitis typically occurs in response to the penetration of a wide range of microorganisms in the mammary gland. It exists in two forms: the clinical one, with overt signs of inflammation, udder swelling, and changes in milk physical composition, and the subclinical one, where usually the only alteration observed is the increase in the somatic cell count (SCC) derived from the proliferation and migration of the immune cells in the udder (Bradley, 2002). Subclinical mastitis is the most challenging form, estimated to be 15-40 times more frequent than its clinical counterpart (Martin et al., 2018). It represents a significant source of infection for the animals within the herds leading to an essential decrease in milk production (Ruegg, 2017) while going unnoticed. The most common causative agents of mastitis, accounting for over 80% of the infections (Ranjan et al., 2006), are *Escherichia coli* (*E. coli*), *Staphylococcus aureus* (*S. aureus*), *Streptococcus agalactiae (S. agalactiae)*, *Streptococcus uberis* (*S. uberis*) and *Streptococcus dysgalactiae* (*S. dysgalactiae*). However, in the last few decades, the spreading of other microorganisms, like microalgae of the genus *Prototheca,* has rapidly become an emerging threat for the dairy sector, mainly because nowadays, there is still no treatment effective towards this type of microorganism (Shave et al., 2021).

The different etiology of invading pathogens can trigger a diverse host immune response, consequently affecting the extent and outcome of the infection (Schukken et al., 2011). For example, Gram-negative bacteria like *E. coli* usually induce an intense stimulation of cytokine production, leading to fully activating both the local and generalized host-immune response (Jensen et al., 2013; Petzl et al., 2018), Gram-positive pathogens elicit a weaker immune reaction with usually no systemic repercussion (Günther et al., 2017). In addition, the immuno-cytofluorimetric study conducted in subclinical mastitis induced naturally by *S. agalactiae* and *Prototheca* (Pegolo et al., 2022b) highlighted a differential immune reaction between the two microorganisms, primarily directed

towards an innate response in the case of *S. agalactiae*, as opposed to the adaptive response triggered by *Prototheca* spp. Therefore, a deeper understanding of the pathogen-specific molecular mechanisms underlying the pathogenesis of mastitis and the induced immune response is pivotal for uncovering new ways of predicting the infection outcome and designing practical diagnostic and therapeutic tools for battling this costly disease.

In this context, RNA sequencing (RNA-Seq) represents a suitable tool for investigating the complexity of the host-pathogen interaction (Wang et al., 2020a). Different transcriptomic studies have evaluated the mammary gland (Günther et al., 2017; Wang et al., 2020a) or hepatic (Heimes et al., 2020) response to intra-mammary infection with different types of pathogens in cows. However, only a few studies evaluated the changes in somatic milk cells (SC) transcriptome (Asselstine et al., 2019; Niedziela et al., 2021) in response to subclinical intra-mammary infection (IMI), and, more importantly, to the best of our knowledge, no studies are currently available on the investigation of milk SC transcriptomic signature of *Prototheca* spp infection in dairy cattle. Finally, most of the previous transcriptomic studies were conducted using experimentally induced models of clinical mastitis (Wang et al., 2020b), while less information is available on naturally occurring subclinical mastitis (Cheng et al., 2021).

Work on identifying putative candidate genes associated with mastitis has already been carried out in genome-wide association studies (GWAS) and transcriptomic profiling (Welderufael et al., 2018; Asselstine et al., 2019; Bakhtiarizadeh et al., 2020). However, the concordance among these studies could be higher, highlighting the difficulty in identifying reliable and reproducible biomarkers for mastitis detection and mastitis resistance. In this context, integrating transcriptomic with phenomic information might represent the ultimate step not only to strengthen the information on the complexity of the molecular system by reinforcing complementary levels of knowledge but also to create more reliable prediction models (Naserkheil et al., 2022).

Therefore, this study aimed to i) evaluate the milk SC transcriptomic signatures upon natural infection of *S. agalactiae* and *Prototheca* spp., ii) integrate transcriptomic and phenomic information

to explain better the complexity underlying the molecular mechanisms of mastitis, and identify hub variables for early mastitis detection and prediction and, iii) perform a meta-analysis using three publicly available datasets to confirm the reproducibility of our results.

METHODS

Animal cohort, housing, and diet

Thirty-one multiparous Holstein cows (ranging from 3 to 7 years of age) between 98 and 448 days in milk were selected from a commercial herd of 450 lactating cows (Veneto region, Italy) regularly monitored for *S. agalactiae* and *Prototheca* spp between January and February 2021. Herd selection was based firstly on a prevalence study conducted by the *Istituto Zooprofilattico delle Venezie* (IZSVe) for the identification of the most common pathogen responsible for mastitis in the Veneto region and on ease of access to the farm location and the cooperation of the dairy farm owners and their associated veterinary practices. For successful participation in the study, we required the following criteria: (i) absence of clinical signs of infection; ii) no antibiotic treatment or anti-inflammatory medications before enrollment; (iii) being multiparous and non-pregnant; and (iv) having > 98 days in milk (DIM). Moreover, we required that animals used as negative control had no previous history of mastitis. Information was collected from the herd management software (Dairy Comp Sata, Alta Italia Srl, Milan, Italy). Based on these criteria, an initial bacteriological screening (time 0, T0) was performed on 188 lactating cows to identify healthy individuals and cows with subclinical mastitis from *S. agalactiae* or *Prototheca* spp. Animals with co-infection were excluded from the experiment. Moreover, cows with chronic mastitis cases (apparently healthy cows with lumps palpable in the udder and milk quality changes) were not enrolled. Following the bacteriological test results, we created three experimental groups from eligible animals: (i) healthy individuals (n = 9) with a negative bacteriological examination in all glands at T0 and time 1 (T1, two weeks after T0); (ii) naturally infected animals with *S. agalactiae* (n = 11) and (iii) naturally

infected animals with *Prototheca* spp. *(n* = 11*)*. At T1, a second bacteriological assessment was made on all the animals enrolled to confirm the bacteriological evaluation made at T0 (Table S1).

Cows were fed a total mixed ration formulated to meet or exceed the requirements of mid-lactation dairy cattle, mainly based on corn silage, sorghum silage, and concentrate. Feed was delivered once a day at 8:00, and the amount fed was adjusted daily to allow for a minimum of 5% refusals. Drinking water was available in automatic water bowls, and cows were milked twice daily, from 2:00 to 6:00, and from 14:00 to 18:00. Individual cow milk yield was recorded at each milking using herd software.

Animal health was managed by the farmers and local veterinarians, who intervened when needed. All cows were subjected to the same management practices and environment to ensure sample homogeneity.

Ethical statement

This study was part of the LATSAN project that aimed to develop innovative tools for evaluating and studying mammary gland health and improving dairy cows' nutritional milk quality and coagulation properties. The research was approved by the Ethical Animal Care and Use Committee (OPBA - Organismo Preposto al Benessere degli Animali) of the Università Cattolica del Sacro Cuore and by the Italian Ministry of Health (protocol number 510/2019-PR of 19/07/2019).

Milk sample collection

Before morning milking, ~ 200 mL of milk from all quarters (pool sample) was aseptically collected from each animal according to the National Mastitis Guidelines ((NMC), 2017). Briefly, teat ends were externally cleaned with commercial pre-milking disinfectant, dried with individual towels, and then washed again with alcohol 70%. Composite milk of the four glands was then collected after discarding the first streams of foremilk from each quarter and stored at 4 °C before microbiological analysis. Four milk aliquots (~ 50 mL) of each milk sample were collected and gently mixed separately into sterile tubes for analysis as follows: (i) microbiological analysis; (ii) evaluation of milk composition, SCC, and differential somatic cell count (DSCC) measurement; (iii) milk flow

cytometry analysis, and (iv) RNA extraction and transcriptomic analysis. All the samples were immediately refrigerated at 4°C to minimize the metabolic activity of cells and enzymes and keep the bacteriological composition as stable as possible. Samples were transported under refrigerated conditions (4°C) to the different laboratories.

Microbiological analysis

Microbiological examination of milk samples was conducted at the IZSVe laboratories (Legnaro, PD, Italy). After reception (within 4h after sample collection), samples were frozen and analyzed within 3 d. Pegolo et al. (2022b) reported specifics of the microbiological analyses in detail. Briefly, 10 µL of every composite sample were inoculated in each of the following selective media: i) Baird Parker agar with rabbit plasma fibrinogen (BP-RPF; Biokar Diagnostics, Beauvais, France), ii) tallium kristalviolette tossin agar (TKT; IZSVe internal production), and iii) *Prothoteca* isolation medium (PIM; IZSVe internal production). Suspected colonies of *S. agalactiae* were confirmed using the Christie– Atkins–Munch-Peterson test (NMC, 2017) after 24h of incubation. At the same time, *Prototheca* isolation medium plates were observed at 24, 48, and 72 h, and the wet mount method confirmed suspected colonies (NMC, 2017).

Milk composition and quality traits

Milk composition (protein, casein, lactose, fat, and urea content), milk conductivity (mS/cm), and milk pH analysis were carried out on fresh samples using an FT6000 Milkoscan infrared analyzer (Foss A/S, Hillerød, Denmark). SCC and DSCC were measured through the Fossomatic 7 DC analyzer (Foss A/S).

Flow cytometry analysis

A 50 mL aliquot of each sample was immediately processed in the Comparative Biomedicine Department (BCA) cell laboratory of the University of Padova (Italy) for flow cytometry analysis. In all cases, analyses were performed within 12 h after sample collection with milk stored at 4ºC. The whole flow cytometry methodology and analysis are reported in the work of Pegolo et al. (2022b)

Briefly, for each sample, flow cytometry analysis was run in four tubes containing: 1) only cells (no antibodies; used as a negative control); 2) cd4pe-cd8alexa fluor 647; 3) cd11bfitc-cd14pe; and 4) cd45fitc-cd21pe-cd18alexa fluor 647. Flow cytometric analyses were performed using a CyFlow Space flow cytometer (Partec-Sysmex, Sysmex Europe GmbH, Norderstedt, Germany) fitted with a blue laser (488 nm), a red laser (635 nm) and a UV laser. The data were analyzed with the FlowMax software version 2.82 (Sysmex- Partec, Sysmex Europe GmbH, Norderstedt, Germany). The morphology and complexity of the cells were evaluated in an FSC vs. SSC dot plot; total white blood cells were identified as CD45 and CD18 positive events; polymorphonuclear cells as CD11b positive CD14 negative events; macrophages as CD11b and CD14 positive events; T-helper lymphocytes as CD4 positive and CD8 negative events; T cytotoxic lymphocytes as CD8 positive and CD4 negative events; B lymphocytes as CD45, CD21, and CD18 positive events. In this study, we considered for the statistical analyses only animals subjected to RNA-seq analyses which are part of the broader cohort of animals previously analyzed (Pegolo et al., 2022b).

The Kruskal Wallis and Dunn test assessed significance among the experimental groups for pairwise comparison of milk production, composition, and flow cytometry variables. The significance was set at $P < 0.05$.

RNA extraction from milk somatic cells

A 50 mL aliquot from each sample was first centrifuged at 2,000 g for 10 min at 4°C. The fat layer and the supernatant were discarded, and the cell pellet was then washed with 50 mL of PBS with ethylenediaminetetraacetic acid (EDTA) at 0.05 mmol/L, pH 7.2. Samples were then re-centrifuged at 1,500 for 10 min at 4°C, the supernatant discarded, and the pellet was re-suspended in 800 µL of Trizol (Invitrogen, Carlsbad, CA) and stored at -80 °C until the RNA extraction.

Total RNA was extracted from the Trizol reagent and purified using a NucleoSPin miRNA kit (Macherey-Nagel, Düren, Germany), following the combined protocol with TRIzol lysis with small and large RNA in one fraction (total RNA). RNA concentration and quality were determined by Agilent 2100 Bioanalyzer (Santa Clara, CA). Extracted RNA was stored at −80 °C until use.

Library preparation

The 31 RNA samples, including control (n = 9), positive for *S. agalactiae* (n = 11), and positive for *Prototheca* (n = 11), were subsequently sent on dry ice to the Nuova Genetica Italiana (NGI, Como, Italy) facility for library preparation and sequencing. MGIEasy rRNA Depletion kit V1.1 (MGI Tech Co., Ltd., Shenzen, China) was used to remove ribosomal rRNA and maximise unique sequencing reads. RNA-seq libraries were then prepared from 500 ng of total RNA using the MGIEasy RNA Library Prep Set V3.1 (MGI TechCo., Ltd, Shenzen, China), according to the manufacturer's protocol. RNA-seq experiments were performed on a DNBSEQ-G400 high throughput machine (MGI Tech Co) using a paired-end approach using the DNBSEQ-G400 sequencing kit (MGI TechCo., Ltd., Shenzen, China).

RNA-seq data processing and analysis

Data pre-processing was made following the consensus pipeline built by Overbey et al. (2021). First, the quality control of RNA sequences was assessed with the FastQC software (v. 0.11.9). Clean reads were obtained by removing low-quality bases and adaptors with the TrimGalore software (v. 0.6.4) (Krueger, 2019). FastQC was used again on the trimmed sequences to check the quality of the reads. MultiQC package (v.1.8) (Ewels et al., 2016) was run to create summary statistics reports that included the sample quality control result categories from FastQC across all experiment samples. The sample information of clean data is shown in Table S2.

The paired-end clean reads were aligned against the *Bos taurus* DNA reference genome (ARS-UCD1.2) from the USDA's Agricultural Research Service with the splice-aware STAR (v.2.7.3a) (Dobin et al., 2013). The genome indexing was performed using ARS-UCD1.2 as the reference FASTA and the Ensembl gene annotation file (Bos_taurus.ARS-UCD1.2.106.gtf.gz; http://ftp.ensembl.org/pub/release-106/gtf/bos_taurus/).

We subsequently used RSEM (v.1.3.3) (Li and Dewey, 2011) to quantify gene expression. Similar to STAR, RSEM was run in two distinct phases. The first phase used the reference genome

and GTF files to prepare indexed genome files. The second phase used the indexed files and the mapped reads from STAR to assign counts to each gene.

Gene expression evaluation and differential expression analysis

Counts filtering, data normalization, and differential expression analysis were performed in R studio (R v.4.1.2, R studio v. 1.4.1103). Only protein-coding genes were considered for the analysis.

We first normalized the transcriptome count matrix with the sequencing depth for each sample by calculating counts per million (CPM). We filtered out genes expressed in less than 10 samples with CPM < 0.5 using the edgeR package (v. 3.36.0) (Robinson et al., 2010). Genes failing these criteria were removed before the exploration and differential expression analysis.

Exploratory analysis of the expressed genes matrix was performed using unsupervised principal component analysis (PCA) and the multidimensional scaling (MDS) analysis after the regularized-logarithm transformation (edgeR) or variance stabilizing transformation in DESeq2 (v. 1.34.0) (Love et al., 2014).

Differentially expressed gene (DEG) analysis was performed pairwise using DESeq2 and edgeR packages: i) negative animals vs. *S. agalactiae* naturally infected animals; ii) uninfected subjects vs. *Prototheca* infected animals, and iii) *S. agalactiae* vs. *Prototheca* infected individuals. Then, the *voom()* function from the limma R package (v.3.50.0) was used to fit a generalized linear regression model to correct the data with the group as a fixed effect. The group factor and the cow dependency were included in the generalized linear model using the *nbinomWaldTest()* function, which estimates and tests the significance of regression coefficients with the following explicit parameter settings: betaPrior = FALSE, maxit = 5,000, useOptim = TRUE, useT = FALSE, useQR = TRUE, minmu = 0.5. The *P*-values were adjusted for multiple testing using the Benjamini and Hochberg procedure (FDR, false discovery rate).

Only DEGs with an adjusted *P*-value < 0.05 and shared between DESeq2 and edgeR approaches were used for the downstream pathway analysis.

Functional pathway analysis

The shared list of DEGs for each comparison was fed to the Cytoscape (v. 3.9.1, http://cytoscape.org) ClueGo plugin (v. 2.5.8) (Bindea et al., 2009) software to identify relevant biological processes and immune systems networks. A minimum of 10 genes were needed to be associated with a term. These genes would represent at least 4% of the total number of related genes. Only pathways with a *P*-value < 0.05 (Bonferroni step-down correction) were retained. Results were illustrated as a functionally grouped network of terms, having the most significant one as a leading term. The edges that show term-to-term interactions were obtained using a Kappa score of 0.4.

Then, coupled with DEG-driven approaches, we used the pathifier algorithm from the pathifier R package (v.1.32.0) (Drier et al., 2013), which by transforming the whole transcriptome expression data into pathway-level information, infers the pathway deregulation scores by measuring how much the gene expression of a sample deviates from normal behavior. Kyoto Encyclopedia of Genes and Genomes (KEGG) annotation was used, and the *quantify_pathway_deregulation()* function was used to quantify the deregulation scores. Euclidean distance was used to calculate samples' distances, then visualized using the Ward.D2 clustering method in a heatmap.

Phenomics: complementary data integration approaches

A global non-metric multidimensional scaling (NMDS) ordination was used to extract, visualize and summarize the variation in the transcriptome (the "response variable") using the vegan R package (v.2.5.7) (Dixon, 2003). Stress values were calculated to determine the number of dimensions for each NMDS. Stress values measure how much the distances in the reduced ordination space depart from those in the original *p*-dimensional space. High-stress values indicate a greater possibility that structuring observations in the ordination space is entirely unrelated to the actual full-dimensional space. Then, the explanatory variables related to milk composition and quality traits, cytometry cell profiles, and host morphometric parameters were fitted to the ordination plots using the *envfit()* function in the vegan R package (Clarke and Ainsworth, 1993) with 10,000 permutations. The *envfit()* function performs multivariate analysis of variance (MANOVA) and linear correlations for categorical and continuous variables. The effect size and significance of each covariate were

determined by comparing the difference in the centroids of each group relative to the total variation, and all *P*-values derived from the *envfit()* function were Benjamini-Hochberg adjusted. The obtained r^2 gives the proportion of variability (that is, the main dimensions of the ordination) that can be attributed to the explanatory variables.

As a second integrative approach, the *N*-integration algorithm DIABLO (Data Integration Analysis for Biomarker discovery using Latent Components) of the mixOmics R package (Rohart et al., 2017) (http://mixomics.org/, v. 6.18.1) was used. We combined host-centered transcriptomics with phenomics data to achieve this integrated perspective, coined holo-omics (Nyholm et al., 2020). It is to be noted that, in the case of the *N*-integration algorithm DIABLO, the variables of all the data sets were also centered and scaled to unit variance before integration. In this case, the relationships among all data sets were studied by adding a different categorical variable, *e.g.,* the infection status of cows. Healthy cows (n = 9) were compared to infected individuals (n = 22). DIABLO seeks to estimate latent components by modeling and maximizing the correlation between pairs of pre-specified datasets to unravel similar functional relationships (Singh et al., 2019). The model was first fine-tuned using leave-one-out cross-validation by splitting the data into training and testing. Then, classification error rates were calculated using balanced error rates (BERs) between the predicted latent variables and the class labels' centroid. Only interactions with $r \geq |0.80|$ were visualized using CIRCOS.

Identification and validation of hub variables

To visualize the high-confidence variable co-associations, only those with $r \geq |0.90|$ and more than 15 connections were automatically visualized using the organic layout algorithm in Cytoscape (version 3.9.1). The Molecular Complex Detection (MCODE) Cytoscape plug-in (version 2.01, Hogue and Groll, 2001) was adopted to detect densely connected modules within the interaction network. MCODE scores ≥ 3 were set as a cut-off criterion with default parameters.

Finally, cytoHubba (version 0.1) (Chin et al., 2014), a Cytoscape plug-in, was used to explore the network modules for identifying hub genes, defined as genes having high correlation in candidate

modules. The top 20 variables were identified and ranked using the Maximal Clique Centrality (MCC) method.

To validate the abovementioned hub genes as putative markers for mastitis infection, we performed the receiver operating characteristics (ROC) and precision-recall analyses using the R package pROC (v. 1.18.0) package to quantify the infection status predictive power of hallmark variables.

The meta-analysis cohort

To confirm the reproducibility of our prediction results in healthy and infected individuals, studies on the transcriptome of the milk somatic cell in dairy cows with high-throughput RNA sequence data in .fastq format deposited in publicly accessible databases and available metadata were retrieved.

We obtained 81 somatic cell transcriptomic samples from three independently published studies as an orthogonal dataset. The three studies were labeled as Seo (Seo et al., 2016), Asselstine (Asselstine et al., 2019), and Niedziela (Niedziela et al., 2021). Data included acute and subclinical infection regarding both naturally and experimentally infected animals. Raw sequence data and metadata from the Seo study were available at GSE60575 in the GEO database. In contrast, Asselstine study raw fastq files and metadata were retrieved from the NCBI under PRJNA544129 Bioproject accession number. Raw fastq files and metadata of the Niedziela study were available in the European Nucleotide Archive (ENA) repository with the project number PRJEB43443. The published data was pre-processed and annotated as described above. In this validation set, "healthy" subjects were defined as those reported as not being infected in the original research; alternatively, "mastitis" subjects were defined as those diagnosed with mastitis infection either by the California Mastitis Test (Asselstine et al., 2019) or after 24 h from the disease onset with two different strains of *Staphylococcus aureus* (Niedziela et al., 2021).

Ten-fold cross-validation sparse Partial Least Squares Discriminant Analysis (sPLS-DA) was employed to evaluate the prediction model's performance and validate the essential genes responsible

for the differences between groups using the mixOmics R package. The DESeq2 R package quantified differences between groups' relative gene abundance.

RESULTS

Animals and data

The 31 Holstein cows enrolled in this study were, on average, 4 years of age. The mean DIM at enrollment for all dairy cattle was 235 days, and the mean parity was 2.5, ranging from 2 to 5.

The average milk yield was 26.97 (± 9.01) kg/day. Milk had 2.19 ± 0.74% of fat, 3.49 ± 0.29% of protein, 2.73 ± 0.27% of casein, and 4.51 ± 0.44% of lactose. Milk pH and conductivity were 6.46 ± 0.08 and 9.94 ± 1.30 mS/cm, respectively (Table S1). A schematic summary of the experimental design and the conducted analyses is reported in Fig. 1.

S. agalactiae and Prototheca infections induced milk quality changes and divergent infiltration of immune cells

S. agalactiae and *Prototheca* infections were not accompanied by any clinical sign (e.g., udder swelling, redness, udder pain) or systemic reaction, but a significant drop in milk production ($P < 0.05$*)* compared with their uninfected counterparts (Fig. S1a). However, in contrast to the bacterial pathogen, *Prototheca* infection affected more the milk quality, by reducing the lactose content ($P < 0.05$; Fig. S1b) and casein index ($P < 0.05$; Fig. S1c). Additionally, algal infection increased milk conductivity ($P < 0.05$; Fig. S1d), mirroring possible changes in the blood-milk barrier permeability. The milk protein, casein, and fat proportions were similar between groups (Fig. S1e-g) and urea concentration and milk pH (Fig. S1h-i). Both types of pathogens, however, increased the amounts of somatic cells (SC, $P < 0.001$; Fig. S2a) and the combined ratio of neutrophils and lymphocytes to total SCC (DSCC) compared to healthy animals ($P < 0.05$; Fig. S2b), reflecting the inflammatory status of the mammary gland. Although both pathogens increased the total leucocyte population ($P < 0.001$, Fig. S2c), the immunological cell content differed between *S. agalactiae* and *Prototheca*

pathogens, evoking distinctly different immune responses to both pathogens. For instance, exposure to *S. agalactiae* primarily triggered the recruitment of nearby macrophages that increased more than 1.55-fold than the healthy animals. In contrast, *Prototheca* decreased macrophage populations by 20% (Fig. S2d) and sharply increased T helper cells (+ 73%), T killer cells (+ 110%), and B cells (+30%) compared to *S. agalactiae* (Fig. S2 e-g). No significant differences were found among the three experimental groups for PMN cells (Fig. S2h), even if their proportion was slightly higher in the *S. agalactiae*-induced mastitis. Importantly, we found large variability of immune cell contents among individuals within each group, as assessed by the principal component analysis (PCA) visualization (Fig. S2i).

Somatic cell transcriptome changes upon Prototheca spp. and S. agalactiae infections

A total of seven billion paired-end reads were obtained from the somatic cells of 31 dairy cows (9 healthy and 22 naturally infected subjects), corresponding to an average of 119 M ± 49.8M per sample. After quality filtering, 88.84 % of high-quality paired reads were mapped, on average, to the bovine reference genome ARS-UCD1.2 and aligned with 27,607 unique genes. After filtering for genes with CPM > 0.5 in at least two samples, we obtained 14,564 abundant genes, henceforth referred to as expressed genes, corresponding to ~ 53% of the transcriptome (Table S3).

The generalized PCA showed that the expression of genes varied according to the infection status, separating infected and uninfected individuals (generalized PCA axis 1: 22 %; Fig. 2a). Altogether, transcriptome signatures fell into different *S. agalactiae, Prototheca,* and control response patterns, except for two *S. agalactiae* and one *Prototheca* spp. samples that grouped with healthy ones. Besides pathogen infection, we determined to what extent differences in host-associated variables could further explain the observed patterns of transcriptional variation. Using NMDS o[r]dinations (Fig. 2b) for visualizing the structure of gene expression (ordination stress = 12.36%, k = 2, non-metric fit r^2 = 0.985, linear fit r^2 = 0.932), the principal contributors explaining the total variance of the transcriptome were flow cytometry variables, including leucocytes (*envfit*, r^2 = 0.5329, $P < 0.001$), T-helper (*envfit*, r^2 = 0.4167, $P < 0.001$), T-killer (*envfit*, r^2 = 0.3999, $P < 0.001$) cells,

and PMN (*envfit*, $r^2 = 0.3512$, $P < 0.01$), together with udder health parameters, such as DSCC (*envfit*, $r^2 = 0.4456$, $P < 0.001$) and SCC (*envfit*, $r^2 = 0.4456$, $P < 0.001$), and milk yield (*envfit*, $r^2 = 0.3569$, $P < 0.01$) (Fig. 2c). Macrophages, lactose, casein index, pH, and conductivity also accounted for the transcriptome variation, albeit of lesser significance.

A total of 1,682 DEGs were detected in *Prototheca*-infected animals (671 downregulated, 1,011 upregulated; $P < 0.05$) compared to the healthy controls, and these differences were even more significant for the *S. agalactiae*-infected group (n = 2,427 DEGs; 890 down- and 1,537 upregulated; $P < 0.05$). DEG lists shared significant similarities between the two types of pathogens, as 40% of *Prototheca*'s DEGs were also expressed in the *S. agalactiae* group. When comparing *S. agalactiae* with *Prototheca* infections, 974 genes were differentially expressed (378 under and 596 over-expressed genes; $P < 0.05$). The list of all the DEGs belonging to each comparison is reported in Table S4. To gain insights into the biological and immune processes in response to the type of infection, we performed a functional enrichment analysis of the differentially up and downregulated genes.

Following *Prototheca* infection, 33 % of DEGs were involved in immune system response (n = 228), antigen processing (n = 41), and response to other organisms' invasion (n = 293) (Fig. 3a). Immune activation involved pathways related to innate response, such as toll-like receptors (*TLR9*) and pro-inflammatory molecules like IL-15, IL-17A, and IL-17F and adaptive immune response. The adaptive immune response mainly guided the defense line against *Prototheca* through the activation of class II MHC molecules (*BoLA-DMA*, *BoLA-DMB*, *BoLA DOA*, *BoLA-DOB*, *BoLA-DRA*, *BoLA-DRB3),* stimulation of IFN-γ and IL4I1 (interleukin 4 induced 1), and proliferation of lymphocytes (n = 86 genes). The induction of *CD48* and *CD80* promoted T cell activation following *Prototheca* infection. At the same time, B lymphocyte upregulation was conducted by B lymphocyte differentiation (*IKZF3*) and B-cells response to antigens (*POU2AF1*).

In the context of *S. agalactiae* infection, the innate immune cells development and differentiation (n = 137) appeared as the first line of defense against the disease, coupled with a reduction of mitochondrial energy metabolism (n = 526, 21% of DEGs), that is, depletion of the TCA

(*n* = 83), oxoacid (*n* = 266), and carbohydrate derivative metabolic processes (*n* =177) (Fig. 3b). At a closer look, the immune response to *S. agalactiae* infection seemed to be led by Notch receptor 1 (*NOTCH1*), NF-κB signaling pathway, and pattern recognition receptors (PRR) like *TLR9* and *NOD2*. Other pathways involved in cell-cell adhesion (*SELL*, *CD274, BOLA-DRA, CSF3R*), pro-inflammatory chemokines, and cytokine secretion (*IL-17A*, *IL-17F*, *OSM*, *LTA*, *LTB*) and activation of the complement system C3 mediated were also enriched. Additionally, *S. agalactiae* modulated the expression of myeloid-derived suppressor cells (MDSCs) by activating the transcription factor *STAT3*. Despite the high number of differentially expressed genes in comparing the two pathogen infections (*S. agalactiae* vs. *Prototheca;* Table S4), no differential enriched pathways resulted from the analysis.

Last, we sought to support further that *S. agalactiae* and *Prototheca* triggered different immune responses using Pathifier. Through this algorithm, we identified 69 KEGG pathways with Pathway Deregulation Scores (PDS) significantly associated with the two types of infection compared to healthy individuals (Fig. 3c). Notably, the peroxisome proliferator-activated receptor (PPAR) pathway, an essential modulator of the immune response, was firmly (more than 2-fold) deregulated in *S. agalactiae*-infected animals compared to healthy counterparts. Additionally, NK cells mediated cytotoxicity path, which can be considered a more innate defense, was almost two times more deregulated in the *S. agalactiae* infection concerning the healthy animals. The energy-related paths (e.g., TCA cycle, carbohydrate metabolism, oxidative phosphorylation) in the *S. agalactiae* infection resulted in 2-fold deregulation compared to the healthy animals. At the same time, no significant alterations were detected for *Prototheca*'s infected animals.

Specific alteration in *Prothotheca*'s PDS involved a clear focal adhesion path, almost 3-fold lower than healthy animals and more than 2-fold lower than *S. agalactiae*. Conversely, the adherens junction path was specifically deregulated in *S. agalactiae* infection. Both disorders' pathways involving extracellular matrix receptor interaction (ECM) and gap junction were mildly deregulated. A closer look at the immune paths showed that both infections displayed strong PDS concerning the

NF-κB signaling pathway. Despite it, the *S. agalactiae* infection showed the most diverging scores, especially when considering innate response-related functions such as leucocyte migration, endocytosis, and neutrophil extracellular trap formation. Interestingly, PDS concerning Th17 cell differentiation and B cell receptor signaling, pathways more explicit for the adaptive immune response, were significantly deregulated (more than 2-fold) in *Prototheca*-infected animals, even if a modest alteration was also observed in *S. aglactiae* infection.

A set of core mastitis-response genes

Even if there were differences in intensity between the pathways regulated by the two microorganisms, the core immune transcriptome did not seem to respond so differently. For this reason, beyond the type of pathogen, we identified 1,954 DEGs in response to mastitis, of which 1,289 were under-expressed and 665 were over-expressed in infected individuals compared to healthy ones. Among them, 681 DEGs were commonly shared with the pathogen-specific DEGs lists. These were considered the core mastitis-response genes (Table S4). Enzymes make up the most significant gene function category (67%), outranking transcription factors (TF, 8.8%), transporters (6.7%), transmembrane receptors (4.4%), kinases (3.8%), and G-protein coupled receptors (2.5%). Around 9.8% of genes were unannotated. Two hundred twenty-nine genes were highly expressed in infected samples. The functional analysis showed that roughly 23% of these genes (n = 156) were directly involved in activating and regulating the immune response. In contrast, 69 genes were associated explicitly with catabolic and oxidative pathways (Fig. S3). Interestingly, inflammasome activation and regulation were enriched upon encountering a pathogenic agent (e.g., *NLRC5, TLR9, GBP5, PLCG2*) and the mitochondrial-related genes (n = 80; hypergeometric test, an adjusted P-value of 5.29×10^{-5}). Moreover, we found the downregulation of pathways involved in the lipid metabolism and synthesis of de-novo fatty acids (*FASN*, *ACACA*).

Integration of core mastitis-response genes and phenomic data

Using DIABLO, we observed the strongest covariation between the core mastitis genes and the immune cells populations (IS) (r^2 = 0.72), followed by the udder health (UH) (r^2 = 0.64) and milk quality (MQ) parameters (r^2 = 0.64; Fig 4a). No important covariation was found between the core mastitis response genes and the host variables (e.g., parity and DIM; r^2 = 0.33). Concomitantly, the immune cells co-varied with the udder health-related variables (r^2 = 0.64). Then, to add biological meaning to the predicted model, we investigated the relationship between the DIABLO-selected features with the highest covariation. The first latent variable of the immune cells data set supported induction of the immune system response in mastitis cows, with increased infiltrations of leucocytes and T-killer cells, and to less extent, PMN, T-helper cells, macrophages, and B cells (Fig 4b). Paired with these immune-related cells, the first latent variable for the udder health parameters pointed at higher levels of DSCC, SCC, and milk conductivity upon mastitis but lower casein index, lactose, and pH (Fig 4c). Regarding the genes selected, the first latent variable of the expected model indicated that subjects with mastitis overexpressed the Prostate Androgen-Regulated Mucin-Like Protein 1 (*PARM1)* gene. Moreover, in infected animals, we observed the induction of genes involved in the transcription of class II MHC molecules (*CIITA*), cell proliferation and apoptosis (*SAMD9*), and adhesion and diapedesis of granulocytes (*SELPLG*) (Fig. 4d). Conversely, healthy subjects were primarily defined by genes related to molecules transportation and transmembrane proteins (*LAPTM4A*, *ANO10*, *GNA11*).

Identification of mastitis hub variables

The co-association of gene expression and phenomic data obtained from CIRCOS (Fig. 4e) resulted in a network construction consisting of 116 interactors (nodes) and 4,430 interactions (edges). This network was further analyzed using the Molecular Complex Detection (MCODE) Cytoscape plug-in, which identified four densely connected modules. The two most significant modules showed MCODE scores of 54.732 and 6.667, respectively. Seven variables (*AGFG1*,

CEMIP2, *ITGB7*, *RRAD*, Urea, T killer cells, and leucocytes) were not assigned to any module (Fig. S4).

Lastly, with cytoHubba, we identified the top 20 hub variables, which are displayed in Fig. 5 and ordered as follows according to the MCC ranking method: milk conductivity, lactose, *P2RY6*, *SPTBN5*, *BoLA-DOA*, *ENSBTAG00000053850*, *CIITA*, *GNA11*, *ENSBTAG0000003367*, casein index, *ENSBTAG0000003408*, *HIP1*, *CLMN*, *RESEF*, *EFHD1*, *LAPTM4A1*, *FCRL5*, milk yield, *PAH*, *ELF5* (Full MCC ranking in Table S5). The genes that were commonly shared by both DIABLO and cytoHubba (*P2RY6, SPTBN5, BOLA-DOA, CIITA, GNA11, ENSBTAG0000003367,* ENSBTAG0000003408*, HIP1, LAPTM4A, FCRL5)* were then submitted to a ROC analysis resulted in having excellent prediction performances in terms of discriminating healthy and mastitis animals with sensitivity > 0.89, specificity > 0.81, accuracy > 0.87 and precision > 0.69. The detailed predictive performances of the hub genes are presented in Table S6.

Meta-analysis of subclinical mastitis

To validate the core mastitis response genes, we additionally gather publicly available somatic cell transcriptome RNA-seq datasets (n = 81) derived from three independent studies in dairy cows: i) pooled milk sampled from 12 healthy cows (Seo et al., 2016), ii) quarter samples from 6 cows (n = 12) (Asselstine et al., 2019); and, 14 cows sampled at five different times (n = 48) (Niedziela et al., 2021). As with the study cohort, the transcriptome profiles could be distinguished based on the infection status (Fig. 6a), with higher transcriptome dispersion in infected individuals. The first principal component accounted for 51 % of the total variance, and the first two accounted for 74%. Consistent with the training set, many genes (n = 7,226) were differentially expressed between the healthy and infected individuals, including 3,912 up- and 3,314 down-regulated. We then intersected the list of core mastitis-response genes and the list of DEGs from these published datasets (Fig. 6b). We found that 53% of these core-mastitis response genes were DEG in the validation set, strengthening the shared transcriptional response to infection, independently on the pathogen,

regional, and other potential differences, such as diet, medication, energy expenditure, age, and DIM of the dairy cows.

To further assess the feasibility of the previously selected hub genes for discriminating healthy and infected animals, we re-performed a ROC analysis on the unified training-validation dataset (*n* =112). The class II transactivator (*CIITA*) had the best prediction performances, having sensitivity, specificity, and accuracy > 0.7 and precision equal to 0.8.

DISCUSSION

Somatic cells are released in milk as the first line of defense against mammary infections, and they are widely applied as an indicator for mastitis screening and detection. Moreover, their expression signatures fit suitably for performing mastitis mammary-gland-specific studies and monitoring the pathogen-specific molecular response.

To the best of our knowledge, this is the first study on the milk somatic cell transcriptome variations in response to the infection of two pathogens: *S. agalactiae*, a Gram-positive bacterium, and *Prototheca* spp., a microscopic alga. *Prototheca*'s molecular mechanism of action is still poorly understood and has only recently been recognized as a non-negligible mastitis agent, with an estimated 11.2% prevalence on Italian territory (Shave et al., 2021). In addition, for the first time, we integrated somatic cell transcriptome data with a wide range of phenotypic traits in a joint analysis to identify hub variables affected by subclinical intramammary infection in dairy cattle.

We tried to minimize any possible source of variation that might modify the transcriptional response by excluding primiparous cows and animals during their periparturient period. In fact, in early lactating cows, the decreased feed intake and the increased energy demand often result in an energy-negative balance condition (Ingvartsen and Moyes, 2013), which can also affect the proper activation of the immune cell's metabolism in response to pathogens' invasion (Wathes et al., 2009).

In terms of production performances, the significant drop in milk yield in the infected animals (regardless of the pathogen) was in line with what was observed by previous studies conducted on subclinical mastitis (Bobbo et al., 2017). Instead, the most quality-impaired milk was the one from *Prototheca* infection. Pegolo et al. (Pegolo et al., 2022a) reported similar results on a broader cohort of animals affected by subclinical IMI from four pathogens: *S. aureus*, *S. agalactiae*, *S. uberis,* and *Prototheca* spp. Both pathogens significantly increased milk SCC compared to the negative control samples but resulted in different leucocyte proportions. Exposure to *S. agalactiae* primarily triggered an innate immune response by recruiting nearby macrophages at the site of infection. At the same time, *Prototheca* seemed to show resistance to the phagosomal defense mechanism and a more adaptive-driven response through the crucial role of T-cells. Even though the diverse leucocyte profile suggests a different immune response, the large variability of the immune cells assessed by the PCA (Fig. S2i) evokes that beyond infection and protection, immune cell patterns could also be driven by a range of genetic and non-genetic factors such as the inflammation stage as well as individual environmental exposures.

SC transcriptome response upon S. agalactiae and Prototheca infection

To further dissect these phenotypic phenomena, the RNAseq approach was applied. Antigen-presenting and processing was a critical pathway enriched in *Prototheca spp*. infected samples, suggesting that the adaptive immune response was mainly guided by the activation of MHC class II molecules. The bovine MHC genes are called the Bovine Lymphocyte antigen (BoLA) (Behl et al., 2012). Specifically, class II is expressed on antigen-presenting cells and activates CD4+ T cells, resulting in the coordination and regulation of effector cells (Wieczorek et al., 2017).

S. agalactiae was found to modulate the expression of MDSCs, a heterogeneous subset of immature monocytes and granulocytes, by activating *STAT3*, which stimulates myelopoiesis, inhibiting myeloid-cell differentiation (Gabrilovich and Nagaraj, 2009). Interestingly, the infection of *S. agalactiae* induced a down-regulation of mitochondrial energy related-pathways like the TCA cycle, oxoacid, and carbohydrate derivative metabolic processes, which might be partially linked to

the relationship existing between macrophages and mitochondria. With the onset of inflammation, macrophages activate, showing a proinflammatory profile metabolically characterized by an increased glycolysis and lactate production (Tur et al., 2017). In contrast, even in the presence of oxygen, mitochondrial oxidative phosphorylation (OXPHOS) is reduced in pro-inflammatory macrophages (the so-called Warburg effect), presumably as an effect of the tricarboxylic acid (TCA) cycle fragmentation (Jha et al., 2015).

Finally, using the Pathifier algorithm, we identified 69 KEGG pathways which explained the differential gene expression profile in the two types of infection for the negative control. In line with our previous findings, we observed the significant deregulation in the TCA cycle, carbohydrate metabolism, oxidative phosphorylation pathways as well as the peroxisome proliferator-activated receptor (PPAR), which is an essential modulator of the immune response tightly linked to mitochondrial metabolism (Di Cara et al., 2019), confirmed the role of mitochondria as major hubs in inflammatory and immune response carried out by *S. agalactiae*.

Pathways involving ECM, focal adhesion, and gap junction, which comprehend the group of genes involved in the communication and integrity of the epithelial cells (Dedhar, 2000), were mildly deregulated in both infections, confirming transcriptomic and iTRAQ-proteomics patterns on Chinese Holstein cows challenged with *S. agalactiae* via nipple tube perfusion (Zhang et al., 2018) . Altered gene expression within these pathways might be related to the reorganization of the actin cytoskeleton, which may represent one way of reducing tissue damage caused by invading pathogens. Our previous functional analysis supported these findings in which we found some integrin family coding genes. In *S. agalactiae*-positive samples, we found *ITGA5* upregulated, similar to the work of Niedziela et al. (2021). This gene encodes the light and heavy chains that create the α5 subunit, which, by associating with the β1 subunit, form the receptors for fibronectin and fibrinogen. These two glycoproteins are essential mediators of the pathogen adhesion (Hauck and Ohlsen, 2006). Their abundance increased in the work of Mudaliar et al. (Mudaliar et al., 2016), where a label-free proteomic approach analyzed the changes in the protein profile of milk whey in a cohort of animals

experimentally infected by S. uberis. In *Prototheca* infection, we found instead *ITGA9*, which encodes integrin subunit β7, necessary for the leukocyte adhesion (Ivaska and Heino, 2000).

Moreover, PDS concerning Th17 cell differentiation and B cell receptor signaling, pathways more specific for the adaptive immune response, were significantly deregulated in *Prototheca*-infected animals, even if a modest alteration was observed in *S. agalactiae* infection.

It is crucial to notice that with this approach, we also found significant differences in PDS within samples belonging to the same group. This is partially linked to the fact that we did not have the species identification for *Prototheca* or the strain identification for *S. agalactiae*, which led to divergent responses. Secondly, as we worked on naturally occurring mastitis, we could have had different stages of inflammation in our samples.

Common transcriptomic signature of subclinical intramammary infection

Despite some differences in the DEGs expression within the activation of the immune system pathways, we did not find extreme differences in the two microorganisms' transcriptomic profiles, despite the great phylogenetic distance between them. This was further confirmed by the enrichment analysis comparing the DEGs among the two pathogens that did not produce significant pathways. It is, however, important to remember that *S. agalactiae* and *Prototheca* are known to induce a weak immune reaction compared to Gram-negative bacteria like *E. coli*, which results in subclinical mastitis with usually no systemic repercussions (Günther et al., 2017). Both pathogens can put in place mechanisms of immune evasion that might explain a more moderate inflammatory response. Indeed, in *S. agalactiae*, synthetase proteins such as FbsA and FbsB are involved in fibrinogen binding that can decrease the risk of opsonization by phagocytic cells. Moreover, the production of the serine protease CspA allows this pathogen to further evade the immune response by cleaving specific chemokines responsible for the neutrophil recruitment (Kabelitz et al., 2021). *Prototheca*, conversely, seems to be able to form a biofilm which could be implicated in its pathogenicity, partial immune invasion (Pegolo et al., 2022b), and ability to persist in the environment (Leonel Gonçalves et al., 2015).

Regardless of the pathogen, we identified several highly expressed genes in animal samples with subclinical IMI. This suggests these "core mastitis-response genes" may represent a typical infection signature and provide a potential therapeutic window for mastitis drug development.

Pathway analysis conducted on the 681 core-mastitis response genes (the ones commonly shared by the two infections) identified that most upregulated pathways were related to the activation of innate and adaptive immune system processes. Interestingly, among the expressed ILs, the significant upregulation of *IL-17A* and *IL-17F* might be specifically linked to the so-called "type 3" immunity, which encompasses innate and adaptive immune response and is characterized both by the recruitment of neutrophils and the stimulation of epithelial antimicrobial defenses at the sites of infection (Rainard et al., 2020). Overexpression of *IL-17A* and *IL-17F* encoding genes was also observed in the mammary gland of cows infected with *E.coli* (Roussel et al., 2015) and in goats infected by *S. aureus* (Rainard et al., 2018). In contrast, the downregulated ones involved oxoacid metabolic processes. These biological processes were commonly shared with the work of Asselstine et al. (2019), one of the few published papers that characterized the milk somatic cell transcriptome in Holstein cows. In addition, they found several unspecific Gene Ontology (GO) terms, which might be explained by the fact that genes expressing alternative transcripts might have been associated with a heterogeneous role in biological functions (Cardoso et al., 2018).

NLRC5, *TLR9*, *GBP5*, and *PLG2* were highly enriched upon encountering the pathogenic agents. These genes are well-known to have a pivotal role in the activation of *NLRP3* (NLR family PYD containing 3), which is considered one of the essential activators and synergic components of the inflammasome (Yu and Lee, 2016) which, in turn, are a class of molecules assembled by the PRRs which are well known to play an important role in innate immunity through the stimulation of pro-inflammatory cytokines and pyroptosis (Backert et al., 2016). Moreover, the identification of several mitochondrial-related DEGs was consistent with the hypothesis that mitochondria activity may be regulated by subclinical intramammary infection.

Within the pathway involved in the fatty acid metabolic process, we found several lipogenic genes (*FASN*, *ACACA*) that were downregulated. The decreased expression of these genes in animals with mastitis compared with negative control was also observed in the work of Moyes et al. (Moyes et al., 2009) and Huma et al. (Huma et al., 2020) might probably be related to the fact that with the onset of inflammation, the energy demand needed to produce new fatty acids is too high (Li et al., 2021).

Interestingly, the most under-expressed gene was the *SLC34A2*, according to what was observed by Asselstine et al. (2019). *SLC34A2* encodes the solute carrier family 34 members 2, a sodium-dependent phosphate transporter, which was upregulated in disease-resistant cows (van Altena et al., 2016). It is therefore considered a putative biomarker for selecting disease resistance in dairy cattle.

Identification of hub variables for subclinical intramammary infection

Among the candidate DE genes identified by DIABLO, the class II transactivator (*CIITA*) had one of the highest loading scores and was significantly upregulated in mastitis animals. *CIITA* has been recognized as one of the master regulators in the gene expression of MHC. Moreover, it is involved in the transcription graduation of over 60 immunologically essential genes, including interleukin 4 (*IL-4*) and *IL-10* (Devaiah and Singer, 2013). *CIITA* was found to be a critical regulator of the immune response of *Cynoglossus semilaevis* towards the infection of *Vibrio harvey*, suggesting its putative involvement also in the molecular inflammatory process of mastitis. Interestingly, it was also proposed as one of the most important candidate genes for bovine paratuberculosis tolerance in the GWAS study conducted by Canive et al. (2022)

All the immune-cells variables selected by DIABLO were positively associated with mastitis, with total leucocytes having the highest loading score. It is already widely established for milk quality traits that SCC, and more recently DSCC, represent the most critical and easy-to-use indicator for identifying inflammation at the udder level (Pilla et al., 2013; Guzzo et al., 2018).

Moreover, lactose and casein index proportion are two traits that are highly influenced by inflammatory processes. The reduction in lactose proportion associated with clinical and subclinical mastitis is related to the reduced secretory activity of the mammary epithelial cells and an increase in the permeability of the mammary epithelium due to tight junction impairment (Televičius et al., 2021). The casein index reduction upon IMI is related to the increased proteolytic activity due to both endogenous and bacterial proteases that particularly damage the casein fractions (Le Roux et al., 2003; Forsbäck et al., 2010). Casein index and lactose content represent a useful additional tool for discriminating against healthy/infected animals.

Among the phenotypic indicators identified by the integrated DIABLO-cytohubba approach, milk conductivity, and lactose were the ones showing the highest ranking. In fact, during the inflammation of the mammary gland, the osmotic balance is maintained through the increase in Na^+ and Cl^-, which are responsible for the rise in milk electrical conductivity (Norberg et al., 2004). In this context, a study by Ebrahimie et al. (Ebrahimie et al., 2018) on Holstein cows identified milk lactose and conductivity, together with SCC, as the most reliable indicators for the detection of subclinical mastitis. Moreover, also a recent work conducted by Antanaitis et al. (Antanaitis et al., 2021) on 5814 cows observed a strong association between lactose levels and subclinical mastitis pathogens, concluding that it might be helpful to include lactose (as well as milk conductivity) as a biomarker of suspected udder inflammation in health prevention programs.

Among the selected hub genes, *BoLA-DOA* encodes the major histocompatibility complex (MHC), class II, DO alpha. It is already well known that molecules linked to MHC play a fundamental role in the antigen recognition, presentations, and activation of the adaptive immune response (Behl et al., 2012). Among the numerous molecules that belong to the BoLA family, *BoLA-DOA* specific mechanism of action has yet to be unraveled. Nonetheless, some studies found this gene upregulated in the presence of mastitis. Chen et al. (Chen et al., 2021), for example, in the transcriptional survey of exosomes derived from *Staphylococcus aureus*-infected bovine mammary epithelial cells, found a significant upregulation of the *BoLA-DOA* gene. Conversely, Cheng et al. (Cheng et al., 2022)

observed the downregulation of *BoLA-DOA* and several other genes involved in antigen presentation and processing in the blood-circulating leucocytes of animals recovering from *E.coli* clinical mastitis. This difference in gene expression might be related to the fact that *E. coli*, unlike *S. agalactia*e or *Prototheca*, generally induces an acute and robust udder inflammation with a more generalized immune response (Petzl et al., 2018).

One of the essential hub genes downregulated in mastitis animals was *GNA11,* which encodes for a type of guanine nucleotide-binding protein (G-protein) functioning as a modulator or transducer in the transmembrane signaling systems. In the recent work of Pan et al. (Pan et al., 2022) on transcriptome evaluation in early calf nutrition, *GNA11* was significantly implicated in energy-related pathways, especially fat metabolism. Therefore, its downregulation in mastitis-positive samples should be further evaluated better to explain the relationships between energetic pathways and immune response.

CONCLUSIONS

This work evaluated for the first time the somatic cell transcriptomic signature derived from naturally occurring subclinical mastitis caused by two different etiological agents: *S. agalactiae* and *Prototheca* spp. Even though we found some differences in the immune-related pathways and gene expression between the two infections (e.g., more robust activation of the antigen and processing complex in *Prototheca* and potent inhibition of energy-related pathways in *S. agalactiae* infection), the core immune response was commonly shared between the two pathogens.

The integrated analysis of core mastitis response genes and phenotypic traits confirmed a strong correlation between the transcriptome and the leucocyte populations determined by flow cytometry and with udder health traits (SCC, lactose, conductivity, and casein index), strengthening the need to systematically include them as screening and diagnostic indicator for the IMI detection. Finally, the predictive performances on the hub genes tested within the meta-analysis highlighted that *CIITA*

might have a crucial role in the molecular mechanism underlying the animals' response to subclinical IMI and need further evaluation in future studies, also taking into consideration a wider cohort of animals.

List of abbreviations

SCC: Somatic cell count; *E. coli*: *Escherichia coli*; *S. aureus*: *Staphylococcus aureus*; *S. agalactiae*: *Streptococcus agalactiae*; RNA-Seq: RNA sequencing; SC: Somatic cells; IMI: Intra-mammary infection; GWAS: Genome-wide association study; IZSVe: Istituto Zooprofilattico Sperimentale delle Venezie; DIM: Days in milk; T0: time 0; T1: time 1; DSCC: Differential somatic cell count; CPM: Counts per million; PCA: Principal component analysis; NMDS: Non-parametric multidimensional scaling; DEG: Differentially expressed gene; KEGG: Kyoto Encyclopedia of Genes and Genomes; FDR: False discovery rate; DIABLO: Data Integration Analysis for Biomarker discovery using Latent Components; MCODE: Molecular complex detection; MCC: Maximal Clique Centrality; ROC: receiver operating characteristics; PDS: pathway deregulation scores; *TLR9*: Toll-like receptor 9; *IL-15*: Interleukin 15; *IL-17A*: Interleukin 17A; *IL-17F*: Interleukin 17F; *BoLA-DOA*: *Bos taurus* major histocompatibility complex, class II, DOA; IFN- γ: Interferon gamma; *IL41I*: Interleukin induced 1; PRR: Pattern recognition receptor; NF-κB: Nuclear factor κB; TCA: tricarboxylic acid cycle; *NOD2*: Nucleotide binding oligomerization domain containing 2; *OSM*: Oncostatin; *LTA*: Lymphotoxin Alpha; *LTB*: Lymphotoxin; MDSC: myloid derived suppressor cells; *STAT3*: Signal transducer and activator of transcription 3; PPAR: peroxisome proliferator-activated receptor; NK: natural killer; ECM: extracellular matrix receptor interaction; Th17: T helper 17 cells; TF: transcription factors; *NLRC5*: NLR Family CARD Domain Containing 5; *FASN*: Fatty acid synthase; *ACACA*: Acetyl-CoA Carboxylase Alpha; IS: Immune cells; UH: Udder health; MQ: Milk quality; PARM1: Prostate androgen-regulated mucine-like protein 1; *CIITA*: Class II Major Histocompatibility Complex Transactivator; *SAMD9*: Sterile Alpha Motif Domain Containing 9; *SELPLG*: Selectin P Ligand; *LAPTM4A*: Lysosomal Protein Transmembrane 4 Alpha; *ANO10*:

Anoctamin 10; *GNA11*: G Protein Subunit Alpha 11; *ITGB7*: Integrin Subunit Beta 7; *RRAD*: Ras Related Glycolysis Inhibitor And Calcium Channel Regulator; *P2RY6*: Pyrimidinergic Receptor P2Y6; SPTBN5: Spectrin Beta, Non-Erythrocytic 5; *HIP1*: Huntingtin Interacting Protein 1; MHC: Major histocompatibility complex; OXPHOS: mitochondrial oxidative phosphorylation; *ITGA5*: Integrin Subunit Alpha 5; *ITGA9*: Integrin Subunit Alpha 9; GO: Gene ontology; ER: endoplasmic reticulum; *IL-10*: Interleukin10;

Declarations

Ethics approval and consent to participate

The research was approved by the Ethical Animal Care and Use Committee (OPBA - Organismo Preposto al Benessere degli Animali) of the Università Cattolica del Sacro Cuore and by the Italian Ministry of Health (protocol number 510/2019-PR of 19/07/2019)

Consent for publication

Not applicable.

Availability of data and material

The sequencing data of this study were deposited in the NCBI's Sequence Read Archive (SRA) under the PRJNA911953 accession number.

Public RNA-seq data were downloaded from the GEO database at GSE60575 accession number (Seo), from NCBI PRJNA544129 Bioproject accession number (Asselstine), and the European Nucleotide Archive (ENA) repository with the project number PRJEB43443 (Niedziela).

Competing interests

The authors do not report any conflict of interest.

Funding

This research was part of the project LATSAN funded by the Ministero delle politiche agricole alimentari, forestali e del turismo (MIPAAF), Rome, Italy. Moreover, the study was conducted within the Agritech National Research Center and received funding from the European Union Next-

GenerationEU (PIANO NAZIONALE DI RIPRESA E RESILIENZA (PNRR) – MISSIONE 4 COMPONENTE 2, INVESTIMENTO 1.4 – D.D. 1032 17/06/2022, CN00000022). This manuscript reflects only the authors' views and opinions; neither the European Union nor the European Commission can be considered responsible for them.

Authors contributions

S.P. conceived the study. A.C. contributed to the experimental design. V.B. and A.V. collected the samples and performed lab analysis. M.E.G. helped with flow cytometry analysis. E.C. performed the RNA extraction. V.B., N.M., and D.G. analyzed the data. V.B., N.M., S.P., D.G., M.E.G, A.C., and P.A.M. contributed to data interpretation. V.B. prepared and wrote the initial draft. All authors reviewed and approved the final manuscript.

ACKNOWLEDGMENTS

The authors acknowledge Rossella Tessari, Alessandro Toscano, Selene Massaro, Martina Piazza, and Giorgia Secchi (DAFNAE, Legnaro PD, Italy) for supporting sampling activities and laboratory work. The CARIPARO foundation founds the Ph.D. grant of Vittoria Bisutti.

REFERENCES

van Altena, S.E.C., B. de Klerk, K.A. Hettinga, R.J.J. van Neerven, S. Boeren, H.F.J. Savelkoul, and E.J. Tijhaar. 2016. A proteomics-based identification of putative biomarkers for disease in bovine milk. Vet. Immunol. Immunopathol. 174:11–18. doi:10.1016/j.vetimm.2016.04.005.

Antanaitis, R., V. Juozaitienė, V. Jonike, W. Baumgartner, and A. Paulauskas. 2021. Milk lactose as a biomarker of subclinical mastitis in dairy cows. Animals 11:1–11. doi:10.3390/ani11061736.

Asselstine, V., F. Miglior, A. Suárez-Vega, P.A.S. Fonseca, B. Mallard, N. Karrow, A. Islas-Trejo, J.F. Medrano, and A. Cánovas. 2019. Genetic mechanisms regulating the host response during mastitis. J. Dairy Sci. 102:9043–9059. doi:10.3168/jds.2019-16504.

Backert, S., H. Dubois, A. Wullaert, and M. Lamkanfi. 2016. Inflammasome signaling and bacterial infections.

Bakhtiarizadeh, M.R., S. Mirzaei, M. Norouzi, N. Sheybani, and M.S. Vafaei Sadi. 2020. Identification of Gene Modules and Hub Genes Involved in Mastitis Development Using a Systems Biology Approach. Front. Genet. 11:1–16. doi:10.3389/fgene.2020.00722.

Behl, J.D., N.K. Verma, N. Tyagi, P. Mishra, R. Behl, and B.K. Joshi. 2012. The Major Histocompatibility Complex in Bovines: A Review. ISRN Vet. Sci. 2012:1–12. doi:10.5402/2012/872710.

Bindea, G., B. Mlecnik, H. Hackl, P. Charoentong, M. Tosolini, A. Kirilovsky, W.H. Fridman, F. Pagès, Z. Trajanoski, and J. Galon. 2009. ClueGO: A Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. Bioinformatics 25:1091–1093. doi:10.1093/bioinformatics/btp101.

Bobbo, T., P.L. Ruegg, G. Stocco, E. Fiore, M. Gianesella, M. Morgante, D. Pasotto, G. Bittante, and A. Cecchinato. 2017. Associations between pathogen-specific cases of subclinical mastitis and milk yield, quality, protein composition, and cheese-making traits in dairy cows. J. Dairy Sci. 100:4868–4883. doi:10.3168/jds.2016-12353.

Bradley, A.J. 2002. Bovine mastitis: An evolving disease. Vet. J. 164:116–128. doi:10.1053/tvjl.2002.0724.

Canive, M., G. Badia-Bringué, P. Vázquez, J.M. Garrido, R.A. Juste, A. Fernandez, O. González-Recio, and M. Alonso-Hearn. 2022. A Genome-Wide Association Study for Tolerance to Paratuberculosis Identifies Candidate Genes Involved in DNA Packaging, DNA Damage Repair, Innate Immunity, and Pathogen Persistence. Front. Immunol. 13:1–14. doi:10.3389/fimmu.2022.820965.

Di Cara, F., P. Andreoletti, D. Trompier, A. Vejux, M.H. Bülow, J. Sellin, G. Lizard, M. Cherkaoui-Malki, and S. Savary. 2019. Peroxisomes in immune response and inflammation. Int. J. Mol. Sci. 20:1–29. doi:10.3390/ijms20163877.

Cardoso, T.F., R. Quintanilla, A. Castelló, R. González-Prendes, M. Amills, and Á. Cánovas. 2018. Differential expression of mRNA isoforms in the skeletal muscle of pigs with distinct growth and fatness profiles. BMC Genomics 19:1–12. doi:10.1186/s12864-018-4515-2.

Chen, Y., H. Jing, M. Chen, W. Liang, J. Yang, G. Deng, and M. Guo. 2021. Transcriptional Profiling of Exosomes Derived from Staphylococcus aureus -Infected Bovine Mammary Epithelial Cell Line MAC-T by RNA-Seq Analysis. Oxid. Med. Cell. Longev. 1–18.

Cheng, Z., L. Buggiotti, M. Salavati, C. Marchitelli, S. Palma-Vera, A. Wylie, H. Takeda, L. Tang, M.A. Crowe, D.C. Wathes, M. Crowe, N. McLoughlin, A. Fahey, E. Matthews, A. Santoro, C. Byrne, P. Rudd, R. O'Flaherty, S. Hallinan, C. Wathes, Z. Cheng, A. Fouladi, G. Pollott, D. Werling, B. Sanz Bernardo, M. Salavati, L. Buggiotti, A. Wylie, M. Bell, M. Vaneetvelde, K. Hermans, G. Opsomer, S. Moerman, J. De Koster, H. Bogaert, J. Vandepitte, L. Vandevelde, B. Vanranst, J. Hoglund, S. Dahl, K. Ingvartsen, M. Sørensen, L. Foldager, S. Ostergaard, J. Rothmann, M. Krogh, E. Meyer, C. Gaillard, J. Ettema, T. Rousing, F. Signorelli, F. Napolitano, B. Moioli, A. Crisa, L. Buttazzoni, J. McClure, D. Matthews, F. Kearney, A. Cromie, M. McClure, S. Zhang, X. Chen, H. Chen, J. Zhao, L. Yang, G. Hua, C. Tan, G. Wang, M. Bonneau, A. Pompozzi, A. Pearn, A. Evertson, L. Kosten, A. Fogh, T. Andersen, M. Lucy,

C. Elsik, G. Conant, J. Taylor, N. Gengler, M. Georges, F. Colinet, M. Ramos Pamplona, H. Hammami, C. Bastin, H. Takeda, A. Laine, A.S. Van Laere, M. Schulze, and S. Palma-Vera. 2021. Global transcriptomic profiles of circulating leucocytes in early lactation cows with clinical or subclinical mastitis. Mol. Biol. Rep. 48:4611–4623. doi:10.1007/s11033-021-06494-8.

Cheng, Z., S. Palma-Vera, L. Buggiotti, M. Salavati, F. Becker, D. Werling, and D.C. Wathes. 2022. Transcriptomic Analysis of Circulating Leukocytes Obtained during the Recovery from Clinical Mastitis Caused by Escherichia coli in Holstein Dairy Cows. Animals 12. doi:10.3390/ani12162146.

Chin, C., S. Chen, H. Wu, C. Ho, M. Ko, and C. Lin. 2014. cytoHubba : identifying hub objects and sub- networks from complex interactome 8:1–7.

Clarke, K.R., and M. Ainsworth. 1993. A method of linking multivariate community structure to environmental variables. Mar. Ecol. Prog. Ser. 92:205–219. doi:10.3354/meps092205.

Dedhar, S. 2000. Cell-substrate interactions and signaling through ILK. Curr. Opin. Cell Biol. 12:250–256. doi:10.1016/S0955-0674(99)00083-6.

Devaiah, B.N., and D.S. Singer. 2013. CIITA and its dual roles in MHC gene transcription. Front. Immunol. 4:1–6. doi:10.3389/fimmu.2013.00476.

Dixon, P. 2003. VEGAN, a package of R functions for community ecology. J. Veg. Sci. 14:927. doi:10.1658/1100-9233(2003)014[0927:VAPORF]2.0.CO;2.

Dobin, A., C.A. Davis, F. Schlesinger, J. Drenkow, C. Zaleski, S. Jha, P. Batut, M. Chaisson, and T.R. Gingeras. 2013. STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29:15–21. doi:10.1093/bioinformatics/bts635.

Drier, Y., M. Sheffer, and E. Domany. 2013. Pathway-based personalized analysis of cancer. Proc. Natl. Acad. Sci. U. S. A. 110:6388–6393. doi:10.1073/pnas.1219651110.

Ebrahimie, E., F. Ebrahimi, M. Ebrahimi, S. Tomlinson, and K.R. Petrovski. 2018. A large-scale study of indicators of sub-clinical mastitis in dairy cattle by attribute weighting analysis of milk

composition features : highlighting the predictive power of lactose and electrical conductivity. J. Dairy Res. 85:193–200. doi:10.1017/S0022029918000249.

Ewels, P., S. Lundin, and K. Max. 2016. Data and text mining MultiQC : summarize analysis results for multiple tools and samples in a single report 32:3047–3048. doi:10.1093/bioinformatics/btw354.

Forsbäck, L., H. Lindmark-Månsson, A. Andrén, and K. Svennersten-Sjaunja. 2010. Evaluation of quality changes in udder quarter milk from cows with low-to-moderate somatic cell counts. Animal 4:617–626. doi:10.1017/S1751731109991467.

Gabrilovich, D.I., and S. Nagaraj. 2009. Myeloid-derived suppressor cells as regulators of the immune system. Nat. Rev. Immunol. 9:162–174. doi:10.1038/nri2506.

Günther, J., W. Petzl, I. Bauer, S. Ponsuksili, H. Zerbe, H.J. Schuberth, R.M. Brunner, and H.M. Seyfert. 2017. Differentiating Staphylococcus aureus from Escherichia coli mastitis: S. aureus triggers unbalanced immune-dampening and host cell invasion immediately after udder infection. Sci. Rep. 7:1–14. doi:10.1038/s41598-017-05107-4.

Guzzo, N., C. Sartori, and R. Mantovani. 2018. Genetic parameters of different measures of somatic cell counts in the Rendena breed. J. Dairy Sci. 101:8054–8062. doi:10.3168/jds.2017-14047.

Halasa, T., K. Huijps, O. Østerås, and H. Hogeveen. 2007. Economic effects of bovine mastitis and mastitis management: A review. Vet. Q. 29:18–31. doi:10.1080/01652176.2007.9695224.

Hauck, C.R., and K. Ohlsen. 2006. Sticky connections: Extracellular matrix protein recognition and integrin-mediated cellular invasion by Staphylococcus aureus. Curr. Opin. Microbiol. 9:5–11. doi:10.1016/j.mib.2005.12.002.

Heimes, A., J. Brodhagen, R. Weikard, H.M. Seyfert, D. Becker, M.M. Meyerholz, W. Petzl, H. Zerbe, M. Hoedemaker, L. Rohmeier, H.J. Schuberth, M. Schmicke, S. Engelmann, and C. Kühn. 2020. Hepatic Transcriptome Analysis Identifies Divergent Pathogen-Specific Targeting-Strategies to Modulate the Innate Immune System in Response to Intramammary Infection. Front. Immunol. 11. doi:10.3389/fimmu.2020.00715.

Hogue, C.W., and M. Groll. 2001. An automated method for finding molecular complexes in large protein interaction networks. BMC Bioinformatics 29:137–140.

Huma, Z.I., N. Sharma, S. Kour, S. Tandon, P.K. Guttula, S. Kour, A.K. Singh, R. Singh, and M.K. Gupta. 2020. Putative biomarkers for early detection of mastitis in cattle. Anim. Prod. Sci. 60:1669–1680. doi:10.1071/AN19539.

Ingvartsen, K.L., and K. Moyes. 2013. Nutrition, immune function and health of dairy cattle. Animal 7:112–122. doi:10.1017/S175173111200170X.

Ivaska, J., and J. Heino. 2000. Adhesion receptors and cell invasion: Mechanisms of integrin-guided degradation of extracellular matrix. Cell. Mol. Life Sci. 57:16–24. doi:10.1007/s000180050496.

Jensen, K., J. Günther, R. Talbot, W. Petzl, H. Zerbe, H.J. Schuberth, H.M. Seyfert, and E.J. Glass. 2013. Escherichia coli- and Staphylococcus aureus-induced mastitis differentially modulate transcriptional responses in neighbouring uninfected bovine mammary gland quarters. BMC Genomics 14. doi:10.1186/1471-2164-14-36.

Jha, A.K., S.C.C. Huang, A. Sergushichev, V. Lampropoulou, Y. Ivanova, E. Loginicheva, K. Chmielewski, K.M. Stewart, J. Ashall, B. Everts, E.J. Pearce, E.M. Driggers, and M.N. Artyomov. 2015. Network integration of parallel metabolic and transcriptional data reveals metabolic modules that regulate macrophage polarization. Immunity 42:419–430. doi:10.1016/j.immuni.2015.02.005.

Kabelitz, T., E. Aubry, K. van Vorst, T. Amon, and M. Fulde. 2021. The role of streptococcus spp. In bovine mastitis. Microorganisms 9:1–21. doi:10.3390/microorganisms9071497.

Krueger, F. 2019. TrimGalore.

Leonel Gonçalves, J., S. Hwa In Lee, E. de Paula Arruda, D. Pedroso Galles, V. Camargo Caetano, C.A. Fernandes de Oliveira, A.M. Fernandes, and M. Veiga dos Santos. 2015. Biofilm-producing ability and efficiency of sanitizing agents against Prototheca zopfii isolates from bovine subclinical mastitis. J. Dairy Sci. 98:3613–3621. doi:10.3168/jds.2014-9248.

Li, B., and C.N. Dewey. 2011. RSEM : accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinformatics 12:1–16.

Li, L., W. Tang, M. Zhao, B. Gong, M. Cao, and J. Li. 2021. Study on the Regulation Mechanism of Lipopolysaccharide on Oxidative Stress and Lipid Metabolism of Bovine Mammary Epithelial Cells. Physiol. Res. 70:777–785. doi:10.33549/physiolres.934682.

Love, M.I., W. Huber, and S. Anders. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome 15:1–21. doi:10.1186/s13059-014-0550-8.

Martin, P., H.W. Barkema, L.F. Brito, S.G. Narayana, and F. Miglior. 2018. Symposium review: Novel strategies to genetically improve mastitis resistance in dairy cattle. J. Dairy Sci. 101:2724–2736. doi:10.3168/jds.2017-13554.

Moyes, K.M., J.K. Drackley, D.E. Morin, M. Bionaz, S.L. Rodriguez-Zas, R.E. Everts, H.A. Lewin, and J.J. Loor. 2009. Gene network and pathway analysis of bovine mammary tissue challenged with Streptococcus uberis reveals induction of cell proliferation and inhibition of PPAR signaling as potential mechanism for the negative relationships between immune response and lipi. BMC Genomics 10:1–29. doi:10.1186/1471-2164-10-542.

Mudaliar, M., R. Tassi, F.C. Thomas, T.N. Mcneilly, S.K. Weidt, M. Mclaughlin, D. Wilson, R. Burchmore, P. Herzyk, P. David, and R.N. Zadoks. 2016. Molecular BioSystems in an experimental model of Streptococcus uberis. Mol. Biosyst. 12:2748–2761. doi:10.1039/C6MB00290K.

Naserkheil, M., F. Ghafouri, S. Zakizadeh, N. Pirany, Z. Manzari, S. Ghorbani, M.H. Banabazi, M.R. Bakhtiarizadeh, M.A. Huq, M.N. Park, H.W. Barkema, D. Lee, and K.S. Min. 2022. Multi-Omics Integration and Network Analysis Reveal Potential Hub Genes and Genetic Mechanisms Regulating Bovine Mastitis. Curr. Issues Mol. Biol. 44:309–328. doi:10.3390/cimb44010023.

Niedziela, D.A., P. Cormican, G. Foucras, F.C. Leonard, and O.M. Keane. 2021. Bovine milk somatic cell transcriptomic response to Staphylococcus aureus is dependent on strain genotype. BMC Genomics 22:1–20. doi:10.1186/s12864-021-08135-7.

N.M.C. 2017. Laboratory Handbook on Bovine Mastitis. third. National Mastitis Council, Inc.

Norberg, E., H. Hogeveen, I.R. Korsgaard, N.C. Friggens, K.H.M.N. Sloth, and P. Løvendahl. 2004. Electrical conductivity of milk: Ability to predict mastitis status. J. Dairy Sci. 87:1099–1107. doi:10.3168/jds.S0022-0302(04)73256-7.

Nyholm, L., A. Koziol, S. Marcos, A.B. Botnen, O. Aizpurua, S. Gopalakrishnan, M.T. Limborg, M.T.P. Gilbert, and A. Alberdi. 2020. Holo-Omics: Integrated Host-Microbiota Multi-omics for Basic and Applied Biological Research. iScience 23:101414. doi:10.1016/j.isci.2020.101414.

Overbey, E.G., A.M. Saravia-Butler, Z. Zhang, K.S. Rathi, H. Fogle, W.A. da Silveira, R.J. Barker, J.J. Bass, A. Beheshti, D.C. Berrios, E.A. Blaber, E. Cekanaviciute, H.A. Costa, L.B. Davin, K.M. Fisch, S.G. Gebre, M. Geniza, R. Gilbert, S. Gilroy, G. Hardiman, R. Herranz, Y.H. Kidane, C.P.S. Kruse, M.D. Lee, T. Liefeld, N.G. Lewis, J.T. McDonald, R. Meller, T. Mishra, I.Y. Perera, S. Ray, S.S. Reinsch, S.B. Rosenthal, M. Strong, N.J. Szewczyk, C.G.T. Tahimic, D.M. Taylor, J.P. Vandenbrink, A. Villacampa, S. Weging, C. Wolverton, S.E. Wyatt, L. Zea, S. V. Costes, and J.M. Galazka. 2021. NASA GeneLab RNA-seq consensus pipeline: standardized processing of short-read RNA-seq data. iScience 24. doi:10.1016/j.isci.2021.102361.

Pan, C., C. Yang, Y. Ma, H. Sheng, Z. Lei, S. Wang, H. Hu, X. Feng, J. Zhang, and Y. Ma. 2022. Identification of Key Genes Associated With Early Calf-Hood Nutrition in Subcutaneous and Visceral Adipose Tissues by Co-Expression Analysis. Front. Vet. Sci. 9:1–16. doi:10.3389/fvets.2022.831129.

Pegolo, S., R. Tessari, V. Bisutti, A. Vanzin, D. Giannuzzi, M. Gianesella, A. Lisuzzo, E. Fiore, A. Barberio, E. Schiavon, E. Trevisi, F. Piccioli Cappelli, L. Gallo, P. Ruegg, R. Negrini, and A. Cecchinato. 2022a. Quarter-level analyses of the associations among subclinical intramammary infection and milk quality, udder health, and cheesemaking traits in Holstein cows. J. Dairy Sci.. doi:10.3168/jds.2021-21267.

Pegolo, S., A. Toscano, V. Bisutti, D. Giannuzzi, A. Vanzin, A. Lisuzzo, F. Bonsembiante, M.E. Gelain, and A. Cecchinato. 2022b. Streptococcus agalactiae and Prototheca spp. induce different mammary gland leukocyte responses in Holstein cows. JDS Commun. 3:270–274. doi:10.3168/jdsc.2022-0216.

Petzl, W., H. Zerbe, J. Günther, H.M. Seyfert, J. Hussen, and H.J. Schuberth. 2018. Pathogen-specific responses in the bovine udder. Models and immunoprophylactic concepts. Res. Vet. Sci. 116:55–61. doi:10.1016/j.rvsc.2017.12.012.

Pilla, R., M. Malvisi, G.G.M. Snel, D. Schwarz, S. König, C.P. Czerny, and R. Piccinini. 2013. Differential cell count as an alternative method to diagnose dairy cow mastitis. J. Dairy Sci. 96:1653–1660. doi:10.3168/jds.2012-6298.

Rainard, P., P. Cunha, R.P. Martins, F.B. Gilbert, P. Germon, and G. Foucras. 2020. Type 3 immunity: a perspective for the defense of the mammary gland against infections. Vet. Res. 51:1–8. doi:10.1186/s13567-020-00852-3.

Rainard, P., C. Gitton, T. Chaumeil, T. Fassier, C. Huau, M. Riou, G. Tosser-Klopp, Z. Krupova, A. Chaize, F.B. Gilbert, R. Rupp, and P. Martin. 2018. Host factors determine the evolution of infection with Staphylococcus aureus to gangrenous mastitis in goats. Vet. Res. 49:1–17. doi:10.1186/s13567-018-0564-4.

Ranjan, R., D. Swarup, R.C. Patra, and D. Nandi. 2006. Bovine protothecal mastitis: A review. CAB Rev. Perspect. Agric. Vet. Sci. Nutr. Nat. Resour. 1:1–7. doi:10.1079/PAVSNNR20061017.

Robinson, M.D., D.J. Mccarthy, and G.K. Smyth. 2010. edgeR : a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26:139–140. doi:10.1093/bioinformatics/btp616.

Rohart, F., B. Gautier, A. Singh, and K.A. Lê Cao. 2017. mixOmics: An R package for 'omics feature selection and multiple data integration. PLoS Comput. Biol. 13:e1005752. doi:10.1371/journal.pcbi.1005752.

Roussel, P., P. Cunha, A. Porcherie, W. Petzl, F.B. Gilbert, C. Riollet, H. Zerbe, P. Rainard, and P.

Germon. 2015. Investigating the contribution of IL-17A and IL-17F to the host response during Escherichia coli mastitis. Vet. Res. 46:1–14. doi:10.1186/s13567-015-0201-4.

Le Roux, Y., F. Laurent, and F. Moussaoui. 2003. Polymorphonuclear proteolytic activity and milk composition change. Vet. Res. 34:629–645. doi:10.1051/vetres:2003021.

Ruegg, P.L. 2017. A 100-Year Review: Mastitis detection, management, and prevention. J. Dairy Sci. 100:10381–10397. doi:10.3168/jds.2017-13023.

Schukken, Y.H., J. Günther, J. Fitzpatrick, M.C. Fontaine, L. Goetze, O. Holst, J. Leigh, W. Petzl, H.J. Schuberth, A. Sipka, D.G.E. Smith, R. Quesnell, J. Watts, R. Yancey, H. Zerbe, A. Gurjar, R.N. Zadoks, and H.M. Seyfert. 2011. Host-response patterns of intramammary infections in dairy cows. Vet. Immunol. Immunopathol. 144:270–289. doi:10.1016/j.vetimm.2011.08.022.

Seo, M., H.J. Lee, K. Kim, K. Caetano-Anolles, J.Y. Jeong, S. Park, Y.K. Oh, S. Cho, and H. Kim. 2016b. Characterizing milk production related genes in holstein using RNA-seq. Asian-Australasian J. Anim. Sci. 29:343–351. doi:10.5713/ajas.15.0525.

Shave, C.D., L. Millyard, and R.C. May. 2021. Now for something completely different: Prototheca, pathogenic algae. PLoS Pathog. 17:1–7. doi:10.1371/JOURNAL.PPAT.1009362.

Singh, A., C.P. Shannon, B. Gautier, F. Rohart, M. Vacher, S.J. Tebbutt, and K.A.L. Cao. 2019. DIABLO: An integrative approach for identifying key molecular drivers from multi-omics assays. Bioinformatics 35:3055–3062. doi:10.1093/bioinformatics/bty1054.

Televičius, M., V. Juozaitiene, D. Malašauskienė, R. Antanaitis, A. Rutkauskas, M. Urbutis, and W. Baumgartner. 2021. Inline milk lactose concentration as biomarker of the health status and reproductive success in dairy cows. Agriculture 11:1–11. doi:10.3390/agriculture11010038.

Tur, J., T. Vico, J. Lloberas, A. Zorzano, and A. Celada. 2017. Macrophages and Mitochondria: A Critical Interplay Between Metabolism, Signaling, and the Functional Activity. 1st ed. Elsevier Inc.

Wang, D., L. Liu, S.M.A. Augustino, T. Duan, T.J. Hall, D.E. MacHugh, J. Dou, Y. Zhang, Y. Wang, and Y. Yu. 2020a. Identification of novel molecular markers of mastitis caused by

Staphylococcus aureus using gene expression profiling in two consecutive generations of Chinese Holstein dairy cattle. J. Anim. Sci. Biotechnol. 11:1–17. doi:10.1186/s40104-020-00494-7.

Wang, X., F. Su, X. Yu, N. Geng, L. Li, R. Wang, M. Zhang, J. Liu, Y. Liu, and B. Han. 2020b. RNA-Seq Whole Transcriptome Analysis of Bovine Mammary Epithelial Cells in Response to Intracellular Staphylococcus aureus. Front. Vet. Sci. 7:1–14. doi:10.3389/fvets.2020.00642.

Wathes, D.C., Z. Cheng, W. Chowdhury, M.A. Fenwick, R. Fitzpatrick, D.G. Morris, J. Patton, and J.J. Murphy. 2009. Negative energy balance alters global gene expression and immune responses in the uterus of postpartum dairy cows. Physiol. Genomics 39:1–13. doi:10.1152/physiolgenomics.00064.2009.

Welderufael, B.G., P. Løvendahl, D.J. de Koning, L.L.G. Janss, and W.F. Fikse. 2018. Genome-wide association study for susceptibility to and recoverability from mastitis in Danish Holstein cows. Front. Genet. 9:1–12. doi:10.3389/fgene.2018.00141.

Wieczorek, M., E.T. Abualrous, J. Sticht, M. Alvaro-Benito, S. Stolzenberg, F. Noé, and C. Freund. 2017. Major Histocompatibility Complex (MHC) Class i and MHC Class ii Proteins : Conformational Plasticity in Antigen Presentation. Front. Immunol. 8:1–16. doi:10.3389/fimmu.2017.00292.

Yu, J.W., and M.S. Lee. 2016. Mitochondria and the NLRP3 inflammasome: physiological and pathological relevance. Arch. Pharm. Res. 39:1503–1518. doi:10.1007/s12272-016-0827-4.

Zhang, H., H. Jiang, Y. Fan, Z. Chen, M. Li, Y. Mao, N.A. Karrow, J.J. Loor, S. Moore, and Z. Yang. 2018. Transcriptomics and iTRAQ-Proteomics Analyses of Bovine Mammary Tissue with Streptococcus agalactiae-Induced Mastitis. J. Agric. Food Chem. 66:11188–11196. doi:10.1021/acs.jafc.8b02386.

TABLES AND FIGURES

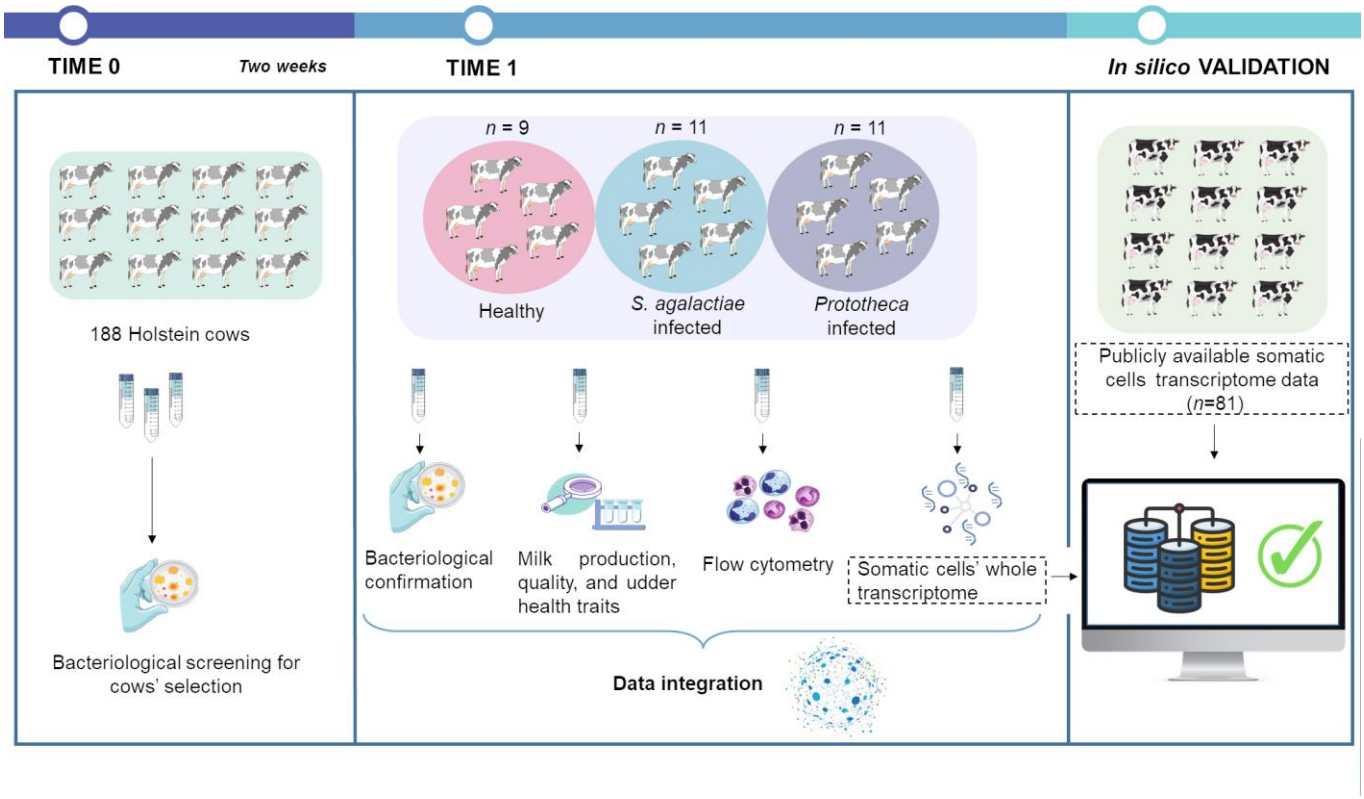


Fig.1. A schematic flow of the experimental design

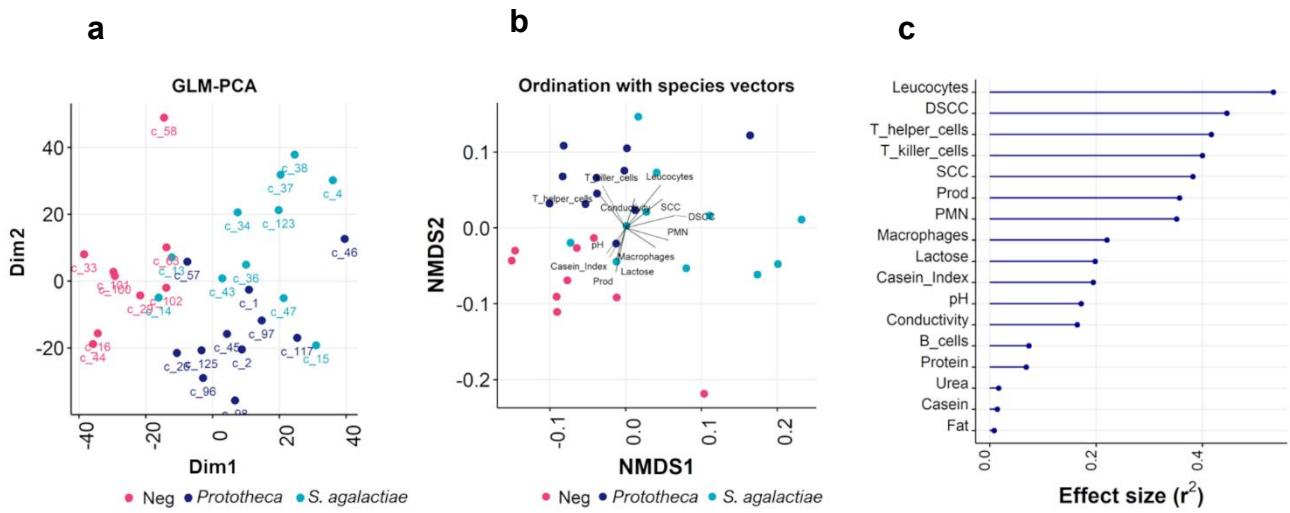


Fig. 2. Plots of samples spatial separation. **(a)** principal component analysis (PCA) of the 14,564 expressed genes; (**b**) non-metric multidimensional scaling (NMDS) for the visualization of the variation of the 14,564 expressed genes according to the phenotypic traits; (**c**) the principal contributors explaining the total variance of the transcriptome. SCC: somatic cell count; prod: milk yield (kg/d); DSCC: differential somatic cell count.

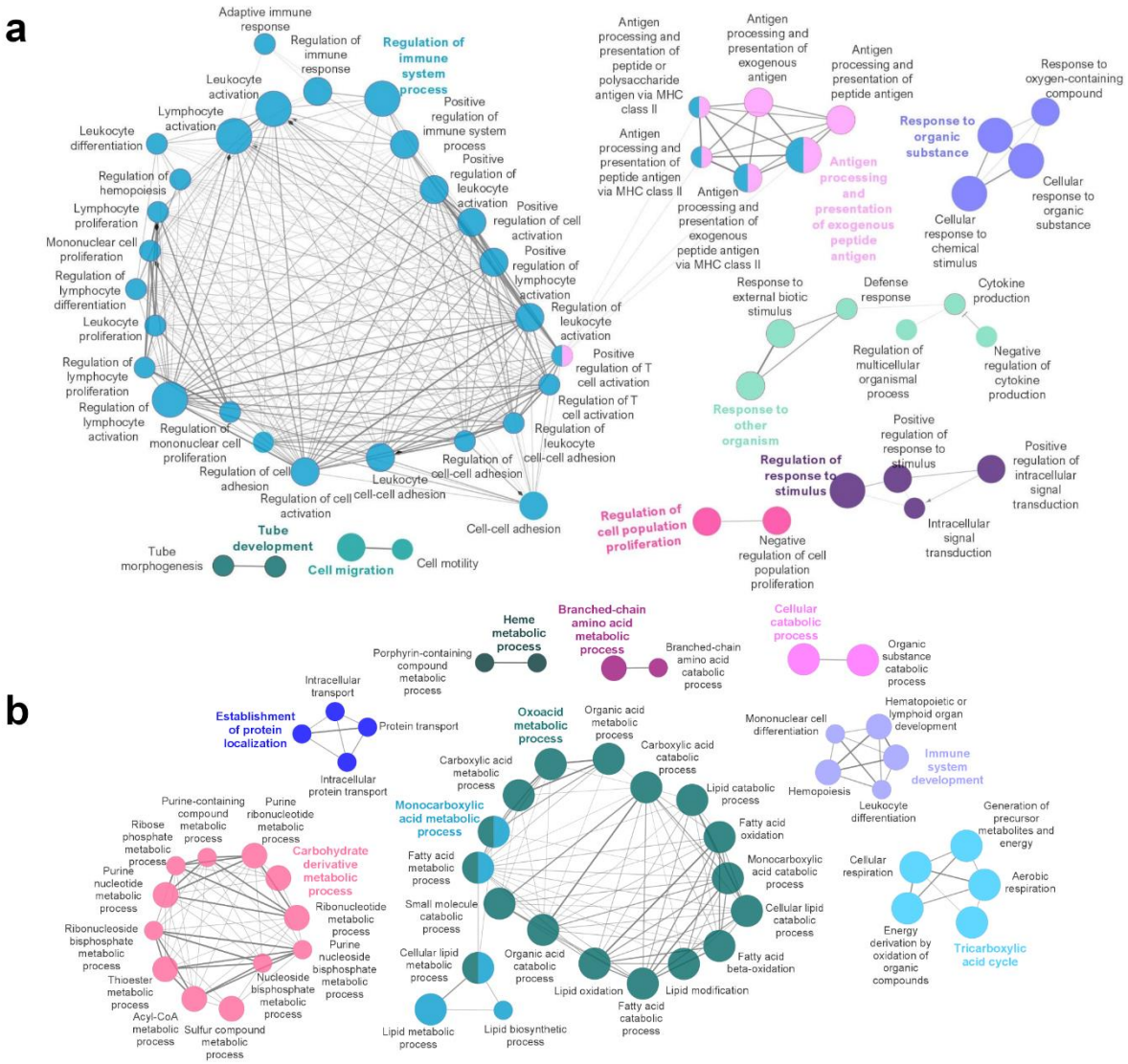

a
Adaptive immune response
Regulation of immune response
Regulation of immune system process
Leukocyte activation
Lymphocyte activation
Leukocyte differentiation
Positive regulation of immune system process
Regulation of hemopoiesis
Positive regulation of leukocyte activation
Lymphocyte proliferation
Positive regulation of cell activation
Mononuclear cell proliferation
Positive regulation of lymphocyte activation
Regulation of lymphocyte differentiation
Regulation of leukocyte activation
Leukocyte proliferation
Positive regulation of T cell activation
Regulation of lymphocyte proliferation
Regulation of T cell activation
Regulation of lymphocyte activation
Regulation of leukocyte cell-cell adhesion
Regulation of mononuclear cell proliferation
Regulation of cell-cell adhesion
Regulation of cell adhesion
Leukocyte cell-cell adhesion
Regulation of cell activation
Cell-cell adhesion
Tube development
Tube morphogenesis
Cell migration
Cell motility
Antigen processing and presentation of peptide or polysaccharide antigen via MHC class II
Antigen processing and presentation of exogenous antigen
Antigen processing and presentation of peptide antigen
Antigen processing and presentation of peptide antigen via MHC class II
Antigen processing and presentation of exogenous peptide antigen
Antigen processing and presentation of exogenous peptide antigen via MHC class II
Response to oxygen-containing compound
Response to organic substance
Cellular response to organic substance
Cellular response to chemical stimulus
Response to external biotic stimulus
Defense response
Cytokine production
Regulation of multicellular organismal process
Negative regulation of cytokine production
Response to other organism
Positive regulation of response to stimulus
Regulation of response to stimulus
Positive regulation of intracellular signal transduction
Intracellular signal transduction
Regulation of cell population proliferation
Negative regulation of cell population proliferation
b
Heme metabolic process
Branched-chain amino acid metabolic process
Cellular catabolic process
Porphyrin-containing compound metabolic process
Branched-chain amino acid catabolic process
Organic substance catabolic process
Intracellular transport
Establishment of protein localization
Protein transport
Intracellular protein transport
Oxoacid metabolic process
Organic acid metabolic process
Carboxylic acid metabolic process
Carboxylic acid catabolic process
Lipid catabolic process
Monocarboxylic acid metabolic process
Fatty acid oxidation
Fatty acid metabolic process
Monocarboxylic acid catabolic process
Small molecule catabolic process
Cellular lipid catabolic process
Cellular lipid metabolic process
Organic acid catabolic process
Fatty acid beta-oxidation
Lipid oxidation
Lipid modification
Fatty acid catabolic process
Lipid metabolic process
Lipid biosynthetic process
Mononuclear cell differentiation
Hematopoietic or lymphoid organ development
Immune system development
Hemopoiesis
Leukocyte differentiation
Generation of precursor metabolites and energy
Cellular respiration
Aerobic respiration
Energy derivation by oxidation of organic compounds
Tricarboxylic acid cycle
Purine-containing compound metabolic process
Purine ribonucleotide metabolic process
Ribose phosphate metabolic process
Carbohydrate derivative metabolic process
Purine nucleotide metabolic process
Ribonucleotide metabolic process
Ribonucleoside bisphosphate metabolic process
Purine nucleoside bisphosphate metabolic process
Thioester metabolic process
Nucleoside bisphosphate metabolic process
Acyl-CoA metabolic process
Sulfur compound metabolic process
c
Color Key
0
0.4
0.8
Value
PDS-Heatmap
Neg
S. agalactiae
Protothecа
c_58
c_4
c_46
c_123
c_15
c_34
c_37
c_38
c_16
c_44
c_102
c_101
c_33
c_100
c_29
c_97
c_117
c_47
c_1
c_36
c_43
c_13
c_63
c_14
c_26
c_57
c_2
c_45
c_125
c_96
c_98

Fig. 3. Pathway analysis of the differentially expressed genes in milk somatic cells. (**a**) ClueGO functionally grouped network of up and downregulated genes within the healthy and *Prototheca*'s infected cows. (**b**) ClueGO functionally grouped network of up and downregulated genes within the healthy and *S. agalactiae*'s infected cows. Terms are represented as nodes linked according to a kappa score ≥ 0.4. The node size means the term enrichment significance. Functionally related groups partially overlap. (**c**) heatmap built on the pathway deregulation scores (PDSs) of the whole transcriptomic data of healthy, *Prototheca*'s, and *S. agalactiae*'s infected animals. Each row corresponds to a pathway, and each column to a sample. Blue corresponds to "no deregulation", and yellow to high deregulation.

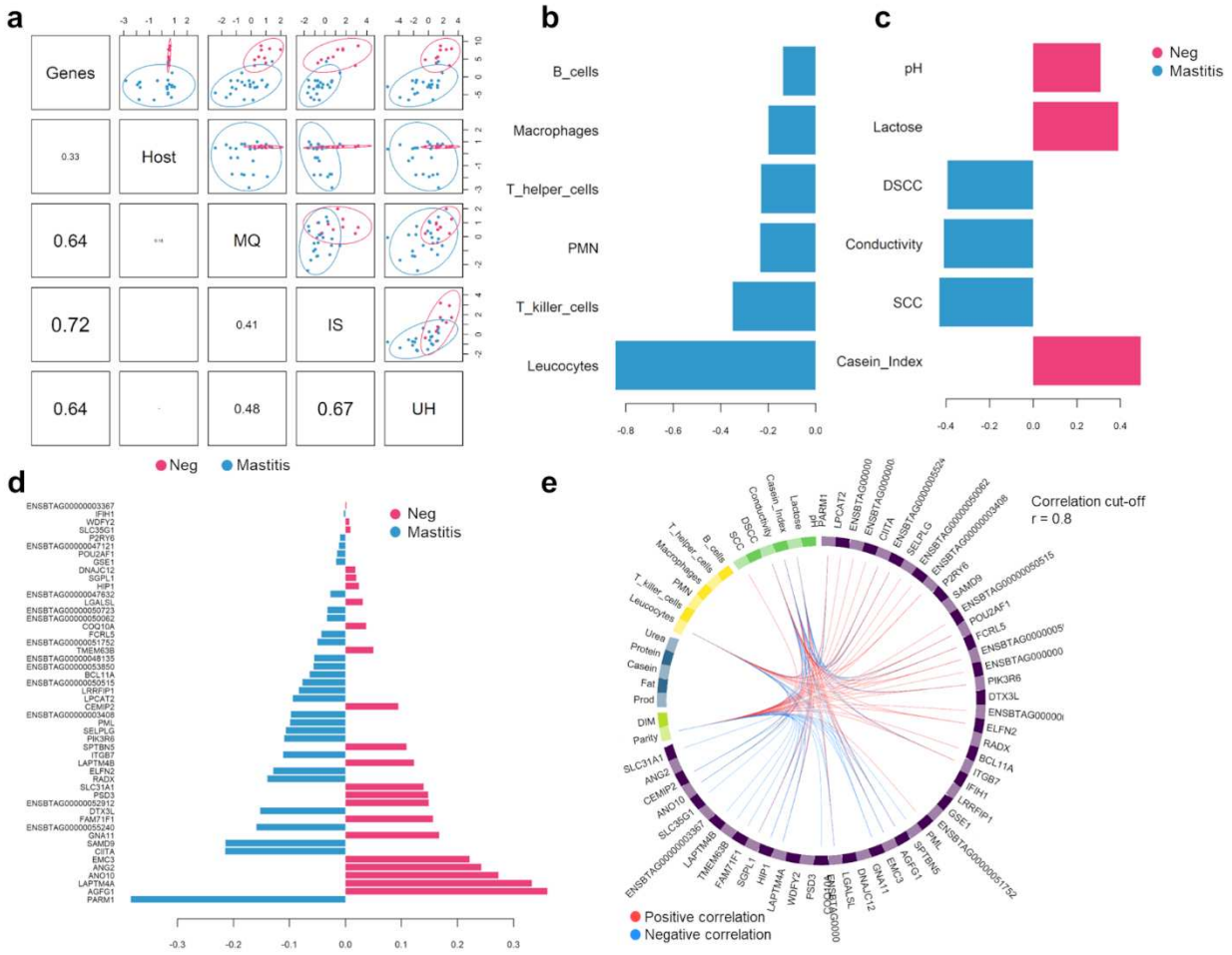


Fig. 4. Integration of transcriptomic and phenomic information using the DIABLO approach. (**a**) correlation plot among the different sets of categories; (**b-d**) the loading plots of flow cytometry immune cells (**b**) udder health traits (**c**) and differentially expressed genes (DEGs) (**d**). Pink and light blue bars represent healthy animals and cows with subclinical intramammary infection (sIMI), respectively. The negative values of the loading weights (light blue bars) signify that the corresponding variables had higher expression/value in infected animals. The positive values (pink bars) mean that related variables had higher expression/values in healthy animals. (**e**) Circos plot showing the correlation between candidate variables. MQ: milk quality, IS: immune system, UH: udder health.

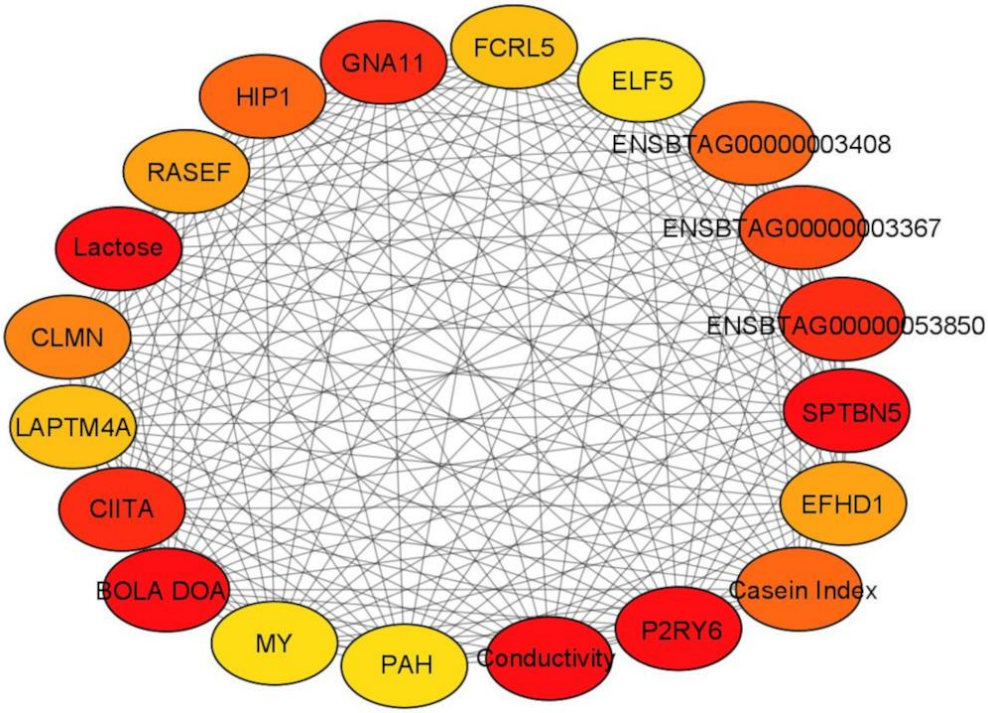


Fig. 5. Cytohubba top 20 hub variables according to the maximal clique centrality method (MCC). Higher ranking is represented by a redder color.

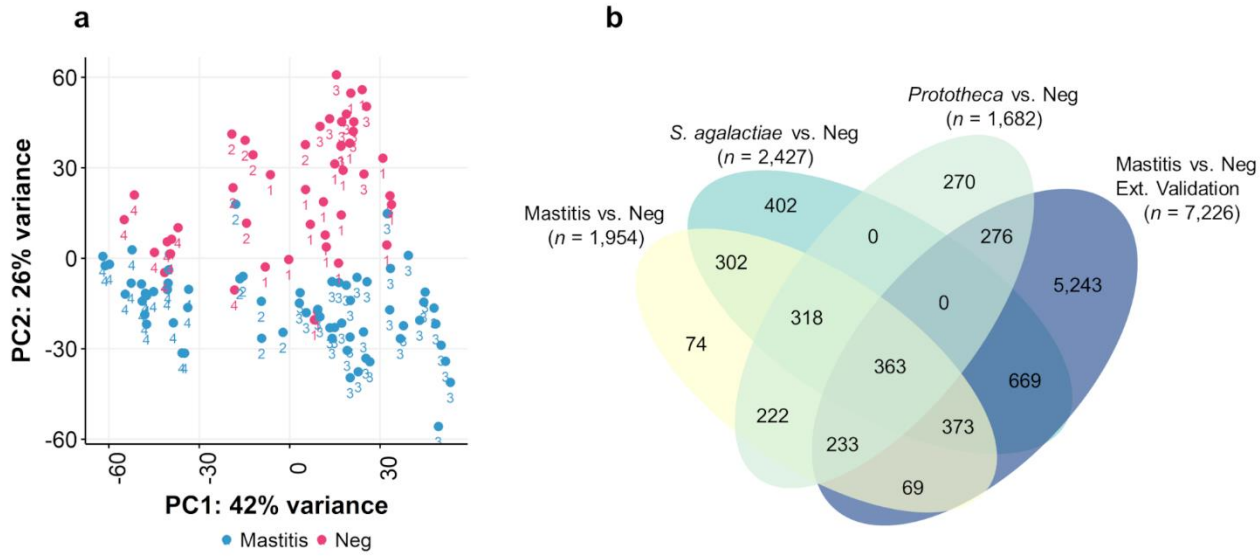


Fig. 6. The meta-analysis. (**a**) principal component analysis (PCA) of the present dataset and the ones downloaded from public repositories on milk somatic cell transcriptome. (**b**) Venn diagram of differentially expressed genes (DEGs) among the different experimental comparisons (*Prototheca* vs. healthy, *S. agalactiae* vs. healthy, mastitis vs. healthy, mastitis vs. healthy external data set). 1: transcriptomic data from the work of Seo et al. [38]; 2: transcriptomic data from the work of Asselstine et al. [14]; 3: transcriptomic data from the work of Niedziela et al. [15]; 4: transcriptomic data from the present dataset.

SUPPLEMENTARY INFORMATION

The Supplementary materials that are not reported in the present document (too large files) can be found at: https:// doi. org/ 10. 1186/ s40104- 023- 00890-9

Additional file 1: Table S1. Metadata of the study

Additional file 2: Table S2. Read mapping statistics of the transcriptome data before and after trimming. Quality control has been conducted using multiQC.

Additional file 3: Fig S1. Variation of milk phenotypic traits in healthy, *Prototheca*'s and *S. agalactiae*'s infected animals. Boxplots of milk yield (**a**), milk lactose (**b**), milk casein index (**c**), milk conductivity (**d**), milk protein (**e**), milk casein (**f**), milk fat (**g**), milk urea (**h**) and milk pH (**i**) according to the three experimental groups.

Additional file 4: Fig. S2. Flow cytometry results for immune cell populations according to the three experimental groups. Violin plots of milk somatic cell score (SCS) (**a**), differential somatic cell count (DSCC) (**b**), leucocytes (**c**), macrophages (**d**), T helper cells (**e**), T killer cells (**f**), B cells (**g**) and, PMN (**h**) in healthy, *Prototheca's* and *S. agalactiae's* infected animals. (**i**) Principal component analysis (PCA) shows the samples' separation according to the flow cytometry variables.

Additional file 5: Table S3. Matrix of the 14,564 expressed genes obtained after filtering for counts per million (CPM) > 0.5 in at least 10 samples. Genes failing these criteria were removed from the exploratory and DEGs analyses.

Additional file 6: Table S4. List of the DEGs for each experimental comparison (*Prototheca* vs. healthy, *S. agalactiae* vs. healthy, *Prototheca* vs. *S. agalactiae*, healthy vs. mastitis, healthy vs. mastitis in the meta-analysis cohort, core mastitis response genes).

Additional file 7: Fig. S3. ClueGo pathway analysis of the 681 "core mastitis response genes" commonly shared between *S. agalactiae* and *Prototheca.*

Additional file 8: Fig. S4. PPI network construction and module analysis carried out with Cytoscape's plug-in MCODE. Nodes belonging to different modules are differently colored. White nodes are variables that were not assigned to any modules. Lines represent the interaction between nodes (red arrows are positive interactions, and blue are negative).

Additional file 9: Table S5. Full Maximal Clique Centrality (MCC) ranking of the 20 hub variables obtained with cytohubba.

Additional file 10: Table S6. Predictive performances obtained through ROC analysis for the hub genes.

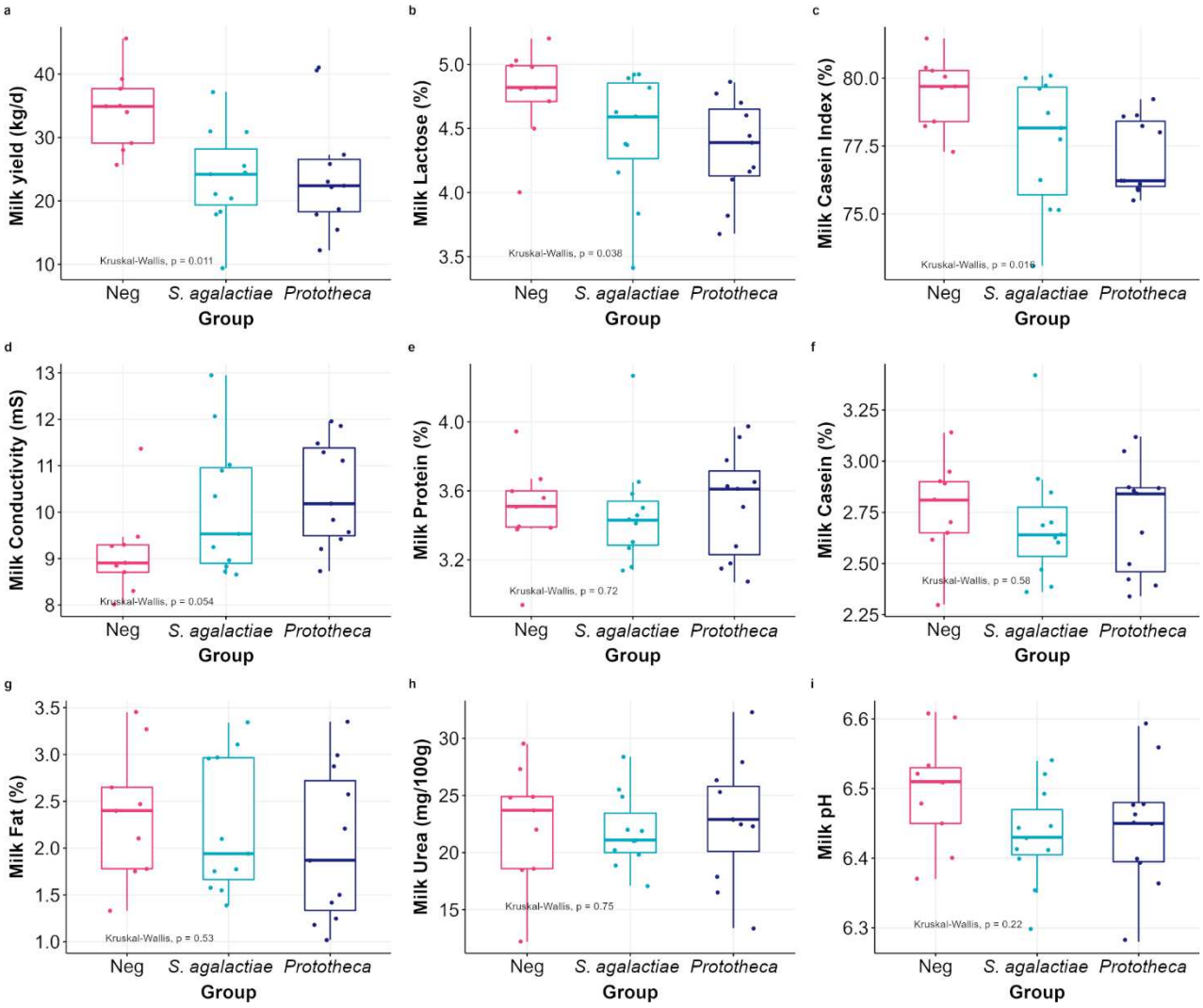


Fig S1. Variation of milk phenotypic traits in healthy, *Prototheca*'s and *S. agalactiae*'s infected animals. Boxplots of milk yield (**a**), milk lactose (**b**), milk casein index (**c**), milk conductivity (**d**), milk protein (**e**), milk casein (**f**), milk fat (**g**), milk urea (**h**) and milk pH (**i**) according to the three experimental groups.

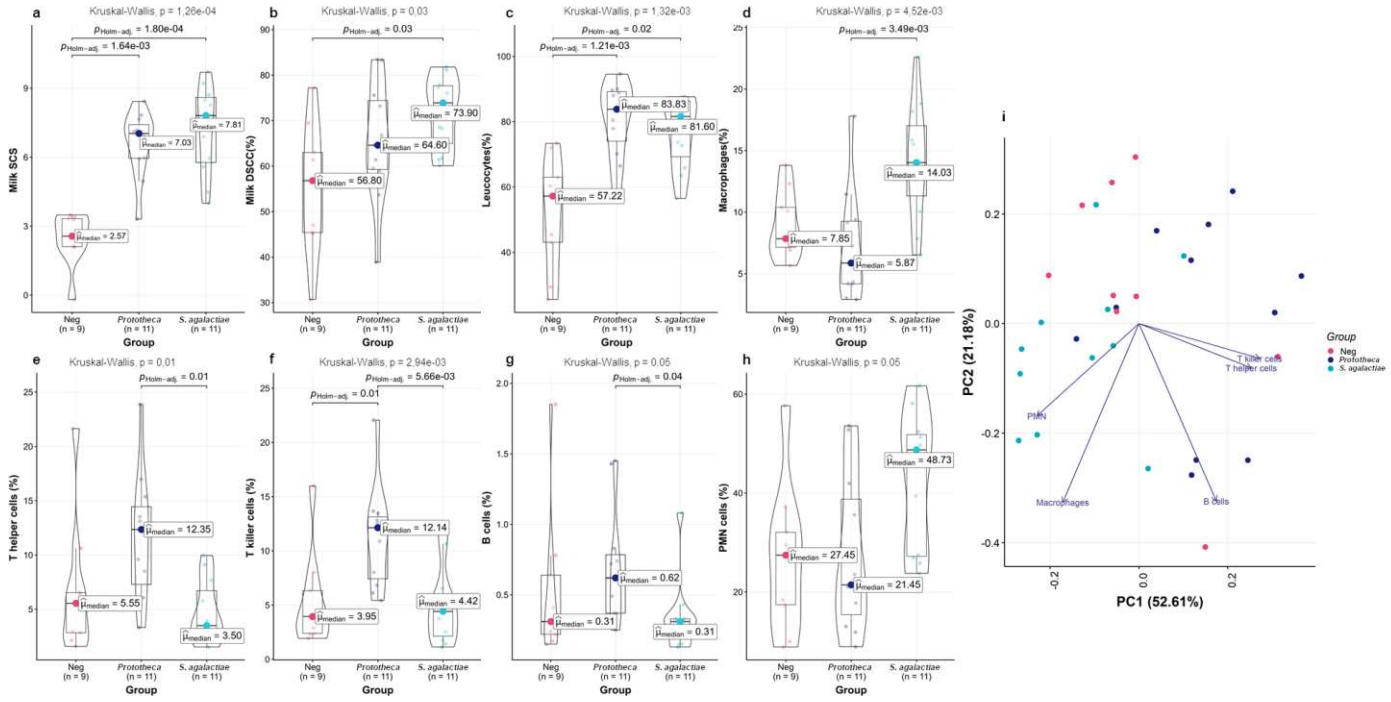


Fig. S2. Flow cytometry results for immune cell populations according to the three experimental groups. Violin plots of milk somatic cell score (SCS) (**a**), differential somatic cell count (DSCC) (**b**), leucocytes (**c**), macrophages (**d**), T helper cells (**e**), T killer cells (**f**), B cells (**g**) and, PMN (**h**) in healthy, *Prototheca's* and *S. agalactiae's* infected animals. (**i**) Principal component analysis (PCA) shows the samples' separation according to the flow cytometry variables.

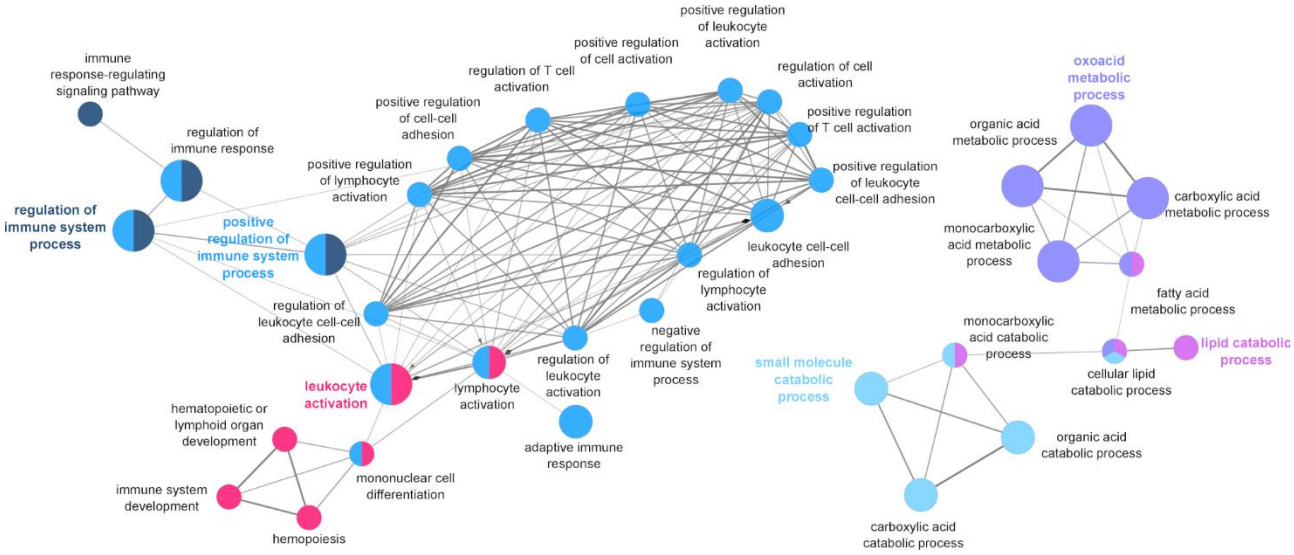


Fig. S3. ClueGo pathway analysis of the 681 "core mastitis response genes" commonly shared between *S. agalactiae* and *Prototheca.*

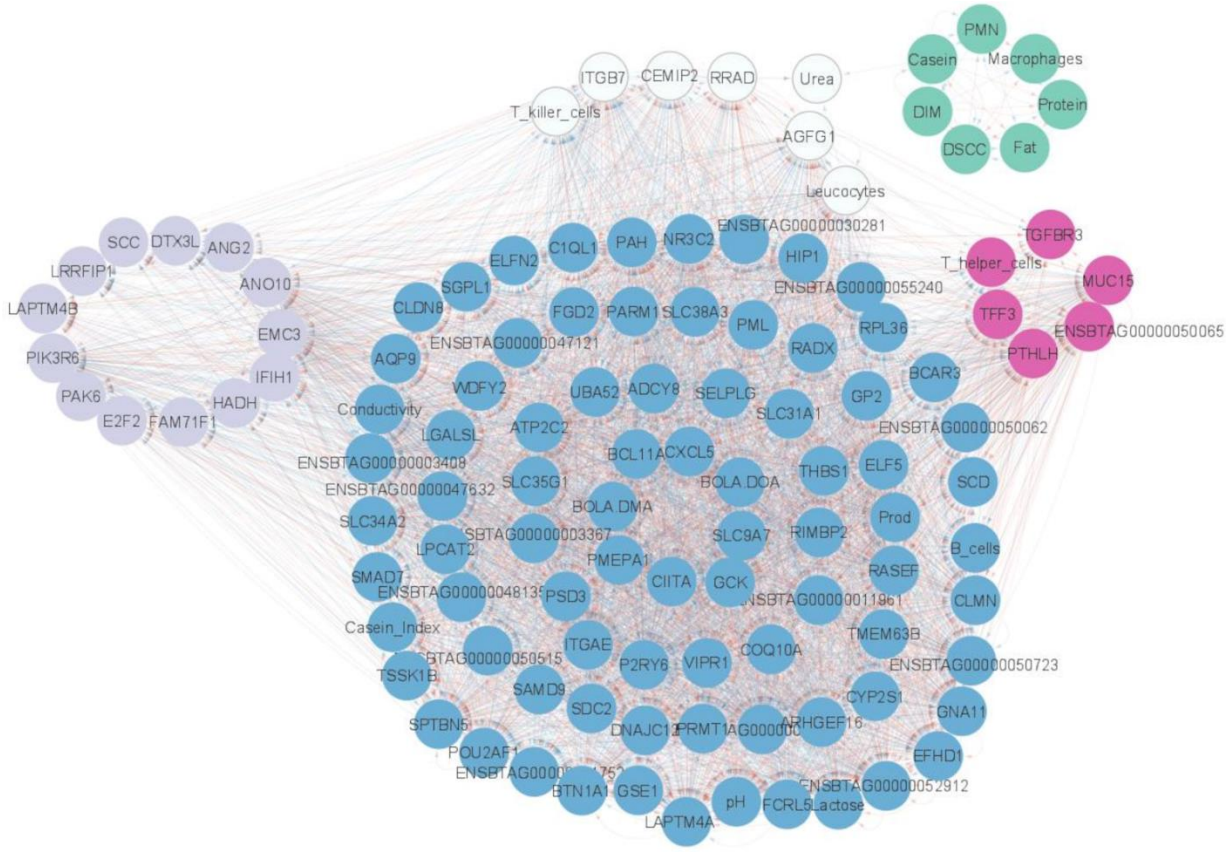


Fig. S4. PPI network construction and module analysis carried out with Cytoscape's plug-in MCODE. Nodes belonging to different modules are differently colored. White nodes are variables that were not assigned to any modules. Lines represent the interaction between nodes (red arrows are positive interactions, and blue are negative).

GENERAL CONCLUSIONS AND FUTURE PERSPECTIVES

Our work provided new insights on the effect of udder inflammation and subclinical IMI on both phenotypic (i.e., milk protein profile) and molecular level. With the first contribution we focused on the evaluation of two phenotypic indicators commonly used for the screening of mastitis, the SCC and DSCC, and we evaluated their effect on the detailed milk protein profile in a cohort of clinically healthy cows. Our results firstly confirmed the unfavorable effect of increased SCC on milk protein fractions, especially on caseins and shed some new light on the association between DSCC and the protein profile. Indeed, when we expressed the DSCC parameter as count of PMN + LYM and MAC we observed that the greater impairment for the protein fractions was related to proteolytic activity derived from macrophages, as they were present in higher number given the non-pathological status of the animals. With the second contribution, we dug deeper on the behavior of milk protein fractions in animals having subclinical IMI by exploring in an extensive way the pattern of alteration at the mammary gland at the quarter level. Our results showed that subclinical IMI significantly impaired particularly the caseins fractions while no important alteration was observed for the whey proteins proportion. However, we have substantiated that the increase in SCC seemed to be the factor associated with the most significant changes in the protein fractions. This suggests that the increase in endogenous proteolytic enzymes related to the onset of inflammation, and not the infection (i.e., the bacteriological status), might be the pivotal aspect driving the alteration of the milk protein profile.

The third contribution, shifting from the phenotypic to the molecular level, allowed us to shed some light on the molecular mechanisms behind subclinical IMI and on the immune reaction of the animal in response to the infection of two different microorganisms: *S. agalactiae* and *Prototheca*. Our results showed that, even though we observed some differences in terms of immune-related pathways and gene expression between the two infections, there was a strong "core" immune response that was commonly shared by the two pathogens, probably linked to the subclinical nature of infection. Finally, with the integration of the transciptomic and phenotypic information we observed

a strong correlation between the transcriptome, the immune cells and some udder traits variables (i.e., lactose and milk conductivity). Also, we identified some putative genes that could be informative for the animal's response to subclinical mastitis. In the future, we will aim at validating these potential biomarkers in a wider cohort of animals.

To conclude, it is worth noting that while individual data layers provide valuable information on their own, much stronger insights could be obtained by integrating different sources of information which could offer a better comprehension of the complexity of this disease. For these reasons, our future challenge will be to perform a multi-layer omic integration including together with the transcriptomic and phenotypic data, also peptidomic, metagenomic and DNA-methylation information to possibly identify hub variables to be used for screening purpose at the herd level.

LIST OF PUBLICATIONS

In press

Bisutti V., Vanzin A., Pegolo S., Toscano A., Gianesella M., Sturaro E., Schiavon S., Gallo L., Tagliapietra F., Giannuzzi D., Cecchinato A. Effect of intramammary infection and inflammation on milk protein profile assessed at the quarter level in Holstein cows. *Accepted in the JDS.*

Vanzin A., Franchin C., Arrigoni G., Battisti I., Masi A., Bisutti V., Giannuzzi D., Cecchinato A., Gallo L., Pegolo S. 2023 Streptococcus agalactiae and Prototheca spp. Infections induce changes in milk peptidome in Holstein cattle. *Accepted in the Journal of Agricoltural and Food Chemistry.*

Lisuzzo A., Laghi L., Fiore E., Cecchinato A., Bisutti V., Pegolo S., Giannuzzi D., Tessari R., Barberio A., Schiavon E., Mazzotto E., Tommasoni C., Gianesella M., Blood metabolome differences associated to subclinical intramammary infection due to Streptococcus agalactiae and Prototheca spp. in dairy cows. *Accepted in the JDS.*

Published

Catellani, A., F. Ghilardelli, E. Trevisi, A. Cecchinato, V. Bisutti, F. Fumagalli, H.V.L.N. Swamy, Y. Han, S. Van Kuijk, and A. Gallo. 2023. Effects of Supplementation of a Mycotoxin Mitigation Feed Additive in Lactating Dairy Cows Fed Fusarium Mycotoxin-Contaminated Diet for an Extended Period. Toxins (Basel). 15:1–20. doi:https://doi.org/10.3390/toxins15090546.

Giannuzzi, D., A. Vanzin, S. Pegolo, A. Toscano, V. Bisutti, L. Gallo, S. Schiavon, and A. Cecchinato. 2023. Novel insights into the associations between immune cell population distribution in mammary gland and milk minerals in Holstein cows. J. Dairy Sci. Revis.. doi:10.3168/jds.2023-23729.

Pegolo, S., A. Toscano, V. Bisutti, A. Vanzin, D. Giannuzzi, L. Gallo, F. Tagliapietra, P. Ruegg, M. Gianesella, M. Mele, S. Schiavon, and A. Cecchinato. 2023. Associations between subclinical intramammary infection and milk fatty acid profile at the quarter level in Holstein cattle. Animal 17:100978. doi:10.1016/j.animal.2023.100978.

Toscano, A., D. Giannuzzi, S. Pegolo, A. Vanzin, V. Bisutti, L. Gallo, E. Trevisi, A. Cecchinato, and S. Schiavon. 2023. Associations between the detailed milk mineral profile, milk composition, and metabolic status in Holstein cows. J. Dairy Sci.. doi:10.3168/jds.2022-23161.

Bisutti, V., N. Mach, D. Giannuzzi, A. Vanzin, E. Capra, R. Negrini, M.E. Gelain, A. Cecchinato, P.A. Marsan, and S. Pegolo. 2023. Transcriptome - wide mapping of milk somatic cells upon subclinical mastitis infection in dairy cattle. J. Anim. Sci. Biotechnol. 9:1–19. doi:10.1186/s40104-023-00890-9.

Macedo Mota, L.F., V. Bisutti, A. Vanzin, S. Pegolo, A. Toscano, S. Schiavon, F. Tagliapietra, L. Gallo, P. Ajmone Marsan, and A. Cecchinato. 2023. Predicting milk protein fraction using infrared spectroscopy and a gradient boosting machine for breeding purposes in Holstein cattle. J. Dairy Sci. 106:1853–1873. doi:10.3168/jds.2022-22119.

Gislon, G., Bava, L., Bisutti, V., Tamburini, A., & Brasca, M. (2023). Bovine beta casein polymorphism and environmental sustainability of cheese production: The case of Grana Padano PDO and mozzarella cheese. Sustainable Production and Consumption, 35, 85–94. https://doi.org/10.1016/j.spc.2022.10.017

Bisutti, V., Vanzin, A., Toscano, A., Pegolo, S., Giannuzzi, D., Gallo, L., Trevisi, E., Negrini, R., Cecchinato, A., Tagliapietra, F., & Schiavon, S. (2022). Impact of somatic cell count combined with differential somatic cell count on milk protein fractions in Holstein cattle. Journal of Dairy Science. https://doi.org/10.3168/jds.2022-2207

Mota, L. F. M., Giannuzzi, D., Bisutti, V., Pegolo, S., Trevisi, E., Schiavon, S., Gallo, L., Fineboym, D., Katz, G., & Cecchinato, A. (2022). Real-time milk analysis integrated with stacking

ensemble learning as a tool for the daily prediction of cheese-making traits in Holstein cattle. Journal of Dairy Science, 105(5), 4237–4255. https://doi.org/10.3168/jds.2021-21426

Pegolo, S., Toscano, A., Bisutti, V., Giannuzzi, D., Vanzin, A., Lisuzzo, A., Bonsembiante, F., Gelain, M. E., & Cecchinato, A. (2022). Streptococcus agalactiae and Prototheca spp. induce different mammary gland leukocyte responses in Holstein cows. JDS Communications, 3(4), 270–274. https://doi.org/10.3168/jdsc.2022-0216

Bisutti, V., Pegolo, S., Giannuzzi, D., Mota, L. F. M., Vanzin, A., Toscano, A., Trevisi, E., Ajmone Marsan, P., Brasca, M., & Cecchinato, A. (2022). The β-casein (CSN2) A2 allelic variant alters milk protein profile and slightly worsens coagulation properties in Holstein cows. Journal of Dairy Science. https://doi.org/10.3168/jds.2021-21537

Pegolo, S., Tessari, R., Bisutti, V., Vanzin, A., Giannuzzi, D., Gianesella, M., Lisuzzo, A., Fiore, E., Barberio, A., Schiavon, E., Trevisi, E., Piccioli Cappelli, F., Gallo, L., Ruegg, P., Negrini, R., & Cecchinato, A. (2022). Quarter-level analyses of the associations among subclinical intramammary infection and milk quality, udder health, and cheesemaking traits in Holstein cows. Journal of Dairy Science. https://doi.org/10.3168/jds.2021-21267

Pegolo, S., Mota, L. F. M., Bisutti, V., Giannuzzi, D., Gallo, L., Schiavon, S., Tagliapietra, F., Chion, A. R., Trevisi, E., Negrini, R., Marsan, P. A., & Cecchinato, A. (2021). Genetic parameters of differential somatic cell count, milk composition, and cheese-making traits measured and predicted using spectral data in Holstein cows. Journal of Dairy Science.

Pegolo, S., Yu, H., Morota, G., Bisutti, V., Rosa, G. J. M., Bittante, G., & Cecchinato, A. (2021). Structural equation modeling for unraveling the multivariate genomic architecture of milk proteins in dairy cattle. Journal of Dairy Science, 104(5), 5705–5718. https://doi.org/10.3168/jds.2020-18321

Pegolo, S., Giannuzzi, D., Bisutti, V., Tessari, R., Gelain, M.E., Gallo, L., Schiavon, S., Tagliapietra, F., Trevisi, E., Ajmone Marsan, P., Bittante, G., Cecchinato, A., 2021. Associations between differential somatic cell count and milk yield, quality, and technological characteristics in Holstein cows. J. Dairy Sci. 104, 4822–4836. https://doi.org/10.3168/jds.2020-19084

Scientific national and international congresses

Bisutti V., Giannuzzi D., Vanzin A., Toscano A., Gianesella M., Pegolo S., Cecchinato A. Effects of intramammary infection on milk protein profile measured at the quarter level in dairy cows. In EAAP2023 congress, Lyon, France, 28/08-01/09/2023 (Poster).

Pegolo S., Giannuzzi D., Vanzin A., Bisutti V., Toscano A., Capra E., Ajmone Marsan P., Cecchinato A. DNA-methylome analysis of milk somatic cells upon subclinical mastitis in Holstein cattle. In EAAP2023 congress, Lyon, France, 28/08-01/09/2023 (Oral communication).

Giannuzzi D., Vanzin A., Toscano A., Bisutti V., Pegolo S., Cecchinato A. Associations between the detailed milk mineral profile and total and differential cell count in Holstein Friesian dairy cows. In: ASPA2023 congress, Monopoli, 13-16/06/2023 (oral communication).

Pegolo S., Giannuzzi D., Bisutti V., Ramirez Mauricio M. A., Macedo Mota L. F., Ajmone Marsan P., Trevisi E., Cecchinato A. Genetic parameters for blood indicators of metabolic distress and milk traits in dairy cattle. In: ASPA2023 congress, Monopoli, 13-16/06/2023 (oral communication).

Bisutti V., Mach N., Giannuzzi D., Toscano A., Capra E., Cecchinato A., Pegolo S. Transcriptomic and phenomic data integration to identify markers of subclinical mastitis in Holstein cattle. In: ASPA2023 congress, Monopoli, Italy, 15-16/06/2023.

Vanzin A., Battisti I., Franchin C., Bisutti V., Toscano A., Giannuzzi D., Masi A., Arrigoni G., Gallo L., Cecchinato A., Pegolo S. Mass spectrometry-based characterization of the bovine milk peptidome upon Streptococcus agalactiae and Prototheca spp. Infection. In: ASPA2023 congress, Monopoli, Italy, 15-16/06/2023.

Bisutti V., Mach N., Giannuzzi D., Cecchinato A., Negrini R., Ajmone-Marsan P., Pegolo S. Transcriptomic map of mastitis in dairy cows: clues on S. agalactiae and Prototheca infections. In EAAP2022 congress, Porto, Portugal 05-09/09/2022 (Oral communication).

Pegolo S., Toscano A., Bisutti V., Gianesella M., Giannuzzi D., Schiavon S., Gallo L., Tagliapietra F., Cecchinato A. Sub-clinical mastitis affects milk fatty acid profile in Holstein cattle. In EAAP2022 congress, Porto, Portugal 05-09/09/2022 (Oral communication).

Toscano A., Giannuzzi D., Pegolo S., Vanzin A., Bisutti V., Cecchinato A., Schiavon S. Milk mineral profile is associated with blood energy metabolites in Holstein dairy cattle. In III Convegno AISSA#under40, Bolzano, 15-15/07/2022 (Poster).

Pegolo S., Toscano A., Bisutti V., Giannuzzi D., Vanzin A., Gelain M.E., Bonsembiante F., Cecchinato A. Cytofluorimetric analyses revealed a pathogen-specific immune modulation in response to intramammary infection in Holstein cows. In: III Convegno AISSA#under40, Bolzano, 15-15/07/2022 (Oral communication).

Bisutti V., Vanzin A., Toscano A., Pegolo S., Trevisi E., Ajmone Marsan P., Negrini R., Cecchinato A. Associations between udder health traits and milk protein fractions in Holstein cows. In: EAAP2021 congress, Davos, Switzerland 30/08-04/09/2021 (Oral communication).

Bisutti V., Vanzin A., Toscano A., L.F.M. Mota. Associations among somatic cell count, differential somatic cell count and milk protein profile in Holstein cattle. In: ASPA2021 congress, Padova, 21-24/09/2021 (oral communication).

Vanzin A., Toscano A., Bisutti V., Pegolo S., Giannuzzi D., Cecchinato A. Associations between udder health traits and milk mineral profile in Holstein Friesian cattle. In: ASPA2021 congress, Padova, 21-24/09/2021 (poster).

S. Pegolo, R. Tessari, V. Bisutti, E. Fiore, M. Gianesella, D. Giannuzzi, S. Schiavon, L. Gallo, E. Trevisi, P. Ajmone-Marsan, R. Negrini, Cecchinato. The impact of sub-clinical mastitis infections on technological traits in Holstein cattle. In: EAAP2021 congress, Davos, Switzerland 30/08-04/09/2021 (Oral communication).

D. Giannuzzi, R. Tessari, E. Fiore, M. Gianesella, V. Bisutti, S. Pegolo, “Association between liver ultrasound measurements and hematochemical indicators for early prediction of metabolic alterations in Holstein cows”. In: ASPA2021 congress, Padova, 21-24/09/2021 (oral comunication).

ACKNOWLEDGMENTS

First, I would like to express my gratitude to my supervisor, Prof. Alessio Cecchinato, and my co-supervisor, Dr. Sara Pegolo, for their support, guidance, and for providing me with the opportunity to participate in this project and join the team. Additionally, I am thankful for the chance to work with Dr. Nuria Mach during my time abroad in France at the ENVT. She taught me a great deal and helped me interpret complex results.

A special thanks goes to Filippo, who supported me throughout this period, and to my family, particularly my mom, who always had my back and inspired me with her hard work and dedication.

But most importanltly, I am grateful for having met my colleagues, with whom I have shared most of my days (both in and out of the office). They have really become true friends and a continuous source of inspiration and laughter.