Gene expression profiling of bovine raw milk as a new tool to monitor the inflammatory status of the udder : a pilot study

C. Gitton[a], Y. Le Vern[a], M. Gaborit[b], R. Prado Martins[a,1], P. Germon[a]*

[a] ISP UMR 1282, INRAE, Université François Rabelais de Tours, Nouzilly, France
[b] UEP UE 0326, INRAE, Borculo, Le Pin au Haras, France.

[1] Present address: PFIE UE 1277, INRAE, Nouzilly, France

* Corresponding author
Email: pierre.germon@inrae.fr

Abstract

Inflammation of the mammary gland, or mastitis, is a common disease of dairy cows with important consequences on the economy of dairy farms, well-being of cows and antibiotic usage. Monitoring of the inflammatory status of mammary gland currently relies on milk somatic cell counting. In order to improve this monitoring, methods aiming at differentiating cell types among milk cells are being developed. The objective of the present study was to explore the use of gene expression profiling from raw milk samples to define the inflammatory status of udder quarters. A pilot study allowed us to select 36 genes among 95 whose expression were associated with one of three milk cell types, i.e. neutrophils, macrophages or T lymphocytes, purified by cell sorting. Based on expression data measured by RT-qPCR for these genes, we were able to identify four different types of milk samples, each with specific properties. One cluster was characterized by increased expression of genes typifying neutrophils, a second one was identified by increased expression of genes typifying macrophages while the two others were distinguished based on the expression of genes associated with T lymphocytes. Clustering of samples based on gene expression was shown to be comparable to that obtained by flow cytometry analysis. Our results highlight the potential of profiling of milk samples based on gene expression for the analysis of the inflammatory status of udders.

Keywords

Mastitis, inflammation, milk, gene expression profiling

Introduction

Mastitis remains a major issue in dairy farming. Upon entry and recognition of bacterial pathogens by the host immune cells, in particular mammary epithelial cells and macrophages, an inflammatory response is triggered which leads to recruitment of immune cells in the mammary gland.

As a consequence, somatic cell counts (SCC) in milk is routinely used as an indirect indicator of the sanitary status of mammary glands.

SCC values below 100,000 cells per ml are generally considered as an indication of a healthy mammary quarter (Hillerton; Harmon, 1994). On the opposite, SCC values above 200,000 cells/ml are generally taken as an indication of an on-going inflammatory response or a quarter recovering from infection (Hillerton; Smith and Hillerton, 2001). In between these two SCC values one cannot state with confidence whether a quarter is healthy or not. In addition, setting SCC threshold values to distinguish healthy from infected quarters should also take into account different criteria such as the lactation stage and parity (Mondini et al., 2025). In any case, these two values are not definite indicators of a health status: mastitic quarters below SCC values of 100,000 could be still be infected by mastitis pathogens (Schwarz et al., 2010).

As a consequence, several improvements have been tested, in particular the differential cell counts that determines the type of cells present in milk upon inflammation (Rivas et al., 2001; Pilla et al., 2013). Cells present in milk belong to different cell types including neutrophils, macrophages, lymphocytes and epithelial cells. The proportion and quantities of each of these cell population is highly variable. For instance, the strong inflammation observed during clinical mastitis is characterized by a majority of neutrophils which can amount to almost 90% of cells present in milk (Sordillo et al., 1997).

Reports on the cell composition of mammary gland secretions from healthy animals show contrasting results. Several studies, using electronic microscope examinations or flow cytometry analyses of milk samples from healthy quarters, revealed mostly macrophages (50- 80%) with 10-20 % lymphocytes (Lee et al., 1980; Lun et al., 2007; Souza et al., 2020; Farschtschi et al., 2021; Silva et al., 2021). Yet, contrasting results have been presented in which T lymphocytes are the most prevalent cell type in the absence of infections with almost 50% T lymphocytes and around 20-30 % macrophages (Park et al., 1992; Dosogne et al., 2003; Schwarz et al., 2011).

These discrepancies might be due to different methodologies, to the potential confusion between macrophages and epithelial cells after staining and light microscopy examination (Leitner et al., 2000). The cell composition is also highly dependent on the inflammatory status of the mammary gland, with increased prevalence of neutrophils and decreasing percentages for lymphocytes, and on the lactation status with increased values in late lactation for macrophages (Leitner et al., 2000; Schwarz et al., 2011).

Flow cytometry analyses identified different subpopulations of T lymphocytes, in particular CD4+, CD8+ as well as $\gamma\delta$ T cells, along with different macrophage classical (CD14+ CD16-) and non-classical (CD14- CD16+) subpopulations (Farschtschi et al., 2022). While CD4 and CD8 T lymphocytes constitute approximately 20-30 % of all T-cells, $\gamma\delta$ T cells are also present in milk in significant numbers (10-30% of all lymphocytes in milk) (Taylor et al., 1994; Leitner et al., 2000; Farschtschi et al., 2022). The proportion of these subpopulations among T lymphocytes depends on the mammary gland health status: indeed, the CD4/CD8 ratio increases upon infection or inflammation (Park et al., 1992; Asai et al., 1998; Rivas et al., 2000; Mehrzad et al., 2008; Tassi et al., 2013) as well as in late lactation (Taylor et al., 1994). While B-cells are only present in limited numbers in healthy mammary glands, their proportion among lymphocytes was shown to increase upon inflammation (Schwarz et al., 2013).

Based on such evidence that the proportion of the different cell types is related to the health status of a mammary gland, differential cell count strategies have been explored in order to not only rely on SCC values to characterize diseased quarters (Rivas et al., 2001; Schwarz et al., 2011). One such Differential Somatic Cell Count (DSCC) method discriminates between neutrophils and macrophages while lymphocytes are counted along with neutrophils (Damm et al., 2017; Halasa and Kirkeby, 2020). Four groups based on SCC and DSCC values were determined. Similar to previous studies, this method highlights variations depending on parity and lactation status (Schwarz et al., 2020b; Zecconi et al., 2023). Although both indicators are linked to the recruitment of leukocytes in mammary gland secretions, their dynamics was found to be slightly different, with DSCC values decreasing faster than SCC in the resolutive phase of inflammation (Kirkeby et al., 2020). DSCC can thus be considered as an additional indicator for the control of mastitis. Yet, the use of DSCC will require the definition of suitable thresholds in order to improve its predictive value (Schwarz et al., 2020b). In fact, a recent survey suggest very little added value of DSCC over SCC in terms of positive- and ngetive- predictive values for the detection of mastitis cases due to major pathogens (Fonseca et al., 2025a). Furthermore, the

definition of suitable thresholds for SCC/DSCC combinations will have to take into account a potential herd effect since authors have also observed significantly different values between different herds, showing that the herd effect could be more important that the pathogen (Kirkeby et al., 2021).

High resolution differential cell count based on flow cytometry has also been proposed recently to better characterize the inflammation/infectious status of mammary glands (Farschtschi et al., 2022). A more global approach has also been tested using single-cell RNAseq highlighting the presence of different cell populations and subpopulations in a healthy milk sample (Becker et al., 2021; Wiarda et al., 2025). Altogether these different methods allow the identification of cell populations present in milk.

We hypothesized that another strategy to characterize the inflammatory status of mammary glands could be based on a functional assessment of the cells present in mammary gland secretion. In order to do so, we first identified a subset of genes specific for macrophages, neutrophils or T lymphocytes, three of the main cell populations present in milk. We then undertook a gene-expression profiling analysis of these selected genes based on RNA samples collected from unprocessed raw milk. Gene expression profiles were determined allowing clustering of milk samples and the clustering obtained was then compared to results obtained by flow cytometry.

Materials and methods

Collection of milk samples

Quarter foremilk samples (40 ml) from all four quarters were collected under aseptic conditions as recommended by the National Mastitis Council guidelines, from 28 Holstein (HF) and 10 Normande (NO) cows at the experimental farm of Le Pin-au-Haras (Unité Expérimentale du Pin-au-Haras). No criteria was applied for the selection of cows from which milk samples were collected. More specifically, after teat cleaning, the first streams of milk were discarded, premilking disinfectant solution was applied and then drier, teat apex was disinfected by scrubing with a 70% ethanol moistened gauze pad and milk was collected in 50ml tubes held at 45° angle to avoid contamination by dust materials. Milk was then transported to the laboratory at room temperature within 3h.

A 500 µL portion of each milk sample was subjected to a Fossomatic model 90 (Foss Electric, Hillerod, Denmark) to determine the milk SCC. The somatic cell score (SCS) was calculated with the formula : SCS = Log2(SCC / 100,000) + 3. One milk sample from each cow was selected, except for four cows for which two samples were collected from two different quarters: the only criteria for selection of milk samples was to have a panel of milk samples representing a large range of SCC values. These quarter milk samples originated from cows that were at different stages of lactation (days-in-milk ranging from 35 to 391 days) and with different SCC values (ranging from 17,000 to 3,603,000 cells/ml). None of the samples were from clinical mastitis cases. Details are presented in **Supplementary table S1**.

Bacteriological examination of the selected quarter milk samples was performed by plating 50 µL of milk on sheep blood agar plates (ThermoFicher Columbia agar with Sheep Blood PLUS #PB5039A) and recording bacterial colonies after 24 and 48 hours of incubation at 37°C. Basic tests (colony morphology and colour, hemolysis, Gram-staining, catalase) were performed in order to qualify the bacteriological status of milk samples. Samples were considered contaminated when more that 3 different colony types were observed.

Isolation and labelling of milk cells

Milk cells were collected from the selected milk samples: An equal volume of 1× DPBS (reference, Sigma D8537) was added to milk samples. The diluted milk sample was left at room temperature for 15 minutes and then centrifuged at room temperature for 10 minutes at 1400g.

The cream layer and the supernatant were removed and the cell pellet was then washed once in 10ml of room temperature (DPBS, 2mM EDTA, 1% horse serum) and then centrifuged 5 minutes at 500g. The cells were then resuspended in FACS buffer (DPBS, 2mM EDTA, 10% horse serum). Cells were then labelled prior to flow cytometry analysis or cell sorting.

Cells were quantified using a LUNA-FL cell counter (Logos Biosystems) according to the manufacturer's instructions. One million cells prepared as indicated above were transferred to a 1.5ml tube, centrifuged and resuspended in FACS buffer. Cells were labeled for 30 min with primary antibodies (30 min, 4°C). Primary antibodies used in the present study were CD14-PE-AF750 (Bio-Rad, reference MCA1568P750, clone Tük4), CD45-PE (Bio-Rad, reference MCA2220PE, clone 1.11.32), G1 (Kingfisher Biotech, reference WSC0608B-100, clone CH138A) and CD3 (Kingfisher Biotech, reference WS0561B-100, clone MM1A). After washing in FACS buffer, cells were incubated (30 min, 4°C) with the secondary antibodies BV421 labelled rat anti-mouse IgM antibody (BD Biosciences, reference 562595, clone R6-60.2) and Alexa Fluor™ 488 labelled Goat anti-Mouse IgG1 (Thermo Fisher Scientific, reference A21121) in FACS buffer, matching respectively G1 and CD3 primary antibodies. After washing in FACS buffer, cells were resuspended in DPBS containing the Zombie Aqua viability dye (BioLegend, reference 423101) for 30 minutes at 4°C. After a final wash in FACS buffer, cells were resuspended in FACS buffer.

Cells sorting of purified specific leucocytes populations

Milk cells from three quarter milk samples were collected and labelled as described above. These three samples were selected based on their SCC, representative of low, intermediate and high SCC quarters. Labelled milk cells were then sorted on a MoFlo Astrios EQ (Beckman Coulter) high speed cell sorter equipped with 4 lasers to obtain, among live cells, purified macrophages (CD45+ G1- CD14+), T lymphocytes (CD45+ CD3+) and neutrophils(CD45+ G1+). The average sorting speed was 10,000 cells/s.

Flow cytometry analysis of milk cells

To determine the proportion of leucocyte subpopulations in milk samples, milk cells were isolated and labelled as described above. Milk cells were then analyzed on a LSR Fortessa™ X-20 Flow cytometer (Becton Dickinson) and results were analyzed with the Kaluza software v2.1 (Beckman Coulter). Among live cells, macrophages were identified as CD45+ G1- CD14+ cells, neutrophiles as CD45+ G1+ cells and T lymphocytes as CD45+ CD3+ cells. Our gating strategy is exemplified in **Supplementary Figure S1**.

RNA extraction and high-throughput RT-qPCR

For sorted cells, immediately after collection, cells were centrifuged 10 minutes at 1000g, resuspended in 1 ml of Trizol and stored at -80°C until RNA was extracted as described below. For milk samples, 200 microliters of milk were mixed with 800 µl of Trizol (Thermofisher) in a lysing matrix D tube (MP Bio) and homogenized for 45 seconds – speed 6 in a FastPrep-24 apparatus. Samples were then stored at -80°C until further processed.

Upon thawing of milk and purified cells samples mixed with Trizol, 200 µl chloroform was added, the tube was vortexed for 10 seconds and then centrifuged at 12000g for 15min at 4°C. The aqueous phase was collected, mixed with an equal volume of 70% ethanol, loaded on a NucleoSpin RNA columns (Macherey Nagel) and RNA was then purified and treated with DNAse I as indicated by the manufacturer. Total RNA (50 ng) was reverse transcribed to cDNA using 5x iScript reverse transcription supermix (Biorad) according to manufacturer's instructions. Ten-fold diluted cDNA were preamplified according to Fluidigm's protocol (quick reference PN 100-5875 B1) using PreAmp master mix and a mix of all forward and reverse primers at 500nM each. Preamplification program was 2 minutes 95°C, followed by 16 cycles of (95°C for 15s, 60°C for 4 min). Preamplified cDNA were then diluted 5-fold with (10 mM Tris, pH 8.0, 1 mM EDTA) buffer. Gene expression level were measured from pre-amplified cDNA on 96.96 GE Dynamic Array IFC

using the Fluidigm BioMark™ HD System. Primers used in this study are listed in **Supplementary Table S1**.

Data analysis

A reference sample was built from the arithmetic mean of Ct for each gene of analyzed samples. Fold changes compared to this reference sample were calculated by the ΔΔCt method using *ACTB* and *GAPDH* as reference genes (Livak and Schmittgen, 2001). Heatmaps were generated using the pheatmap package (v1.0.10) in RStudio (1.1). Likelyhood Ratio Test was performed using the GTest function of the R package DescTools (v0.99). PCA analyses were performed using the PCA function of the FactoMineR package (v2.9). Mann-Whitney and Kruskall-Wallis tests were performed using the stats package (v4.5.1). Relatedness of dendrogram based on gene expression and flow cytometry proportions were analyzed using the dendextend package (v1.19.1). Scripts used in the analysis and raw qPCR data are available at https://github.com/pgermon37/milk_RNA_profile.

Results

Identification of a gene panel able to discriminate macrophages, neutrophils and T lymphocytes

The aim of our study was to investigate if profiling of milk samples based on gene expression could provide relevant information regarding the inflammatory status of mammary gland quarters. For such an analysis to take into account the contribution of different leucocyte populations to the gene expression profile of raw milk samples, we first needed to select a panel of genes that showed different levels of expression depending on the cell type.

Therefore, we sorted macrophages neutrophils (live CD45+ G1+ cells), (live CD45+ G1-CD14+ cells), and T lymphocytes (live CD45+ CD3+ cells) by flow cytometry from three quarter milk samples having somatic cell counts ranging from 45,000 to 389,000 cells/ml (**Supplementary Table S2**).

The expression of a panel of 95 genes by these purified sorted cells was investigated by high-throughput RT-qPCR. Genes were selected based on their expected role in the function of cells types studied and based on the laboratory's expertise in bovine immunity. Only genes yielding a specific amplification with no primer dimers, as evaluated by melting curve analysis, in all samples were selected which restricted the number of genes to 78. Expression results for these 78 genes in the three selected cell types were first analyzed by hierarchical clustering (**Figure 1A**) and Principal Component Analysis (**Figure 1B**). Both analyses showed clear separation of samples based on cell types. The hierarchical clustering analysis allowed the identification of 6 clusters of genes with specificities for one or two of the cell types. For example, the cluster identified as "Lc" contains genes that are more expressed in T lymphocytes than in the two other cell types. A second cluster specifically more expressed in neutrophils, labeled "PMN2", can also be clearly observed, as well as a cluster labeled "Macro" associated with higher expression in macrophages. The list of genes from these six clusters are presented in **Supplementary Table S3**.

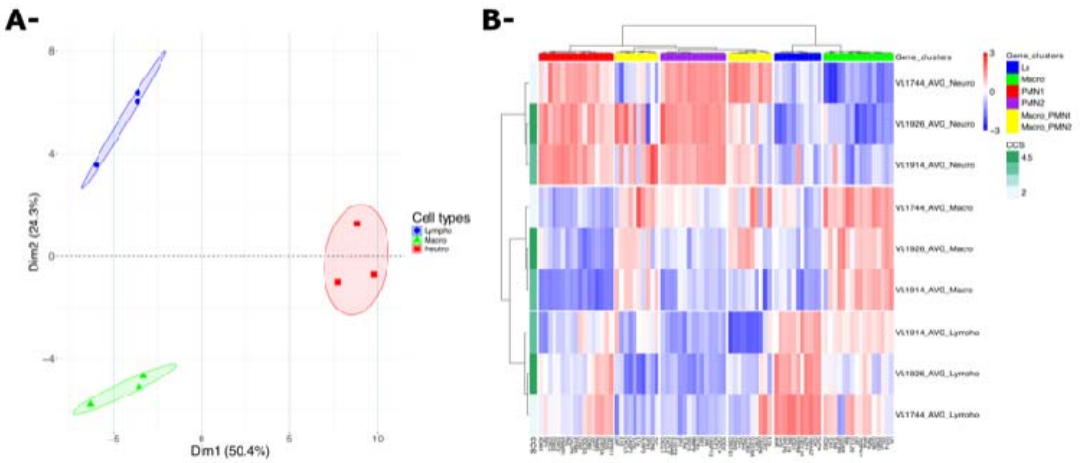


Figure 1: Heatmap and PCA of gene expression of sorted milk macrophages, neutrophils and T lymphocytes. Cells from three milk samples from three different cows (one smaple per cow) were collected by centrifugation, labelled with antibodies against CD3 (T lymphocytes), CD14 (macrophages) and G1 (neutrophils) and sorted on a MoFlo Astrios EQ cell sorter. RNA was extracted from purified cell populations and analyzed by high-throughput RT-qPCR. Gene expression values were calculated against a theoretical control sample and analyzed by principal component analysis (A) and hierarchical clustering (B).

Based on these results, because they showed clearly distinct patterns of expression between the three leucocyte populations investigated, we selected the 41 genes belonging to clusters “Lc” (11 genes), “PMN2” (15 genes) and “Macro” (15 genes) for the investigation of expression profiles in RNA from raw milk samples.

Comparison of clustering based on gene expression and flow cytometry profiles

We then analyzed quarter milk samples from 28 Holstein (HF), which included the three quarter milk samples used for sorting the purified cell types, and from 10 Normande (NO) cows. One sample per cow was selected so that the panel of samples covered a range of SCC value from 17,000 to 3,603,000 cells/ml, with only 6 samples above the 400,000 cells/ml threshold. Cows were at different stages of lactation (days-in-milk ranging from 35 to 391 days). Detailed description of the different milk samples are given in **Supplementary Table S2**.

SCC distribution of selected milk samples are depicted in **Supplementary Figure S2**. We could not detect any statistically significant difference in terms of SCC between samples from Holstein and Normande cows (Mann-Whitney p value = 0.29). The breed criteria was thus not explored further in downstream analyses.

For each sample, the gene expression profile after RNA extraction was determined by high-throughput RT-qPCR. Among the 41 genes selected above, expression values were obtained for all samples for only 36 genes (excluded genes were *CD4*, *CXCL10*, *CCL20*, *IFNB1* and *TNFa*). Analysis by hierarchical clustering of expression values for these 36 genes, performed without any a priori, clearly identified four clusters of samples (**Figure 2A**). The PCA analysis showed a clear separation of these four clusters (**Figure 2B)**. Clusters 1 and 2 encompass samples with SCC values generally below 100,000 cells/ml whereas cluster 3 and 4 samples are characterized by significantly higher SCC (**Figure 3**).

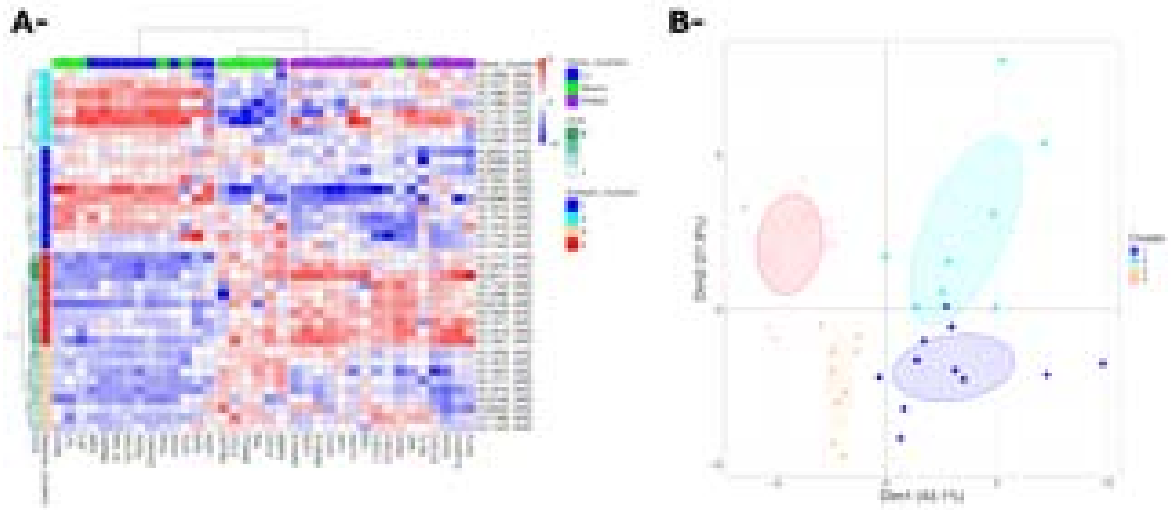


Figure 2: Heatmap and PCA of gene expression of raw milk samples. Milk samples were aseptically collected from 38 individual quarters. RNA was extracted directly from raw milk samples and analyzed by high-throughput RT-qPCR. Gene expression values were calculated against a theoretical control sample and analyzed by hierarchical clustering (A) and principal component analysis (B). Color code for clusters is indicated on the figure.

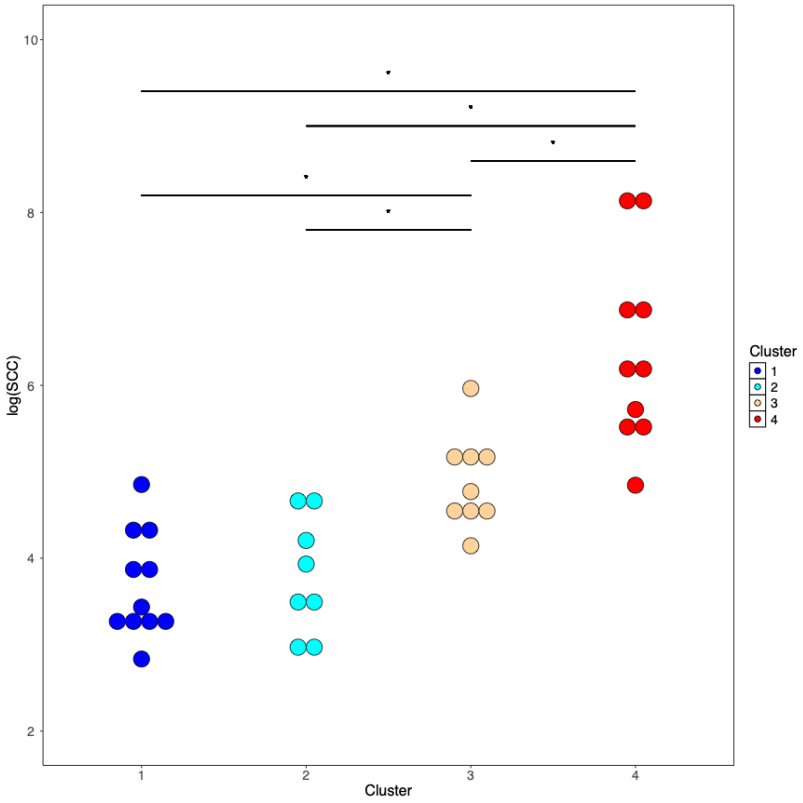


Figure 3: log(SCC) values of milk samples from the four different clusters. The log(SCC) of each of the milk samples analyzed is represented and samples are grouped and colored depending on the cluster to which they were allocated in the hierarchical clustering analysis. Color code is identical to that of Figure 2. *

indicates statistical significance (p < 0,05). p-values were calculated using Mann and Whitney after global comparison using a Kruskal and Wallis test.

As can be seen in **Figure 2**, our method was able to differentiate different categories of samples in the SCC range of 50,000 to 200,000 with samples belonging to clusters 1, 2 and 3. In this SCC range we thus had samples with significantly different immune status: even though the SCC of cluster 3 samples is low, these samples are characterized by a higher expression of "Macro" genes and are thus quarters where macrophages could be present in higher proportions or be more transcriptionally active. As expected, high SCC milk samples belonged to cluster 4 characterized by higher expression of "PMN2" genes associated with the neutrophil cell population. Samples belonging to clusters 1 and 2 are characterized by a higher level of expression of "Lc" genes but can be distinguished from one another by different levels of expression of "Macro" and "PMN2" genes maybe suggesting different immune status.

We could not detect any statistically significant difference in lactation stage between the different clusters (**Supplementary Figure S3** – Kruskal-Wallis p value = 0.35).

For each of these milk samples, the proportion of macrophages, neutrophils and T lymphocytes were obtained by flow cytometry (see **Supplementary Table S2**). Analysis of these results by Principal Component Analysis showed separation of samples from the 4 clusters identified previously by gene-expression profiling with some outliers (**Figure 4**).

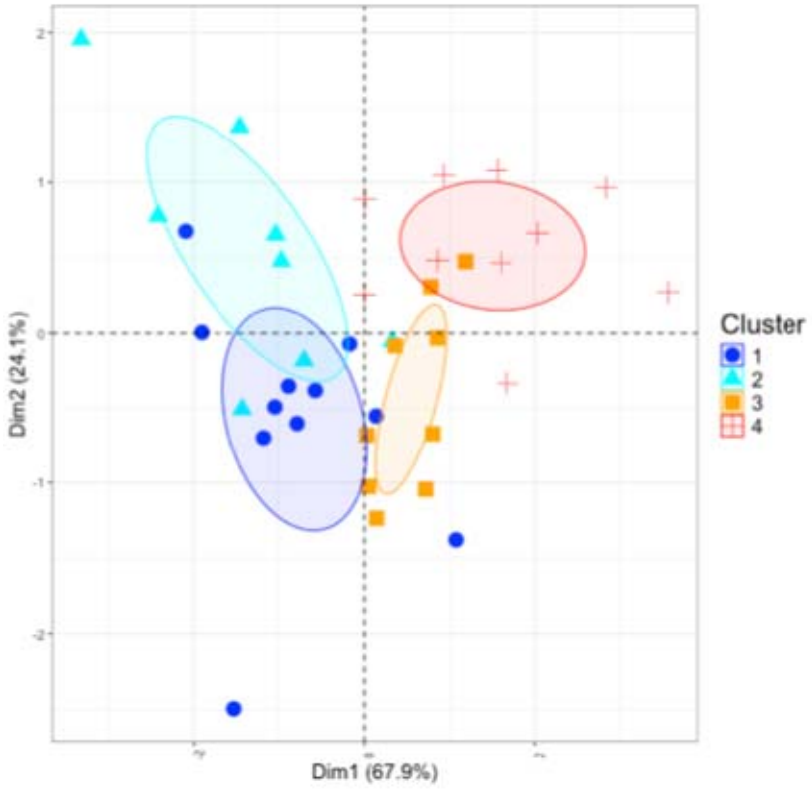


Figure 4: Principal component analysis of proportions of milk samples analyzed by flow cytometry. Proportions of macrophages, neutrophils and T-lymphocytes among cells from collected milk samples was determined by flow cytometry. Principal Component Analysis of the cell composition was performed on the percent values for each of the three cell types analyzed. Dots are colored depending on the cluster to

which they were allocated based on the hierarchical clustering analysis of gene expression data. Color code is identical to that of Figure 2 and is indicated on the figure.

To quantitatively estimate the relatedness of analyses based on gene expression and flow cytometry, we analyzed the relatedness of dendrograms obtained with both methods calculating Baker's index (data not shown). This analysis showed that the two trees share significant similarities: the null hypothesis that the two trees are not similar can be rejected with a risk below 0.01.

Clustering based on gene expression profile differentiates samples with similar SCC.

Classical bacteriological analysis was performed on all selected milk samples. Absence of bacteria was observed in 13 samples while, when bacteria were present, only Gram-positive pathogens were detected (16 samples). Bacteriological status could not be determined for 5 samples and 4 samples could not be examined due to technical reasons (**Supplementary Table S2**). When the clustering was analyzed relative to the presence of bacteria in the milk, the number of "No growth" sample decreases while the number of samples showing growth of "Gram-positive" bacteria increases (**Figure 5**). As expected, cluster 4 samples have a tendency to contain bacteria (Likelihood Ratio Test p-value = 0.06).

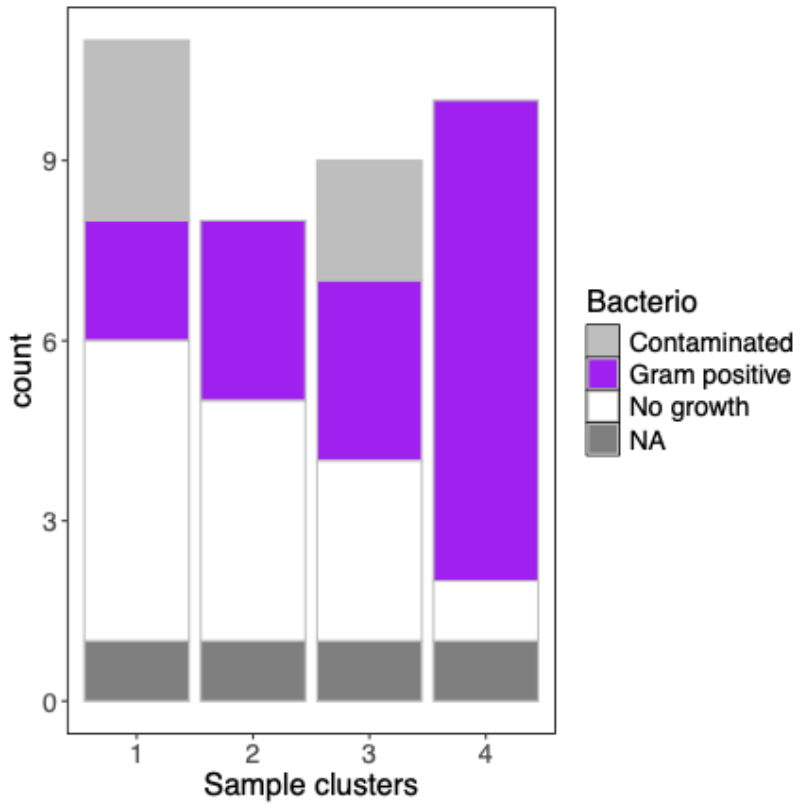


Figure 5: Barplot representing the number of samples in bacteriology positive or negative samples. Bacteriological status of each milk sample was analyzed by plating 50 microliter of each milk sample on sheep-blood agar plates. Samples with more that 3 types of colonies were considered as contaminated. Gram staining was performed on colonies growing on plates. NA indicates absence of bacteriological analysis.

Discussion

The objective of the present study was to investigate a new method for the functional characterization of the inflammatory status of mammary gland quarters. As an alternative to the characterization of cells present in milk, we investigated the gene expression profile of milk samples as a surrogate for the functionality of cells present in milk.

In order to determine if this strategy would allow us to characterize the inflammatory status of low cell counts samples, for which current strategies provide low positive predictive values of being infected, we focused our analyses on samples with rather low SCC values: the median value for SCC is 95 500 (Q1-Q3 : 38 250 – 218 000). In order to improve the diagnostic of inflammatory quarters, analyses were done on quarter milk samples. Indeed, analyses of composite samples show reduced predictive power compared to analyses based on quarter milk samples (Dufour and Dohoo, 2012; Fonseca et al., 2025c). Importantly, it should be noted that none of the milk samples analyzed were from clinical cases, even the two samples with the high SCC values.

Based on the expression level of 36 genes from which expression could be measured in all samples, we identified four different clusters of milk samples, each with specific gene expression profiles. Cluster 4 samples are characterized by increased expression from neutrophil and macrophage related genes. Cluster 3 samples are typified by increased expression of macrophage related genes while clusters 1 and 2 samples are characterized by higher expression of genes typifying T lymphocytes. This clustering based on gene expression actually overlaps with that obtained from flow cytometry: consistent with the increased expression of neutrophil gene set, cluster 3 and 4 samples are the ones with the highest number of neutrophils. Compared to clusters 1 and 2, clusters 3 and 4 can therefore be considered as inflamed quarters. Nevertheless, cluster 3 and 4 samples are different: cluster 3 samples show higher values for macrophage related genes but less expression of neutrophil related genes compared to cluster 4 samples. Hence, cluster 3 samples could originate from mammary gland quarters in the process of resolution with increased proportions of dead neutrophils, could be representatives of chronically infected cows or be specific to a particular pathogen. With the available data, we cannot decide which hypothesis is the most likely. The number of samples we analyzed is clearly not sufficient to firmly establish a link between a particular profile and a specific pathogen.

Our results also clearly distinguished cluster 1 from cluster 2 samples, two types of milk samples with low somatic cell counts. These low SCC milk samples are indeed not identical in terms of inflammatory status. Although they show similar expression of genes typifying T lymphocytes, they showed different expression of genes typifying macrophages or neutrophils. In particular cluster 2 samples express higher levels of the neutrophil-associated genes while cluster 1 samples showed higher expression of macrophage-associated genes. They consequently represent two different inflammatory statuses. Because of higher expression of pro-inflammatory markers (CXCL8 for example), whose expression is clearly induced in inflamed mammary glands (Bannerman et al., 2004), we hypothesize that cluster 2 samples are samples where a recruitment of neutrophils has been initiated. Overall, cluster 2 samples indeed have higher percentage of neutrophils (71 % +/- 25% vs. 39% +/- 25%, $p < 0.01$). On the contrary, we hypothesize that cluster 1 samples correspond to healthy samples with limited inflammation. Compared to cluster 2 samples, cluster 1 samples have higher expression levels of macrophage-associated genes: considering these samples as healthy is consistent with the observation of higher proportions of macrophages in healthy mammary glands (Lee et al., 1980; Damm et al., 2017; Halasa and Kirkeby, 2020).

The method described in this pilot study has several advantages over current methods to qualify the inflammatory status of mammary gland quarters. Compared to clustering strategies proposed in other studies, such as grouping samples based on SCC values and bacteriological status or on SCC and DSCC values (Schwarz et al., 2020a; Souza et al., 2020), the clusters we

have identified were obtained without any a priori, are only based on intrinsic properties of the milk samples linked to macrophages, neutrophils and T lymphocytes and are not conditioned by user-defined thresholds. It should be noted that even in low SCC milk samples a clear gene expression pattern could be observed. Therefore, contrary to DSCC which requires SCC values above 50,000 cells/ml, a low SCC is not a limit to the use of our method (Damm et al., 2017).

In addition, the protocol presented here does not rely on preliminary centrifugation of milk samples which, as can be the case for flow cytometry methods, might fail to recover all cells; in particular macrophages that can remain in the cream layer after centrifugation (Lee et al., 1980).

We therefore believe that it could bring an added value to the definition of inflammatory status of mammary gland quarters. More specifically, this method could be used to investigate in more details the response of mammary gland quarters to infection. For instance, neutrophil subtypes in different inflammatory conditions could be investigated which would extend our observations of increased proportions of a neutrophil subset expressing MHCII surface molecules during subclinical mastitis (Rambault et al., 2023). Moreover, in vaccine based studies were a local immune response is expected, the gene panel could be adapted to monitor more precisely the response of T-lymphocytes such as the Th17 response that has been shown to be beneficial to the clearance of bacteria (Herry et al., 2017; Cebron et al., 2020).

For these more detailed analyses to be possible, new developments will be required to increase the panel of genes analyzed in order to better distinguish additional cell types or subtypes. Single-cell RNAseq analyses have recently been performed on milk cells (Wiarda et al., 2025). These data confirmed some of the associations described herein: for example, CXCL8 and TLR4 were also identified as specific for granulocytes/neutrophils. CD14 was identified as associated to macrophages, as expected for a macrophage specific gene (Grandoni et al., 2023). More detailed analyses of these data and of the cell clusters observed could provide usefull information regarding additionnal genes specific for different cell types or even subtypes.

The method described herein nevertheless has certain limitations. The main limit is that it is not easily accessible to on-farm diagnostic of mammary gland inflammation and requires specific equipments. Its interest is limited for the moment to laboratory studies trying to better qualify the inflammatory status of the udder. We also failed to clearly quantify epithelial cells by flow cytometry. Although these cells are likely to represent only a low proportion of milk cells, their proportions could potentially modify the clustering observed by flow cytometry. Nevertheless, it is likely that the gene expression profiling of milk samples is actually susceptible to the presence of epithelial cells in milk (Boutinaud et al., 2015). It is also possible, since fat globules have been shown to contain RNA molecules originating from the milk producing epithelial cell, that the gene expression profile also reflects the functionality of mammary epithelial cells (Canovas et al., 2014). As a consequence, RNA quantification from raw milk might also reflect both the presence and/or functionality of epithelial cells.

The present report should be considered as a pilot study that was limited to only a small number of samples. In order to validate the usefulness of our method, more broader studies would be required. A more detailed investigation of the relationship between gene expression clusters and the bacteriological status of mammary gland quarters would definitely support the predictive value of these analyses. So far, only an tendency for an increased proportion of samples with Gram+ bacteria in cluster 4 samples was observed, consistent with the fact that these samples are the ones with the highest SCC. It also remains to be determined whether the gene expression profile varies depending on parity, days-in-milk as these parameters are clearly associated with variations in SCC and DSCC (Fonseca et al., 2025b). Complementary analyses should also take into account the variability of milk cellular composition at different stages during milking (Ostensson et al., 1988). The current profiles identified were based on foremilk samples which are likely to be enriched in SCC in samples with more than 100,000 cells/ml (Sarikaya and Bruckmaier, 2006).

To conclude, we set up a new way of monitoring the inflammatory status of mammary gland quarters during lactation based on the profiling of gene expression from raw milk with minimal processing before RNA extraction. We were able to distinguish four types of milk samples that corresponded to different inflammatory status. Potential refinements include the addition of new genes to better distinguish quarters with low SCC values, improved qualification of the inflammatory status of mammary glands along with more in-depth characterization of bacteria present in milk samples.

Although not applicable as such in commercial farms, this method paves the way to the development of new medium- or high- throughput phenotyping strategies. Such methods will provide useful information for the implementation of selective dry cow therapy or for the routine analysis of health status of mammary glands.

Acknowledgements

The authors would like to thank the animal staff of the “Unité Expérimentale du Pin-au-Haras” (INRAE, Gouffern-en-Auge, France) and of the “Unité Expérimentale de Physiologie animale de l'Orfrasière” (INRAE, Nouzilly, France) for their help in collect ing milk samples. We are grateful to Aude Remot for critical reading of the manuscript.

Ethics approval

Not applicable. Milk samples were collected from cows as part of a routine milking.

Funding

This work received funding from APIS-GENE (Masticells project).

Conflict of interest disclosure

The authors declare that they comply with the PCI rule of having no financial conflicts of interest in relation to the content of the article.

Data, scripts, code, and supplementary information availability

Data are available online: https://doi.org/10.5281/zenodo.18633548
Supplementary information is available online: https://doi.org/10.5281/zenodo.18631973

Author ORCIDs

Christophe Gitton 0000-0001-7986-0414
Yves Le Vern 0009-0009-3930-6107
Marie Gaborit 0009-0008-5461-1196
Rodrigo Prado Martins 0000-0002-8213-4015
Pierre Germon 0000-0003-3078-876X

References

Asai, K., Kai, K., Rikiishi, H., Sugawara, S., Maruyama, Y., Yamaguchi, T., Ohta, M., Kumagai, K., 1998. Variation in CD4+ T and CD8+ T lymphocyte subpopulations in bovine mammary gland secretions

475 during lactating and non-lactating periods. Veterinary Immunology and Immunopathology 65, 51-
476 61. doi:10.1016/s0165-2427(98)00176-7.
477 Bannerman, D.D., Paape, M.J., Lee, J.W., Zhao, X., Hope, J.C., Rainard, P., 2004. Escherichia coli and
478 *Staphylococcus aureus* elicit differential innate immune responses following intramammary
479 infection. Clinical and Diagnostic Laboratory Immunology 11, 463-472. doi:10.1128/CDLI.11.3.463-
480 472.2004.
481 Becker, D., Weikard, R., Hadlich, F., Kuhn, C., 2021. Single-cell RNA sequencing of freshly isolated
482 bovine milk cells and cultured primary mammary epithelial cells. Sci Data 8, 177.
483 doi:10.1038/s41597-021-00972-1.
484 Boutinaud, M., Herve, L., Lollivier, V., 2015. Mammary epithelial cells isolated from milk are a
485 valuable, non-invasive source of mammary transcripts. Front Genet 6, 323.
486 doi:10.3389/fgene.2015.00323.
487 Canovas, A., Rincon, G., Bevilacqua, C., Islas-Trejo, A., Brenaut, P., Hovey, R.C., Boutinaud, M.,
488 Morgenthaler, C., VanKlompenberg, M.K., Martin, P., Medrano, J.F., 2014. Comparison of five
489 different RNA sources to examine the lactating bovine mammary gland transcriptome using RNA-
490 Sequencing. Sci Rep 4, 5297. doi:10.1038/srep05297.
491 Cebron, N., Maman, S., Walachowski, S., Gausseres, B., Cunha, P., Rainard, P., Foucras, G., 2020.
492 Th17-related mammary immunity, but not a high systemic Th1 immune response is associated
493 with protection against *E. coli* mastitis. npj Vaccines 5, 108. doi:10.1038/s41541-020-00258-4.
494 Damm, M., Holm, C., Blaabjerg, M., Bro, M.N., Schwarz, D., 2017. Differential somatic cell count-A
495 novel method for routine mastitis screening in the frame of Dairy Herd Improvement testing
496 programs. Journal of Dairy Science 100, 4926-4940. doi:10.3168/jds.2016-12409.
497 Dosogne, H., Vangroenweghe, F., Mehrzad, J., Massart-Leen, A.M., Burvenich, C., 2003. Differential
498 leukocyte count method for bovine low somatic cell count milk. Journal of Dairy Science 86, 828-
499 834. doi:10.3168/jds.S0022-0302(03)73665-0.
500 Dufour, S., Dohoo, I.R., 2012. Monitoring dry period intramammary infection incidence and
501 elimination rates using somatic cell count measurements. Journal of Dairy Science 95, 7173-7185.
502 doi:10.3168/jds.2012-5839.
503 Farschtschi, S., Hildebrandt, A., Mattes, M., Kirchner, B., Pfaffl, M.W., 2022. Using High-Resolution
504 Differential Cell Counts (HRDCCs) in Bovine Milk and Blood to Monitor the Immune Status over
505 the Entire Lactation Period. Animals (Basel) 12 doi:10.3390/ani12111339.
506 Farschtschi, S., Mattes, M., Hildebrandt, A., Chiang, D., Kirchner, B., Kliem, H., Pfaffl, M.W., 2021.
507 Development of an advanced flow cytometry based high-resolution immunophenotyping method
508 to benchmark early immune response in dairy cows. Sci Rep 11, 22896. doi:10.1038/s41598-021-
509 02405-w.
510 Fonseca, M., Kurban, D., Roy, J.-P., Santschi, D.E., Molgat, E., Dufour, S., 2025a. Usefulness of
511 differential somatic cell count for udder health monitoring: Diagnostic performance of somatic
512 cell count and differential somatic cell count for diagnosing intramammary infections in dairy
513 herds with automated milking systems. Journal of Dairy Science 108, 3929-3941.
514 doi:10.3168/jds.2024-25404.
515 Fonseca, M., Kurban, D., Roy, J.P., Santschi, D.E., Molgat, E., Dufour, S., 2025b. Usefulness of
516 differential somatic cell count for udder health monitoring: Effect of intramammary infections,
517 days in milk, quarter location, and parity on quarter-level differential somatic cell count and
518 somatic cell score in apparently healthy dairy cows. Journal of Dairy Science 108, 3878-3899.
519 doi:10.3168/jds.2024-25401.
520 Fonseca, M., Kurban, D., Roy, J.P., Santschi, D.E., Molgat, E., Yang, D.A., Dufour, S., 2025c. Usefulness
521 of differential somatic cell count for udder health monitoring: Identifying referential values for
522 differential somatic cell count in healthy quarters and quarters with subclinical mastitis. Journal of
523 Dairy Science 108, 3917-3928. doi:10.3168/jds.2024-25403.
524 Grandoni, F., Fraboni, D., Canonico, B., Papa, S., Buccisano, F., Schuberth, H.-J., Hussen, J., 2023. Flow
525 Cytometric Identification and Enumeration of Monocyte Subsets in Bovine and Water Buffalo
526 Peripheral Blood. Current Protocols 3, e676. doi:https://doi.org/10.1002/cpz1.676.

Halasa, T., Kirkeby, C., 2020. Differential Somatic Cell Count: Value for Udder Health Management. Front Vet Sci 7, 609055. doi:10.3389/fvets.2020.609055.

Harmon, R.J., 1994. Physiology of mastitis and factors affecting somatic cell counts. Journal of Dairy Science 77, 2103-2112. doi:10.3168/jds.S0022-0302(94)77153-8.

Herry, V., Gitton, C., Tabouret, G., Reperant, M., Forge, L., Tasca, C., Gilbert, F.B., Guitton, E., Barc, C., Staub, C., Smith, D.G.E., Germon, P., Foucras, G., Rainard, P., 2017. Local immunization impacts the response of dairy cows to *Escherichia coli* mastitis. Sci Rep 7, 3441. doi:10.1038/s41598-017-03724-7.

Hillerton, J., Redefining mastitis based on somatic cell count. Book of abstracts of the p. 4–6.

Kirkeby, C., Schwarz, D., Denwood, M., Farre, M., Nielsen, S.S., Gussmann, M., Toft, N., Halasa, T., 2021. Dynamics of somatic cell count (SCC) and differential SCC during and following intramammary infections. Journal of Dairy Science 104, 3427-3438. doi:10.3168/jds.2020-19378.

Kirkeby, C., Toft, N., Schwarz, D., Farre, M., Nielsen, S.S., Zervens, L., Hechinger, S., Halasa, T., 2020. Differential somatic cell count as an additional indicator for intramammary infections in dairy cows. Journal of Dairy Science 103, 1759-1775. doi:10.3168/jds.2019-16523.

Lee, C.S., Wooding, F.B., Kemp, P., 1980. Identification, properties, and differential counts of cell populations using electron microscopy of dry cows secretions, colostrum and milk from normal cows. Journal of Dairy Research 47, 39-50. doi:10.1017/s0022029900020860.

Leitner, G., Shoshani, E., Krifucks, O., Chaffer, M., Saran, A., 2000. Milk leucocyte population patterns in bovine udder infection of different aetiology. J Vet Med B Infect Dis Vet Public Health 47, 581-589. doi:10.1046/j.1439-0450.2000.00388.x.

Livak, K.J., Schmittgen, T.D., 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods 25, 402-408. doi:10.1006/meth.2001.1262.

Lun, S., Astrom, G., Magnusson, U., Ostensson, K., 2007. Total and differential leucocyte counts and lymphocyte subpopulations in lymph, afferent and efferent to the supramammary lymph node, during endotoxin-induced bovine mastitis. Reprod Domest Anim 42, 126-134. doi:10.1111/j.1439-0531.2006.00741.x.

Mehrzad, J., Janssen, D., Duchateau, L., Burvenich, C., 2008. Increase in Escherichia coli inoculum dose accelerates CD8+ T-cell trafficking in the primiparous bovine mammary gland. Journal of Dairy Science 91, 193-201. doi:10.3168/jds.2007-0096.

Mondini, S., Gislon, G., Zucali, M., Sandrucci, A., Tamburini, A., Bava, L., 2025. Factors influencing somatic cell count and leukocyte composition in cow milk: A field study. Journal of Dairy Science 108, 2721-2733. doi:10.3168/jds.2024-25357.

Ostensson, K., Hageltorn, M., Astrom, G., 1988. Differential cell counting in fraction-collected milk from dairy cows. Acta Veterinaria Scandinavica 29, 493-500. doi:10.1186/BF03548647.

Park, Y.H., Fox, L.K., Hamilton, M.J., Davis, W.C., 1992. Bovine mononuclear leukocyte subpopulations in peripheral blood and mammary gland secretions during lactation. Journal of Dairy Science 75, 998-1006. doi:10.3168/jds.S0022-0302(92)77842-4.

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

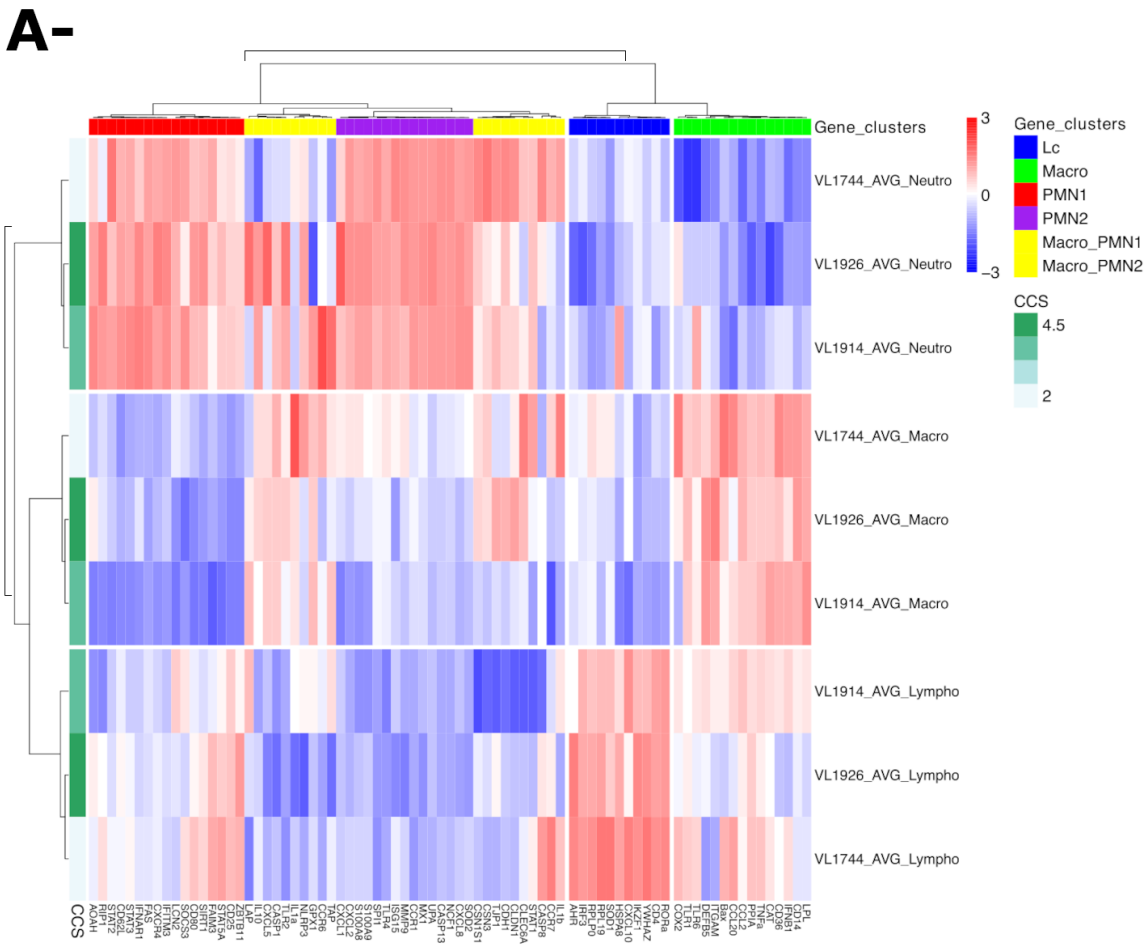
Rambault, M., Gilbert, B.F., Roussel, P., Tessier, A., David, V., Germon, P., Winter, N., Remot, A., 2023. MHC-II expressing neutrophils circulate in blood and milk during mastitis and show high microbicidal activity. Journal of Dairy Science

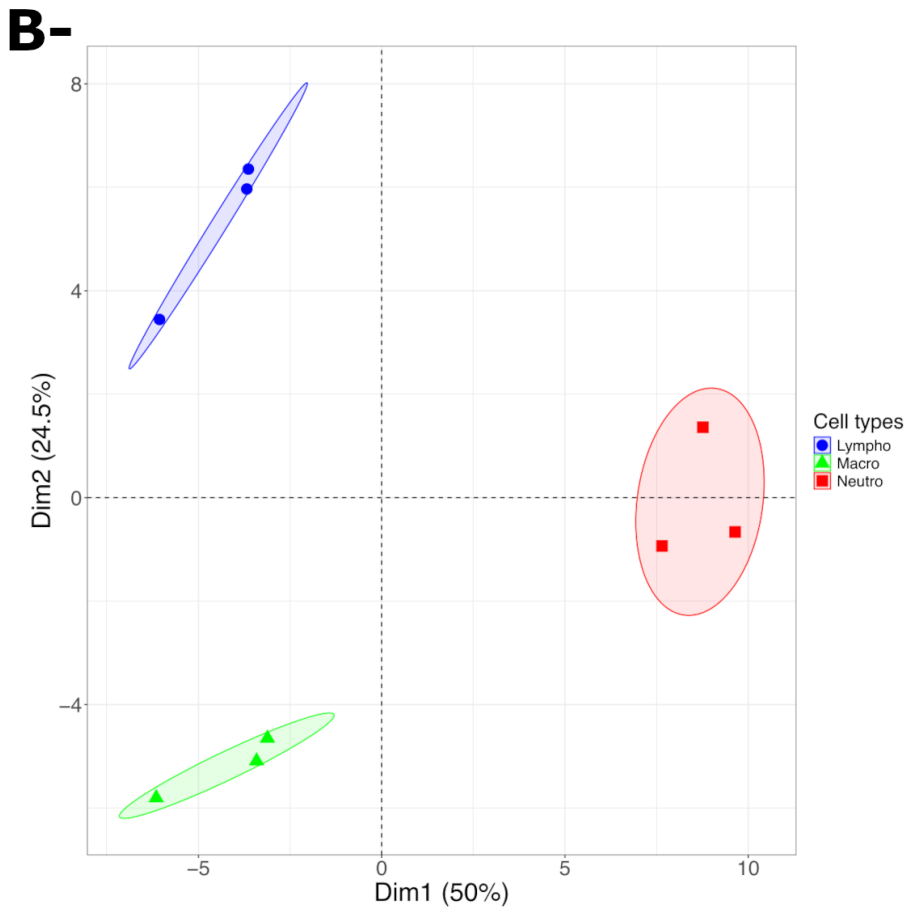
Rivas, A.L., Quimby, F.W., Blue, J., Coksaygan, O., 2001. Longitudinal evaluation of bovine mammary gland health status by somatic cell counting, flow cytometry, and cytology. Journal of Veterinary Diagnostic Investigation 13, 399-407. doi:10.1177/104063870101300506.

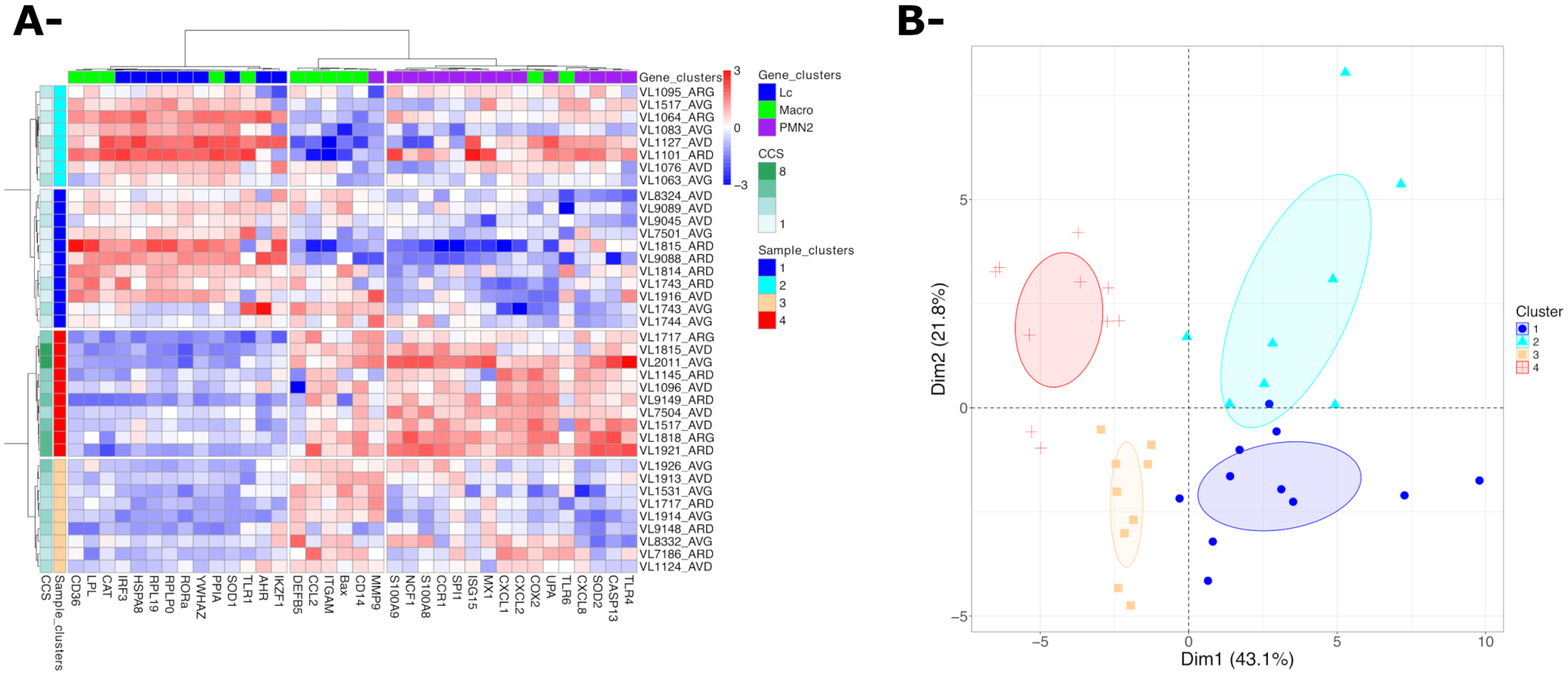
Rivas, A.L., Quimby, F.W., Coksaygan, O., Olmstead, L., Lein, D.H., 2000. Longitudinal evaluation of CD4+ and CD8+ peripheral blood and mammary gland lymphocytes in cows experimentally inoculated with Staphylococcus aureus. Canadian Journal of Veterinary Research 64, 232-237.

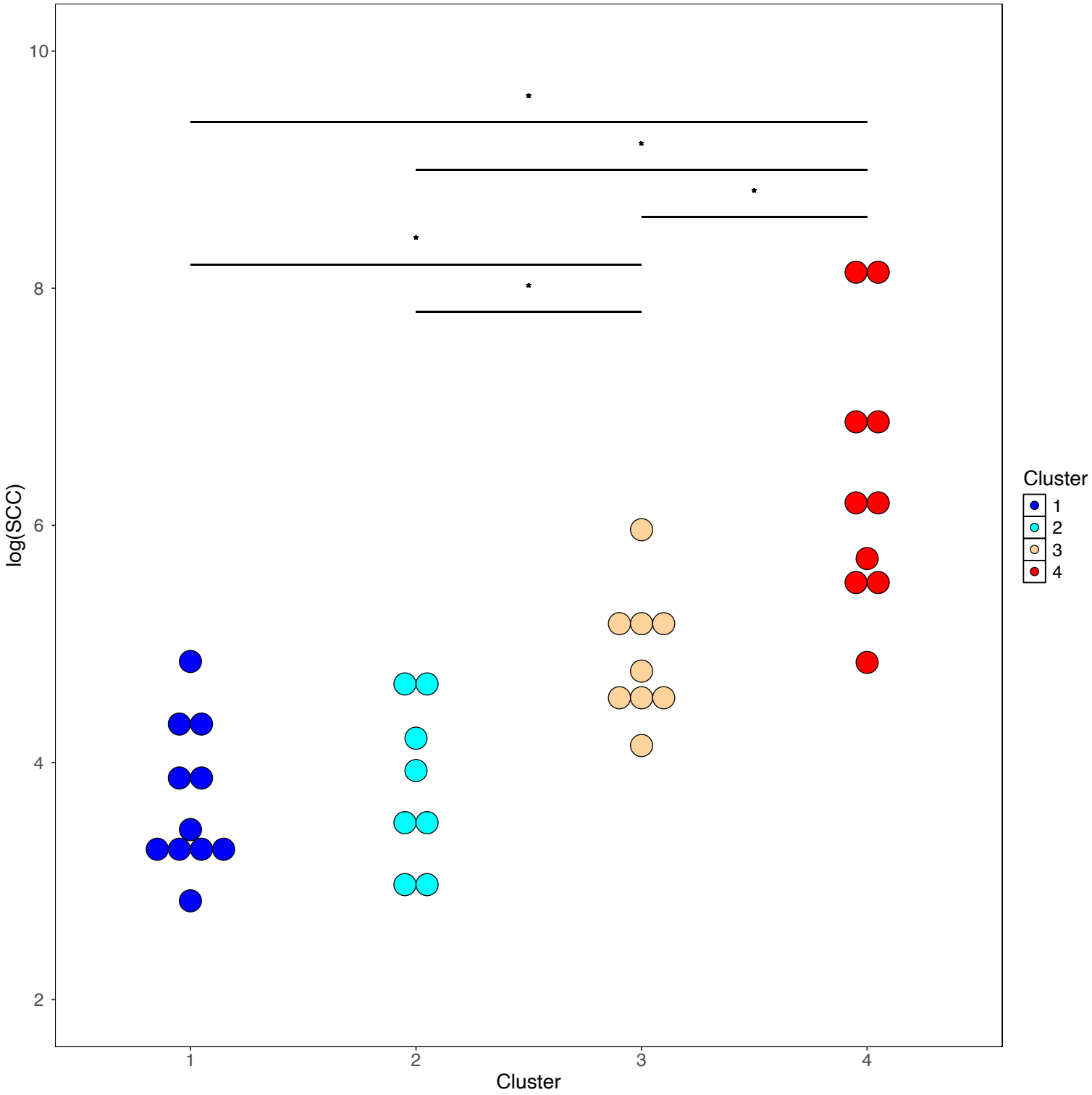
578 Sarikaya, H., Bruckmaier, R.M., 2006. Importance of the sampled milk fraction for the prediction of
579 total quarter somatic cell count. Journal of Dairy Science 89, 4246-4250. doi:10.3168/jds.S0022-
580 0302(06)72470-5.
581 Schwarz, D., Diesterbeck, U.S., Failing, K., Konig, S., Brugemann, K., Zschock, M., Wolter, W., Czerny,
582 C.P., 2010. Somatic cell counts and bacteriological status in quarter foremilk samples of cows in
583 Hesse, Germany--a longitudinal study. Journal of Dairy Science 93, 5716-5728.
584 doi:10.3168/jds.2010-3223.
585 Schwarz, D., Diesterbeck, U.S., Konig, S., Brugemann, K., Schlez, K., Zschock, M., Wolter, W., Czerny,
586 C.P., 2011. Flow cytometric differential cell counts in milk for the evaluation of inflammatory
587 reactions in clinically healthy and subclinically infected bovine mammary glands. Journal of Dairy
588 Science 94, 5033-5044. doi:10.3168/jds.2011-4348.
589 Schwarz, D., Kleinhans, S., Reimann, G., Stuckler, P., Reith, F., Ilves, K., Pedastsaar, K., Yan, L., Zhang,
590 Z., Valdivieso, M., Barreal, M.L., Fouz, R., 2020a. Investigation of dairy cow performance in
591 different udder health groups defined based on a combination of somatic cell count and
592 differential somatic cell count. Prev Vet Med 183, 105123. doi:10.1016/j.prevetmed.2020.105123.
593 Schwarz, D., Rivas, A.L., Konig, S., Diesterbeck, U.S., Schlez, K., Zschock, M., Wolter, W., Czerny, C.P.,
594 2013. CD2/CD21 index: a new marker to evaluate udder health in dairy cows. Journal of Dairy
595 Science 96, 5106-5119. doi:10.3168/jds.2013-6804.
596 Schwarz, D., Santschi, D.E., Durocher, J., Lefebvre, D.M., 2020b. Evaluation of the new differential
597 somatic cell count parameter as a rapid and inexpensive supplementary tool for udder health
598 management through regular milk recording. Prev Vet Med 181, 105079.
599 doi:10.1016/j.prevetmed.2020.105079.
600 Silva, V.M., Souza, M.T., Blagitz, M.G., Souza, F.N., Batista, C.F., Alves, A.J., Fernandes, A.C.C.,
601 Sanchez, E.M.R., Ordinola-Ramirez, C.M., da Costa, L., Della Libera, A., 2021. Milk lymphocyte
602 profile and macrophage functions: new insights into the immunity of the mammary gland in
603 quarters infected with Corynebacterium bovis. BMC Vet Res 17, 282. doi:10.1186/s12917-021-
604 02989-5.
605 Smith, K.L., Hillerton, J.E., 2001. Guidelines on Normal and Abnormal Raw Milk Based on Somatic Cell
606 Counts and Signs of Clinical Mastitis. National Mastitis Council Research Committee
607 Sordillo, L.M., Shafer-Weaver, K., DeRosa, D., 1997. Immunobiology of the mammary gland. Journal
608 of Dairy Science 80, 1851-1865. doi:10.3168/jds.S0022-0302(97)76121-6.
609 Souza, F.N., Blagitz, M.G., Batista, C.F., Takano, P.V., Gargano, R.G., Diniz, S.A., Silva, M.X., Ferronatto,
610 J.A., Santos, K.R., Heinemann, M.B., De Vliegher, S., Della Libera, A., 2020. Immune response in
611 nonspecific mastitis: What can it tell us? Journal of Dairy Science 103, 5376-5386.
612 doi:10.3168/jds.2019-17022.
613 Tassi, R., McNeilly, T.N., Fitzpatrick, J.L., Fontaine, M.C., Reddick, D., Ramage, C., Lutton, M.,
614 Schukken, Y.H., Zadoks, R.N., 2013. Strain-specific pathogenicity of putative host-adapted and
615 nonadapted strains of *Streptococcus uberis* in dairy cattle. Journal of Dairy Science 96, 5129-5145.
616 doi:10.3168/jds.2013-6741.
617 Taylor, B.C., Dellinger, J.D., Cullor, J.S., Stott, J.L., 1994. Bovine milk lymphocytes display the
618 phenotype of memory T cells and are predominantly CD8+. Cellular Immunology 156, 245-253.
619 doi:10.1006/cimm.1994.1169.
620 Wiarda, J.E., Davila, K.M.S., Trachsel, J.M., Loving, C.L., Boggiatto, P., Lippolis, J.D., Putz, E.J., 2025.
621 Single-cell RNA sequencing characterization of Holstein cattle blood and milk immune cells during
622 a chronic Staphylococcus aureus mastitis infection. Sci Rep 15, 12689. doi:10.1038/s41598-025-
623 96657-5.
624 Zecconi, A., Zaghen, F., Meroni, G., Sora, V., Martino, P.A., Laterza, G., Zanini, L., 2023. Early Milk
625 Total and Differential Cell Counts as a Diagnostic Tool to Improve Antimicrobial Therapy
626 Protocols. Animals (Basel) 13 doi:10.3390/ani13071143.

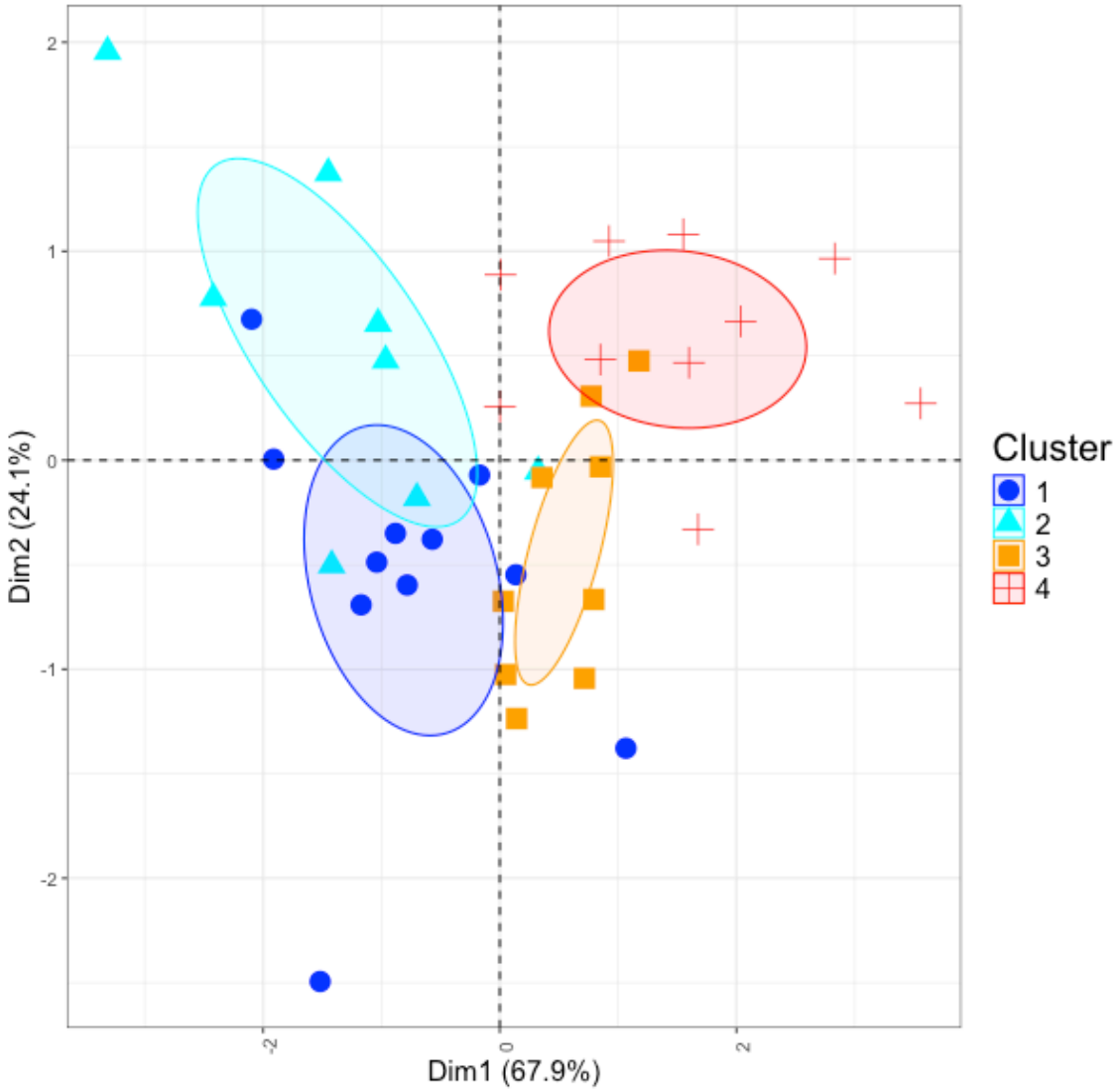
627

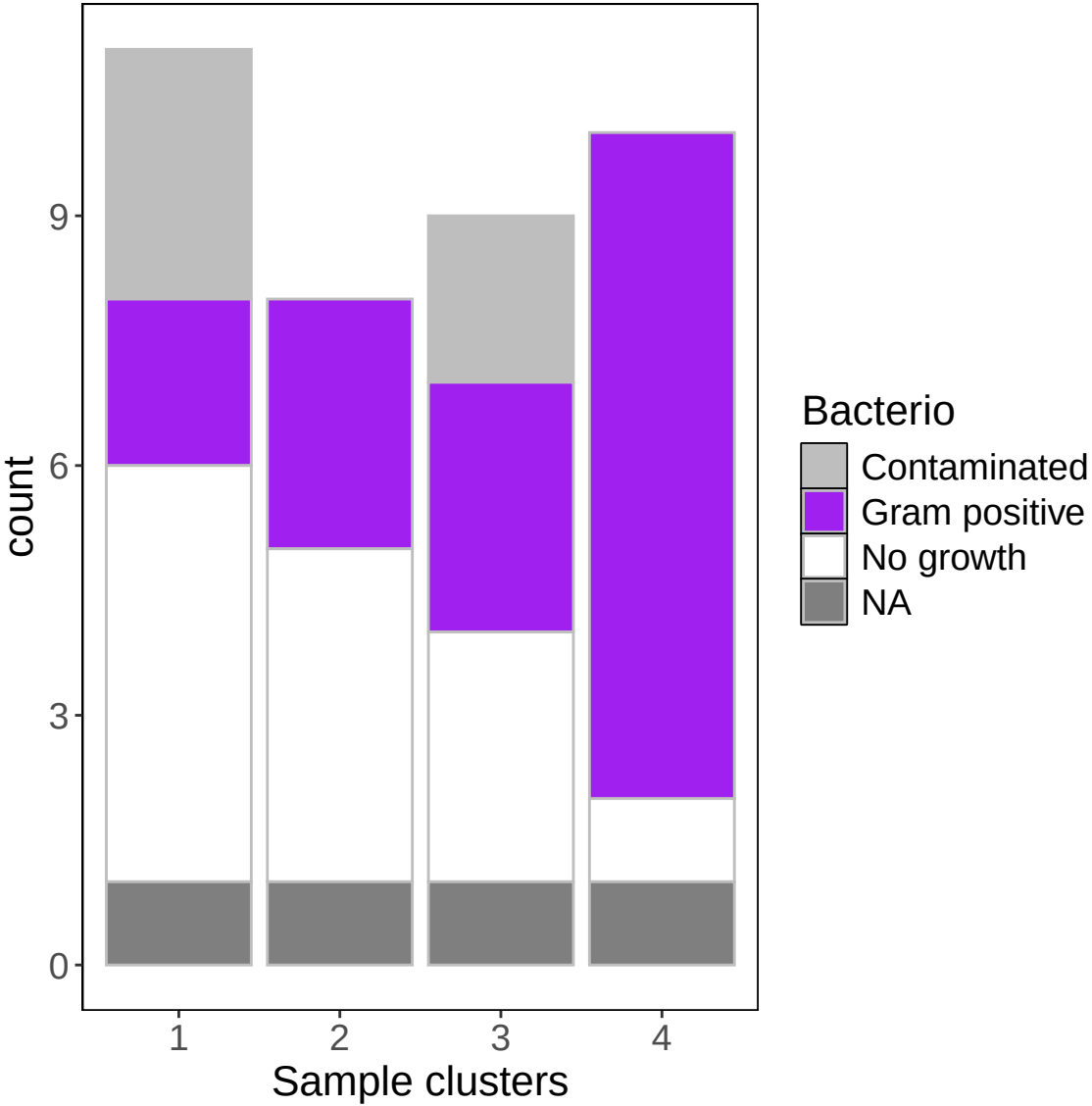

A-
Gene_clusters
VL1744_AVG_Neutro
VL1926_AVG_Neutro
VL1914_AVG_Neutro
VL1744_AVG_Macro
VL1926_AVG_Macro
VL1914_AVG_Macro
VL1914_AVG_Lympho
VL1926_AVG_Lympho
VL1744_AVG_Lympho
CCS
3
0
−3
Gene_clusters
Lc
Macro
PMN1
PMN2
Macro_PMN1
Macro_PMN2
CCS
4.5
2


B-
Dim2 (24.5%)
Dim1 (50%)
8
4
0
−4
−5
0
5
10
Cell types
Lympho
Macro
Neutro


A-
B-
Gene_clusters
Lc
Macro
PMN2
CCS
8
1
Sample_clusters
1
2
3
4
3
0
−3
VL1095_ARG
VL1517_AVG
VL1064_ARG
VL1083_AVG
VL1127_AVD
VL1101_ARD
VL1076_AVD
VL1063_AVG
VL8324_AVD
VL9089_AVD
VL9045_AVD
VL7501_AVG
VL1815_ARD
VL9088_ARD
VL1814_ARD
VL1743_ARD
VL1916_AVD
VL1743_AVG
VL1744_AVG
VL1717_ARG
VL1815_AVD
VL2011_AVG
VL1145_ARD
VL1096_AVD
VL9149_ARD
VL7504_AVD
VL1517_AVD
VL1818_ARG
VL1921_ARD
VL1926_AVG
VL1913_AVD
VL1531_AVG
VL1717_ARD
VL1914_AVG
VL9148_ARD
VL8332_AVG
VL7186_ARD
VL1124_AVD
CCS
Sample_clusters
CD36
LPL
CAT
IRF3
HSPA8
RPL19
RPLP0
RORa
YWHAZ
PPIA
SOD1
TLR1
AHR
IKZF1
DEFB5
CCL2
ITGAM
Bax
CD14
MMP9
S100A9
NCF1
S100A8
CCR1
SPI1
ISG15
MX1
CXCL1
CXCL2
COX2
UPA
TLR6
CXCL8
SOD2
CASP13
TLR4
Dim2 (21.8%)
Dim1 (43.1%)
5
0
−5
−5
0
5
10
Cluster
1
2
3
4


10
8
6
4
2
log(SCC)
1
2
3
4
Cluster
Cluster
1
2
3
4


Dim2 (24.1%)
Dim1 (67.9%)
2
1
0
-1
-2
-2
0
2
Cluster
1
2
3
4


count
0
3
6
9
1
2
3
4
Sample clusters
Bacterio
Contaminated
Gram positive
No growth
NA