Gene expression profiling of bovine raw milk is linked to the inflammatory status of the udder

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 37 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

Implications

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. Improved profiling of the inflammatory status of udders is important in modern farm management. We propose a precise monitoring of the inflammatory status based on gene expression profiling from raw milk. Milk samples are classified in four clusters that differ in terms of cell composition. The method will help the identification of new on-farm criteria for precise monitoring of inflammatory status at the udder quarter level.

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 is a grey zone where one cannot state with confidence whether a quarter is healthy or not. 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 colonized by bacteria (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. While electronic microscope examinations of milk samples revealed mostly macrophages (50- 80%) with 10-20 % lymphocytes (Lee et al., 1980; Lun et al., 2007), flow cytometry analyses have identified T lymphocytes as 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 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). The definition of such thresholds 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 transcriptomic 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 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. Milk was then transported to the lab 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, exceptionally two, from each cow was selected: the only criteria for selection 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 and recording bacterial colonies after 24 hours of incubation at 37°C.

Isolation of milk cells

Milk cells were collected from the selected milk samples: milk was diluted with one volume of DPBS (reference, Sigma D8537), 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)

Flow cytometry analysis of milk cells

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). 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). 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. Data were acquired with 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+ CD14+ cells, neutrophiles as CD45+ G1+ cells and T lymphocytes as CD45+ CD3+ cells.

Cells sorting of purified specific leucocytes populations

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

Immediately after collection, cells were centrigufed 10 minutes at 1000g, resuspended in 1 ml of Trizol and stored at -80°C until RNA was extracted as described below.

RNA extraction and high-throughput RT-qPCR

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. Primers used in this study are listed in **Table 1**. Ten-fold diluted cDNA were preamplified according to Fluidigm's protocol (quick reference PN 100-5875 B1) and preamplified

cDNA were diluted 5-fold with TE. Gene expression level were measured on 96.96 GE Dynamic Array IFC using the Fluidigm BioMark™ HD System.

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).

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 (live CD45+ CD14+), neutrophils (live CD45+ G1+) and T lymphocytes (live CD45+ CD3+) by flow cytometry from three quarter milk samples having somatic cell counts ranging from 45,000 to 389,000 cells/ml (**Supplementary table S1**).

The expression of a panel of 95 genes by these purified sorted cells was investigated by high-throughput RT-qPCR. Only genes yielding a correct amplification in all samples were selected which restricted the number of genes to 79. Expression results for these 79 genes in the three selected cell types were first analyzed by Principal Component Analysis showing that the three populations were clearly separated on the PCA plot (**Figure 1A)**. A hierarchical clustering was then performed and confirmed that samples could be clearly distinguished depending on the cell-type and that 6 clusters of genes could be identified (**Figure 1B**). Each of these 6 clusters of genes shows 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 S2**.

Based on these results, because they showed clearly distinct patterns of expression between the three leucocyte populations investigated, we selected the 42 genes belonging to clusters “Lc” (11 genes), “PMN2” (15 genes) and “Macro” (16 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 S1**.

SCC distribution of selected milk samples are depicted in **Supplementary figure S1**. 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 42 genes selected above, expression values were obtained for all samples for only 37 genes (excluded genes were *CD4*, *CXCL10*, *CCL2*, *IFNB1* and *TNFa*). Analysis by hierarchical clustering of expression values for these 37 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**).

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 is low, samples belonging to cluster 3 are characterized

by a higher expression of “Macro” genes and are thus quarters that are likely responding to some sort of pro-inflammatory stimulus. 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 S2** – 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 S1**). Analysis of these results by Principal Component Analysis showed a clear separation of samples from the 4 clusters identified previously by gene-expression profiling (**Figure 4**).

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. 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 S1**). 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).

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.

Milk cell composition is tightly linked to the sanitary status of the mammary gland (Halasa and Kirkeby, 2020). Hence, the SCC has been, and still is, taken as an essential criteria for dairy herd management programs (Harmon, 1994). Recent developments included a more precise definition of the different cell types present in milk samples, providing more detailed information on the inflammatory status of the mammary gland (Damm et al., 2017).

Another alternative for the identification of healthy mammary gland would be to not only focus on the proportion of different cell types but also consider their functionality. By doing so, one would expect to clearly define a profile of what is a healthy mammary

gland, in addition to the criteria of low SCC. Indeed, even quarters with an SCC below 100 000 are sometimes infected (Schwarz et al., 2010).

We thus chose to investigate 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 to characterize the inflammatory status of low cell counts samples, 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). 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. Analyses were done on quarter milk samples which are likely to improve the diagnostics of inflammatory quarters. Indeed, analyses on composite samples show reduced predictive power compared to analyses on quarter milk samples (Dufour and Dohoo, 2012; Fonseca et al., 2025b).

We focused our analyses on three major cell types: macrophages, neutrophils and T lymphocytes. Among a set of 95 genes, a preliminary screen allowed us to select 79 genes that typifies these three cells types. Of these 79 genes, we restricted our analyses to 42 genes from three clusters clearly distinguishing neutrophils, macrophages and T lymphocytes. The three other gene clusters were not kept because of reduced specificity. For instance, the two clusters labeled “Macro_PMN1” and “Macro_PMN2” led to higher expression values in both macrophages and neutrophils. The PMN1 cluster also gave higher expression values in T lymphocytes samples.

Among the 42 genes in clusters “Lc”, “PMN2” and “Macro”, only 37 were expressed in all milk samples. Despite this slight reduction, our results clearly distinguished two types of milk samples with low somatic cell counts, i.e. samples from clusters 1 and 2. These low SCC milk samples are 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 neutrophils-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). 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 this method (Damm et al., 2017).

Cluster 3 and cluster 4 samples have statistically significant higher SCC values than cluster 1 and 2 samples: they can thus be considered as inflamed quarters but they show clearly different gene expression patterns. Despite similar profiles in terms of expression of macrophages-associated genes, cluster 4 samples show higher levels of expression of neutrophils-associated genes. Different hypotheses can be raised: cluster 3 samples could be 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.

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 and are not conditioned by user-defined thresholds. We therefore believe that this will bring an added value to the definition of inflammatory status of mammary gland quarters.

Complementary analyses would have to 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).

The clustering obtained with gene expression data was compared to clustering of samples based on flow-cytometry data. One caveat of the cytometry approach is the possibility that not all cells are collected by centrifugation. In particular, macrophages might remain in the cream layer after centrifugation (Lee et al., 1980). Although this is a possibility, we should mention that similar proportions of cell types were obtained using an alternative method to collect cells: in this method, milk cells were labeled by

adding directly Vybrant DyeCycle Ruby to raw milk and then sorted with a high-speed cell sorter (data not shown).

Anyhow, the clustering based on gene expression actually overlaps that obtained from flow cytometry: consistent with the increased expression of PMN gene set, cluster 3 and 4 samples are the ones with the highest number of neutrophils while the increased expression of macrophage gene set in cluster 2 samples is consistent with higher counts of macrophages in these samples.

The method presented could be improved by increasing 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. In particular, it would be interesting to investigate genes specific for neutrophil subtypes since our recent data showed an increased proportion of a neutrophil subset expressing MHCII surface molecules during subclinical mastitis (Rambault et al., 2023).

One cell type that we failed to clearly quantify by flow cytometry in our samples is the epithelial cell type. Although they are likely to represent only a low proportion of milk cells, 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.

Because the gene expression profile presented here involves direct extraction of RNA from raw milk without any prior processing and is linked to macrophage, PMN, T lymphocytes and potentially epithelial cells, we believe that this method will provide more detailed information on the cell composition of milk samples than current DSCC methods only based on the proportion of macrophages in milk.

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. More extensive studies will be required to determine the predictive value of these analyses.

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., 2025a).

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 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.

Ethics approval

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

Data and model availability statement

All data and models are available upon request.

Declaration of generative AI and AI-assisted technologies in the writing process

No generative AI nor AI-assisted technologies were used in the writing process.

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

Declaration of interest

The authors have not stated any conflicts of interest.

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 collecting milk samples. We are grateful to Aude Remot for critical reading of the manuscript.

Financial support statement

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

Authors' contributions

This study was conceived by PG, RPM and CG. CG and PG performed experiments. PG designed, obtained funding, analyzed RT-qPCR data, prepared all figures and wrote the manuscript. All authors read and approved the manuscript before publication.

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 during lactating and non-lactating periods. Veterinary Immunology and Immunopathology 65, 51-61. doi:10.1016/s0165-2427(98)00176-7.

Bannerman, D.D., Paape, M.J., Lee, J.W., Zhao, X., Hope, J.C., Rainard, P., 2004. Escherichia coli and *Staphylococcus aureus* elicit differential innate immune responses following intramammary infection. Clinical and Diagnostic Laboratory Immunology 11, 463-472. doi:10.1128/CDLI.11.3.463-472.2004.

Becker, D., Weikard, R., Hadlich, F., Kuhn, C., 2021. Single-cell RNA sequencing of freshly isolated bovine milk cells and cultured primary mammary epithelial cells. Sci Data 8, 177. doi:10.1038/s41597-021-00972-1.

Boutinaud, M., Herve, L., Lollivier, V., 2015. Mammary epithelial cells isolated from milk are a valuable, non-invasive source of mammary transcripts. Front Genet 6, 323. doi:10.3389/fgene.2015.00323.

Canovas, A., Rincon, G., Bevilacqua, C., Islas-Trejo, A., Brenaut, P., Hovey, R.C., Boutinaud, M., Morgenthaler, C., VanKlompenberg, M.K., Martin, P., Medrano, J.F., 2014. Comparison of five different RNA sources to examine the lactating bovine mammary gland transcriptome using RNA-Sequencing. Sci Rep 4, 5297. doi:10.1038/srep05297.

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

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

Dufour, S., Dohoo, I.R., 2012. Monitoring dry period intramammary infection incidence and elimination rates using somatic cell count measurements. Journal of Dairy Science 95, 7173-7185. doi:10.3168/jds.2012-5839.

Farschtschi, S., Hildebrandt, A., Mattes, M., Kirchner, B., Pfaffl, M.W., 2022. Using High-Resolution Differential Cell Counts (HRDCCs) in Bovine Milk and Blood to Monitor the Immune Status over the Entire Lactation Period. Animals (Basel) 12 doi:10.3390/ani12111339.

Fonseca, M., Kurban, D., Roy, J.P., Santschi, D.E., Molgat, E., Dufour, S., 2025a. Usefulness of differential somatic cell count for udder health monitoring: Effect of intramammary infections, days in milk, quarter location, and parity on quarter-level differential somatic cell count and somatic cell score in apparently healthy dairy cows. Journal of Dairy Science 108, 3878-3899. doi:10.3168/jds.2024-25401.

Fonseca, M., Kurban, D., Roy, J.P., Santschi, D.E., Molgat, E., Yang, D.A., Dufour, S., 2025b. Usefulness of differential somatic cell count for udder health monitoring: Identifying referential values for differential somatic cell count in healthy quarters and quarters with subclinical mastitis. Journal of Dairy Science 108, 3917-3928. doi:10.3168/jds.2024-25403.

Grandoni, F., Fraboni, D., Canonico, B., Papa, S., Buccisano, F., Schuberth, H.-J., Hussen, J., 2023. Flow Cytometric Identification and Enumeration of Monocyte Subsets in Bovine and Water Buffalo 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.

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.

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.

Sarikaya, H., Bruckmaier, R.M., 2006. Importance of the sampled milk fraction for the prediction of total quarter somatic cell count. Journal of Dairy Science 89, 4246-4250. doi:10.3168/jds.S0022-0302(06)72470-5.

Schwarz, D., Diesterbeck, U.S., Failing, K., Konig, S., Brugemann, K., Zschock, M., Wolter, W., Czerny, C.P., 2010. Somatic cell counts and bacteriological status in quarter foremilk samples of cows in Hesse, Germany--a longitudinal study. Journal of Dairy Science 93, 5716-5728. doi:10.3168/jds.2010-3223.

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

Schwarz, D., Kleinhans, S., Reimann, G., Stuckler, P., Reith, F., Ilves, K., Pedastsaar, K., Yan, L., Zhang, Z., Valdivieso, M., Barreal, M.L., Fouz, R., 2020a. Investigation of dairy cow performance in different udder health groups defined based on a combination of somatic cell count and differential somatic cell count. Prev Vet Med 183, 105123. doi:10.1016/j.prevetmed.2020.105123.

Schwarz, D., Rivas, A.L., Konig, S., Diesterbeck, U.S., Schlez, K., Zschock, M., Wolter, W., Czerny, C.P., 2013. CD2/CD21 index: a new marker to evaluate udder health in dairy cows. Journal of Dairy Science 96, 5106-5119. doi:10.3168/jds.2013-6804.

Schwarz, D., Santschi, D.E., Durocher, J., Lefebvre, D.M., 2020b. Evaluation of the new differential somatic cell count parameter as a rapid and inexpensive supplementary tool for udder health management through regular milk recording. Prev Vet Med 181, 105079. doi:10.1016/j.prevetmed.2020.105079.

Smith, K.L., Hillerton, J.E., 2001. Guidelines on Normal and Abnormal Raw Milk Based on Somatic Cell Counts and Signs of Clinical Mastitis. National Mastitis Council Research Committee

Sordillo, L.M., Shafer-Weaver, K., DeRosa, D., 1997. Immunobiology of the mammary gland. Journal of Dairy Science 80, 1851-1865. doi:10.3168/jds.S0022-0302(97)76121-6.

Souza, F.N., Blagitz, M.G., Batista, C.F., Takano, P.V., Gargano, R.G., Diniz, S.A., Silva, M.X., Ferronatto, J.A., Santos, K.R., Heinemann, M.B., De Vliegher, S., Della Libera, A., 2020. Immune response in nonspecific mastitis: What can it tell us? Journal of Dairy Science 103, 5376-5386. doi:10.3168/jds.2019-17022.

Tassi, R., McNeilly, T.N., Fitzpatrick, J.L., Fontaine, M.C., Reddick, D., Ramage, C., Lutton, M., Schukken, Y.H., Zadoks, R.N., 2013. Strain-specific pathogenicity of putative host-adapted and nonadapted strains of *Streptococcus uberis* in dairy cattle. Journal of Dairy Science 96, 5129-5145. doi:10.3168/jds.2013-6741.

Taylor, B.C., Dellinger, J.D., Cullor, J.S., Stott, J.L., 1994. Bovine milk lymphocytes display the phenotype of memory T cells and are predominantly CD8+. Cellular Immunology 156, 245-253. doi:10.1006/cimm.1994.1169.

Wiarda, J.E., Davila, K.M.S., Trachsel, J.M., Loving, C.L., Boggiatto, P., Lippolis, J.D., Putz, E.J., 2025. Single-cell RNA sequencing characterization of Holstein cattle blood and milk immune cells during a chronic Staphylococcus aureus mastitis infection. Sci Rep 15, 12689. doi:10.1038/s41598-025-96657-5.

Zecconi, A., Zaghen, F., Meroni, G., Sora, V., Martino, P.A., Laterza, G., Zanini, L., 2023. Early Milk Total and Differential Cell Counts as a Diagnostic Tool to Improve Antimicrobial Therapy Protocols. Animals (Basel) 13 doi:10.3390/ani13071143.

Table 1 : Primers used in the RT-qPCR analysis

Gene	NCBI reference	Forward	Reverse
ABCG2	NM_001037478	GGAGTCATGAAACCTGGCC	GGATCCTTCCTTGCAGCTAA
ACTB	BT030480	ACGGGCAGGTCATCACCATC	AGCACCGTGTTGGCGTAGAG
AHR	NM_001206026	GTGCAGAAAACTGTCAAGCCA	AACATCTGGTGGGAAAGGCAG
AOAH	NM_001078096.2	AACGACAGCAACAAGATGGC	AACACCCCAAATGCCGTTAC
Bax	NM_173894	AGAAGCTGAGCGAGTGTCTGAA	CGCTCTCGAAGGAAGTCCAA
CASP1	XM_002692921	CTCCCACCTGGCAGGAATAC	AGGAGCTGGAAAAGGAGGGA
CASP13	NM_176638.5	TCCGGACATTCAACAACCGT	ACCCACAATTCCCCACGATT
CASP8	NM_001045970.2	AATATTGGGGAGCAGCTGGG	AGGCATCCTTGATGGGTTCC
CAT	NM_001035386.2	AGATGGACACAGGCACATGA	ACTGCCTCTCCATTTGCATT
CCL2	NM_174006	GCTCGCTCAGCCAGATGCAA	GGACACTTGCTGCTGGTGACTC
CCL20	NM_174263	TTCGACTGCTGTCTCCGATA	GCACAACTTGTTTCACCCACT
CCR1	NM_001077839.1	ATGTCTTTGTCGGCGAGAGG	TCTGTGGACAGGAAGGGGAA
CCR6	NM_001194961	AGAAAGCCAAGTCCATGACG	TGGCGTTGCTGAAAATCCAC
CCR7	NM_001024930.3	AGCAAGCAACTCAACATCGC	AGGCGTACAAGAAAGGGTTG
CD14	NM_174008	TCCACAGTCCAGCCGACAAC	AACGGCGCTAGACCAGTCAG
CD25	NM_174358	AACACACAGATGCGCAGAAC	TTCACGTTCGTGTTCCCATG
CD36	NM_001278621	CTGGCTGTGTTTGGAGGGATTC	ACTGTCACTTCATCTGGATTCTGC
CD4	NM_001103225.1	GGTGTTCAGACTGACTGCCA	TTCCATTGCACGGAAGGGTT
CD62L	NM_174182	ATGCCTGCCACAAAGCAAAG	TTCCACACATTGTCCATGGC
CD80	NM_001206439.1	ACTTCTGTTCAGGCATCACC	AGATCCGAAGGCTTGTCAGTT
CDH1	NM_001002763	GAAGAACAGCACGTACACGGC	CCTCTGGCAGAAGTCCATGGT
CLDN1	NM_001001854	AAGACGACGAGGCACAGAAG	AGCCCAGCCAATGAAGAGA
CLEC6A	NM_001034479.1	TACCTGGAAGCCGTTTGGTT	CAAGTGAGCTCCCATCCCAA
COX-2	AF004944	CATGGGTGTGAAAGGGAGGAA	ATTTGTGCCCTGGGGATCAG
CSN1S1	NM_181029.2	GTTCCCAATAGTGCTGAGGAACGACT	GGGTAGAAGTAGGCCAGTTCCTGAT
CSN3	NM_174294	TGCAATGATGAAGAGTTTTTTCCTAG	GATTGGGATATATTTGGCTATTTTGT
CXCL1	NM_175700.2	CCAAACCGAAGTCATAGCC	TCAGTTGGCACTAGCCTTGTTTAGC
CXCL10	NM_001046551	TTCAGGCAGTCTGAGCCTAC	ACGTGGGCAGGATTGACTTG
CXCL2	NM_174299.3	GTGTCTCAACCCCGCCGCTC	TCCAGATGGCCTTAGGAGGTGG
CXCL3	NM_001046513.2	AGCGATGCTGCTCCTGCTCCT	CCATGGGAGCTTCAGGGTTGAG
CXCL5	NM_174300.2	TGTTTAACCACCACACCGGG	AGGTGGCTATCACTTCCACC
CXCL8	NM_173925.2	TGAAGCTGCAGTTCTGTCAAG	TTCTGCACCCACTTTTCCTTGG
CXCR4	NM_174301	GCAGGTAGCAAAGTGACCCT	CGGAAGCAGGGTTCCTTCAT
DEFB5	NM_001130761	TCGTGCTCCTCTTCCTAGTC	GGCACGAGATCGGAATACAG
FAIM3	XM_002693895	TCCTTCCAGTAGTCGGAGCC	TGGTGCATCTTCCAGACTCG
FAS	NM_174662.2	GCCCACATGGCTGGTATCAA	TTTTTCCGTTTGCCAGGAGG
GAPDH	DQ403066	GGCATCGTGGAGGGACTTATG	GCCAGTGAGCTTCCCGTTGAG
GPX1	NM_174076	GCAAGGTGCTGCTCATTGAG	CGCTGCAGGTCATTCATCTG
HSPA8	NM_174345	CGAATCATCAATGAGCCAACTG	TGCCACCCCCTAAATCAAAG
IFITM3	NM_001078141	CCTGAACATCTGCTCCCTGG	CTCGGAGACTGCTTGAACGA

IFNAR1	NM_174552.2	TCCTTTGCCACGTGTCAAGT	AGTAGCGTGAGGGAGACAGA
IFNB1	NM_174350.1	GACAGTCGTGCAAGTGCAAA	CAGTCACGGACGTAACCTGT
IFNL3	NM_001281901.1	GCCAGGGATGCCTTTGAAGA	GAGTTAGCCATGGCCTCCAG
IKZF1	NM_001105257.1	GCCTGTGAAGTGAATGGGGA	CTTTGCTGCCTTGGACACTG
IL1-alpha	NM_174092	CTGAAGAAGAGACGGTTGAG	ATGCATTCCTGGTGGATGAC
IL1-beta	EU276067	CTCTCACAGGAAATGAACCGAG	GCTGCAGGGTGGGCGTATCACC
IL10	NM_174088.1	GTGATGCCACAGGCTGAGAA	TGCTCTTGTTTTCGCAGGGCAG
IL6	EU276071	TGCTGGTCTTCTGGAGTATC	GTGGCTGGAGTGGTTATTAG
IRF3	NM_001029845.3	GGAAGGATAAGCCCGACCTG	GAGTCCTTGCTGTGGTCCTC
ISG15	NM_174366	CGCCCAGAAGATCAATGTGC	TCCTCACCAGGATGGAGATG
ITGAM	NM_001039957.1	TTGAGGCGACGATGGAGTTC	ACTTTCACCTGCCCAGCAAT
LAP	NM_203435	TGCTCCTTGCGCTCCTCTTC	CTCCGAGACAGGTGCCAATC
LCN2	XM_002691670	TGGCGGGGAATGCAATTAAG	TAACAGGATGGAGGTGACGTTG
LPL	NM_001075120	AGTTTATGAACTGGATGGCGGATG	TCAGGAGAAAGGCGACTTGGAG
MIF	NM_001033608	GCAAGCCGGCACAGTACATC	CCGCGTTCATGTCGCAGAAG
MMP9	NM_174744	CGTTCCGACGACATGCTCTG	CATTGCCGTCCTGGGTGTAG
MUC1	NM_174115	CTCTCCAGGCCATGATAGTG	AAGTGACCATGGAGCTTGAC
MX1	NM_173940.2	GGCCACATCCCCTTGATCAT	CGTACTGGTCCTTGTCCTGG
NCF1	NM_174119.4	GCAGGACCTGTCTGAGAAGG	GATGTCCCCCGCCTCAATAG
NFE2L2	NM_001011678	AGGACATGGATTTGATTGAC	TACCTGGGAGATGTTGGCA
NLRP3	NM_001102219.1	CTCAGTGGCAATACCCTGGG	AGCACTGTCCCAACCACAAT
NOS2	NM_001076799	CTTGAGCGAGTGGTGGATGG	ATCTGAGGGCTGGCATAGGG
OAS1Z	NM_001029846.2	CCAATGGTTCTTCTGCCCCT	GGCAGGAGGTGGTCTTTGAT
OASL	NM_001081659.1	GCTCAAAGAAAAGAGGCCCATC	CCATCTGTAGCCTGTTGCCA
OCLN	NM_001082433	GGTAACTTGGAGACGCTTTC	GTTCTGCTTGTAGGCTCTTAT
PPIA	BC105173	TCCGGGATTTATGTGCCAGGG	GCTTGCCATCCAACCACTCAG
RIPK1	NM_001035012	GCGAATCTCTCGGGTTGTGT	GCCTGCTCCAGGAAGTCTG
RORa	NM_001192861	GCAATGCCACCTACTCCTGT	GCGGCCAAATTTTACAGCATC
RPL19	NM_001040516	TACTGCCAATGCTCGAATGC	TGATACATGTGGCGGTCAATC
RPLP0	NM_001012682	CAACCCTGAAGTGCTTGACAT	AGGCAGATGGATCAGCCA
S100A8	NM_001113725	CTCCCTGATTGACGTCTACC	TCCAGGCCCACCTTTATCAC
S100A9	NM_001046328	TGACACCCTGATCCAGAAAG	GCCACCAGCATAATGAACTC
SAA3	NM_181016	CCTCAAGGAAGCTGGTCAAG	TACCTGGTCCCTGGTCATAC
SIRT1	NM_001192980	ATACACTGGAGCAGGTT	TTCATCAGCTGGGCATCTAG
SOCS3	NM_174466.2	GATCCCTCTGGTGTTGAGCC	CAGCTGGGTGACTTTCTCGT
SOD1	NM_174615	TGTTGCCATCGTGGATATTGTAG	CCCAAGTCATCTGGTTTTTCATG
SOD2	NM_201527.2	GCAATTCCCCTTGGGGTTCT	CAGCCTGCGTTGAAGTTCAA
SPI1	NM_001192133.2	CACCCCCGACATCTGATCTG	AAGCGTCAACCATCCTGAGG
STAT1	NM_001077900.1	CAAAGGAAGCCCCAGAGCCTAT	GCCACTCTTCTGTGTTCACTTAC
STAT2	NM_001205689.1	AGCCCGTTTCAGGATCAGC	CAGTGCAGCTTTCTGCCAGTTC
STAT3	NM_001012671	GACCGGTGTCCAGTTCACAA	AAATTTCCGGGACCCTCTGA
STAT5A	NM_001012673.1	CATGTACCCACAGAACCCTGAC	GGGAGAGAGGGCTCCAGACT
TAP	AF014106	GTAGGAAATCCTGTAAGCTGTG	GTGTCTTGGCCTTCTTTTAC

TGFB1	XM_024977949	CCTGAGCCAGAGGCGGACTAC	GCTCGGACGTGTTGAAGAAC
TJP1	XM_024982002	AATGCATCCTGACCACCAGG	GATGGTGCCGGGTTTGTTTC
TLR1	NM_001046504.1	ACCCTACTCTGAACCTCAAG	GACTGCACACTGGATTTCTG
TLR2	NM_174197.2	ACTGGGTGGAGAACCTCATGGTCC	ATCTTCCGCAGCTTACAGAAGC
TLR4	NM_174198.6	GCATGGAGCTGAATCTCTAC	CAGGCTAAACTCTGGATAGG
TLR6	NM_001001159	CTCCGGGAGATAGTCACTTC	GGCCCTGGATTCTATTATGG
TNF alpha	NM_173966.3	TCTTCTCAAGCCTCAAGTAACAAGC	CCATGAGGGCATTGGCATAC
TSLP	XM_024995349	AGAGAGCTACCGGAACATCA	GGGCTGGTCTTCACAGTAGA
TXNRD1	NM_174625	GACCAAACCATCGAAGGAGAGTAT	CTCGTGCAAGCATCTCTTCCT
UPA	NM_174147	GCCAGGAGTCTACACAAGGG	TGGGGTCCTTCAGAGGACAA
YWHAZ	NM_174814	GTAGGAGCCCGTAGGTCATC	CCTCAGCCAAGTAGCGGTAG
ZBTB11	NM_001205556.2	CGTCGAAGAGCACTCACTGT	GGCAGGGAAGGAAGACAGAG

Figure captions

Figure 1: Heatmap and PCA of gene expression of sorted milk macrophages, neutrophils and T lymphocytes.

Cells from three milk samples 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).

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.

Figure 4: Principal component analysis of proportions of milk samples analyzed by flow cytometry.

Milk samples were aseptically collected from 38 individual quarters. Milk was mixed with one volume of DPBS and cells were recovered by centrifugation. Cells were labeled with antibodies against CD3 (T lymphocytes), CD14 (macrophages) and G1 (neutrophils) and cell composition was determined by flow cytometry. A Principal Component Analysis of the cell composition of milk samples 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.

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.

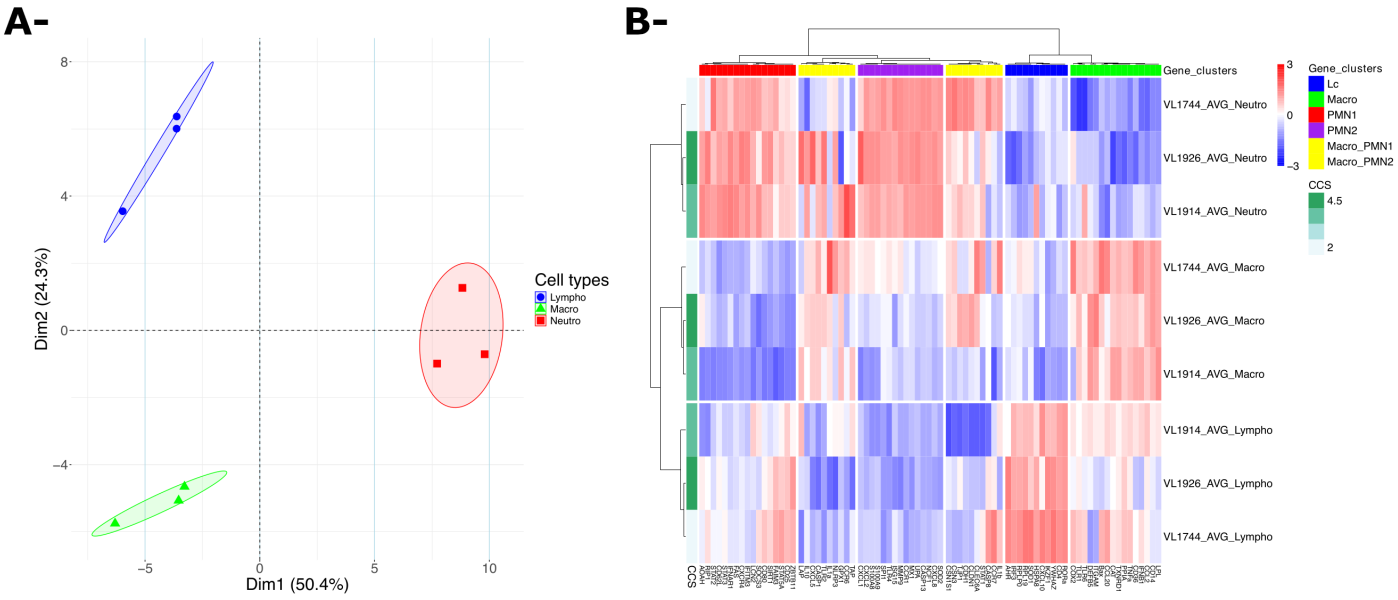

A-
B-
Dim1 (50.4%)
Dim2 (24.3%)
Cell types
Lympho
Macro
Neutro
Gene_clusters
VL1744_AVG_Neutro
VL1926_AVG_Neutro
VL1914_AVG_Neutro
VL1744_AVG_Macro
VL1926_AVG_Macro
VL1914_AVG_Macro
VL1914_AVG_Lympho
VL1926_AVG_Lympho
VL1744_AVG_Lympho
CCS
Lc
Macro
PMN1
PMN2
Macro_PMN1
Macro_PMN2

Figure 1

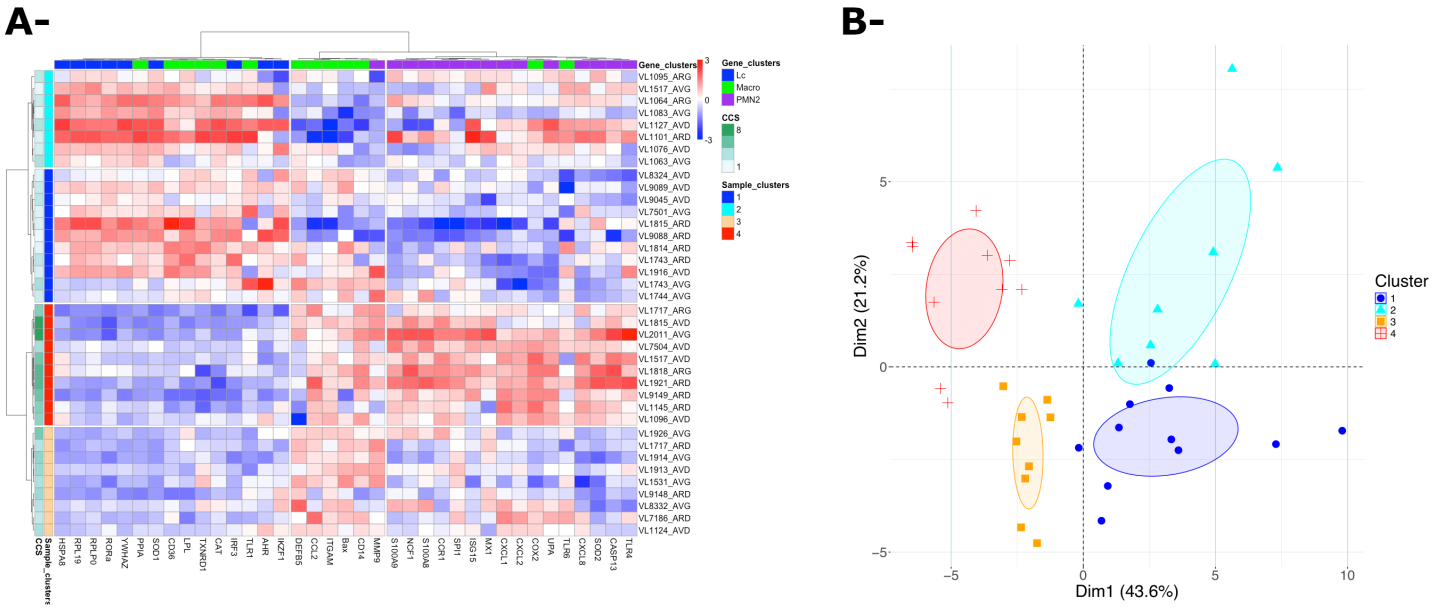

A-
B-
Gene_clusters
Lc
Macro
PMN2
CCS
8
1
Sample_clusters
1
2
3
4
Cluster
Dim1 (43.6%)
Dim2 (21.2%)

Figure 2

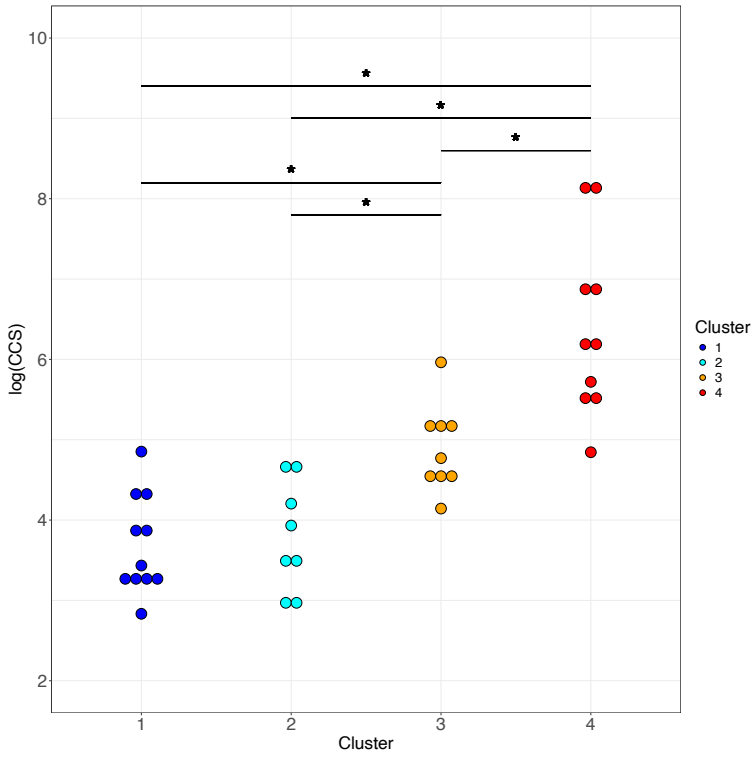


Figure 3

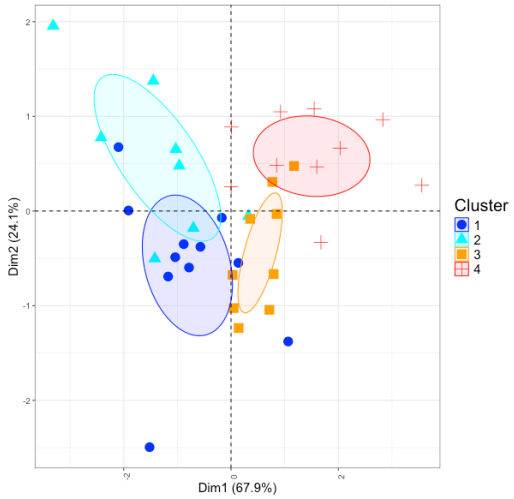


Figure 4

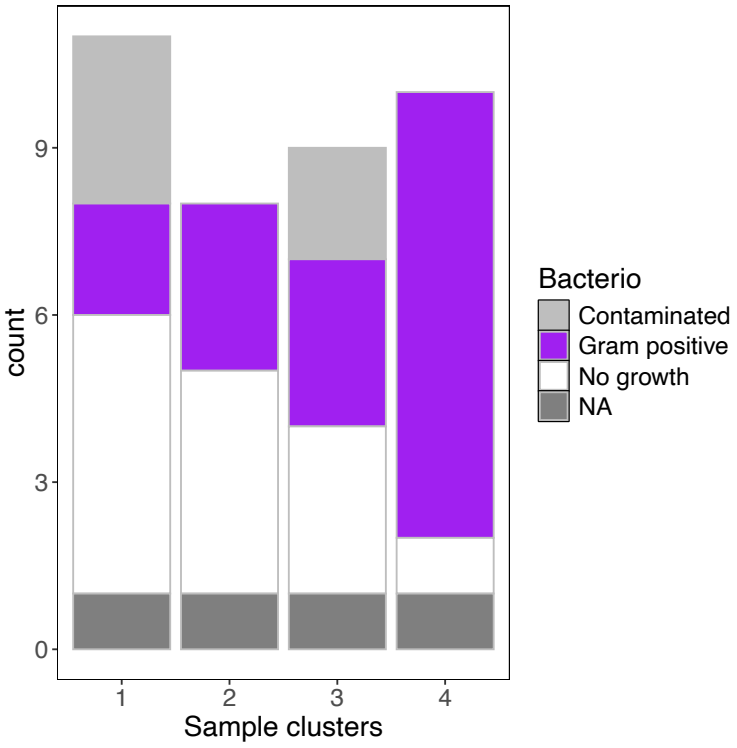

count
9
6
3
0
1
2
3
4
Sample clusters
Bacterio
Contaminated
Gram positive
No growth
NA

Figure 5

Gene expression profiling of raw milk is linked to the inflammatory status of the udder

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

animal journal

Supplementary material

Supplementary Table S1: description of milk samples used in this study

Quarter_ID[a]	Breed	SCC	Percent_Macrophage	Percent_Lymphocytes	Percent_Neutrophils	Counts_Macrophages	Counts_Lymphocytes	Counts_Neutrophils	Cluster	Days_in_milk	Bacteriology
VL1743_ARD	HF	75	22,3	19,3	31,4	17	14	24	1	243	Contaminated
VL1744_AVG[b]	HF	45	16,0	13,0	26,0	7	6	12	1	215	Contaminated
VL1814_ARD	HF	23	26,6	25,9	22,8	6	6	5	1	123	Contaminated
VL9045_AVD	HF	76	9,6	15,2	64,8	7	12	49	1	391	Gplus
VL9089_AVD	NO	51	7,0	28,3	55,0	4	14	28	1	294	Gplus
VL1743_AVG	HF	128	42,9	33,1	10,8	55	42	14	1	243	no_growth
VL1815_ARD	HF	30	4,0	5,9	83,8	1	2	25	1	284	no_growth
VL7501_AVG	NO	27	15,6	28,5	39,9	4	8	11	1	261	no_growth
VL8324_AVD	NO	25	28,1	61,6	0,9	7	15	0	1	265	no_growth
VL9088_ARD	NO	17	9,1	19,4	65,5	2	3	11	1	299	no_growth
VL1916_AVD	HF	31	27,6	23,4	27,8	9	7	9	1	214	ND
VL1063_AVG	NO	111	13,1	12,2	61,0	14	13	68	2	293	Gplus
VL1064_ARG	NO	51	11,0	6,2	75,2	6	3	38	2	291	Gplus
VL1095_ARG	HF	67	5,4	5,0	87,5	4	3	59	2	273	Gplus
VL1076_AVD	HF	36	26,6	14,7	44,9	10	5	16	2	252	no_growth
VL1083_AVG	HF	19	7,5	5,0	83,9	1	1	16	2	259	no_growth
VL1101_ARD	HF	30	3,1	0,5	96,0	1	0	29	2	298	no_growth
VL1127_AVD	HF	101	3,0	1,2	94,3	3	1	95	2	293	no_growth
VL1517_AVG	HF	20	29,4	19,0	24,2	6	4	5	2	341	ND
VL1913_AVD	HF	170	16,0	10,0	60,0	27	17	102	3	201	Contaminated

VL1914_AVG[b]	HF	200	17,0	5,0	63,0	34	10	126	3	199	Contami-nated
VL1124_AVD	HF	118	8,7	15,5	72,3	10	18	85	3	316	Gplus
VL1926_AVG[b]	HF	389	8,0	4,0	79,0	31	16	307	3	193	Gplus
VL7186_ARD	HF	63	13,5	48,0	26,0	9	30	16	3	282	Gplus
VL1531_AVG	HF	107	32,6	26,0	16,3	35	28	17	3	249	no_growth
VL8332_AVG	NO	83	8,5	55,8	19,2	7	46	16	3	269	no_growth
VL9148_ARD	NO	155	5,8	36,6	51,0	9	57	79	3	307	no_growth
VL1717_ARD	HF	90	63,9	6,7	8,5	58	6	8	3	276	ND
VL1096_AVD	NO	224	4,7	2,3	85,7	10	5	192	4	223	Gplus
VL1145_ARD	NO	277	9,7	4,1	75,8	27	11	210	4	212	Gplus
VL1517_AVD	HF	470	14,9	3,3	71,9	70	16	338	4	341	Gplus
VL1717_ARG	HF	305	49,4	6,5	25,2	151	20	77	4	276	Gplus
VL1815_AVD	HF	3231	3,9	1,1	89,7	126	34	2899	4	284	Gplus
VL1921_ARD	HF	956	5,7	1,1	90,6	55	10	866	4	213	Gplus
VL7504_AVD	HF	127	6,2	8,2	75,1	8	10	95	4	271	Gplus
VL9149_ARD	HF	506	6,7	1,3	87,8	34	7	444	4	296	Gplus
VL2011_AVG	HF	3603	3,6	4,2	89,1	129	151	3210	4	35	no_growth
VL1818_ARG	HF	976	9,9	1,9	71,0	97	19	693	4	153	ND

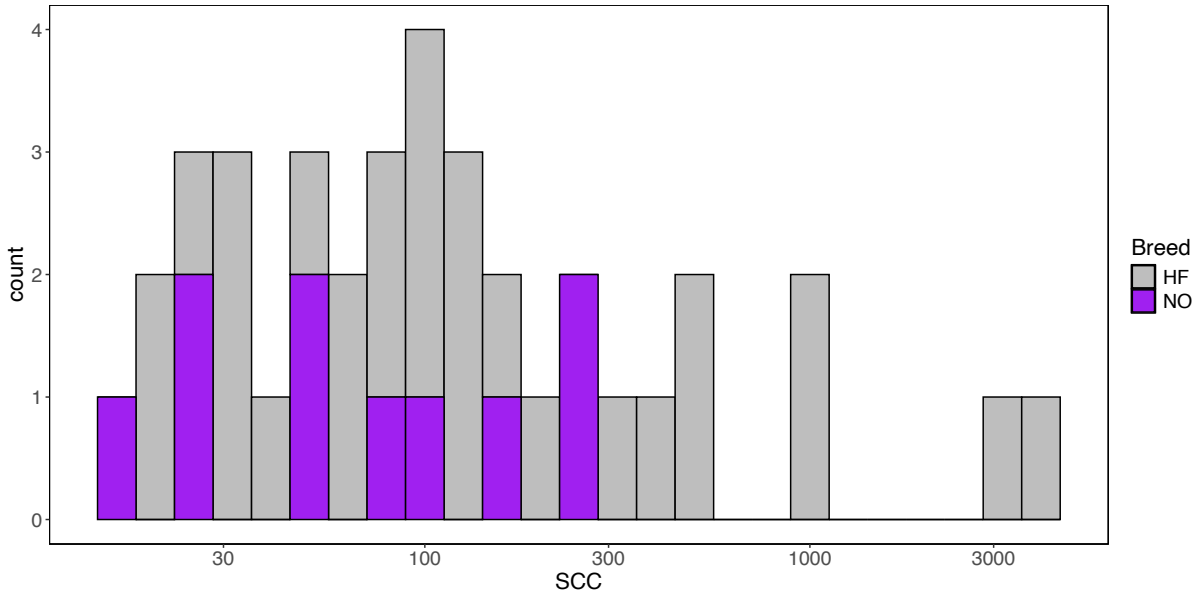
[a] AVD : front right quarter; AVG: front left quarter; ARD: rear right quarter; ARG: rear left quarter
[b] indicates samples included in the preliminary analysis

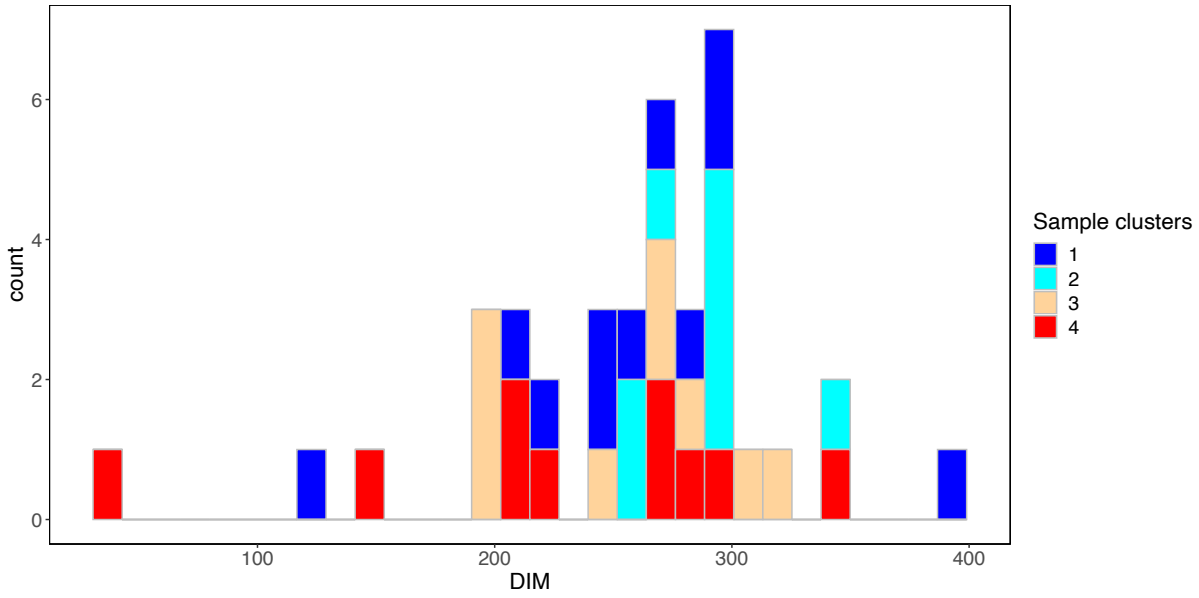
Supplementary Table S2 : genes in the different gene expression clusters

Gene	Cluster	Name_cluster
AHR	1	Lc
CD4	1	Lc
CXCL10	1	Lc
HSPA8	1	Lc
IKZF1	1	Lc
IRF3	1	Lc
RORa	1	Lc
RPL19	1	Lc
RPLP0	1	Lc
SOD1	1	Lc
YWHAZ	1	Lc
AOAH	2	PMN1
CD25	2	PMN1
CD62L	2	PMN1
CD80	2	PMN1
CXCR4	2	PMN1
FAIM3	2	PMN1
FAS	2	PMN1
IFITM3	2	PMN1
IFNAR1	2	PMN1
LCN2	2	PMN1
RIP1	2	PMN1
SIRT1	2	PMN1
SOCS3	2	PMN1
STAT2	2	PMN1
STAT3	2	PMN1
STAT5A	2	PMN1
ZBTB11	2	PMN1
Bax	3	Macro
CAT	3	Macro
CCL2	3	Macro
CCL20	3	Macro
CD14	3	Macro
CD36	3	Macro
COX2	3	Macro
DEFB5	3	Macro
IFNB1	3	Macro
ITGAM	3	Macro
LPL	3	Macro
PPIA	3	Macro
TLR1	3	Macro
TLR6	3	Macro
TNFa	3	Macro
TXNRD1	3	Macro
CASP1	4	Macro_PMN1
CCR6	4	Macro_PMN1
CXCL5	4	Macro_PMN1
GPX1	4	Macro_PMN1
IL1a	4	Macro_PMN1
IL10	4	Macro_PMN1
LAP	4	Macro_PMN1
NLRP3	4	Macro_PMN1
TAP	4	Macro_PMN1
TLR2	4	Macro_PMN1
CASP13	5	PMN2
CCR1	5	PMN2
CXCL1	5	PMN2
CXCL2	5	PMN2

CXCL8	5	PMN2
ISG15	5	PMN2
MMP9	5	PMN2
MX1	5	PMN2
NCF1	5	PMN2
S100A8	5	PMN2
S100A9	5	PMN2
SOD2	5	PMN2
SPI1	5	PMN2
TLR4	5	PMN2
UPA	5	PMN2
CASP8	6	Macro_PMN2
CCR7	6	Macro_PMN2
CDH1	6	Macro_PMN2
CLDN1	6	Macro_PMN2
CLEC6A	6	Macro_PMN2
CSN1S1	6	Macro_PMN2
CSN3	6	Macro_PMN2
IL1b	6	Macro_PMN2
STAT1	6	Macro_PMN2
TJP1	6	Macro_PMN2

Supplementary Figure S1:
Distribution of milk samples based on SCC values. Samples from 28 Holstein (grey) and 10 Normande (purple) cows are represented.



Supplementary Figure S2

Distribution of milk samples based on days-in-milk (DIM) values. Samples are color-coded based on the gene-expression cluster to which they belong.