

# Age Related Changes in the Circulating Leukocyte Transcriptome in Dairy Cows

Laura Buggiotti (✉ [lbuggiotti@rvc.ac.uk](mailto:lbuggiotti@rvc.ac.uk))

The Royal Veterinary College <https://orcid.org/0000-0002-7404-7800>

Zhangrui Cheng

The Royal Veterinary College

Mazdak Salavati

The University of Edinburgh The Roslin Institute

Claire D. Wathes

The Royal Veterinary College

---

## Research article

**Keywords:** ageing, leukocytes, cow, primiparous, multiparous, longevity.

**Posted Date:** December 28th, 2020

**DOI:** <https://doi.org/10.21203/rs.3.rs-133583/v1>

**License:** © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

# Abstract

**Background:** Previous studies of ageing have identified many pathways which are consistently altered in humans and model organisms. Dairy cows are often culled at quite young ages due to an inability to cope adequately with metabolic and infectious diseases, resulting in reduced milk production and infertility. Improved longevity is therefore a desirable trait which would benefit both farmers and their cows. This study analysed the transcriptome derived from RNA-seq data of leukocytes obtained from Holstein cows in early lactation with respect to lactation number.

**Results:** Samples were divided into three age groups for analysis: i) primiparous (PP, n = 53), ii) multiparous in lactations 2-3 (MP 2-3, n = 121), and iii) MP in lactations 4-7 (MP>3, n = 55). Leukocyte expression was compared between PP vs MP>3 cows with MP 2-3 as background using DESeq2 followed by weighted gene co-expression network analysis (WGCNA). Seven modules were significantly correlated ( $r \geq 0.25$ ) to the trait age. Genes from the modules which were more highly expressed in either the PP or MP>3 cows were pooled, and the gene lists subjected to David functional annotation cluster analysis. The top three clusters from modules more highly expressed in the PP cows all involved regulation of gene transcription, particularly zinc fingers. Another cluster included genes encoding enzymes in the mitochondrial beta-oxidation pathway. Top clusters up-regulated in MP>3 cows included the terms *Glycolysis/Gluconeogenesis*, *C-type lectin*, and *Immunity*. Differentially expressed candidate genes for ageing previously identified in the human blood transcriptome up-regulated in PP cows were mainly associated with T-cell function (*CCR7*, *CD27*, *IL7R*, *CAMK4*, *CD28*), mitochondrial ribosomal proteins (*MRPS27*, *MRPS9*, *MRPS31*), and DNA replication and repair (*WRN*). Those up-regulated in MP>3 cows encoded immune defence proteins (*LYZ*, *CTSZ*, *SREBF1*, *GRN*, *ANXA5*, *ADARB1*).

**Conclusions:** Genes and pathways associated with age in cows were identified for the first time to date, and we found that many were comparable to those known to be associated with ageing in humans and model organisms. We also detected changes in energy utilization and immune responses in leukocytes with age.

## Background

Recent studies in human populations have highlighted age-associated changes in leucocyte functionality affecting innate and adaptive immune functions [1]. The causes and consequences of ageing on the human blood transcriptome have, however, proved difficult to dissect due to interactions with environmental influences, genetic factors and a large number of age-related diseases [2]. Studies on model organisms have highlighted that ageing is characterized by many alterations at molecular, cellular and tissue level [3]. Studies of the ageing transcriptome have been performed in species including *C. elegans* [4], flies [5], rodents [6] and humans [7]. This approach has identified various signatures found to occur repeatedly across different tissues and organisms. Candidate genes whose expression is consistently associated with cellular ageing have been classified into six categories: i) downregulation of genes encoding mitochondrial proteins, ii) downregulation of the protein synthesis machinery, iii)

dysregulation of immune system genes, iv) a reduction in growth factor signalling, v) constitutive responses to stress and DNA damage, and vi) dysregulation of gene expression and mRNA processing [8].

Although all living creatures age at some point, our knowledge on the biology of ageing is still not sufficient. The physiological process of ageing in humans is associated with a progressive loss of function and increased vulnerability to disease, frailty, and disability [9]. Because the incidence of adult diseases increases with age, a better understanding of the biology of ageing could greatly improve our efforts to elucidate the pathophysiology of such conditions [10].

Longevity is an economically important trait in dairy cows, which also has welfare implications. Cattle can potentially live for over 20 years, but this is rare in practice and the average lifespan in dairy cows is currently around 4.5 to 7 years [11, 12]. To maximise economic potential, heifers should first calve at 24 months of age so becoming primiparous cows, starting their first lactation, and beginning to pay back their rearing costs through the production of saleable milk [13]. Ideally, they should continue to calve at annual intervals, with their milk production potential increasing progressively until the fourth or fifth lactation [14]. Cows which survive for longer achieve greater lifetime milk production associated with higher profitability [12, 15] together with the benefit of reduced greenhouse gas emissions [16]. Many cows do not achieve optimum longevity due to premature culling, for which the main reasons are mastitis, infertility and lameness [17, 18]. There are major alterations in the metabolic profiles associated with the start of lactation [19]. These also change with age as primiparous cows are still growing, so their energetic demands for growth compete with those of milk production, whereas in older cows the higher milk yields are associated with greater mobilisation of body tissue, leading to higher circulating concentrations of nonesterified fatty acids (NEFA) and beta-hydroxybutyrate (BHB) [20]. Cows with greater milk production potential face an increased risk of glucose shortages in their immune cells, which contributes to immune dysfunction in the peripartum period [21]. Both metabolic and immune dysfunction therefore impact on the transcriptomic changes in leukocytes in early lactation.

In this study we have compared the leucocyte transcriptome between young (primiparous, PP) and older (multiparous, MP) cows in order to shed light on the genes and related genomic pathways involved in age-related symptoms arising during the different phases of a cow's life. This has the potential to inform both breeding and management practices, so providing a significant gain to both agricultural production and animal welfare.

## Results

### **Weighted gene co-expression network analysis (WGCNA) to determine the relationships of leucocyte gene transcription between cow parity**

Whole blood transcriptomes of 229 cow samples from six dairy herds were obtained in early lactation. These had an average mapping rate against the reference genome of 96.2%, resulting in an average total number of reads of 34,439,525 (Supplementary Table 1). As we were most interested in understanding

the effect of ageing, we contrasted the leukocyte expression between the two extreme age groups, PP vs MP > 3, using those classified as MP2-3 as background. Firstly, we performed a differential expression analysis using DESeq2. Over 5,000 genes out of the 17,216 genes initially mapped were found to be differentially expressed between age groups following Benjamini-Hochberg false discovery rate adjustment ( $p_{adj} < 0.1$ ). This list was narrowed down to 2,925 differentially expressed genes (DEG) when contrasting young vs old cows (Supplementary Fig. 1). We then applied WGCNA to identify gene modules from our dataset associated to the trait of age. From an initial total of 21,207 genes, 13,769 (64.9%) genes passed the DESeq2 filtering steps. Samples from three cows were removed at this stage as outliers (two from MP3 and one from MP2-3 age group, respectively; Supplementary Fig. 2), leaving a total of 226 samples. This unsupervised technique identified 32 interconnected gene modules from the filtered gene list (Fig. 1A). The number of genes per module is shown in Fig. 1B. Of these, the violet and turquoise modules were significantly related to the trait age in both the PP and MP > 3 age groups (Fig. 1C). As expected, modules with a significant positive correlation in the younger cows were negatively correlated with age in the older ones, and vice versa, while no modules showed a significant correlation with the medium age group. Within these seven significant modules, 2,274 genes were more highly expressed in the leukocytes of the PP cows and 2,721 genes were more highly expressed in the leukocytes of the MP > 3 cows. The top 20 genes with the highest differential expression between PP and MP > 3 cows are listed in Table 1.

Table 1

Top 20 genes with the highest differential expression between PP and MP > 3 cows. Positive fold change (FC), more highly expressed in PP cows; negative FC (shaded), more highly expressed in MP > 3 cows.

| Gene Symbol         | Gene name   | logFC  | adj.p.value |
|---------------------|---|--------|-------------|
| <i>NELL2</i>        | neural EGFL like                                    | -1.581 | 1.40E-63    |
| <i>ZNF462</i>       | zinc finger protein 462                             | 1.931  | 1.79E-38    |
| <i>PLCL1</i>        | phospholipase C like 1 (inactive)                   | 1.042  | 2.45E-37    |
| <i>IGF2BP3</i>      | insulin like growth factor 2 mRNA binding protein 3 | 2.268  | 6.36E-37    |
| <i>CNTNAP1</i>      | contactin associated protein 1                      | 1.675  | 2.85E-35    |
| <i>SOX13</i>        | SRY-box transcription factor 13                     | 1.925  | 8.17E-35    |
| <i>BLK</i>          | BLK proto-oncogene, Src family tyrosine kinase      | 0.881  | 1.23E-32    |
| <i>MYBL1</i>        | MYB proto-oncogene like 1                           | -0.771 | 1.93E-32    |
| <i>CD163L1</i>      | CD163 molecule like 1                               | 1.463  | 6.55E-32    |
| <i>WNT5A</i>        | Wnt family member 5A                                | 2.673  | 2.56E-31    |
| <i>SCRN1</i>        | secernin 1  | 1.129  | 5.03E-31    |
| <i>LOC515828</i>    | uncharacterised                                     | 1.207  | 6.12E-30    |
| <i>CSPG4</i>        | chondroitin sulfate proteoglycan 4                  | 2.831  | 8.30E-30    |
| <i>DBN1</i>         | drebrin 1   | 0.942  | 1.95E-29    |
| <i>PTPRK</i>        | protein tyrosine phosphatase receptor type K        | -3.301 | 2.00E-29    |
| <i>LAMA4</i>        | laminin subunit alpha 4                             | 2.998  | 4.15E-29    |
| <i>TCAF1</i>        | TRPM8 channel associated factor 1                   | 1.050  | 1.17E-28    |
| <i>LOC751811</i>    | WC1 isolate CH211                                   | 1.389  | 2.85E-28    |
| <i>SOX4</i>         | SRY-box transcription factor 4                      | 0.820  | 3.69E-28    |
| <i>LOC104969122</i> | uncharacterised                                     | -1.396 | 9.97E-28    |

## Significant modules in leukocytes of both PP and MP > 3 COWS

The violet module containing 87 genes had the highest correlations with age (+ 0.59 in PP and - 0.43 in MP > 3 cows, Fig. 1C). Within this module the *LAMA4* gene was the most highly correlated gene (Table 1, gene significance =  $9.05 \times 10^{-22}$ ), with lower expression in the MP > 3 cows, representing its module membership. This is a laminin gene, part of the family of extracellular matrix glycoproteins which form

the major non-collagenous constituent of basement membranes. GO enrichment analysis of the genes in this module revealed the terms *receptor binding* and *cell periphery* (Fig. 2A).

The turquoise module represents the second module which had significant correlations in opposite directions with PP and MP > 3 cows, of -0.29 and 0.33, respectively. This module contained 1,818 genes of which *NELL2* was the most highly correlated with age (Table 1). This was also overall the most differentially expressed gene in leukocytes with higher expression in MP > 3 cows. It was a large module containing 1,818 genes with a wide variety of functions (Fig. 2B, Supplementary Table 2). The top six most significant GO terms (all  $< 1.36 \times 10^{-25}$  and containing between 389 and 959 genes) were *extracellular region*, *membrane*, *single-organism process*, *vesicle* and *cell periphery*.

## Modules up-regulated in leukocytes of PP cows

The modules midnightblue, darkred, and orange were all significantly positive in PP cows with correlations spanning from 0.27 to 0.25, implying lower expression in leukocytes as cows aged (Fig. 2A, Supplementary Table 2). The most significant GO terms in the midnightblue module were all biological processes: *protein autophosphorylation*, *immune system process*, *cell activation* and *regulation of cell-cell adhesion*. The darkred module had enrichment of biological regulation and the molecular function term *cell periphery*, while the orange module was mainly enriched with genes involved in *cilium* and *cilium organisation*; *RNA processing and regulation of gene expression*, and *microtubule organizing center*.

## Black Module: down-regulated in leukocytes of PP cows

The black module was down-regulated in leukocytes of PP (correlation of -0.25) and almost significantly up-regulated in MP > 3 (correlation of 0.24). Genes were therefore more highly expressed in the older cows. GO enrichment analysis of genes in this module (Fig. 2B, Supplementary Table 2) showed that the most significant terms all related to cellular components (*membrane*, *vesicle*, *cytoplasm*, *extracellular region* and *cell*). Within the biological process category *single-organism process*, *monocarboxylic acid metabolic process*, *small molecule catabolic process* and *negative regulation of cellular process* were significant.

## Yellow Module: down-regulated in leukocytes of MP > 3 COWS

Lastly, the yellow module was significantly down-regulated in leukocytes as cows aged (MP > 3; correlation - 0.27). In this module *IGF2BP3* (insulin like growth factor 2 mRNA binding protein 3) was the most highly correlated gene (Table 1). Within this module the two cellular component terms *binding* and *nucleic acid binding transcription factor activity* were most highly represented, together with three biological processes (*regulation of gene expression*, *regulation of nitrogen compound metabolic process* and *regulation of cellular macromolecule biosynthetic process*) (Fig. 2A, Supplementary Table 2).

## Functional annotation cluster analysis

The next stage of the analysis involved pooling all the genes from the modules which were more highly expressed in the PP cows (yellow, orange, midnightblue, darkred, violet) and those which were more highly expressed in the MP > 3 cows (black and turquoise). These two gene lists were then subjected to David functional annotation cluster analysis [22]. All the results are provided in Supplementary Table 3. The top six most significant clusters in each category are summarised in Table 2, with the genes in each of these clusters listed in Supplementary Tables 4A and 4B.

Table 2

Summary of the six most significant clusters obtained from DAVID functional annotation cluster analysis of the DEG in the modules which were significantly positively or negatively correlated to age in primiparous (PP) cows, indicating either lower or higher gene expression in older multiparous MP > 3 cows.

| Positive Modules   | Negative Modules   |
|--|--|
| <p><b>Cluster 1: Enrichment Score: 16.18</b></p> <ul style="list-style-type: none"> <li>• Krueppel-associated box</li> <li>• Zinc finger C2H2-type/integrase DNA-binding domain</li> <li>• Zinc finger, C2H2</li> <li>• Nucleic acid binding</li> <li>• Zinc finger, C2H2-like</li> <li>• KRAB</li> <li>• ZnF_C2H2</li> <li>• Regulation of transcription, DNA-templated</li> <li>• Metal ion binding</li> </ul> | <p><b>Cluster 1: Enrichment Score: 29.73</b></p> <ul style="list-style-type: none"> <li>• Membrane</li> <li>• Transmembrane helix</li> <li>• Transmembrane</li> <li>• integral component of membrane</li> </ul>  |
| <p><b>Cluster 2: Enrichment Score: 7.01</b></p> <ul style="list-style-type: none"> <li>• Zinc-finger</li> <li>• Metal ion binding</li> <li>• Zinc</li> <li>• Metal-binding</li> </ul>  | <p><b>Cluster 2: Enrichment Score: 19.87</b></p> <ul style="list-style-type: none"> <li>• Signal</li> <li>• Disulfide bond</li> <li>• Glycoprotein</li> <li>• Signal peptide</li> <li>• Glycosylation site:N-linked (GlcNAc...)</li> <li>• Disulfide bond</li> </ul> |
| <p><b>Cluster 3: Enrichment Score: 6.83</b></p> <ul style="list-style-type: none"> <li>• Transcription regulator SCAN</li> <li>• Retrovirus capsid, C-terminal</li> <li>• SCAN</li> </ul>  | <p><b>Cluster 3: Enrichment Score: 7.32</b></p> <ul style="list-style-type: none"> <li>• Transmembrane region</li> <li>• Topological domain:Cytoplasmic</li> <li>• Topological domain:Extracellular</li> </ul>   |

| Positive Modules  | Negative Modules  |
|---|---|
| <b>Cluster 4: Enrichment Score: 5.05</b> <ul style="list-style-type: none"> <li>• Nucleotide binding</li> <li>• Nucleotide-binding, alpha-beta plait</li> <li>• RNA recognition motif domain</li> <li>• RRM</li> </ul>  | <b>Cluster 4: Enrichment Score: 3.94</b> <ul style="list-style-type: none"> <li>• Glycolysis / Gluconeogenesis</li> <li>• Glycolysis</li> <li>• Glycolytic process</li> <li>• Carbon metabolism</li> <li>• Biosynthesis of amino acids</li> </ul> |
| <b>Cluster 5: Enrichment Score: 2.02</b> <ul style="list-style-type: none"> <li>• BTB/POZ fold</li> <li>• BTB/POZ-like</li> <li>• BTB</li> </ul>  | <b>Cluster 5: Enrichment Score: 3.70</b> <ul style="list-style-type: none"> <li>• C-type lectin fold</li> <li>• C-type lectin-like</li> <li>• C-type lectin</li> <li>• CLECT</li> <li>• C-type lectin, conserved site</li> </ul>                  |
| <b>Cluster 6: Enrichment Score: 1.88</b> <ul style="list-style-type: none"> <li>• Immunoglobulin-like domain</li> <li>• Immunoglobulin-like fold</li> <li>• Immunoglobulin subtype 2</li> <li>• Immunoglobulin I-set</li> <li>• Immunoglobulin subtype</li> <li>• IGc2</li> <li>• IG</li> </ul> | <b>Cluster 6: Enrichment Score: 3.28</b> <ul style="list-style-type: none"> <li>• Immunity</li> <li>• Innate immunity</li> <li>• Innate immune response</li> </ul>  |

Clusters 1, 2 and 3 from the modules which were positively correlated in the PP cows were all involved in the regulation of gene transcription. Cluster 1 consisted almost entirely of 56 genes encoding zinc finger proteins. The most significant term in Cluster 1 was *Krüppel-associated box*. All but five of the Cluster 1 genes were included in the much larger Cluster 2 with 142 genes, for which the most significant terms were *Zinc-finger* and *metal ion binding*. Cluster 2 also contained all 18 genes from Cluster 3, for which the three significant terms were *Transcription regulator SCAN*, *Retrovirus capsid*, *C-terminal* and *SCAN*. Cluster 4 contained 61 genes involved in *nucleotide binding*. Cluster 5 contained 31 genes for which the most significant term was *BTB/POZ fold*. Finally, Cluster 6 contained 50 genes with the most significant term being *Immunoglobulin-like domain*. Overall, this analysis shows that the most significant functions which were more highly expressed in the younger animals all related to the regulation of gene transcription.

Turning to the genes which were more highly expressed in the MP > 3 cows, Clusters 1, 2 and 3 were all large clusters containing 471, 435, and 254 genes, respectively, with quite general functions relating to the cell surface, transport across it and signalling. In Cluster 1 the top three terms were *Membrane*, *Transmembrane helix* and *Transmembrane*, in Cluster 2 they were *Signal*, *Disulfide bond* and *Glycoprotein* and in Cluster 3 *transmembrane region*, *topological domain: Cytoplasmic* and *topological domain: Extracellular*. The top term in Cluster 4 was *Glycolysis/Gluconeogenesis* and this contained 27 genes, nearly all of which encoded enzymes involved in the glycolytic pathway (Fig. 3).

The top term in Cluster 5 was *C-type lectin fold*, and 11 of the 25 genes in this cluster encoded C-lectins. Finally, Cluster 6 contained 45 genes involved in *Immunity* and *Innate immunity*. These are illustrated in Fig. 4, which shows that they mainly encoded proteins for pattern recognition receptors (PRR), the bovine MHC complex, the complement system and several with antimicrobial activity. Another noteworthy cluster which was negatively related to age was Cluster 16 (Supplementary Table 3). The top pathway in this was *Fatty acid degradation* ( $P < 0.0013$ ). This cluster contained a number of genes encoding enzymes in the mitochondrial beta-oxidation pathway, which converts fatty acids into acetyl CoA to enter the TCA cycle (*ACSL1*, *ACSL4*, *ACSL5*, *ACSL6*, *CPT1*, *CPT2*, *ACOX1*, *ACADVL*, *ECHS1*, *ACAD5*, *ACAA1*, *ACAT2*, *ACADS*, *ACADVL*). This implies greater use of this pathway to supply energy in the leukocytes of the PP cows.

## Comparison with other species

Finally, we investigated the overlap between genes significantly associated with age in leucocytes in our study and 170 candidate genes identified from previous studies of ageing in humans and model organisms (mainly mice and *C. elegans* [2]). Almost one quarter of these candidate genes (38/170, 22.3%) were present in our list of DEG, of which 14 and 24 respectively were from positive and negative significantly correlated modules, indicating reduced or increased expression in older cows. Of these 16 of the 38 DEG were previously identified as being age-related in the human blood transcriptome ([2]; Table 3).

Table 3

List of genes overlapping candidate age-related genes found in humans and with differential expression between PP (highly expressed) and MP > 3 (highly expressed) cows.

|        | Gene name | Module       | Description  | Reference |
|--------|-----------|--------------|--|-----------|
| PP     | CCR7      | darkred      | Control the migration of memory T cells to inflamed tissues, as well as stimulating dendritic cell maturation                                      | [62]      |
|        | CD27      | midnightblue | Required for generation and long-term maintenance of T cell immunity   | [62]      |
|        | FOXO1     | midnightblue | Transcription factor which may play a role in myogenic growth and differentiation  | [63]      |
|        | IL7R      | midnightblue | Plays a critical role during lymphocyte development (age and longevity expression association)   | [64]      |
|        | CAMK4     | midnightblue | Implicated in transcriptional regulation in lymphocytes, neurons and male germ cells (longevity genetics candidate)                                | [65]      |
|        | CD28      | midnightblue | Essential for T-cell proliferation and survival, cytokine production, and T-helper type-2 development  | [62]      |
|        | WRN       | yellow       | RecQ helicase important in the maintenance of genome stability. Mutations cause Werner syndrome (premature ageing)                                 | [66]      |
|        | MRPS27    | yellow       | Mitochondrial ribosomal protein assisting protein synthesis within the mitochondrion (mitochondrial ageing)  | [33]      |
|        | MRPS9     | yellow       | Mitochondrial ribosomal protein assisting protein synthesis within the mitochondrion (mitochondrial ageing)  | [33]      |
|        | MRPS31    | orange       | Mitochondrial ribosomal protein assisting protein synthesis within the mitochondrion (mitochondrial ageing)  | [33]      |
| MP > 3 | SREBF1    | black        | Transcription factor involved in sterol biosynthesis   | [63]      |
|        | ANXA5     | turquoise    | Potential role in cellular signal transduction, inflammation, growth and differentiation   | [67]      |
|        | LYZ       | turquoise    | Antimicrobial peptide  | [67]      |
|        | CTSZ      | turquoise    | Lysosomal cysteine protease with role in immune defense  | [67]      |
|        | ADARB1    | turquoise    | Enzyme in the A(adenosine) to I (inosine) RNA editing pathway, involved in a general maintenance of cellular health (longevity genetics candidate) | [49]      |

| Gene name | Module    | Description  | Reference |
|-----------|-----------|--|-----------|
| GRN       | turquoise | Granulin precursor protein which regulates cell growth. Also, co-ordinates CpG trafficking to TLR9, promoting activation of signalling cascade | [67]      |

Those up-regulated in PP cows were mainly associated with T-cell development and function (*CCR7*, *CD27*, *IL7R*, *CAMK4*, *CD28*) and protein synthesis within the mitochondria (*MRPS27*, *MRPS9*, *MRPS31*), and also included *WRN*, a gene associated with premature ageing in humans [23]. Those up-regulated in MP > 3 cows included genes encoding proteins involved in immune defence (*LYZ*, *CTSZ*, *SREBF1*, *GRN*, *ANXA5* and *ADARB1*).

## Discussion

Livestock species are raised primarily for their economic benefit to humans. Most dairy cows are culled before they reach the end of their potential lifespan due to poor milk production or fertility and/or an increased prevalence of diseases such as those causing mastitis or lameness [17, 18]. The main focus of previous studies into longevity in cows has been to increase the average survival time in the milking herd, so improving the profitability of the dairy industry. For example, one genome wide association study (GWAS) which investigated longevity found genes such as *NPFFR2*, previously identified as a candidate gene for mastitis resistance and two zinc finger proteins (*ZNF613*, *ZNF717*), which have been associated with calving difficulties [24]. Information about the age-related morbidities and causes of death that afflict cattle due to natural ageing is, however, limited. In contrast, there is a growing body of previous work into the underlying causes of cellular ageing which has been based on studies of human populations and model organisms.

This study is the first, to our knowledge, to explore age-related changes in a cattle population by measuring global gene expression in their leukocytes. For this we compared first lactation PP cows versus older multiparous MP > 3 cows using WCGNA analysis in order to identify potential genes and related pathways involved in the ageing process. The blood samples were all collected in early lactation, a time when lactating cows are placed under metabolic stress due to the nutrient requirements of milk production [19, 25] but none are pregnant. These aspects improved comparability between samples but meant that we did not include any data from younger growing heifers of < 2 years. Despite the many differences in physiology, lifestyle and lifespan between species, many of the genes and pathways identified in cows were the same as those highlighted in previous studies of ageing in other species [2–4, 7].

All the samples analysed were whole blood leukocytes. Many aspects of the immune system alter during ageing, eventually leading to immunosenescence [8]. During this process different immune cell subsets are affected in different ways [1]. There is also an increase in baseline systemic inflammation with age, termed “inflammaging” [26, 27]. This may arise as a result of multiple mechanisms, including the

accumulation of misfolded proteins, impaired clearance of senescent cells and obesity. Most individual leukocytes have a short lifespan in the circulation, measured in days or weeks, before they are destroyed by the lymphatic system, although there is a small pool of long-lived T and B lymphocytes which can survive for years, providing immunological memory. In humans, the relative abundance of naive T-cells decreases with chronological age while the population of memory T-cells increases [2, 28, 29]. The leukocyte transcriptome provides a reflection of the basic functions required for cell survival together with the various responses of the different cell types to the changing environment within the body, which alters with key factors such as disease exposure and nutrition. The age-related changes reported here will, to some extent, reflect altered abundance of different cell types as well as their changing expression patterns. This is partly mitigated by the use of WGCNA, which identifies networks of co-expressed genes whose expression is highly correlated. The use of whole blood also avoids the pitfall of potential artefacts which can be induced during cell separation procedures.

Previous work based on studies of human populations and model organisms has identified a number of transcriptional signatures for cellular ageing which occur repeatedly across different tissues and organisms and which segregate into six main groups [8, 30]. In brief these are i) downregulation of genes encoding mitochondrial proteins; ii) downregulation of the protein synthesis machinery (including ribosome biogenesis); iii) dysregulation of immune system genes, immune senescence; iv) reduction in growth factor signalling; v) constitutive responses to stress and accumulated DNA damage, and vi) dysregulation of processes regulating gene expression and mRNA processing (transcription and translation). We have obtained evidence for all of these within our population of cows.

## Mitochondria and oxidative stress

Mitochondria regulate a multitude of different metabolic and signalling pathways and also play an important role in programmed cell death [31]. Oxidative metabolism causes endogenous production of free radical molecules and oxidative damage accumulates in multiple tissues and species with age [30]. Ageing in the model species *Drosophila melanogaster* was associated with down-regulation of numerous mitochondrial genes, including those involved in the electron-transport-chain and mitochondrial metabolism [28, 32]. Knockdown of mitochondrial ribosomal proteins was shown to reduce mitochondrial respiration and activate the mitochondrial unfolded protein response [33] while accumulated mutations in somatic mitochondrial DNA (mtDNA) and respiratory chain dysfunction were associated with ageing in mice [34]. In our study, genes encoding proteins involved in fatty acid beta-oxidation were more highly expressed in leukocytes in PP cows. There were also age-related changes in expression of genes encoding mitochondrial ribosomal proteins, with higher expression of *MRPS9*, *MRPS27* and *MRPS31* in leukocytes of PP cows whereas *MRPL17* and *MRPL38* showed enriched expression in MP > 3 cows.

## Protein synthesis

Protein homeostasis is essential to maintain protein structure and function, but the control of this process declines during ageing [30]. In our study, GO enrichment analysis of all the genes in the violet module, which were more highly expressed in the younger cows, revealed their involvement in *protein*

*autophosphorylation* and *catalytic activity* (enzyme activity). Autophosphorylation is a type of post-translational modification of proteins. In eukaryotes, this process occurs by the addition of a phosphate group to serine, threonine or tyrosine residues within protein kinases, normally to regulate the catalytic activity. Genes involved in *regulation of nitrogen compound metabolic process* were identified in the yellow module as being down-regulated in the older MP > 3 cows. This is in line with the expectation that leukocytes in ageing cows might have an overall downregulation of protein synthesis.

## Immune system

As we were studying a leukocyte population, changes in gene expression relating to immune function were expected. As animals age they are exposed to an increasing variety of disease causing microorganisms, while a progressive loss of function of the immune system increases their vulnerability to infection [9]. Notable age-related changes within the immune cell population include reduced cytokine signalling, diminished production of nitric oxide and peroxide, decreased phagocytic ability and reduced ability of dendritic cells to migrate and process antigens [1]. Our results are in accord with the study by Peters et al [2], who investigated leucocytes from ageing human populations, in finding pathways which were either up- or down-regulated with increasing age. The GO term *immune system process* was enriched in both PP and MP > 3 cows, with genes involved in adaptive immunity up-regulated in the PP cows (*FYN, ITK, LCK*, etc) and genes related to innate immunity up-regulated in MP > 3 cows (*CTSS, CTSH, IRAK2, TLR2*, etc). The black module was down-regulated in the PP cows and contained genes involved in the *regulation of cytokine secretion* (*IL10, CD14, FGR, IL17RC*). Expression of genes in the turquoise module increased in older cows, containing genes involved in *neutrophil degranulation* and *innate immune system*. The darkred and midnightblue modules were both more highly expressed in the PP cows and contained the terms *disease* and *immune system*.

There was also a significant overlap between the DEG with immune functions and candidate age-related genes in the human transcriptome [2]. Those up-regulated in PP cows were mainly associated with T-cell development and function (*CCR7, CD27, IL7R, CAMK4, CD28*) while those up-regulated in MP > 3 cows included genes encoding proteins involved in immune defence. Of these *LYZ* encodes lysozyme, an antimicrobial peptide, *CTSZ* encodes cathepsin Z a lysosomal cysteine protease with multiple roles in host immune defense mechanisms, *SREBF1* encodes a transcription factor involved in TLR4 signalling, while *GRN* encodes granulins, involved in TLR9 signalling, *ANXA5* is involved in T-cell activation and *ADARB1* encodes a deaminase enzyme with A-to-I RNA editing activity, which is important for the maintenance of cellular health but may also play a role in response to viral infection. The results from our study therefore suggest that the older cows are more actively engaged in combatting disease pathogens through activation of the innate immune system and also support a higher level of inflammation with ageing. Parturition is itself an inflammatory process and older cows are also more vulnerable to metabolic disorders after calving. Our previous study showed that metabolic disorders led to prolonged uterine inflammation by up-regulating the genes and pathways related to immune and inflammatory processes [35].

## Growth factor signalling

The relationship between ageing and metabolic regulation is bidirectional, as ageing impairs the activity of key metabolic signalling pathways and the ensuing metabolic dysregulation results in accelerated ageing [3]. Cell signalling pathways that sense the availability of nutrients and the energy status of the cells communicate with hormonal and growth factor signalling pathways to co-ordinately regulate whole body metabolic homeostasis. Ageing results in a gradual deterioration of various cellular functions including metabolic regulation [36]. The turquoise module was the second highest correlated module with ageing, containing genes more highly expressed in MP > 3 cows. The *insulin-like growth factor binding protein* pathway was indeed enriched with genes like *INSR*, *IGF1R*, *IGF1*, *LDLR*, *HTRA1*, *IGFBP7* etc. (Cluster 8; Supplementary Table 3). The latter pathway is interlinked with metabolic pathways to ensure coordinate regulation and fine-tuning of cellular metabolic responses in line with cellular energy status, nutrient availability and hormonal/growth factor signalling input [36].

## Stress and DNA damage

Accumulation of genetic damage represent one of the major contributions to ageing of cells and organisms. Cellular DNA is constantly exposed to exogenous and endogenous DNA-damaging agents like reactive oxygen species, nitric oxide metabolites, and alkylating agents [37] leading to accumulation of mutations in the genome aggravated by loss of capacity in the DNA repair systems [38]. DNA damage is tightly linked to various ageing stresses, such as oxidative stress, telomere shortening, inflammation, irradiation, exposure to chemicals, and mitotic stress [39]. In our dataset, *NELL2* was the most highly correlated gene with age in the turquoise module and was also overall the most differentially expressed gene between the PP and MP > 3 cows, with greater leucocyte expression in the older animals. The encoded protein Neural EGFL Like 2 is highly conserved in mammals and is a glycoprotein containing several von Willebrand factor C and epidermal growth factor (EGF)-like domains. This has a variety of possible roles but amongst these a cell survival-promoting effect mediated by an intracellular mitogen-activated protein kinase (MAPK) pathway has been relatively well studied in neural tissues [40, 41]. *NELL2* is also important in protecting cells from death caused by endoplasmic reticulum (ER) stress resulting in the accumulation of unfolded proteins which trigger the unfolded protein response. Within this context, overexpression of *NELL2* decreased expression of ER stress-induced C/EBP homologous protein (CHOP) and the pro-apoptotic caspases 3 and 7 while increasing expression of ER chaperones and anti-apoptotic Bcl-xL [42].

Another relevant gene to consider is *MTOR*, which encodes Mechanistic Target of Rapomycin, a protein belonging to a family of phosphatidylinositol kinase-related kinases which mediate cellular responses to stresses such as DNA damage and nutrient deprivation. This gene is a central regulator of metabolic homeostasis and is associated with lifespan in many species [43, 44]. *MTOR* is a component of two distinct complexes of which mTORC1 controls protein synthesis, cell growth and proliferation. Genes which are part of the *MTOR* pathway include the transcription factors *FOXO1* and *FOXP1*. A GWAS for longevity, based on a population of Holstein cows, previously identified a region on Bta16 containing *MTOR* [45]. In our study *FOXP1* expression was positively related to increasing age, whereas *FOXO1* expression was negatively related. Also, of potential relevance here is the opposing roles of these two

FOX genes in the regulation of glucose homeostasis, through competition in binding to the insulin response element in gene promoters. Up-regulation of *FOXP1* in mice inhibited the hepatic expression of key gluconeogenic genes, including *PGC-1 $\alpha$* , *PEPCK* and *G6PC* [46]. *LAMTOR1*, -2 and -3 all also featured in the list of DEG which were negatively related to age in our study. These genes encode late endosomal/lysosomal adaptor, MAPK and MTOR activator-1, -2 and -3 respectively, all subunits of the Ragulator complex. This functions as a lysosome anchor, which recruits Rag GTPase and its associated mTORC1 complex to the lysosomal surface prior to MTOR activation [47].

## Regulation of gene expression

The control of gene expression becomes more dysregulated with cellular ageing. A large number of genes and pathways identified in this study are involved in regulation of gene expression. The violet module was the most highly correlated with age and contained a group of genes involved in transcriptional regulation including *ZNF462*, *SOX13*, and *SOX4*, all of which featured in the top 20 genes whose expression was most highly correlated with age (Table 1). Genes in the orange module were more highly expressed in PP cows and this module was enriched with genes involved in *gene expression* (*CHTOP*, *CPSF6*, *THOC1*, *UPF3B*, etc.) and *metabolism of RNA* (*AMDHD1*, *SRSF2*, *SRSF4*, *SRSF5*, etc.). In the yellow module *IGF2BP3* was down-regulated in the older cows and was the most highly correlated gene. IGF2BP family members were initially identified as post-transcriptional regulators of IGF2. They are RNA-binding proteins which direct nuclear RNA export and translation/degradation rates, so playing a major role as regulators of the RNA life cycle. *IGF2BP3* has recently risen to prominence as a potential oncogene [48]. Other genes in this module were also grouped under *regulation of gene expression* (*TGFB3*, *NEUROG2*, *ZNF554*, etc.). The yellow module also contained the gene *WRN*, which exhibited lower transcript abundance in older cows. A mutation in this gene in humans causes Werner Syndrome, an autosomal recessive disorder characterized by the premature development of ageing features. The encoded protein is a member of the RecQ family of proteins and is involved in DNA replication and repair, and telomere maintenance, so playing a crucial role in genome stability [23]. Expression of *WRN* was similarly reduced in leucocytes of older humans [2]. *ADARB1* was another candidate gene from previous studies which was up-regulated in the older MP > 3 cows. It encodes a deaminase enzyme with A-to-I RNA editing activity, was previously identified in a study of men aged 90–119 years, and is also associated with longevity in *C. elegans* [49].

## Metabolism

One key difference between PP and MP > 3 is that milk production potential in dairy cows increases with age [14]. The liver coordinates the extensive metabolic changes required for milk production and these are reflected in circulating metabolite concentrations [19, 25]. Milk synthesis has a high requirement for glucose. In ruminants this demand is met almost exclusively through hepatic gluconeogenesis and cows are at risk of glucose insufficiency during early lactation, the time period we investigated [21]. NEFA are released from lipid stores as an alternative energy source and are either used by the udder to provide milk triglycerides, fully oxidized in the liver to provide energy, or partially oxidised resulting in the production of ketone bodies, in particular BHB. Circulating BHB concentrations are thus an index of fatty acid oxidation

and concentrations are significantly higher in older cows [20]. BHB, NEFA and glucose concentrations can all influence leukocytes directly. Immune cells require an adequate supply of nutrients including glucose to mount an effective immune defense [50]. Neutrophils from cows with more elevated NEFA and BHB concentrations after calving had reduced expression of genes important for granulocyte recruitment, IFN signaling and apoptosis [51]. This suggested that neutrophil survival time was longer in the circulation when exposed to pro-inflammatory conditions. Another study of the circulating leucocyte transcriptome in early lactation found that expression of genes in KEGG pathways relating to DNA replication, cell cycle, homologous recombination, base excision repair, and valine, leucine, and isoleucine biosynthesis were all inhibited as plasma BHB increased, whereas genes involved with vitamin metabolism, the endocrine system, signalling molecules and the immune system were activated [52].

In this study genes in both black and turquoise modules were negatively significantly correlated to the trait age in PP and David functional annotation cluster analysis found enrichment of the terms *fatty acid metabolism* and *glycolysis*. Interestingly, we found that the majority of 27 genes in Cluster 4, with higher expression in older cows, were involved in the *Glycolysis/Gluconeogenesis* pathway (Fig. 3). Although the underlying metabolic background is very different, up-regulation of a small cluster of genes relating to “Fatty acid metabolism, peroxisome activity” was also associated with ageing in the human blood transcriptome [2].

## Other key age-related genes

Within the violet module *LAMA4* was the most highly correlated gene overall, with greater expression in the PP cows. This encodes a subunit of laminin, part of the family of extracellular matrix glycoproteins which form the major non-collagenous constituent of basement membranes. Laminin is thought to mediate the attachment, migration and organization of cells into tissues by interacting with other extracellular matrix components, suggesting that these activities may be reduced in older cows. There is a second MTOR complex mTORC2, which acts as a regulator of the actin cytoskeleton. This is a network of actin and actin binding proteins which are important for a range of essential cellular processes including organelle transport, cell migration, phagocytosis, and cell cycle progression. This is in agreement with the blood transcriptomic study in humans [2] in which expression of a cluster of genes relating to the actin cytoskeleton and focal adhesion also increased with ageing (*ACTA2*, *ACTN4*, *CSRP1*, *ILK*, *LPP*, *TAGLIN*, *TLN1*, *VCL* and *WDR1*). In contrast, *ARHGAP15* was more highly expressed in younger cows (this study) and in humans [2]. This gene encodes Rho GTPase activating protein 15 which is involved in T- and B-cell signalling and promotes an increase in actin stress fibres and cell contraction [53].

## Conclusions

The samples collected for this study provided the opportunity to analyse the transcriptomic profile of blood leukocytes in a large number of early lactating cows and we were thus able to capture the complex and temporally dynamic biological pathways potentially involved in the ageing process in this species. We have, we believe for the first time, identified genes and pathways associated with age in cows and

found that many were comparable to those known to be associated with ageing in humans and model organisms. Immune-related pathways were differentially expressed between primiparous and older cows, including genes involved in innate and adaptive immunity, with many immune defence genes being more highly expressed in older animals. In addition, we found age-associated changes in mitochondrial and metabolic function, ribosome biogenesis, transcriptional regulation and DNA replication, elongation and repair. Pathways supplying leukocytes with energy changed with age, with higher expression of genes encoding enzymes involved in beta-oxidation of fatty acids in the PP cows whereas genes involved in glycolysis were up-regulated in the older cows. These changes may be particularly relevant to understanding how dairy cows respond to metabolic stress during early lactation, when they are short of energy and there is competition for nutrient supply between the demands of milk production and the need for immune defence. An improved understanding of these processes may help dairy farmers to improve both genetic selection and nutritional management to increase longevity, so benefitting agricultural production and animal welfare.

## Methods

### GPlusE study design and sample collection

The samples used in this study were collected as a part of the Genotype plus Environment (GplusE) FP7-Project (<http://www.gpluse.eu>). The animals were located on six experimental dairy farms belonging to members of the consortium (Supplementary Table 1). One 3 ml blood sample from each cow was drawn from the tail vein into a Tempus blood collection tubes (Thermo Fischer, UK) in early lactation, at around 14 days after calving and stored at -20°C. This sample was taken with the informed consent and ethical approval of the organisations involved and complied with the relevant national and EU legislation under the European Union (Protection of Animals used for Scientific Purposes) Regulations 2012 (S.I. No. 543 of 2012). All cows were subsequently released back into the herd. Details of the management of each herd are provided in Krogh et al. [54].

A total of 229 adult Holstein Friesian cows ranging between parities 1 and 7 were recruited. These were subsequently divided into three age groups for analysis: i) primiparous (PP, n = 53), ii) multiparous in lactations 2–3 (MP 2–3, n = 121), and iii) MP in lactations 4–7 (MP > 3, n = 55).

### RNA isolation, library preparation, and sequencing

The individuals performing the sample processing were blinded to the age groups. Total RNA was isolated from whole blood leukocytes using the Tempus Spin RNA isolation Kit (Thermo Fischer, Loughborough, UK) following the manufacturer's instructions. Sample purity and RNA quantity were measured using both a NanoDrop 1000 (Thermo Fischer, UK) and an Agilent BioAnalyzer 2000, using the Agilent RNA 6000 Nano Kit (Agilent, Cheadle, UK). Library preparation was conducted at the University of Liege (GIGA Research Facility, Liege, Belgium), using 0.75 ug of total RNA with the Illumina TruSeq Stranded Total RNA Ribo-Zero Gold Sample Preparation kit (Illumina, San Diego, USA) and sequenced on

the Illumina NextSeq 500 sequencer, producing on average 30 million single end reads of 75 nucleotides length per sample.

## Transcriptomic analyses

Reads with poor quality were trimmed or removed according to base quality using Trimmomatic v.0.36 [55]. The quality of raw and trimmed FASTQ files was assessed with FastQC (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>). The latest version of the *Bos taurus* assembly (ARS-UCD1.2), and its corresponding gene set, was used as reference to map reads using the splice aware aligner HISAT2 [56]. Next, SAM files were converted to BAM files and coordinate sorted with SAMtools [57]. BAM files were processed further with Picard Tools (<http://picard.sourceforge.net/>) in order to remove PCR duplicates, add read group information, sort by chromosome and create indexes. Reads per gene were counted with StringTie [58].

Differential leukocyte gene expression analysis between the three age groups was conducted with the package DESeq2 [59]. Herd effect was considered and removed using *limma* remove batch effect (`limma::removeBatchEffect`; Supplementary Fig. 3). Weighted gene co-expression network analysis (WGCNA, R package; [60]) was used to construct a co-expression network on the DESeq2 normalized data. WGCNA follows a 6-step process to predict which genes are connected to each other, cluster them into gene networks and test which gene networks are associated with phenotypic status, leading to the selection of hub genes.

Genes with variance  $< 0.05$  were filtered out, and the results (total of 13,769 genes) were used as input to the signed WGCNA network construction. We generated a “signed weighted” adjacency matrix for downstream analyses in which the direction of a pair of genes’ correlation is preserved, and a positive correlation may indicate “activation” whereas a negative correlation may indicate “repression”. The adjacency matrix reported a correlation value between every pair of gene expression measurement across all 229 samples. The next step was to raise the adjacency matrix to a software-determined exponential power, thereby reducing noise by pushing weaker pairwise connection values closer to zero relative to stronger values. The exponential power used is the lowest value needed to ensure the network approximates scale-free topology and was set to seven. The adjacency matrix was then transformed into a “topological overlap” matrix by calculating topological overlap (TOM) scores for each gene. This score accounts for each pair of genes’ connection strength (adjacency value) to each other as well as their connection strengths (adjacency values) to every other gene in the adjacency matrix. WGCNA identifies gene co-expression networks via average linkage hierarchical clustering using a TOM-based dissimilarity measure as a measure of distance. The resultant dendrogram of clustered genes is segregated into individual modules with at least 35 genes using WGCNA’s dynamic tree-cutting algorithm. WGCNA calculates each module’s “eigengene” (first principle component), using all samples’ gene expression values for all genes in that module. A module eigengene is considered a summarized expression profile representative of that module for all samples. Finally, each module eigengene was tested for statistical association to the phenotypic trait of age (PP, MP2-3 and MP > 3).

Cattle genes based on the Ensembl v.100 database were retrieved using the Ensembl BioMart online tool [61] and gene ontology (GO) analysis was performed using genes found in significantly correlated modules to trait. Moreover, a David functional annotation cluster analysis [22] was performed; all 13,769 genes were used as background against all significantly correlated modules.

## Abbreviations

MP: multiparous; PP: primiparous; WGCNA: weighted gene co-expression network analysis; NEFA: nonesterified fatty acid; BHB: beta-hydroxybutyrate; DEG: differentially expressed genes; GO: gene ontology; PRR: pattern recognition receptors.

## Declarations

### Ethics approval and consent to participate

The data in this study were collected as a part of the Genotype plus Environment (GplusE) FP7-Project (<http://www.gpluse.eu>). All procedures had local ethical approval and complied with the relevant national and EU legislation under the European Union (Protection of Animals used for Scientific Purposes) Regulations 2012 (S.I. No. 543 of 2012).

***University College Dublin (UCD), Schools of Veterinary Medicine; and Agriculture and Food Science), Belfield, Dublin 4, Ireland.*** The University farm is a licenced premise for the conduct of animal based research under the Cruelty to Animals Act 1876 as amended by European Communities regulations 2002 and 2005 (as required under the European Union Directive 2010/63/EU). Animal experiments were approved by the University College Dublin animal research ethics committee. The licence to perform this experiment was ref. B100/2177.

***The Agri-Food and Biosciences Institute, Hillsborough, Co. Down, UK, BT26 6DR.*** Animal experimentation was performed under the Animals (Scientific Procedures) Act 1986 and covered by a Home Office Project Licence and a Certificate of Designation for the Establishment. The work was also approved by the AFBI Ethical Review Committee.

***Aarhus University (AU), Department of Animal Science, Denmark.*** The work was approved by the Danish Animal Experiments Inspectorate and complied with the Danish and EU Laws concerning animal experimentation and care of experimental animals.

***Consiglio per la Ricerca e la Sperimentazione in Agricoltura, Italy (CRA).*** The cows involved in this project were supervised in compliance with the Italian law and regulations regarding experimental work on animals. The CRA-PCM is authorized to use farm animals for experimental design (as stated in DM 26/96-4 of Italian Welfare Ministry).

***Walloon Agricultural Research Centre, Department for Quality of Agricultural Products (CRA-W).*** All the procedures were approved by the regional ethics committee in animal experimentation (Walloon Region,

Belgium) and were in compliance with European Communities regulations 2002 and 2005 (as required under the European Union Directive 2010/63/EU).

***Leibniz Institute for Farm Animal Biology - FBN Dummerstorf.*** All procedures were approved by the research ethical committee of FBN based on the German and EU legislation (European Union Directive 2010/63/EU) concerning the use of research animals in relation to the welfare and care of experimental animals.

### **Consent for publication**

Not applicable.

### **Availability of data and materials**

The datasets generated and analysed during the current study are available in the GenBank (SRA) repository, and it will be released upon acceptance of the manuscript.

### **Competing interests**

The authors declare that they have no competing interests

### **Funding**

This project received funding from the European Union's Seventh Framework Programme (Brussels, Belgium).

### **Authors' contributions**

LB and DCW designed the analyses; LB, ZC, MZ, and DCW were involved in the interpretation of data; LB and ZC analysed the phenotype and gene expression data; LB wrote the first draft of the manuscript. LB, ZC, MZ, and DCW all edited the final version. Members of the GplusE consortium undertook the animal experiments, collected the samples and performed the RNA sequencing.

### **Acknowledgements**

This project received funding from the European Union's Seventh Framework Programme (Brussels, Belgium) for research, technological development, and demonstration under grant agreement no. 613689. The views expressed in this publication are the sole responsibility of the authors and do not necessarily reflect the views of the European Commission. List of authors in the GplusE Consortium: Niamh McLoughlin, Alan Fahey, Elizabeth Matthews, Andreia Santoro, Colin Byrne, Pauline Rudd, Roisin O'Flaherty, Sinead Hallinan, Claire Wathes, Zhangrui Cheng, Ali Fouladi, Geoff Pollott, Dirk Werling, Beatriz Sanz Bernardo, Mazdak Salavati, Laura Buggiotti, Alistair Wylie, Matt Bell, Mieke Vaneetvelde, Kristof Hermans, Geert Opsomer, Sander Moerman, Jenne De Koster, Hannes Bogaert, Jan Vandepitte, Leila Vandeveld, Bonny Vanranst, Johanna Hoglund, Susanne Dahl, Soren Ostergaard, Janne Rothmann,

Mogens Krogh, Else Meyer, Charlotte Gaillard, Jehan Ettema, Tine Rousing, Federica Signorelli, Francesco Napolitano, Bianca Moioli, Alessandra Crisa, Luca Buttazzoni, Jennifer McClure, Daragh Matthews, Francis Kearney, Andrew Cromie, Matt McClure, Shujun Zhang, Xing Chen, Huanchun Chen, Junlong Zhao, Liguang Yang, Guohua Hua, Chen Tan, Guiqiang Wang, Michel Bonneau, Andrea Pompozzi, Armin Pearn, Arnold Evertson, Linda Kosten, Anders Fogh, Thomas Andersen, Matthew Lucy, Chris Elisk, Gavin Conant, Jerry Taylor, Nicolas Gengler, Michel Georges, Frederic Colinet, Marilou Ramos Pamplona, Hedi Hammami, Catherine Bastin, Haruko Takeda, Aurelie Laine, Anne-Sophie Van Laere, Martin Schulze, Cinzia Marchitelli, Frank Becker and Sergio Palma Vera.

## Authors' information

Department of Pathobiology and Population Sciences, Royal Veterinary College, Hawkshead Lane, North Mymms, Hatfield, Herts, AL9 7TA, United Kingdom.

## References

1. Nikolich-Žugich J. The twilight of immunity: emerging concepts in aging of the immune system. *Nat Immunol.* 2018;19:10–9.
2. Peters MJ, Joehanes R, Pilling LC, Schurmann C, Conneely KN, Powell J, et al. The transcriptional landscape of age in human peripheral blood. *Nat Commun.* 2015;6:1–14.
3. Kenyon CJ. The genetics of ageing. *Nature.* 2010;464:504–12.
4. Tarkhov AE, Alla R, Ayyadevara S, Pyatnitskiy M, Menshikov LI, Shmookler Reis RJ, et al. A universal transcriptomic signature of age reveals the temporal scaling of *Caenorhabditis elegans* aging trajectories. *Sci Rep.* 2019;9:7368.
5. Moskalev AA, Shaposhnikov MV, Zemskaya NV, Koval L, Schegoleva EV, Guvatova ZG, et al. Transcriptome Analysis of Long-lived *Drosophila melanogaster* E(z) Mutants Sheds Light on the Molecular Mechanisms of Longevity. *Sci Rep.* 2019;9. doi:10.1038/s41598-019-45714-x.
6. Shavlakadze T, Morris M, Fang J, Wang SX, Zhu J, Zhou W, et al. Age-Related Gene Expression Signature in Rats Demonstrate Early, Late, and Linear Transcriptional Changes from Multiple Tissues. *Cell Rep.* 2019;28:3263–73.e3.
7. Lai RW, Lu R, Danthi PS, Bravo JI, Goumba A, Sampathkumar NK, et al. Multi-level remodeling of transcriptional landscapes in aging and longevity. *BMB Rep.* 2019;52:86–108.
8. Frenk S, Houseley J. Gene expression hallmarks of cellular ageing. *Biogerontology.* 2018;19:547–66.
9. Kirkwood TBL. Understanding the odd science of aging. *Cell.* 2005;120:437–47.
10. Barzilai N, Gabriely I. Genetic Studies Reveal the Role of the Endocrine and Metabolic Systems in Aging. *J Clin Endocrinol Metab.* 2010;95:4493–500.
11. Pritchard T, Coffey M, Mrode R, Wall E. Understanding the genetics of survival in dairy cows. *J Dairy Sci.* 2013;96:3296–309.

12. De Vries A, Marcondes MI. Review. Overview of factors affecting productive lifespan of dairy cows. *Anim Int J Anim Biosci.* 2020;14:s155–64.
13. Boulton AC, Rushton J, Wathes DC. An empirical analysis of the cost of rearing dairy heifers from birth to first calving and the time taken to repay these costs. *Anim Int J Anim Biosci.* 2017;11:1372–80.
14. Ray DE, Halbach TJ, Armstrong DV. Season and lactation number effects on milk production and reproduction of dairy cattle in Arizona. *J Dairy Sci.* 1992;75:2976–83.
15. Dennis NA, Stachowicz K, Visser B, Hely FS, Berg DK, Friggens NC, et al. Combining genetic and physiological data to identify predictors of lifetime reproductive success and the effect of selection on these predictors on underlying fertility traits. *J Dairy Sci.* 2018;101:3176–92.
16. Grandl F, Furger M, Kreuzer M, Zehetmeier M. Impact of longevity on greenhouse gas emissions and profitability of individual dairy cows analysed with different system boundaries. *Anim Int J Anim Biosci.* 2019;13:198–208.
17. Esslemont RJ, Kossaibati MA. Culling in 50 dairy herds in England. *Vet Rec.* 1997;140:36–9.
18. Bell MJ, Wall E, Russell G, Roberts DJ, Simm G. Risk factors for culling in Holstein-Friesian dairy cows. *Vet Rec.* 2010;167:238–40.
19. Drackley JK, Overton TR, Douglas GN. Adaptations of Glucose and Long-Chain Fatty Acid Metabolism in Liver of Dairy Cows during the Periparturient Period. *J Dairy Sci.* 2001;84:E100–12.
20. Wathes DC, Fenwick M, Cheng Z, Bourne N, Llewellyn S, Morris DG, et al. Influence of negative energy balance on cyclicity and fertility in the high producing dairy cow. *Theriogenology.* 2007;68:232–41.
21. Habel J, Sundrum A. Mismatch of Glucose Allocation between Different Life Functions in the Transition Period of Dairy Cows. *Anim Open Access J MDPI.* 2020;10. doi:10.3390/ani10061028.
22. Huang DW, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc.* 2009;4:44–57.
23. Luo J. WRN protein and Werner syndrome. *North Am J Med Sci.* 2010;3:205–7.
24. Zhang Q, Guldbbrandtsen B, Thomasen JR, Lund MS, Sahana G. Genome-wide association study for longevity with whole-genome sequencing in 3 cattle breeds. *J Dairy Sci.* 2016;99:7289–98.
25. Bell AW, Bauman DE. Adaptations of glucose metabolism during pregnancy and lactation. *J Mammary Gland Biol Neoplasia.* 1997;2:265–78.
26. Franceschi C, Garagnani P, Vitale G, Capri M, Salvioli S. Inflammaging and “Garb-aging. *Trends Endocrinol Metab TEM.* 2017;28:199–212.
27. Akbar AN, Gilroy DW. Aging immunity may exacerbate COVID-19. *Science.* 2020;369:256–7.
28. Landis G, Shen J, Tower J. Gene expression changes in response to aging compared to heat stress, oxidative stress and ionizing radiation in *Drosophila melanogaster*. *Aging.* 2012;4:768–89.
29. Moro-García MA, Alonso-Arias R, López-Larrea C. Molecular mechanisms involved in the aging of the T-cell immune response. *Curr Genomics.* 2012;13:589–602.

30. Campisi J, Kapahi P, Lithgow GJ, Melov S, Newman JC, Verdin E. From discoveries in ageing research to therapeutics for healthy ageing. *Nature*. 2019;571:183–92.
31. Shigenaga MK, Hagen TM, Ames BN. Oxidative damage and mitochondrial decay in aging. *Proc Natl Acad Sci U S A*. 1994;91:10771–8.
32. Reynolds KA, Boudoures AL, Chi MM-Y, Wang Q, Moley KH. Adverse effects of obesity and/or high-fat diet on oocyte quality and metabolism are not reversible with resumption of regular diet in mice. *Reprod Fertil Dev*. 2015;27:716–24.
33. Houtkooper RH, Mouchiroud L, Ryu D, Moullan N, Katsyuba E, Knott G, et al. Mitonuclear protein imbalance as a conserved longevity mechanism. *Nature*. 2013;497:451–7.
34. Bratic A, Larsson N-G. The role of mitochondria in aging. *J Clin Invest*. 2013;123:951–7.
35. Wathes DC, Cheng Z, Chowdhury W, Fenwick MA, Fitzpatrick R, Morris DG, et al. Negative energy balance alters global gene expression and immune responses in the uterus of postpartum dairy cows. *Physiol Genomics*. 2009;39:1–13.
36. Bettedi L, Foukas LC. Growth factor, energy and nutrient sensing signalling pathways in metabolic ageing. *Biogerontology*. 2017;18:913–29.
37. De Bont R, van Larebeke N. Endogenous DNA damage in humans: a review of quantitative data. *Mutagenesis*. 2004;19:169–85.
38. Kenyon J, Gerson SL. The role of DNA damage repair in aging of adult stem cells. *Nucleic Acids Res*. 2007;35:7557–65.
39. Hoeijmakers JHJ. DNA damage, aging, and cancer. *N Engl J Med*. 2009;361:1475–85.
40. Aihara K, Kuroda S, Kanayama N, Matsuyama S, Tanizawa K, Horie M. A neuron-specific EGF family protein, NELL2, promotes survival of neurons through mitogen-activated protein kinases. *Brain Res Mol Brain Res*. 2003;116:86–93.
41. Munemasa Y, Chang C-S, Kwong JMK, Kyung H, Kitaoka Y, Caprioli J, et al. The neuronal EGF-related gene Nell2 interacts with Macf1 and supports survival of retinal ganglion cells after optic nerve injury. *PloS One*. 2012;7:e34810.
42. Kim DY, Kim HR, Kim KK, Park JW, Lee BJ. NELL2 Function in the Protection of Cells against Endoplasmic Reticulum Stress. *Mol Cells*. 2015;38:145–50.
43. Johnson SC, Rabinovitch PS, Kaeberlein M. mTOR is a key modulator of ageing and age-related disease. *Nature*. 2013;493:338–45.
44. Smith HJ, Sharma A, Mair WB. Metabolic Communication and Healthy Aging: Where Should We Focus Our Energy? *Dev Cell*. 2020;54:196–211.
45. Steri R, Moioli B, Catillo G, Galli A, Buttazzoni L. Genome-wide association study for longevity in the Holstein cattle population. *Anim Int J Anim Biosci*. 2019;13:1350–7.
46. Zou Y, Gong N, Cui Y, Wang X, Cui A, Chen Q, et al. Forkhead Box P1 (FOXP1) Transcription Factor Regulates Hepatic Glucose Homeostasis. *J Biol Chem*. 2015;290:30607–15.

47. Mu Z, Wang L, Deng W, Wang J, Wu G. Structural insight into the Ragulator complex which anchors mTORC1 to the lysosomal membrane. *Cell Discov.* 2017;3:1–10.
48. Mancarella C, Scotlandi K. IGF2BP3 From Physiology to Cancer: Novel Discoveries, Unsolved Issues, and Future Perspectives. *Front Cell Dev Biol.* 2019;7:363.
49. Sebastiani P, Montano M, Puca A, Solovieff N, Kojima T, Wang MC, et al. RNA editing genes associated with extreme old age in humans and with lifespan in *C. elegans*. *PLoS One.* 2009;4:e8210.
50. Walls J, Sinclair L, Finlay D. Nutrient sensing, signal transduction and immune responses. *Semin Immunol.* 2016;28:396–407.
51. Crookenden MA, Moyes KM, Kuhn-Sherlock B, Lehnert K, Walker CG, Looor JJ, et al. Transcriptomic analysis of circulating neutrophils in metabolically stressed periparturient grazing dairy cows. *J Dairy Sci.* 2019;102:7408–20.
52. Minuti A, Jahan N, Lopreiato V, Piccioli-Cappelli F, Bomba L, Capomaccio S, et al. Evaluation of circulating leukocyte transcriptome and its relationship with immune function and blood markers in dairy cows during the transition period. *Funct Integr Genomics.* 2020;20:293–305.
53. Seoh ML, Ng CH, Yong J, Lim L, Leung T. ArhGAP15, a novel human RacGAP protein with GTPase binding property. *FEBS Lett.* 2003;539:131–7.
54. Krogh MA, Hostens M, Salavati M, Grelet C, Sorensen MT, Wathes DC, et al. Between- and within-herd variation in blood and milk biomarkers in Holstein cows in early lactation. *Anim Int J Anim Biosci.* 2020;14:1067–75.
55. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinforma Oxf Engl.* 2014;30:2114–20.
56. Kim D, Paggi JM, Park C, Bennett C, Salzberg SL. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat Biotechnol.* 2019;37:907–15.
57. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. *Bioinforma Oxf Engl.* 2009;25:2078–9.
58. Perteau M, Perteau GM, Antonescu CM, Chang T-C, Mendell JT, Salzberg SL. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat Biotechnol.* 2015;33:290–5.
59. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 2014;15:550.
60. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics.* 2008;9:559.
61. Smedley D, Haider S, Ballester B, Holland R, London D, Thorisson G, et al. BioMart—biological queries made easy. *BMC Genom.* 2009;10:22.
62. Chou JP, Ramirez CM, Wu JE, Effros RB. Accelerated aging in HIV/AIDS: novel biomarkers of senescent human CD8 + T cells. *PLoS One.* 2013;8:e64702.
63. Harries LW, Fellows AD, Pilling LC, Hernandez D, Singleton A, Bandinelli S, et al. Advancing age is associated with gene expression changes resembling mTOR inhibition: evidence from two human

populations. *Mech Ageing Dev.* 2012;133:556–62.

64. Passtoors WM, Boer JM, Goeman JJ, Akker EB, van den, Deelen J, Zwaan BJ, et al. Transcriptional profiling of human familial longevity indicates a role for ASF1A and IL7R. *PLoS One.* 2012;7:e27759.

65. Newman AB, Murabito JM. The epidemiology of longevity and exceptional survival. *Epidemiol Rev.* 2013;35:181–97.

66. Yu C-E, Oshima J, Fu Y-H, Wijsman EM, Hisama F, Alisch R, et al. Positional Cloning of the Werner's Syndrome Gene. *Science.* 1996;272:258–62.

67. de Magalhães JP, Curado J, Church GM. Meta-analysis of age-related gene expression profiles identifies common signatures of aging. *Bioinformatics.* 2009;25:875–81.

## Figures

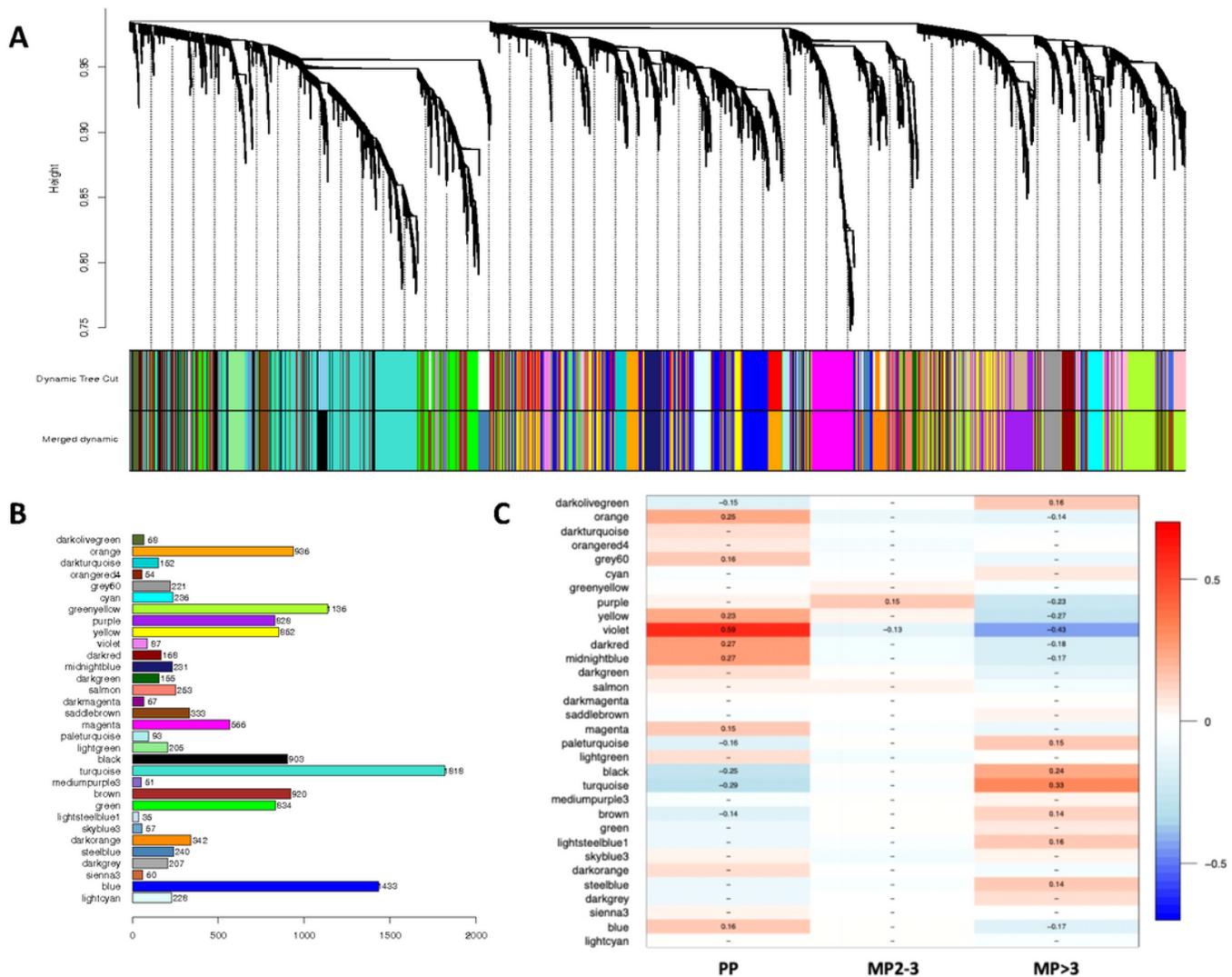
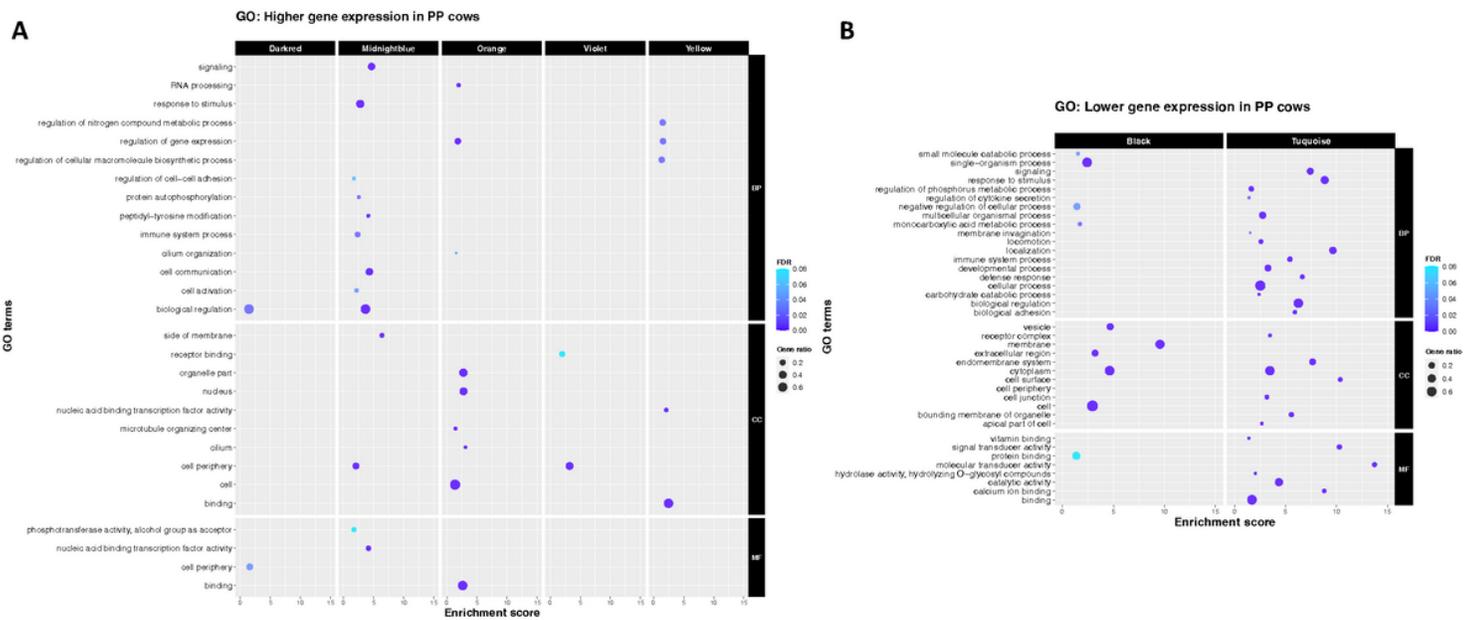


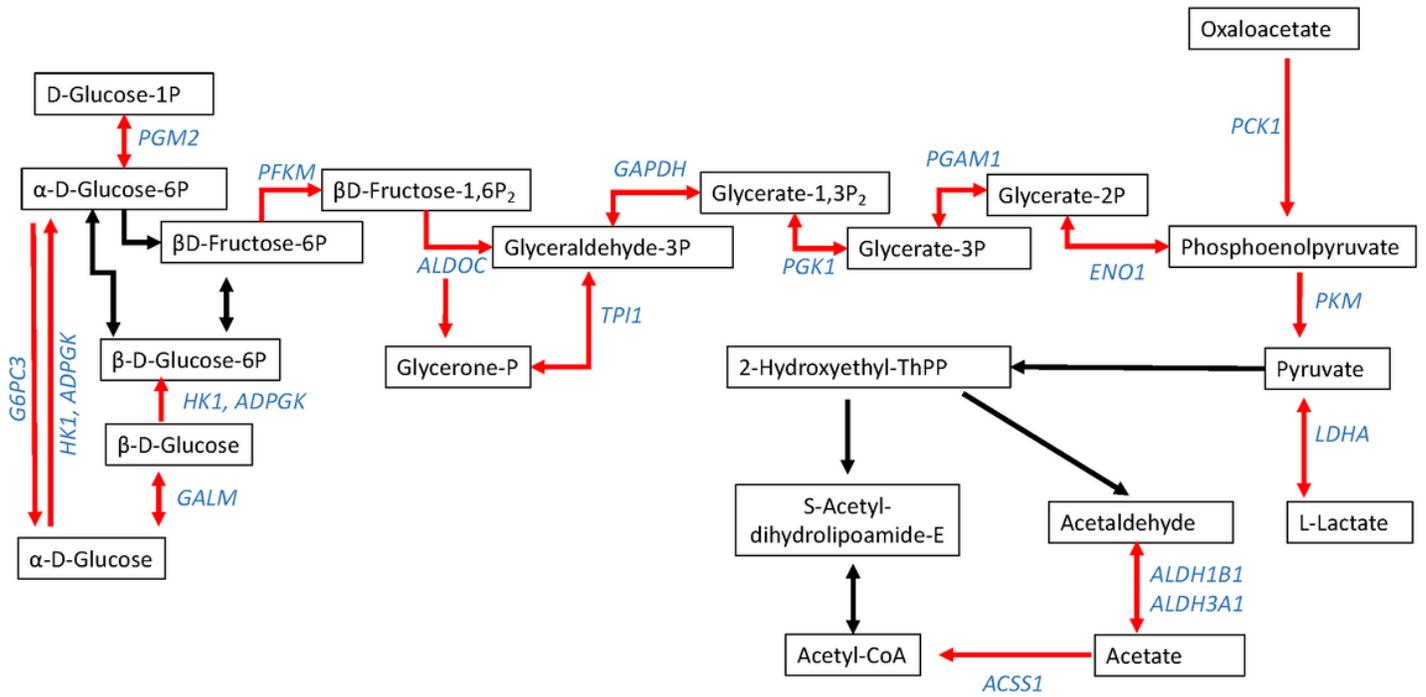
Figure 1

(A) Cluster dendrogram of 13,769 genes (individual black lines at top) clustered by their topological overlap dissimilarity scores. The multi-coloured panel next to “Dynamic Tree Cut” shows 42 identified modules using the Dynamic Tree Cut algorithm. The second multi-coloured panel shows 32 larger modules identified after highly correlated smaller modules were merged together (agreement of 0.8). (B) Number of genes per module identified using WGCNA. (C) Module-trait correlations according to the age of the cow: i) primiparous (PP, n = 53), ii) multiparous in lactations 2-3 (MP 2-3, n = 121), and iii) MP in lactations 4-7 (MP>3, n = 55). Red, positive correlation, blue, negative correlation. Numbers indicate the r value: significance was taken at  $P = \pm 0.25$ .



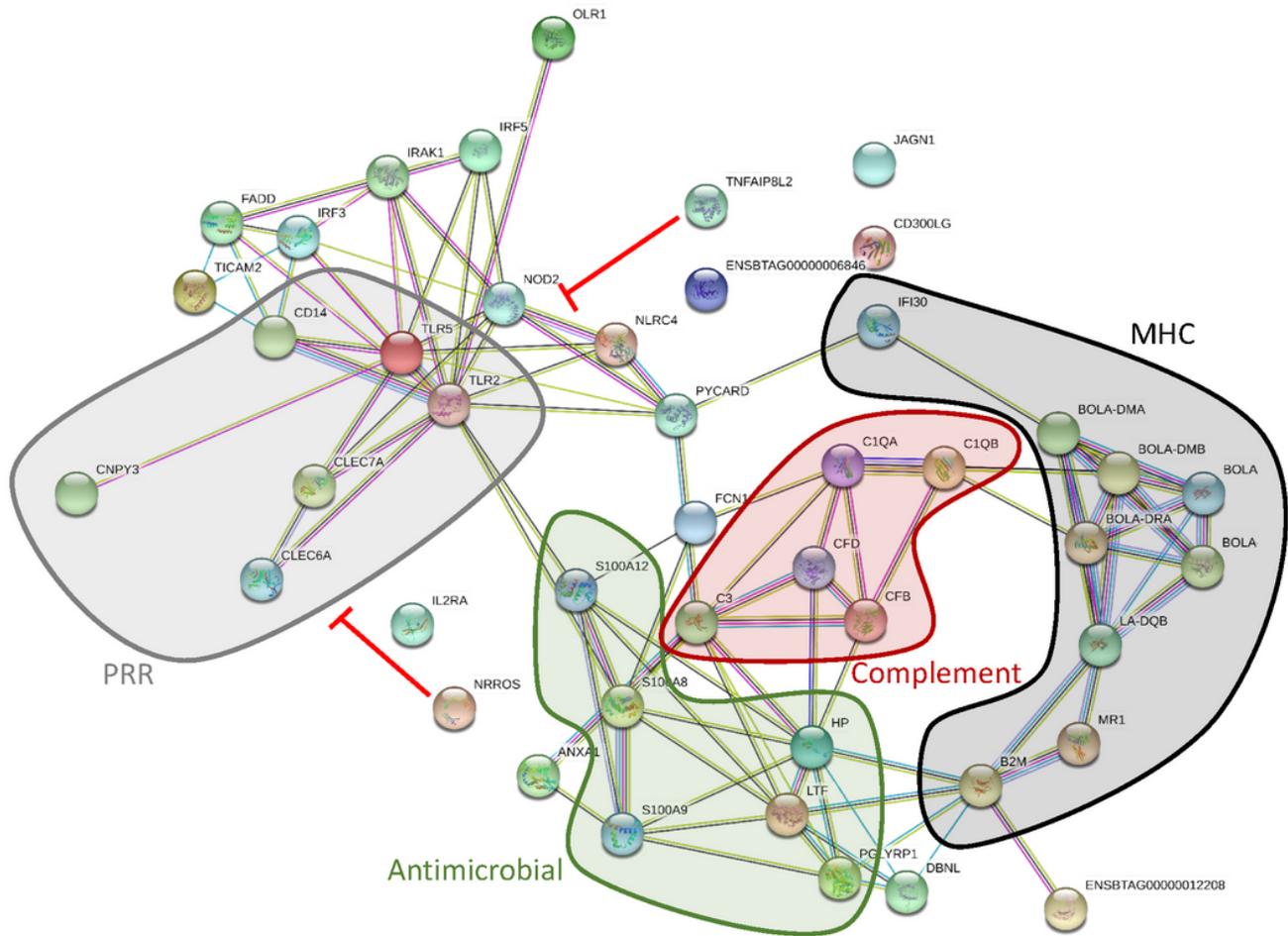
**Figure 2**

GO enrichment analysis of significantly correlated modules ( $FDR \leq 0.1$ ). (A) Positively correlated in PP cows: violet, darkred, midnightblue, orange, yellow; (B) negatively correlated in PP cows: black, turquoise.



**Figure 3**

David functional annotation Cluster 4 contained many genes (shown in blue) in the Glycolysis/Gluconeogenesis KEGG pathway which were more highly expressed in MP>3 vs PP cows. This would potentially maximise the use of glucose as a precursor of acetyl CoA for use in the TCA cycle.



**Figure 4**

David functional annotation Cluster 6 contained genes involved in immunity which were more highly expressed in MP>3 vs PP cows. This included genes encoding proteins for pattern recognition receptors (PRR), the bovine MHC complex, the complement system and several with antimicrobial activity. NRROS and TNFAIP8L2 can both act as negative regulators of TLR signalling mediated via inhibition of NF-kappa-B activation.

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [LBuggiottiAdditionalfile1.xlsx](#)
- [LBuggiottiAdditionalfile2.pdf](#)
- [LBuggiottiARRIVEchecklist.pdf](#)