

OPEN

# RNA-Seq analysis of ileocecal valve and peripheral blood from Holstein cattle infected with *Mycobacterium avium* subsp. *paratuberculosis* revealed dysregulation of the CXCL8/IL8 signaling pathway

Marta Alonso-Hearn<sup>1</sup>, Maria Canive<sup>1</sup>, Cristina Blanco-Vazquez<sup>2</sup>, Rosana Torremocha<sup>3</sup>, Ana Balseiro<sup>2</sup>, Javier Amado<sup>4</sup>, Endika Varela-Martinez<sup>5</sup>, Ricardo Ramos<sup>3</sup>, Begoña M. Jugo<sup>5</sup> & Rosa Casais<sup>2</sup>

Paratuberculosis is chronic granulomatous enteritis of ruminants caused by *Mycobacterium avium* subsp. *paratuberculosis* (MAP). Whole RNA-sequencing (RNA-Seq) is a promising source of novel biomarkers for early MAP infection and disease progression in cattle. Since the blood transcriptome is widely used as a source of biomarkers, we analyzed whether it recapitulates, at least in part, the transcriptome of the ileocecal valve (ICV), the primary site of MAP colonization. Total RNA was prepared from peripheral blood (PB) and ICV samples, and RNA-Seq was used to compare gene expression between animals with focal or diffuse histopathological lesions in gut tissues versus control animals with no detectable signs of infection. Our results demonstrated both shared, and PB and ICV-specific gene expression in response to a natural MAP infection. As expected, the number of differentially expressed (DE) genes was larger in the ICV than in the PB samples. Among the DE genes in the PB and ICV samples, there were some common genes irrespective of the type of lesion including the *C-X-C motif chemokine ligand 8 (CXCL8/IL8)*, *apolipoprotein L (APOLD1)*, and the *interferon inducible protein 27 (IFI27)*. The biological processes (BP) enriched in the PB gene expression profiles from the cows with diffuse lesions included the killing of cells of other organism, defense response, immune response and the regulation of neutrophil chemotaxis. Two of these BP, the defense and immune response, were also enriched in the ICV from the cows with diffuse lesions. Metabolic analysis of the DE genes revealed that the N-glycan biosynthesis, bile secretion, one-carbon pool by folate and purine metabolism were significantly enriched in the ICV from the cows with focal lesions. In the ICV from cows with diffuse lesions; the valine, leucine and isoleucine degradation route, purine metabolism, vitamin digestion and absorption and the cholesterol routes were enriched. Some of the identified DE genes, BP and metabolic pathways will be studied further to develop novel diagnostic tools, vaccines and immunotherapeutics.

<sup>1</sup>NEIKER-Instituto Vasco de Investigación y Desarrollo Agrario, Animal Health Department, Derio, Bizkaia, Spain. <sup>2</sup>SERIDA, Servicio Regional de Investigación y Desarrollo Agroalimentario, Centro de Biotecnología Animal, Deva, Asturias, Spain. <sup>3</sup>Scientific Park of Madrid, Genomic Unit, Campus de Cantoblanco, Madrid, Spain. <sup>4</sup>LSAPA, Animal Health Laboratory of the Principality of Asturias, Department of Microbiology, Gijón, Asturias, Spain. <sup>5</sup>University of the Basque Country (UPV-EHU), Department of Genetics, Physical Anthropology and Animal Physiology, Faculty of Science and Technology, Bilbao, Spain. Correspondence and requests for materials should be addressed to M.A.-H. (email: [malonso@neiker.eus](mailto:malonso@neiker.eus))

Johne's disease or paratuberculosis (PTB) is a chronic granulomatous enteritis of ruminant animals caused by *Mycobacterium avium* subsp. *paratuberculosis* (MAP). Transmission of MAP primarily occurs by the fecal-oral route through the ingestion of MAP contaminated feces, colostrum, or milk. Infection usually occurs within the first months of life of the animal but remains subclinical for an average of 2–5 years. After being ingested, MAP crosses the intestinal mucosa where it is phagocytosed by sub-epithelial macrophages establishing a chronic infection<sup>1</sup>. MAP is able to survive and proliferate within phagosomes by inhibiting apoptosis and phagosomal acidification and by preventing presentation of antigens to the immune system<sup>2</sup>. As the infection progresses, the lesions in the intestine and mesenteric lymph nodes become more severe. Rather than localized, the granulomatous infiltrate becomes diffuse disrupting the mucosal structure and function and affecting jejunum and ileum<sup>3,4</sup>. Although understanding the immunological mechanisms leading to the various outcomes of MAP infection remains problematic, there is no doubt that the interaction of Mycobacteria with the innate immune system plays a central role in the disease pathogenesis<sup>5</sup>.

PTB is responsible for significant economic losses in dairy herds worldwide due to decreased milk production, increased management costs and premature culling from clinical disease<sup>6–9</sup>. Furthermore, there is some evidence suggesting that MAP infection might be associated with Crohn's disease (CD) in humans<sup>10</sup>. Several studies have demonstrated that more than 50% of the dairy cattle herds are positive for MAP antibodies in USA and in Europe and, therefore, bovine PTB can be considered endemic in these areas<sup>11,12</sup>. Commercial inactivated vaccines against bovine PTB are very successful in reducing MAP presence in feces and tissues and in increasing both milk production and cattle productive life in infected farms<sup>13,14</sup>. However, PTB vaccination with heat-killed inactivated vaccines is not allowed in most European countries due to its interference with *Mycobacterium bovis* detection tests<sup>15</sup>. Therefore, PTB control is currently based on testing and culling plus preventing MAP transmission to susceptible animals using appropriate hygienic-sanitary strategies<sup>16,17</sup>. However, the efficiency of the control programs based on the “test and cull” policy is strongly conditioned by the sensitivity of the tests used to identify early infections. Presently, fecal culture is considered “the gold standard” test for the ante-mortem diagnosis of MAP infection<sup>18</sup>. However, individual fecal culture is expensive and time consuming and only detects advanced infections due to the relatively late onset of fecal shedding during the natural course of MAP infection. In fact, the sensitivity of the fecal culture is 70% in animals with PTB-associated clinical signs, but only 23–29% in animals with subclinical infection, which may shed MAP intermittently and in lower numbers in feces and milk contaminating the environment and infecting other animals<sup>19</sup>. Early stage diagnostics such as the *IFN-γ* release assay detects whether a T-cell mediated immune response has been elicited in response to mycobacterial antigens but only reflects MAP exposure, and thus cannot discriminate between individuals with controlled infection from those with subclinical disease<sup>20</sup>. It is clear that the detection of subclinical infections before the bacterium is shed and transmitted to herd mates and possibly to humans remains a challenge and, therefore, novel tools are needed to detect MAP-infected animals at early stages of the infection.

Transcriptomic profiling is a promising source of novel biomarkers for early MAP infection in cattle, for monitoring the progress of the disease, and for the search of immune markers that can be used in the design of therapeutics and vaccines<sup>21</sup>. RNA-Sequencing (RNA-Seq) gives a quantitative and qualitative view of the transcriptomic profile of the tissues or cells being studied. Previous RNA-Seq analysis examined the transcriptome of bovine monocyte-derived macrophages (MDM) infected with MAP and revealed putative biomarkers for early infection<sup>22,23</sup>. To date only one study has used RNA-Seq to examine the transcriptome of ICV samples collected from dairy cows naturally infected with MAP<sup>24</sup>. Some of the subclinical animals included in this study did not show any specific evidence of MAP infection and no pathways controlling the immune response appeared to be affected in this animal group. The authors concluded that perhaps with deeper sequencing in further experiments, potential biomarkers could be ascertained. Since the peripheral blood (PB) is the most available physiological fluid for the detection of biomarkers, in the current study we examined whether it recapitulates, at least in part, the transcriptome of the ileocecal valve (ICV), the primary site of MAP colonization. For this purpose, RNA-Seq was used to identify host genes differentially expressed (DE) in ICV and PB samples collected from PTB-infected animals with focal or diffuse lesions in gut tissues versus control animals. Gene expression data were interpreted using gene ontology (GO) categories and metabolic pathways containing multiple DE genes. Overall, our study provides insights into how cattle respond to a natural MAP infection in two targets, ICV and PB, and revealed common and unique biological processes activated in animals with early and more advanced PTB-associated lesions.

## Results

**Summary of the RNA-Seq data.** The animals included in our study were tested by multiple diagnostic tests including histopathological analysis of gut tissues, acid-fast stain, specific antibody response to MAP, fecal qPCR and tissue and fecal bacteriological culture. The results of the premortem (at the annual sampling time) and post-mortem (at the time of slaughter) diagnostic tests are showed in Table 1. All control animals (n = 3) showed a negative result for all diagnostic tests. Culling of control animals was due to age, accident, and coxofemoral luxation. All the infected cows included in our study had detectable lesions in gut tissues with distinct severity, focal or diffuse. At the slaughter time, the animals with focal lesions in gut tissues (n = 6) showed a negative fecal culture result but were fecal PCR positive (1/6), ELISA positive (1/6), and/or had evidence of MAP infection in gut tissues by ZN (1/6), PCR (5/6) and/or bacteriological culture (2/6). None of the animals with focal lesions had a heavy bacterial load (>50 cfu/gr) in feces and tissues and none of them was culled due to paratuberculosis-associated clinical signs. In contrast, the cows with diffuse histopathological lesions in gut tissues (n = 5) were culled due to a positive ELISA or PCR result (2/5), or because they showed PTB-associated clinical signs, including low milk production (1/5) or profuse diarrhea and weight loss (2/5). At the slaughter time, all of the cows with diffuse lesions had a positive ZN and ELISA result, a PCR-positive result in feces and tissue samples, and had a positive bacteriological culture from gut tissues. The mean amount of MAP estimated by qPCR in the gut tissues of all the

ID	Histopathological analysis			ELISA (OD)		Fecal PCR		Fecal culture (CFUs)		Gut tissues	
	Microscopic	Macroscopic	ZN	Pre-mortem	At slaughter	Pre-mortem	At slaughter	Pre-mortem	At slaughter	PCR	Culture
1	Negative	No	Neg	Neg (1.26)	Neg (1.93)	Neg	Neg	Neg	Neg	Neg	Neg
2	Negative	No	Neg	Neg (2.77)	Neg (8.54)	Neg	Neg	Neg	Neg	Neg	Neg
3	Negative	No	Neg	Neg (8.99)	Neg (2.45)	Neg	Neg	Neg	Neg	Neg	Neg
4	Focal	No	Neg	ND	Neg (2.71)	ND	Neg	ND	Neg	Pos (0.38)	Neg
5	Focal	No	Neg	Neg (1.00)	Neg (1.03)	Neg	Neg	Neg	Neg	Pos (7.50)	Low
6	Focal	No	Neg	Neg (4.39)	Neg (5.51)	Neg	Neg	Neg	Neg	Pos (305.28)	Neg
7	Focal	No	Neg	Neg (4.92)	Neg (0.74)	Pos (0.03)	Neg	Neg	Neg	Pos (12.26)	Neg
8	Focal	No	Neg	Pos (95.85)	Pos (133.17)	Neg	Pos	Neg	Neg	Pos (70.54)	Medium
9	Multifocal	Yes	Pos	Neg (0.00)	Neg (2.97)	Neg	Neg	Neg	Neg	Neg	Neg
10	Diffuse paucibacillary	Yes	Pos	Neg (53.00)	Pos (283.13)	Neg	Pos	Neg	Neg	Pos (667.65)	Heavy
11	Diffuse intermediate	Yes	Pos	Pos (142.31)	Pos (187.69)	Pos (0.14)	Pos (0.14)	Neg	Neg	Pos (3896.19)	Heavy
12	Diffuse intermediate	Yes	Pos	Pos (174.01)	Pos (241.18)	Pos (7.50)	Pos (114.00)	Neg	Heavy	Pos (104032)	Heavy
13	Diffuse intermediate	Yes	Pos	Neg (32.19)	Pos (288.75)	Pos (1.41)	Pos (39.88)	Neg	Neg	Pos (1572.24)	Low
14	Diffuse multibacillary	Yes	Pos	Pos (286.51)	Pos (255.55)	Pos (11.70)	Pos (2832.00)	Neg	Neg	Pos (173316.00)	Heavy

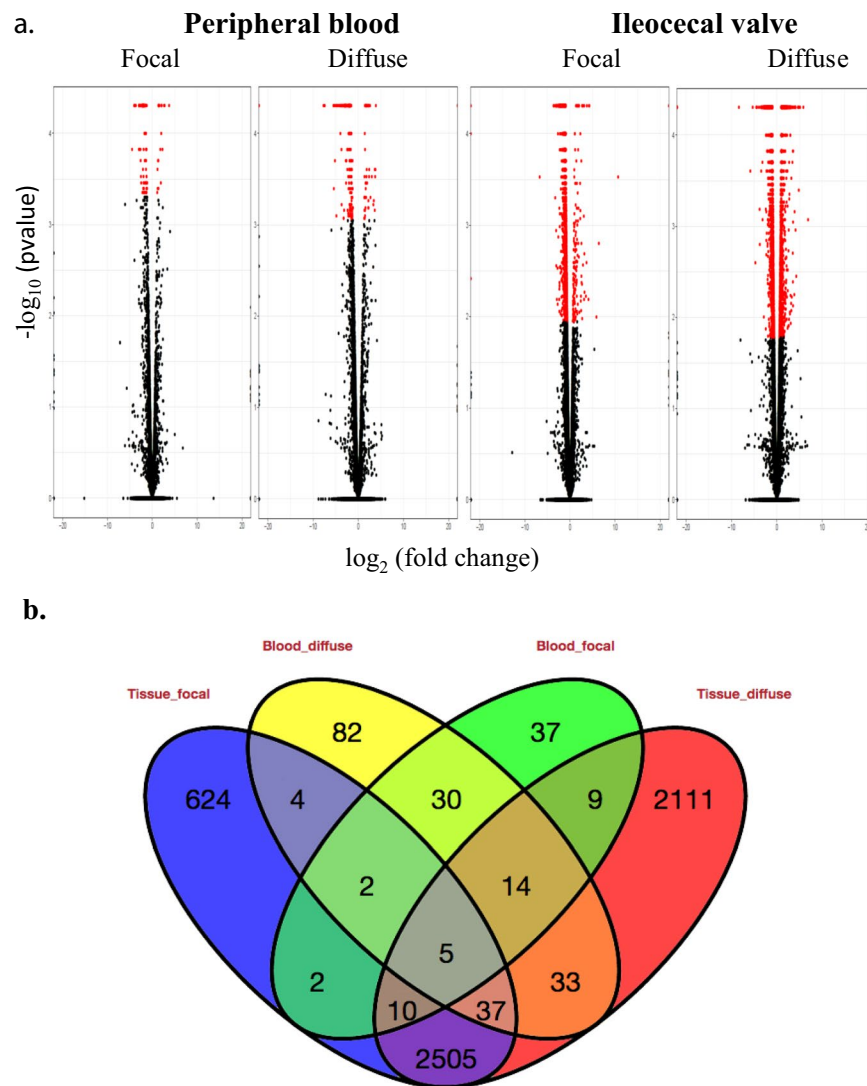
**Table 1.** Histopathological analysis, ZN stain, ELISA, PCR and bacteriological culture results from all the animals included in the current study. Neg, Negative; Pos, Positive; ZN, Ziehl-Neelsen; OD, Optical Density; CFU, Colony Forming Units; ND, No Determined. DNA samples with a PCR-positive result using the LSI VetMax Triplex real-time PCR were quantified using the ParaTB Kuantu-VK kit. qPCR results are expressed as MAP DNA copies per gram of feces or tissues  $\times 10^2$ .

animals with diffuse lesions was  $>66765$  copies of MAP/gr tissue while the mean bacterial load in the gut tissue samples of the animals with focal lesions was  $<30528$  copies of MAP/gr tissue.

RNA-Seq libraries were prepared from PB and ICV of 14 animals within the three categories; with focal or diffuse histopathological lesions in gut tissues and control. Duplicate RNA extractions and corresponding RNA-Seq libraries were generated from PB and ICV samples of five of the 14 cows (ID 2, 6, 8, 11, 12) to check the reproducibility of the RNA-Seq technology. The RNA-Seq data summary for each biological sample and experimental replicate including number of reads per individual RNA-Seq library at each phase of the analysis is provided in Supplementary Table I. All the libraries were sequenced generating an average of de 22.31 million raw reads per library with a Phred quality score  $>30$  which is considered a benchmark for quality in next-generation sequencing. Following filtering of reads based on quality score, minimum size of 60 bp and ambiguous base N percentage less than 5%, an average of 21.44 million reads remained. Alignment of the filtered RNA-Seq reads to the *Bos taurus* reference genome yielded mean values per library of 19.96 million reads. From the mapped reads, an average of 5% of the reads mapped to multiple locations in the genome and were excluded for gene expression analysis. A more detailed analysis of the reads mapping to unique locations was obtained using bedtools and featureCounts, two software packages developed for counting reads to genomic features such as upstream and downstream regions of genes, exons, introns and intergenic regions. Analysis of the individual library reads from the PB samples revealed that 57% of the reads aligned to exons, 28% to introns, 3% to intergenic regions, 2% to upstream regions and 10% to downstream regions. In the ICV samples, introns accounted for 24% of reads, exons for 63% of the reads, 2% of the aligned reads mapped intergenic regions, 1% upstream regions and 10% downstream regions.

**Analysis of differential gene expression from RNA-Seq data.** Although high technical reproducibility has been claimed for the RNA-Seq technology<sup>25,26</sup>, we performed duplicate RNA extractions and corresponding RNA-Seq library preparations from PB and ICV samples of some of the animals (ID 2, 6, 8, 11, 12). Variations in this experiment could arise from the RNA preparation, fragmentation or priming method, ligation efficiencies, amplification and other variables in the technical steps of library construction and amplification. After performing gene expression analysis as describe in materials and methods, no significant differences were observed between each sample and corresponding replicate which support the reproducibility of gene expression studies using RNA-Seq (Supplementary Fig. 1).

The DE genes (FDR  $< 0.05$ ) between cows with focal or diffuse histopathological lesions versus control cows are presented as red dots in Fig. 1a. The number of DE genes was higher in the ICV samples than in the PB samples and increased in animals with diffuse histopathological lesions when compared with the group of animals with focal lesions. This was expected since diffuse lesions represent a more advanced stage of the disease. In the PB samples, 109 and 207 genes were DE in the comparisons between animals with focal or diffuse lesions and control cows, respectively (Fig. 1b). Data analysis revealed 51 DE genes that were common in the PB samples regardless of the type of lesion. The transcriptomic analysis of the ICV samples showed that 3189 and 4724 genes were DE in the animals with focal or diffuse lesions when compared with control animals, respectively. Between these two comparisons, 2557 genes were DE in the ICV samples of the MAP-infected cows regardless of the comparison. A total of 19 DE genes were common among the animals with focal lesions regardless of the tested sample, ICV or PB. Similarly, 89 genes appeared DE in the PB and ICV samples of the animals with diffuse lesions.



**Figure 1.** RNA-seq analysis. **(a)** Volcano plots showing the DE genes ( $\log_2$  fold change) versus  $-\log$  (p value) for each comparison. The red spots represent the DE genes in each comparison. **(b)** Venn diagram showing transcriptional changes between the groups comparisons based on RNA-Seq data.

### Gene expression comparison between animals with PTB-associated histopathological lesions versus control cows.

A total of 109 and 207 were DE in the PB samples of the animals with focal or diffuse lesions when compared with the control animals, respectively. Among these DE genes, 21 and 48 were upregulated in each comparison, respectively (Table 2). Table 3 shows the top five upregulated and top-five downregulated genes in the PB samples of the animals with focal lesions versus control samples ranked by fold-change. The top five upregulated genes included the *ATP binding cassette subfamily A member 13 (ABCA13)*, a receptor for class I MHC antigens named *Leukocyte immunoglobulin-like receptor subfamily A member 6 (LOC618416)* and three still uncharacterized proteins. The top five genes downregulated in the PB samples from the animals with focal lesions versus control cows were: the *C-C motif chemokine ligand 14 (CCL14)*, *Insulin-like growth factor 2 mRNA binding protein 3 (IGF2BP3)*, *Osteoglycin (OGN)*, *Glycogen phosphorylase muscle associated (PYGM)*, and the *CD5 antigen-like precursor (CD5L)*. Many of the top ranked DE genes have immune-related functions such *LOC618416*, *CD5L* and *CCL14*; being the last two genes downregulated. The *CD5L* is a secreted protein expressed by macrophages in lymphoid and inflamed tissues that acts as a key regulator of lipid synthesis. *CCL14* is a chemotactic factor that attracts T-cells and monocytes, but not neutrophils, eosinophils, or B-cells. Fold change in the downregulated genes ranged from  $-4.5$  to  $-1.1$ , with a mean fold change of  $-2.0$ . For the upregulated genes, the fold change of expression ranged between 3.7 to 1.0 with a mean log fold of 1.7.

RNA-Seq analysis detected 207 DE genes in the PB samples of the cows with diffuse lesions compared to the control group. A total of 48 out of the 207 DE genes were upregulated (mean fold change of 2.1, range between 3.8 to 1.3) and 159 downregulated (mean fold change of  $-2.3$ , range between  $-1.1$  to  $-7.7$ ). The top five upregulated genes in the PB samples of the animals with diffuse lesions when compared with control animals were: *Two pore channel 3 (TPC3)*, *Creatine kinase M-type (CKM)*, *Desmin (DES)*, *Myosin light chain 2 (MYL2)* and an



Sample	Comparison	DE genes (N)	Up-regulated genes (N)	Down-regulated genes (N)	Comparison specific DE genes (N)	Up-regulated genes (N)	Down-regulated genes (N)
PB	Focal vs Control	109	21	88	37	11	26
PB	Diffuse vs. Control	207	48	159	82	21	61
PB	Focal vs. Control + Diffuse vs. Control	51	6	45	30	4	26
ICV	Focal vs. Control	3189	168	3021	620	68	552
ICV	Diffuse vs. Control	4724	1473	3251	2105	1321	784
ICV	Focal vs. Control + Diffuse vs. Control	2557	96	2446	2505	93	2397
PB, ICV	Focal vs. Control	19	0	18	2	0	2
PB, ICV	Diffuse vs. Control	89	20	49	33	17	12
PB, ICV	Focal vs. Control + Diffuse vs. Control	5	0	4	5	0	4

**Table 2.** Differentially expressed (DE) genes and specific DE genes for each comparison.

uncharacterized protein (Table 3). The top five downregulated genes in this comparison were: the *Cathelicidins* 3 and 4 (*CATHL3* and *CATHL4*), *CD177* molecule (*CD177*), *C-C motif Chemokine ligand 14* (*CCL14*), and the *Multidrug resistance-associated protein 4-like*. *CATHL3* and *CATHL4* exert a potent antimicrobial activity due to an impairment of the function of the respiratory chain and of energy-dependent activities in the inner membrane of susceptible microorganisms. *CD177* is a neutrophil G-protein-linked surface glycoprotein.

RNA-Seq analysis revealed that 3189 and 4724 genes were dysregulated in the ICV samples from the animals with focal or diffuse lesions versus de control group, respectively (Table 2). The top five upregulated genes in the ICV of cows with focal lesions were the *Intelectin 2 precursor* (*ITLN2*), *Cholecystokinin* (*CCK*), *Homeobox B13* (*HOXB13*), *Fatty acid binding protein 6* (*FABP6*) and *Heat shock protein family A member 6* (*HSPA6*). The upregulated genes mean fold change of expression was 1.5 but, in this case there were larger differences in the expression levels between genes, having the most upregulated gene a fold change of 10.6 and the less upregulated one a fold change of 0.6. *CCK* is a peptide hormone that induces gall bladder contraction and release of pancreatic enzymes in the gut. *HOXB3* is a transcriptional factor previously implicated with inflammation. *FABP6* is involved in the gastric acid and pepsinogen secretion and is also required for efficient apical to basolateral transport of conjugated bile acids in ileal enterocytes. The top five downregulated genes were two uncharacterized proteins, the *Duodenase-1-like* with both trypsin-like and chymotrypsin-like activities, *Tumor suppressor candidate 5* (*TUSC5*), and the *Aquaporin 8* (*AQP8*) (Table 3). Mean fold change of the downregulated genes was  $-1.1$  with values ranging between  $-6.7$  and  $-0.6$ .

In the ICV samples, a total of 4724 DE genes were identified in the comparison between the animals with diffuse lesions versus the control group. From these 4724 DE genes, 1473 genes were upregulated and 3251 were downregulated (Table 2). The level of gene expression was similar between the up- and downregulated genes as demonstrated by the  $\log_2$  fold changes. The upregulated genes on average showed a 1.2 fold increase between the cows with diffuse lesions and the control group (range between 6.8 to 0.5), whereas the downregulated genes averaged a  $-1.2$  fold decrease between groups (range between  $-8.3$  to  $-0.5$ ). The top five upregulated genes included the *Intelectin 2 precursor* (*ITLN2*), *Heat shock protein family A member 6* (*HSPA6*), *Immunoglobulin superfamily member 23* (*IGSF23*), *Apolipoprotein B* (*APOB*) and the *Carbamoyl-phosphate synthase 1* (*CPS1*). On the other hand, the top five downregulated genes were the *Lysozyme C milk isozyme 1* (*LYZ1*), *Adiponectin C1Q* (*ADIPOQ*), *Lysozyme C intestinal isozyme* (*LYSB*), *Acyl-CoA synthetase medium chain family member 1* (*ACSM1*) and the *Aquaporin 8* (*AQP8*). Several of the downregulated genes in this comparison such as the *ITLN2*, *LYZ1*, *LYSB* and *ADIPOQ* play important roles in the defense response against pathogens. Lysozymes such as *LYZ1* and *LYSB* have primarily a bacteriolytic function. *ADIPOQ* is an adiponectin involved in the control of the fat metabolism and it has an anti-inflammatory activity by negatively regulating *TNF- $\alpha$*  expression in macrophages, and also by counteracting its effects.

**Common gene expression signatures in the PB and ICV samples of MAP-infected animals.** DE gene analysis revealed that 51 genes were DE in the PB of the MAP-infected animals regardless of the severity of the lesion (Table 2). Similarly, 2557 genes were DE in the ICV gene expression profiles from the infected cows versus the control group. The PB and ICV gene expression profiles from the cows with focal lesions shared 19 DE genes, and 89 DE genes were common in the PB and ICV expression profiles of the cows with diffuse lesions. In Table 4 we show the top three upregulated and downregulated genes in the PB and/or ICV gene expression profiles from MAP-infected animals. The *Oxidized low density lipoprotein receptor* (*ORL1*) and the *Tweety family protein 2* (*TTYH2*) appeared both upregulated in the PB samples of the MAP-infected cows, regardless of the severity of their lesions in gut tissues. The *CCL14* chemokine, an epithelial-derived chemokine with antibacterial properties, was the most downregulated gene in the PB gene expression profile from MAP-infected cows versus the control group.

In the ICV samples, the *ITLN2* precursor showed the highest transcription increase in the cows with both focal and diffuse lesions;  $\log_2$  fold-change was 10.6 and 6.8, respectively. Immunohistochemistry provided further evidence of the *ITLN2* production and localization in ICV. As seen in Fig. 2, *ITLN2* expression increases upon infection with MAP compared to control samples. Goblet cells and Paneth cells along the ICV from the infected animals were intensively labeled with the bovine *ITLN2* antibody. *FABP6*, a gastrotropin involved in the efficient apical to basolateral transport of conjugated bile acids in ileal enterocytes was highly upregulated in the ICV gene

Sample	Comparison	Gene ID	Description	Fold Change (Log <sub>2</sub> ) <sup>a</sup>
PB	Focal vs. Control	ENSBTAG00000003531	<i>ATP binding cassette A member 13 (ABCA13)</i>	3.7
		ENSBTAG00000039691	Uncharacterized protein	2.7
		ENSBTAG00000039086	Uncharacterized protein	2.6
		ENSBTAG00000000930	Uncharacterized protein	2.2
		ENSBTAG00000019348	<i>leukocyte Ig-like receptor subfamily A member 6</i>	2.0
		ENSBTAG00000022514	<i>CD5 antigen-like precursor (CD5L)</i>	-3.6
		ENSBTAG00000001032	<i>Glycogen phosphorylase, muscle associated (PYGM)</i>	-3.8
		ENSBTAG000000011824	<i>Osteoglycin (OGN)</i>	-3.9
		ENSBTAG000000019406	<i>Insulin growth factor 2 binding protein 3 (IGF2BP3)</i>	-4.0
	ENSBTAG000000010738	<i>C-C motif chemokine ligand 14 (CCL14)</i>	-4.5	
	Diffuse vs. Control	ENSBTAG000000046419	<i>Two pore channel 3 (TPC3)</i>	3.8
		ENSBTAG000000013921	<i>Creatine kinase, M-type (CKM)</i>	3.8
		ENSBTAG00000005353	<i>Desmin (DES)</i>	3.7
		ENSBTAG000000046611	Uncharacterized protein	3.6
		ENSBTAG000000018369	<i>Myosin light chain 2 (MYL2)</i>	3.6
		ENSBTAG000000047764	<i>multidrug resistance-associated protein 4-like</i>	-5.1
		ENSBTAG000000010738	<i>C-C motif chemokine ligand 14 (CCL14)</i>	-5.3
		ENSBTAG00000002976	<i>CD177 molecule (CD177)</i>	-5.4
ENSBTAG000000031599		<i>Cathelicidin-3 (CATHL3)</i>	-7.4	
ENSBTAG000000020072	<i>Cathelicidin-4 (CATHL4)</i>	-7.7		
ICV	Focal vs. Control	ENSBTAG000000022937	<i>Intelectin 2 precursor (ITNL2)</i>	10.6
		ENSBTAG000000013027	<i>Cholecystokinin (CCK)</i>	6.3
		ENSBTAG000000014663	<i>Homeobox B13 (HOXB13)</i>	5.8
		ENSBTAG000000010632	<i>Fatty acid binding protein 6 (FABP6)</i>	4.2
		ENSBTAG000000039035	<i>Heat shock protein 70 family A member 6 (HSPA6)</i>	4.2
		ENSBTAG000000020184	<i>Aquaporin 8 (AQP8)</i>	-3.2
		ENSBTAG000000019379	<i>Tumor suppressor candidate 5 (TUSC5)</i>	-3.2
		ENSBTAG000000046611	Uncharacterized protein	-3.6
		ENSBTAG000000013055	<i>Duodenase-1-like</i>	-3.6
	ENSBTAG000000047008	Uncharacterized protein	-6.7	
	Diffuse vs. Control	ENSBTAG000000022937	<i>Intelectin 2 precursor (ITNL2)</i>	6.8
		ENSBTAG000000039035	<i>Heat shock protein family 70A member 6 (HSPA6)</i>	5.8
		ENSBTAG000000048075	<i>Ig superfamily member 23 (IGSF23)</i>	5.1
		ENSBTAG000000008505	<i>Apolipoprotein B (APOB)</i>	4.8
		ENSBTAG000000016662	<i>Carbamoyl-phosphate synthase 1 (CPS1)</i>	4.8
		ENSBTAG000000020184	<i>Aquaporin 8 (AQP8)</i>	-4.5
		ENSBTAG000000001417	<i>Acyl-CoA synthetase medium chain 1 (ACSM1)</i>	-4.7
		ENSBTAG000000026323	<i>Lysozyme C, intestinal isozyme (LYSB)</i>	-5.3
ENSBTAG000000019813		<i>Adiponectin, C1Q (ADIPOQ)</i>	-5.8	
ENSBTAG000000011941	<i>Lysozyme C, milk isozyme (LYZ1)</i>	-8.3		

**Table 3.** The top five up- and downregulated DE genes in samples from MAP-infected cows versus control samples. <sup>a</sup>A negative fold-change value reports downregulated gene expression.

expression profile from the infected cows irrespective of the severity of the histopathological lesions. The stress response protein, *HSPA6*, was also highly expressed in the ICV from MAP-infected cows with log<sub>2</sub> fold-changes of 4.2 and 5.8 in the gene expression profiles of the cows with focal or diffuse lesions versus control cows, respectively. A significant reduction in the expression of the *AQP8* gene, a protein involved in water/solute intestine homeostasis, was observed in the ICV gene expression profile from the cows with focal or diffuse lesions versus the control group; log<sub>2</sub> fold-change = -3.2 and -4.5, respectively.

Several genes involved in the control of infectious diseases such as the *Basic leucine zipper ATF transcription factor 2 (BATF2)* and the *C-X-C motif chemokine ligand 10 (CXCL10)*, were upregulated in the PB and ICV gene expression profiles from the animal with diffuse lesions. *BATF2* belongs to the activator protein 1 (*AP-1*) transcription factor family and interacts with the *IFN regulatory factor 1* to mediate downstream proinflammatory immune responses. *CXCL10* is chemotactic for monocytes and T-lymphocytes by binding to *CXCR3*. *C4b-binding protein alpha-like, epiregulin (ERG)* and the *CXCL8/IL8* were downregulated in the PB and ICV gene expression profiles from cows with diffuse lesions versus the control group.

Five genes were DE in the PB and ICV gene expression profiles from the animals with PTB-associated histopathological lesions irrespective of the type of lesion including the *CXCL8/IL8*, *Apolipoprotein L domain*

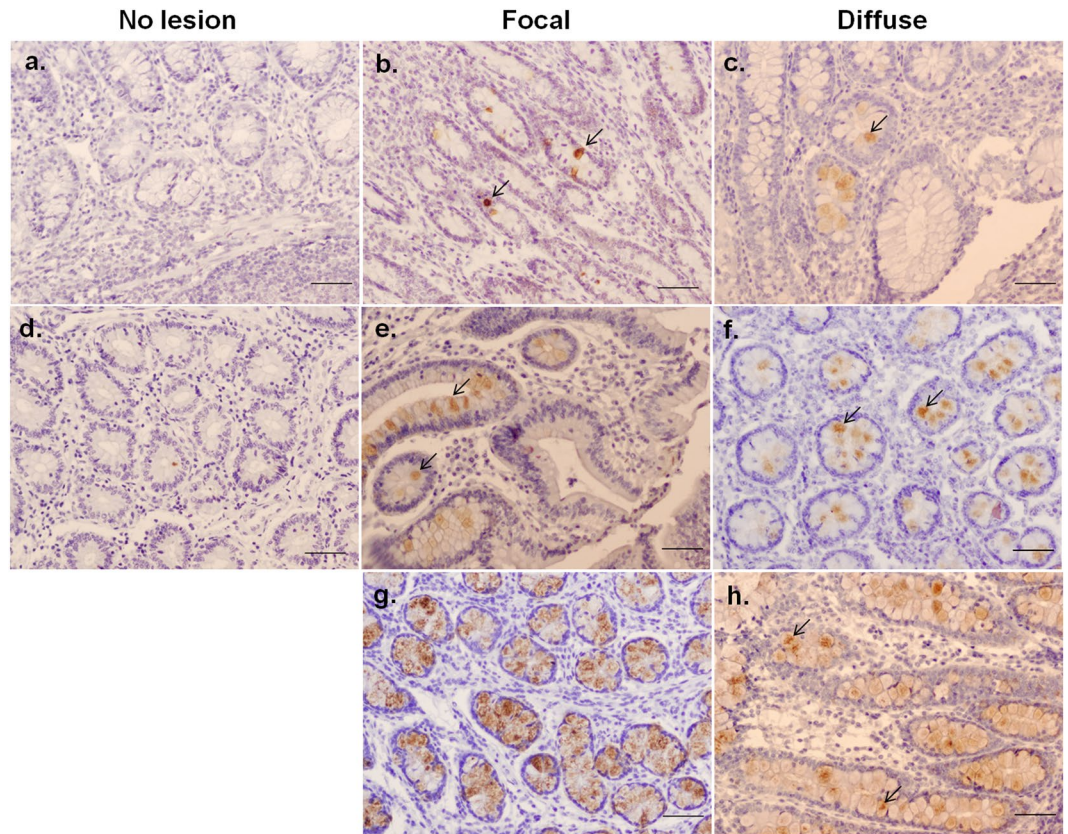
Sample	Comparison	Gen ID	Description (Gene Symbol)	Lesion	Fold Change (Log2) <sup>a</sup>	
PB	Focal vs. Control + Diffuse vs. Control	ENSBTAG00000019348	Uncharacterized protein	Focal	2.0	
				Diffuse	1.6	
		ENSBTAG00000004547	<i>Oxidized low density lipoprotein receptor1 (OLR1)</i>	Focal	1.1	
				Diffuse	2.3	
		ENSBTAG00000011007	<i>Tweety family member 2 (TTYH2)</i>	Focal	1.3	
				Diffuse	1.7	
		ENSBTAG0000001032	<i>Glycogen phosphorylase (PYGM)</i>	Focal	-3.8	
				Diffuse	-2.7	
		ENSBTAG00000019406	<i>Insulin growth factor2 mRNA binding prot3 (IGF2BP3)</i>	Focal	-4.0	
				Diffuse	-2.8	
		ENSBTAG00000010738	<i>C-C motif chemokine ligand 14 (CCL14)</i>	Focal	-4.5	
			Diffuse	-5.3		
ICV	Focal vs. Control + Diffuse vs. Control	ENSBTAG00000022937	<i>Intelectin 2 precursor (ITLN2)</i>	Focal	10.6	
				Diffuse	6.8	
		ENSBTAG00000010632	<i>Fatty acid binding protein 6 (FABP6)</i>	Focal	4.2	
				Diffuse	4.7	
		ENSBTAG00000039035	<i>Heat shock protein family A member 6 (HSPA6)</i>	Focal	4.2	
				Diffuse	5.8	
		ENSBTAG00000047181	<i>Multidrug resistance-associated protein 4</i>	Focal	-3.2	
				Diffuse	-2.3	
		ENSBTAG00000020184	<i>Aquaporin 8 (AQP8)</i>	Focal	-3.2	
			Diffuse	-4.5		
		ENSBTAG00000013055	<i>Duodenase-1-like</i>	Focal	-3.6	
			Diffuse	-2.0		
PB + ICV	Focal vs. Control	ENSBTAG00000036154	<i>TCR gamma alternate reading frame protein (TARP)</i>	Blood	-2.8	
				ICV	-1.1	
		ENSBTAG00000002362	<i>Apolipoprotein L domain containing 1 (APOLD1)</i>	Blood	-2.6	
				ICV	-1.1	
		ENSBTAG00000038080	<i>Duodenase-1</i>	Blood	-2.2	
				ICV	-1.9	
		Diffuse vs. Control	ENSBTAG00000011467	<i>Basic ATF transcription factor2 (BATF2)</i>	Blood	2.6
					ICV	2.8
	ENSBTAG0000001725		<i>C-X-C motif chemokine ligand 10 (CXCL10)</i>	Blood	2.0	
				ICV	2.2	
	ENSBTAG00000046803		<i>Growth arrest specific 1 (GAS1)</i>	Blood	2.4	
				ICV	1.1	
	ENSBTAG00000039196		<i>C4b-binding protein alpha-like</i>	Blood	-4.8	
				ICV	-2.0	
			ENSBTAG00000010273	<i>Epiregulin (EREG)</i>	Blood	-4.8
				ICV	-2.0	
		ENSBTAG00000019716	<i>C-X-C motif chemokine ligand 8 (CXCL8)</i>	Blood	-3.9	
			ICV	-1.9		

**Table 4.** The top 3 up- and downregulated genes in the PB and/or ICV samples of infected animals with focal and/or diffuse lesions versus control samples. <sup>a</sup>A negative fold-change value reports downregulated gene expression.

*containing 1 (APOLD1), Interferon  $\alpha$ -inducible protein 27 (IFI27), KIAA1324 like (KIAA1324L) and ArfGAP with RhoGAP domain, ankyrin repeat and PH domain 2 (ARAP2). CXCL8/IL8, APOLD1, KIAA1324L and ARAP2 were downregulated in all the comparisons. The ARAP2 gene is thought to play a major role in the activation of the major histocompatibility complex class II genes expression<sup>27</sup>. The IFI27 was downregulated in the PB samples and upregulated in the ICV samples from the cows with focal ( $\log_2$  fold = 1.6) and diffuse lesions ( $\log_2$  fold = 2.6) when compared with the control group. The IFI27 (*ISG12*) mediates IFN-induced apoptosis characterized by a rapid and robust release of cytochrome C from the mitochondria and activation of caspases 2,3,6,8 and 9 which may in turn influence the antiviral activities of the IFN<sup>28</sup>.*

**Gene ontology (GO) analysis.** Functional categorization of the DE genes in each comparison was performed using the Bioconductor Goseq package to identify biological process (BP), cellular component (CC) and molecular function (MF) (Table 5). While in the PB samples of the cows with focal lesions we did not identify





**Figure 2.** Localization of the *Intelectin 2* (*ITLN2*) precursor in the ICV of MAP-infected versus control cows by immunohistochemistry. Sections of the ICV from animal without lesions (**a,d**), and with focal (**b,e,g**) or diffuse lesions (**c,f,h**) were labeled with an anti-bovine *ITLN2* antibody and stained as explained in materials and methods. The arrows show sites of antibody binding. Bars are 50  $\mu$ M.

any significantly overexpressed GO, eleven GOs were significantly enriched in the PB samples of the animals with diffuse lesions. Out of these eleven GO; 5 were BP, 4 were CC and 2 were MF (Fig. 3a). The five identified BP were related to the immune response, including killing of cells of other organisms (GO:0031640), defense response (GO:0006952, GO:0050832), immune response (GO:0006955) and positive regulation of neutrophil chemotaxis (GO:0090023). The enriched CC identified in the PB samples of cows with diffuse lesions were the extracellular space and cell membrane.

We identified 83 and 80 overrepresented GO in the ICV samples of the cows with focal and diffuse lesions, respectively. The top 11 GOs identified in the ICV samples of the animals with focal and diffuse lesions were CC. Enriched CC in the ICV samples of the cows with both focal and diffuse lesions included the cytosol and the cytoplasm being these the top dysregulated CC. Among the top five BP enriched in the ICV samples of the cows with focal lesions were the cellular response to DNA damage stimulus (GO:0006974), protein transport (GO:0015031), immune system process (GO:00023746), DNA repair (GO:0006281) and the neutrophil chemotaxis GO (GO:0030593) (Fig. 3b). Three of the top five overexpressed BP in the ICV from the cows with diffuse lesions were associated with the innate immune and inflammatory response (Fig. 3c). Interestingly, the defense response (GO:0006952) and immune response (GO:0006955), were both enriched in the PB and ICV gene expression profiles from the cows with diffuse lesions (Fig. 4). Our results revealed that the chemokines *CXCL8/IL8* and *CXCL10* were regulated in opposite directions in cows with the more advanced lesions. While *CXCL8/IL8* was downregulated in the PB and ICV samples, *CXCL10* was upregulated in both MAP-targeted samples from the cows with diffuse lesions. This differential modulation of *CXCL10* and *CXCL8/IL8* may suggest a different role of both chemokines in the infection. While *CXCL8/IL8* attracts neutrophils, basophils and T-cells to the site of the infection, *CXCL10* is chemotactic for monocytes and T-lymphocytes by binding to CXCR3. Expression changes in genes with antibacterial activity such as the *Cathelicidin 6* (*CATHL6*) and  $\beta$ -defensin (*DEF $\beta$ 4A*) were also observed. Seven DE genes included in the immune response pathway (GO:0006955) were dysregulated in the PB and ICV gene expression profiles from the cows with diffuse lesions including the *CCL14*, *ENPP2*, *CD36*, *CXCL8/IL8*, *BOLA-DQ $\beta$* , *MHC class I heavy chain* and one uncharacterized protein (ENSBTAG00000040323). While *MHC class I heavy chain* was upregulated; *ENPP2*, *CD36*, *CXCL8/IL8*, and *BOLA-DQ $\beta$*  were downregulated in the PB and ICV expression profiles from the cows with diffuse lesions when compared with the control group.

**Metabolic analysis.** Enriched metabolic routes related to the DE genes in each comparison were analyzed. Although no metabolic pathways referenced in the KEGG database were enriched in the PB samples, 78 and



Sample	Comparison	Total enriched GO (N)	Enriched Biological process (N)	Enriched Cellular component (N)	Enriched Molecular function (N)	Specific enriched GO (N)	Enriched Biological process (N)	Enriched Cellular component (N)	Enriched Molecular function (N)
PB	Focal vs. Control	0	0	0	0	0	0	0	0
PB	Diffuse vs. Control	11	5	4	2	7	3	2	2
PB	Focal vs. Control + Diffuse vs. Control	0	0	0	0	0	0	0	0
ICV	Focal vs. Control	83	32	31	19	34	13	12	9
ICV	Diffuse vs. Control	80	38	27	15	30	17	9	4
ICV	Focal vs. Control + Diffuse vs. Control	48	19	18	11	47	19	17	11
PB, ICV	Focal vs. Control	0	0	0	0	0	0	0	0
PB, ICV	Diffuse vs. Control	3	2	1	0	2	2	0	0
PB, ICV	Focal vs. Control + Diffuse vs. Control	0	0	0	0	0	0	0	0

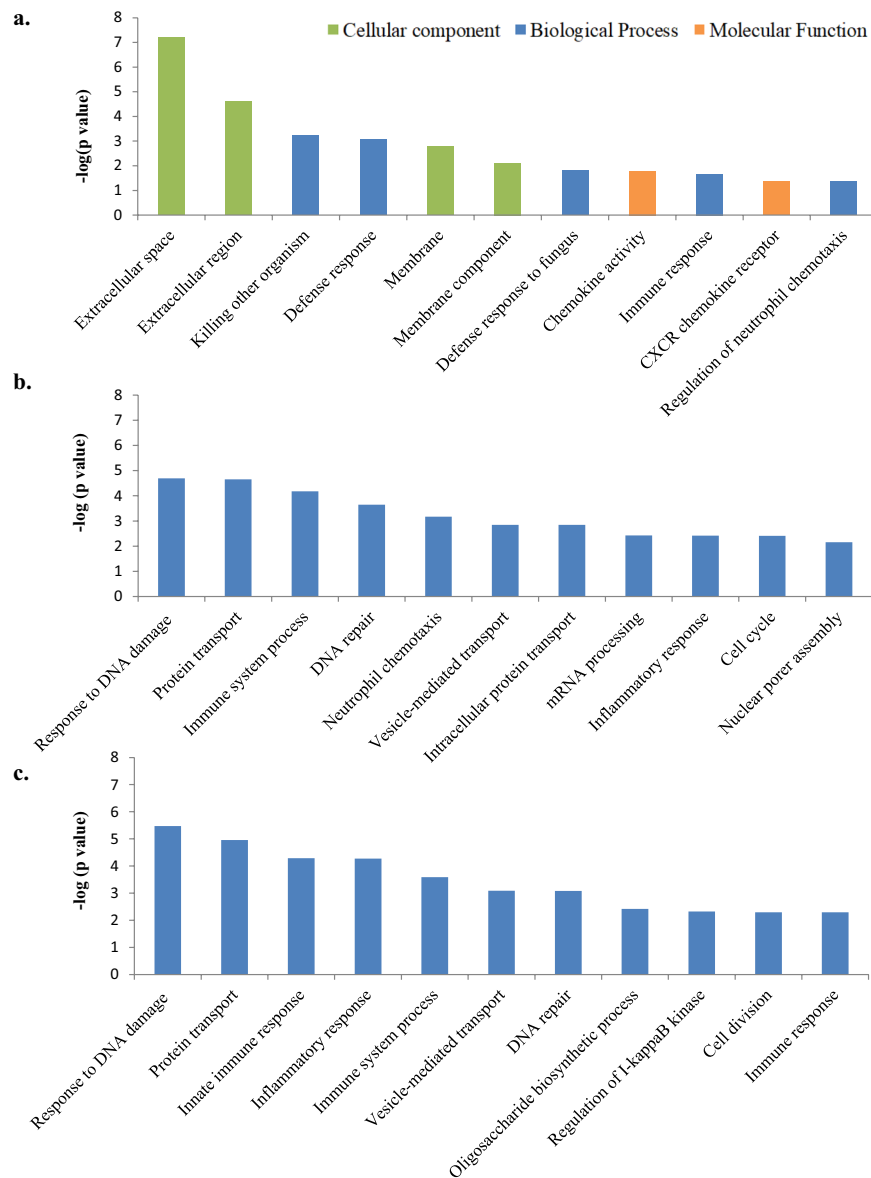
**Table 5.** Number of enriched gene ontologies (GOs) and specific GOs for each comparison.

82 metabolic routes were dysregulated in the ICV samples of the cows with focal and diffuse lesions, respectively. Only three of the 78 metabolic pathways dysregulated in the ICV gene expression profiles from the cows with focal lesions were significantly enriched including the N-Glycan biosynthesis (bta00510,  $P = 0.004$ ), purine metabolism (bta00230,  $P = 0.015$ ) and the one carbon pool by folate pathway (bta00670,  $P = 0.021$ ). The metabolic analysis using the DE genes in the ICV from cows with diffuse lesions versus the control group revealed that the DE genes in this comparison affected metabolic processes such as the valine, leucine and isoleucine degradation (bta00280,  $P = 0.008$ ) and the purine metabolism (bta00230,  $P = 0.003$ ) with sequences corresponding to 10 and 34 enzymes, respectively. Therefore, our metabolic analysis revealed bta00230 as a common metabolic route significantly enriched in the ICV from the infected cows versus the control cows. Using the STRING database, the bile secretion metabolic route (bta04976) was significantly enriched in the ICV from the cows with focal lesions (enrichment score = 244.99,  $FDR = 0.001$ ) with 18 proteins involved in this network. In the ICV from the cows with diffuse lesions, the vitamin digestion and absorption (enrichment score = 599.07,  $FDR = 0.001$ ) and the cholesterol (enrichment score = 677.38,  $FDR = 0.001$ ) routes were significantly enriched with six and four proteins matching each route, respectively.

**Protein to protein interaction analysis.** Protein to protein interaction analysis using the DE genes in the PB samples from the animals with focal lesions revealed a *HBEGF-CXCL8/IL8* functional association (Fig. 5a). Using the DE genes in the PB samples from the cows with diffuse lesions, 11 functional associations and two *CXCL8/IL8* and *Collagen type 1,  $\alpha 2$  chain (COL1A2)* networks were identified (Fig. 5b). The *COL1A2* was linked to a secreted protein, the *Acidic and rich in cysteine protein (SPARC)*, the *Syndecan 4 (SDC4)* and to the *Integrin subunit  $\alpha 9$  (ITGA9)*. When the DE genes in the ICV gene expression profile from the cows with focal lesions were used, 13 functional interactions and one *CXCL8/IL8* centered network were retrieved (Fig. 5c). With a confidence cuff of 0.7, 90 functional interactions and a centered mitogen-activated *MAP kinase 8 (MAPK8)* network were retrieved when the DE genes in the ICV of the animals with diffuse lesions were used (Fig. 5d). Our protein-protein interaction analysis revealed that the interaction between *MAPK8* and *CXCL8/IL8* might be mediated by two leucine zipper protein members of the AP1 transcription factor complex, *FOS* and *FOSB*.

## Discussion

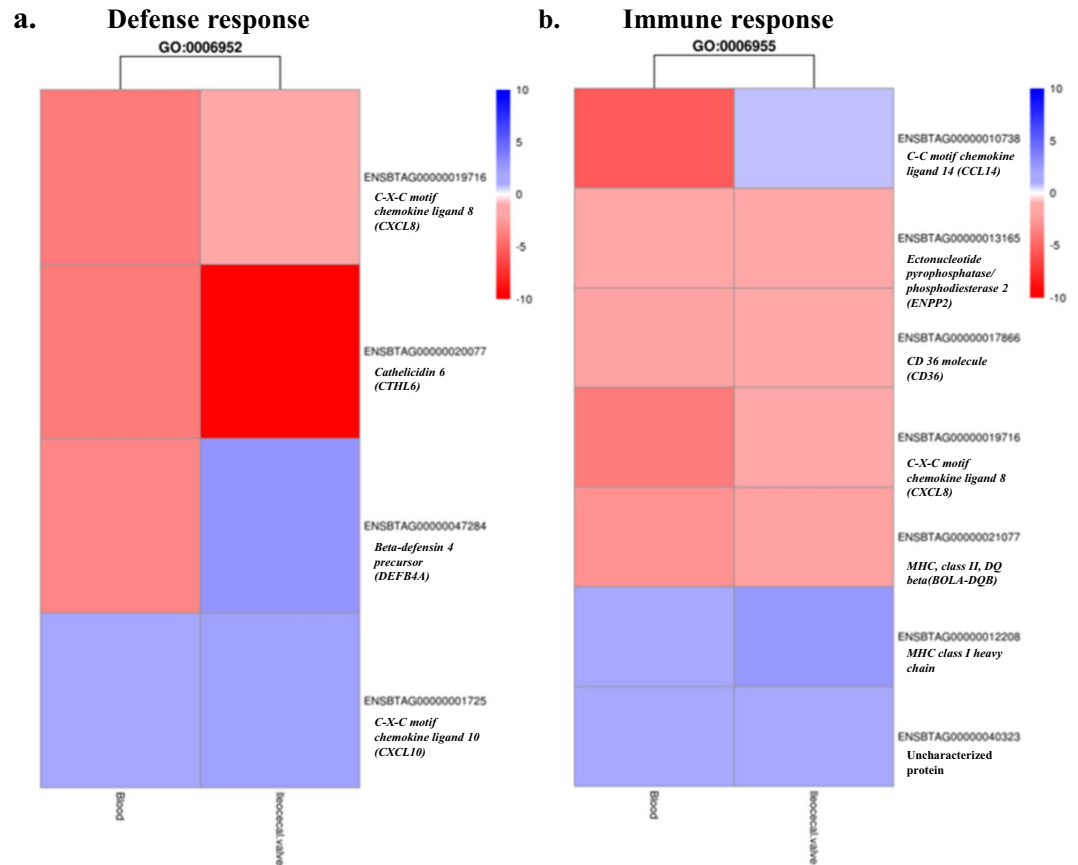
Transcriptomic profiling of the host response to MAP infection has the potential to advance our understanding of MAP-host interactions and to reveal the mechanisms involved in the establishment and progression of bovine PTB. This information can help to set up the basis for novel diagnosis tools and for the development of novel preventive and therapeutic approaches which could improve cattle health. Host transcriptomic studies have highlighted the importance to compare whole blood versus other tissues. To our knowledge, our study represents the first attempt to study the specificities and common signatures of the ICV and PB gene expression profiles of PTB-infected cattle using RNA-Seq. All the infected cows included in our study had detectable lesions in gut tissues with distinct severity, focal or diffuse. Although the number of control animals is limited, our results showed extensive expression differences between infected and control cows. We found that the number of DE genes was higher in the ICV than in the PB samples and increased with the severity of the lesion. More specifically, our study revealed 19 genes DE in the PB and ICV gene expression profiles from the cows with focal lesions versus the control group. Similarly, 89 genes were DE in the PB and ICV gene expression profiles from the cows with diffuse lesions versus the control group. Interestingly, the most upregulated gene in the ICV samples of the infected animals was the *ITLN2* precursor;  $\log_2$  fold change = 10.6 and 6.8 in the cows with focal or diffuse lesions versus the control group, respectively. The *ITLN2* is an extracellular lectin able to bind *Mycobacterium tuberculosis* and *Mycobacterium bovis* bacillus Calmette-Guerin (BCG)<sup>29</sup>. The *FABP6*, an ileum-specific bile acid (BA) transporter, was also highly upregulated in the ICV gene expression profile from the infected cows versus control cows which suggests a dysregulation of the BA secretion in MAP-infected animals. We also observed that MAP infection downregulated the expression of the BA receptor FXR in the ICV from the cows with focal and diffuse lesions versus control cows;  $\log_2$  fold = -0.8 and -1.0, respectively. In agreement with our results, Hempel *et al.*



**Figure 3.** Gene ontology (GO) analysis using the DE genes in the two groups of infected cows compared to the control group. **(a)** The eleven GO enriched in the PB samples collected from the group of animals with diffuse lesions in gut tissues versus the control group. **(b)** The 11 top biological processes overrepresented in the ICV samples collected from the animals with focal lesions versus the control group. **(c)** The 11 top biological processes enriched in the ICV samples from the group of cows with diffuse lesions versus the control group.

found the *FABP6* transporter as the top upregulated gene in the comparison of clinical versus uninfected control animals<sup>24</sup>. Apart from *FABP6*, we found other genes highly dysregulated in the ICV gene expression profile from the cows with diffuse lesions such as the *Immunoglobulin superfamily member 23 (IGSF23)* ( $\log_2$  fold = 5.1), *Apolipoprotein B (APOB)* ( $\log_2$  fold = 4.8), *Solute carrier family 10 member 2 (SLC10A2)* ( $\log_2$  fold = 2.7), and the *Matrix metalloproteinase 13 (MMP13)* ( $\log_2$  fold = 2.0) which had been previously reported by Hempel *et al.* in the comparison of clinical versus control cows. Little is known about the role of *IGSF23*, but other *IGSF* members are cell adhesion molecules that perform important immunological functions, including recognizing a variety of counterpart molecules on the cell surface or extracellular matrix<sup>30</sup>.

Several genes involved in the control of infectious diseases such as the *basic leucine zipper transcription factor 2 (BATF2)* and *CXCL10* chemokine were found in our study highly upregulated in the PB and ICV gene expression profiles from the cows with diffuse lesions versus the control group but were unaffected in the cows with focal lesions. *BATF2* is predominantly expressed in monocytes and macrophages and has a central role in macrophage activation by regulating inflammatory responses during mycobacterial infection. In human tuberculosis (TB), *BATF2* has been suggested as a biomarker and a potential host directed drug target<sup>31,32</sup>. The *CXCL10* chemokine, also known as *IFN- $\gamma$  inducible protein-10 (IP-10)*, is chemotactic for monocytes and T-lymphocytes by binding to CXCR3 receptor. *CXCL10* is expressed mainly by activated CD14+ memory T cells, which produce a Th1 pattern

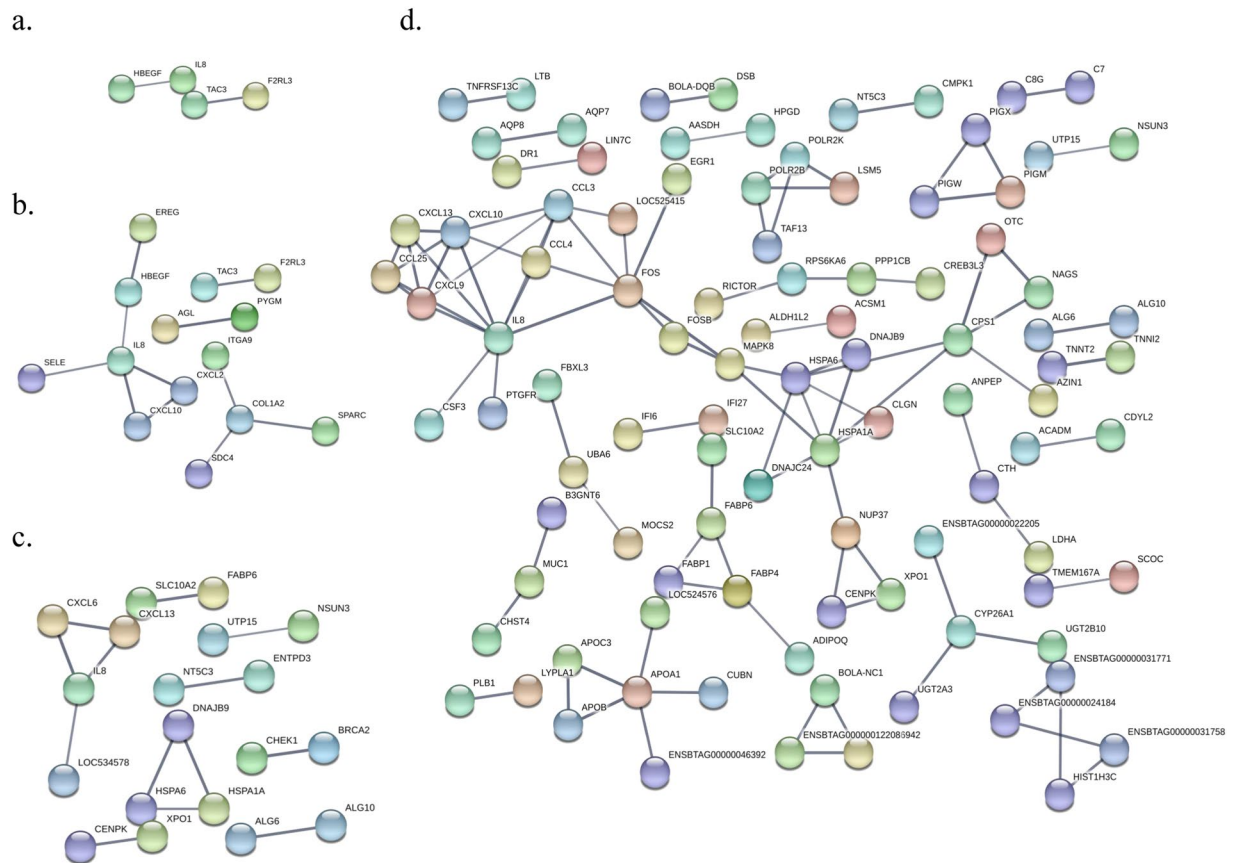


**Figure 4.** Enriched biological process in PB and ICV from the cows with diffuse lesions versus the control group. The defense response (GO:0006952) and immune response (GO:0006955) biological processes were both enriched in the PB and ICV gene expression profiles from the cows with diffuse lesions. The heatmaps represent the fold change ( $\log_2$  fold) of the up- and downregulated genes included in the enriched GO:0006952 and GO:0006955. Colors and intensities depend on expression levels. Red indicated gene downregulation and blue upregulation.

of cytokine production<sup>33</sup>. Several studies have demonstrated that *CXCL10* can be used as a biomarker of TB, where significantly greater levels of mycobacterial antigen-induced *CXCL10* protein were detected in whole blood of TB patients compared to healthy controls<sup>34–36</sup>. More recently, the simultaneous measurement of *CXCL10* and *IFN- $\gamma$*  enhanced test sensitivity for bovine TB identification in cattle in clinical stages of the infection<sup>37</sup>.

Our study revealed enrichment of the following GOs in the ICV from the cows with focal lesions: cellular response to DNA damage stimulus (GO:0006974), protein transport (GO:0015031), immune system process (GO:0002376), DNA repair (GO:0006281) and neutrophil chemotaxis (GO:0030593). Three of the top five GO enriched in the ICV gene expression profile from the cows with diffuse lesions were associated with the innate immune response (GO:0045087), inflammatory response (GO:0006954) and the immune system process (GO:0002376). Interestingly, the GOs related to the defense response (GO:0006952) and immune response (GO:0006955) were enriched in the PB and ICV from the cows with diffuse lesions. The DE genes included in the GO:0006952 were the *CXCL8/IL8*, *CXCL10*, *CATHL6*, and *DEFB4A*. Seven DE genes corresponding to the immune response (GO:0006955) were dysregulated in both the PB and ICV gene expression profiles from the cows with diffuse lesions including *CCL14*, *ENPP2*, *CD36*, *CXCL8/IL8*, *BOLA-DQB*, *MHC class I heavy chain* and the uncharacterized protein ENSBTAG000000040323. *ENPP2*, *CD36*, *CXCL8/IL8*, and *BOLA-DQB* were downregulated in the PB and ICV from the cows with diffuse lesions when compared with the control group. Downregulation of members of the bovine major histocompatibility (*MHC*) class II genes can potentially alter antigen presentation and activation of T helper cells during MAP infection<sup>38</sup>. Our results showed reduced expression of *BOLA-DO $\beta$*  and *BOLA-DQB* genes in the PB from the cows with focal and diffuse lesions, respectively. In the ICV samples, the *BOLA-DR $\alpha$* , *BOLA-DMS*, *BOLA3*, and the *BOLA-DOB* genes were downregulated in cows with focal lesions when compared with the control group. In contrast, in the ICV samples of the cows with diffuse lesions upregulation of the *BOLA-DO $\alpha$* , *BOLA-DMB*, *BOLA-DMA* genes was detected which might suggest the activation of T helper cells as the infection progresses.

Interestingly, our study revealed *CXCL8/IL8* as a common chemokine highly downregulated in our database and involved in the following pathways; neutrophil chemotaxis (GO:0030593), defense response (GO:0006952) and immune response (GO:0006955). Similarly, *CXCL8/IL8* production was reduced in stimulated PB mononuclear cells from CD patients in comparison to healthy subjects and in MDM from *Mycobacterium bovis*-infected



**Figure 5.** Protein-protein network analysis using the DE genes ( $\log_2$  fold  $> 2$  and  $< 2$ ) in the transcriptomic profiles of the cows with PTB-associated lesions versus the control group. **(a)** Using the DE genes in the PB samples from the animals with focal lesions, two *HBEGF-IL8* and *TAC3-F2RL3* functional interactions were detected. **(b)** In the PB samples from the cows with diffuse lesions, two *IL8* and *COL1A2* centered networks were observed. **(c)** When the DE genes in the ICV transcriptomic profiles of the cows with focal lesions were used, a *IL8* centered network was also obtained. **(d)** A *MAPK8* centered network was retrieved when the DE genes in the ICV transcriptomic profile of the cows with diffuse lesions were searched in the STRING database. Individual nodes represent proteins with relationships represented by edges. The confidence score ( $> 0.7$ ) of each interaction is mapped to the edge thickness. The proteins with no associations to other proteins in the networks were hidden.

cattle versus healthy control cattle<sup>39,40</sup>. Since the *CXCL8/IL8* is the major chemokine responsible for neutrophil recruitment and activation by binding to CXCR3, decreased *CXCL8/IL8* expression may reflect impairment of neutrophil recruitment and activation during a natural MAP infection. Other *CXCL8/IL8*-associated biological functions include plasma exudation, general granulocytophilia, degranulation of neutrophils, respiratory burst response, and mobilization of intra-cellular  $Ca^{2+}$ <sup>41</sup>. In our study, the *Epi-regulin (ERG)* and *CXCL8/IL8* genes were both downregulated in the PB and ICV gene expression profiles from the cows with diffuse lesions versus control cows. *ERG* belongs to the *epidermal growth factor (EGF)* family together with the *HBEGF*; acting both as mitogenic stimulators via binding to *EGF* receptors (*EGFRs*). Transactivation of *EGFR* leads to downstream signaling events including *MAPK8* phosphorylation and activation which leads to apoptosis induction<sup>42</sup>. Induced *CXCL8/IL8* release is predominantly dependent on *EGF* binding to the *EFGR* leading to activation of *MAPK8*<sup>43</sup>. Our protein-protein interaction analysis using the DE genes in the cows with diffuse lesions revealed a central role for *CXCL8/IL8* and *MAPK8* in MAP pathogenesis. This kinase was downregulated in the ICV gene expression profiles from the cows with focal and diffuse lesions when compared with the control group;  $-1.7$  and  $-2.16938$ , respectively. Our results support the idea that MAP infection limits signaling via the *MAPK8-IL8* expression pathway which may favor MAP survival. Although the role of *CXCL8/IL8* in MAP pathogenesis has not been characterized before, the *CXCL8/IL8* has been shown to play an important role in the pathogenesis of the human inflammatory bowel disease (IBD)<sup>33</sup>.

Our protein-protein interaction analysis using the DE genes in the PB from the cows with focal and diffuse lesions identified a *CXCL8/IL8* network in both stages of the disease. In other hand, using the DE genes in the PB from the cows with diffuse lesions a *COL1A2* centered network was identified containing a *COL1A2-SPARC* functional interaction. Interestingly, a *SPARC*-centered network was expressed more strongly in *Mycobacterium bovis*-challenged MDM from bovine TB infected cows that in uninfected cows<sup>44</sup>. *SPARC*, also known as



osteonectin, is a matrix protein that binds collagen, and it is required for the development of granuloma-like structures during chronic infections<sup>45</sup>. In our study, *SPARC* and *COL1A2* were upregulated in the PB gene expression profiles from the cows with diffuse lesions which suggests that the expression of both proteins could lead to a bad prognosis in MAP-infected cows.

Recently, metabolomic analysis yielded a clear separation between non-infected and MAP-infected cattle, indicating changes in general metabolism, nutrition uptake and energy balance during the early stage of the disease<sup>46</sup>. Using RNA-Seq, we identified three metabolic pathways significantly enriched in the ICV from the cows with focal lesions including the N-Glycan biosynthesis (map00510,  $P = 0.004$ ), purine metabolism (map00230,  $P = 0.015$ ) and the one carbon pool by folate route (map00670,  $P = 0.021$ ). Altered concentrations of N-glycans, such as mannose, were previously observed by the Buck *et al.* early after MAP infection<sup>46</sup>. In addition, it has been recently proposed that *Mycobacterium tuberculosis* infection manipulates the glycosylation machinery and the N-Glycoproteome of human macrophages<sup>47</sup>. The metabolic analysis performed using the DE genes in the ICV from the cows with diffuse lesions revealed enrichment of the branched amino acids (BCAA; valine, leucine and isoleucine) degradation (map00280,  $p = 0.008$ ) and of the purine metabolism (map00230,  $p = 0.003$ ). The degradation of the BCAA in advanced stages of the infection is consistent with the fact that PTB eventually leads to intestinal malabsorption and hypoproteinemia in the final stages of the infection<sup>4</sup>. Interestingly, our study identified the purine metabolism as a common enriched route in the ICV gene expression profiles of MAP-infected cows regardless of the type of lesion. In agreement with our findings, a previous study showed that MAP infection increased intracellular ATP in bovine monocytes<sup>48</sup>. Using the STRING database, the bile secretion metabolic route (bta04976) was significantly enriched in the ICV gene expression profiles from the cows with focal lesion (enrichment score = 244.99, FDR = 0.001). In the ICV gene expression profile from the cows with diffuse lesions, the cholesterol and the vitamin digestion and absorption routes were significantly enriched, respectively. In agreement with these findings, recent reports have demonstrated that MAP is able to manipulate the host lipid metabolism and accumulate cholesterol within macrophages which might favor MAP persistence<sup>49</sup>. In addition, decreases in serum 25-hydroxyvitamin D<sub>3</sub> (25OHD<sub>3</sub>) levels were significantly lower in cows in the clinical stage of disease compared with either cows in the subclinical stage and non-infected control cows<sup>50</sup>. It is generally accepted that cattle in advanced stages of MAP infection are unable to metabolize and absorb vitamin D<sub>3</sub> across the intestinal wall, and therefore suffer from weight loss and cachexia.

This study is the first to simultaneously describe transcriptomic changes in PB and ICV samples of cattle naturally infected with MAP. Several important observations were made. First, by comparing the transcriptomic profiles of two MAP-targeted tissues we described both unique and overlapping changes in the transcriptome of the infected cows versus the control group. Second, our study highlighted the ability of the RNA-Seq technology to reveal roles for genes that have not been previously implicated in the host response to MAP infection. Third, our transcriptomic analysis provided potential biomarkers for the development of future diagnostic tools, vaccines and therapeutics.

## Materials and Methods

**Ethic statement.** Experimental procedures performed on the animals used in this study were approved by the Animal Ethics Committee of the Servicio Regional de Investigación y Desarrollo Agroalimentario (SERIDA) and authorized by the Regional Consejería de Agroganadería y Recursos Autoctonos of the Principality of Asturias (approval code PROAE 29/2015). All the procedures were carried out in accordance with the European Guidelines for the Care and Use of Animals for Research Purposes (2012/63/EU). PB and fecal samples were collected by trained personnel and in accordance with good veterinary practice.

**Animal population.** Fourteen Holstein Friesian cows from a single commercial dairy farm in Asturias (Spain) were monitored annually by obtaining PB and fecal samples to assess disease status by ELISA, and fecal bacteriological culture and PCR. The mean prevalence of the disease in the farm estimated by ELISA was 6.29% in the sampling period (2016–2018).

### Tissue and fecal sampling for histopathology, immunohistochemistry and bacterial culture.

For histopathological analysis, samples from ileocecal lymph nodes, distal jejunal lymph node, ICV, and distal jejunum were collected aseptically from each animal and placed in formalin. The collected samples were fixed in 10% neutral buffered formalin, and dehydrated through graded alcohols and xilol before being embedded in paraffin wax. Several sections were cut from each tissue sample using a microtome Leica RM2035 (Leica Microsystems, Barcelona, Spain), mounted on treated microscope slides (Fisher Scientific Co) and subsequently stained with hematoxylin-eosin (HE) and Ziehl-Neelsen (ZN). The stained sections were examined under a microscope Olympus BX51 equipped with an Olympus U-CMAD3 digital camera for pathological lesions and for the presence of acid-fast bacteria (AFB). According to their location and extension, inflammatory cell type, and Mycobacterial load, PTB-associated histopathological lesions were classified as focal (focal and multifocal), and diffuse (paucibacillary, intermediate or multibacillary)<sup>3</sup>.

For immunohistochemistry, the Avidin-Biotin kit with peroxidase-based detection was used (Vector Laboratories, California, USA). Briefly, ICV sections were deparaffinised, hydrated and rinsed with tap water. Afterwards, slides were treated to quench the endogenous peroxidase by incubation with methanol containing 3% H<sub>2</sub>O<sub>2</sub> for 10 min at room temperature and washed with water for 10 min. The tissue sections were incubated overnight at 4 °C with an anti-*ITLN2* polyclonal antibody (Aviva Systems Biology, San Diego, California), washed with TBS and incubated with a goat biotinylated anti-rabbit IgG (1:200 dilution). The sections were then incubated with VECTASTAIN Elite ABC Reagent for 30 min at room temperature. Finally, the sections were incubated with 3,3'-diaminobenzidine tetrahydrochloride (DAB, Sigma, St. Louis, MO, USA) for 5 min and washed with TBS. The stained slides were dehydrated, mounted and studied under light microscopy.

For bacteriological culture, a pool (2 gr) of ileocecal lymph nodes, distal jejunal lymph node, ICV, and distal jejunum were decontaminated with 38 mL of hexa-decyl pyridinium chloride at a final concentration of 0.75% (Sigma, St. Louis, MO) and homogenized in a stomacher blender. After 30 min of incubation at room temperature, 15 mL of the suspension was transferred to a new tube and incubated overnight for decontamination and sedimentation. Approximately, 200 µl of the suspension was taken from the layer near the sediment and inoculated into two slants of Herrolds egg yolk medium (HEYM; Becton Dickinson, Sparks, MD) and into two slants of Lowenstein-Jensen medium (LJ; Difco, Detroit, MI), both supplemented with Mycobactin J (Allied Monitor, Fayette, MO) as previously described<sup>51</sup>. Bacterial load in tissues was classified as low (<10 cfu; estimated average 2 cfu/tube), medium (between 10 to 50 cfu, estimated average 20 cfu/tube), or heavy (>50 cfu; estimated average 200 cfu/tube). At the time of slaughter, feces were taken from the rectum of each animal and processed within 48 h after arrival at the laboratory. The fecal samples (2 g each) were decontaminated, blended in a stomacher, and cultured in HEYM and LJ, as previously described for tissue culture.

**Enzyme-linked immunosorbent assay (ELISA) for the detection of MAP-specific antibodies.** Blood was collected from the coccygeal vein of each animal into 4.5 ml serum clot activator Vacutainer® tubes (Vacuette, Kremsmunster, Austria). Serum was separated after clotting by centrifugation (2500 × g for 20 min) and stored at −20 °C until use. The serum samples were tested using the Mycobacterium Paratuberculosis Antibody test kit (IDEXX laboratories, Oofddorp, The Netherlands) according to the manufacturer's instructions. The optical density (OD) in each well was measured at 450 nm by an ELISA plate reader (model 680, Sigma, St. Louis, MO). The measured ODs were normalized and the results were expressed as a percentage of the positive control OD according to the following formula: % relative OD sample/OD positive control = 100 × [(OD sample<sub>+Ag</sub> − OD sample<sub>−Ag</sub>)/(OD mean positive control<sub>+Ag</sub> − OD mean positive control<sub>−Ag</sub>).

**Fecal real-time polymerase chain reaction (PCR).** Isolation of genomic DNA from feces was performed using the MagMax Total Nucleic Acid Isolation kit according to the manufacturer's instructions (ThermoFisher Scientific, Lissieu, France). For detection of MAP DNA, the LSI VetMax Triplex real-time PCR was used according to the manufacturer's instructions (ThermoFisher Scientific, Lissieu, France). The kit enables real-time PCR detection of *Map* IS900 and F57 genes in DNA extracted from feces, liquid cultures, and tissues or colonies. Real-time PCR amplifications were performed using the MX3000P Real-Time PCR system (Stratagene, San Diego, USA) detection system with the following conditions: 1 cycle at 50 °C for 2 min, 1 cycle of 95 °C for 10 min, 45 cycles of denaturation at 95 °C for 15 s, and annealing/extension at 60 °C for 60 s.

**Fecal quantitative real-time PCR (qPCR).** PCR-positive samples were also tested by real-time qPCR using the ParaTB Kuantif-VK kit following the manufacturer's instructions (Vacunek, Bizkaia, Spain). The kit uses a F57 TagMan probe labeled with the fluorescent reporter dye 5-carboxyfluorescein (FAM) at the 5' end and primers that specifically amplify the single-copy F57 insertion sequence of MAP. Inhibition of the amplification reaction is ruled out by including in the master mix an internal plasmid control with specific primers and an internal hybridization probe labeled with 6-carboxy-4',5'-dichloro-2',7'-dimethoxyfluorescein succinimidyl ester (JOE) at the 5' end. This internal amplification control molecule is co-amplified alongside the F57 diagnostic target in a duplex format. Quantification of MAP titer (F57 copy numbers per gram of feces) was accomplished by preparing a standard curve using serial dilutions of a standard sample containing a known number of MAP DNA copies. Real-time qPCR amplifications were performed using the Step One Plus detection system (Applied Biosystems, Carlsbad, CA) with the following conditions: 1 cycle at 95 °C for 10 min, 45 cycles of denaturation at 95 °C for 15 s, and annealing/extensions at 60 °C for 60 s. The results were analyzed with the ABI Prism software version 1.4.

**RNA extraction and RNA-Seq library preparation.** At the time of slaughter, PB samples were collected from the coccygeal vein of all the cows included in the study in PAXgene Blood RNA tubes (2.5 ml) (Qiagen, Hilden, Germany). Total RNA was extracted from the PB samples using the PAXgene blood RNA kit according to the manufacturer's instructions (Qiagen, Hilden, Germany). For RNA isolation, 150–200 mg of ICV of all the animals were harvested and immediately submerged in 2 ml of RNA later (Sigma, St. Louis, MO). Purification of RNA was performed using the RNeasy Mini Kit according to the manufacturer's instructions (Qiagen, Hilden, Germany). Residual genomic DNA was removed using DNase digestion with RNase-free DNase I Amplification grade following the recommended protocol (Invitrogen, Spain). Concentration and quality of the total RNAs were measured using an Agilent Bioanalyzer 2100 (Agilent Technologies, Santa Clara, CA, US). All samples had an RNA integrity value of 7 or greater. Approximately 250 ng of RNA were used for each RNA-Seq library creation using the Illumina NEBNext® Ultra Directional RNA Library preparation kit following the manufacturer's instructions (Illumina Inc, CA, US). All RNA-Seq libraries were quantified using a Qubit® Fluorometer and double stranded DNA high Sensitivity Assay Kit (Invitrogen, Spain). RNA-Seq libraries quality was assessed using an Agilent Bioanalyzer and Agilent high sensitivity DNA chip to confirm that the insert sizes were 200–250 bp for all the individual libraries.

**RNA sequencing and bioinformatic analysis of RNA-Seq data.** RNA-Seq libraries were single-end sequenced in a 1 × 75 format using an Illumina NextSeq500 sequencer at the Genomic Unit of the Scientific Park of Madrid, Spain. The raw reads were filtered by their length (minimum size 60 bp long) and percentage of ambiguous base N less than 5% using Prinseq-lite<sup>52</sup>. Trimmed reads were subsequently mapped to the *Bos Taurus* reference genome (*Bos\_taurus*.UMD3.1. version 87) with TopHat mapper<sup>53</sup>. Reads were assigned to a gene if they were not multi-hit reads. The resulting alignment files were provided to Cufflinks to generate a transcriptome assembly for each condition. These assemblies were then merged together using Cuffmerge, which is included in

the Cufflinks package. This merged assembly provided a uniform basis for calculating gene and transcript expression levels in each condition. The reads and merged assembly were fed to Cuffdiff which calculates expression levels and tested the statistical significance of each observed change in expression<sup>54</sup>. The fold change (log<sub>2</sub> scale), P-values and false discovery rates (FDR) for each gene were obtained. The genes with a FDR-adjusted threshold <0.05 were considered differentially expressed (DE) when compared to the control group. To visualize and integrate all the data produced by the Cufflink analysis, CummeRbund was used for cluster analysis and for handling the transformation of Cuffdiff data into R statistical computing environment. RNA-Seq data have been deposited in the NCBI Gene Expression Omnibus (GEO) database under the accession number (GSE137395).

**Gene ontology (GO) and metabolic analysis of the DE genes obtained using RNA-Seq.** To identify GO and metabolic routes involved in MAP infection, genes with a significant DE in each comparison were evaluated using the Goseq R Bioconductor<sup>55–57</sup>. To begin the analysis, Goseq quantifies the length bias present in the DE genes by calculating a probability weighting function (PWF) which gives the probability that a gene will be DE based on its length alone. Next, Goseq analysis requires the use of the Wallenius non-central hypergeometric distribution to approximate the null distribution for GO category membership and to calculate representation of GO categories amongst the set of DE genes. Using the Wallenius approximation, P-values for representation of the DE genes in each GO were generated and a P-value threshold <0.05 selected. GO analysis provided categories of genes involved in different biological processes (BP), molecular functions (MF) and those integral for different cell compartments (CC). KEGG pathways enrichment analysis was also performed using the STRING v11.0<sup>58</sup>.

**Protein-protein association networks.** The DE genes with log<sub>2</sub> fold changes >2 or <2 were analyzed using the STRING database v10.5<sup>59</sup>. The STRING database aims to collect, score and integrate all known and predicted protein-protein association data. The basic interaction unit in the STRING database is the functional association, i.e. a link between two proteins that both contribute jointly to a specific function. Network nodes represent proteins and edges protein-protein interactions. For each interaction a combined and final score is computed based on seven evidence channels including neighborhood in the genome, gene fusions, co-occurrence across genomes, co-expression, experimental/biochemical data, association in curated databases, and co-mentioned in PubMed abstracts. Only functional interactions with a high confidence score (>0.7) were included in the networks.

## Data Availability

The datasets generated during the current study are available from the corresponding author on reasonable request.

## References

- Bermudez, L. E., Petrofsky, M., Sommer, S. & Barletta, R. G. Peyer's patch-deficient mice demonstrate that *Mycobacterium avium* subsp. *paratuberculosis* translocates across the mucosal barrier via both M cells and enterocytes but has inefficient dissemination. *Infect. Immun.* **78**, 3570–3577 (2010).
- Abendaño, N., Juste, R. A. & Alonso-Hearn, M. Anti-inflammatory and antiapoptotic responses to infection: A common denominator of human and bovine macrophages infected with *Mycobacterium avium* subsp. *paratuberculosis*. *Biomed Res. Int.* **2013**, 1–7 (2013).
- González, J. *et al.* Histopathological classification of lesions associated with natural paratuberculosis infection in cattle. *J. Comp. Pathol.* **133**, 184–196 (2005).
- Sweeney, R. W. Pathogenesis of paratuberculosis. *Vet. Clin. North Am. - Food Anim. Pract.* **27**, 537–546 (2011).
- Cheng, A. C. *et al.* CLEC9A modulates macrophage-mediated neutrophil recruitment in response to heat-killed *Mycobacterium tuberculosis* H37Ra. *PLoS One* **12**, 1–20 (2017).
- Ott, S. L., Wells, S. J. & Wagner, B. A. Herd-level economic losses associated with John's disease on US dairy operations. *Prev. Vet. Med.* **40**, 179–192 (1999).
- Hasonova, L. & Pavlik, I. Economic impact of paratuberculosis in dairy cattle herds: A review. *Vet. Med. (Praha)*. **51**, 193–211 (2006).
- Smith, R. L. *et al.* Effect of John's disease status on reproduction and culling in dairy cattle. *J. Dairy Sci.* **93**, 3513–3524 (2010).
- McAloon, C. G. *et al.* The effect of paratuberculosis on milk yield—A systematic review and meta-analysis. *J. Dairy Sci.* **99**, 1449–1460 (2015).
- Robertson, R., Hill, B., Cerf, O., Jordan, K. & Venter, P. A commentary on current perspectives on *Mycobacterium avium* subsp. *paratuberculosis*, John's disease and Crohn's disease: A review by Over *et al.* (2011). *Crit. Rev. Microbiol.* **38**, 183–184 (2012).
- Nielsen, S. S. & Toft, N. A review of prevalences of paratuberculosis in farmed animals in Europe. *Prev. Vet. Med.* **88**, 1–14 (2009).
- Lombard, J. E. Epidemiology and Economics of Paratuberculosis. *Vet. Clin. North Am. - Food Anim. Pract.* **27**, 525–535 (2011).
- Juste, R. A. *et al.* Significant reduction in bacterial shedding and improvement in milk production in dairy farms after the use of a new inactivated paratuberculosis vaccine in a field trial. *BMC Res. Notes* **2**, 233 (2009).
- Alonso-Hearn, M. *et al.* Immunization of adult dairy cattle with a new heat-killed vaccine is associated with longer productive life prior to cows being sent to slaughter with suspected paratuberculosis. *J. Dairy Sci.* **95**, 618–629 (2012).
- Garrido, J. M. *et al.* Paratuberculosis vaccination causes only limited cross-reactivity in the skin test for diagnosis of bovine tuberculosis. *PLoS One* **8**, 2–8 (2013).
- Groenendaal, H. *et al.* A simulation of John's disease control. *Prev. Vet. Med.* **54**, 225–245 (2002).
- Garry, F. Control of paratuberculosis in dairy herds. *Vet. Clin. North Am. - Food Anim. Pract.* **27**, 599–607 (2011).
- Kalis, C. H. J., Hesselink, J. W., Barkema, H. W. & Collins, M. T. Culture of strategically pooled bovine fecal samples as a method to screen herds for paratuberculosis. *J. Vet. Diagnostic Investig.* **12**, 547–551 (2000).
- Nielsen, S. S. & Toft, N. Ante mortem diagnosis of paratuberculosis: A review of accuracies of ELISA, interferon- $\gamma$  assay and faecal culture techniques. *Vet. Microbiol.* **129**, 217–235 (2008).
- Capsel, R. T. *et al.* Composition and potency characterization of *Mycobacterium avium* subsp. *paratuberculosis* purified protein derivatives. *PLoS One* **11**, 1–19 (2016).
- Seth, M. *et al.* Biomarker discovery in subclinical mycobacterial infections of cattle. *PLoS One* **4**, 4(5), e5478 (2009).
- Casey, M. E. *et al.* Analysis of the bovine monocyte-derived macrophage response to *Mycobacterium avium* subspecies *paratuberculosis* infection using RNA-seq. *Front. Immunol.* **6**, 1–15 (2015).

23. Marino, R. *et al.* John's disease in cattle: an *in vitro* model to study early response to infection of *Mycobacterium avium* subsp. *paratuberculosis* using RNA-seq. *Mol. Immunol.* **91**, 259–271 (2017).
24. Hempel, R. J., Bannantine, J. P. & Stabel, J. R. Transcriptional profiling of ileocecal valve of Holstein dairy cows infected with *Mycobacterium avium* subsp. *paratuberculosis*. *PLoS One* **11**, 1–19 (2016).
25. Marioni, J. C., Mason, C. E., Mane, S. M., Stephens, M. & Gilad, Y. RNA-seq: an assessment of technical reproducibility and comparison with gene expression arrays. *Genome Res.* **18**, 1509–17 (2008).
26. Mortazavi, A., Williams, B. A., McCue, K., Schaeffer, L. & Wold, B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat. Methods* **5**, 621–8 (2008).
27. Casas, E., Garcia, M. D., Wells, J. E. & Smith, T. P. L. Association of single nucleotide polymorphisms in the *ANKRA2* and *CD180* genes with bovine respiratory disease and presence of *Mycobacterium avium* subsp. *paratuberculosis*. *Anim. Genet.* **42**, 571–577 (2011).
28. Cheriya, V., Leaman, D. W. & Borden, E. C. Emerging roles of FAM14 family members (G1P3/ISG 6–16 and ISG12/IFI27) in innate immunity and cancer. *J. Interf. Cytokine Res.* **31**, 173–181 (2010).
29. Tsuji, S. *et al.* Capture of heat-killed *Mycobacterium bovis* bacillus Calmette-Guérin by intelectin-1 deposited on cell surfaces. *Glycobiology* **19**, 518–526 (2009).
30. Xu, Z. & Jin, B. A novel interface consisting of homologous immunoglobulin superfamily members with multiple functions. *Cell. Mol. Immunol.* **7**, 11–19 (2010).
31. Guler, R., Roy, S., Suzuki, H. & Brombacher, F. Targeting *Batf2* for infectious diseases and cancer. *Oncotarget* **6**, 26575–26582 (2015).
32. Guler, R. *et al.* *Batf2* differentially regulates tissue immunopathology in Type 1 and Type 2 diseases. *Mucosal Immunol.* **12**, 390–402 (2019).
33. Sunil, Y., Ramadori, G. & Raddatz, D. Influence of *NFκB* inhibitors on *IL-1β*-induced chemokine *CXCL8* and *-10* expression levels in intestinal epithelial cell lines: Glucocorticoid ineffectiveness and paradoxical effect of PDTC. *Int. J. Colorectal Dis.* **25**, 323–333 (2010).
34. Jenum, S. *et al.* Added value of *IP-10* as a read-out of *Mycobacterium tuberculosis* specific immunity in young children. *Pediatr. Infect. Dis. J.* **35**, 1336–1338 (2016).
35. Wang, S. *et al.* Evaluation of the diagnostic potential of *IP-10* and *IL-2* as biomarkers for the diagnosis of active and latent tuberculosis in a BCG-vaccinated population. *PLoS One* **7**, 1–10 (2012).
36. Wang, S. *et al.* Evaluation of a whole blood chemiluminescent immunoassay of interferon-gamma inducible protein 10 (*IP-10*) for diagnosis of tuberculosis patients. *Clin. Lab.* **62**, 165–172 (2016).
37. Coad, M. *et al.* Simultaneous measurement of antigen-induced *CXCL10* and *IFN-γ* enhances test sensitivity for bovine TB detection in cattle. *Vet. Microbiol.* **230**, 1–6 (2019).
38. Purdie, A. C., Plain, K. M., Begg, D. J., de Silva, K. & Whittington, R. J. Expression of genes associated with the antigen presentation and processing pathway are consistently regulated in early *Mycobacterium avium* subsp. *paratuberculosis* infection. *Comp. Immunol. Microbiol. Infect. Dis.* **35**, 151–162 (2012).
39. Gijssbers, K. *et al.* *CXCR1*-binding chemokines in inflammatory bowel diseases: Down-regulated *IL-8/CXCL8* production by leukocytes in Crohn's disease and selective *GCP-2/CXCL6* expression in inflamed intestinal tissue. *Eur. J. Immunol.* **34**, 1992–2000 (2004).
40. Shukla, S. K. *et al.* Differential gene expression in *Mycobacterium bovis* challenged monocyte-derived macrophages of cattle. *Microbial Pathogenesis* **113**, 480–489 (2017).
41. Manna, S. K. & Ramesh, G. T. Interleukin-8 induces nuclear transcription factor- $\kappa$ B through a *TRAF6*-dependent pathway. *J. Biol. Chem.* **280**, 7010–7021 (2005).
42. Luppi, F., Longo, A. M., de Boer, W. I., Rabe, K. F. & Hiemstra, P. S. Interleukin-8 stimulates cell proliferation in non-small cell lung cancer through epidermal growth factor receptor transactivation. *Lung Cancer* **56**, 25–33 (2007).
43. Itoh, Y. *et al.* *IL-8* promotes cell proliferation and migration through metalloproteinase-cleavage proHB-EGF in human colon carcinoma cells. *Cytokine* **29**, 275–282 (2005).
44. Lin, J. *et al.* Transcriptome changes upon *in vitro* challenge with *Mycobacterium bovis* in monocyte-derived macrophages from bovine tuberculosis-infected and healthy cows. *Vet. Immunol. Immunopathol.* **163**, 146–156 (2015).
45. Rotta, G. *et al.* Contrasting roles of SPARC-related granuloma in bacterial containment and in the induction of anti-*Salmonella typhimurium* immunity. *J. Exp. Med.* **205**, 657–667 (2008).
46. De Buck, J., Shaykhtudinov, R., Barkema, H. W. & Vogel, H. J. Metabolomic profiling in cattle experimentally infected with *Mycobacterium avium* subsp. *paratuberculosis*. *PLoS One* **9**, e111872 (2014).
47. Hare, N. J. *et al.* *Mycobacterium tuberculosis* infection manipulates the glycosylation machinery and the N-Glycoproteome of human macrophages and their microparticles. *J. Proteome Res.* **16**, 247–263 (2017).
48. Woo, S. R., Barletta, R. G. & Czuprynski, C. J. ATP release by infected bovine monocytes increases the intracellular survival of *Mycobacterium avium* subsp. *paratuberculosis*. *Comp. Immunol. Microbiol. Infect. Dis.* **32**, 365–377 (2009).
49. Johansen, M. D., de Silva, K., Plain, K. M., Whittington, R. J. & Purdie, A. C. *Mycobacterium avium* subspecies *paratuberculosis* is able to manipulate host lipid metabolism and accumulate cholesterol within macrophages. *Microb. Pathog.* **130**, 44–53 (2019).
50. Stabel, J. R., Reinhardt, T. A. & Hempel, R. J. Short communication: Vitamin D status and responses in dairy cows naturally infected with *Mycobacterium avium* ssp. *paratuberculosis*. *J. Dairy Sci.* **102**, 1594–1600 (2019).
51. Juste, R. A., Marco, J. C., Carmen, S. & Oc, D. Comparison of different media for the isolation of small ruminant strains of *Mycobacterium paratuberculosis*. *Vet. Microbiol.* **28**, 385–390 (1991).
52. Schmeider, R. & Edwards, R. Quality control and preprocessing of metagenomic datasets. *Bioinformatics* **27**, 863–864 (2011).
53. Trapnell, C., Pachter, L. & Salzberg, S. L. TopHat: Discovering splice junctions with RNA-Seq. *Bioinformatics* **25**, 1105–1111 (2009).
54. Trapnell, C. *et al.* Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat. Protoc.* **7**, 562–578 (2012).
55. Oshlack, A. & Wakefield, M. J. Transcript length bias in RNA-seq data confounds systems biology. *Biol. Direct* **4**, 1–10 (2009).
56. Young, M. D., Wakefield, M. J., Smyth, G. K. & Oshlack, A. Gene ontology analysis for RNA-seq: accounting for selection bias. *Genome Biol.* **11**, R14 (2010).
57. Rivals, I., Personnaz, L., Taing, L. & Potier, M. C. Enrichment or depletion of a GO category within a class of genes: Which test? *Bioinformatics* **23**, 401–407 (2007).
58. Szklarczyk, D. *et al.* STRING v11: Protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res.* **47**, D607–D613 (2018).
59. Szklarczyk, D. *et al.* The STRING database in 2017: Quality-controlled protein-protein association networks, made broadly accessible. *Nucleic Acids Res.* **45**, D362–D368 (2017).

## Acknowledgements

This work was supported by grants from the Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria (INIA) and by European Funds for Regional Development (FEDER) (INIA RTA2014-00009-C02 and RTA2018-094192). The study is partially funded by the Principado de Asturias (PCTI 2018–2020, GRUPIN: IDI2018-000237). Maria Canive and Cristina Blanco-Vazquez are recipients of INIA fellowships. The authors thank ASTEGA Veterinary services for their assistance on sample collection. We are also grateful to Kyle Hearn for the careful editing of the manuscript.



### Author Contributions

M.A.H. coordinated the project. M.A.H. and R.C. designed the experiments. R.C. and C.B.V. performed animal sampling. R.C., C.B.V., A.B. and J.A. performed diagnostic tests. R.T. and R.R. performed the RNA-sequencing at the Parque Científico de Madrid sequencing platform. B.M.J. and E.V.M. helped with the bioinformatics analysis. M.A.H. and M.C. interpreted all the results and drafted the manuscript. All the authors carefully read and approved the final version of the manuscript.

### Additional Information

**Supplementary information** accompanies this paper at <https://doi.org/10.1038/s41598-019-51328-0>.

**Competing Interests:** The authors declare no competing interests.

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2019