Transcriptional response of ovine lung to infection with jaagsiekte sheep retrovirus

Anna Eleonora Karagianni[a,*], Deepali Vasoya[b], Jeanie Finlayson[a], Henny M. Martineau[a,‡], Ann R. Wood[a], Chris Cousens[a], Mark P. Dagleish[a], Mick Watson[b], David J. Griffiths[a#]

[a]*Moredun Research Institute, Edinburgh, UK;*

[b]*Roslin Institute and Royal (Dick) School of Veterinary Studies, University of Edinburgh, UK*

Address correspondence to David J. Griffiths (david.griffiths@moredun.ac.uk)

Running title: RNA-Seq of ovine pulmonary adenocarcinoma

Keywords: JSRV, jaagsiekte, RNA-Seq, lung adenocarcinoma, sheep, cancer

* Present address: Roslin Institute and Royal (Dick) School of Veterinary Studies, Easter Bush, Midlothian, UK

‡ Present address: Royal Veterinary College, Hertfordshire, UK

Word counts: Abstract (225) Main text (5013) Materials and methods (1147)

Abstract

Jaagsiekte sheep retrovirus (JSRV) is the etiologic agent of ovine pulmonary adenocarcinoma (OPA), a neoplastic lung disease of sheep. OPA is an important economic and welfare issue for sheep farmers and a valuable naturally-occurring animal model for human lung adenocarcinoma. Here, we used RNA sequencing to study the transcriptional response of ovine lung tissue to infection by JSRV. We identified 1,971 ovine genes differentially-expressed in JSRV-infected lung compared to non-infected lung, including many genes with roles in carcinogenesis and immunomodulation. The differential expression of selected genes was confirmed using immunohistochemistry and RT-qPCR. A key finding was the activation of anterior-gradient-2, yes-associated protein-1 and amphiregulin in OPA tumor cells, indicating a role for this oncogenic pathway in OPA. In addition, there was differential expression of genes related to innate immunity including genes encoding cytokines, chemokines and complement system proteins. In contrast, there was little evidence for upregulation of genes involved in T-cell immunity. Many genes related to macrophage function were also differentially expressed, reflecting the increased abundance of these cells in OPA-affected lung tissue. Comparison of the genes differentially regulated in OPA with transcriptional changes occurring in human lung cancer revealed important similarities and differences between OPA and human lung adenocarcinoma. This study provides valuable new information on the pathogenesis of OPA and strengthens the use of this naturally occurring animal model for human lung adenocarcinoma.

43 **Importance**

44 Ovine pulmonary adenocarcinoma is a chronic respiratory disease of sheep caused by jaagsiekte
45 sheep retrovirus (JSRV). OPA is a significant economic problem for sheep farmers in many
46 countries and is a valuable animal model for some forms of human lung cancer. Here, we
47 examined changes in host gene expression that occur in the lung in response to JSRV infection.
48 We identified a large number of genes with altered expression in infected lung, including factors
49 with roles in cancer and immune system function. We also compared the data from OPA to
50 previously published data from human lung adenocarcinoma and found a large degree of overlap
51 in the genes that were dysregulated. The results of this study provide exciting new avenues for
52 future studies of OPA and may have comparative relevance for understanding human lung
53 cancer.

54

55 **Introduction**

56 Ovine pulmonary adenocarcinoma (OPA) is an infectious lung disease of sheep that is a
57 significant economic problem and welfare concern for sheep producers in many countries (1, 2).
58 The main clinical features of OPA are loss of condition and respiratory distress and in many
59 advanced cases there is an accumulation of fluid in the lungs that drains from the sheep's nostrils
60 when its head is lowered (3). Lung function is often further compromised by the presence of
61 bacterial or parasitic co-infection (1). By the time clinical signs become apparent, tumor growth
62 is typically extensive and the disease is invariably fatal.

63

64 OPA is caused by jaagsiekte sheep retrovirus (JSRV), an exogenous betaretrovirus (4). Despite
65 its etiological role in OPA, infected sheep produce only a limited adaptive immune response to

JSRV antigens and this has precluded the development of an effective preclinical serological diagnostic test or vaccine to control the spread of disease (2). The mechanism underlying the poor immune responsiveness of sheep to JSRV is not completely understood. It appears likely that it is largely due to immunological tolerance elicited by the developmental expression of closely related endogenous JSRV (enJSRV) proteins in the ovine fetal thymus (5, 6). However, local immunomodulatory mechanisms are also proposed to contribute (7).

JSRV has a specific tropism for differentiated epithelial cells of the distal lung and OPA tumor cells predominantly express markers of type II alveolar epithelial cells (AEC2) (8-10). An unusual feature of JSRV is that the Env glycoprotein functions as a viral oncoprotein to drive neoplastic transformation *in vitro* (11, 12) and *in vivo* (13-16). JSRV Env expression activates a number of signaling pathways that control cellular proliferation, including PI3K-Akt and MEK-ERK1/2 (17, 18). Several cellular factors that bind Env have been identified and are proposed to be involved in transformation (19-21) but further work is necessary to provide a complete model for Env-mediated tumorigenesis and to explain how this leads to the unique clinical presentation of OPA.

In addition to its veterinary importance, OPA represents a valuable animal model for some forms of human lung cancer due to similarities in histological appearance and the activation of common oncogenic signaling pathways (22-25). In its early stages; such as in subclinical natural disease and in experimentally-infected lambs, OPA resembles a minimally invasive adenocarcinoma with a predominantly lepidic growth pattern (22, 24). In advanced natural disease, OPA is more closely similar to adenocarcinoma with papillary or acinar predominant growth with or without

mucinous features (22, 25). The similarity of OPA to human lung adenocarcinoma suggests that this naturally occurring sheep tumor could be valuable for understanding lung carcinogenesis, particularly at the early stages of disease, which are difficult to diagnose and study in humans.

In order to examine the pathogenesis of OPA, we determined changes in host gene expression in the lungs of lambs following experimental infection with JSRV. Many genes were identified to have altered expression and we confirmed the upregulation of some of these using immunohistochemistry and RT-qPCR. We also compared the differential gene expression of OPA-affected animals with previously published data on the two most common types of non-small-cell lung carcinoma (NSCLC) in humans; lung adenocarcinoma (LUAD) and lung squamous cell carcinoma (LUSC). Collectively, this study provides new information that greatly enhances our understanding of the host response to JSRV and provides a number of exciting new avenues for future work on OPA.

RESULTS

JSRV-induced gene expression in sheep lung tissue.

Whole transcriptome profiling (RNA-Seq) was performed on lung samples derived from JSRV infected (n=4) and mock-infected (n=4) specific pathogen-free (SPF) lambs. The samples used were obtained from a previous study (9) in which 6-day old lambs were infected with JSRV by intratracheal injection and euthanized when signs of respiratory distress appeared (66 – 85 days post-inoculation). Mock-infected lung tissue came from age and sex-matched control lambs that had been inoculated with cell culture medium (9). OPA tumor lesions were observed in

hematoxylin and eosin-stained lung tissue sections from infected lambs and confirmed with immunohistochemistry (IHC) for JSRV Env (SU) protein. Such lesions were not present in lung tissue from mock-infected lambs (9). Infected tissues contained many tumor foci distributed throughout the lung that otherwise appeared histologically normal, a pattern typical of experimentally-induced OPA (22). Samples for RNA-Seq were generated by pooling tissue from 7 distinct sites of each lung. In order to capture specimens that were representative of the whole tissue the samples analyzed were not specifically enriched for tumor cells.

RNA-Seq generated over 60 million reads per sample, of which approximately 80% mapped uniquely onto the ovine and JSRV genomes (Table 1). A total of 15,149 sheep genes were identified and reads were quantified to identify those that were differentially expressed between mock-infected and JSRV-infected samples. Principal component analysis of the normalized counts of sheep genes clearly separated the mock-infected and the JSRV-infected groups (Fig. 1A). JSRV infection produced a radical change in sheep gene expression with 1,971 differentially expressed transcripts identified between the two groups (Fig. 1B) (1237 upregulated and 734 downregulated). We defined differentially expressed genes as those showing up or down regulation following JSRV infection with a false discovery rate (FDR) below 0.05, regardless of the observed fold-change. The complete lists of mapped genes and differentially expressed genes are presented in Supplementary Dataset S1. Hierarchical clustering of all differentially expressed genes is shown in Fig. 1C. Fig. 1D summarizes the 25 most significantly upregulated and downregulated genes.

Although the RNA-Seq analysis revealed clear differences in gene expression between the infected and control groups, one infected lamb (Infected_F_85days) showed an intermediate overall expression pattern (Fig 1A, 1C). Of the four infected lambs, this animal had the lowest percentage of reads mapping to JSRV in the RNA-Seq data (Table 1) and therefore likely had the smallest proportion of tumor-affected tissue and the largest contribution from 'healthy' uninfected tissue. Nevertheless, hierarchical clustering grouped this animal with the other infected lambs (Fig. 1C, D) and so it was included in subsequent analyses.

Functional analysis of RNA-Seq data.

KEGG pathway enrichment analysis of the upregulated and downregulated genes was performed using the DAVID annotation software (https://david.ncifcrf.gov/ (26)) (Supplementary Dataset S2). Upregulated pathways in the OPA-affected lung included genes involved in metabolic pathways, epithelial cell differentiation, cell cycle and wound healing. Pathways involved in immune response and inflammation were down-regulated. Ingenuity Pathway Analysis (IPA) (27) was also used to functionally analyze the list of differentially expressed genes. The main 'Diseases and Biofunctions' related to the differentially expressed genes are shown in Fig. 2. As with DAVID analysis, cancer related functions, such as cell proliferation and tissue development, were identified as highly enriched whereas those related to immune response and inflammation were less strongly represented. These results revealed key signaling networks in cancer, including neovascularization, cell viability and tumor growth, to be activated in JSRV-infected tissue compared to mock-infected tissue (Fig. 3). Based on these analyses, we used IHC to validate selected upregulated markers, focusing in particular on pathways of relevance to OPA pathogenesis, as discussed below.

Cancer-related gene expression in OPA-affected lung tissue.

Many of the most significantly upregulated genes in OPA have previously been associated with important aspects of tumorigenesis in mice and humans. These include functions such as cell proliferation and differentiation (*CLDN2*, *CTSL*, *PROM2*, *AGR2*, *EPHA7*, *SOX9*, *KRT18*), angiogenesis (*MMP9*, *CCLA2*, *EPHB2*, *CXCL8*), metastasis (*CCLA2*, *EPHB2*, *AGR2*, *HS6ST2*, *MUC1*) and signal transduction, including EGF family members (*AREG*, *EREG*, *NRG2*) and Wnt signaling components (e.g., *WNT10B*, *PPARG*, *KRT18*, *LEF1*, *AXIN2*, *FZD3*, *FZD5*, *ROR1*, *CTNNB1* and *CDH1*). In addition, factors previously identified as specific markers of lung adenocarcinoma were upregulated, including *PROM2*, *CLDN3* and *TJP3*, as were genes proposed to have potential diagnostic or prognostic value in human cancers, including *MMP9*, *AGR2*, *SULF1*, *NHSL1*, *LGR5*, *MUC1* and *PIK3C2G*.

Although several factors related to angiogenesis were found to be upregulated using RNA-Seq, a previous study reported that vascular endothelial growth factor (VEGF) signaling was downregulated in OPA (28). The RNA-Seq data is consistent with that study, finding reduced expression of VEGFD, VEGFA and VEGFR2 (*KDR*) in JSRV-infected lambs, although only VEGFD passed our cut-off for statistical significance. The upregulation of MMP9 in OPA-affected lung tissue found by RNA-Seq is also in agreement with the previous study (28).

The upregulation of AGR2 (anterior gradient-2) was of particular interest as this protein has been shown to promote oncogenesis in adenocarcinomas of several tissues, including the lung (29). To analyze the expression of AGR2 in OPA, IHC was performed on lung tissue sections from

natural and experimental OPA and mock-infected lambs. The cytoplasm of bronchiolar epithelial cells of mock-infected lung labeled strongly for AGR2 but there was no labeling of cells in the alveolar compartment (Fig. 4A, D). In contrast, in JSRV-infected lung, there were many regions of strong AGR2 labeling and these correlated closely with areas of JSRV-Env labeling in serial adjacent sections (Fig. 4B, C, E, F). This was evident even on relatively small tumor foci suggesting that AGR2 is activated early following JSRV infection. Immunofluorescent labeling confirmed the co-expression of JSRV Env and AGR2 in infected lung tissue, although not every infected cell cluster labeled positively for AGR2 (Fig. 4G-J). AGR2 has been shown to stimulate the expression of the EGFR ligand amphiregulin (*AREG*) in adenocarcinoma cells (30). AREG was also upregulated in OPA and, as for AGR2, IHC demonstrated positive labeling for AREG in OPA tumor cells from both natural and experimentally-induced disease (Fig. 4K-M).

The activation of AREG by AGR2 is mediated by Yes-associated protein (YAP1), a nuclear effector of the Hippo signaling pathway (30), which led us to ask whether Hippo signaling is involved in OPA. YAP1 was not found to be differentially expressed by RNA-Seq but to examine this question further, IHC was performed on tissue sections of mock-infected and OPA-affected lung using antibodies to the phosphorylated form of YAP1 and to total YAP1 (Fig. 5). IHC detected both the phosphorylated and unphosphorylated forms of YAP1 in the bronchiolar epithelium and in the alveolar compartment of mock-infected lung (Fig. 5A, D). As expected, the phosphorylated (inactive) form was predominantly detected in the cytoplasm, whereas an antibody to total YAP1 (which detects active and inactive forms) exhibited prominent nuclear labeling that was more intense than labeling of the cytoplasm. Tumor cells in experimental OPA

had strong nuclear labeling for total YAP1 (Fig. 5B) but this appeared less intense in natural OPA (Fig. 5C). IHC for regulators of the Hippo pathway, MST1/2 and LATS1/2, found these proteins were readily detectable in AEC2s in healthy sheep lung (Fig. 5G, J), as previously reported in mouse lung (31). OPA tumor cells also labeled strongly for both markers (Fig. 5H, I, K, L). Interestingly, the distribution of labeling for LATS1/2 differed between experimentally-induced tumors and natural cases of OPA, being predominantly cytoplasmic in experimental disease and predominantly nuclear in natural disease. While a complete analysis of the activation state of the Hippo pathway is outside the scope of this study, these data indicate that the AGR2-YAP1-AREG axis is active in OPA and may contribute to oncogenesis in this disease.

Immune response to JSRV infection and oncogenesis.

The RNA-Seq analysis identified altered expression of many genes related to immune responses including various cytokines and chemokines (e.g., *CSF2, CCL2, CXCL6, CXCL8, CXCL14, IL1A, IL6, TGFB3* and *TNFSF18* were upregulated and *CSF1, CCL4, CCL18, CCL26, CXCL12, IL15* and *TNF* were downregulated), along with several members of the complement pathway (*C3, C5, C6, C7, MBL1 (ENSOARG00000010165), CFI, CD46, PTX3, C4BPA, C4BPB* upregulated and *C1QA* and *C1QB* downregulated), confirming a substantial host response to infection and tumor growth. Selected cytokines were also evaluated by RT-qPCR in experimentally-induced and natural OPA and this confirmed their differential expression (Fig. 6). Interestingly, the expression data does not support strong induction of type I interferon responses in JSRV-infected lung. Although IFNA was not identified using RNA-Seq, RT-qPCR analysis found it to be downregulated in natural OPA (Fig. 6). Expression of IFNB1 was not significantly changed in JSRV-infected tissue. Furthermore, when considering a recently

described panel of 90 highly conserved mammalian type-1 interferon-stimulated genes (32), only one (IRF7) was significantly upregulated in our data.

IFNG expression was not significantly upregulated in cases of experimental OPA suggesting little or no activation of cellular Th1 responses in early disease. However, RT-qPCR found that IFNG was increased in tissue from cases of advanced natural OPA (Fig. 6B), consistent with a previous report of IFNG-positive macrophages in natural OPA (7). Markers of T-lymphocytes were either not significantly changed (CD3 (gamma, delta and epsilon chains) and CD4) or downregulated (CD8, alpha and beta chains). Interestingly, expression of *FOXP3* and *CD44*, both markers of regulatory T-cells (Tregs), was significantly increased in OPA. Tregs negatively regulate T-cell immune responses through a number of mechanisms and have been shown to reduce the ability of CD8 T-cells to produce TNF (33). Consistent with this, *TNF* expression was downregulated in experimentally and naturally-infected animals (Fig. 6).

In agreement with previous studies (7, 34), macrophages were observed in many sections of the experimental OPA cases studied here. In addition, approximately 10% (198 of 1971) of the differentially expressed genes in our analysis have macrophage-related functions (Fig. 7; Supplementary Dataset S3), confirming that macrophage-related gene expression is altered in JSRV-infected tissue. Upregulated genes include some previously associated with an inflammatory phenotype (e.g., *HIF1A*, *IL1A*, *CXCL8*, *CSF2*, *IRAK1*, *IRF7*) or an immunoregulatory phenotype (e.g., *CD163*, *CCL2*, *LGMN*, *MMP9*, *MMP14*), suggesting that a complex pattern of macrophage activation exists in OPA. Selected macrophage markers were also studied by IHC. Labeling for CD68, a commonly-used marker of monocytes and tissue

macrophages, identified cells located in normal alveolar lung tissue and along the margin of tumor foci (Fig. 8A-C). In comparison, an antibody to CD163, a marker for macrophages with an immunomodulatory phenotype (35), labeled many more cells, including cells that surrounded or infiltrated the tumor foci (Fig. 8D-F). Macrophages in or around OPA foci also labeled positively for legumain (LGMN), another marker associated with an immunoregulatory phenotype (Fig. 8G-I) and hypoxia-inducible factor 1-alpha (HIF1A, which is associated with an inflammatory macrophage phenotype (Fig. 8J-L). LGMN was also detected in alveolar macrophages in mock-infected tissue, while HIF1A labeling in mock-infected lung was limited to the bronchiolar epithelium. Interestingly, the antibodies to HIF1A and LGMN also labeled some tumor cells whereas anti-CD163 and anti-CD68 did not. Consistent with the greater abundance of macrophages in OPA-affected lung, RNA-Seq identified increased expression of several myeloid cell chemoattractants (e.g., *CCL2, CSF2, S100A8, S100A9, and CXCL8)* in JSRV-infected tissues.

Comparison of gene expression of ovine pulmonary adenocarcinoma with human lung adenocarcinoma.

We next compared the RNA-Seq data from OPA with previously published expression data from the two most common types of NSCLC (LUAD and LUSC). Gene counts data from RNA-Seq studies on tumors from patients with LUAD and LUSC were downloaded from The Cancer Genome Atlas (TCGA) database (portal.gdc.cancer.gov) (36). Data were selected from patients for whom matched normal and tumor data were available (57 LUAD and 49 LUSC; Supplementary Dataset S4 shows their clinical annotation). The genes differentially expressed between tumor and normal samples were identified in each clinical disease stage (LUAD I – IV

and LUSC I - III); (Supplementary Dataset S5). To compare differentially expressed genes in sheep and human cancers, we considered only those sheep genes with one-to-one orthologues in humans (1726 out of 1971 (87%)). We then compared the lists of differentially expressed orthologous genes between OPA and the different stages of human LUAD and LUSC by calculating the Pearson correlation coefficient between the log ratios of genes in common between the two lists. The correlation plot of orthologous genes between OPA and the different stages of human LUAD and LUSC (Fig. 9A) revealed that the gene expression profile of OPA is more similar to LUAD than it is to LUSC and that OPA is more closely correlated with stage I LUAD than with later stages. The main Diseases and Biofunctions associated with the differentially expressed orthologous genes in OPA and stage I LUAD were identified with IPA (Fig. 9B). Supplementary Dataset S6 lists the fold-change of the genes differentially expressed in OPA with their human orthologues in LUAD and LUSC and identifies those genes with similar or divergent changes in expression between the two species. Fig. 9C summarizes the fold-change of the 50 genes most significantly deregulated in OPA with their human orthologue from stage I LUAD. Of these, 37 were consistent in their direction of change and 13 were discordant.

DISCUSSION

The pathogenesis of OPA is intriguing. How does JSRV infection and expression of its oncogenic Env protein elicit the striking clinical and pathological phenotype typical of this disease? As the mechanisms underlying this process remain largely undetermined, here we used a global RNA-Seq approach to gain insight into the gene pathways and networks that are altered during the early stages of OPA. We identified 1971 differentially expressed sheep genes and

validated several of these by IHC or RT-qPCR. Some changes revealed the altered cellular composition of infected and uninfected lung. For example, the increased number of macrophages and the presence of tumor cells derived from AEC2s in infected tissue is reflected in the upregulation of markers of these cell types. However, many of the genes and pathways that have altered expression in JSRV infection have not been previously studied in OPA but are associated with important aspects of OPA pathogenesis including tumor pathways and local immune responses. Collectively, this study provides important new information on the host response to JSRV infection that will be valuable for studies of OPA directly and for exploitation of the sheep disease as a lung cancer model.

In this study, we utilized tissue from experimentally-infected SPF lambs in order to avoid the potential confounding effects of additional respiratory infections that commonly occur in natural cases of OPA (1). For welfare reasons, experimentally-infected lambs must be culled when respiratory signs first appear and therefore the tissues studied represent an early stage of disease compared to the advanced stage of natural cases of OPA. We used sections of total lung tissue because we were interested in analyzing the gene expression of the whole infected tissue. A consequence of this approach is that the experimental OPA samples contained a greater proportion of histologically normal lung tissue than tumor tissue. We anticipated that this excess of 'healthy' tissue would reduce the sensitivity of our study to detect differentially expressed genes, in particular down-regulated genes (37, 38). This would appear to be the case as only four genes were found to have downregulated expression by more than two-fold in OPA. Nevertheless, while this approach is expected to underestimate the fold-changes of differentially-expressed genes it should still provide the correct order of significance. Future work will focus

318 on analysis of gene expression in specific cell types such as tumor cells and myeloid cells, which
319 should increase the sensitivity of detection of differentially expressed genes in those cell subsets.
320
321 **Oncogenic signaling pathways in OPA.**
322 The mechanisms involved in JSRV Env-mediated tumorigenesis are not completely understood
323 and multiple cellular pathways appear to be activated (15, 17, 18). Several studies have shown
324 that Env expression leads to activation of the PI3K/Akt and MAPK/ERK1/2 signaling pathways
325 (14, 18, 39, 40), while additional pathways including EGFR and Wnt signaling may also play a
326 role (15, 19, 41, 42). The RNA-Seq data provided evidence supporting a variety of changes in
327 gene expression related to carcinogenesis, including increased expression of several ligands of
328 EGFR and other ERBB family receptors and upregulation of a variety of ligands and receptors
329 related to Wnt signaling in JSRV-infected lung tissue.
330
331 Interestingly, the major components of the PI3K/Akt and MAPK/ERK1/2 signaling pathways
332 were not significantly upregulated. Indeed, DAVID pathway analysis reported that both of these
333 pathways were downregulated in OPA. There are several possible explanations for this apparent
334 discrepancy. For example, activation of the MAPK and PI3K/Akt pathways is mediated by
335 protein phosphorylation cascades, which may not correspond directly to changes in transcription
336 of those genes. In addition, while previous studies have shown roles for Akt and ERK signaling
337 in tumor cells or transfected cell lines, the RNA-Seq analysis presented here reflects global
338 transcriptional changes within the whole tissue, which comprises multiple cell types. Most
339 importantly, the MAPK and PI3K/Akt pathways are involved in many cellular processes in
340 addition to oncogenesis, including immune responses (43, 44), and the list of genes that DAVID

and similar tools uses for identifying members of a pathway is very broad. For example, several of the genes identified by DAVID as involved in the PI3K/Akt and MEK/ERK1/2 pathways are involved in immune responses (e.g., *TLR2, CSF1*, *COL6A3, NFKB1 and TNF)*, which is consistent with the other changes in immune and inflammatory responses observed in our study.

An important novel finding in our study was the activation of AGR2 in OPA tumor cells (Fig. 4). AGR2 is a protein disulfide isomerase localized in the endoplasmic reticulum in normal cells, but is upregulated in a variety of human adenocarcinomas where it may also be present in secreted and cell-surface-bound forms (45, 46). AGR2 is associated with poor prognosis in several cancer types and appears to mediate its oncogenic effect through the regulation of other genes including *TP53* (47) and *AREG* (30) and through extracellular functions such as promoting angiogenesis and extracellular matrix remodeling (46, 48). In humans, AGR2 expression is significantly higher in LUAD than in LUSC (49). The activation of AREG by AGR2 has been shown to be dependent on the Hippo pathway effector protein YAP1 in human adenocarcinoma cell lines (30). Consistent with this, IHC analysis confirmed the presence of nuclear YAP1 in OPA tumor cells (Fig. 5), suggesting a possible role for Hippo signaling in OPA pathogenesis. While further work is required, the transcriptional upregulation of AGR2, EGFR ligands and other oncogenic factors identified by RNA-Seq provides new opportunities for understanding oncogenic signaling in OPA. In turn, OPA provides a model for studying the function of these pathways in a naturally occurring tumor, particularly at the early stages of tumorigenesis.

Local immune responses in OPA-affected lung tissue.

363 The absence of a significant adaptive immune response to JSRV in sheep is commonly attributed
364 to enJSRV expression in the fetal thymus during immune development, which is proposed to
365 lead to immune tolerance through the deletion of T-lymphocytes that recognize the closely
366 related JSRV proteins (5). The results of the RNA-Seq analysis suggest that JSRV infection
367 induces substantial immunological changes in lung tissue, including altered expression of
368 numerous cytokines, chemokines and complement factors together with an increase in
369 macrophages associated with tumor foci. These changes suggest that local immune-modulatory
370 mechanisms active within the OPA-affected lung might also suppress the immune response to
371 JSRV.
372
373 The activation of complement factors and complement regulatory factors in OPA is consistent
374 with a previous microarray analysis of gene expression in a mouse model of JSRV Env-mediated
375 transformation that also found evidence of complement upregulation (50). Complement has
376 been shown previously to inhibit infection of a number of viruses including retroviruses (51),
377 and studies on human and murine cancers have established roles for complement in modulating
378 tumor growth (52-55). Therefore, complement might also play an immunomodulatory role in
379 OPA tumors.
380
381 Macrophages exhibit significant functional plasticity (56, 57) and gene expression studies
382 suggest the existence of numerous subpopulations of macrophages, including some with
383 important roles in cancer (56-58). In addition, studies of mouse and human cancers have shown
384 that macrophages and other myeloid cells in the tumor microenvironment are essential for tumor
385 survival and metastasis (58, 59). Tumor-associated macrophages are also abundant in OPA (7)

but the phenotype of these cells and their functional significance is unclear. In experimentally-induced OPA, we identified changes in the expression of 198 genes related to macrophage function, including genes previously associated with an inflammatory or immunoregulatory phenotype. Interestingly, the expression of macrophage markers in OPA appears to vary depending on their location within the affected tissue, with cells on the periphery of tumor foci being CD68-positive and cells within the tumor being CD163-positive and CD68-negative (Fig. 8). CD163 is regarded as a reliable marker for immunomodulatory macrophages that are associated with poor prognosis in human tumors (60, 61). The positive CD163 labeling of OPA-associated macrophages therefore suggests that these cells might also promote tumor growth in OPA.

In contrast, the transcriptome analysis provided little evidence to support a strong adaptive immune response in OPA. For example, there was no activation of T-cell markers other than a modest but statistically significant increase in markers of Tregs (FOXP3 and CD44; although note that CD44 is also a marker for macrophages and AEC2s). In addition, there was no evidence of substantial activation of type 1 interferon responses in OPA, whereas IFNG expression was not significantly changed in JSRV-infected lambs but was upregulated in natural cases studied by RT-qPCR (Fig. 6). This is possibly due to the presence of concurrent bacterial infections in the natural cases but demonstrates an important difference between early experimental OPA and advanced natural cases.

Collectively, the gene expression data reported here suggest the presence of an immunomodulatory environment within the OPA lung, which has the potential to suppress the

409 immune response to JSRV and tumor cells and to actively promote tumor growth. A more
410 detailed quantitative analysis of the role of tumor-associated myeloid cells, complement and
411 Tregs in OPA is necessary to determine their contribution to tumor growth and development of
412 clinical disease. This could reveal insights into the relevance of the tumor microenvironment to
413 the apparent immune tolerance of sheep to JSRV and may inform the design of vaccine strategies
414 for controlling OPA.
415
416 **Comparison of the transcriptomes of OPA and human LUAD.**
417 As OPA is frequently cited as an animal model for human lung adenocarcinoma, we compared
418 the transcriptome data from OPA with published data on human lung tumors in an attempt to
419 identify the similarities and differences between the diseases in the two species. The results
420 indicate closer similarity of the transcriptome of OPA with human LUAD than with LUSC and
421 in particular with early (stages I and II) compared to more advanced disease (stages III and IV)
422 (Fig. 9A, C; Supplementary Dataset S6). This is consistent with the histological appearance of
423 OPA which resembles a minimally invasive adenocarcinoma in its early stages (2, 22, 24, 25). It
424 would be interesting in future studies to compare gene expression in the early stage of disease
425 studied here with that of more advanced OPA.
426
427 Many genes showed a common pattern of differential expression in LUAD, LUSC and OPA but
428 there were also many that did not. Some of these reflect differences between LUAD and LUSC
429 and highlight the diversity of gene expression in the different tumor types. For example, 142
430 genes were upregulated in both OPA and stage I LUAD but downregulated or not significantly
431 altered in stage I LUSC. Notably, this includes *AGR2* and the Wnt pathway effector protein

beta-catenin. Similarly, 86 genes were upregulated in OPA and stage I LUSC but not upregulated in stage I LUAD. In addition, there were 134 genes upregulated in OPA that were downregulated in both LUAD and LUSC. These include *SFTPC* and *LAMP3* which are known markers of AEC2s in sheep and humans (62). LAMP3 has been detected in OPA previously but is rarely expressed in human tumors except in bronchioloalveolar-adenocarcinoma (lepidic-prominent adenocarcinoma-in-situ) (62), which were not represented in the LUAD cases in the TGCA data studied. Interestingly, the upregulation of complement factors observed in OPA was not evident in either LUAD or LUSC and several other key immune response related genes, such as *TLR10*, *CCL4*, *CCL26*, *CD8A*, *CD8B*, *TNF* and *IDO1*, were specifically downregulated in OPA but not in human LUAD, suggesting that innate immune responses in the sheep and human diseases may differ in important ways. Collectively, this analysis provides support for the similarity of experimental OPA to early stage LUAD but highlights that there remain many differences between the sheep and human diseases.

In summary, the findings from this first large-scale analysis of host gene expression in OPA significantly increase our understanding of the disease pathogenesis at a transcriptional level and will inform future research directed at improving OPA disease control. Moreover, the interspecies comparative data between sheep and humans provide additional support for the use of OPA as a model for early stages of LUAD, particularly non-invasive forms. Finally, a deeper understanding of the pathological changes of early tumors could help to identify novel biomarkers for the early detection of cancer lesions in both species.

MATERIALS AND METHODS

455

456 **Animals and tissues.**

457 Tissues were available from a previous study in which four specific pathogen free (SPF) lambs
458 were inoculated through intra-tracheal injection with JSRV at six days of age (9). Four
459 additional lambs received cell culture supernatant (mock-infected control). Each group
460 contained one female and three male lambs. All lambs were caesarean-derived and housed in
461 SPF conditions to minimize the risk of acquiring additional respiratory infections. Once the
462 clinical signs of respiratory disease were apparent in the JSRV-infected animals (66d, 71d and
463 85d (n=2) post inoculation), lambs were culled. Each time a JSRV-infected lamb was
464 euthanized, a healthy animal from the mock-infected control group was culled to provide age and
465 sex-matched control tissues. Tissues were collected from 24 locations in each lung and stored in
466 liquid nitrogen for RNA extraction and in 10% buffered formalin for IHC. To study tissues from
467 animals with natural disease, lung samples were taken from four farm-raised adult sheep in the
468 advanced stages of clinical OPA and four clinically healthy adult sheep. Cases were selected
469 which had no gross appearance of bacterial or parasitic infection. All protocols involving animal
470 handling and the use of post-mortem material were approved by the Animal Welfare and Ethical
471 Review Body of Moredun Research Institute in accordance with the U.K. Animals (Scientific
472 Procedures) Act 1986.

473

474 **RNA extraction and sequencing.**

475 For RNA-Seq, RNA was extracted from frozen tissue samples from seven distinct sites of the
476 lungs of each animal and pooled in equal amounts. Total RNA from cryosectioned tissue was
477 extracted using an RNeasy mini kit (Qiagen) according to the manufacturer's instructions. RNA

concentration and purity were measured (ND-1000 Nanodrop) and RNA integrity was confirmed (Agilent RNA 6000 Nano kit with Agilent 2100 Bioanalyzer). All samples had an RNA integrity number (RIN) greater than 8.0. Prior to sequencing, frozen lung sections derived from the same lung sites were examined by histology, which confirmed the presence of OPA tumor lesions in the JSRV-infected animals and the absence of lesions in the mock-infected negative controls. Total RNA was processed (TruSeq RNA Library kit) to generate cDNA libraries according to the manufacturer's instructions and subsequently sequenced with an Illumina HiSeq2000 instrument using 100 bp paired-end sequencing (Edinburgh Genomics, University of Edinburgh, UK).

Processing of next generation sequencing data and differential expression analysis. Raw sequencing reads were processed to remove adaptors and poor-quality sequences (Q25 and below) using Cutadapt 1.10 (63). Non-redundant reads were then mapped to the sheep (Oar_v3.1, Ensembl FTP release 74) and JSRV (GenBank accession AF105220.1) genomes using HISAT2-2.0.4 (64). The quantification of gene expression was calculated using HTSeq-counts (65). Transcripts with fewer than 100 total reads across the eight samples were excluded. The sheep annotation (GTF) was obtained from Ensembl (Oar_v3.1.79). The virus and sheep gene counts were imported into the edgeR package (66) and counts were normalized using a trimmed mean of M-values (67) and fitted to a negative binomial generalized log-linear model to calculate the dispersion factor for each gene (68). Differentially expressed genes were then identified by applying a FDR cutoff of 0.05 (69). Principal component analysis of normalized counts was performed using only sheep gene expression (i.e., counts from viral genes were removed) in order to see the variation between the two groups.

Gene function annotation and pathway analysis.

Identification of enriched KEGG pathways in the upregulated and downregulated gene lists was performed with DAVID (Database for Annotation, Visualization, and Integrated Discovery) v6.8 (26). Pathway analysis was performed using Ingenuity Pathway Analysis (IPA), version 01-04 (27) to infer the functional roles and relationships of the differentially expressed genes based on the log2 fold-change value of each gene.

Comparative analysis with human datasets.

To compare significantly-expressed genes in JSRV-infected lung with data from human NSCLC, HTSeq-counts from RNA-Seq data for LUAD and LUSC were obtained from the GDC portal (portal.gdc.cancer.gov). 57 LUAD and 49 LUSC patients were selected for which corresponding normal and tumor data were available. Differentially expressed genes in human tumors were identified using the edgeR pipeline as for the sheep data. The list of differentially expressed genes in human data was filtered for only those that have an orthologous gene in sheep (identified using Ensembl Biomart) and were differentially expressed in sheep. The comparison of expression of these short-listed genes was made using the correlation between expression profile (i.e., fold-change) in humans and in sheep. The gene lists obtained are presented in Supplementary Data S5.

RT-qPCR

RT-qPCR was performed according to MIQE recommendations (70). Primers and probes for reference and target genes are summarized in Table 2. RT-qPCR was performed using an ABI 7000 Sequence Detection System in 96-well plates (Applied Biosystems) with either TaqMan

one-step RT-PCR reagents (Applied Biosystems; JSRV, CCL2) or Power SYBR-Green RNA-to CT 1-Step Kit (Applied Biosystems; all other target genes). Each sample (5 sites per animal) was tested in duplicate using 100 ng of RNA in a 20 µl final reaction volume. All experiments with SYBR-green included a melting curve analysis to confirm the specificity of the amplicons (95°C for 15 s, 60°C for 20 s and 95°C for 15 s). Standard curves constructed from 10-fold serial dilutions of positive control RNA were used to determine efficiency and replicate quality (R2) of each primer set. In addition, the level of gene expression in experimental samples was ensured to lay within the limits of the standard curve. For comparison of RNA transcription levels between samples, results from the RT-qPCR experiments were normalized to two reference genes, succinate dehydrogenase (SDHA) and β-actin (ACTB) (pre-determined as stable reference genes using geNORM (71)). Statistical analysis for quantitative PCR was performed by group-wise comparison based on PCR efficiencies and the mean crossing point deviation between the sample and control group using Relative Expression Software Tool (REST) (72).

Immunohistochemistry and immunofluorescence

Tissue sections were processed routinely through graded alcohols, embedded in paraffin-wax and IHC was performed on sections (4 µm) mounted on charged glass microscope slides as previously described (9). The dilutions and sources of the primary antibodies used are described in Table 3. Isotype controls were used in semi-serial tissue sections for each primary antibody. Additionally, the primary antibody was omitted to check for nonspecific labeling by the secondary antibody or the visualization system. Images for bright-field microscopy were examined using an Olympus BX51 microscope, and photographs were captured with an

Olympus DP70 camera with analySIS software (Soft Imaging System GmbH, Munster, Germany). Immunofluorescence was performed as previously described (9) using primary antibodies to AGR2 (1:100, Abcam ab76473), and JSRV Env (1:50) (73) with appropriate secondary antibodies conjugated with Alexa-488 (A11008, Molecular Probes) or Alexa-555 (A31622, Molecular Probes). Slides were mounted with medium containing 4′,6-diamidino-2-phenylindole (DAPI) (Vectashield, Vector Laboratories). Images were analyzed using a Zeiss Axio Imager 2 fluorescence microscope with Apotome, and AxioVision Software.

Data Availability

The raw RNA-Seq reads (fastq data) of each sample are present in the European Nucleotide Archive with the accession ID PRJEB27638.

ACKNOWLEDGEMENTS

We thank Edinburgh Genomics (https://genomics.ed.ac.uk/) for conducting the RNA-Seq experiments and Sarah Wootton for providing the anti-JSRV SU antibody. We thank staff of the Moredun Research Institute Clinical Division for exceptional animal care and the farmers who support our work through the donation of OPA-affected sheep. This project was supported by the Biotechnology and Biological Sciences Research Council (BB/L009129/1, BB/L008505/1), including institute strategic program and national capability awards to The Roslin Institute (BB/P013740/1, BB/P013759/1, BB/P013732/1, BB/J004235/1, BB/J004243/1, BBS/E/D/20002173, BBS/E/D/20002174) and the Scottish Government Rural and Environment Science and Analytical Services Division (RESAS).

570 **REFERENCES**

571 1. De Las Heras M, González L, Sharp JM. 2003. Pathology of ovine pulmonary
572 adenocarcinoma. Curr Top Microbiol Immunol. 275:25-54.
573 2. Griffiths DJ, Martineau HM, Cousens C. 2010. Pathology and pathogenesis of ovine
574 pulmonary adenocarcinoma. J Comp Pathol 142:260-283.
575 3. Cousens C, Thonur L, Imlach S, Crawford J, Sales J, Griffiths DJ. 2009. Jaagsiekte sheep
576 retrovirus is present at high concentration in lung fluid produced by ovine pulmonary
577 adenocarcinoma-affected sheep and can survive for several weeks at ambient
578 temperatures. Res Vet Sci 87:154-156.
579 4. Palmarini M, Sharp JM, de las Heras M, Fan H. 1999. Jaagsiekte sheep retrovirus is
580 necessary and sufficient to induce a contagious lung cancer in sheep. J Virol 73:6964-
581 6972.
582 5. Palmarini M, Mura M, Spencer TE. 2004. Endogenous betaretroviruses of sheep:
583 teaching new lessons in retroviral interference and adaptation. J Gen Virol 85:1-13.
584 6. Spencer TE, Mura M, Gray CA, Griebel PJ, Palmarini M. 2003. Receptor usage and fetal
585 expression of ovine endogenous betaretroviruses: implications for coevolution of
586 endogenous and exogenous retroviruses. J Virol 77:749-753.
587 7. Summers C, Norval M, De Las Heras M, Gonzalez L, Sharp JM, Woods GM. 2005. An
588 influx of macrophages is the predominant local immune response in ovine pulmonary
589 adenocarcinoma. Vet Immunol Immunopathol 106:285-294.
590 8. De las Heras M, de Martino A, Borobia M, Ortin A, Alvarez R, Borderias L, Gimenez-
591 Mas JA. 2014. Solitary tumours associated with Jaagsiekte retrovirus in sheep are
592 heterogeneous and contain cells expressing markers identifying progenitor cells in lung
593 repair. J Comp Pathol 150:138-147.
594 9. Martineau HM, Cousens C, Imlach S, Dagleish MP, Griffiths DJ. 2011. Jaagsiekte sheep
595 retrovirus infects multiple cell types in the ovine lung. J Virol 85:3341-3355.
596 10. Murgia C, Caporale M, Ceesay O, Di Francesco G, Ferri N, Varasano V, de las Heras M,
597 Palmarini M. 2011. Lung adenocarcinoma originates from retrovirus infection of
598 proliferating type 2 pneumocytes during pulmonary post-natal development or tissue
599 repair. PLoS Pathog 7:e1002014.

11. Maeda N, Palmarini M, Murgia C, Fan H. 2001. Direct transformation of rodent fibroblasts by jaagsiekte sheep retrovirus DNA. Proc Natl Acad Sci U S A 98:4449-4454.
12. Rai SK, Duh FM, Vigdorovich V, Danilkovitch-Miagkova A, Lerman MI, Miller AD. 2001. Candidate tumor suppressor HYAL2 is a glycosylphosphatidylinositol (GPI)-anchored cell-surface receptor for jaagsiekte sheep retrovirus, the envelope protein of which mediates oncogenic transformation. Proc Natl Acad Sci U S A 98:4443-4438.
13. Caporale M, Cousens C, Centorame P, Pinoni C, De las Heras M, Palmarini M. 2006. Expression of the jaagsiekte sheep retrovirus envelope glycoprotein is sufficient to induce lung tumors in sheep. J Virol 80:8030-8037.
14. Cousens C, Maeda N, Murgia C, Dagleish MP, Palmarini M, Fan H. 2007. In vivo tumorigenesis by Jaagsiekte sheep retrovirus (JSRV) requires Y590 in Env TM, but not full-length orfX open reading frame. Virology 367:413-421.
15. Linnerth-Petrik NM, Santry LA, Yu DL, Wootton SK. 2012. Adeno-associated virus vector mediated expression of an oncogenic retroviral envelope protein induces lung adenocarcinomas in immunocompetent mice. PLoS One 7:e51400.
16. Wootton SK, Halbert CL, Miller AD. 2005. Sheep retrovirus structural protein induces lung tumours. Nature 434:904-907.
17. Hofacre A, Fan H. 2010. Jaagsiekte sheep retrovirus biology and oncogenesis. Viruses 2:2618-2648.
18. Liu SL, Miller AD. 2007. Oncogenic transformation by the jaagsiekte sheep retrovirus envelope protein. Oncogene 26:789-801.
19. Danilkovitch-Miagkova A, Duh FM, Kuzmin I, Angeloni D, Liu SL, Miller AD, Lerman MI. 2003. Hyaluronidase 2 negatively regulates RON receptor tyrosine kinase and mediates transformation of epithelial cells by jaagsiekte sheep retrovirus. Proc Natl Acad Sci U S A 100:4580-4585.
20. Hsu T, Phung A, Choe K, Kim JW, Fan H. 2015. Role for a zinc finger protein (Zfp111) in transformation of 208F rat fibroblasts by jaagsiekte sheep retrovirus envelope protein. J Virol 89:10453-10466.
21. Monot M, Erny A, Gineys B, Desloire S, Dolmazon C, Aublin-Gex A, Lotteau V, Archer F, Leroux C. 2015. Early steps of jaagsiekte sheep retrovirus-mediated cell

transformation involve the interaction between Env and the RALBP1 cellular protein. J Virol 89:8462-8473.

22. Youssef G, Wallace WA, Dagleish MP, Cousens C, Griffiths DJ. 2015. Ovine pulmonary adenocarcinoma: a large animal model for human lung cancer. ILAR J 56:99-115.

23. Palmarini M, Fan H. 2001. Retrovirus-induced ovine pulmonary adenocarcinoma, an animal model for lung cancer. J Natl Cancer Inst 93:1603-1614.

24. Mornex JF, Thivolet F, De las Heras M, Leroux C. 2003. Pathology of human bronchioloalveolar carcinoma and its relationship to the ovine disease. Curr Top Microbiol Immunol 275:225-248.

25. Perk K, Hod I. 1982. Sheep lung carcinoma: an endemic analogue of a sporadic human neoplasm. J Natl Cancer Inst 69:747-749.

26. Huang DW, Sherman BT, Lempicki RA. 2009. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc 4:44-57.

27. Kramer A, Green J, Pollard J, Jr., Tugendreich S. 2014. Causal analysis approaches in Ingenuity Pathway Analysis. Bioinformatics 30:523-530.

28. Gomes M, Archer F, Girard N, Gineys B, Dolmazon C, Bobet Erny A, Mornex JF, Leroux C. 2017. Blocked expression of key genes of the angiogenic pathway in JSRV-induced pulmonary adenocarcinomas. Vet Res 48:76.

29. Brychtova V, Mohtar A, Vojtesek B, Hupp TR. 2015. Mechanisms of anterior gradient-2 regulation and function in cancer. Semin Cancer Biol 33:16-24.

30. Dong A, Gupta A, Pai RK, Tun M, Lowe AW. 2011. The human adenocarcinoma-associated gene, AGR2, induces expression of amphiregulin through Hippo pathway co-activator YAP1 activation. J Biol Chem 286:18301-18310.

31. Chung C, Kim T, Kim M, Kim M, Song H, Kim TS, Seo E, Lee SH, Kim H, Kim SK, Yoo G, Lee DH, Hwang DS, Kinashi T, Kim JM, Lim DS. 2013. Hippo-Foxa2 signaling pathway plays a role in peripheral lung maturation and surfactant homeostasis. Proc Natl Acad Sci U S A 110:7732-7737.

32. Shaw AE, Hughes J, Gu Q, Behdenna A, Singer JB, Dennis T, Orton RJ, Varela M, Gifford RJ, Wilson SJ, Palmarini M. 2017. Fundamental properties of the mammalian innate immune system revealed by multispecies comparison of type I interferon responses. PLoS Biol 15:e2004086.

33. Fulton RB, Meyerholz DK, Varga SM. 2010. Foxp3+ CD4 regulatory T cells limit pulmonary immunopathology by modulating the CD8 T cell response during respiratory syncytial virus infection. J Immunol 185:2382-2392.

34. Sozmen M, Beytut E. 2012. An investigation of growth factors and lactoferrin in naturally occurring ovine pulmonary adenomatosis. J Comp Pathol 147:441-451.

35. Komohara Y, Hirahara J, Horikawa T, Kawamura K, Kiyota E, Sakashita N, Araki N, Takeya M. 2006. AM-3K, an anti-macrophage antibody, recognizes CD163, a molecule associated with an anti-inflammatory macrophage phenotype. J Histochem Cytochem 54:763-771.

36. Chang JT, Lee YM, Huang RS. 2015. The impact of the Cancer Genome Atlas on lung cancer. Transl Res 166:568-585.

37. Kim HK, Kim J, Korolevich S, Choi IJ, Kim CH, Munroe DJ, Green JE. 2011. Distinctions in gastric cancer gene expression signatures derived from laser capture microdissection versus histologic macrodissection. BMC Med Genomics 4:48.

38. Klee EW, Erdogan S, Tillmans L, Kosari F, Sun Z, Wigle DA, Yang P, Aubry MC, Vasmatzis G. 2009. Impact of sample acquisition and linear amplification on gene expression profiling of lung adenocarcinoma: laser capture micro-dissection cell-sampling versus bulk tissue-sampling. BMC Med Genomics 2:13.

39. De Las Heras M, Ortin A, Benito A, Summers C, Ferrer LM, Sharp JM. 2006. In-situ demonstration of mitogen-activated protein kinase Erk 1/2 signalling pathway in contagious respiratory tumours of sheep and goats. J Comp Pathol 135:1-10.

40. Cousens C, Alleaume C, Bijsmans E, Martineau HM, Finlayson J, Dagleish MP, Griffiths DJ. 2015. Jaagsiekte sheep retrovirus infection of lung slice cultures. Retrovirology 12:31.

41. Varela M, Golder M, Archer F, de las Heras M, Leroux C, Palmarini M. 2008. A large animal model to evaluate the effects of Hsp90 inhibitors for the treatment of lung adenocarcinoma. Virology 371:206-215.

42. Vaughan AE, Halbert CL, Wootton SK, Miller AD. 2012. Lung cancer in mice induced by the jaagsiekte sheep retrovirus envelope protein is not maintained by rare cancer stem cells, but tumorigenicity does correlate with Wnt pathway activation. Mol Cancer Res 10:86-95.

43. Arthur JS, Ley SC. 2013. Mitogen-activated protein kinases in innate immunity. Nat Rev Immunol 13:679-692.

44. Collins DC, Chenard-Poirier M, Lopez JS. 2018. The PI3K pathway at the crossroads of cancer and the immune system: strategies for next generation immunotherapy combinations. Curr Cancer Drug Targets 18:355-364.

45. Dumartin L, Whiteman HJ, Weeks ME, Hariharan D, Dmitrovic B, Iacobuzio-Donahue CA, Brentnall TA, Bronner MP, Feakins RM, Timms JF, Brennan C, Lemoine NR, Crnogorac-Jurcevic T. 2011. AGR2 is a novel surface antigen that promotes the dissemination of pancreatic cancer cells through regulation of cathepsins B and D. Cancer Res 71:7091-7102.

46. Fessart D, Domblides C, Avril T, Eriksson LA, Begueret H, Pineau R, Malrieux C, Dugot-Senant N, Lucchesi C, Chevet E, Delom F. 2016. Secretion of protein disulphide isomerase AGR2 confers tumorigenic properties. Elife 5:e13887.

47. Hrstka R, Bouchalova P, Michalova E, Matoulkova E, Muller P, Coates PJ, Vojtesek B. 2016. AGR2 oncoprotein inhibits p38 MAPK and p53 activation through a DUSP10-mediated regulatory pathway. Mol Oncol 10:652-662.

48. Guo H, Zhu Q, Yu X, Merugu SB, Mangukiya HB, Smith N, Li Z, Zhang B, Negi H, Rong R, Cheng K, Wu Z, Li D. 2017. Tumor-secreted anterior gradient-2 binds to VEGF and FGF2 and enhances their activities by promoting their homodimerization. Oncogene 36:5098-5109.

49. Pizzi M, Fassan M, Balistreri M, Galligioni A, Rea F, Rugge M. 2012. Anterior gradient 2 overexpression in lung adenocarcinoma. Appl Immunohistochem Mol Morphol 20:31-36.

50. Chang HW, Lin Z-M, Wu M-J, Wang L-Y, Chow Y-H, Jiang SS, Ch'ang H-J, Chang VHS. 2017. Characterization of a transgenic mouse model exhibiting spontaneous lung adenocarcinomas with a metastatic phenotype. PLoS One 12:e0175586.

51. Agrawal P, Nawadkar R, Ojha H, Kumar J, Sahu A. 2017. Complement evasion strategies of viruses: an overview. Front Microbiol 8:1117.

52. Reis ES, Mastellos DC, Ricklin D, Mantovani A, Lambris JD. 2018. Complement in cancer: untangling an intricate relationship. Nat Rev Immunol 18:5-18.

53. Okroj M, Hsu Y-F, Ajona D, Pio R, Blom AM. 2008. Non-small cell lung cancer cells produce a functional set of complement factor I and its soluble cofactors. Molecular Immunol 45:169-179.

54. Markiewski MM, DeAngelis RA, Benencia F, Ricklin-Lichtsteiner SK, Koutoulaki A, Gerard C, Coukos G, Lambris JD. 2008. Modulation of the antitumor immune response by complement. Nat Immunol 9:1225-1235.

55. Medler TR, Murugan D, Horton W, Kumar S, Cotechini T, Forsyth AM, Leyshock P, Leitenberger JJ, Kulesz-Martin M, Margolin AA, Werb Z, Coussens LM. 2018. Complement C5a fosters squamous carcinogenesis and limits T cell response to chemotherapy. Cancer Cell 34:561-578 e6.

56. Qian B, Pollard JW. 2010. Macrophage diversity enhances tumor progression and metastasis. Cell 141:39-51.

57. Zilionis R, Engblom C, Pfirschke C, Savova V, Zemmour D, Saatcioglu HD, Krishnan I, Maroni G, Meyerovitz CV, Kerwin CM, Choi S, Richards WG, De Rienzo A, Tenen DG, Bueno R, Levantini E, Pittet MJ, Klein AM. 2019. Single-cell transcriptomics of human and mouse lung cancers reveals conserved myeloid populations across individuals and species. Immunity 50: 1317-1334.

58. Kitamura T, Qian BZ, Pollard JW. 2015. Immune cell promotion of metastasis. Nat Rev Immunol 15:73-86.

59. Lewis CE, Pollard JW. 2006. Distinct role of macrophages in different tumor microenvironments. Cancer Res 66:605-612.

60. Medrek C, Ponten F, Jirstrom K, Leandersson K. 2012. The presence of tumor associated macrophages in tumor stroma as a prognostic marker for breast cancer patients. BMC Cancer 12:306.

61. Yang L, Wang F, Wang L, Huang L, Wang J, Zhang B, Zhang Y. 2015. CD163+ tumor-associated macrophage is a prognostic biomarker and is associated with therapeutic effect on malignant pleural effusion of lung cancer patients. Oncotarget 6:10592-10603.

62. Salaun B, de Saint-Vis B, Pacheco N, Pacheco Y, Riesler A, Isaac S, Leroux C, Clair-Moninot V, Pin JJ, Griffith J, Treilleux I, Goddard S, Davoust J, Kleijmeer M, Lebecque S. 2004. CD208/dendritic cell-lysosomal associated membrane protein is a marker of normal and transformed type II pneumocytes. Am J Pathol 164:861-871.

63. Martin M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnetjournal 17:10-12.
64. Kim D, Langmead B, Salzberg SL. 2015. HISAT: a fast spliced aligner with low memory requirements. Nat Methods 12:357-360.
65. Anders S, Pyl PT, Huber W. 2015. HTSeq--a Python framework to work with high-throughput sequencing data. Bioinformatics 31:166-169.
66. Robinson MD, McCarthy DJ, Smyth GK. 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26:139-140.
67. Robinson MD, Oshlack A. 2010. A scaling normalization method for differential expression analysis of RNA-seq data. Genome Biology 11:R25-R25.
68. McCarthy DJ, Chen Y, Smyth GK. 2012. Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. Nucleic Acids Res 40:4288-4297.
69. Lun AT, Chen Y, Smyth GK. 2016. It's DE-licious: a recipe for differential expression analyses of RNA-seq experiments using quasi-likelihood methods in edgeR. Methods Mol Biol 1418:391-416.
70. Bustin SA, Benes V, Garson JA, Hellemans J, Huggett J, Kubista M, Mueller R, Nolan T, Pfaffl MW, Shipley GL, Vandesompele J, Wittwer CT. 2009. The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. Clin Chem 55:611-622.
71. Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A, Speleman F. 2002. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. Genome Biol 3:RESEARCH0034.
72. Pfaffl MW, Horgan GW, Dempfle L. 2002. Relative expression software tool (REST) for group-wise comparison and statistical analysis of relative expression results in real-time PCR. Nucleic Acids Res 30:e36.
73. Wootton SK, Metzger MJ, Hudkins KL, Alpers CE, York D, DeMartini JC, Miller AD. 2006. Lung cancer induced in mice by the envelope protein of jaagsiekte sheep retrovirus (JSRV) closely resembles lung cancer in sheep infected with JSRV. Retrovirology 3:94.

783 74. Garcia-Crespo D, Juste RA, Hurtado A. 2006. Differential gene expression in central
784 nervous system tissues of sheep with natural scrapie. Brain Res 1073-1074:88-92.
785 75. Li H, Cunha CW, Davies CJ, Gailbreath KL, Knowles DP, Oaks JL, Taus NS. 2008.
786 Ovine herpesvirus 2 replicates initially in the lung of experimentally infected sheep. J
787 Gen Virol 89:1699-1708.
788 76. Budhia S, Haring LF, McConnell I, Blacklaws BA. 2006. Quantitation of ovine cytokine
789 mRNA by real-time RT-PCR. J Immunol Methods 309:160-172.
790 77. McNeilly TN, Baker A, Brown JK, Collie D, Maclachlan G, Rhind SM, Harkiss GD.
791 2008. Role of alveolar macrophages in respiratory transmission of visna/maedi virus. J
792 Virol 82:1526-1536.
793 78. Leutenegger CM, Alluwaimi AM, Smith WL, Perani L, Cullor JS. 2000. Quantitation of
794 bovine cytokine mRNA in milk cells of healthy cattle by real-time TaqMan polymerase
795 chain reaction. Vet Immunol Immunopathol 77:275-287.
796 79. Kawashima K, Meyerholz DK, Gallup JM, Grubor B, Lazic T, Lehmkuhl HD,
797 Ackermann MR. 2006. Differential expression of ovine innate immune genes by preterm
798 and neonatal lung epithelia infected with respiratory syncytial virus. Viral Immunol
799 19:316-323.
800 80. Dunphy J, Horvath A, Barcham G, Balic A, Bischof R, Meeusen E. 2001. Isolation,
801 characterisation and expression of mRNAs encoding the ovine CC chemokines,
802 monocyte chemoattractant protein (MCP)-1alpha and -2. Vet Immunol Immunopathol
803 82:153-164.
804 81. Clark EL, Bush SJ, McCulloch MEB, Farquhar IL, Young R, Lefevre L, Pridans C,
805 Tsang HG, Wu C, Afrasiabi C, Watson M, Whitelaw CB, Freeman TC, Summers KM,
806 Archibald AL, Hume DA. 2017. A high resolution atlas of gene expression in the
807 domestic sheep (Ovis aries). PLoS Genet 13:e1006997.

808

809
810 **Table 1 Read mapping statistics**

811

	Lamb ID [a]							
	Infected 1M_66days	Infected 2F_85days	Infected 3M_71days	Infected 4M_85days	Control 5M_66days	Control 6M_85days	Control 7M_71days	Control 8F_85days
Total reads	64,879,262	72,907,048	84,205,226	68,631,006	60,417,140	69,875,636	64,020,136	72,061,316
Total quality trimmed	63,766,561	71,678,589	82,667,841	67,438,835	59,431,332	68,705,449	62,964,748	70,830,855
Quality trimmed (%)	98.28%	98.32%	98.17%	98.26%	98.37%	98.33%	98.35%	98.29%
Total uniquely mapped	51,720,272	58,431,025	67,009,475	55,039,775	48,778,008	55,995,176	51,447,365	58,063,922
Uniquely mapped (%)	79.72%	80.14%	79.58%	80.20%	80.74%	80.14%	80.36%	80.58%
Viral reads	109,952	68,763	342,032	196,608	54[b]	62[b]	128[b]	49[b]
Viral reads (%)	0.17%	0.09%	0.41%	0.29%	0.00%	0.00%	0.00%	0.00%

812
813
814 [a] Lamb ID indicates the infection status, sex (male (M) or female (F)) and days post-inoculation
815 when culled.
816 [b] Reads mapping to the JSRV genome represented 0.09 – 0.41% of total reads in the four
817 infected lambs. A small number of reads mapping to JSRV (49 – 128 per sample) were also
818 detected in the tissue from mock-infected control lambs. All of these reads mapped to regions of
819 very high similarity between JSRV and enJSRV (data not shown) and so can be attributed to the
820 transcription of endogenous viruses in the samples studied.

Table 2 Primers and probes used in this study

Target[a]	Primer sequences (5'-3')	Conc. (nM)	Size (bp)	GenBank Accession No. (Reference)
ACTB	F: CTGAGCGCAAGTACTCCGTGT R: GCATTTGCGGTGGACGAT	300 300	125	NM_001009784 (74)
SDHA	F: CATCCACTACATGACGGAGCA R: ATCTTGCCATCTTCAGTTCTGCTA	200 200	90	AY970969 (74)
IFN-α	F: GCACTGGATCAGCAGCTCACTG R: CTCAAGACTTCTGCTCTGACAACCT	200 200	188	AY 802984 (75)
IFN-γ	F: TTCTTGAACGGCAGCTCTGAG R: TGGCGACAGGTCATTCATCA	300 300	127	X52640 (76)
CSF2	F: GATGGATGAAACAGTAGAAGTCG R: CAGCAGTCAAAGGGAATGAT	500 500	261	NM_001009805 (77)
TGF-β1	F: GAACTGCTGTGTTCGTCAGC R: GGTTGTGCTGGTTGTACAGG	500 500	169	NM_001009400 (77)
TNF-α	F: GCCCTGGTACGAACCCATCTA R: CGGCAGGTTGATCTCAGCAC	200 200	82	NM_001024860 (75)
IL-1β	F: CCTAACTGGTACATCAGCACTTCTCA R:TCCATTCTGAAGTCAGTTATATCCTG	200 200	95	NM_001009465 (75)
IL-6	F: TCCAGAACGAGTTTGAGG R: CATCCGAATAGCTCTCAG	500 500	236	NM_ 001009392 (77)
IL-8	F: ACTGCGAAAATTCAGAAATCATTGTTA R: CTTCAAAAATGCCTGCACAACCTTC	500 500	53	S74436 (78)
IL-10	F: AGCAAGGCGGTGGAGCAG R: GATGAAGATGTCAAACTCACTCATGG	200 200	90	NM_001009327 (75)
IL-18	F: ACTGTTCAGATAATGCACCCCAG R: TTCTTACACTGCACAGAGATGGTTAC	200 200	100	NM_001009263 (75)
CCL2	F: GCTGTGATTTTCAAGACCATCCT R: GGCGTCCTGGACCCATT Probe: AAAGAGTTTTGTGCAGACCCCAACC	900 400 900	72	DY503036 (79, 80)

[a] Amplification conditions were 30 min 45°C, 10 min 95°C, 40 cycles of (15s 95°C, 1 min 60 °C), except for CSF2 (30 min 45°C, 10 min 95°C, 40 cycles of (20s 94°C, 30s 57°C, 30s 72°C)), CXCL8 (30 min 45°C, 10 min 95°C, 40 cycles of (20s 94°C, 30s 55°C, 30s 72°C)), TGF-beta (30 min 45°C, 10 min 95°C, 40 cycles of (20s 94°C, 30s 55°C, 30s 72°C)), IL-6 (30 min 45°C, 10 min 95°C, 40 cycles of (20s 94°C, 30s 52°C, 30s 72°C)) and CCL2 (30 min 45°C, 10 min 95°C, 40 cycles of (20s 95°C, 30s 58°C, 30s 60°C)).

831 **Table 3 Summary of antibodies used for IHC**

832

Target Antigen	IHC Dilution	Antibody type[a]	Source
JSRV Env (SU)	1/200	Mouse mab	(73)
AGR2	1/200	Rabbit mab	Abcam, ab134167
AREG	1/200	Rabbit mab	Abcam, ab224350
MST1/2	1/5000	Rabbit pab	Abcam, ab87322
LATS1/2	1/2000	Rabbit pab	Abcam, ab70565
Total YAP1	1/100	Rabbit mab	Abcam, ab52771
Phosphorylated (s127) YAP1	1/300	Rabbit pab	Abcam, ab76252
CD68	1/150	Mouse mab	DAKO, M0718
CD163	1/200	Mouse mab	Bio-rad, MCA1853
LGMN	1/400	Mouse mab	Abcam, ab125286
HIF1A	1/50	Mouse mab	ThermoFisher MA1-16511

833

834 [a]mab: monoclonal, pab; polyclonal

835

836 **FIGURE LEGENDS**

837

838 **Figure 1. Host gene expression is modified in JSRV-infected lung tissue.**

839 A. Principal component analysis was performed using read counts from JSRV-infected and
840 mock-infected (control) lambs for 15,149 genes mapping to the sheep genome. Each circle
841 represents a mock-infected lamb and each triangle indicates a JSRV-infected lamb. The values
842 on each axis (PC1 and PC2) represent the percentage of variance explained by each component.
843 The principal component 1 (PC1) separates the infected and mock samples into two clusters with
844 the highest variance of 63%. Within these clusters, there is greater variance in the JSRV-
845 infected sheep than the mock-infected sheep and this likely corresponds to the proportion of
846 tumor-affected tissue in the samples. For example, the lamb 'Infected_F_85days' is the closest
847 to the mock-infected cluster and had the lowest proportion of reads mapping to JSRV (see Table
848 1), while the samples furthest from the mock-infected animals had the highest proportion of viral
849 reads. B. The gene expression data visualized as a 2D scatter plot of the $\log_2$ ratio of expression
850 values between infected and mock-infected tissue (i.e., the fold-change) versus the mean
851 expression across all samples. Each dot represents one gene and the red color indicates the 1,971
852 genes identified as differentially expressed between the infected group and mock-infected
853 controls using a false discovery rate (FDR) < 0.05. C. Hierarchical clustering based on
854 normalized gene counts of differentially expressed genes in samples derived from JSRV-treated
855 and mock-infected lambs (1237 were upregulated and 734 downregulated). D. Hierarchical
856 clustering of the 25 most significantly upregulated and downregulated genes.

857

Figure 2. Main Diseases and Biofunctions associated with differentially expressed genes in JSRV-infected sheep lung.

The figure shows the Biofunction enrichment profiles identified by Ingenuity Pathway Analysis, plotted by relative statistical significance. Significance values were calculated based on a right-tailed Fisher's exact test and the -log(p-value) is displayed on the horizontal axis of the bar chart. The taller the bar, the more significant the pathway effect.

Figure 3. Mechanistic network analysis predicted by Ingenuity Pathway Analysis.

The figure shows mechanistic network analysis by IPA of genes differentially expressed in experimentally-induced OPA compared to mock-infected lung. A. Neovasculization network. B. Network for cell viability and tumor growth. The regulators are colored by their predicted activation state: activated (orange) or inhibited (blue). Darker colors indicate higher scores. The edges connecting the nodes are colored orange when leading to activation of the downstream node, blue when leading to its inhibition, and yellow if the findings underlying the relationship are inconsistent with the state of the downstream node. Pointed arrowheads indicate that the downstream node is expected to be activated if the upstream node connected to it is activated, while blunt arrowheads indicate that the downstream node is expected to be inhibited if the upstream node that connects to it is activated.

Figure 4. Immunohistochemical detection of AGR2 and AREG in JSRV-infected cells.

Immunohistochemical labeling of sheep lung with antibodies to JSRV SU (A, B, C) and AGR2 (D, E, F). A and D, serial adjacent sections of mock-infected sheep lung. B and E, serial sections of OPA lesions in the lung of a lamb experimentally-infected with JSRV. C and F, serial sections of lung tissue from a natural case of OPA. Brown pigment indicates positive labeling. G, H, I, J,

882 immunofluorescence labeling of lung tissue of an experimentally-infected lamb. Panel H, AGR2
883 (green); panel I, JSRV (red). Panel G shows staining with DAPI (blue) used to visualize nuclei
884 and panel J shows the merged image. K, L and M, IHC labeling with an antibody to AREG. K,
885 mock-infected lung shows labeling of some epithelial cells (arrows) and macrophages
886 (arrowheads). L, JSRV experimentally-infected lamb and M, naturally OPA-affected sheep lung
887 show labeling of OPA lesions.
888
889 **Figure 5. Immunohistochemical detection of Hippo pathway regulators in OPA-affected**
890 **lung.**
891 IHC was performed on sections of ovine lung tissue with antibodies to proteins involved in
892 Hippo pathway signaling. Left column, mock-infected sheep lung; middle column,
893 experimentally-infected lamb; right column, natural OPA. Brown pigment indicates positive
894 labeling. A-C, total YAP1; D-F, phosphorylated (inactive) YAP1 (P-YAP1); G-I, MST1/2; J-L,
895 LATS1/2. Note P-YAP expression is localized in the cytoplasm of tumor cells (panels E and F),
896 whereas total YAP1 expression is shown the nuclei and the cytoplasm of tumor cells (panels B
897 and C). Black arrows in panels B and C indicate examples of nuclear locating YAP1.
898
899 **Figure 6. RT-qPCR detection of cytokine expression in OPA.**
900 The figure shows box-whisker plots of relative expression of selected genes in OPA-affected
901 lung tissue compared to healthy control tissue measured by RT-qPCR. A. Experimentally-
902 induced OPA relative to lung tissue from age and sex matched mock-infected lambs. B. Natural
903 OPA relative to healthy adult sheep lung. Data analysis was performed using the software tool
904 REST (72), which calculates the significance of the expression ratio based on the use of two
905 reference genes and the amplification efficiency for each gene. Significant changes in expression

are indicated by asterisks (*; standard error <0.005 at P=0.05). The dotted line within each box represents the median value, the boxed area encompasses the interquartile range and the whiskers indicate the maximum and minimum data points. Note the difference in scales on the Y axis between the two charts. The data are consistent with the results of RNA-Seq analysis for these genes in experimentally-infected lambs and controls, where *CSF2*, *CCL2*, *CXCL8* and *IL6* were upregulated, *TNF* was downregulated and *IL1B*, *IL10*, *IL18* and *IFNG* were not significantly changed. *TGFB1* and *IFNA* were not detected by RNA-Seq although *TGFB3* was upregulated. (Note that *IFNA* was not examined in experimental OPA.)

Figure 7. Expression of macrophage-related genes is altered in OPA.

To identify macrophage-related functions in our data, a list of macrophage-related genes was compiled based on Gene Ontology (GO) terms (891 genes) and those present in macrophage coexpression clusters from the recently published sheep genome atlas (450 genes) (81) (see Supplementary Dataset S3). Combining these two lists gave a total of 1255 genes (86 were common to both sets), of which 1076 were present in our RNA-Seq data and 198 were differentially expressed in JSRV-infected lung tissue. A. Hierarchical clustering of 198 markers related to macrophage function that are differentially expressed in OPA. B. Hierarchical clustering of the 25 most significantly upregulated and downregulated macrophage-related genes in OPA. Orange coloring indicates upregulation, blue indicates downregulation.

Figure 8. Immunohistochemical detection of macrophage markers in OPA-affected lung tissue.

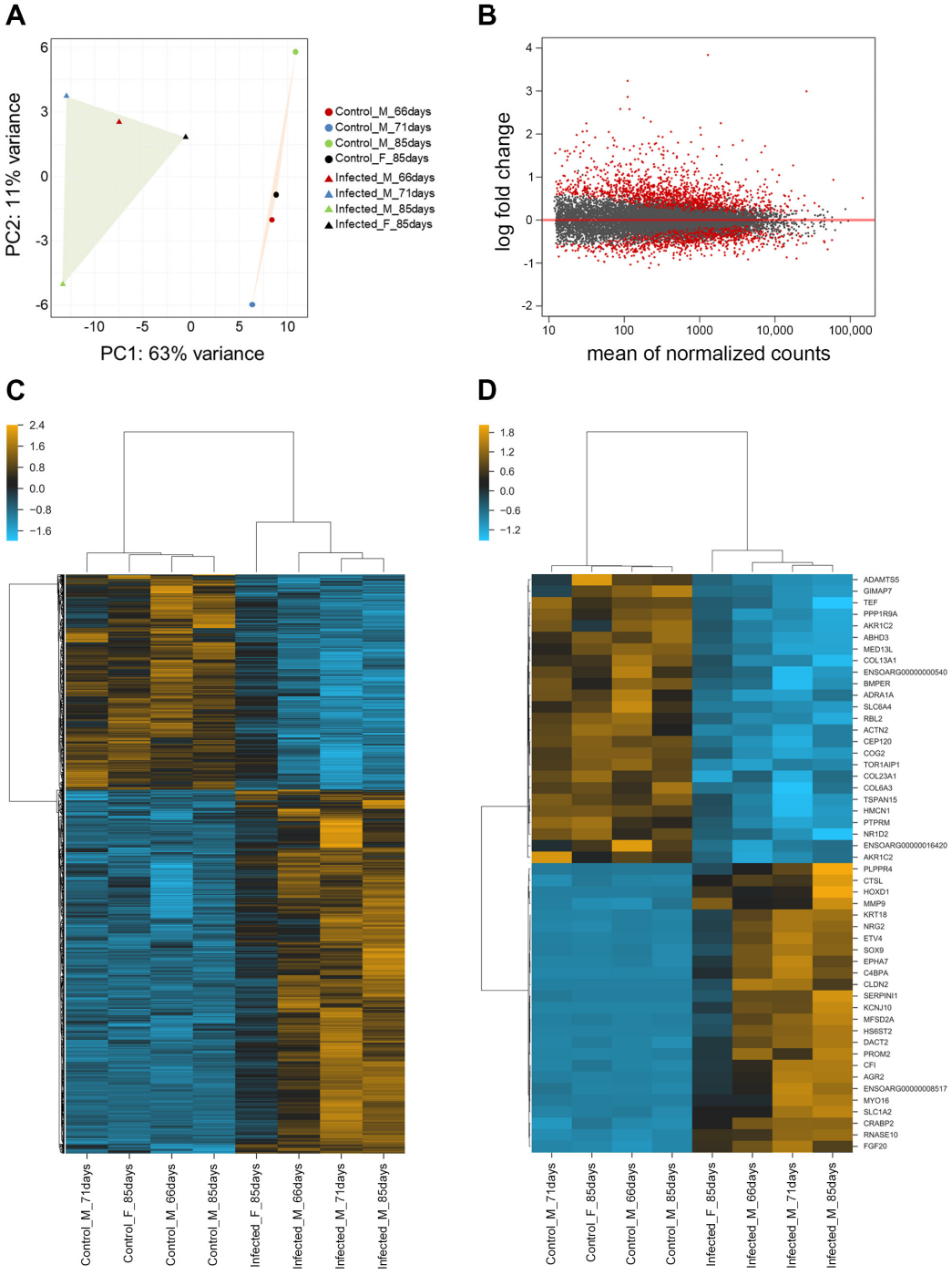
Immunohistochemistry of sheep lung labeled with antibodies to markers of macrophages. Left column, mock-infected lung; middle column, experimentally-infected lung; right hand column,

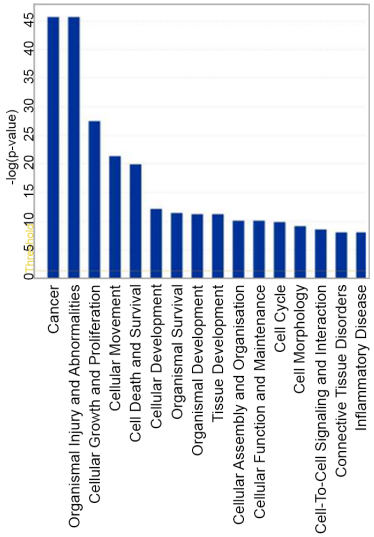
natural OPA-affected lung. Brown pigment indicates positive labeling. A-C, labeling with an antibody to CD68. D-F, serial sections of tissues shown in panels A-C labeled with an antibody to CD163. Note the difference in distribution of labeling with the two antibodies. G-I, labeling with an antibody to LGMN; J-L, labeling with an antibody to HIF1A. Note that antibodies to LGMN and HIF1A label some tumor cells in addition to macrophages. In addition, LGMN exhibits a different pattern of labeling of tumor cells in experimental and natural OPA, where experimental cases have intense cytoplasmic labeling and natural cases have prominent labeling of the apical region.

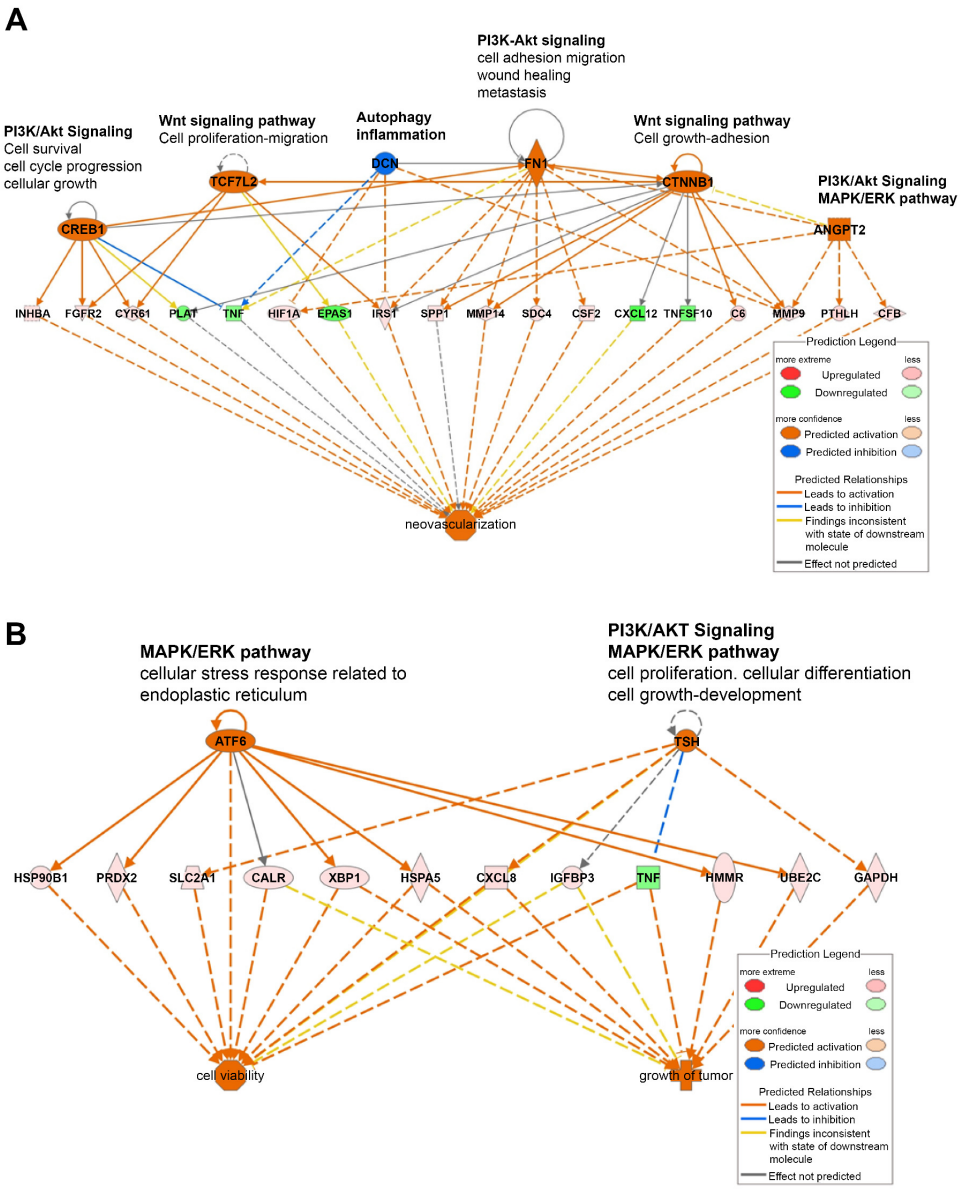
Figure 9. Comparison of gene expression in OPA and human lung cancer.

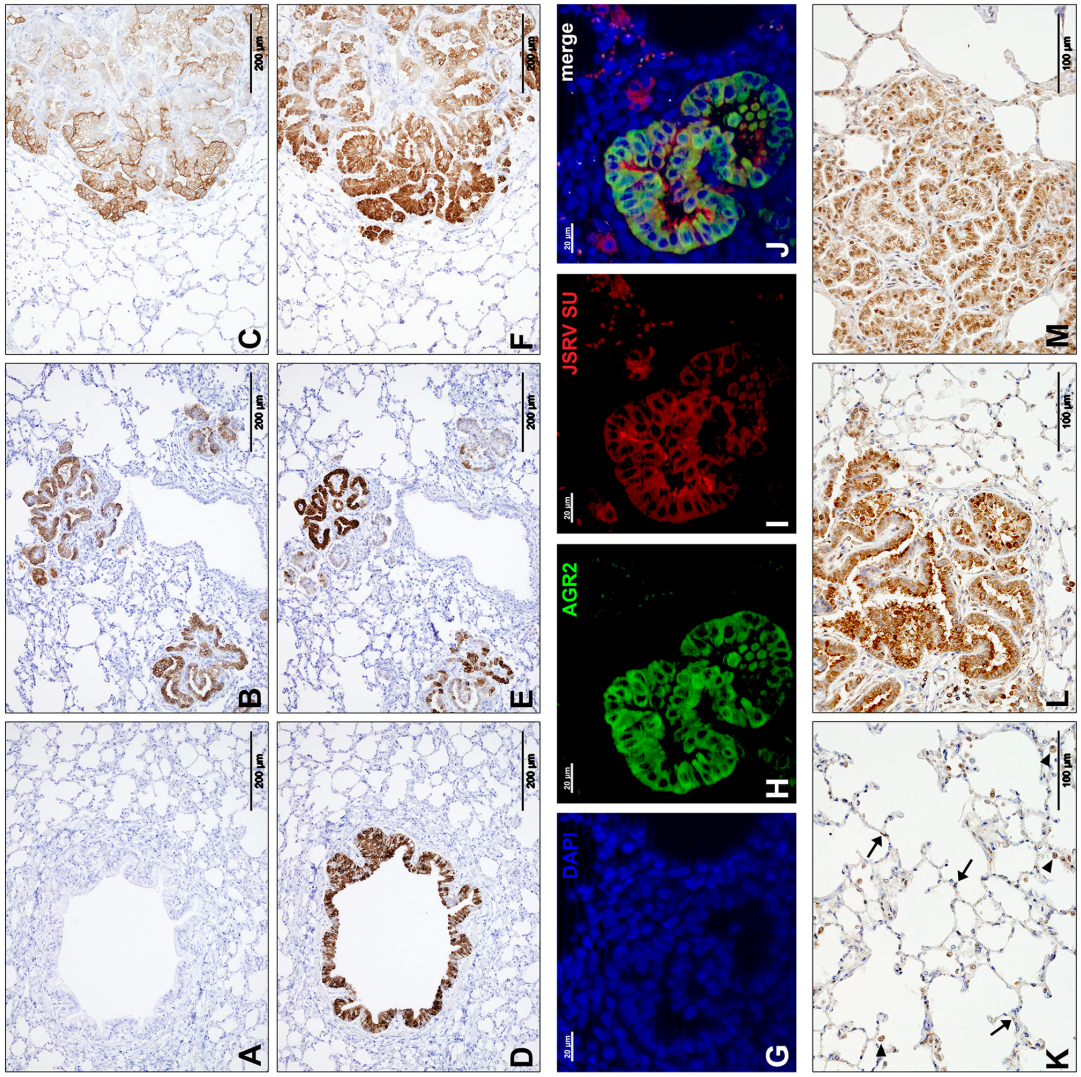
Genes differentially expressed in experimentally-induced OPA were compared with genes differentially expressed in human lung cancer, including clinical stages I - IV of LUAD and I - III of LUSC. (LUSC stage IV was not analyzed as there were only two samples present in the data.) A. Correlation plot comparing differentially expressed genes in OPA with the different clinical stages of human LUAD and LUSC. The height of each bar shows the percentage of differentially expressed sheep genes that have differentially expressed orthologues in the human data set. The shading indicates the Pearson correlation coefficient between the log ratios of genes in common between the two lists. The data confirm closer similarity of OPA to LUAD than LUSC and to stage I LUAD in particular. B. The differentially expressed orthologous genes in both sheep and human stage I LUAD were analyzed with Ingenuity Pathway Analysis software and the figure shows the enriched Diseases and Biofunctions associated with those genes. C. The plot shows the changes in gene expression in experimentally-induced OPA and human stage I LUAD for the 50 most significantly changed sheep genes with deregulated human orthologues. Note that 37 of the 50 genes show a similar direction of change in expression, whereas 13 show

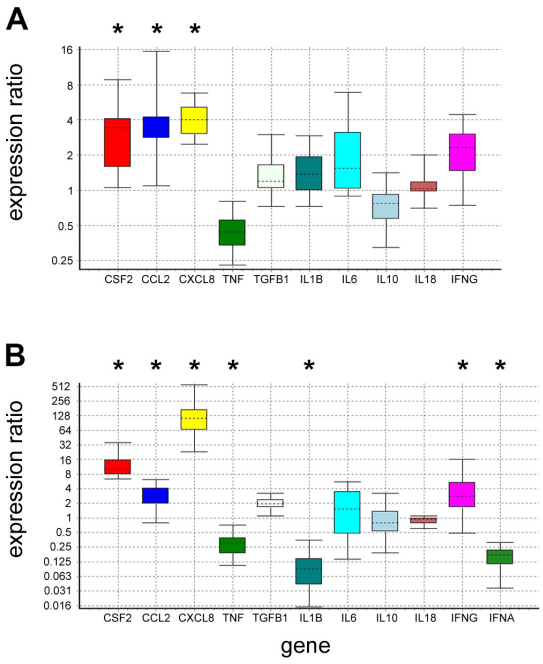
954 divergent changes, highlighting similarities and differences between the sheep and human
955 diseases.











Journal of Virology