



OPEN Mesothelioma cell heterogeneity identified by single cell RNA sequencing

Licun Wu^{1,2}, Zongjie Wang³, Amin Zia^{1,4}, Shana O. Kelley^{3,5,6,7,8,9,10,11} & Marc de Perrot^{1,2,12,13}✉

Mesothelioma cell heterogeneity encompasses diverse morphological and molecular characteristics observed within tumors, significantly impacting disease progression, treatment outcomes, and the development of targeted therapies. This heterogeneity has long posed challenges for accurate diagnosis and effective treatment, but understanding its complexities offers the potential for novel diagnostic modalities and therapeutic interventions. This study employed single-cell RNA sequencing (scRNA-seq) to investigate mesothelioma cell heterogeneity from various sources, including cell culture (CC), peritoneal lavage (Lav) from the tumor microenvironment, and circulating tumor cells (CTC) in murine models. Gene set enrichment analysis was used to identify distinct gene signatures for each subpopulation. The results revealed unique characteristics for mesothelioma cells depending on their origin. In the CC group, up-regulated genes were primarily involved in tumor cell cycle control, proliferation, and apoptosis. In the CTC group, up-regulated genes were associated with cancer cell stemness. The Lav group showed up-regulated genes facilitating interactions between tumor cells and the microenvironment, such as epithelial–mesenchymal transition and immune responses mediated by IFN- α and IFN- γ . Some pathways were shared among all tumor cells, suggesting the potential for transitioning between functional states under specific conditions. This may be the first study to explore circulating mesothelioma cell heterogeneity using scRNA-seq. The distinct gene signatures identified in each mesothelioma cell subpopulation likely play critical roles in tumor initiation and progression, offering potential novel targets for therapeutic intervention. These findings could help inform the development of more effective, personalized treatments for mesothelioma, ultimately improving patient outcomes.

Keywords Tumor cell heterogeneity, Single cell RNA sequencing (scRNA-seq), Mesothelioma (MESO), Gene set enrichment analysis, Signaling pathway

Pleural mesothelioma (PM) is a notorious cancer characterized by an escalating incidence and formidable clinical management challenges, often culminating in a grim prognosis¹. Recognized as heterogeneous both histologically and molecularly, MESO manifests diverse cell populations within tumors, a phenomenon termed tumor cell heterogeneity. Histological diversity in PM encompasses three primary types—epithelioid, sarcomatoid and biphasic morphologies, each demonstrating substantial associations with clinical outcomes². Mesothelioma cell heterogeneity embodies a multifaceted interplay of varied cellular subpopulations within tumors, characterized by disparities in genetic, morphological, and functional attributes^{3,4}. This inherent diversity poses formidable hurdles in devising effective diagnostic modalities and targeted therapeutic approaches for

¹Latner Thoracic Surgery Research Laboratories, Division of Thoracic Surgery, Toronto General Hospital, Toronto General Hospital Research Institute, University Health Network, University of Toronto, Toronto, ON M5G 1L7, Canada. ²Princess Margaret Cancer Centre, University Health Network, Toronto, ON M5G 1L7, Canada. ³Department of Biomedical Engineering, Northwestern University, Evanston, IL, USA. ⁴dYcode Inc., Toronto, ON L6C 2R9, Canada. ⁵Department of Chemistry, Weinberg College of Arts and Sciences, Northwestern University, Evanston, IL, USA. ⁶Department of Pharmaceutical Sciences, Leslie Dan Faculty of Pharmacy, University of Toronto, Toronto, ON, Canada. ⁷International Institute for Nanotechnology, Northwestern University, Evanston, IL, USA. ⁸Department of Biochemistry, Feinberg School of Medicine, Northwestern University, Chicago, IL, USA. ⁹Robert H. Lurie Comprehensive Cancer Center, Northwestern University, Chicago, IL, USA. ¹⁰Simpson Querrey Institute, Northwestern University, Chicago, IL, USA. ¹¹Chan Zuckerberg Biohub Chicago, Chicago, IL, USA. ¹²Department of Immunology, University of Toronto, Toronto, ON M5S 1A8, Canada. ¹³Division of Thoracic Surgery, Toronto General Hospital, 9N-961, 200 Elizabeth Street, Toronto, ON M5G 2C4, Canada. ✉email: marc.deperrot@uhn.ca

MESO. A comprehensive analysis of the nuances of mesothelioma cell heterogeneity holds significant potential for advancing personalized treatment paradigms and ultimately enhancing patient outcomes.

The tumor microenvironment plays a pivotal role in driving tumor progression, invasion, and metastasis across various cancers, including PM⁵. The intricate interplay between tumor cells and their surrounding microenvironment has been recognized for decades. Particularly, the epithelial–mesenchymal transition (EMT) process stands out as a key contributor to the dismal prognosis associated with mesothelioma. Through our previous investigations, we have identified a panel of EMT genes that appears to be unique to mesothelioma tumors⁶. Notably, the up-regulation of this specific EMT gene signature correlates strongly with diminished survival rates among PM patients⁷. These findings underscore the critical importance of elucidating the underlying molecular mechanisms to inform potential therapeutic strategies and prognostic assessments in PM.

There are a lot of studies that have been dedicated to exploring cancer cell heterogeneity including mesothelioma^{8–10}. Notably, a study identified 12 expression programs exhibiting heterogeneity across various cancer cell lines. These programs were associated with diverse biological processes, including the cell cycle, senescence, stress and interferon responses, EMT, and protein metabolism. This highlights the intricate molecular landscape underlying cancer cell heterogeneity^{8,11}. A recent study provides the first comprehensive single-cell transcriptomic atlas of the human parietal pleura, offering unprecedented resolution of its cellular composition, which identifies novel pleural-specific fibroblast subtypes, characterizes in vitro models of mesothelial cells, and compares them with in vivo data, enhancing understanding of pleural biology^{9,12}. Another study used scRNA-seq, paired with other genomic and histologic analyses, to explore the EMT of PM malignant cells and their tumor microenvironment. It identified distinct malignant cell programs for epithelioid and sarcomatoid histologies, a new uncommitted EM phenotype in biphasic tumors, and signaling pathways as potential drivers of PM cell fate. These findings offer valuable insights into PM biology and highlight non-malignant cell signals as contributors to EMT and tumor progression¹⁰. This comprehensive analysis sheds light on the diverse molecular profiles within MESO cell populations, emphasizing the importance of considering heterogeneity in understanding cancer biology and developing targeted therapies^{13,14}.

MESO has limited treatment options, and its subtypes can influence treatment responses. The processes of EMT and the composition of immune cell populations are closely associated with the effectiveness of immunotherapy¹⁵. Up to date, there is a scarcity of specific studies investigating circulating mesothelioma cell heterogeneity at the transcriptomic level¹⁶.

In a ground-breaking study by Mangiante et al., a thorough examination of whole-genome sequencing data, coupled with transcriptomic and epigenomic data, was conducted using multiomics factor analysis. The results revealed four distinct dimensions that were found to be complementary. These dimensions effectively captured significant interpatient molecular disparities by emphasizing extreme phenotypes indicative of interdependent tumor specialization. Importantly, these findings shed light on the intricate interplay between the functional biology of PM and its genomic background, thereby offering valuable insights into the diverse clinical manifestations observed among PM patients¹³.

Tumor cell heterogeneity presents significant challenges in the realm of cancer treatment, emphasizing the critical need to comprehend and characterize this complexity to advance the development of efficacious cancer therapeutics. Leveraging technological breakthroughs, such as scRNA-seq, has emerged as a pivotal tool in deciphering the intricacies and dynamics of tumor heterogeneity. This technology enables the identification of potential therapeutic targets and the formulation of personalized treatment strategies.

In the context of this study, our objective is to delineate mesothelioma cell heterogeneity using scRNA-seq technique in both in vitro and in vivo conditions. By employing this advanced sequencing method, we aim to shed light on the diverse cell populations within PM tumors, providing a comprehensive understanding of their molecular landscape, including in circulating tumor cells. These new findings are anticipated to contribute to the development of novel strategies for prognosis and therapeutics not only in PM but also other types of cancer.

Ultimately, the insights gained from this research endeavour have the potential to pave the way for targeted and more effective treatments, improving patient outcomes and advancing the field of cancer therapeutics. Through a deeper understanding of mesothelioma cell heterogeneity, we aim to make significant strides in the ongoing efforts to combat this challenging and notorious cancer.

Results

Study design and data integration

Three distinct groups of mesothelioma RN5 cells were prepared from separate experimental conditions: cultured cells (CC) derived from in vitro cell line culture, circulating tumor cells (CTC), and peritoneal lavage tumor cells (Lav), both obtained in vivo from peripheral blood or peritoneal lavage, respectively, from tumor-bearing mice at 4 weeks post RN5 cell intraperitoneal (ip) injection (Fig. 1A). Enriched CTC (MSLN⁺CD45[−]) from peripheral blood were prepared using a MACS column and a microfluidic chip (Fig. 1S).

Subsequently, merged data were analyzed to elucidate the transcriptomic characteristics of tumor cell clusters. Among the total of 32,371 cells analyzed, 8208 were identified as tumor cells based on the expression of mesothelioma markers *Msln*, *Wt1*, and *Sparc* (Fig. 1B). Expression levels of tumor cell markers varied across different clusters. Six clusters of merged tumor cells were characterized by heatmap and volcano plots illustrating top up- and down-regulated genes (Figs. 2S & 3S). Additionally, all genes exhibiting significant changes within each cluster were documented (Table 1S).

The subsequent analysis focused on discerning differences among the three groups, as well as elucidating distinct features following reclustering of each group.

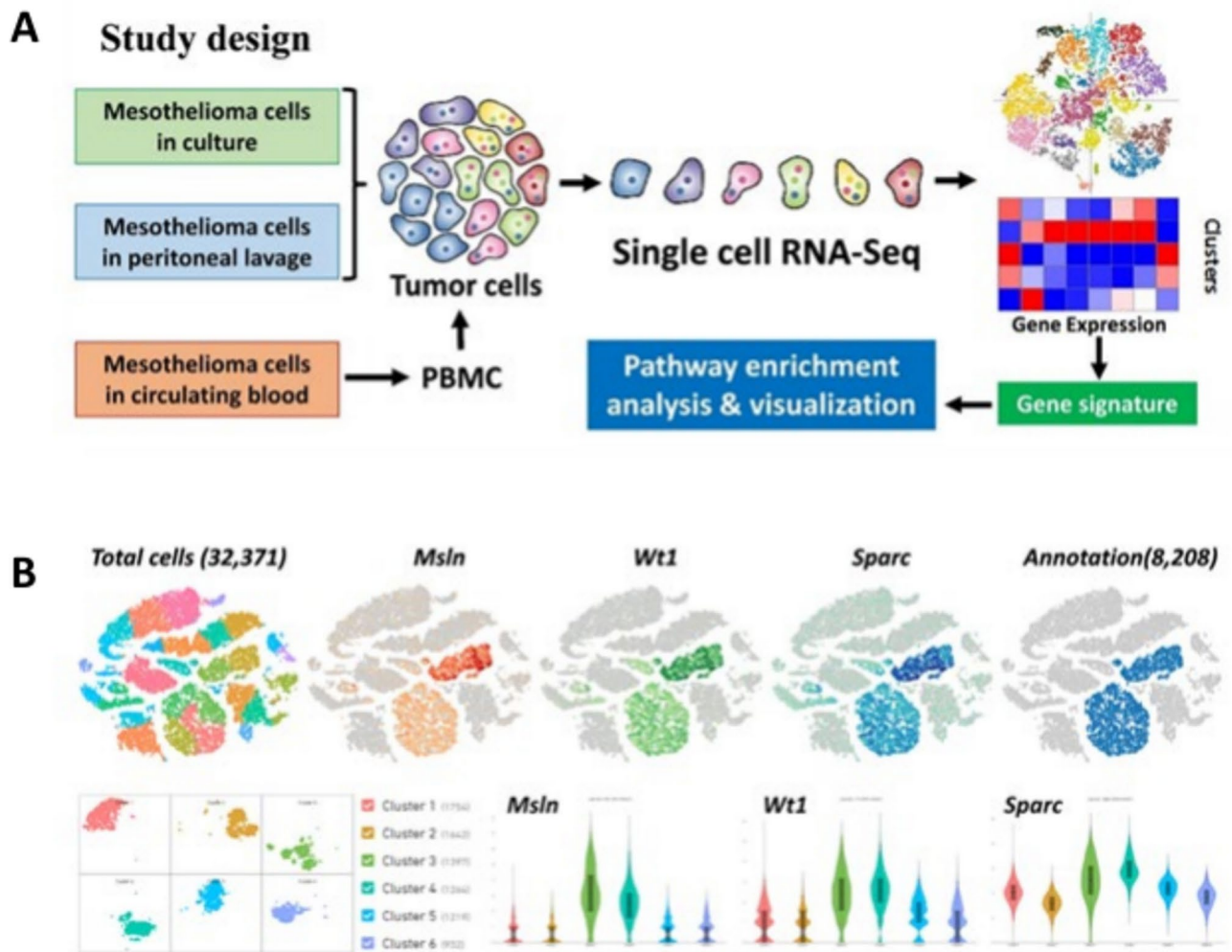


Fig. 1. Workflow of study design. (A) Single cells obtained from different sources for RNA sequencing to identify mesothelioma cell heterogeneity. Cultured cells (CC): murine mesothelioma RN5 cells were harvested approximately at 75% confluence, circulating tumor cells (CTC) and peritoneal lavage tumor cells (Lav) were obtained from tumor-bearing mice at 4 weeks post RN5 cell intraperitoneal (*ip*) injection. (B) Annotation of mesothelioma cells identified by gene expression of mesothelin (*Msln*), Wilm's tumor-1 (*Wt1*) and *Sparc*. Tumor cell clusters were reclustered into 6 subpopulations with various expression of *Msln*, *Wt1* and *Sparc* genes.

Tumor cell identification and reclustering

The total number of tumor cells (8208) annotated from merged RN5 cells underwent analysis through t-distributed stochastic neighbor embedding (tSNE) clustering. This analysis was performed on cultured cells (CC), lavage cells derived from a mouse *ip* model (Lav), and circulating tumor cells within peripheral blood mononuclear cells (PBMC) of a mouse *ip* model (CTC). These cells expressed the genes *Msln*, *Wt1*, and *Sparc*. Each group, divided by sample ID, was then reclustered into subpopulations: CC into 6 clusters, CTC into 4 clusters, and Lav into 5 clusters (Fig. 2A).

Top 50 up-regulated genes within each group of RN5 cells from CC, CTC, and Lav were identified (Fig. 2B). Additionally, the top 10 genes exhibiting significant changes (Log_2 Fold change > 1 or < -1 , and $p < 0.05$) from each subcluster in the three groups were visualized in heatmaps (Fig. 2C). Top 30 up- or down-regulated genes were presented in each cluster of the three groups (Fig. 4S), while volcano plots displayed all genes with significant changes in each cluster (Fig. 5S). Interestingly, the most upregulated genes in CTC were related to angiogenesis and platelets activation (Pbbp, Gp9, Clec11b), while the most upregulated genes in Lav were related to complement activation Q1qa, C1qb), suggesting that platelets could play a particularly important role in tumor metastasis in mesothelioma.

Furthermore, all up-regulated genes meeting the criteria (Log_2 Fold change > 1.0 and $p < 0.05$) were identified in CC (258 genes), CTC (147 genes), and Lav (105 genes) (Table 2S).

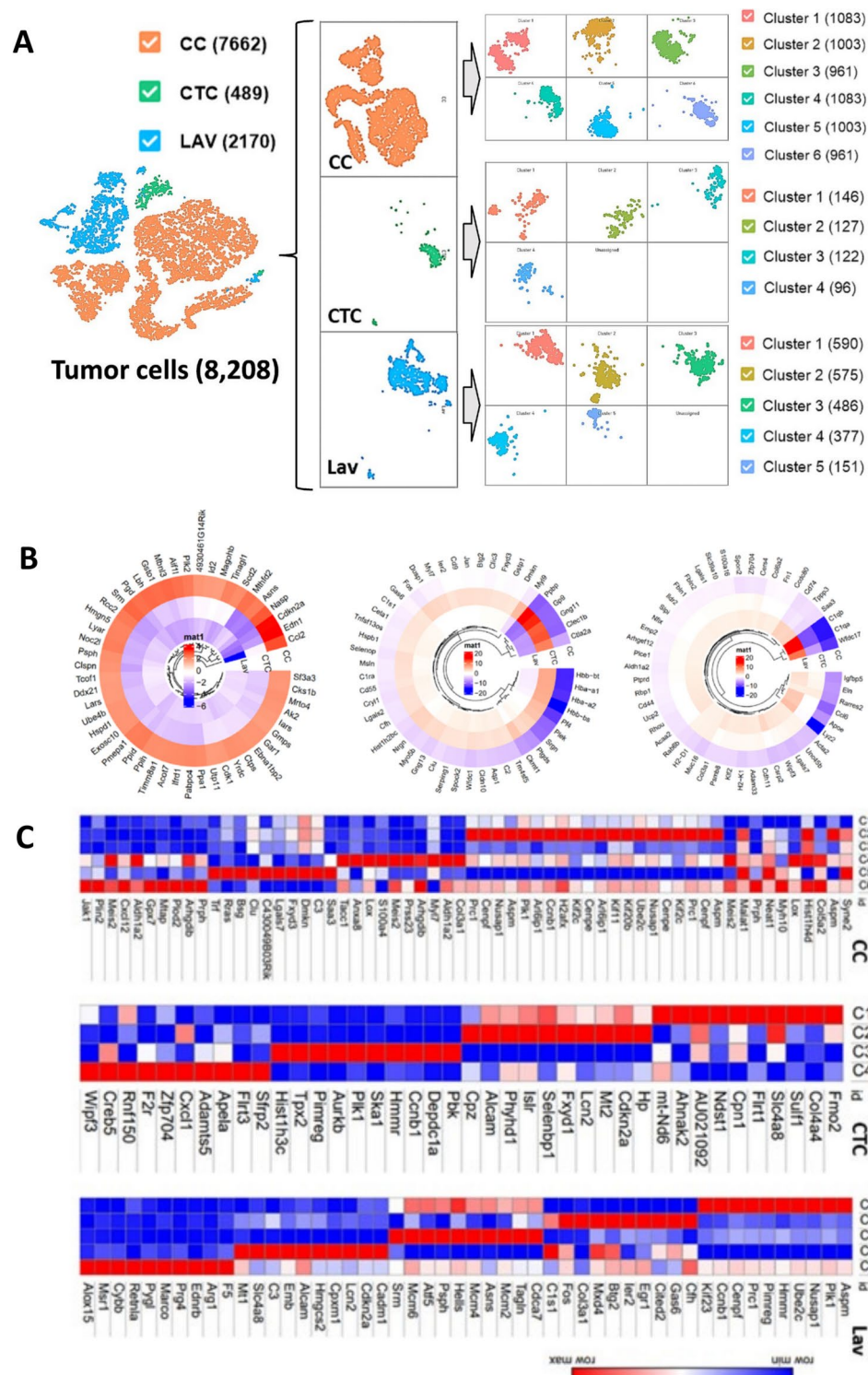


Fig. 2. Distinct subpopulations of all groups of merged tumor cells. **(A)** Total number of tumor cells (8208) annotated from merged tSNE clusters of cultured RN5 cells (CC), circulating tumor cells in peripheral blood monocytes (PBMC) and lavage cells derived from mouse *ip* model (LAV) of mouse *ip* model (CTC). Each group divided by sample ID was reclustered into subpopulations of CC 6 clusters, CTC 4 clusters, and Lav 5 clusters, respectively. **(B)** Gene expression in circular heatmap of top 50 up-regulated genes in each group: CC; CTC; and Lav. **(C)** Heatmaps of top 10 genes with significant change (Log2 Fold change > 1 or < -1, and p < 0.05) of each cluster from three groups.

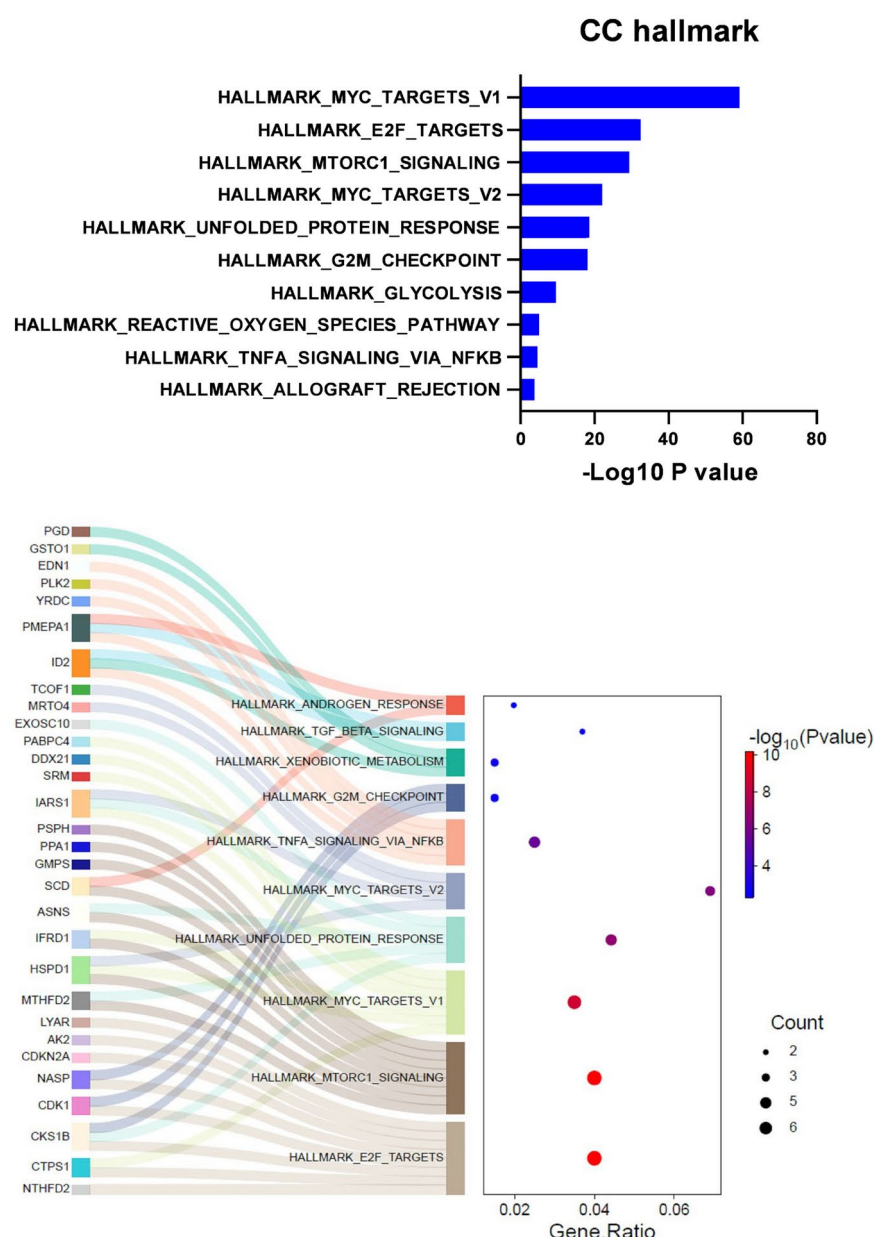
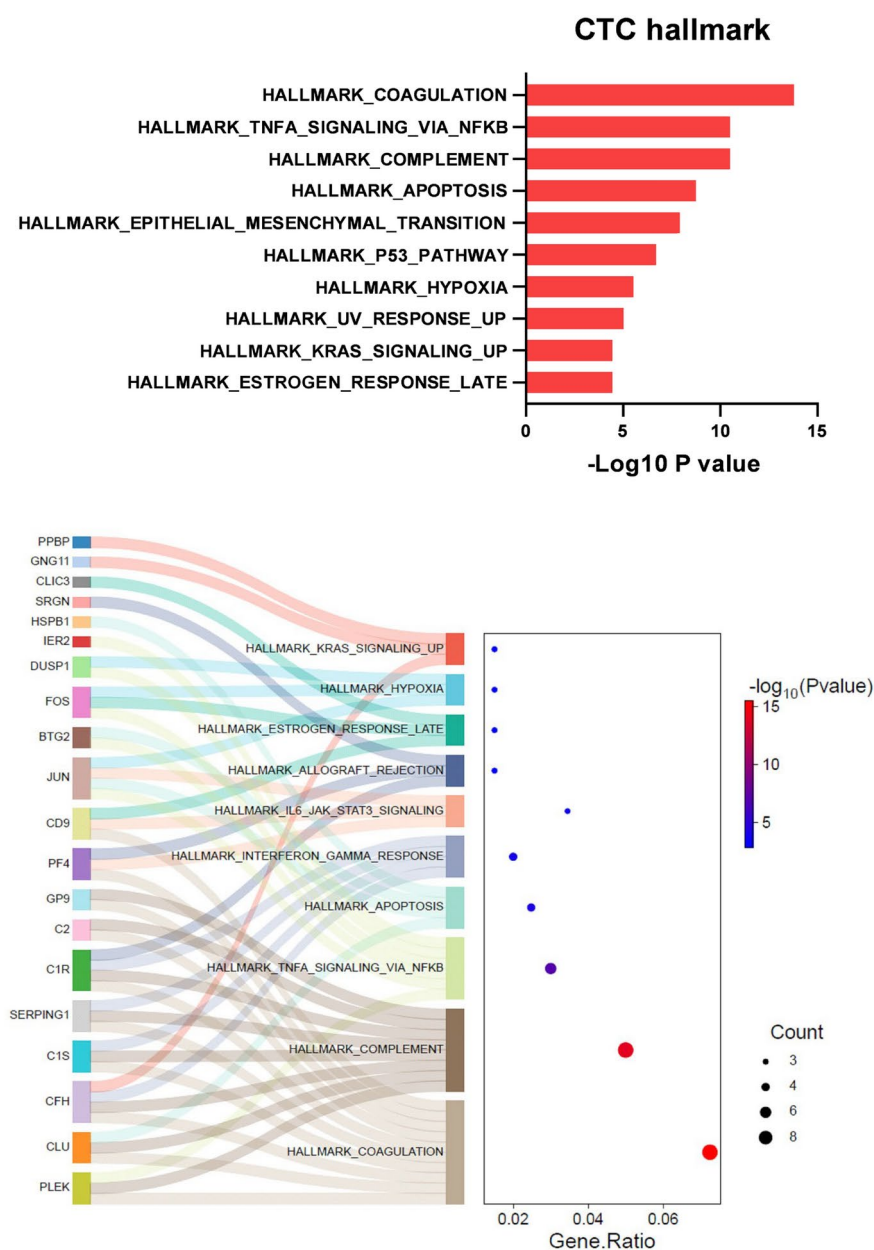
A

Fig. 3. Hallmark gene sets of top 50 up-regulated genes in CC, CTC and Lav groups determined by GSEA. **(A)** CC top 50 up-regulated genes involved hallmarks; **(B)** CTC top 50up involved hallmarks; **(C)** Lav top 50up involved hallmarks. Top 10 hallmaks and specific genes involved were shown in Sankey diagram on the left, the same pathways were shown in bar graph on the right, from top to bottom represents the hallmark numbers 1–10. On the right panel **(A)–(C)**, left part is Sankey plot, representing genes within each pathway; right part is dot plot, dot sizes represent gene numbers, and dot colors represent p values.

Hallmark pathways associated with up-regulated genes

The hallmark pathways prominently enriched among the top 50 up-regulated genes in the CC group include MYC targets v1 and v2, E2F targets, mTORC1 signaling, unfolded protein response, and G2M checkpoint. These pathways are known for their canonical roles in regulating cell cycle progression and proliferation (Fig. 3A). In contrast, the CTC group exhibits significant enrichment in hallmark pathways such as coagulation, TNF α signaling via NF κ B, complement, apoptosis, and epithelial–mesenchymal transition (EMT). This suggests that the overexpression of genes in CTCs may be associated with the promotion of cancer cell stemness in mesothelioma (Fig. 3B). Notably, among the top 10 hallmark pathways, EMT emerges as the most significant pathway in the context of tumor microenvironment. The panel of genes associated with EMT was predominantly related to the extracellular matrix (EMC) including genes that are characteristic of cancer associated myofibroblasts

B**Figure 3.** (continued)

(myCAF) such as *ACTA2*, *CD44*, and *FN1*. EMT was also associated with the emergence of an IFN- α and IFN- γ response, suggesting the EMT process may be associated with an immunogenic response. Overall, this supports a notion that the tumor microenvironment, and in particular, myCAFs may contribute to the EMT process as a mechanism of escape from the immunogenic response (Fig. 3C).

Gene ontology (GO) annotations of the up-regulated genes

Gene Ontology (GO) term annotations in biological process (BP), cellular component (CC), and molecular function (MF) categories were identified using GSEA (<https://www.gsea-msigdb.org/gsea/login.jsp>). The top 10 important GO term annotations for cultured cells (CC) (Fig. 4A), circulating tumor cells (CTC) (Fig. 4B), and peritoneal lavage cells (Lav) (Fig. 4C) were determined. Interestingly, GO terms in each group are unique, with no overlap among the three groups. Specifically, the CC group does not share any GO terms with either the CTC or Lav group. However, the CTC and Lav groups exhibit some overlaps: 4 overlaps in biological processes (GO BP) (Fig. 4D), 3 overlaps in cellular components (GO CC) (Fig. 4E), and 6 overlaps in molecular functions (GO MF) (Fig. 4F). This highlights distinct functional annotations associated with each cell type and underscores the heterogeneous nature of mesothelioma cells in different microenvironments.

C

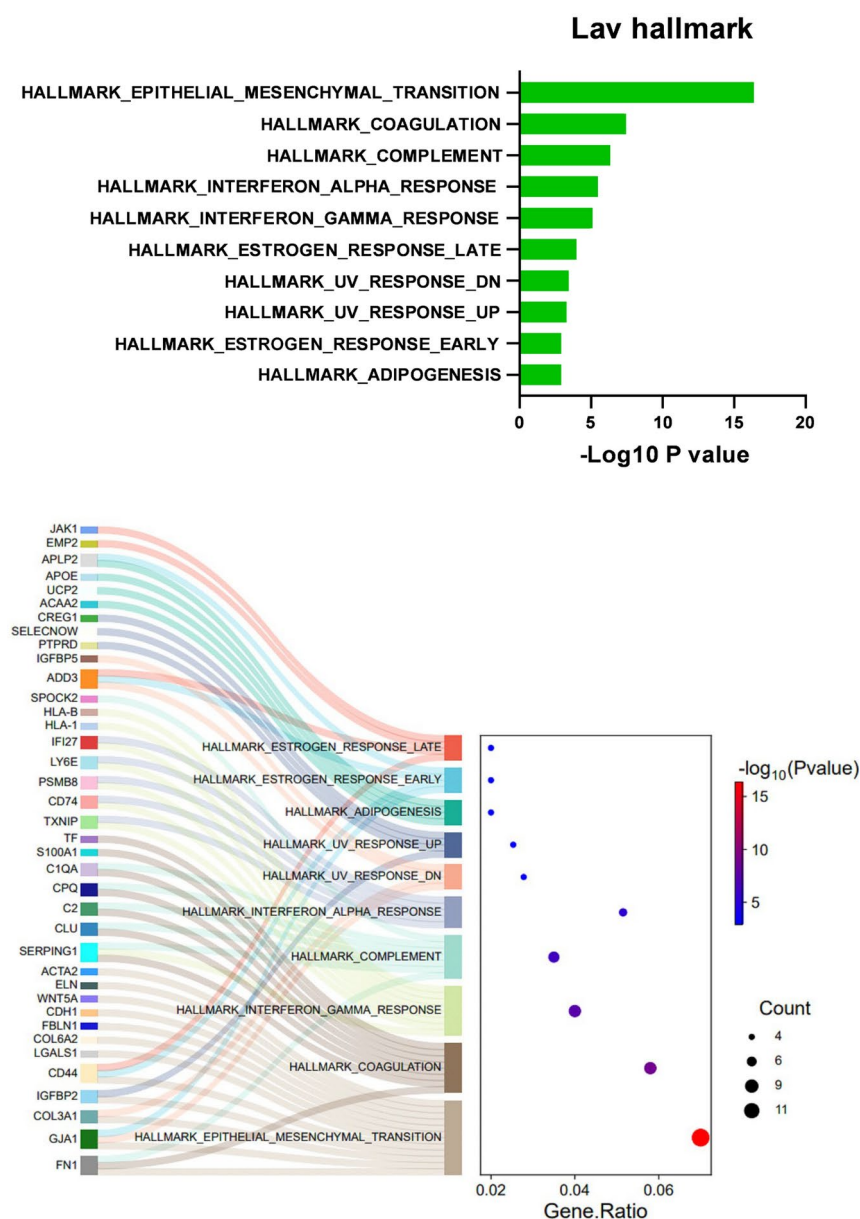


Figure 3. (continued)

Up-regulated genes in subclusters of cultured tumor cells with more tendency to promote cell proliferation

As anticipated, tumor cells cultured under optimized conditions exhibit a propensity for rapid proliferation. Gene set enrichment analysis (GSEA) revealed significant enrichment of hallmark gene sets associated with cell proliferation, including MYC targets v1 and v2, E2F, MTORC1 signaling, unfolded protein response, and G2M checkpoint, within the CC group with up-regulated genes (Fig. 3A). Gene sets of MYC targets v1 and v2, E2F, MTORC1 signaling had a large number of overlapped genes with CC up-regulated genes, however, only a few or no genes were found overlapped with CTC or Lav up-regulated genes (Fig. 6S). Velocity analysis did not show significant change in the trend of pseudo time along the cell population (Fig. 7S).

To further explore the heterogeneity within the CC group, we conducted reclustering, resulting in the identification of 6 subpopulations (Fig. 2A). Subsequently, the up-regulated genes within each subpopulation were analyzed for overlaps with cell proliferation gene sets, expressed as percentages.

Remarkably, the percentages of overlapping genes in CC clusters 2 and 3 exceeded 40%, which was notably higher than the average overlaps observed in other clusters within the CTC and Lav groups (Fig. 5A). This

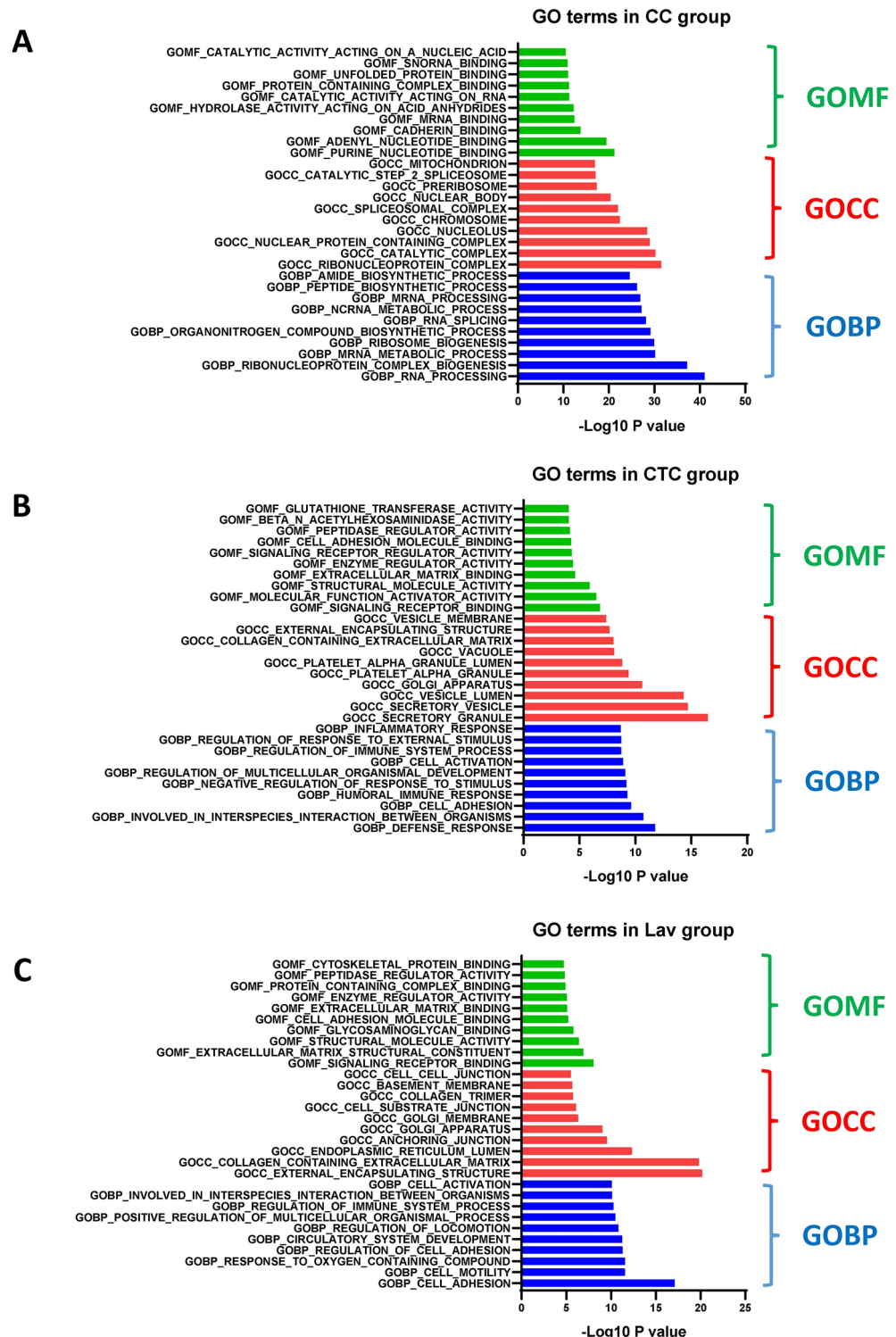


Fig. 4. Top 10 important Gene Ontology terms categories of the up-regulated genes identified by GSEA. (A)–(C) Gene Ontology (GO) term annotations in biological process (BP), cellular component (CC), and molecular function (MF) categories were identified using GSEA (<https://www.gsea-msigdb.org/gsea/login.jsp>). The top 10 important GO term annotations for cultured cells (CC) (A), circulating tumor cells (CTC) (B), and peritoneal lavage cells (Lav) (C) were determined. GO terms in each group are unique, with no overlap among the three groups. The CC group does not share any GO terms with either the CTC or Lav group. (D)–(F) The CTC and Lav groups exhibit some overlaps: 4 overlaps in biological processes (GO BP) (D), 3 overlaps in cellular components (GO CC) (E), and 6 overlaps in molecular functions (GO MF) (F). Each color represents specific pathways in the group no overlapping with other groups, same as shown in the venn diagrams.

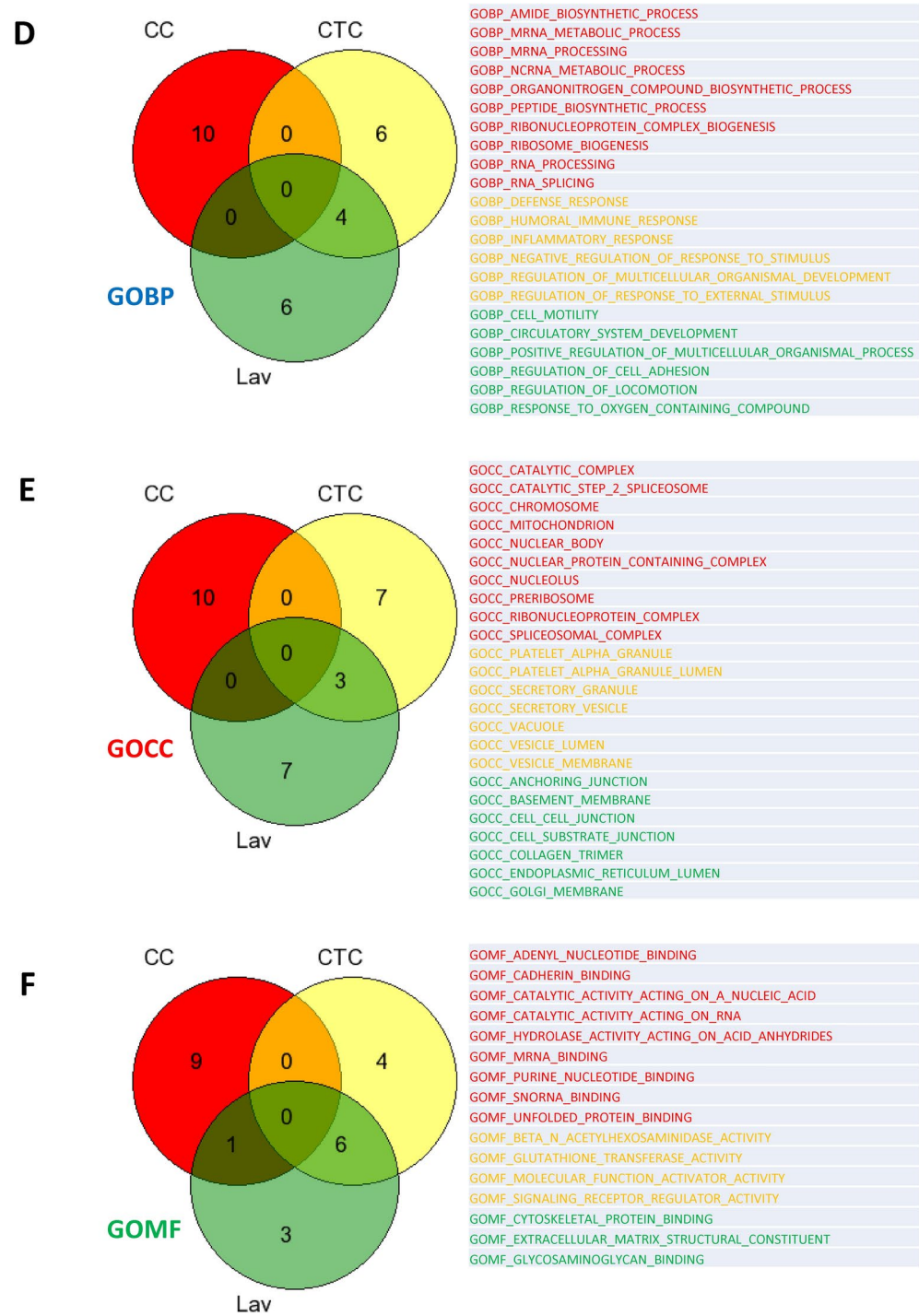


Figure 4. (continued)

suggests that subpopulations within the CC group, particularly clusters 2 and 3, may exhibit heightened proliferative potential compared to other clusters and groups.

Circulating tumor cells appear to possess more stemness property determined by stemchecker

Compared to other groups, the overlaps of up-regulated genes in each cluster of the CC group exhibited the least overlaps with stem cell gene sets, with the exception of the spermatogonial stem cell set, which showed approximately 2% overlaps in clusters 4 and 5. Conversely, in the CTC group, cluster 3 displayed significantly higher overlaps with embryonic stem cell, neural stem cell, intestinal stem cell, and hematopoietic stem cell gene sets compared to the other clusters. However, in the Lav group, clusters 4 and 5 exhibited the highest overlaps

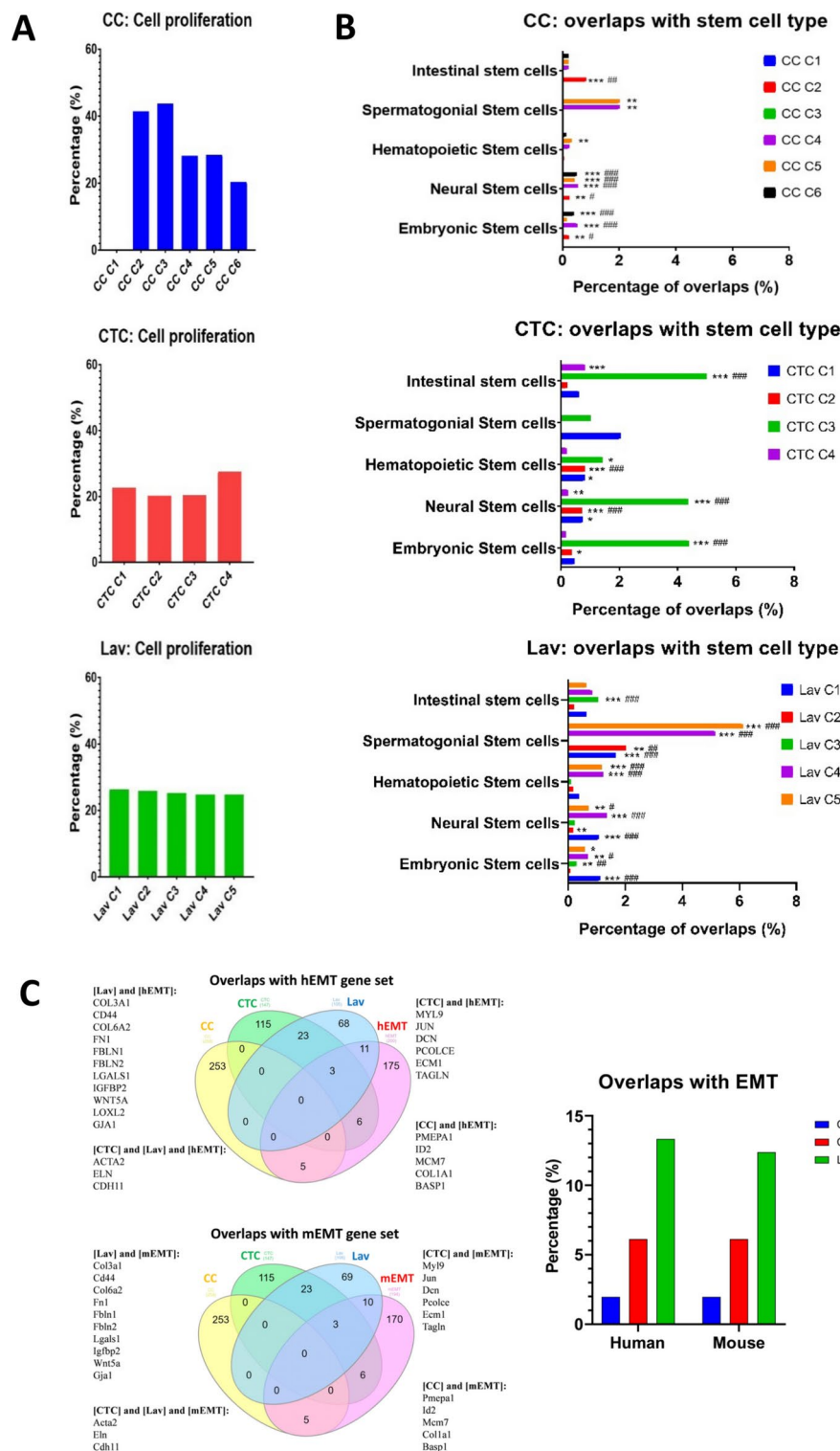


Fig. 5. Overlaps of up-regulated genes in subclusters of each group with cell proliferation gene set in Gene Ontology term BP identified by GSEA. (A) Percentage of the gene list from each cluster of CC, CTC, and Lav groups overlapped with cell proliferation gene set in BP category (%). (B) Overlaps of the gene list from each cluster of CC, CTC, and Lav groups with stem cell gene sets by StemChecker. (C) Overlaps of the gene list from each cluster of CC, CTC, and Lav groups with human and mouse gene sets of hallmark EMT. Top hallmark gene set enrichment in Lav group was presented.

A Top100 up-regulated genes involved cancer-related pathways

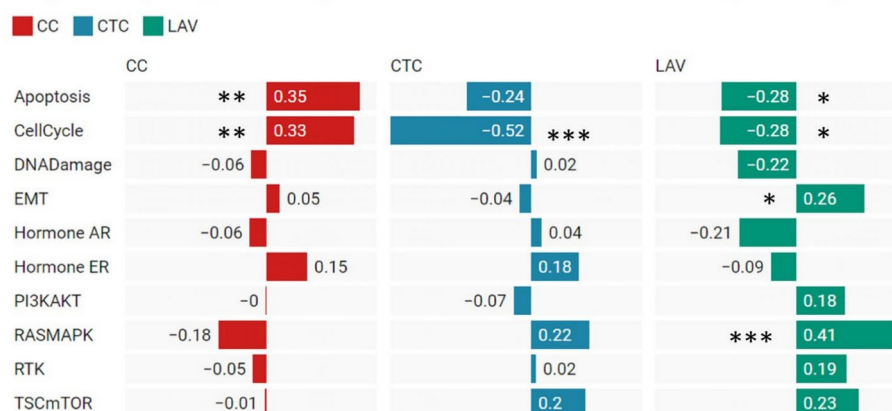


Chart: Licun Wu • Source: Dr. de Perrot Lab • Created with Datawrapper

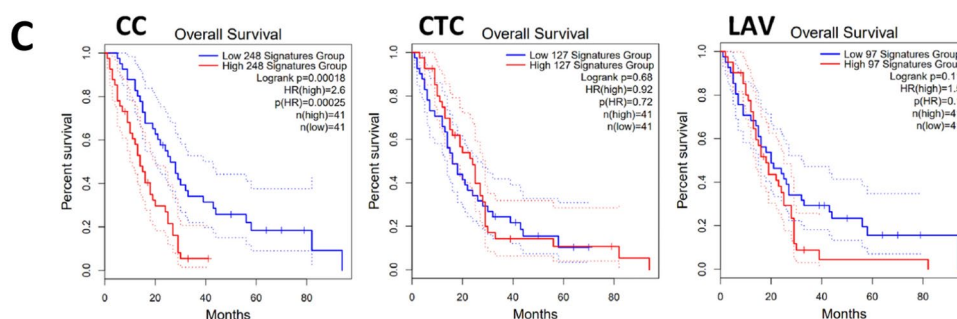
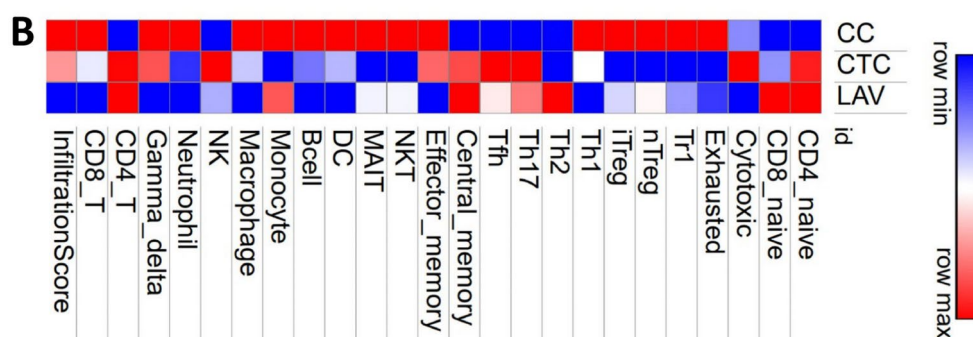


Fig. 6. The association of GSEA score with activity of cancer-related pathways and overall survival in MESO cohort TCGA. **(A)** Top up-regulated genes (Log2 Fold change > 1) involved cancer-related pathways in MESO, TCGA; **(B)** Tumor infiltrating lymphocytes in MESO, TCGA; and **(C)** Importing the gene list as a gene signature of each group to the GEPIA online platform to analyze the survival difference between low and high expression levels in MESO, TCGA.

(5.15% and 6.12%, respectively) with the spermatogonial stem cell gene set, while showing lower overlaps with other stem cell types (Fig. 5B).

Given that EMT ranks as the top upregulated pathway in the Lav group, we specifically investigated the overlaps of up-regulated genes in each group with the hallmark EMT gene sets in both human and mouse.

EMT pathway ranks top hallmark gene enrichment in Lav group

The human and mouse gene sets of hallmark EMT comprise 200 and 194 genes, respectively. Using the InteractiVenn program, we calculated the overlaps of up-regulated genes in each group with the human or mouse EMT gene sets. Among the up-regulated genes, 5 genes (1.98%) in CC, 9 genes (6.12%) in CTC, and 14 genes (13.33%) in Lav overlapped with human EMT gene identifiers. Similarly, 5 genes (1.98%) in CC, 9 genes (6.12%) in CTC, and 13 genes (12.38%) in Lav overlapped with mouse EMT gene identifiers. The gene names of those overlapping with each group were also included (Fig. 5C). Notably, the Lav group exhibited the highest overlaps with the hallmark EMT gene set, including genes such as *COL3A1*, *CD44*, *COL6A2*, *FN1*, *FBLN1*,

FBLN2, *LGALS1*, *IGFBP2*, *WNT5A*, *LOXL2*, *GJA1*, *ACTA2*, *ELN*, and *CDH11*. These genes are associated with EMT processes and indicate a potential role for EMT in the Lav group.

The association of GSVA score with activity of cancer-related pathways and overall survival in MESO cohort TCGA

The differences in cancer-related pathway activity between high and low GSVA score groups in mesothelioma were analyzed using the GSVA platform (<https://guolab.wchscu.cn/GSCA/#/>). The association between GSVA score and activity of cancer-related pathways revealed distinct patterns among the CC, CTC, and Lav groups.

In the CC group, the GSVA score of up-regulated genes was positively correlated with cell cycle and apoptosis pathways, while being negatively correlated with the RAS-MAPK pathway. Conversely, in the CTC and Lav groups, the GSVA scores of up-regulated genes were negatively correlated with cell cycle and apoptosis pathways, but positively correlated with the RAS-MAPK pathway in the Lav group. Notably, in the Lav group, both the EMT and RAS-MAPK pathways were significantly up-regulated (Fig. 6A).

We also looked at the correlation of the GSVA score of up-regulated genes in each group with tumor-infiltrating lymphocytes in MESO, TCGA, and showed that the overall immune cell infiltration score is higher in CC genes compared with CTC and Lav groups (Fig. 6B), indicating that up-regulated gene expression of Lav and CTC may result in more immunosuppressive microenvironment.

Furthermore, survival analysis using TCGA database revealed interesting findings. The up-regulated genes in the CC group exhibited a highly significant impact on the survival of patients with mesothelioma. Higher levels of gene expression in the gene signature were associated with poorer prognosis compared to the group with lower levels (Logrank p value = 0.00018. However, the other two groups, CTC and Lav, did not show a significant impact on patient survival, with Logrank p values of 0.68 and 0.11, respectively (Fig. 6C).

Taken together, the high level of gene expression driving cell cycle and proliferation in the CC group may indicate significant prognostic value in PM.

Discussion

Mesothelioma exhibits considerable heterogeneity in both cellular and molecular biology, contributing to variations in tumor behavior and treatment response. Tumor cell heterogeneity poses challenges in predicting prognosis and designing effective therapeutic approaches. Therefore, to better understand the molecular and cellular diversity in mesothelioma would be crucial for accurate prognosis. Patients with more heterogeneous tumors may experience different clinical outcomes, treatment responses, and survival rates. Identifying tumor cell heterogeneity through advanced molecular profiling techniques is able to tailor personalized therapy and improve prognosis^{13,17}.

This is a series of studies using murine mesothelioma cell line RN5, following determination of EMT gene signature from the up-regulated genes at all-time points in RN5-bearing mice⁵. RN5 cell line was established in Nf2 heterogeneous C57BL/6 mice with a characteristics of biphasic subtype. We used this cell line to identify a panel of EMT genes specifically in mesothelioma. In this study, we observed remarkable enrichment of key pathways, including MYC targets v1 and v2, E2F targets, mTORC1 signaling, unfolded protein response, and G2M checkpoint, in cultured RN5 cells. These pathways are well-known for their canonical roles in controlling cell cycle progression and proliferation. The MYC gene, in particular, has garnered significant attention in cancer research due to its pivotal role in cell cycle regulation and proliferation. MYC is a proto-oncogene that encodes transcription factors involved in regulating the expression of genes critical for cell growth, proliferation, and apoptosis. Its dysregulation has been implicated in various aspects of cancer development and progression¹⁸.

E2F is a family of transcription factors pivotal in orchestrating the expression of genes essential for DNA synthesis and cell proliferation. Central to cell cycle regulation, E2F plays a critical role in facilitating the transition from the G1 phase to the S phase. Among the transcriptional targets regulated by E2F are cyclins, cyclin-dependent kinases (CDK), checkpoint regulators, as well as DNA repair and replication proteins. Extensive evidence underscores the crucial involvement of E2F transcription factors in modulating cell proliferation¹⁹. Tumors characterized by high E2F scores exhibit a significant enrichment in the expression of numerous cell cycle-related hallmark gene sets, including the G2M checkpoint, MYC targets v1 and v2, MTORC1 signaling, and unfolded protein response. The E2F pathway score serves as a reflection of underlying rapid cell proliferation within the tumor microenvironment²⁰.

Our findings reveal that the up-regulated genes in the CC group prominently engage in MYC targets v1 and v2 as well as E2F targets, primarily governing cell cycle regulation and proliferation. Furthermore, the mammalian target of rapamycin complex 1 (MTORC1) emerges as a pivotal regulator of cell growth and proliferation. MTORC1 represents a protein complex that integrates diverse signals, including nutrient availability, energy status, and growth factors, to meticulously orchestrate cell cycle progression and cellular metabolism^{21,22}. In addition, our analysis identified up-regulation of both the unfolded protein response and MTORC1 signaling pathways in the CC group.

The up-regulated genes identified in the CTC group exhibited significant overlap with gene sets associated with various stem cell types, suggesting their potential involvement in maintaining cancer cell stemness. Remarkably, oncogenic KRAS has been implicated in augmenting classical stemness signaling pathways, with KRAS overexpression, alone or in combination with TP53 alterations, playing a pivotal role in MESO development and progression^{23,24}. KRAS signaling is known to be crucial for stemness maintenance, as well as for regulating coagulation and complement pathways, which are vital for resolving inflammatory processes and facilitating wound healing, thus underscoring their regenerative capacity²⁵.

MESO tumor encompasses epithelioid, biphasic, and sarcomatoid subtypes, each exhibiting distinct EMT phenotypes. Previous research has highlighted the association of certain EMT genes, such as *COL5A2*, *SPARC*, and *ACTA2*, with the upregulation of TGF- β 1 signaling, hedgehog signaling, and IL-2-STAT5 signaling pathways

7. Among mesothelioma cells from CC, CTC and Lav groups, the up-regulated genes in the Lav group appear to predominantly govern the interaction between tumor cells and the microenvironment, including pathways associated with EMT.

Dongre et al. elucidated new insights into the mechanisms underlying EMT and its implications for cancer. They highlighted how EMT can confer increased tumor-initiating and metastatic potential to cancer cells, as well as render them more resistant to certain therapeutic regimens²⁶.

The murine mesothelioma RN5 cell line, characterized by biphasic morphology, exemplifies this phenomenon. Mesenchymal phenotypic tumor cells are known to drive the EMT process. Conventional therapies, such as cisplatin-based chemotherapy and gamma ray radiation, have been shown to result in the enrichment of mesenchymal stem cells^{27,28}. EMT has emerged as one of the mechanisms contributing to therapy resistance in PM²⁷.

More recently, cancer-associated fibroblasts (CAF) have been identified as key stromal cells driving the EMT process in cancer. Hu et al. discovered that CAF exosome LINC00355 promotes EMT and chemoresistance in colorectal cancer, further highlighting the intricate interplay between tumor cells and the tumor microenvironment in driving cancer progression and therapeutic resistance²⁹.

Mesenchymal stromal cells (MSC) and fibroblasts often exhibit similar morphology, leading to challenges in distinguishing between the two cell types. However, studies have revealed that cell subsets with the MSC phenotype also display characteristics of fibroblasts. Notably, these cell subsets with the fibroblast phenotype do not necessarily demonstrate the MSC phenotype, suggesting a unidirectional relationship where fibroblasts may originate from MSC subsets³⁰.

Recent advancements in scRNA-seq have significantly enhanced our understanding of mesothelioma's cellular complexity, offering valuable information that can be leveraged to refine treatment selection and develop more effective, individualized therapeutic strategies. In conclusion, tumor cell heterogeneity is a fundamental characteristic of most cancers arising from the accumulation of genetic, epigenetic and functional changes in tumor cells. Phenotypic heterogeneity of tumor cells may be driven by genetic alterations, epigenetic modifications, or the influence of the tumor microenvironment. Tumor cells within a heterogeneous population can exhibit distinct functional properties. We mainly aimed to identify the unique gene signature of each mesothelioma cell cluster and the distinct functional properties of tumor cells within a heterogeneous subpopulation, showing that some tumor cells may possess stem cell-like characteristics and can self-renew and differentiate into other cell types, while others may have different proliferative potential. Functional heterogeneity can also involve variances in cellular behavior, such as migration, invasion, angiogenesis, and response to therapy. Our findings may provide clues to better understand the specific functions contributing to tumorigenesis and progression, so as to search potential novel targets for therapeutic strategies.

Materials and methods

Murine mesothelioma RN5 cell culture and mouse models

The murine mesothelioma cell line RN5 was initially derived from C57BL/6 mice subsequent to asbestos exposure by our research team³¹. These RN5 cells exhibited biphasic morphology. Culturing conditions involved maintenance in RPMI1640 medium supplemented with 10% fetal bovine serum and 1% penicillin–streptomycin at 37 °C in a 5% CO₂ atmosphere. To ensure cell line integrity, prophylactic treatment with 5 µg/ml Plasmocin™ (Invivogen) was administered for a minimum of 2 weeks, confirming mycoplasma-free status.

For experimental procedures, exponentially growing RN5 cells (approximately 90% confluence) were prepared as follows: 1) For scRNA-seq, RN5 cells (2×10^6 cells in 500 µl PBS) were submitted; 2) For intraperitoneal (*ip*) injection into 6–8 week-old C57BL/6 mice obtained from the Jackson Laboratories, RN5 cells (2×10^6 cells in 200 µl PBS) were administered. Over an 8-week observation period, five mice were sacrificed weekly, with naive mice serving as controls. Total cells were harvested via peritoneal lavage with PBS. Briefly, the peritoneal cavity was exposed and rinsed with 5 ml PBS per mouse, collecting the lavage. Tumor spheroids were removed by filtration through a 40 µm cell strainer (ThermoFisher). Fresh single cells obtained were utilized for scRNA-seq analysis, and 3) Peripheral blood collection occurred at 4 weeks post-tumor cell injection. Upon CO₂ inhalation-induced euthanasia, approximately 8 ml of pooled blood was collected from 20 tumor-bearing mice. Enrichment of circulating tumor cells was performed utilizing the obtained blood samples.

All experimental protocols were approved by the Committee of Animal Resources Centre, Animal Use Protocol (AUP#3399), University Health Network (UHN). All methods were carried out in accordance with ARRIVE guidelines and regulations. All methods were performed in accordance with the relevant guidelines and regulations in UHN.

Enrichment of circulating tumor cells (CTC) from peripheral blood

The enrichment of MSLN⁺CD45[−] CTC from blood was conducted using a MACS column and a microfluidic chip using a protocol that we previously reported^{32,33}. In brief, fresh blood collected from tumor-grafted mice underwent gradient centrifugation using the peripheral blood mononuclear cells (PBMC) isolation protocol with Leucosep tubes³⁴. In this protocol, 15 ml of Ficoll was added to a 50 ml Leucosep tube at room temperature (RT). The tube was then centrifuged for 1 min at 1000 × g. Following this, the tube was filled with anticoagulated blood, and subsequently rinsed with a balanced salt solution (5% sodium citrate in PBS). The blood was diluted at a 1:3 ratio with the balanced salt solution and subjected to centrifugation for 25 min at 800 × g at RT, with the brake off. After gradient centrifugation, the buffy coat was isolated and incubated with rat anti-mouse CD45 microbeads (#130-052-301, Miltenyi Biotech) for 15 min in the refrigeration. The incubated samples were then processed through a MACS LD column (#130-042-901, Miltenyi), and fractions of CD45[−] cells were collected for further processing. Subsequently, these CD45[−] cells were incubated with rat anti-mouse MSLN antibody (#D233-3, MBL International) for 20 min in the refrigerator, followed by incubation with anti-rat

IgG microbeads (#130-048-502, Miltenyi) for 15 min in the refrigerator. The labeled samples were processed by a microfluidic immunomagnetic cell sorting (MICS) device to capture MSLN⁺ populations. Post-capture, the MSLN⁺CD45⁻ fraction was resuspended in 500 µL of PBS and immediately submitted for scRNA-seq. In some experiments, a small portion of cells, both pre- and post-capture, at each stage were stained with rat anti-mouse CD45 (#550994, BD Bioscience) and human recombinant anti-microbead antibodies (#130-122-219, Miltenyi) for 20 min in the refrigerator. These stained samples were subjected to analysis using an Attune NxT acoustic flow cytometer to evaluate the purity of MSLN⁺CD45⁻ cells within the samples.

Single cell RNA sequencing (scRNA-Seq) analysis

Fresh single cells, encompassing cultured cells (CC), circulating tumor cells (CTC), and total cells from peritoneal lavage (Lav), were prepared as described previously. These cells were subsequently processed by the Princess Margaret Genomic Centre at the University Health Network (UHN) following standard protocols available at www.pmgenomics.ca. Analysis of single-cell gene expression in clusters was conducted utilizing Loupe Cell Browser v5.0.0, provided by 10× Genomics, as well as CReSCENT: Cancer Single Cell Expression Toolkit, an online platform accessible at <https://crescent.cloud/>. CReSCENT is populated with public datasets and preconfigured pipelines that are accessible to computational biology non-experts, and user-editable to allow for optimization, comparison, and re-analysis on the fly. CReSCENT is under an open-source license via Github (General Public License v3.0).

10× Genomics Chromium v2 was used library preparation. Sequencing is performed on Illumina NextSeq platform. The threshold 3659 per barcode (linear) was selected by unique molecular identifiers (UMIs), and 6% (506/7868) cells were removed. Barcodes with unexpectedly high counts of UMIs may represent multiplets, and barcodes with very few genes may represent low-quality cells or empty droplets, especially those with fewer than 3 are unavailable for reclustering. To set thresholds for mitochondrial UMIs, we selected a reference genome to use pre-selected mitochondrial gene set mouse or human genome (mm10) whose threshold of cells with mitochondrial read percentage was selected 10%. Gene expression counts were log-normalized. This threshold was used to identify potential over-expression of mitochondrial genes, which could indicate poor cell quality or cells undergoing cellular stress or death.

Data acquisition and analysis

Differential gene expression analysis was conducted based on the predefined threshold criteria, which included a log2 fold change greater than 1 (equivalent to a twofold change) and a p-value less than 0.05.

For pathway analysis, total tumor cells obtained from the merged data of the three groups (CC, CTC, and Lav) were identified using tumor cell marker genes such as Msln, Wt1, and Sparc. Subsequently, tumor cell clusters were subjected to reclustering based on sample ID to identify globally distinguishing genes. Further reclustering of tumor cells within each group allowed for the investigation of specific functions within subpopulations. Top up-regulated genes associated with hallmark pathways and gene ontology (GO) annotation terms (BP: biological process; CC: cellular components; MF: molecular function) within each group were analyzed using the GSEA online platform available at <https://www.gsea-msigdb.org/gsea> (Versions: MSigDB 2024.1; GSEA 4.3.3).

Additionally, survival analysis was performed using The Cancer Genome Atlas (TCGA) data. The expression levels of top up-regulated genes were evaluated to determine their association with the prognosis of mesothelioma patients. This analysis was carried out using the TCGA data analysis platform accessible at <http://gepia.cancer-pku.cn/> (Version: GEPIA2.0).

Hallmark gene sets enrichment analysis

Gene set enrichment analysis (GSEA) was conducted to investigate functionally enriched pathways and hallmark gene sets associated with the identified subgroups. The hallmark gene sets utilized in the analysis were obtained from the Molecular Signatures Database (MSigDB), accessible at <http://software.broadinstitute.org/gsea/msigdb/>. A significance threshold of $p < 0.05$ was applied to determine significantly enriched pathways³⁵. Particular genes that are involved in hallmark pathways are shown in Sankey plots³⁶.

For the analysis of cell proliferation, the gene lists from each cluster within the groups were uploaded to the web-based gene set analysis toolkit available at <https://www.webgestalt.org/> (WebGestalt V1.0). *Mus musculus* was selected as the organism to obtain biological process (BP) genes from the Gene Ontology (GO) database. This allowed for the calculation of the percentage of overlaps with cell proliferation within the BP category.

For stemness analysis, stem cell type annotation was conducted using the online platform developed by SysBio Lab at the University of Algarve, Portugal, accessible at <http://stemchecker.sysbiolab.eu/>³⁷. Upon importing the gene list, *Mus musculus* (Mouse) was selected from the "Checkerboard Options," with masking of both "Mask Cell Proliferation Genes" and "Mask Cell Cycle Genes" to ensure specific focus on stem cell-related genes.

The analysis encompassed 25 stemness signatures and 73 transcription factors gene sets. The statistical details table provided significance of enrichment for genes included in composite gene sets associated with different stem cell types among the input genes identified in StemChecker. Composite gene sets for various cell types represent the unions of all selected stemness signatures corresponding to each cell type. Significance (p-value) was calculated via the hypergeometric test, assessing enrichment against the full annotated genome of the selected organism. Additionally, adjusted p-values were calculated using Bonferroni correction to account for multiple comparisons.

For EMT enrichment analysis, the up-regulated genes from each group were compared with the EMT hallmark gene set. This gene set, consisting of 200 genes, was downloaded from the Molecular Signatures Database (MSigDB) at https://www.gsea-msigdb.org/gsea/msigdb/cards/HALLMARK_EPITHELIAL_MESENCHYMAL_TRANSITION.html.

To analyze the overlaps between any two comparisons, the InteractiVenn platform was utilized. This platform allows for the visualization of shared genes among different gene lists. The InteractiVenn platform is accessible at <https://www.interactivenn.net/>³⁸. By comparing the up-regulated genes from each group with the EMT hallmark gene set, the analysis focuses on determining the number of genes from each group that may participate in the EMT process.

Association of the up-regulated genes with overall survival of MESO patients in TCGA

For gene set variation analysis (GSVA) and survival analysis, the Gene Set Cancer Analysis (GSCA) tool estimates the association between GSVA score and overall survival (OS) in MESO. GSVA scores and clinical survival data are merged by sample barcode. Tumor samples are then divided into high and low GSVA score groups based on the median GSVA score. Subsequently, the R package survival is employed to fit the survival time and survival status of the two groups. Cox proportional-hazards model and logrank tests are performed to generate Kaplan–Meier curves for the survival comparison between the high and low GSVA score groups in MESO³⁹. It's important to note that GSVA score represents the variation of gene set activity over a specific cancer sample population in an unsupervised manner. The GSVA score reflects the integrated level of expression of a gene set, and it is positively correlated with the expression of the gene set. Additional information on GSVA score and its interpretation can be found at <https://guolab.wchscu.cn/GSCA/#/expression> (Version2024).

Gene expression correlated with cancer-related pathways in TCGA data

Using the same platform (<https://guolab.wchscu.cn/GSCA/#/expression>) and selecting the cancer type MESO, we computed the correlation between GSVA score and cancer-related pathways, as well as immune cell infiltration. The GSVA and pathway activity module presents the correlation between GSVA score and pathway activity, which is defined by pathway scores. This analysis provides insights into the relationship between the expression level of gene sets and the activity of cancer-related pathways. In this analysis, statistical significance is denoted by "∗: P value < 0.05 " and "∗∗: $FDR < 0.05$ ", indicating results with p -values or false discovery rates (FDR) below the specified thresholds.

Statistical analysis

Statistical analysis was conducted using GraphPad Prism 8.0 (GraphPad Inc., San Diego, CA, USA). For comparisons between two groups, an unpaired two-tailed Student's t -test was employed. A p -value less than 0.05 was considered statistically significant. Results were presented as mean \pm SEM. Significance levels were indicated as follows: ∗, $p < 0.05$; ∗∗, $p < 0.01$; ∗∗∗, $p < 0.001$ in all figures.

For survival analysis comparing overall survival (OS) between low and high-risk groups, Kaplan–Meier analysis was performed with the log-rank test. All tests were two-tailed, and a p -value < 0.05 and/or $FDR < 0.05$ were considered significant.

Data availability

Expression profiles of all three groups of mesothelioma cells and subclusters of each group are available in the supplementary data. All other raw data generated in this study could be available upon request from the corresponding author

Received: 5 November 2024; Accepted: 28 February 2025

Published online: 13 March 2025

References

- Kindler, H. L. et al. Treatment of malignant pleural mesothelioma: American Society of Clinical Oncology Clinical Practice Guideline. *J. Clin. Oncol.* **36**(13), 1343–1373. <https://doi.org/10.1200/JCO.2017.76.6394> (2018).
- Beasley, M. B., Galateau-Salle, F. & Dacic, S. Pleural mesothelioma classification update. *Virchows Arch.* **478**(1), 59–72. <https://doi.org/10.1007/s00428-021-03031-7> (2021).
- Meiller, C. et al. Multi-site tumor sampling highlights molecular intra-tumor heterogeneity in malignant pleural mesothelioma. *Genome Med.* **13**(1), 113. <https://doi.org/10.1186/s13073-021-00931-w> (2021).
- Zhang, M. et al. Clonal architecture in mesothelioma is prognostic and shapes the tumour microenvironment. *Nat. Commun.* **12**(1), 1751. <https://doi.org/10.1038/s41467-021-21798-w>. Erratum In: *Nat Commun.* **2021**;12(1):3569 (2021).
- Wu, L. et al. Defining and targeting tumor-associated macrophages in malignant mesothelioma. *Proc. Natl. Acad. Sci. USA* **120**(9), e2210836120. <https://doi.org/10.1073/pnas.2210836120> (2023).
- Wu, L., Amjad, S., Yun, H., Mani, S. & de Perrot, M. A panel of emerging EMT genes identified in malignant mesothelioma. *Sci. Rep.* **12**(1), 1007. <https://doi.org/10.1038/s41598-022-04973-x> (2022).
- Wu, L. et al. Prognostic value of EMT gene signature in malignant mesothelioma. *Int. J. Mol. Sci.* **24**(5), 4264. <https://doi.org/10.3390/ijms24054264> (2023).
- Kinker, G. S. et al. Pan-cancer single-cell RNA-seq identifies recurring programs of cellular heterogeneity. *Nat. Genet.* **52**(11), 1208–1218. <https://doi.org/10.1038/s41588-020-00726-6> (2020).
- Obacz, J. et al. Single-cell transcriptomic analysis of human pleura reveals stromal heterogeneity and informs *in vitro* models of mesothelioma. *Eur. Respir. J.* **63**(1), 2300143. <https://doi.org/10.1183/13993003.00143-2023> (2024).
- Severson, D. T. et al. Comprehensive multi-site profiling of the malignant pleural mesothelioma micro-environment identifies candidate molecular determinants of histopathologic type. *bioRxiv* <https://doi.org/10.1101/2024.05.24.595690> (2024).
- Liu, J. et al. A gene signature linked to fibroblast differentiation for prognostic prediction of mesothelioma. *Cell Biosci.* **14**(1), 33. <https://doi.org/10.1186/s13578-023-01180-7> (2024).
- Shamseddin, M. et al. Use of preclinical models for malignant pleural mesothelioma. *Thorax* **76**(11), 1154–1162. <https://doi.org/10.1136/thoraxjnl-2020-216602> (2021).
- Mangiante, L. et al. Multiomic analysis of malignant pleural mesothelioma identifies molecular axes and specialized tumor profiles driving intertumor heterogeneity. *Nat. Genet.* **55**(4), 607–618. <https://doi.org/10.1038/s41588-023-01321-1> (2023).

14. Carbone, M., Adusumilli, P. S., Alexander, H. R. Jr, Baas, P., Bardelli, F., Bononi, A., Bueno, R., Felley-Bosco, E., Galateau-Salle, F., Jablons, D., Mansfield, A. S., Minaai, M., de Perrot, M., Pesavento, P., Rusch, V., Severson, D. T., Taioli, E., Tsao, A., Woodard, G., Yang, H., Zauderer, M. G. & Pass, H. I. Mesothelioma: Scientific clues for prevention, diagnosis, and therapy. *CA Cancer J. Clin.* **69**(5), 402–429. <https://doi.org/10.3322/caac.21572> (2019). Erratum in: *CA Cancer J. Clin.* 2020;70(4):313–314. <https://doi.org/10.3322/caac.21607>.
15. Kuryk, L. et al. novel insights into mesothelioma therapy: Emerging avenues and future prospects. *Front. Oncol.* **12**, 916839. <https://doi.org/10.3389/fonc.2022.916839> (2022).
16. Laure, A. et al. Genomic and transcriptomic analyses of malignant pleural mesothelioma (MPM) samples reveal crucial insights for preclinical testing. *Cancers* **15**(10), 2813. <https://doi.org/10.3390/cancers15102813> (2023).
17. Torricelli, F. et al. Spatially resolved, high-dimensional transcriptomics sorts out the evolution of biphasic malignant pleural mesothelioma: New paradigms for immunotherapy. *Mol. Cancer* **22**(1), 114. <https://doi.org/10.1186/s12943-023-01816-9> (2023).
18. Llombart, V. & Mansour, M. R. Therapeutic targeting of “undruggable” MYC. *EBioMedicine* **75**, 103756. <https://doi.org/10.1016/j.ebiom.2021.103756> (2022).
19. Müller, H. & Helin, K. The E2F transcription factors: Key regulators of cell proliferation. *Biochim Biophys. Acta* **1470**(1), M1–12. [https://doi.org/10.1016/s0304-419x\(99\)00030-x](https://doi.org/10.1016/s0304-419x(99)00030-x) (2000).
20. Chida, K. et al. E2F target score is associated with cell proliferation and survival of patients with hepatocellular carcinoma. *Surgery* **174**(2), 307–314. <https://doi.org/10.1016/j.surg.2023.04.030> (2023).
21. Hetz, C. The unfolded protein response: Controlling cell fate decisions under ER stress and beyond. *Nat. Rev. Mol. Cell Biol* **13**(2), 89–102. <https://doi.org/10.1038/nrm3270> (2012).
22. Yu, J. et al. PCYT1A suppresses proliferation and migration via inhibiting mTORC1 pathway in lung adenocarcinoma. *Biochem. Biophys. Res. Commun.* **529**(2), 353–361. <https://doi.org/10.1016/j.bbrc.2020.05.164> (2020).
23. Marazioti, A. et al. KRAS signaling in malignant pleural mesothelioma. *EMBO Mol. Med.* **14**(2), e13631. <https://doi.org/10.15252/emmm.202013631> (2022).
24. Manoharan, G. B., Laurini, C., Bottone, S., Ben Fredj, N. & Abankwa, D. K. K-Ras binds calmodulin-related Centrin1 with potential implications for K-Ras driven cancer cell stemness. *Cancers* **15**(12), 3087. <https://doi.org/10.3390/cancers15123087> (2023).
25. Budeus, B., Unger, K., Hess, J., Sentek, H. & Klein, D. Comparative computational analysis to distinguish mesenchymal stem cells from fibroblasts. *Front. Immunol.* **14**, 1270493. <https://doi.org/10.3389/fimmu.2023.1270493> (2023).
26. Dongre, A. & Weinberg, R. A. New insights into the mechanisms of epithelial–mesenchymal transition and implications for cancer. *Nat. Rev. Mol. Cell Biol.* **20**(2), 69–84. <https://doi.org/10.1038/s41580-018-0080-4> (2019).
27. Ebrahimi, N. et al. Harnessing function of EMT in cancer drug resistance: A metastasis regulator determines chemotherapy response. *Cancer Metastasis Rev.* <https://doi.org/10.1007/s10555-023-10162-7> (2024).
28. Wu, L. et al. Putative cancer stem cells may be the key target to inhibit cancer cell repopulation between the intervals of chemoradiation in murine mesothelioma. *BMC Cancer* **18**(1), 471. <https://doi.org/10.1186/s12885-018-4354-1> (2018).
29. Hu, J. H., Tang, H. N. & Wang, Y. H. Cancer-associated fibroblast exosome LINC00355 promotes epithelial–mesenchymal transition and chemoresistance in colorectal cancer through the miR-34b-5p/CRKL axis. *Cancer Gene Ther.* **31**(2), 259–272. <https://doi.org/10.1038/s41417-023-00700-4> (2024).
30. Fan, C. et al. Single-cell transcriptome integration analysis reveals the correlation between mesenchymal stromal cells and fibroblasts. *Front. Genet.* **13**, 798331. <https://doi.org/10.3389/fgene.2022.798331> (2022).
31. Rehrauer, H. et al. How asbestos drives the tissue towards tumors: YAP activation, macrophage and mesothelial precursor recruitment, RNA editing, and somatic mutations. *Oncogene* **37**(20), 2645–2659. <https://doi.org/10.1038/s41388-018-0153-z> (2018).
32. Duong, B. T. V. et al. A liquid biopsy for detecting circulating mesothelial precursor cells: A new biomarker for diagnosis and prognosis in mesothelioma. *EBioMedicine* **61**, 103031. <https://doi.org/10.1016/j.ebiom.2020.103031> (2020).
33. Wang, Z. et al. Isolation of tumour-reactive lymphocytes from peripheral blood via microfluidic immunomagnetic cell sorting. *Nat. Biomed. Eng.* **7**(9), 1188–1203. <https://doi.org/10.1038/s41551-023-01023-3> (2023).
34. Herzig, M. C. et al. A streamlined proliferation assay using mixed lymphocytes for evaluation of human mesenchymal stem cell immunomodulation activity. *J. Immunol. Methods* **488**, 112915. <https://doi.org/10.1016/j.jim.2020.112915> (2021).
35. Zheng, H. et al. Characterization of stem cell landscape and identification of stemness-relevant prognostic gene signature to aid immunotherapy in colorectal cancer. *Stem Cell Res. Ther.* **13**(1), 244. <https://doi.org/10.1186/s13287-022-02913-0> (2022).
36. Tang, D. et al. SRplot: A free online platform for data visualization and graphing. *PLoS ONE* **18**(11), e0294236. <https://doi.org/10.1371/journal.pone.0294236> (2023).
37. Pinto, J. P. et al. StemChecker: A web-based tool to discover and explore stemness signatures in gene sets. *Nucleic Acids Res.* **43**(W1), W72–W77. <https://doi.org/10.1093/nar/gkv529> (2015).
38. Heberle, H., Meirelles, G. V., da Silva, F. R., Telles, G. P. & Minghim, R. InteractiVenn: A web-based tool for the analysis of sets through Venn diagrams. *BMC Bioinform.* **16**, 169. <https://doi.org/10.1186/s12859-015-0611-3> (2015).
39. Liu, C. J. et al. GSCA: An integrated platform for gene set cancer analysis at genomic, pharmacogenomic and immunogenomic levels. *Brief Bioinform.* **24**(1), bbac558. <https://doi.org/10.1093/bib/bbac558> (2023).

Acknowledgements

We gratefully acknowledge the financial support from Princess Margaret Cancer Foundation (PMCF), Canadian Mesothelioma Foundation (CMF), and Ontario Institute for Cancer Research (OICR).

Author contributions

LW, DZW, AZ, SOK and MdP designed the experiments, performed most of these experiments, analyzed the data, and wrote the manuscript; DZW performed some of the experiments; SOK provided scientific expertise and the interpretation of data for the work; and LW and MdP contributed to the idea generation, experimental design, and manuscript writing and conceived the project.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-025-92542-3>.

Correspondence and requests for materials should be addressed to M.P.

Reprints and permissions information is available at www.nature.com/reprints.

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-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, 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 licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence 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 licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025