

The dynamic dysregulated network identifies stage-specific markers during lung adenocarcinoma malignant progression and metastasis

Li Wang,^{1,2} Wangyang Liu,^{1,2} Kailai Liu,¹ Lixia Wang,¹ Xiangzhe Yin,¹ Lin Bo,¹ Haotian Xu,¹ Shihua Lin,¹ Ke Feng,¹ Xinyu Zhou,¹ Lin Lin,¹ Meiting Fei,¹ Caiyu Zhang,¹ Shangwei Ning,¹ and Hongying Zhao¹

¹College of Bioinformatics Science and Technology, Harbin Medical University, Harbin 150081, China

Brain metastasis occurs in approximately 30% of patients with lung adenocarcinoma (LUAD) and is closely associated with poor prognosis, recurrence, and death. However, dynamic gene regulation and molecular mechanism driving LUAD progression remain poorly understood. In this study, we performed a comprehensive single-cell transcriptome analysis using data from normal, early stage, advanced stage, and brain metastasis LUAD. Our single-cell-level analysis reveals the cellular composition heterogeneity at different stages during LUAD progression. We identified stage-specific risk genes that could contribute to LUAD progression and metastasis by reprogramming immune-related and metabolic-related functions. We constructed an early advanced metastatic dysregulated network and revealed the dynamic changes in gene regulations during LUAD progression. We identified 6 early advanced (*HLA-DRB1*, *HLA-DQB1*, *SFTPB*, *SFTPC*, *PLA2G1B*, and *FOLR1*), 8 advanced metastasis (*RPS15*, *RPS11*, *RPL13A*, *RPS24*, *HLA-DRB5*, *LYPLA1*, *KCNJ15*, and *PSMA3*), and 2 common risk genes in different stages (*SFTPD* and *HLA-DRA*) as prognostic markers in LUAD. Particularly, decreased expression of *HLA-DRA*, *HLA-DRB1*, *HLA-DQB1*, and *HLA-DRB5* refer poor prognosis in LUAD by controlling antigen processing and presentation and T cell activation. Increased expression of *PSMA3* and *LYPLA1* refer poor prognosis by reprogramming fatty acid metabolism and RNA catabolic process. Our findings will help further understanding the pathobiology of brain metastases in LUAD.

INTRODUCTION

Lung cancer is the cancer with the fastest increasing morbidity and mortality, and it is also one of the malignant tumors threatening the health and life of the population. The 2 main subtypes of lung cancer are small cell lung cancer and non-small cell lung cancer (NSCLC),¹ histologically divided into lung adenocarcinoma (LUAD), squamous cell carcinoma, and large cell carcinoma, among which LUAD is the most common histological subtype.² If detected early, surgical resection of NSCLC has a good prognosis.³ However, approximately 75% of patients have advanced

disease (stage III/IV) at the time of diagnosis.⁴ Because of the mutation of cancer cells, the invasion degree is increased.^{5,6} The cells separate from the primary tumor and travel through blood vessels to other organs. Brain metastasis is the leading cause of prognosis, recurrence, and death in NSCLC patients, of whom approximately 20%–40% eventually develop brain metastasis.⁷ Advanced cancer and brain metastasis can shorten overall survival and decrease quality of life. However, the specific aspects of the different stages of LUAD progression remain poorly understood. It has become an urgent task to study the molecular characteristics of different stages of LUAD progression and further analyze the molecular mechanisms of cancer in search of new therapeutic targets.

Cancer is a multicellular community composed of malignant epithelial cells and different types of non-malignant immune and stromal cells, which show dynamic interactions.⁸ Numerous studies have shown significant tumor heterogeneity of both genetic and cellular compositions between primary and metastatic tumors.⁹ Advanced tumors consist of subclones with various genetic alterations and functional roles.¹⁰ Tumor heterogeneity and the complex cellular architecture play a crucial role in cancer progression and response to treatment.¹¹ However, the dynamic molecular features that characterize the contributions of tumor heterogeneity to malignant progression, metastasis, and poor survival are largely unknown. Many different cell types in our bodies express unique transcriptomes. Traditional bulk sequencing can only provide an average expression signal for a collection of cells, whereas single-cell RNA sequencing (scRNA-seq) isolates all cells in a tissue sample and sequences each

Received 8 June 2022; accepted 17 November 2022;

<https://doi.org/10.1016/j.omtn.2022.11.019>.

²These authors contributed equally

Correspondence: Li Wang, College of Bioinformatics Science and Technology, Harbin Medical University, Harbin 150081, China.

E-mail: wangli@hrbmu.edu.cn

Correspondence: Shangwei Ning, College of Bioinformatics Science and Technology, Harbin Medical University, Harbin 150081, China.

E-mail: ningsw@ems.hrbmu.edu.cn

Correspondence: Hongying Zhao, College of Bioinformatics Science and Technology, Harbin Medical University, Harbin 150081, China.

E-mail: zhaohongying@hrbmu.edu.cn



cell in turn. scRNA-seq can be used to carry out non-biased high-throughput studies with the smallest samples, measure gene expression levels more accurately, and have greater sensitivity in quantifying rare mutations and transcripts.¹² There is increasing evidence that gene expression is heterogeneous, even in similar cell types.^{13,14} scRNA-seq can reveal the complicated and rare cells, adjust the relationship between genes, and reveal the trajectory in the development of different cell lineages.¹⁵

In summary, we integrated scRNA-seq datasets from lung tissues, early stage, advanced stage, and brain metastasis LUAD tissues to characterize dynamic dysregulated networks, identify stage-specific risk genes and explore their prognostic value during LUAD tumor progression and metastasis (Figure 1). This study revealed the dynamic molecular characteristics and functions of LUAD in the progression of early advanced brain metastasis, providing an important theoretical basis for the molecular mechanism research and targeted therapy of LUAD.

RESULTS

Identification of cellular components at different stages of LUAD progression

To explore the changes in cell composition during the progression of LUAD, we obtained scRNA-seq data from different tissue sources, including normal lung tissue (normal lung), early primary LUAD tissue (early LUAD), advanced primary LUAD tissue (advanced LUAD), and brain metastasis tissue (metastatic LUAD) (Figure 2A). We cataloged the tissue cells into 9 cell lineages based on the expression levels of canonical marker genes¹⁶ (Figure 2B, Table S1). These include epithelial cells, stromal cells (fibroblasts, endothelial cells), immune cells (T lymphocytes, natural killer [NK] cells, B lymphocytes, myeloid cells, and mast cells), and oligodendrocytes. Except for oligodendrocytes, which were only present in metastatic LUAD, other cell types were recognized in normal lung, early LUAD, advanced LUAD, and metastatic LUAD (Figures 2C–2E). Compared with normal lung tissue, T lymphocytes and B lymphocytes were significantly enriched and NK cells were decreased in primary LUAD tissues (early LUAD and advanced LUAD, respectively). The results indicated that the adaptive immune process of the body was gradually activated with the occurrence and development of cancer. In addition, compared with normal lung tissue, it was found that the percentage of epithelial cells increased significantly in primary LUAD tissue and metastatic LUAD, which were 8.0%, 18.9%, and 45.4%, respectively (Figure 2F). The results indicated that the tumor microenvironment showed high cellular heterogeneity and dynamically changing after LUAD progression and metastasis. By analyzing the changes in cell composition caused by the occurrence, development, and invasion of LUAD, it was found that the components of different cell types were significantly different in the process of normal primary LUAD brain metastasis, revealing the difference of tissue-specific cell population. The signature genes of all cell types showed stable expression patterns, and each classic marker gene was only expressed in the corresponding cell types identified by us, independent of the stage of cancer development (Figure 2G). These

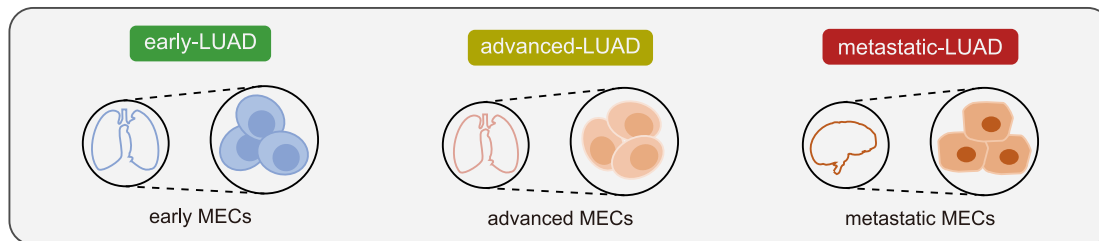
results indicate that there are significant differences in the proportion of cell components in different stages of LUAD progression.

Heterogeneity of tumor cells in the pathological process of early stage, advanced stage to brain metastasis of LUAD

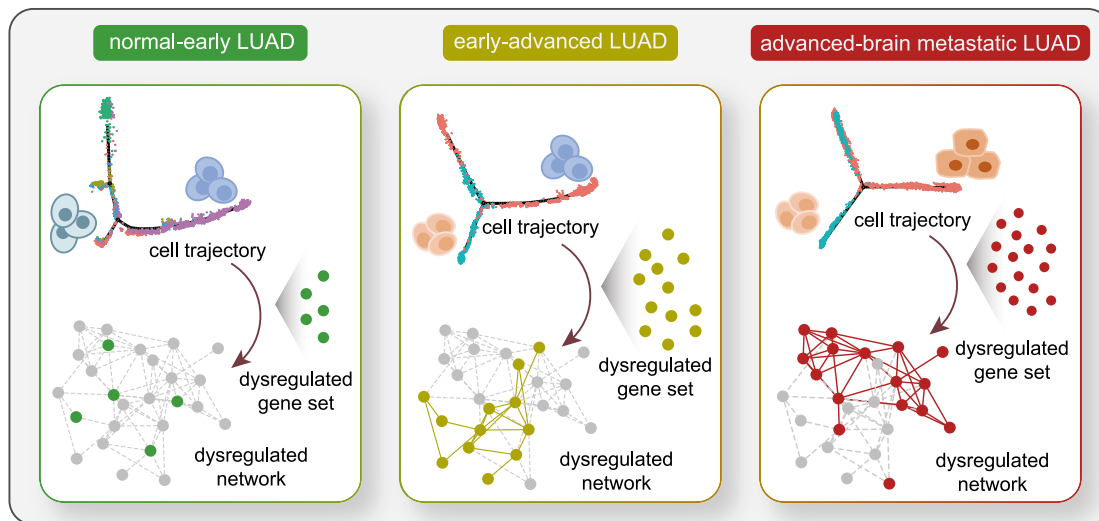
To characterize the heterogeneity of tumor cells and identify stage-specific risk genes in LUAD tissues, we identified the malignant tumor cells in the early stage, advanced stage, and brain metastasis LUAD (Figures 3A and 3B). Using stromal cells (fibroblasts and endothelial cells) and immune cells (T cells, NK cells, B cells, myeloid cells, and mast cells) as reference cells, inferCNV analysis was applied to infer large-scale copy number variations based on the scRNA-seq data. We identified 2,071, 4,103, and 7,448 LUAD malignant epithelial cells in early stage (early MECs), advanced stage (advanced MECs), and brain metastatic stage (metastatic MECs), respectively. We divided epithelial cells in lung tissue into 4 subpopulations based on canonical marker gene expression, namely, alveolar type I cells (*AGER*), alveolar type II cells (*SFTPC/LAMP3*), Club cells (*SCGB1A1*), and ciliated cells (*FOXJ1/RFX2*) (Figures 3C and 3D). Alveolar cell type II is the cell type that is enriched in normal lung tissue (Figure 3D). Alveolar type II cells are components of the pulmonary epithelium and serve as progenitor cells of alveolar type I cells to promote the repair and regeneration of the alveolar epithelium.¹⁷ A trajectory analysis was performed on epithelial cells in normal lung tissue and early MECs to assess inter-cellular heterogeneity (Figure 4A). We found that alveolar type I cells were mainly located in state 4, alveolar type II cells were mainly located in state 5, ciliated cells were located in state 3, and state 3 was almost full of ciliated cells. Most of the early MECs were in state 1, and most of the cells in state 1 also belonged to the early MECs (Figure 4B). By comparing early MECs with normal epithelial cells in cell state 1, we identified 133 normal-early risk genes, including 78 significantly upregulated genes and 55 significantly downregulated genes (fold change [FC] > 2, false discovery rate [FDR] < 0.05). Functional analysis showed that normal-early risk genes were mainly enriched in respiratory gaseous exchange by respiratory system, antigen processing and presentation and T cell activation (Figure S1A). Downregulated genes in early MECs were mainly enriched in respiratory gaseous exchange by respiratory system and tissue homeostasis (Figure 4C). The significantly upregulated genes in early MECs were mainly enriched in immune-related functions, such as leukocyte activation, T cell activation, negative regulation of endopeptidase activity, and negative regulation of peptidase activity. Protease activity has been reported to be pivotal in executing antigen receptor responses and lymphocyte function.¹⁸ The dysregulation of protease activity and immune-related genes were related with cancer risk.^{19,20}

Trajectory analysis was performed on early MECs and advanced MECs to assess inter-cellular heterogeneity (Figure 4D). The results showed that the early MECs were mainly located in cell state 2, and most of the cell types in cell state 2 were early MECs (Figure 4E). Advanced MECs were mainly located in cell state 1, and most of the cells in cell state 1 were also advanced MECs. Therefore, cell state 2 is the specific cell state in the early stage of LUAD, and cell state 1 is

Step 1 Obtain single-cell data and identify malignant epithelial cells(MECs) at different stages of LUAD progression



Step 2 Construction of cell trajectory and dynamic dysregulated network



Step 3 Dynamic dysregulated network reconstruction and analysis

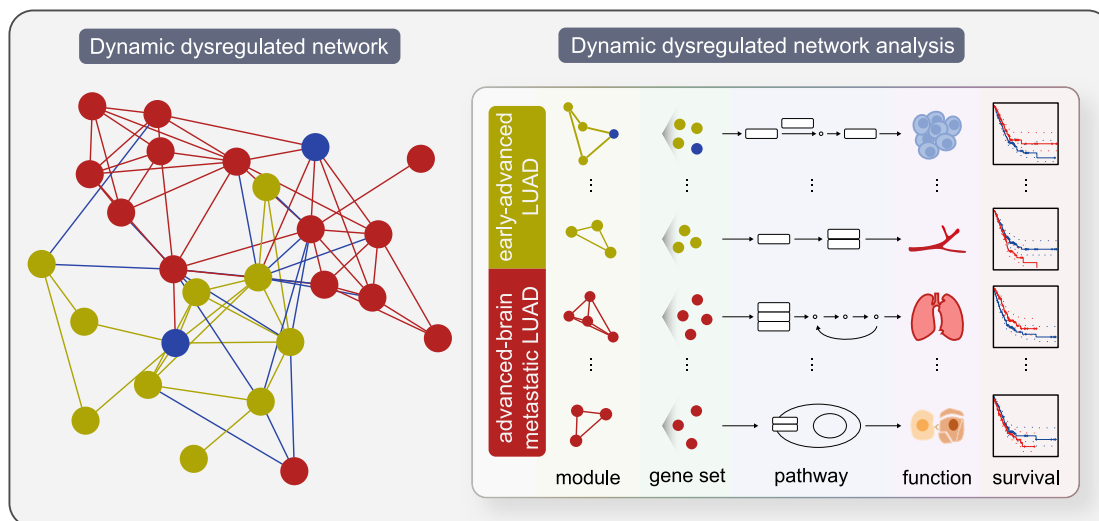


Figure 1. Molecular mechanism and functional analysis of different stages of LUAD progression

the specific cell state in the advanced stage of LUAD. By comparing advanced MECs with early MECs cells, we identified 151 early advanced risk genes. Early advanced risk genes were mainly enriched

in immune-related functions, such as antigen processing and presentation, humoral immune response, and T cell activation (Figure S1B). Furthermore, we identified 124 DEGs in state 1, 33 DEGs in state 2,

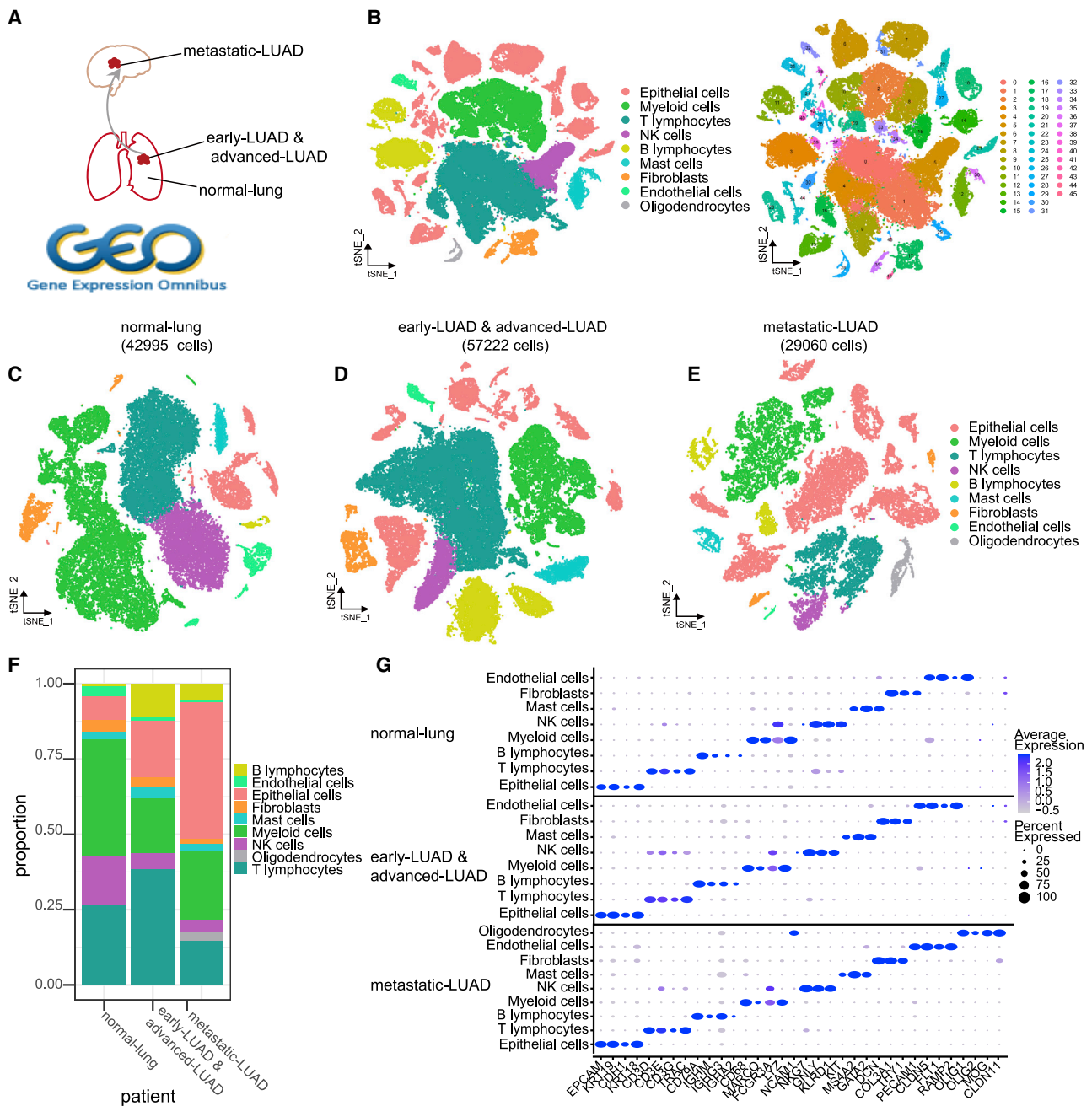


Figure 2. Identification of cellular components of scRNA-seq data from different stage of LUAD progression

(A) Overview of single-cell data from different tissues obtained from GEO. (B) tSNE map of 129,277 single cells. (C–E) tSNE plots of single cells in different tissues, colored by cell types. (F) The proportion of each cell types in different tissues. (G) Bubble plots of the average expression levels of classic marker genes in nine cell types from different tissue sources.

and 283 DEGs in state 3. We found that the DEGs of cell state 1 are mainly associated with neutrophil degranulation. The DEGs of cell state 2 were mainly related to regulation of endogenous peptidase activity, regulation of peptidase activity, cell chemotaxis, and cell-cell adhesion. The DEGs of cell state 3 were mainly related to cellular

detoxification, response to hydrogen peroxide, and cell response to toxic substances (Figure 4F). State 1 is a specific branch of advanced MECs in the early advanced trajectory, which is significantly enriched in neutrophil degranulation. Neutrophils have been reported to play an important role in cancer occurrence and progression.²¹ Granule

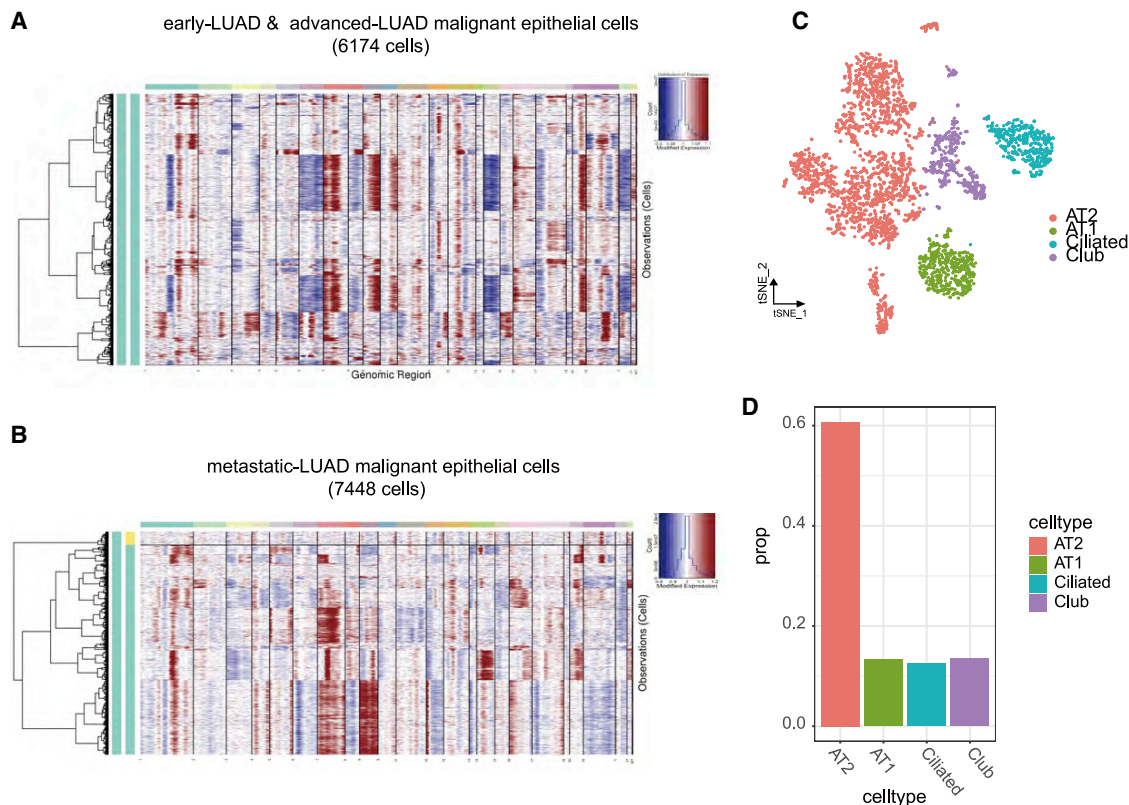


Figure 3. Identification of MECs at different stages of LUAD progression

(A, B) Heatmap showing large-scale CNV of epithelial cells in primary LUAD (A) and metastatic LUAD (B). (C) tSNE plots of normal epithelial subclusters, colored by cell types. Alveolar type I cells (AT1), alveolar type II cells (AT2), club cells (Club), and ciliated cells (Ciliated). (D) The bar chart details the proportions of different cell types.

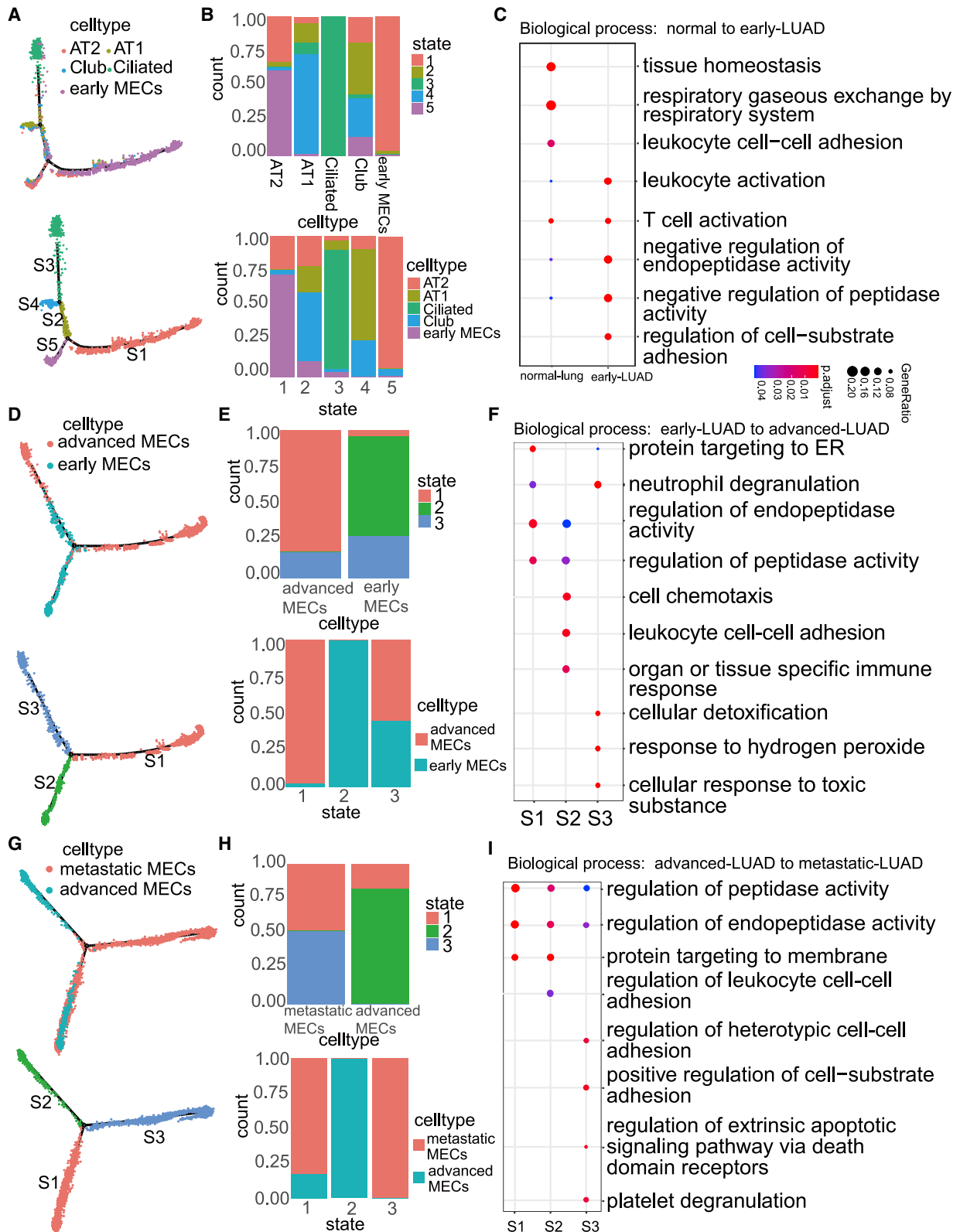
protein released by neutrophil degranulation is associated with tumor progression and may promote the development of LUAD.²²

Intercellular heterogeneity was evaluated by a trajectory analysis of advanced MECs and metastatic MECs (Figure 4G). The results showed that advanced MECs were mainly located in cell state 2, and most of the cell types in cell state 2 were advanced MECs. Metastatic MECs were mainly located in cell state 3, and most of the cells in cell state 3 were also metastatic MECs (Figure 4H). Therefore, cell state 2 is the specific cell state of advanced stage of LUAD, and cell state 3 is the specific cell state of brain metastasis of LUAD. By comparing advanced MECs with metastatic MECs cells, we identified 279 advanced metastatic risk genes. Functional analysis showed that advanced metastatic risk genes significantly influenced immune response-related and metabolic-related functions, such as antigen processing and presentation, T cell activation, fatty acid metabolic process, RNA catabolic process, and response to hypoxia (Figure S1C). For example, LYPLA1 was significantly upregulated from advanced to metastatic LUAD, which is involved in fatty acid metabolic processes. Recent studies have reported that reprogramming fatty acid metabolism could promote lymph node metastasis in cervical cancer.²³ Furthermore, we identified 83 DEGs in cell state 1, 73 DEGs in cell state 2, and 97 DEGs in cell state 3. We found that DEGs

in cell state 1 and cell state 2 are mainly related to regulation of endogenous peptidase activity, regulation of peptidase activity, and protein targeting to membrane. The DEGs of cell state 3 are mainly related to the regulation of exogenous signaling pathways through death domain receptors, positive regulation of cell substrate adhesion, and platelet degranulation (Figure 4I). State 3 is a unique branch of metastatic MECs in the advanced metastatic trajectory, which is significantly enriched in biological functions such as platelet degranulation. The interaction between cancer cells and activated platelets has been reported to promote the metastatic behavior of tumor cells.^{24–26} Metastatic MECs in state 3 may contribute to brain metastasis of LUAD through modulation of platelet degranulation. Our results demonstrated differences in transcriptional trajectories, revealing different cell expression patterns at different stages of LUAD progression.

Construction and functional analysis of early advanced metastatic dynamic dysregulated network in LUAD

To characterize the dynamic changes during LUAD progression at the single-cell level, we constructed early advanced and advanced metastatic dysregulated networks by integrating interaction relationships mediated by risk genes (Figures 5A and 5B). The early advanced dysregulated network involved 151 risk genes and 83 dysregulated interactions. The advanced metastatic dysregulated network involved



(legend on next page)

279 risk genes and 570 dysregulated interactions. Then, an early advanced metastatic dynamic dysregulated network was constructed by comparing early advanced and advanced metastatic dysregulated interactions (Figure 5C). As compared with early advanced dysregulated network, the advanced metastatic dysregulated network gained interactions 567 (73.6%) and lost 80 interactions (10.4%). These results showed the dynamic rewiring of network connections during different stages of LUAD progression.

We used the MCODE method to cluster the closely related nodes in the dynamic dysregulated network of early advanced brain metastasis, obtaining a total of 11 significant modules (Figure 5D). We found that module 7 contained only DEGs from advanced LUAD to brain metastasis, and module 11 contained only DEGs from the early stage to the advanced stage (Figure 5E). Functional enrichment analysis was performed for genes in the module (Figure 5F). The results show that module 1 is mainly composed of RPL and RPS family genes, and plays an important role in tumor cell growth and proliferation. Module 2 correlated with antigen processing and presentation, consisting of both MHC-I (*HLA-A*, *HLA-B*, and *HLA-C*) and MHC-II (*HLA-DRA*, *HLA-DBQ1*, *HLA-DRB1*, and *HLA-DRB5*) molecules. These genes are downregulated in LUAD progression, indicating increased tumor immune escape and decreased T cell recognition and activation, further confirming the decreased immunogenicity of cancer cells. Module 3 correlated with cellular responses, in which chemical signals bind to the corresponding receptors to induce events within the cell that ultimately change its behavior. Module 4 correlated with the regulation of nerve cell maturation, and most of module 4 genes are DEGs from advanced LUAD to brain metastasis. Module 5 correlated with the function of normal epithelial cells, such as gas exchange through the respiratory system and tissue homeostasis, suggesting that epithelial cells still function as normal epithelial cells during the progression of LUAD. Module 8 correlated with ATP synthesis, oxygen electron transport chains, and oxidative phosphorylation. Module 9 correlated with the importation of inorganic ions through the plasma membrane, and the transmembrane transport of inorganic ions is closely related to a variety of life processes.

Later, DAVID was used to enrich the KEGG pathway of module genes and identify the sub-pathway where the module genes were located, allowing us to reconstruct the risk pathway dynamically related to early advanced brain metastasis of LUAD (Figure 6). In many pathological conditions, such as during autoimmune, cancer and infection, major histocompatibility complex (MHC) provides peptides to T cells leading to the initiation of adaptive immune responses.²⁷ The results showed that MHC-I (*HLA-A*, *HLA-B*, and *HLA-C*) and *B2M* were significantly enriched in the upstream of the T cell receptor signaling pathway to regulate CD8 cell killing target cells and regulate

the activity of natural killer cells. MHC-II (*HLA-DQB1*, *HLA-DRB1*, *HLA-DRB5*, and *HLA-DRA*) molecules are significantly enriched in the upstream of the T cell receptor signaling pathway and regulate the secretion of cytokines by CD4 T cells and the activation of other immune cells. In this study, low expression of MHC-I (*HLA-A*, *HLA-B*, and *HLA-C*) molecules and MHC-II (*HLA-DQB1*, *HLA-DRB1*, *HLA-DRB5*, and *HLA-DRA*) molecules inhibited antigen processing and presentation, indicating increased tumor immune escape and thus promoting tumor progression. Dysregulation of the cell cycle leads to uncontrolled cell proliferation, one of the molecular hallmarks of tumorigenesis. The cell cycle is a well-organized mechanism and is tightly controlled by regulatory molecules such as cyclins, cyclin-dependent kinases and their inhibitors. Cyclin D1 (*CCND1*), a key regulatory protein, controls the transition from G1 to S phase during cell division.²⁸ In this study, *CCND1* is highly expressed in the progression of LUAD and is enriched in the downstream of the pathway, which can promote cell cycle progression and tumor development. Ferroptosis is a novel iron-dependent programmed cell death mode.²⁹ It plays a key role in inhibiting tumorigenesis by removing cells deficient in key nutrients or damaged by infection or environmental stress in the environment.³⁰ Ferroptosis-protection genes³¹ (*AKRIC1*, *AKRIC2*, and *AKRIC3*) are highly expressed in advanced metastatic LUAD and are significantly enriched in the upstream of chemical carcinogenesis-reactive oxygen pathway, promoting the progression of lung cancer. *GNB2*, *JUN*, and *FOS* are significantly enriched in MAPK signaling pathway, and *JUN* and *FOS* are enriched in the downstream of the pathway, promoting cancer cell proliferation, angiogenesis, and cancer metastasis. These results showed reveals the dynamic changes in gene regulation and function during different stages of LUAD progression.

Identification of prognostic markers in the dynamic dysregulated network at different stages of LUAD progression

To identify potential prognostic markers at different stages of LUAD progression, we performed survival analysis using 88 module genes based on the GEPIA2 database. As a result, we identified 16 stage-specific risk genes as prognostic markers of LUAD, including 6 early advanced risk genes (*HLA-DRB1*, *HLA-DQB1*, *SFTPB*, *SFTPC*, *PLA2G1B*, and *FOLR1*) and 8 advanced metastatic risk genes (*RPS15*, *RPS11*, *RPL13A*, *RPS24*, *HLA-DRB5*, *LYPLA1*, *KCNJ15*, and *PSMA3*), 1 common risk gene *SFTPD* in normal early and advanced metastatic stages, and 1 common risk gene *HLA-DRA* in normal early, early advanced, and advanced metastatic stages (Figures 7 and 8). For example, *HLA-DRB1*, *SFTPC*, and *FOLR1* showed significantly decreased expressions from the early to the advanced stage of LUAD (Figures 7 and 8). Decreased expression of *HLA-DRB1* ($p = 0.01$), *SFTPC* ($p = 3.0e-04$), and *FOLR1* ($p = 0.02$) were associated with poorer prognosis in LUAD (Figure 7). *HLA-DRB1*

Figure 4. Pseudotemporal trajectory analysis revealing heterogeneity of tumor cells in different stages of LUAD

(A) Developmental trajectory representation of normal epithelial cells and early MECs along inferred pseudotime by Monocle. Each point corresponds to a single cell, colored according to cell type (top) and cell state (bottom). (B, E, and H) Major cell state composition of each cell type (top) and cell types distribution in each cell state (bottom). (C, F, and I) The bubble plots show significantly enriched biological processes based on the analysis of the DEGs in normal-early LUAD (C), early-advanced LUAD (F), and advanced-metastatic LUAD (I). (D) Developmental trajectory of early MECs and advanced MECs. (G) Developmental trajectory of advanced MECs and metastatic MECs.

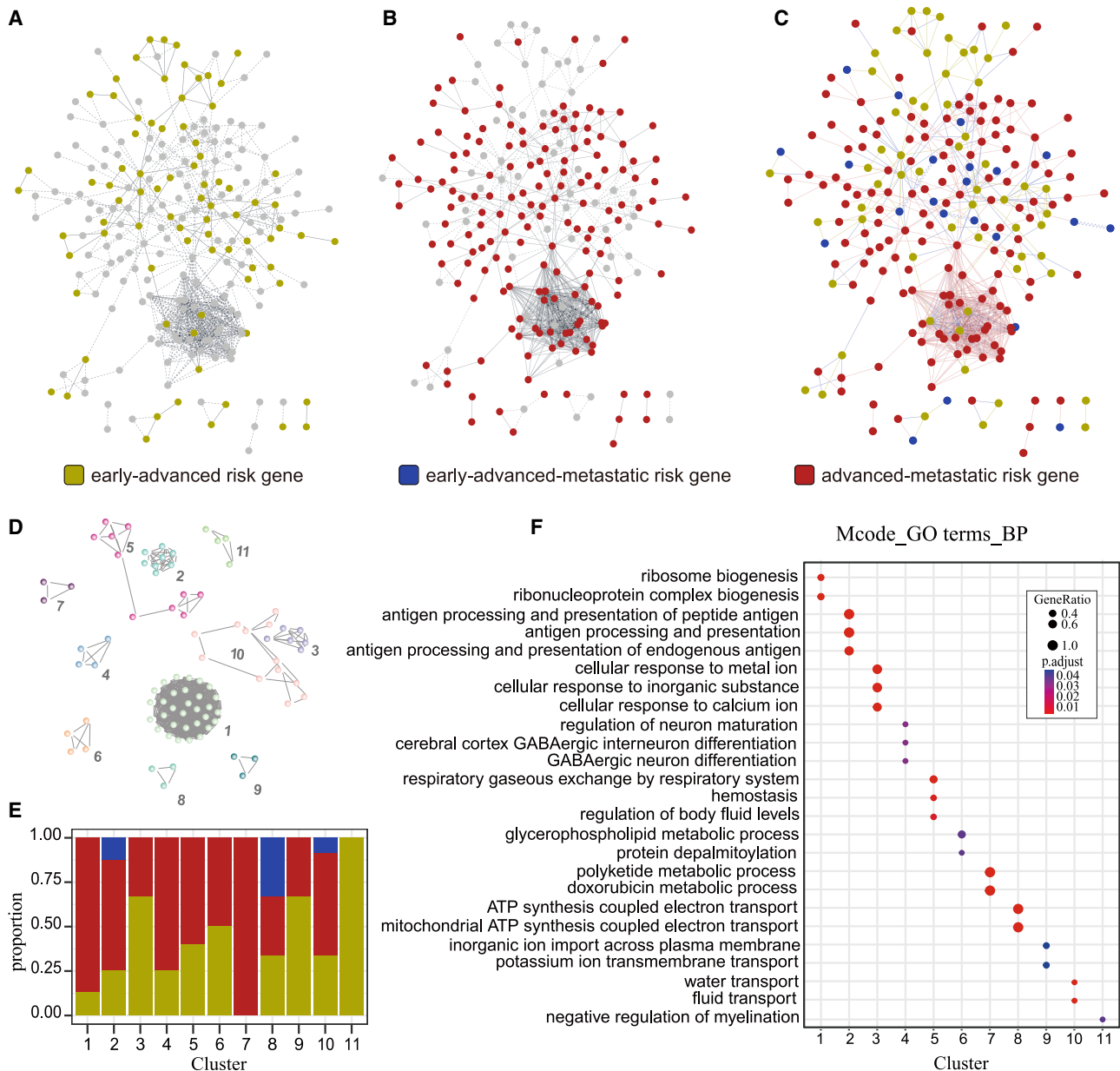


Figure 5. Construction and functional analysis of early advanced metastatic dynamic dysregulated network

(A) Early advanced dysregulated network. (B) Advanced metastatic dysregulated network. (C) Dynamic network of early advanced brain metastasis in LUAD. In the network, the greenyellow node is early advanced risk gene, the red node is advanced metastatic risk gene, and the blue nodes is early advanced metastatic risk gene. (D) Key modules identified in the network by MCODE. (E) Histogram of the proportion of different types of genes in key modules. (F) Significant enrichment of the GO terms for module genes. Filled colors from blue to red represent adjusted p values.

can bind to receptors on CD4⁺ T cells to regulate the differentiation and proliferation of cytotoxic T lymphocytes and other lymphocyte subsets.^{32,33} Previous studies have shown that lower expression of *HLA-DRB1* in patients with advanced stage than early stage of NSCLC.³⁴ Previous studies have demonstrated that the loss of *SFTPC* is more likely to be a marker associated with advanced lung tumors. The downregulation of *SFTPC* can promote cell proliferation and pre-

dict a poor survival rate in LUAD.³⁵ *FOLR1* was reported to be decreased in advanced p-stage than in early p-stage of NSCLC and its higher expression is associated with a better prognosis in early stage NSCLC.³⁶ We found that *RPS15*, *RPS11*, *RPL13A*, and *RPS24* were significantly downregulated from advanced to metastatic LUAD. These genes are involved in the EIF2 signaling pathway that related to angiogenesis and tumor metastasis.³⁷ Patients with the

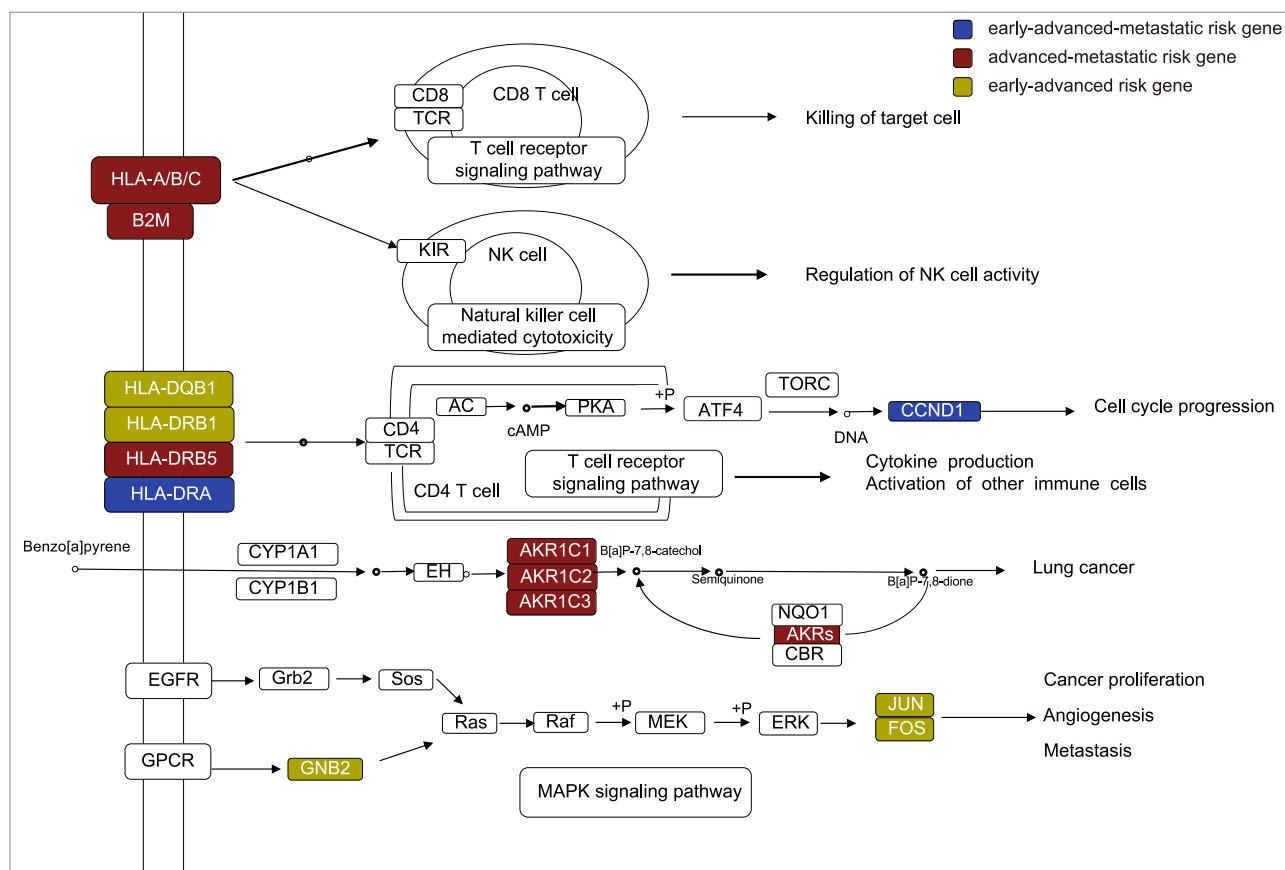


Figure 6. Stage-specific risk genes were associated with pathways contributing to LUAD progression

Stage-specific risk genes affect multiple important biological pathways. Filled colors represent stage-specific risk genes, with green/yellow for early advanced risk genes, red for advanced metastatic risk genes, and blue for early advanced metastatic risk genes.

lower expression of *RPS15*, *RPS11*, *RPL13A*, and *RPS24* genes had a better prognosis. *HLA-DRB5* is downregulated in the progression of brain metastasis, which is relevant to the development of cancer and metastatic progression,³⁸ and the low expression of *HLA-DRB5* is associated with poor prognosis. We found that *LYPLA1*, *KCNJ15*, and *PSMA3* were significantly upregulated from advanced to metastatic LUAD. High expression of *LYPLA1* and *PSMA3* were associated with a poor prognosis in LUAD. *LYPLA1* has been reported to be significantly differentially expressed between early and late recurrence groups in colorectal cancer.³⁹ Inhibition of *LYPLA1* expression can inhibit cell proliferation and migration in NSCLC.⁴⁰ *KCNJ15* has been reported to be associated with the epithelial-mesenchymal transition.⁴¹ *PSMA3* is involved in some metabolic pathways and processes, and plays an important role in the occurrence, development, and metastasis of cancer.⁴² The expression of *HLA-DRA* demonstrated downregulation from normal tissue, early stage to advanced stage, and lowest in brain metastasis of LUAD (Figure 8). A lower expression of *HLA-DRA* in LUAD tumors was associated with reduced patient survival (log-rank test, $p = 1.5e-3$) (Figure 7). Recent studies have shown that reduced *HLA-DRA* correlated with a lower immune enrichment score of $CD4^+$ T cells⁴³ and could predict the

response to anti-programmed death-1 immunotherapy in NSCLC.⁴⁴ Furthermore, we examined the association of these prognostic genes with the drug sensitivity profiles of NSCLC chemotherapeutic drugs in 135 NSCLC cell lines. We found that prognostic genes showed significant differential expression in mRNA and protein level in sensitive vs. resistant NSCLC cell lines (2-sample t-tests; $p < 0.05$). For example, *HLA-DRA* mRNA expression was associated with sensitivity to gemcitabine and *HLA-DRB1* mRNA expression was associated with resistance to carboplatin. The protein expression of *HLA-DRA* was associated with resistance to paclitaxel and sensitivity to gefitinib.⁴⁵ Altogether, stage-specific risk genes could serve as prognostic markers and play important roles in different stages of LUAD progression.

DISCUSSION

LUAD is the most common subtype of lung cancer, which has a poor prognosis. In recent years, increasing research progress on LUAD have been achieved by facilitating the analysis of single-cell transcriptome sequencing. For example, single-cell transcriptomic studies revealed substantial heterogeneity in cellular composition of the tumor microenvironment^{8,46,47} and molecular properties in primary and

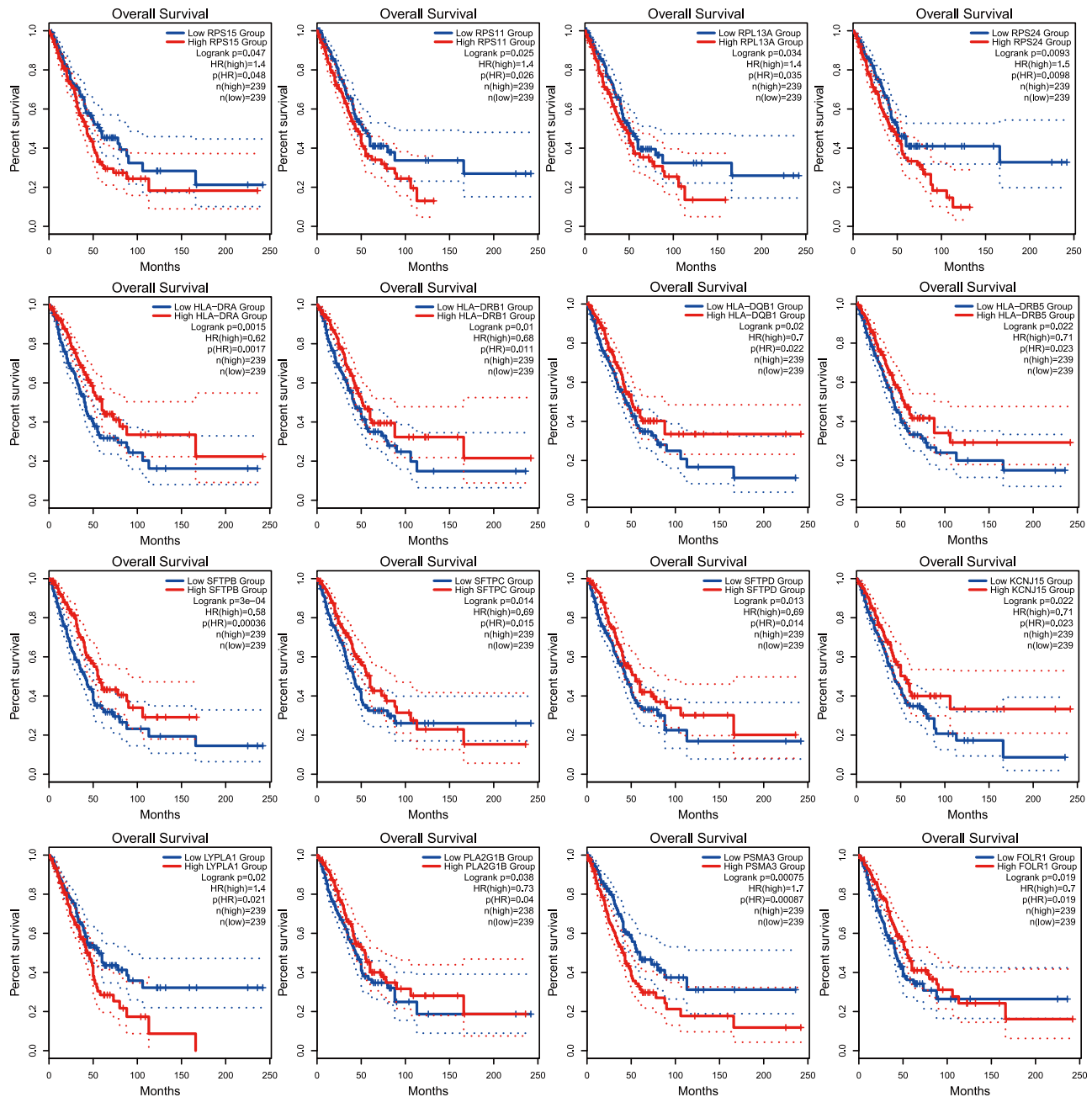


Figure 7. Identification of stage-specific prognostic genes during LUAD metastatic progression

Comparison of overall survival among patients with high (red) or low (blue) expression levels of risk genes by Kaplan-Meier analysis (with log rank values) in the cohort of LUAD cancer patients from TCGA.

metastatic LUAD.^{48,49} Liu et al. demonstrated intra-patient and intra-tumor heterogeneity in the regulation of pathways related to LUAD tumor progression.⁹ Many studies have characterized the cell differentiation trajectory and topological architecture of LUAD evolution.^{50,51} Several studies have started to characterize the molecular and cellular dynamics at different stages of LUAD progression.^{16,50,52–54} However, the dynamic gene regulation and molecular

mechanism driving the progression of LUAD from early stage, advanced stage to metastatic stage have not been comprehensively elucidated.

Therefore, we provided an integrative approach to characterize dynamic dysregulated networks and identify stage-specific risk genes during LUAD progression from early stage, advanced stage, to

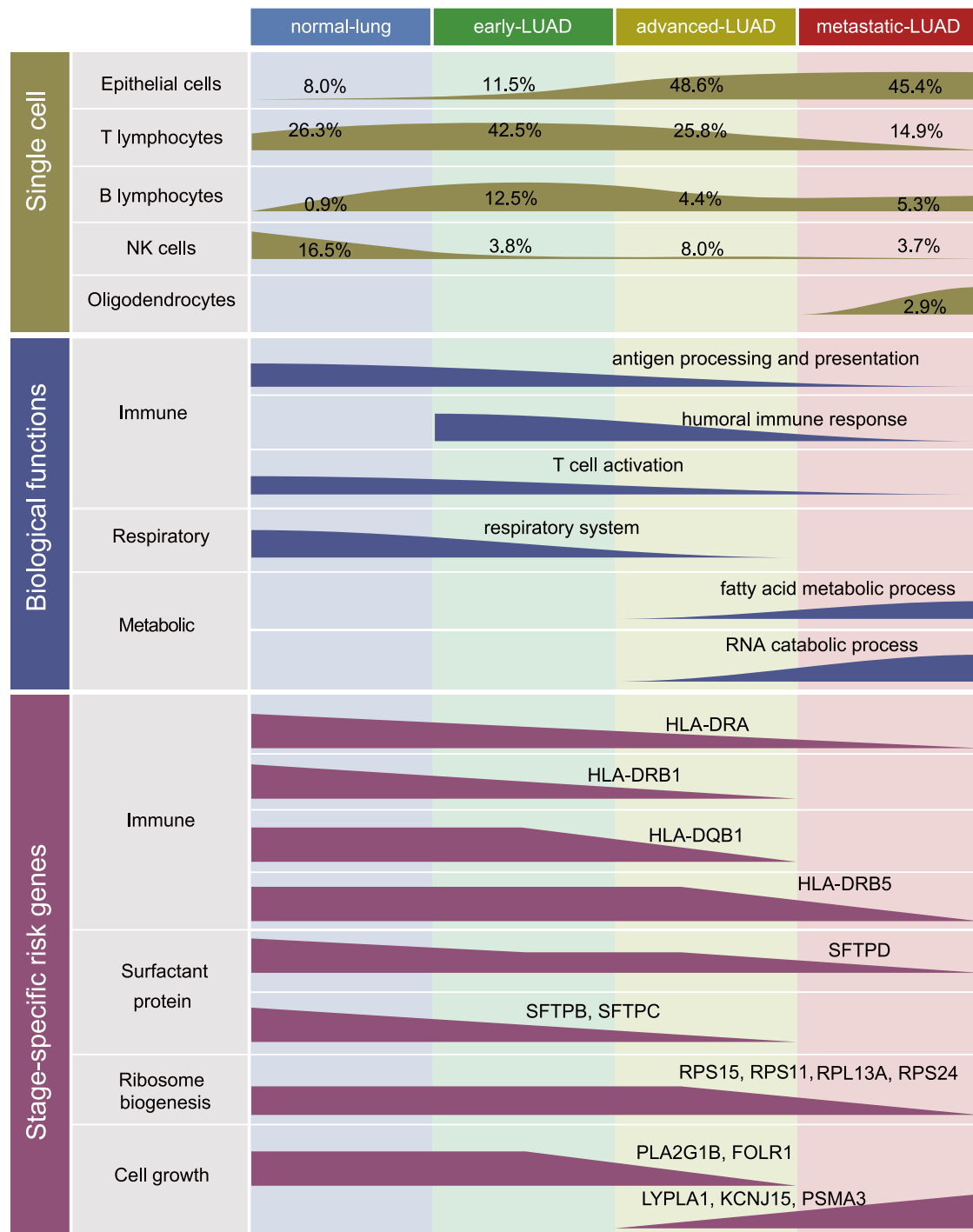


Figure 8. The dynamic changes in different stages of LUAD progression

Dynamic changes in tumor microenvironment cell composition (top), dysregulated biological functions (middle), and stage-specific risk genes (bottom) in different stages of normal lung, early LUAD, advanced LUAD, and metastatic LUAD.

metastatic stage. First, a trajectory analysis of MECs at different stages showed the increased transcriptional heterogeneity during LUAD progression and metastasis. We also identified LUAD stage-specific risk genes by comparing the tumor cell populations from different LUAD stage. Next, we constructed and comprehensively analyzed the early advanced metastatic dynamic dysregulated network in LUAD. We identified several functional modules, which involved in many distinct cancer-related biological functions and pathways. Previous studies have shown that various signaling pathways are associated with the occurrence and metastasis of NSCLC, including MAPK signaling pathway^{55,56} and T cell receptor signaling pathway,^{57–59} which was also confirmed by our study. Finally, we identified 6 early advanced risk genes, 8 advanced metastatic risk genes, and 2 common risk genes in different stages, which can be used as prognostic biomarkers in LUAD. For example, *SFTPB* showed significantly decreased expressions from early to advanced stage of LUAD. Decreased expression of *SFTPB* was associated with poorer prognosis in LUAD ($p = 3.0e-04$) (Figure 7). *SFTPB* has been reported to be used as potential blood biomarkers for distinguishing LUAD from lung squamous cell carcinoma.⁶⁰ These results highlight a crucial role of stage-specific risk genes in the development of LUAD.

In summary, we integrated scRNA-seq datasets from lung tissues, early stage, advanced stage, and brain metastasis LUAD tissues to characterize dynamic dysregulated networks, identify stage-specific risk genes and explores their prognostic value during LUAD tumor progression and metastasis. Our results reveal tumor heterogeneity in cellular composition, gene expression and biological functions at different stages of LUAD progression. We also report 16 stage-specific risk genes that could serve as prognostic biomarkers for LUAD (Figure 8). This study provides an in-depth analysis of the dynamic molecular characteristics and functions in the progression of early-advanced-brain metastasis in LUAD, providing an important theoretical basis for the molecular mechanism research and targeted therapy of LUAD.

MATERIALS AND METHODS

scRNA-seq data preprocessing

Single-cell transcriptome data of LUAD were obtained based on GEO database (GEO: GSE131907).¹⁶ The sequencing data contained 208,506 cells from 58 LUAD samples from 44 patients. These cells were derived from normal lung tissue, lymph nodes, primary tumor tissue, brain metastasis, and pleural effusion. In this study, cells from normal lung tissue (normal lung), early primary LUAD tissue (early LUAD), advanced primary LUAD tissue (advanced LUAD), and brain metastasis tissue (metastatic LUAD) were selected, including 42,995 cells from normal lung tissue, 45,149 cells from early primary LUAD tissue, 12,073 cells from advanced primary LUAD tissue, and 29,060 cells from brain metastasis. scRNA-seq data obtained from GEO were preprocessed according to different tissue sources using R package Seurat (<https://satijalab.org/seurat/>),⁶¹ which is widely used in preprocessing, dimension reduction, and cell clustering for scRNA-seq data. R package Seurat was used to calculate the percentage of mitochondrial reads. Cells that had unique feature counts

higher than 5,000 or less than 200 or a single cell with more than 10% mitochondrial genes were filtered out. After moving the low-quality cells, the filtered single cell data was normalized by the `NormalizeData` function.

Unsupervised reduction and clustering

We used the `RunPCA` function of the Seurat R package to determine the principal components (PCs) of cells at different stages of LUAD, and used the `ElbowPlot` function to visualize PCs. Furthermore, cells with similar expressions were clustered by the `FindClusters` function, and the shared nearest neighbor algorithm was used to identify cell clusters. Finally, the identity of each cluster was identified through t-distributed random neighborhood embedding visualization of cell clusters, combined with the expression values of known marker genes of various cell clusters.

Copy number variation was inferred based on scRNA-seq data

Cancer cells were identified by calculating single-cell copy number variations (CNVs) using the `inferCNV` algorithm,⁶² which is used to explore the expression intensity of genes across positions of tumor genomes in comparison to a set of reference normal cells to determine somatic changes.⁶³ First, non-malignant cells, including stromal cells and immune cells, were selected as reference cells to estimate the CNV of malignant cells. For epithelial cells treated by `inferCNV`, we selected 2 parameters to summarize CNV signals¹⁶: (1) mean squares: the mean squares of estimated CNV signals across all genomic locations; (2) copy number correlation coefficient: the Pearson correlation coefficient between CNV signal and the average CNV signal of the top 5% of cells. Epithelial cells with abnormal CNV signals (mean square of >0.02 and copy number correlation coefficient of >0.4) were defined as malignant tumor cells.

Single-cell trajectory construction and analysis

Monocle is an R package for analyzing scRNA-seq, which can arrange cells in simulated chronological order to show their developmental trajectories such as cell differentiation and other biological processes.⁶⁴ In our study, Monocle was used to construct cell trajectories by integrating epithelial cells from normal lung tissues and MECs from `inferCNV` inferred early primary LUAD, advanced primary LUAD, and brain metastasis tissues. The first step is to select genes that provide important information for defining trajectory progression, also known as feature selection. Then, the above gene sets were used to construct the data trajectory. We took the UMI counts of the gene-cell matrix as input to Monocle, and then its `newCellDataSet` function was called to create an object with the parameter `expressionFamily = negbinomial.size`. Then the `differentialGeneTest` function is used to calculate the differentially expressed genes (DEGs) of each cell type as the gene set to define the progression of the trajectory, and dimension reduction and cell ordering are carried out. Finally, the dynamic trajectory of epithelial cells in LUAD progression is deduced. Trajectory analysis will reveal the heterogeneity in gene expression and biological functions among MECs at different stages of LUAD, which is termed as different cell expression patterns.

Differential expression analysis and functional analysis

The FindMarkers function of the Seurat package was used to identify DEGs with the statistical threshold of FC of more than 2 and FDR of less than 0.05. R package clusterprofiler⁶⁵ and DAVID (<https://david.ncicrf.gov/>) to gene ontology (GO) enrichment analysis and Kyoto Encyclopedia of Genes and Genomes pathway enrichment analysis, respectively.

Identification of stage-specific risk genes

Genes that were significantly differentially expressed normal relative to early MECS with state 1 in the trajectory were defined as normal early risk genes. Genes that were significantly differentially expressed in state 1 relative to state 2 in the trajectory of early advanced LUAD were defined as early advanced risk genes. Genes that were significantly differentially expressed in state 3 relative to state 2 in the trajectory of advanced metastatic LUAD were defined as advanced metastatic risk genes. The common DEGs between early advanced stage and advanced metastatic stage were defined as early advanced metastatic risk genes.

Construct a dynamic dysregulated network of early advanced brain metastasis of LUAD

The early advanced dysregulated network is constructed by integrating interaction relationships mediated by early advanced risk genes. The advanced metastatic dysregulated network is constructed by integrating interaction relationships mediated by advanced metastatic risk genes. The early advanced metastatic dynamic dysregulated network is constructed by integrating early advanced and advanced metastatic dysregulated networks. The interaction relationships are obtained from the STRING database,⁶⁶ and Cytoscape⁶⁷ software was used to visualize the network.

Identification of key modules in the dynamic dysregulated network

Based on the dynamic network of early-advanced-brain metastasis, we used molecular complex detection (MCODE) (a plugin in Cytoscape) to identify the most significant modules in the network, and the parameters of MCODE are as follows: degree cutoff = 2, node score cutoff = 0.2, K-core = 2, and max depth = 100.

Survival analysis

We used the "Survival Analysis" component of GEPIA2⁶⁸ (<http://gepia2.cancer-pku.cn/#index>), which can be used to assess the relationship between gene expression and cancer prognosis based on gene expression levels and to evaluate hypotheses using log-rank tests, to detect OS (overall Survival) of risk genes in TCGA. The thresholds for distinguishing low expression group and high expression group were defined as cutoff-high (50%) and cutoff-low (50%), respectively, and genes with p-values of less than 0.05 were selected as prognostic markers.

DATA AVAILABILITY STATEMENT

All data and codes in this study are available under proper request; please contact wangli@hrbmu.edu.cn.

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.omtn.2022.11.019>.

ACKNOWLEDGMENTS

The authors are grateful to all those who contributed to this study and also thank all the funding that provided financial support for this study.

Supported by the Training Program for Young Innovative Talents in Regular Undergraduate colleges and universities of Heilongjiang Province (UNPYSCT-2020174); the China Postdoctoral Science Special Foundation (2019T120280 and 2020T130161); the Hei Long Jiang Postdoctoral Special Foundation (LBH-TZ1018 and LBH-Z17218); the China Postdoctoral Science Foundation (2018M631943 and 2018M641860); and the Hei Long Jiang Postdoctoral Foundation (LBH-Z17110).

AUTHOR CONTRIBUTIONS

H.Y.Z., S.W.N., and L.W. designed the study, implemented the algorithm, and performed the analysis. H.Y.Z., W.Y.L., L.W., and L.L. wrote and revised the manuscript. Y.X.Y., K.L.L., L.X.W., L.B., H.T.X., S.H.L., K.F., X.Y.Z., M.T.F., and C.Y.Z. helped to collect the data and prepare the figures and tables. All authors read, reviewed, and approved the final manuscript.

DECLARATION OF INTERESTS

None declared.

REFERENCES

1. Sher, T., Dy, G.K., and Adjei, A.A. (2008). Small cell lung cancer. *Mayo Clin. Proc.* 83, 355–367. <https://doi.org/10.4065/83.3.355>.
2. Herbst, R.S., Heymach, J.V., and Lippman, S.M. (2008). Lung cancer. *N. Engl. J. Med.* 359, 1367–1380. <https://doi.org/10.1056/NEJMra0802714>.
3. Shah, R., Sabanathan, S., Richardson, J., Mearns, A.J., and Goulden, C. (1996). Results of surgical treatment of stage I and II lung cancer. *J. Cardiovasc. Surg.* 37, 169–172.
4. Walters, S., Maringe, C., Coleman, M.P., Peake, M.D., Butler, J., Young, N., Bergström, S., Hanna, L., Jakobsen, E., Kölbek, K., et al.; ICBP Module 1 Working Group (2013). Lung cancer survival and stage at diagnosis in Australia, Canada, Denmark, Norway, Sweden and the UK: a population-based study, 2004–2007. *Thorax* 68, 551–564. <https://doi.org/10.1136/thoraxjnl-2012-202297>.
5. You, H., Baluszek, S., and Kaminska, B. (2019). Immune microenvironment of brain metastases—are microglia and other brain macrophages little helpers? *Front. Immunol.* 10, 1941. <https://doi.org/10.3389/fimmu.2019.01941>.
6. Muliaditan, T., Caron, J., Okesola, M., Opzoomer, J.W., Kosti, P., Georgouli, M., Gordon, P., Lall, S., Kuzeva, D.M., Pedro, L., et al. (2018). Macrophages are exploited from an innate wound healing response to facilitate cancer metastasis. *Nat. Commun.* 9, 2951. <https://doi.org/10.1038/s41467-018-05346-7>.
7. Bhatt, V.R., Kedia, S., Kessinger, A., and Ganti, A.K. (2013). Brain metastasis in patients with non-small-cell lung cancer and epidermal growth factor receptor mutations. *J. Clin. Oncol.* 31, 3162–3164. <https://doi.org/10.1200/JCO.2013.49.8915>.
8. Bischoff, P., Trinks, A., Obermayer, B., Pett, J.P., Wiederspahn, J., Uhlitz, F., Liang, X., Lehmann, A., Jurmeister, P., Elsner, A., et al. (2021). Single-cell RNA sequencing reveals distinct tumor microenvironmental patterns in lung adenocarcinoma. *Oncogene* 40, 6748–6758. <https://doi.org/10.1038/s41388-021-02054-3>.
9. Liu, Y., Ye, G., Huang, L., Zhang, C., Sheng, Y., Wu, B., Han, L., Wu, C., Dong, B., and Qi, Y. (2020). Single-cell transcriptome analysis demonstrates inter-patient and intra-tumor heterogeneity in primary and metastatic lung adenocarcinoma. *Aging* 12, 21559–21581. <https://doi.org/10.18632/aging.103945>.

10. Fan, M., Xia, P., Clarke, R., Wang, Y., and Li, L. (2020). Radiogenomic signatures reveal multiscale intratumour heterogeneity associated with biological functions and survival in breast cancer. *Nat. Commun.* *11*, 4861. <https://doi.org/10.1038/s41467-020-18703-2>.
11. Cook, D.J., Kallus, J., Jörnsten, R., and Nielsen, J. (2020). Molecular natural history of breast cancer: leveraging transcriptomics to predict breast cancer progression and aggressiveness. *Cancer Med.* *9*, 3551–3562. <https://doi.org/10.1002/cam4.2996>.
12. Lin, Y., Chen, D., Ding, Q., Zhu, X., Zhu, R., and Chen, Y. (2021). [Progress in single-cell RNA sequencing of lung adenocarcinoma]. *Zhongguo Fei Ai Za Zhi* *24*, 434–440. <https://doi.org/10.3779/j.issn.1009-3419.2021.102.19>.
13. Li, L., and Clevers, H. (2010). Coexistence of quiescent and active adult stem cells in mammals. *Science* *327*, 542–545. <https://doi.org/10.1126/science.1180794>.
14. Shalek, A.K., Satija, R., Shuga, J., Trombetta, J.J., Gennert, D., Lu, D., Chen, P., Gertner, R.S., Gaubblomme, J.T., Yosef, N., et al. (2014). Single-cell RNA-seq reveals dynamic paracrine control of cellular variation. *Nature* *510*, 363–369. <https://doi.org/10.1038/nature13437>.
15. Hwang, B., Lee, J.H., and Bang, D. (2018). Single-cell RNA sequencing technologies and bioinformatics pipelines. *Exp. Mol. Med.* *50*, 1–14. <https://doi.org/10.1038/s12276-018-0071-8>.
16. Kim, N., Kim, H.K., Lee, K., Hong, Y., Cho, J.H., Choi, J.W., Lee, J.I., Suh, Y.L., Ku, B.M., Eum, H.H., et al. (2020). Single-cell RNA sequencing demonstrates the molecular and cellular reprogramming of metastatic lung adenocarcinoma. *Nat. Commun.* *11*, 2285. <https://doi.org/10.1038/s41467-020-16164-1>.
17. Ruaro, B., Salton, F., Braga, L., Wade, B., Confalonieri, P., Volpe, M.C., Baratella, E., Maiocchi, S., and Confalonieri, M. (2021). The history and mystery of alveolar epithelial type II cells: focus on their physiologic and pathologic role in lung. *Int. J. Mol. Sci.* *22*, 2566. <https://doi.org/10.3390/ijms22052566>.
18. Yu, J.W., Hoffman, S., Beal, A.M., Dykon, A., Ringenberg, M.A., Hughes, A.C., Dare, L., Anderson, A.D., Finger, J., Kasparcova, V., et al. (2015). MALT1 protease activity is required for innate and adaptive immune responses. *PLoS One* *10*, e0127083. <https://doi.org/10.1371/journal.pone.0127083>.
19. Chakrabarty, S., Kahler, J.P., van de Plassche, M.A.T., Vanhoutte, R., and Verhelst, S.H.L. (2019). Recent advances in activity-based protein profiling of proteases. *Curr. Top. Microbiol. Immunol.* *420*, 253–281. https://doi.org/10.1007/82_2018_138.
20. Ma, X.B., Xu, Y.Y., Zhu, M.X., and Wang, L. (2020). Prognostic signatures based on thirteen immune-related genes in colorectal cancer. *Front. Oncol.* *10*, 591739. <https://doi.org/10.3389/fonc.2020.591739>.
21. Kargl, J., Busch, S.E., Yang, G.H.Y., Kim, K.H., Hanke, M.L., Metz, H.E., Hubbard, J.J., Lee, S.M., Madtes, D.K., McIntosh, M.W., and Houghton, A.M. (2017). Neutrophils dominate the immune cell composition in non-small cell lung cancer. *Nat. Commun.* *8*, 14381. <https://doi.org/10.1038/ncomms14381>.
22. Mollinedo, F. (2019). Neutrophil degranulation, plasticity, and cancer metastasis. *Trends Immunol.* *40*, 228–242. <https://doi.org/10.1016/j.it.2019.01.006>.
23. Zhang, C., Liao, Y., Liu, P., Du, Q., Liang, Y., Ooi, S., Qin, S., He, S., Yao, S., and Wang, W. (2020). FABP5 promotes lymph node metastasis in cervical cancer by reprogramming fatty acid metabolism. *Theranostics* *10*, 6561–6580. <https://doi.org/10.7150/thno.44868>.
24. Weber, M.R., Zuka, M., Lorgner, M., Tschan, M., Torbett, B.E., Zijlstra, A., Quigley, J.P., Staffin, K., Eliceiri, B.P., Krueger, J.S., et al. (2016). Activated tumor cell integrin α v β 3 cooperates with platelets to promote extravasation and metastasis from the blood stream. *Thromb. Res.* *140* (Suppl 1), S27–S36. [https://doi.org/10.1016/S0049-3848\(16\)30095-0](https://doi.org/10.1016/S0049-3848(16)30095-0).
25. Egan, K., Crowley, D., Smyth, P., O'Toole, S., Spillane, C., Martin, C., Gallagher, M., Canney, A., Norris, L., Conlon, N., et al. (2011). Platelet adhesion and degranulation induce pro-survival and pro-angiogenic signalling in ovarian cancer cells. *PLoS One* *6*, e26125. <https://doi.org/10.1371/journal.pone.0026125>.
26. Philippe, C., Philippe, B., Fouqueray, B., Perez, J., Le Bret, M., and Baud, L. (1993). Protection from tumor necrosis factor-mediated cytotoxicity by platelets. *Am. J. Pathol.* *143*, 1713–1723.
27. Jongsma, M.L.M., Guarda, G., and Spaepen, R.M. (2019). The regulatory network behind MHC class I expression. *Mol. Immunol.* *113*, 16–21. <https://doi.org/10.1016/j.molimm.2017.12.005>.
28. Stacey, D.W. (2003). Cyclin D1 serves as a cell cycle regulatory switch in actively proliferating cells. *Curr. Opin. Cell Biol.* *15*, 158–163. [https://doi.org/10.1016/s0955-0674\(03\)00008-5](https://doi.org/10.1016/s0955-0674(03)00008-5).
29. Dixon, S.J. (2017). Ferroptosis: bug or feature? *Immunol. Rev.* *277*, 150–157. <https://doi.org/10.1111/imr.12533>.
30. Fearnhead, H.O., Vandenabeele, P., and Vanden Berghe, T. (2017). How do we fit ferroptosis in the family of regulated cell death? *Cell Death Differ.* *24*, 1991–1998. <https://doi.org/10.1038/cdd.2017.149>.
31. Wohlhieter, C.A., Richards, A.L., Uddin, F., Hulton, C.H., Quintanal-Villalonga, À., Martin, A., de Stanchina, E., Bhanot, U., Asher, M., Shah, N.S., et al. (2020). Concurrent mutations in STK11 and KEAP1 promote ferroptosis protection and SCD1 dependence in lung cancer. *Cell Rep.* *33*, 108444. <https://doi.org/10.1016/j.celrep.2020.108444>.
32. Mangalam, A.K., Taneja, V., and David, C.S. (2013). HLA class II molecules influence susceptibility versus protection in inflammatory diseases by determining the cytokine profile. *J. Immunol.* *190*, 513–518. <https://doi.org/10.4049/jimmunol.1201891>.
33. Zhao, H., Shi, J., Zhang, Y., Xie, A., Yu, L., Zhang, C., Lei, J., Xu, H., Leng, Z., Li, T., et al. (2020). LncTarD: a manually-curated database of experimentally-supported functional lncRNA-target regulations in human diseases. *Nucleic Acids Res.* *48*, D118–D126. <https://doi.org/10.1093/nar/gkz985>.
34. Bulut, I., Meral, M., Kaynar, H., Pirim, I., Bilici, M., and Gorguner, M. (2009). Analysis of HLA class I and II alleles regarding to lymph node and distant metastasis in patients with non-small cell lung cancer. *Lung Cancer* *66*, 231–236. <https://doi.org/10.1016/j.lungcan.2009.01.012>.
35. Li, B., Meng, Y.Q., Li, Z., Yin, C., Lin, J.P., Zhu, D.J., and Zhang, S.B. (2019). MiR-629-3p-induced downregulation of SFTPC promotes cell proliferation and predicts poor survival in lung adenocarcinoma. *Artif. Cells Nanomed. Biotechnol.* *47*, 3286–3296. <https://doi.org/10.1080/21691401.2019.1648283>.
36. Iwakiri, S., Sonobe, M., Nagai, S., Hirata, T., Wada, H., and Miyahara, R. (2008). Expression status of folate receptor alpha is significantly correlated with prognosis in non-small-cell lung cancers. *Ann. Surg. Oncol.* *15*, 889–899. <https://doi.org/10.1245/s10434-007-9755-3>.
37. Alexandrova, E., Pecoraro, G., Sellitto, A., Melone, V., Ferravante, C., Rocco, T., Guacci, A., Giurato, G., Nassa, G., Rizzo, F., et al. (2020). An Overview of candidate therapeutic target genes in ovarian cancer. *Cancers* *12*, E1470. <https://doi.org/10.3390/cancers12061470>.
38. Yang, X.X., Pan, H.Z., Li, P.Y., Li, F.X., Xu, W.W., Wu, Y.S., Yao, G.Y., and Li, M. (2011). HLA class II variants in Chinese breast cancer patients. *Asian Pac. J. Cancer Prev.* *12*, 3075–3079.
39. Snoeren, N., Emmink, B.L., Koerkamp, M.J.G., van Hooff, S.R., Goos, J.A.C.M., van Houdt, W.J., de Wit, M., Prins, A.M., Piersma, S.R., Pham, T.V., et al. (2013). Maspin is a marker for early recurrence in primary stage III and IV colorectal cancer. *Br. J. Cancer* *109*, 1636–1647. <https://doi.org/10.1038/bjc.2013.489>.
40. Mohammed, A., Zhang, C., Zhang, S., Shen, Q., Li, J., Tang, Z., and Liu, H. (2019). Inhibition of cell proliferation and migration in nonsmall cell lung cancer cells through the suppression of LYPLA1. *Oncol. Rep.* *41*, 973–980. <https://doi.org/10.3892/or.2018.6857>.
41. Nakamura, S., Kanda, M., Koike, M., Shimizu, D., Umeda, S., Hattori, N., Hayashi, M., Tanaka, C., Kobayashi, D., Yamada, S., et al. (2020). KCNJ15 expression and malignant behavior of esophageal squamous cell carcinoma. *Ann. Surg. Oncol.* *27*, 2559–2568. <https://doi.org/10.1245/s10434-019-08189-8>.
42. Wong, Y.H., Chen, R.H., and Chen, B.S. (2014). Core and specific network markers of carcinogenesis from multiple cancer samples. *J. Theor. Biol.* *362*, 17–34. <https://doi.org/10.1016/j.jtbi.2014.05.045>.
43. Wang, A.R., Khullar, S., Brown, J., Baschnagel, A., Buehler, D., Kendzierski, C., and Iyer, G. (2022). Abstract 3859: remodeling the extracellular matrix environment enables the dissemination of primary tumor cells through a chemokine gradient to establish brain metastasis in non-small cell lung cancer adenocarcinoma. *Cancer Res.* *82*, 3859.
44. Mei, J., Jiang, G., Chen, Y., Xu, Y., Wan, Y., Chen, R., Liu, F., Mao, W., Zheng, M., and Xu, J. (2022). HLA class II molecule HLA-DRA identifies immuno-hot tumors and predicts the therapeutic response to anti-PD-1 immunotherapy in NSCLC. *BMC Cancer* *22*, 738. <https://doi.org/10.1186/s12885-022-09840-6>.

45. Ye, Q., and Guo, N.L. (2022). Single B cell gene Co-expression networks implicated in prognosis, proliferation, and therapeutic responses in non-small cell lung cancer bulk tumors. *Cancers* 14, 3123. <https://doi.org/10.3390/cancers14133123>.
46. He, D., Wang, D., Lu, P., Yang, N., Xue, Z., Zhu, X., Zhang, P., and Fan, G. (2021). Single-cell RNA sequencing reveals heterogeneous tumor and immune cell populations in early-stage lung adenocarcinomas harboring EGFR mutations. *Oncogene* 40, 355–368. <https://doi.org/10.1038/s41388-020-01528-0>.
47. Zhang, L., Zhang, Y., Wang, C., Yang, Y., Ni, Y., Wang, Z., Song, T., Yao, M., Liu, Z., Chao, N., et al. (2022). Integrated single-cell RNA sequencing analysis reveals distinct cellular and transcriptional modules associated with survival in lung cancer. *Signal Transduct. Target. Ther.* 7, 9. <https://doi.org/10.1038/s41392-021-00824-9>.
48. Ma, K.Y., Schonesen, A.A., Brock, A., Van Den Berg, C., Eckhardt, S.G., Liu, Z., and Jiang, N. (2019). Single-cell RNA sequencing of lung adenocarcinoma reveals heterogeneity of immune response-related genes. *JCI Insight* 4, 121387. <https://doi.org/10.1172/jci.insight.121387>.
49. Li, Q., Wang, R., Yang, Z., Li, W., Yang, J., Wang, Z., Bai, H., Cui, Y., Tian, Y., Wu, Z., et al. (2022). Molecular profiling of human non-small cell lung cancer by single-cell RNA-seq. *Genome Med.* 14, 87. <https://doi.org/10.1186/s13073-022-01089-9>.
50. Sinjab, A., Han, G., Treekitkarnmongkol, W., Hara, K., Brennan, P.M., Dang, M., Hao, D., Wang, R., Dai, E., Dejima, H., et al. (2021). Resolving the spatial and cellular architecture of lung adenocarcinoma by multiregion single-cell sequencing. *Cancer Discov.* 11, 2506–2523. <https://doi.org/10.1158/2159-8290.CD-20-1285>.
51. Luo, Y., Deng, X., Que, J., Li, Z., Xie, W., Dai, G., Chen, L., and Wang, H. (2022). Cell trajectory-related genes of lung adenocarcinoma predict tumor immune microenvironment and prognosis of patients. *Front. Oncol.* 12, 911401. <https://doi.org/10.3389/fonc.2022.911401>.
52. Chen, Z., Huang, Y., Hu, Z., Zhao, M., Li, M., Bi, G., Zheng, Y., Liang, J., Lu, T., Jiang, W., et al. (2021). Landscape and dynamics of single tumor and immune cells in early and advanced-stage lung adenocarcinoma. *Clin. Transl. Med.* 11, e350. <https://doi.org/10.1002/ctm2.350>.
53. He, Y., Liu, X., Wang, H., Wu, L., Jiang, M., Guo, H., Zhu, J., Wu, S., Sun, H., Chen, S., et al. (2021). Mechanisms of progression and heterogeneity in multiple nodules of lung adenocarcinoma. *Small Methods* 5, e2100082. <https://doi.org/10.1002/smt.202100082>.
54. Wang, Z., Li, Z., Zhou, K., Wang, C., Jiang, L., Zhang, L., Yang, Y., Luo, W., Qiao, W., Wang, G., et al. (2021). Deciphering cell lineage specification of human lung adenocarcinoma with single-cell RNA sequencing. *Nat. Commun.* 12, 6500. <https://doi.org/10.1038/s41467-021-26770-2>.
55. Zhu, Z.J., and He, J.K. (2018). TINCER facilitates non-small cell lung cancer progression through BRAF-activated MAPK pathway. *Biochem. Biophys. Res. Commun.* 497, 971–977. <https://doi.org/10.1016/j.bbrc.2018.02.059>.
56. Pradhan, R., Singhvi, G., Dubey, S.K., Gupta, G., and Dua, K. (2019). MAPK pathway: a potential target for the treatment of non-small-cell lung carcinoma. *Future Med. Chem.* 11, 793–795. <https://doi.org/10.4155/fmc-2018-0468>.
57. Chang, W.A., Tsai, Y.M., Tsai, Y.C., Wu, C.Y., Chang, K.F., Lien, C.T., Hung, J.Y., Hsu, Y.L., and Kuo, P.L. (2019). Differential expression profiles of the transcriptome in bone marrow-derived cells in lung cancer revealed by next generation sequencing and bioinformatics. *Oncol. Lett.* 17, 4341–4350. <https://doi.org/10.3892/ol.2019.10085>.
58. Brownlie, R.J., and Zamoyska, R. (2013). T cell receptor signalling networks: branched, diversified and bounded. *Nat. Rev. Immunol.* 13, 257–269. <https://doi.org/10.1038/nri3403>.
59. Zhao, H., Liu, X., Yu, L., Lin, S., Zhang, C., Xu, H., Leng, Z., Huang, W., Lei, J., Li, T., et al. (2021). Comprehensive landscape of epigenetic-dysregulated lncRNAs reveals a profound role of enhancers in carcinogenesis in BC subtypes. *Mol. Ther. Nucleic Acids* 23, 667–681. <https://doi.org/10.1016/j.omtn.2020.12.024>.
60. Hanjani, N.A., Esmaelizad, N., Zanganeh, S., Gharavi, A.T., Heidarizadeh, P., Radfar, M., Omid, F., MacLoughlin, R., and Doroudian, M. (2022). Emerging role of exosomes as biomarkers in cancer treatment and diagnosis. *Crit. Rev. Oncol. Hematol.* 169, 103565. <https://doi.org/10.1016/j.critrevonc.2021.103565>.
61. Butler, A., Hoffman, P., Smibert, P., Papalex, E., and Satija, R. (2018). Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat. Biotechnol.* 36, 411–420. <https://doi.org/10.1038/nbt.4096>.
62. Chung, W., Eum, H.H., Lee, H.O., Lee, K.M., Lee, H.B., Kim, K.T., Ryu, H.S., Kim, S., Lee, J.E., Park, Y.H., et al. (2017). Single-cell RNA-seq enables comprehensive tumour and immune cell profiling in primary breast cancer. *Nat. Commun.* 8, 15081. <https://doi.org/10.1038/ncomms15081>.
63. Patel, A.P., Tirosh, I., Trombetta, J.J., Shalek, A.K., Gillespie, S.M., Wakimoto, H., Cahill, D.P., Nahed, B.V., Curry, W.T., Martuza, R.L., et al. (2014). Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma. *Science* 344, 1396–1401. <https://doi.org/10.1126/science.1254257>.
64. Cao, J., Spielmann, M., Qiu, X., Huang, X., Ibrahim, D.M., Hill, A.J., Zhang, F., Mundlos, S., Christiansen, L., Steemers, F.J., et al. (2019). The single-cell transcriptional landscape of mammalian organogenesis. *Nature* 566, 496–502. <https://doi.org/10.1038/s41586-019-0969-x>.
65. Yu, G., Wang, L.G., Han, Y., and He, Q.Y. (2012). clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* 16, 284–287. <https://doi.org/10.1089/omi.2011.0118>.
66. Szklarczyk, D., Morris, J.H., Cook, H., Kuhn, M., Wyder, S., Simonovic, M., Santos, A., Doncheva, N.T., Roth, A., Bork, P., et al. (2017). The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible. *Nucleic Acids Res.* 45, D362–D368. <https://doi.org/10.1093/nar/gkw937>.
67. Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B., and Ideker, T. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* 13, 2498–2504. <https://doi.org/10.1101/gr.1239303>.
68. Tang, Z., Kang, B., Li, C., Chen, T., and Zhang, Z. (2019). GEPIA2: an enhanced web server for large-scale expression profiling and interactive analysis. *Nucleic Acids Res.* 47, W556–W560. <https://doi.org/10.1093/nar/gkz430>.