

Cross-Regional Transcriptome Data Reveal Transcriptional Abnormalities Associated with Lung Adenocarcinoma

Ling Zhang^{#1}, Cong Fan^{#2}, Huanxia Shang¹, Xiaojing Wang¹, Xin Zhang¹,
Qingzhe Qiao¹, Jiajia Ni^{*3}, Shucai Wu^{*1}

Abstract

Background: Lung cancer is the leading cause of cancer-related deaths worldwide, yet there has been little attention given to the correlation between the cancer transcriptome and the incidence and mortality of lung cancer across different geographic regions.

Methods: To analyze this correlation, we screened the transcriptome datasets of stage I lung adenocarcinoma (LAC) patients from the Lung Cancer Explorer and examined their correlation with the age-standardized incidence rate (ASIR), age-standardized mortality rate (ASMR), and mortality-to-incidence ratio (MIR).

Results: The expression difference rates (DRs) of certain genes (SPARCL1, SRPX, PMP22, MSR1, BST1, AKAP12, MAOB, vimentin, serglycin, ILK, ESD, transgelin, NCOA1, and PLPP1) were significantly negatively correlated with the ASIR of female LAC. Additionally, the DR of KRT19 was significantly positively correlated with the ASIR of female LAC. Furthermore, the DRs of COL10A1, SMAD7, COL3A1, and AQP1 were significantly positively correlated with the ASMR and MIR of female LAC, while the DR of KRT15 was significantly negatively correlated with the ASMR and MIR of female LAC. In male LAC patients, the DR of RGS2 was significantly negatively correlated with the ASIR, while the DRs of SPARCL1, COX7A1, IL3RA, and ADH1B were significantly positively correlated with the ASMR and MIR. Additionally, the DR of AIMP2 was significantly negatively correlated with the ASMR and MIR.

Conclusions: Our findings suggest that the expression levels of serglycin, ILK, ESD, and PLPD1 may play a significant role in the development of LAC. This information can be valuable for identifying potential treatment targets for lung cancer.

Keywords: Incidence Rate, Lung Adenocarcinoma, Transcriptome, Mortality Rate, Regional Difference.

Introduction

Lung cancer is the leading cause of cancer-related death worldwide (1,2). The incidence and mortality rates of different types of lung cancer vary across different regions of the world (3,4). Non-small cell lung cancer (NSCLC) is the most common type of lung cancer, accounting for approximately 85% of

all lung cancer cases worldwide (5,6). Lung adenocarcinoma is the most common subtype of NSCLC and is characterized by higher levels of gene transcription compared to other subtypes (7).

Multiple comparative studies have been conducted on the transcriptome of lung cancer

1: Respiratory Department, Hebei Chest Hospital, Shijiazhuang 050041, China.

2: Geriatric Respiratory Medicine Department, Hebei General Hospital, Shijiazhuang 050051, China.

3: Guangzhou Laboratory, Guangzhou 510005, China.

#The first and the second authors contributed equally to this work.

*Corresponding author: Jiajia Ni; Tel: +86 13925151156; E-mail: nijiajia2005@126.com and Shucai Wu; Tel: +86 13925151156; E-mail: shucaiwu2009@163.com.

Received: 3 Jan, 2024; Accepted: 5 Aug, 2025

and adjacent tissues (or healthy lung tissue) (7-10). These research results are valuable for understanding transcriptome changes in lung cancer. However, little attention has been given to the correlations between these transcriptome results and the incidence and mortality rates of lung cancer across geographic regions. Analyzing these correlations could help us more effectively explore the proteins and cytokines closely related to the occurrence and development of lung cancer, and provide important information for identifying potential treatment targets. Therefore, in this study, we analyzed the correlation between lung adenocarcinoma (LAC) transcriptome data and the age-standardized incidence rate (ASIR), age-standardized mortality rate (ASMR), and mortality-to-incidence ratio (MIR). We also analyzed the closely correlated proteins and cytokines.

Materials and Methods

Data collection

The procedures followed were in accordance with the ethical standards of the responsible committee on human experimentation (institutional or regional) and with the Helsinki Declaration of 1975, as revised in 2000 (available at http://www.wma.net/e/policy/17-c_e.html). The gene expression datasets of LAC with paired non-involved healthy lungs were screened from the Lung Cancer Explorer (LCE, <https://lce.biohpc.swmed.edu/lungcancer/>) (8). All datasets are standardized microarray data. ASIR, ASMR, and MIR data were obtained from Sharma (1).

Inclusion/exclusion criteria

Inclusion criteria: Patients with stage I LAC and clear sex information, along with paired healthy controls. Exclusion criteria: (1) Patients with lung cancer that was not stage I LAC; (2) Patients with stage I LAC without clear sex information and paired healthy controls; (3) Patients with LAC at a stage other than stage I. In total, 11 datasets were selected for further analysis, consisting of 867 LAC samples and 208 paired control samples.

Data analysis

Significantly different genes were detected using White's non-parametric t-test in the Statistical Analysis of Taxonomic and Functional Profiles (STAMP, version 2.1.3) (11). Traditional fold change (FC) is not suitable for measuring expression differences due to the negative value of standardized microarray data (12,13). Instead, we used the difference ratio (DR) to measure gene expression differences, calculated as:

$$DR = \frac{A_A - A_C}{A_C}$$

where DR is the difference ratio, and A_A and A_C represent the average levels of the adenocarcinoma and control samples, respectively. To further analyze the data, we conducted correlation analysis, principal component analysis (PCA), and K-means clustering analysis using R packages basicTrendline, vegan, and factoextra, respectively. Welch two sample t-test was used to detect the difference in age between male and female participants. Heatmap profiles were generated using the R package pheatmap. Statistical significance was determined at $P < 0.05$.

Results

Except for one dataset with no participant age information and one dataset with only female data, all other datasets contain both male and female age information, and except for one dataset with significant age difference between male and female (Welch two sample t-test, $P < 0.05$; Table 1). The consistency of DR in male and female patients can be determined by a significant linear correlation between the two. However, the results of the correlation analysis showed that the DRs in male patients from two European countries (Italy and Switzerland) were not significantly correlated with those in female patients ($P > 0.05$; Figs. 1a and 1b). In contrast, the DRs in male patients from six North American datasets (one Canadian and five USA) were significantly correlated with those in paired female patients ($P < 0.05$; Fig. 1c-h). Interestingly, both DRs in male patients from the two Asian countries showed a significant correlation with those in female

patients ($P < 0.05$; Figs. 1i and 1j). It is worth noting that only one correlation line had a slope of approximately 1 (Fig. 1e). While the P values of the two North American datasets were less than 0.05, the slopes were approximately 0, indicating that the DRs in male patients did not significantly correlate with those in female

patients (Fig. 1g and 1h). These findings suggest that the pathogenesis of LAC in males and females in Asia and some parts of North America may be consistent, while there may be significant differences in the pathogenesis of LAC between males and females in Europe.

Table 1. Age difference between male and female participants in each group. The age data is displayed in mean \pm standard error. These data were statistically tested using Welch two sample t-test. * $P < 0.05$.

Dataset	Age				
	Female	Male	<i>t</i>	<i>df</i>	<i>p</i>
All datasets	66.312 \pm 0.463	66.708 \pm 0.424	-0.631	910.83	0.528
2_Landi_2008	67.714 \pm 1.960	67.290 \pm 1.090	0.189	21.397	0.852
6_Baty_2010	64.667 \pm 4.240	58.067 \pm 3.404	1.214	11.774	0.249
7_Takeuchi_2006	62.346 \pm 1.945	63.885 \pm 1.433	-0.637	51.986	0.527
8_Xi_2008	69.75 \pm 1.109	69.75 \pm 3.707	0	8.277	1
18_Lu_2010					
19_Dehan_2007					
36_Selamet_2012	68.560 \pm 1.486	73.222 \pm 1.501	-2.208	49.996	0.032 *
55_Beer_2002	64.340 \pm 1.704	63.560 \pm 1.766	0.305	59.357	0.761
56_Bhattacharjee_2001	62.773 \pm 1.584	66.094 \pm 1.680	-1.439	70.469	0.155
60_TCGA_LUAD_2016	66.424 \pm 0.760	65.696 \pm 0.911	6.613	242.16	0.541
61_TCGA_LUSC_2016	68.267 \pm 0.881	68.409 \pm 0.608	-0.133	148.25	0.894

Differences in transcriptional levels of LAC in patients from different areas

The PCA results also showed that there was no obvious sex difference in the DR profiles (Fig. 2a). K-Means cluster analysis results indicated that DR profiles were clustered based on dataset rather than sex or patient location (Figs. 2b and 2c), which was likely due to the differences in microarrays and subsequent analysis methods. Heatmap profiles showed that the main female DR profiles indicated a regional clustering trend (Fig. 2d), whereas the main male DR profiles did not show a similar trend (Fig. 2e).

Spearman correlation between main different genes in female and male LAC patients

The transcription levels of several genes, including collagen type X alpha 1 (COL10A1; Entrez ID: 1300), polo like kinase (PLK1;

Entrez ID: 5347), cytochrome P450 family 24 subfamily A member 1 (CYP24A1; Entrez ID: 1591), solute carrier family 7 member 5 (SLC7A5; Entrez ID: 8140), kinetochore associated 1 (KNTC1; Entrez ID: 9735), H2A.X variant histone (H2AX; Entrez ID: 3014), and insulin like growth factor 2 mRNA binding protein 3 (IGF2BP3; Entrez ID: 10643), were significantly upregulated in LAC tissue compared to controls. Conversely, the transcription levels of lipoprotein lipase (LPL; Entrez ID: 4023), SPARC like 1 (SPARCL1, Entrez ID: 8404), angiopoietin 1 (ANGPT1; Entrez ID: 284), AOC3, H2.0 like homeobox (HLX; Entrez ID: 3142), CD69 (Entrez ID: 969), and forkhead box F1 (FOXF1; Entrez ID: 2294) were significantly downregulated (Fig. 3).

Lung Adenocarcinoma Transcriptome

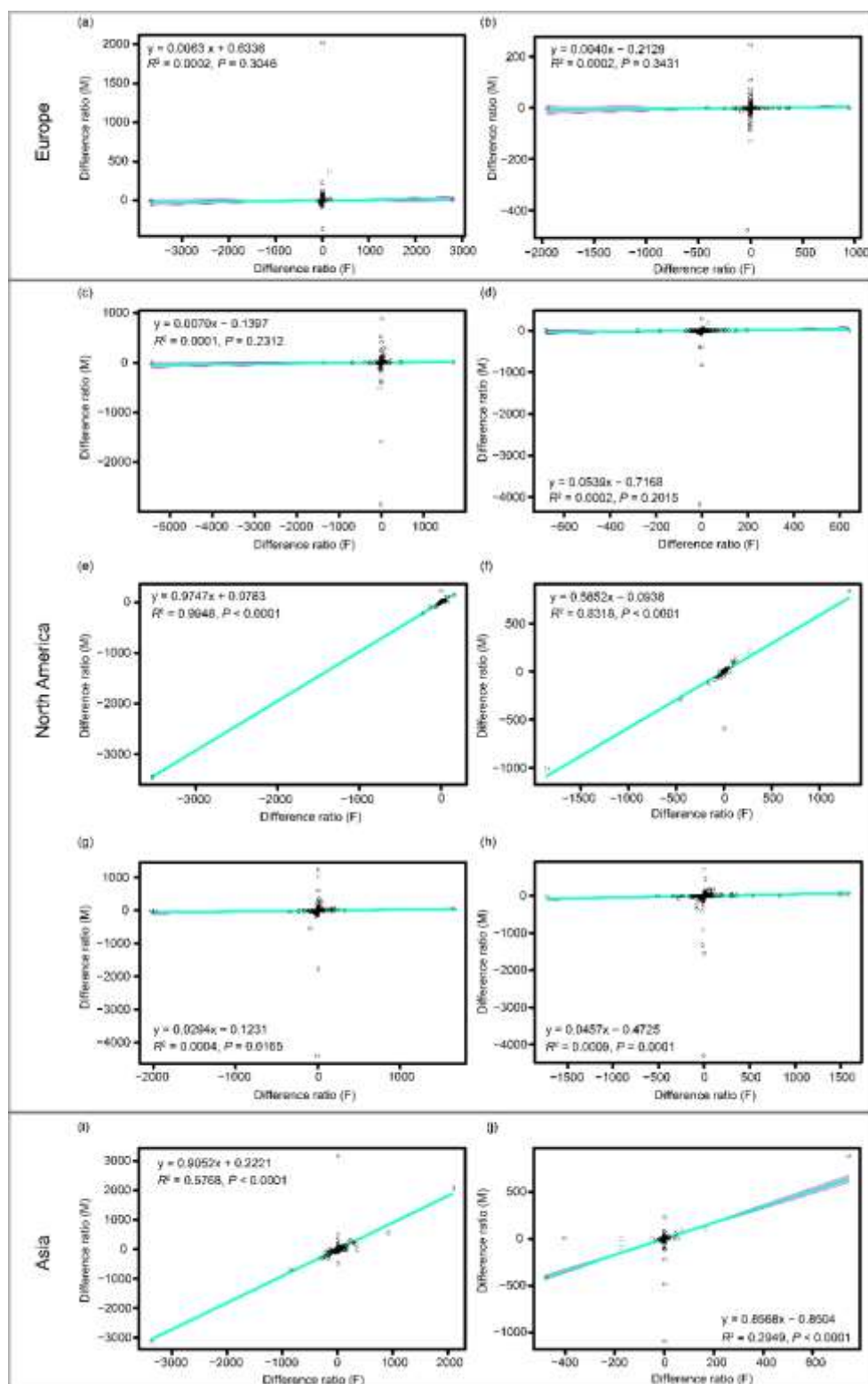


Fig. 1. Correlation between the different ratios of male and female adenocarcinoma samples collected from various regions. (a) 2_Landi_2008 dataset from Italy; (b) 6_Baty_2010 dataset from Switzerland; (c) 36_Selamat_2012 dataset from Canada; (d) 8_Xi_2008 dataset from the USA; (e) 55_Beer_2002 dataset from the USA; (f) 56_Bhattacharjee_2001 dataset from the USA; (g) 60_TCGA_LUAD_2016 dataset from the USA; (h) 61_TCGA_LUSC_2016 dataset from the USA; (i) 22_Fujiwara_2012a dataset from Japan; and (j) 19_Dehan_2007 dataset from Israel. These datasets were obtained from the Lung Cancer Explorer (<https://lce.biohpc.swmed.edu/lungcancer/>).

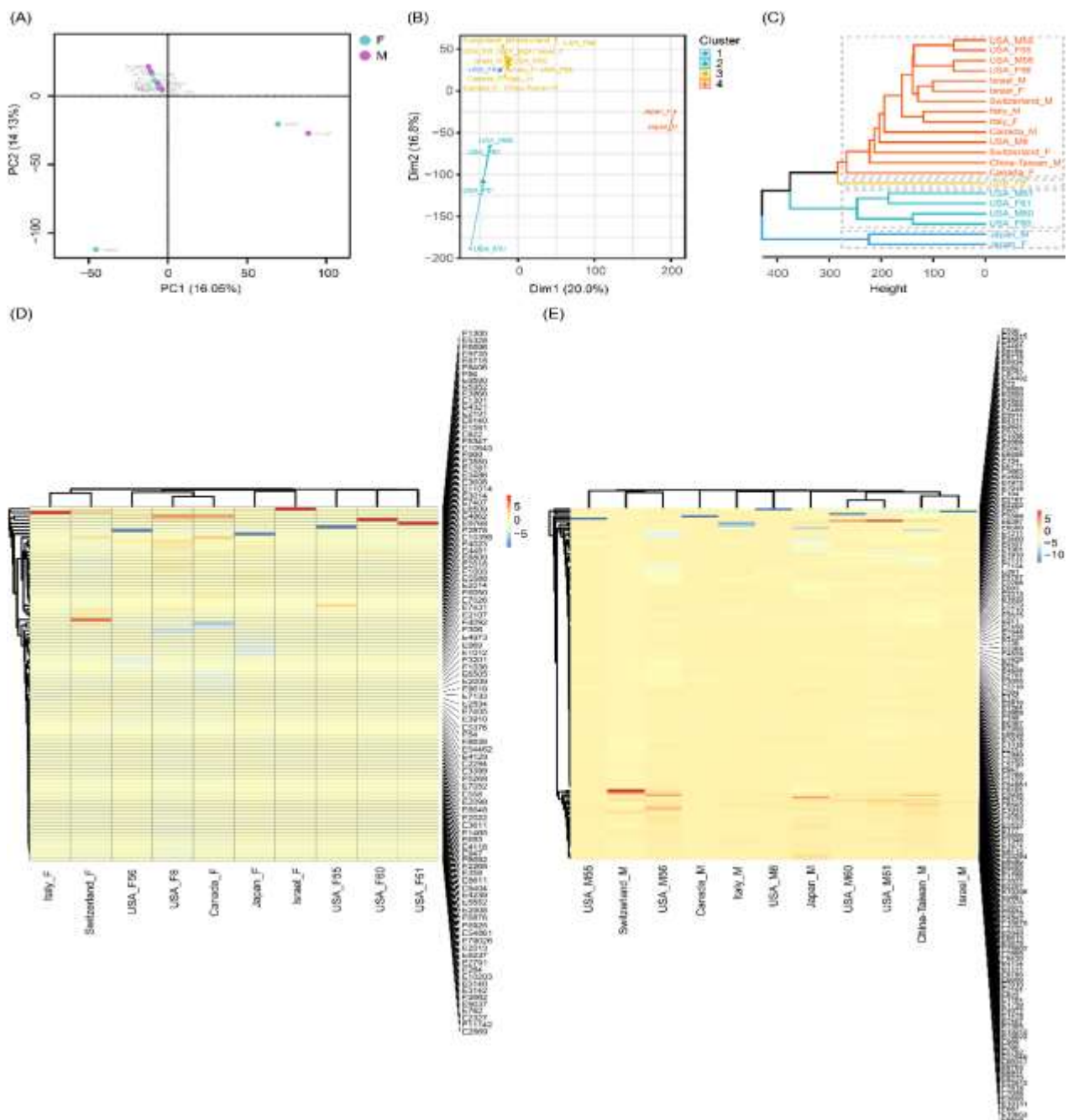


Fig. 2. Differences in the transcriptional levels of lung adenocarcinoma in patients from various regions. (a) Principal component analysis (PCA) profile; (b) K-means cluster plot; (c) K-means cluster dendrogram; (d) heatmap profile displaying the difference ratios of the main differential genes in females; (e) heatmap profile displaying the difference ratios of the main differential genes in males.

Spearman correlation analysis revealed significant positive correlations between *PLK1*, *CYP24A1*, *KNTC1*, and *IGF2BP3* in female LAC patients (Spearman correlation coefficient ≥ 0.6 and $P < 0.05$; Fig. 3a). Additionally, *COL10A1* was significantly positively correlated with *ANGPT1*

(Spearman correlation coefficient ≥ 0.6 and $P < 0.05$; Fig. 3a). On the other hand, *SLC7A5* was significantly negatively correlated with *SPARCL1* (Spearman correlation coefficient ≤ -0.6 , $P < 0.05$). *H2AX* was significantly negatively correlated with *HLX*, *CD69*, and *FOXF1*, while the latter three were

Lung Adenocarcinoma Transcriptome

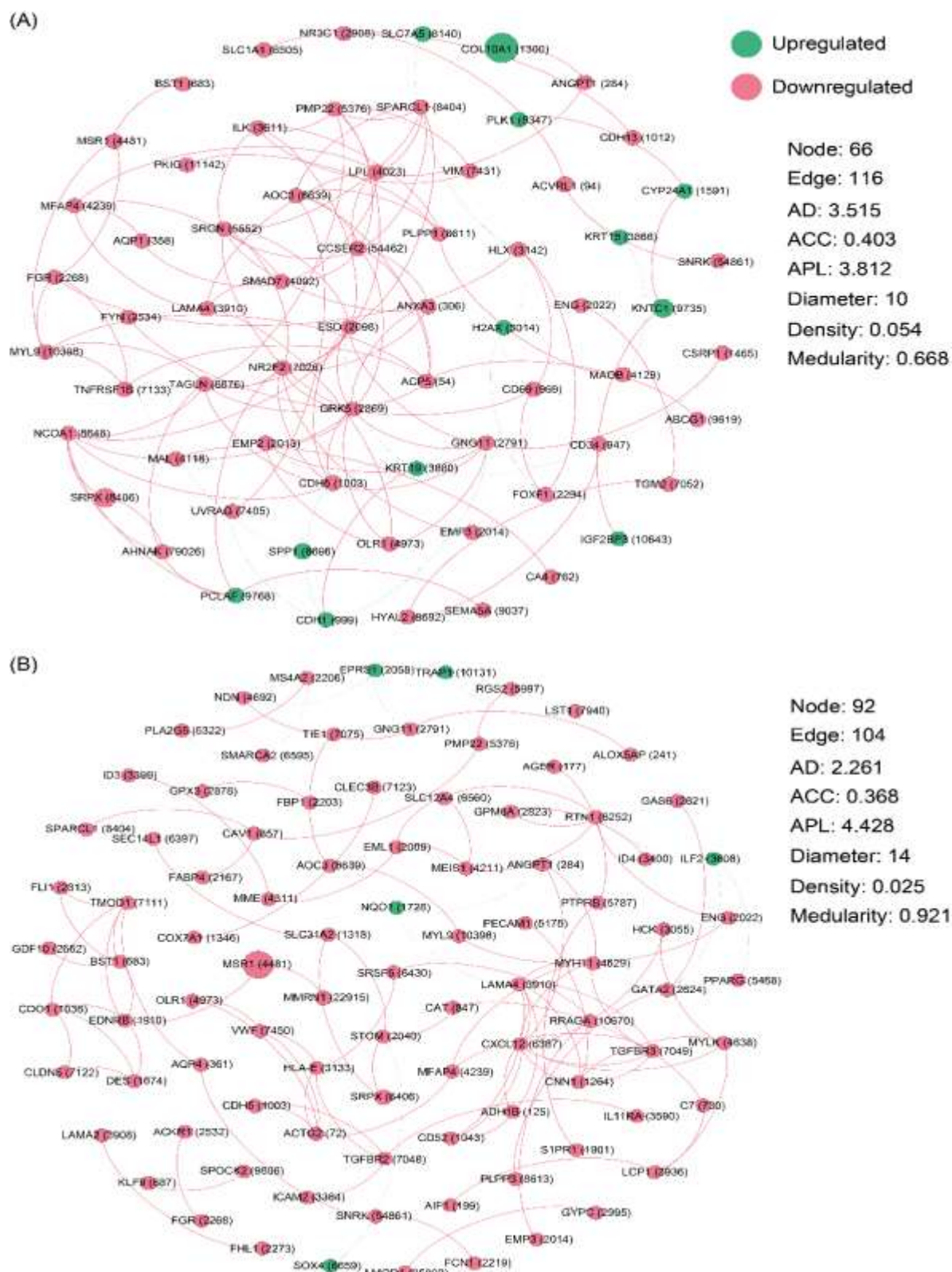


Fig. 3. Co-occurrence networks indicating the Spearman correlation between different genes in female (a) and male (b) lung adenocarcinoma patients. Correlations with $|\text{Spearman correlation coefficient}| \geq 0.6$ and $P < 0.05$ were considered significant. The node diameters represent the different ratios. The red and green edges represent positive and negative correlations, respectively. ACC, average clustering coefficient; AD, average degree; APL, average path length.

significantly positively correlated with each other (Fig. 3a). Furthermore, the transcription levels of SPARCL1, peripheral myelin protein 22 (PMP22; Entrez ID: 5376), integrin linked kinase (ILK; Entrez ID: 3611), LPL, AOC3, coiled-coil serine rich protein 2 (CCSER2; Entrez ID: 54462), serglycin (Entrez ID: 5552), SMAD family member 7 (SMAD7; Entrez ID: 4092), annexin A3 (ANXA3; Entrez ID: 306), esterase D (ESD; Entrez ID: 2098), nuclear receptor subfamily 2 group F member 2 (NR2F2; Entrez ID: 7026), G protein-coupled receptor kinase 5 (GRK5; Entrez ID: 2869), and transgelin (Entrez ID: 6876) were significantly downregulated in LAC tissue and were also significantly positively correlated with each other ($P < 0.05$; Fig. 3a). In contrast, the transcription level of keratin 19 (KRT19; Entrez ID: 3880) was significantly upregulated in the LAC tissue and was significantly negatively correlated with G protein subunit gamma 11 (GNG11; Entrez ID: 2791), ESD, epithelial membrane protein 2 (EMP2; Entrez ID: 2013), cadherin 5 (CDH5; Entrez ID: 1003), and cysteine and glycine rich protein 1 (CSR1; Entrez ID: 1465), which were all significantly downregulated in LAC tissue ($P < 0.05$; Fig. 3a). Additionally, the transcription level of cadherin 1 (CDH1; Entrez ID: 999) was significantly upregulated in LAC tissue and was significantly negatively correlated with CDH5, UV radiation resistance associated (UVRAG; Entrez ID: 7405), and GNG11.

In male LAC patients, the transcription level of glutamyl-prolyl-tRNA synthetase 1 (EPRS1; Entrez ID: 2058) was significantly upregulated compared to controls and was significantly negatively correlated with necdin (Entrez ID: 4692) and GNG11. The transcription level of TNF receptor associated protein 1 (TRAP1; Entrez ID: 10131) was significantly upregulated compared to controls and was significantly negatively correlated with leukocyte-specific transcript 1 (LST1; Entrez ID: 940). Similarly, the transcription level of interleukin enhancer binding factor 2 (ILF2; Entrez ID: 3608) was significantly upregulated compared to controls, and was significantly

negatively correlated with advanced glycosylation end-product specific receptor (AGER; Entrez ID: 177), peroxisome proliferator activated receptor gamma (PPARG; Entrez ID: 5468), and GATA binding protein 2 (GATA2; Entrez ID: 2624) ($P < 0.05$; Fig. 3b).

The transcription level of NAD(P)H quinone dehydrogenase 1 (NQO1; Entrez ID: 1728) was significantly upregulated and negatively correlated with ANGPT1, and laminin subunit alpha 4 (LAMA4; Entrez ID: 3910). The transcription level of SRY-box transcription factor 4 (SOX4; Entrez ID: 6659) was significantly upregulated and significantly negatively correlated with stomatin (Entrez ID: 2040) ($P < 0.05$; Fig. 3b). In addition, the transcription levels of C-X-C motif chemokine ligand 12 (CXCL12; Entrez ID: 6387), LAMA4, Ras related GTP binding A (RRAGA; Entrez ID: 10670), calponin 1 (CNN1; Entrez ID: 1264), and transforming growth factor beta receptor 3 (TGFBR3; Entrez ID: 7049) were significantly downregulated and formed a close significant positive correlation ($P < 0.05$; Fig. 3b). Similarly, the transcription levels of RTN1 (Entrez ID: 6252), myosin light chain kinase (MYLK; Entrez ID: 4638), transforming growth factor beta receptor 2 (TGFBR2; Entrez ID: 7048), actin gamma 2 (ACTG2; Entrez ID: 72), and von Willebrand factor (VWF; Entrez ID: 7450) were significantly downregulated and were also closely correlated with other downregulated genes (edge ≥ 5 , $P < 0.05$; Fig. 3b). These results suggest that there are sex differences in the changes in transcription levels of the main differentially expressed genes in LAC tissue.

Correlations between transcriptional difference ratio of LAC and ASIR, ASMR, and MIR

The results of the Spearman correlation analysis showed that the DRs of certain genes were significantly correlated with the ASIR, ASMR, and MIR of female and male LAC. Specifically, the DRs of SPARCL1, sushi repeat containing protein X-linked (SRPX, Entrez ID: 8406), PMP22, macrophage scavenger receptor 1

(MSR1; Entrez ID: 4481), bone marrow stromal cell antigen 1 (BST1; Entrez ID: 683), A-kinase anchoring protein 12 (AKAP12; Entrez ID: 9590), monoamine oxidase B (MAOB; Entrez ID: 4129), vimentin (Entrez ID: 7431), serglycin, ILK, ESD, transgelin, nuclear receptor coactivator 1 (NCOA1; Entrez ID: 8648), and phospholipid phosphatase 1 (PLPP1, Entrez ID: 8611) were negatively correlated with the ASIR of female LAC, while the DR of KRT19 was positively correlated. Additionally, the DRs of COL10A1, ANGPT1, SMAD7, collagen type III alpha 1 chain (COL3A1; Entrez ID: 1281), and aquaporin 1 (AQP1; Entrez ID: 358) were positively correlated with the ASMR of female LAC, while the DR of keratin 15 (KRT15; Entrez ID: 3866) was negatively correlated. Furthermore, the DRs of COL10A1, microfibril associated protein 4 (MFAP4, Entrez ID: 4239), arginase 2 (ARG2; Entrez ID: 384), LPL, SMAD7, COL3A1, MAOB, and AQP1 were positively correlated with the MIR of female LAC, while the DR of KRT15 was negatively correlated. These correlations were all statistically significant ($P < 0.05$; Fig. 4a). The DRs of COL10A1, MFAP4, and ANGPT1 were significantly correlated with the MIR of female LAC, while the DRs of SPARCL1, SRPX, PMP22, ANGPT1, serglycin, ILK, ESD, transgelin, and PLPP1 were significantly correlated with the ASIR of female LAC ($P < 0.05$; Fig. 4b). Interestingly, the DR of regulator of G protein signaling 2 (RGS2; Entrez ID: 5997) was significantly negatively correlated with the ASIR of male LAC, while the DRs of SPARCL1, allograft inflammatory factor 1 (AIF1; Entrez ID: 199), cytochrome c oxidase subunit 7A1 (COX7A1; Entrez ID: 1346), interleukin 3 receptor subunit alpha (IL3RA; Entrez ID: 3563), alcohol dehydrogenase 1B (ADH1B; Entrez ID: 125), TEK receptor tyrosine kinase (TEK, Entrez ID: 7010), CXCL12, Src like adaptor (SLA; Entrez ID: 6503), and CD52 (Entrez ID: 1043) were

significantly positively correlated with the ASMR of male LAC. Additionally, the DR of aminoacyl tRNA synthetase complex interacting multifunctional protein 2 (AIMP2; Entrez ID: 7965) was significantly negatively correlated with the ASMR of male LAC; while the DRs of SPARCL1, COX7A1, IL3RA, glypican 3 (GPC3; Entrez ID: 2719), and ADH1B were significantly positively correlated with the MIR of male LAC. Additionally, the DR of AIMP2 was significantly negatively correlated with the MIR of male LAC ($P < 0.05$; Fig. 4c). Furthermore, the DRs of SPARCL1 and AIMP2 were significantly correlated with the MIR of male LAC, while the DRs of TEK and CD52 were significantly correlated with the ASMR of male LAC ($P < 0.05$; Fig. 4d).

Discussion

Lipid metabolism reprogramming is a hallmark of cancer and plays an important role in shaping the tumor microenvironment and cancer cell phenotype, contributing to the occurrence and development of tumors (14). Therefore, reprogramming of lipid metabolism is an essential link in tumor metabolism (15). LPL catalyzes the hydrolysis of chylomicrons and very-low-density lipoprotein triglycerides and plays a key role in lipid metabolism and transport (16). Lu et al. (16) found that the LPL gene is expressed at a low level, with an average ratio 0.26 in LAC tissues compared to controls. They inferred that LPL may play an important role in LAC development. Our results also indicated that the transcription level of LPL in LAC tissues was significantly lower than that in the controls. Amine oxidases have been linked to leukocyte migration and tumorigenesis. In a study by Chang et al. (17), it was found that amine oxidase copper containing 3 (AOC3; Entrez ID: 8639) was significantly decreased in lung tumor tissue compared to normal tissue.

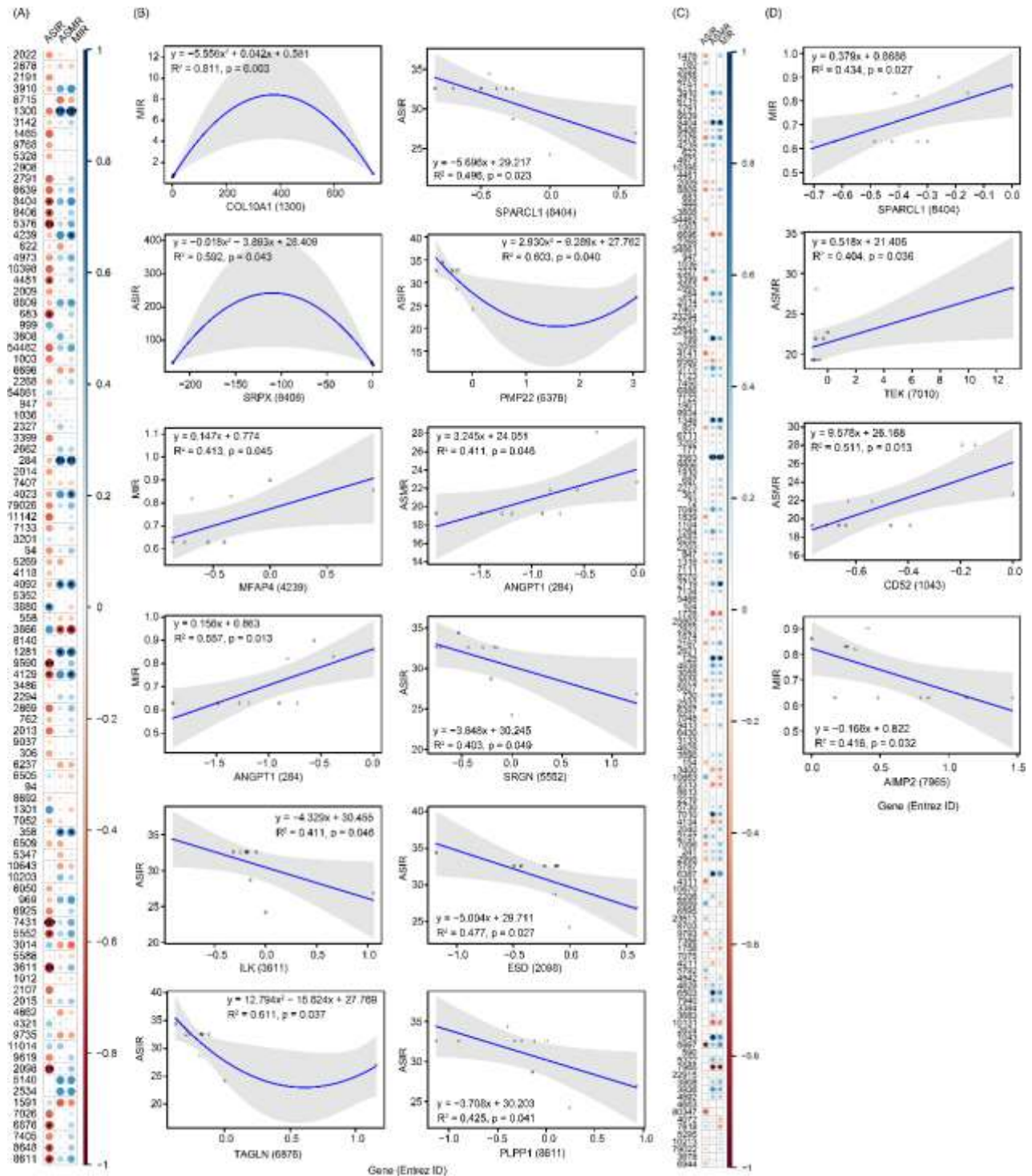


Fig. 4. Correlations between the transcriptional difference ratios of male and female lung adenocarcinoma (LAC) to healthy control tissues and age-standardized incidence rate (ASIR), age-standardized mortality rate (ASMR), and mortality-to-incidence ratio (MIR). (a) Bubble chart displaying Spearman correlations between the transcriptional difference ratio of female LAC to healthy control tissues and ASIR, ASMR, and MIR; (b) correlation curves/lines between the transcriptional difference ratio of female LAC to healthy control tissues and ASIR, ASMR, and MIR; (c) bubble chart displaying Spearman correlations between the transcriptional difference ratio of male LAC to healthy control tissues and ASIR, ASMR, and MIR; (d) correlation curves/lines between the transcriptional difference ratio of male LAC to healthy control tissues and ASIR, ASMR, and MIR. * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$.

The researchers also discovered that knocking down AOC3 resulted in reduced CD4⁺ T-cell attachment to lung cancer cells, decreased trans-endothelial migration *in vitro*, and reduced CD4⁺ T-cell trafficking to the lung *in vivo* (17). Our own results also showed a significant decrease in the transcription level of AOC3 in LAC tissues compared to controls (Fig. 3). Additionally, a previous study has reported an upregulation of COL10A1 transcription in gastric cancer (18). Similarly, our results showed an upregulation of COL10A1 transcription in LAC tissues compared to controls. SPARCL1 is a matricellular protein with anti-adhesive, anti-proliferative, and anti-tumorigenic functions, and is frequently downregulated in tumors such as colorectal carcinoma and NSCLC (19). Our results showed a significant downregulation of SPARCL1 transcription in LAC tissues compared to controls, and a significant negative correlation with the ASIR, suggesting that low SPARCL1 expression may indicate a higher risk of LAC.

The expressions of certain proteins and cytokines, including serglycin, ILK, and PMP22, has been found to be closely linked to the development of lung cancer (20-22). In fact, a study has shown that suppressing PMP22 expression can inhibit the proliferation, invasion, and apoptosis of lung cancer cells (20). Therefore, the downregulation of PMP22 in this study was probably a mechanism that hinders, rather than promotes, tumor growth. Additionally, Guo et al. (21) found that serglycin is often overexpressed in LAC and can promote malignant behavior or through a CD44-dependent pathway, and that increased serglycin expression is associated with a poor prognosis in primary LAC. Similarly, Nikou et al. (22) reported that ILK is overexpressed in human LAC and in KRAS-driven LAC in mice. However, our results showed that serglycin and ILK were actually downregulated in LAC compared to controls. Furthermore, ESD activity has been linked to tumor development, with lower ESD activity in lung adenocarcinoma patients indicating a

higher tumor grade (23). Our findings suggest that the downregulation of ESD may be a contributing factor in the development of LAC. In contrast, Sun et al. (24) found that transgelin promotes lung cancer progression by activating cancer-associated fibroblasts and increasing IL-6 release. Therefore, the downregulation of transgelin in LAC in our study may be a mechanism that inhibits the development of LAC. Additionally, our results showed a significant negative correlation between the DRs of serglycin, ILK, ESD, and PLPD1 and ASIR (Fig. 4), indicating that the expression levels of these genes are closely linked to the occurrence of LAC.

Kuo et al. (25) discovered that ANGPTL1 expression was inversely correlated with invasion, lymph node metastasis, and poor clinical outcomes in a study of 102 patients with lung cancer. They also found that ANGPTL1 suppressed the migratory, invasive, and metastatic abilities of lung and breast cancer cell lines *in vitro* and reduced metastasis in mice injected with cancer cell lines that overexpressed ANGPTL1. These findings with our results suggest that the downregulation of ANGPT1 may be a contributing factor in the development of LAC. However, our results also showed a significant positive correlation between ANGPT1 and ASMR and MIR, indicating that the overexpression of ANGPT1 may be linked to the deterioration of LAC. The mechanism behind this contradiction is currently unclear.

The pathogenesis of LAC in males and females in Asia and some parts of North America was likely to be consistent, but there may be significant differences in the pathogenesis of LAC between males and females in Europe. Our study found that the transcription levels of LPL and AOC3 were significantly lower in LAC tissues compared to controls, and downregulation of ESD may play a key role in the development of LAC. Additionally, our results showed that the expression levels of serglycin, ILK, ESD, and PLPD1 were closely linked to the occurrence of LAC. Furthermore, our findings suggest that there were sex differences in the changes of

transcription levels of the main differentially expressed genes in LAC tissues.

Ethical statement

The study was reviewed and approved by the Ethics Committee of Hebei Chest Hospital. The procedures followed were in accordance with the ethical standards of the responsible committee on human experimentation (institutional or regional) and with the Helsinki Declaration of 1975, as revised in 2000 (available at http://www.wma.net/e/policy/17-c_e.html).

We confirmed that all the data were anonymized and maintained with confidentiality. The requirement for informed consent has been waived by the Ethics Committee of Hebei Chest Hospital because of the retrospective nature of the current study.

References

- Sharma R. Mapping of global, regional and national incidence, mortality and mortality-to-incidence ratio of lung cancer in 2020 and 2050. *Int J Clin Oncol.* 2022;27:665-675.
- Esfandi F, Fallah H, Arsang-Jang S, Taheri M, Ghafouri-Fard S. The expression of *CCAT2*, *UCA1*, *PANDA* and *GHET1* long non-coding RNAs in lung cancer. *Rep Biochem Mol Biol.* 2019;8:36-41.
- Barta JA, Powell CA, Wisnivesky JP. Global epidemiology of lung cancer. *Ann Glob Health.* 2019;85:1-16.
- Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin.* 2021;71:209-249.
- Cridelli C, Rossi A, Carbone DP, Guarize J, Karachaliou N, et al. Non-small-cell lung cancer. *Nat Rev Dis Primers.* 2015;1:15009.
- Bakhtiyari N, Sharifi A, Aftabi Y, Gilani N, Zafari V, Ansarin A, Seyedrezazadeh E. Association between *NOX4* and *Nrf2* genes in non-small-cell lung carcinoma: A case-control study. *Rep Biochem Mol Biol.* 2021;10:327-333.

Funding

Self-funded.

Conflict of interest

There are no conflicts of interest.

Acknowledgement

We would like to thank Guangdong Meilikang Bio-Science Ltd., China for help with data analysis.

Authors' Contribution

Conceptualization, J.N., and S.W.; Investigation, L.Z., C.F., H.S., X.W., X.Z., Q.Q., and J.N.; Methodology, L.Z., X.W., J.N., and S.W.; Formal analysis, L.Z., C.F., and J.N.; Project administration, S.W.; Resources, X.Z., J.N., and S.W.; Funding acquisition, J.N., and S.W.; Visualization, L.Z., H.S., Q.Q., and J.N.; Writing - original draft, L.Z., and C.F.; Writing - review and editing, J.N. and S.W.

- The Cancer Genome Atlas Research Network. Comprehensive molecular profiling of lung adenocarcinoma. *Nature.* 2014;511:543-550.
- Cai L, Lin S, Girard L, Zhou Y, Yang L, et al. LCE: an open web portal to explore gene expression and clinical associations in lung cancer. *Oncogene.* 2019;38:2551-2564.
- Zhu X, Luo H, Xu Y. Transcriptome analysis reveals an important candidate gene involved in both nodal metastasis and prognosis in lung adenocarcinoma. *Cell Biosci.* 2019;9:92.
- Wang Z, Li Z, Zhou K, Wang C, Jiang L, et al. Deciphering cell lineage specification of human lung adenocarcinoma with single-cell RNA sequencing. *Nat Comm.* 2021;12:6500.
- Parks DH, Tyson GW, Hugenholtz P, Beiko RG. STAMP: statistical analysis of taxonomic and functional profiles. *Bioinformatics.* 2014;30:3123-3124.
- Dembélé D, Kastner P. Fold change rank ordering statistics: a new method for detecting differentially expressed genes. *BMC Bioinformatics.* 2014;15:14.
- Adler M, Alon U. Fold-change detection in biological systems. *Curr Opin Syst Biol.* 2018; 8:81-89.

14. Hao Y, Li D, Xu Y, Ouyang J, Wang Y, et al. Investigation of lipid metabolism dysregulation and the effects on immune microenvironments in pan-cancer using multiple omics data. *BMC Bioinformatics*. 2019;20:195.
15. Wang W, Bai L, Li W, Cui J. The lipid metabolic landscape of cancers and new therapeutic perspectives. *Front Oncol*. 2020;10:605154.
16. Lu J, Li J, Ji C, Yu W, Xu Z, et al. Expression of lipoprotein lipase associated with lung adenocarcinoma tissues. *Mol Biol Rep*. 2008;35:59-63.
17. Chang C-Y, Wu K-L, Chang Y-Y, Tsai P-H, Huang J-Y, et al. Amine oxidase, copper containing 3 exerts anti-mesenchymal transformation and enhances CD4⁺ T-cell recruitment to prolong survival in lung cancer. *Oncol Rep*. 2021;46:203.
18. Chen S, Wei Y, Liu H, Gong Y, Zhou Y, et al. Analysis of collagen type X alpha 1 (COL10A1) expression and prognostic significance in gastric cancer based on bioinformatics. *Bioengineered*. 2021;12:127-137.
19. Klingler A, Regensburger D, Tenkerian C, Britzen-Laurent N, Hartmann A, et al. Species-, organ- and cell-type-dependent expression of SPARCL1 in human and mouse tissues. *PLoS ONE*. 2020;15:e0233422.
20. Qu H, Zhu M, Tao Y, Zhao Y. Suppression of peripheral myelin protein 22 (PMP22) expression by miR29 inhibits the progression of lung cancer. *Neoplasma*. 2015;62:881-886.
21. Guo JY, Hsu HS, Tyan SW, Li FY, Shew JY, et al. Serglycin in tumor microenvironment promotes non-small cell lung cancer aggressiveness in a CD44-dependent manner. *Oncogene*. 2017;36:2457-2471.
22. Nikou S, Arbi M, Dimitrakopoulos FID, Sirinian C, Chadla P, et al. Integrin-linked kinase (ILK) regulates KRAS, IPP complex and Ras suppressor-1 (RSU1) promoting lung adenocarcinoma progression and poor survival. *J Mol Histol*. 2020;51:385-400.
23. Wiedl T, Arni S, Roschitzki B, Grossmann J, Collaud S, et al. Activity-based proteomics: Identification of ABHD11 and ESD activities as potential biomarkers for human lung adenocarcinoma. *J Proteom*. 2011;74:1884-1894.
24. Sun C, Zhang K, Ni C, Wan J, Duan X, et al. Transgelin promotes lung cancer progression via activation of cancer-associated fibroblasts with enhanced IL-6 release. *Oncogenesis*. 2023;12:18.
25. Kuo T-C, Tan C-T, Chang Y-W, Hong C-C, Lee W-J, et al. Angiotensin-like protein 1 suppresses SLUG to inhibit cancer cell motility. *J Clin Inv*. 2013;123:1082-1095.