The PDZ-binding motif of SARS-CoV envelope protein induces cancerization and poor prognosis of lung adenocarcinoma

Long Chen[1] and Li Zhong[1]*

1. Bioengineering Institute of Chongqing University, 174 Shazheng Street, Chongqing, China

* Correspondence: Julia Li Zhong, Bioengineering Institute of Chongqing University, 174 Shazheng Street, Chongqing 400000, China. Tel:+86 15802371097. E-mail: jlzhong@cqu.edu.cn (JLZ)

Abstract

Both in lung adenocarcinoma (LUAD) and severe acute respiratory syndromes (SARS) uncontrolled inflammation could be detected in lung tissue. Whether the similarity mechanism exists is still unknown. PDZ-binding motif (PBM) in SARS-CoV E protein has been demonstrated as virulence factor induce inflammation storm. Study function of PBM in LUAD is significant for mechanism exploration of carcinogenesis mediated by SARS-CoV. To identify gene expression fluctuation induced by PBM, a microarray sequencing data of lung tissue infected by wild type (SARS-CoV-E-wt) and recombinant virus (SARS-CoV-E-mutPBM) was analysis followed by functional enrichment analysis. To understand the role of screened specific genes in LUAD, overall survival and immune correlation were calculated. A total of 12 genes (MAPK1, PRKCA, FGFR4, KDR, PTPRD, BCL2L15, UBD, MAMDC2, LTBP4, PTPRB, LGI3 and ITGA8) might participate in initial and development stage of LUAD through expression variation and mutation. Meanwhile, a total of 12 genes (CARHSP1, EIF4E2, HMGA1, IL1R2, MAGOHB, PVR, ADCY9, ELF5, ESYT3, SCML4, SECL14L4 and THRA) could lead to poorer prognosis via dysregulation. In addition, MAMDC2 and ITGA8 down-regulated by PBM could also alter prognosis. Though the conservative PBM (-D-L-L-V-) could be found at the end of carboxyl terminal in multi E proteins of coronaviruses, the specific function of each one depend on the whole amino acid sequence simultaneously. In conclusion, PBM of SARS-CoV E protein could promote carcinogenesis of LUAD by dysregulating important gene expression profiles followed

by influence immune response and overall prognosis. The results in present study also provided reference for the therapy of SARS-CoV-2 in LUAD patients.

Keywords: SARS-CoV; PBM; envelope protein; LUAD; bioinformatics

Introduction

Coronaviruses (CoVs) are pathogens responsible for a wide range of deadly epidemics in mammals [1]. In 2002, a novel coronavirus causing severe acute respiratory syndrome (SARS-CoV) was identified. Although SARS-CoV has not reappeared, some similar CoVs are still widely disseminated all over the world. In December 2019, there was an outbreak of coronavirus pneumonia (COVID19) caused by SARS-CoV-2 in Wuhan, Hubei province in China. As SARS-CoV, SARS-CoV-2 also encodes spike protein (S), nucleocapsid protein (N), envelope (E) and matrix protein (M) and uses the same cell entry receptor and manner result in infection. Thus, the study of SARS-CoV is helpful to better understand the pathogenesis of SARS-CoV-2 and screen biomarkers and therapeutic targets.

SARS-CoV, an enveloped virus enveloped a positive sense RNA genome (29.7 kb), belongs to the Coronavirinae subfamily, genus β [1]. E protein of SARS-CoV is a small integral membrane protein with 76 amino acids that contains a short hydrophilic N-terminus and a C-terminus, and a hydrophobic transmembrane region in-between position [2]. The transmembrane region contains one amphipathic a-helix at least and oligomerizes to form an ion channel (IC) in the membrane [3-5]. Normally, E protein is abundant in the infected cells [6], and mainly localized in endoplasmic reticulum Golgi intermediate compartment (ERGIC), where it could participate in virus budding, morphogenesis and trafficking actively [7-9]. Previously study has confirmed that SARS-CoV lacking the E protein (a virulence factor) was attenuated [10-13]. E protein could also increase the apoptosis and reduce the stress response induced by SARS-CoV infection [14]. Moreover, E protein plays an important role in promoting ion imbalances within cells and disrupting cellular pathways through the PDZ-binding motif (PBM) mediated protein-protein interactions [15-18].

PDZ domains are protein recognition sequences, consisting of ~85 amino acids which could bind with a specific peptide sequence (PBM), usually located at the end of the C-terminus [19-21]. Proteins with PDZ domains, typically found in the cell cytoplasm and plasma membrane, play key roles in various cellular response processes of viruses, such as cellular polarity, signal transduction pathways and cell-cell junctions influencing viral replication, dissemination in the host or pathogenesis [22]. In the previous study, it has been identified the PBM of SARS-CoV E protein was a virulence determinant in host. The recombinant viruses with mutated PBM were attenuated in mice, leading to a decreased expression of inflammatory cytokines and a substantially increased survival. Furthermore, the E protein PBM could interact with the syntenin protein, affecting p38-mitogen activated protein kinase (MAPK) activation, a protein participated in inflammatory cytokines expression, responsible for the SARS-CoV pathogenicity [23].

In addition to viral infections, lung cancer has also been interrelated to pulmonary inflammation closely. Lung adenocarcinoma (LUAD) is the most common primary lung cancer which falls under the umbrella of non-small cell lung cancer (NSCLC). LUAD usually occurs in the lung periphery, and in many cases, may be found in scars or areas of chronic inflammation [19, 20]. Thus, we assumed that pneumonia caused by PBM of coronavirus might lead to increased LUAD susceptibility and poorer prognosis. However, the mechanism of PBM in LUAD has not been elucidated. In our study, via bioinformatics analysis, the function and pathology of PBM in SARS-CoV E protein were studied comprehensively. In addition, through calculating overall survival and tumor immune score, demonstrating the potential mechanism of PBM in LUAD.

Methods

Ethics statement

Animal experimental protocols were approved by the Ethical Committee of The Center for Animal Health Research (CISAINIA) (permit numbers: 2011-009 and 2011-09) in strict accordance with Spanish National Royal Decree (RD 1201/2005) and international EU guidelines 2010/63/UE about protection of animals used for

experimentation and other scientific purposes and Spanish National law 32/2007 about animal welfare in their exploitation, transport and sacrifice and also in accordance with the Royal Decree (RD 1201/2005). Infected mice were housed in a ventilated rack (Allentown, NJ).

Sequencing Datasets

Dataset GSE52920 provided by Luis Enjuanes was acquired from GEO Database and all experimental methods refer to the original study [23]. 8 week-old specific pathogen free BALB/c female mice were purchased from Harlan Lab. BALB/c mice were infected at 16 weeks-old with 100,000 plaque forming units (pfu). All protocols were approved by the Ethical Review Committee at The Center for Animal Health Research (CISA-INIA). Both infected and mock-infected mice were housed in a ventilated rack (Allentown, NJ).

At 2 days after infection, lung tissues from infected mice were collected and homogenized by gentleMACS Dissociator (Miltenyibiotec). Then, according to the manufacturer's instructions, total RNA was extracted using the RNeasy purification kit (Qiagen).Each transcriptomic comparison contained three biological replicates. About 200 ng total RNA was purified with RNeasy Mini Kit (Qiagen) and amplified by One Color Low Input Quick Amp Labeling Kit (Agilent Technologies). As described in One-Color Microarray Based Gene Expression Analysis Manual Ver. 6.5, Agilent Technologies, the probes and hybridization were prepared. In brief, each 600 ng hybridization of Cy3 probes were mixed with 1 ul of 25x Fragmentation Buffer, 5 ul of 10x Blocking Agent and Nuclease free water in a 25 ml reaction system, incubated at 60°C (30 mins)to fragment RNA and stopped with 25 ul of 2x Hybridization Buffer. The samples were placed on ice and quickly loaded onto arrays, hybridized at 65°C for 17 hours in a Hybridization oven rotator and then washed using GE wash buffer 1 at 25°C and GE Wash Buffer 2 at 37°C for 1 minute, respectively. Samples were dried by centrifugation at 2000 rpm for 2 minutes. Slides were Sure Print G3 Agilent 8660K Mouse (G4852A-028005) Images were captured by an Agilent Microarray Scanner and spots quantified using Feature Extraction Software (Agilent Technologies). Background correction by normexp method (with an offset of 50) and normalization of

expression data by adjustment of quantiles were performed using LIMMA package [24]. In addition, sequencing data of 586 LUAD patients were obtained from the cancer genome atlas (TCGA) database.

Microarray data and enrichment analysis

Differentially expressed genes were determined using linear model methods. Each probe was tested for alterations of expression over replicates via the empirical Bayes moderated t-statistic. To control the false discovery rate (FDR), defined as the expected proportion of false positives among the significant tests, p-values were corrected by Benjamini and Hochberg method [25]. The expected FDR was controlled to be less than 5%. Genes were considered differentially expressed when the FDR was ≤ 0.01. In addition, only genes with the $|\log \text{fold change}| \geq 1$ were considered for further analysis. Enrichment analysis of differently expressed genes (DEGs) was performed with DAVID version 6.7 (p-value ≤ 0.05).

Immune correlation of DEGs

Relations between the expression level of DEGs and abundance of tumor-infiltrating lymphocytes (TILs), three kinds of immunomodulators (24 immunoinhibitor, 45 immunostimulator, 21 MHC molecule) and 41 chemokines (or 18 receptors) were calculated from TISIDB, a web portal for tumor and immune system interaction, which integrates multiple heterogeneous data types. A total of 28 TIL types from Charoentong's study [26] were inferred by using gene set variation analysis (GSVA).

Overall surveil of specific genes

The overall surveil of the DEGs which also affect prognosis were calculated by using gene expression profiling interactive analysis (GEPIA) based on gene expression. The median group cutoff (cutoff-high=50%, cutoff-low=50%) was applied. The hazard ratio based on Cox PH model was calculated and the 95% confidence interval was also added. The clinical data were also obtained from TCGA database. All databases used are listed in Table 1.

Table1. List of Database

Database ID	URL
GEO Dataset	https://www.ncbi.nlm.nih.gov/gds/?term=
TCGA	https://www.cancer.gov/
cBioportal of cancer genomics	https://www.cbioportal.org/
The Human Protein Atlas	https://www.proteinatlas.org/
Linked Omics	http://www.linkedomics.org/
DAVID	https://david.ncifcrf.gov/
GEPIA	http://gepia.cancer-pku.cn/index.html
TISIDB	http://cis.hku.hk/TISIDB/index.php

Statistical analyses

All results were presented as mean values ± standard error of the mean (SEM). Unless mentioned otherwise, the statistical comparison between groups was performed via using t-test, a maximum of three comparisons were performed per panel, and robustness of statistical significance was verified after correction for multiple testing. Probability was considered to be significant at p-value ≤ 0.05.

Result

Gene dysregulation induced by PBM

As described in Luis Enjuanes' study [23], to evaluate the role of SARS-CoV E protein PBM in virus pathogenesis, a recombinant SARS-CoVs with mutated E protein PBM (SARS-CoV-E-mutPBM) were generated by mutating the last 4 amino acids (-D-L-L-V-) to 4 glycines (-G-G-G-G-) in carboxy-terminal, maintaining the full-length E protein of SARS-CoV **(Figure 1A)**. Infection of BALB/c mice with both SARS-CoV-E-wt and SARS-CoV-E-mutPBM, a total of 225 down-regulated and 547 up-regulated genes were detected in mutPBM group ($|\log FC| \geq 1$, adj. p-value ≤ 0.05) **(Figure 1B)**. Followed by enrichment analysis in DAVID database respectively, 37 pathways, 87 biological processes (BPs), 6 cellular components (CCs) and 16 molecular functions (MFs) were enriched by down-regulated gene set while 47 pathways, 136 BPs, 63 CCs and 55 MFs were enriched by up-regulated gene set. The top 10 items of each kind were

shown (**Figure 1C and D**). Interestingly, the down-regulated genes were enriched in the Pertussis pathway, an acute respiratory infectious disease. Also, the BP items named inflammatory response, immune response and defense response to virus were enriched. The results may suggest that PBM in SARS-CoV E protein participates in respiratory disease via inflammatory response to virus. Meanwhile, pathways associated with cancerization (PI3K-Akt signaling pathway and pathway in cancer) were enriched by up-regulated genes. It was indicated that PBM participates in cancer development via regulated gene expression.

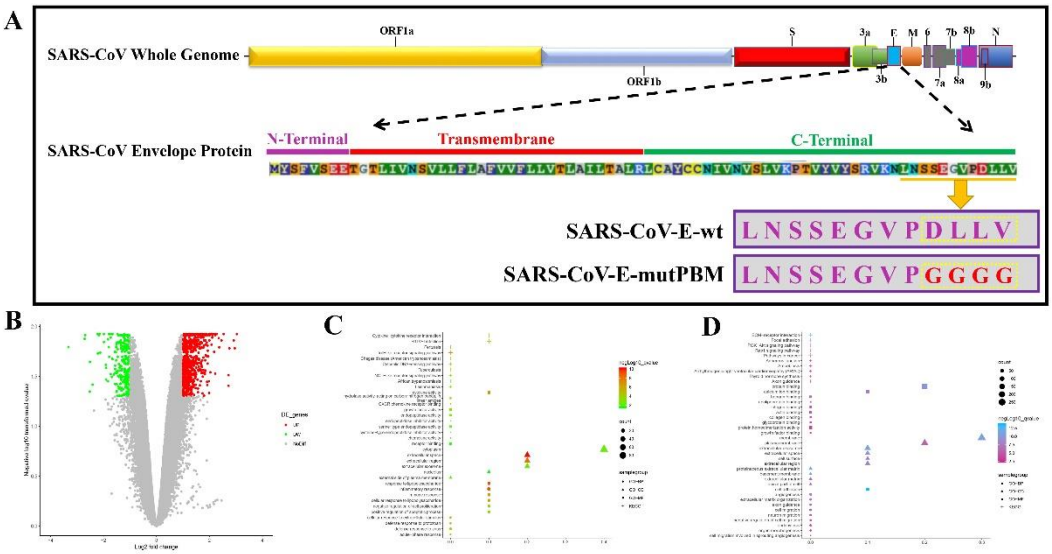

Figure 1. Enrichment analysis of DEGs induced by PBM. A, The SARS-CoV genome is shown at the top, and the expanded region shows the E protein sequence with different regions. Below, sequences corresponding to the end of E protein are shown in boxes for the different viruses. SARS-CoV-E-wt represents the wild type sequence. In SARS-CoV-E-mutPBM, E protein PBM was eliminated by point mutations, maintaining the full-length protein. Mutations are shown in red. **B,** The DEGs screened from SARS-CoV-E-mutPBM infected mouse lung tissue were shown in a volcano plot. The red point represents up-regulated genes while the green one represents down-regulated genes. The gray one represents a gene with non-significant difference. **C and D,** The top 10 pathways, BPs, CCs and MFs of down- and up-regulated genes in scatter plot. The legends localized on the right.

PBM function in the budding stage of LUAD

As described above, the DEGs regulated by PBM of SARS-CoV E protein might play important roles in cancerization. However, more than 700 DEGs were screened and not all DEGs were involved in cancer. To identify the DEGs actually participated in the carcinogenesis of LUAD, the functions of specific DEGs were confirmed by comparison with MSigDB gene sets.

LUAD is the common specie of NSCLC. Molecular mechanisms altered in NSCLC include activation of oncogenes, such as K-RAS, EGFR and EML4-ALK, and inactivation of tumor suppressor genes, such as p53, p16INK4a, RAR-beta, and RASSF1. A total of 54 genes were demonstrated to influence NSCLC development. Among them, the expression level of MAPK1 and PRKCA were also mediated by PBM. Strangely, both MAPK1 and PRKCA have no significant difference between normal and tumor tissues **(Figure 2A and B)**, which suggested they just play a key role in budding stage of LUAD. To confirm the function mode of MAPK1 and PRKCA, both positively and negatively correlated significant genes were screened **(Figure 2C and D)**. Furthermore, the gene set enrichment analysis was performed with redundancy reduction (weighted set cover). Notably, the enrichment analysis reminded both 2 genes involved in cell adhesion and junction **(Figure S1)**.

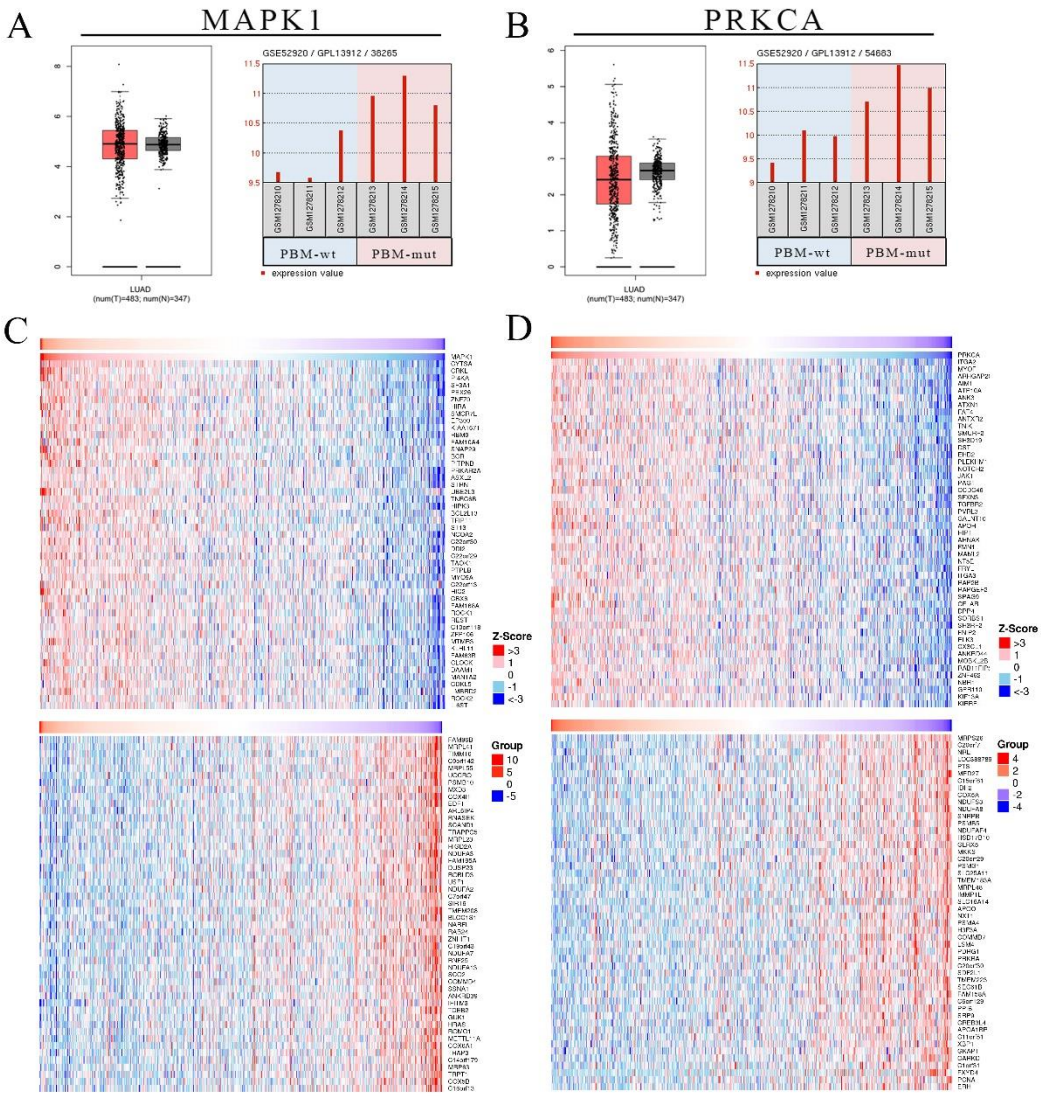

Figure 2. MAPK1 and PRKCA in LUAD. A and B, The expression profiles of MAPK1 and PRKCA in comparison of normal tissue vs. LUAD tissue and PBM-wt vs. PBM-mut. **C and D**, The heatmap of significant correlated genes of MAPK1 and PRKCA in LUAD. Top, the positively correlated genes while below, the negatively correlated genes. The legends localized on the right.

In addition to the NSCLC pathway gene set, there are 26 most significant mutated genes were identified in LUAD patients **(Figure S2)**. Among them, FGFR4, KDR and PTPRD were mutated in 9%, 15% and 22% patients respectively which were also regulated by PBM **(Figure 3 and Table S1)**. Moreover, it was identified that the expression of FGFR4, KDR and PTPRD was positively correlated with mutation **(Figure 3B)** and significantly decreased in tumor tissue of LUAD **(Figure 3D)**. Further study indicated that only FGFR4 was associated with TCGA pathway named RTK-RAS.

Furthermore, FGFR4 was the upstream regulator of MAPK1 **(Figure S3)**.

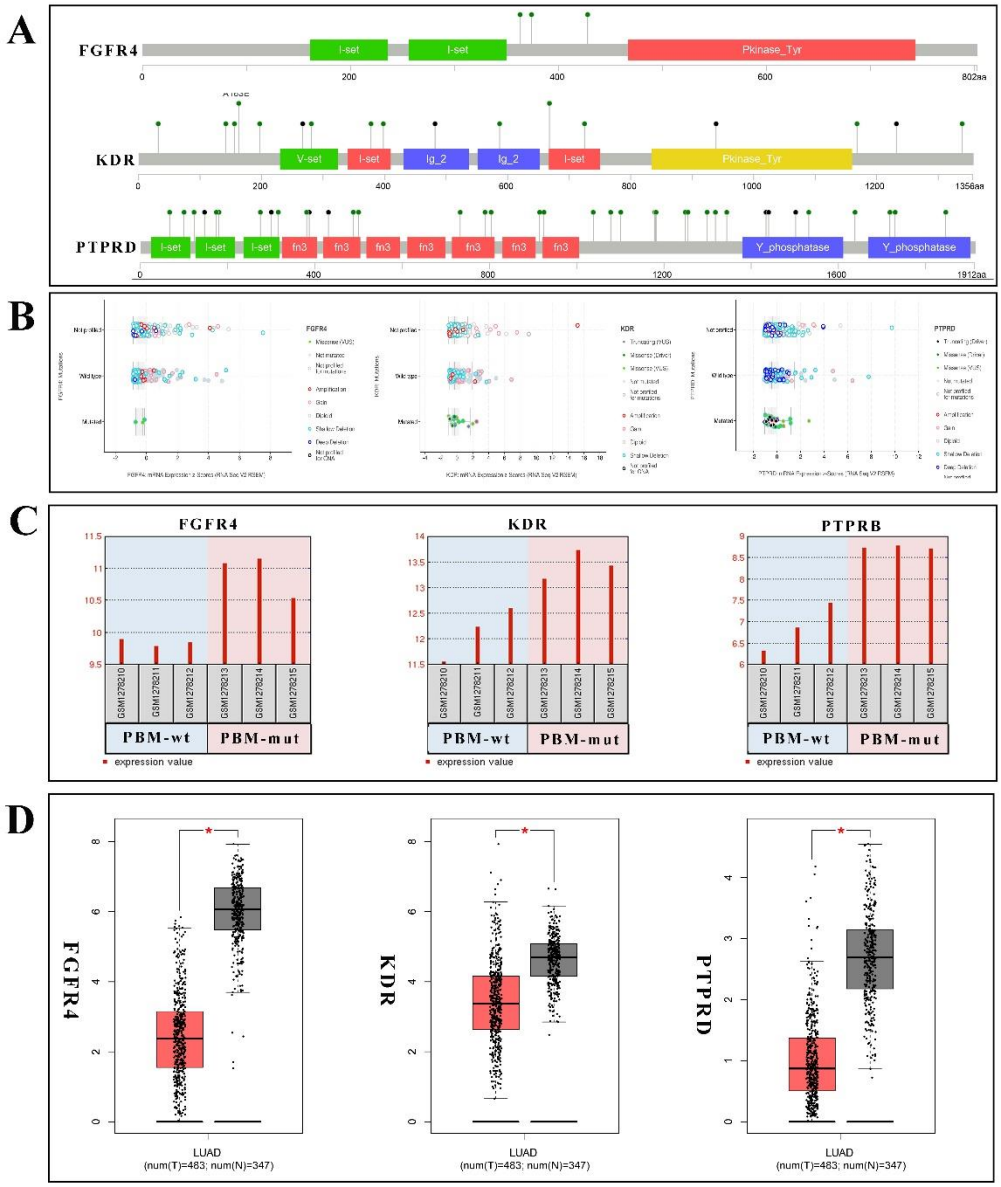

Figure 3. Mutation and expression profile of FGFR4, KDR and PTPRD. A, The sketch map of the mutation site. The height of each dot represents patient's number of a specific mutation. **B,** The correlation between mRNA expression and mutation of FGFR4, KDR and PTPRD were shown in scatter plots. The legends localized on the right. **C,** The expression level of FGFR4, KDR and PTPRD in PBM-wt and PBM-mut infected group. **D,** The expression level of FGFR4, KDR and PTPRD in normal and LUAD tissues.

PBM function in the development stage of LUAD

In LUAD, except for FGFR4, KDR and PTPRD, more than 900 DEGs (703 down-regulated genes and 246 up-regulated genes, |log FC| ≥ 2, adj. p-value ≤ 0.05) were detected in LUAD. After comparison among all DEG sets, a total of 7 genes (BCL2L15, UBD, MAMDC2, LTBP4, PTPRB, LGI3 and ITGA8) own both similar mRNA and protein expression profiles in both SARS-CoV-E-wt infected mouse lung tissues and LUAD tissues **(Figure 4A, B and D)**. Among them, only MAMDC2 and ITGA8 could lead to poorer prognosis with a lower expression level in LUAD **(Figure 4C)**. Thus, it was suggested that once LUAD patients infected by SARS-CoV-E-wt, MAMDC2 and ITGA8 would be decreased by PBM and increase mortality. Though the role of MAMDC2 in LUAD development is still unclear, the function of ITGA8 in LUAD and lung fibrosis has been proved [27, 28]. Due to both LUAD and SARS-CoV infection could result in berserk lung inflammation, the distribution of expression level across immune subtypes and immune correlation (lymphocyte, immune-modulators and chemokines) of these 7 genes were calculated **(Figure 5A and B)**. The result might confirm that all 7 DEGs affect LUAD development by immune response.

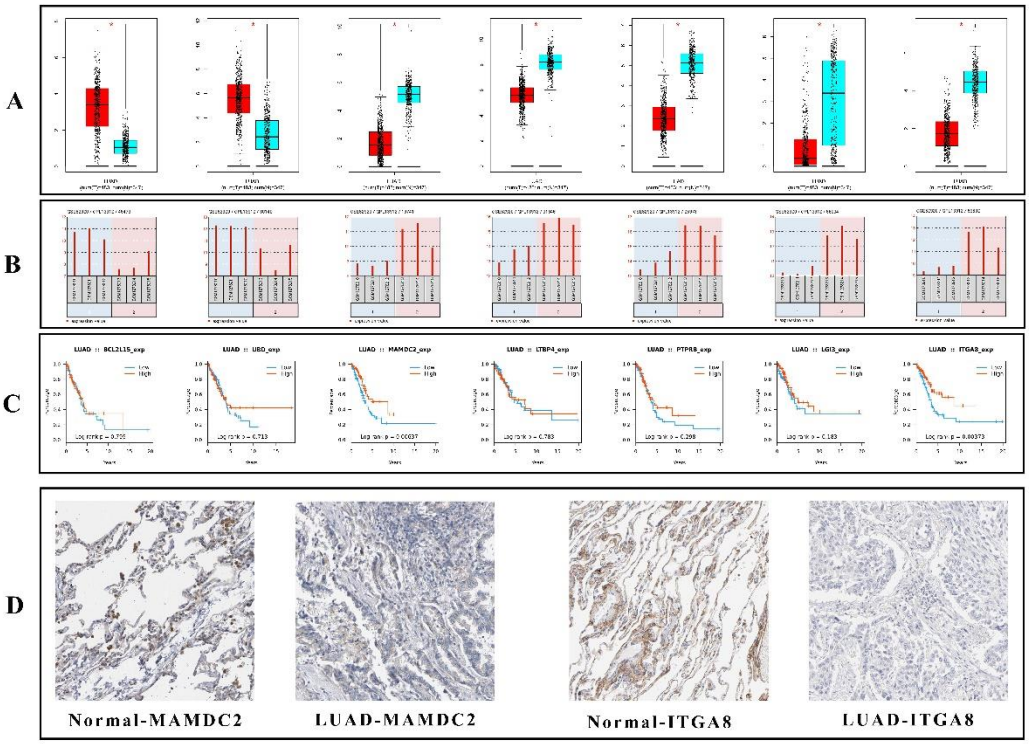

Figure 4. The expression profile and prognosis of BCL2L15, UBD, MAMDC2, LTBP4, PTPRB,

LGI3 and ITGA8 in LUAD. A and B, The expression level of BCL2L15, UBD, MAMDC2, LTBP4, PTPRB, LGI3 and ITGA8 in normal tissue vs. LUAD tissue and PBM-wt vs. PBM-mut groups. **C,** The overall survival of BCL2L15, UBD, MAMDC2, LTBP4, PTPRB, LGI3 and ITGA8 in LUAD patients. The red line represents a high expression level while the blue line represents a low expression level. Differences in prognosis of MAMDC2 and ITGA8 were statistically significant (p-value＜0.05). **D,** The pathological section of MAMDC2 and ITGA8 in normal and LUAD tissue. The brown represents gene expression level.

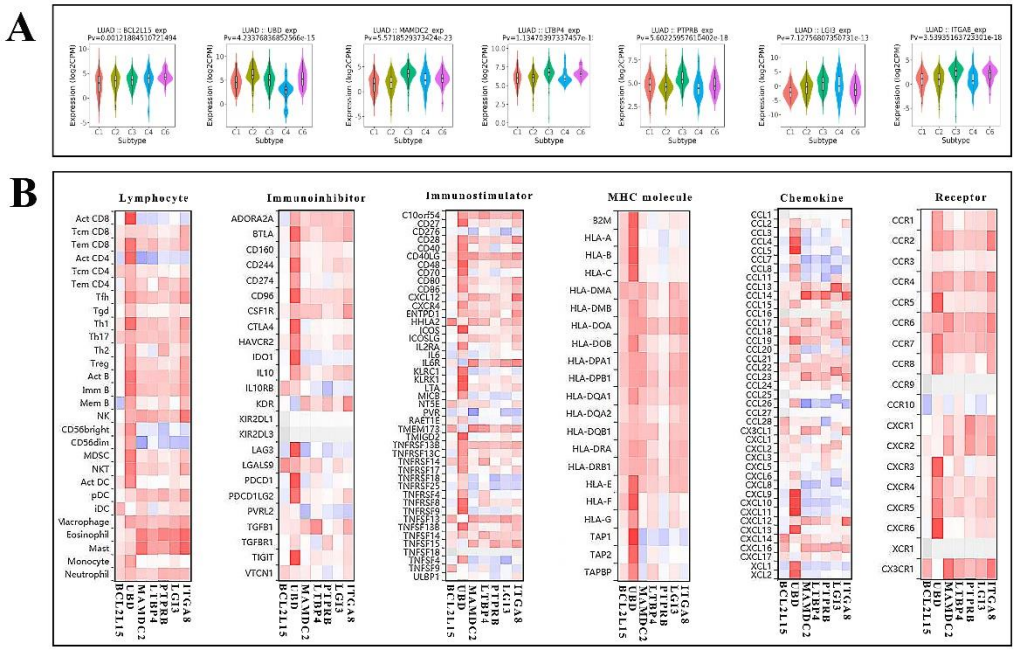

Figure 5. Immune correlation of BCL2L15, UBD, MAMDC2, LTBP4, PTPRB, LGI3 and ITGA8 in LUAD. A, The correlation of immune subtype **vs.** mRNA expression of BCL2L15, UBD, MAMDC2, LTBP4, PTPRB, LGI3 and ITGA8 in LUAD. C1 (wound healing); C2 (IFN-gamma dominant); C3 (inflammatory); C4 (lymphocyte depleted); C5 (immunologically quiet); C6 (TGF-b dominant). **B,** The heatmap showed the correlation of 6 immunological factors (lymphocyte, immune-inhibitor, immune-stimulator, MHC molecule, chemokine, and chemokine receptor) vs. mRNA expression of BCL2L15, UBD, MAMDC2, LTBP4, PTPRB, LGI3 and ITGA8 in LUAD.

PBM alter the prognosis of LUAD

As mentioned above, the PBM of SARS-CoV E protein was a virulence determinant in the host which could lead to an inflammation storm with increased

expression of inflammatory cytokines and a substantially reduced survival. According to MSigDB and TCGA database, total of 20 genes were detected to influence prognosis of LUAD. Among them, CARHSP1, EIF4E2, HMGA1, IL1R2, MAGOHB and PVR which were up-regulated by PBM could lead to poor prognosis with higher expression levels. In contrast, ADCY9, ELF5, ESYT3, SCML4, SECL14L4 and THRA which were down-regulated by PBM could lead to poor prognosis with lower expression levels (Figure 6). Thus, it was concluded that PBM might alter patients' survival by regulating all these 12 DEGs.

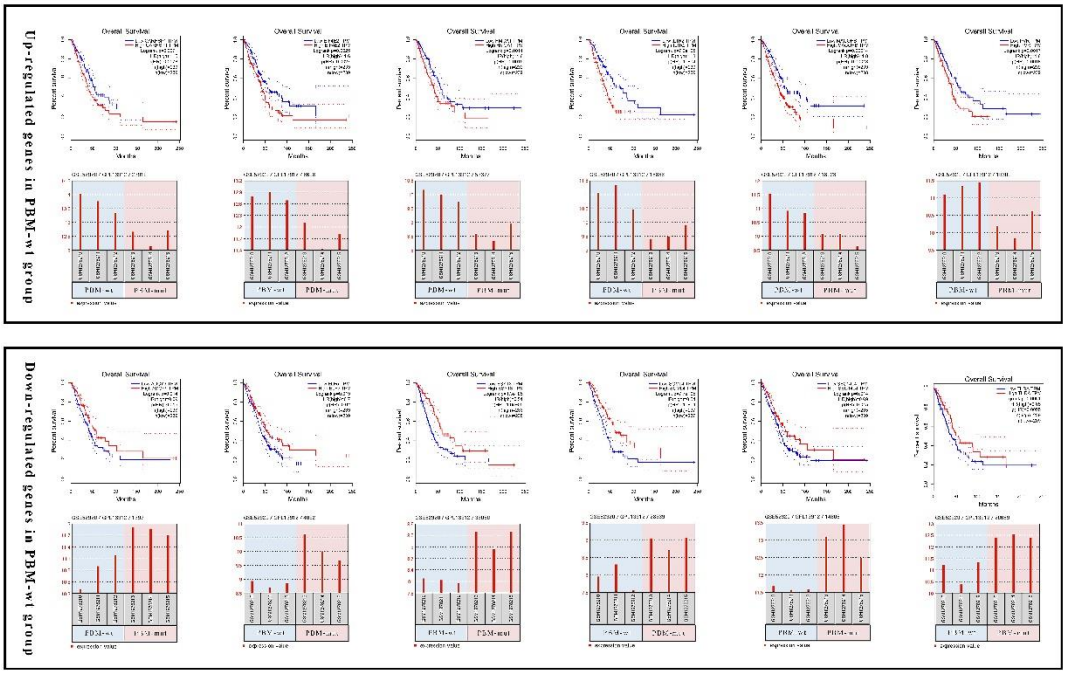

Figure 6. PBM induced DEGs alter prognosis of LUAD. A and B, The overall survival of all 12 DEGs (6 up-regulated and 6 down-regulated) in LUAD patients. The red line represents high TPM while the blue line represents low TPM. Differences in the prognosis of DEGs were statistically significant (p-value＜0.05). The expression level of 12 DEGs in PBM-wt vs. PBM-mut groups.

PBM in coronavirus

Coronavirus is a group of viruses which could cause acute respiratory disease. To confirm that the PBM of E protein is the main pathological cause of epidemic,

homology alignment of E protein sequence was performed. 7 different E proteins of SARS-CoV-2, SARS-related coronavirus, Bat SARS-like coronavirus RsSHC014, Coronavirus BtR1-BetaCoV/SC2018, SARS coronavirus GD01, SARS coronavirus ExoN1 and were similar to SARS-CoV (homology＞90%) with type II PBM (-D-L-L-V-) except for Bat coronavirus BM48-31/BGR/2008 **(Figure 7)**. In addition, only E protein of SARS-CoV-2 lacks an amino acid in hydrophilic carboxyl terminal where the variation of amino acid sequence often appeared. The alignment result further confirmed that the PBM (-D-L-L-V-) may be as a key virulence factor in coronavirus infection.

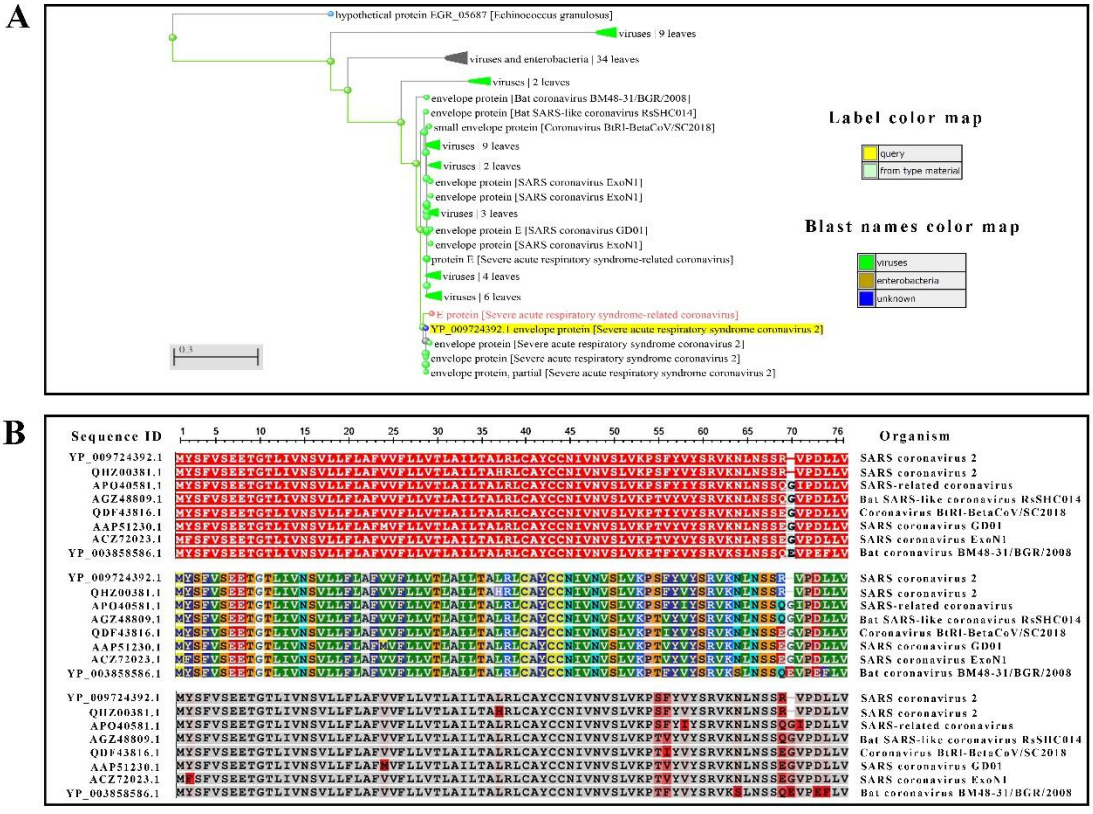

Figure 7. Homologous sequence alignment of SARS-CoV E protein. A, Phylogenetic trees of SARS-CoV E protein sequence. The tree method is fast minimum evolution and the maximum sequencing difference is 0.75. **B,** From top to bottom, amino acid were colored by conservation, RasMol amino acid colors and frequency-based difference.

Discussion

In our previous study, it was confirmed that LUAD patients own increased susceptibility of SARS-CoV due to the higher expression level of angiotensin-converting enzyme 2 (ACE2), the receptor for SARS-CoV in respiratory organs [29, 30]. Nevertheless, the action is always reciprocal. The PBM of SARS-CoV E protein was presently verified to increase the carcinogenesis of LUAD and lead to a poorer prognosis.

MAPK1 is a member of MAP kinase family which are also called extracellular signal-regulated kinases (ERKs) and function as binding points for numerous biochemical signals [31, 32]. Former researches have indicated that MAPK1 participates in various tumor processes [33-35]. In SARS-CoV-E-wt infected lung tissue, MAPK1 was down-regulated while it has been demonstrated MAPK1 could promote the tumorigenesis of LUAD[36]. Thus, we believe there are other regulators participate in LUAD tumorigenesis induced by PBM of SARS-CoV E protein. After bioinformatics analysis, it was found that FGFR4, one of the most significant mutated genes with lower expression level in LUAD, is an upstream regulator of MAPK1 and decreased in SARS-CoV-E-wt infected lung tissue (Figure S3). Furthermore, in LUAD, 8.3% of MAPK1 was mutated. Therefore, mutation of MAPK1 was the primary cause of tumorigenesis.

SARS-CoV-2, a novel coronavirus which also triggers severe acute respiratory syndrome. Similar to SARS-CoV, an open reading frame (ORF) encodes E protein (75 aa) was also located in genome of SARS-CoV-2. Interestingly, through homologous alignment, 2 different sequences of SARS-CoV-2 E proteins were detected (the 37^{th} aa was mutated from L to H). In addition, the 37^{th} aa was located at the transmembrane domain where could form an IC. E protein IC activity is important for SARS-CoV fitness and pathogenesis via promoting ion imbalances within cells [5]. Previously it was reported that the IC of E protein activates the inflammasome through calcium release from intracellular stores [5, 37]. IC of SARS-CoV E protein is exerted in ERGIC membrane which facilitates $Ca2^{+}$ release and contributes to the activation of the

inflammasome complex, leading to the release of pro-inflammatory cytokines such as tumor necrosis factor alpha (TNF-α), IL-1β, and IL-6. The accumulation of these cytokines will lead to an exacerbated pro-inflammatory response, which leads to death. Therefore, it was speculated that there are 2 kinds of SARS-CoV-2 with different pathogenicity actually which results from mutation in E protein. Furthermore, the carboxyl terminal sequences of different coronavirus were poorly conserved even most have PBM which would cause conformational and functional alteration of E proteins.

In summary, SARS-CoV-wt infection will increase the susceptibility of LUAD and cause poor prognosis. The COVID19 caused by SARS-CoV-2 is a more widespread epidemic disease. However, sequencing data about clinical samples of COVID19 is still scarce. Through a retrospective and comparative study of SARS-CoV vs. SARS-CoV-2, it will be helpful to better understand the pathogenesis of SARS-CoV-2 and find therapeutic targets.

Acknowledgement statement

This work was supported by the Key projects of Chongqing Municipal Science and Technology Commission (cstc2017jcyjbx0044) and Fundamental Research Funds for the Central Universities (2019CDQYGD038).

Author contributions

All authors contributed equally to this work. LC and LZ contributed to the conception and design of the work. LC analyzed data and wrote this manuscript. LZ contributed to interpretation of data for the work. All authors contributed to final approval of the version to be published. All authors agree to be accountable for all aspects of the work.

Competing interests

The authors declare that they have no competing interests.

Reference

[1] Enjuanes L, Dediego ML, Alvarez E, Deming D, Sheahan T, Baric R. Vaccines to prevent severe acute respiratory syndrome coronavirus-induced disease. Virus research. 2008;133(1):45-62.

[2] Torres J, Parthasarathy K, Lin X, Saravanan R, Kukol A, Liu DX. Model of a putative pore: the pentameric alpha-helical bundle of SARS coronavirus E protein in lipid bilayers. Biophysical journal. 2006;91(3):938-947.

[3] Verdia-Baguena C, Nieto-Torres JL, Alcaraz A, DeDiego ML, Torres J, Aguilella VM, Enjuanes L. Coronavirus E protein forms ion channels with functionally and structurally-involved membrane lipids. Virology. 2012;432(2):485-494.

[4] Verdia-Baguena C, Nieto-Torres JL, Alcaraz A, Dediego ML, Enjuanes L, Aguilella VM. Analysis of SARS-CoV E protein ion channel activity by tuning the protein and lipid charge. Biochimica et biophysica acta. 2013;1828(9):2026-2031.

[5] Nieto-Torres JL, DeDiego ML, Verdia-Baguena C, Jimenez-Guardeno JM, Regla-Nava JA, Fernandez-Delgado R, Castano-Rodriguez C, Alcaraz A, Torres J, Aguilella VM, Enjuanes L. Severe acute respiratory syndrome coronavirus envelope protein ion channel activity promotes virus fitness and pathogenesis. PLoS pathogens. 2014;10(5):e1004077.

[6] Nal B, Chan C, Kien F, Siu L, Tse J, Chu K, Kam J, Staropoli I, Crescenzo-Chaigne B, Escriou N, van der Werf S, Yuen KY, Altmeyer R. Differential maturation and subcellular localization of severe acute respiratory syndrome coronavirus surface proteins S, M and E. The Journal of general virology. 2005;86(Pt 5):1423-1434.

[7] Nieto-Torres JL, Dediego ML, Alvarez E, Jimenez-Guardeno JM, Regla-Nava JA, Llorente M, Kremer L, Shuo S, Enjuanes L. Subcellular location and topology of severe acute respiratory syndrome coronavirus envelope protein. Virology. 2011;415(2):69-82.

[8] Ruch TR, Machamer CE. A single polar residue and distinct membrane topologies impact the function of the infectious bronchitis coronavirus E protein. PLoS pathogens. 2012;8(5):e1002674.

[9] Ye Y, Hogue BG. Role of the coronavirus E viroporin protein transmembrane domain in virus assembly. Journal of virology. 2007;81(7):3597-3607.

[10] DeDiego ML, Alvarez E, Almazan F, Rejas MT, Lamirande E, Roberts A, Shieh WJ, Zaki SR, Subbarao K, Enjuanes L. A severe acute respiratory syndrome coronavirus that lacks the E gene is attenuated in vitro and in vivo. Journal of virology. 2007;81(4):1701-1713.

[11] Dediego ML, Pewe L, Alvarez E, Rejas MT, Perlman S, Enjuanes L. Pathogenicity of severe acute respiratory coronavirus deletion mutants in hACE-2 transgenic mice. Virology. 2008;376(2):379-389.

[12] Netland J, DeDiego ML, Zhao J, Fett C, Alvarez E, Nieto-Torres JL, Enjuanes L, Perlman S. Immunization with an attenuated severe acute respiratory syndrome coronavirus deleted in E protein protects against lethal respiratory disease. Virology. 2010;399(1):120-128.

[13] Fett C, DeDiego ML, Regla-Nava JA, Enjuanes L, Perlman S. Complete protection against severe acute respiratory syndrome coronavirus-mediated lethal respiratory disease in aged mice by immunization with a mouse-adapted virus lacking E protein. Journal of virology. 2013;87(12):6551-6559.

[14] DeDiego ML, Nieto-Torres JL, Jimenez-Guardeno JM, Regla-Nava JA, Alvarez E, Oliveros JC, Zhao J, Fett C, Perlman S, Enjuanes L. Severe acute respiratory syndrome coronavirus envelope protein regulates cell stress response and apoptosis. PLoS pathogens. 2011;7(10):e1002315.

[15] Ichinohe T, Pang IK, Iwasaki A. Influenza virus activates inflammasomes via its intracellular M2 ion channel. Nature immunology. 2010;11(5):404-410.

[16] Ito M, Yanagi Y, Ichinohe T. Encephalomyocarditis virus viroporin 2B activates NLRP3 inflammasome. PLoS pathogens. 2012;8(8):e1002857.

[17] Andrew A, Strebel K. HIV-1 Vpu targets cell surface markers CD4 and BST-2 through distinct mechanisms. Molecular aspects of medicine. 2010;31(5):407-417.

[18] Hyser JM, Collinson-Pautz MR, Utama B, Estes MK. Rotavirus disrupts calcium homeostasis by NSP4 viroporin activity. mBio. 2010;1(5).

[19] Hung AY, Sheng M. PDZ domains: structural modules for protein complex assembly. The Journal of biological chemistry. 2002;277(8):5699-5702.

[20] Munz M, Hein J, Biggin PC. The role of flexibility and conformational selection in the binding promiscuity of PDZ domains. PLoS computational biology. 2012;8(11):e1002749.

[21] Gerek ZN, Keskin O, Ozkan SB. Identification of specificity and promiscuity of PDZ domain interactions through their dynamic behavior. Proteins. 2009;77(4):796-811.

[22] Javier RT, Rice AP. Emerging theme: cellular PDZ proteins as common targets of pathogenic viruses. Journal of virology. 2011;85(22):11544-11556.

[23] Jimenez-Guardeno JM, Nieto-Torres JL, DeDiego ML, Regla-Nava JA, Fernandez-Delgado R, Castano-Rodriguez C, Enjuanes L. The PDZ-binding motif of severe acute respiratory syndrome coronavirus envelope protein is a determinant of viral pathogenesis. PLoS pathogens. 2014;10(8):e1004320.

[24] Smyth GK. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. Statistical applications in genetics and molecular biology. 2004;3:Article3.

[25] Reiner A, Yekutieli D, Benjamini Y. Identifying differentially expressed genes using false discovery

rate controlling procedures. Bioinformatics (Oxford, England). 2003;19(3):368-375.

[26] Charoentong P, Finotello F, Angelova M, Mayer C, Efremova M, Rieder D, Hackl H, Trajanoski Z. Pan-cancer Immunogenomic Analyses Reveal Genotype-Immunophenotype Relationships and Predictors of Response to Checkpoint Blockade. Cell reports. 2017;18(1):248-262.

[27] Yang X, Deng Y, He RQ, Li XJ, Ma J, Chen G, Hu XH. Upregulation of HOXA11 during the progression of lung adenocarcinoma detected via multiple approaches. International journal of molecular medicine. 2018;42(5):2650-2664.

[28] Hung CF, Wilson CL, Chow YH, Schnapp LM. Role of integrin alpha8 in murine model of lung fibrosis. PloS one. 2018;13(5):e0197937.

[29] Hamming I, Timens W, Bulthuis ML, Lely AT, Navis G, van Goor H. Tissue distribution of ACE2 protein, the functional receptor for SARS coronavirus. A first step in understanding SARS pathogenesis. The Journal of pathology. 2004;203(2):631-637.

[30] Kuba K, Imai Y, Rao S, Gao H, Guo F, Guan B, Huan Y, Yang P, Zhang Y, Deng W, Bao L, Zhang B, Liu G, Wang Z, Chappell M, Liu Y, Zheng D, Leibbrandt A, Wada T, Slutsky AS, Liu D, Qin C, Jiang C, Penninger JM. A crucial role of angiotensin converting enzyme 2 (ACE2) in SARS coronavirus-induced lung injury. Nature medicine. 2005;11(8):875-879.

[31] Wu LK, Liu YC, Ma G, Shi LL, He XM. High levels of glucose promote the activation of hepatic stellate cells via the p38-mitogen-activated protein kinase signal pathway. Genetics and molecular research : GMR. 2016;15(3).

[32] Jung YC, Han S, Hua L, Ahn YH, Cho H, Lee CJ, Lee H, Cho YY, Ryu JH, Jeon R, Kim WY. Kazinol-E is a specific inhibitor of ERK that suppresses the enrichment of a breast cancer stem-like cell population. Biochemical and biophysical research communications. 2016;470(2):294-299.

[33] Wang H, Ke J, Guo Q, Barnabo Nampoukime KP, Yang P, Ma K. Long non-coding RNA CRNDE promotes the proliferation, migration and invasion of hepatocellular carcinoma cells through miR-217/MAPK1 axis. 2018;22(12):5862-5876.

[34] Jin LL, Zhang SJ. miR-574-3p inhibits proliferation and invasion in esophageal cancer by targeting FAM3C and MAPK1. 2019.

[35] Diao L, Wang S, Sun Z. Long noncoding RNA GAPLINC promotes gastric cancer cell proliferation by acting as a molecular sponge of miR-378 to modulate MAPK1 expression. OncoTargets and therapy. 2018;11:2797-2804.

[36] Wang M, Liao Q. PRKCZ-AS1 promotes the tumorigenesis of lung adenocarcinoma via sponging miR-766-5p to modulate MAPK1. 2020:1-8.

[37] Nieto-Torres JL, Verdia-Baguena C, Jimenez-Guardeno JM, Regla-Nava JA, Castano-Rodriguez C, Fernandez-Delgado R, Torres J, Aguilella VM, Enjuanes L. Severe acute respiratory syndrome coronavirus E protein transports calcium ions and activates the NLRP3 inflammasome. Virology. 2015;485:330-339.