Integrative Bioinformatics Analysis Identified Hub Genes in Association with Development of Lung Adenocarcinoma

Xilin Hu

The Affiliated Hospital of Qingdao University

Hanlin Xu

The Affiliated Hospital of Qingdao University

Wenjie Jiao

The Affiliated Hospital of Qingdao University

Kaihua Tian (✉ tiankaihua@qdu.edu.cn)

The Affiliated Hospital of Qingdao University

Research article

Keywords: lung adenocarcinoma, Weighted Gene Co-expression Network Analysis, differential gene expression analysis, differential co-expression genes, targets.

Posted Date: October 6th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-76206/v1>

License: This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background: Cancer-related death is mainly caused by lung adenocarcinoma (LUAD), an aggressive malignant tumor. Therefore, the identification of important LUAD-related genes and the further analysis of its prognostic significance are critical for the survival of LUAD patients.

Methods: Weighted Gene Co-expression Network Analysis (WGCNA) and differential gene expression analysis methods were adopted to screen out TCGA-LUAD database and the gene expression profiles of GSE32863 from GEO. Functional annotation analysis and protein-protein interaction (PPI) network were conducted on differential co-expression genes. Furthermore, survival analysis was carried out on twelve hub genes that were identified by applying the CytoHubba plugin of Cytoscape. Finally, we validated the protein level of survival-related gene in HPA database.

Results: A total of 358 differential co-expression genes were extracted from the database of TCGA and GEO. These genes were mainly enriched in extracellular structure organization, cell-substrate adhesion and response to acid chemical (biological process), cell-cell junction and collagen-containing extracellular matrix (cellular component), DNA-binding transcription activator activity, glycosaminoglycan binding, and growth factor binding (molecular function). In the KEGG analysis, the main pathways were Drug metabolism-cytochrome P450. Moreover, in a PPI network, the 12 hub genes (GNG11, ADRB2, ADCY4, TGFBR2, IL6, GPC3, VIPR1, GRK5, CAV1, RAMP3, RAMP2, and CALCRL) were identified. The expression level of ADCY4, VIPR1, and TGFBR2 was corresponded with clinical stages and overall survival (OS) in LUAD patients. Finally, the protein level of ADCY4 and VIPR1 was down-regulated consistently with mRNA levels in LUAD samples.

Conclusion: Perhaps, ADCY4, VIPR1 and TGFBR2 play important role in tumorigenesis, so they will serve as prognostic biomarkers and therapeutic targets of LUAD in the future.

Background

Among various cancers, lung cancer accounts for the highest incidence and is the leading cause of cancer-related death worldwide⁽¹⁾. Lung cancer can be classified into two pathological forms, including non-small cell lung carcinoma (NSCLC) and small cell lung carcinoma (SCLC). The main subtypes of NSCLC refer to lung adenocarcinoma (LUAD), lung squamous cell carcinoma (LSCC), and large-cell carcinoma⁽²⁾. Recently, it is demonstrated that the morbidity of LUAD is higher than that of LSCC in the world⁽³⁾. Molecular targeted therapy and immunotherapy gradually improved the survival rate of LUAD patients. For instance, tyrosine kinase inhibitors (TKIs) is used as the first-line clinical treatment for advanced LUAD patients suffering from sensitive epidermal growth factor receptor (EGFR) gene mutations⁽⁴⁾. What's more, the purpose of the immunotherapy for lung cancer reverses immune checkpoints, mainly including programmed death protein-1 (PD-1) and programmed death ligand-1 (PD-L1), and obvious and therapeutic effect was observed in specific lung cancer patients⁽⁵⁾. However, the 5-year survival rate of LUAD patients with pessimistic prognosis remains at a low level (6, 7). Due to the

lack of studies on the specific molecular mechanisms underlying the pathogenesis of LUAD, no effective targeted therapies were discovered. Therefore, accurate indicators for the tumorigenesis of LUAD are considered to have the capability to significantly improve the prognosis and treatment of LUAD.

With the development of genomic technologies, it is increasingly popular to explore the molecular mechanisms of cancer and discover the accurate biomarkers with gene expression profiles analysis of bioinformatics(8). A systems biology algorithm of Weighted Gene Co-expression Network Analysis (WGCNA) can be adopted to make a significant association analysis with clinical traits and identify the potential biological functions of these genes(9). One important aim of WGCNA is to construct co-expression modules of highly correlated genes and evaluate the association between interested modules and clinical traits(10). In addition, differential gene expression analyses were utilized to explore molecular mechanisms underlying genome regulation and discover changes of expression levels between tumor and normal tissues(11). Differential gene expression analyses can contribute to the discovery of potential cancer-related biomarkers. Therefore, the results obtained by WGCNA and differential gene expression analysis, were combined to select highly related genes that can be regarded as candidate biomarkers.

In this study, the LUAD-related mRNA expression profiles from the TCGA and GEO database were analyzed to discriminate different co-expression genes using WGCNA and differential gene expression analysis. The LUAD development was further explored through protein-protein interaction (PPI) analysis, functional enrichment analysis and survival analysis. Our study provided a new method to understand potential molecular mechanisms of LUAD with the analysis of differential co-expression genes.

Methods

The bioinformatics analysis process of hub gene is shown in Figure 1.

Data collection

The gene expression profiles of LUAD were downloaded from database of The Cancer Genome Atlas (TCGA, <https://portal.gdc.cancer.gov/>) and the Gene Expression Omnibus (GEO) database of the National Center for Biotechnology Research (<http://www.ncbi.nlm.nih.gov/geo/>). The TCGA data containing information of 535 LUAD samples and 44 normal tissues, were downloaded by R package TCGA biolinks(12). In this study, genes were kept with a cpm (count per million) ≥ 1 . We use function `rpkms` in `edgeR` package(13) to filter a total of 15,127 genes with RPKM values for next analysis. Gene expression profile of GSE32863 were downloaded from and performed on Illumina HumanWG-6 v3.0 expression beadchip (GPL6884). GSE32863 included 58 lung adenocarcinoma and 58 adjacent non-tumor lung tissues from LUAD patients. Data-set GSE32863 was used as a training set to construct co-expression networks and identify hub genes in this study.

Identification of Functional Modules In A Co-expression Network

WGCNA package in R was used to construct co-expression network for data profiles of TCGA-LUAD and GSE32863(9). For relating modules to external sample traits, the modules of highly correlated genes among samples were explored by WGCNA. In order to build a scale-free network, we selected soft powers $\beta = 3$ and 20 according to the pickSoftThreshold function. Next, the following formula: $a_{ij} = |S_{ij}|^\beta$ (a_{ij} : adjacency matrix between gene i and gene j , S_{ij} : similarity matrix which is done by Pearson correlation of all gene pairs, β : soft power value) was used to create the adjacency matrix. Then, topological overlap matrix (TOM) and the corresponding dissimilarity (1-TOM) were transformed. Afterwards, the similar gene expressions were classified into different gene co-expression modules by constructing a hierarchical clustering of genes based on the 1-TOM matrix. According to the previous study, we calculated the module-trait associations to identify functional modules(14). Therefore, high-related coefficient modules were relevant to clinical traits, and were considered candidates for subsequent analysis. It has a specific description of these method in the previous study(14).

Differentially Expressed Genes (DEGs) Screening and Interaction With the Functional Modules

Differential expression analyses on RNA-Sequencing and microarray data were integrated using the R package limma(15). Limma was respectively applied in the TCGA-LUAD and GSE32863 dataset to select out differentially expressed genes (DEGs) between LUAD and normal tissues. To control for the false discovery Rate, the Benjamini–Hochberg method was used to adjusted p -value. The threshold for identifying DEGs was set at $|\log FC| \geq 1.0$ and $\text{adj. } p < 0.05$. Finally, in order to identify potential prognostic genes, we use R package VennDiagram to construct the overlapping genes between DEGs and co-expression genes.

Enrichment Analysis for the potential prognostic genes

A Gene Ontology (GO) analysis contains three domains: molecular functions (MFs), cellular components (CCs) and biological processes (BPs). The potential functions of the aberrantly expressed intersecting genes involved in signaling pathways were analyzed using the Kyoto Encyclopedia of Genes and Genomes (KEGG). R package clusterProfiler package(16) was used to explore the functions among genes of interest, with a cut-off criterion of adjusted $p < 0.05$.

Construction of PPI and Screening of Hub Genes

Protein-protein network (PPI) of these candidate hub genes were constructed by STRING online tool(<https://string-db.org>)(17). In this study, we set a combined score ≥ 0.4 as the cutoff value to build a visualized network model using Cytoscape (v3.7.2)(18). Degree algorithm, calculated by CytoHubba (a plugin in Cytoscape)(19), was used to find hub nodes in a co-expression network. Subsequently, we selected the gene with the top 12 degree values (with a connectivity degree of ≥ 6) as hub genes.

Clinical relevance and Prognostic Values of Hub Gene

Based on the TCGA dataset, we analyzed the relative expression of hub genes among tumor individual cancer stages. Survival analysis for the hub genes was performed by Kaplan–Meier univariate survival

analysis with the TCGA database, using the survival package in R software. Moreover, GEPIA (<http://gepia.cancerpku.cn/index.html>) was used to evaluate the association between disease-free survival (DFS) and hub genes expressed in LUAD patients(20). Log-rank $p < 0.05$ were considered as statistically significant.

Protein Expressions of Survival-Related Hub Genes

The Human Protein Atlas database (HPA, <https://www.proteinatlas.org/>) were used to validate the expression differences of the survival-related genes between LUAD and normal tissues by immunohistochemistry (IHC).

Result

Identification of Co-Expression Gene Modules

We used the WGCNA package on R to construct gene co-expression networks from the TCGA-LUAD and GSE32863 datasets. Totally, 15 modules in the TCGA-LUAD and 7 modules in the GSE32863 (Figure 2A and Figure 3A) were identified, and a gray module was excluded, because it was assigned into no cluster. Then, the heatmap of module-trait relationships was plotted to evaluate the relevance between the two clinical traits and each module. The results in Figure 2B and 3B, revealed that the turquoise module in the TCGA-LUAD and brown module in the GSE32863 were proved to be highest associated with normal lung tissue (turquoise module: $r = 0.79$, $p = 5e-127$, and brown module: $r = 0.97$, $p = 9e-71$). Moreover, compared with normal tissues, the expression of the genes in the turquoise module and brown module was all downregulated in LUAD tissues.

Extraction of Overlapping Genes between the DEGs and Co-expression Networks

We discovered that 3,583 DEGs in the TCGA-LUAD dataset (Figure 4A) and 957 DEGs in the GSE32863 dataset (Figure 4B) were dysregulated in tumor tissues using the limma package (cut-off criteria of $|\log FC| \geq 1.0$, adj. $p < 0.05$). A total of 5313 and 1699 co-expression genes were enrolled in the turquoise module of TCGA-LUAD dataset and the brown module in GSE32863, respectively. Totally, as shown in Figure 4C, 358 overlapping genes were extracted to verify the genes among co-expression networks.

Functional Enrichment Analysis

In order to further explore the potential functions of the 358 overlapping genes, GO and KEGG enrichment analyses were performed by the cluster profiler package. Several GO-enriched gene sets were displayed in Figure 5. The results of the biological process (BP) revealed that these genes are mainly enriched in extracellular structure organization, cell-substrate adhesion and response to acid chemical. The cellular component (CC) of these genes were mainly involved in cell-cell junction and collagen-containing extracellular matrix. Most importantly, molecular function (MF) analysis proved that DNA-binding transcription activator activity, glycosaminoglycan binding and growth factor binding were suggested to be associated with these genes. In the KEGG analysis, the main pathways were enriched by these genes

such as Drug metabolism – cytochrome P450, Ether lipid metabolism and Arachidonic acid metabolism (Figure 6).

Construction of PPI Network and Identification of Hub Genes

We applied the STRING database to establish the PPI network among the overlapped genes (Figure 7A). Figure 7B gives information that hub genes were extracted from the PPI network with a connectivity degree ranking. Finally, the top 12 highest-scored genes, including GNG11, ADRB2, ADCY4, TGFBR2, IL6, GPC3, VIPR1, GRK5, CAV1, RAMP3, RAMP2, and CALCRL were selected as hub genes.

Clinical relevance of hub gene expression in LUAD patients

After the 12 hub genes were screened out, based on the dataset in TCGA, we validated that ADCY4 expression had a strong positive correlation with tumor individual cancer stages by applying package *cliCor* in R ($p < 0.05$; Figure 8A). Then with VIPR1, stage I had a significant expression difference with stage IV ($p = 0.033$, Figure 8E). Finally, TGFBR2 expression of stage I and stage III had significant differences ($p = 0.024$, Figure 8I).

Prognostic Values and Protein Expression of 12 Hub Genes

We applied the R survival package and the GEPIA2 database to perform OS (Figure 9) and DFS (Figure 10) analyses of the twelve hub genes by Kaplan–Meier plotter, respectively. The Kaplan–Meier analysis revealed that the lower expression level of ADCY4, VIPR1, and TGFBR2 was obviously related to the worse OS of the LUAD patients (ADCY4: $p = 0.003$, VIPR1: $p = 0.019$, and TGFBR2: $p = 0.002$) (Figure 9A, 9E, and 9I). No significant difference was found in LUAD patients with the expression level of ADCY4, VIPR1 and TGFBR2 in DFS ($p > 0.05$) (Figure 10A, 10E, and 10I). Furthermore, based on the HPA database, the protein level of the ADCY4 and VIPR1 gene was extremely lower in LUAD tissues, compared with normal tissues (Figure 11). The expression level of TGFBR2 is not discovered in the HPA database.

Discussion

Although molecular targeted therapy and immunotherapy have been gradually applied in LUAD patients, the prognosis of LUAD is generally poor. Different cancer subtypes present different molecular targets. Thus, it is vital to explore better biomarkers for specific prognosis and progression of LUAD. In this study, through integrated bioinformatic analysis of DEGs and co-expression genes that are associated with LUAD, we finally obtained 358 genes with the same expression trends in the TCGA and GSE32863 databases. GO and KEGG pathway enrichment analyses suggested that differential co-expression genes were mainly enriched in the pathways that are related to cell proliferation, invasion and migration. Furthermore, according to the connectivity degree ranking from the CytoHubba plugin in Cytoscape, we screened out the top 12 LUAD-related genes, including GNG11, ADRB2, ADCY4, TGFBR2, IL6, GPC3, VIPR1, GRK5, CAV1, RAMP3, RAMP2, and CALCRL. Among them, the expression of ADCY4, VIPR1 and TGFBR2 were found to be associated with clinical stages of LUAD. Moreover, the down expression of

ADCY4, VIPR1, and TGFBR2 was closely related to poor OS in LUAD. Finally, we carried out survival and immunohistochemical analysis for ADCY4, VIPR1 and TGFBR2.

The cAMP signaling pathway can be activated by specific adenylyl cyclases. Due to the distinct tissue distribution of specific AC isoforms, cAMP can selectively regulate effector proteins. Therefore, ACs often integrate cellular signals and are regarded as key enzymes in these signaling pathways(21). For example, high expression of ADCY7 resulted in growth, stimulating apoptosis and increased c-Myc expression in leukemic cells(22). One study illustrated that, served as an oncogene for gastric cancer, ADCY3 could be regulated by DNA methylation(23, 24). It was reported that ADCY4 was significantly downregulated in breast cancer. ADCY4 was strongly related to G protein coupled receptors and the cAMP signaling pathway. ADCY4 expression was correlated with OS and tumor stages of patients suffering from breast cancer (25). Recent studies reported that VIPR1 can inhibit the development of several cancers, including medulloblastoma(26), prostate cancer(27) and lymphoblastoma(28), suggesting that VIPR1 plays an important role in the inhibition of the growth and development of cancer cells. In hepatocellular carcinoma, DNA methylation of VIPR1 significantly inhibited gene transcription and low expression of VIPR1 was related to poor survival and histological differentiation(29). It was reported that VIPR1 was down-regulated in lung adenocarcinoma tissues compared with normal tissues(30). The expression of VIPR1 was lower in lung cancer cells H1299 than that in normal lung epithelial cells BEAS-2B. Moreover, the growth, migration, and invasion of H1299 cells were remarkably inhibited by VIPR1 (30). TGFBR2 can inhibit the malignant progression of various cancers, such as glioblastoma(31), cervical cancer(32), bladder cancer(33) and so on. In the TGF- β signaling, phosphorylation and activation of TGFBR1 begins with the binding of TGF- β ligands to TGFBR2(34). Afterwards, the activated TGFBR1 activates Smad proteins that can suppress EMT in cancer(34–36). Li's study proved that the lower expression of TGFBR2 was related to up-regulated abilities of proliferation, migration and invasion in lung cancer cell (A549) (37).

In our work, several limitations were also pointed out. Although the potential diagnostic genes between LUAD and normal tissues were identified using comprehensive bioinformatics analysis, it was hard to be very accurate for each LUAD patient. Furthermore, the specific molecular mechanisms of ADCY4, VIPR1 and TGFBR2 which had a significant effect in the prognosis of LUAD patients should be verified through several experiments in the future.

Conclusion

This study aimed at identifying survival-related genes that are associated with LUAD, thus further developing a new therapeutic strategy for targeted genes. Through integrated bioinformatic analysis of the DEGs and co-expression genes associated with LUAD, we finally obtained 358 genes with same expression trends. Further analysis suggested that three hub genes, including ADCY4, VIPR1 and TGFBR2, were significantly related to tumor clinical stages and overall survival. Therefore, ADCY4, VIPR1 and TGFBR2 may be important in the tumorigenesis and serve as prognostic biomarkers and therapeutic targets of LUAD in the future.

Abbreviations

LUAD

lung adenocarcinoma

PPI

protein-protein interaction

NSCLC

non-small cell lung carcinoma

SCLC

small cell lung carcinoma

LUSC

lung squamous cell carcinoma

TKIs

tyrosine kinase inhibitors

EGFR

epidermal growth factor receptor

PD-1

programmed death protein-1

PD-L1

programmed death ligand-1

IHC

immunohistochemistry

HPA

Human Protein Atlas database

GEO

Gene Expression Omnibus

TCGA

The Cancer Genome Atlas

WGCNA

Weighted gene co-expression network analysis

FPKM

Fragments per kilobase of transcript per million

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and material

The datasets analyzed during the current study are available in the The Cancer Genome Atlas (<https://portal.gdc.cancer.gov>), The Gene Expression Omnibus database (<http://www.ncbi.nlm.nih.gov/geo>) and The Human Protein Atlas database(<https://www.proteinatlas.org>).

Competing interests

The authors declare that they have no potential conflicts of interest.

Funding

Not applicable.

Authors' contributions

XLH has made substantial contributions to the design of the work, the analysis and interpretation of data as well as drafted the work. HLX has agreed to be personally accountable for the author's own contributions and to ensure that questions related to the accuracy or integrity of any part of the work. WJJ and KHT have approved the submitted version All authors have read and approved the manuscript.

Acknowledgements

Not applicable.

References

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *Cancer J Clin.* 2018;68(6):394–424.
2. Vuong HG, Nguyen TQ, Nguyen HC, Nguyen PT, Ho ATN, Hassell L. Efficacy and Safety of Crizotinib in the Treatment of Advanced Non-Small-Cell Lung Cancer with ROS1 Rearrangement or MET Alteration: A Systematic Review and Meta-Analysis. *Targeted oncology.* 2020.
3. Cancer Genome Atlas Research N. Comprehensive molecular profiling of lung adenocarcinoma. *Nature.* 2014;511(7511):543–50.
4. Zhou C, Yao LD. Strategies to Improve Outcomes of Patients with EGRF-Mutant Non-Small Cell Lung Cancer: Review of the Literature. *Journal of thoracic oncology: official publication of the International Association for the Study of Lung Cancer.* 2016;11(2):174–86.
5. Villanueva N, Bazhenova L. New strategies in immunotherapy for lung cancer: beyond PD-1/PD-L1. *Therapeutic advances in respiratory disease.* 2018;12:1753466618794133.

6. Riaz SP, Luchtenborg M, Coupland VH, Spicer J, Peake MD, Moller H. Trends in incidence of small cell lung cancer and all lung cancer. *Lung cancer*. 2012;75(3):280–4.
7. Ni M, Shi XL, Qu ZG, Jiang H, Chen ZQ, Hu J. Epithelial mesenchymal transition of non-small-cell lung cancer cells A549 induced by SPHK1. *Asian Pacific journal of tropical medicine*. 2015;8(2):142–6.
8. Ayyildiz D, Piazza S. Introduction to Bioinformatics. *Methods in molecular biology*. 2019;1986:1–15.
9. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinform*. 2008;9:559.
10. Zhang B, Horvath S. A general framework for weighted gene co-expression network analysis. *Stat Appl Genet Mol Biol*. 2005;4:Article17.
11. Segundo-Val IS, Sanz-Lozano CS. Introduction to the Gene Expression Analysis. *Methods in molecular biology*. 2016;1434:29–43.
12. Colaprico A, Silva TC, Olsen C, Garofano L, Cava C, Garolini D, et al. TCGAbiolinks: an R/Bioconductor package for integrative analysis of TCGA data. *Nucleic acids research*. 2016;44(8):e71.
13. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*. 2010;26(1):139–40.
14. Wang CCN, Li CY, Cai JH, Sheu PC, Tsai JJP, Wu MY, et al. Identification of Prognostic Candidate Genes in Breast Cancer by Integrated Bioinformatic Analysis. *Journal of clinical medicine*. 2019;8(8).
15. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic acids research*. 2015;43(7):e47.
16. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. *Omics: a journal of integrative biology*. 2012;16(5):284–7.
17. Nagy J, Per B, Trinn C, Nagy G, Burger T. [Incidence of IgA (monomer-dimer, IgA1-IgA2) and IgG producing cells in the tonsils of patients with IgA nephropathy]. *Orv Hetil*. 1988;129(28):1481–5.
18. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome research*. 2003;13(11):2498–504.
19. Chin CH, Chen SH, Wu HH, Ho CW, Ko MT, Lin CY. cytoHubba: identifying hub objects and sub-networks from complex interactome. *BMC systems biology*. 2014;8(Suppl 4):11.
20. Tang Z, Kang B, Li C, Chen T, Zhang Z. GEPIA2: an enhanced web server for large-scale expression profiling and interactive analysis. *Nucleic acids research*. 2019;47(W1):W556-W60.
21. Pierre S, Eschenhagen T, Geisslinger G, Scholich K. Capturing adenylyl cyclases as potential drug targets. *Nature reviews Drug discovery*. 2009;8(4):321–35.
22. Li C, Xie J, Lu Z, Chen C, Yin Y, Zhan R, et al. ADCY7 supports development of acute myeloid leukemia. *Biochem Biophys Res Commun*. 2015;465(1):47–52.
23. Hong SH, Goh SH, Lee SJ, Hwang JA, Lee J, Choi IJ, et al. Upregulation of adenylate cyclase 3 (ADCY3) increases the tumorigenic potential of cells by activating the CREB pathway. *Oncotarget*.

2013;4(10):1791–803.

24. Xiang TX, Yuan Y, Li LL, Wang ZH, Dan LY, Chen Y, et al. Aberrant promoter CpG methylation and its translational applications in breast cancer. *Chinese journal of cancer*. 2013;32(1):12–20.
25. Fan Y, Mu J, Huang M, Imani S, Wang Y, Lin S, et al. Epigenetic identification of ADCY4 as a biomarker for breast cancer: an integrated analysis of adenylate cyclases. *Epigenomics*. 2019;11(14):1561–79.
26. Fruhwald MC, O'Dorisio MS, Fleitz J, Pietsch T, Reubi JC. Vasoactive intestinal peptide (VIP) and VIP receptors: gene expression and growth modulation in medulloblastoma and other central primitive neuroectodermal tumors of childhood. *International journal of cancer*. 1999;81(2):165–73.
27. Collado B, Carmena MJ, Sanchez-Chapado M, Ruiz-Villaespesa A, Bajo AM, Fernandez-Martinez AB, et al. Expression of vasoactive intestinal peptide and functional VIP receptors in human prostate cancer: antagonistic action of a growth-hormone-releasing hormone analog. *Int J Oncol*. 2005;26(6):1629–35.
28. Goursaud S, Pineau N, Becq-Giraudon L, Gressens P, Muller JM, Janet T. Human H9 cells proliferation is differently controlled by vasoactive intestinal peptide or peptide histidine methionine: implication of a GTP-insensitive form of VPAC1 receptor. *J Neuroimmunol*. 2005;158(1–2):94–105.
29. Abad C, Jayaram B, Becquet L, Wang Y, O'Dorisio MS, Waschek JA, et al. Erratum to: VPAC1 receptor (Vipr1)-deficient mice exhibit ameliorated experimental autoimmune encephalomyelitis, with specific deficits in the effector stage. *J Neuroinflamm*. 2017;14(1):157.
30. Zhao L, Yu Z, Zhao B. Mechanism of VIPR1 gene regulating human lung adenocarcinoma H1299 cells. *Medical oncology*. 2019;36(11):91.
31. Xu H, Zhang B, Yang Y, Li Z, Zhao P, Wu W, et al. LncRNA MIR4435-2HG potentiates the proliferation and invasion of glioblastoma cells via modulating miR-1224-5p/TGFBR2 axis. *J Cell Mol Med*. 2020;24(11):6362–72.
32. Yuan J, Yi K, Yang L. TGFBR2 Regulates Hedgehog Pathway and Cervical Cancer Cell Proliferation and Migration by Mediating SMAD4. *J Proteome Res*. 2020;19(8):3377–85.
33. Li Y, Qiao L, Zang Y, Ni W, Xu Z. Circular. RNA FOXO3 Suppresses Bladder Cancer Progression and Metastasis by Regulating MiR-9-5p/TGFBR2. *Cancer management research*. 2020;12:5049–56.
34. Kim JH, Jang YS, Eom KS, Hwang YI, Kang HR, Jang SH, et al. Transforming growth factor beta1 induces epithelial-to-mesenchymal transition of A549 cells. *J Korean Med Sci*. 2007;22(5):898–904.
35. Keklikoglou I, Koerner C, Schmidt C, Zhang JD, Heckmann D, Shavinskaya A, et al. MicroRNA-520/373 family functions as a tumor suppressor in estrogen receptor negative breast cancer by targeting NF-kappaB and TGF-beta signaling pathways. *Oncogene*. 2012;31(37):4150–63.
36. Harazono Y, Muramatsu T, Endo H, Uzawa N, Kawano T, Harada K, et al. miR-655 Is an EMT-suppressive microRNA targeting ZEB1 and TGFBR2. *PloS one*. 2013;8(5):e62757.
37. Li G, Wu F, Yang H, Deng X, Yuan Y. MiR-9-5p promotes cell growth and metastasis in non-small cell lung cancer through the repression of TGFBR2. *Biomedicine & pharmacotherapy = Biomedecine & pharmacotherapie*. 2017;96:1170-8.

Figures

Figure 1

The bioinformatics analysis process.

Figure 2

Identification of gene modules and the association with clinical traits of LUAD in the TCGA dataset. (A) The Cluster dendrogram of co-expression network modules was ordered by a hierarchical clustering of genes based on the 1-TOM matrix. Each module was assigned different colors. (B) Module-trait relationships. Each row corresponds to a color module and column corresponds to a clinical trait (cancer and normal). Each cell contains the corresponding correlation and p-value.

Figure 3

Identification of gene modules and the association with clinical traits of LUAD in the GSE32863 dataset. (A) The Cluster dendrogram of co-expression network modules was ordered by a hierarchical clustering of genes based on the 1-TOM matrix. Each module was assigned different colors. (B) Module-trait relationships. Each row corresponds to a color module and each column correlates to a clinical trait (cancer and normal). Each cell contains the corresponding correlation and p-value.

Figure 4

Identification of differentially expressed genes (DEGs) among the TCGA and GSE32863 datasets of LUAD with the cut-off criteria of $|\log FC| \geq 1.0$ and $\text{adj. } p < 0.05$. (A) Volcano plot of DEGs in the TCGA dataset. (B) Volcano plot of DEGs in the GSE32863 dataset. (C) The Venn diagram of genes among DEG lists and co-expression module. In total, 358 overlapping genes in the intersection of DEG lists and two co-expression modules.

Figure 5

Gene Ontology (GO) enrichment analysis for overlapping genes. The color represents the adjusted p-values.

Figure 6

Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis for overlapping genes. The color represents the adjusted p-values.

Figure 7

Visualization of the protein-protein interaction (PPI) network and the candidate hub genes. (A) PPI network of the genes between DEG lists and two co-expression modules. The blue nodes represent the genes. Edges indicate interaction associations between nodes. (B) Identification of the hub genes from the PPI network using degree algorithm. Edges represent the protein-protein associations. The red nodes represent genes with a high degree scores, while the yellow node represent genes with a low degree score.

Figure 8

Aberrant expression of the twelve hub genes among differential clinical stages in the TCGA-LUAD database. (A) Gene expression of ADCY4 among clinical stages. (B) Gene expression of RAMP3 among clinical stages. (C) Gene expression of RAMP2 among clinical stages. (D) Gene expression of IL6 among clinical stages. (E) Gene expression of VIPR1 among clinical stages. (F) Gene expression of TGFBR2 among clinical stages. (G) Gene expression of CAV1 among clinical stages. (H) Gene expression of GNG11 among clinical stages. (I) Gene expression of CALCRL among clinical stages. (J) Gene expression of ADRB2 among clinical stages.

Figure 9

Overall survival (OS) analysis of 12 hub genes in LUAD patients from the TCGA database. (A) Survival analysis for ADCY4 in LUAD. (B) Survival analysis for RAMP3 in LUAD. (C) Survival analysis for RAMP2 in LUAD. (D) Survival analysis for IL6 in LUAD. (E) Survival analysis for VIPR1 in LUAD. (F) Survival analysis for TGFBR2 in LUAD. (G) Survival analysis for CAV1 in LUAD. (H) Survival analysis for GNG11 in LUAD. (I) Survival analysis for CALCRL in LUAD. (J) Survival analysis for ADRB2 in LUAD. The patients were stratified into high-level group (red) and low-level group (blue) according to median expression of the gene. Log-rank $p < 0.05$ was considered to be a statistically significant difference.

Figure 10

Disease-free survival (DFS) analysis of 12 hub genes in LUAD patients from the GEPIA2 database. (A) Survival analysis for ADCY4 in LUAD. (B) Survival analysis for RAMP3 in LUAD. (C) Survival analysis for RAMP2 in LUAD. (D) Survival analysis for IL6 in LUAD. (E) Survival analysis for VIPR1 in LUAD. (F) Survival analysis for TGFBR2 in LUAD. (G) Survival analysis for CAV1 in LUAD. (H) Survival analysis for GNG11 in LUAD. (I) Survival analysis for CALCRL in LUAD. (J) Survival analysis for ADRB2 in LUAD. The patients were stratified into high-level group (red) and low-level group (blue) according to median expression of the gene. Log-rank $p < 0.05$ was considered to be a statistically significant difference.

Figure 11

Immunohistochemistry of the ADCY4 and VIPR1 gene in LUAD and normal tissues from the human protein atlas (HPA) database. (A) Protein levels of ADCY4 in LUAD tissues (antibody CAB005070; staining: low; intensity: weak; quantity: >75%). (B) Protein levels of ADCY4 in normal lung tissues (antibody CAB005070; staining: medium; intensity: moderate; quantity: >75%). (C) Protein levels of VIPR1 in LUAD tissues (antibody HPA026777; staining: low; intensity: weak; quantity: >75%). (D) Protein levels

of VIPR1 in normal lung tissues (antibody HPA026777; staining: medium; intensity: moderate; quantity: >75%).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [datasets.docx](#)