

Potential for the blood- biomarker cell free UPK2 RNA to detect recurrence during monitoring after surgical resection of adenocarcinoma of the lung

Ji Zhu

First Affiliated Hospital of the Second Military Medical University

Qijue Lu

First Affiliated Hospital of the Second Military Medical University

Bin Li

Shanghai Chest Hospital

Huafei Li

School of Life Sciences, Shanghai University

Cong Wu

First Affiliated Hospital of the Second Military Medical University

Chunguang Li

First Affiliated Hospital of the Second Military Medical University

Hai Jin (✉ jinhai_wenzhang@163.com)

First Affiliated Hospital of the Second Military Medical University

Research article

Keywords: Lung adenocarcinoma, TCGA database, WGCNA, Plasma, UPK2

Posted Date: July 29th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-25690/v2>

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

[Read Full License](#)

Abstract

Background: Lung cancer is one of the most common malignant tumors in the world. Non-small-cell lung cancer (NSCLC) accounting for top of all diagnosed lung cancers. Lung adenocarcinoma(LUAD) is the most common subtype in NSCLC. This study attempts to identify and validate biomarkers that can be used to monitor recurrence after LUAD surgery.

Methods: In this research, we downloaded lung adenocarcinoma data from the TCGA database and selected postoperative recurrence samples, and then performed WGCNA analysis to find key co-expression gene modules. Enrichment analysis of key gene modules was also performed using the DAVID database. Finally, we performed a survival analysis of the most interesting biomarker UPK2 obtained from the TCGA analysis in the Oncomine database and evaluated its impact on prognosis. We collected 132 blood samples from patients with early-stage lung adenocarcinoma who were tested the expression level of free mRNA in the plasma.

Results: The results revealed that UPK2, KLHDC3, GALR2 and TYRP1 occupied a central position in the co-expression network which were also significantly correlated with the survival of patients. The expression level of free UPK2 in the plasma relative to GADPH in non-relapsed patients was 0.1623, and in relapsed patients, it was 0.2763. ROC was used to evaluate the effectiveness of free UPK2 mRNA in the blood for monitoring postoperative recurrence, with an AUC of 0.767 and a 95% confidence interval of 0.675-0.858. At the same time, patients with high expression of free UPK2 mRNA had significantly poorer survival than those with low expression of UPK2.

Conclusions: The expression level of free UPK2 mRNA in plasma has the potential as an indicator of postoperative recurrence in patients with early LUAD. This may have guiding significance for the subsequent clinical treatment of patients.

Introduction

Non-small-cell lung cancer (NSCLC) accounting for approximately 85% of all diagnosed lung cancers(1). Lung adenocarcinoma (LUAD) is the most common subtypes of NSCLC and the rate of morbidity and mortality have risen in recent years(2,3).But progress in science and technology has promoted new treatment options for patients greatly. Han et al (4)discovered that pemetrexed plus carboplatin combined with gefitinib extended survival for patients with LUAD harboring sensitive EGFR mutations. Also, Zhuang et al (5) showed that combine nadroparin with radiotherapy induces stronger synergistic antitumor effects in LUAD A549 cells. Meanwhile, new studies showing that current treatment strategies can be further improved(6-8).

Despite advances in surgery, molecular subtyping and targeted therapy, prognosis of LUAD remains relatively poor(9). Patients with LUAD often relapse and develop metastases after surgery, chemotherapy and radiation therapy(10). Because of those malignant features, patients just have the 50%-70% 5-year survival rate even in early stage of LUAD(11). Moreover, patients with advanced LUAD, which is often

resistant to conventional chemotherapies or targeted therapeutic drugs(12). Almost without exception, all of the high-potency anticancer drugs are not long term effective. Cancer cells become resistant to anticancer drugs as they rapidly mutant. Although the surgical methods of cancer are constantly improved and chemotherapy drugs are updated, the problem of completely removing residual cancer cell still cannot be solved. So, the risk of recurrence is still high. The recurrence of cancer can make a big challenge to the further treatment and the deterioration is so rapid. Also, the recurrent correlation mechanism have not yet been specific. Among 289 patients with stage I and II lung adenocarcinoma, 85 had found a distant recurrence in five years(13). So the study of cancer recurrence may have a very important clinical value.

With the development of precision medicine, more and more molecules, the so-called biomarker molecules, have been found to have a high correlation with the certain biological events. Norman et al. have found that Ink4a/Arf expression is a biomarker of aging, and the increasing expression of Ink4a/Arf is related to advancing age in all rodent tissue (14). In addition, CSF A β 1–42 is found and verified as the biomarker for Alzheimer's disease in the autopsy cohort of CSF sample with high sensitivity for detection of 96.4%. (15), the biomarkers are not just limited to the proteins, and also mRNA can act as the biomarker for some biomolecular issues. Naoharu Iwai et al. have already found that miR-208 can be used as the useful indicator of myocardial injury (16). Obviously, the discovery and verification of tumor molecular markers is one of most popular focuses in this field, and significant progress has been made, especially in the discovery and identification of molecular markers associated with the clinic effects of tumor therapy. In 2011, Human epididymis protein 4 (HE4) was approved by the Food and Drug Administration to monitor recurrence or progressive disease in epithelial ovarian cancer in conjunction with CA125 (17). Chen et al. reported that abnormalities of AIB1 was tightly associated with clinical/prognostic significance of urothelial carcinoma, and high AIB1 expression was associated with increased hazard ratios for 5-year CSS (80.6% vs. 55.8%, $p= 0.008$) and OS (78.1% vs. 54.8%, $p= 0.006$) (18). The HOXB13/IL17BR (H/I) biomarker predicted recurrence risk in ER-positive, lymph node–negative breast cancer patients (19). There were also a considerable number of reports on this field, and more and more biomarkers were verified and applied to clinics, leading to the corresponding patients benefit from this progress.

In this study, the bioinformatical approaches were employed to analyze a robust of sequence data, and several biomarkers were delved and then verified by using the clinic samples. The obtained biomarkers in this study may play a key role in the cancer recurrence.

Methods

Data analysis:

The datasets used in this study were downloaded from the TCGA database. The softConnectivity function in the WGCNA package was used to analyze the effects of different power values on the co-expression network and co-expression modules in the scale independence and average connectivity . The

"randomly selected gene" parameter was set to 5000, and the other parameters were set to default values. Next, we summarized the expression values by using the collapseRows function in the WGCNA package, and then performed cluster analysis using flashClust, and the interaction/correlation of each module was visualized through heat map. In addition, the interested gene modules were cluster analyzed by the GO in the DAVID database(DAVID <https://david.ncifcrf.gov/summary.jsp>). And the survival analysis of the corresponding genes were performed by using Kaplan-Meier analysis and Oncomine database.

Patient recruitment

We collected blood samples from patients with early-stage lung adenocarcinoma who went to the Department of Oncology, Shanghai Changhai Hospital for surgery from February 2006 to March 2010. All patients received systematic treatment after surgery, including but not limited to the optimal local treatments, such as radiochemotherapy and targeted therapy. Recurrence of the disease was based on imaging findings (CXR or CT). This study was in accordance with the Declaration of Helsinki and Good Clinical Practice and was reviewed by Shanghai Changhai Hospital Ethics Committee. All patients signed informed consent with respect to the use of their blood for scientific purposes.

Sample collection and processing

The blood samples of 132 patients were collected for the first time at 90 days after surgery. Thereafter, the second, third and fourth collections were performed on the 180th day, the first year and the second year after surgery, respectively. If disease recurrence was found during repeated examinations, blood samples collection would be stopped. Blood were collected in 5 ml heparin anticoagulation tubes and immediately used to extract free mRNA (Geneseed, China). The free mRNA was reverse transcribed using a reverse transcription kit (TAKARA). A quarter of the RT product was mixed with a pre-formulated 2×SYBR Green PCR mix containing UPK2 qRT-PCR primers and then supplemented with water to 20 ul of system (Roche). Amplification was performed to collect the dissolution curve as follows: 94°C for 5 Minutes; 40 cycles at 94°C for 2 minutes, 60°C for 1 minute and 72°C for 2 minutes; last cycle at 72°C for 2 minutes. We also detected GAPDH mRNA as an internal reference and calculated the relative expression of UPK2 by the $2^{-\Delta CT}$ method.

Statistical analysis

This analysis included only 132 samples of patients with lung adenocarcinoma admitted to our hospital. The mean serum UPK2 level of recurrence was calculated. The ROC analysis of lung adenocarcinoma recurrence was assessed using the pROC package in the R language and the area under the curve was calculated. The survival curve is plotted using the ggsurvplot package in the R language.

Result

Exploration of specific co-expression modules associated with lung adenocarcinoma recurrence

To establish a co-expressed gene network associated with postoperative recurrence of lung adenocarcinoma, we used WGCNA to analyze gene expression profile data from patients with recurrent lung adenocarcinoma in the TCGA database. Finally, we screened the transcriptome data of 24 relapsed patients, including 14 males and 10 females. Based on the correlation between each two genes, the gene expression data of these patients with recurrent lung adenocarcinoma were classified into 39 gene modules using unsupervised average linkage hierarchical clustering, and labelled in a heat map with different colors (Figure 1A). Gene modules of different colors contained mutually exclusive co-expressed genes. For some genes that could not be classified into a particular module, we incorporated them into gray modules. WGCNA could analyze the correlation between gene modules and a series of phenotypes, thus this method was used to analyze the correlation between the specific gene modules of patients with postoperative recurrent early lung adenocarcinoma and a series of phenotypes, such as age, gender, survival, recurrence, recurrence type and pathological stage. Without any phenotypic and genetic preferences for module partitioning, we found that the purple module had a significant relationship with survival and recurrence, with correlation coefficients greater than 0.7 (Figure 1B). Therefore, we believed that these genes and their co-expression patterns may be associated with the recurrence of lung adenocarcinoma.

Biological insights from module purple

WGCNA classifies co-expressed genes of all patient samples into specific modules related to a series of traits regulated by the same mechanism. In the previous section, we obtained the purple module most relevant to postoperative recurrence. To verify the relationship between the co-expressed genes contained in the purple module and lung adenocarcinoma, we further constructed a heat map of the gene expression in 24 recurrent tumor tissues and 53 paracancerous tissues. The results showed significant difference in the expression pattern of the purple module between paracancerous tissues and recurrent tumor tissues (Figure 2A). However, we observed that the expression pattern of the purple module gene in patients with recurrent tumors was not as consistent as in paracancerous tissues, which exhibited three expression patterns of light red, light blue, and deep red, suggesting different mechanisms of relapse. Of the 117 genes in the purple module, there were 68 genes with significant differences between recurrent tumor tissues and paracancerous tissues ($|\log_{2}FC| < 0.6$, $FDR < 0.05$) (Figure 2B). These 68 genes were used to map the expression heat maps of the two tissues (recurrent tumor tissues and paracancerous tissues), which showed that the expression patterns of the 68 genes were clearly more uniform than in the heat maps constructed with 117 genes. To further analyze the biological function of the genes in the purple module, the genes of the purple module (117) were then analyzed by GO in the DAVID, and the most significant Go term was "Cytosol" (p value=0.0356) (Figure 2C).

Further clarification of key genes associated with adenocarcinoma recurrence in the purple module

Gene significance (GS) has a high correlation with gene connectivity, which means that nodes with higher connectivity in the co-expression network also play an important role in the process of performing biological functions. Therefore, we also constructed a co-expression network of genes for lung

adenocarcinoma recurrence, and obtained a total of 2840 edge and 879 nodes (power=8) (Figure 3A). We found that there were four genes, UPK2, KLHDC3, GALR2, and TYRP1, in the co-expression network with more nodes linked appearing in the purple module, which was highly correlated with survival and recurrence (Table1). Among these genes, the expression levels of UPK2, KLHDC3 and GALR in tumor tissues were higher than those in paracancerous tissues, and the expression level of TYRP1 in tumor tissues was lower than that in paracancerous tissues. Then, we further analyzed the function of these four genes in lung adenocarcinoma using more clinical data in the Oncomine database. The results showed that the survival outcomes of patients with low expression of UPK2, KLHDC3 and GALR2 were significantly better than those of patients with high expression (Figure 3B, C and D), with P values of $4.9e-05$, 0.009 , and $1.7e-05$, respectively; while patients with high expression of TYRP1 had significantly better prognosis than those with low expression, with P value of $2e-07$ (Figure 3E).

Demographic information and clinical characteristics of patients with surgically treated early-stage lung adenocarcinoma receiving UPK2 plasma free mRNA testing

Recent studies have shown that plasma free mRNA has the potential to act as a tumor marker. Table 2 shows the demographic information and clinical characteristics of 105 patients who meet the study criteria out of 132 patients with early-stage lung adenocarcinoma admitted to our hospital. Of these ADC patients, 58 are male (55%), 47 are female (45%), and the average age of all patients is 58 years (39-83 years), indicating that the patients admitted do not have age or gender bias. The pathological stage of most patients is stage I or stage II (83%), and that of the remaining patients is stage IIIa (17%). After surgery, 43 patients received adjuvant therapy (41%), including 39 patients receiving radiation therapy and 8 patients receiving adjuvant chemotherapy.

Diagnostic performance of UPK2

We began to collect the patient's blood, test the free UPK2 level, and then perform imaging examination from the time of first repeated examination, which was on the 90th day after surgery. If the imaging examination indicated that the patient had relapsed, then he or she would be classified into the relapsed group, and the relative expression level of UPK2 mRNA detected would be recorded. If the patient had no recurrence during the follow-up period, then the mean value of multiple testing would be recorded as the corresponding UPK2 expression level. We found that there were no significant differences in UPK2 between lung adenocarcinoma patients of different ages and genders (Figure 4A and B). Interestingly, for non-relapsed patients, UPK2 was maintained at a lower expression level. The expression level of UPK2 relative to GAPDH in relapsed patients was 0.2763, while the average UPK2 expression level in non-relapsed patients was 0.1623, which was significantly lower than that of relapsed patients ($P < 0.001$; Figure 4C). More interestingly, for the same patient, the level of UPK2 expression at relapse was higher than that when there was no recurrence (Figure 4D). In addition, we plotted the ROC curve and calculated the AUC to determine whether the expression level of UPK2 in plasma could be used to distinguish between relapsed and non-relapsed patients (Figure 4E). The results showed that, when plasma UPK2 expression levels were used alone as diagnostic biomarkers, the AUC was 0.767 with a 95% confidence

interval of 0.675-0.858. Moreover, ADC patients included in the study were also divided into UPK2 high expression group and low expression group, and their survival curves were plotted, respectively. The results indicated that patients with high plasma UPK2 mRNA expression had poorer survival, while those with low plasma UPK2 mRNA expression had a better prognosis (Figure 4F).

Discussion

Lung cancer has become a global health problem due to high morbidity and high mortality (20). Lung adenocarcinoma (ADC) is a histological type of non-small cell lung cancer that is becoming a major component of lung cancer (20-22). Despite significant advances in cancer treatment in recent years, 5-year survival rate of ADC is still not satisfactory (23,24). With the advent of precision medicine concepts, molecular biomarkers and molecular drug targets have become hotspots in cancer research, thus targeted treatment of certain targets has enabled tumor patients to get different degrees of treatment benefits. For example, lung cancer patients with EGFR mutations can benefit from the treatment of TKIs, such as Gefitinib and Erlotinib (25). Other potential biomarkers are mainly oncogene-driven mutations, including ALK translocation and ROS1 gene rearrangement (26,27). Therefore, there is an urgent need to identify and validate clinically relevant and effective prognostic markers for lung ADC to complement existing molecular biomarkers and further guide treatment decisions.

Uroplakin 2 (UPK2) is a highly specific marker of bladder transitional cell carcinoma. As early as 1999, it was found that UPK2 mRNA was detected in blood samples from two patients with metastatic bladder cancer who did not receive chemotherapy and 1/8 of patients with metastatic bladder cancer who received chemotherapy, but it was not detected in 50 patients with non-metastatic bladder cancer or in normal control group, indicating that detection of peripheral blood UPK2 was associated with metastatic spread of bladder cancer cells (28). UPK2 specificity and sensitivity testing may be potential means of detecting bladder cancer metastasis, staging, and monitoring chemotherapy response. Lotan et al. (29) have tested 11 immunohistochemical markers at the primary sites of various micropapillary carcinomas, and found that urinary tract protein (UPs) can also be used as the best marker for identifying urinary mesothelial IMC. Li et al. (30) have found that UPK2 is expressed in 63% of plasma cell samples, significantly higher than UPK3 (6%), and they further indicate that UPK2 is a valuable marker and should be included in immunohistochemical markers to facilitate the differential diagnosis of tumors with plasmacytoid features. Further studies by Matuszewski et al. (31) have found that with the development of bladder cancer, the concentration of UPK2 in the urine is decreased, which further confirms the diagnostic value of UPK2 concentration in plasma and urine for bladder cancer. Based on clinical diagnostic needs, Tian et al. (32) have evaluated the expression of UPK2 by bladder tissue microarray and find that UPK2 is highly specific (100%). It can therefore be used as a marker to identify urothelial lineage tumors and to help distinguish between bladder and prostate cancers, or used in combination with GATA3 as potential markers for metastatic breast cancer.

Hoang et al. have found that the positive rate of UPK2, GATA3 and p40 antibody combined testing was 94.2% (97/103) in invasive urothelial carcinomas, indicating combination of these three antibodies has a

high sensitivity to the differential diagnosis of invasive urothelial carcinoma. But the combination testing of UPK2, GATA3 and p40 is negative in lung adenocarcinoma, colon adenocarcinoma and renal cell carcinoma (33). However, so far, there is still no report on the expression and role of UPK2 in patients with ADC recurrence.

In this study, we first verified the expression level of UPK2, and found that the expression of this gene was significantly increased in patients with ADC recurrence and the prognosis of patients with high expression of UPK2 was poor. These different prognostic trends are consistent with previous studies in different types of cancers, suggesting that the differential expression of UPK2 in patients with ADC recurrence may have clinical implications. Enrichment analysis showed that the function of this gene was mainly related to post-translational modification of proteins. Thereafter, we collected blood samples from 105 patients with ADC, 35 of whom were patients with ADC recurrence, and the other 70 patients with no recurrence. RT-qPCR showed that the expression of UPK2 mRNA in the blood of relapsed patients was significantly higher than that of patients without recurrence, indicating that the difference in UPK2 expression was closely related to the recurrence of ADC and UPK2 was expected to be a biomarker for recurrence of ADC patients. In future studies, we will examine the expression of UPK2 in ADC patients with different genders, stages, and lymph node metastasis, and further clarify the clinicopathological correlation of UPK2 expression in order to provide a promising new strategy for targeted therapy of ADC.

Conclusions

After bioinformatics analysis of the data, we obtained genes related to the prognosis of patients, and then tested the expression level of free mRNA in the blood of patients to verify our prediction. We found that the survival rate of patients with high expression of free upk2 mRNA was significantly poorer than that of patients with low expression of free upk2 mRNA. Therefore, The expression level of free UPK2 mRNA in plasma has the potential as an indicator of postoperative recurrence in patients with early LUAD.

Highlights

1. Free UPK2 mRNA in plasma may be an effective biomarker for monitoring recurrence after early LUAD surgery.
2. Detection of free UPK2 mRNA in plasma is convenient for manipulation.
3. The prognosis of LUAD patients with low UPK2 expression in cancer tissues is better than that of patients with high UPK2 expression. Increased free UPK2 mRNA in plasma means recurrence of LUAD.

Abbreviations

Non-small-cell lung cancer (NSCLC)

Lung adenocarcinoma (LUAD)

Human epididymis protein 4 (HE4)

Gene significance (GS)

Uroplakin 2 (UPK2)

urinary tract protein (UPs)

Declarations

Ethics approval and consent to participate

Not applicable.

Availability of data and materials

The datasets during and/or analysed during the current study available from the corresponding author on reasonable request.

Competing interests

The authors declare that they have no competing interests.

Funding

No funding was received.

AUTHORS CONTRIBUTION

JZ performed all experiments, prepared the figures and wrote the manuscript. QL and BL performed clinic sample collection and preparation. HL and CW performed bioinformatics analysis. CL and HJ provided the laboratory support, discussion, reviewed manuscript and provided financial support.

All authors have read and approved the manuscript

Acknowledgements

Not applicable.

Patient consent for publication to Consent to publication

Consent to publication.

References

1. Zhu Y, Xing P, Wang S, Ma D, Mu Y, Li X, Xu Z, Li J: Evaluation of calculating carboplatin dosage in carboplatin-pemetrexed therapy as the first-line therapy for Chinese patients with advanced lung adenocarcinoma. *THORAC CANCER*.9(3):400-7, 2018.
2. Liu AN, Qu HJ, Yu CY, Sun P: Knockdown of LINC01614 inhibits lung adenocarcinoma cell progression by up-regulating miR-217 and down-regulating FOXP1. *J CELL MOL MED*.22(9):4034-44, 2018.
3. Zhou X, Zhang P, Luo W, Zhang L, Hu R, Sun Y, Jiang H: Ketamine induces apoptosis in lung adenocarcinoma cells by regulating the expression of CD69. *Cancer Med*.7(3):788-95, 2018.
4. Han B, Jin B, Chu T, Niu Y, Dong Y, Xu J, Gu A, Zhong H, Wang H, et al.: Combination of chemotherapy and gefitinib as first-line treatment for patients with advanced lung adenocarcinoma and sensitive EGFR mutations: A randomized controlled trial. *INT J CANCER*.141(6):1249-56, 2017.
5. Zhuang X, Qiao T, Xu G, Yuan S, Zhang Q, Chen X: Combination of nadroparin with radiotherapy results in powerful synergistic antitumor effects in lung adenocarcinoma A549 cells. *ONCOL REP*.36(4):2200-6, 2016.
6. Noronha V, Zanwar S, Joshi A, Patil VM, Mahajan A, Janu A, Agarwal JP, Bhargava P, Kapoor A, et al.: Practice Patterns and Outcomes for Pemetrexed Plus Platinum Doublet as Neoadjuvant Chemotherapy in Adenocarcinomas of Lung: Looking Beyond the Usual Paradigm. *Clin Oncol (R Coll Radiol)*.30(1):23-9, 2018.
7. Marini KD, Croucher DR, McCloy RA, Vaghjiani V, Gonzalez-Rajal A, Hastings JF, Chin V, Szczepny A, Kostyrko K, et al.: Inhibition of activin signaling in lung adenocarcinoma increases the therapeutic index of platinum chemotherapy. *SCI TRANSL MED*.10(451), 2018.
8. Tao H, Meng Q, Li M, Shi L, Tang J, Liu Z: Outcomes of bevacizumab combined with chemotherapy in lung adenocarcinoma-induced malignant pleural effusion. *THORAC CANCER*.9(2):298-304, 2018.
9. Wu K, Zhang X, Li F, Xiao D, Hou Y, Zhu S, Liu D, Ye X, Ye M, et al.: Frequent alterations in cytoskeleton remodelling genes in primary and metastatic lung adenocarcinomas. *NAT COMMUN*.6:10131, 2015.
10. Wang X, Xu W, Zhan P, Xu T, Jin J, Miu Y, Zhou Z, Zhu Q, Wan B, et al.: Overexpression of geranylgeranyl diphosphate synthase contributes to tumour metastasis and correlates with poor prognosis of lung adenocarcinoma. *J CELL MOL MED*.22(4):2177-89, 2018.
11. Zhao S, Guo W, Li J, Yu W, Guo T, Deng W, Gu C: High expression of Y-box-binding protein 1 correlates with poor prognosis and early recurrence in patients with small invasive lung adenocarcinoma. *Onco Targets Ther*.9:2683-92, 2016.
12. Zhao S, Guo W, Li J, Yu W, Guo T, Deng W, Gu C: High expression of Y-box-binding protein 1 correlates with poor prognosis and early recurrence in patients with small invasive lung adenocarcinoma. *Onco Targets Ther*.9:2683-92, 2016.
13. Aramini B, Casali C, Stefani A, Bettelli S, Wagner S, Sangale Z, Hughes E, Lanchbury JS, Maiorana A, et al.: Prediction of distant recurrence in resected stage I and II lung adenocarcinoma. *LUNG*

CANCER.101:82-7, 2016.

14. Krishnamurthy J, Torrice C, Ramsey MR, Kovalev GI, Al-Regaiey K, Su L, Sharpless NE:Ink4a/Arf expression is a biomarker of aging. J CLIN INVEST.114(9):1299-307, 2004.
15. Shaw LM, Vanderstichele H, Knapik-Czajka M, Clark CM, Aisen PS, Petersen RC, Blennow K, Soares H, Simon A, et al.:Cerebrospinal fluid biomarker signature in Alzheimer's disease neuroimaging initiative subjects. ANN NEUROL.65(4):403-13, 2009.
16. Ji X, Takahashi R, Hiura Y, Hirokawa G, Fukushima Y, Iwai N:Plasma miR-208 as a biomarker of myocardial injury. CLIN CHEM.55(11):1944-9, 2009.
17. Plotti F, Capriglione S, Terranova C, Montera R, Aloisi A, Damiani P, Muzii L, Scaletta G, Benedetti-Panici P, et al.:Does HE4 have a role as biomarker in the recurrence of ovarian cancer? Tumour Biol.33(6):2117-23, 2012.
18. Huang Y, Cen J, Wei J, Chen Z, Fang Y, Feng Z, Lu J, Liang Y, Luo J, et al.:Impact of AIB1 expression on the prognosis of upper tract urothelial carcinoma after radical nephroureterectomy. CANCER BIOMARK., 2019.
19. Sgroi DC, Carney E, Zarrella E, Steffel L, Binns SN, Finkelstein DM, Szymonifka J, Bhan AK, Shepherd LE, et al.:Prediction of late disease recurrence and extended adjuvant letrozole benefit by the HOXB13/IL17BR biomarker. J Natl Cancer Inst.105(14):1036-42, 2013.
20. Siegel RL, Miller KD, Jemal A:Cancer Statistics, 2017. CA Cancer J Clin.67(1):7-30, 2017.
21. Youlden DR, Cramb SM, Baade PD:The International Epidemiology of Lung Cancer: geographical distribution and secular trends. J THORAC ONCOL.3(8):819-31, 2008.
22. Rosell R, Bivona TG, Karachaliou N:Genetics and biomarkers in personalisation of lung cancer treatment. LANCET.382(9893):720-31, 2013.
23. Hung JJ, Yeh YC, Jeng WJ, Wu KJ, Huang BS, Wu YC, Chou TY, Hsu WH:Predictive value of the international association for the study of lung cancer/American Thoracic Society/European Respiratory Society classification of lung adenocarcinoma in tumor recurrence and patient survival. J CLIN ONCOL.32(22):2357-64, 2014.
24. Subramaniam S, Thakur RK, Yadav VK, Nanda R, Chowdhury S, Agrawal A:Lung cancer biomarkers: State of the art. J Carcinog.12:3, 2013.
25. Remon J, Moran T, Reguart N, Majem M, Carcereny E, Lianes P:Beyond EGFR TKI in EGFR-mutant non-small cell lung cancer patients: main challenges still to be overcome. CANCER TREAT REV.40(6):723-9, 2014.
26. Soda M, Choi YL, Enomoto M, Takada S, Yamashita Y, Ishikawa S, Fujiwara S, Watanabe H, Kurashina K, et al.:Identification of the transforming EML4-ALK fusion gene in non-small-cell lung cancer. NATURE.448(7153):561-6, 2007.
27. Bergethon K, Shaw AT, Ou SH, Katayama R, Lovly CM, McDonald NT, Massion PP, Siwak-Tapp C, Gonzalez A, et al.:ROS1 rearrangements define a unique molecular class of lung cancers. J CLIN ONCOL.30(8):863-70, 2012.

28. Li SM, Zhang ZT, Chan S, McLenan O, Dixon C, Taneja S, Lepor H, Sun TT, Wu XR:Detection of circulating uroplakin-positive cells in patients with transitional cell carcinoma of the bladder. J Urol.162(3 Pt 1):931-5, 1999.
29. Lotan TL, Ye H, Melamed J, Wu XR, Shih I, Epstein JI:Immunohistochemical panel to identify the primary site of invasive micropapillary carcinoma. AM J SURG PATHOL.33(7):1037-41, 2009.
30. Li W, Liang Y, Deavers MT, Kamat AM, Matin SF, Dinney CP, Czerniak B, Guo CC:Uroplakin II is a more sensitive immunohistochemical marker than uroplakin III in urothelial carcinoma and its variants. AM J CLIN PATHOL.142(6):864-71, 2014.
31. Matuszewski M, Szymanska B, Dlugosz A, Malkiewicz B, Dembowski J, Piwowar A:Preliminary Evaluation of the Diagnostic Usefulness of Uroplakin 2 with an Assessment of the Antioxidant Potential of Patients with Bladder Cancer. BIOMED RES INT.2018:8693297, 2018.
32. Tian W, Guner G, Miyamoto H, Cimino-Mathews A, Gonzalez-Roibon N, Argani P, Li X, Sharma R, Subhawong AP, et al.:Utility of uroplakin II expression as a marker of urothelial carcinoma. HUM PATHOL.46(1):58-64, 2015.
33. Hoang LL, Tacha D, Bremer RE, Haas TS, Cheng L:Uroplakin II (UPII), GATA3, and p40 are Highly Sensitive Markers for the Differential Diagnosis of Invasive Urothelial Carcinoma. Appl Immunohistochem Mol Morphol.23(10):711-6, 2015.

Tables

Table1. UPK2, KLHDC3, GALR2 and TYRP1 highly correlated with survival and recurrence.

gene	Nodes_num power=8	KM probe (JetSet best)	P value
UPK2	15	207862_at	4.9-e06
GALR2	14	211226_at	0.0093
KLHDC3	12	208784_s_at	1.7e-06
TYRP1	11	205694_at	2e-07

	All (N=105)	Recurrence (N=35)	No Recurrence (N=70)
Gender, n(%)			
Male	58(55)	23(66)	35(50)
Female	47(45)	12(34)	35(50)
Median age, years(range)	58(39-83)	59(39-80)	58(42-83)
Mortality, n(%)			
Dead	74(70)	26(74)	48(69)
Alive	23(22)	7(20)	16(23)
Unknown	8(8)	2(6)	6(9)
Smoking status, n(%)			
Former	70(67)	22(63)	48(69)
Active	25(24)	11(31)	14(20)
Never	8(8)	2(6)	6(9)
Unknown	2(2)	0(0)	2(3)
Stage at diagnosis, n(%)			
I	65(62)	21(60)	44(63)
II	22(21)	10(29)	12(17)
IIIA	18(17)	4(11)	14(20)
Treatment, n(%)			
No Adjuvant treatment	62(59)	16(46)	46(66)
Adjuvant treatment	43(41)	19(54)	24(34)
Adjuvant_chemotherapy	8(8)	4(11)	4(6)
Radiation	39(37)	17(49)	22(31)
UPK2 Mean expression, value(range)	0.20(0.05-0.80)	0.28(0.08-0.61)	0.16(0.05-0.80)

Table2. Demographic information and clinical characteristics of 105 patients.

Figures

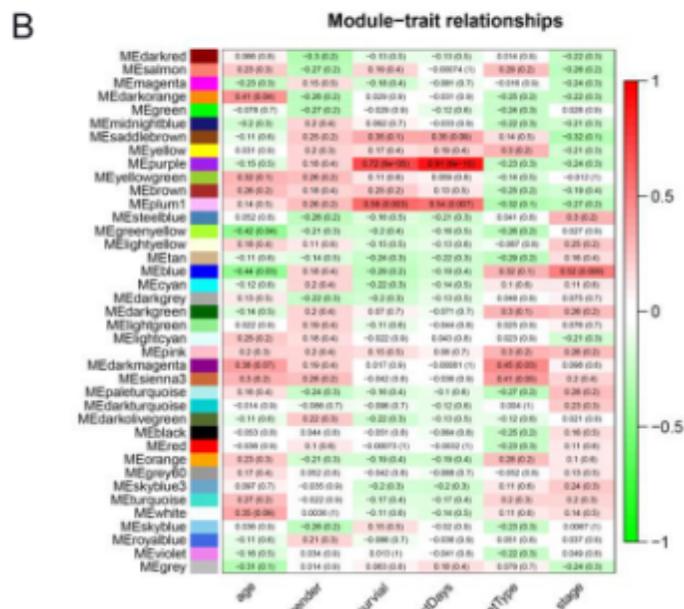
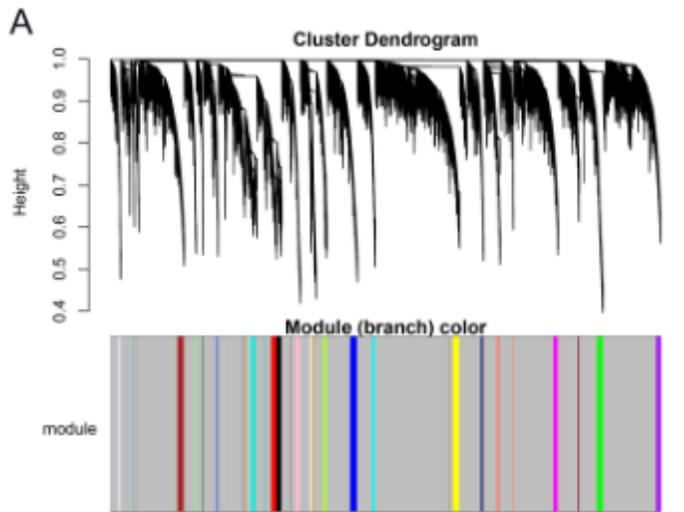


Figure 1

A: Cluster the genes and label different gene modules with specific colors. Gene dendrogram obtained by average linkage hierarchical clustering. B: Heat map of the correlation between gene modules and phenotypes. The correlation weights are labeled in separate boxes, and the negative numbers represent a negative correlation between the gene module and the phenotype. The greater the absolute value of the weight, the stronger the correlation, and the darker the blue or red color of the square. The values in parentheses indicate the P value of the correlation coefficient.

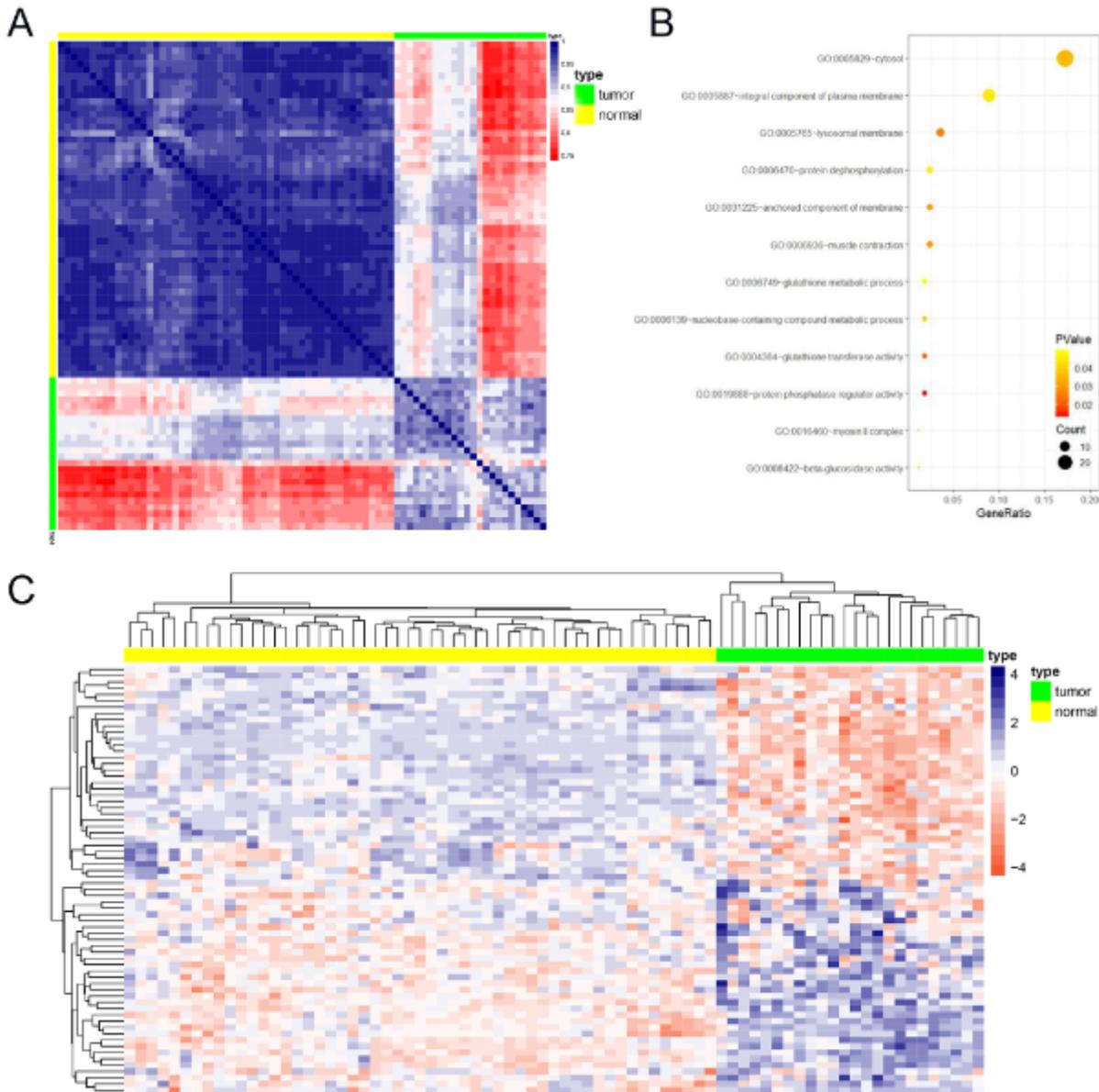


Figure 2

A: Gene expression heat map of the purple module most relevant to recurrence in tumor tissues and paracancerous tissue. The yellow-labeled sample is paracancerous sample and the green-labeled sample is recurrent tissue sample. B: GO enrichment analysis for the 117 genes comprising the purple module identifies multiple biological processes. The original significance outputted from DAVID for GO biological processes are transformed in to "-log (P-value)" for plotting. C: Expression Heat map of differentially expressed genes in the purple module most relevant to recurrence in tumor tissues and paracancerous tissue. The yellow-labeled sample is paracancerous sample and the green-labeled sample is recurrent tissue sample.

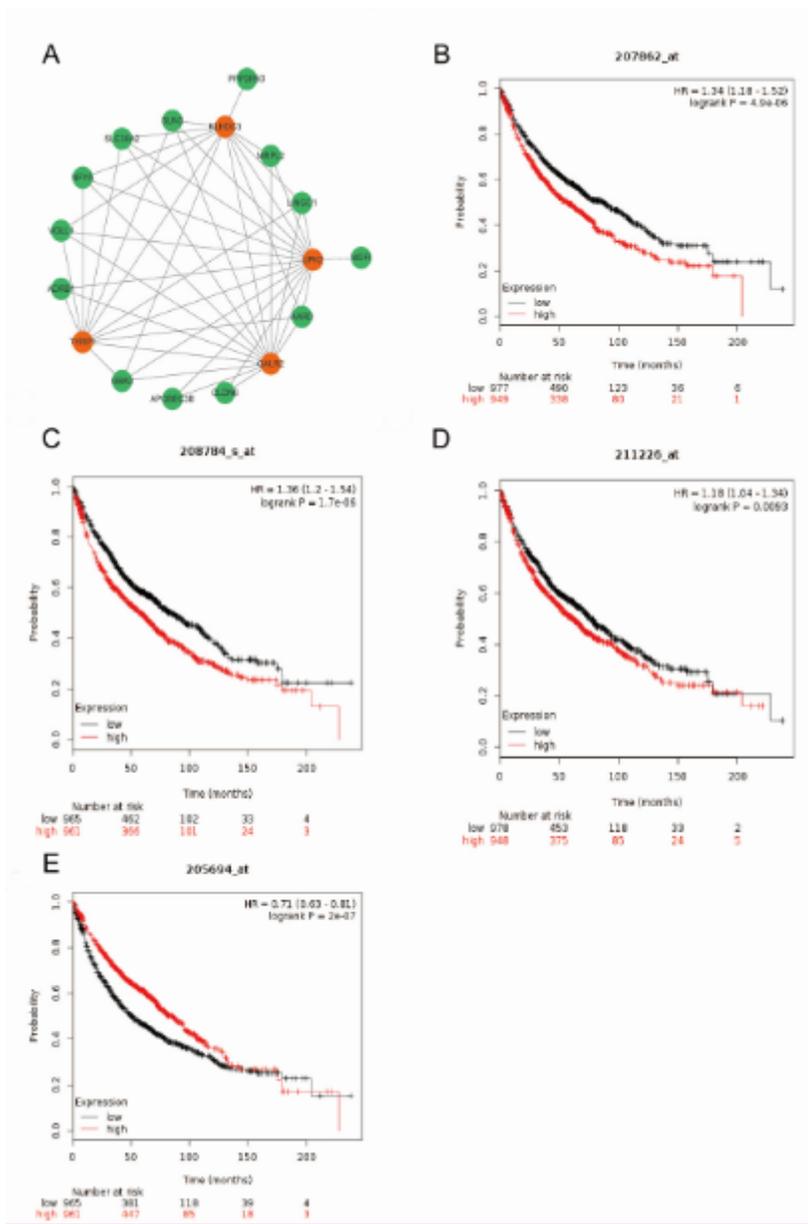


Figure 3

A: Cytoscape that visualizes the co-expression network. The four orange circles show four genes with the highest degree of connectivity in the purple module. B-E: Kaplan-Meier survival plots for OS of UPK2 (B), KLHDC3 (C), GALR2 (D) and TYRP1 (E). The red line in the survival curve represents the survival of patients with high expression, and the black line represents the survival of those with low expression.

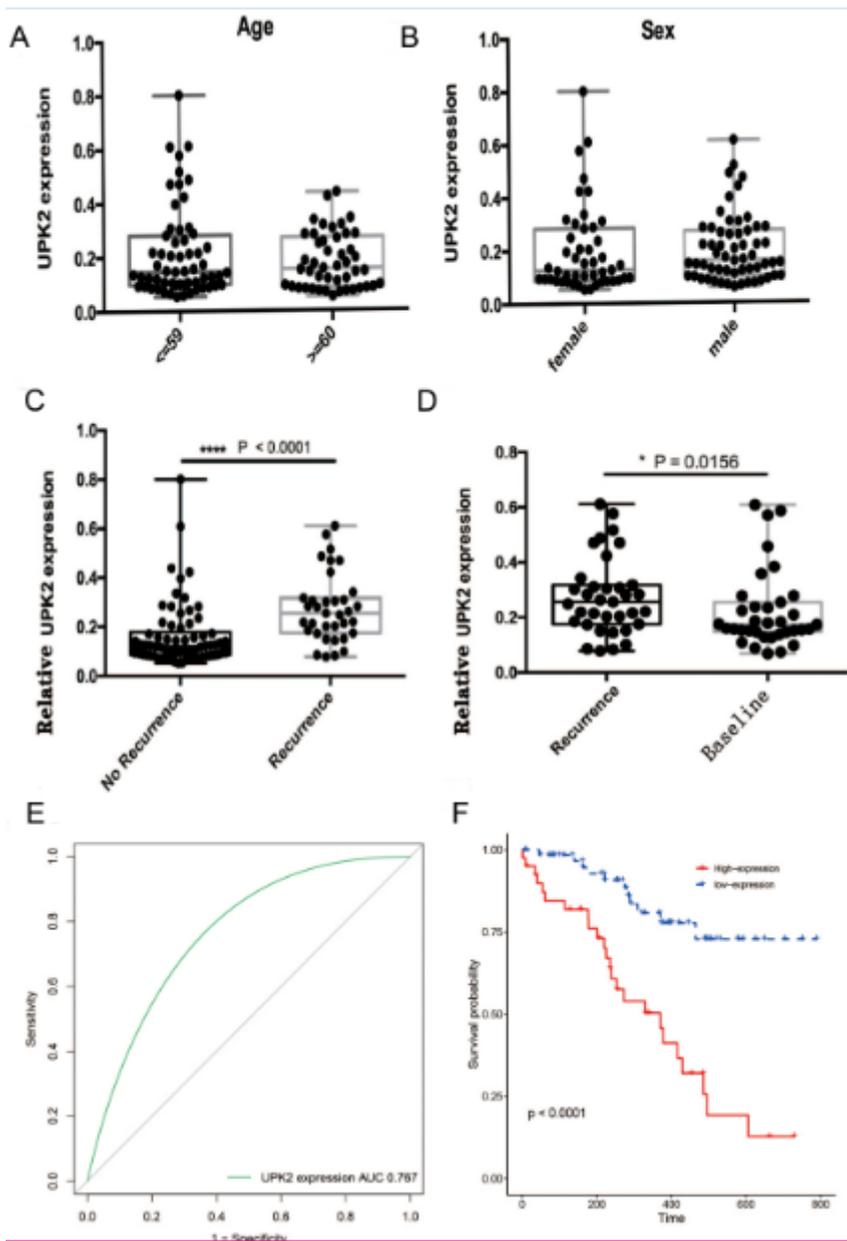


Figure 4

A: Scatter plot of UPK2 expression in lung adenocarcinoma patients of different ages; B: Scatter plot of UPK2 expression in lung adenocarcinoma patients of different genders; C: Scatter plot of UPK2 expression in the relapsed and non-relapsed lung adenocarcinoma patients; D: Scatter plot of UPK2 expression in patients with lung adenocarcinoma before and after relapse; E: Diagnostic performance of plasma cell-free UPK2 mRNA for recurrence detection in LUAD; F: Kaplan-Meier Survival plots for OS of lung adenocarcinoma patients with different UPK2 expression