Comprehensive Analysis of the Expression and Prognosis for IRXs in Non-small Cell Lung Cancer

Qiongzi Wang

  Xinxiang Medical University   https://orcid.org/0000-0002-2335-2346

Xueshan Qiu (✉ xsqiu@cmu.edu.cn)

  Henan Key Laboratory of Immunology and Targeted Drugs

Research article

Abstract

Iroquoishomeobox transcription factor family IRXshave been increasingly reported to play roles in suppressing or promoting a variety of cancers, however, little is known about their expression and prognostic value in terms of human lung cancer. In this study, Oncomine, GEPIA, Kaplan-Meier plotter, and cBioPortal databases were used to analyze the different expression patterns and prognostic values of six IRXs in NSCLC and examine their related functions and pathways using GO enrichment. Compared with normal lung cancer tissues, the expression of IRX1 and IRX2 in NSCLC tissues was significantly lower and was positively correlated with the 10-year survival rate of patients. Higher expression of IRX4 was related to terminal tumors, and suggested a poor prognosis. It was also found that IRXs may play a tumor-suppressive role in the localization of cytoplasm in NSCLC, while localization in the nucleus suggests a more malignant behavior. Together these results suggest that IRX1 and IRX2 may be prognostic indicators of LUAD, and that IRX4 could be a potential target for LUAD treatment.

Introduction

Lung cancer is the second leading cause of death worldwide, as defined by the World Health Organization (WHO)Bray, et al.2018. Global statistics show that lung cancer is still the most common type of cancers, with 11.6% of the total cancer cases diagnosed and 18.4% of the total cancer deaths.Bray, Ferlay, Soerjomataram, Siegel, Torre, Jemal2018. More than half of lung cancer patients died within one year after diagnosis, and the 5-year survival rate was only 17.8%Zappa, Mousa2016. Lung cancer cells originate from respiratory epithelial cells, and the most common type of lung cancer is non-small cell lung cancer (NSCLC), accounting for 85% of lung cancer patients. NSCLC can be divided into three main pathological subtypes: adenocarcinoma (LUAD), squamous cell carcinoma (LUSC), and large cell carcinomaDela Cruz, et al.2011. LUAD is the most common subtype, accounting for 40% of NSCLC cases, followed by LUSC(25-30%) and large cell (undifferentiated) carcinoma (5-10%)Zappa, Mousa2016. In the development of NSCLC, the survival of patients and response to treatment, the discovery and recognition of different biomarkers is of great significance for the development of early detection tools and treatment options.

IRX genes were first identified in the neuro-developmental system of Drosophila melanogaster, and encode highly conserved transcription factors (TFs)Gomez-Skarmeta, Modolell2002. Genes from this family were expressed in the embryonic tissues of invertebrates and vertebrates, and were found to participate in the development of various organs as well as other crucial functions for survival Gomez-Skarmeta, Modolell1996Leyns, et al.1996Netter, et al.1998. In vertebrates, six IRX genes have been identified and divided into two categories, where IRXA (IRX1, IRX2, IRX4) and IRXB (IRX3, IRX5, IRX6), IRX1 and 3, IRX2 and 5, and IRX4 and 6 are respectively homologous genesHouweling, et al.2001McDonald, et al.2010Ogura, et al.2001. In the human genome, IRXA and IRXB are located at 5p15.33 and 16ql1.2-q13, respectively Ogura, Matsumoto, Kuroiwa, Isobe, Otoguro, Jurecic, Baldini, Matsuda, Ogura2001. Recently, it has been reported that IRX genes play important roles in the occurrence and development of a variety of tumors.

Overexpression of IRX1 may lead to pulmonary dysplasia by inducing pulmonary interstitial thickening （Doi, et al.，2011）. Hypomethylation of IRX1 is related to a high risk of lung metastasis of tumors （Lu, et al.，2015）, and IRX1 methylation is negatively correlated with LUSC prognosis （Gao, et al.，2019）. In LUAD, the DNA methylation level of IRX2 was found to decrease, and was negatively correlated with invasion （Sato, et al.，2014）. Whole genome analysis indicated that IRX4 and IRX5 were potential carcinogenic genes for NSCLC, as well as that IRX5 was negatively correlated with the survival rate of lung adenocarcinoma in nonsmokers. Silencing IRX5 can inhibit tumor formation and cause G1 phase cell arrest, as CyclinD1 is its downstream target （D. L. Zhang, et al.，2018）.

Despite this understanding, the function of IRXs in NSCLC and their basic mechanism of promoting or inhibiting cancer have not been fully elucidated. The abnormal expression of IRXs and their relationship with clinicopathological features and prognosis have been reported in human NSCLC, however, as far as we know, bioinformatics has not been used to explore the role of IRXs in NSCLC. RNA and DNA research has become increasingly important to biomedical study as tools to understand pathological processes. Based on the analysis of thousands of published gene expression datasets and variant copy numbers, we analyzed the results of differential IRX factor expression and mutations in NSCLC patients in order to determine IRXs expression patterns, potential functions, and different prognostic values in NSCLC.

Materials And Methods

Oncomine Analysis

The transcriptional levels of IRXs in various tumor tissues were analyzed by tumor datasets in Oncomine database（https://www.Oncomine.org/resource/login.html）, and the expression levels of IRXs mRNA in normal tissues and tumor tissues were statistically analyzed. According to the definition of statistical significance, the cutoffs of P value and fold change were set to 0.05 and 2.0 respectively.

GEPIA Dataset

GEPIA （Tang, et al.，2017）(http://gepia.cancer-pku.cn/) is an interactive web server, it can be used to analyze the whole transcriptome datas of 9736 tumors and 8587 normal samples from TCGA databases and GTEX projects. It provides the differential expression analysis of various genes in tumor tissues relative to normal tissues, different tumor types or pathological stages, and carries out survival analysis of patients combined with clinical information, and can analyze the correlation between genes.

Prognosis Analysis

In order to analyze the overall survival rate of patients with non-small cell lung cancer, patients were divided into two groups according to the median expression (high expression and low expression), and Kaplan Meier survival curve was used for evaluation. Hazard ratio (HR), 95% confidence intervals (CIs) and log rank p values are indicated in the figures. The number of samples was shown at the bottom of the main chart. (http://kmplot.com/analysis/index.php?p=service&cancer=lung).

The Cancer Genome Atlas Data and cBioPortal

All non-small cell lung cancer data sets were screened out in cBioPortal database, and the data from 3854 pathological reports were included. The cBioPorta online tool was used to analyze the IRXs genome map, including mutations, putative copy number alterations (CNAs) for genome identification of important cancer targets, and Z-score of mRNA expression.

Immunohistochemistry

The immunohistochemical data of IRXs were from The Human Protein Atlas (http://www.proteinatlas.org/), which showed the expression and location of IRXs in different lung cancer tissue types. IRX1 antibody (HPA04316), IRX2 antibody (HPA054669), IRX6 antibody (HPA018332).

Results

Expression of IRXs in Different Types of Cancer and Transcriptional Levels in NSCLC

Examining the expression of IRXs in pan-cancer obtained from the TCGA database (Figure 1), IRX1, IRX4, and IRX6 are expressed in certain cancers, while IRX2, IRX3, and IRX5 are generally expressed across most cancers. In NSCLC, the total expression of IRX1 and IRX2 in both LUAD and LUSC was moderate, with no significant difference; in contrast, the expression of IRX3 and IRX5 in LUAD was higher than in LUSC, and the expression of IRX4 and IRX6 in LUSC was significantly higher than that of LUAD (especially IRX4). The transcriptional levels of IRXs in NSCLC compared with those in normal tissues was also analyzed by Oncomine(Figure 2A). IRX1 and IRX2 were significantly under-expressed in NSCLC, and the mRNA expression levels of IRX1 were low in two data sets, expression of IRX2 was also low in three datasets. IRX4 and IRX5 were over-expressed in NSCLC, and the mRNA expression levels of those were high in one datasets. There were no significant results in terms of differential IRX3 and IRX6 expression in NSCLC. Looking further at the various types of datasets statistics for NSCLC⬚Table 1⬚, it can be seen that the expression of IRX1 and IRX2 in LUAD, LUSC, and large cell lung cancer are significantly lower than in normal lung tissue, and that the logFC value is more than double. The expression of IRX3 in two datasets was not statistically significant in either cancer or adjacent cancer in LUAD, but it was lower than in normal lung tissue in LUSC. Interestingly, we found that the expressions of IRX4, IRX5 and IRX6 were more concentrated in one lung cancer subtype⬚IRX4 expression in LUSC was statistically higher, while IRX5 expression was statistically higher in LUAD and the expression of IRX6 in large cell lung cancer was low.

Relationship between mRNA Expression Levels of IRXs and Clinicopathological Factors in NSCLC

The GEPIA database was used to compare the mRNA expression level of IRXs between NSCLC and normal lung tissues (Figure 2B, C). The expression of IRX1 and IRX2 in LUAD and LUSC were found to be lower than that seen in normal tissues. Expression of IRX4 in LUSC was significantly higher than in normal tissues, but almost no expression was seen in LUAD. The expression of IRX3 and IRX5 in LUAD

was higher than normal, but was lower than normal in LUSC. IRX6 expression was lower than normal in LUAD. The expression and significance of IRXs in different stages of NSCLC was also examined, and it was found that IRX1, IRX2, and IRX4 had significant differences in expression in different stages of LUAD, while IRX3 and IRX5 were differentially expressed in different stages of LUSC. The expression of IRX6 in different stages of both LUAD and LUSC was not statistically different.

The immunohistochemical information of IRX1, IRX2, and IRX6 in lung cancer was obtained from the Human Protein Atlas database, but unfortunately, there were no immunohistochemical results available forIRX3, IRX4, and IRX5(Figure 4). From this data, the existing positive expression rate of IRX1, IRX2, and IRX6 in lung cancer tissue was found to be relatively low, positive results for expression were mainly found in the cytoplasm and membrane. Due to the lack of immunohistochemical results for normal pulmonary bronchial epithelium and alveolar epithelium, the expression trend of IRXs in NSCLC tissues could not be fully determined from this analysis.

The Effect of mRNA Expression Levels of IRXs on the Prognosis of NSCLC

The survival rate of IRX mRNA expression in different NSCLC patients was analyzed using the Kaplan-Meier and GEPIA databases (Figure 5). The results showed that the expression of IRX1, IRX2, IRX3, and IRX6 were positively correlated with the prognosis of LUAD patients，while the expression of IRX4 was negatively correlated. Interestingly, survival analysis showed that IRXs expression had a more significant effect on the prognosis of patients with LUAD (whether for high expression of IRX1, IRX2, IRX3 or IRX6, or for low expression of IRX4). In the analysis of LUSC prognosis, only patients with high IRX2 expression had a higher survival rate, which contrasted the prognosis correlation of IRX2 in LUAD.

Analysis of Gene Alteration of IRXs in NSCLC Patients

The genetic alternation of IRXs in NSCLC was analyzed with cBioportal for cancer genomics (Figure 6 and Supplementary Table1). In 3854 samples of 3797 NSCLC patients, 49% of IRXs were altered, among which 15% of IRX1, IRX2, and IRX4 were altered, with less alteration observed for IRX3, IRX5, and IRX6. The alternation of IRX1, IRX2, and IRX4 in NSCLC was mainly gene amplification mutations.

Predicted Functions and Pathways of IRXs in NSCLC

We further analyzed the proteins that may interact with IRXs by using the String database, and constructed an interaction network of the first 30 proteins according to the combined score (Figure 7A and Supplementary Table 2). Following this, the g: Profiler online analysis program was used to enrich the functions of these interacting genes (Figure 7B-D). GO enrichment analysis predicts the function of IRXs in three aspects including biological process, molecular function, and cellular components. In terms of molecular function, IRXs were mainly related to DNA binding, transcriptional activation, and could also participate in Smad binding and BMP (bone morphogenetic protein) receptor binding(Supplementary Table 3). In terms of biological process, IRXs were mainly related to organ development and cell differentiation. Through the analysis of KEGG (Kyoto Encyclopedia of Genes and Genomes) enrichment

of IRX-interacting genes, there were seven pathways related to IRXs found in NSCLC(Figure 7E and Supplementary Table 4), including the TGF-β(transforming growth factor-β) signaling pathway (hsa04350) and Hippo signaling pathway (hsa04390).

Discussion

Various known IRXs have been found in human tumors, and are known to participate in suppressing or promoting cancer in various different tumor backgrounds. However, the roles IRXs play in NSCLC have rarely been reported. This study examined for the first time the expression of IRXs in NSCLC and their relationship with clinicopathological factors and prognosis from the level of transcription and protein expression. The results from this research could provide indications for more accurate diagnosis and prognosis predictions of NSCLC.

At present, IRX1 has been reported to act as a tumor suppressor gene in gastric cancer, bladder cancer, and head and neck squamous cell carcinoma(Bennett, et al.2010Kitchen, et al.2016Liu, et al.2018).It has also been reported to promote cancer in human glioma and osteosarcoma(Lu, Song, Tang, Zou, Han, Zhao, Yong, Yin, Xu, Xie, Kang, Lam, Yang, Shen, Wang2015P. Zhang, et al.2018). In NSCLC, the methylation level of IRX1 is thought to be positively correlated with the prognosis of patients with LUSC (Gao, Zhuang, Zhou, Ma, Zhao, Liu, Liu, Li, Feng, Sun2019), hypomethylation of IRX1 can promote tumor lung metastasis through CXCL14/NF-κB signaling(Lu, Song, Tang, Zou, Han, Zhao, Yong, Yin, Xu, Xie, Kang, Lam, Yang, Shen, Wang2015). In gastric cancer cells tudies, increasing IRX1 expression is able to inhibit peritoneal expansion and lung metastasis through anti-vasculogenic mimicry mechanisms(Jiang, et al.2011). In head and neck squamous cell carcinoma, IRX1 inhibits mitotic activity and promotes apoptosis through the TGF-β pathway(Bennett, et al.2009). In this study, expression of IRX1 mRNA in both LUAD and LUSC was significantly lower than that seen in normal tissues. We also found that patients with high expression of IRX1 had a higher ten years survival rate, and that this increased expression was negatively correlated with tumor stage, though there was no significant correlation between IRX1 and the prognosis or tumor stage of patients with LUSC.

Current research on IRX2 shows that it regulates MMP-2 and MMP-9 by activating the AKT pathway to promote tumor migration and invasion in osteosarcoma(Liu, et al.2015Liu, et al.2014). In soft tissue sarcoma, IRX2 affects development through the Wnt pathway(Adamowicz, et al.2006). Loss of IRX2 expression can also lead to early hematogenous spread in breast cancer(Werner, et al.2015). In NSCLC, it has been reported that methylation of IRX2 CpG islands in LUAD may be related to malignancy (Rauch, et al.2012). In this study it was found that IRX2, similar to IRX1, had low expression in both LUAD and LUSC cancer tissues, it is interesting that when patient's prognoses were analyzed, LUAD patients with high IRX2 expression had higher survival rates, while LUSC patients with similar expression of IRX2 had lower rates of survival. Immunohistochemical information indicated that positive expression of IRX2 in LUSC was localized inside the nucleus, while mainly in cytoplasm in LUAD. It has been reported that IRX1is localized in the cytoplasm in normal brain tissue and low-grade gliomas, but was found in the nucleus in higher grade gliomas (Zhang, Liu, Xu, Wang, Cheng, Jin, Wang, Yang, Liu, Zhang, Tu2018).

IRX5 is able to promote cancer cell proliferation through cyclinD1 in LUAD, and immunohistochemical data shows that it is localized in the nucleus in LUAD（Zhang, Qu, Ma, Zhou, Wang, Zhao, Zhang, Zhang, Wang, Zhang, Yu, Sun, Gao, Cheng, Guo, Huang, Zhou，2018）. From this, we can speculate that the effect of IRXs on the tumor process is related to cell localization, such that expression of IRX2 in the nucleus of LUSC cells determines a poor prognosis.

Studies of IRX3 in nephroblastoma have shown that it can inhibit tumor growth through the classic Wnt/β-catenin pathway（Holmquist Mengelbier, et al.，2019）. There are also reports that IRX3 interacts with the NOTCH pathway and Rho signaling during the formation of renal tumors（Mengelbier, et al.，2010；Scarlett, et al.，2015）. IRX3, as the target of mir-377, can also promote the occurrence of liver cancer（Wang, et al.，2016）, however, there is no related research on the role of IRX3 in lung cancer. Here, the expression of IRX3 in LUAD tumor tissue was found to be higher than that in normal tissue, while expression in tumor tissue of LUSC patients was slightly lower. When analyzing the relationship of expression with tumor stage, the expression of IRX3 in LUSC patients was significantly different between tumor stages. However, prognostic statistics show that IRX3 was positively correlated with the prognosis of LUAD patients, which leads to the opposite conclusion of the expression pattern data. In order to determine why these differences exist, the composition of the IRX3 protein structure was examined (Supplementary Figure1), and it was found that the structure of IRX3 and IRX1 are very similar, both have a homeobox TALE-type domain and a homeobox KN domain in the same location. The difference between the two is that the lenghth of IRX3 protein is 501aa, while IRX1 is 480aa, a leucine to proline change exists in the 422nd amino acid of IRX3 protein, along with a glutamine to histidine change at the 479th. From this, we can speculate that the different roles of IRX1 and IRX3 in NSCLC are not related to their homeobox TALE-type and homeobox KN domains, but should be related to their different N-terminal structures, though the detailed mechanism behind this differential behavior must be studied further

IRX4 has been reported to inhibit the growth of prostate cancer cells by interacting with vitamin D receptors (VDR)（Nguyen, et al.，2012）. Differential gene analysis of NSCLC showed that IRX4 may also act as a tumor-promoting gene in NSCLC（Zhang, Qu, Ma, Zhou, Wang, Zhao, Zhang, Zhang, Wang, Zhang, Yu, Sun, Gao, Cheng, Guo, Huang, Zhou，2018）. Here, IRX4 was scarcely expressed in LUAD patients, but was expressed significantly higher in tumor tissues of LUSC patients. Interestingly, the expression of IRX4 was correlated to a poor prognosis in LUAD, but was not significantly correlated with the prognosis in LUSC. Therefore, although the expression of IRX4 in LUAD patients is low, it may still act as a good prognostic indicator. It is important to note, however, that the above results were all based on mRNA expression, as there was a lack of protein expression information available, because of this, the expression and function of IRX4 in NSCLC still needs further study.

IRX5 can inhibit tumor development via the Hippo and non-classical Wnt pathways in nephroblastoma （Holmquist Mengelbier, Lindell-Munther, Yasui, Jansson, Esfandyari, Karlsson, Lau, Hui, Bexell, Hopyan, Gisselsson，2019）. In prostate cancer, IRX5 plays a cancer-promoting role, and its deletion can cause P21 increase, G2/M phase cell arrest, induce apoptosis, and promote P53 expression（Myrthue, et al.，2008）. IRX5 promotes cell migration through the RhoA/Rock1 pathway in colon cancer（Zhu, et al.，2019）, and

has also been reported to promote hepatocarcinogenesis（Zhu, et al.，2018）. In LUAD, IRX5 can cause G1 phase cell arrest and promote cell proliferation through cyclinD1（Zhang, Qu, Ma, Zhou, Wang, Zhao, Zhang, Zhang, Wang, Zhang, Yu, Sun, Gao, Cheng, Guo, Huang, Zhou，2018）. In this study, the expression of IRX5 in LUAD tumor tissues was found to be higher than that in normal tissues, which is consistent with known reports. Interestingly, IRX5 expression patterns were significantly different from different tumor stages in LUSD patients, with a significant increase seen in stage IV patients, suggesting further that high expression of IRX5 indicates a malignant outcome with LUSC. IRX5 is similar in structure to IRX2, whose role in cancer is mostly related to cell migration and distant spread, so it is reasonable to speculate that high expression of IRX5 in patients with stage IV LUSC may be affected by migration invasion as well.

IRX6 has rarely been studied in tumors, it was reported, however, that IRX6 plays a role as a oncogene in colorectal cancer, and that its expression suggests a poor prognosis（Zuo, et al.，2019）. Here, IRX6 expression in tumor tissues of LUAD patients was found to be lower than that in normal tissues and patients with high expression had better prognoses, suggesting that it plays a tumor-suppressive role in LUAD. The known immunohistochemical data also shows that IRX6 is expressed in the cytoplasm in tumor cells of lung cancer patients, which is consistent with our previous conjectures, the members of the IRX family seem to play a suppressive role in cancer when located in the cytoplasm, but may play a role in promoting cancer when localized to the nucleus.

This study demonstrates important insights into the expression patterns and prognostic analysis of IRX members in NSCLC. We believe that in NSCLC, IRX1 and IRX2 play a tumor suppressive role in LUAD and may become prognostic indicators. IRX4 seems to have a carcinogenic role in LUAD, and may become a potential target for LUAD treatment. Differential localization of IRXs in lung cancer cells was also found to have opposing effects in terms of tumor development, the discovery of this phenomenon still needs more study in order to gain a deeper understanding of the mechanism of IRX action and the mechanism of nuclear access，which would be helpful in terms of understanding how to regulate these proteins to change the malignant biological behavior of tumor cells.

Declarations

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable

Data Availability Statement

The raw data supporting the conclusions of this manuscript will be made available by the authors, without undue reservation, to any qualified researcher.

References

1. Adamowicz M, Radlwimmer B, Rieker RJ, Mertens D, Schwarzbach M, Schraml P, et al. Frequent amplifications and abundant expression of TRIO, NKD2, and IRX2 in soft tissue sarcomas. *Genes, chromosomes & cancer.* 2006;45(9):829-838.

2. Bennett KL, Lee W, Lamarre E, Zhang X, Seth R, Scharpf J, et al. HPV status-independent association of alcohol and tobacco exposure or prior radiation therapy with promoter methylation of FUSSEL18, EBF3, IRX1, and SEPT9, but not SLC5A8, in head and neck squamous cell carcinomas. *Genes, chromosomes & cancer.* 2010;49(4):319-326.

3. Bennett KL, Romigh T, Eng C. Disruption of transforming growth factor-beta signaling by five frequently methylated genes leads to head and neck squamous cell carcinoma pathogenesis. *Cancer research.* 2009;69(24):9301-9305.

4. Bhattacharjee A, Richards WG, Staunton J, Li C, Monti S, Vasa P, et al. Classification of human lung carcinomas by mRNA expression profiling reveals distinct adenocarcinoma subclasses. *Proceedings of the National Academy of Sciences of the United States of America.* 2001;98(24):13790-13795.

5. 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. *CA: a cancer journal for clinicians.* 2018;68(6):394-424.

6. Dela Cruz CS, Tanoue LT, Matthay RA. Lung cancer: epidemiology, etiology, and prevention. *Clinics in chest medicine.* 2011;32(4):605-644.

7. Doi T, Lukosiute A, Ruttenstock E, Dingemann J, Puri P. Expression of Iroquois genes is up-regulated during early lung development in the nitrofen-induced pulmonary hypoplasia. *Journal of pediatric surgery.* 2011;46(1):62-66.

8. Gao C, Zhuang J, Zhou C, Ma K, Zhao M, Liu C, et al. Prognostic value of aberrantly expressed methylation gene profiles in lung squamous cell carcinoma: A study based on The Cancer Genome

Atlas. *Journal of cellular physiology.* 2019;234(5):6519-6528.

9. Garber ME, Troyanskaya OG, Schluens K, Petersen S, Thaesler Z, Pacyna-Gengelbach M, et al. Diversity of gene expression in adenocarcinoma of the lung. *Proceedings of the National Academy of Sciences of the United States of America.* 2001;98(24):13784-13789.

10. Gomez-Skarmeta JL, Modolell J. araucan and caupolican provide a link between compartment subdivisions and patterning of sensory organs and veins in the Drosophila wing. *Genes & development.* 1996;10(22):2935-2945.

11. Gomez-Skarmeta JL, Modolell J. Iroquois genes: genomic organization and function in vertebrate neural development. *Current opinion in genetics & development.* 2002;12(4):403-408.

12. Holmquist Mengelbier L, Lindell-Munther S, Yasui H, Jansson C, Esfandyari J, Karlsson J, et al. The Iroquois homeobox proteins IRX3 and IRX5 have distinct roles in Wilms tumour development and human nephrogenesis. *The Journal of pathology.* 2019;247(1):86-98.

13. Hou J, Aerts J, den Hamer B, van Ijcken W, den Bakker M, Riegman P, et al. Gene expression-based classification of non-small cell lung carcinomas and survival prediction. *PloS one.* 2010;5(4):e10312.

14. Houweling AC, Dildrop R, Peters T, Mummenhoff J, Moorman AF, Ruther U, et al. Gene and cluster-specific expression of the Iroquois family members during mouse development. *Mechanisms of development.* 2001;107(1-2):169-174.

15. Jiang J, Liu W, Guo X, Zhang R, Zhi Q, Ji J, et al. IRX1 influences peritoneal spreading and metastasis via inhibiting BDKRB2-dependent neovascularization on gastric cancer. *Oncogene.* 2011;30(44):4498-4508.

16. Kitchen MO, Bryan RT, Emes RD, Glossop JR, Luscombe C, Cheng KK, et al. Quantitative genome-wide methylation analysis of high-grade non-muscle invasive bladder cancer. *Epigenetics.* 2016;11(3):237-246.

17. Leyns L, Gomez-Skarmeta JL, Dambly-Chaudiere C. iroquois: a prepattern gene that controls the formation of bristles on the thorax of Drosophila. *Mechanisms of development.* 1996;59(1):63-72.

18. Liu T, Zhou W, Cai B, Chu J, Shi G, Teng H, et al. IRX2-mediated upregulation of MMP-9 and VEGF in a PI3K/AKT-dependent manner. *Molecular medicine reports.* 2015;12(3):4346-4351.

19. Liu T, Zhou W, Zhang F, Shi G, Teng H, Xiao J, et al. Knockdown of IRX2 inhibits osteosarcoma cell proliferation and invasion by the AKT/MMP9 signaling pathway. *Molecular medicine reports.* 2014;10(1):169-174.

20. Liu X, Zhang J, Liu L, Jiang Y, Ji J, Yan R, et al. Protein arginine methyltransferase 5-mediated epigenetic silencing of IRX1 contributes to tumorigenicity and metastasis of gastric cancer. *Biochimica et biophysica acta Molecular basis of disease.* 2018;1864(9 Pt B):2835-2844.

21. Lu J, Song G, Tang Q, Zou C, Han F, Zhao Z, et al. IRX1 hypomethylation promotes osteosarcoma metastasis via induction of CXCL14/NF-kappaB signaling. *The Journal of clinical investigation.* 2015;125(5):1839-1856.

22. McDonald LA, Gerrelli D, Fok Y, Hurst LD, Tickle C. Comparison of Iroquois gene expression in limbs/fins of vertebrate embryos. *Journal of anatomy.* 2010;216(6):683-691.

23. Mengelbier LH, Karlsson J, Lindgren D, Ora I, Isaksson M, Frigyesi I, et al. Deletions of 16q in Wilms tumors localize to blastemal-anaplastic cells and are associated with reduced expression of the IRXB renal tubulogenesis gene cluster. *The American journal of pathology.* 2010;177(5):2609-2621.

24. Myrthue A, Rademacher BL, Pittsenbarger J, Kutyba-Brooks B, Gantner M, Qian DZ, et al. The iroquois homeobox gene 5 is regulated by 1,25-dihydroxyvitamin D3 in human prostate cancer and regulates apoptosis and the cell cycle in LNCaP prostate cancer cells. *Clinical cancer research : an official journal of the American Association for Cancer Research.* 2008;14(11):3562-3570.

25. Netter S, Fauvarque MO, Diez del Corral R, Dura JM, Coen D. white+ transgene insertions presenting a dorsal/ventral pattern define a single cluster of homeobox genes that is silenced by the polycomb-group proteins in Drosophila melanogaster. *Genetics.* 1998;149(1):257-275.

26. Nguyen HH, Takata R, Akamatsu S, Shigemizu D, Tsunoda T, Furihata M, et al. IRX4 at 5p15 suppresses prostate cancer growth through the interaction with vitamin D receptor, conferring prostate cancer susceptibility. *Human molecular genetics.* 2012;21(9):2076-2085.

27. Ogura K, Matsumoto K, Kuroiwa A, Isobe T, Otoguro T, Jurecic V, et al. Cloning and chromosome mapping of human and chicken Iroquois (IRX) genes. *Cytogenetics and cell genetics.* 2001;92(3-4):320-325.

28. Okayama H, Kohno T, Ishii Y, Shimada Y, Shiraishi K, Iwakawa R, et al. Identification of genes upregulated in ALK-positive and EGFR/KRAS/ALK-negative lung adenocarcinomas. *Cancer research.* 2012;72(1):100-111.

29. Rauch TA, Wang Z, Wu X, Kernstine KH, Riggs AD, Pfeifer GP. DNA methylation biomarkers for lung cancer. *Tumour biology : the journal of the International Society for Oncodevelopmental Biology and Medicine.* 2012;33(2):287-296.

30. Sato T, Arai E, Kohno T, Takahashi Y, Miyata S, Tsuta K, et al. Epigenetic clustering of lung adenocarcinomas based on DNA methylation profiles in adjacent lung tissue: Its correlation with smoking history and chronic obstructive pulmonary disease. *International journal of cancer.* 2014;135(2):319-334.

31. Scarlett K, Pattabiraman V, Barnett P, Liu D, Anderson LM. The proangiogenic effect of iroquois homeobox transcription factor Irx3 in human microvascular endothelial cells. *The Journal of biological chemistry.* 2015;290(10):6303-6315.

32. Stearman RS, Dwyer-Nield L, Zerbe L, Blaine SA, Chan Z, Bunn PA, Jr., et al. Analysis of orthologous gene expression between human pulmonary adenocarcinoma and a carcinogen-induced murine model. *The American journal of pathology.* 2005;167(6):1763-1775.

33. Su LJ, Chang CW, Wu YC, Chen KC, Lin CJ, Liang SC, et al. Selection of DDX5 as a novel internal control for Q-RT-PCR from microarray data using a block bootstrap re-sampling scheme. *BMC genomics.* 2007;8:140.

34. Tang Z, Li C, Kang B, Gao G, Li C, Zhang Z. GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses. *Nucleic acids research.* 2017;45(W1):W98-W102.

35. Wachi S, Yoneda K, Wu R. Interactome-transcriptome analysis reveals the high centrality of genes differentially expressed in lung cancer tissues. *Bioinformatics.* 2005;21(23):4205-4208.

36. Wang P, Zhuang C, Huang D, Xu K. Downregulation of miR-377 contributes to IRX3 deregulation in hepatocellular carcinoma. *Oncology reports.* 2016;36(1):247-252.

37. Weiss J, Sos ML, Seidel D, Peifer M, Zander T, Heuckmann JM, et al. Frequent and focal FGFR1 amplification associates with therapeutically tractable FGFR1 dependency in squamous cell lung cancer. *Science translational medicine.* 2010;2(62):62ra93.

38. Werner S, Stamm H, Pandjaitan M, Kemming D, Brors B, Pantel K, et al. Iroquois homeobox 2 suppresses cellular motility and chemokine expression in breast cancer cells. *BMC cancer.* 2015;15:896.

39. Yamagata N, Shyr Y, Yanagisawa K, Edgerton M, Dang TP, Gonzalez A, et al. A training-testing approach to the molecular classification of resected non-small cell lung cancer. *Clinical cancer research : an official journal of the American Association for Cancer Research.* 2003;9(13):4695-4704.

40. Zappa C, Mousa SA. Non-small cell lung cancer: current treatment and future advances. *Translational lung cancer research.* 2016;5(3):288-300.

41. Zhang DL, Qu LW, Ma L, Zhou YC, Wang GZ, Zhao XC, et al. Genome-wide identification of transcription factors that are critical to non-small cell lung cancer. *Cancer letters.* 2018;434:132-143.

42. Zhang P, Liu N, Xu X, Wang Z, Cheng Y, Jin W, et al. Clinical significance of Iroquois Homeobox Gene - IRX1 in human glioma. *Molecular medicine reports.* 2018;17(3):4651-4656.

43. Zhu L, Liu Y, Chen Q, Yu G, Chen J, Chen K, et al. Long-Noncoding RNA Colorectal Neoplasia Differentially Expressed Gene as a Potential Target to Upregulate the Expression of IRX5 by miR-136-5P to Promote Oncogenic Properties in Hepatocellular Carcinoma. *Cellular physiology and biochemistry : international journal of experimental cellular physiology, biochemistry, and pharmacology.* 2018;50(6):2229-2248.

44. Zhu Q, Wu Y, Yang M, Wang Z, Zhang H, Jiang X, et al. IRX5 promotes colorectal cancer metastasis by negatively regulating the core components of the RHOA pathway. *Molecular carcinogenesis.* 2019;58(11):2065-2076.

45. Zuo S, Dai G, Ren X. Identification of a 6-gene signature predicting prognosis for colorectal cancer. *Cancer cell international.* 2019;19:6.

Tables

Table.1 The significant changes of *IRXs* expression in transcription level between different types of NSCLC (ONCOMINE database)

	Type of NSCLC vs. Normal Lung Tissue	P-value	Fold Change	t-Test	Source and Reference
IRX1	Lung Adenocarcinoma vs. Normal	2.25E-04	-2.186	-4.357	Garber Lung Proc Natl Acad Sci U S A 2001〚Garber, et al.〛2001〛
	Squamous Cell Lung Carcinoma vs. Normal	0.007	-1.743	-2.826	Garber Lung Proc Natl Acad Sci U S A 2001
	Large Cell Lung Carcinoma vs. Normal	0.01	-2.061	-3.248	Garber Lung Proc Natl Acad Sci U S A 2001
	Lung Adenocarcinoma vs. Normal	1.42E-07	-2.147	-5.696	Hou Lung PLoS One 2010〚Hou, et al.〛2010〛
	Squamous Cell Lung Carcinoma vs. Normal	9.85E-10	-2.706	-7.602	Hou Lung PLoS One 2010
	Large Cell Lung Carcinoma vs. Normal	6.39E-07	-2.819	-6.358	Hou Lung PLoS One 2010
	Lung Adenocarcinoma vs. Normal	1.82E-04	-1.666	-3.821	Okayama Lung Cancer Res 2012〚Okayama, et al.〛2012〛
IRX2	Lung Adenocarcinoma vs. Normal	1.12E-04	-3.367	-5.262	Garber Lung Proc Natl Acad Sci U S A 2001
	Squamous Cell Lung Carcinoma vs. Normal	2.54E-04	-3.17	-4.557	Garber Lung Proc Natl Acad Sci U S A 2001
	Large Cell Lung Carcinoma vs. Normal	6.73E-04	-4.223	-5.151	Garber Lung Proc Natl Acad Sci U S A 2001
	Lung Adenocarcinoma vs. Normal	7.39E-09	-3.518	-6.725	Hou Lung PLoS One 2010
	Squamous Cell Lung Carcinoma vs. Normal	5.50E-10	-3.903	-8.485	Hou Lung PLoS One 2010
	Large Cell Lung Carcinoma vs. Normal	5.03E-06	-7.842	-5.967	Hou Lung PLoS One 2010
	Lung Adenocarcinoma vs. Normal	0.21	-1.386	-0.936	Yamagata Lung Clin Cancer Res 2003〚Yamagata, et al.〛2003〛
	Squamous Cell Lung Carcinoma vs. Normal	0.069	-2.035	-2.063	Yamagata Lung Clin Cancer Res 2003
	Large Cell Lung Carcinoma vs. Normal	0.163	-1.571	-1.101	Yamagata Lung Clin Cancer Res 2003
IRX3	Lung Adenocarcinoma vs. Normal	0.882	1.165	1.208	Garber Lung Proc Natl Acad Sci U S A 2001
	Squamous Cell Lung Carcinoma vs. Normal	0.005	-1.521	-2.926	Garber Lung Proc Natl Acad Sci U S A 2001

	Large Cell Lung Carcinoma vs. Normal	0.025	-1.802	-2.826	Garber Lung Proc Natl Acad Sci U S A 2001
	Lung Adenocarcinoma vs. Normal	0.469	-1.011	-0.077	Hou Lung PLoS One 2010
	Squamous Cell Lung Carcinoma vs. Normal	1.50E-04	-1.837	-3.934	Hou Lung PLoS One 2010
	Large Cell Lung Carcinoma vs. Normal	0.002	-2.223	-3.216	Hou Lung PLoS One 2010
IRX4	Lung Adenocarcinoma vs. Normal	1.62E-14	1.2	9.143	Weiss Lung Statistics Sci Transl Med 2010 [Weiss, et al. 2010]
	Squamous Cell Lung Carcinoma vs. Normal	2.20E-26	1.171	12.363	Weiss Lung Statistics Sci Transl Med 2010
	Squamous Cell Lung Carcinoma vs. Normal	0.009	1.639	3.289	Wachi Lung Bioinformatics 2005 [Wachi, et al. 2005]
	Lung Adenocarcinoma vs. Normal	0.006	1.118	2.598	Hou Lung PLoS One 2010
	Squamous Cell Lung Carcinoma vs. Normal	9.90E-06	1.57	5.163	Hou Lung PLoS One 2010
	Large Cell Lung Carcinoma vs. Normal	0.023	1.14	2.126	Hou Lung PLoS One 2010
IRX5	Lung Adenocarcinoma vs. Normal	2.55E-04	1.743	3.808	Stearman Lung Am J Pathol 2005 [Stearman, et al. 2005]
	Squamous Cell Lung Carcinoma vs. Normal	0.044	1.923	1.748	Bhattacharjee Lung Proc Natl Acad Sci U S A 2001 [Bhattacharjee, et al. 2001]
	Lung Adenocarcinoma vs. Normal	0.004	1.65	2.749	Su Lung BMC Genomics 2007 [Su, et al. 2007]
IRX6	Lung Adenocarcinoma vs. Normal	0.115	-1.407	-1.319	Garber Lung Proc Natl Acad Sci U S A 2001
	Squamous Cell Lung Carcinoma vs. Normal	0.058	-1.684	-1.704	Garber Lung Proc Natl Acad Sci U S A 2001
	Large Cell Lung Carcinoma vs. Normal	0.02	-2.598	-2.558	Garber Lung Proc Natl Acad Sci U S A 2001

1. Garber ME, Troyanskaya OG, Schluens K, Petersen S, Thaesler Z, Pacyna-Gengelbach M, *et al.* Diversity of gene expression in adenocarcinoma of the lung. *Proceedings of the National Academy of Sciences of the United States of America* (2001)98:13784-9.doi: 10.1073/pnas.241500798.
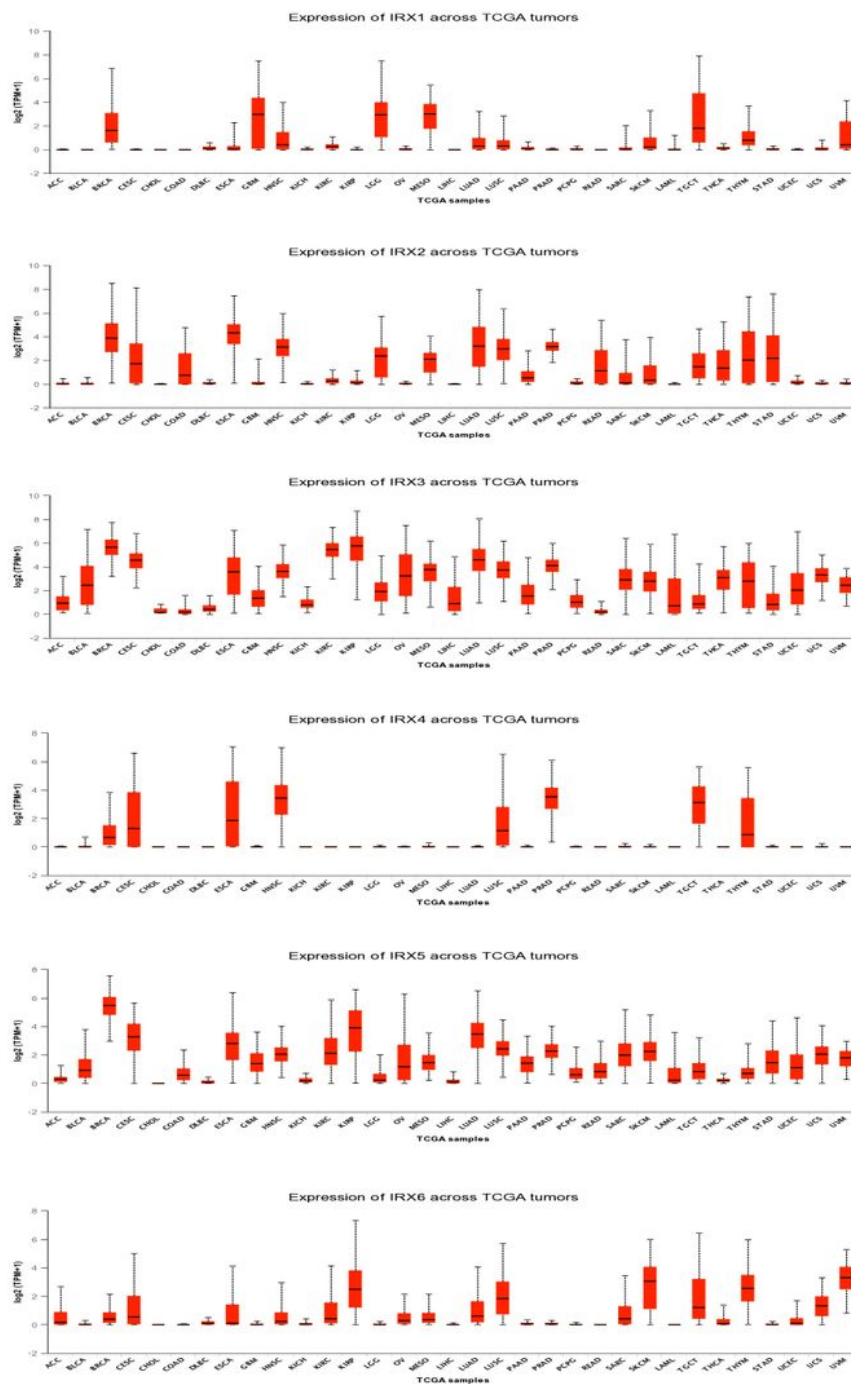
Figures

Figure 1

Analysis of mRNA expression levels of IRXs in different cancers Analysis of mRNA expression level of IRXs in different cancers, data from TCGA database.
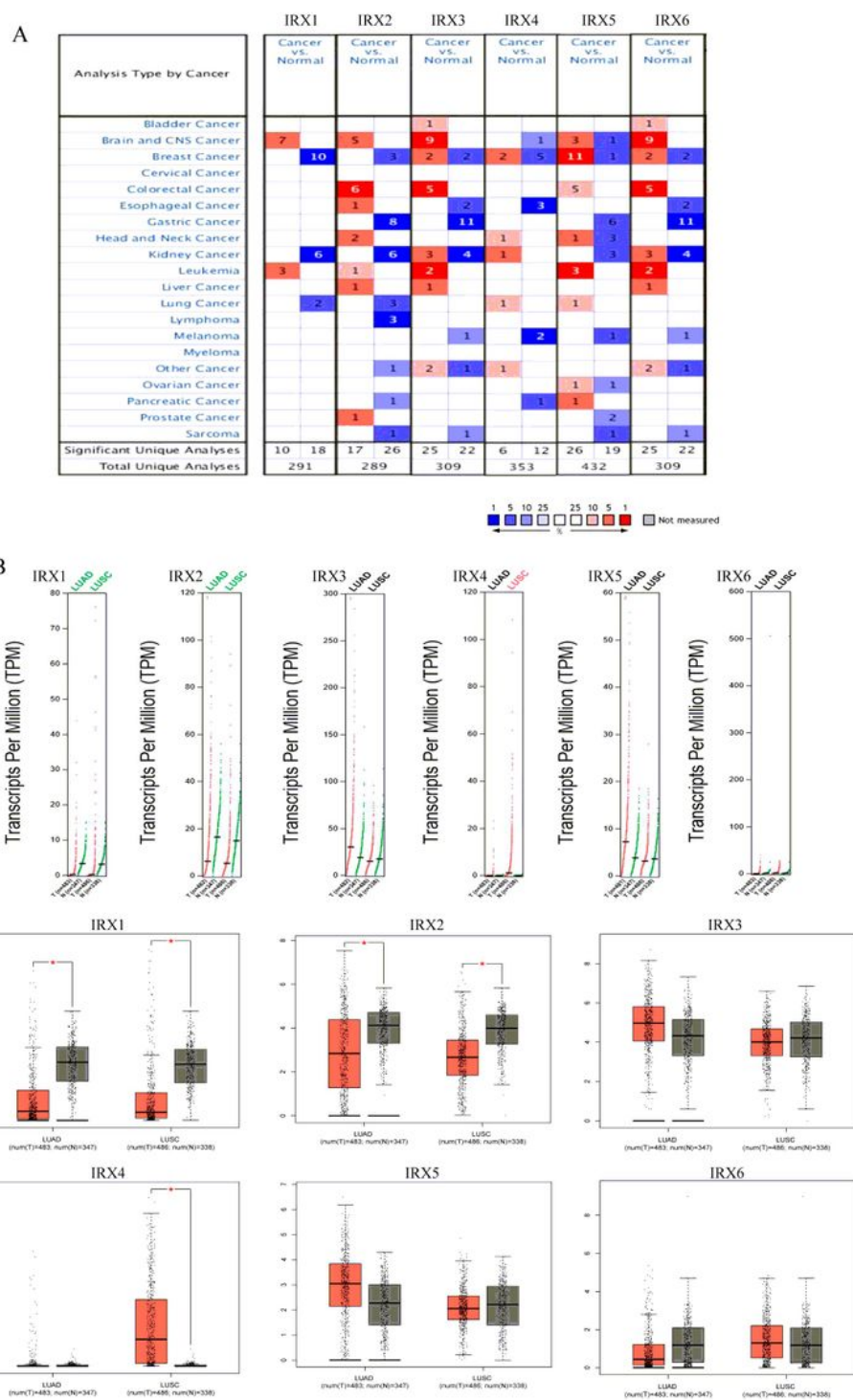
A

Analysis Type by Cancer	IRX1 Cancer vs. Normal		IRX2 Cancer vs. Normal		IRX3 Cancer vs. Normal		IRX4 Cancer vs. Normal		IRX5 Cancer vs. Normal		IRX6 Cancer vs. Normal	
Bladder Cancer					1						1	
Brain and CNS Cancer	7		5		9			1	3	1	9	
Breast Cancer		10		3	2	2	2	5	11	1	2	2
Cervical Cancer												
Colorectal Cancer			6		5				5		5	
Esophageal Cancer			1			2	3					2
Gastric Cancer				8		11				6		11
Head and Neck Cancer			2				1		1	3		
Kidney Cancer		6		6	3	4	1			3	3	4
Leukemia	3		1		2				3		2	
Liver Cancer			1		1						1	
Lung Cancer	2				5		1		1			
Lymphoma				3								
Melanoma						1		2		1		1
Myeloma												
Other Cancer				1	2	1	1		1		2	1
Ovarian Cancer			1						1		1	
Pancreatic Cancer			1					1	1			
Prostate Cancer			1							2		
Sarcoma				1		1				1		1
Significant Unique Analyses	10	18	17	26	25	22	6	12	26	19	25	22
Total Unique Analyses	291		289		309		353		432		309	

1 5 10 25 25 10 5 1 ☐ Not measured

B

C

Figure 2

The Expression of IRXs in NSCLC A. The Transcription Levels of E2F Factors in Different Types of Cancers (Oncomine). B. The expression of IRXs in LUAD and LUSC(scatter diagram). C. The expression of IRXs in LUAD and LUSC (box plot).
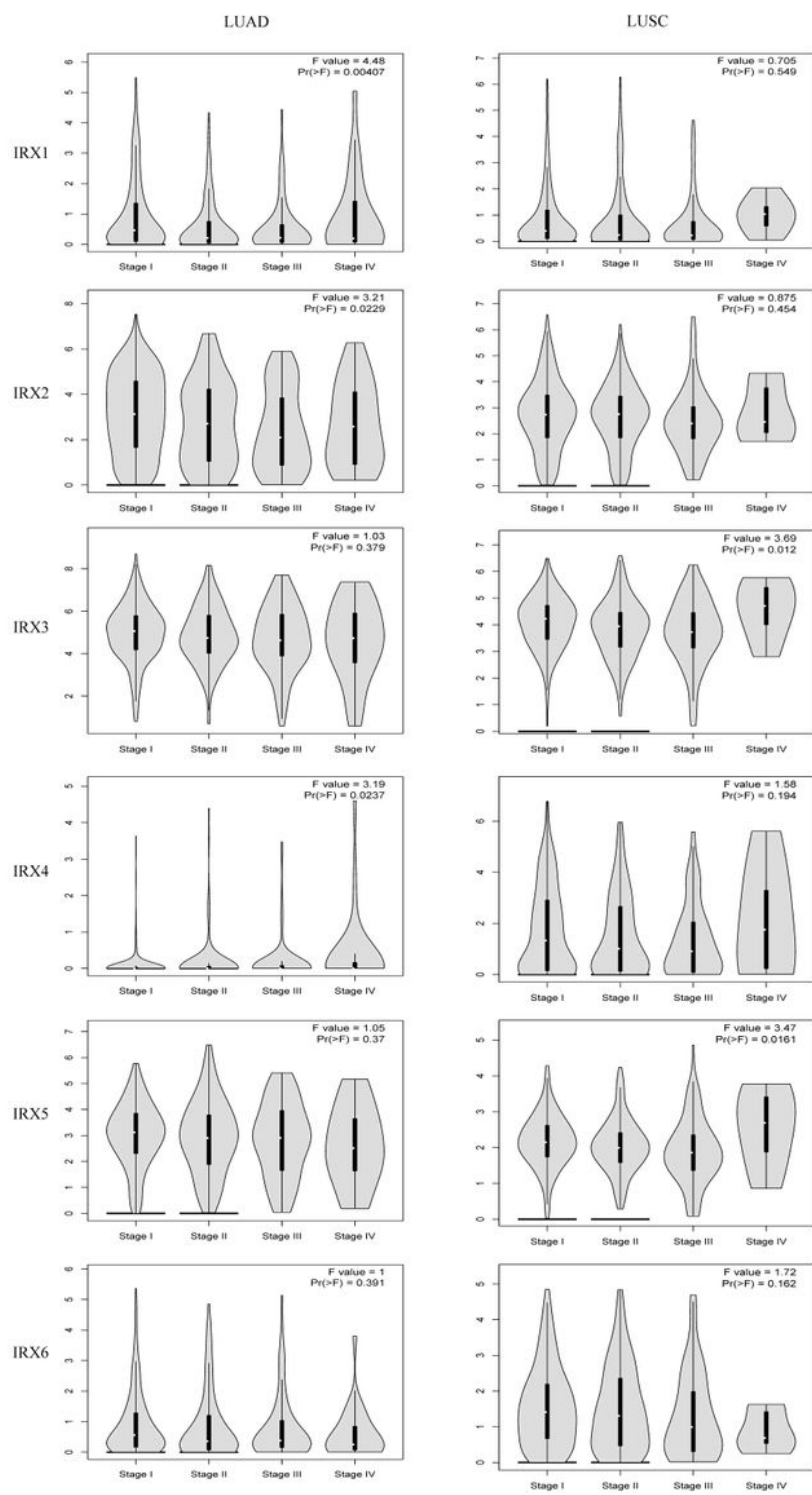
Figure 3

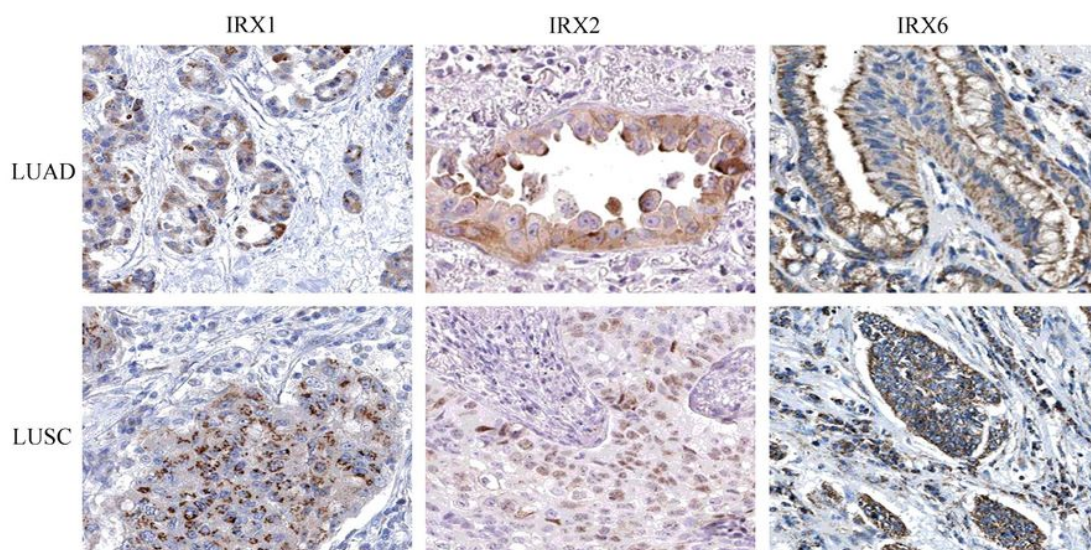Correlation between IRXs Expression and Tumor Stage in NSCLC Patients (GEPIA)
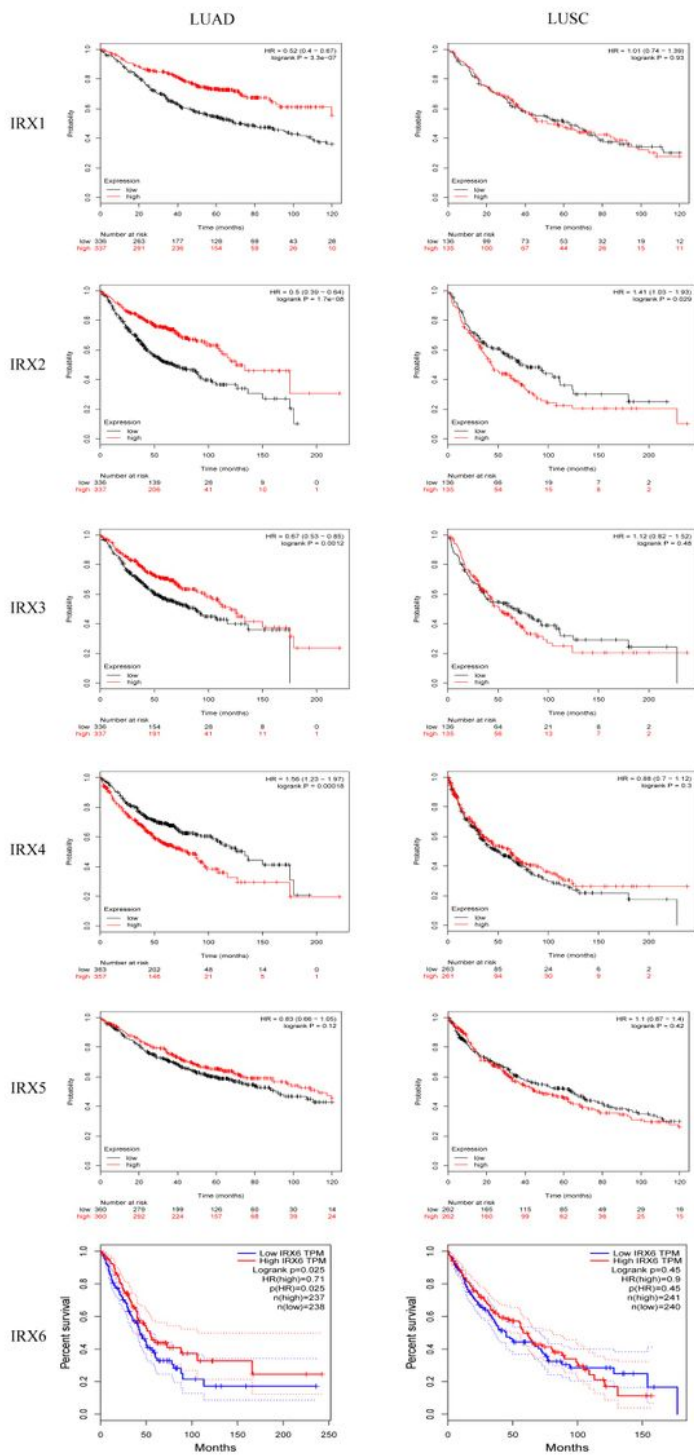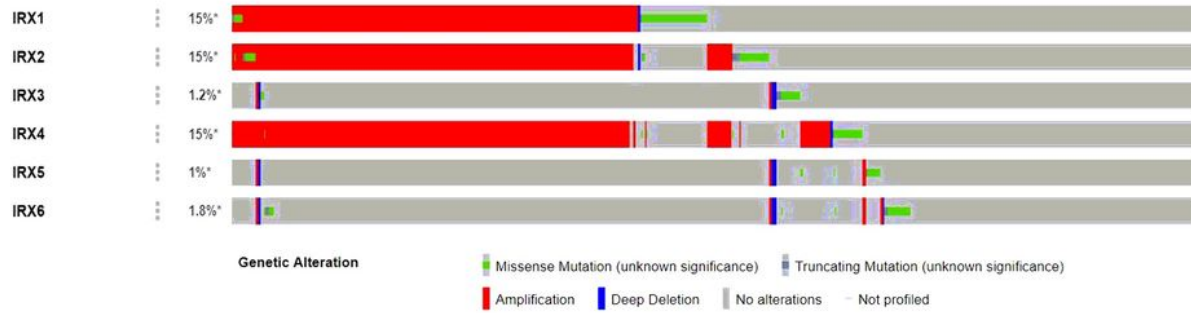
Figure 4

The Expression of IRXs in NSCLC (IHC)

Figure 5

The Prognostic Value of mRNA Level of IRXs in NSCLC Patients (Kaplan-Meier Plotter)
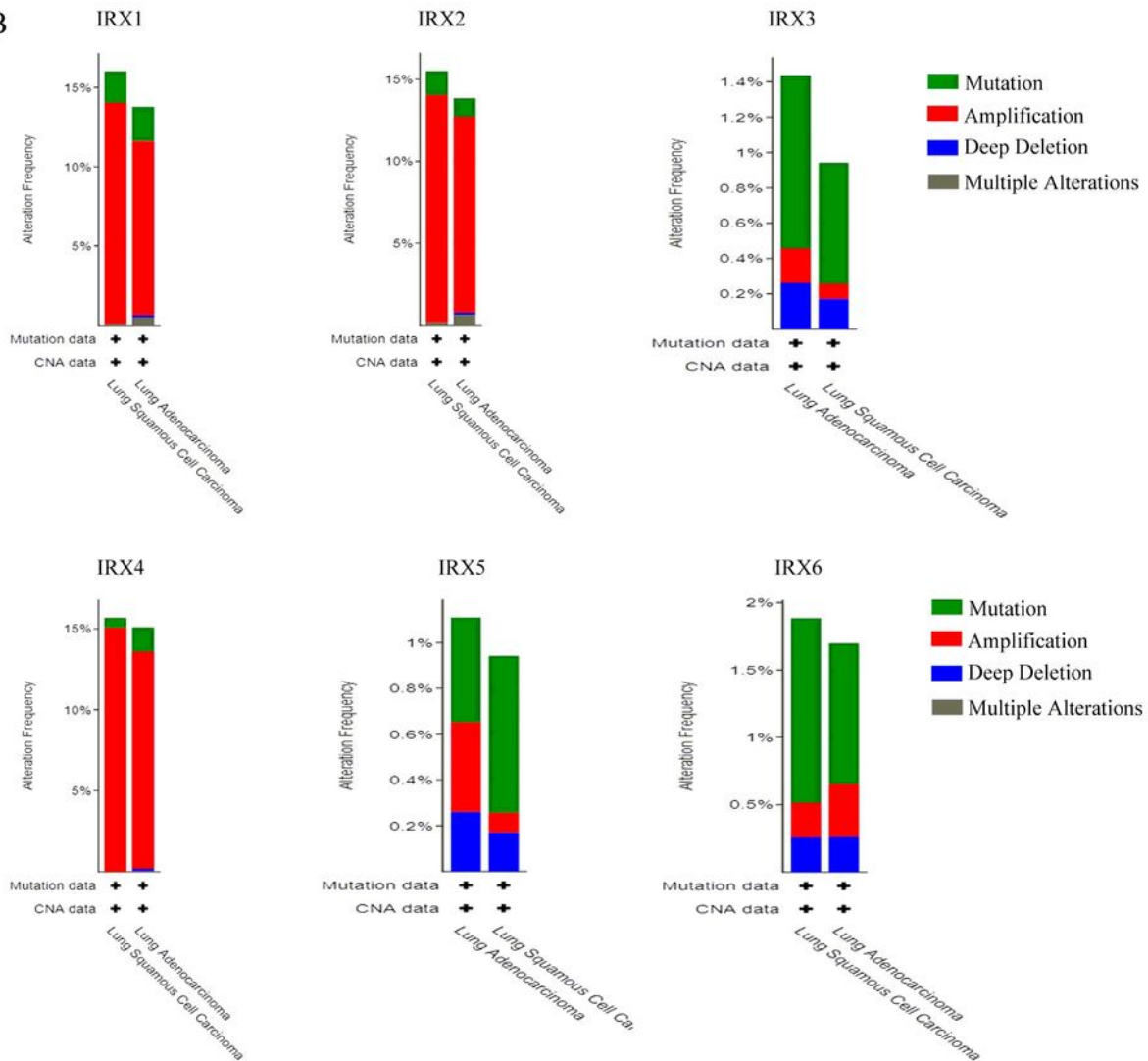
Figure 6

IRXs Gene Expression and Mutation Analysis in NSCLC (cBioPortal) A. IRXs gene expression and mutation analysis in NSCLC (cBioPortal). B. IRXs gene expression and mutation analysis in LUAD and LUSC. (cBioPortal)
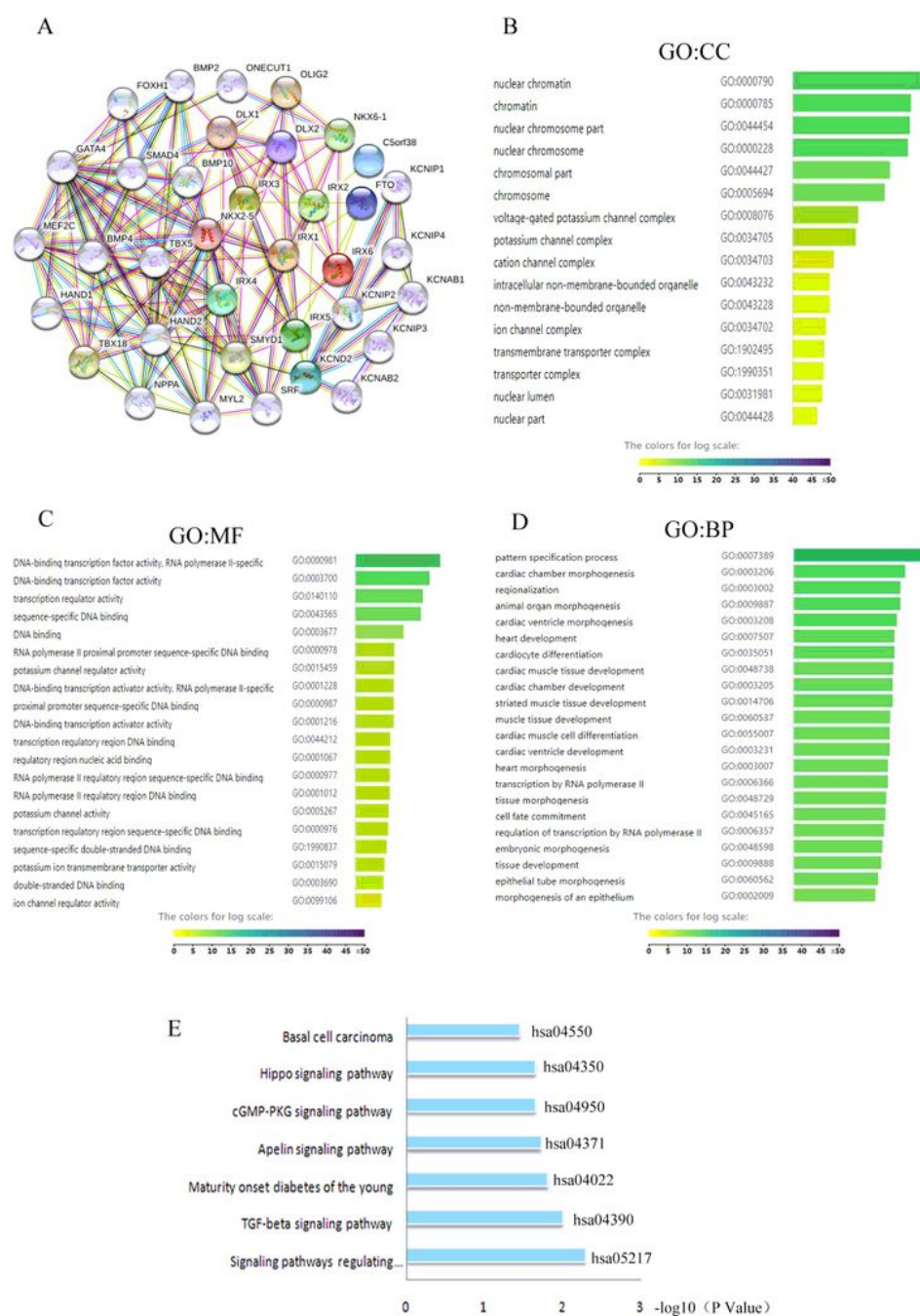
Figure 7

The Functions of IRXs and genes that interact with IRXs A. A network of genes that interact with IRXs(String). The functions of IRXs and genes interact with IRXs were predicted by the analysis of gene ontology (GO) by g:Profiler tools(https://biit.cs.ut.ee/gprofiler/gost) GO enrichment analysis predicted the functional roles of target host genes based on three aspects, including cellular components (B), molecular

functions (C), biological processes(D). E. The functions of IRXs and genes interact with IRXs were predicted by the analysis of Kyoto Encyclopedia of Genes and Genomes (KEGG) by g:Profiler tools.

Supplementary Files

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

- SupplementaryTable4.KEGGanalysis.xlsx
- SupplementaryTable3.FunctionsofIRXsbyGOanalysis.csv
- SupplementaryTable2.GenesinteractwithIRXsfromString.xlsx
- SupplementaryTable1.MutationAnalysisofIRXs.xlsx
- SupplementaryFigure.1.jpg