



Long Noncoding RNA Expression Profiles of Lung Adenocarcinoma Ascertained by Microarray Analysis

Gang Xu¹*, Jie Chen³*, Qinshi Pan¹, Keta Huang², Jingye Pan³, Wenhui Zhang¹, Jian Chen¹, Fangyou Yu¹, Tieli Zhou¹, Yumin Wang¹*

1 Department of Laboratory Medicine, the First Affiliated Hospital of Wenzhou Medical University, Wenzhou, China, **2** Department of Pathology, the First Affiliated Hospital of Wenzhou Medical University, Wenzhou, China, **3** Department of Intensive Care Unit, the First Affiliated Hospital of Wenzhou Medical University, Wenzhou, China

Abstract

Background: Long noncoding RNAs (lncRNAs) have been shown to be involved in the development and progression of lung cancer. However, the roles of lncRNAs in lung cancer are not well understood.

Methodology/Principal Findings: We used a high-throughput microarray to compare the lncRNA and messenger RNA (mRNA) expression profiles in lung adenocarcinoma and normal tissue (NT) samples. Several candidate adenocarcinoma-associated lncRNAs were verified by real-time quantitative reverse transcription polymerase chain reaction (PCR) analysis. Using abundant and varied probes, we were able to assess 30,586 lncRNAs and 26,109 mRNAs in our microarray. We found that 2,420 lncRNAs and 1,109 mRNAs were differentially expressed (≥ 2 -fold change) in lung adenocarcinoma samples and NT, indicating that many lncRNAs were significantly upregulated or downregulated in lung adenocarcinoma. We also found, via quantitative PCR, that 19 lncRNAs were aberrantly expressed in lung adenocarcinoma compared with matched histologically normal lung tissues. Among these, LOC100132354 and RPLPOP2 were the most aberrantly expressed lncRNAs, as estimated by quantitative PCR in 100 pairs of lung adenocarcinoma and NT samples.

Conclusions/Significance: Our study ascertained the expression patterns of lncRNAs in lung adenocarcinoma by microarray. The results revealed that many lncRNAs were differentially expressed in lung adenocarcinoma tissues and NT, suggesting that they may play a key role in tumor development.

Citation: Xu G, Chen J, Pan Q, Huang K, Pan J, et al. (2014) Long Noncoding RNA Expression Profiles of Lung Adenocarcinoma Ascertained by Microarray Analysis. PLoS ONE 9(8): e104044. doi:10.1371/journal.pone.0104044

Editor: Muy-Teck Teh, Barts & The London School of Medicine and Dentistry, Queen Mary University of London, United Kingdom

Received: March 7, 2014; **Accepted:** July 7, 2014; **Published:** August 4, 2014

Copyright: © 2014 Xu et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This work was supported by the National Natural Science Foundation of China (81171614, 81271906), Wenzhou Municipal Science and Technology Bureau of China (Y20110041, Y20130170). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* Email: wym0577@163.com

† These authors contributed equally to this work.

Introduction

Lung cancer has the highest mortality rate of all cancers, and its incidence is gradually growing [1]. Non-small cell lung cancer (NSCLC) is the most common type of lung cancer (accounting for 80% of all cases) and includes squamous cell carcinoma, adenocarcinoma, and large cell carcinoma. Even though surgical resection, radiation therapy, and chemotherapy technologies are continuously improving, patients with lung cancer remain exceedingly vulnerable to relapse and death [2]. The cure rate for lung cancer is low and the average 5-year survival is less than 15% [3–6]. In recent years, a growing proportion of lung adenocarcinomas have been diagnosed as NSCLC that are attributable to causes such as environmental pollution. However, the mechanisms underlying lung cancer development have not been elucidated; thus, the study of lung cancer remains extremely important.

Long noncoding RNAs (lncRNAs) are transcript RNA molecules longer than 200 nucleotides that do not encode a protein and

reside in the nucleus or cytoplasm [7]. lncRNAs are usually divided into five categories: sense, antisense, bidirectional, intronic, and intergenic. In recent years, a large number of lncRNAs have been identified, prompting the creation of the human lncRNA database, which provides lncRNA expression profiles and other important information [8]. The abnormal expression of lncRNAs has been implicated in a range of diseases, including cancer [9].

Studies have shown that lncRNAs known to be differentially expressed in normal cells and tumor cells are important for the regulation of gene expression; therefore, aberrant expression of lncRNAs can lead to abnormalities of gene expression and tumorigenesis. The altered expression of lncRNAs is a feature of many types of cancers and has been shown to promote the development, invasion, and metastasis of tumors by a variety of mechanisms [9,10]. lncRNAs regulate expression on the epigenetic, transcriptional, and post-transcriptional levels [11–13].

lncRNAs have been shown to be involved in the development and progression of lung cancer; however, research concerning

lncRNA involvement in lung cancer is in its infancy. Known lung cancer-associated lncRNAs are few and include HOTAIR, H19, ANRIL, MALAT1 (lung adenocarcinoma associated transcript 1) [14,15], SCAL1 (smoke and cancer-related long-chain non-coding RNA1) [16], lncRNA AK126698 [17], and lncRNA GAS6-AS1 (GAS6 antisense RNA1) [18]. The identification of additional lung cancer-associated lncRNAs and their mechanism of action require further investigation.

In this study, we analyzed the expression patterns of lncRNAs and mRNAs in lung adenocarcinoma samples and compared them with the corresponding patterns in adjacent nontumorous tissue (NT) samples. Several of the differentially expressed lncRNAs were evaluated by SYBR RT-PCR in 100 pairs of tissue samples. Our results suggest that lncRNA expression patterns may provide new molecular biomarkers for the diagnosis of lung adenocarcinoma.

Materials and Methods

Patient samples

The lung adenocarcinoma samples and corresponding NT samples were prospectively collected from 105 patients of the First Affiliated Hospital of Wenzhou Medical University, China, from April 2012 to August 2013. Of these patients, five (the basic medical records see Table S1 in File S1) were used for microarray analysis of lncRNAs and 100 were used for additional evaluations (Table S2 in File S1). The diagnosis of adenocarcinoma was confirmed by histopathology. The lung adenocarcinoma and matched NT samples were snap-frozen in liquid nitrogen immediately after resection. This study was approved by the Institutional Ethics Review Board of the First Affiliated Hospital of Wenzhou Medical University, and all patients provided written informed consent for this study.

RNA extraction

Lung adenocarcinoma cells were obtained by laser microdissection; the proportion of cancer cells in the tissue sections was 100%. We combined the five lung adenocarcinoma samples and the five corresponding NT normal samples, and the two grouped samples were subjected to RNA extraction. Total RNA was extracted using Trizol reagent (Invitrogen, Carlsbad, CA, USA), according to the manufacturer's protocol. The integrity of the RNA was assessed by electrophoresis on a denaturing agarose gel. A NanoDrop ND-1000 spectrophotometer was used for the accurate measurement of RNA concentration (OD_{260}), protein contamination (OD_{260}/OD_{280} ratio), and organic compound contamination (OD_{260}/OD_{230} ratio).

Microarray and computational analysis

For microarray analysis, an Agilent Array platform (Agilent Technologies, Santa Clara, CA, USA) was employed. Sample preparation and microarray hybridization were performed based on the manufacturer's standard protocols with minor modifications. Briefly, mRNA was purified from total RNA after removal of rRNA by using an mRNA-ONLY Eukaryotic mRNA Isolation Kit (Epicentre Biotechnologies, USA). Then, each sample was amplified and transcribed into fluorescent cRNA along the entire length of the transcripts without 3' bias by using a random priming method. The labeled cRNAs were hybridized onto a Human lncRNA Array v3.0 (8×60 K; Arraystar), which was designed for 30,586 lncRNAs and 26,109 coding transcripts. The lncRNAs were carefully constructed using the most highly respected public transcriptome databases (RefSeq, UCSC Known Genes, GENCODE, etc.) as well as landmark publications. Each transcript was accurately identified by a specific exon or splice junction probe. Positive probes for housekeeping genes and negative probes were also printed onto the array for hybridization quality control. After

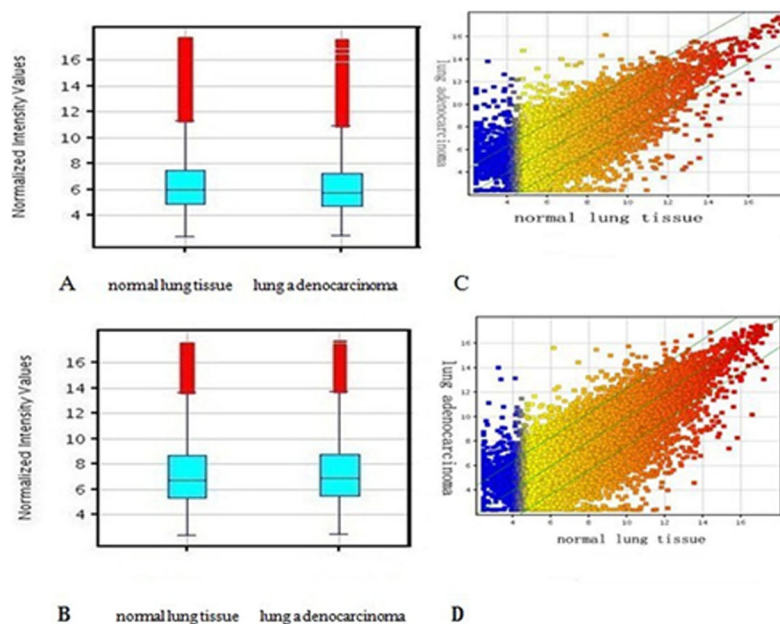


Figure 1. Box plots and Scatter plots showing the variation in lncRNA and mRNA expression between the lung adenocarcinoma and normal lung tissue arrays. (A–B) Box plots showing the distribution of the lncRNA (A) and mRNA (B) profiles. After normalization, the distributions of the \log_2 -ratios among the tested samples were nearly the same. (C–D) Scatter plots showing the variation in lncRNA (C) and mRNA (D) expression between the lung adenocarcinoma and normal lung tissue arrays. The values of the X and Y axes in the scatter plot are averaged normalized values in each group (\log_2 -scaled). The lncRNAs above the top green line and below the bottom green line are those with a >3 -fold change in expression between tissues. doi:10.1371/journal.pone.0104044.g001

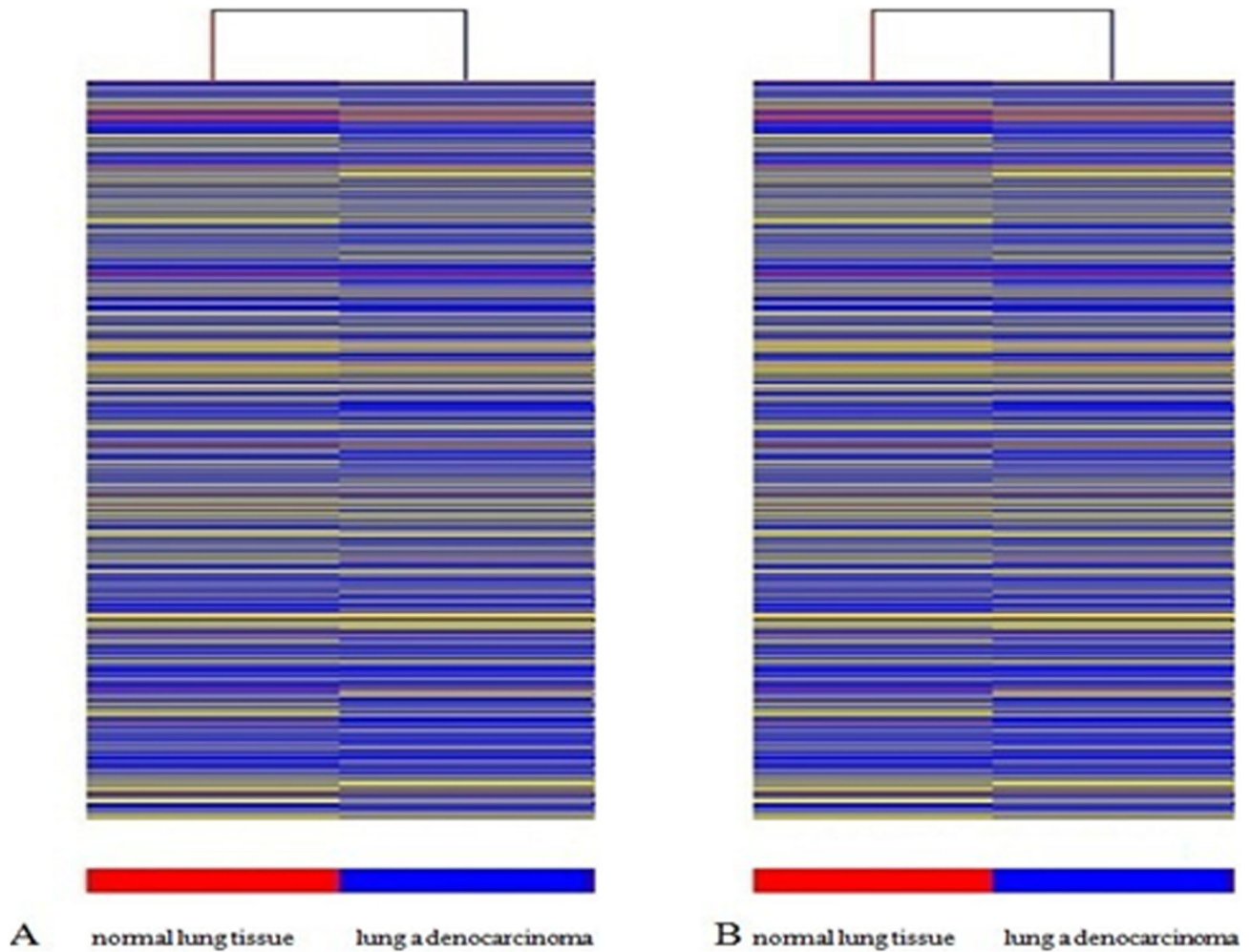


Figure 2. Heat map and hierarchical clustering of lncRNA and mRNA profile comparison between the lung adenocarcinoma and normal lung samples. (A) lncRNA (B) mRNA.
doi:10.1371/journal.pone.0104044.g002

washing the slides, the arrays were scanned using an Agilent G2505C scanner, and the acquired array images were analyzed with Agilent Feature Extraction software (version 11.0.1.1). Quantile normalization and subsequent data processing were performed using the GeneSpring GX v12.0 software package

(Agilent Technologies). The microarray work was performed by KangChen Bio-tech, Shanghai, People’s Republic of China.

Functional group analysis

GO analysis was derived from Gene Ontology (www.geneontology.org), which provides three structured networks of

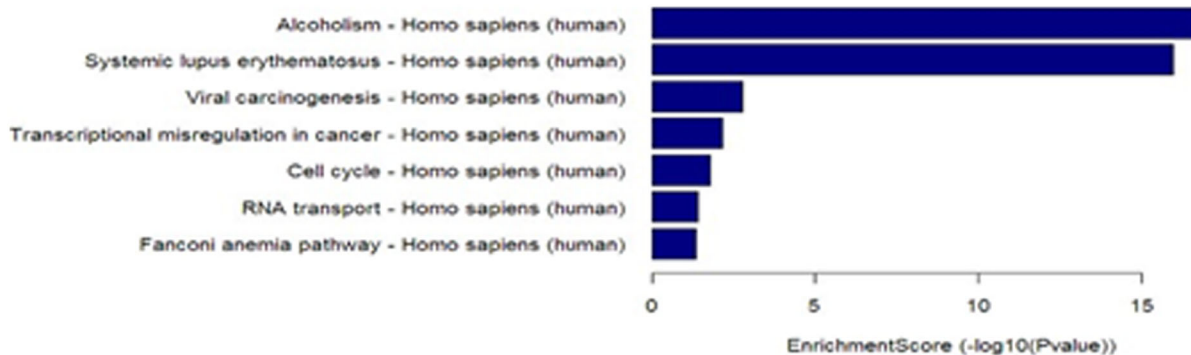


Figure 3. Pathway analysis of upregulated mRNAs in lung adenocarcinoma. Seven upregulated pathways were identified, including ethanol metabolism, viral carcinogenesis, RNA transduction, and cell cycle pathways.
doi:10.1371/journal.pone.0104044.g003

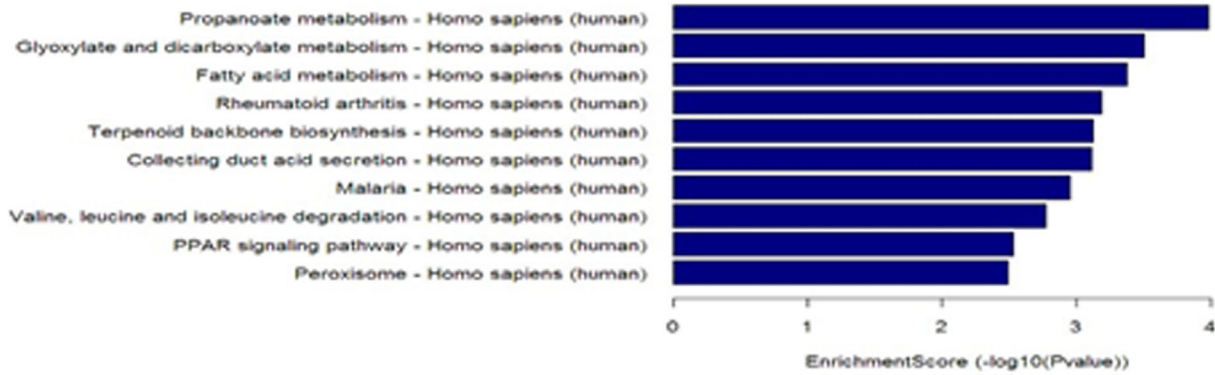


Figure 4. Pathway analysis of downregulated mRNAs in lung adenocarcinoma. Twenty-four downregulated pathways were identified, including propanoate metabolism and fatty acid metabolism pathways. doi:10.1371/journal.pone.0104044.g004

defined terms that describe gene product attributes. The P-value denotes the significance of GO Term enrichment in the differentially expressed mRNA list ($P \leq 0.05$ was considered statistically significant). We also performed pathway analysis for the differentially expressed mRNAs based on the latest KEGG (Kyoto Encyclopedia of Genes and Genomes) database. This analysis allowed us to determine the biological pathways for which a significant enrichment of differentially expressed mRNAs existed ($P \leq 0.05$ was considered statistically significant).

Construction of the Coding-non-coding Gene Coexpression Network

The weighted coexpression network is constructed by calculating a pairwise correlation matrix between all probe sets across microarray samples. The resulting Pearson correlation matrix was transformed into an adjacency matrix. Microarray data can be noisy, and the number of samples is often small, so we weighed the Pearson.

correlations by taking their absolute value and raising them to the power. The nodes of coexpression network correspond to gene expressions, and edges between genes are determined by the correlations between gene expression Pearson correlation coeffi-

cients (Pearson correlation coefficient, PCC) was calculated between the coding and non-coding/coding and coding with the use of R statistical analysis. $PCC > 0.98$ was considered statistically significant in both cases. The CNC (coding-non-coding)/CC (coding-coding) network was drawn with the Cytoscape (v2.8.1) software, which green nodes represent coding gene (mRNA), red nodes represent non-coding (lncRNA). The solid line represents a positive correlation between the two nodes, the dotted line represents a negative correlation.

Quantitative PCR

Total RNA was extracted from frozen lung adenocarcinoma tissues by using TRIzol reagent (Invitrogen) and then reverse-transcribed using an RT Reagent Kit (Thermo Scientific), according to the manufacturer’s instructions. LncRNA expression in lung adenocarcinoma tissues was measured by quantitative PCR by using SYBR Premix Ex Taq and an ABI 7000 instrument. Some candidate lncRNAs were validated by SYBRP PCR, these gene primers in the study for Q-PCR see Table S3 in File S1. Among these, two lncRNAs that were significantly differentially expressed in adenocarcinoma and normal tissues (LOC100132354, RPLP0P2) were evaluated in all of the patients

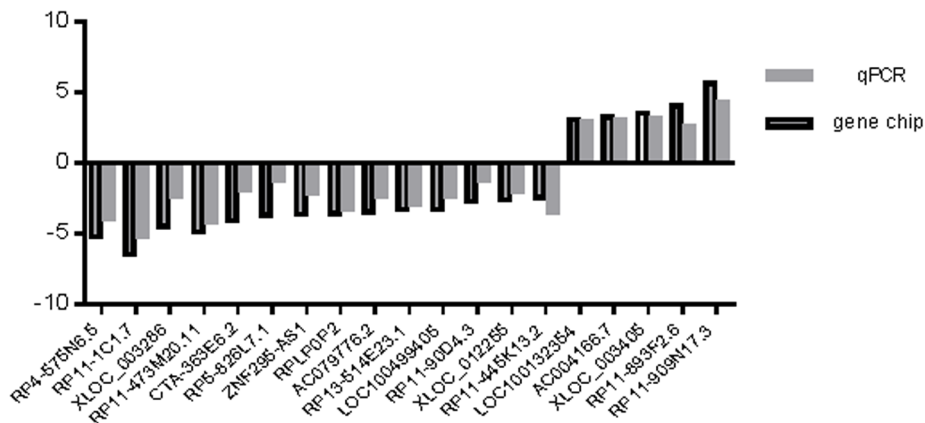


Figure 5. Comparison between gene chip data and qPCR result. RP11-1C1.7, RP4-575N6.5, RP11-473M20.11, XLOC_003286, CTA-363E6.2, RP5-826L7.1, ZNF295-AS1, RPLP0P2, AC079776.2, RP13-514E23.1, LOC100499405, RP1-90D4.3, XLOC_012255, RP11-445K13.2, LOC100132354, AC004166.7, XLOC_003405, RP11-893F2.6, RP11-909N17.3 determined to be differentially expressed in lung adenocarcinoma samples compared with NT samples in six patients by microarray were validated by qPCR. The heights of the columns in the chart represent the log-transformed median fold changes (T/N) in expression across the six patients for each of the four lncRNAs validated; the bars represent standard errors. The validation results of the 19 lncRNAs indicated that the microarray data correlated well with the qPCR results. doi:10.1371/journal.pone.0104044.g005

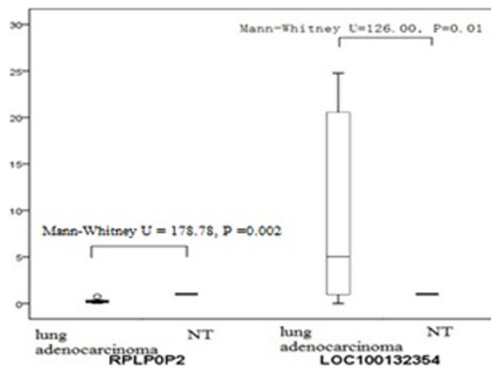


Figure 6. The expression level of RPLP0P2 and LOC100132354 between one hundred lung adenocarcinoma sample and the normal lung sample. LOC100132354 expression of lung adenocarcinoma was significantly higher than the adjacent tissues (Mann-Whitney $U = 126.00$, $P = 0.01$), while RPLP0P2 expression of lung adenocarcinoma was significantly lower than the adjacent tissues (Mann-Whitney $U = 178.78$, $P = 0.002$). doi:10.1371/journal.pone.0104044.g006

included in this study. Total RNA (2 mg) was transcribed to cDNA. PCR was performed in a total reaction volume of 20 μ l, including 10 μ l of SYBR Premix (2 \times), 2 μ l of cDNA template, 1 μ l of PCR forward primer (10 mM), 1 μ l of PCR reverse primer (10 mM), and 6 μ l of double-distilled water. The quantitative real-time PCR reaction included an initial denaturation step of 10 min at 95°C; 40 cycles of 5 s at 95°C, 30 s at 60°C; and a final extension step of 5 min at 72°C. All experiments were performed in triplicate, and all samples were normalized to GAPDH. The median in each triplicate was used to calculate relative lncRNA concentrations ($\Delta\text{Ct} = \text{Ct median lncRNA} - \text{Ct median GAPDH}$), and the fold changes in expression were calculated [19].

Statistical methods

All results are expressed as mean \pm standard deviation. Statistical analysis was performed for the comparison of two groups in the microarray, and analysis of variance for multiple comparisons was performed using the Student's *t*-test. Differences with $P < 0.05$ were considered statistically significant in both cases. The fold change and the Student's *t*-test were used to analyze the statistical significance of the microarray results. The false discovery rate (FDR) was calculated to correct the *P*-value. The threshold value used to designate differentially expressed lncRNAs and mRNAs was a fold change of ≥ 2.0 or ≤ 0.5 ($P < 0.05$).

Results

Overview of lncRNA profiles

To study the potential biological functions of lncRNAs in lung adenocarcinoma, we examined the lncRNA and mRNA expression profiles in human lung adenocarcinoma through microarray analysis (Fig. 1). For this analysis, authoritative data sources containing more than 30,586 lncRNAs were used. The expression profiles of 2,420 lncRNAs indicated that they were differentially expressed (fold change ≥ 2.0 or ≤ 0.5 ; $P < 0.05$) between lung adenocarcinoma and normal lung samples. Among these, 1,213 lncRNAs were found to be upregulated more than two-fold in the lung adenocarcinoma group compared to the normal lung group, while 1,207 lncRNAs were downregulated more than two-fold ($P < 0.05$; Table S4, S5 in File S1 and Fig. 1).

lncRNA classification and subgroup analysis

The expression profiles of 564 intergenic lncRNAs indicated that they were differentially expressed (fold change ≥ 2.0 , $P < 0.05$) between lung adenocarcinoma and normal lung samples. Among these, 338 were upregulated and 226 were downregulated. We also identified some nearby coding genes that may be regulated by these lncRNAs (Table S6 in File S1). lncRNAs with enhancer-like functions (lncRNA-a) were identified using GENCODE annotation. The expression profiles of 77 enhancer-like lncRNAs indicated that they were differentially expressed (fold change ≥ 2.0 , $P < 0.05$) between lung adenocarcinoma and normal lung samples. Among these, 43 were upregulated and 34 were downregulated. We also identified some nearby coding genes that may be regulated by these enhancer-like lncRNAs (Table S7 in File S1). HOX cluster profiling: This data table contains 83 HOX clusters (Table S8 in File S1).

Overview of mRNA profiles

In total, 1,109 mRNAs were found to be differentially expressed between the lung adenocarcinoma and normal lung samples, including 278 that were upregulated and 831 that were downregulated (Table S9–S10 in File S1, Figs. 1–2).

GO analysis

The genes corresponding to the downregulated mRNAs included 357 genes involved in biological processes, 98 genes involved in cellular components and 96 genes involved in molecular functions. The genes corresponding to the upregulated mRNAs included 244 genes involved in biological processes, 80 genes involved in cellular components, and 69 genes involved in molecular functions (Table S11–S12 in File S1).

Pathway analysis

Seven upregulated pathways were identified, including ethanol metabolism, viral carcinogenesis, RNA transduction, and cell cycle pathways. Twenty-four downregulated pathways were identified, including propionate metabolism and fatty acid metabolism pathways (Tables S13 in File S1 and figure 3–4).

lncRNA-mRNA co-expression network

We build the lncRNA-mRNA co-expression network of four important lncRNAs, including RP11-1C1.7 (ENST00000416930), RPLP0P2 (ENST00000492786), LOC100132354 (NR_024478), RP11-909N17.3 (ENST00000519762), See Table S14 in File S1.

Real-time quantitative PCR validation

Based on the features of the differentially expressed lncRNAs, such as fold difference, gene locus, and nearby encoding gene, we initially identified a number of interesting candidate lncRNAs for further analysis (including RP11-1C1.7, RP4-575N6.5, RP11-473M20.11, XLOC_003286, CTA-363E6.2, RP5-826L7.1, ZNF295-AS1, RPLP0P2, AC079776.2, RP13-514E23.1, LOC100499405, RP11-90D4.3, XLOC_012255, RP11-445K13.2, LOC100132354, AC004166.7, XLOC_003405, RP11-893F2.6, and RP11-909N17.3). We verified the expression of these lncRNAs by real-time quantitative RT-PCR by using GAPDH as a reference gene and by calculating the $2^{-\Delta\Delta\text{Ct}}$ values. We found that the microarray results for several of the lncRNAs were consistent with the results of RT-PCR (Fig. 5). Of these, LOC100132354 and RPLP0P2 exhibited the most significantly changed expression in our analysis of 100 pairs of lung adenocarcinoma and normal lung tissue samples. The expression of LOC100132354 was significantly higher in lung adenocarcinoma

ma than in the adjacent tissues (Mann-Whitney $U = 126.00$, $P = 0.01$), while the expression of RPLP0P2 was significantly lower in lung adenocarcinoma than in the adjacent tissues (Mann-Whitney $U = 178.78$, $P = 0.002$) (Fig. 6).

Discussion

According to the 2012 China Oncology Annual Report, in 2009, the incidence and mortality of lung cancer was the highest among cancers in male patients and the second highest among cancers in female patients in China. Lung adenocarcinoma is a type of NSCLC that has rising incidence rates in women and non-smokers. However, the pathogenesis of lung cancer remains unclear; therefore, further study of lung cancer is of great importance.

LncRNAs play an important role in many biological processes, including X-chromosome inactivation, gene imprinting, and stem cell maintenance [20,21]. Furthermore, lncRNAs are important factors in the control of gene expression in cancer [22], and lncRNAs such as HOTAIR have been shown to play an important role in the development and progression of tumors [9]. It has also been demonstrated that lncRNAs are differentially expressed in normal cells and tumor cells. As lncRNAs constitute an important class of gene expression regulatory factors, their aberrant expression would inevitably lead to abnormal gene expression levels, which may result in tumorigenesis. Promoters bind to many transcription factors by mechanisms such as chromosomal rearrangements and transfer elements [23].

In this study, we analyzed lncRNA expression profiles in the tissues of lung adenocarcinoma patients to uncover the potential role of lncRNAs in the pathogenesis of this disease. High-throughput microarray techniques revealed a set of differentially expressed lncRNAs, including 1,213 that were upregulated and 1,207 that were downregulated in lung adenocarcinoma tissue compared to normal lung tissue. LncRNAs are usually divided into five categories: sense, antisense, bidirectional, intronic, and intergenic. LncRNAs are known to function via a variety of mechanisms; however, a common and important function of lncRNAs is to alter the expression of nearby encoding genes by affecting the process of transcription [24] or directly playing an enhancer-like role [25,26]. In the present study, we increased the accuracy of target prediction by comparing differentially expressed mRNAs with differentially expressed lncRNAs. The lncRNA expression profiles indicated that 564 lncRNAs were differentially expressed (338 upregulated and 226 downregulated) between lung adenocarcinoma and normal lung samples. The expression profiles included 77 differentially expressed enhancer-like lncRNAs, with 43 upregulated and 34 downregulated. We also identified some nearby coding genes that may be regulated by lncRNAs and enhancer-like lncRNAs. Moreover, we also performed HOX cluster profiling of lncRNAs and coding transcripts.

Otherwise, we build the lncRNA-mRNA co-expression network and we analyzed four important lncRNAs, including RP11-1C1.7, RPLP0P2, LOC100132354, RP11-909N17.3. It shown that RP11-1C1.7, RPLP0P2, LOC100132354, RP11-909N17.3 were individually relation to some mRNAs.

References

- Jemal A, Murray T, Ward E, Samuels A, Tiwari RC, et al. (2005) Cancer statistics, 2005. *CA Cancer J Clin* 55: 10–30.
- Gridelli C, Rossi A, Maione P (2003) Treatment of non-small-cell lung cancer: state of the art and development of new biologic agents. *Oncogene* 22: 6629–6638. doi:10.1038/sj.onc.1206957.
- Stewart DJ (2010) Tumor and host factors that may limit efficacy of chemotherapy in non-small cell and small cell lung cancer. *Crit Rev Oncol Hematol* 75: 173–234. doi:10.1016/j.critrevonc.2009.11.006.
- Chen CH, Lai JM, Chou TY, Chen CY, Su LJ, et al. (2009) VEGFA upregulates FLJ10540 and modulates migration and invasion of lung cancer via PI3K/AKT pathway. *PLoS One* 4: e5052. doi:10.1371/journal.pone.0005052.

In order to obtain insights into lncRNA target gene function, GO analysis and KEGG pathway annotation were applied to the lncRNA target gene pool. GO analysis revealed that the number of genes corresponding to downregulated mRNAs was larger than that corresponding to upregulated mRNAs. KEGG annotation showed that there were seven upregulated pathways (including ethanol metabolism, viral carcinogenesis, RNA transduction, and cell cycle pathways) and 24 downregulated pathways (including propionate metabolism and fatty acid metabolism pathways). These pathways might play important roles in the occurrence and development of lung adenocarcinoma. We found that 19 of the lncRNAs identified in the microarray analysis were confirmed by RT-PCR to be aberrantly expressed in lung cancer tissues. Among these lncRNAs, LOC100132354 was the most significantly upregulated and RPLP0P2 was the most significantly downregulated. This result suggests that LOC100132354 and RPLP0P2 might contribute to the development of lung adenocarcinoma; further study of the biological function of LOC100132354 and RPLP0P2 will be required to confirm this notion.

To summarize, our study revealed a set of lncRNAs with differential expression in lung adenocarcinoma compared with normal lung tissue. Furthermore, potential roles for these lncRNAs in the regulation of ethanol metabolism and propionate metabolism signaling pathways were identified. Moreover, we found that LOC100132354 and RPLP0P2 might contribute to the development of lung adenocarcinoma. Further investigation of the lncRNAs identified in this study will likely shed light on their biological functions and their association with lung cancer.

Supporting Information

File S1 Contains the files: Table S1, The demographics characteristics of 5 cases lung adenocarcinoma for microarray analysis. **Table S2**, The demographics characteristics of 100 cases lung adenocarcinoma. **Table S3**, LncRNAs gene primers in the study for Q-PCR. **Table S4**, Upregulated lncRNAs in lung adenocarcinoma. **Table S5**, Downregulated lncRNAs in lung adenocarcinoma. **Table S6**, LincRNAs profile of lung adenocarcinoma. **Table S7**, Enhancer lncRNAs profile of lung adenocarcinoma. **Table S8**, HOX profile of lung adenocarcinoma. **Table S9**, Upregulated mRNA in lung adenocarcinoma. **Table S10**, Downregulated mRNA in lung adenocarcinoma. **Table S11**, GO analysis upregulated mRNAs. **Table S12**, GO analysis downregulated mRNAs. **Table S13**, Pathway analysis. **Table S14**, CNC_network. (RAR)

Acknowledgments

We thank all of the donors who donated to the Microarray Service at KangChen Bio-technology Company in Shanghai.

Author Contributions

Conceived and designed the experiments: YW GX QP KH. Performed the experiments: YW GX QP KH WZ. Analyzed the data: Jie Chen FY TZ JP. Contributed reagents/materials/analysis tools: KH WZ Jian Chen. Wrote the paper: KH WZ Jie Chen FY. Sequence analysis: KH WZ.

5. Ogawa E, Takenaka K, Katakura H, Adachi M, Otake Y, et al. (2008) Perimembrane Aurora-A expression is a significant prognostic factor in correlation with proliferative activity in non-small-cell lung cancer (NSCLC). *Ann Surg Oncol* 15: 547–554. doi:10.1245/s10434-007-9653-8.
6. Rachev B, Woods LM, Mitry E, Riga M, Cooper N, et al. (2008) Cancer survival in England and Wales at the end of the 20th century. *Br J Cancer* 99 Suppl 1: S2–10. doi:10.1038/sj.bjc.6604571.
7. Ponting CP, Oliver PL, Reik W (2009) Evolution and functions of long noncoding RNAs. *Cell* 136: 629–641. doi:10.1016/j.cell.2009.02.006.
8. Dinger ME, Pang KC, Mercer TR, Crowe ML, Grimmond SM, et al. (2009) NRED: a database of long noncoding RNA expression. *Nucleic Acids Res* 37: D122–126. doi:10.1093/nar/gkn617.
9. Gupta RA, Shah N, Wang KC, Kim J, Horlings HM, et al. (2010) Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. *Nature* 464: 1071–1076. doi:10.1038/nature08975.
10. Fu X, Ravindranath L, Tran N, Petrovics G, Srivastava S (2006) Regulation of apoptosis by a prostate-specific and prostate cancer-associated noncoding gene, PCGEM1. *DNA Cell Biol* 25: 135–141. doi:10.1089/dna.2006.25.135.
11. Zhang H, Chen Z, Wang X, Huang Z, He Z, et al. (2013) Long non-coding RNA: a new player in cancer. *J Hematol Oncol* 6: 37. doi:10.1186/1756-8722-6-37.
12. Hauptman N, Glavac D (2013) Long non-coding RNA in cancer. *Int J Mol Sci* 14: 4655–4669. doi:10.3390/ijms14034655.
13. Chen G, Wang Z, Wang D, Qiu C, Liu M, et al. (2013) LncRNADisease: a database for long-non-coding RNA-associated diseases. *Nucleic Acids Res* 41: D983–986. doi:10.1093/nar/gks1099.
14. Gibb EA, Brown CJ, Lam WL (2011) The functional role of long non-coding RNA in human carcinomas. *Mol Cancer* 10: 38. doi:10.1186/1476-4598-10-38.
15. Ji P, Diederichs S, Wang W, Boing S, Metzger R, et al. (2003) MALAT-1, a novel noncoding RNA, and thymosin beta4 predict metastasis and survival in early-stage non-small cell lung cancer. *Oncogene* 22: 8031–8041. doi:10.1038/sj.onc.1206928.
16. Thai P, Statt S, Chen CH, Liang E, Campbell C, et al. (2013) Characterization of a novel long noncoding RNA, SCAL1, induced by cigarette smoke and elevated in lung cancer cell lines. *Am J Respir Cell Mol Biol* 49: 204–211. doi:10.1165/rcmb.2013-0159RC.
17. Yang Y, Li H, Hou S, Hu B, Liu J, et al. (2013) The noncoding RNA expression profile and the effect of lncRNA AK126698 on cisplatin resistance in non-small-cell lung cancer cell. *PLoS One* 8: e65309. doi:10.1371/journal.pone.0065309.
18. Han L, Kong R, Yin DD, Zhang EB, Xu TP, et al. (2013) Low expression of long noncoding RNA GAS6-AS1 predicts a poor prognosis in patients with NSCLC. *Med Oncol* 30: 694. doi:10.1007/s12032-013-0694-5.
19. Ren S, Peng Z, Mao JH, Yu Y, Yin C, et al. (2012) RNA-seq analysis of prostate cancer in the Chinese population identifies recurrent gene fusions, cancer-associated long noncoding RNAs and aberrant alternative splicings. *Cell Res* 22: 806–821. doi:10.1038/cr.2012.30.
20. Mercer TR, Dinger ME, Mattick JS (2009) Long non-coding RNAs: insights into functions. *Nat Rev Genet* 10: 155–159. doi:10.1038/nrg2521.
21. Wang KC, Chang HY (2011) Molecular mechanisms of long noncoding RNAs. *Mol Cell* 43: 904–914. doi:10.1016/j.molcel.2011.08.018.
22. Khachane AN, Harrison PM (2010) Mining mammalian transcript data for functional long non-coding RNAs. *PLoS One* 5: e10316. doi:10.1371/journal.pone.0010316.
23. Loh YH, Wu Q, Chew JL, Vega VB, Zhang W, et al. (2006) The Oct4 and Nanog transcription network regulates pluripotency in mouse embryonic stem cells. *Nat Genet* 38: 431–440. doi:10.1038/ng1760.
24. Mattick JS, Gagen MJ (2001) The evolution of controlled multitasked gene networks: the role of introns and other noncoding RNAs in the development of complex organisms. *Mol Biol Evol* 18: 1611–1630.
25. Mattick JS (2010) Linc-ing Long noncoding RNAs and enhancer function. *Dev Cell* 19: 485–486. doi:10.1016/j.devcel.2010.10.003.
26. Orom UA, Derrien T, Beringer M, Gumireddy K, Gardini A, et al. (2010) Long noncoding RNAs with enhancer-like function in human cells. *Cell* 143: 46–58. doi:10.1016/j.cell.2010.09.001.