HYPOTHESIS

Comprehensive analysis of the long noncoding RNA HOXA11-AS gene interaction regulatory network in NSCLC cells

Yu Zhang1†, Rong-quan He2†, Yi-wu Dang1, Xiu-ling Zhang1, Xiao Wang3, Su-ning Huang4, Wen-ting Huang1, Meng-tong Jiang1, Xiao-ning Gan1, You Xie1, Ping Li1, Dian-zhong Luo1, Gang Chen1*‡ and Ting-qing Gan2*‡

Abstract

Background: Long noncoding RNAs (lncRNAs) are related to different biological processes in non-small cell lung cancer (NSCLC). However, the possible molecular mechanisms underlying the effects of the long noncoding RNA HOXA11-AS (HOXA11 antisense RNA) in NSCLC are unknown.

Methods: HOXA11-AS was knocked down in the NSCLC A549 cell line and a high throughput microarray assay was applied to detect changes in the gene profiles of the A549 cells. Bioinformatics analyses (gene ontology (GO), pathway, Kyoto Encyclopedia of Genes and Genomes (KEGG), and network analyses) were performed to investigate the potential pathways and networks of the differentially expressed genes. The molecular signatures database (MSigDB) was used to display the expression profiles of these differentially expressed genes. Furthermore, the relationships between the HOXA11-AS, de-regulated genes and clinical NSCLC parameters were verified by using NSCLC patient information from The Cancer Genome Atlas (TCGA) database. In addition, the relationship between HOXA11-AS expression and clinical diagnostic value was analyzed by receiver operating characteristic (ROC) curve.

Results: Among the differentially expressed genes, 277 and 80 genes were upregulated and downregulated in NSCLC, respectively (fold change ≥ 2.0, P < 0.05 and false discovery rate (FDR) < 0.05). According to the degree of the fold change, six upregulated and three downregulated genes were selected for further investigation. Only four genes (RSPO3, ADAMTS8, DMBT1, and DOCK8) were reported to be related with the development or progression of NSCLC based on a PubMed search. Among all possible pathways, three pathways (the PI3K-Akt, TGF-beta and Hippo signaling pathways) were the most likely to be involved in NSCLC development and progression. Furthermore, we found that HOXA11-AS was highly expressed in both lung adenocarcinoma and squamous cell carcinoma based on TCGA database. The ROC curve showed that the area under curve (AUC) of HOXA11-AS was 0.727 (95% CI 0.663–0.790) for lung adenocarcinoma and 0.933 (95% CI 0.906–0.960) for squamous cell carcinoma patients. Additionally, the original data from TCGA verified that ADAMTS8, DMBT1 and DOCK8 were downregulated in both lung adenocarcinoma and squamous cell carcinoma, whereas RSPO3 expression was upregulated in lung adenocarcinoma and downregulated in lung squamous cell carcinoma. For the other five genes (STMN2, SPINK6, TUSC3, LOC100128054, and C8orf22), we
Background
Lung cancer is the most common cancer worldwide and the first leading cause of cancer death [1, 2]. More than 1.8 million lung cancer patients are diagnosed each year, accounting for approximately 13% of newly diagnosed cancer cases [3]. Lung cancer can be divided into two categories based on the histological type [small cell lung cancer (SCLC) and non-small cell lung cancer (NSCLC)]. NSCLC accounts for 80–85% of new lung cancers. NSCLC can be divided into different subgroups, such as adenocarcinoma, squamous cell carcinoma, adenosquamous carcinoma, undifferentiated carcinoma and large cell carcinoma. More than 70% of NSCLC cases are advanced disease and the 5-year survival rate for NSCLC is only 16% [4]. Hence, research into the etiology and mechanism is of great significance for the diagnosis and treatment of lung cancer.

Long non-coding RNAs (lncRNAs) represent RNAs more than 200 nucleotides in length that lack a protein-coding capacity. Many lncRNAs have been reported to be associated with transcriptional regulation, disease development or epigenetic gene regulation [5–7]. Additionally, lncRNAs are involved in numerous biological functions, such as tumorigenesis, immune responses, cell differentiation and other biological processes [8–11]. To date, many lncRNAs have been reported to play important roles in NSCLC, such as lncRNA-TATDN1, PVT1 and MALAT1, which may influence the NSCLC cell proliferation, invasion and metastasis, respectively [12–14]. However, the biological and molecular mechanisms underlying the actions of HOXA11-AS in NSCLC have not been fully explored.

HOXA11-AS (also known as HOXA11S, HOXA-AS5, HOXA11AS, HOXA11-AS1, and NCRNA00076) is located on 7p15.2 (NCBI Gene ID: 221883). HOXA11-AS is a member of the homeobox (HOX) family of genes with a length of 3885 nt. To date, only 2 studies have reported a relationship between HOXA11-AS and cancer. Richards et al. [15] demonstrated that HOXA11-AS inhibited the oncogenic phenotype of epithelial ovarian cancer by analyzing genome-wide association study data and performing a series of functional experiments. Wang et al. [16] confirmed that HOXA11-AS was a cell cycle-associated IncRNA and could serve as a biomarker of glioma progression using a high-throughput microarray and gene set enrichment analysis. However, the expression and function of HOXA11-AS in NSCLC tissues is unknown. We designed this study to explore expression profile changes after HOXA11-AS knock-down and the possible molecular mechanisms of HOXA11-AS in NSCLC development and progression. A flow chart of this study was shown in Fig. 1.

Methods
Knock-down of HOXA11-AS in the NSCLC cell A549 and transfection with HOXA11-AS-siRNA
The human NSCLC A549 cell line was purchased from the Type Culture Collection of the Chinese Academy of Sciences (Shanghai, China). NSCLC A549 cells were cultivated with 10% heat-inactivated fetal bovine serum (Invitrogen Corp, Grand Island, NY, USA) in a humidified 5% CO2 atmosphere with 2 mM glutamine and gentamicin at 37 °C. Three DcR3-specific siRNAs were synthesized by GenePharma (Shanghai, China) and merged into one siRNA pool (Table 1). The NSCLC A549 cell line was transfected with the HOXA11-AS-siRNA. The CombiMAG magnetofection reagent (OZ BIOSCIENCES, Marseille, France) was used for the transfection according to the manufacturer’s instructions.

Microarray analysis and computational analysis
The sample analysis and microarray hybridization were performed by Kangchen Bio-tech (Shanghai, China). Briefly, RNA was purified and extracted from 1 mg of total RNA after removing the rRNA (mRNA-ONLY Eukaryotic mRNA Isolation Kit, Epicentre Biotechnologies, Madison, USA). Then, each sample was transcribed and amplified into fluorescent cRNA using a random

Conclusions: Our results showed that the gene profiles were significantly changed after HOXA11-AS knock-down in NSCLC cells. We speculated that HOXA11-AS may play an important role in NSCLC development and progression by regulating the expression of various pathways and genes, especially DOCK8 and TGF-beta pathway. However, the exact mechanism should be verified by functional experiments.

Keywords: HOXA11-AS, NSCLC, Microarray assay, GO, KEGG, Pathway
priming method. The cRNAs were labeled and hybridized onto the Human MRNA Array v3.0 (8 × 60 K, Arraystar, Rockville, MD, USA). After washing the slides, the arrays were scanned with the Agilent Scanner G2505C. The Agilent Feature Extraction software (version 11.0.1.1) was used to analyze the acquired array images. Quantile normalization and subsequent data processing were implemented by the GeneSpring GX v11.5.1 software package (Agilent Technologies). Differentially expressed genes were identified based on fold change filtering (fold change ≥2.0 or ≤0.5), a paired t test (p < 0.05) and multiple hypothesis testing (FDR < 0.05). The P values and FDR were calculated with Microsoft Excel and MATLAB, respectively. Differentially expressed genes between the RNAi and control samples were identified with an absolute fold change >2 as the cut-off. The molecular signatures database (MSigDB, http://www.broadinstitute.org/msigdb) was applied to visualize the expression profiles of these differentially expressed genes (Figs. 2, 3).

**GO analysis and pathway analysis**
To better understand the potential roles of the differentially expressed genes, gene ontology (GO) analysis and pathway analysis were performed as previously described [17]. In this process, we included the following three independent categories derived from the GO Consortium website (http://www.geneontology.org): biological process (BP), cellular component (CC) and molecular function (MF) [17]. The enrichment of the upregulated and downregulated coding genes was analyzed by uploading the datasets to the database for annotation, visualization and integrated discovery (DAVID, http://david.abcc.ncifcrf.gov/). The Kyoto Encyclopedia of Genes and Genomes (KEGG) database (http://www.genome.jp/kegg/) was used to analyze the biological pathways where there was an obvious enrichment of differentially expressed genes [18].

**Additional analysis of 9 de-regulated genes in NSCLC from TCGA**
TCGA is a collection of exome sequencing, DNA methylation, SNP array, miRNA-seq, and RNA-seq data [19]. TCGA can be used to analyze complicated clinical profiles and cancer genomics [20, 21]. In this study, original expression data for HOXA11-AS and the 9 genes de-regulated in lung adenocarcinoma and squamous cell carcinoma were extracted from TCGA and analyzed. Additionally, original data for cancerous or

---

**Table 1 The HOXA11-AS-siRNA sequences**

| ID            | Target sequence | GC % |
|---------------|-----------------|------|
| HOXA11-AS-RNAi(32154-2) | CTACCATCCTGAGCCTTA | 52.63 |
| HOXA11-AS-RNAi(32155-1) | TGACATCCGAGGAGACTTC | 52.63 |
| HOXA11-AS-RNAi(32156-1) | CGTAATCGCCGGTGTAACT | 52.63 |
non-cancerous lung tissues were downloaded and analyzed. Also, the relationship between HOXA11-AS expression and clinical diagnostic value was analyzed by receiver operating characteristic (ROC) curve. Besides, we extracted the co-genes of HOXA11-AS from TCGA through R Project for Statistical Computing (https://www.r-project.org/). Genes with a FDR < 0.05 was considered for co-expressed relationship.
Statistical analysis
SPSS 20.0 was applied for the statistical analysis. The Mann–Whitney U test was used to compare the expression of the four deregulated genes in terms of different clinical features (age, gender, TNM stage, tumor size, distant metastasis and lymph node metastasis). P < 0.05 was considered statistically significant (two-sides).

Results
Gene expression profiles regulated by the HOXA11-AS IncRNA
A high throughput microarray assay was applied to detect differential expression profiles between HOXA11-AS and HOXA11-AS RNAi in three paired A549 cell cultures. Thirteen thousand three hundred and twenty-three upregulated genes and 14,384 downregulated genes were differentially expressed in the HOXA11-AS-control and HOXA11-AS-RNAi groups. A summary of these differentially expressed genes is presented in Fig. 4. The fold changes (HOXA11-AS-control vs HOXA11-AS-RNAi) and P values were calculated using the normalized expression values. Using microarray analysis, 357 genes were identified as significantly differentially expressed in NSCLC compared with the RNAi control samples (fold change ≥ 2.0, P < 0.05 and FDR < 0.05). Among them, 277 genes were upregulated in all three NSCLC groups and 80 genes were downregulated. Furthermore, the number of aberrantly expressed genes varied with the different fold changes (Table 2). Among them, 16 genes were upregulated by more than sixfold in the HOXA11-AS compared to the HOXA11-AS RNAi samples and 3 genes were downregulated by more than fourfold. 6 of the 15 upregulated genes were upregulated by more than tenfold (Table 2). The top 6 upregulated and top 3 downregulated genes are shown in Table 3. Among these 9 aberrantly expressed genes, the expression of RSPO3 (NM_032784, fold change = 41.610487, P = 8.0502E−09) was dramatically upregulated and the expression of LOC100128054 (NR_033969, fold change = 4.6652225, P = 4.45517E−05) was significantly downregulated. When we searched PubMed (http://www.ncbi.nlm.nih.gov/pubmed) to identify reported functions for these differentially expressed genes, we found that only 4 genes (RSPO3, ADAMTS8, DMBT1, and DOCK8) were reported to be associated with NSCLC. RSPO3 was reported to promote tumor aggressiveness in Keap1-deficient lung adenocarcinomas [22]. ADAMTS8 was related to promoter hypermethylation in early-stage NSCLCs [23, 24]. DMBT1 was a candidate tumor suppressor gene; DMBT1 expression is often lost in lung cancer, indicating that DMBT1 inactivation may have a significant influence on lung tumorigenesis [25]. DOCK8 was suggested to be involved in the development and/or progression of lung cancer [26]. Thus, these genes may play essential roles in the occurrence and development of NSCLC.

GO and pathway analysis
The GO analysis identified biological processes, molecular functions and cellular components in which the differentially expressed genes may be involved. The top five most enriched GO terms are shown in Table 4. The GO analysis results clarified the most significant functional
groups, such as single-organism process, cellular response to stimulus, biological regulation, and cellular component organization (Figs. 5, 6). To better understand the relevant functions of these genes, a function network was constructed based on the GO analysis (Figs. 7, 8). We constructed only the BP ontology for the downregulated genes because only 80 genes were identified (Fig. 8).

The KEGG analysis showed that the aberrantly expressed genes might be related to different pathways. A total of 21 upregulated pathways and only 1 downregulated pathway were available through the pathway analysis. The most important enriched pathway terms are shown in Table 5 ($P_{\text{upregulated}} < 0.01$). Three pathways (PI3K-Akt signaling pathway, TGF-beta signaling pathway and Hippo signaling pathway) were previously reported to be involved in NSCLC development and progression. As reported, the PI3K-Akt signaling pathway was related to NSCLC cell proliferation, apoptosis and autophagy [27–29]. The TGF-beta signaling pathway could be associated with the NSCLC cell DNA damage response, radiation sensitivity, viability and invasion capacity [30, 31]. The Hippo signaling pathway was involved in NSCLC cell migration and invasion [32].

A gene network of these 357 genes was constructed in the present study (Fig. 9). The relationships between HOXA11-AS and the differentially expressed genes were easily observed from this network analysis.

Supplementary information from the TCGA

In order to explore the relationship between HOXA11-AS expression and NSCLC, we performed a clinical study with the original data in TCGA. We found that HOXA11-AS was upregulated in both lung adenocarcinoma and

Table 2 Number of aberrantly expressed genes in the microarray

|        | Fold change 2–4 | Fold change 4–6 | Fold change > 6 | Total |
|--------|-----------------|-----------------|-----------------|-------|
| Upregulated | 238             | 23              | 16              | 277   |
| Downregulated | 77              | 3               |                 | 80    |

Table 3 The top 6 upregulated and top 3 downregulated genes

| Genbank accession | Gene symbol | Fold change | P       |
|-------------------|-------------|-------------|---------|
| NM_032784         | RSPO3       | 41.610487   | 8.05E−09|
| NM_007037         | ADAMTS8     | 21.38143    | 8.074E−06|
| NM_007029         | STMN2       | 20.713397   | 3.617E−05|
| NM_007329         | DMBT1       | 12.747211   | 7.941E−07|
| NM_205841         | SPINK6      | 10.662449   | 8.553E−08|
| NM_178234         | TUSC3       | 10.584962   | 1.619E−05|
| NR_033969         | LOC100128054| 4.6652225   | 4.455E−05|
| NM_203447         | DOCK8       | 4.4005987   | 7.142E−06|
| NM_001007176      | C8orf22     | 4.0156588   | 0.000753|

Fig. 4 Gene clip after HOXA11-AS knock-down in NSCLC. a Volcano plot; b box-scatter plot
squamous cell carcinoma compared to non-cancerous lung tissues (both $P < 0.0001$, Fig. 10a, b). And the ROC curve revealed that the area under curve (AUC) of HOXA11-AS was 0.727 (95% CI 0.663–0.790) for lung adenocarcinoma patients and 0.933 (95% CI 0.906–0.960) for squamous cell carcinoma patients (both $P < 0.0001$), which could gain high diagnostic value of HOXA11-AS level in NSCLC (Fig. 10c, d).

To elucidate the relationships between the 9 de-regulated genes and NSCLC, we searched the original data from 514 adenocarcinoma cases and 501 squamous cell carcinoma cases in TCGA. We also compared gene expression between adenocarcinoma, squamous cell carcinoma and non-cancerous lung tissues. We found that RSPO3, ADAMTS8, DMBT1 and DOCK8 were all down-regulated in squamous cell carcinoma tissues compared to non-cancerous lung tissues, whereas STMN2, TUSC3 and C8orf22 were upregulated in squamous cell carcinoma (all $P < 0.001$, Fig. 11). Additionally, ADAMTS8, DMBT1 and DOCK8 were down-regulated in adenocarcinoma and STMN2 and TUSC3 were up-regulated in lung adenocarcinoma (all $P < 0.01$, Fig. 12). RSPO3 was overexpressed in adenocarcinoma but not squamous cell carcinoma ($P = 0.023$).

Furthermore, we compared the correlation between HOXA11-AS and de-regulated genes in NSCLC based on TCGA. The results showed that the HOXA11-AS expression was negatively correlated with DOCK8 in squamous cell carcinoma ($r = -0.124$, $P = 0.048$) and lung adenocarcinoma ($r = -0.176$, $P = 0.005$). No obviously correlation was found between HOXA11-AS and other de-regulated genes (Table 6). Besides, the co-genes

| GO.ID      | Term                                | Ontology | Enrichment score | P               |
|------------|--------------------------------------|----------|------------------|-----------------|
| GO:0071294 | Cellular response to zinc ion        | BP       | 8.557911479      | 2.76751E–09     |
| GO:0044707 | Single-multicellular organism process| BP       | 8.514757717      | 3.05663E–09     |
| GO:0048731 | System development                   | BP       | 8.040063137      | 3.98049E–09     |
| GO:0032501 | Multicellular organisam process      | BP       | 8.203661346      | 6.2566E–09      |
| GO:0048856 | Anatomical structure development     | BP       | 8.143712557      | 7.182E–09       |
| GO:0005578 | Proteinaceous extracellular matrix   | CC       | 8.300810864      | 5.00263E–09     |
| GO:001012  | Extracellular matrix                 | CC       | 8.221562341      | 6.00396E–09     |
| GO:0005604 | Basement membrane                    | CC       | 6.79110942       | 1.61769E–07     |
| GO:0044420 | Extracellular matrix part            | CC       | 5.570158827      | 2.69055E–06     |
| GO:0005576 | Extracellular region                 | CC       | 5.516943306      | 3.04128E–06     |
| GO:0005102 | Receptor binding                     | MF       | 8.517557576      | 3.03698E–09     |
| GO:0008083 | Growth factor activity               | MF       | 7.9661485        | 1.08106E–08     |
| GO:0005178 | Integrin binding                     | MF       | 4.513658828      | 3.06432E–05     |
| GO:0030414 | Peptidase inhibitor activity         | MF       | 4.359131435      | 4.37206E–05     |
| GO:0061134 | Peptidase regulator activity         | MF       | 4.350560953      | 4.46107E–05     |

Table 4 Top 5 enriched GO terms (BP, CC, and MF) from the microarray data
Fig. 5 Distribution of gene ontology (GO) terms for the upregulated genes in NSCLC. The pie plot showing the gene ontology classification for the upregulated genes in NSCLC. The graph does not contain all upregulated genes because the majority do not have assigned GOs. 

**a** Biological process (BP)  
**b** Cellular component (CC)  
**c** Molecular function (MF)
Fig. 6 Distribution of gene ontology (GO) terms for the downregulated genes in NSCLC. The pie plot showing the gene ontology classification for the downregulated genes in NSCLC. The graph does not contain all downregulated genes because the majority do not have assigned GOs. 

- **Biological process (BP)**
- **Cellular component (CC)**
- **Molecular function (MF)**
Fig. 7  A function network of gene ontology (GO) terms for the upregulated genes in NSCLC.  

- **a** Biological process (BP)  
- **b** Cellular component (CC)  
- **c** Molecular function (MF)
Fig. 8 A function network (BP) of Gene Ontology (GO) terms for the downregulated genes in NSCLC. BP biological process
of HOXA11-AS in TCGA was extracted through R Project for Statistical Computing. We found that RSPO3, ADAMTS8, DMBT1, DOCK8, STMN2, SPINK6 and TUSC3 were the co-genes of HOXA11-AS in lung adenocarcinoma whereas RSPO3, ADAMTS8, DMBT1, DOCK8, STMN2, SPINK6, TUSC3 and C8orf22 were the co-genes of HOXA11-AS in squamous cell carcinoma.

In addition, we also investigated the relationship between the expression levels of the de-regulated genes and clinicopathological parameters or patient survival. Only ADAMTS8 was related to the TNM stage ($t = 0.041, P = 0.032$) in squamous cell carcinoma. In lung adenocarcinoma tissues, RSPO3 was obviously more highly expressed in the advanced stages (III and IV) than the early stages (I–II, $t = −2.462, P = 0.015$). When lymph node metastasis was analyzed, higher RSPO3 expression was found in cases with lymph node metastasis than in cases without ($t = −2.346, P = 0.020$). We also found that higher ADAMTS8 expression was more common in females ($t = −2.924, P = 0.004$) and cases with distant metastasis ($P = 0.045$). Higher DMBT1 and DOCK8 expression was also more common in females than males (all $P < 0.05$). DOCK8 was significantly more highly expressed in the advanced stages (III and IV, $t = 3.482, P = 0.001$) and cases with lymph node metastasis ($t = 2.087, P = 0.037$). Additionally, TUSC3 was related to age ($P = 0.037$). The upregulated expression of RSPO3, ADAMTS8 and DOCK8 was associated with the overall survival (all $P < 0.05$) and disease-free survival of adenocarcinoma patients (all $P < 0.05$, Fig. 13), which indicated that RSPO3, ADAMTS8 and DOCK8 might influence the prognosis of adenocarcinoma. Based on the aforementioned results, we speculated that HOXA11-AS may play an important role in NSCLC development and progression by regulating the expression of various pathways and genes, especially DOCK8 and TGF-beta pathway. However, the exact mechanism should be verified by functional experiments.

**Discussion**

Lung cancer is the most common malignancy in humans and accounts for approximately 13% of newly diagnosed cancer cases per year [1–3]. NSCLC accounts for 80–85% of all lung cancers. Over the past decades, the possible molecular mechanism underlying NSCLC has been extensively explored. However, the particular pathogenesis of NSCLC is still vague. Growing evidence indicates that lncRNAs may play important roles in regulating gene expression in NSCLCs. For example, lncRNA-TATDN1 is associated with NSCLC invasion and metastasis by influencing E-cadherin, HER2, β-catenin and Ezrin expression [12], lncRNA-PVT1 promotes NSCLC cell proliferation by epigenetically regulating LATS2 expression [13] and lncRNA-MALAT1 influences tumor invasion in NSCLC by regulating DNA methylation [14]. In this study, we explored the possible biological and molecular mechanisms of HOXA11-AS in NSCLC. A microarray assay, various bioinformatics analyses (GO, pathway, KEGG, and network analyses) and the original data in TCGA were used to study differentially expressed genes and their relationships with NSCLC. Also, ROC curve showed that HOXA11-AS expression might have an important value in diagnosis of lung cancer. Moreover, we searched Oncomine (https://www.oncomine.org/resource/login.html) and gene expression omnibus

**Table 5 The most important enriched pathway terms from the microarray data**

| Pathway ID | Definition                               | Enrichment_Score | P     |
|------------|------------------------------------------|------------------|-------|
| hsa04978   | Mineral absorption—*Homo sapiens* (human)| 4.539769         | 2.88557E-05 |
| hsa04610   | Complement and coagulation cascades—*Homo sapiens* (human)| 3.791003 | 0.000161807 |
| hsa05200   | Pathways in cancer—*Homo sapiens* (human)| 3.329344 | 0.000468443 |
| hsa05020   | Prion diseases—*Homo sapiens* (human)    | 3.073759         | 0.000843803 |
| hsa05144   | Malaria—*Homo sapiens* (human)           | 2.569023         | 0.002697597 |
| hsa05217   | Basal cell carcinoma—*Homo sapiens* (human)| 2.385839 | 0.00411302 |
| hsa05323   | Rheumatoid arthritis—*Homo sapiens* (human)| 2.34211 | 0.004548726 |
| hsa04060   | Cytokine-cytokine receptor interaction—*Homo sapiens* (human)| 2.310768 | 0.004889133 |
| hsa05020   | Transcriptional misregulation in cancer—*Homo sapiens* (human)| 2.264309 | 0.005441153 |
| hsa04151   | PI3K-Akt signaling pathway—*Homo sapiens* (human)| 2.098056 | 0.007978919 |

| Downregulated genes | Pathway ID | Definition                               | Enrichment_Score | P     |
|---------------------|------------|------------------------------------------|------------------|-------|
| hsa04350            | TGF-beta signaling pathway—*Homo sapiens* (human)| 1.561561 | 0.02744347 |
(GEO; http://www.ncbi.nlm.nih.gov/geo/) database, but no positively relationship was found. In addition, through the above-mentioned bioinformatics analyses, 4 genes (RSPO3, ADAMTS8, DMBT1, and DOCK8) and 3 pathways (PI3K-Akt signaling pathway, TGF-beta signaling pathway and Hippo signaling pathway) were identified as related to NSCLC. The original data from TCGA verified that ADAMTS8, DMBT1 and DOCK8 were down-regulated in adenocarcinoma and squamous cell carcinoma, whereas RSPO3 was overexpressed in adenocarcinoma and down-regulated in squamous cell carcinoma. Furthermore, RSPO3, ADAMTS8 and DOCK8 were also related to the overall survival and disease-free survival of lung adenocarcinoma patients in the TCGA data. Besides, we found that the HOXA11-AS expression was negatively correlated with DOCK8 both in squamous cell carcinoma and lung adenocarcinoma. Therefore, we hypothesized that HOXA11-AS might play an essential role in NSCLC development and progression by regulating DOCK8 expression through TGF-beta pathway.
However, the real mechanism should be verified by functional experiments.

During the process of researching the relationship between these 4 de-regulated genes and 3 pathways, we found that DOCK8 and the TGF-beta signaling pathway played significant roles in the metastasis of lung adenocarcinoma [33]. Yu et al. [33] used RNA and protein analysis, Rac1 activity, imaging, cellular assays, public data set analysis and xenograft mouse models to show that DOCK4 played an important role in mediating TGF-beta-driven lung adenocarcinoma cell extravasation and metastasis. Thus, DOCK4 may act as a key component of the TGF-beta pathway. Additionally, we found that DOCK8 and the Hippo signaling pathway could play a role in neuroblastoma relapse [34]. DOCK8 mutations and YAP activation were reported to be associated with neuroblastoma relapse in one study. YAP is a member of the Hippo signaling pathway [35]; however, whether the expression of DOCK8 plays a role in NSCLC through the Hippo signaling pathway is unknown. DOCK8 (also known as MRD2, ZIR8 and HEL-205) is located on 9p24.3 (NCBI Gene ID: 81704). DOCK family proteins have been confirmed to play roles in the regulation of cell morphology, adhesion, migration and growth [36–39]. DOCK8 was reported to
Fig. 11  Differential expression of genes between squamous cell carcinoma and normal lung tissues based on The Cancer Genome Atlas (TCGA) database. a RSPO3; b ADAMTS8; c DMBT1; d DOCK8; e SPINK6; f TUSC3; g C8orf22
Fig. 12  Differential expression of RSPO3, ADAMTS8, DMBT1 and DOCK8 between lung adenocarcinoma and normal lung tissues based on The Cancer Genome Atlas (TCGA) database. a RSPO3, b ADAMTS8, c DMBT1, d DOCK8, e SPINK6, f TUSC3
exhibited in different cancers, such as hepatocellular carcinoma and some epithelial cancers [40, 41]. However, to date only 2 papers have reported roles for DOCK8 in lung cancer. Kang et al. [42] analyzed 22 lung squamous cell carcinoma cases and found that the loss of chromosome 9 p was specific for lung squamous cell carcinoma; thus, the DOCK8 gene might be a potential target for therapeutic measures against lung squamous cell carcinoma. Takahashi et al. [26] found that genetic and epigenetic inactivation of DOCK8 was related to the development and/or progression of lung cancer using an array-CGH analysis. The original data from TCGA verified that higher DOCK8 expression was related to gender, TNM stage, lymph node metastasis and survival, which indicated that DOCK8 might play a significant role in NSCLC. We found that the TGF-beta signaling pathway was related to radiation sensitivity, extravasation, metastasis and apoptosis [30, 33, 43]. Additionally, the deregulation of the Hippo signaling pathway induced tumors in model organisms and occurred in different human carcinomas, including lung, ovarian, colorectal and liver cancers [44]. The Hippo signaling pathway controls organ size by regulating the cell cycle, proliferation, and apoptosis [45, 46]. However, numerous in vivo and in vitro experiments need to be performed to verify whether HOXA11-AS plays a role in NSCLC development and progression by regulating DOCK8 expression through the TGF-beta or Hippo signaling pathway.

Other differentially expressed genes and pathways were investigated. Several studies have reported the functions of these genes and pathways. Gong et al. [22] found that RSPO3 was aberrantly overexpressed in half of Keap1-deficient lung adenocarcinomas and that RSPO3 overexpression resulted in much poorer survival. In vitro experiments verified that RSPO3 overexpression was related to cell proliferation and migration. The findings of these authors suggest that RSPO3 overexpression may potentially act as a driving mechanism behind the aggressiveness of Keap1-deficient lung adenocarcinomas. Dunn et al. [24] performed a microarray analysis combined with comparative multiplex RT-PCR, immunohistochemical studies and DNA methylation analysis and found that ADAMTS8 was down-regulated in primary NSCLC. ADAMTS8 down-regulation was related to promoter hypermethylation, which might be associated with NSCLC development. Mollenhauer et al. [47] explored DMBT1 expression in normal and lung cancer tissues using reverse-transcription PCR and immunohistochemical studies and found DMBT1 down-regulation in the lung cancer cell lines. However, this finding was controversial because up-regulated expression was detected in the tumor-flanking epithelium and upon respiratory inflammation. The authors found that a switch took place during lung carcinogenesis. Finally, they hypothesized that the sequential changes in DMBT1 expression in different locations reflected a time course that might indicate a possible mechanism in epithelial cancer. In addition, we also further researched the relationships between the other 5 de-regulated genes (STMN2, SPINK6, TUSC3, LOC100128054, and C8orf22) and disease progression. As reported, STMN2 could be a novel developmentally-associated marker and STMN2 could contribute to regulating the adipocyte/osteoblast balance [48]. Also STMN2 could be a novel target of beta-catenin/TCF-mediated carcinogenesis in hepatoma cells [49]. SPINK6

| Gene symbol | Squamous cell carcinoma | Lung adenocarcinoma |
|-------------|-------------------------|---------------------|
|             | R          | P          | R          | P          |
| RSPO3      | −0.022    | 0.724     | −0.025    | 0.656     |
| ADAMTS8    | −0.107    | 0.087     | −0.020    | 0.756     |
| STMN2      | 0.000     | 0.996     | 0.049     | 0.434     |
| DMBT1      | −0.061    | 0.330     | −0.064    | 0.309     |
| SPINK6     | −0.007    | 0.911     | −0.034    | 0.589     |
| TUSC3      | −0.014    | 0.828     | 0.013     | 0.839     |
| DOCK8      | −0.176    | 0.005     | −0.124    | 0.048     |
| C8orf22    | −0.007    | 0.910     | −0.025    | 0.656     |
could be a prognostic indicator in nasopharyngeal carcinoma patients, and SPINK6 could play a critical role in promoting metastasis of nasopharyngeal carcinoma patients [50]. Moreover, TUSC3 was reported to related to the development of different cancers, such as glioblastoma, colorectal cancer, pancreatic cancer, and so on [51–53]. No items of LOC100128054 and C8orf22 were found from pubmed.
In addition, many studies have researched the different mechanisms of the PI3K-Akt signaling pathway. The PI3K/AKT/mTOR signaling pathway is well-known to play essential roles in cell proliferation, invasion, apoptosis, and angiogenesis in lung cancer [54–56]. However, numerous experiments are required to identify the real mechanisms underlying the roles of HOXA11-AS and its corresponding differentially expressed genes in NSCLC.

Conclusion
In summary, because HOXA11-AS may be an important factor in different biological processes of lung cancer, we performed bioinformatics analyses (GO, pathway, KEGG, and network analyses) to identify differentially expressed genes and potential pathways. In this work, we systematically analyzed HOXA11-AS-related genes and their functional categorization, pathways and networks. Original data from TCGA was used to verify the relationships between the expression levels of HOXA11-AS and the de-regulated genes and clinicopathological parameters or patient survival. Based on the results, we speculated that HOXA11-AS may play an important role in NSCLC development and progression by regulating the expression of various pathways and genes, especially DOCK8 and TGF-beta pathway. However, the exact mechanism should be verified by functional experiments.

Abbreviations
LncRNAs: long noncoding RNAs; NSCLC: non-small cell lung cancer; GO: gene ontology; KEGG: Kyoto Encyclopedia of Genes and Genomes; MSigDB: molecular signatures database; TCGA: the cancer genome atlas; FDR: false discovery rate; BP: biological process; CC: cellular component; MF: molecular function; DAVID: database for annotation, visualization and integrated discovery.

Authors’ contributions
Y2 and RH participated in clinical data analysis and drafted the manuscript. YD and XW participated in the statistical analysis and corrected the manuscript. XZ and SH prepared for the specimens, carried out the siRNA isolation. WH, MJ and NG performed the statistical analysis, prepared for the figures and revised the manuscript. YX, PL, DL, GC and TG conceived of the study, participated in design and coordination and corrected the manuscript. All authors read and approved the final manuscript.

Author details
1 Department of Pathology, First Affiliated Hospital of Guangxi Medical University, No. 6 Shuangyong Road, Guangxi, Zhum Autonomous Region, Nanning 530021, People’s Republic of China. 2 Department of Medical Oncology, First Affiliated Hospital of Guangxi Medical University, No. 6 Shuangyong Road, Guangxi Zhum Autonomous Region, Nanning 530021, People’s Republic of China. 3 Department of Orthopedics, China-Japan Union Hospital of Jilin University, 2 Sendai Street, Changchun 130033, People’s Republic of China. 4 Department of Radiotherapy, First Affiliated Hospital of Guangxi Medical University, No. 6 Shuangyong Road, Guangxi Zhum Autonomous Region, Nanning 530021, People’s Republic of China.

Acknowledgements
The study was supported by a Fund of the Guangxi Provincial Health Bureau Scientific Research Project (Z2013201, Z2014055), a Fund of the National Natural Science Foundation of China (NSFC81360327, NSFC81560469), the Natural Science Foundation of Guangxi, China (2015GXNSFCA139009) and the Scientific Research Project of the Basic Ability Promoting for Middle Age and Youth Teachers of Guangxi Universities (KY2016YB077). The funders had no role in the study design, the data collection and analysis, the decision to publish, or the preparation of the manuscript.

Competing interests
The authors declare that they have no competing interests.

Availability of data and materials
Data sharing not applicable to this article as no datasets were generated or analysed during the current study.

Funding
Fund of the Guangxi Provincial Health Bureau Scientific Research Project (Z2015201, Z2014055) was used for the design of the study and collection. Fund of the National Natural Science Foundation of China (NSFC81360327, NSFC81560469) was used for data analysis, and the Natural Science Foundation of Guangxi, China (2015GXNSFCA139009) and the Scientific Research Project of the Basic Ability Promoting for Middle Age and Youth Teachers of Guangxi Universities (KY2016YB077) was used for interpretation of data and in writing the manuscript.

Received: 17 May 2016 Accepted: 23 November 2016
Published online: 01 December 2016

References
1. Xu YJ, Du Y, Fan Y. Long noncoding RNAs in lung cancer: what we know in 2015. Clin Transl Oncol. 2016;18(7):660–5.
2. Kang CG, Lee HJ, Kim SH, Lee EO. Zerumbone suppresses osteopontin-induced cell invasion through inhibiting the FAK/AKT/ROCK pathway in human non-small cell lung cancer A549 cells. J Nat Prod. 2016;79(1):156–60.
3. Torre LA, Bray F, Siegel RL, Ferlay J, Ervik M, Lortet‑Tieulent J, Jemal A. Global cancer statistics, 2012. CA A Cancer J Clin. 2015;65(2):87–108.
4. Chen G, Umelo IA, LV S, Teugels E, Foster K, Kronenberg P, Devaee A, Sadones J, Geers C, De Greve J. miR-146a inhibits cell growth, cell migration and induces apoptosis in non-small cell lung cancer cells. PLoS ONE. 2013;8(3):e60317.
5. Wilusz JE. Long noncoding RNAs: re-writing dogmas of RNA processing and stability. Bba-Gen Rel Mech. 2016;1859(1):128–38.
6. Yuan X, Wang J, Tang X, Li Y, Xia F, Gao X. Berberine ameliorates nonalcoholic fatty liver disease by a global modulation of hepatic miRNA and lncRNA expression profiles. J Trans Med. 2015;13:24.
7. Ponting CP, Oliver PL, Reik W. Evolution and functions of long noncoding RNAs. Cell. 2009;136(4):629–41.
8. Mercer TR, Dinger ME, Mattick JS. Long non-coding RNAs: insights into functions. Nat Rev Genet. 2009;10(3):155–9.
9. Wang KC, Zhang HY. Molecular mechanisms of long noncoding RNAs. Mol Cell. 2011;43(6):904–14.
10. He Y, Meng XM, Huang C, Wu BM, Zhang L, Lv XW, Li J. Long noncoding RNAs: novel insights into hepatocellular carcinoma. Cancer Lett. 2014;344(1):120–7.
11. Tafj R, Pang KC, Mercer TR, Dinger M, Mattick JS. Non-coding RNAs: regulators of disease. J Pathol. 2010;220(2):126–39.
12. Zeequin N, Xuemei Z, Wei L, Zongjuan M, Yueji Y, Yanli H, Yiping Z, Xia M, Wei W, Wenjing D, et al. The role and potential mechanisms of LncRNA-TATDN1 on metastasis and invasion of non-small cell lung cancer. Oncotarget. 2016;7(14):18219–28.
13. Wan L, Sun M, Liu JJ, Wei CC, Zhang EB, Kong R, Xu TP, Huang MD, Wang ZX. Long non-coding RNA PVT1 promotes non-small cell lung cancer cell proliferation through epigenetically regulating LAT52 expression. Mol Cancer Ther. 2016;15(5):1082–94.
14. Guo F, Guo L, Li Y, Zhou Q, Li Z. MALAT1 is an oncogenic long non-coding RNA associated with tumor invasion in non-small cell lung cancer regulated by DNA methylation. Int J Clin Exp Pathol. 2015;8(12):15903–10.
15. Richards EJ, Permuth-Wey J, Li Y, Chen YA, Coppola D, Reid BM, Lin HY, Teer JK, Berchuck A, Birrer MJ, et al. A functional variant in HOXA11-AS,
a novel long non-coding RNA, inhibits the oncogenic phenotype of epithelial ovarian cancer. Oncotarget. 2015;6(33):34745–57.

16. Wang Q, Zhang J, Liu Y, Zhang W, Zhou J, Duan R, Pu P, Kang C, Han L. A novel cell cycle-associated IncRNA, HOXA11-AS, is transcribed from the 5′-prime end of the HOXA transcript and is a biomarker of progression in glioma. Cancer Lett. 2016;373(2):251–9.

17. Ashburner M, Ball CA, Blake DA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, et al. Gene ontology: tool for the unification of biology. The gene ontology consortium. Nat Genet. 2000;25(1):25–9.

18. Kanehisa M, Goto S, Furumichi M, Tanabe M, Hirakawa M. KEGG for representation and analysis of molecular networks involving diseases and drugs. Nucleic Acids Res. 2010;38:D355–60.

19. Bornstein S, Schmidt M, Choonoo G, Levin T, Gray J, Thomas CR Jr, Waterman DD. Attenuation of the DNA damage response by transforming growth factor-β pathway. Mol Med Rep. 2016;13(4):3243–8.

20. Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA, Jacobsen A, Byrne CJ, Heuer ML, Larsson E, et al. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. Cancer Discov. 2012;2(5):401–4.

21. Gao J, Aksoy BA, Dogrusoz U, Drescher G, Grosse B, Sumer SO, Sun Y, Jacobsen A, Sinha R, Larson E, et al. Integrative analysis of complex cancer genomics and molecular pathways using the cBioPortal for Cancer Genomics. Sci Signal. 2013;6:pl1.

22. Gong X, Yi J, Carmon KS, Crumbley CA, Xiong W, Thomas A, Fan X, Guo S, An Z, Chang JT, et al. Aberrant RSP03-LGR4 signaling in Keap1-deficient lung adenocarcinomas promotes tumor aggressiveness. Oncogene. 2013;34(36):4692–701.

23. Drilon A, Sugita H, Sima CS, Zauderer M, Rudin CM, Kris MG, Rusch VW, Azooz CG. A prospective study of tumor suppressor gene methylation as a prognostic biomarker in surgically resected stage I to IIIA non-small cell lung cancers. J Thorac Oncol. 2014;9(9):1272–7.

24. Dunn JR, Panutopoulos D, Shaw MW, Heighway J, Dormer R, Salmo EN, Watson SG, Field JK, Liloglou T. MET-2 silencing and promoter hypermethylation in NSCLC. Br J Cancer. 2004;91(6):1149–54.

25. Wu W, Kemp BL, Proctor ML, Gazdar AF, Minna JD, Hong WK, Mao L. Expression of DMBT1, a candidate tumor suppressor gene, is frequently lost in lung cancer. Cancer Res. 1999;59(8):1846–51.

26. Takahashi K, Kohno T, Ajima R, Sasaki H, Minna JD, Fujiwara T, Tanaka N, Ajin Y, Wang SJ, Cui HY, et al. SPINK6 promotes metastasis of nasopharyngeal carcinoma via transforming growth factor-beta1. Tumour Biol. 2015;36(4):2657–65.

27. Zhang Q, Cui HY, Li Y, Zhao P, Zhang Y, Fu ZG, Chen ZN, Jiang JL. CD147 promotes Src-dependent activation of Rac1 signaling through STAT5/DOCK8 during the motility of hepatocellular carcinoma cells. Oncotarget. 2015;6(1):243–57.

28. Jin H, Oka F, Wang Y, Xu Y, Shang Y. Curcumin inhibits cell proliferation and induces apoptosis and autophagy via the PI3K/Akt/mTOR/p70S6K pathway in human non-small cell lung cancer cells. Biochem Biophys Res Commun. 2015;465(1):64–70.

29. Sun B, Hou C, Jiang Y, Zheng J, Yang N, Ji C, Liang Z, Shi J, Zhang R, Liu Q, Jin H, Qiao F, Wang Y, Xu Y, Shang Y. Curcumin inhibits cell proliferation and induces apoptosis of human non-small cell lung cancer cells. Biochem Biophys Res Commun. 2016;465(1):64–70.

30. Cheng T, Hu C, Yang H, Cao L, An J. Transforming growth factor-beta-induced mir143 expression in regulation of non-small cell lung cancer cell viability and invasion capacity in vitro and in vivo. Int J Radiat Oncol Biol Phys. 2015;91(1):91–9.

31. Cheng T, Hu C, Yang H, Cao L, An J. Transforming growth factor-beta-induced mir143 expression in regulation of non-small cell lung cancer cell viability and invasion capacity in vitro and in vivo. Int J Oncol. 2014;44(5):1597–601.

32. You L, Wu YQ, Chen J, Mao JH, Tetsu O, Li H, Jablons DM, You L. Correction: decreased TUSC3 promotes pancreatic cancer progression and epithelial-mesenchymal transition (EMT) through WNT/beta-catenin and MAPK signalling. J Pathol. 2016;239(1):60–71.

33. Feng N, Luo DH, Zou X, et al. ZSPINK6 promotes metastasis of nasopharyngeal carcinoma via binding and activation of epithelial growth factor receptor. Cancer Res. 2016. doi:10.1158/0008-5472.CAN-16-1281.

34. Zheng LS, Yang JP, Cao Y, Peng LX, Sun R, Xie P, Wang MY, Meng DF, Luo DH, Zou X, et al. Phosphatidylethanolamine-binding protein 4 promotes lung cancer cells proliferation and invasion via PI3K/Akt/mTOR axis. J Thorac Dis. 2015;7(10):1806–16.

35. Hindley CJ, Conduit AL, Menon V, Thomas R, Azmitia LM, Davis JA, Pruszak J. The Hippo pathway member YAP enhances human neural crest cell fate and migration. Sci Rep. 2016;6:23208.

36. Yajima V, Paulding C, Sondella R, McClatchey AI, Sato M, Wahner DC, Reynolds R, Bell DW, Lake R, van den Heuvel S, et al. DOCK4, a GTPase activator, is disrupted during tumorigenesis. Cell. 2003;112(5):673–84.

37. Moller N, Irani-Tehrani M, Kiosses WB, Del Pozo MA, Schwartz MA, Zizimin A, zmolin, a novel Cdc42 activator, reveals a new GEF domain for Rho proteins. Nat Cell Biol. 2002;4(9):639–47.

38. Cote JF, Vuori K. Identification of an evolutionarily conserved superfamily of DOCK180-related proteins with guanine nucleotide exchange activity. J Cell Sci. 2002;115(Pt 24):4901–13.

39. Rossman KL, Der CJ. GEF. J GEFS: going to tumour on RHO GTPases with guanine nucleotide-exchange factors. Nat Rev Mol Cell Biol. 2005;6(2):167–80.

40. Wang SJ, Cui HY, Li Y, Zhao P, Zhang Y, Fu ZG, Chen ZN, Jin H, Jiang JL. CD147 promotes Src-dependent activation of Rac1 signaling through STAT5/DOCK8 during the motility of hepatocellular carcinoma cells. Oncotarget. 2015;6(1):243–57.

41. Aydin SE, Klici SS, Aytekin C, Kumar A, Porras O, Kainulainen L, Kostyuchenko L, Genel F, Kutucukler N, Karaca N, et al. DOCK8 deficiency: clinical and immunological phenotype and treatment options—a review of 136 patients. J Clin Immunol. 2015;35(2):189–98.

42. Moller N, Irani-Tehrani M, Kiosses WB, Del Pozo MA, Schwartz MA, Zizimin A, zmolin, a novel Cdc42 activator, reveals a new GEF domain for Rho proteins. Nat Cell Biol. 2002;4(9):639–47.