



Integrative Analysis Reveals Enhanced Regulatory Effects of Human Long Intergenic Non-Coding RNAs in Lung Adenocarcinoma

You Zhou^a, Kai Wu^a, Jianping Jiang^b, Jinfei Huang^a, Peiwei Zhang^a, Yufei Zhu^a, Guohong Hu^a, Jingyu Lang^a, Yufang Shi^a, Landian Hu^a, Tao Huang^{a,*}, Xiangyin Kong^{a,*}

^a State Key Laboratory of Medical Genomics, Institute of Health Sciences, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences and Shanghai Jiao Tong University School of Medicine, Shanghai 200031, China

^b State Key Laboratory of Genetic Engineering, School of Life Sciences, Fudan University, Shanghai 200433, China

Received 13 April 2015; revised 29 June 2015; accepted 2 July 2015

Available online 10 July 2015

ABSTRACT

Although there is an accumulating appreciation of the key roles that long intergenic non-coding RNAs (lincRNAs) play in diverse cellular processes, our knowledge of how lincRNAs function in cancer remains sparse. Here, we present a comprehensive landscape of RNA-seq transcriptome profiles of lung adenocarcinomas and their paired normal counterparts to unravel gene regulation rules of lincRNAs. Consistent with previous findings of co-expression between neighboring protein-coding genes, lincRNAs were typically co-expressed with their neighboring genes, which was found in both cancerous and normal tissues. By building a mathematical model based on correlated gene expression, we distinguished an additional subset of lincRNAs termed “regulatory lincRNAs”, representing their dominant roles in gene regulation. The number of regulatory lincRNAs was significantly higher in cancerous compared to normal tissues, and most of them positively regulated protein-coding genes in *trans*. Functional validation, using knockdown, determined that regulatory lincRNA, GAS5, affected its predicted protein-coding targets. Moreover, we discovered hundreds of differentially expressed regulatory lincRNAs with inclusion of some cancer-associated lincRNAs. Our integrated analysis reveals enhanced regulatory effects of lincRNAs and provides a resource for the study of regulatory lincRNAs that play critical roles in lung adenocarcinoma.

KEYWORDS: Human long intergenic non-coding RNA; Gene regulation; Lung adenocarcinoma; RNA-seq

INTRODUCTION

Lung cancer is one of the leading causes of cancer mortality worldwide, accounting for over one million deaths per year (Jemal et al., 2011). Conventionally, lung cancer is classified into non-small cell lung cancer (NSCLC) and small cell lung cancer (SCLC) (Cagle and Allen, 2012). NSCLC is further assigned to three histological subtypes: large cell carcinoma (LCC), squamous cell carcinoma (SCC) and adenocarcinoma

(ADC), of which the last subtype is the most frequent and is attributed to more than half of all lung cancer cases (Rapp et al., 1988). Previous lung cancer studies intensively focused on the dysregulations of protein-coding genes as well as pathways to identify oncogenes, tumor-suppressor genes and other lung cancer-related genes which could serve as diagnostic targets for clinical therapy (Rikova et al., 2007; Fehring et al., 2012). To date, thousands of low-abundance long intergenic non-coding RNAs (lincRNAs) have been identified owing to technological advances (such as RNA sequencing) and further functional analysis has revealed that many of them play critical roles in triggering a wide spectrum of human cancers including lung cancer (Prensner and Chinnaiyan, 2011; Gutschner and Diederichs, 2012).

* Corresponding authors. Tel: +86 21 5492 0605, fax: +86 21 6467 8976 (X. Kong); Tel: +86 21 5492 3262, fax: +86 21 6467 8976 (T. Huang).

E-mail addresses: tohuangtao@126.com (T. Huang); xykong@sibs.ac.cn (X. Kong).

<http://dx.doi.org/10.1016/j.jgg.2015.07.001>

1673-8527/Copyright © 2015, Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, and Genetics Society of China. Published by Elsevier Limited and Science Press. All rights reserved.

Since the discovery of small regulatory RNAs now known as microRNAs (Wightman et al., 1993), which modulate approximately 30% of protein-coding genes and have provided a new perspective on the role of RNA in the regulation of gene expression, increasing studies have focused on the vast number of lincRNAs, which avoid overlap with other types of genes and participate in numerous cellular processes including epigenetic modification (*HOTAIR*) (Rinn et al., 2007), transcriptional regulation (*DLX6-AS1*) (Feng et al., 2006), and post-transcriptional regulation (*ZEB2-AS1*) (Beltran et al., 2008). In addition, a few well-studied examples indicate that lincRNAs can influence human diseases such as cancer by promoting tumor survival and metastasis (Calin et al., 2007; Gupta et al., 2010). For example, *CDKN1A-AS1* (Huarte et al., 2010) and *PCAT1* (Prensner et al., 2011) have been reported to modulate both oncogenic and tumor-suppressor pathways. The aberrant expression of several lincRNAs is associated with malignancy grade in human cancers, indicating their huge potential value in diagnosis and prognostication (Zhang et al., 2013b; Ziats and Rennert, 2013). Yet, in contrast to these well-described examples, only a limited number of lincRNAs have been labeled with documented roles in tumorigenesis (Matouk et al., 2007; Kotake et al., 2011) and even fewer have been implicated in lung adenocarcinoma. The most well-characterized lung cancer associated lincRNA is *MALAT1*, whose elevated level is found in NSCLC metastatic tumors and closely correlated with poor prognosis in early-stage patients (Ji et al., 2003; Gutschner et al., 2013). Recent studies have also revealed the role of *lincRNA-LET* in the regulation of hypoxia-induced metastasis in SCC patients (Yang et al., 2013a), the effect of intronic non-coding RNA *AK126698* on cisplatin resistance by targeting Wnt pathway in NSCLC cells (Yang et al., 2013b), and the up-regulation of *SCAL1* in numerous lung cancer cell lines (Thai et al., 2013). However, these individual studies are merely considered as the tip of the iceberg in regard to unraveling the precise mechanisms of lincRNA-mediated cancer biology, and the principal properties and gene regulation functions of most lincRNAs remain poorly determined.

In this study, we harnessed the unbiased view of the transcriptome offered by next-generation sequencing platforms to explore the roles that lincRNAs play in the regulation of gene expression in lung adenocarcinomas and their paired normal counterparts from 72 Korean patients (Seo et al., 2012). The well-defined lincRNA database we used was produced by the GENCODE consortium comprising 11,904 human lincRNA transcripts (Derrien et al., 2012). We first estimated the impacts of lincRNAs on the correlated expression of nearby genes, and found that they were typically co-expressed with their neighboring genes in both cancerous and normal tissues. Then, we built a mathematical model based on the correlated gene expression patterns to predict a subset of lincRNAs termed “regulatory lincRNAs” and classified their target protein-coding genes. Our computational predictions showed a sharp rise in the number of regulatory lincRNAs in cancerous tissues compared to normal counterparts, which was also validated in an independent cohort of lung tissues

from The Cancer Genome Atlas (TCGA). More than 90% of the regulatory lincRNAs functioned in *trans*. One of the most widely expressed regulatory lincRNAs, *GAS5*, was functionally determined to affect its predicted protein-coding targets *CPS1* and *AKR1C2*. Additionally, we observed hundreds of lincRNAs with altered roles in normal cells compared to cancerous cells as well as some differentially expressed regulatory lincRNAs. In conclusion, our integrated analysis uncovered enhanced regulatory effects of lincRNAs in lung adenocarcinoma and identified a class of regulatory lincRNAs that may serve as novel biomarkers and targets for clinical therapy.

RESULTS

Characterization of lincRNAs

To investigate the functions of lincRNAs in human lung adenocarcinoma, we first annotated all of the lincRNAs identified in our 72 tumor and corresponding normal tissues using GENCODE v19 (Derrien et al., 2012) set comprised of 11,904 human lincRNA transcripts. We only selected transcripts (both coding and non-coding) expressed at least in 20% ($n = 15$) of normal or tumor samples in order to exclude the difference among individuals and ensure the robustness of gene expression. On average, each tumor sample showed 547 expressed lincRNA transcripts, which was comparable with those in each normal sample (561) (Fig. S1A, $P = 0.2495$, one-sided pairwise Student's *t*-test). For each transcript or gene, the number of reads aligned to it (raw read count) was normalized by FPKM (fragments per kilobase per million mapped reads) (Trapnell et al., 2010), such that a higher FPKM represented a higher expression level. Three categories were defined for further analysis, namely “normal specific lincRNAs”, which were detected only in normal samples, “tumor specific lincRNAs”, which were detected only in tumor samples, and “shared lincRNAs”, which were detected in both, and their numbers were 126, 272, and 712, respectively. We also calculated the ratio of the number of lincRNA transcripts divided by the number of total expressed genes (FPKM > 0) in each sample. Similarly, there was no difference between normal and tumor samples (Fig. S1B, $P = 0.5353$, one-sided pairwise Student's *t*-test). Detailed information for all of the lincRNA transcripts found in our samples is listed in Table S1.

Next, we analyzed the properties of lincRNAs detected in our samples, comparing them with protein-coding genes when appropriate. From the view of transcript lengths distribution, normal and tumor specific lincRNAs were shorter than shared lincRNAs (Fig. 1A, $P < 0.001$, one-sided Wilcoxon signed-rank test). Consistent with prior research (Cabili et al., 2011), lincRNAs had a lower coding potential (Fig. 1B, $P < 2.2 \times 10^{-16}$, one-sided Wilcoxon signed-rank test) and fewer exons relative to protein-coding genes (Fig. 1C). Moreover, the expression levels of lincRNAs were lower than those of protein-coding genes (Fig. 1D, $P < 2.2 \times 10^{-16}$, one-sided Wilcoxon signed-rank test), with an approximately 10-

Download English Version:

<https://daneshyari.com/en/article/2787441>

Download Persian Version:

<https://daneshyari.com/article/2787441>

[Daneshyari.com](https://daneshyari.com)