A three long non-coding RNA signature to improve survival prediction in patients with Wilms' tumor

PING REN and MEILING HU

1Department of Medicine, The Affiliated Hospital of Ningbo University School of Medicine, Ningbo, Zhejiang 315020; 2Department of Surgery, Cixi People's Hospital of Zhejiang, Ningbo, Zhejiang 315300, P.R. China

Received March 22, 2019; Accepted September 17, 2019

DOI: 10.3892/ol.2019.10990

Abstract. The prognostic significance of long non-coding RNAs (lncRNAs) in Wilms' tumor (WT) remains unclear. The present study identified differential lncRNA expression patterns between WT and normal tissues; and evaluated the prognostic value of these lncRNAs. Multidimensional data from 136 samples were retrieved from the Therapeutically Applicable Research to Generate Effective Treatments (TARGET) database. Computational data analyses were performed to detect survival-associated molecular signatures. Kaplan-Meier survival curves revealed the prognostic prediction values for three of these lncRNAs, namely DLGAP1 antisense RNA 2 (DLGAP1-AS2), RP11-93B14.6 and RP11554F20.1. Cox regression analysis revealed that the three-lncRNA signature was significantly associated with patient survival (P<0.01). Functional enrichment analyses suggested that the target genes of these three lncRNAs may be involved in known cancer-associated pathways. The present study revealed a three-lncRNA signature consisting of DLGAP1-AS2, RP11-93B14.6 and RP11554F20.1 that may be used as a prognostic marker for patients with WT.

Introduction

Wilms tumor (WT) is the most common type of renal cancer in children between the ages 0 and 14 years, with an incidence rate of 9 in 1 million children, based on November 2017 SEER data submission (1,2). As such, close attention is being paid to this major global public health issue (3,4). The incidence of WT in children <15 years of age in the United States of America is ~7-8 cases per million, with ~500 new cases per year (5). Although the prognosis of children with WT has improved due to multimodal therapies, WT continue to recur even five years post-diagnosis (6-8). Thus, considering the disease's severity, the early diagnosis of WT and investigation of the molecular mechanisms associated with the development of the disease are of great importance. Effective biomarkers for the early detection of WT are urgently needed to improve the quality of life and survival of patients with WT.

Technological advances in genomic and transcriptomic analysis have led to the identification of various types of non-coding RNAs (ncRNAs) (9,10). Long non-coding RNAs (lncRNAs), which consist of >200 nucleotides, account for a large number of ncRNAs (11). Emerging evidence indicates that lncRNAs are important in genetic, post-transcriptional and epigenetic regulation (12,13). In addition, a growing number of studies suggest that lncRNAs are involved in cell proliferation, migration and apoptosis (14,15). Therefore, lncRNAs may serve as potential diagnostic and prognostic biomarkers for WT (16,17). However, knowledge of the association between lncRNAs and prognosis in patients with WT is limited (18). The present study investigated differential lncRNA expression patterns between WT and normal tissues, and revealed a three-lncRNA signature that may predict the survival time of patients with WT.

Materials and methods

Acquisition of therapeutically applicable research to generate effective treatments(TARGET) data. Raw lncRNA expression data and corresponding clinical information were downloaded from the TARGET database (ocg.cancer.gov/programs/target) Version: December, 2018. Samples obtained from patients with an overall survival (OS) of more than one month were included in the present study. Additionally, the patient clinical information, differentially expressed lncRNAs and prognosis information were downloaded. A total of 136 samples, including 130 WT tissues and six normal
tissues, were investigated in the present study. The detailed clinical characteristics of the patients are presented in Table I. R language package (edgeR, Release 3.9; http://www.bioconductor.org/packages/release/bioc/html/edgeR.html) was used to interpret lncRNA sequencing data and the limma package (Release 3.9; http://www.bioconductor.org/packages/release/bioc/html/limma.html) was used to analyze differentially expressed lncRNAs between WT and normal tissues. Fold-changes (FCs) in the expression of individual lncRNAs were calculated, and differentially expressed lncRNAs with log2|FC| >1.0 and P<0.05 were considered significant.

Patient prognosis and target gene prediction. Differentially expressed lncRNA profiles were normalized following log2 transformation. The Kaplan-Meier method and the log-rank test were used to evaluate the prognostic value of each differentially expressed lncRNA. Three lncRNAs that were significantly associated with OS were identified as prognostic lncRNAs and were subjected to receiver operating characteristic (ROC) and Cox regression analyses. The Cox model was used to investigate the association between the expression of each lncRNA and OS according to age, ethnicity, gender and disease stage. lncRNAs with hazard ratios (HRs) <1 were defined as protective, while those with HRs >1 were defined as high-risk. Eventually, a prognostic lncRNA signature was constructed, and patients with WT were classified into low- and high-risk groups using the median risk score as the cut-off value. Kaplan-Meier analysis and log-rank test were used to evaluate differences in patient survival between the two groups. ROC analysis was performed to compare the sensitivity and specificity of the survival prediction based on the lncRNA risk score. Protein-coding genes (PCGs) genes correlated with the differentially expressed lncRNAs were identified using the co-expression method. The PCGs with a Pearson's correlation coefficient >0.40 and P<0.01 were considered to be associated with the lncRNAs. Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway (https://david.ncifcrf.gov/) and Gene Ontology (GO; https://david.ncifcrf.gov/) enrichment analyses were subsequently performed for the target genes. P<0.05 were considered to indicate a statistically significant difference.

Statistical analysis. Univariate and multivariate Cox proportional hazard regression and Kaplan-Meier survival analyses with log-rank test were used to compare each lncRNA (low vs. high expression level) and their prognostic signatures (low vs. high risk). Hazard ratio (95% CI) was expressed in the Cox regression analysis. The data for categorical variables were presented as percentages (%). Pearson's correlation test and log-rank test were used in this study. P<0.05 was considered to indicate a statistically significant difference. Statistical analyses were performed using R (version 3.5.1; www.r-project.org) (19) and SPSS software (version 17.0; SPSS Inc.).

Results

Characteristics of the differentially expressed lncRNAs. The present study investigated 136 samples, including 130 WT and 6 normal tissues. The detailed clinical characteristics, including gender, ethnicity, age at diagnosis and disease stage, are presented in Table I. A total of 1,833 differentially expressed lncRNAs, including 1,087 upregulated and 746 downregulated lncRNAs, were identified between WT and normal tissues (Fig. 1).

Association between the three lncRNAs and OS in patients with WT. The Kaplan-Meier method and log-rank test were used to evaluate the association between OS and lncRNA expression patterns. As depicted in Fig. 2, two upregulated lncRNAs (RP11-93B14.6 and RP11554F20.1) and one downregulated lncRNA (DLGAP1-AS2), were significantly associated with OS rate (all P<0.05).

Prognostic value of the three-lncRNA signature risk scores for WT. A prognostic signature was identified by integrating three
lncRNA expression profiles and the corresponding estimated regression coefficient. The 130 patients investigated in the present study were subsequently divided into high- and low-risk groups (n=65 per group) according to the median risk score. Patients in the high-risk group had a significantly worse OS than those in the low-risk group (P<0.001; Fig. S1). The prognostic ability of the three-lncRNA signature was evaluated using ROC curve analysis. Results revealed that the area under the ROC curve of the three-lncRNA signature was 0.80 (Fig. S2).

Univariate and multivariate Cox regression analyses were used to determine the effects of the three-lncRNA signature (high vs. low risk) on OS (Tables II and III). In univariate models, gender (HR=1.94; P=0.015), disease stage (HR=2.65; P<0.001), DLGAP1-AS2 signature (HR=0.60; P<0.001), RP11-93B14.6 signature (HR=1.37; P<0.001) and RP11-554F20.1 signature (HR=1.92; P<0.001) were all independent factors associated with OS in patients with WT. In multivariate models, gender (HR=3.10; P<0.001), stage (HR=3.36; P<0.001), DLGAP1-AS2 signature (HR=0.60; P=0.001), RP11-93B14.6 signature (HR=1.30; P=0.006) and RP11-554F20.1 signature (HR=2.26; P<0.001) were associated with OS in patients with WT.

Patients in the low-risk group expressed increased levels of the protective lncRNA (DLGAP1-AS2) compared with patients in the high-risk group (P<0.05; Fig. 3). Patients in the high-risk group expressed increased levels of RP11-93B14.6 and RP11-554F20.1 compared with the low-risk group (P<0.05; Fig. 3). In addition, patients in the low-risk group had a significantly longer survival time than those in the high-risk group (P<0.05; Fig. 3).

lncRNA target prediction and functional GO and KEGG analysis. A co-expression method was used to predict the potential target of the three lncRNAs of interest. GO enrichment and KEGG pathway analyses were performed to elucidate the biological functions of the associated target genes (Fig. 4). GO biological process, molecular function and cellular component terms were mainly enriched in ‘coenzyme binding’, ‘cofactor
In addition, KEGG pathways were significantly enriched in 'fatty acid metabolism', 'histidine metabolism', 'peroxisome', 'AGE-RAGE signaling pathway in diabetic complication' and 'insulin resistance' (P<0.05).

Discussion

WT is currently the most common primary malignant renal tumor in children (20). Although outcomes have improved due to multidisciplinary treatments, chemotherapeutic drugs and radiotherapy, the incidence and recurrence of WT remains high, and presents a heavy burden for patients, families and medical institutions (21-23). The prognosis of patients with WT may be considerably improved if the characteristics of the tumor, including clinical symptoms and genetic phenotypes, are reliably predicted at the time of initial diagnosis (18). Therefore, there is a requirement for the identification of prognostic biomarkers as well as the investigation of the molecular mechanisms underlying the development of WT.

A number of previously studies have revealed that genetic factors may contribute to the development of WT (2,18). Furthermore, studies have indicated that altered lncRNA expression levels are associated with disease development; however, their prognostic values have not been thoroughly investigated (18,24). Therefore, the present study identified three differentially expressed lncRNAs, which were correlated with OS in patients with WT. The lncRNAs, including DLGAP1-AS2, RP11-93B14.6 and RP11-554F20.1, were validated as independent prognostic factors for patients with WT. Additionally, the target genes of the three lncRNAs, as well as their enrichment binding', 'flavin adenine dinucleotide binding', 'acetylgalactosaminyltransferase activity' and 'steroid binding' (P<0.05). In addition, KEGG pathways were significantly enriched in 'fatty acid metabolism', 'histidine metabolism', 'peroxisome', 'AGE-RAGE signaling pathway in diabetic complication' and 'insulin resistance' (P<0.05).
Figure 3. Three lncRNAs based risk score distribution, patients' event-free survival time and a heatmap of the expression profiles of the three lncRNA. Red dots represent the high-risk group and green dots represent the low-risk group.

Figure 4. (A) The Gene Ontology and (B) Kyoto Encyclopedia of Genes and Genomes pathway enrichment analyses.
pathways and biological functions, were investigated using bioinformatics. The results suggested that these three lncRNAs may participate in the molecular pathogenesis, clinical progression and prognosis of WT, highlighting the potential of lncRNA profiling to improve the clinical prognosis in patients with WT.

Previous studies have demonstrated that functional lncRNA expression plays an important role in carcinogenesis and strongly correlates with gene expression and pathway regulation (25,26). lncRNAs promote tumor formation, progression and metastasis by regulating various major pathways, including cancer cell proliferation, cell cycle arrest, apoptosis, differentiation, adhesion, migration, invasion and survival (27,28). Zhu et al (24) reported that upregulated expression of long intergenic non-protein coding RNA 473 was associated with the molecular pathogenesis of WT via the miR-195/IKK complex α signaling pathway. However, the sample size, sample types and number of lncRNAs assessed were limited.

The present study analyzed high-throughput data and identified two upregulated lncRNAs (RP11-93B14.6 and RP11554F20.1) and one downregulated lncRNA (DLGAPI-AS2) associated with the clinical outcomes of patients with WT. Therefore, the expression levels of these three lncRNAs may serve as prognostic markers for WT. Furthermore, to the best of our knowledge, the functions of these lncRNAs have not been previously investigated. The results of the present study revealed that RP11-93B14.6, RP11554F20.1 and DLGAPI-AS2 were enriched in ‘coenzyme binding’, ‘cofactor binding’, ‘fatty acid metabolism’, ‘histidine metabolism’, ‘peroxisome’, ‘AGE-RAGE signaling pathway in diabetic complication’ and ‘insulin resistance’, all of which are classical signaling pathways closely associated with tumorigenesis and the progression of malignancies (29,30). However, further molecular investigations in patients with WT are required to validate these results and to inform new therapeutic interventions.

The present study had several limitations. Firstly, the mechanisms underlying the prognostic values of the three lncRNAs were not investigated and warrant experimental studies. Secondly, a single published dataset with a small sample size was used. Therefore, the results obtained may differ from the general population. Thirdly, this study is not an intervention and treatment experiment, therefore data on the therapeutic effect of WT cannot be obtained. Finally, the WT cohort had a relatively high censor rate, which may have affected the reliability of the Kaplan-Meier estimates. Thus, a larger and multicenter clinical cohort study is required to validate the results obtained in the present study.

Acknowledgements

Not applicable.

Funding

No funding was received.

Availability of data and materials

The datasets generated and/or analyzed during the present study are available in the TARGET repository, https://cancer.gov/programs/target.

Authors' contributions

PR contributed to the conception, design and final approval of the submitted version. MH contributed to the interpretation of data and completion of figures and tables. Both authors have read and approved the final version of the manuscript.

Ethics approval and consent to participant

The study was performed in accordance with the guidelines of the Declaration of Helsinki. The study protocol was approved by the TARGET publication guidelines.

Patient consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

References


