Identification and Validation of a Four-Long Non-coding RNA Signature Associated With Immune Infiltration and Prognosis in Colon Cancer

Yanbo Wang[1], Jing Liu[2], Fenghai Ren[1], Yanjie Chu[2] and Binbin Cui[3]*

[1]Department of Thoracic Surgery, Harbin Medical University Cancer Hospital, Harbin, China, [2]Department of Gastroenterology and Hepatology, The Second Affiliated Hospital of Harbin Medical University, Harbin, China, [3]Department of Colorectal Surgery, Harbin Medical University Cancer Hospital, Harbin, China

The emerging evidence has demonstrated the critical roles of long non-coding RNAs (lncRNAs) as regulators in the tumor immune microenvironment (TIME). However, the tumor immune infiltration-associated lncRNAs and their clinical significance in colon cancer have not yet been thoroughly investigated. This study performed an integrative analysis of lncRNA expression profiles and immune cell infiltration profiles and identified 258 immune infiltration-associated lncRNAs. Of them, four lncRNAs (AC008494.3, LINC00926, AC022034.1, and SNHG26) were significantly and independently associated with the patient's overall survival. Finally, we developed a tumor immune infiltration-associated lncRNA signature (TIILncSig) comprising of these four lncRNAs, which can divide colon cancer patients of The Cancer Genome Atlas (TCGA) into high-risk and low-risk groups with a significantly different outcome [Hazard ratio (HR) = 2.718, 95% CI = 1.955–3.779, $p < 0.001$]. Prognostic performance of the TIILncSig was further validated in another independent colon cancer cohort (HR = 1.832, 95% CI = 1.045–3.21, $p = 0.034$). Results of multivariate Cox regression and stratification analysis demonstrated that the TIILncSig is an independent predictive factor from other clinical features (HR = 2.687, 95% CI = 1.912–3.776, $p < 0.001$ for TCGA cohort and HR = 1.837, 95% CI = 1.047–3.223, $p = 0.034$ for GSE17538 cohort). Literature analysis provided experimental evidence supporting roles of the TIILncSig in cancer carcinogenesis and progression and immune regulation. Summary, our study will help to understand the mechanisms of lncRNAs in immune regulation in the tumor microenvironment and provide novel biomarkers or targets for prognosis prediction and therapy decision-making for patients with colon cancer.

Keywords: long non-coding RNAs, colon cancer, immune microenvironment, tumor immune infiltration, integrative analysis

INTRODUCTION

Colorectal cancer (CRC) is the most common type of gastrointestinal cancer and is the second leading cause of cancer-related death (Siegel et al., 2020a,b). CRC cases were classified into colon tumors and rectum tumors according to codes from the International Classification of Diseases for Oncology (Siegel et al., 2020a). It is estimated to be 104,610 individuals newly diagnosed with colon cancer in the United States in 2020 (Siegel et al., 2020a). The tumor, lymph node, metastasis (TNM) staging system, histological differentiation grade, and tumor sidedness have been widely used for colon cancer classification, prognosis prediction, and therapy decision-making. Surgical removal of the primary tumor followed by adjuvant chemotherapy is the main treatment for colon cancer patients. However, in clinical practice, the TNM staging system was not enough to predict prognosis and make therapeutic decisions for the patient with colon cancer patients. There is growing evidence that molecular biomarkers have become indispensable in the more personalized precision medicine era (Vacante et al., 2018; Koncina et al., 2020).

Increasing evidence has suggested that the tumor immune microenvironment (TIME) plays either tumor-promoting or tumor-suppressive roles in various cancers and profoundly influences tumor prognosis and therapy response (Chew et al., 2012; Binnewies et al., 2018). The tumor microenvironment of CRC has been characterized using different methods in several studies and highlighted its prognostic, predictive, and therapeutic implications. For example, tumor-infiltrating lymphocytes (TILs), an essential component of TIME, have been recognized as an essential histopathologic feature associated with the prognosis of patients with colorectal cancer (Idos et al., 2020). Recent studies have found that transcriptomic or epigenetic features are associated with specific immune cell subpopulations of TIME. They, therefore, could be used to infer the composition of tumor-infiltrating immune cells (Zhang et al., 2020b). Long non-coding RNAs, constituting the major class of non-coding RNAs (ncRNAs), have been recognized as an essential regulator involved in nearly all biological progress. The emerging roles of lncRNAs as essential regulators of the human immune system have been recognized in recent studies (Turner et al., 2014; Chen et al., 2017). Increasing evidence suggested that lncRNAs played vital roles in the development, differentiation, activation of different immune cells and contributed to the modulation of innate and adaptive immunity. Furthermore, lncRNAs could act as communicators and mediators between the tumor microenvironment and cancer cells, highlighting their potential as immunotherapy targets and biomarkers. For example, Sun et al. (2020) identified 57 tumor immune infiltration-associated lncRNAs in non-small cell lung cancer and found that seven of them are associated with patient's survival and response to immune checkpoint inhibitor (ICI) immunotherapy. Zhou et al. (2020) developed a lncRNA signature of tumor-infiltrating B lymphocytes that are predictive of prognosis and immunotherapy response in bladder cancer. Although continued efforts are coming, the tumor immune infiltration-associated lncRNAs and their clinical significance in colon cancer have not yet been thoroughly investigated.

In this study, we performed an integrative analysis of lncRNA expression profiles, immune cell infiltration profiles, and clinical profiles to infer tumor immune infiltration-associated lncRNAs in colon cancer and explored their value in predicting prognosis.

MATERIALS AND METHODS

Clinical and Transcriptomic Data of Colon Cancer Patients

Clinical and transcriptomic data of 512 colon cancer patients were obtained from UCSC Xena Browser (GDC TCGA Colon Cancer cohort).[1] The RNA-seq gene expression level 3 data [log2(RPKM +1) transformed] were used for this study. After removed those patients with overall survival <30 days, a total of 419 was left for further study. Clinical and transcriptomic data of 219 colon cancer patients were derived from Gene Expression Omnibus (GEO) database.[2] Microarray data profiled by Affymetrix Human Genome U133 Plus 2.0 Array was used and then background correction, quantile normalization using the R package "affy."

Acquisitions of lncRNA Expression Profiles of Colon Cancer Patients

Approved human lncRNA information was obtained from HUGO Gene Nomenclature Committee (HGNC).[3] After cross-referenced and removed lncRNAs with 0 value in more than 20% of samples, a total of 3,959 lncRNAs were retained for RNA-seq data. For microarray data, a total of 5,919 lncRNAs were retained through repurposing array probes into the human genome and HGNC database. After cross-validation analysis among different cohorts based on different platforms, 2,235 overlapped lncRNA among different platforms were kept for further analysis.

Calculation of Infiltrating Immune Cell Abundance in the Tumor Microenvironment

The relative proportions of 22 types of infiltrating immune cells in the tumor microenvironment were inferred using the CIBERSORT algorithm[4] with the default signature matrix at 500 permutations (Newman et al., 2015).

Statistical Analysis

The association between lncRNA expression and the abundance of infiltrating immune cells was measured using the Pearson correlation coefficient (PCC). The association between lncRNA expression and overall survival were evaluated using the univariate and multivariate Cox regression analysis. The lncRNA-based

---

[1]http://xena.ucsc.edu/
[2]https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE17536
[3]https://www.genenames.org/
[4]http://cibersort.stanford.edu/

scoring model was developed using the linear combination of lncRNA biomarkers' transformed expression values with the multivariate Cox regression coefficient as the weight. The differences in overall survival between the high-risk and low-risk groups were assessed using the Kaplan-Meier survival curves and log-rank tests. Hazard ratios (HR) and 95% CI were calculated. The time-dependent receiver operating characteristic (ROC) curve and area under the time-dependent ROC curve (AUC) was calculated for assessing the predictive performance for survival at 3- and 5-years. All statistical analyses were conducted with R software and Bio-conductor (version 3.6.3).

RESULTS

Identification of Prognostic Tumor Immune Infiltration-Associated lncRNAs in Colon Cancer

We first inferred infiltrating levels of 22 immune cell types in the tumor microenvironment through the CIBERSORT algorithm based on the RNA-seq data of patients in the TCGA cohort. Then, we measured the association of lncRNAs with different infiltrating immune cells by calculating the PCC between expression levels of each lncRNA and infiltrating abundance of each infiltrating immune cells in the TME. Finally, we identified 258 lncRNAs that are highly correlated with infiltrating abundance of at least one of 22 immune cell types (**Supplementary File 1**), as shown in **Figure 1A**. To determine the prognostic value of these 258 immune infiltration-associated lncRNAs, we performed univariate Cox regression analysis to examine the association between expression levels of these 258 immune infiltration-associated lncRNAs and overall survival of patients in the TCGA cohort, and found that 29 of 258 immune infiltration-associated lncRNAs are significantly associated with overall survival as shown in **Figure 1B**. To further investigate the independence of these 29 prognostic immune infiltration-associated lncRNAs in predicting overall survival, we then conducted multivariate Cox regression analysis on the expression level of 29 prognostic immune infiltration-associated lncRNAs with overall survival as a dependent variable and other individual clinical features as explanatory variables, and identified four of 29 prognostic immune infiltration-associated lncRNAs as independent prognostic factors that still maintained a significant association with overall survival after adjusted by other clinical features as shown in **Figure 1C** (**Table 1**).

Development and Evaluation of a Tumor Immune Infiltration-Associated lncRNA Signature in Predicting Survival

To construct a tumor immune infiltration-associated lncRNA signature to predict overall survival, we fitted these four independent prognostic lncRNAs in a multivariate Cox regression model with overall survival as a dependent variable to measure relative contributions. We then developed a tumor immune infiltration-associated lncRNA signature (TIILncSig) by a linear combination of expression levels of these four independent prognostic lncRNAs, weighted by the corresponding coefficient derived from above multivariate analysis according to previous studies, as follows: TIILncSig = expression(AC008494.3)*(−2.334119617)+expression(LINC00926)*0.289309684+expression(AC022034.1)*0.283336978+expression(SNHG26)*0.579001809. Then patients of the TCGA cohort were assigned a risk score based on TIILncSig and subsequently were classified into high-risk groups ($n = 210$) and low-risk groups ($n = 209$) according to the median value of risk score. As shown in **Figure 2A**, patients in the high-risk group were observed to have significantly poor overall survival than those in the low-risk group (log-rank test $p < 0.001$; **Figure 2A**). The three- and five-survival rates of patients in the high-risk group were 73.9 and 58.1%, respectively, whereas the corresponding rates of patients in the low-risk were 86.5 and 72.9%, respectively. Furthermore, the TIILncSig achieved AUC values of 0.659 and 0.584 in predicting 3 and 5-years survival (**Figure 2B**). The distribution of TIILncSig-based risk score, the survival status and expression pattern of prognostic lncRNAs of patients in the TCGA cohort were plotted. As shown in **Figure 2C**, more deaths tended to be enriched in the high-risk group relative to the low-risk group. Moreover, AC008494.3 tended to be a protective factor whose high expression is associated with low risk. On the contrary, the other three lncRNAs (LINC00926, AC022034.1, and SNHG26) tended to be risk factors whose high expression is associated with high risk. As shown in **Figure 2D**, three risky lncRNAs (LINC00926, AC022034.1, and SNHG26) revealed significantly higher expression in the high-risk groups compared to in the low-risk group, and one protective lncRNA AC008494.3was observed to be expressed at low levels in the high-risk group compared to in the low-risk group (**Figure 2D**).

Independent Validation of the Prognostic Performance of the TIILncSig

To further confirm the robustness and reliability of the TIILncSig in predicting survival, we further applied the TIILncSig into another independent GEO COAD cohort (GSE17538). The TIILncSig classified 219 patients of the GSE17538 cohort into the high-risk group ($n = 110$) and low-risk group ($n = 109$) with obvious different overall survival (log-rank test $p = 0.08$; **Figure 3A**). As shown in **Figure 3A**, the overall survival of patients in the high-risk group patients was shorter than that of patients in the low-risk group patients. The three- and five-survival rates of patients in the high-risk group were 67.3 and 58.4%, respectively, whereas the corresponding rates of patients in the low-risk group were 77.7 and 63.0%. Furthermore, the TIILncSig achieved AUC values of 0.578 and 0.542 in predicting 3 and 5-years survival (**Figure 3B**). The distribution of TIILncSig-based risk score, the survival status and expression pattern of prognostic lncRNAs of patients in the independent GSE17538 cohort were similar to those observed in the TCGA cohort. As shown in **Figures 3C,D**, higher expression of three risky lncRNAs (LINC00926, AC022034.1, and SNHG26) and a lower expression of one protective lncRNA (AC008494.3) were observed in patients of the high-risk group compared to those in the low-risk group.
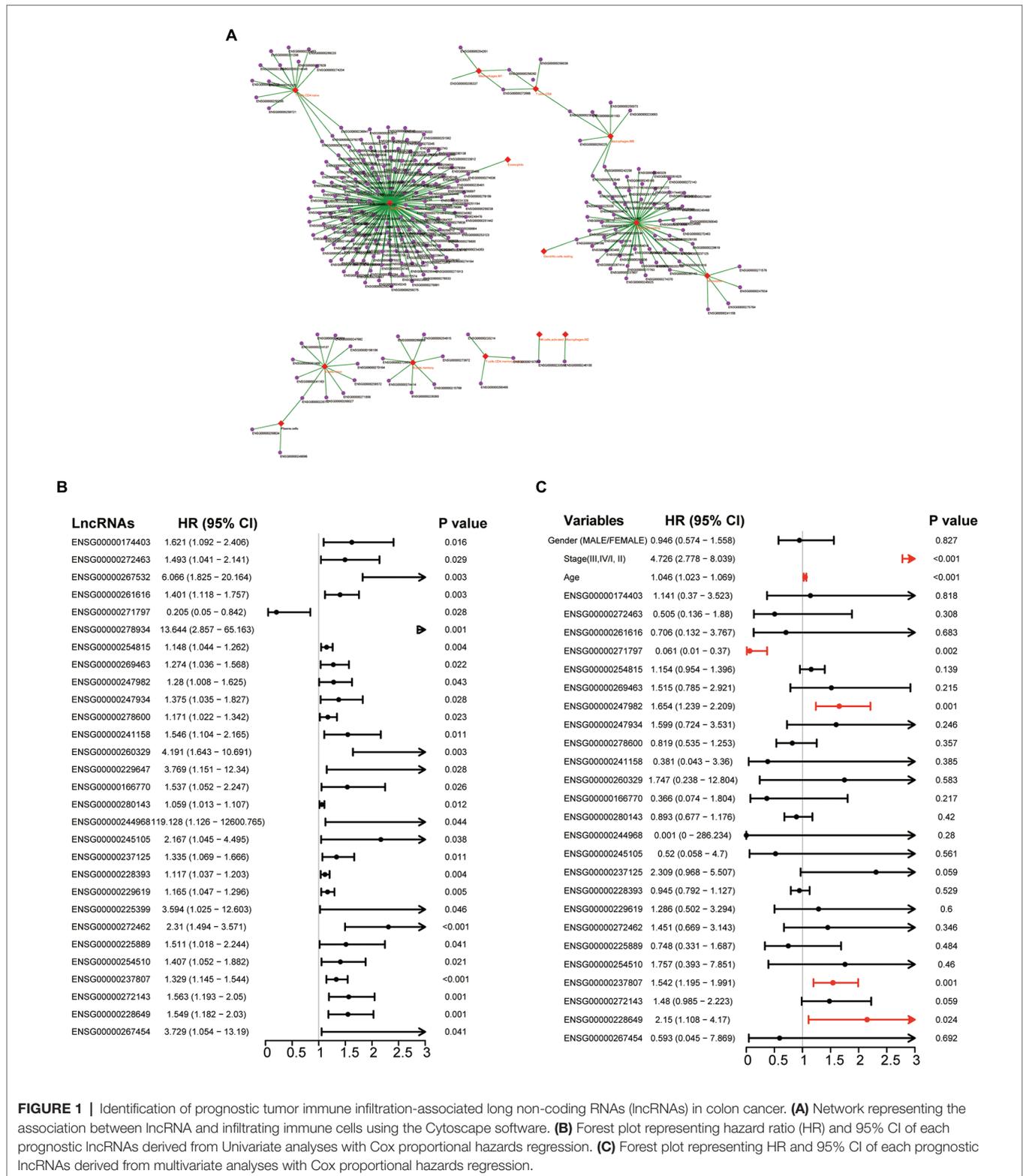
FIGURE 1 | Identification of prognostic tumor immune infiltration-associated long non-coding RNAs (lncRNAs) in colon cancer. **(A)** Network representing the association between lncRNA and infiltrating immune cells using the Cytoscape software. **(B)** Forest plot representing hazard ratio (HR) and 95% CI of each prognostic lncRNAs derived from Univariate analyses with Cox proportional hazards regression. **(C)** Forest plot representing HR and 95% CI of each prognostic lncRNAs derived from multivariate analyses with Cox proportional hazards regression.

Independence of the TIILncSig From Other Clinical Features

Results of the univariate analysis showed that although the TIILncSig is significant with overall survival in the TCGA cohort (HR = 2.718, 95% CI = 1.955–3.779, $p < 0.001$) and GSE17538 cohort (HR = 1.832, 95% CI = 1.045–3.21, $p = 0.034$), the stage is also significant in the univariate analysis (HR = 2.518, 95% CI = 1.658–3.824, $p < 0.001$ for TCGA cohort HR = 3.695,

TABLE 1 | The detailed information of four prognostic tumor immune infiltration-associated lncRNAs in colon cancer.

Ensemble ID	Gene name	Genomic location	HR	95% CI	Coef	p-value
ENSG00000271797	AC008494.3	Chr5: 115,262,505–115,263,448(+)	0.205	0.05–0.842	−1.584	0.028
ENSG00000247982	LINC00926	Chr15: 57,300,365–57,307,769(+)	1.28	1.008–1.625	0.247	0.043
ENSG00000237807	AC022034.1	Ch8: 53,493,523–53,524,336(−)	1.329	1.145–1.544	0.285	<0.001
ENSG00000228649	SNHG26	Chr7: 22,854,126–22,872,945(+)	1.549	1.182–2.03	0.438	0.001

FIGURE 2 | Performance evaluation of the tumor immune infiltration-associated lncRNA signature (TIILncSig) for predicting overall survival in the TCGA cohort. **(A)** Kaplan-Meier survival curves of overall survival between high- and low-risk groups. **(B)** The time-dependent receiver operating characteristic (ROC) analysis of the sensitivity and specificity for survival prediction by the TIILncSig. **(C)** The risk score distribution, survival status, and expression pattern of the TIILncSig. **(D)** Boxplot representing expression differences of lncRNAs between high- and low-risk groups.

95% CI = 2.189–6.236, p < 0.001 for GSE17538 cohort; **Figure 4A**). Therefore, we further performed multivariate Cox regression analyses with overall survival as the dependent variable and the TIILncSig and other clinical features as explanatory variables in each cohort to investigate whether the prognostic performance of the TIILncSig is independent of

FIGURE 3 | Performance validation of the TIILncSig for predicting overall survival in the GSE17538 cohort. **(A)** Kaplan-Meier survival curves of overall survival between high- and low-risk groups. **(B)** The time-dependent ROC analysis of the sensitivity and specificity for survival prediction by the TIILncSig. **(C)** The risk score distribution, survival status, and expression pattern of the TIILncSig. **(D)** Boxplot representing expression differences of lncRNAs between high- and low-risk groups.

other important clinical features of COAD patients. As shown in **Figure 4B**, the TIILncSig still had significant associations with overall survival in the TCGA cohort (HR = 2.687, 95% CI = 1.912–3.776, $p < 0.001$) and GSE17538 cohort (HR = 1.837, 95% CI = 1.047–3.223, $p = 0.034$). We further performed a stratification analysis for stage and examined the prognostic value of the TIILncSig within the early stage or late stage. All patients were classified into the early-stage group (Stage I and II) and late-stage group (Stage III and IV). When the TIILncSig was applied to patients with stage I and II, we found that the TIILncSig could effectively classify patients into a high-risk group and low-risk group with significantly different overall survival ($p < 0.001$, log-rank test; **Figure 4C**). Similar prediction results were observed for patients in the late-stage group. As shown in **Figure 4D**, patients with stage III and

IV were separated into two risk subgroups with significantly different overall survival ($p < 0.001$, log-rank test; **Figure 4D**). These multivariate and stratification analysis results demonstrated the independence of the TIILncSig from other clinical features in predicting overall survival.

DISCUSSION

Colon cancer is a well-known heterogeneous disease characterized by distinct clinical, pathological, molecular features, prognosis, and therapy response (Linnekamp et al., 2015; Zhai et al., 2017). Various molecular characteristics have been widely observed to contribute to colon cancer heterogeneity. Different gene expression patterns have been used to define

FIGURE 4 | Independence of the TIILncSig from other clinical features. **(A)** Univariable Cox regression analysis of the TIILncSig and other clinical features with overall survival in each cohort. **(B)** Multivariable Cox regression analysis of the TIILncSig and other clinical features with overall survival in each cohort. **(C)** Kaplan-Meier survival curves of overall survival between high- and low-risk groups for patients with stage I and II. **(D)** Kaplan-Meier survival curves of overall survival between high- and low-risk groups for patients with stage III and IV.

molecularly distinct subtypes with different clinical behavior (Perez-Villamil et al., 2012; Schlicker et al., 2012; Budinska et al., 2013). MicroRNA expression profiles have also been reported to show different expression patterns associated with those obtained by mRNA expression profiling; therefore, they can classify CRC tumors in the same way as mRNA (Sarver et al., 2009; Slattery et al., 2016). Recently, a new ncRNA molecule termed lncRNAs, also revealed different expression profiles associated with distinct clinical features, prognosis, and therapy response and have been used to defined lncRNA-derived molecular subtype (Chen et al., 2014; Zhou et al., 2018a).

Increasing efforts in colon cancer immunogenomics have shown that the type, density, and location of infiltrating immune cells within a tumor have a significant effect on clinical outcome and subtype classification (Zhang et al., 2018; Catalano et al., 2019; Shen et al., 2019). The functional roles and molecular mechanisms of lncRNAs involved in the regulation of tumor immune microenvironment have been reported in some studies (Zhang et al., 2020a). Some known immune-related lncRNAs have been proven to contribute to tumor initiation, growth, and metastasis and acted as potential biomarkers and targets (Zhou et al., 2017, 2018b). However, the tumor immune infiltration-associated lncRNAs and their clinical significance in colon cancer have not yet been fully investigated. Many studies have shown that cellular composition of immune infiltrates in the tumor could be quantified through traditional microscopy-based, but also could be inferred using immunoinformatics

algorithms and computational approaches based on genomic, transcriptomic, or DNA methylation data, which provided a very convenient way to unraveled tumor-immune interactions (Zhang et al., 2020b). Therefore, in this study, we first inferred infiltrating profiles of 22 immune cell types in the tumor microenvironment through the CIBERSORT algorithm based on the RNA-seq data of patients in the TCGA cohort. Then, we performed an integrative analysis of lncRNA expression profiles and immune cell infiltration profiles and identified 258 immune infiltration-associated lncRNAs. Combining survival data and Cox regression analysis, four of 258 immune infiltration-associated lncRNAs was found to be independent prognostic factors associated with overall survival after adjusted by other clinical features. To facilitate clinical application, we constructed an lncRNAs-based scoring model based on the expression levels of these four lncRNAs, and applied this model to different patient cohorts. Results from different cohorts demonstrated that the TIILncSig is not only a robust and reliable prognostic factor, but also in independent of other clinical features.

Although more and more lncRNAs have been discovered and identified using the experimental or computational approaches, only a small fraction of them have well been functionally characterized. Abnormal expression of LINC00926 has been observed in several cancers. For example, Wang et al. (2018) found that overexpression of LINC00926 was observed to be associate with improved overall survival in acute myeloid leukemia. Wu and colleagues found that dysregulated expression of

LINC00926 was associated with prostate cancer-related fatigue during localized radiation therapy by constructing mRNA and lncRNA regulatory networks. Functional analysis revealed the association of LINC00926 with inflammatory response and immune response-related biological processes (Ye et al., 2018). Further study of Liang et al. (2020) demonstrated the upregulated expression of LINC00926 in late relapsed Hodgkin lymphoma. Recently, Ma et al. (2020) performed a co-expression analysis of lncRNA and immune genes in breast cancer and found that LINC00926 is positively correlated with TNFRSF13C and CD19. The interaction between LINC00926 and the histone H3K4 methyltransferase, MLL1, was confirmed to lead to elevated proinflammatory cytokines and inflammatory states (Bam et al., 2019). Another lncRNAs in this TIILncSig, SNHG26, have also been associated with carcinogenesis and clinical outcome in several cancers. For example, SNHG26 has significantly upregulated in bladder urothelial carcinoma and associated with poor survival (Bao et al., 2017). Overexpression of SNHG6 was found to promote cell proliferation and metastasis in clear cell renal cell carcinoma (ccRCC) by interacting with YBX1 (Zhao et al., 2021). These existing experimental evidence further support the predicted TIILncSig in cancer carcinogenesis and progression and immune regulation. The other two lncRNAs have not been reported to associate with immune progress. Therefore they may be needed to be experimentally and functionally investigated. Although, our results uncovered TIILncSig as a potentially useful biomarker, the TIILncSig would need to be validated in a prospective immunotherapy clinical trial as a companion science trial. Furthermore, whether or not the TIILncSig would be useful in other cancers also should be tested. Finally, our study lacks a validation process for the TIILncSig. Further biological experiments should be made to elucidate the function of the TIILncSig.

In conclusion, we identified several novel lncRNAs involved in the regulation of tumor immune microenvironment associated with different patient clinical outcomes by performing integrative analysis of lncRNA expression profiles, immune cell infiltration profiles, and clinical profiles. Our study will help in understanding the mechanisms of lncRNA in immune regulation in the tumor microenvironment and provide novel biomarkers or targets for prognosis prediction and therapy decision-making for patients with colon cancer.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found at: clinical and transcriptome data of colon cancer patients was downloaded from UCSC Xena Browser (http://xena.ucsc.edu/, GDC TCGA Colon Cancer cohort) and Gene Expression Omnibus (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE17536).

AUTHOR CONTRIBUTIONS

BC conceived and designed the experiments. YW, JL, FR, and YC performed the experiments and analyzed the data. YW and JL wrote the paper. All authors contributed to the article and approved the submitted version.

FUNDING

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2021.671128/full#supplementary-material

REFERENCES

Bam, M., Yang, X. M., Nagarkatti, P. S., and Nagarkatti, M. (2019). Inflammation in PTSD is a consequence of dysregulated WNT signaling orchestrated by long non-coding RNA LINC00926 and lysine methyltransferase MLL1. *J. Immunol.* 202:182.20.

Bao, Z., Zhang, W., and Dong, D. (2017). A potential prognostic lncRNA signature for predicting survival in patients with bladder urothelial carcinoma. *Oncotarget* 8, 10485–10497. doi: 10.18632/oncotarget.14441

Binnewies, M., Roberts, E. W., Kersten, K., Chan, V., Fearon, D. F., Merad, M., et al. (2018). Understanding the tumor immune microenvironment (TIME) for effective therapy. *Nat. Med.* 24, 541–550. doi: 10.1038/s41591-018-0014-x

Budinska, E., Popovici, V., Tejpar, S., D'ario, G., Lapique, N., Sikora, K. O., et al. (2013). Gene expression patterns unveil a new level of molecular heterogeneity in colorectal cancer. *J. Pathol.* 231, 63–76. doi: 10.1002/path.4212

Catalano, I., Grassi, E., Bertotti, A., and Trusolino, L. (2019). Immunogenomics of colorectal tumors: facts and hypotheses on an evolving saga. *Trends Cancer* 5, 779–788. doi: 10.1016/j.trecan.2019.10.006

Chen, Y. G., Satpathy, A. T., and Chang, H. Y. (2017). Gene regulation in the immune system by long noncoding RNAs. *Nat. Immunol.* 18, 962–972. doi: 10.1038/ni.3771

Chen, H., Xu, J., Hong, J., Tang, R., Zhang, X., and Fang, J. Y. (2014). Long noncoding RNA profiles identify five distinct molecular subtypes of colorectal cancer with clinical relevance. *Mol. Oncol.* 8, 1393–1403. doi: 10.1016/j.molonc.2014.05.010

Chew, V., Toh, H. C., and Abastado, J. P. (2012). Immune microenvironment in tumor progression: characteristics and challenges for therapy. *J. Oncol.* 2012:608406. doi: 10.1155/2012/608406

Idos, G. E., Kwok, J., Bonthala, N., Kysh, L., Gruber, S. B., and Qu, C. (2020). The prognostic implications of tumor infiltrating lymphocytes in colorectal cancer: a systematic review and meta-analysis. *Sci. Rep.* 10:3360. doi: 10.1038/s41598-020-60255-4

Koncina, E., Haan, S., Rauh, S., and Letellier, E. (2020). Prognostic and predictive molecular biomarkers for colorectal cancer: updates and challenges. *Cancer* 12:319. doi: 10.3390/cancers12020319

Liang, Y., Zhu, H., Chen, J., Lin, W., Li, B., and Guo, Y. (2020). Construction of relapse-related lncRNA-mediated ceRNA networks in Hodgkin lymphoma. *Arch. Med. Sci.* 16, 1411–1418. doi: 10.5114/aoms.2020.98839

Linnekamp, J. F., Wang, X., Medema, J. P., and Vermeulen, L. (2015). Colorectal cancer heterogeneity and targeted therapy: a case for molecular disease subtypes. *Cancer Res.* 75, 245–249. doi: 10.1158/0008-5472.CAN-14-2240

Ma, W., Zhao, F., Yu, X., Guan, S., Suo, H., Tao, Z., et al. (2020). Immune-related lncRNAs as predictors of survival in breast cancer: a prognostic signature. *J. Transl. Med.* 18:442. doi: 10.1186/s12967-020-02522-6

Newman, A. M., Liu, C. L., Green, M. R., Gentles, A. J., Feng, W., Xu, Y., et al. (2015). Robust enumeration of cell subsets from tissue expression profiles. *Nat. Methods* 12, 453–457. doi: 10.1038/nmeth.3337

Perez-Villamil, B., Romera-Lopez, A., Hernandez-Prieto, S., Lopez-Campos, G., Calles, A., Lopez-Asenjo, J. A., et al. (2012). Colon cancer molecular subtypes identified by expression profiling and associated to stroma, mucinous type and different clinical behavior. *BMC Cancer* 12:260. doi: 10.1186/1471-2407-12-260

Sarver, A. L., French, A. J., Borralho, P. M., Thayanithy, V., Oberg, A. L., Silverstein, K. A., et al. (2009). Human colon cancer profiles show differential microRNA expression depending on mismatch repair status and are characteristic of undifferentiated proliferative states. *BMC Cancer* 9:401. doi: 10.1186/1471-2407-9-401

Schlicker, A., Beran, G., Chresta, C. M., Mcwalter, G., Pritchard, A., Weston, S., et al. (2012). Subtypes of primary colorectal tumors correlate with response to targeted treatment in colorectal cell lines. *BMC Med. Genomics* 5:66. doi: 10.1186/1755-8794-5-66

Shen, R., Li, P., Li, B., Zhang, B., Feng, L., and Cheng, S. (2019). Identification of distinct immune subtypes in colorectal cancer based on the stromal compartment. *Front. Oncol.* 9:1497. doi: 10.3389/fonc.2019.01497

Siegel, R. L., Miller, K. D., and Jemal, A. (2020b). Cancer statistics, 2020. *CA Cancer J. Clin.* 70, 7–30. doi: 10.3322/caac.21590

Siegel, R. L., Miller, K. D., Sauer, A. G., Fedewa, S. A., Butterly, L. F., Anderson, J. C., et al. (2020a). Colorectal cancer statistics, 2020. *CA Cancer J. Clin.* 70, 145–164. doi: 10.3322/caac.21601

Slattery, M. L., Herrick, J. S., Mullany, L. E., Wolff, E., Hoffman, M. D., Pellatt, D. F., et al. (2016). Colorectal tumor molecular phenotype and miRNA: expression profiles and prognosis. *Mod. Pathol.* 29, 915–927. doi: 10.1038/modpathol.2016.73

Sun, J., Zhang, Z., Bao, S., Yan, C., Hou, P., Wu, N., et al. (2020). Identification of tumor immune infiltration-associated lncRNAs for improving prognosis and immunotherapy response of patients with non-small cell lung cancer. *J. Immunother. Cancer* 8:e000110. doi: 10.1136/jitc-2019-000110

Turner, M., Galloway, A., and Vigorito, E. (2014). Noncoding RNA and its associated proteins as regulatory elements of the immune system. *Nat. Immunol.* 15, 484–491. doi: 10.1038/ni.2887

Vacante, M., Borzi, A. M., Basile, F., and Biondi, A. (2018). Biomarkers in colorectal cancer: current clinical utility and future perspectives. *World J. Clin. Cases* 6, 869–881. doi: 10.12998/wjcc.v6.i15.869

Wang, F., Tian, X., Zhou, J., Wang, G., Yu, W., Li, Z., et al. (2018). A three lncRNA signature for prognosis prediction of acute myeloid leukemia in patients. *Mol. Med. Rep.* 18, 1473–1484. doi: 10.3892/mmr.2018.9139

Ye, Y., Li, S. L., and Wang, S. Y. (2018). Construction and analysis of mRNA, miRNA, lncRNA, and TF regulatory networks reveal the key genes associated with prostate cancer. *PLoS One* 13:e0198055. doi: 10.1371/journal.pone.0198055

Zhai, Z., Yu, X., Yang, B., Zhang, Y., Zhang, L., Li, X., et al. (2017). Colorectal cancer heterogeneity and targeted therapy: clinical implications, challenges

and solutions for treatment resistance. *Semin. Cell Dev. Biol.* 64, 107–115. doi: 10.1016/j.semcdb.2016.08.033

Zhang, Z., Bao, S., Yan, C., Hou, P., Zhou, M., and Sun, J. (2020b). Computational principles and practice for decoding immune contexture in the tumor microenvironment. *Brief. Bioinform.* bbaa075. doi: 10.1093/bib/bbaa075 [Epub ahead of print]

Zhang, Y., Liu, Q., and Liao, Q. (2020a). Long noncoding RNA: a dazzling dancer in tumor immune microenvironment. *J. Exp. Clin. Cancer Res.* 39:231. doi: 10.1186/s13046-020-01727-3

Zhang, L., Zhao, Y., Dai, Y., Cheng, J. N., Gong, Z., Feng, Y., et al. (2018). Immune landscape of colorectal cancer tumor microenvironment from different primary tumor location. *Front. Immunol.* 9:1578. doi: 10.3389/fimmu.2018.01578

Zhao, P., Deng, Y., Wu, Y., Guo, Q., Zhou, L., Yang, X., et al. (2021). Long noncoding RNA SNHG6 promotes carcinogenesis by enhancing YBX1-mediated translation of HIF1alpha in clear cell renal cell carcinoma. *FASEB J.* 35:e21160. doi: 10.1096/fj.202000732RR

Zhou, M., Hu, L., Zhang, Z., Wu, N., Sun, J., and Su, J. (2018a). Recurrence-associated long non-coding RNA signature for determining the risk of recurrence in patients with colon cancer. *Mol. Ther. Nucleic Acids* 12, 518–529. doi: 10.1016/j.omtn.2018.06.007

Zhou, M., Zhang, Z., Bao, S., Hou, P., Yan, C., Su, J., et al. (2020). Computational recognition of lncRNA signature of tumor-infiltrating B lymphocytes with potential implications in prognosis and immunotherapy of bladder cancer. *Brief. Bioinform.* bbaa047. doi: 10.1093/bib/bbaa077 [Epub ahead of print]

Zhou, M., Zhang, Z., Zhao, H., Bao, S., Cheng, L., and Sun, J. (2018b). An immune-related six-lncRNA signature to improve prognosis prediction of glioblastoma multiforme. *Mol. Neurobiol.* 55, 3684–3697. doi: 10.1007/s12035-017-0572-9

Zhou, M., Zhao, H., Xu, W., Bao, S., Cheng, L., and Sun, J. (2017). Discovery and validation of immune-associated long non-coding RNA biomarkers associated with clinically molecular subtype and prognosis in diffuse large B cell lymphoma. *Mol. Cancer* 16:16. doi: 10.1186/s12943-017-0580-4