



Long Non-coding RNA Expression Patterns in Stomach Adenocarcinoma Serve as an Indicator of Tumor Mutation Burden and Are Associated With Tumor-Infiltrating Lymphocytes and Microsatellite Instability

Dongdong Yang^{1†}, Jinling Yu^{2†}, Bing Han³, Yue Sun³, Steven Mo^{4,5} and Jing Hu^{3*}

OPEN ACCESS

Edited by:

Palmiro Poltronieri,
Italian National Research Council, Italy

Reviewed by:

Renato Franco,
University of Campania Luigi Vanvitelli,
Italy

Massimo Mallardo,
University of Naples Federico II, Italy

*Correspondence:

Jing Hu
hujing@ems.hrbmu.edu.cn

[†] These authors have contributed
equally to this work and share first
authorship

Specialty section:

This article was submitted to
Molecular and Cellular Oncology,
a section of the journal
Frontiers in Cell and Developmental
Biology

Received: 16 October 2020

Accepted: 21 January 2021

Published: 12 February 2021

Citation:

Yang D, Yu J, Han B, Sun Y, Mo S
and Hu J (2021) Long Non-coding
RNA Expression Patterns in Stomach
Adenocarcinoma Serve as an
Indicator of Tumor Mutation Burden
and Are Associated With
Tumor-Infiltrating Lymphocytes
and Microsatellite Instability.
Front. Cell Dev. Biol. 9:618313.
doi: 10.3389/fcell.2021.618313

¹ Department of Surgical Oncology, The Fourth Affiliated Hospital of Harbin Medical University, Harbin, China, ² Department of General Practice, The Fourth Affiliated Hospital of Harbin Medical University, Harbin, China, ³ Department of Internal Medical Oncology, Harbin Medical University Cancer Hospital, Harbin, China, ⁴ YDLife Academy of Sciences, Nanning, China, ⁵ Chinese Academy of Sciences Center for Excellence in Brain Science and Intelligence Technology, Shanghai, China

Long non-coding RNAs (lncRNAs) are crucial in controlling important aspects of tumor immunity. However, whether the expression pattern of lncRNAs in stomach adenocarcinoma (STAD) reflects tumor immunity is not fully understood. We screened differentially expressed lncRNAs (DElncRNAs) between high and low tumor mutation burden (TMB) STAD samples. Using the least absolute shrinkage and selection operator method, 33 DElncRNAs were chosen to establish a lncRNA-based signature classifier for predicting TMB levels. The accuracy of the 33-lncRNA-based signature classifier was 0.970 in the training set and 0.950 in the test set, suggesting the expression patterns of the 33 lncRNAs may be an indicator of TMB in STAD. Survival analysis showed that a lower classifier index reflected better prognosis for STAD patients, and the index showed correlation with expression of immune checkpoint molecules (PD1, PDL1, and CTLA4), tumor-infiltrating lymphocytes, and microsatellite instability. In conclusion, STAD samples with different tumor mutation burdens have different lncRNA expression patterns. The 33-lncRNA-based signature classifier index may be an indicator of TMB and is associated expression of immune checkpoints, tumor-infiltrating lymphocytes, and microsatellite instability.

Keywords: lncRNAs, tumor immunity, stomach adenocarcinoma, the Cancer Genome Atlas, immune checkpoint molecules

INTRODUCTION

Stomach adenocarcinoma (STAD) accounts for 90% of gastric cancers (Hoffmann, 2015). Surgical treatment remains the optimal treatment for early-stage STAD, but most STAD patients are diagnosed in an advanced stage (Song et al., 2017). Development of targeted drugs has benefited many patients with specific genetic mutations (Shih et al., 2005; Sequist et al., 2007), but short-term drug resistance is still a limitation of this treatment (Nishio et al., 1999). Recently, immune

checkpoint inhibitors (ICIs) showed promising effects against STAD (Atkins et al., 2017; Farina et al., 2017; Tosoni et al., 2017; Alsharedi and Katz, 2018; Killock, 2018; Li et al., 2018; Suresh et al., 2018; Togasaki et al., 2018; Arias Ron et al., 2019). However, ICIs are not effective for all STAD patients, so it would be helpful to identify indicators that can detect which patients are most likely to respond to ICIs.

Tumor immunogenicity, expression of immune checkpoint molecules and tumor-infiltrating lymphocytes are indicators of ICI treatment efficacy (Schreiber et al., 2011; Galon and Bruni, 2019). Tumor mutation burden (TMB), the number of somatic mutations in tumors as determined by whole-exon sequencing (Chalmers et al., 2017), is a widely used indicator of tumor immunogenicity and is also related to the therapeutic efficacy of ICIs (Alexandrov et al., 2013; Carbone et al., 2017; Hellmann et al., 2018). Efficacy of ICIs in STAD cancer patients is connected to high TMB levels (Kim et al., 2014; Goodman et al., 2019; Samstein et al., 2019; Wang et al., 2019). In patients with chemo-refractory advanced gastric cancer, overall survival rate was significantly higher in the high TMB group than in the low TMB group (Wang et al., 2019). However, the cost of determining TMB is high, and detection methods are not widely available in the clinic.

Long non-coding RNA (lncRNA) plays an important role in the development of various immune cells and the dynamic transcription process that controls markers that activate immune cells (Atianand et al., 2017). Under normal physiological conditions, expression of lncRNAs is tissue-specific and tightly

regulated, whereas lncRNA expression in cancer is aberrant (Camacho et al., 2018). Therefore, we hypothesized that the expression pattern of lncRNA may be an indicator of TMB. We analyzed lncRNA expression patterns in low- and high-TMB STAD samples and defined a 33-lncRNA-based signature classifier that could accurately distinguish them. Our present may provide a cost-effective way to identify STAD patients more likely to respond to ICIs, thereby helping personalize treatment.

RESULTS

LncRNAs Differentially Expressed Between High- or Low-TMB STAD Samples

The workflow of the present study is shown in **Figure 1**. We analyzed the lncRNA expression profiles and mutation data of 348 STAD samples, comprising 64 with high TMB and 284 with low TMB. These samples were randomly assigned to the training set (261 samples) and the test set (87 samples). The training set contained 48 high- and 213 low-TMB samples. There were no significant differences in routine clinicopathological characteristics between the training and test sets (**Table 1**). A total of 309 lncRNAs were differentially expressed in the high-TMB sample compared to low-TMB samples (**Figure 2A**). Among these differentially expressed lncRNAs (DELncRNAs), 205 were up-regulated, and 104 were down-regulated. Hierarchical clustering showed that expression

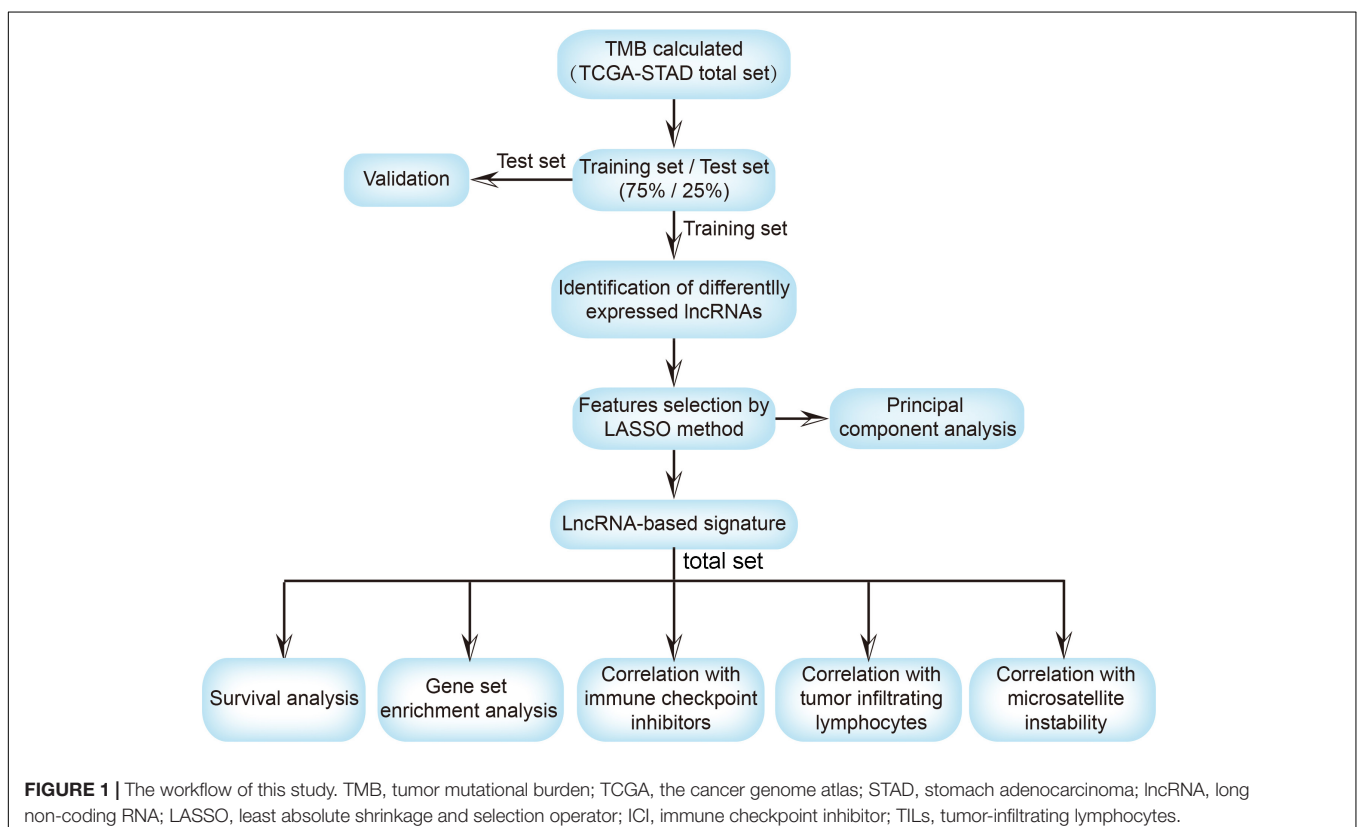


TABLE 1 | Summary of patient cohort information.

Factor	Cohort		P-value
	Training (N = 261)	Test (N = 87)	
Gender			0.797
Male	165	57	
Female	96	30	
Age (year)			0.050
<65	117	27	
≥65	141	59	
Not available	3	1	
T			0.571
T1–2	65	25	
T3–4	196	62	
N			0.852
N0	79	30	
N1–3	177	56	
Nx	4	1	
Not available	1	0	
M			0.091
M0	235	73	
M1	17	6	
Mx	9	8	
Pathological stage			0.803
I–II	123	39	
III–VI	138	48	
Histological type			0.599
Diffuse type	45	13	
Mucinous type	16	3	
Not otherwise specified	141	49	
Papillary type	4	1	
Signet ring type	8	2	
Tubular type	47	18	
Not available	0	1	
Microsatellite			0.904
MSI-H	46	15	
MSS-L	41	12	
MSS	173	60	
Not available	1	0	
Grade			0.840
G1	5	3	
G2	93	30	
G3	157	52	
Gx	6	2	

MSI-H, microsatellite instability-high; MSI-L, microsatellite instability-low; MSS, microsatellite stability.

of DElncRNAs could distinguish STAD samples with different levels of TMB (Figure 2B).

The 33-LncRNA-Based Classifier Index

The least absolute shrinkage and selection operator (LASSO) logistic regression method developed a lncRNA-based classifier using the expression profiles of the DElncRNAs in the training set. A total of 33 DElncRNAs were identified as optimal lncRNAs with non-zero regression coefficients (Figure 3A). Principal

component analysis demonstrated that the expression patterns of the 33 DElncRNAs could distinguish samples with high or low TMB (Figure 3B). The accuracy of the 33-lncRNA-based classifier was 0.970 in the training set and 0.950 in the test set, and the sample recognition efficiency of the classifier was high (Table 2). Additionally, receiver operating characteristic (ROC) curve analysis confirmed that the 33-lncRNA-based classifier could predict TMB levels of STAD samples, with an area under the curve (AUC) of 0.999 (Figure 3C) in the training set and 0.974 (Figure 3D) in the test set. Moreover, the difference in AUCs between the training and test set was not significant ($P = 0.190$). Higher values of the classifier index were associated with lower TMB.

Biological Characteristics of STAD With High LncRNA-Based Classifier Index

Stomach adenocarcinoma (STAD) patients with a high lncRNA-based classifier index had worse overall survival (Figure 4A) and recurrence-free survival (Figure 4B). Biological processes (BPs) enriched in samples with a high lncRNA-based classifier index included branched chain amino acid catabolic process, positive regulation of protein export from nucleus, and signal transduction involved in cell cycle checkpoint (Figure 4C). Meanwhile, BPs enriched in samples with low values of the lncRNA-based classifier index were cardiac cell development, cardiac myofibril assembly, cardiocyte differentiation, and regulation of adherens junction organization (Figure 4D). KEGG pathways enriched in high-index groups included cytosolic DNA sensing pathway, p53 signaling pathway, and pyrimidine metabolism (Figure 4E), while those enriched in low-index groups were melanogenesis, neuroactive ligand receptor interaction, and vascular smooth muscle contraction (Figure 4F).

The 33-LncRNA-Based Classifier Index Associate Expression of Immune Checkpoint

To explore the roles of the classifier index in STAD, we assessed correlations of the 33-lncRNA based classifier index with the expression of three immune checkpoint molecules (PDCD1 also known as PD1, CD274 also known as PDL1, and CTLA4). The 33-lncRNA based classifier index showed significant correlation with CD274 (Pearson $R = -0.36$, $P = 2.4e-12$), CTLA4 (Pearson $R = -0.21$, $P = 6.7e-05$) and PDCD1 (Pearson $R = -0.22$, $P = 2.7e-05$) (Figures 5A–C). Moreover, the 33-lncRNA based classifier index showed significant negative correlation with TMB (Figure 5D).

The Correlation Between the 33-LncRNA-Based Classifier Index and Tumor Infiltrating Lymphocytes

We explored the correlation of the 33 lncRNAs and tumor-infiltrating lymphocytes (Figure 6A). The lncRNA LINC02345 was positively correlated with resting dendritic cells, M2 macrophages and activated memory CD4 T cells. Activated memory CD4 T cells were also positively correlated with lncRNAs DGCR11, DEPDC1-AS1, LINC00909, MIR210HG,

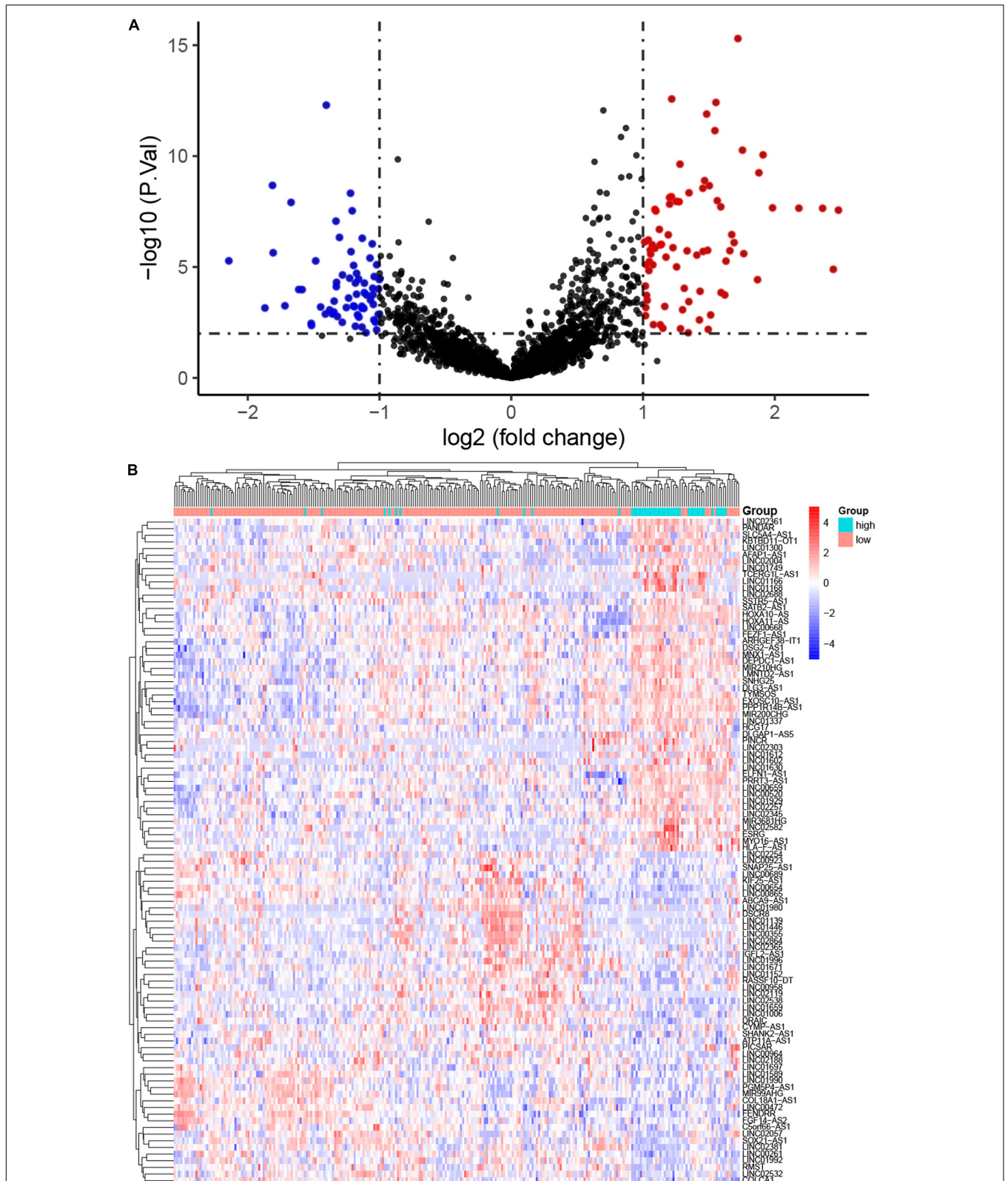


FIGURE 2 | Differentially expressed long non-coding RNAs (DElncRNAs) between stomach adenocarcinoma (STAD) samples with high and low tumor mutation burden (TMB). **(A)** Volcano plot showing DElncRNAs. Red points represent up-regulated; blue points represent down-regulated RNAs; and black points represent no significant difference. **(B)** The expression pattern of DElncRNAs can basically distinguish the level of TMB in STAD.

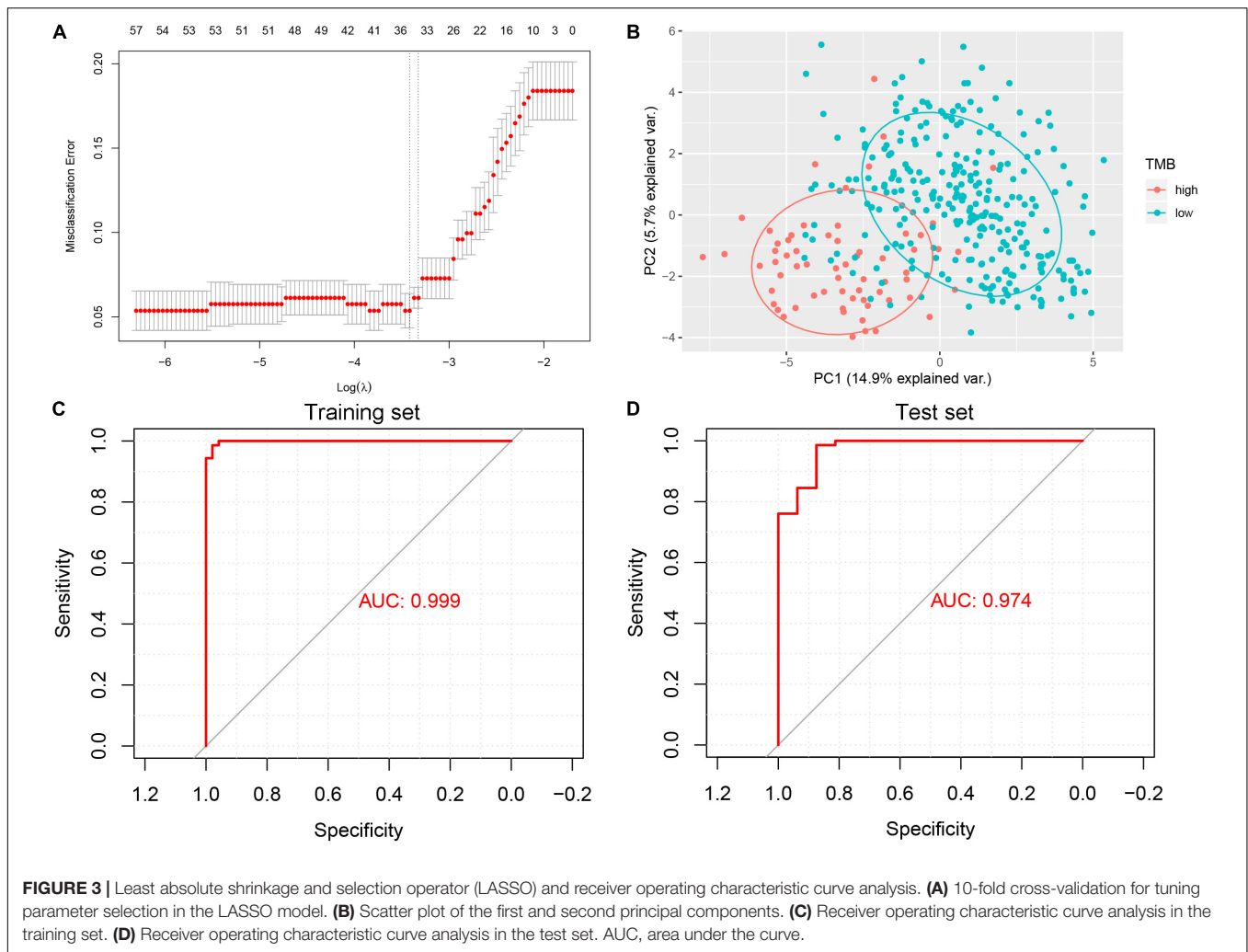


TABLE 2 | Performance of the 33-lncRNA-based classifier index in predicting tumor mutation burden (TMB) in stomach adenocarcinoma (STAD) samples.

Group	Se	Sp	PPV	NPV	Accuracy	AUC
Training set	0.830	0.960	1.000	0.960	0.970	0.999
Test set	0.560	0.910	1.000	0.910	0.950	0.974

AUC, area under the receiver operating characteristic curve; NPV, negative predictive value; PPV, positive predictive value; Se, sensitivity; Sp, specificity.

DSG2-AS1, and RPARP-AS. These most of these 33 lncRNAs were negatively correlated with endothelial cells, fibroblasts and resting mast cells. The 33-lncRNA based classifier index was positively correlated with endothelial cells, fibroblasts and resting mast cells. It was negatively correlated with the following cells: Memory B cells, resting dendritic cells, M1 macrophages, activated mast cells, activated NK cells, activated memory CD4 T cells, CD8 T cells, and gamma delta T cells.

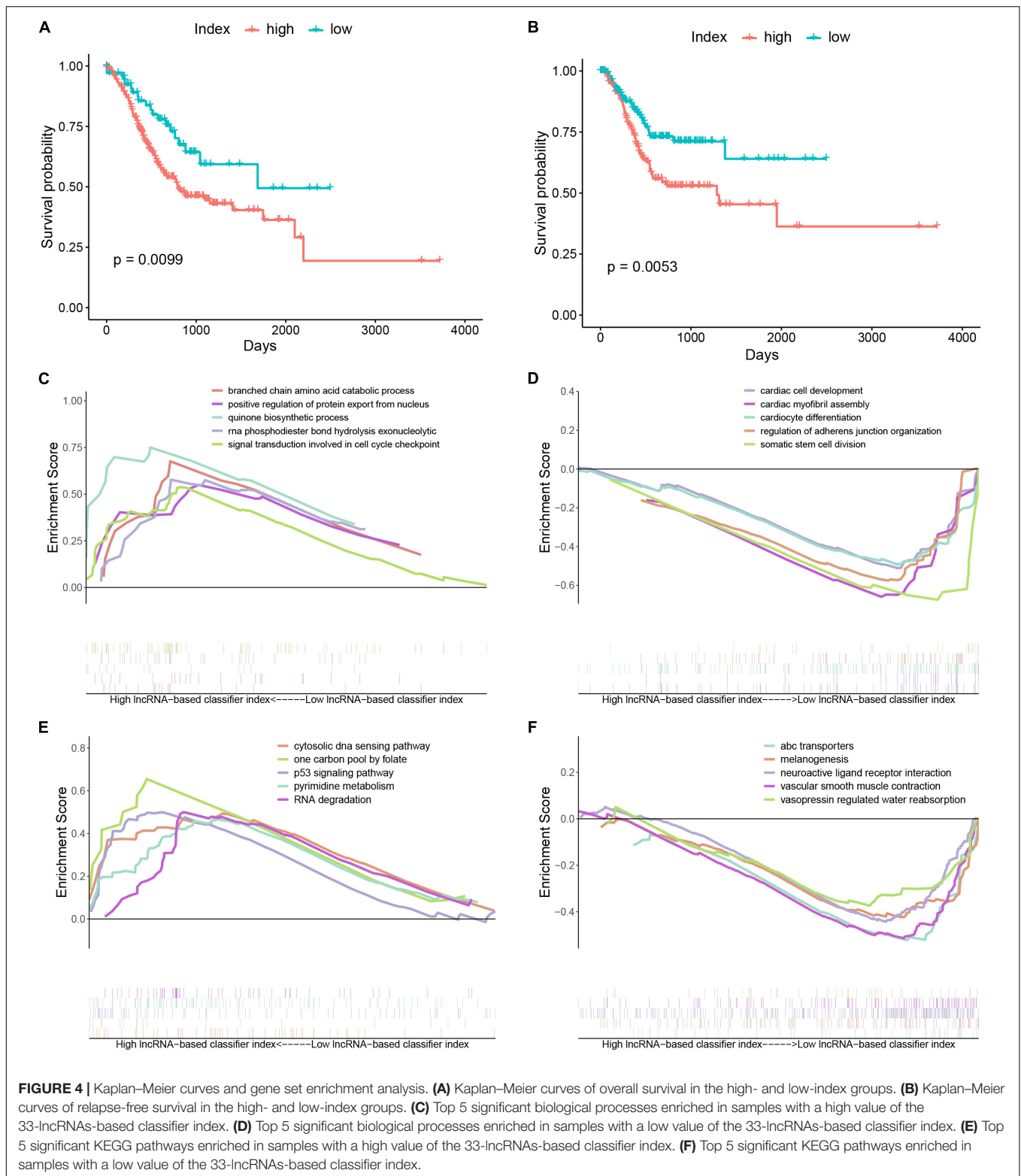
The LncRNA-Based Index Associates With Microsatellite Instability

Microsatellite instability-high (MSI-H) is associated with response to treatment of ICIs in STAD (Fuchs et al., 2018). We found that STAD patients with MSI-H have a lower

lncRNA base index (**Figure 6B**) than those with MSI-low or microsatellite stable (MSS).

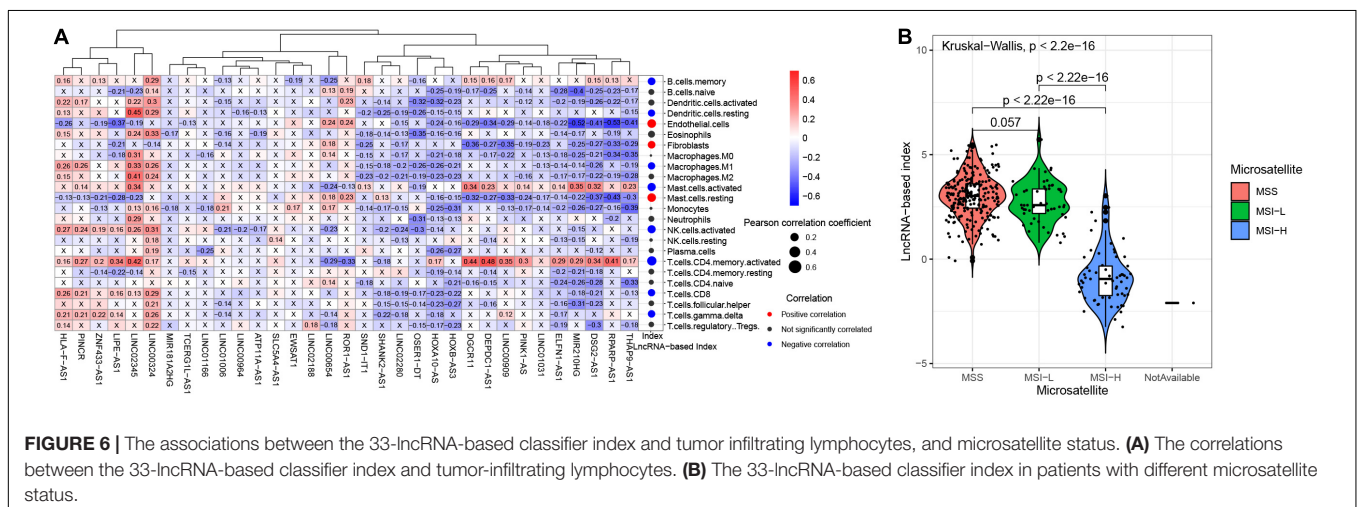
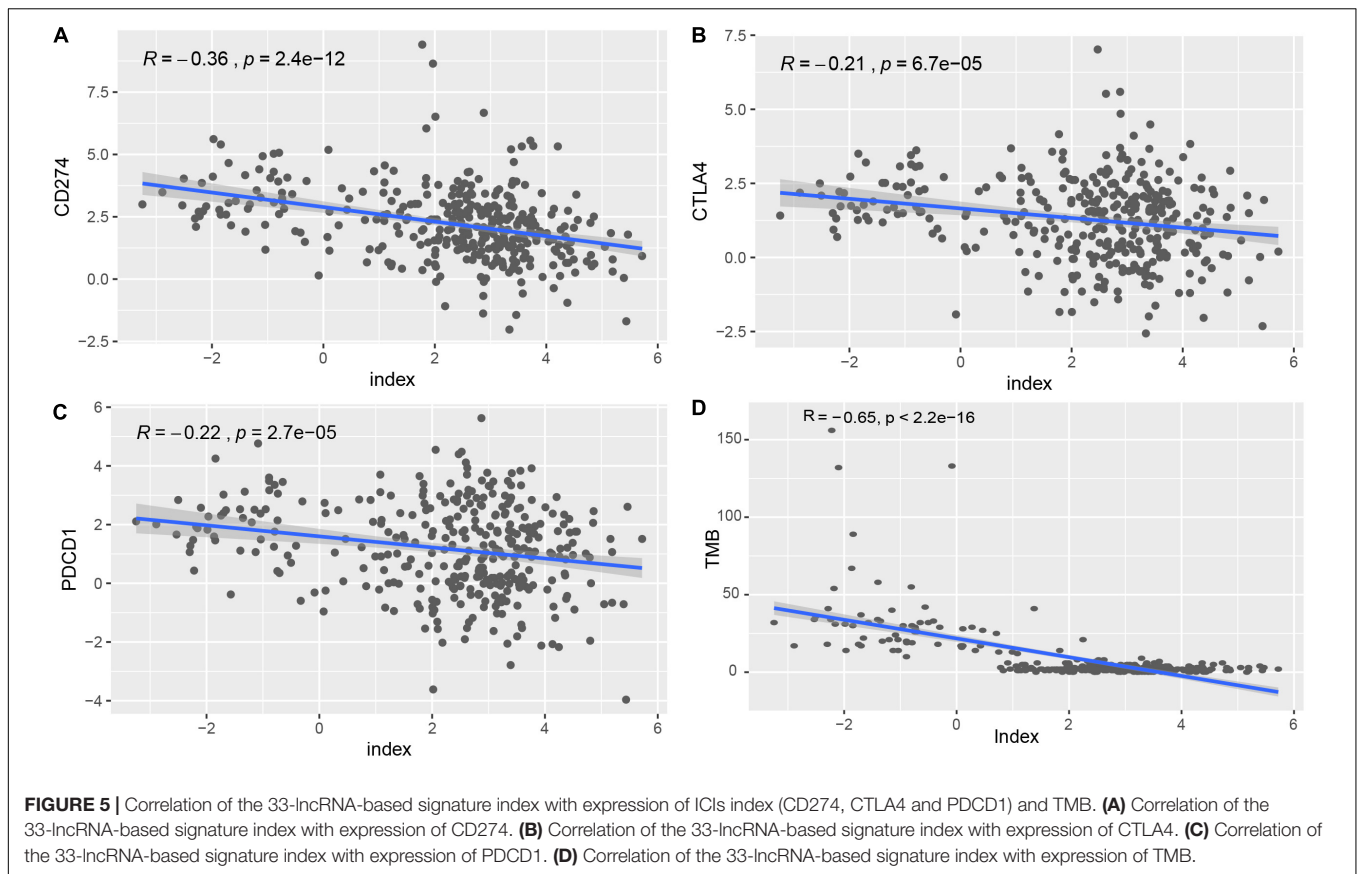
DISCUSSION

Tumor mutation burden (TMB) is used in clinical practice for predicting ICI treatment efficacy, and high TMB is related to better prognosis in patients treated with immunotherapy (Rizvi, 2018; Samstein, 2019). However, TMB detection is too expensive to be widely used in clinic. In the present study, we screened the DELncRNAs between STAD samples with high or low TMB, and identified a 33-lncRNA-based signature classifier that can assess TMB in a cost-effective manner. The accuracy of the 33-lncRNA-based classifier was high and robust.



Moreover, the classifier had a high specificity, positive predictive value and negative predictive value. The expression pattern of these 33 lncRNAs may be a novel indicator for predicting ICI treatment efficacy.

Tumor mutation burden (TMB) is a robust predictor of the efficacy of ICI treatment (Cao et al., 2019) and a potential prognostic biomarker (Romero, 2019). We found that a low value of our 33-lncRNA-based index was associated with better overall



and recurrence-free survival in STAD patients. Interestingly, immune-related biological pathways and KEGG pathways were not enriched in STAD samples regardless of the value of the lncRNA-based index. This may be because the molecular function of most lncRNAs is still unknown. The biological mechanism underlying these 33 lncRNAs associated with TMB levels required further exploration.

Tumor mutation burden (TMB), expression of immune checkpoint molecules and tumor-infiltrating lymphocytes are

associated with ICI efficacy. However, TMB may predict the efficacy of ICI independently of the expression level of immune checkpoint molecule CD274 (Carbone et al., 2017). Our 33-lncRNA-based signature classifier index showed significant low correlation with PDCD1, CD274, and CTLA4. This suggests that the signature classifier index cannot replace the measurement of immune checkpoint expression for predicting ICI efficacy. ICIs appear to be more effective against tumors when they have been infiltrated by lymphocytes (Maleki Vareki, 2018). Our

33 lncRNAs were associated with numbers of resting dendritic cells, M1 macrophages, resting mast cells, activated NK cells and activated memory CD4 T cells, while the 33-lncRNA-based classifier index was mainly associated with numbers of memory B cells, resting dendritic cells, endothelial cells, fibroblasts, M1 macrophages, activated NK cells and activated memory CD4 T cells. In gastric cancer, high levels of tumor-infiltrating lymphocytes are related to better prognosis (Morihira et al., 2019; Zhang et al., 2019). Moreover, STAD patients with MSI-H score have a lower lncRNA-based index. Thus, the 33-lncRNA-based signature classifier index is associated with four factors that can predict ICI efficacy: TMB, expression of immune checkpoint molecules, tumor-infiltrating lymphocytes, and microsatellite instability. Future prospective studies should investigate whether the index can be used in the clinic to predict ICI efficacy.

Although this study may provide an exciting potential surrogate measure for TMB in STAD, there are limitations worth mentioning. Firstly, the classifier needs to be validated in a larger independent data set. Second, the molecular mechanism of these 33 lncRNAs in the immune response is not yet clear. Thirdly, the relationship between the lncRNA-based index and other molecular typing systems were not clear.

In conclusion, STAD samples with different TMB levels have different lncRNA expression patterns. The 33-lncRNA-based signature classifier index may be an indicator of TMB and is associated expression of immune checkpoints, tumor-infiltrating lymphocytes, and microsatellite instability.

EXPERIMENTAL PROCEDURES

Data Processing

The *GDCquery_Maf* function of the *TCGAbiolinks* package (Colaprico et al., 2016) was used to download STAD mutation annotation files from the Cancer Genome Atlas (TCGA) database¹. The *MuTect2* pipeline (Cibulskis et al., 2013) called somatic mutations. The *read.maf* function in the *maftools* package (Mayakonda et al., 2018) downloaded maf data, which included variation samples and the variation count of each sample. In our study, TMB was defined as the number of somatic cell variations per megabase (MB) in the genome (Hellmann et al., 2018). High TMB was defined as ≥ 10 mutations/MB, and low TMB as < 10 mutations/MB (Hellmann et al., 2018; Lv et al., 2019). The estimated size of the exome was 38 MB (Chalmers et al., 2017). The lncRNA and mRNA were identified according to the comprehensive gene annotation (version 33) file downloaded from gencode². STAD stomach samples, with both lncRNA expression profiles and mutations, were randomly assigned to the training set (75%) and the test set (25%).

Screening DElncRNAs in High-TMB Compared to Low-TMB STAD Samples

In the training set, the *voom* function (Law et al., 2014) was used to normalize the RNA sequence data of STAD.

The *limma* package (Smyth, 2004; Law et al., 2014; Ritchie et al., 2015) was used to screen the DElncRNAs in high-TMB compared to low-TMB STAD samples. LncRNAs with a *P* (adjusted by the false discovery rate) < 0.01 (Lv et al., 2019) were considered significant. Based on Euclidean distance, we performed bidirectional hierarchical clustering (Szekely and Rizzo, 2005) using the expression profiles of DElncRNAs and displayed the results as a heat map using the *pheatmap* R package³.

Feature Selection and Principal Component Analysis

Using the *glmnet* package (Friedman et al., 2010), LASSO logistic regression identified optimal lncRNAs with non-zero coefficient from DElncRNAs, and these lncRNAs were used to create a LASSO signature classifier. The corresponding parameters are set to “*nfolds* = 10” representing 10-fold cross-validation, and “*family* = ‘binomial’” representing two-class logistic regression analysis. In addition, the expression profiles of the optimal DElncRNAs were used to perform principal component analysis. The first and second component was applied to visualize the result displayed as a two-dimensional scatter plot using *ggbiplot* package⁴.

Construction of the LncRNA-Based Signature Classifier for Predicting TMB

Through LASSO logistic regression, we obtained the regression coefficient (Coef) of the lncRNA and calculated the classifier index for each STAD sample according to formula:

$$\text{Index} = \text{Expr}_{\text{lncRNA1}} * \text{Coef}_{\text{lncRNA1}} + \text{Expr}_{\text{lncRNA2}} * \text{Coef}_{\text{lncRNA2}} + \text{Expr}_{\text{lncRNA3}} * \text{Coef}_{\text{lncRNA3}} + \dots + \text{Expr}_{\text{lncRNA}_n} * \text{Coef}_{\text{lncRNA}_n}$$

where “Expr” represents the expression value of the lncRNA. The classifier’s accuracy, positive predictive value, negative predictive value, sensitivity, specificity and area under the receiver operating characteristic curve (AUC) were determined. The receiver operating characteristic curves were generated using the *pROC* package (Robin et al., 2011) in R, and compared using the DeLong method (DeLong et al., 1988).

Survival Analysis and Gene Set Enrichment Analysis

To further explore prognostic value of the lncRNA-based signature classifier index in STAD, survival analysis was performed using *survival* and *survminer* package⁵. The *surv_cutpoint* function in the *survminer* package was used to determine the optimal cutoff to divide patients high or low-index groups. GSEA2-2.2.4 (Java) version (Mootha et al., 2003; Subramanian et al., 2005) was used to explore potential biological characteristics in the high- or low-index groups. According to default parameters, the *c5.bp.v6.2.symbols.gmt* and *c2.cp.kegg.v6.2.symbols.gmt* datasets in MsigDB V6.2 database

³<https://CRAN.R-project.org/package=pheatmap>

⁴<https://github.com/vqv/ggbiplot>

⁵<https://CRAN.R-project.org/package=survival>

¹<https://gdc.cancer.gov/>

²<https://www.gencodegenes.org>

(Liberzon et al., 2015) were used as reference gene sets. Nominal $P < 0.05$ was considered significant.

Correlation of the LncRNA-Based Signature Classifier Index With Expression of Immune Checkpoints

The expression of immune checkpoint molecules is another factor currently considered to affect ICI efficacy (Schreiber et al., 2011). Therefore, in our study, we investigated the correlation of the lncRNA-based signature classifier index with the expression of three immune checkpoints (PDCD1, CD274, and CTLA4). The correlation between the lncRNA-based signature classifier index and TMB was also explored.

Estimation of the Correlations of Tumor-Infiltrating Lymphocytes and the LncRNA-Based Signature Classifier Index

Firstly, we integrated two gene signatures, CIBERSORT (Newman et al., 2015) and MCP-Counter (Becht et al., 2016), to identify the mark gene set of the tumor-infiltrating lymphocytes and stromal cells according to a previous study (Xiao et al., 2019). Subsequently, we used single sample gene set enrichment analysis to calculate the abundance of each cell subset in each sample based on gene expression data. Then, the Pearson correlation between the lncRNA-based signature classifier index and the cell abundances were calculated. $P < 0.05$ was considered significant.

Statistical Analysis

All the analyses were performed using R software⁶. The χ^2 -test was used for categorical data. An unpaired t -test was used to screen DELncRNAs. Kaplan–Meier survival curve analysis with log-rank method was used to compare survival between the two

⁶<https://www.r-project.org/>

REFERENCES

- Alexandrov, L. B., Nik-Zainal, S., Wedge, D. C., Aparicio, S. A., Behjati, S., Biankin, A. V., et al. (2013). Signatures of mutational processes in human cancer. *Nature* 500, 415–421. doi: 10.1038/nature12477
- Alsharedi, M., and Katz, H. (2018). Check point inhibitors a new era in renal cell carcinoma treatment. *Med. Oncol.* 35:85. doi: 10.1007/s12032-018-1147-y
- Arias Ron, D., Labandeira, C. M., Areses Manrique, M. C., Sampedro Domarco, P., Abdulkader, I., Garcia-Mata, J., et al. (2019). Dramatic response of leptomeningeal carcinomatosis to nivolumab in PD-L1 highly expressive non-small cell lung cancer: a case report. *Front. Oncol.* 9:819. doi: 10.3389/fonc.2019.00819
- Atianand, M. K., Caffrey, D. R., and Fitzgerald, K. A. (2017). Immunobiology of long noncoding RNAs. *Annu. Rev. Immunol.* 35, 177–198. doi: 10.1146/annurev-immunol-041015-055459
- Atkins, M. B., Clark, J. I., and Quinn, D. I. (2017). Immune checkpoint inhibitors in advanced renal cell carcinoma: experience to date and future directions. *Ann. Oncol.* 28, 1484–1494. doi: 10.1093/annonc/mdx151
- Becht, E., Giraldo, N. A., Lacroix, L., Buttard, B., Elarouci, N., Petitprez, F., et al. (2016). Estimating the population abundance of tissue-infiltrating immune and stromal cell populations using gene expression. *Genome Biol.* 17:218. doi: 10.1186/s13059-016-1070-5
- Camacho, C. V., Choudhari, R., and Gadad, S. S. (2018). Long noncoding RNAs and cancer, an overview. *Steroids* 133, 93–95. doi: 10.1016/j.steroids.2017.12.012
- Cao, D., Xu, H., Xu, X., Guo, T., and Ge, W. (2019). High tumor mutation burden predicts better efficacy of immunotherapy: a pooled analysis of 103078 cancer patients. *Oncoimmunology* 8:e1629258. doi: 10.1080/2162402X.2019.1629258
- Carbone, D. P., Reck, M., Paz-Ares, L., Creelan, B., Horn, L., Steins, M., et al. (2017). First-line nivolumab in stage IV or recurrent non-small-cell lung cancer. *N. Engl. J. Med.* 376, 2415–2426. doi: 10.1056/NEJMoa1613493
- Chalmers, Z. R., Connelly, C. F., Fabrizio, D., Gay, L., Ali, S. M., Ennis, R., et al. (2017). Analysis of 100,000 human cancer genomes reveals the landscape of tumor mutational burden. *Genome Med.* 9:34. doi: 10.1186/s13073-017-0424-2

groups of patients. Kruskal–Wallis method was used compared the continuous variable between three groups. Unless otherwise stated, we considered $P < 0.05$ to be statistically significant.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

DY, JY, SM, and BH performed the material preparation, data collection, and analysis. YS and JH wrote the first draft of the manuscript. All authors commented on previous versions of the manuscript, contributed to the study conception and design, read and approved the final manuscript.

FUNDING

This work was supported in part by the National Natural Science Fund of China (Grant Nos. 81501960 and 81972162 to JH), the Natural Science Fund for Outstanding Youth of Heilongjiang Province (Grant No. YQ2019H026 to JH), and the Major Program of Haiyan Fund of the Harbin Medical University Cancer Hospital (Grant No. JJZD2019-04 to JH).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fcell.2021.618313/full#supplementary-material>

Supplementary Table 1 | The 33 long non-coding RNA selected by LASSO method.

- Cibulskis, K., Lawrence, M. S., Carter, S. L., Sivachenko, A., Jaffe, D., Sougnez, C., et al. (2013). Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. *Nat. Biotechnol.* 31, 213–219. doi: 10.1038/nbt.2514
- Colaprico, A., Silva, T. C., Olsen, C., Garofano, L., Cava, C., Garolini, D., et al. (2016). TCGAAbiolinks: an R/Bioconductor package for integrative analysis of TCGA data. *Nucleic Acids Res.* 44:e71. doi: 10.1093/nar/gkv1507
- DeLong, E. R., DeLong, D. M., and Clarke-Pearson, D. L. (1988). Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics* 44, 837–845.
- Farina, M. S., Lundgren, K. T., and Bellmunt, J. (2017). Immunotherapy in urothelial cancer: recent results and future perspectives. *Drugs* 77, 1077–1089. doi: 10.1007/s40265-017-0748-7
- Friedman, J., Hastie, T., and Tibshirani, R. (2010). Regularization paths for generalized linear models via coordinate descent. *J. Stat. Softw.* 33, 1–22.
- Fuchs, C. S., Doi, T., Jang, R. W., Muro, K., Satoh, T., Machado, M., et al. (2018). Safety and efficacy of pembrolizumab monotherapy in patients with previously treated advanced gastric and gastroesophageal junction cancer: phase 2 clinical KEYNOTE-059 trial. *JAMA Oncol.* 4:e180013. doi: 10.1001/jamaoncol.2018.0013
- Galon, J., and Bruni, D. (2019). Approaches to treat immune hot, altered and cold tumours with combination immunotherapies. *Nat. Rev. Drug Discov.* 18, 197–218. doi: 10.1038/s41573-018-0007-y
- Goodman, A. M., Sokol, E. S., Frampton, G. M., Lippman, S. M., and Kurzrock, R. (2019). Microsatellite-stable tumors with high mutational burden benefit from immunotherapy. *Cancer Immunol. Res.* 7, 1570–1573. doi: 10.1158/2326-6066.CIR-19-0149
- Hellmann, M. D., Ciuleanu, T. E., Pluzanski, A., Lee, J. S., Otterson, G. A., Audigier-Valette, C., et al. (2018). Nivolumab plus ipilimumab in lung cancer with a high tumor mutational burden. *N. Engl. J. Med.* 378, 2093–2104. doi: 10.1056/NEJMoa1801946
- Hoffmann, W. (2015). Current status on stem cells and cancers of the gastric epithelium. *Int. J. Mol. Sci.* 16, 19153–19169. doi: 10.3390/ijms160819153
- Killock, D. (2018). New window of opportunity with ICIs in melanoma. *Nat. Rev. Clin. Oncol.* 15:723. doi: 10.1038/s41571-018-0124-x
- Kim, T. M., Jung, S. H., Kim, M. S., Baek, I. P., Park, S. W., Lee, S. H., et al. (2014). The mutational burdens and evolutionary ages of early gastric cancers are comparable to those of advanced gastric cancers. *J. Pathol.* 234, 365–374. doi: 10.1002/path.4401
- Law, C. W., Chen, Y., Shi, W., and Smyth, G. K. (2014). voom: Precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biol.* 15:R29. doi: 10.1186/gb-2014-15-2-r29
- Li, P., Yang, X., Feng, Y., Wu, L., Ma, W., Ding, G., et al. (2018). The impact of immunosenescence on the efficacy of immune checkpoint inhibitors in melanoma patients: a meta-analysis. *Onco. Targets Ther.* 11, 7521–7527. doi: 10.2147/OTT.S165368
- Liberzon, A., Birger, C., Thorvaldsdottir, H., Ghandi, M., Mesirov, J. P., and Tamayo, P. (2015). The molecular signatures database (MSigDB) hallmark gene set collection. *Cell Syst.* 1, 417–425. doi: 10.1016/j.cels.2015.12.004
- Lv, Y., Huang, Z., Lin, Y., Fang, Y., Chen, Z., Pan, L., et al. (2019). MiRNA expression patterns are associated with tumor mutational burden in lung adenocarcinoma. *Oncoimmunology* 8:e1629260. doi: 10.1080/2162402X.2019.1629260
- Maleki Vareki, S. (2018). High and low mutational burden tumors versus immunologically hot and cold tumors and response to immune checkpoint inhibitors. *J. Immunother. Cancer* 6:157. doi: 10.1186/s40425-018-0479-7
- Mayakonda, A., Lin, D. C., Assenov, Y., Plass, C., and Koeffler, H. P. (2018). Maftools: efficient and comprehensive analysis of somatic variants in cancer. *Genome Res.* 28, 1747–1756. doi: 10.1101/gr.239244.118
- Mootha, V. K., Lindgren, C. M., Eriksson, K. F., Subramanian, A., Sihag, S., Lehar, J., et al. (2003). PGC-1 α -responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nat. Genet.* 34, 267–273. doi: 10.1038/ng1180
- Morihiro, T., Kuroda, S., Kanaya, N., Kakiuchi, Y., Kubota, T., Aoyama, K., et al. (2019). PD-L1 expression combined with microsatellite instability/CD8+ tumor infiltrating lymphocytes as a useful prognostic biomarker in gastric cancer. *Sci. Rep.* 9:4633. doi: 10.1038/s41598-019-41177-2
- 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
- Nishio, K., Nakamura, T., Koh, Y., Suzuki, T., Fukumoto, H., and Saijo, N. (1999). Drug resistance in lung cancer. *Curr. Opin. Oncol.* 11, 109–115. doi: 10.1097/00001622-199903000-00006
- Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W., et al. (2015). limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* 43:e47. doi: 10.1093/nar/gkv007
- Rizvi, H. (2018). Molecular determinants of response to anti-programmed cell death (PD)-1 and anti-programmed death-ligand 1 (PD-L1) blockade in patients with non-small-cell lung cancer profiled with targeted next-generation sequencing. *J. Clin. Oncol.* 36, 633–641. doi: 10.1200/JCO.2017.75.3384
- Robin, X., Turck, N., Hainard, A., Tiberti, N., Lisacek, F., Sanchez, J. C., et al. (2011). pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics* 12:77. doi: 10.1186/1471-2105-12-77
- Romero, D. (2019). TMB is linked with prognosis. *Nat. Rev. Clin. Oncol.* 16:336. doi: 10.1038/s41571-019-0206-204
- Samstein, R. M. (2019). Tumor mutational load predicts survival after immunotherapy across multiple cancer types. *Nat. Genet.* 51, 202–206. doi: 10.1038/s41588-018-0312-8
- Samstein, R. M., Lee, C. H., Shoushtari, A. N., Hellmann, M. D., Shen, R., Janjigian, Y. Y., et al. (2019). Tumor mutational load predicts survival after immunotherapy across multiple cancer types. *Nat. Genet.* 51, 202–206.
- Schreiber, R. D., Old, L. J., and Smyth, M. J. (2011). Cancer immunoeediting: integrating immunity's roles in cancer suppression and promotion. *Science* 331, 1565–1570. doi: 10.1126/science.1203486
- Sequist, L. V., Bell, D. W., Lynch, T. J., and Haber, D. A. (2007). Molecular predictors of response to epidermal growth factor receptor antagonists in non-small-cell lung cancer. *J. Clin. Oncol.* 25, 587–595. doi: 10.1200/JCO.2006.07.3585
- Shih, J. Y., Gow, C. H., and Yang, P. C. (2005). EGFR mutation conferring primary resistance to gefitinib in non-small-cell lung cancer. *N. Engl. J. Med.* 353, 207–208. doi: 10.1056/NEJM200507143530217
- Smyth, G. K. (2004). Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Stat. Appl. Genet. Mol. Biol.* 3:Article3. doi: 10.2202/1544-6115.1027
- Song, Z., Wu, Y., Yang, J., Yang, D., and Fang, X. (2017). Progress in the treatment of advanced gastric cancer. *Tumour. Biol.* 39:1010428317714626. doi: 10.1177/1010428317714626
- Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., et al. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U.S.A.* 102, 15545–15550. doi: 10.1073/pnas.0506580102
- Suresh, K., Naidoo, J., Lin, C. T., and Danoff, S. (2018). Immune checkpoint immunotherapy for non-small cell lung cancer: benefits and pulmonary toxicities. *Chest* 154, 1416–1423. doi: 10.1016/j.chest.2018.08.1048
- Szekely, G. J., and Rizzo, M. L. (2005). Hierarchical clustering via joint between-within distances: extending ward's minimum variance method. *J. Classif.* 22, 151–183. doi: 10.1007/s00357-005-0012-9
- Togasaki, K., Sukawa, Y., Kanai, T., and Takaishi, H. (2018). Clinical efficacy of immune checkpoint inhibitors in the treatment of unresectable advanced or recurrent gastric cancer: an evidence-based review of therapies. *Onco. Targets Ther.* 11, 8239–8250. doi: 10.2147/OTT.S152514
- Tosoni, A., Franceschi, E., Pasquini, E., Lanese, A., Donini, E., Foschini, M. P., et al. (2017). Immunotherapy in head and neck cancer: evidence and perspectives. *Immunotherapy* 9, 1351–1358. doi: 10.2217/imt-2017-0125

- Wang, F., Wei, X. L., Wang, F. H., Xu, N., Shen, L., Dai, G. H., et al. (2019). Safety, efficacy and tumor mutational burden as a biomarker of overall survival benefit in chemo-refractory gastric cancer treated with toripalimab, a PD-1 antibody in phase Ib/II clinical trial NCT02915432. *Ann. Oncol.* 30, 1479–1486. doi: 10.1093/annonc/mdz197
- Xiao, Y., Ma, D., Zhao, S., Suo, C., Shi, J., Xue, M. Z., et al. (2019). Multi-omics profiling reveals distinct microenvironment characterization and suggests immune escape mechanisms of triple-negative breast cancer. *Clin. Cancer Res.* 25, 5002–5014. doi: 10.1158/1078-0432.CCR-18-3524
- Zhang, D., He, W., Wu, C., Tan, Y., He, Y., Xu, B., et al. (2019). Scoring system for tumor-infiltrating lymphocytes and its prognostic value for gastric cancer. *Front. Immunol.* 10:71. doi: 10.3389/fimmu.2019.00071

Conflict of Interest: SM was employed by the YDILife Academy of Sciences.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Yang, Yu, Han, Sun, Mo and Hu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.