Article

Construction of an M1 macrophage-related lncRNA signature for predicting the tumor immune microenvironment
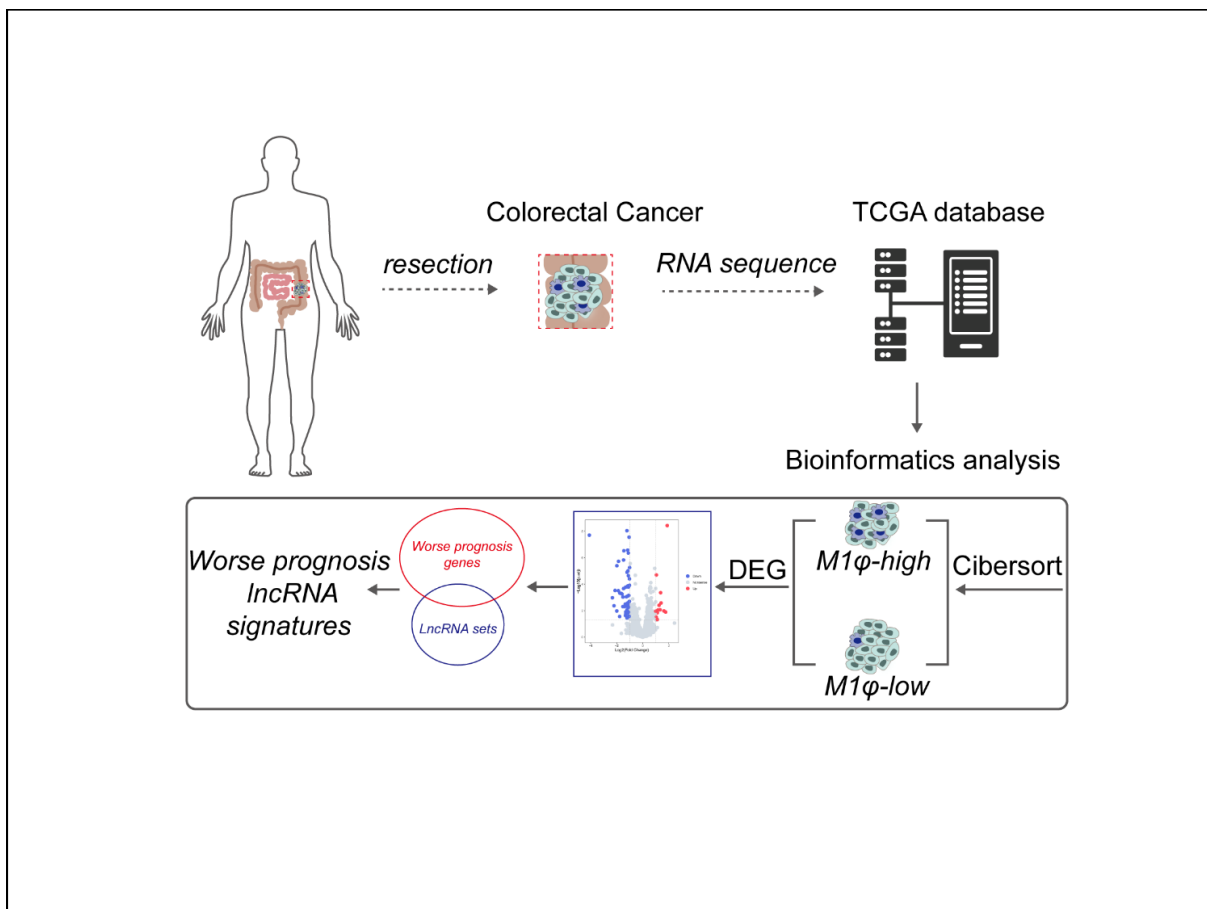
Qi Wu, Yiming Liu, Qingsong Hu, and Huihui Wu ✉

School of Basic Medical Sciences, Division of Life Sciences and Medicine, University of Science and Technology of China, Hefei 230026, China

✉Correspondence: Huihui Wu, E-mail: wuhuihui@ustc.edu.cn

Graphical abstract

Bioinformatic analysis screened a 7-lncRNA signature that predicted M1 macrophage infiltration and prognosis.

Public summary

■ A total of 7 prognosis-related lncRNAs were screened and constructed as a prediction signature.

■ The 7-lncRNA signature precisely predicts the tumor-associated immune microenvironment.

■ The tumor mutation landscape was depicted based on the 7-lncRNA signature.

Construction of an M1 macrophage-related lncRNA signature for predicting the tumor immune microenvironment

Qi Wu, Yiming Liu, Qingsong Hu, and Huihui Wu ✉

School of Basic Medical Sciences, Division of Life Sciences and Medicine, University of Science and Technology of China, Hefei 230026, China

✉Correspondence: Huihui Wu, E-mail: wuhuihui@ustc.edu.cn

Abstract: Long noncoding RNAs (lncRNAs) are considered crucial molecules associated with the tumor microenvironment (TME) and tumor immune microenvironment (TIM). Macrophages are important members of the immune system, and M1 macrophage function-associated lncRNAs still need to be further investigated. In this study, a lncRNA signature was constructed based on transcriptome differences between high and low M1 macrophage infiltration cohorts. This lncRNA signature included seven lncRNAs: LINC01494, ZDHHC20-IT1, LINC01450, LINC00871, EVX1-AS, KIF25-AS and AADACL2-AS1, and all of them were upregulated in patients lacking M1 macrophages, indicating their roles in inhibiting macrophage infiltration and polarizing to the M1 subtype, leading to an immune exclusion TME, which has been demonstrated to be closely correlated with poor prognosis. This lncRNA signature not only predicted undesirable clinical outcomes but was also associated with the immunosuppressive environment of the tumor region, which is mediated by hindering antigen presentation and processing progress. In addition, the predictive value of this lncRNA signature for immune checkpoint inhibition (ICI) therapy was also evaluated, which further enriched and strengthened the power of lncRNAs in predicting the immunotherapy response rate.

Keywords: colon cancer; M1 macrophage; lncRNA; tumor immune microenvironment; immune checkpoint therapy

CLC number: R735.3   **Document code:** A

1   Introduction

A total of 1.9 million colon cancer patients were newly diagnosed in 2020, causing over 900,000 deaths, ranking second in all cancer-related deaths[1]. The incidence rate continues to rise around the world. Studies have identified that bad lifestyles and habits, such as smoking[2, 3], high-fat diet[4] and alcoholism[5], increase the incidence risk of colon cancer. Apart from external environmental factors, approximately 2% to 5% of colon cancer patients are derived from intestinal familial genetic diseases such as Lynch syndrome and familial adenomatous polyposis (FAP)[6, 7]. Gene mutations in hMSH2 and hMLH1, which are responsible for mismatch repair (MMR), were detected in most Lynch syndrome patients, which greatly increases the risk of colon cancer oncogenesis if left untreated[8]. APC loss or mutation is a hazard factor that contributes to Wnt-β-catenin signaling pathway dysregulation, further leading to FAP or even colon cancer[8, 9].

Long noncoding RNAs (lncRNAs) are defined as >200 nt noncoding RNAs that cannot be translated into functional proteins. Recently, lncRNAs, as regulatory molecules, have displayed attractive prospects in the functional regulation of cells, especially in the field of cancer research. Abnormal lncRNA transcription has been found in various types of cancer. For instance, PCGEM1 and PRNCR1 can bind to the androgen receptor and promote prostate cancer progression[10]. In pace with the development and rise of immunotherapy,

light has been shed on the relationships between lncRNAs and the tumor immune microenvironment[11−14].

Macrophages play a crucial role in innate and adaptive immune processes, which work for pathogen or senescent cell clearance and homeostasis maintenance. A classic classification of macrophages is based on specific functions, categorizing macrophages into M1 and M2 subtypes. M1 macrophages are generally responsible for pro-inflammatory functions and kill cancerous cells[15, 16], while M2 macrophages mainly exert repairing or anti-inflammatory activities in injured tissue[15, 17, 18] and pro-tumor function in the background of cancer[19, 20]. Tumor-associated macrophages (TAMs) are now attracting increasing attention. Studies have demonstrated that a complex TME can reprogram macrophages into different subtypes[21−23]. At the same time, the question about where TAMs originate also arose, and many studies have described that TAMs are recruited by chemotactic factors secreted from tumor cells[24], such as CCL2[25], CCL5[26], and CXCL12[27].

In this study, we mainly focused on TME M1 macrophages, and bioinformatic approaches were used to find differences in transcriptome levels between different COAD patient populations, which were divided by M1 proportions. M1 macrophage-related lncRNAs were identified and further screened by survival analysis. The remaining lncRNAs were collected into a signature to test its prognostic value and performance in the TCGA-COAD cohort and other independent

validation datasets.

2 Materials and methods

2.1 Data collection

Count and fragments per kilobase of per million (FPKM) RNA-Seq data of colon cancer (TCGA-COAD, n=453) and corresponding clinical phenotype information were obtained from UCSC XENA (http://xena.ucsc.edu/). Independent validation microarray datasets with clinical information and survival data were downloaded from the GEO database (https://www.ncbi.nlm.nih.gov/geo/) under accession numbers GSE14333 (n=226, https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE14333), GSE17537 (n=55, https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE17537), GSE17536 (n=177, https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE17536), GSE17538 (n=244, https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE17538), and GSE39582 (n=585, https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE39582).

2.2 Tumor microenvironment infiltrated immune cell proportion

The CIBERSORT algorithm[28] was used to deconvolute immune cell proportions from bulk RNA-Seq data based on the LM22 immune cell signature, which contains B cells (naive and memory), plasma cells, CD8 T cells, CD4 T cells, follicular helper T cells, regulatory T cells (Tregs), gamma delta T cells, NK cells (resting and activated), monocytes, macrophages (M0, M1, M2), dendritic cells (resting and activated), mast cells (resting and activated), neutrophils and eosinophils. M1 macrophages were extracted for further evaluation and analysis.

2.3 Differentially expressed lncRNA detection and lncRNA signature construction

DESeq2[29] was used to find differentially expressed genes between the high (M1-Mφ^{high}) and low (M1-Mφ^{low}) M1 macrophage infiltration groups, which were stratified by the survival package and survminer package in R. |logFC| > 1 and adjusted p value ⩽ 0.05 were set to filter upregulated and downregulated genes. LncRNA information was downloaded from the LNCipedia database (https://lncipedia.org/)[30]. The upregulated genes in the M1-Mφ^{low} group were intersected with the lncRNA list to obtain upregulated lncRNAs in the M1-Mφ^{low} group. These lncRNAs were further narrowed by survival analysis to filter survival-correlated lncRNAs for lncRNA signature construction.

2.4 Gene set enrichment analysis and functional enrichment analysis

The score of the lncRNA signature was calculated by GSVA, and differential biological processes (BPs) between the lncRNA score^{high} and lncRNA score^{low} groups were analyzed through gene set enrichment analysis (GSEA) software developed by the Broad Institute[31, 32]. Gene sets were downloaded from Molecular Signatures Database v7.5.1, GSEA official website (http://www.gsea-msigdb.org/gsea/msigdb/index.jsp), C5 ontology gene sets and H hall mark gene sets

were queried for downstream analysis. In addition, mRNAs negatively correlated (cor ⩽ −0.3, p value ⩽ 0.05) with the lncRNA signature were selected for further gene ontology (GO) and functional enrichment analysis, which were performed by the online enrichment tool g:Profiler (https://biit.cs.ut.ee/gprofiler/gost)[33].

2.5 Genomic mutation profile

Mutation information of TCGA-COAD was downloaded from the TCGA official website, and the package maftools[34] was used to visualize the mutation profile. The top 30 mutated genes in lncRNA score^{high} and lncRNA score^{low} patients were extracted for further analyses.

2.6 Statistical analysis

Kaplan–Meier survival plots and log rank p values were carried out to identify different survival probabilities between the M1^{high} and M1^{low} groups, which were determined by the "surv_cutpoint" function in the "survminer" R package. Multivariate Cox regression analyses were performed by the "coxph" function in the "survival" R package. The Spearman correlation coefficient and corresponding p values were calculated by the "cor.test" function. Mann–Whitney tests were performed to distinguish differences between two different groups. Statistical analyses were performed using R v4.1.3 with the necessary R packages and GraphPad Prism 7.0.

3 Results

3.1 Identification of differentially expressed lncRNAs between M1-Mφ^{high} and M1-Mφ1^{low}

To discover the relationship between infiltration of M1 macrophages in the tumor microenvironment and patient prognosis, the M1 macrophage (M1-Mφ) proportion of each TCGA-COAD patient calculated by Cibersort was grouped into two cohorts, the M1-Mφ^{low} (n=47) group and the M1-Mφ^{high} (n=369) group, which were divided by the "surv_cutpoint" function from the survival package in R (cutpoint=0.007690675). Patients with lower M1 macrophage infiltration had significantly shorter overall survival times than patients with higher M1 macrophage infiltration (Fig. 1a, log rank p=0.0082). The same result was obtained in another independent validation dataset (Fig. 1b, GSE14333, log rank p=0.0068, cutpoint=0.0453332; Fig. 1c, integrated dataset of GSE17536 and GSE17538, log rank p=0.0472; Fig. 1d, GSE39582, cutpoint=0.003866829, log rank p=0.0088, cutpoint=0.005733291). To investigate the differences between the M1-Mφ^{low} and M1-Mφ^{high} groups, differentially expressed genes were calculated by the DESeq2 package, and a total of 1634 DEGs were detected, which contained 194 upregulated genes and 1440 downregulated genes in the M1-Mφ^{low} group. Then, differentially expressed lncRNAs (DElncRNAs) were extracted by intersecting DEGs with the lncRNA list from the LNCipedia database, and a total of 71 DElncRNAs were obtained (Fig. 1e), which contained 13 upregulated lncRNAs and 58 downregulated lncRNAs (Fig. 1f, Wilcoxon rank sum test). Here, upregulated lncRNAs in the M1-Mφ^{low} group were selected for further analysis, and the overexpression of these lncRNAs may
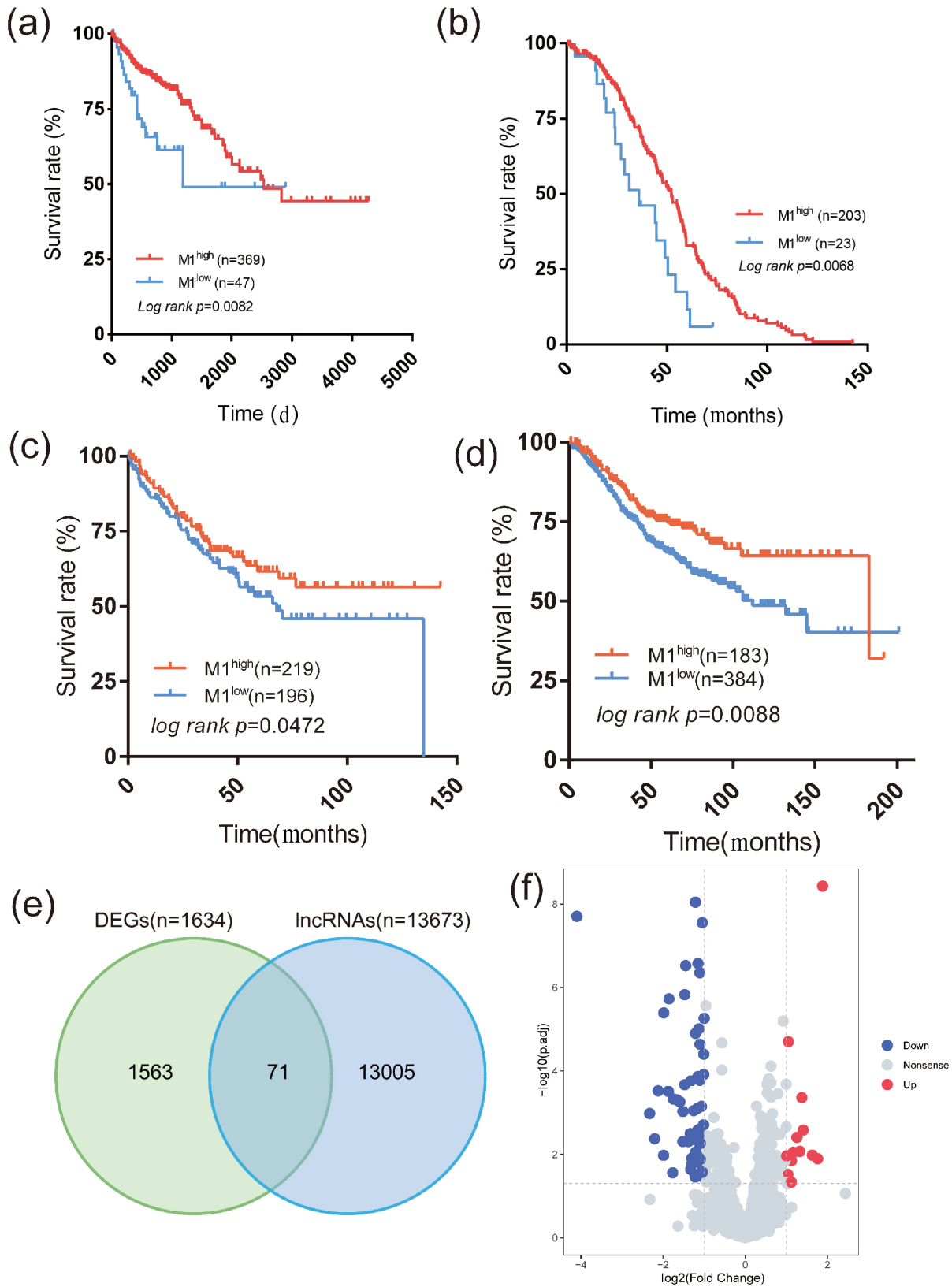
Fig. 1. M1-Mφ predicts the clinical prognosis of colon cancer patients and DElncRNA detection. (a–d) Kaplan–Meier curve of TCGA-COAD dataset, GSE14333, integrated dataset of GSE17536 and GSE17538, GSE39582, respectively, stratified by M1 macrophage proportion. (e) Venn diagram shows DElncRNAs through intersecting DEGs (n=1634) with the lncRNA list (n=13076), 71 DElncRNAs containing 58 downregulated lncRNAs and 13 upregulated lncRNAs. (f) Volcano plot of DElncRNAs, with upregulated lncRNAs shown in red (n=13), downregulated lncRNAs shown in blue (n=58, Wilcoxon rank sum test) and nonsense lncRNAs shown in gray (n=1180, Wilcoxon rank sum test).

be responsible for lower M1 macrophage infiltration in the TME.

3.2 lncRNA signature construction and evaluation

Survival analysis of 13 upregulated lncRNAs was carried out. 7 lncRNAs stood out (Fig. 2a–g): LINC01494 (log rank p=0.0011, cutpoint=0.04112853), ZDHHC20-IT1 (p=0.0043, cutpoint=0.3504403), LINC01450 (p=0.012, cutpoint= 0.01601357), LINC00871 (p=0.0013, cutpoint=0.081546), EVX1-AS (p=0.034, cutpoint=1.016255), KIF25-AS1 (p=0.0013, cutpoint=0.07708173), AADACL2-AS1 (p=0.0037, cutpoint=0.1166067), as their overexpression significantly correlated with shorter survival time. These lncRNAs were collected into a lncRNA signature for further evaluations and studies.

To assess the lncRNA signature performance on prognosis prediction, the lncRNA signature score of each TCGA-COAD patient was calculated by the GSVA package in R. All scored patients were stratified into 2 groups based on the median lncRNA signature score (median score = −0.04486603), and the Kaplan–Meier curve showed that patients with higher lncRNA signature scores had significantly shorter overall survival times (Fig. 2h, left, log rank p=0.00024). Additionally, in the lncRNA score[high] group, patients had lower infiltration of M1 macrophages, suggesting that the lncRNA signature score is negatively correlated with M1 macrophages in the TME (Fig. 2i, p<0.0001, Mann–Whitney test). In the validation cohort of GSE17537, the lncRNA signature score also correlated with a shorter overall survival time (Fig. 2h, right, log rank p=0.011, cutpoint=−0.4367792).

Multivariate Cox regression analysis of the lncRNA signature showed that it served as an independent risk factor that significantly correlated with patient prognosis (Fig. 2j). A lower lncRNA score acted as a protective factor, which is consistent with earlier survival analysis results. In addition, with the progression of colon cancer stage, the lncRNA score increased (Fig. 2k). The analysis above indicated that the lncRNA signature is a significant risk factor correlated with patient prognosis.

3.3 lncRNA signature correlates with immune inertness

GSEA was performed to detect differentially activated pathways or biological processes between the lncRNA score[high] and lncRNA score[low] groups. Notably, in 50 hallmark gene sets, IL2-STAT5 signaling (NES=−2.12, FDR q val=0.001), interferon gamma response (NES=−2.11, FDR q val=0.000), IL6-JSK-STAT3 signaling (NES=−2.08, FDR q val=0.001), interferon alpha response (NES=−1.90, FDR q val=0.008) and complement (NES=−2.09, FDR q val=0.001) were significantly enriched in the lncRNA score[low] group (Fig. 3a), indicating that the lncRNA score[high] cohort was weak in antigen presentation-associated pathways.

The tumor immune microenvironment was assessed via the ESTIMATE package in R. The tumor immune score of TCGA-COAD patients was calculated, and its correlation with the lncRNA signature score was also tested. The results suggested that a higher lncRNA signature score was negatively correlated with the immune score (Fig. 3b, Rs= −0.302, p=3.935×10⁻¹⁰), indicating a lack of immune cells within the

TME of colon cancer. Commonly, lncRNAs exert their functions by manipulating mRNAs to regulate corresponding downstream pathways. Hence, we calculated the relationships between protein-coding gene expression and the lncRNA signature score. mRNAs with a Spearman correlation coefficient ⩽ −0.3 and p ⩽ 0.05 (n=141) were removed for GO enrichment and functional enrichment analyses. Notably, the results showed that these mRNAs were largely enriched in immune-associated reactome pathways (Fig. 3c, Fisher's exact test) and biological processes (Fig. 3d, Fisher's exact test), indicating that members in the lncRNA signature may impede immune-associated gene expression and function.

3.4 lncRNA signature negatively correlates with antigen presentation and processing and immune cell chemotaxis

Immune cell recruitment to the TME depends on tumor-related antigens that can be recognized by antigen-presenting cells such as DCs and macrophages. In addition, chemokines secreted by tumor cells or stromal cells in the TME mediate extratumor boundary immune cell chemotaxis. Antigen processing- and presentation-related gene sets were included to evaluate their relationship with the lncRNA signature. The results showed that the higher the lncRNA set score was, the lower the score for antigen presentation and processing (Fig. 4a, Rs = −0.307, p=2.054×10⁻¹⁰). Additionally, the chemotaxis score of each sample in the lncRNA score[high] and lncRNA score[low] groups was assessed, and the results showed that the lncRNA score[high] group had a lower chemotaxis score, which indicated an immune exclusion microenvironment (Fig. 4b, Mann–Whitney test). These results provide additional evidence that patients with a high lncRNA score may also have an immune-cold TME. In addition, we tested the relationship between the lncRNA score and genes that positively regulate antigen processing and presentation and positively regulate chemotaxis. We found that these immune-positive-regulating genes were mostly negatively correlated with the lncRNA score (Fig. 4c, d), which further confirmed the inertness of the immune-related pathway in lncRNA score[high] patients.

3.5 lncRNA signature predicted immune checkpoint inhibition therapy

We then investigated the performance of the lncRNA signature in predicting immune checkpoint inhibition (ICI) therapy outcomes. The immune checkpoint genes PD1 and PD-L1 were extracted, and their correlations with the lncRNA signature score were calculated. The differences in PD1 and PD-L1 expression between the lncRNA score[high] and lncRNA score[low] groups were tested, which revealed that patients with lower lncRNA scores also had higher PD1 and PDL1 expression (Fig. 5a, Mann–Whitney test). To verify the predictive value of the lncRNA score for ICI responses, we queried the immunophenoscore (iPS) of TCGA-COAD patients in The Cancer Immunome Atlas (TCIA) database. LncRNA set score[high] patients had significantly lower PD1 blocker scores (Fig. 5b, Student's t test), indicating an undesirable response to ICI therapy compared to the other group of patients. Then,
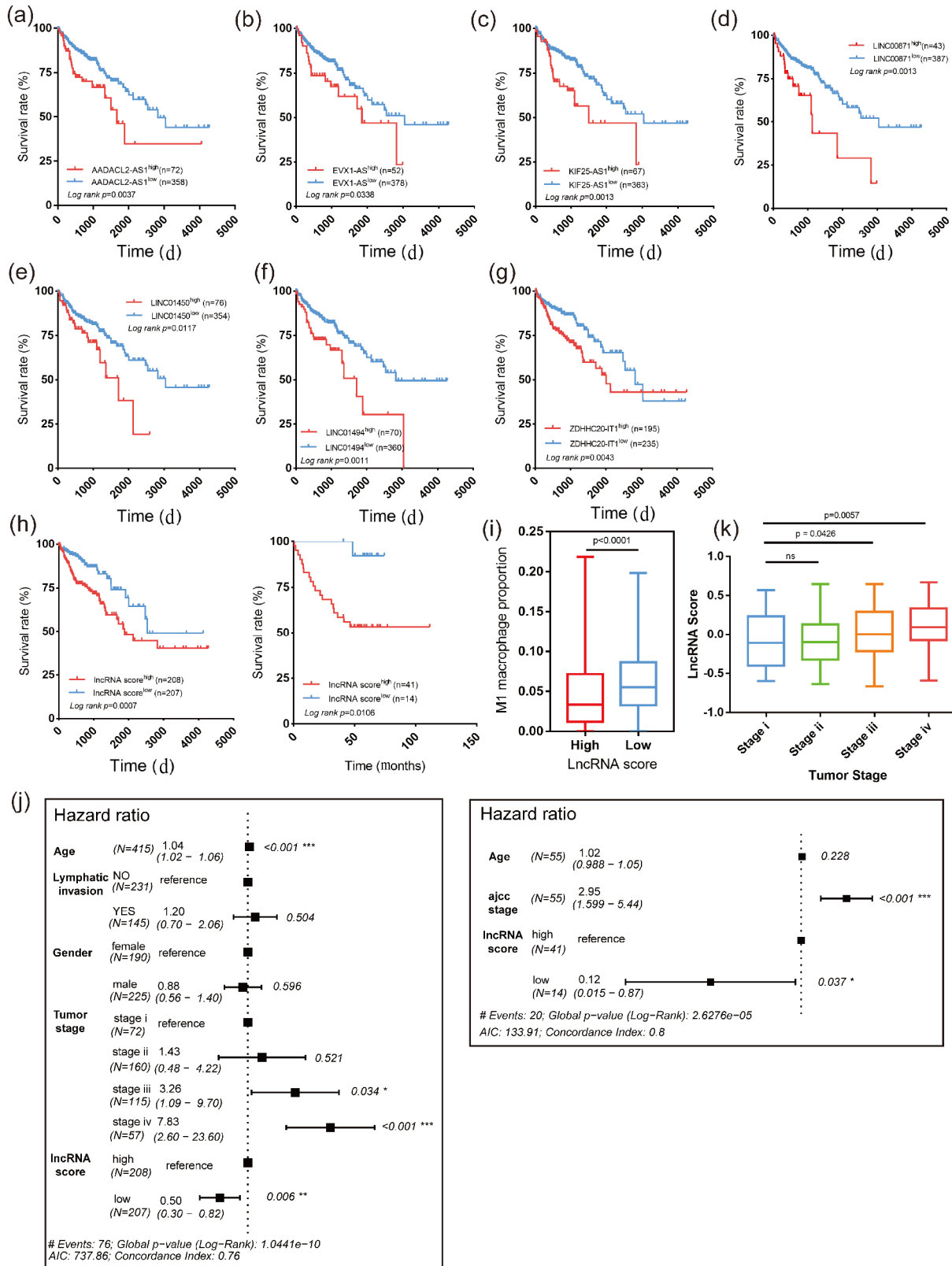
Fig. 2. The constructed lncRNA signature is closely associated with clinical outcomes. (a–g) Kaplan–Meier curve of 7 survival-related lncRNAs. (a) AA-DACL2-AS (log rank *p*=0.0037); (b) EVX1-AS (log rank *p* =0.0338); (c) KIF25-AS1 (log rank *p* =0.0013); (d) LINC00871 (log rank *p* =0.0013); (e) LINC01450 (log rank *p* =0.0117); (f) LINC01494 (log rank *p* =0.0011); (g) ZHDDC20-IT1 (log rank *p* = 0.0043). (h) Kaplan–Meier curve of the TCGA-COAD cohort (left, log rank *p*=0.0007) and an independent validation dataset GSE17537 (log rank *p*=0.0106). (i) Box plot of the M1 macrophage infiltration proportion in the lncRNAhigh and lncRNAlow groups (*p*<0.0001, Mann–Whitney test, data are represented as the mean ± SD). (j) Forest plot of the multivariate Cox regression analysis of the TCGA-COAD cohort (left) and the independent validation cohort GSE17537 (right). (k) Box plot of the lncRNA score of different COAD tumor stages (Mann–Whitney test, compared to stage i COAD tumors, data are represented as the mean ± SD).
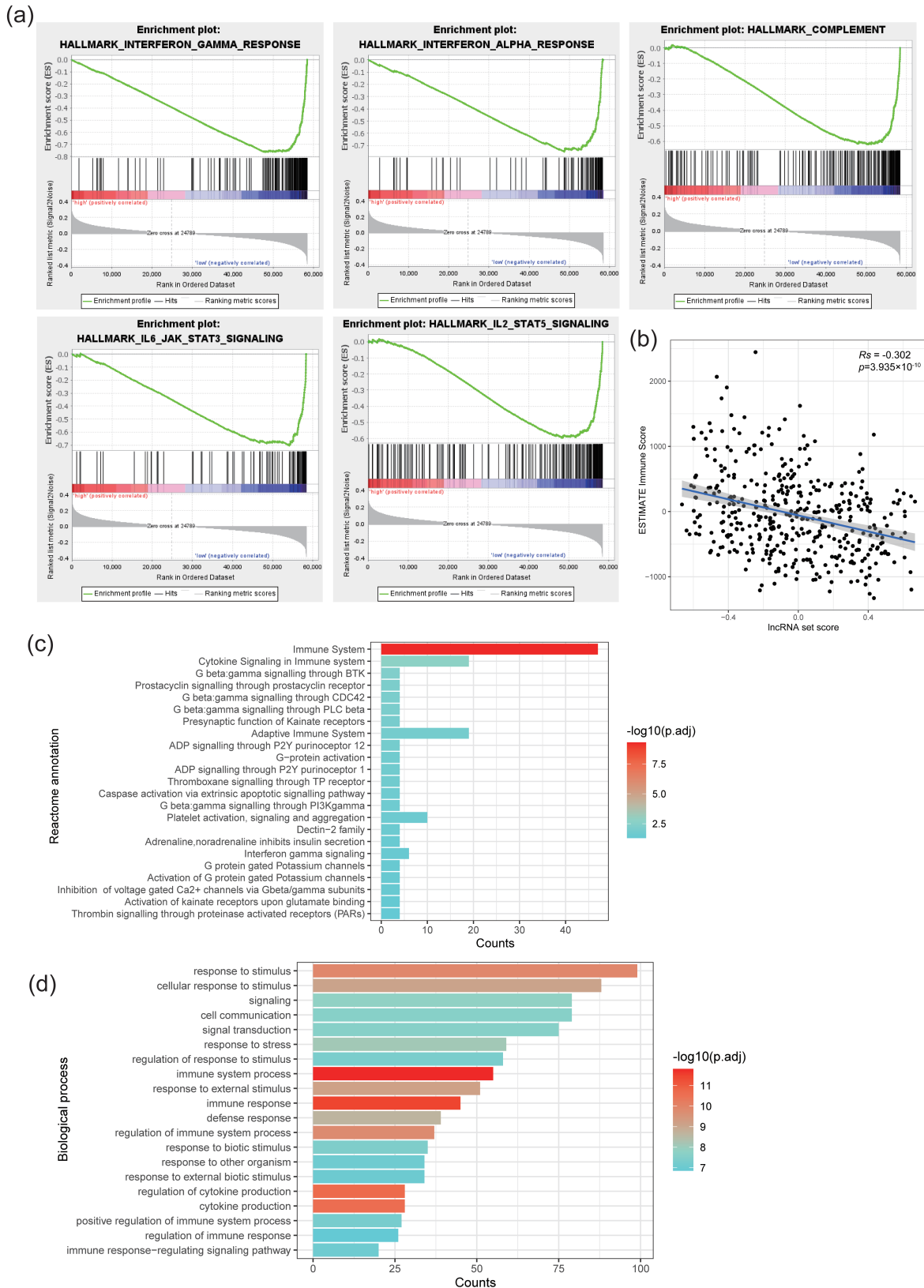
Fig. 3. The tumor immune microenvironment is negatively correlated with the lncRNA signature. (a) GSEA plots of differentially enriched cancer hallmark pathways between the lncRNA score[high] and lncRNA score[low] cohorts. (b) Spearman's correlation between lncRNA score and ESTIMATE immune score (Rs=−0.302, p=3.935×10[−10]). (c) Reactome enrichment plot of mRNAs that negatively correlated with the lncRNA set (n=23, Fisher's exact test). (d) GO:BP enrichment plot of mRNAs that negatively correlated with the lncRNA signature. The top 20 enriched biological processes are displayed (Fisher's exact test).
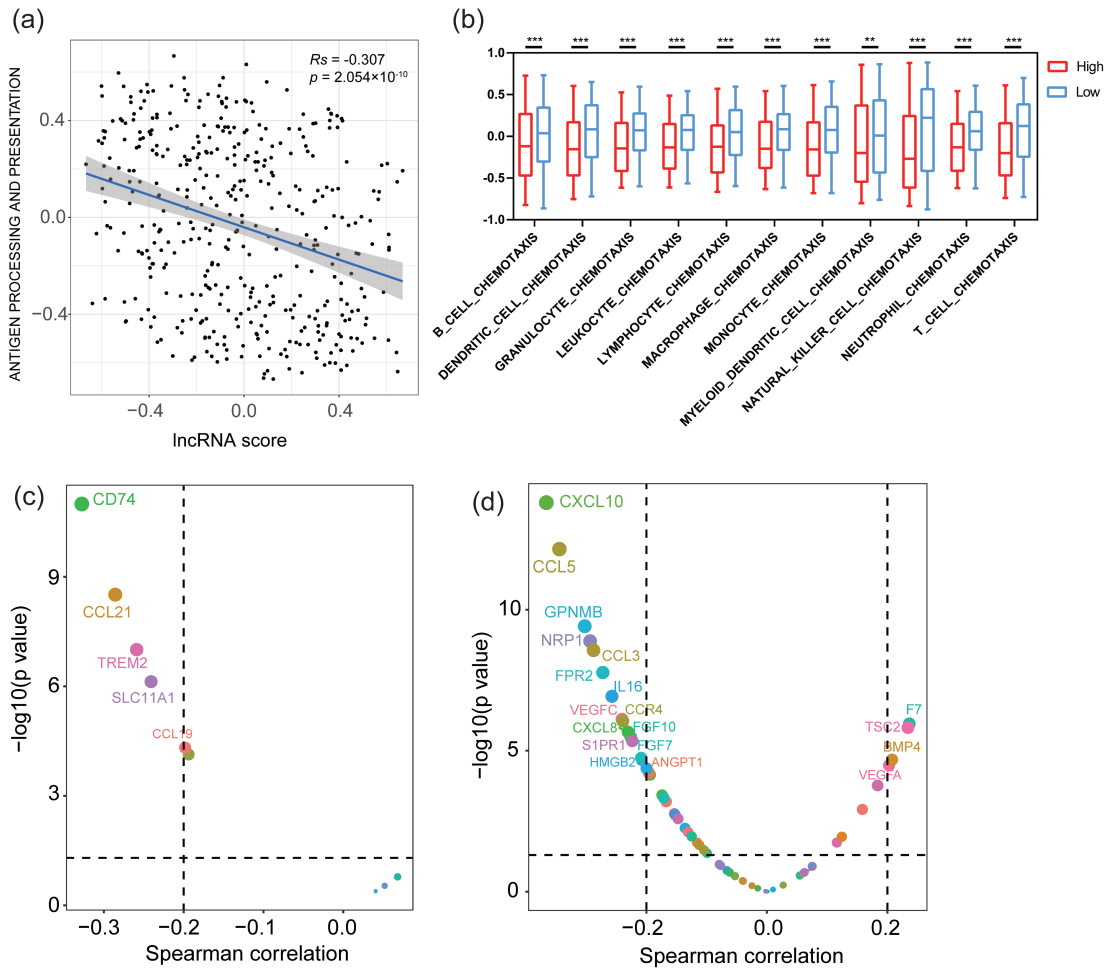
Fig. 4. Antigen-presenting processes were downregulated in the lncRNA score[high] group. (a) Spearman correlation between lncRNA score and antigen presenting and processing score (Rs=−0.307, p=2.054×10⁻¹⁰). (b) Box plot of differences in chemotaxis score between the high lncRNA group and the low lncRNA group (**: p<0.05, ***: p<0.0001, Mann–Whitney test, data are represented as the mean ± SD). (c) Spearman correlation of lncRNA score and genes that positively regulate antigen processing and presentation. (d) Spearman correlation of lncRNA score and genes that positively regulate chemotaxis.

we grouped TCGA-COAD patients into PD1[high] and PD1[low] groups and PD-L1[high] and PD-L1[low] groups based on the median expression of PD1 and PD-L1, respectively. Then, we combined the PD1/PD-L1 groups with the previous lncRNA score groups to investigate the cross impact of the lncRNA score and immune checkpoint gene expression on patient prognosis. The results suggested that patients with low lncRNA scores and high PD1 (Fig. 5c, left, log rank p=0.0042) or PD-L1 (Fig. 5c, right, log rank p=0.0046) expression had prolonged survival times compared with the remaining 3 groups, while for patients with high lncRNA scores, even if they had higher immune checkpoint expression, their prognosis was still undesirable.

Studies have reported the relationship between MSI subtypes in colon cancer and their potential value in predicting ICI therapy outcomes[35−37]. MSI information of TCGA-COAD patients was downloaded and combined with the lncRNA signature score. In the lncRNA score[high] group, only 10 patients had the MSI-H phenotype, while in the lncRNA score[low] group, 59 patients had the MSI-H phenotype, suggesting that the lncRNA score[high] and lncRNA score[low] groups showed different MSI phenotypes (Fig. 5d, p<0.05, chi-square test).

These results further support that the lncRNA signature score might serve as a predictor for ICI therapy response.

3.6 LncRNA score indicates tumor mutation landscape differences

Tumor mutation burden (TMB) is considered a biomarker for predicting ICI therapy outcomes[38, 39], and tumors with higher TMB tend to load more neoantigens that can be recognized by immune cells such as T cells[40]. The TMB of lncRNA score[high] and lncRNA score[low] samples was calculated via the maftools package, and the results showed that TMB in the lncRNA score[low] group was significantly higher than that in the lncRNA score[high] group (Fig. 6a, p = 0.0029, Mann–Whitney test). This feature was consistent with the analyses above that a lower lncRNA score is correlated with a higher immune score, indicating a better immune cell infiltration level due to more loading of neoantigens. The genome mutation landscapes of lncRNA score[high] and lncRNA score[low] were inspected and visualized. The top 30 mutated genes were selected and are displayed in Fig. 6b. After excluding shared mutated genes (n=21, APC, TP53, KRAS, TTN, PIK3CA, SYNE1, MUC16, FAT4, RYR2, DNAH5, FLG, ZFHX4,
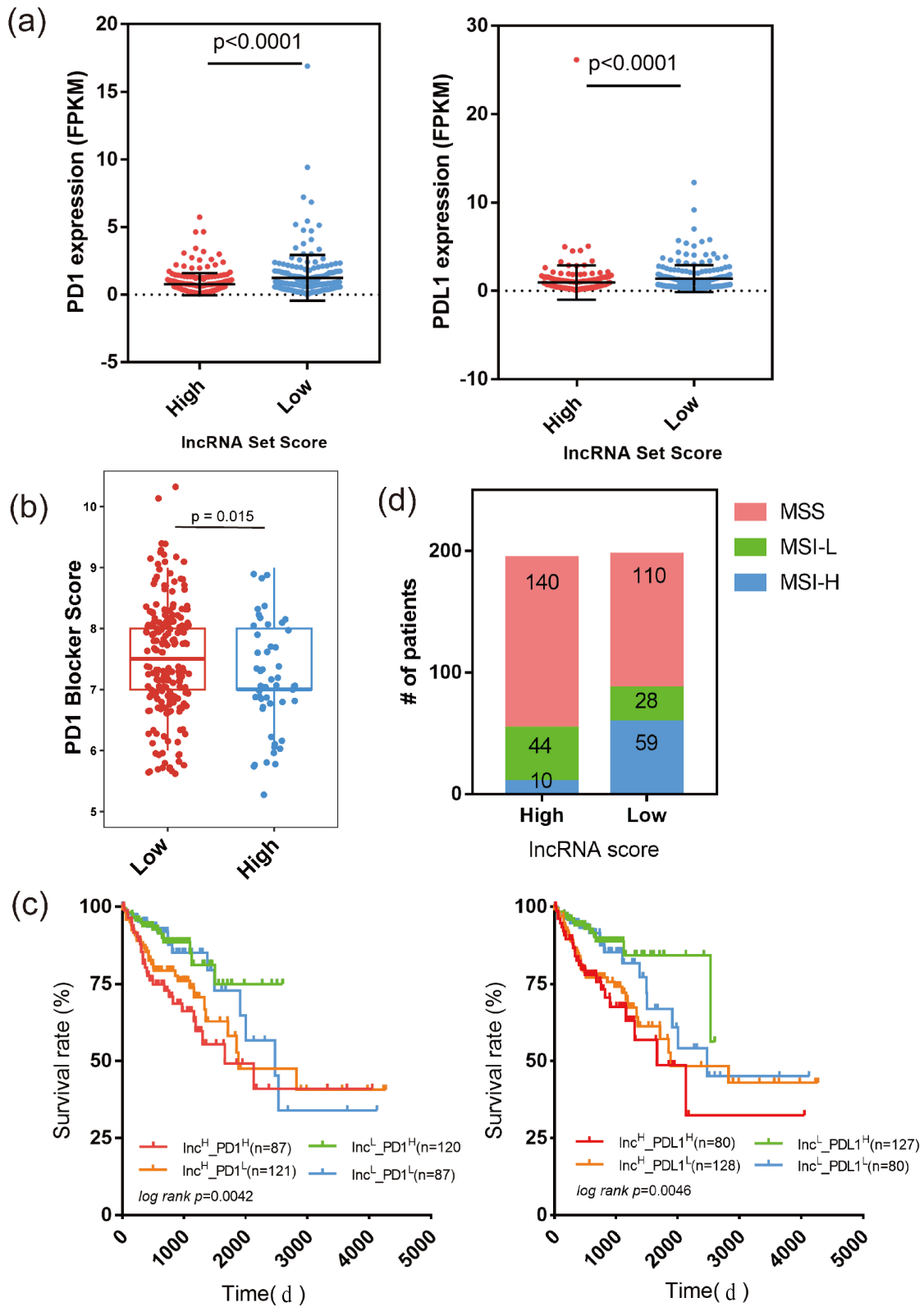
Fig. 5. LncRNA signature predicts ICI therapy outcomes. (a) Swarm diagram shows the differences in PD1 (left) and PD-L1 (right) between the lncRNA score[high] and lncRNA score[low] groups (data are represented as the mean ± SD, Mann–Whitney test). (b) Predictive value of PD1 blocker efficiency between the lncRNA set score[high] and score[low] groups (p=0.015, Student's t test) (c) Kaplan–Meier curve of four patient groups stratified by lncRNA score and PD1 expression (left) and lncRNA score and PD-L1 expression (right). (d) Stacked histograms of MSS, MSI-L and MSI-H patient numbers in the lncRNA scorehigh and lncRNA scorelow groups (p<0.05, chi-square test).
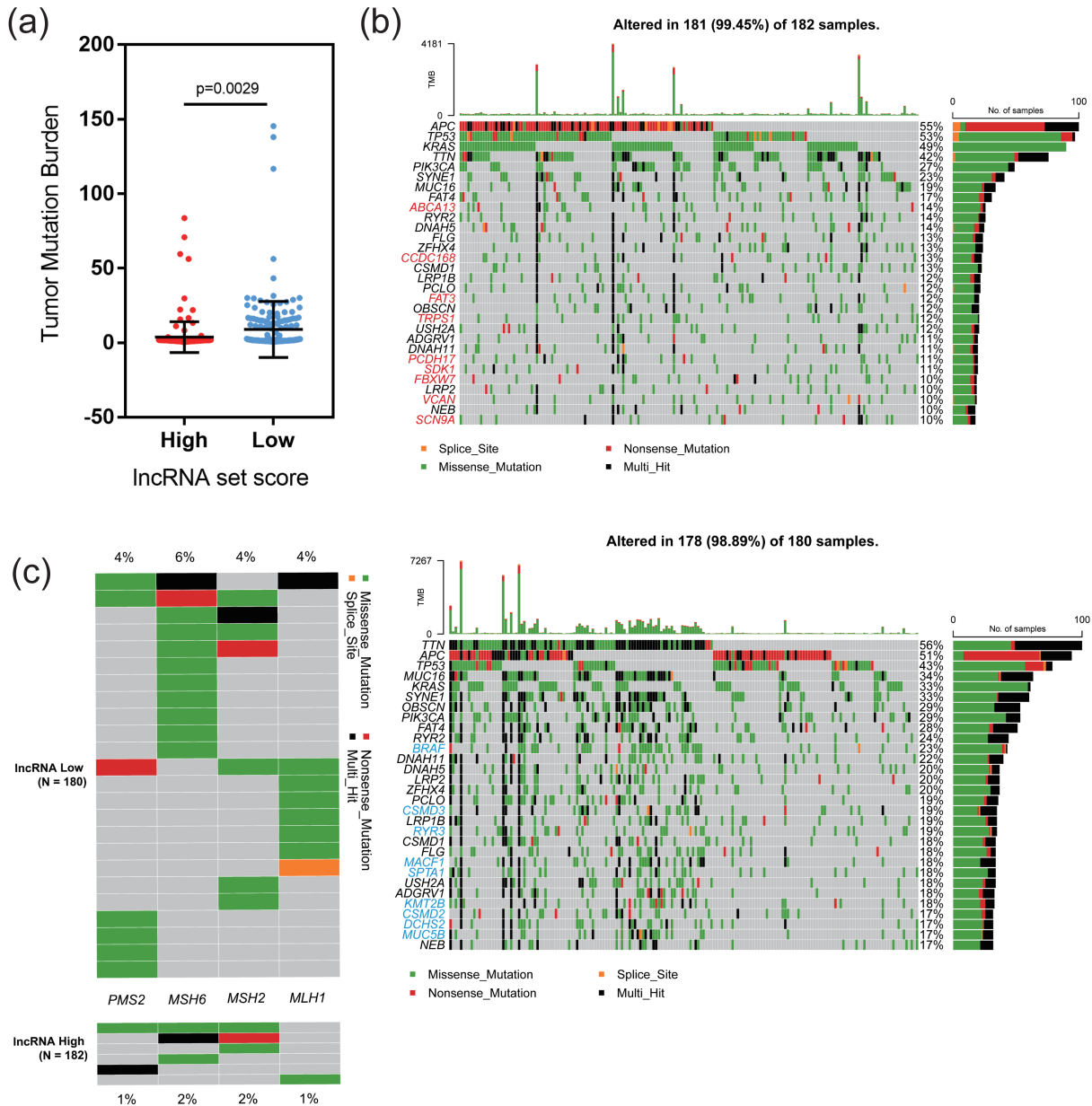
Fig. 6. Relationship between tumor mutation burden and lncRNA signature. (a) Swarm diagram shows the differences in TMB between the lncRNA scorehigh and lncRNA scorelow groups (data are represented as the mean ± SD, Mann–Whitney test). (b) Waterfall plot shows the top 30 mutated genes in the lncRNA scorehigh (upper panel) and lncRNA scorelow groups (lower panel). Genes in black suggest shared mutated genes across the lncRNA scorehigh and lncRNA scorelow groups. Genes in red were highly mutated in the lncRNA scorehigh group. Genes in cyan were highly mutated in the lncRNA scorelow group. (c) Comparing MSI-related gene mutations between the lncRNA scorehigh and lncRNA scorelow groups.

CSMD1, LRP1B, PCLO, USH2A, OBSCN, DNAH11, AD-GRV1, LRP2 and NEB), the remaining genes were extracted. In the lncRNA score^{high} group (n=9): ABCA13, CCDC168, FAT3, TRPS1, PCDH17, SDK1, FBXW7, VCAN, SCN9A. lncRNA score^{low} group (n=9): BRAF, CSMD3, RYR3, MACF1, SPTA1, KMT2B, DCHS2, CSMD2, MUC5B. MSI-related gene mutations rates were queried as well. As expected, MSI-associated genes were more frequently mutated in the lncRNA score^{low} group (Fig. 6c, MLH1: 4% in the lncRNA score^{low} group, 1% in the lncRNA score^{high} group; MSH2: 4% in the lncRNA score^{low} group, 2% in the lncRNA score^{high} group; MSH6: 6% in the lncRNA score^{low} group, 2%

in the lncRNA score^{high} group; PMS2: 4% in the lncRNA score^{low} group, 1% in the lncRNA score^{high} group). This result further proved that frequent MSI-associated gene mutations induced MSI features in the lncRNA score^{low} group, leading to a better prognosis.

4 Discussion

Immune cell infiltration in the TME is closely correlated with tumor progression and has been widely recognized as an important factor in predicting the prognosis of cancer patients. However, clinical identification of immune cells basic-

ally depends on pathologists' diagnosis via hematoxylin and eosin (HE) staining or immunohistochemistry (IHC) staining, which introduces diagnostic bias of individual pathologists. With the rapid increase in transcriptome data, immune cell-specific signatures have been discovered, and their relationship with patient prognosis has been described in detail. In addition, with the development of sequencing techniques, some noncoding RNAs have been detected, and their regulatory functions on protein-coding mRNAs have attracted increasing attention. lncRNA molecular markers have been reported and validated. Recent studies have also discovered that lncRNAs are correlated with tumor-infiltrating immune cells. Huang et al.[12] reported that the lncRNA NKILA induced T-cell death to escape immune killing. Zhou et al.[41] built a B-cell-based lncRNA signature as a prognostic predictor for bladder cancer.

M1 macrophage infiltration in the tumor region is considered a factor that predicts better clinical outcomes[42−44]. In addition, the role of noncoding RNAs in macrophage polarization cannot be ignored[43]. In our work, patients with lower M1 macrophage residence were closely focused on, and their differentially expressed lncRNAs were queried and investigated. Seven significantly upregulated lncRNAs were identified in the M1[low] cohort. As we assumed, this 7-lncRNA signature correlates with worse patient prognosis. When the lncRNA signature score was combined with other clinical characteristics, it still acted as an independent prognostic factor. The functions of individual members in the lncRNA signature were also reported in some articles. LINC01494 promotes glioma development via the miR-122-5p-CCGN1 axis[45], and the overexpression of LINC01494 can inhibit the sponge of miR-122-5p and further upregulate CCGN1, a cell cycle-related gene, causing glioma progression. The relationship between EVX1-AS and colon cancer was predicted in the lncRNADisease v2.0 database (http://www.rnanut.net/lncrnadisease/index.php/home)[46]. However, the role of these lncRNAs in macrophage development and polarization is still unknown, and more experimental evidence is needed to confirm this hypothesis. According to the canonical function of lncRNAs, they may affect macrophage function in a direct way, which regulates macrophage differentiation and polarization[43,46], or in an indirect way, lncRNAs affect tumor cell metabolism or export by exosomes to modulate macrophage functions[47,48].

The colon cancer immune environment was reported as an important feature that predicts ideal outcomes[49, 50]. The ESTIMATE immune score, an index indicating the immune cell infiltration level, was negatively correlated with the lncRNA signature, suggesting an immunodeficient environment in the highly scored lncRNA signature group. We investigated factors that affect TME immune cell infiltration, and the potential cause of such an immune-excluded TME might be mediated by impeding antigen presentation and weakening the recruitment of various immune cells.

ICI therapy is a novel and prospective approach for cancer treatment, and even if limited patients could benefit from ICI[51], it still attracts many researchers' attention. Meanwhile, biomarkers used to evaluate ICI response still need further discovery and assessment. Currently, the immune checkpoint molecules PD1 and PD-L1 and their inhibitors are widely used and investigated in clinical and research situations. Hence, an effect evaluation signature is crucial for predicting ICI therapy outcomes. By combining PD1 or PD-L1 expression with the lncRNA signature score, we identified that low lncRNA score cohort patients had significantly higher expression of PD1 and PD-L1 than patients with higher lncRNA scores. By comparing survival time among the lncRNA score[high]-PD1[high]/PD-L1[high], lncRNA score[high]-PD1[low]/PD-L1[low], lncRNA score[low]-PD1[high]/PD-L1[high] and lncRNA score[low]-PD1[low]/PD-L1[low] groups, patients with low lncRNA scores and high PD1/PD-L1 expression showed better prognosis than the other three groups. Notably, patients with low lncRNA score, even if higher PD1/PD-L1 expression detected, their overall survival time still significantly shorter than low lncRNA score patients, which provided evidence that PD1/PD-L1 expression alone was not enough for predicting ICI therapy outcomes, linking PD1/PD-L1 expression level with lncRNA score could serve a robust predicting result with enhanced precision. MSI, especially MSI-H, as a predictor of better prognosis of colon cancer has been reported[52−54]. MSI colon cancers deficient in the DNA mismatch repair (dMMR) system are responsible for monitoring and rectifying microsatellite-associated genome errors. MSI colon cancer generally possesses mutations in MLH1, MSH2, MSH6 and PMS2, which are crucial for the dMMR system[54]. Loss of function of these four genes leads to mismatch repair dysfunction, and cancer cells are prone to display more neoantigens that can be detected by immune cells. In addition to colon cancer, studies have found that MSI is a general phenomenon across various tumor types, such as endometrial cancer[55], gastric cancer[56], and ovarian cancer[57]. Grouping patients via lncRNA signature score, we found that MSI patients, especially MSI-H patients, were mostly grouped into the lncRNA score[low] group (n=59), while only 10 patients were characterized as MSI-H in the lncRNA score[high] group, indicating better clinical outcomes in lncRNA score[low] patients.

Tumor mutation burden, which is considered a better prognostic factor, was significantly higher in lncRNA score[low] group patients, concordant with previous survival analyses. Additionally, four MSI-related genes (MLH1, MSH2, MSH6 and PMS2) were more frequently mutated in the lncRNA score[low] group, causing MSI in colon cancer and further linked to an ideal clinical prognosis. Notably, FXBW7, a gene that was previously reported to be closely related to anti-PD1 therapy resistance in melanoma, was highly mutated in the lncRNA score[high] group (16% vs 10%)[58]. Loss of function of FXBW7 changed TIM by inhibiting IFN-I and MHC-I expression, and restoration of FXBW7 rescued PD-1 blockade therapy. This result provided evidence that lncRNA score[low] patients may be resistant to ICI therapy due to high mutation of FXBW7, which is consistent with the analysis that the lncRNA signature and PD-1/PD-L1 expression together could predict ICI therapy outcomes.

With the development of single-cell RNA sequencing, the classification of M1 and M2 macrophages is now overly simple and rough. More newly discovered and defined subpopulations of macrophages have been discovered and investigated. For example, SPP1⁺ macrophages together with FAP⁺

fibroblasts promote the progression of colorectal cancer[59]. The infiltration of TREM2[+] macrophages remodels an immunosuppressive microenvironment in the tumor region[60−61]. In contrast, FOLR2[+] macrophages could boost the immune response, leading to better clinical outcomes[62−63]. However, the roles of lncRNAs in macrophage development and differentiation are still unclear. Therefore, more experiments and bioinformatic analyses are needed to elucidate how lncRNAs regulate the development and transformation of macrophage subtypes.

Acknowledgements

Conflict of interest

The authors declare that they have no conflict of interest.

Biographies

Qi Wu is currently a master's student at the School of Basic Medical Sciences, Division of Life Sciences and Medicine, University of Science and Technology of China, under the supervision of Prof. Qingsong Hu. His research mainly focuses on bioinformatics and tumor infiltrated macrophages.

Huihui Wu received her Ph.D. degree in Cell Biology from the University of Science and Technology of China. She is currently a postdoctor at the University of Science and Technology of China. Her research interests include the tumor microenvironment and its role in tumor progression.

References

[1] Sung H, Ferlay J, Siegel R L, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.,* **2021**, *71* (3): 209–249.

[2] Gram I T, Park S Y, Wilkens L R, et al. Smoking-related risks of colorectal cancer by anatomical subsite and sex. *Am. J. Epidemiol.,* **2020**, *189* (6): 543–553.

[3] Akter S, Islam Z, Mizoue T, et al. Smoking and colorectal cancer: A pooled analysis of 10 population-based cohort studies in Japan. *Int. J. Cancer,* **2021**, *148* (3): 654–664.

[4] Tarasiuk A, Mosińska P, Fichna J. The mechanisms linking obesity to colon cancer: An overview. *Obes. Res. Clin. Pract.,* **2018**, *12* (3): 251–259.

[5] Tverdal A, Høiseth G, Magnus P, et al. Alcohol consumption, HDL-cholesterol and incidence of colon and rectal cancer: a prospective cohort study including 250,010 participants. *Alcohol Alcoholism,* **2021**, *56* (6): 718–725.

[6] Lynch H T, Snyder C L, Shaw T G, et al. Milestones of Lynch syndrome: 1895-2015. *Nat. Rev. Cancer,* **2015**, *15* (3): 181–194.

[7] Li J Y, Wang R, Zhou X, et al. Genomic and transcriptomic profiling of carcinogenesis in patients with familial adenomatous polyposis. *Gut,* **2020**, *69* (7): 1283–1293.

[8] Jasperson K W, Tuohy T M, Neklason D W, et al. Hereditary and familial colon cancer. *Gastroenterology,* **2010**, *138* (6): 2044–2058.

[9] Bian J, Dannappel M, Wan C, et al. Transcriptional regulation of Wnt/β-catenin pathway in colorectal cancer. *Cells,* **2020**, *9* (9): 2125.

[10] Yang L, Lin C, Jin C, et al. lncRNA-dependent mechanisms of androgen-receptor-regulated gene activation programs. *Nature,* **2013**, *500* (7464): 598–602.

[11] Li G, Kryczek I, Nam J, et al. LIMIT is an immunogenic lncRNA in cancer immunity and immunotherapy. *Nat. Cell Biol.,* **2021**, *23* (5): 526–537.

[12] Huang D, Chen J, Yang L, et al. NKILA lncRNA promotes tumor immune evasion by sensitizing T cells to activation-induced cell death. *Nat. Immunol.,* **2018**, *19* (10): 1112–1125.

[13] Jiang W X, Pan S Y, Chen X, et al. The role of lncRNAs and circRNAs in the PD-1/PD-L1 pathway in cancer immunotherapy. *Mol. Cancer,* **2021**, *20* (1): 116.

[14] Xu H, Jiang Y, Xu X, et al. Inducible degradation of lncRNA Sros1 promotes IFN-γ-mediated activation of innate immune responses by stabilizing Stat1 mRNA. *Nat. Immunol.,* **2019**, *20* (12): 1621–1630.

[15] Mills C D, Lenz L L, Harris R A. A breakthrough: Macrophage-directed cancer immunotherapy. *Cancer Res.,* **2016**, *76* (3): 513–516.

[16] Choo Y W, Kang M, Kim H Y, et al. M1 macrophage-derived nanovesicles potentiate the anticancer efficacy of immune checkpoint inhibitors. *ACS Nano,* **2018**, *12* (9): 8977–8993.

[17] Li M, Sun X, Zhao J, et al. CCL5 deficiency promotes liver repair by improving inflammation resolution and liver regeneration through M2 macrophage polarization. *Cell. Mol. Immunol.,* **2020**, *17* (7): 753–764.

[18] Vogel D Y S, Glim J E, Stavenuiter A W D, et al. Human macrophage polarization in vitro: Maturation and activation methods compared. *Immunobiology,* **2014**, *219* (9): 695–703.

[19] Lan J Q, Sun L, Xu F, et al. M2 macrophage-derived exosomes promote cell migration and invasion in colon cancer. *Cancer Res.,* **2019**, *79* (1): 146–158.

[20] Yamaguchi T, Fushida S, Yamamoto Y, et al. Tumor-associated macrophages of the M2 phenotype contribute to progression in gastric cancer with peritoneal dissemination. *Gastric Cancer,* **2016**, *19* (4): 1052–1065.

[21] Sharma A, Seow J J W, Dutertre C A, et al. Onco-fetal reprogramming of endothelial cells drives immunosuppressive macrophages in hepatocellular carcinoma. *Cell,* **2020**, *183* (2): 377–394.e21.

[22] Cassetta L, Fragkogianni S, Sims A H, et al. Human tumor-associated macrophage and monocyte transcriptional landscapes reveal cancer-specific reprogramming, biomarkers, and therapeutic targets. *Cancer Cell,* **2019**, *35* (4): 588–602.e10.

[23] Morrissey S M, Zhang F, Ding C L, et al. Tumor-derived exosomes drive immunosuppressive macrophages in a pre-metastatic niche through glycolysis dominant metabolic reprogramming. *Cell Metab.,* **2021**, *33* (10): 2040–2058.e10.

[24] Franklin R A, Liao W, Sarkar A, et al. The cellular and molecular origin of tumor-associated macrophages. *Science,* **2014**, *344* (6186): 921–925.

[25] Zigmond R E, Echevarria F D. Macrophage biology in the peripheral nervous system after injury. *Prog. Neurobiol.,* **2019**, *173*: 102–121.

[26] Dan H, Liu S, Liu J, et al. RACK1 promotes cancer progression by increasing the M2/M1 macrophage ratio via the NF-κB pathway in oral squamous cell carcinoma. *Mol. Oncol.,* **2020**, *14* (4): 795–807.

[27] Boimel P J, Smirnova T, Zhou Z N, et al. Contribution of CXCL12 secretion to invasion of breast cancer cells. *Breast Cancer Res.,* **2012**, *14* (1): R23.

[28] Newman A M, Steen C B, Liu C L, et al. Determining cell type abundance and expression from bulk tissues with digital cytometry. *Nat. Biotechnol.,* **2019**, *37* (7): 773–782.

[29] Love M I, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.,* **2014**, *15* (12): 550.

[30] Volders P J, Anckaert J, Verheggen K, et al. LNCipedia 5: towards a reference set of human long non-coding RNAs. *Nucleic Acids Res.,* **2019**, *47* (D1): D135–D139.

[31] Subramanian A, Tamayo P, Mootha V K, et al. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U.S.A.,* **2005**, *102* (43): 15545–15550.

[32] Mootha V K, Lindgren C M, Eriksson K F, et al. PGC-1α-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nat. Genet.,* **2003**, *34* (3): 267–273.

[33] Raudvere U, Kolberg L, Kuzmin I, et al. g:Profiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update). *Nucleic Acids Res.,* **2019**, *47* (W1): W191–W198.

[34] Mayakonda A, Lin D C, Assenov Y, et al. Maftools: efficient and comprehensive analysis of somatic variants in cancer. *Genome Res.,* **2018**, *28* (11): 1747–1756.

[35] Diao Z L, Han Y X, Chen Y Q, et al. The clinical utility of microsatellite instability in colorectal cancer. *Crit. Rev. Oncol. Hematol.,* **2021**, *157*: 103171.

[36] Schrock A B, Ouyang C, Sandhu J, et al. Tumor mutational burden is predictive of response to immune checkpoint inhibitors in MSI-high metastatic colorectal cancer. *Ann. Oncol.,* **2019**, *30* (7): 1096–1103.

[37] Ganesh K, Stadler Z K, Cercek A, et al. Immunotherapy in colorectal cancer: Rationale, challenges and potential. *Nat. Rev. Gastroenterol. Hepatol.,* **2019**, *16* (6): 361–375.

[38] Chida K, Kawazoe A, Kawazu M, et al. A low tumor mutational burden and PTEN mutations are predictors of a negative response to PD-1 blockade in MSI-H/dMMR gastrointestinal tumors. *Clin. Cancer Res.,* **2021**, *27* (13): 3714–3724.

[39] Chalmers Z R, Connelly C F, Fabrizio D, et al. Analysis of 100,000 human cancer genomes reveals the landscape of tumor mutational burden. *Genome Med.,* **2017**, *9* (1): 34.

[40] Jardim D L, Goodman A, de Melo Gagliato D, et al. The challenges of tumor mutational burden as an immunotherapy biomarker. *Cancer Cell,* **2021**, *39* (2): 154–173.

[41] Zhou M, Zhang Z C, Bao S Q, et al. Computational recognition of lncRNA signature of tumor-infiltrating B lymphocytes with potential implications in prognosis and immunotherapy of bladder cancer. *Brief. Bioinform.,* **2021**, *22* (3): bbaa047.

[42] Xia Y, Rao L, Yao H, et al. Engineering macrophages for cancer immunotherapy and drug delivery. *Adv. Mater.,* **2020**, *32* (40): 2002054.

[43] Mohapatra S, Pioppini C, Ozpolat B, et al. Non-coding RNAs regulation of macrophage polarization in cancer. *Mol. Cancer,* **2021**, *20* (1): 24.

[44] Gunassekaran G R, Poongkavithai Vadevoo S M, Baek M C, et al. M1 macrophage exosomes engineered to foster M1 polarization and target the IL-4 receptor inhibit tumor growth by reprogramming tumor-associated macrophages into M1-like macrophages. *Biomaterials,* **2021**, *278*: 121137.

[45] Li C, Hu G, Wei B, et al. lncRNA LINC01494 promotes proliferation, migration and invasion in glioma through miR-122-5p/CCNG1 axis. *OncoTargets Ther.,* **2019**, *12*: 7655–7662.

[46] Bao Z, Yang Z, Huang Z, et al. LncRNADisease 2.0: an updated database of long non-coding RNA-associated diseases. *Nucleic Acids Res.,* **2019**, *47* (D1): D1034–D1037.

[47] Liang Z X, Liu H S, Wang F W, et al. LncRNA RPPH1 promotes colorectal cancer metastasis by interacting with TUBB3 and by promoting exosomes-mediated macrophage M2 polarization. *Cell Death Dis.,* **2019**, *10* (11): 829.

[48] Xu Z J, Chen Y, Ma L, et al. Role of exosomal non-coding RNAs from tumor cells and tumor-associated macrophages in the tumor microenvironment. *Mol. Ther.,* **2022**, *30* (10): 3133–3154.

[49] Zhou R, Zhang J W, Zeng D Q, et al. Immune cell infiltration as a biomarker for the diagnosis and prognosis of stage I–III colon cancer. *Cancer Immunol. Immunother.,* **2019**, *68* (3): 433–442.

[50] Galon J, Costes A, Sanchez-Cabo F, et al. Type, density, and location of immune cells within human colorectal tumors predict clinical outcome. *Science,* **2006**, *313* (5795): 1960–1964.

[51] Bagchi S, Yuan R, Engleman E G. Immune checkpoint inhibitors for the treatment of cancer: Clinical impact and mechanisms of response and resistance. *Annu. Rev. Pathol.,* **2021**, *16*: 223–249.

[52] Popat S, Hubner R, Houlston R S. Systematic review of microsatellite instability and colorectal cancer prognosis. *J. Clin. Oncol.,* **2005**, *23* (3): 609–618.

[53] Gryfe R, Kim H, Hsieh E T K, et al. Tumor microsatellite instability and clinical outcome in young patients with colorectal cancer. *N. Engl. J. Med,* **2000**, *342* (2): 69–77.

[54] Vilar E, Gruber S B. Microsatellite instability in colorectal cancer —the stable evidence. *Nat. Rev. Clin. Oncol.,* **2010**, *7* (3): 153–162.

[55] Bellone S, Roque D M, Siegel E R, et al. A phase II evaluation of pembrolizumab in recurrent microsatellite instability-high (MSI-H) endometrial cancer patients with Lynch-like versus *MLH*-1 methylated characteristics (NCT02899793). *Ann. Oncol.,* **2021**, *32* (8): 1045–1046.

[56] Puliga E, Corso S, Pietrantonio F, et al. Microsatellite instability in Gastric Cancer: Between lights and shadows. *Cancer Treat. Rev.,* **2021**, *95*: 102175.

[57] Pal T, Permuth-Wey J, Kumar A, et al. Systematic review and meta-analysis of ovarian cancers: Estimation of microsatellite-high frequency and characterization of mismatch repair deficient tumor histology. *Clin. Cancer Res.,* **2008**, *14* (21): 6847–6854.

[58] Gstalder C, Liu D, Miao D, et al. Inactivation of Fbxw7 impairs dsRNA sensing and confers resistance to PD-1 blockade. *Cancer Discov.,* **2020**, *10* (9): 1296–1311.

[59] Qi J, Sun H, Zhang Y, et al. Single-cell and spatial analysis reveal interaction of FAP⁺ fibroblasts and SPP1⁺ macrophages in colorectal cancer. *Nat. Commun.,* **2022**, *13* (1): 1742.

[60] Zhou L, Wang M, Guo H, et al. Integrated analysis highlights the immunosuppressive role of TREM2⁺ macrophages in hepatocellular carcinoma. *Front. Immunol.,* **2022**, *13*: 848367.

[61] Katzenelenbogen Y, Sheban F, Yalin A, et al. Coupled scRNA-seq and intracellular protein activity reveal an immunosuppressive role of TREM2 in cancer. *Cell,* **2020**, *182* (4): 872–885.e19.

[62] Bugatti M, Bergamini M, Missale F, et al. A population of TIM4⁺FOLR2⁺ macrophages localized in tertiary lymphoid structures correlates to an active immune infiltrate across several cancer types. *Cancer Immunol. Res.,* **2022**, *10* (11): 1340–1353.

[63] Nalio Ramos R, Missolo-Koussou Y, Gerber-Ferder Y, et al. Tissue-resident FOLR2⁺ macrophages associate with CD8⁺ T cell infiltration in human breast cancer. *Cell,* **2022**, *185* (7): 1189–1207.e25.