Molecular Therapy
Nucleic Acids
Original Article

# Systematic analysis of DNA methylation-mediated TF dysregulation on lncRNAs reveals critical roles in tumor immunity

Jiaqi Yin,[1,4] Na Ding,[1,4] Jiaxin Yu,[1,4] Zishan Wang,[3] Limei Fu,[1] Yongsheng Li,[2] Xia Li,[1] and Juan Xu[1]

[1]College of Bioinformatics Science and Technology, Harbin Medical University, Harbin 150081, China; [2]School of Interdisciplinary Medicine and Engineering, Harbin Medical University, Harbin 150081, China; [3]Department of Genetics and Genomic Sciences, Center for Transformative Disease Modeling, Tisch Cancer Institute, Icahn Institute for Data Science and Genomic Technology, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA

**Emerging evidence suggests that DNA methylation affects transcriptional regulation and expression perturbations of long non-coding RNAs (lncRNAs) in cancer. However, a comprehensive investigation into the transcriptional control of DNA methylation-mediated dysregulation of transcription factors (TFs) on lncRNAs has been lacking. Here, we integrated the transcriptome, methylome, and regulatome across 21 human cancers and systematically identified the transcriptional regulation of DNA methylation-mediated TF dysregulations (DMTDs) on lncRNAs. Our findings reveal that TF regulation of lncRNAs is significantly impacted by DNA methylation. Comparative analysis of DMTDs on mRNAs revealed a conserved pattern of TFs involvement. Pan-cancer Methylation TFs (MethTFs) and Methylation LncRNAs (MethLncRNAs) were identified, and were found to be closely associated with cancer hallmarks and clinical features. In-depth analysis of co-expressed mRNAs with pan-cancer MethLncRNAs unveiled frequent disruptions in cancer immunity, particularly in the context of inflammatory response. Furthermore, we identified five immune-related network modules that contribute to immune cell infiltration in cancer. Immune-related subtypes were subsequently classified, characterized by high levels of immune cell infiltration, expression of immunomodulatory genes, and relevant immune cytolytic activity score, major histocompatibility complex score, response to chemotherapy, and prognosis. Our findings provide valuable insights into cancer immunity from the epigenetic and transcriptional regulation perspective.**

## INTRODUCTION

Gene expression is regulated by transcription factors (TFs).[1,2] Therefore, it is important to identify the target genes of TFs and further explore their regulatory pathways and underlying mechanisms. With the development of high-throughput sequencing technology, it is possible to identify genome-wide TF binding sites (TFBSs) by chromatin immunoprecipitation sequencing (ChIP-seq).[3–5] It is well known that abnormalities of transcription regulation are closely related to carcinogenesis; however, the underlying molecular mechanism remains to be elucidated.

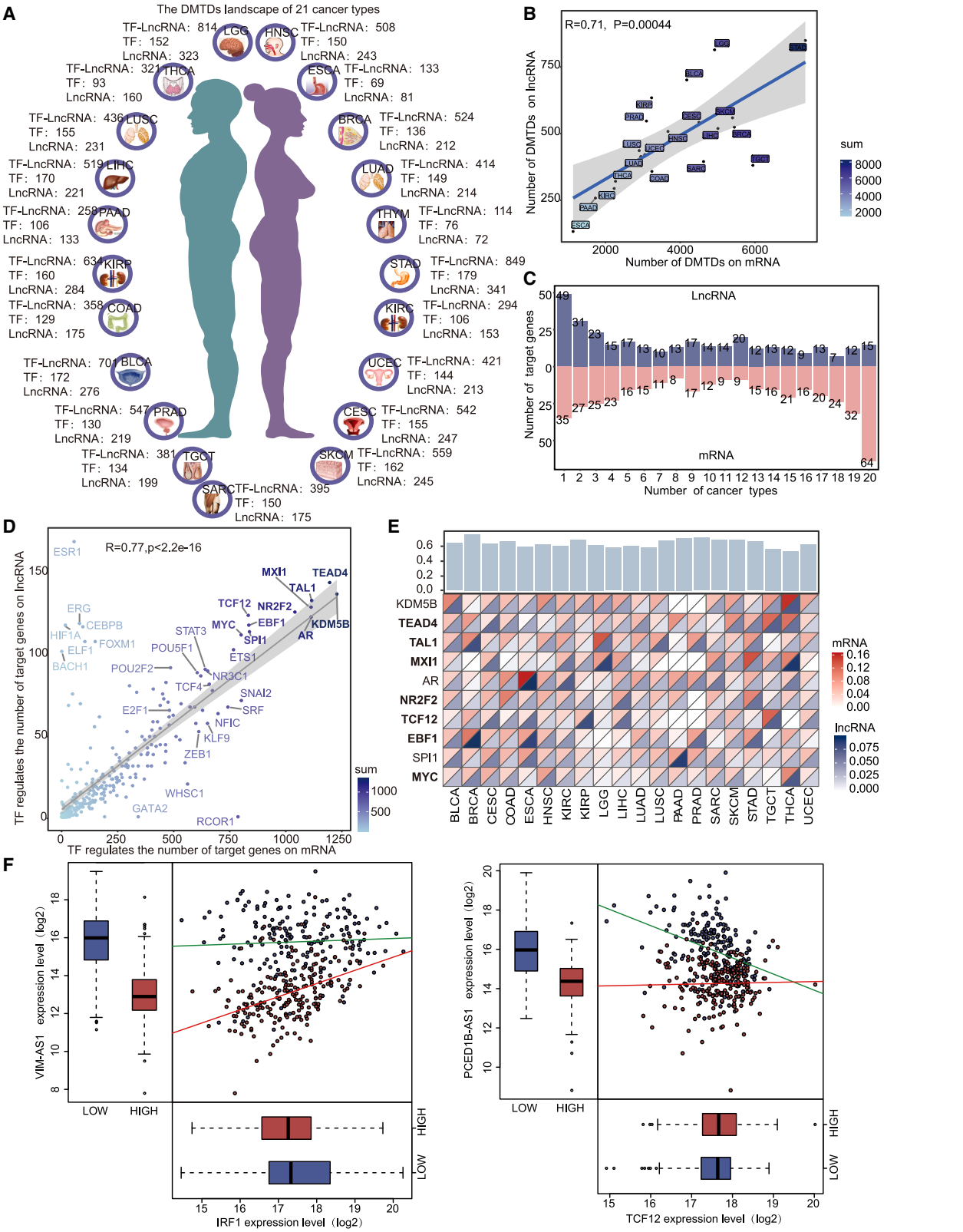
DNA methylation, as a type of important epigenetic modification, can silence a wide range of genes in various cancer types.[6–9] Increasing evidence suggests that DNA methylation in TFBSs can regulate the expression of target genes by affecting TF-binding efficiency[10] and many TFs have been discovered that bind CpG methylated sequences.[11] If the regulatory activity of a TF to its given target was affected by the DNA methylation of the target gene, this regulation was considered as a DNA methylation-mediated transcriptional dysregulation (DMTD). Considering the widespread changes to DNA methylation and transcriptional regulation across cancers, identifying DMTDs is crucial for human diseases. TFs whose regulatory activity are affected by DNA methylation are also defined as MethTFs. In our previous study, widespread regulation disruptions of TFs on mRNAs by DNA methylation were discovered, and we showed that they regulated several hallmarks of cancer.[12] Moreover, the cooperative regulation modules of MethTFs were closely related to their prognostic potential.

DNA methylation also has been shown to repress the expression of long non-coding RNAs (lncRNAs), and extensive methylation abnormalities occurring in the promoter region of lncRNAs have been discovered across cancers.[13–15] Here, methylation-mediated lncRNAs were defined as MethLncRNAs. There has been no systematic study on DNA methylation-mediated lncRNAs at the pan-cancer level. In contrast, current studies have shown that lncRNAs act as critical regulators in cancer immunity, but their precise role and the underlying regulation of tumor immune response have not been fully elucidated.[16,17] The MethLncRNA EPIC1 was found to suppress tumor

Received 7 March 2023; accepted 12 October 2023;
https://doi.org/10.1016/j.omtn.2023.102058.

[4]These authors contributed equally

**Correspondence:** Yongsheng Li, School of Interdisciplinary Medicine and Engineering, Harbin Medical University, Harbin 150081, China.
**E-mail:** liyongsheng@ems.hrbmu.edu.cn
**Correspondence:** Xia Li, College of Bioinformatics Science and Technology, Harbin Medical University, Harbin 150081, China.
**E-mail:** lixia@hrbmu.edu.cn
**Correspondence:** Juan Xu, College of Bioinformatics Science and Technology, Harbin Medical University, Harbin 150081, China.
**E-mail:** xujuanbiocc@ems.hrbmu.edu.cn




*(legend on next page)*

cell antigen presentation, leading to resistance to anti-programmed cell death protein 1 treatment.[18] These findings support the idea that some lncRNAs regulate the human immune response in the tumor microenvironment; however, most functions of MethLncRNAs and their role in tumor immunity are still unknown.

To systematically investigate DMTDs on lncRNAs in cancers, we integrated the transcriptome, methylome, and regulatome across 21 human cancers and systematically screened for the DMTDs on lncRNAs. The DMTD regulation patterns on lncRNAs were compared with those on mRNAs across cancers. Functional analysis was further performed and we found that DMTDs were closely correlated with hallmarks of cancer, especially cancer immunity. In addition, we identified co-regulatory immune modules, which were validated in independent verification sets. Thus, comprehensive analysis of methylation-mediated regulation between MethTFs and MethLncRNAs will contribute to our enhanced understanding of the complex transcriptional regulatory mechanism in cancer. Our analysis highlights the importance of methylation-mediated transcription regulation to lncRNA expression and also reveals critical roles of lncRNAs in cancer immunity.

## RESULTS

### DNA methylation widely mediated the transcriptional regulation on lncRNAs across cancers

Several studies have found that transcriptional regulation could be mediated by DNA methylation in the promoter regions of target genes. We have previously found that DMTDs for mRNAs are prevalent across cancers and the regulation patterns are complex.[12] lncRNAs are an important type of regulatory RNA and with a comparable amount as mRNA. Thus, we further systematically identified DMTDs for lncRNAs across 21 cancers. The detailed sample number is shown in Figure S1. Sample sizes varied from 121 in THYM to 189 in BRCA. Based on the proposed computational framework, we identified DMTDs for lncRNAs ranging from 114 to 849 DMTDs in 21 cancers, suggesting that DNA methylation could also influence the transcriptional regulation of lncRNAs (Figure 1A). Moreover, DMTDs for lncRNAs in each cancer formed a connecting and complex regulatory network, indicating the complex regulation patterns of DNA methylation (Figure S2).

We compared the transcriptional dysregulation patterns of lncRNAs mediated by DNA methylation to those of mRNAs, which were obtained from our previous studies.[12] After comparing with DMTDs on mRNAs in 20 cancer types, we found a significant positive correlation (Figure 1B) ($R = 0.71$; $p = 0.00044$). If there were more DMTDs for mRNAs in a cancer type, DNA methylation was more likely to mediate the regulation of lncRNAs. Similar results were found for the number of lncRNAs/mRNAs, as well as the number of targets for TFs (Figure S3). These results suggested that DMTDs were extensive and conserved in cancer, and the regulation patterns of DMTDs for lncRNAs were similar to mRNAs across cancers.

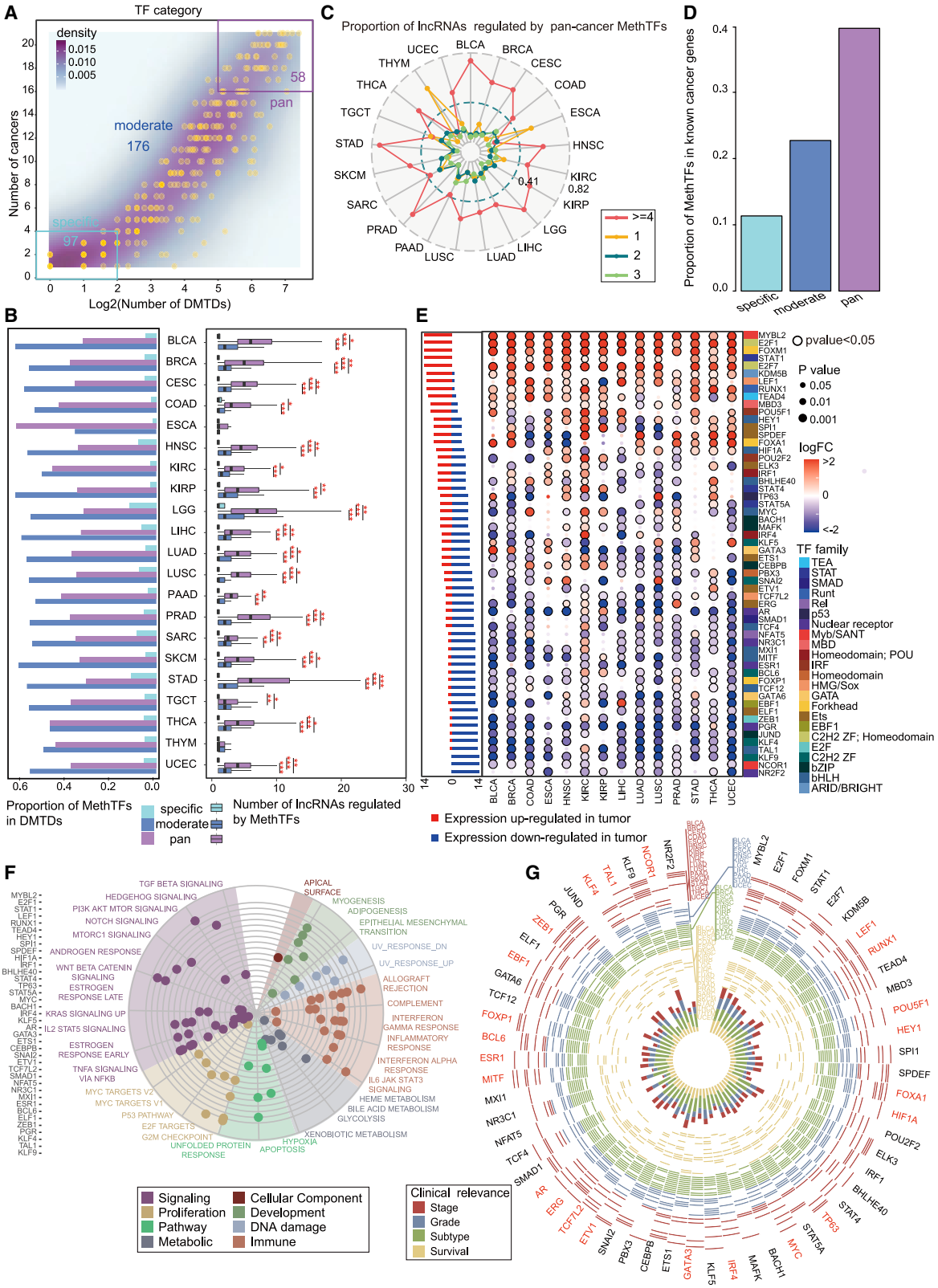
We further calculated the DMTD number of MethTFs across cancers and found that DNA methylation-mediated transcriptional regulation of mRNAs and lncRNAs are comparable across 20 cancers (Figure 1C) and exhibit strong positive correlation (Figure 1D) ($R = 0.77$; $p < 2.2e\text{-}16$), such as TEAD4, TAL1, and KDM5B. Similar results were observed in individual cancers (Figure S4). To remove the effects of promoter length, we re-identified the DMTDs of mRNAs by defining mRNA promoters as ±1 kb around transcription start sites (TSSs). As described above, the results also indicate that DNA methylation widely mediated transcriptional regulation on both mRNAs and lncRNAs across cancer types, and the regulation patterns of DMTDs on lncRNAs were similar to those of mRNAs (Figure S5).

In addition, some MethTFs tended to regulate more lncRNAs than mRNAs, including ESR1, ERG, HIF1A, and FOXM1. For instance, ESR1 regulated about 169 lncRNAs, whereas only 56 mRNAs were regulated. Taken together, these results indicated that the regulatory roles of these TFs should not be ignored from the viewpoint of lncRNAs.

It has been demonstrated that MethTFs have different binding motifs when the target sequence is methylated versus when it is not,[19] and DMTDs for mRNAs tend to enrich in methylated motifs. To analyze the transcriptional regulation of lncRNAs mediated by DNA methylation, we also assessed whether DMTDs for lncRNAs tend to be regulated by known methylated motifs, which were obtained from MeDReaders.[20] As shown in Figure 1E, an average of approximately 63.6% of DMTDs for lncRNAs showed binding by at least one methylated motif, and 7 of 10 MethTFs with the highest number of target genes had known methylated motifs, which regulate a considerable number of target genes across cancers. For example, TCF12 has one methylated motif, and all regulations were influenced by its methylated motif, whose overexpression in ovarian cancer is associated with higher histological grade and metastasis.[21] Additionally, we found that, in BRCA, the regulatory activities of TCF12 and IRF1 on their target lncRNAs varied between hypomethylated and hypermethylated groups (Figure 1F). Finally, we found that 91.9% of

**Figure 1. Identification of DMTDs in cancer**

(A) The number of DMTDs, MethTFs, and lncRNAs in each cancer type. (B) The number of DMTDs in each cancer. The horizontal and vertical axes represent the amount of DMTD that regulates mRNA and lncRNA in cancer, respectively. (C) The number of TFs involved in various cancers. The blue bar plot represents the number of TFs generated by regulating lncRNA, and the pink bar plot represents the number of TFs generated by regulating mRNA. (D) TFs regulate the number of target genes. The horizontal axis represents the number of mRNAs regulated by TFs, and the vertical axis represents the number of lncRNAs regulated by TFs. (E) The top of the figure represents the percentages of individual cancers in which DMTD is recognized by MEME. The bottom of the figure represents the proportion of the number of target genes regulated by the top 10 TFs to the total number of DMTD-related genes in the cancer. The upper triangle represents mRNA and the lower triangle represents lncRNA. The red letters represent TFs with known methylation sites. (F) The scatterplots showing the expression correlation between TF IRF1 and lncRNA VIM-AS1 in BRCA (left), and another example of TF TCF12 in BRCA (right). The boxplots showing the distribution of expression across cancer patients. The red and blue lines were fitted based on the patients with high (red) and low (blue) methylation levels of target lncRNAs.




*(legend on next page)*

DMTDs were only observed in one cancer (Figure S6A), and the high heterogeneity was also observed between cancer pairs (Figure S6B). These results indicate a strong cancer-specific regulation. Taking into account the heterogeneity between different cancer subtypes, we further compared the DMTDs in two breast cancer subtypes, including LuminaA and triple-negative breast cancer (TNBC). As a result, only 11.09% of DMTDs were shared. For example, the changes of regulatory activities of ETV4 or MYBL1 on their target lncRNAs were just observed in TNBC, which have been validated to be critical in TNBC.[22,23] Moreover, although a small proportion of DMTDs was observed in more than three cancer types, they formed a connecting regulatory network, suggesting a conserved "neuronal" DMTD network for lncRNAs that may maintain the network architecture across cancers. These results highlighted that systematic analysis of DMTDs for lncRNAs across cancers could further deepen our understanding about transcriptional dysregulations of lncRNAs mediated by DNA methylation.

### Pan-cancer MethTFs play important roles in carcinogenesis

It has been found that regulators with more targets tend to localize at the center of the regulatory network and have important biological functions. As described above, a few MethTFs indeed regulated hundreds of lncRNAs; thus, we attempted to identify these MethTFs across cancers. The pan-cancer MethTFs were defined as those observed in more than 15 cancer types and regulated more than 32 target genes. In total, 58 pan-cancer MethTFs were identified, accounting for 17.5% of MethTFs (Figure 2A). In contrast, 97 cancer-specific MethTFs were identified, which were only identified in no more than two cancers. Moreover, 45.6%–73.7% of DMTDs for lncRNAs were regulated by pan-cancer MethTFs, and these pan-cancer MethTFs were likely to regulate higher numbers of lncRNAs compared with other MethTFs in each cancer type (Figure 2B). We also found that high proportions of pan-cancer MethTFs regulated at least four lncRNAs in each cancer, particularly in BLCA, STAD, HNSC, and LGG, indicating that lncRNAs regulated by pan-cancer MethTFs tend to have more DMTDs (Figure 2C). Moreover, we found that approximately 40% pan-cancer MethTFs were significantly enriched in known cancer genes ($p = 0.0009175$, Fisher's exact test) (Figure 2D). We further discovered that the pan-cancer MethTFs showed significant differential expression in cancers, suggesting that expression perturbations might further change their regulation of lncRNAs (Figure 2E). For example, pan-cancer MethTF MYBL2 in Myb/SANT family was significantly upregulated in all analyzed cancers (Figure S7); this was consistent with previous findings that upregulation of MYBL2 was associated with poor prognosis in a variety of cancers.[24,25] In contrast, the pan-cancer MethTFs, NR2F2, and PGR in the nuclear receptor family were significantly downregulated in most cancers. Previous studies have found that NR2F2 was a tumor suppressor, which can effectively regulate a variety of signaling pathways, and control tumor cell growth and angiogenesis in the tumor microenvironment.[26] Overall, these results suggest that pan-cancer MethTFs can be used as cancer diagnostic biomarkers.

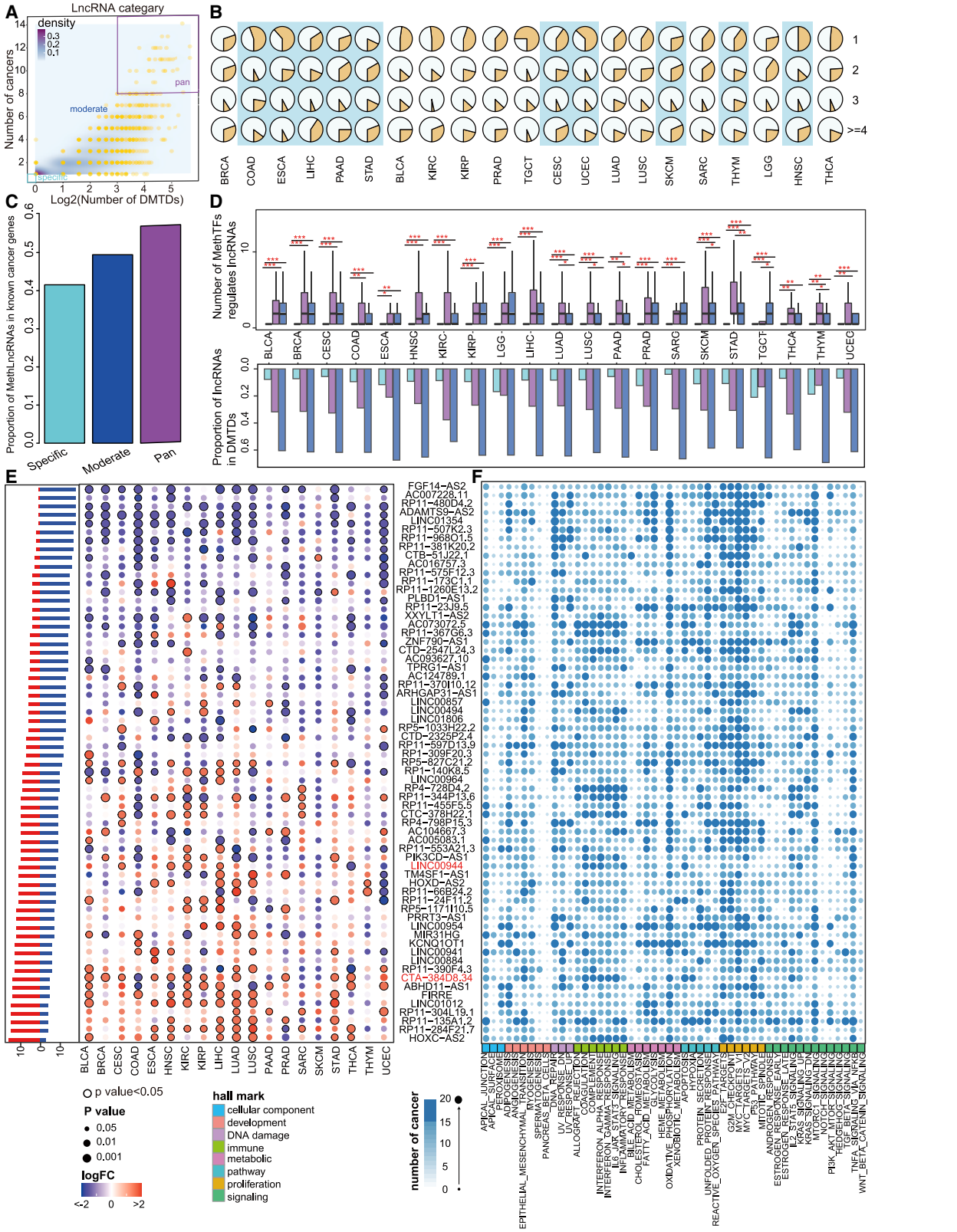
We further found that most pan-cancer MethTFs are significantly associated with cancer as revealed by their contribution to functions and clinical relevance. For example, the immune-related hallmark was significantly regulated by HIF1A and IRF1, whereas the differentiation-related hallmark was significantly regulated by the pan-cancer MethTF MYC (Figure 2F). We also observed that many immune-related functions were affected, such as inflammatory response, interferon gamma response, and interferon alpha response. Inflammation in the tumor microenvironment promotes tumor growth by increasing malignant cell proliferation and survival, stimulating angiogenesis and metastasis, and altering adaptive immune responses.[27] Type I interferons (IFN-α and IFN-β) have numerous direct and indirect effects on tumors.[28] HIF1A upregulation also has been shown to be associated with enhanced tumor immunity and stromal characteristics in 10 cancers, as well as aggressive phenotypes in human cancers.[29] By analyzing the functions of specific MethTFs, we found that specific MethTFs in different cancers participate in different functions. For example, SARC mainly participates in the UV response and UP and tumor necrosis factor-α signaling via the nuclear factor-κB pathway. LGG mainly participates in the unfolded protein response and transforming growth factor-β (TGF-β) signaling, whereas KIRP only participates in TGF-β signaling (Figure S8). Studies have found that TGF-β signaling regulates cell function and plays a key role in cell development and carcinogenesis.[30] In contrast, a notable statistically significant link between the expression of pan-cancer MethTFs and clinically relevant events was also found (Figure 2G). For example, the expression of IRF1 was significantly correlated with clinical features of multiple cancers, including different expressions across different stages or different subtypes; it was also a factor linked to patient survival (Figure S9). Similarly, STAT4 is another example that was closely associated with clinical features (Figure S10). Taken together, our comprehensive analysis of pan-cancer MethTFs indicates the overall changes and their potential prognostic value in cancer.

### Pan-cancer MethLncRNAs regulate cancer hallmarks and are closely correlated with clinical features

Since pan-cancer MethLncRNAs are well known regulators in cancer, we next explored the functions of pan-cancer MethLncRNAs

**Figure 2. Characterization of MethTFs in cancer**

(A) The number of TFs involved in regulatory relationships and the number of cancers. (B) The bar plot on the left represents the proportion of three types of TFs in different cancers, and the boxplot on the right shows the number of lncRNAs regulated by three types of TFs. (C) Pan-cancer MethTFs regulate the proportion of lncRNA numbers in each cancer. (D) Proportion of TFs in known cancer genes across the three TF categories. Known cancer genes were obtained from the Cancer Gene Census. (E) The heatmap shows the differential expression of pan-cancer MethTFs and the bar plot represent the number of cancers with high and low expression; red represents high expression and blue represents low expression. (F) Hallmarks of cancer enriched by pan-cancer MethTFs. (G) Clinically relevant pan-cancer MethTFs across human cancers. Colored cells in outer circle denote significant events. Colored bars in middle circle denote number of significant cases. Red, blue, green, and yellow denote significant cases in the following categories: grade, stage, subtype, and survival, respectively.



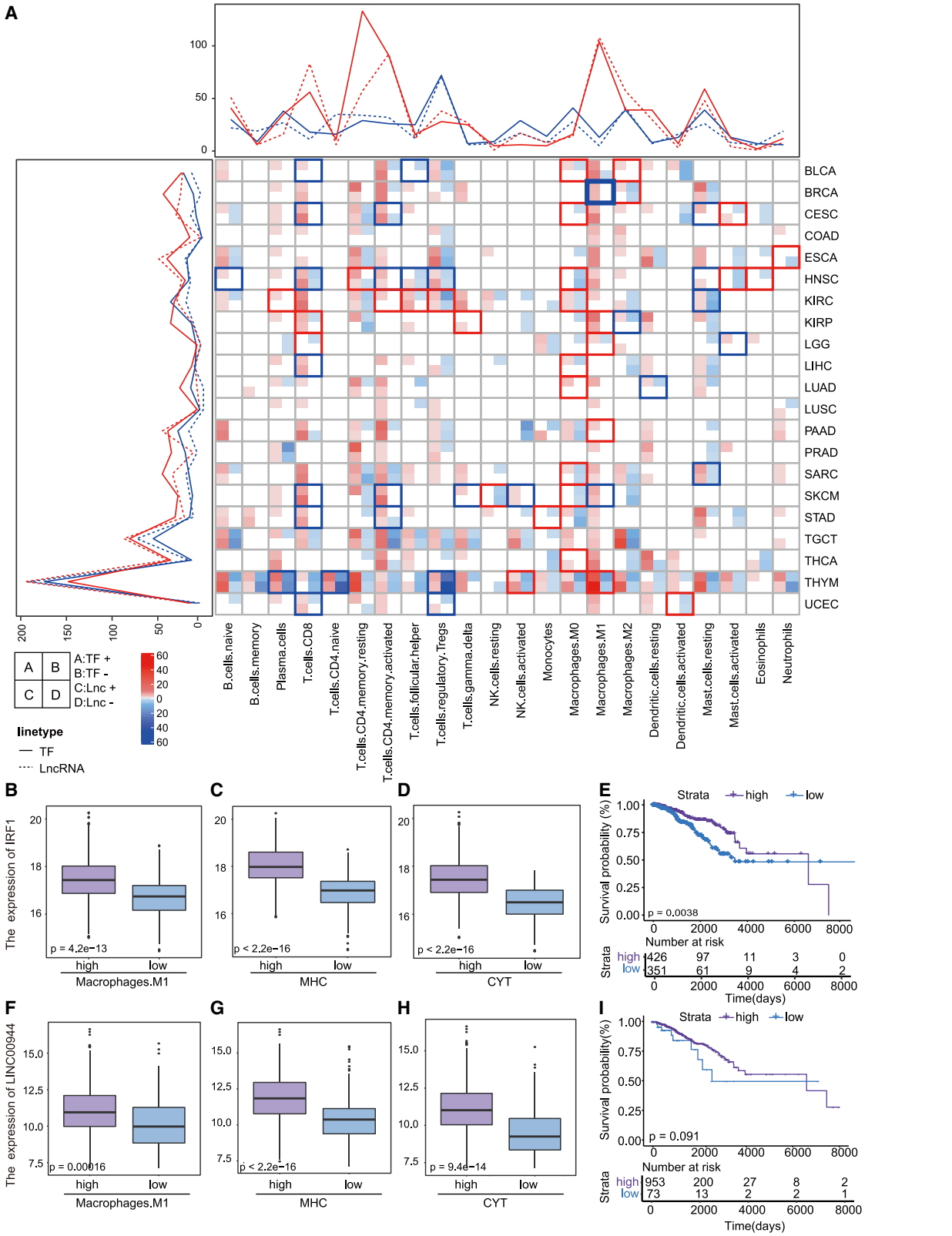

*(legend on next page)*

regulated by MethTFs in cancer, which were required to be observed in more than eight cancer types and regulated more than eight genes. In total, 65 pan-cancer MethLncRNAs were identified (Figure 3A), accounting for 3.7% of all lncRNAs and 6.3%–28.6% of all DMTDs. After annotating the biological types of pan-cancer MethLncRNAs, we found that most were antisense lncRNAs, and the proportion was significantly higher when compared with the whole genome (32/65; $p = 0.012$). Indeed, previous studies have reported that antisense lncRNAs were affected by DNA methylation in tumor progression,[31–33] and further disruptions of transcriptional regulation by DNA methylation on lncRNAs were revealed here. In addition, 496 cancer-specific MethLncRNAs were also newly identified in one cancer (Table S1). We also found that pan-cancer MethLncRNAs were likely to be more regulated by TFs compared with other lncRNAs in individual cancers (Figure 3B). Moreover, about 57% of pan-cancer MethLncRNAs are known cancer genes derived from Lnc2Cancer,[33] Cancer LncRNA Census,[34] and LncRNADisease2.0[35] (Figure 3C). In addition, we found that an average of 49.8% of target lncRNAs were known cancer genes in 21 cancers. The lncRNAs were differentially expressed in a variety of cancers, especially in BRCA, COAD, LUSC, and LUAD, suggesting that the target lncRNAs in the DMTDs we identified could serve as potential targets or biomarkers (Figure S11). We also found that higher proportions of pan-cancer MethLncRNAs were regulated by relatively more TFs in each cancer (Figure 3D). In addition, we found that pan-cancer MethLncRNAs tended to be differentially expressed, which might be acted as oncogenes or tumor suppressors (Figure 3E). Notably, the expression levels of pan-cancer MethLncRNAs were significantly negatively correlated with their corresponding methylation levels. Compared with all lncRNAs, the negative correlation ratios of pan-cancer MethLncRNA reached 86.6%, on average, which was significantly higher than all lncRNAs (Figures S12A and S12B). Moreover, a considerable percentage of pan-cancer MethLncRNAs negatively correlated with DNA methylation were likely to be differentially expressed; however, the directions of expression changes were different across cancer types (Figure S12C). For example, the proportions of downregulated lncRNAs were high in COAD and UCEC, while a large proportions of lncRNAs were upregulated in BLCA, KIRC, LIHC, and LUAD. These results indicate that DNA methylation might affect the expression of pan-cancer MethLncRNAs. MethLncRNA FIRRE was upregulated in most cancers, which was consistent with previous findings; it was highly differentially expressed in colorectal cancer,[36] and higher levels of FIRRE, which acts as an oncogene in DLBCL by promoting cell proliferation and reducing apoptosis, were associated with poorer overall survival.[37] In addition, we found that RP11−135A1.2, LINC01012, and CTA−384D8.34 were also upregulated in most cancers. LINC01354 is accompanied by downregulation of hypermethylation, which is associated with a poor prognosis. Moreover, LINC01354 may play an important role in predicting the prognosis of LUAD, thereby acting as a tumor suppressor gene regulated by DNA methylation. In contrast, FGF14−AS2, AC007228.11, RP11−968O1.5, and LINC01354 were significantly downregulated in most cancers. Previous studies have found that the lncRNA FGF14-AS2 was significantly downregulated in breast cancer and patients with lower FGF14-AS2 expression had advanced clinical stage.[38,39]

To assess the function of pan-cancer MethLncRNAs, we first performed a functional enrichment analysis of cancer hallmarks for each pan-cancer MethLncRNA based on its co-expressed mRNAs. Then, functional annotation was also performed on individual TFs, which regulated the same lncRNA, and the overlapping functions were removed to rule out the influence of those genes that have a coincident correlative expression pattern. As shown in Figure 3F, pan-cancer MethLncRNAs were significantly enriched in cancer hallmarks. The LINC00944 expression had a strong relationship with immune signaling pathways and positively correlated with tumor-infiltrating T lymphocytes and pro-apoptotic markers, in agreement with a previous study.[40] In addition, the other study found that MIR31HG overexpression was evidently correlated with high immune infiltrate levels of $CD8^+$ T cells, macrophages, neutrophils, myeloid dendritic cells, and B cells in thyroid cancer. KCNQ1OT1 was shown to regulate cancer cell proliferation, cell cycle, migration and invasion, metastasis, glucose metabolism, and immune evasion, in a previous study.[41] Abnormal expression of ADAMTS9-AS2 in different tumors was found to be closely related to tumor proliferation, invasion, migration, and apoptosis inhibition. ADAMTS9-AS2 is involved in DNA methylation, mediates the phosphatidylinositol 3-kinase/Akt/mammalian target of rapamycin signaling pathway, and regulates miRNAs and proteins, indicating its important therapeutic potential in cancer.[42] In contrast, specific MethLncRNAs were also significantly enriched in cancer hallmarks. We found specific MethLncRNAs involved in the same function in some cancers. Specifically, in most cancers they are mainly involved in immune, developmental, and signal-related functions. Some specific MethLncRNAs involved in specific functions in cancers, such as THYM mainly participates in proliferation-related functions (Figure S13). Moreover, the expression of pan-cancer MethLncRNAs was associated with clinically relevant events across cancers (Figure S14). LINC00944 is considered to be an important gene involved in cancer immunity.[16,38] As an example, we observed that the high expression of LINC00944 was significantly

**Figure 3. Characterization of MethLncRNAs in cancer**

(A) The number of lncRNAs involved in regulatory relationships and the number of cancers. (B) The number of TFs obtained by regulating lncRNAs in each cancer. (C) Proportion of lncRNAs in known cancer lncRNAs across the three lncRNAs categories. Known cancer lncRNAs were obtained from the Lnc2Cancer, Cancer LncRNA Census, and LncRNADisease2.0. (D) The bar plot below represents the proportion of three types of lncRNAs in different cancers, and the boxplot above shows the TF quantity obtained by regulating the three types of lncRNA. (E) The heatmap shows the differential expression of pan-cancer MethLncRNAs. The bar plot represents the number of cancers with high and low expression. Red represents high expression and blue represents low expression. (F) Cancer hallmarks enriched by pan-cancer MethLncRNAs. Both the size and color of the dots represent the number of cancers enriched. (F) Shares the gene names with (E) to visualize the links between differential expression and molecular function for each pan-cancer MethLncRNA.




*(legend on next page)*

correlated with tumor stage, grade, and tumor subtype, as well as patient survival (Figures S15A–S15D). Indeed, it has been reported that LINC00944 was involved in cancer immunity and participated in several immune-related pathways.[16,38] In addition, although the p value of survival analysis is not statistically significant by analyzing all patients of primary skin cutaneous melanoma (SKCM), a weak significance was found by just considering patients with relatively low TNM stages (Figure S15D). Its expression was also upregulated in SKCM patients with metastasis (Figure S15E). Moreover, when analyzing an immunotherapy cohort of SKCM(PRJEB23709), we found that the expression of LINC00944 was downregulated in responded patients compared with no-responded ones (Figure S15F). Notably, SKCM patients with high LINC00944 expression levels had significantly worse overall survivals (Figure S15G). These results suggest that there was an underlying correlation of LINC00944 expression with cancer immunity, as well as immunotherapy response. At the same time, we found similar results for CTA-384D8.34 (Figure S16). Therefore, these results indicate that pan-cancer MethLncRNAs plays important roles in carcinogenesis.

### Pan-cancer genes are closely involved in cancer immunity

Tumor cells live in a complex tumor microenvironment that is necessary for tumor growth and survival. Previous studies have shown that DNA methylation is an important epigenetic mechanism involved in controlling T cell responses.[43] To further investigate the relationship between pan-cancer MethTFs/MethLncRNAs and the immune microenvironment, we first used CIBERSORT to calculate the proportions of 22 immune cell types based on RNA-seq data across cancer types. Furthermore, we calculated the correlation between these pan-cancer MethLncRNAs/MethTFs and the infiltration degrees of different immune cells and found that the majority of pan-cancer genes involved in DNA methylation-mediated regulations were significantly correlated with immune cells, especially in THYM. This was consistent with previous studies that showed that the thymus is a central lymphatic organ responsible for many immune functions, including the production of mature, functional T cells and the induction of self-tolerance.[44] In addition, we found a significant positive correlation between pan-cancer gene expression and immune cell infiltration in various cancers. Meanwhile, in most cancers, genes were significantly positively correlated with CD4 memory resting T cells, CD4 memory activated T cells, M1 macrophages, and resting mast cells, while genes were significantly negatively correlated with regulatory T cells (Tregs) (Figure 4A).

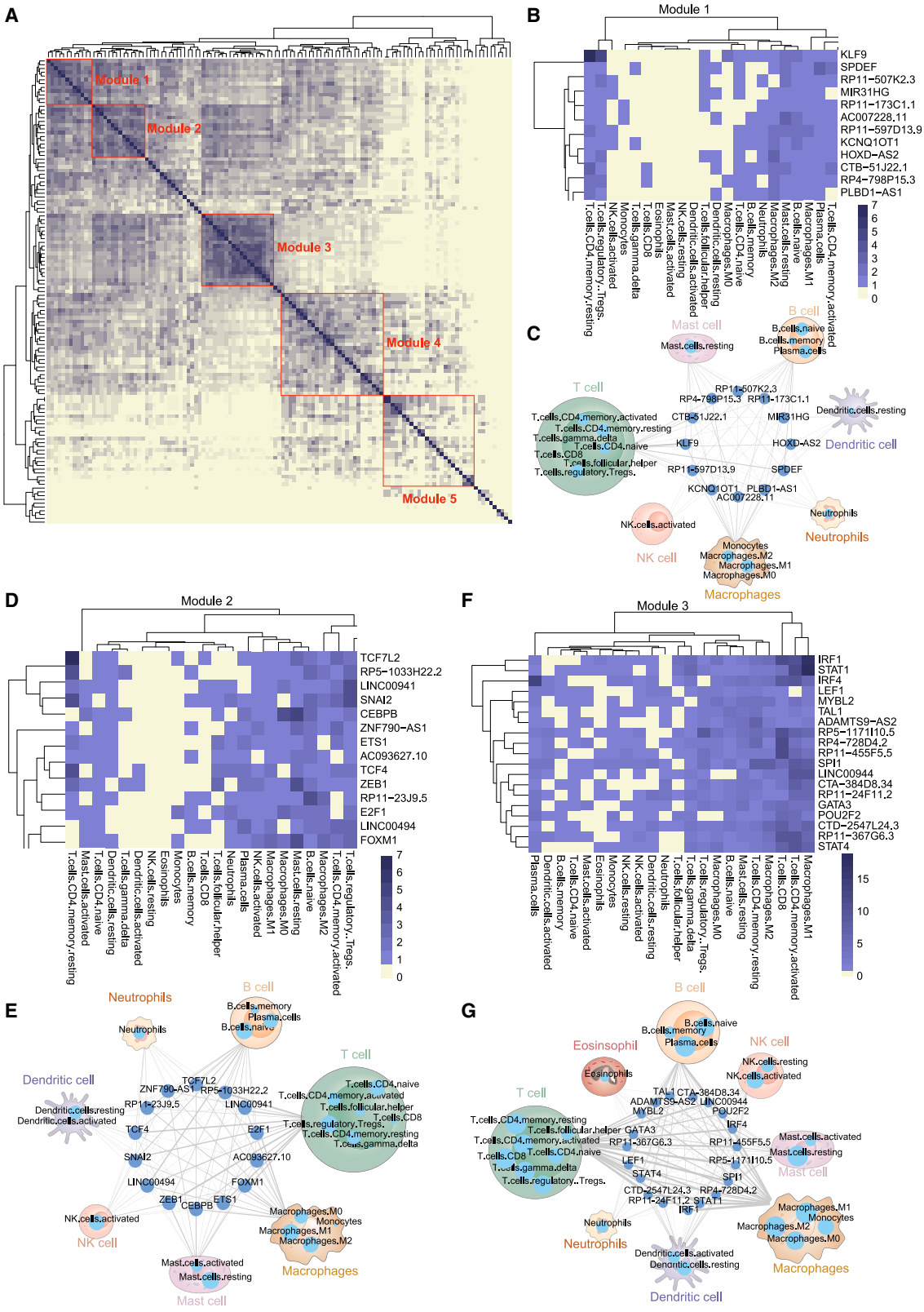
Previous studies have discovered that some immune cells act as risk factors in cancer, while some act as protective factors. We also found that cell types associated with pan-cancer MethLncRNAs/MethTFs are potential prognostic markers. For instance, the M1 macrophage, has been identified as a protective factor in BRCA ($p = 0.013$; $\beta = -4.19$, Cox regression analysis). Macrophages can be influenced by a variety of factors that change their phenotype and, thus, affect their function, including pan-cancer MethTFs IRF1 and STAT4, whose expression were significantly increased in the group with a high infiltration score of M1 macrophages (Figures 4B and S17A). Moreover, the expression increases of IRF1 and STAT4 were found in the group with high major histocompatibility complex (MHC) (Figures 4C and S17B) scores or CYT scores (Figures 4D and S17C), and significant positive correlation were also revealed (Table S2). Activated macrophages are usually divided into two categories, M1-like macrophages and M2-like macrophages. Both M1 macrophages and M2 macrophages are closely related to inflammatory responses, but M1 macrophages are mainly involved in pro-inflammatory responses, while M2 macrophages are mainly involved in anti-inflammatory responses.[45] Indeed, it has been found that IRF1 is an important tumor suppressor in breast cancer and plays a major role in controlling the transcriptional program of macrophages both at the basal level and after IFN-γ activation.[46] Our research provides evidence that upregulated IRF1 and STAT4 are associated with better overall survival (OS) in BRCA (Figures 4E and S17D). Previous studies also discovered that overexpression of STAT4 mRNA was significantly associated with a favorable OS in breast cancer patients.[47] In addition, the pan-cancer MethLncRNAs, LINC00944 was significantly positively correlated with M1 macrophages in BRCA (Figure 4F). Studies have shown that the expression of LINC00944 is strongly correlated with immune signaling pathways, tumor-infiltrating T lymphocytes, and pro-apoptotic markers,[40] and the low expression of LINC00944 was correlated with poor prognosis in breast cancer patients. Here, we discovered that high expression of LINC00944 not only significantly contributed to high immune infiltration scores of M1 macrophages, but also higher MHC scores and CYT scores (Figures 4G and 4H). LINC00944 was found to be a protective factor for OS in BRCA (Figure 4I). Similar results were discovered for another pan-cancer MethLncRNA CTA-384D8.34 (Figures S17E–S17H). In addition, M2 macrophages have been identified as a risk factor in breast cancer, and we found that M2 macrophage-related upregulation of ESR1 is associated with poorer OS in BRCA (Figures S17I–S17L), with evidence that ESR1 is an important prognostic marker in BRCA.[48–50] These results support the contention that pan-cancer genes play critical roles in cancer immunity.

### Pan-cancer MethTF-LncRNA modules related to tumor immunity

As one of the main factors of cancer development, we further explored pan-cancer MethTFs/MethLncRNAs that cooperatively contributed to the immune system dysregulation. First, we used pan-cancer genes to perform unsupervised clustering on the intersection divided by a maximum of significantly correlated immune cells. As a result, five

**Figure 4. Immune infiltration of pan-cancer genes in cancer**

(A) Correlation between pan-cancer MethTFs and pan-cancer MethLncRNAs and 22 types of immune cells in each cancer. Red represents the number of positive correlations, blue represents the number of negative correlations ($|corr| > 0.3$; $p < 0.05$). The red border represents risk factors and the blue border represents protective factors. (B–D) The expression of IRF1 was correlated with macrophage M1, MHC score, and CYT score. (E) Association between IRF1 expression and patient survival. The number of patients in different risk groups is also provided. (F–I) The similar result for lncRNA LINC00944.




*(legend on next page)*

co-regulatory immune modules were detected, including 12 to 24 pan-cancer MethTFs/MethLncRNAs (Figure 5A), and these regulations occurred in multiple cancers (Figures 5B–5G), especially module 3. In module 3, there were 10 TFs and 9 lncRNAs that cooperatively regulate 22 types of immune cells (Figures 5F and 5G). Notably, macrophages M1/M2, $CD4^+$, and $CD8^+$ T cells were regulated by almost all genes in this module in no less than one cancer type. Pan-cancer MethTF IRF1 and STAT1 tended to regulate M1 macrophages in 20 cancers. IRF1 involves regulatory molecular networks that play a key role in M1 macrophage inflammatory responses and viral defense.[51] Intratumoral $CD4^+$ T cells are believed to be responsible for the production of antitumor immune responses and inflammatory mediators that induce tumor growth, invasion, angiogenesis, and metastasis.[52,53] We found that there are 50 pan-cancer MethTFs and 46 pan-cancer MethLncRNAs associated with CD4 memory resting T cells in 17 cancers. For example, KLF9 in module 1, TCF7L2 in module 2, NR2F2, NR3C1, and ELK3 in module 4; and NFAT5 in module 5 tended to regulate CD4 memory resting T cells in multiple cancers (Figure S18). The above results indicated that these pan-cancer MethTFs and MethLncRNAs synergistically regulate some immune cells in cancers.

### Pan-cancer MethTF/MethLncRNA modules help classification of cancer subtypes

In addition to identifying key genes associated with immune infiltration in cancer, cancer subtyping is key to improving personalized treatment.[54–56] Therefore, we classified tumor molecular subtypes according to pan-cancer MethTF-LncRNA modules. Considering the maximum number of genes, module 3 was used to classify cancer samples into four subtypes based on the expression of pan-cancer MethTF/MethLncRNA in module 3 by consensus cluster method (Figure S19; Table S3). Notably, the majority of module 3 genes, immune checkpoint genes, and immunology regulons were highly expressed in C3 patients and underexpressed in C1 patients. Moreover, the immune subtype C3 we identified also had higher immune cell infiltration scores estimated by both TIMER and xCell algorithms (Figure 6A). Next, we identified the molecular functions related to different subtypes by gene set variation analysis. We found that patients in the C3 subtype exhibited higher T cell and Treg activities, which were lower in C1 patients (Figure 6A). It is well-known that Treg cells are key regulatory cells in inflammation and act as immune suppressor T cells, which prevent tissue damage caused by excessive autoimmunity through their inhibitory function.[57] Further, we mapped samples of TCGA onto TME subtypes based on previous studies,[58] and found C3 subtypes tended to be enriched in immune enriched, fibrotic and immune enriched, non-fibrotic, whereas C1 subtypes tended to be enriched in desert. Next, the immune microenvironment of four subtypes were compared, and indeed C3 patients showed higher infiltration of immune cells, such as B cells, CD4 T cells, CD8 T cells, and macrophages (Figure 6B). Moreover, we found that the C1 subtype had a better prognosis than the C3 subtype ($p = 0.0047$), as well as a relatively high response rate ($p = 0.0001153$) to radiotherapy and chemotherapy (Figure S20). In addition, CYT scores and MHC scores were higher in C3 patients (Figure 6B). In addition, the other four immune modules were analyzed and cancer samples were classified into different groups. Interestingly, we found that many subtypes corresponded with one main subtype classified by module 3 (Figure S21). For example, samples in the subtype C1 of module 2 was mainly grouped in the subtype C1 of module 3. These results indicate an underlying consistency of subtypes distinguished by different immune modules of pan-cancer MethTF/MethLncRNA. Moreover, more stringent samples in high- or low-immunity subtypes remained based on the subtypes grouped by other modules. We found that the results were consistent with those of module 3, and the difference between the two subtypes was even more obvious (Figure S22).

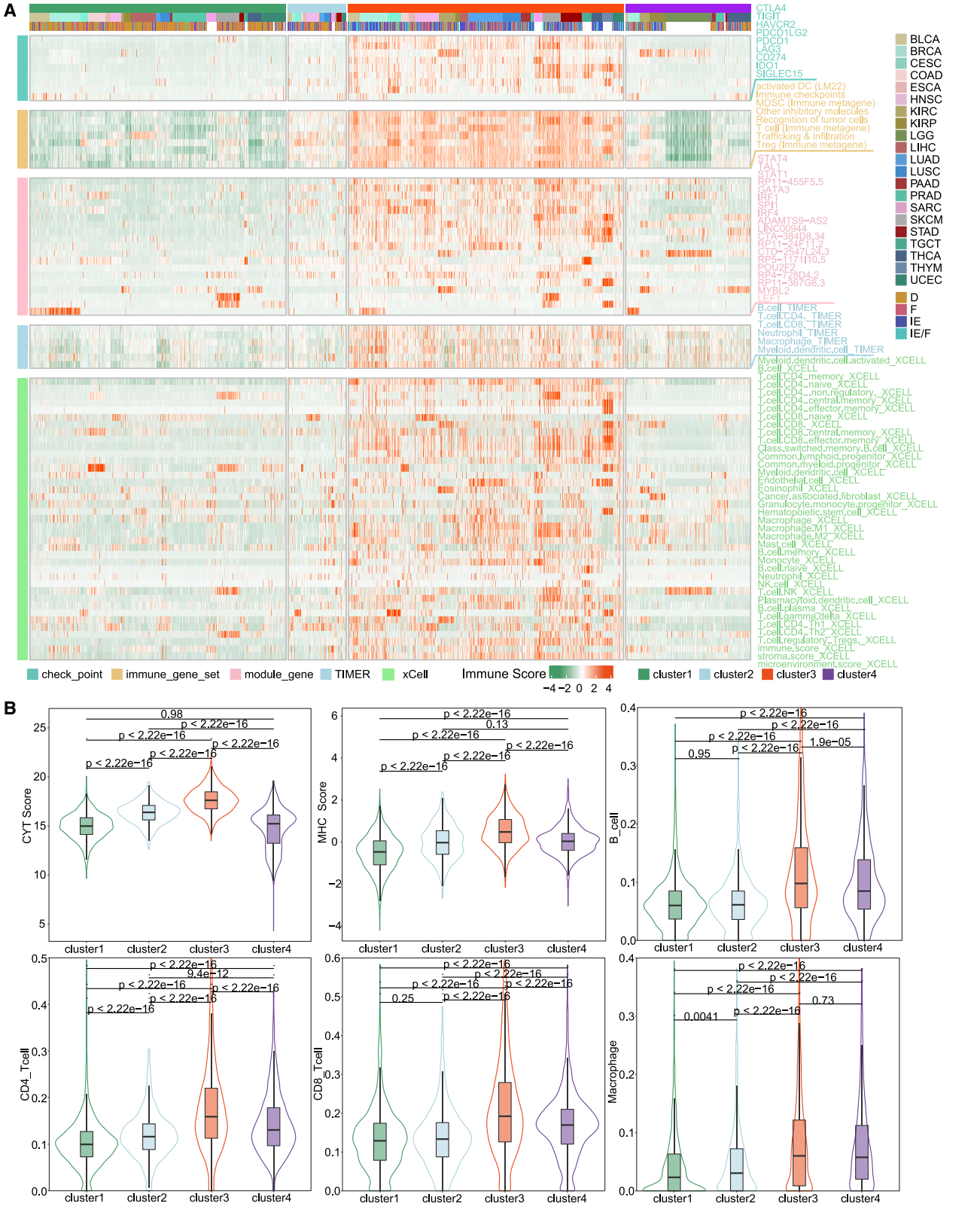
We finally validated the efficiency of module 3 to identify the immune subtype in two independent datasets, including the IMvigor210 cohort and GSE91061 (Figures 7A and 7B). The immune-related sample clusters were discovered and closely associated with the tumor microenvironment, with the highest proportion of patients who responded well to immune checkpoint blockade therapy (Figures 7C and 7D). We also found that C3 had a better survival rate than the other two clusters (Figures 7C and 7D). Considering the underlying correlation between expression and activity, we evaluated whether pan-cancer MethTFs/MethLncRNAs tended to be expressed to a greater extent in immune cells than in malignant cells by analyzing 17 single-cell datasets. We found that the expression of pan-cancer MethTFs/LncRNAs in module 3 was higher in immune cells than in malignant cells (Figures 7E and 7F). These results indicated that pan-cancer MethTFs/LncRNAs tend to have high activity in immune cells. Taken together, these results suggested that the immune-related co-regulatory genes identified here can provide valuable insights into tumor classification and have independent prognostic value.

## DISCUSSION

TFs, cofactors, and chromatin regulators control gene expression programs, and misregulation of gene expression programs can lead to a number of diseases.[59] Our study proposed a three-step method to identify the DMTDs in cancer. We identified DMTDs in 21 cancers and discovered that transcriptional regulation of both mRNAs and lncRNAs was widely affected by DNA methylation, and a considerable proportion of DMTDs occur in multiple cancers. In particular, the regulation of known cancer TFs tend to be affected by DNA methylation, which also regulated more lncRNAs, indicating that the identified pan-cancer MethTFs played central roles in cancer.

**Figure 5. Co-regulatory immune modules in cancer**

(A) Pan-cancer MethTFs and pan-cancer MethLncRNAs co-regulate immune modules. (B–G) The heatmap above represents pan-cancer MethTFs and pan-cancer MethLncRNAs that were significantly associated with immune cells in the number of participating cancers for co-modulated immune modules. The network diagram below represents co-regulatory immune module networks; the size of the immune cell spot corresponds to the number of associated genes; the thickness of the line corresponds to the number of cancers involved.




*(legend on next page)*

Moreover, we found that pan-cancer MethTFs involved in multiple cancers tended to be differentially expressed and involved in many carcinogenesis-related biological processes, such as immune and cancer hallmarks. As a master regulator in TNBC, IRF1 has been reported to be important for immune suppression by inducing effectors.[60] In our study, we found that HIF1A and IRF1 were also involved in inflammatory response, IFN-γ response, IFN-α response, and other immune-related hallmarks. These pan-cancer MethTFs were also significantly correlated with clinical features, including tumor prognosis. However, it is difficult to determine the causal relationship between DNA methylation and TF binding. The candidate DMTDs identified in this study need to be validated in further experiments.

As important factors regulating gene expression, lncRNAs have been proven to have carcinogenesis-related or tumor suppressor effects.[61] In this study, we identified pan-cancer MethLncRNAs whose regulations were mainly mediated by DNA methylation, and their expression tended to be significantly different across cancers. To evaluate the associations between DMTDs and prognosis, we performed univariate Cox regression analyses for each DMTD in cancers, where the integrative risk score of a DMTD was calculated as in previous studies.[62,63] As results, large proportions of DMTDs were related to cancer prognosis (Figure S23). After multivariate Cox analysis by considering tumor mutation load, tumor purity, and immune checkpoint inhibitors, many DMTDs as independent prognostic factors were found. Indeed, many pan-cancer MethLncRNAs are known cancer genes. In the case of LINC01012, it was shown to be differentially expressed in 12 cancers in our study and has also been previously confirmed to play a role in colon cancer.[64] These differentially expressed pan-cancer MethLncRNAs were discovered to be involved in cancer and act as immune hallmarks, as well as prognostic factors.

Based on their strong association with immune functions, we further found their role in immune cell infiltration, and five key modules composed of pan-cancer MethTFs/MethLncRNAs were detected that cooperatively regulated immune functions. Finally, these key modules could effectively distinguish cancer subtypes with high immune activity, and were highly correlated with various immune indicators. We propose here that these modules might shed light on future cancer immunotherapies.

In this study, we comprehensively integrated gene expression profiles and DNA methylation profiles to identify DMTDs on lncRNAs across cancer types. The developed computational framework can be extended to other complex diseases, which will provide insights into the underlying mechanisms of epigenetic regulation. Moreover, our analyses revealed the regulatory patterns of DMTDs across various cancer types. We found that pan-cancer MethTFs involved in multiple cancers tended to be differentially expressed and involved in numerous cancer-related biological processes. In particular, five critical modules composed of pan-cancer MethTFs/MethLncRNAs were detected that cooperatively regulated immune-related functions. These key modules could effectively distinguish cancer subtypes with high immune activity and were highly correlated with various immune indicators. This study provides valuable insights that deepen our understanding of the complex epigenetic regulation in the tumor microenvironment and provides potential new candidates that may be useful in cancer immunotherapy. The identified lncRNAs, TFs, and target genes provide valuable candidates for future experimental validation.

In conclusion, we have systematically identified DMTD on lncRNAs in human cancer types, and further revealed critical regulatory modules involved in tumor immunity and immunology subtypes. Our study provides valuable insights that deepen our understanding of the complex epigenetic regulation in the tumor microenvironment and provides new candidates that that may be useful in cancer immunotherapy.

## MATERIALS AND METHODS

### Sample-paired gene expression profiles and DNA methylation profiles across 21 cancer types

Genome-wide gene expression profiles and DNA methylation profiles across 21 cancers were downloaded from The Cancer Genome Atlas (TCGA, https://portal.gdc.cancer.gov). The expression levels were quantified as fragments per kilobase per million reads, and genes with expression in at least 30% samples were retained. In total, 462 TFs and 12,113 lncRNAs were considered for further analysis. DNA methylation datasets were assayed by the Illumina Infinium HumanMethylation450 BeadChip array. Similarly, probes with β-values of greater than 0 in at least 30% samples were reserved. The remaining probes were mapped to gene promoter regions, which were defined as ±1-kb regions around the TSSs.[65–67] Next, the level of DNA methylation of each gene was the average β-values of probes mapped to its promoter region. We only retained cancers with at least 150 samples for subsequent analysis.

### Clinical information

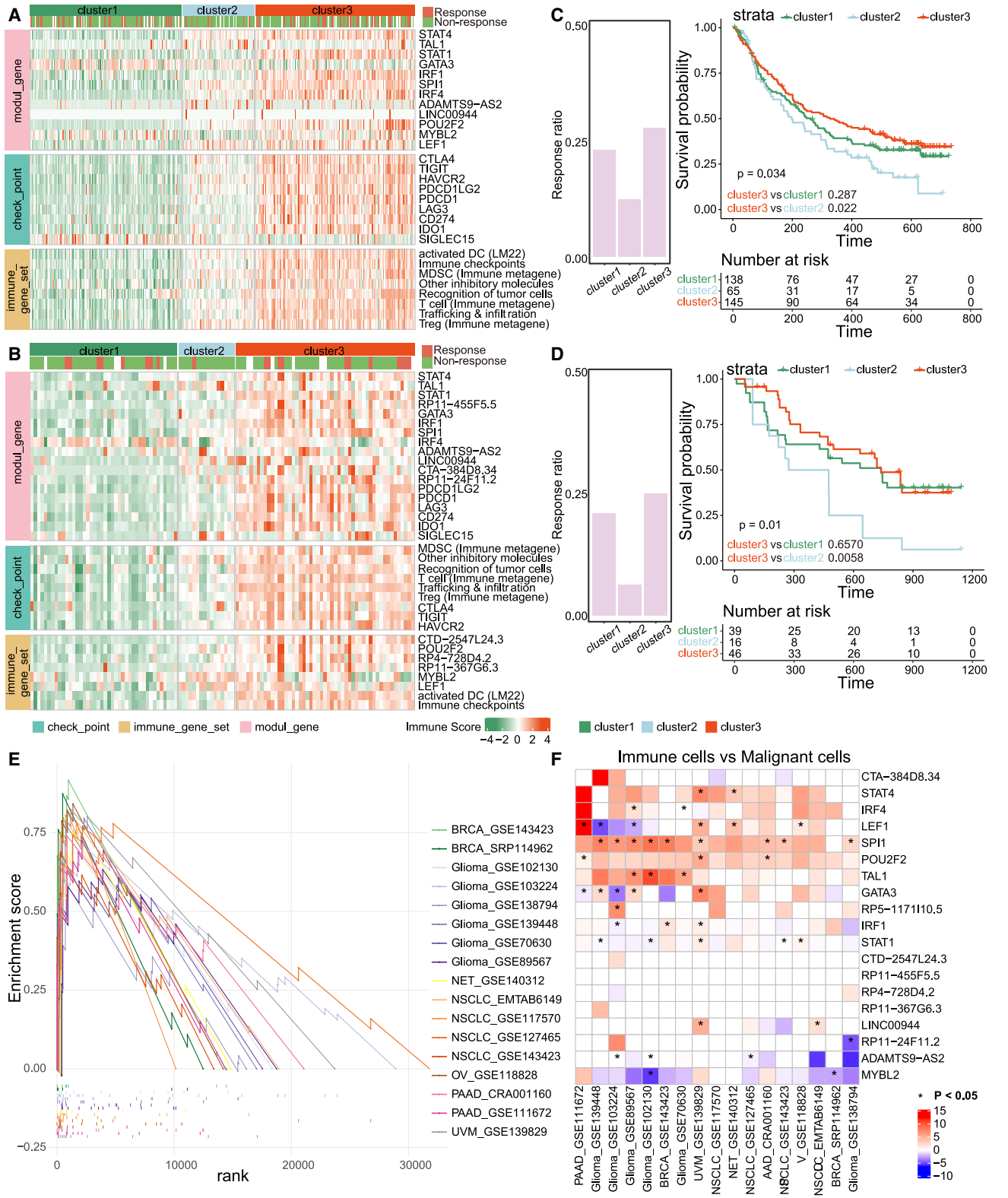
The clinical information of cancer patients was also downloaded from the TCGA project, including survival status, tumor stage, grade, survival time, and response to chemotherapy. Survival analysis was performed using R-package survival (Version 3.2–13) and survminer (Version 0.4.9).

### Identification of DMTDs on lncRNAs across cancer types

We used a three-step computational framework similar to one of our previous studies on the identification of cancer-context transcriptional dysregulation mediated by DNA methylation, to further identify the DMTDs on lncRNAs. Briefly, we intersected the ChIP-Seq peak regions (1 kb upstream to 1 kb downstream of the TSS) of 471

**Figure 6. Pan-cancer MethTF-LncRNA modules differentiate cancer subtypes**

(A) The heatmap shows the expression of genes in module 3, the expression of immune checkpoint genes, the single sample gene set enrichment analysis scores of eight immune-related signatures, and the immune cell infiltration levels of TIMER and xCell. Annotation bars indicate sample subtypes, cancer types, and immunoenrichment types of samples in TCGA cohort. (B) CYT scores, MHC scores, and estimates for different cells in the four clusters.




*(legend on next page)*

TFs from ChIPBase v2.0 with lncRNA promoter regions, the corresponding lncRNAs with peaks were identified as candidate targets of TFs.[68] There were 1,745,038 transcriptional regulatory pairs for 15,549 lncRNAs. Next, we used univariate linear regression to identify cancer-context TF-lncRNA transcriptional regulations. TF-lncRNA regulation with a Benjamini and Hochberg-adjusted p value of less than 0.01 was obtained. A goodness-of-fit test was further used to assess the TF-lncRNA regulations, and TF-lncRNA regulation with a coefficient of greater than 0.95 was retained as the cancer context TF-lncRNA regulatory pair. Finally, the DNA methylation mediated TF-lncRNA dysregulation was identified based on the TF-lncRNA regulation alterations in different DNA methylation levels, and the same thresholds were used similar as a previous study.[12]

#### Identification of differentially expressed genes in cancer

Differentially expressed MethTFs and lncRNAs in each cancer type were identified by limma (Version 3.46.0) R packages, and genes with at least 4-fold changes and a p value of less than 0.05 were regarded as differentially expressed. Here, 14 cancer types with at least 5 normal samples were analyzed.

#### Functional analysis

To predict the biological functions of pan-cancer MethLncRNAs, mRNAs were first ranked based on their Spearman correlation coefficients with each pan-cancer MethLncRNA. Functional gene sets were obtained from the MSigDB database, and we particularly analyzed gene sets associated with cancer hallmark and immune-related functions. Functional enrichment analysis was realized by the fgsea package (Version 1.16.0). In addition, the cancer markers denoted by pan-cancer MethTFs were obtained via functional annotation.

#### Quantitative analysis of immune cell infiltration

CIBERSORT was used to estimate tumor immune infiltration in each cancer sample, and 22 immune cell types were analyzed, including B cells, T cells, natural killer cells, and macrophages. In addition, TIMER[69] and xCell[70] algorithms were used to calculate the global infiltration levels of immune cells. Then, the Spearman correlation coefficient was calculated between the expression level of genes and infiltration levels of each immune cell type. The gene cell pair was reserved if its absolute value of correlation coefficient was greater than 0.3 and the p value was less than 0.05.

#### Statistical analysis

ANOVA was used to assess the statistical significance of clinically relevant events, including cancer subtype, stage, and grade. In addition, we used the Cox regression model to evaluate the association between survival time and the expression level of genes of interest. Fisher's exact test was used to evaluate the bias of MethLncRNAs, biotypes, negative correlation between DNA methylation and expression of panMethLncRNA, and whether patients of C3 subtype responded to radiotherapy and chemotherapy.

## DATA AND CODE AVAILABILITY

The data that support the findings of this study are available upon reasonable request to the corresponding authors.

## SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.omtn.2023.102058.

## ACKNOWLEDGMENTS

This work was supported by the National Natural Science Foundation of China [32322020, 32170676, 31970646, 32060152, and 32300541]; Natural Science Foundation of Heilongjiang Province (Key Program) [ZD2023C007]; Natural Science Foundation of Heilongjiang Province [LH2023C063]; the China Postdoctoral Science Foundation [2021M693831]; and the China Brain Project [2021ZD0202403] and Heilongjiang Touyan Innovation Team Program.

## AUTHOR CONTRIBUTIONS

J.Q.Y. and N.D. contributed study design, data analysis; and paper writing; J.X.Y. and L.M.F. collected samples and generated data; Z.S.W. carried out data interpretations and helped data discussion; J.X., X.L., and Y.S.L contributed paper revisions.

## DECLARATION OF INTERESTS

No potential conflicts of interest were disclosed.

## REFERENCES

1. Nie, L., Vázquez, A.E., and Yamoah, E.N. (2009). Identification of transcription factor-DNA interactions using chromatin immunoprecipitation assays. Methods Mol. Biol. *493*, 311–321.
2. Popp, A.P., Hettich, J., and Gebhardt, J.C.M. (2021). Altering transcription factor binding reveals comprehensive transcriptional kinetics of a basic gene. Nucleic Acids Res. *49*, 6249–6266.
3. Pavesi, G. (2017). ChIP-Seq Data Analysis to Define Transcriptional Regulatory Networks. Adv. Biochem. Eng. Biotechnol. *160*, 1–14.
4. Vorontsov, I.E., Fedorova, A.D., Yevshin, I.S., Sharipov, R.N., Kolpakov, F.A., Makeev, V.J., and Kulakovskiy, I.V. (2018). Genome-wide map of human and mouse transcription factor binding sites aggregated from ChIP-Seq data. BMC Res. Notes *11*, 756.

**Figure 7. Pan-cancer MethTF-LncRNA modules differentiate cancer subtypes in validation cohort**

(A) The heatmap shows the expression of genes in module 3, the expression of immune checkpoint genes, and the single sample gene set enrichment analysis scores of 8 immune-related signatures in Imv210 cohort. Annotation bars indicate sample subtypes, and response of samples. (B) The heatmap shows the expression of genes in module 3, the expression of immune checkpoint genes, and the single sample gene set enrichment analysis scores of eight immune-related signatures in GSE91061 cohort. Annotation bars indicate sample subtypes, cancer types, and response of samples. (C and D) The analysis results for IMvigor210 and GSE91061 cohorts, respectively. Proportion of response samples in three clusters and survival of three clusters; bar plot represents patient response to immune checkpoint blockade therapy. (E) Genes in module 3 are enriched in single-cell data (gene rank is fold change value expressed in immune cells and malignant cells). (F) Heatmap of differential expression of module 3 gene in single cell datasets.

5. Wang, X., Sterr, M., Burtscher, I., Chen, S., Hieronimus, A., Machicao, F., Staiger, H., Häring, H.U., Lederer, G., Meitinger, T., et al. (2018). Genome-wide analysis of PDX1 target genes in human pancreatic progenitors. Mol. Metab. *9*, 57–68.
6. Kulis, M., and Esteller, M. (2010). DNA methylation and cancer. Adv. Genet. *70*, 27–56.
7. Meng, H., Cao, Y., Qin, J., Song, X., Zhang, Q., Shi, Y., and Cao, L. (2015). DNA methylation, its mediators and genome integrity. Int. J. Biol. Sci. *11*, 604–617.
8. Bogdanović, O., and Lister, R. (2017). DNA methylation and the preservation of cell identity. Curr. Opin. Genet. Dev. *46*, 9–14.
9. Choi, S.J., Jung, S.W., Huh, S., Chung, Y.S., Cho, H., and Kang, H. (2017). Alteration of DNA Methylation in Gastric Cancer with Chemotherapy. J. Microbiol. Biotechnol. *27*, 1367–1378.
10. Sun, X., Yi, J., Yang, J., Han, Y., Qian, X., Liu, Y., Li, J., Lu, B., Zhang, J., Pan, X., et al. (2021). An integrated epigenomic-transcriptomic landscape of lung cancer reveals novel methylation driver genes of diagnostic and therapeutic relevance. Theranostics *11*, 5346–5364.
11. Yin, Y., Morgunova, E., Jolma, A., Kaasinen, E., Sahu, B., Khund-Sayeed, S., Das, P.K., Kivioja, T., Dave, K., Zhong, F., et al. (2017). Impact of cytosine methylation on DNA binding specificities of human transcription factors. Science *356*, eaaj2239.
12. Wang, Z., Yin, J., Zhou, W., Bai, J., Xie, Y., Xu, K., Zheng, X., Xiao, J., Zhou, L., Qi, X., et al. (2020). Complex impact of DNA methylation on transcriptional dysregulation across 22 human cancer types. Nucleic Acids Res. *48*, 2287–2302.
13. Wang, Z., Yang, B., Zhang, M., Guo, W., Wu, Z., Wang, Y., Jia, L., Li, S., Cancer Genome Atlas Research Network, Xie, W., and Yang, D. (2018). lncRNA Epigenetic Landscape Analysis Identifies EPIC1 as an Oncogenic lncRNA that Interacts with MYC and Promotes Cell-Cycle Progression in Cancer. Cancer Cell *33*, 706–720.e9.
14. Li, Y., Zhang, Y., Li, S., Lu, J., Chen, J., Wang, Y., Li, Y., Xu, J., and Li, X. (2015). Genome-wide DNA methylome analysis reveals epigenetically dysregulated non-coding RNAs in human breast cancer. Sci. Rep. *5*, 8790.
15. Xu, J., Wang, Z., Li, S., Chen, J., Zhang, J., Jiang, C., Zhao, Z., Li, J., Li, Y., and Li, X. (2018). Combinatorial epigenetic regulation of non-coding RNAs has profound effects on oncogenic pathways in breast cancer subtypes. Brief. Bioinform. *19*, 52–64.
16. Li, Y., Jiang, T., Zhou, W., Li, J., Li, X., Wang, Q., Jin, X., Yin, J., Chen, L., Zhang, Y., et al. (2020). Pan-cancer characterization of immune-related lncRNAs identifies potential oncogenic biomarkers. Nat. Commun. *11*, 1000.
17. Wang, Y., Zhao, Y., Guo, W., Yadav, G.S., Bhaskarla, C., Wang, Z., Wang, X., Li, S., Wang, Y., Chen, Y., et al. (2022). Genome-wide gain-of-function screening characterized lncRNA regulators for tumor immune response. Sci. Adv. *8*, eadd0005.
18. Guo, W., Wang, Y., Yang, M., Wang, Z., Wang, Y., Chaurasia, S., Wu, Z., Zhang, M., Yadav, G.S., Rathod, S., et al. (2021). LincRNA-immunity landscape analysis identifies EPIC1 as a regulator of tumor immune evasion and immunotherapy resistance. Sci. Adv. *7*, eabb3555.
19. Hu, S., Wan, J., Su, Y., Song, Q., Zeng, Y., Nguyen, H.N., Shin, J., Cox, E., Rho, H.S., Woodard, C., et al. (2013). DNA methylation presents distinct binding sites for human transcription factors. Elife *2*, e00726.
20. Wang, G., Luo, X., Wang, J., Wan, J., Xia, S., Zhu, H., Qian, J., and Wang, Y. (2018). MeDReaders: a database for transcription factors that bind to methylated DNA. Nucleic Acids Res. *46*, D146–D151.
21. Gao, S., Bian, T., Zhang, Y., Su, M., and Liu, Y. (2019). TCF12 overexpression as a poor prognostic factor in ovarian cancer. Pathol. Res. Pract. *215*, 152527.
22. Yuan, Z.Y., Dai, T., Wang, S.S., Peng, R.J., Li, X.H., Qin, T., Song, L.B., and Wang, X. (2014). Overexpression of ETV4 Protein in Triple-Negative Breast Cancer Is Associated with a Higher Risk of Distant Metastasis. Onco. Targets Ther. *7*, 1733–1742.
23. Player, A., Abraham, N., Burrell, K., Bengone, I.O., Harris, A., Nunez, L., Willaims, T., Kwende, S., and Walls, W. (2017). Identification of candidate genes associated with triple negative breast cancer. Genes Cancer *8*, 659–672.
24. Xin, Z., Li, Y., Meng, L., Dong, L., Ren, J., and Men, J. (2022). Elevated expression of the MYB proto-oncogene like 2 (MYBL2)-encoding gene as a prognostic and predictive biomarker in human cancers. Math. Biosci. Eng. *19*, 1825–1842.
25. Li, M., Sun, Q., and Wang, X. (2017). Transcriptional landscape of human cancers. Oncotarget *8*, 34534–34551.
26. Xu, M., Qin, J., Tsai, S.Y., and Tsai, M.J. (2015). The role of the orphan nuclear receptor COUP-TFII in tumorigenesis. Acta Pharmacol. Sin. *36*, 32–36.
27. Mantovani, A., Allavena, P., Sica, A., and Balkwill, F. (2008). Cancer-related inflammation. Nature *454*, 436–444.
28. Borden, E.C. (2019). Interferons alpha and beta in cancer: therapeutic opportunities from new insights. Nat. Rev. Drug Discov. *18*, 219–234.
29. Chen, B., Li, L., Li, M., and Wang, X. (2020). HIF1A expression correlates with increased tumor immune and stromal signatures and aggressive phenotypes in human cancers. Cell. Oncol. *43*, 877–888.
30. Derynck, R., and Zhang, Y.E. (2003). Smad-dependent and Smad-independent pathways in TGF-beta family signalling. Nature *425*, 577–584.
31. Wong, C.H., Li, C.H., Man Tong, J.H., Zheng, D., He, Q., Luo, Z., Lou, U.K., Wang, J., To, K.F., and Chen, Y. (2022). The establishment of CDK9/RNA PolII/H3K4me3/DNA methylation feedback promotes HOTAIR expression by RNA elongation enhancement in cancer. Mol. Ther. *30*, 1597–1609.
32. Zhang, L., Li, Z., Skrzypczynska, K.M., Fang, Q., Zhang, W., O'Brien, S.A., He, Y., Wang, L., Zhang, Q., Kim, A., et al. (2020). Single-Cell Analyses Inform Mechanisms of Myeloid-Targeted Therapies in Colon Cancer. Cell *181*, 442–459.e29.
33. Gao, Y., Shang, S., Guo, S., Li, X., Zhou, H., Liu, H., Sun, Y., Wang, J., Wang, P., Zhi, H., et al. (2021). Lnc2Cancer 3.0: an updated resource for experimentally supported lncRNA/circRNA cancer associations and web tools based on RNA-seq and scRNA-seq data. Nucleic Acids Res. *49*, D1251–D1258.
34. Vancura, A., Lanzós, A., Bosch-Guiteras, N., Esteban, M.T., Gutierrez, A.H., Haefliger, S., and Johnson, R. (2021). Cancer LncRNA Census 2 (CLC2): an enhanced resource reveals clinical features of cancer lncRNAs. NAR Cancer *3*, zcab013.
35. Bao, Z., Yang, Z., Huang, Z., Zhou, Y., Cui, Q., and Dong, D. (2019). LncRNADisease 2.0: an updated database of long non-coding RNA-associated diseases. Nucleic Acids Res. *47*, D1034–D1037.
36. Li, M., Zhao, L.M., Li, S.L., Li, J., Gao, B., Wang, F.F., Wang, S.P., Hu, X.H., Cao, J., and Wang, G.Y. (2018). Differentially expressed lncRNAs and mRNAs identified by NGS analysis in colorectal cancer patients. Cancer Med. *7*, 4650–4664.
37. Shi, X., Cui, Z., Liu, X., Wu, S., Wu, Y., Fang, F., and Zhao, H. (2019). LncRNA FIRRE is activated by MYC and promotes the development of diffuse large B-cell lymphoma via Wnt/beta-catenin signaling pathway. Biochem. Biophys. Res. Commun. *510*, 594–600.
38. Jin, Y., Zhang, M., Duan, R., Yang, J., Yang, Y., Wang, J., Jiang, C., Yao, B., Li, L., Yuan, H., et al. (2020). Long noncoding RNA FGF14-AS2 inhibits breast cancer metastasis by regulating the miR-370-3p/FGF14 axis. Cell Death Discov. *6*, 103.
39. Li, R., Yang, Y.E., Yin, Y.H., Zhang, M.Y., Li, H., and Qu, Y.Q. (2019). Methylation and transcriptome analysis reveal lung adenocarcinoma-specific diagnostic biomarkers. J. Transl. Med. *17*, 324.
40. de Santiago, P.R., Blanco, A., Morales, F., Marcelain, K., Harismendy, O., Sjöberg Herrera, M., and Armisén, R. (2021). Immune-related lncRNA LINC00944 responds to variations in ADAR1 levels and it is associated with breast cancer prognosis. Life Sci. *268*, 118956.
41. Cagle, P., Qi, Q., Niture, S., and Kumar, D. (2021). KCNQ1OT1: An Oncogenic Long Noncoding RNA. Biomolecules *11*, 1602.
42. Xu, W., Wang, B., Cai, Y., Chen, J., Lv, X., Guo, C., and Yuan, C. (2021). ADAMTS9-AS2: A Functional Long Non-coding RNA in Tumorigenesis. Curr. Pharm. Des. *27*, 2722–2727.
43. Schmidl, C., Delacher, M., Huehn, J., and Feuerer, M. (2018). Epigenetic mechanisms regulating T-cell responses. J. Allergy Clin. Immunol. *142*, 728–743.
44. Shelly, S., Agmon-Levin, N., Altman, A., and Shoenfeld, Y. (2011). Thymoma and autoimmunity. Cell. Mol. Immunol. *8*, 199–202.
45. Yunna, C., Mengru, H., Lei, W., and Weidong, C. (2020). Macrophage M1/M2 polarization. Eur. J. Pharmacol. *877*, 173090.
46. Langlais, D., Barreiro, L.B., and Gros, P. (2016). The macrophage IRF8/IRF1 regulome is required for protection against infections and is associated with chronic inflammation. J. Exp. Med. *213*, 585–603.

47. Wang, S., Yu, L., Shi, W., Li, X., and Yu, L. (2018). Prognostic roles of signal transducers and activators of transcription family in human breast cancer. Biosci. Rep. *38*.

48. Carausu, M., Bidard, F.C., Callens, C., Melaabi, S., Jeannot, E., Pierga, J.Y., and Cabel, L. (2019). ESR1 mutations: a new biomarker in breast cancer. Expert Rev. Mol. Diagn. *19*, 599–611.

49. Dustin, D., Gu, G., and Fuqua, S.A.W. (2019). ESR1 mutations in breast cancer. Cancer *125*, 3714–3728.

50. Heeke, A.L., Elliott, A., Feldman, R., O'Connor, H.F., Pohlmann, P.R., Lynce, F., Swain, S.M., Nunes, M.R., Magee, D., Oberley, M.J., et al. (2022). Molecular characterization of ESR1 variants in breast cancer. Breast Cancer Res. Treat. *196*, 279–289.

51. Chu, Y.B., Li, J., Jia, P., Cui, J., Zhang, R., Kang, X., Lv, M., and Zhang, S. (2021). Irf1- and Egr1-activated transcription plays a key role in macrophage polarization: A multiomics sequencing study with partial validation. Int. Immunopharmacol. *99*, 108072.

52. Matsumoto, H., Thike, A.A., Li, H., Yeong, J., Koo, S.L., Dent, R.A., Tan, P.H., and Iqbal, J. (2016). Increased CD4 and CD8-positive T cell infiltrate signifies good prognosis in a subset of triple-negative breast cancer. Breast Cancer Res. Treat. *156*, 237–247.

53. Mu, L., Yang, C., Gao, Q., Long, Y., Ge, H., DeLeon, G., Jin, L., Chang, Y.E., Sayour, E.J., Ji, J., et al. (2017). CD4+ and Perivascular Foxp3+ T Cells in Glioma Correlate with Angiogenesis and Tumor Progression. Front. Immunol. *8*, 1451.

54. Jackson, S.E., and Chester, J.D. (2015). Personalised cancer medicine. Int. J. Cancer *137*, 262–266.

55. Wang, Q., Li, M., Yang, M., Yang, Y., Song, F., Zhang, W., Li, X., and Chen, K. (2020). Analysis of immune-related signatures of lung adenocarcinoma identified two distinct subtypes: implications for immune checkpoint blockade therapy. Aging (Albany NY) *12*, 3312–3339.

56. Naylor, E.C., Desani, J.K., and Chung, P.K. (2016). Targeted Therapy and Immunotherapy for Lung Cancer. Surg. Oncol. Clin. N. Am. *25*, 601–609.

57. Barbi, J., Pardoll, D., and Pan, F. (2014). Treg functional stability and its responsiveness to the microenvironment. Immunol. Rev. *259*, 115–139.

58. Bagaev, A., Kotlov, N., Nomie, K., Svekolkin, V., Gafurov, A., Isaeva, O., Osokin, N., Kozlov, I., Frenkel, F., Gancharova, O., et al. (2021). Conserved pan-cancer microenvironment subtypes predict response to immunotherapy. Cancer Cell *39*, 845–865.e7.

59. Lee, T.I., and Young, R.A. (2013). Transcriptional regulation and its misregulation in disease. Cell *152*, 1237–1251.

60. Yamashita, N., Long, M., Fushimi, A., Yamamoto, M., Hata, T., Hagiwara, M., Bhattacharya, A., Hu, Q., Wong, K.K., Liu, S., and Kufe, D. (2021). MUC1-C integrates activation of the IFN-gamma pathway with suppression of the tumor immune microenvironment in triple-negative breast cancer. J. Immunother. Cancer *9*, e002115.

61. Peng, W.X., Koirala, P., and Mo, Y.Y. (2017). LncRNA-mediated regulation of cell signaling in cancer. Oncogene *36*, 5661–5667.

62. Yang, Z., Xu, F., Wang, H., Teschendorff, A.E., Xie, F., and He, Y. Pan-cancer characterization of long non-coding RNA and DNA methylation mediated transcriptional dysregulation. EBioMedicine. 2021 Jun;68:103399

63. Li, Y., Li, L., Wang, Z., Pan, T., Sahni, N., Jin, X., Wang, G., Li, J., Zheng, X., Zhang, Y., et al. LncMAP: Pan-cancer atlas of long noncoding RNA-mediated transcriptional network perturbations. Nucleic Acids Res., 46, 1113–1123

64. Zhang, B., Lin, J., Zhang, J., Wang, X., and Deng, X. (2022). Integrated Chromatin Accessibility and Transcriptome Landscapes of 5-Fluorouracil-Resistant Colon Cancer Cells. Front. Cell Dev. Biol. *10*, 838332.

65. Al Akoum, C., Akl, I., Rouas, R., Fayyad-Kazan, M., Falha, L., Renno, T., Burny, A., Lewalle, P., Fayyad-Kazan, H., and Badran, B. (2015). NFAT-1, Sp-1, Sp-3, and miR-21: New regulators of chemokine C receptor 7 expression in mature human dendritic cells. Hum. Immunol. *76*, 307–317.

66. Sehra, B., and Franks, R.G. (2017). Redundant CArG Box Cis-motif Activity Mediates SHATTERPROOF2 Transcriptional Regulation during Arabidopsis thaliana Gynoecium Development. Front. Plant Sci. *8*, 1712.

67. Stiers, P.J., van Gastel, N., Moermans, K., Stockmans, I., and Carmeliet, G. (2017). Regulatory elements driving the expression of skeletal lineage reporters differ during bone development and adulthood. Bone *105*, 154–162.

68. Zhou, K.R., Liu, S., Sun, W.J., Zheng, L.L., Zhou, H., Yang, J.H., and Qu, L.H. (2017). ChIPBase v2.0: decoding transcriptional regulatory networks of non-coding RNAs and protein-coding genes from ChIP-seq data. Nucleic Acids Res. *45*, D43–D50.

69. Li, T., Fu, J., Zeng, Z., Cohen, D., Li, J., Chen, Q., Li, B., and Liu, X.S. (2020). TIMER2.0 for analysis of tumor-infiltrating immune cells. Nucleic Acids Res. *48*, W509–W514.

70. Aran, D., Hu, Z., and Butte, A.J. (2017). xCell: digitally portraying the tissue cellular heterogeneity landscape. Genome Biol. *18*, 220.