Pan-Cancer Methylated Dysregulation of Long Non-coding RNAs Reveals Epigenetic Biomarkers

Ning Zhao1, Maozu Guo2*, Chunlong Zhang3, Chunyu Wang4 and Kuanquan Wang1,4

1School of Life Science and Technology, Harbin Institute of Technology, Harbin, China, 2School of Electrical and Information Engineering, Beijing University of Civil Engineering and Architecture, Beijing, China, 3College of Information and Computer Engineering, Northeast Forest University, Harbin, China, 4School of Computer Science and Technology, Harbin Institute of Technology, Harbin, China

Different cancer types not only have common characteristics but also have their own characteristics respectively. The mechanism of these specific and common characteristics is still unclear. Pan-cancer analysis can help understand the similarities and differences among cancer types by systematically describing different patterns in cancers and identifying cancer-specific and cancer-common molecular biomarkers. While long non-coding RNAs (lncRNAs) are key cancer modulators, there is still a lack of pan-cancer analysis for lncRNA methylation dysregulation. In this study, we integrated lncRNA methylation, lncRNA expression and mRNA expression data to illuminate specific and common lncRNA methylation patterns in 23 cancer types. Then, we screened aberrantly methylated lncRNAs that negatively regulated lncRNA expression and mapped them to the ceRNA relationship for further validation. 29 lncRNAs were identified as diagnostic biomarkers for their corresponding cancer types, with lncRNA AC027601 was identified as a new KIRC-associated biomarker, and lncRNA ACTA2-AS1 was regarded as a carcinogenic factor of KIRP. Two lncRNAs HOXA-AS2 and AC007228 were identified as pan-cancer biomarkers. In general, the cancer-specific and cancer-common lncRNA biomarkers identified in this study may aid in cancer diagnosis and treatment.

Keywords: lncRNA, pan-cancer, DNA methylation, biomarker, ceRNA

INTRODUCTION

Cancer is a general term that refers to malignant tumors, and its world-wide incidence and mortality have been high for many years. In 2020, there were 19.29 million new cancer cases and 9.96 million cancer deaths world-wide (Sung et al., 2021). There is no cure for cancer at the moment. Not only do different types of cancer have common biological characteristics such as abnormal cell differentiation and proliferation, lack of growth control, invasion and metastasis, but they also have many specifically biological characteristics, respectively. Therefore, it is necessary to conduct pan-cancer research on a variety of cancer types to ascertain the similarities and differences in molecular characteristics across different cancers.

DNA methylation is an important epigenetic modification that plays an important role in many physiological processes (Martin-Subero, 2011; Li et al., 2013; Neri et al., 2017), and its aberrant behavior can result in gene instability, proto-oncogene activation and tumor suppressor gene inactivation (Li et al., 2012; Hou et al., 2021). Aberrant DNA methylation occurs in almost all
cancers, where unmethylated promoters become methylated or methylated sequences lose their methylation. Focusing on the pan-cancer analysis of DNA methylation can inform research and therapy. Saghainia et al. (Saghainia et al., 2018) described an algorithmic strategy for identifying pan-cancer-related DNA methylation alteration affecting gene expression. Numerous DNA methylation dysregulation were discovered to be associated with patient prognosis and therapeutic response. Methylation research in cancer patients has been shown to improve and maintain the efficiency of cancer treatment (Pauken et al., 2016; Sen et al., 2016; Marwitz et al., 2017).

Long non-coding RNAs (lncRNA) are non-coding RNAs with a length of more than 200 nucleotides. In comparison to protein-coding genes, they have a high degree of tissue specificity in their expression (Cabili et al., 2011). LncRNAs play an important role in the occurrence and development of cancer and many other complex diseases (Zhou et al., 2019). LncRNAs are associated with practically every major cancer type and contribute to all ten hallmarks of cancer (Huarte, 2015; Bartonicek et al., 2016; Espósito et al., 2019; Bao et al., 2020). Additionally, lncRNAs play an important role in immune regulation (Chen et al., 2017; Esposito et al., 2019; Bao et al., 2020). Moreover, lncRNAs (Pauken et al., 2016; Sen et al., 2016; Marwitz et al., 2017).

Zhang et al. (Zhang et al., 2021b) identified lncRNA MTIJP as a ceRNA for the tumor suppressor FBXW7 in gastric cancer by demonstrating competitive binding with MiR-92A-3p. Xu et al. (Xu et al., 2019) discovered that lncRNA SNHG6 regulated the expression of the oncogene EZH2 in colorectal cancer via the ceRNA sponge-associated with MiR-26a/b and MiR-214. Chen et al. (Chen et al., 2019) discovered that the lncRNA PVT1 promoted tumor development in gallbladder cancer by regulating the miR-143/HK2 axis. Wang et al. (Wang et al., 2017) identified the lncRNA HOXD-AS1 as a ceRNA that regulated SOX4 and promoted liver cancer metastasis. The above studies established the role of lncRNAs in corresponding cancer types, and identified the cancer-associated lncRNAs for each cancer type (Dong et al., 2019). Pan-cancer analysis of lncRNAs can help in identifying the similarities and differences between distinct cancer types and identifying potential therapeutic targets for cancer treatment.

Numerous pan-cancer studies have been carried out on lncRNAs. Li et al. (Yongsheng Li et al., 2020) identified multiple pan-cancer immune-associated lncRNAs as potential oncogenic biomarkers. Martens-Uzunova et al. (Martens-Uzunova et al., 2014) summarized the role of lncRNAs in the diagnosis and treatment of urinary tumors, and concluded that lncRNAs could be used as new biomarkers for prostate cancer, kidney cancer and bladder cancer. Zhang et al. (Zhang et al., 2021) identified clinically distinct tumor subtypes by characterizing pan-cancer lncRNA modifiers of the immune microenvironment. Bao et al. (Bao et al., 2021) proposed a framework for identifying lncRNA signatures associated with pan-cancer prognosis.

Previous studies have established a correlation between lncRNAs and epigenetic regulation (Spizzo et al., 2012; Xu et al., 2018), suggesting that they regulate chromatin state and epigenetic inheritance (Tsai et al., 2010). Lu et al. (Lu et al., 2020) found that DNA methylation-mediated lncRNA activation improved temozolomide resistance in glioblastoma, implying that SNHG12 could be a therapeutic target for overcoming temozolomide tolerance.

Detecting the dynamic pattern of lncRNA methylation during cancer development across pan-cancer may help highlight epigenetic changes and aid in cancer diagnosis and treatment. Yang et al. (Yang et al., 2021) presented a novel integrative analysis framework, termed MeLncTRN for integrating data on gene expression, copy number variation, methylation and lncRNA expression. They identified epigenetically-driven lncRNA-gene regulation circuits across 18 cancer types. Wei et al. (Wei et al., 2019) constructed a systematic biological framework to evaluate the co-methylation events between two lncRNAs in nine cancer types. The lncRNA prognostic signatures were identified to significantly correlate with overall survival in cancers. Wang et al. (Wang et al., 2018) characterized the epigenetic landscape of genes encoding lncRNAs associated with pan-cancer and identified EPIC1 as an oncogenic lncRNA. Xu et al. (Xu et al., 2021) constructed networks of lncRNA-associated dysregulated ceRNA across eight cancer types. They screened nine pan-cancer epigenetically related lncRNAs.

However, no research has been conducted to systematically compare methylation changes of lncRNAs in pan-cancer to identify the specific methylation-related lncRNAs. In this study, we used pan-cancer lncRNA methylation data from the TCGA to examine the lncRNA methylation patterns of 23 cancer types and identified differentially methylated lncRNAs (DMlncs). Subsequently, we examined differentially methylated lncRNAs

| Cancer | No. of tumor samples | No. of normal samples | No. of lncRNAs |
|--------|----------------------|-----------------------|---------------|
| BLCA   | 412                  | 21                    | 4,317         |
| BRCA   | 782                  | 96                    | 4,314         |
| CESC   | 307                  | 3                     | 4,314         |
| CHOL   | 36                   | 9                     | 4,309         |
| COAD   | 296                  | 38                    | 4,317         |
| ESCA   | 185                  | 16                    | 4,317         |
| GBM    | 141                  | 2                     | 4,317         |
| HNSC   | 528                  | 50                    | 4,317         |
| KIRC   | 319                  | 160                   | 4,317         |
| KIRP   | 275                  | 45                    | 4,317         |
| LIHC   | 377                  | 50                    | 4,317         |
| LUAD   | 458                  | 32                    | 4,316         |
| LUSC   | 370                  | 42                    | 4,317         |
| PAAD   | 184                  | 10                    | 4,317         |
| PCPG   | 179                  | 3                     | 4,315         |
| PRAD   | 498                  | 50                    | 4,317         |
| READ   | 98                   | 7                     | 4,316         |
| SARC   | 261                  | 4                     | 4,310         |
| SKCM   | 470                  | 2                     | 4,317         |
| STAD   | 395                  | 2                     | 4,314         |
| THCA   | 507                  | 58                    | 4,314         |
| THYM   | 124                  | 2                     | 4,316         |
| UCEC   | 431                  | 46                    | 4,314         |
from different cancer types to identify cancer-specific and cancer-common differentially methylated IncRNAs. Further, combining IncRNA expression data with survival data, IncRNAs with a negative correlation between methylation and expression dysregulation were found as diagnostic biomarkers for each cancer. Finally, the IncRNAs were mapped into the ceRNA network to establish ceRNA relationships with mRNAs confirming their important roles in cancer.

**MATERIALS AND METHODS**

**Data**

Pan-cancer DNA methylation data for IncRNAs from Infinium 450k arrays were downloaded from the TCGA database. The cancer types with normal samples were selected, and a total of 7,634 tumor samples and 746 normal samples from 23 cancer types were retained. Table 1 shows the number of tumor and normal samples, as well as the number of IncRNAs. The methylation status of each probe in each sample was measured using the \( \beta \)-value (Aryee et al., 2014). The \( \beta \)-value denoted the ratio of methylation intensity of the probe to total intensity, with a range of 0 (low methylation) to 1 (high methylation). The probes with \( \beta \)-values greater than 0 in more than 50% of the samples were retained in the methylation profile for each cancer, and the missing values were filled with the average of all non-zero values on the probes. The average \( \beta \)-value of the promoter region was used to determine the methylation level of each IncRNA.

Expression data of IncRNAs and mRNAs for 13 cancer types were downloaded from the TANRIC and the TCGA databases, respectively (Li et al., 2015). The number of tumor and normal samples and the number of IncRNAs and mRNAs are shown in Supplementary Tables S1, S2. Each IncRNA and mRNA expression value was defined as its reads per kilobase per million mapped reads (RPKM) (Mortazavi et al., 2008). Subsequently, we transformed the expression data by \( \log_2 (\text{RPKM}+1) \), reserved the IncRNAs and mRNAs with expression values in more than 70% of the samples, and filled their missing values using the average expression values of these RNAs in the samples.

The expression data of mRNA and IncRNA was downloaded from different databases, so we got the human gene annotation files from the GENCODE database (https://www.gencodegenes.org/) to obtain the corresponding relations of ENSG IDs and gene symbols. Then, using Entrez IDs as the main reference, different versions of human gene names (Entrez IDs, gene symbols and ENSG IDs) were converted to standard human gene names.

**Identification of DMlncs**

The DMlncs for each cancer were first screened using the following formula \( \Delta \beta \):

\[
\Delta \beta = |\bar{\beta}_t - \bar{\beta}_n|
\]  

Where, \( \bar{\beta}_t \) and \( \bar{\beta}_n \) denoted the average level of methylation in tumor and normal samples, respectively. \( \Delta \beta \) was the subtraction difference in average methylation levels between tumor and normal samples. IncRNAs with \( \Delta \beta > 0.1 \) were selected as candidates for DMlncs.

Additionally, the “limma” package (Ritchie et al., 2015) in R language was used to measure the degree of difference between tumor and normal samples. The IncRNAs with \( |\log_2 (\text{fold change})| > 1 \) (PCAWG Transcriptome Core Group et al., 2020) and FDR<0.05 were identified as DMlncs.

**Identification of Differentially Expressed IncRNAs**

The “limma” package was used to calculate differential expression between tumor and normal samples for IncRNA expression data. We took the IncRNAs with \( |\log_2 (\text{fold change})| > 1 \) and FDR<0.05 as differentially expressed IncRNAs.

**Functional Enrichment Analysis**

The GREAT software (McLean et al., 2010) was used to conduct functional enrichment analysis on IncRNAs. We took the IncRNA BED data as input. The IncRNA BED information included the chromosome, start site and end site extracted from GENCODE database. Gene ontology (GO) functions of the output results were selected for subsequent analysis.

**Recognition of ceRNAs**

LncRNA-miRNA and mRNA-miRNA targeted relationships were downloaded from the ENCORI platform (Li et al., 2014). For each pair of IncRNA and mRNA, the intersection of their target miRNAs should be more than two (Zhang et al., 2021a). The intersections of their miRNA lists were subjected to the hypergeometric test, and the lncRNA-mRNA pairs with a \( p \)-value less than 0.05 were considered. A total of 4,739,668 pairs were obtained.

Subsequently, the Pearson correlation coefficient between IncRNA expression and mRNA expression was calculated. LncRNA-miRNA pairs with Pearson correlation coefficient \( > 0.3 \) and \( p \)-value\(<0.05 \) were selected as ceRNAs. Supplementary Table S3 shows the number of lncRNA-mRNA pairs, IncRNAs and mRNAs in each cancer.

**Survival Analysis**

Survival analysis of patients was carried out using the “survival” package in R language, in which the maxstat model was used to evaluate the best cut-off point to divide high-risk and low-risk groups. Kaplan-Meier curves were then drawn to depict the survival of patients of high-risk and low-risk groups.

**Immunological Score**

Three scores were used to assess the immunological effect of IncRNAs: Major Histocompatibility Complex (MHC), Cytolytic Activity (CYT) and Cytotoxic T Lymphocyte (CTL). The MHC score of each sample was calculated as:

\[
\text{Score}_{\text{MHC}} = \left( \sum \exp p_n \right) / 9
\]

Where \( i \) denoted the sample, \( \exp p_n \) represented the expression of gene \( n \), and \( n \) was one of the nine genes (HLA-A, PSMB9, HLA-B,
PSMB8, HLA-C, B2M, TAP2, NLRC5, and TAP1). These nine genes had a strong correlation and were the core gene set of MHC-I (Rooney et al., 2015; Lauss et al., 2017).

The CYT score of each sample was calculated as follows:

\[
\text{Score}_{\text{CYT}} = \left( \exp \text{GZMA} + \exp \text{PRF1} \right) / 2
\]  

In which, \(i\) represented the sample, \(\exp \text{GZMA}\) and \(\exp \text{PRF1}\) represented the expression of GZMA and PRF1, respectively. These two genes were key factors of cytolysis and were upregulated in activated CD8+T cells and strongly responded to CTLA4 and PDCD1 immunotherapy (Narayanan et al., 2018).

The CTL score of each sample was calculated as follows:

\[
\text{Score}_{\text{CTL}} = \left( \exp \text{GZMA} + \exp \text{PRF1} + \exp \text{GZMB} \right) / 3
\]  

In which, \(i\) represented the sample, \(\exp \text{GZMA}\), \(\exp \text{PRF1}\), and \(\exp \text{GZMB}\) represented the expression of GZMA, PRF1, and GZMB, respectively. These three genes were important factors to measure T cell toxicity and immune cell effector function (Basu et al., 2016).

RESULTS

The Workflow of DMlncs Identification

In this study, the identification flow of methylation-related lncRNA biomarkers for pan-cancer is shown in Figure 1. We conducted a study on a total of 23 cancer types. Firstly, DMlncs between tumor and normal samples were obtained using lncRNA methylation data. Specific and common lncRNAs were identified for pan-cancer. Subsequently, lncRNA methylation data and lncRNA expression data were integrated to identify DMlncs whose methylation changes were negatively correlated with expression changes. Following that, prognostic lncRNAs were screened and mapped into the ceRNA network. The ceRNA network was constructed by combining lncRNA and mRNA expression data. Finally, pan-cancer lncRNA biomarkers were identified by analyzing the lncRNAs of the ceRNA network.

Identification of Pan-Cancer DMlncs

To identify DMlncs, we downloaded cancer methylation profiles of 23 cancer types from the TCGA, including both tumor and normal samples. The data of 7,634 tumor samples and 746 normal samples was downloaded. The proportion of tumor samples for each cancer is shown in Figure 2A. BRCA had the largest number of samples, which was nearly double that of other cancer types, while CHOL had the lowest number of samples. A total of 7,542 tumor samples had corresponding clinical data of the patients. As shown in Figure 2B, the age, sex, and survival status of patients were analyzed. The male to female ratio of the patients was 1:1, and the age was concentrated among the elderly, which was consistent with the law of the general onset age of cancer. The majority of patients survived following treatment. Figure 2C shows the mortality rate of tumor patients, three-quarters of the patients survived after surgery. GBM had the highest mortality rate, followed by CHOL. PRAD had the lowest mortality rate.

After standardizing the data, DMlncs for each cancer were identified. The number of DMlncs is shown in Table 2. A total of 2,286 DMlncs were obtained. The majority of the cancer types had a large number of DMlncs, and only a few cancer types had a small number of DMlncs.

Additionally, DMlncs were divided into up-methylated and down-methylated groups. Up-methylated lncRNAs were those whose methylation level was elevated in cancer samples compared with normal samples, while down-methylated lncRNAs were the inverse. Table 2 shows the number of patients in each of the two groups. There were overlaps in the genes of different cancers. In total, there were 1,229 DMlncs in the up-methylated group and 1,654 DMlncs in the down-methylated group. Among them, 597 lncRNAs were up-methylated in some cancers but down-methylated in some other cancers, demonstrating uneven regulation tendencies across different cancers. As shown in Figure 3, the proportion of up-methylated and down-methylated lncRNAs varied between cancer types. There were 13 types of cancer that had a higher number of down-methylated lncRNAs and ten types of cancer had a higher number of up-methylated lncRNAs. The first few cancer types of most DMlncs had obvious higher number of down-methylated lncRNAs and most of the other cancer types had more up-methylated lncRNAs. Certain cancer types contained only a single type.
of DMlncs. For example, all 74 DMlncs in SKCM were up-methylated, and all three DMlncs in THYM were up-methylated as well.

**Cancer-Specific IncRNA Biomarkers**

As a complex disease, cancer has a high heterogeneity and distinct pathogenesis. In this study, we searched for specific DMlncs for cancer detection and treatment.

**TABLE 2** | The number of DMlncs of each cancer.

| Cancer   | Total IncRNAs | Down-methylated IncRNAs | Up-methylated IncRNAs | Cancer-specific up-methylated | Cancer-specific down-methylated |
|----------|---------------|--------------------------|-----------------------|-------------------------------|--------------------------------|
| BLCA     | 1,126         | 931                      | 195                   | 2                             | 45                             |
| BRCA     | 753           | 411                      | 342                   | 23                            | 9                              |
| CESC     | 395           | 99                       | 298                   | 20                            | 8                              |
| CHOL     | 433           | 80                       | 383                   | 65                            | 23                             |
| COAD     | 762           | 464                      | 298                   | 28                            | 14                             |
| ESCA     | 569           | 263                      | 306                   | 28                            | 1                              |
| GBM      | 48            | 7                        | 41                    | 13                            | 2                              |
| HNSC     | 854           | 581                      | 273                   | 9                             | 12                             |
| KIRC     | 435           | 286                      | 149                   | 5                             | 24                             |
| KIRP     | 406           | 154                      | 252                   | 57                            | 4                              |
| LIHC     | 1,074         | 936                      | 138                   | 10                            | 114                            |
| LUAD     | 563           | 349                      | 234                   | 5                             | 1                              |
| LUSC     | 952           | 674                      | 258                   | 11                            | 18                             |
| PAAD     | 344           | 150                      | 194                   | 11                            | 9                              |
| PCPG     | 174           | 112                      | 62                    | 26                            | 47                             |
| PRAD     | 604           | 236                      | 368                   | 77                            | 13                             |
| READ     | 575           | 375                      | 200                   | 5                             | 6                              |
| SARC     | 7             | 1                        | 6                     | 2                             | 1                              |
| SKCM     | 74            | 0                        | 74                    | 22                            | 0                              |
| STAD     | 8             | 7                        | 1                     | 0                             | 7                              |
| THOA     | 118           | 99                       | 19                    | 3                             | 5                              |
| THYM     | 3             | 0                        | 3                     | 2                             | 0                              |
| UCEC     | 1,015         | 651                      | 364                   | 32                            | 47                             |
each cancer. Figure 4A shows the proportion of cancer-specific lncRNAs, and only a fraction of the DMlncs were cancer-specific. LIHC had the most specifically down-methylated lncRNAs, whereas PRAD had the most specifically up-methylated lncRNAs. The proportion of both specifically up and down lncRNAs in PCPG was about 42%. SARC, STAD, and THCA had DMlncs less than ten, preventing them from being compared with other cancers in terms of lncRNA proportion. Except for SARC, STAD, and THCA, PCPG had the highest proportion of specific lncRNAs. STAD had no up-methylated lncRNAs, while all its seven down-methylated lncRNAs were specific. SARC possessed a single down-methylated lncRNA, and it was specific. Two of the three DMlncs of THYM were specific. Supplementary Table S4 shows the specific DMlncs of each cancer.

Subsequently, we performed functional enrichment analysis of specifically up-methylated and down-methylated lncRNAs in each cancer. For each cancer, the functions enriched by specifically up-methylated and down-methylated lncRNAs were analyzed separately, and the number of functions enriched by each cancer is shown on the right side of Figure 4B. There were significant differences in the number of enriched functions for the cancers. LIHC, COAD, and PAAD enriched in more than 200 functions, while the enriched
functions of THYM, STAD, and HNSC were less than five. There also a difference between the number of functions of up-methylated and down-methylated lncRNAs for each cancer. The specifically down-methylated lncRNAs of COAD were enriched in the majority of functions (261), whereas up-methylated lncRNAs of COAD were enriched in only a few functions (14), and LIHC demonstrated a similar pattern. The specifically up-methylated lncRNAs of ESCA were found to be enriched in a variety of functions (194), but the down-methylated lncRNAs were enriched in no functions. The details of enriched functions are shown in Supplementary Table S5.

Among the functions, we selected the most significant function for each group and displayed them on the left side of Figure 4B. The function names, enrichment p-values and other information are shown in Table 3. The lncRNAs of most cancer types were enriched in the “regulation” or “response” functions. Both the up-methylated group of CHOL and the down-methylated group of LIHC were enriched in “depyrimidination”. The specifically up-methylated lncRNAs of GBM were enriched in “methyltransferase activity”, specifically down-methylated lncRNAs of COAD were enriched in “dimethylation”, and the specifically up-methylated lncRNAs of BRCA were enriched in “epigenetic”.

These results established the important role of lncRNAs in the epigenetic process.

In this study, we hypothesized that up-methylated lncRNAs would exhibit decreased expression and vice versa. That is, there is a negative correlation between changes in methylation and expression levels. In order to screen out negative correlated lncRNAs in DMlncs, we used lncRNA expression data to identify differentially expressed lncRNAs in various cancers.

LncRNA expression data of 13 cancer types were downloaded, and the number of differentially expressed lncRNAs in each cancer is shown in Supplementary Table S6. We identified 2,887 over-expressed lncRNAs and 2,375 low-expressed lncRNAs in different cancers. The total number of these lncRNAs was 4,155, and several genes were overlapped between two groups and displayed conflicting regulation patterns across different cancers. The numerical distribution of differentially expressed lncRNAs in various cancers is shown in Figure 5A. Although the number of over-expressed and low-expressed lncRNAs was similar in the majority of cancer types, there were significantly more low-expressed lncRNAs in BRCA and THCA, and significantly more over-expressed lncRNAs in LIHC and STAD.

### TABLE 3 | The functions of specific lncRNAs for each cancer.

| Group | Cancer | ID | Function | FDR q-value |
|-------|--------|----|----------|-------------|
| down  | BLCA   | GO:0.036,207 | Interstrand cross-link repair | 3.34E-03 |
| up    | BLCA   | GO:0.010,729 | Positive regulation of hydrogen peroxide biosynthetic process | 2.07E-05 |
| down  | BRCA   | GO:0.048,284 | Organelle fusion | 6.43E-03 |
| up    | BRCA   | GO:0.040,030 | Regulation of molecular function, epigenetic | 1.15E-10 |
| down  | CESC   | GO:0.001,887 | Selenate compound metabolic process | 2.47E-04 |
| up    | CESC   | GO:0.048,846 | Axon extension involved in axon guidance | 4.69E-03 |
| down  | CHOL   | GO:0.034,587 | PiRNA metabolic process | 1.32E-02 |
| up    | CHOL   | GO:0.045,008 | Depyrimidination | 1.29E-12 |
| down  | COAD   | GO:0.018,016 | N-terminal peptidyl-proline dimethylation | 4.41E-13 |
| up    | COAD   | GO:0.034,214 | Protein hexamerization | 1.23E-05 |
| up    | ESCA   | GO:0.010,157 | Response to chloride | 4.36E-08 |
| down  | GBM    | GO:2.001,162 | Positive regulation of histone H3-K79 methylation | 7.71E-03 |
| up    | HNSC   | GO:0.030,157 | Pancreatic juice secretion | 2.35E-02 |
| down  | HNSC   | GO:0.005,993 | Threonine catabolic process | 3.22E-03 |
| down  | KIRC   | GO:0.046,597 | Negative regulation of viral entry into host cell | 3.49E-18 |
| down  | KIRP   | GO:0.007,144 | Female meiosis I | 8.51E-03 |
| up    | KIRP   | GO:0.046,878 | Positive regulation of saliva secretion | 1.03E-02 |
| down  | LIHC   | GO:0.045,008 | Depyrimidination | 2.81E-13 |
| up    | LIHC   | GO:0.021,897 | Forebrain astrocyte development | 1.48E-05 |
| up    | LUAD   | GO:0.035,814 | Negative regulation of renal sodium excretion | 3.03E-03 |
| down  | LUSC   | GO:0.019,358 | Nicotinate nucleotide salvage | 1.34E-03 |
| up    | LUSC   | GO:0.045,006 | DNA demethylation | 1.65E-02 |
| down  | PAAD   | GO:0.070,585 | Protein localization to mitochondrion | 1.63E-08 |
| up    | PAAD   | GO:0.046,329 | Negative regulation of JNK cascade | 1.04E-05 |
| down  | PCPG   | GO:0.021,723 | Medullary reticular formation development | 3.25E-05 |
| up    | PCPG   | GO:0.006,335 | DNA replication-dependent nucleosome assembly | 1.72E-04 |
| down  | PRAD   | GO:0.001,957 | Intramembranous ossification | 5.15E-03 |
| up    | PRAD   | GO:0.045,053 | Protein retention in Golgi apparatus | 6.23E-09 |
| down  | READ   | GO:0.060,123 | Regulation of growth hormone secretion | 6.57E-05 |
| up    | READ   | GO:0.042,742 | Defense response to bacterium | 3.49E-04 |
| down  | SARC   | GO:0.010,324 | Membrane invagination | 7.98E-05 |
| up    | SARC   | GO:0.002,931 | Response to ischemia | 8.74E-05 |
| down  | STAD   | GO:0.080,184 | Response to phenylpropanoid | 1.85E-04 |
| down  | THCA   | GO:0.007,161 | Calcium-dependent cell-matrix adhesion | 6.69E-05 |
| down  | UCEC   | GO:0.000,978 | Negative regulation of forebrain neuron differentiation | 1.50E-03 |
| up    | UCEC   | GO:0.070,208 | Protein heterotrimerization | 1.10E-07 |
Subsequently, negatively correlated lncRNAs (NClnCs) were screened from differentially expressed lncRNAs. The number of NClnCs in each cancer is shown in Figure 5A. NClnCs were divided into two types: up-methylated-low-expressed lncRNAs (UMLElncs) and down-methylated-over-expressed lncRNAs (DMOElncs) based on the changes in expression and methylation levels. The two types of lncRNAs in each cancer are shown in Figure 5B–N. As seen from the figure, the number of NClnCs was proportional to the number of differentially expressed lncRNAs. However, the proportion of differential lncRNAs was not balanced in several cancers. For example, STAD and THCA had many differentially expressed lncRNAs, but a small number of DMIlnCs, implying a small number of NClnCs. The proportion of UMLElncs and DMOElncs was different in each cancer. For example, both types of lncRNAs for LUSC were abundant. HNSC and LIHC had significantly more DMOElncs, while BRCA and PRAD had significantly more UMLElncs.

The cancer-specific negatively correlated lncRNAs (CSNClncs) were then identified, and the results are shown in Table 4. Among the cancer-specific DMIlnCs, 49 NClnCs were identified, with LIHC having the highest CSNClncs (11), followed by PRAD (7), BLCA (6) and KIRC (6). LUAD, THCA, and UCEC did not contain any CSNClnc.

| Cancer | Negatively correlated lncRNAs |
|--------|--------------------------------|
| BLCA   | X0bac-B476C20, RP5-943J3, RP11-390P2, AC006116, AC073046, RP11-514P8 |
| BRCA   | RP11-667K14, RP11-1094M14, RP11-497H17, LINC00619 |
| CESC   | MEP2C-AS1, CTA-384D8, CTD-2035E11 |
| HNSC   | LA16c-390E6, CTC-548K16, LA16c-325D7 |
| KIRC   | RP11-488L18, AC027601, RP11-118J21, HLA-F-AS1, SNHG312, EPB41L4A-AS1 |
| KIRP   | ACTA2-AS1, RP11-77H9, RP11-126K1 |
| LIHC   | AC007879, AC025535, CTC-246B18, HULC, LINC00665, RP11-215P8, RP11-890B15, RP11-968A15, RP11-973H7, RP9-3954R20 and TEX41 |
| LUSC   | Z83851, RP11-311F12, RP11-757G1, RP11-12L8 |
| PRAD   | MIR205HG, LINC01018, JAZF1-AS1, RP4-639F20, RP1-223B1, RP11-597D13, LINC00115 |
| STAD   | MNX1-AS1, RP11-29893 |
Additionally, survival-correlated lncRNAs were identified by analyzing the CSNClncs of each cancer. As shown in Table 5, 29 lncRNAs were identified to be specifically diagnostic markers for corresponding cancers. They not only exhibited synergistic alterations in expression and methylation, but were also closely associated with the survival of cancer patients.

**TABLE 5** | Significantly survival associated lncRNAs in negatively correlated lncRNAs for each cancer.

| Cancer | Count | lncRNAs |
|--------|-------|---------|
| BLCA   | 5     | RP11-390P2, XXbac-B476C20, RP5-943J3, AC073048, AC006116 |
| BRCA   | 3     | LINC00619, RP11-497H17, RP11-667K14 |
| CESC   | 2     | MEF2C-AS1, CTD-2035E1 |
| HNSC   | 3     | LA16c-390ES, LA16c-325D7, CTC-548K16 |
| KIRC   | 5     | SNHG12, EPB41L4A-AS1, RP11-488L18, AC027601, RP1-118J21, ACTA2-AS1, RP11-77H9 |
| KIRP   | 2     | RP11-215P8, RP11-968A15, RP11-973H7, CTC-246B18 |
| LIHC   | 4     | RP11-757G1, RP11-311F12 |
| LUAD   | 2     | RP11-223B1, LINC01018 |
| STAD   | 1     | MNX1-AS1 |

**FIGURE 6** | The negatively correlated lncRNAs which specific to each cancer in the ceRNA network (A) The ceRNAs of BLCA-specific negatively correlated lncRNAs (B) The ceRNAs of KIRC-specific negatively correlated lncRNAs (C) The ceRNAs of KIRP-specific negatively correlated lncRNAs (D–I) Kaplan-Meier curves for the lncRNAs associated with each cancer types. The red line represents the group with a high level of expression, while the blue line represents the group with a low level of expression. Additionally, the “+” on the lines represents patients who were lost to follow-up. At this point, the number of patients decreases but the overall survival rate remains stable.
The main function of lncRNAs was to combine miRNAs competing with mRNAs, thereby increasing mRNA expression. As the methylation of lncRNAs increased, their expression decreased, which indirectly led to the decrease in mRNA expression, and vice versa. Therefore, we screened the ceRNAs of DMlncs. First, the lncRNA-miRNA and mRNA-miRNA regulatory relationships were integrated. A lncRNA and a mRNA sharing more than two miRNAs were considered to have a ceRNA relationship. The correlation between lncRNA and mRNA expression was calculated for each cancer to confirm the ceRNA relationship. We maintained the ceRNA relationships that showed a positive correlation between lncRNA and mRNA expression. Each survival-related CSNCInc was put into the ceRNA network to search for associated mRNAs. The visualized results of ceRNAs for BLCA, KIRC and KIRP are shown in Figures 6A–C.

BLCA had two lncRNAs mapped into the ceRNA network (AC006116 and AC073046). Both lncRNAs formed a ceRNA relationship with the mRNA ATXN7L3B and formed ceRNA relationships with some other mRNAs, respectively. ATXN7L3B have been confirmed to be associated with an increased risk of colorectal cancer (Leberfarb et al., 2020), and cytoplasmic ATXN7L3B interfered with the nuclear functions of the SAGA deubiquitinate complex (Li et al., 2016). The SAGA complex was composed of two enzymatic modules, which house histone acetyltransferase (HAT) and deubiquitinate (DUB) activities. The DUB module was important for normal embryonic development (Glinsky, 2006; Lin et al., 2012), and alterations in the expression or structure of component proteins were linked to cancer (Lan et al., 2015). Therefore, AC006116 and AC073046, as its ceRNAs, could regulate the expression of ATXN7L3B and were also closely related to cancer. The survival correlation of AC006116 and AC073046 in BLCA is shown in Figures 6D,E. Both lncRNAs were significantly associated with the survival of BLCA patients.

Three lncRNAs in KIRC were mapped into the ceRNA network, among which SNHG12 and AC027601 shared more than 30 mRNAs, whereas EPB41L4A-AS1 did not share any mRNAs with the other two lncRNAs. The survival analysis results for the three lncRNAs in KIRC are shown in Figures 6F–H. The high expression of SNHG12 and AC027601 both showed a worse prognosis. However, the low expression group of EPB41L4A-AS1 showed a worse prognosis. Therefore, we inferred that in KIRC, lncRNAs SNHG12 and AC027601 had a carcinogenic effect, whereas EPB41L4A-AS1 had a tumor-suppressive effect, which explains why it shared no mRNAs with the other two lncRNAs. Numerous studies have established that SNHG12 is associated with cancers (Zhang et al., 2020), and could be used as a potential therapeutic target and biomarker for human cancers (Tamang et al., 2019). DNA-methylation-mediated activation of SNHG12 promoted temozolomide resistance in glioblastoma (La et al., 2020). SNHG12 promoted tumor progression and sunitinib resistance by upregulating CDCA3 in renal cell carcinoma (Liu et al., 2020). EPB41L4A-AS1 was a repressor of the Warburg effect and played an important role in the metabolic reprogramming of cancer (Liao et al., 2019), and EPB41L4A-AS1 has been identified as a potential biomarker in non-small cell lung cancer (Wang et al., 2020). At present, AC027601 has been identified as a survival signature in renal clear cell carcinoma (Qi-Dong et al., 2020), but has not been reported in other cancers. Given that the three lncRNAs were simultaneously identified as KIRC-related lncRNAs in this study, we believed that AC027601 should be closely associated with the occurrence and development of KIRC, and it is a newly identified cancer-related lncRNA.

Only ACTA2-AS1 was mapped into the ceRNA network in KIRP, where it established a ceRNA relationship with PPP1R12B. PPP1R12B has been shown to inhibit tumor growth and metastasis by regulating Grb2/P13K/Akt signaling in colorectal cancer (Ding et al., 2019). ACTA2-AS1 plays an important role in a variety of cancers, for example, ACTA2-AS1 is significantly associated with overall survival in ovarian cancer patients (Li and Zhan, 2019). ACTA2-AS1 plays different roles in different cancers. ACTA2-AS1 knockdown promotes liver cancer cell proliferation, migration and invasion (Zhou and Lv, 2019), while ACTA2-AS1 suppresses lung adenocarcinoma progression (Yang et al., 2020), implying an inhibitory effect on the two cancers. However, ACTA2-AS1 promotes cervical cancer progression (Luo et al., 2020), suggesting its carcinogenic role in cancer. Figure 6I shows the survival analysis result for ACTA2-AS1 in KIRP. We believed that ACTA2-AS1 had a carcinogenic effect in KIRP.

Common lncRNA Biomarkers in Cancers

All cancer types exhibited infinite proliferation, transformation and ease of metastasis. Therefore, we sought to identify DMlncs common to different cancers to help understand the mechanisms underlying the occurrence of common features in cancers. First, the intersection of DMlncs were searched in cancers, and the results are shown in Figure 7. The upper triangle and lower triangle reflected the intersection of up-methylated lncRNAs and the intersection of down-methylated lncRNAs, respectively. The findings were consistent with the hypothesis that the larger the lncRNA set, the greater the overlap with other cancers. The intersections of DMlncs of SARC, STAD, and THYM with other cancers were small. The down-methylated lncRNAs of GBM and SKCM had small intersections with other cancers, whereas the up-methylated lncRNAs of THCA had small intersections with other cancers. There were no intersections of up-methylated lncRNAs in PCPG (62), but the overlaps with other cancers were small. Among the down-methylated lncRNAs, the intersection of KIRC and KIRP was the largest of KIRC, but only ranked 10th of the KIRP. Among the up-methylated lncRNAs, the intersection of KIRC and KIRP was the largest of KIRC, while was the second largest of KIRP.

Subsequently, we extracted DMlncs which were common in various cancers. The findings indicated that there were 19 common DMlncs in more than 15 cancers (RP4-792G4, RP5-855F14, OTX2-AS1, RP11-52L5, CYP1B1-AS1, RP11-175E9, RP11-552E20, HCCAT3, RP11-718011, HOXA-AS2, RP3-326L13, AC007228, RP11-297B11, CTC-523E23, LINC00100, RP11-227D2, EVX1-AS, AC018730, and RP11-465L10). Fourteen of them were up-methylated in all the cancers, indicating their carcinogenic potential, whereas three lncRNAs were down-methylated in the majority of cancers and may act as
Figure 8A shows the comparison of methylation levels of the 19 lncRNAs in tumor and normal samples. It is intuitive to conclude that there were significant differences in lncRNAs methylation levels between tumor and normal samples. Except for LINC01010, RP11-552E20, and RP5-855F14, the lncRNAs had a higher methylation level in tumor samples.

Among the 19 lncRNAs, OXT2-AS1 was shown to be significantly down-methylated in lung squamous cell carcinoma and was closely associated with poor prognosis of cancer (Zheng et al., 2021). CYP1B1-AS1 has been confirmed to play an important role in triple-negative breast cancer, lung adenocarcinoma and acute myeloid leukemia, and was associated with the prognosis of these cancers (Cheng et al., 2021; Ren et al., 2021; Vishnubalaji and Alajez, 2021). Abnormal methylation and low expression of CTC-523E23 led to poor prognosis in patients with lung squamous cell carcinoma (Rui Li et al., 2020). Inhibition of LINC01010 may promote the migration and invasion of lung cancer cells (Cao et al., 2020) and help in the prediction of neuroblastoma prognosis (Gao et al., 2020).

Then, we investigated the functions of the 19 lncRNAs and performed functional enrichment analysis on these lncRNAs using the GREAT software. Figure 8C shows that these lncRNAs were enriched in processes required for organisms such as metabolism and biosynthesis, as well as those closely associated with the occurrence and development of cancer, such as gene expression and transcription.

NClnCs in common lncRNAs were investigated in combination with the results of differentially expressed lncRNAs. Seven of the 19 lncRNAs were found to be differentially expressed, and four (CYP1B1-AS1 (BLCA, BRCA, and LUSC), AC007228 (BLCA, HNSC, LIHC, LUAD, LUSC and PRAD), HOXA-AS2 (BRCA) and LINC01010 (HNSC)) of them had a negative correlation in multiple cancers. Additionally, we analyzed the four NClnCs shared by cancers for their correlation with survival, and survival-related lncRNAs were identified. The four lncRNAs were associated with patients’ survival in a variety of cancers, including CYP1B1-AS1 in eight types of cancers (BRCA, HNSC, KIRC, KIRP, LIHC, LUAD, PRAD and THCA), AC007228 in seven types of cancers (BLCA, HNSC, KIRC, KIRP, LIHC, LUAD and THCA), LINC01010 in five types of cancers (BRCA, LIHC, LUAD, LUSC and THCA) and HOXA-AS2 in four types of cancers (BRCA, KIRC, KIRP and THCA) (Supplementary Figures S1–S4).

To further validate the cancer-common lncRNAs identified in this study, significantly survival-related lncRNAs were mapped to the ceRNA network, and only mRNAs shared by more than four types of cancers were selected for further study. Finally, a subnetwork comprising 33 mRNAs and three lncRNAs

FIGURE 7 | The intersection of DMlncs for cancers. The upper red triangle represents the intersection of up-methylated lncRNAs, while the lower blue triangle represents the intersection of down-methylated lncRNAs.
The ceRNA distribution of the 33 mRNAs in cancers is shown in Figure 9A. Although these mRNAs were shared by multiple cancers, they generally formed ceRNA relationships with a greater number of lncRNAs and had higher correlation coefficients in KIRP. HOXA3 had ceRNA lncRNAs in ten cancers and showed high correlation coefficients in BRCA, CESC, HNSC and LUSC. Figure 9B shows the ceRNA subnetwork, in which HOXA-AS2 and AC007228 were shared by ten cancers, and the two lncRNAs shared some ceRNA relationships with some mRNAs (DMTF1, HOXB3, NOD1, RABL2A, TRIOBP, ZNF443, and ZNF789) but formed ceRNA relationships with some other mRNAs, respectively (AC007228: ZNF10, ZNF211, ZNF229, ZNF471, ZNF583, ZNF614, ZNF649, ZNF673, ZNF793, and ZNF879; HOXA-AS2: DHR53 and HOXA3).

The major histocompatibility complex (MHC) region was one of the regions with the highest gene density and polymorphism. High-throughput sequencing and other technologies confirmed the role of MHC in disease and showed that MHC was associated with cancer and neurological diseases in addition to infection and autoimmune diseases. MHC is involved in antigen recognition during the immune response and is capable of inducing immune cells to participate in immune response (Trowsdale and Knight, 2013). Studies have also reported that immune cytolytic activity (CYT) is positively correlated with the presence of inhibitory receptors (PDCD1, PDL1, CTLA4, LAG3, TIM3, and IDO1), and the presence of CYT is more responsive to immune checkpoint inhibition, suggesting that it can be used as a key marker for immune checkpoint therapy (Narayanan et al., 2018; Wang et al., 2019). Immune control of tumor lesions requires local antigen recognition, activation and amplification of tumor-specific cytotoxic T lymphocytes (CTL). The activated CTL infiltrate the tumor microenvironment and scan the tumor tissue, where they directly interact with the target cells, inducing tumor cell apoptosis and atrophy (Basu et al., 2016). First, effector T lymphocytes are required to migrate to the tumor foci, a process referred to as immune cell infiltration. Following that, they have to make physical contact with the tumor...
FIGURE 9 | The ceRNA network of common DMlncs of cancers (A) The ceRNA distribution of mRNAs in each cancer type. The color of the dots represents the mean of the correlation coefficient for the IncRNAs that form ceRNA relationships in corresponding cancer. The size of the dots represents the number of IncRNAs that have a ceRNA relationship with the corresponding mRNA in cancer (B) The ceRNA network. Square nodes represent IncRNAs, triangle nodes represent mRNAs, and round nodes represent cancer types. The size of IncRNA nodes is proportional to the number of cancer types sharing it, and nodes of different cancer types are distinguished by different colors.

FIGURE 10 | The characteristics of HOXA-AS2 and AC007228 (A) The genomic location of AC007228 (B) The genomic location of HOXA-AS2 (C) Immune response scores of HOXA-AS2 and AC007228 (D) TMB of HOXA-AS2 and AC007228.
cells and scan their MHC. Finally, by releasing perforin or fas/fasl to bind target cells, CTL activates and induces apoptosis (Weigel et al., 2011). Therefore, we assess the immunological effects of HOXA-AS2 and AC007228 using MHC, CYT, and CTL scores.

To improve the evaluation of the effect of lncRNAs on immunity, the R package “ConsensusClusterPlus” (Wilckerson and Hayes, 2010) was used to cluster samples of each cancer based on their expression profiles of HOXA-AS2 and AC007228. We varied the parameter k from two to six, and then selected the optimal subtypes for subsequent immune score evaluation. The score comparison shown in Figure 10C demonstrates that all three scores were consistent across cancer subtypes, indicating that the identified key lncRNAs may aid in predicting the immunological status of cancer patients and provide a basis for tumor treatment. Subsequently, the Wilcoxon rank-sum test was used to compare the immunity scores of different subtypes, and significant differences were observed in the immunity scores of different subtypes in KIRC, LIHC, LUAD, and PRAD.

Finally, we assessed the tumor mutation burden (TMB) among subtypes. TMB was a novel biological target for which therapeutic impact may be predicted. Previous research has demonstrated that the more somatic mutations a cancer patient possesses, the more likely it is that new antigens are produced. Antigen peptides could be loaded onto the MHC and displayed on the cell surface, aiding in their recognition by T cells (Jiang et al., 2018). Therefore, cancer patients with high TMB levels responded better to immune checkpoint blockade therapy. As shown in Figure 10D, the TMB of KIRC, KIRP, LIHC, LUAD, and UCEC corresponded to the immunity score. Subtypes of key lncRNAs played an important role in the immune effect in KIRC, LIHC, and LUAD.

**DISCUSSION**

LncRNAs have been implicated in the occurrence and development of cancer. This study aimed to identify the specific and common lncRNAs with aberrant methylation in pan-cancer. After searching for DMlncs in a variety of cancers, the pan-cancer results were compared. Subsequently, data on lncRNA expression, lncRNA methylation and mRNA expression was integrated to identify lncRNAs with a negative correlation between methylation and expression changes, and survival analysis was performed to further verify the results. Following that, survival-related lncRNAs were mapped to the ceRNA network, and pan-cancer biomarkers were identified by examining the connection characteristics of the network. Finally, the immune effect of the lncRNAs was verified.

In this study, DMlncs for 23 cancers were acquired, and cancer-specific lncRNAs and cancer-common lncRNAs were identified. The NClnCs were further screened for ten cancers, and the correlation between the lncRNAs and mRNAs, as well as the association with survival were verified. Cancer-specific lncRNAs may be used as diagnostic biomarkers for corresponding cancers. In clinical application, these lncRNAs could apply to make detection kits of corresponding cancers. Common lncRNAs in pan-cancers could be used to understand the mechanism underlying common features in cancers. These lncRNAs can be used for the development of targeting drugs for the remission of general symptoms and treatment of cancer.

This study yielded significant results. Not only did we validate several previously known cancer-associated lncRNAs, but we also identified new cancer-related lncRNAs. AC027601 was identified as a novel KIRC-associated lncRNA, and ACTA2-AS1 was discovered to be carcinogenic in KIRP. Additionally, two lncRNAs, HOXA-AS2, and AC007228 were identified as pan-cancer lncRNAs.

However, there are some limitations to this study. Because the number of DMlncs for SARC, STAD, and THYM was less than ten, the more systematic comparisons for these cancers were impossible. The number of normal samples with methylation data for these three cancers was less than ten. The sample proportion was skewed when differences were calculated, resulting in less statistically significant results. Additionally, the DNA methylation data only included 450k arrays and did not cover the entire genome, which might have contributed to the study's insufficiency outcomes. We only identified the epigenetically dysregulated lncRNAs based on DNA methylation, although N6-methyladenosine (m6A) as the RNA post-transcriptional modification has been shown to influence the function of RNAs as well. We did not evaluate the relationship between m6A and lncRNA due to the lack of data.

In the future, more complete data sets on cancers may be collected to allow for more rigorous comparisons. We may use copy number data to analyze the change in copy number of lncRNA and conduct a more comprehensive investigation of the change and function of lncRNA in cancer. Additionally, other types of omics data may be integrated, and factors affecting lncRNAs and gene expression could be evaluated more comprehensively bringing the research process closer to the way molecules interact in the human body. With the development of sequencing techniques, additional methylation data sets such as HM850K, whole-genome bisulfite sequencing (WGBS), and reduced representation bisulfite sequencing (RRBS), as well as other types of methylation data sets, will be used to analyze the function of DNA methylation for lncRNAs.

In general, this study screened cancer-related lncRNA biomarkers based on their methylation alterations and their competing mRNAs. This study considered multiple omics data more comprehensively and used more stringent screening criteria, which effectively eliminated of data deviation errors. The lncRNA biomarkers identified in this study may aid in the investigation of cancer mechanisms.
DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found here: The datasets ANALYZED for this study can be found in the TCGA (https://portal.gdc.cancer.gov/) and TANRIC (https://www.tanric.org/).

FUNDING

This work was supported by the National Natural Science Foundation of China (Grant No. 62031003), and the High Level Innovation Team Construction Project of Beijing Municipal Universities (Grant No. IDHT20190506).

ACKNOWLEDGMENTS

We are grateful to all patients for their dedicated participation in the current study. We thank the TCGA database for sharing the multi-omics data.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fcell.2022.882698/full#supplementary-material

REFERENCES

Aryee, M. J., Jaffe, A. E., Corrada-Bravo, H., Ladd-Acosta, C., Feinberg, A. P., Hansen, K. D., et al. (2014). Minfi: a Flexible and Comprehensive Bioconductor Package for the Analysis of Infinium DNA Methylation Microarrays. Bioinformatics 30 (10), 1363–1369. doi:10.1093/bioinformatics/btu409

Bao, Z., Yang, Z., Huang, Z., Zhou, Y., Cui, Q., and Dong, D. (2019). LncRNA Disease 2.0: an Updated Database of Long Non-coding RNA-Associated Diseases. Nucleic Acids Res. 47 (D1), D1034–D1037. doi:10.1093/nar/gky905

Bao, S., Zhao, H., Yuan, J., Fan, D., Zhang, Z., Su, J., et al. (2020). Computational Identification of Mutator-Derived LncRNA Signatures of Genome Instability For Improving the Clinical Outcome of Cancers: A Case Study in Breast Cancer. Brief Bioinform 21 (5), 1742–1755. doi:10.1093/bib/bbab118

Bao, G., Xu, R., Wang, X., Ji, J., Wang, L., Li, W., et al. (2021). Identification of LncRNA Signature Associated with Pan-Cancer Prognosis. IEEE J. Biomed. Health Inf. 25 (6), 2317–2328. doi:10.1109/JBHI.2020.3027680

Bartonicek, N., Maag, J. L. V., and Dinger, M. E. (2016). Long Noncoding RNAs in Cancer: Mechanisms of Action and Technological Advancements. Mol. Cancer 15 (1), 43. doi:10.1186/s12943-016-0530-6

Basu, R., Whitlock, B. M., Husson, J., Le Floch, A., Jin, W., Oyler-Yaniv, A., et al. (2016). Cytotoxic T Cells Use Mechanical Force to Potentiate Target Cell Killing. Cell 165 (1), 100–110. doi:10.1016/j.cell.2016.01.021

Cabili, M. N., Trapnell, C., Goff, L., Koziol, M., Tazon-Vega, B., Regev, A., et al. (2011). Integrative Annotation of Human Large Intergenic Noncoding RNAs Reveals Global Properties and Specific Subclasses. Genes Dev. 25 (18), 1915–1927. doi:10.1101/gad.17446611

Cao, Q., Dong, Z., Liu, S., An, G., Yan, B., and Lei, L. (2020). Construction of a Metastasis-Associated eCancer ceRNA Network Reveals a Prognostic Signature in Lung Cancer. Cancer Cell Int. 20, 208. doi:10.1186/s12935-020-01295-8

Chen, Y. G., Satpathy, A. T., and Chang, H. Y. (2017). Gene Regulation in the Immune System by Long Noncoding RNAs. Nat. Immunol. 18 (9), 962–972. doi:10.1038/nri.3771

Chen, J., Yu, Y., Li, H., Hu, Q., Chen, X., He, Y., et al. (2019). Long Non-coding RNA PVT1 Promotes Tumor Progression by Regulating the miR-143/HK2 axis in Gallbladder Cancer. Mol. Cancer 18 (1), 33. doi:10.1186/s12943-019-0947-9

Cheng, Y., Wang, X., Qi, P., Liu, C., Wang, S., Wan, Q., et al. (2021). Tumor Microenvironmental Competitive Endogenous RNA Network and Immune Cells Act as Robust Prognostic Predictor of Acute Myeloid Leukemia. Front. Oncol. 11, 584884. doi:10.3389/fonc.2021.584884

Ding, C., Tang, W., Wu, H., Fan, X., Luo, J., Feng, J., et al. (2019). The PEAK1-Ppp1r12b axis Inhibits Tumor Growth and Metastasis by Regulating Grb2/P13K/Akt Signalling in Colorectal Cancer. Cancer Lett. 442, 383–395. doi:10.1016/j.canlet.2018.11.014

Dong, Y., Xiao, Y., Shi, Q., and Jiang, C. (2019). Dysregulated IncRNA-miRNA-mRNA Network Reveals Patient Survival-Associated Modules and RNA Binding Proteins in Invasive Breast Carcinoma. Front. Genet. 10, 1284. doi:10.3389/fgene.2019.01284

Esposito, R., Bosch, N., Lanzós, A., Polidori, T., Pulido-Quétglas, C., and Johnson, R. (2019). Hacking the Cancer Genome: Profiling Therapeutically Actionable Long Non-coding RNAs Using CRISPR-Cas9 Screening. Cancer Cell 35 (4), 545–557. doi:10.1016/j.ccell.2019.01.019

Gao, L., Lin, P., Chen, P., Gao, R. Z., Yang, H., He, Y., et al. (2020). A Novel Risk Signature that Combines 10 Long Noncoding RNAs to Predict Neuroblastoma Prognosis. J. Cell Physiol. 235 (4), 3823–3834. doi:10.1002/jcp.29277

Gao, M., Guo, Y., Xiao, Y., and Shang, X. (2021). Comprehensive Analyses of Correlation and Survival Reveal Informative IncRNA Prognostic Signatures in Colon Cancer. World J. Surg. Oncol. 19 (1), 104. doi:10.1186/s12957-021-02196-4

Glinsky, G. V. (2006). Genomic Models of Metastatic Cancer: Functional Analysis of Death-From-Cancer Signature Genes Reveals Aneuploid, Anoikis-Resistant, Metastasis-Enabling Phenotype with Altered Cell Cycle Control and Activated P53 Protein Chromatin Silencing Pathway. Cell Cycle 5 (11), 1208–1216. doi:10.4161/cc.5.11.2796

Hou, P., Bao, S., Fan, D., Yan, C., Su, J., Qu, J., et al. (2021). Machine Learning-Based Integrative Analysis of Methyhome and Transcriptome Identifies Novel Prognostic DNA Methylation Signature in Uveal Melanoma. Brief Bioinform. 22 (4). doi:10.1093/bib/bbaa371

Huarte, M. (2015). The Emerging Role of LncRNAs in Cancer. Nat. Med. 21 (11), 1253–1261. doi:10.1038/nm.3981

Jiang, P., Gu, S., Pan, D., Fu, J., Sahu, A., Hu, X., et al. (2018). Signatures of T Cell Dysfunction and Exclusion Predict Cancer Immunotherapy Response. Nat. Med. 24 (10), 1550–1558. doi:10.1038/nm.4591-018-0136-1

Lin, X., Koutelou, E., Schibler, A. C., Chen, Y. C., Grant, P. A., and Dent, S. Y. R. (2016). Cytotoxic T Cells Change the Niagara DNA Methylation Landscape in the Metastatic Microenvironment. Cancer Med. 5 (11), 2942–2954. doi:10.1002/cam4.951

Leberfarb, E. Y., Degtyareva, A. O., BrusentsovII, Maximov, V. N., Voevoda, M. I., et al. (2020). Pan-Cancer IncRNA Epigenetic Biomarkers. Front. Oncol. 10, 15. doi:10.3389/fonc.2020.584884

Zhao et al. Pan-Cancer IncRNA Epigenetic Biomarkers

Funding

This work was supported by the National Natural Science Foundation of China (Grant No. 62031003), and the High Level Innovation Team Construction Project of Beijing Municipal Universities (Grant No. IDHT20190506).

Acknowledgments

We are grateful to all patients for their dedicated participation in the current study. We thank the TCGA database for sharing the multi-omics data.

Supplementary Material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fcell.2022.882698/full#supplementary-material
Properties of Glioma. *Front. Immunol.* 10, 1756. doi:10.3389/fimmu.2019.01756

Wang, M., Zheng, S., Li, X., Ding, Y., Zhang, M., Lin, L., et al. (2020). Integrated Analysis of lncRNA-miRNA-mRNA ceRNA Network Identified lncRNA EPB41L4A-AS1 as a Potential Biomarker in Non-small Cell Lung Cancer. *Front. Genet.* 11, 511676. doi:10.3389/fgene.2020.511676

Wei, Y., Dong, S., Zhu, Y., Zhao, Y., Wu, C., Zhu, Y., et al. (2019). DNA Cytomethylation Analysis of lincRNAs across Nine Cancer Types Reveals Novel Potential Epigenetic Biomarkers in Cancer. *Epigenomics* 11 (10), 1177–1190. doi:10.2217/epi-2018-0138

Weigelin, B., Krause, M., and Friedl, P. (2011). Cytotoxic T Lymphocyte Migration and Effector Function in the Tumor Microenvironment. *Immunol. Lett.* 138 (1), 19–21. doi:10.1016/j.imlet.2011.02.016

Willerson, M. D., and Hayes, D. N. (2010). ConsensusClusterPlus: a Class Discovery Tool with Confidence Assessments and Item Tracking. *Bioinformatics* 26 (12), 1572–1573. doi:10.1093/bioinformatics/btq170

Xu, M., Chen, X., Lin, K., Zeng, K., Liu, X., Xu, X., et al. (2019). LncRNA SNHG6 Regulates EZH2 Expression by Sponging miR-26a/b and miR-214 in Colorectal Cancer. *J. Hematol. Oncol.* 12 (1). 3. doi:10.1186/s13045-018-0690-5

Xu, D., Wang, L., Pang, S., Cao, M., Wang, W., Yu, X., et al. (2021). The Functional Characterization of Epigenetically Related lncRNAs Involved in Dysregulated CeRNA-CeRNA Networks across Eight Cancer Types. *Front. Cell Dev. Biol.* 9, 649755. doi:10.3389/fcell.2021.649755

Xu, J., Wang, Z., Li, S., Chen, J., Zhang, J., Jiang, C., et al. (2018). Combinatorial Epigenetic Regulation of Non-coding RNAs has Profound Effects on Oncogenic Pathways in Breast Cancer Subtypes. *Brief Bioinform.* 19 (1), 52–64. doi:10.1093/bib/bbx099

Yang, Z., Xu, F., Wang, H., Teschendorff, A. E., Xie, F., and He, Y. (2021). Pan-cancer Characterization of Long Non-coding RNA and DNA Methylation Mediated Transcriptional Dysregulation. *Elife* 68, 103399. doi:10.1016/j.eldis.2021.103399

Ying, K., Wang, L., Long, G., Lian, C., Chen, Z., and Lin, W. (2020). ACTA2-AS1 Suppresses Lung Adenocarcinoma Progression via Sequestering miR-378a-3p and miR-4428 to Elevate SOX7 Expression. *Cell Biol. Int.* 44 (12), 2438–2449. doi:10.1002/cbin.11451

Li, Y., Jiang, T., Zhou, W., Li, J., Li, X., Wang, Q., et al. (2020). Pan-Cancer Characterization of Immune-Related lncRNAs Identifies Potential Oncogenic Biomarkers. *Nat. Commun.* 11 (1), 1000. doi:10.1038/s41467-020-14802-2

Zhang, G., Li, S., Lu, J., Ge, Y., Wang, Q., Ma, G., et al. (2018). LncRNA MT1J Functions as a ceRNA in Regulating FBXW7 through Competitively Binding to miR-92a-3p in Gastric Cancer. *Mol. Cancer* 17 (1), 87. doi:10.1186/s12943-018-0829-6

Zhang, C., Ren, X., Zhang, W., He, L., Qi, L., Chen, R., et al. (2020). Prognostic and Clinical Significance of Long Non-coding RNA SNHG12 Expression in Various Cancers. *Bioengineered* 11 (1), 1112–1123. doi:10.1080/21655979.2020.1831361

Zheng, H., Bian, C., Tu, S., Yin, F., Guo, P., Zhang, J., et al. (2021a). Integrated Analysis of lncRNA-miRNA-mRNA ceRNA Network in Human Aortic Dissection. *BMC Genomics* 22 (1), 724. doi:10.1186/s12864-021-08012-3

Zhang, Z., Yan, C., Li, K., Bao, S., Li, L., Chen, L., et al. (2021b). Pan-cancer Characterization of lncRNA Modifiers of Immune Microenvironment Reveals Clinically Distinct Depo Tumor Subtypes. *npj Genom. Med.* 6 (1), 52. doi:10.1038/s41525-021-00215-7

Zhao, H., Zhang, X., Frazão, J. R., Condino-Neto, A., and Newburger, P. E. (2013). HOX Antisense lincRNA HOXA-AS2 Is an Apoptosis Repressor in allTransretinoic Acid Treated NB4 Promyelocytic Leukemia Cells. *J. Cell. Biochem.* 114 (10), 2375–2383. doi:10.1002/jcb.24586

Zheng, R., Zheng, M., Wang, M., Lu, F., and Hu, M. (2021). Identification of a Prognostic Long Noncoding RNA Signature in Lung Squamous Cell Carcinoma: a Population-Based Study with a Mean Follow-Up of 3.5 Years. *Arch. Public Health* 79 (1), 61. doi:10.1186/s13390-021-00588-2

Zhou, R. J., and Lv, H. Z. (2019). Knockdown of ACTA2AS1 Promotes Liver Cancer Cell Proliferation, Migration and Invasion. *Mol. Med. Rep.* 19 (3), 2263–2270. doi:10.2217/mmr.19.9836

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

**Publisher’s Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

**Copyright © 2022 Zhao, Guo, Zhang, Wang and Wang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.**
GLOSSARY

BLCA Bladder Urothelial Carcinoma
BRCA Breast Invasive Carcinoma
CESC Cervical Squamous Cell Carcinoma and Endocervical Adenocarcinoma
CHOL Cholangiocarcinoma
COAD Colon Adenocarcinoma
ESCA Esophageal Carcinoma
GBM Glioblastoma Multi-Forme
HNSC Head and Neck Squamous Cell Carcinoma
KIRC Kidney Renal Clear Cell Carcinoma
KIRP Kidney Renal Papillary Cell Carcinoma
LIHC Liver Hepatocellular Carcinoma
LUAD Lung Adenocarcinoma
LUSC Lung Squamous Cell Carcinoma
PAAD Pancreatic Adenocarcinoma
PCPG Pheochromocytoma and Paraganglioma
PRAD Prostate Adenocarcinoma
READ Rectum Adenocarcinoma
SARC Sarcoma
SKCM Skin Cutaneous Melanoma
STAD Stomach Adenocarcinoma
THCA Thyroid Carcinoma

THYM Thymoma
UCEC Uterine Corpus Endometrial Carcinoma
IncRNA Long non-coding RNAs
DMInc differentially methylated IncRNA
NCInc negatively correlated IncRNA
UMLEInc up-methylated-low-expressed IncRNA
DMOEInc down-methylated-over-expressed IncRNA
CSNClncs cancer-specific negatively correlated IncRNAs
MHC major histocompatibility complex
CYT immunolytic cell activity
CTL cytotoxic T cell
TMB tumor mutation burden
IncRNA Long non-coding RNA
DMInc Differentially Methylated IncRNA
NCInc Negatively Correlated IncRNA
UMLEInc Up-Methylated-Low-Expressed IncRNA
DMOEInc Down-Methylated-Over-Expressed IncRNA
CSNClnc Cancer-Specific Negatively Correlated IncRNA
MHC Major Histocompatibility Complex
CYT Cytolytic Activity
CTL Cytotoxic T Lymphocyte
TMB Tumor Mutation Burden