Computational and Structural Biotechnology Journal 21 (2023) 5309–5325

Contents lists available at ScienceDirect

Computational and Structural Biotechnology Journal

journal homepage: www.elsevier.com/locate/csbj

Research article

Pan-cancer landscape of immunology PIWI-interacting RNAs

Dongyi Wan [a,1], Ran Li [b,1], Haohao Huang [c,*], Xiaohua Zhu [a,*], Ganxun Li [d,*]

[a] Department of Nuclear Medicine, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, Hubei, China
[b] Department of Neurosurgery, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, Hubei, China
[c] Department of Neurosurgery, General Hospital of Central Theater Command of Chinese People's Liberation Army, Wuhan 430070, China
[d] Hepatic Surgery Center and Hubei Key Laboratory of Hepato-Biliary-Pancreatic Diseases, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, Hubei, China

ARTICLE INFO

Keywords:
PIWI-interacting RNAs
Tumor microenvironment
ImmPI
Pan-cancer

ABSTRACT

PIWI-interacting RNAs (piRNAs), an emergent type of non-coding RNAs during oncogenesis, play critical roles in regulating tumor microenvironment. Systematic analysis of piRNAs' roles in modulating immune pathways is important for tumor immunotherapy. In this study, in-depth analysis of piRNAs was performed to develop an integrated computational algorithm, the immunology piRNA (ImmPI) pipeline, for uncovering the global expression landscape of piRNAs and identifying their regulatory roles in immune pathways. The immunology piRNAs show a tendency towards overexpression patterns in immune cells, causing perturbations in tumors, being significantly associated with infiltration of immune cells, and having prognostic value. The ImmPI score can contribute to prioritizing tumor-related piRNAs and distinguish two subtypes of SKCM (immune-cold and hot phenotypes), as characterized by different prognoses, immunogenicity and antitumor immunity. Finally, we developed an interactive web resource (ImmPI portal: http://www.hbpding.com/ImmPi) for the biomedical research community, with several useful modules to browse, visualize, and download the results of immunology piRNAs analysis. Overall, our work provides a comprehensive landscape of piRNAs across multiple cancer types and sheds light on their regulatory and functional roles in tumor immunity. These findings pave the way for future research and development of piRNA-based immunotherapies for cancer treatment.

1. Introduction

Aberrant gene expression has been demonstrated to result in a multifarious series of diseases in humans, including tumors [1]. Thousands of regulators control gene expression in mammals, including transcription factors (TFs), chromatin modifiers, and non-coding RNAs (ncRNAs) [2]. piRNAs, a type of small ncRNA with 24–31 nucleotides, primarily are found in the germline of mammals and play crucial roles, including inhibiting transposable element activity through binding to PIWI-related molecules [3–5]. Moreover, recent studies have demonstrated that piRNAs can be expressed and function in somatic tissues of humans [6], and exhibit dysregulated expression in several specific cancer types [7–10]. However, knowledge about their regulatory roles in tumorigenesis is still largely unavailable.

Various studies have concentrated on the biological roles of piRNAs and their underlying mechanisms are reported to be highly diversified [11]. Previous studies have principally reported that piRNAs correlate with distinct cellular responses, including cell proliferation, differentiation, and apoptosis [12]. It has also been shown that piRNAs play important roles in the initiation and development of various cancers, including urogenital carcinoma [13], and diffuse large B-cell lymphoma [14]. Moreover, piRNAs can act as potential diagnostic indicators for early detection of lung [15], colorectal [16], and gastric carcinomas [17]. piRNAs are also crucial prognosis-related biomarkers for colorectal and kidney carcinoma after anti-tumor therapy [8,13]. Recent studies have demonstrated that piRNAs impact the survival activity of cancer cells under exposure of chemotherapies and might function as potential therapeutic targets for treating human malignancy [18]. On the other hand, accumulating research has demonstrated that the tumor microenvironment (TIME) plays crucial functions in oncogenesis [19]. A dysregulated anti-tumor immune system might be one of the foremost causes of tumor occurrence, and immunotherapy has emerged as a useful antitumor therapy [20]. Therefore, precise regulation of immune pathways is crucial for a powerful anti-tumor immunity

* Corresponding authors.
E-mail addresses: huanghaohao@hust.edu.cn (H. Huang), evazhu@vip.sina.com (X. Zhu), ganxunli@tjh.tjmu.edu.cn (G. Li).
[1] The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint First Authors

https://doi.org/10.1016/j.csbj.2023.10.042
Received 7 June 2023; Received in revised form 19 October 2023; Accepted 19 October 2023
Available online 21 October 2023

microenvironment [21]. Nevertheless, the main focus of recent works is on protein molecules, particularly the role of cell-surface ligands, receptors, cytokines, and TFs [22]. In addition, the key roles of piRNAs investigated in previous studies were confined to their impacts on tumor phenotypes [23], and whether they may be involved in regulating tumor immunity remains unclear. Meanwhile, the function of piRNAs in regulating the efficiency of tumor immune checkpoint inhibitors has not even been evaluated. Consequently, further in-depth investigations on piRNAs and their impacts in regulating tumor immunity will contribute to development of effective immunotherapies.

To detect the underlying immunology piRNAs (namely, immune-related piRNAs) involved in oncogenesis, multi-omics profiles from 32 cancer types were integrated to characterize the comprehensive transcriptional landscape, and a novel ImmPI score pipeline was built. Our work identified multiple piRNAs significantly associated with the activity of immune pathways, and its robustness was validated in other independent cohorts. These immunology piRNAs tend to present overexpression patterns within immune cells, exhibit perturbed expression in tumor samples, be significantly associated with infiltration of immune cells and show prognostic value. With tumor-related piRNA prioritization and tumor clustering, our work indicates that the ImmPI score pipeline and ImmPI web portal are beneficial resources for exploring the roles of piRNAs in regulating tumor immunity. In-depth recognition of the thorough repertoire of underlying piRNAs regulating immune pathways is a crucial requirement for comprehending the architecture of the integrated regulatory network within the tumor immune microenvironment (TIME).

2. Results

2.1. Dynamic expression landscape of piRNAs in human cancers

To comprehensively depict the landscape of piRNAs across cancers, we included genome-wide small RNA-seq results of 9788 tumor and 1289 paracancerous samples across 32 cancer types (Fig. 1A, Table S2). We mapped TCGA raw small RNA-seq reads to piRNA referential transcriptomics based on the piRBase database [46] and retained piRNAs with TPM more than 1 and available expression in more than 70% of samples for subsequent analysis. Our analysis recognized a total of 12952 detectable piRNAs across human cancers (Fig. 1B and Fig. S2A). The count of identifiable piRNAs varied from 2539 in sarcomas (SARC) to 7897 in Testicular Germ Cell Cancer (TGCT). These detectable piRNAs were classified into three subgroups: 3525 ubiquitous piRNAs expressed in more than 20 cancers, 6063 intermediately specific piRNAs expressed in 2–20 cancers, and 3364 cancer-specific piRNAs expressed in only one cancer (Fig. 1B). Ubiquitous piRNAs accounted for 43.9% of detectable piRNAs in OV, but for 96.9% of piRNAs in SARC. More interestingly, ubiquitous piRNAs presented higher expression patterns than intermediately specific piRNAs and cancer-specific piRNAs (Fig. S2C). This situation mirrors the traits of protein-coding genes, where housekeeping genes typically exhibit higher expression levels compared to their tissue-specific counterparts [24]. The numbers of cancer-specific piRNAs exhibited an expansive range, from zero in six cancers, including BRCA, CESC, CHOL, COAD, SARC, and STAD, to 873 in OV (Fig. 1B). This cancer-specific expression pattern of piRNAs was still observed with an even more stringent cut-off (TPM $\geq$ 10, Fig. S2D), demonstrating that it is not due to their expression level. The transcriptional similarity between any two tumors was calculated, and the results indicated that a robust cancer-specific pattern of piRNAs and patients from the same tumor type clustered together on the basis of k-means (Fig. 1C). In addition, we utilized another dimensionality reduction algorithm, t-SNE, to validate this cancer-specific phenomenon obtained from k-means method. After dimensionality reduction using t-SNE, the LGG in the bottom left is distinctly separated in terms of piRNA expression from PRAD directly above and OV to the bottom right (Fig. S2B). The expression of piRNAs in cancers with analogous histology were more similar, whereas the expression patterns between piRNAs in different cancers were relatively distinct, such as pan-kidney tumors (KIRC, KIRP and KICH) [25], indicating that piRNAs might act as potential biomarkers with clinical utility in identifying specific tumor types.

2.2. Identification of immunology piRNAs across cancer types

In order to detect potential piRNAs associated with immune pathways, a computational pipeline named the ImmPI score was developed (Fig. 2A). This pipeline is able to comprehensively infer potential piRNAs regulating immune pathways on the basis of a large quantity of sample-matched protein coding genes and piRNA transcriptomics. It is reasoned that if piRNAs play crucial roles in immune biology, their associated protein coding genes should be enriched in relevant immune pathways. In brief, our pipeline identified potential immunology piRNAs in three procedures (Fig. 2A). Firsty, we collected genome-wide protein coding gene and piRNA transcriptomics of the same sample. Secondly, the tumor purity for each sample was estimated, and protein coding genes were ranked on the basis of the rank score (RS) for each candidate piRNA. Thirdly, the estimated level of piRNA-related immune pathways (ImmPI score) was calculated on the basis of modified GSEA [26]. We converted the *P*-value of GSEA into an ImmPI score, and piRNA–pathway pairs with absolute ImmPI scores greater than 0.995 and FDR less than 0.05 were identified as significant for subsequent analysis.

Taking advantage of the multi-omics in TCGA database, the ImmPI score pipeline was applied to detect immunology piRNA in more than 10,400 samples across 32 cancers (Fig. 1A). In particular, our analysis mainly concentrated on immunology gene sets representing diverse immune pathways originating from ImmPort [27], one of the largest public repositories of immunological resources. The number of genes involved in the aforementioned pathways ranged from three to 516 (Fig. 2B). When the ImmPI score pipeline was employed for 32 cancer types, genome-wide screening generated a series of piRNA-related pathways. On average, each cancer type generated ~1840 piRNAs associated with immune pathways (Fig. 2C). Higher amounts of immunology piRNAs were recognized in cancer types with more piRNAs available (Fig. S3). These immunology piRNAs accounted for ~39.7% of all piRNAs (Fig. 2C). In particular, “antigen processing and presentation”, “TCR signaling”, “cytokine receptors” and “cytokines” pathways, emerging as promising targets for immunotherapy [28], were found to be regulated by more piRNAs across cancers (Fig. 2D and Fig. S4), and these immunology piRNAs will be a valuable resource for further understanding potential regulation within the tumor microenvironment. All of the above results indicate that the ImmPI score pipeline can identify important piRNAs involved in immune biological processes and provide an invaluable resource for precision medicine.

2.3. Expression perturbation of immunology piRNAs across cancers

To identify the biological function of immunology piRNAs, we evaluated the 360 top-ranked piRNA–pathway pairs identified across cancer types (Table S3), and further constructed a piRNA–pathway regulatory network. As a result, there were 212 piRNAs, which were available in all 17 cancer types, and 8 immune pathways involved in this network (Fig. 3A), and the majority of these piRNAs were significantly associated with the “Antigen Processing and Presentation” pathway (Fig. S5A). Furthermore, the expression pattern of these piRNAs was investigated and the results demonstrated most of these piRNAs to be upregulated in tumor (Fig. 3A). For instance, five specific piRNAs showed more than two-fold upregulation in 15 cancer types (Fig. S5B). In order to further evaluate which subtype of the “antigen processing and presentation” pathway correlated with these five piRNAs, we calculated the activity of its multiple subtypes by GSVA analysis, and the results demonstrated five piRNAs to be significantly associated with several subtypes, including T and dendritic cells (Fig. S5C). These results

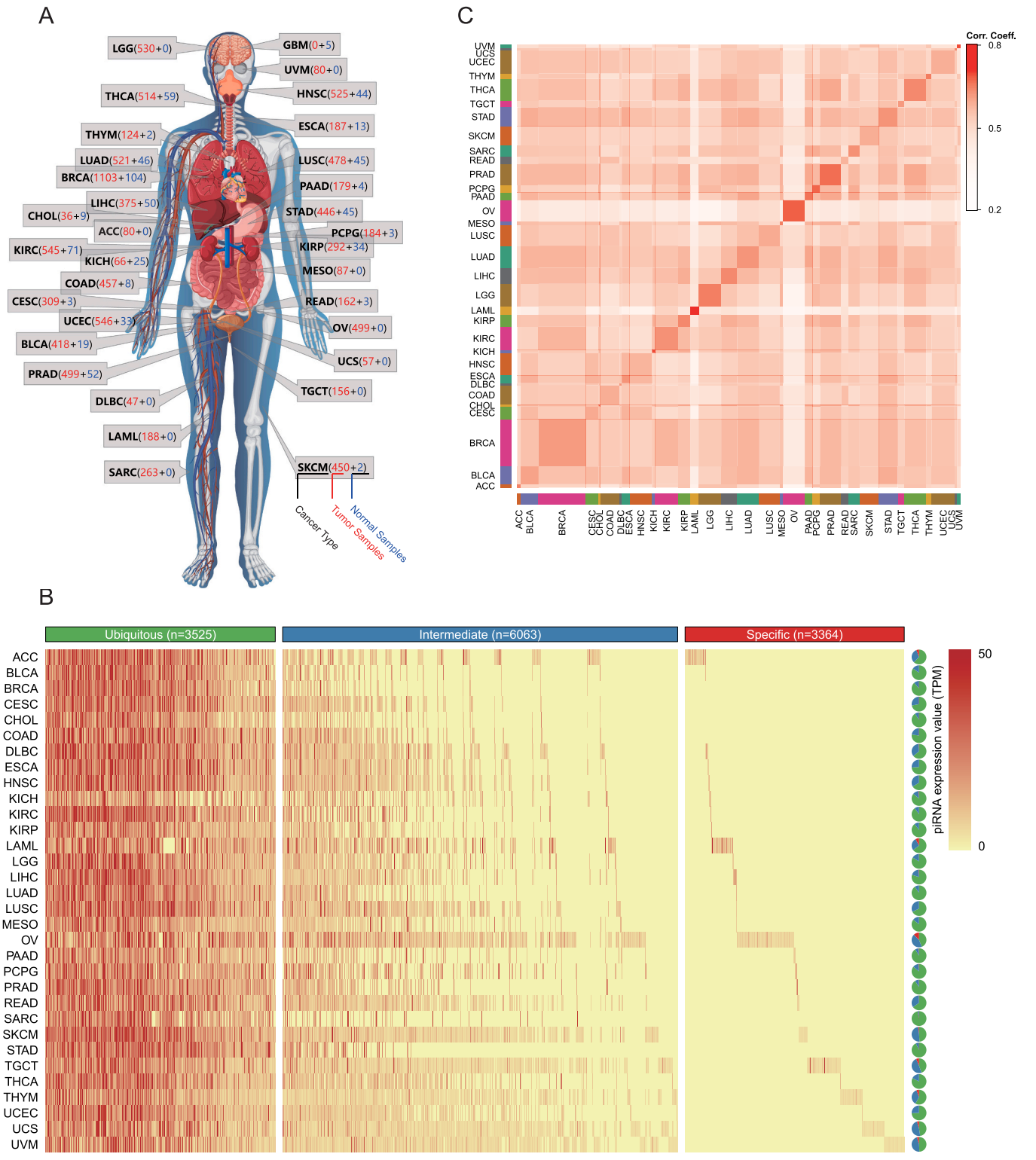


Fig. 1. Expression landscape of piRNAs across cancer types. (A) Distribution of samples across 32 different cancer types, with red and blue numbers indicating the count of cancer and normal samples, respectively. (B) piRNA expression patterns in various human cancers. The green, blue, and red bars represent ubiquitous, intermediately specific, and cancer-type-specific piRNAs, respectively. The pie charts show the proportion of each piRNA category within each cancer type. Refer to Table S2 for abbreviations of the cancer types. The scale bar indicates the expression levels of piRNA in terms of TPM. (C) The heatmap illustrates the similarity in piRNA expression across all tumor samples. Each cell reflects the average correlation coefficient of piRNA expression between samples from the corresponding pair of cancer types. The color bar represents each type of cancer. The scale bar signifies the degree of piRNA expression similarity among TCGA tumor samples, with the intensity of the red hue indicating the correlation coefficient level.

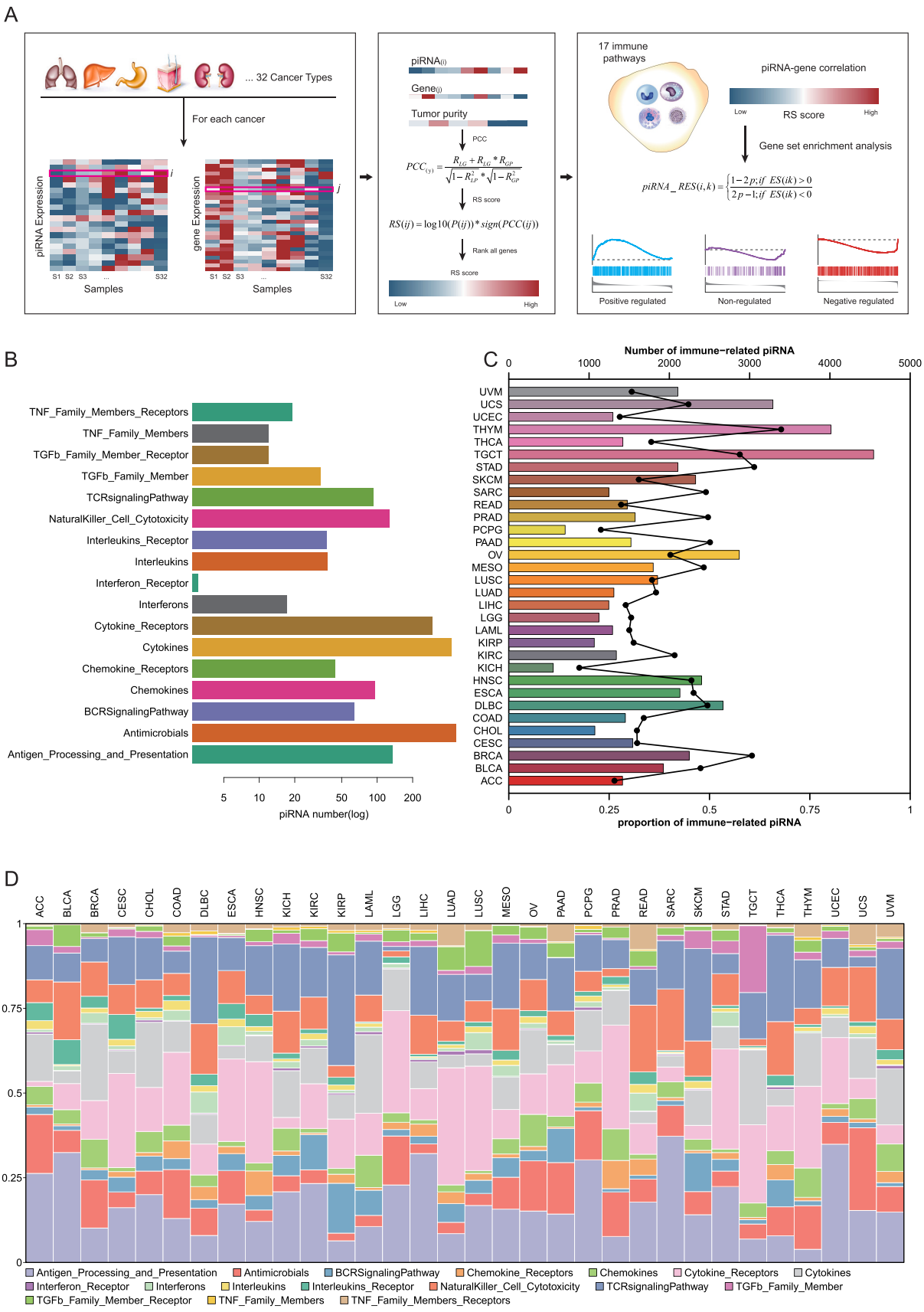
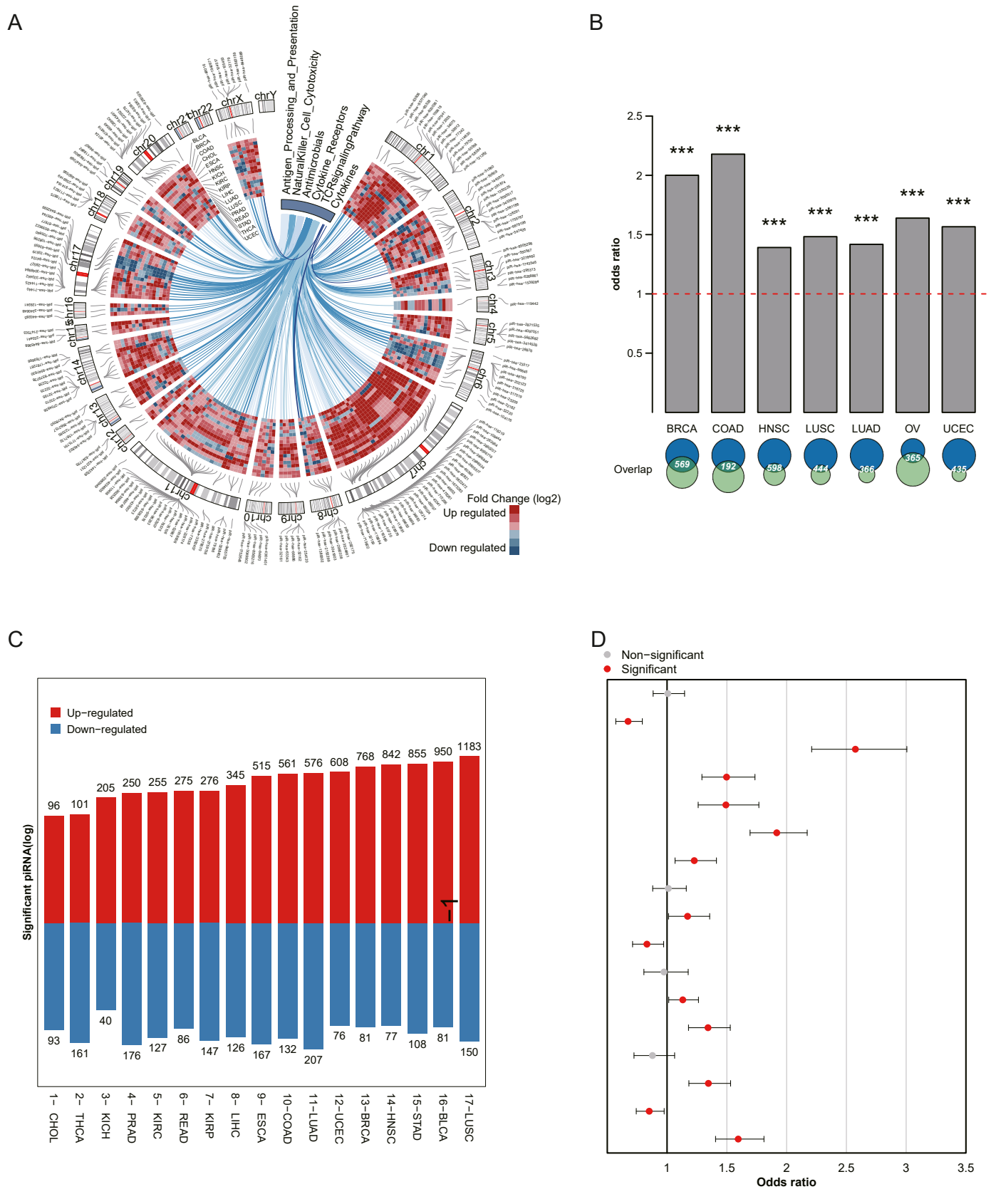


Fig. 2. Identification of immunology piRNAs across cancer types. (A) Schematic illustration of three steps of ImmPI for identifying piRNA regulators. (B) The count of genes present in 17 immune-associated pathways. (C) The tally of immunology piRNAs detected for each cancer type. The upper y-axis indicates the number of piRNAs, while the lower y-axis represents the percentage of immunology piRNAs. (D) Overview of piRNAs identified by extending ImmPI across cancer types. The panel shows the proportion of piRNAs enriched in each immune pathway across cancers.



(caption on next page)

Fig. 3. Validation of immunology piRNAs. (A) The Circos diagram displays the top 360 piRNA–pathway associations across various cancer types. Inside the chart, a heatmap depicts the differential expression of piRNAs among cancers. Shades of red and blue indicate up-regulated and down-regulated expression, respectively, with color intensity reflecting the extent of expression variations. (B) The congruence of immune–piRNA pathways in the separate CPTAC database for identical cancer types is displayed. The bar graphs at the top illustrate the observed/expected ratios from the hypergeometric test, with * ** denoting a p-value below 0.001. The Venn diagrams below detail the count of shared piRNA–pathway associations: the top blue section signifies numbers derived from TCGA, while the bottom green portion indicates numbers sourced from the CPTAC database. (C) The bar chart illustrates the count of significantly differentially expressed immunology piRNAs across various cancers in the TCGA database. Red and blue hues denote up-regulated and down-regulated significantly differentially expressed immunology piRNAs, respectively. The Y-axis displays the log-transformed counts for each piRNA. (D) A stratified Chi-square analysis was employed to compare the proportion of differentially expressed immunology piRNAs to their non-immunology counterparts. The forest plot, referencing Fig. 3C, illustrates odds ratios for each cancer type. Odds ratios depict the likelihood of an event's occurrence relative to its non-occurrence. Red signifies significance, while gray denotes non-significance. Error bars represent the 95% confidence intervals for the odds ratios.

imply that these piRNAs might play crucial roles in antigen processing and presentation-related biological processes.

To evaluate the robustness of the results obtained from the ImmPI score pipeline, we then investigated the reproducibility of the piRNA–pathway pairs in different piRNA datasets for the same tumor. We downloaded another seven sample-matched mRNA and piRNA expression profiles (Fig. S6, and Table S4) from Clinical Proteomic Tumor Analysis Consortium (CPTAC) [29]. The piRNA–pathway pairs were again recognized by the ImmPI score pipeline in each tumor type separately, and significant piRNA-pathway overlap was found for the same tumor type (Fig. 3B). These results demonstrate that our pipeline is efficient in detecting immunology piRNAs.

Dysregulation of immune pathways has been identified as correlating with the initiation and development of multiple tumors. Next, we evaluated whether perturbations of immunology piRNAs can be observed in tumor. The results suggested hundreds of immunology piRNAs to be dysregulated across multiple cancers (Fig. 3C and Fig. S7, $|\log2FC| > 1$ and FDR <0.05). Although some immunology piRNAs might present an opposite trend of expression pattern in different tumor types, a number of immunology piRNAs presented concordant up- or down-regulated expression in more than two tumor types (Fig. S8). By comparing the proportion of differentially expressed immunology piRNAs to their non-immunology counterparts, we found that they were more likely to exhibit expression perturbations (Fig. 3D), especially in some cancer types suitable for immunotherapy, including gastrointestinal [30] and lung cancers [31]. For instance, 1183 and 150 immunology piRNAs presented higher or lower expression patterns in LUSC (Fig. 3C), which was approximately twice as high as the expected proportion. Collectively, the immunology piRNAs identified by the ImmPI score pipeline shows widespread expression perturbations across multiple cancer types, differentially expressed piRNAs are more likely to be immunology-related piRNAs.

2.4. piRNA are associated with immune cell infiltration

Infiltration of immune cells within the tumor immune microenvironment (TIME) plays critical roles in tumor immunity and response efficiency to immunotherapy [32]. Consequently, we hypothesized that piRNAs identified by the ImmPI score pipeline play crucial roles in regulating the tumor immune microenvironment. Therefore, immunology piRNAs will present significant differences between the TIME and paracancerous tissues and correlate with infiltration of immune cells. We thus evaluated infiltration of nine primary immune cells (namely B cells, CD4/8 T cells, Tregs, T helper cells, macrophages, neutrophils, NK cells, and dendritic cells) in each patient on the basis of gene expression by several computational algorithms, including CIBERSORT [33], EPIC [34], MCP-counter [35], quanTIseq [36], TIMER [37], xCell[38], and ImmuCellAI [39]. Relationships between expression of piRNAs and infiltration of immune cells were calculated by Spearman's rank correlation analysis ($|\rho| > 0.2$ and $P < 0.05$), suggesting that a great number of piRNAs associated with infiltration of immune cells are also immunologically associated (Fig. 4A and Table S5). For instance, of the piRNAs with expression associated with immunology in BLCA, 55.61% correlated with T helper cell infiltration. The proportion of piRNAs associated with infiltration of CD4 T cells was much higher in multiple gastrointestinal cancers (Figs. 4A, 55.70% for STAD, 50.34% for READ, 60.91% for CHOL, 61.26% for ESCA, and 57.16% for PAAD). These results indicate that immunology piRNAs identified by the ImmPI score pipeline might play crucial roles in regulation of tumor immunity. Additionally, Fisher's exact test was used to evaluate whether these immunology piRNAs are more likely to correlate with infiltration of immune cells compared to their non-immunology counterparts. The results suggested that a significantly higher proportion of immunology piRNAs correlate with infiltration of immune cells in the majority of cancer types (Fig. 4B, and Table S6).

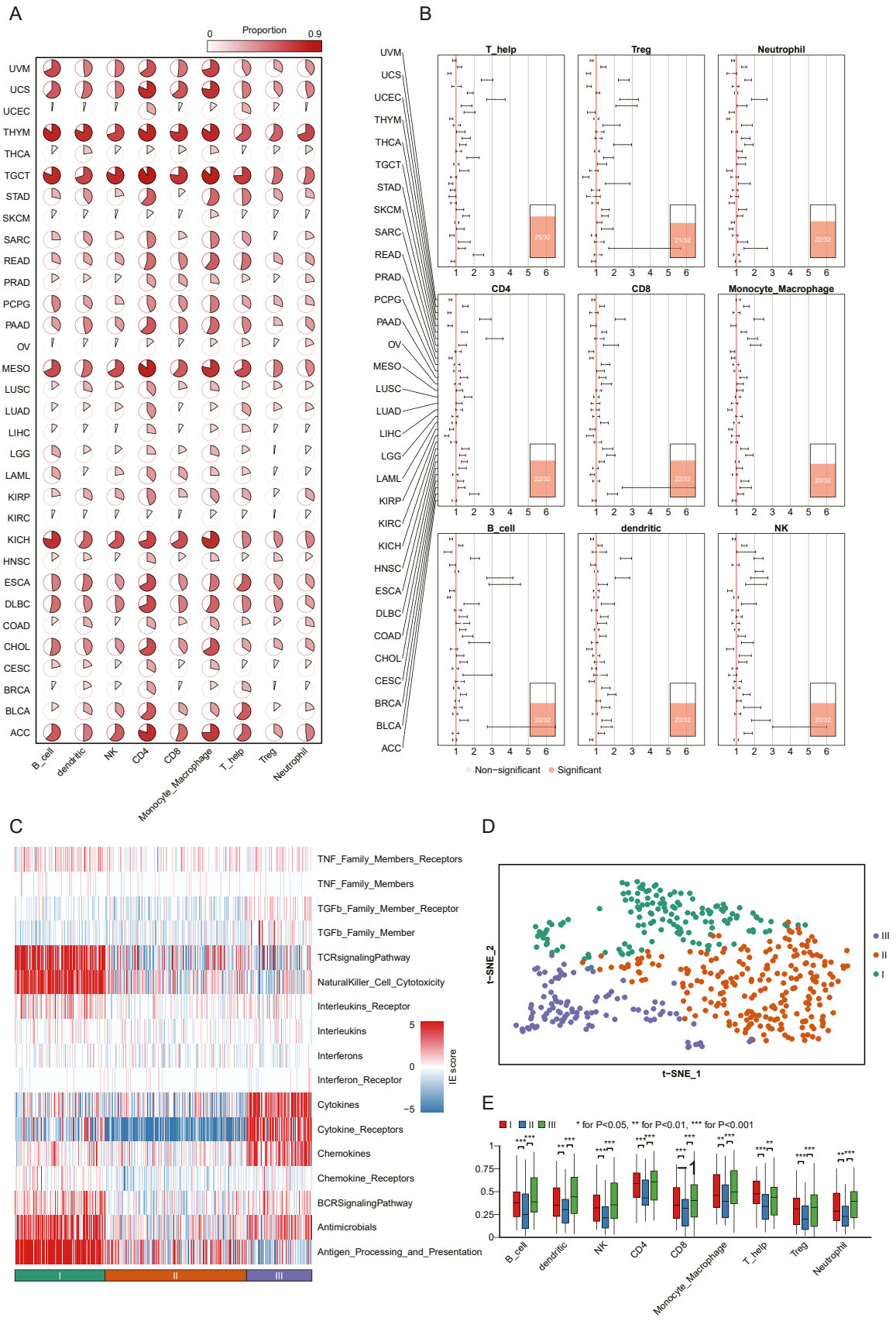
A growing number of analyses have indicated that biomarkers generally present tissue-specific expression patterns. Consequently, expression of the immunology piRNAs generated from the public small RNA-seq profiles of several immune cells was analyzed [40]. The number of specific immunology piRNAs identified in corresponding immune cells from human peripheral blood was significantly higher than that in their non-immunology counterparts (GSE100467, Fig. S9A). In addition, the immunology piRNAs associated with infiltration of one immune cell were more likely to present significantly higher expression in corresponding immune cells from human peripheral blood, including B, CD4/8, NK, monocyte, and neutrophil cells (Fig. S9B).

To further evaluate the potential subtype of immunology piRNAs and corresponding biological roles across cancer types, the Immune Enrichment (IE) score of each piRNA identified by the ImmPI score pipeline was calculated, which was able to estimate the correlation between each piRNA and the 17 immune-related pathways from a pan-cancer perspective. As a result, 481 piRNAs were defined as tumor-immunity-related piRNAs on the basis of specific conditions. We classified these 481 piRNAs into 3 clusters, which included 146 piRNAs in Cluster C1, 229 piRNAs in Cluster C2, and 106 piRNAs in Cluster C3, using the consensus clustering method on the basis of the IE scores of each piRNA (Fig. 4C), and the clustering result was further validated by t-SNE analysis (Fig. 4D). Immune pathways, including antigen processing and presentation, antimicrobials, NK cell cytotoxicity, and the TCR signaling pathway, were enriched in Cluster C1, and other immune pathways, including chemokines, cytokines, and cytokine receptors, were enriched in Cluster C3 (Fig. 4C).

Next, the association between immune cell infiltration and each piRNA across the three clusters was calculated, and the results suggested that the piRNAs in Cluster C2 are significantly least associated with infiltration of all nine immune cells across cancer types (Fig. 4E, and Fig. S10), which further validated the rare immune pathways enriched in Cluster C2 compared to others (Fig. 4C). These results indicate that the subtypes of immunology piRNAs are robust and present contrasting biological roles across cancer types.

2.5. Tumor-related piRNAs, as evidenced by ranking analysis

With the knowledge that immunology piRNAs are associated with infiltration of immune cells, we further evaluated the potential effects of immunology piRNAs on the clinical characteristics and tumor immunity. We hypothesized that the correlations identified in various cancer types between the immunology piRNAs and immune-related pathways play



(caption on next page)

Fig. 4. The immunology piRNAs correlate with immune cell infiltration in cancers. (A) The proportion of immunology piRNAs correlated with immune cell infiltration across cancers. The deeper the red, the larger the portion of the pie chart, indicating a greater relevant proportion. (B) The immunology piRNAs were likely to be enriched in piRNAs correlated with the infiltration of nine main immune cells (namely, B cells, CD4 T cells, CD8 T cells, Tregs, T helper cells, macrophages, neutrophils, NK cells, and dendritic cells). The dots represent the odds ratio (OR) of Fisher's exact test and the error bars show the 95% confidence intervals of the OR. The numbers inside bar plots show the proportion of cancers, where orange indicates cancers with *P*-values < 0.05 in two-sided Fisher's exact test. (C) Heatmap showing the mean IE (Immune Enrichment) score of 481 tumor-immunity-related piRNAs in 3 clusters, which included 146 piRNAs in cluster C1, 229 piRNAs in cluster C2, and 106 piRNAs in cluster C3, using the consensus clustering method on the basis of IE scores of each piRNAs. (D) The t-SNE analysis used to validate the robustness of the clustering analysis, these three piRNA clusters could be distinctly differentiated. (E) Three piRNA clusters showed the significantly different correlation with the infiltration of all nine immune cells across cancer types. As consistent with the enrichment analysis results, and Cluster II group had the lowest proportion related to immune cells. * for $P < 0.05$, * * for $P < 0.01$, * ** for $P < 0.001$.

crucial roles in regulating cancer development. Therefore, we ranked all ImmpiRNAs on the basis of the correlation between piRNA alterations and immune-related pathways across cancer types among 17 immune-related pathways, and the expression difference of ImmpiRNAs between precancerous and tumor tissues (Table S3 and Table S7). Then, we conducted GSEA analysis on the basis of the ranking of associations of ImmpiRNAs and gene sets originating from immune pathways. We found that the top ranking ImmpiRNAs are significantly associated with most immune pathways in multiple cancer types (Fig. 5A). We analyzed the network diagram of the top immunology piRNAs and discovered their intricate relationships. Various piRNAs played important roles in multiple types of cancer by regulating immune pathways. For instance, expression of piR-hsa-8465703 and piR-hsa-312787 presented a significant association with immune-related pathways in multiple tumor types (Fig. 5B, Table S3). The expression level of piR-hsa-8465703 correlated positively with most immune pathways in multiple tumor types, such as antigen processing and presentation in 14 cancer types, and with NK cell cytotoxicity in 10 cancer types (Fig. 5B, Table S3). In essence, these findings suggest that incorporating the ImmpiRNA outcomes can aid in highlighting cancer-associated piRNAs and offer a pathway-centric perspective to deepen our insights into their regulatory influence on cancer development.

2.6. Correlation of immunology piRNAs with prognosis

To evaluate the prognostic relevance of the potential immunology piRNAs, we detected all piRNAs that correlated with patient prognosis, including overall survival (OS), disease-free interval (DFS), disease-specific survival (DSS), and progression-free interval (PFI), across cancer types with $p < 0.05$. Hundreds of risk and protective piRNAs were identified in each cancer type (Fig. 6A and Table S8). Then, the proportion of risk or protective immunology piRNAs was calculated in each cancer type, with the immunology piRNAs exhibited significantly higher numbers of prognostic piRNAs than their non-immunology counterparts across multiple cancer types (30/32 for OS, 19/21 for DFI, 29/32 for DSS, 29/32 for PFI).

Next, the proportion of each piRNA correlating with prognosis was also calculated in 32 cancer types, the majority of prognosis-related immunology piRNAs were identified in more than one cancer type (Fig. 6B). For instance, a number of immunology piRNAs correlated with worse prognosis in more than 10 cancer types (piR-hsa-5775896 for OS, piR-hsa-33063 for PFI and piR-hsa-5775896 for DSS, Fig. 6C-E). Together, our analysis demonstrates that immunology piRNAs are more likely to be associated with patient prognosis and rendered promising therapeutic strategies for tumor.

2.7. Cancer subtyping on the basis of immunology piRNAs

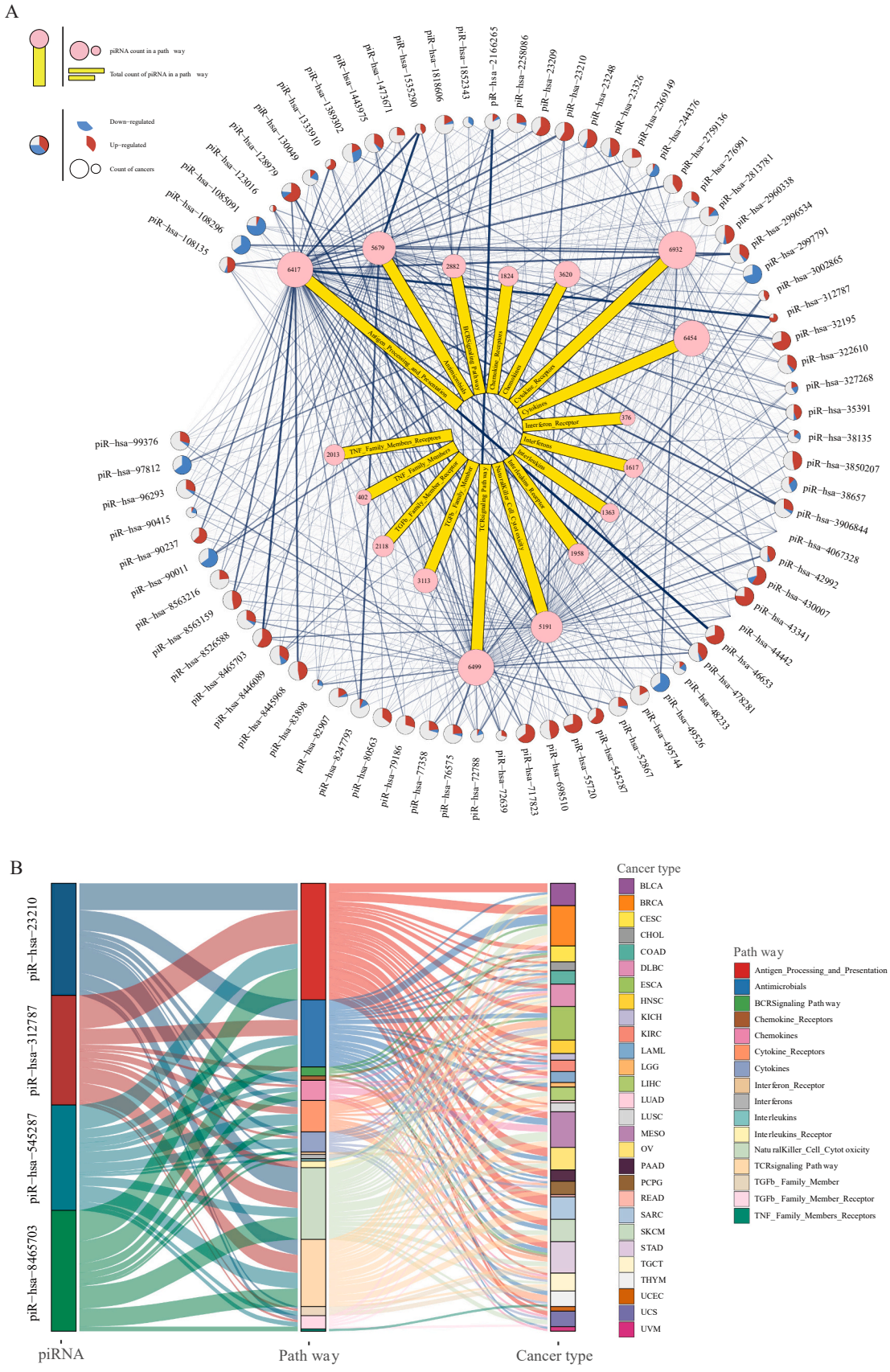
In addition to identifying crucial biomarkers in tumor, cancer subtyping is critical for personalized therapies. Consequently, we evaluated the extent to which the piRNAs detected by the ImmPI score pipeline can be employed for cancer subtyping. Skin cutaneous melanoma (SKCM) is the most prevalent type of skin malignancy and is characterized by a high level of immune cell infiltration. We thus evaluated the roles of immunology piRNAs in SKCM in detail. First, we identified 34 immunology piRNAs by the ImmPI score pipeline, which were associated with infiltrations of all nine immune cell types in SKCM (Fig. S11A). Particularly, all these piRNAs were significantly correlated with prognosis (Fig. S11B), and the expression levels of these piRNAs were significantly associated with each other (Fig. S11C), indicating that they might play critical roles in a module pattern during tumorigenesis. Furthermore, all SKCM samples were classified into two subtypes based on expression of these 34 key immunology piRNAs (Fig. 7A). The majority of these immunology piRNAs exhibited higher expression patterns in the C1 subtype.

Next, we compared the TIME between the two cancer subtypes, and the results indicated that C1 patients had higher infiltration of multiple immune cells, including CD4/8 T cells, B cells, and NK cells (Fig. 7B). In addition, checkpoints showed higher expression in the C1 cluster (Fig. 7C). In addition to validating the difference in immune pathway activities in SKCM patients, the results of GSVA analysis based on gene signatures revealed no difference in angiogenesis and fibrosis signatures between the two cancer subtypes (Fig. S11D). Our results indicated that patients in C1 are likely to have 'hot' tumors whereas the C2 cluster tends towards a 'cold' phenotype. Univariate and multivariate Cox regression analyses were used to investigate the independent predictive value of our subtype. Univariate Cox regression analysis indicated that the C2 subtype had worse OS, PFI, and DSS than the C1 subtype (Fig. 7D). Multivariate Cox regression analysis indicated that our subtype was an independent prognostic factor correlating with OS, PFI, and DSS (Fig. 7E-G).

We investigated its underlying intrinsic immune response mechanisms to reveal the potential mechanism underlying the predictive value of our cancer subtype. The C1 subtype had higher expression of MHC-related antigen-presenting and costimulatory molecules (Fig. 7H), which correlate strongly with enhanced tumor immunogenicity and a relatively hotter immune microenvironment. In addition to intrinsic immune response mechanisms, several extrinsic components, including higher concentrations of immunostimulatory chemokines, T-cell receptor (TCR) diversity, cytolytic activity, and high immunogenicity, were identified in the C1 subtype (Fig. 7I), suggesting different extrinsic immune responses between the two subtypes. These results indicated that our pipeline is able to identify different cancer clusters with notable molecular and immunological diversity, which will contribute to personalized tumor treatment.

2.8. An interactive web portal for tumor immunology piRNAs

We constructed a comprehensive and interactive ImmPI web resource (http://www.hbpding.com/ImmPi) that will aid the biomedical research community in visualizing, searching and browsing immunology piRNAs (Fig. 8). On the basis of immunology piRNAs, we provide several entryways for browsing and querying the piRNA-pathway, piRNA-immune_cell, piRNA-prognosis, and piRNA-cancer relationships across 32 cancer types. Users can enter different browsing pages by clicking the bar plots on the homepage, (Fig. 8A). In addition, users can enter querying pages to query piRNA-pathway, piRNA-immune_cell, piRNA-prognosis, and piRNA-cancer of interest from the menus of the homepage. Our web portal offers a search section to browse data on the basis of piRNAs, tumor type, immune pathways, prognostic type,



(caption on next page)

Fig. 5. Prioritization of cancer-related piRNAs on the basis of immune regulation. (A) The top-ranked piRNAs are enriched in multiple immune signaling pathways. Number of cancer types with piRNAs alterations in tumors (outer loop). Yellow bar denotes number of piRNA enriched in signaling pathways across all types of cancer (inner loop). The longer the yellow bar, the greater the number is, and the corresponding count was marked inside the pink circle on the edge of the yellow bar. Blue lines link the piRNAs and their related immune pathway. The thicker the blue line, the more types of cancer in which a correlation were observed. The size of the outermost circle indicated that the larger the circle, the more types of cancer in which the piRNA was detected. A greater proportion of red indicated a higher proportion of up-regulated piRNAs, while a greater proportion of blue indicated a higher proportion of down-regulated piRNAs. (B) Four examples of piRNAs (piR−hsa−23210, piR−hsa−312787, piR−hsa−545287 and piR−hsa−8465703) in Sankey diagram significantly enriched in multiple immune signaling pathways across cancer types.

immune cells or significant levels (Fig. 8B). The tables and diagram results of all browsing or querying records are offered to users through our portal (Fig. 8C). The GSEA plot, scatter diagram, survival curve, and box-whisker plot are all embedded in each record to present the piRNA-pathway, piRNA-immune_cell, piRNA-prognosis, and piRNA-cancer correlations (Fig. 8C). The results generated in our study can be downloaded in the download menu, and the corresponding figures can also be downloaded as PDF files (Fig. 8D). The help menu provides an introduction and interpretation guidance for our web portal (Fig. 8D). Our valuable web resource will be of enormous interest to the tumor immune community, and will provide innovative viewpoints into tumor immunotherapy.

3. Discussion

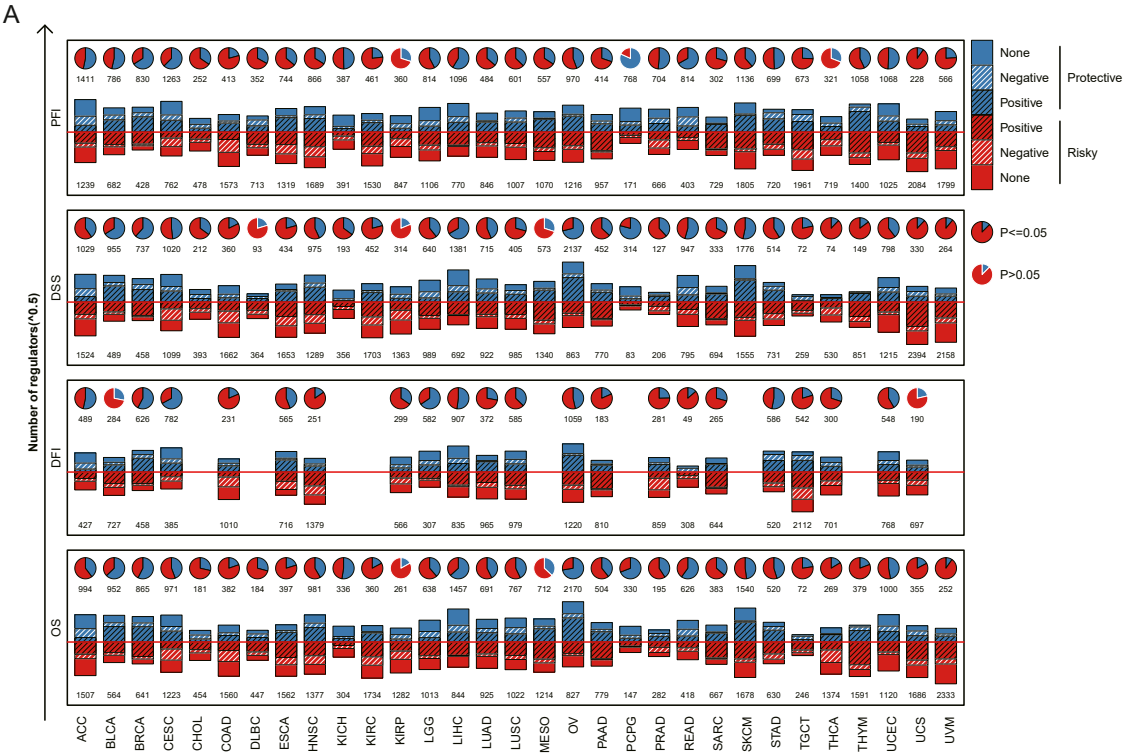
Growing evidence has have demonstrated that piRNAs are critical for tumorigenesis, especially in the tumor microenvironment. However, rare instances have been investigated thus far. In our current study, we investigated the global expression landscape of piRNAs across multiple cancer types and confirmed a vigorous tumor-type-specific expression characteristic of various piRNAs. Our results suggested that piRNAs are promising prognostic and/or diagnostic biomarkers for tumor. By using the ImmPI score pipeline, thousands of piRNAs were identified to play crucial roles in regulating immune pathways in each cancer type, immunology piRNAs are more likely to exhibit distinct expression perturbations in tumor, which correlate significantly with prognosis and infiltration of immune cells. In addition, we demonstrated that ImmPI contributes to prioritizing cancer-related piRNAs and ascertaining cancer subtypes with different immune patterns. The web-based ImmPI resource is the first data portal in the immune-related piRNA field, providing a beneficial resource to comprehensively comprehend the potential immunology impacts of piRNAs in tumor.

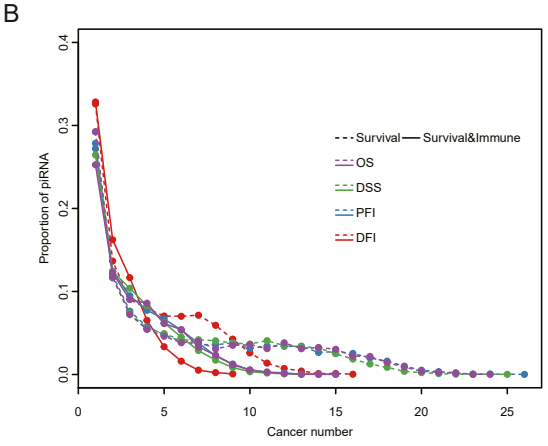
The results of previous studies have indicated that biomarkers generally showed tissue-specific expression patterns. Therefore, the expression characteristics of immunology piRNAs were investigated on the basis of public small RNA-seq profiles of human peripheral blood immune cells [40]. Our analysis demonstrated significantly higher expression levels of immunology piRNAs than those of non-immunology piRNAs, implying that immunology piRNAs have higher expression patterns in human immune cells. In particular, the piRNAs that potentially modulated CD4/8 T or B cell signaling pathways also presented higher expression patterns in CD4/8 T or B cells. These results strongly indicate their crucial roles in the regulation of immunity. In addition to our present pipeline, which can be used for bulk small RNA-seq data for multiple tumor types, increasing expression profiles of immune cells captured by FACS/MACS as well as computational algorithms for evaluating infiltration of the immune cells, including ImmCellAI [39], and CIBERSORT [33], promote our understanding of TIME regulation. We speculate that integration of specific immune cell sequencing profiles and the ImmPI pipeline will improve our comprehension of immune regulation and tumoral genetic heterogeneity.

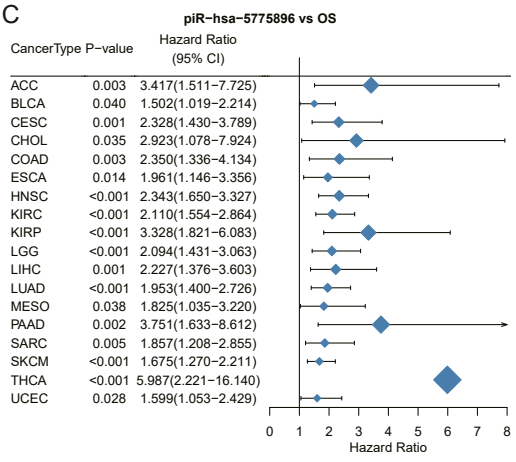
The ImmPI score pipeline is a model-free approach for detecting immunology piRNAs in malignancy. Although various machine-learning approaches have been proposed to detect biomarkers in malignancy [41], it is arduous to apply in immune regulation due to the restricted amounts of experimentally verified immune piRNAs reported in previous studies. To preliminarily confirm the results generated from ImmPI, we investigated whether piRNA–pathway pairs could be reproduced in additional cohorts of the same tumor type (Fig. S6, and Table S4, from CPTAC cohort). Our analysis showed that the identified piRNA–pathway pairs overlapped significantly, indicating that the performance of our approach is robust. Expression-associated protein-coding genes rather than direct targets were used because the overwhelming majority of piRNAs lack known targets. In addition, there is no web-based tool for predicting piRNA-targeting sites within a given mRNA or spliced DNA sequence in humans. Consequently, we found that about 50% of the recognized immunology piRNA-pathway pairs with leading-edge genes had an enrichment in predicted targets, a percentage considerably higher than other piRNA-pathway pairs (Fig. S12). This suggests that the association-based methodology could potentially identify a greater number of immune-related piRNAs for functional verification. The identified leading-edge genes might also aid in anticipating the piRNAs' targets. A growing array of biotechnologies has been employed to pinpoint potential piRNA targets. However, piRNA regulation is context specific, and it is still challenging to identify specific targets in the tumor context. In addition, piRNAs might play biological roles in various patterns, and it is difficult to capture their targets. Thus, we identified protein-coding genes associated with expression of piRNAs and then evaluated their enrichment in immune-related pathways. Owing to the advancement of high-throughput sequencing approaches, it is possible to detect context-specific immunology piRNAs and subsequently enrich our understanding of their biological roles.

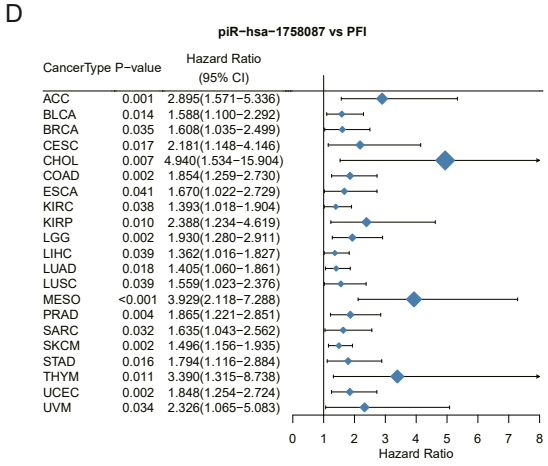
In addition, our analysis indicated that immunology piRNAs can aid in prioritizing tumor-related biomarkers and subtyping. Two subtypes of SKCM were identified on the basis of the expression of immunology piRNAs, which were found to be indepedently associated with prognosis, including OS, PFI, and DSS. Patients in the C1 (immune-hot) and C2 (immune-cold) subgroups had contrasting clinical patterns. The abundance of immune cells in the TIME is known to impact prognosis [42]. In our study, higher infiltration of immune cells, including CD4 and CD8 T cells, B cells, and NK cells was identified in the C1 subtype characterized by stronger immunogenicity. The immune score calculated by the whole infiltration level of immune cells in cancers was higher in C1 patients, who show elevated antitumor immune activity. In addition, checkpoint genes showed higher expression in C1 patients. Because of elevated tumor immunogenicity, enhanced antitumor immunity, and elevated checkpoint expression, it is apparent that the patients in the C1 subgroup tend towards a 'hot' cancer phenotype, indicating that they are more likely to be sensitive to immunotherapy. Altogether, these results indicate that the model based on immunology piRNAs can be utilized to predict prognostic events and response efficiency to immunotherapy.

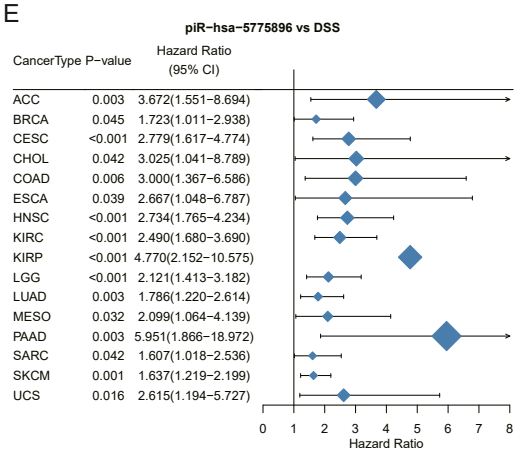
In conclusion, our work reveals the comprehensive landscape of piRNAs across multiple cancer types and highlights their functional and regulatory roles in regulating tumor immunity for the first time. It is crucial to comprehend the impact of immunology piRNAs in cancer. Identifying potential piRNAs offers the first perspective of immunotherapy, and continued evaluation of the immunology piRNAs detected in this study will extraordinarily improve our comprehension of the underlying mechanism of tumors and promote more effective immunotherapies.











(caption on next page)

Fig. 6. The association between immunology piRNAs and clinical outcomes. (A) The correlation between immunology piRNAs and patient prognosis was determined using the log-rank test (P-value < 0.05). Bar charts display the counts of protective (blue bars) and hazardous (red bars) piRNAs across cancer types, with shading illustrating the immunology piRNAs' proportions. The overhead pie charts indicate the percentage of prognosis-linked immunology piRNAs in relation to non-immunology piRNAs. Figures beneath the pie chart depict the total tally of either risk or protective piRNAs. The association between prognosis-related piRNAs and immunology piRNAs was assessed using a hypergeometric test, with significant outcomes marked by the pie charts' black border (P-value < 0.05). OS stands for Overall Survival, DFI for Disease-Free Interval, DSS for Disease-Specific Survival, and PFI for Progression-Free Interval. (B) Prognosis-related piRNAs spread across varying numbers of cancers, with solid and dashed lines indicating those piRNAs associated with prognosis and those also enriched in immune pathways, respectively. (C) piR-hsa-5775896 was correlated with patients' worse OS in more than 10 cancer types. (D) piR-hsa-33063 was correlated with patients' worse PFI in more than 10 cancer types. **(E)** piR-hsa-5775896 was correlated with patients' worse PFI in more than 10 cancer types.

4. Materials and Methods

4.1. Collection of immune-related genes

Seventeen immune pathways were primarily collected on the basis of the ImmPort database [43], which has been extensively applied in recent immunology studies [21,22,44]. All genes involved in these pathways were mapped to Ensembl IDs, and 1811 genes were adopted for subsequent analyses (Table S1).

4.2. Collection and processing of piRNA transcriptomics across cancer types

We downloaded raw small RNA-seq profiles (Level 2) from the TCGA database (http://cancergenome.nih.gov/) and recreated the raw FASTQ files on the basis of BAM files via bedtools2 [45]. In total, 32 divergent TCGA cohorts, each representing a specific tumor, were included in our analysis. Subsequently, we trimmed the recreated FASTQ files on the basis of the qualification of 'Phred quality score more than twenty' and 'reads length more than twenty-one nucleotides' to acquire high-quality sequencing reads corresponding to piRNAs through FASTX-Toolkit (http://hannonlab.cshl.edu/fastx_toolkit/) [8]. All sequences were further remapped to the human reference genome (hg37) via STAR with customized piRNA transcript annotation on the basis of the piRBase database (version 3.0, http://bigdata.ibp.ac.cn/piRBase) [46]. Expression of each piRNA was quantified using featureCounts [47]. Based on piRBase, all piRNA reads were aligned to the newest reference genome through Bowtie to identify underlying piRNAs. When the piRNA genome loci overlapped with RefSeq genes or repeat-associated elements, they were regarded as gene- or repeat-related piRNAs [48]. Because several piRNAs might have different locations, TPM was used to quanlify piRNA transcript levels. We retained piRNAs with a TPM greater than 1 and expression in more than 70% of samples for subsequent analysis.

4.3. Genome-wide mRNA expression across pan-cancer

Genome-wide mRNA profiles across pan-cancer were also downloaded from the TCGA database. FPKM and the raw read count of 32 cancers, quantifying the expression level of mRNAs, were used for subsequent analysis. On the basis of the gene annotations generated from the *HGNC* database [49], expression levels of 19,814 protein-coding genes were obtained for each cancer type.

4.4. Construction of the ImmPI score to identify potential immunology piRNAs

A computational algorithm, named the ImmPI score, integrating piRNA and protein-coding gene transcriptional profiling, was developed to identify potential piRNAs regulating immune pathways (Fig. 2A). In brief, we sorted all protein-coding genes on the basis of their association in expression with a particular piRNA. The sorted gene list for a specific piRNA was then presented to each immune-related pathway to investigate whether the genes in the pathway were enriched in the top or bottom of this gene list. We calculated the ImmPI score of each piRNA–pathway pair and repeated this computational operation for all combinations of piRNAs and immune-related pathways. On the basis of the permutation test, we identified all piRNA–pathway pairs with significantly higher ImmPI scores across cancers.

Discovery of piRNA-correlated genes. For each piRNA, all protein-coding genes were firstly ranked according to the association of their transcriptional level with that piRNA. Transcriptional levels of piRNA x as well as protein-coding gene y across cancer samples were defined as $P(x) = (p_1, p_2, p_3, \ldots, p_i, \ldots, p_n)$ and $G(y) = (g_1, g_2, g_3, \ldots, g_j, \ldots, g_n)$, respectively. We defined the tumor purity across n samples as $T = (t_1, t_2, t_3, \ldots, t_i, \ldots, t_n)$. The partial correlation coefficient (PCC) between the transcriptional level of piRNA i and protein coding gene j was estimated with consideration of the tumor purity acting as the covariable, i.e.,

$$PCC(ij) = \frac{R_{PG} - R_{PT} * R_{GT}}{\sqrt{1 - R_{PT}^2} * \sqrt{1 - R_{GT}^2}}$$

where R_{PG}, R_{PT}, and R_{GT} are the Pearson correlation quotiety between the transcriptional level of piRNA x and protein coding gene y, the transcriptional level of piRNA x and tumor purity, and the transcriptional level of protein coding gene y and tumor purity, respectively. Additionally, the P value of the PCC was also defined as $P(xy)$, and the rank score (RS) of each piRNA-gene pair was estimated as follows:

$$RS(xy) = -\log10(P(xy)) * sign(PCC(xy))$$

All protein-coding genes were sorted based on RS scores and then used for enrichment analysis.

piRNA modulators of immune-related pathways. Inspired by the *GSEA* algorithm, protein-coding genes from immune-related pathways were mapped to the sorted genelist. Furthermore, the enrichment score (ES) was calculated on the basis of GSEA. If N genes existed in the sorted genelist $GL = \{l_1, l_2, l_3, \ldots, l_N\}$, the sorted score was calculated by $RS(l_j) = r_j$. Firstly, the protein-coding genes in immune-related pathway H ("hits") weighted via the RS score and not in S ("misses") present up to a particular position i in GL was assessed as follows:

$$P_{hit}(H, i) = \sum_{l_i \in H\, j \le i} \frac{|r_j|^p}{N_R}, \text{where} N_R = \sum_{l_i \in H} |r_j|^p,$$

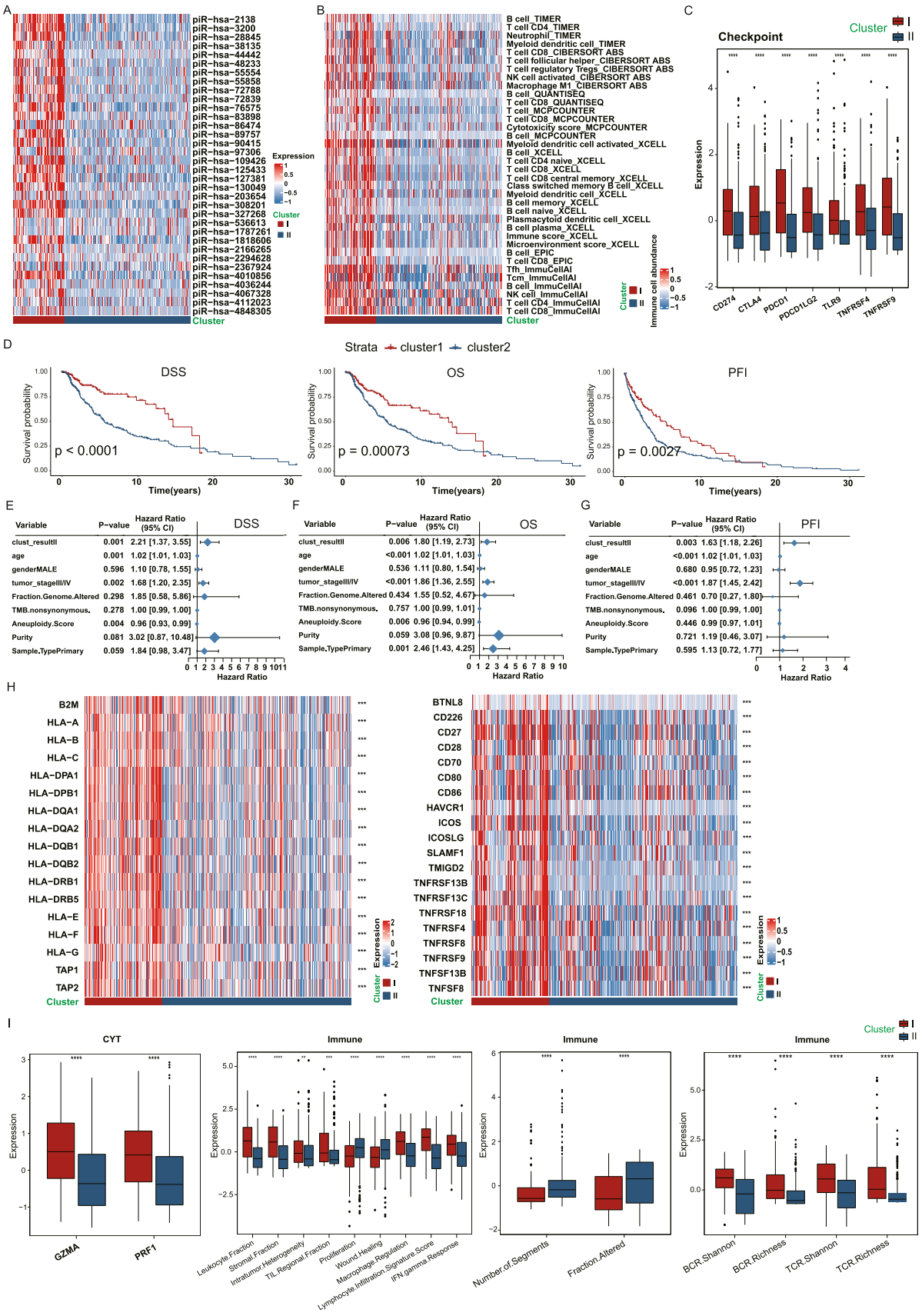
$$P_{miss}(H, i) = \sum_{l_i \notin H\, j \le i} \frac{1}{(N - N_I)}$$

We calculated the ES score on the basis of the maximum deviation from zero of $P_{hit} - P_{miss}$. Additionally, the P-value was estimated for each immune-related pathway including N_I protein-coding genes, as follows:

$$p(\mathrm{ES}(N, N_I) < ES_{ik}) = \sum_{q=-\infty}^{\infty} (-1)^q \exp\left(-2q^2 ES_{ik}^2 n\right),$$

$$n = \frac{(N - N_I)N_I}{N},$$

where ES_{ik} stands for the ES between immune pathway k and piRNA i, the number of protein coding genes in the sorted genelist is denoted by N, and the amount of genes in the particular immune pathway is denoted by N_I. The FDR method was adopted to adjust P-values. In addition, accoding to a recent study, the ES score and P-value were combined to construct an ImmPI score model, i.e.,



(caption on next page)

Fig. 7. Immunology subtypes of SCKM patients with distinct patterns. (A) The classification of SCKM patients on the basis of the expression of the 34 key immunology piRNAs from TCGA. The Heatmap shows the expression level of the screened 34 key immunology piRNAs in two clustering groups. The majority of these immunology piRNAs exhibited higher expression pattern in the C1 subtype. (B) The Heatmap shows the comparison of the infiltration of multiple immune cells estimated by multiple methods based on RNA-sequencing data between two immunology subtypes. Patients in the C1 subtype presented higher infiltration of multiple immune cells. (C) The boxplot shows the comparison of the expression of immune checkpoints between two immunology subtypes. The checkpoints presented higher expression in the C1 subtype. (D) Kaplan-Meier plot indicating that high immunology subtype was significant correlated with better prognosis, including OS, PFI, and DSS. (E-G) The Forest plots show independent predictive value of imunology subtypes for DSS (E), OS (F), and DFI (G), after adjusting accustomed clinical characteristics. (H) Heatmaps show comparison of the expression of MHC-related antigen-presenting and costimulatory molecules between two immunology subtypes. (I) Boxplots show comparison of the expression of immunostimulatory chemokines, T cell receptors (TCR) diversity, cytolytic activity, and the immunogenicity between two immunology subtypes. Wilcox test was used in the comparison between subgroups. * * for P < 0.01, * ** for P < 0.001.

$$\mathrm{ImmPI\_score}(i,k) = \begin{cases} 1-2p; \mathrm{if} ES(i,k) > 0, \\ 2p-1; \mathrm{if} ES(i,k) < 0. \end{cases}$$

Therefore, the range of ImmPI scores was from − 1–1. Significantly positive piRNA–pathway pairs were identified by ImmPI scores greater than 0.995 and FDR less than 0.05, and significantly negative pairs were identified by ImmPI scores less than − 0.995 and FDR more than 0.05.

4.5. Detection of tumor-immunity-associated piRNA

Based on the ImmPI score of each piRNA–pathway pair across cancers, piRNAs correlating with any immune pathways in no less than 15 cancer types were included for subsequent analysis. In addition, we excluded any piRNA whose total number of associations between specific piRNAs and any immune pathways in all cancer types was no more than 25 (Fig. S1). Ultimately, 481 piRNAs were defined as tumor-immunity-associated piRNAs in our analysis. The Immune Enrichment (*IE*) score of each tumor immunity-associated piRNA across 32 cancer types was estimated on the basis of the ImmPI score. We calculated the *IE* score as follows:

$$IE(i,k) = \sum_{c=1}^{32} I(\mathrm{ImmPI\_score}(i,k)) * sign(\mathrm{ImmPI\_score}(i,k)),$$

$$\mathrm{where} I(x) = \begin{cases} 0; if |\mathrm{ImmPI\_score}(i,k)| < 0.995, \\ 1; if |\mathrm{ImmPI\_score}(i,k)| \geq 0.995. \end{cases}$$

4.6. Infiltration of immune cell across cancer types

In order to evaluate the correlation of candidate piRNAs and tumor immunity, immune cell infiltration across cancer types was estimated by six algorithms, including TIMER [50], EPIC [34], MCP-counter [35], xCell [38], QUANTISEQ [36], and CIBERSORT [33]. Our major focus was on B cells, CD4/8 T cells, T helper cells, Tregs, neutrophils, dendritic cells, macrophages/Monocytes, and NK cells. The Spearman correlation coefficient (ρ) was firstly estimated between the expression of piRNAs and abundances of immune cells. Significant pairs of piRNA-immune cells within distinct algorithms were identified on the basis of $P < 0.05$ & $|\rho| > 0.2$. In comprehensive consideration of the relationships between immune cells and piRNAs, piRNA-immune cell pairs (union sets) within the same cell type resulting from distinct computational algorithms were merged in our analysis.

4.7. Differential expression of piRNAs and mRNAs

Tumor types with no less than five paracancerous tissues were used for the differential expression analysis. Differentially expressed piRNAs and mRNAs based on the raw read count between cancer and paracancerous samples were identified using the *DESeq2* R package [51]. We estimated the adjusted *P*-values of each piRNA and mRNA using the false-discovery rate (FDR) approach. An FDR of < 0.05 and absolute $\log_2$-fold change of > 1 were regarded as the cut-offs to identify differentially expressed piRNAs and mRNAs [52].

4.8. Clinical relevance analysis of piRNAs

Clinical information across 32 cancer types, including overall survival (OS), disease-specific survival (DSS), disease-free interval (DFI) or progression-free interval (PFI) [53], stage, grade, sex, age, fraction genome altered (FGA), tumor mutation burden (TMB), and aneuploidy score, was downloaded from the TCGA database. The optimal cut-off of each piRNA was ascertained using the *surv_cutpoint* function of the *survminer* R package to separate the samples into high- or low-expression subgroups. Cox proportional hazards regression analysis was applied to investigate the prognostic power of all piRNAs. Siginificant prognostic piRNAs with positive β values of 'coxph' were regarded as risk piRNAs, and the negative piRNAs were regarded as protective piRNAs.

4.9. Subtype analysis on the basis of immunology piRNAs

To identify piRNAs that might be utilized for categorizing tumor patients, we first screened the key immunology piRNAs siginificantly associated with infiltration of all nine immune cells in skin cutaneous melanoma (SKCM). Next, the *ConsensusClusterPlus* R package was used to divide patients into optimum subgroups by referencing the transcriptional level of these key immunology piRNAs. Expression of piRNA was first scaled by the Z-score and then used in subsequent analysis.

We compared several immune and clinical characteristics of patients to characterize subtypes. Immunogenomic indices were accessed from the pan-cancer immune landscape analysed by Thorsson et al. [54]. Briefly, TCR diversity scores (Shannon entropy and richness) were assessed from corresponding transcriptomics [55]. The cytolytic activity score (CYT) was estimated by calculating the geometric mean of the transcriptional levels of granzyme A (*GZMA*) and perforin 1 (*PRF1*) [56]. For clinical features, multivariate Cox regression analysis was performed to evaluate whether a subtype is an independent prognostic factor.

4.10. Construction of the ImmPI web portal

The ImmPI web portal was built with JS and HTML5 as the front-end, Rscript 4.2.1 and PHP 7.4 as the back-end. All data resources in the ImmPI web portal are stored as * .gz or * . RData. Data retrieval or processing is performed by Rscript in our server. The httpd-service is run on Apache 2.4. One can visit our portal freely at http://www.hbpding.com/ImmPi.

Funding

This research was supported by the National Natural Science Foundation of China (82203879 and 82302249), Natural Science Foundation of Hubei Province (2022CFB883) and China Postdoctoral Science Foundation (2021MD703960).

CRediT authorship contribution statement

Study concept and design: D.W., R.L., G.L. Methodology, acquisition, analysis, or interpretation of the data: D.W., R.L., H.H., X.Z., G.L. Drafting of the manuscript: D.W., R.L., G.L. Statistical analysis: H.H, X.Z, G.L. Study supervision: H.H, X.Z, G.L. All authors read and approved the

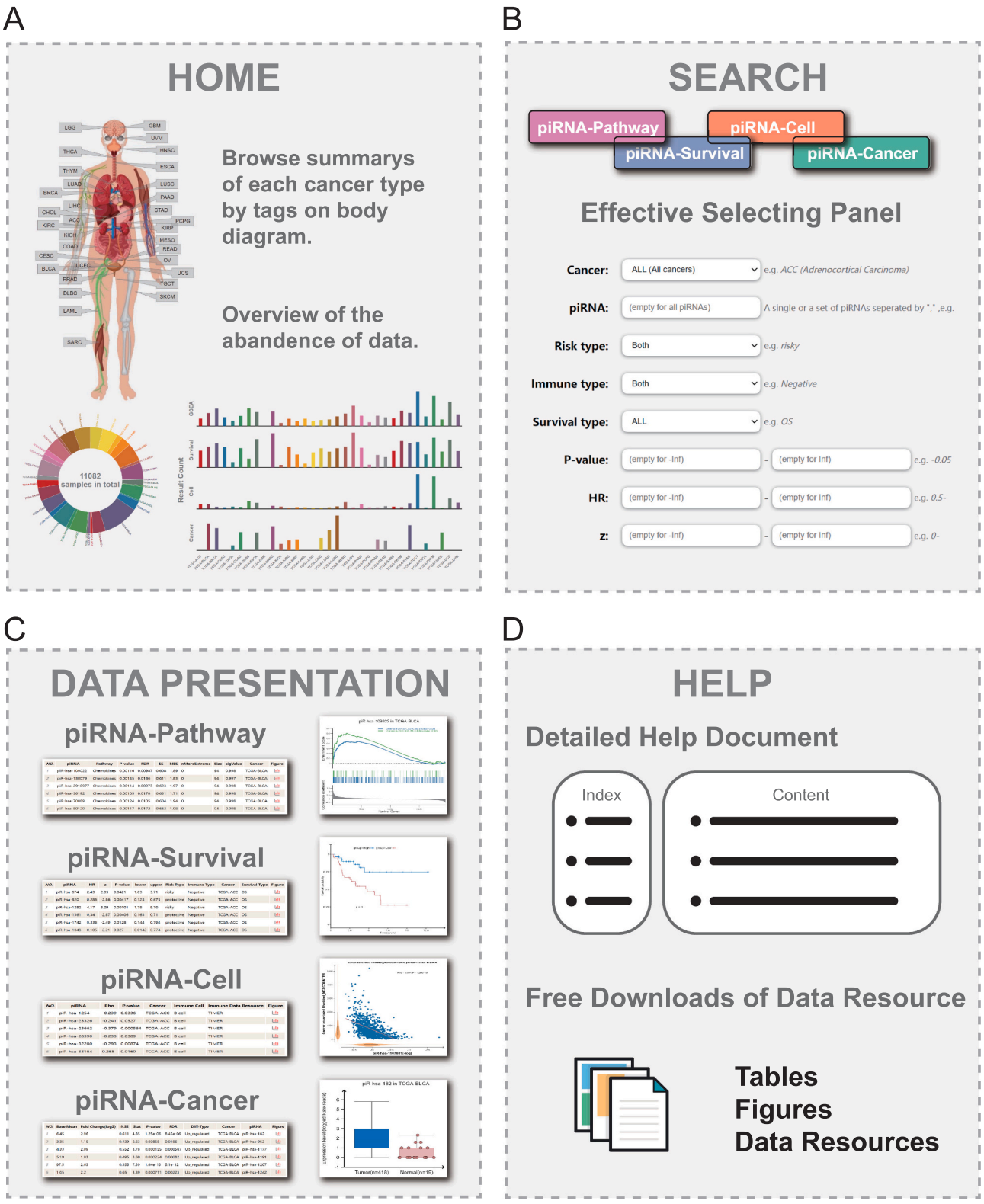


Fig. 8. Introduction to the ImmPI Online Platform: (A) From the homepage, visitors can navigate through immunology piRNAs based on cancers, pathways, immune cells, or indices. (B) The search interface offers various categories like 'piRNA-Pathway', 'piRNA-Cell', 'piRNA-Prognosis', and 'piRNA-Cancer'. There are options for users to fine-tune their search through advanced filtering features. (C) On the results page, a list of piRNAs with corresponding details is displayed. The detailed view offers users insights into enrichment plots, correlation diagrams, KM-curves, and differential expression box plots. (D) The 'Download' section allows users to access all the results generated during our study. For assistance and understanding of the ImmPI platform, users can refer to the 'Help' section, which offers a comprehensive guide.

final manuscript.

Declaration of Competing Interest

The authors declare no competing interests.

Data Availability

The raw genotype and small RNA sequencing data have been deposited in The Cancer Genome Atlas (TCGA, http://cancergenome.nih.gov/) and Clinical Proteomic Tumor Analysis Consortium (CPTAC, https://proteomics.cancer.gov/programs/cptac) program. Small RNA sequencing data of immune cells from human peripheral blood supporting the current study are deposited under GSE100467 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE100467). The ImmPI platform offers an intuitive interface for the biomedical research community to visualize, search, and navigate through immunology piRNAs data (http://www.hbpding.com/ImmPi).

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.csbj.2023.10.042.

References

[1] Delgado MD, Leon J. **Gene expression regulation and cancer**. Clinical & translational oncology: official publication of the Federation of Spanish Oncology Societies and of the National Cancer Institute of Mexico 2006;8(11):780–7.
[2] Lee TI, Young RA. Transcriptional regulation and its misregulation in disease. Cell 2013;152(6):1237–51.
[3] Girard A, Sachidanandam R, Hannon GJ, Carmell MA. A germline-specific class of small RNAs binds mammalian Piwi proteins. Nature 2006;442(7099):199–202.
[4] Lau NC, Seto AG, Kim J, Kuramochi-Miyagawa S, Nakano T, Bartel DP, et al. Characterization of the piRNA complex from rat testes. Sci (N Y, NY) 2006;313(5785):363–7.
[5] Vagin VV, Sigova A, Li C, Seitz H, Gvozdev V, Zamore PD. A distinct small RNA pathway silences selfish genetic elements in the germline. Sci (N Y, NY) 2006;313(5785):320–4.
[6] Yan Z, Hu HY, Jiang X, Maierhofer V, Neb E, He L, et al. Widespread expression of piRNA-like molecules in somatic tissues. Nucleic Acids Res 2011;39(15):6596–607.
[7] Muller S, Raulefs S, Bruns P, Afonso-Grunz F, Plotner A, Thermann R, et al. Next-generation sequencing reveals novel differentially regulated mRNAs, lncRNAs, miRNAs, sdRNAs and a piRNA in pancreatic cancer. Mol Cancer 2015;14:94.
[8] Mai D, Ding P, Tan L, Zhang J, Pan Z, Bai R, et al. PIWI-interacting RNA-54265 is oncogenic and a potential therapeutic target in colorectal adenocarcinoma. Theranostics 2018;8(19):5213–30.
[9] Chu H, Hui G, Yuan L, Shi D, Wang Y, Du M, et al. Identification of novel piRNAs in bladder cancer. Cancer Lett 2015;356(2 Pt B):561–7.
[10] Liu Y, Dou M, Song X, Dong Y, Liu S, Liu H, et al. The emerging role of the piRNA/piwi complex in cancer. Mol Cancer 2019;18(1):123.
[11] Ozata DM, Gainetdinov I, Zoch A, O'Carroll D, Zamore PD. PIWI-interacting RNAs: small RNAs with big functions. Nat Rev Genet 2019;20(2):89–108.
[12] Wang X, Ramat A, Simonelig M, Liu MF. Emerging roles and functional mechanisms of PIWI-interacting RNAs. Nat Rev Mol Cell Biol 2022.
[13] Li Y, Wu X, Gao H, Jin JM, Li AX, Kim YS, et al. Piwi-interacting RNAs (piRNAs) are dysregulated in renal cell carcinoma and associated with tumor metastasis and cancer-specific survival. Mol Med (Camb, Mass) 2015;21(1):381–8.
[14] Han H, Fan G, Song S, Jiang Y, Qian C, Zhang W, et al. piRNA-30473 contributes to tumorigenesis and poor prognosis by regulating m6A RNA methylation in DLBCL. Blood 2021;137(12):1603–14.
[15] Yao J, Wang YW, Fang BB, Zhang SJ, Cheng BL. piR-651 and its function in 95-D lung cancer cells. Biomed Rep 2016;4(5):546–50.
[16] Iyer DN, Wan TM, Man JH, Sin RW, Li X, Lo OS, et al. Small RNA profiling of piRNAs in colorectal cancer identifies consistent overexpression of piR-24000 that correlates clinically with an aggressive disease phenotype. Cancers 2020;12(1).
[17] Zhou X, Liu J, Meng A, Zhang L, Wang M, Fan H, et al. Gastric juice piR-1245: a promising prognostic biomarker for gastric cancer. J Clin Lab Anal 2020;34(4):e23131.
[18] Tan L, Mai D, Zhang B, Jiang X, Zhang J, Bai R, et al. PIWI-interacting RNA-36712 restrains breast cancer progression and chemoresistance by interaction with SEPW1 pseudogene SEPW1P RNA. Mol Cancer 2019;18(1):9.
[19] Bagaev A, Kotlov N, Nomie K, Svekolkin V, Gafurov A, Isaeva O, et al. Conserved pan-cancer microenvironment subtypes predict response to immunotherapy. Cancer Cell 2021;39(6):845–65. e847.
[20] Wang J, Li R, Cao Y, Gu Y, Fang H, Fei Y, et al. Intratumoral CXCR5(+)CD8(+)T associates with favorable clinical outcomes and immunogenic contexture in gastric cancer. Nat Commun 2021;12(1):3080.
[21] Li Y, Jiang T, Zhou W, Li J, Li X, Wang Q, et al. Pan-cancer characterization of immune-related lncRNAs identifies potential oncogenic biomarkers. Nat Commun 2020;11(1):1000.
[22] Jiang T, Zhou W, Chang Z, Zou H, Bai J, Sun Q, et al. ImmReg: the regulon atlas of immune-related pathways across cancer types. Nucleic Acids Res 2021;49(21):12106–18.
[23] Chen S, Ben S, Xin J, Li S, Zheng R, Wang H, et al. The biogenesis and biological function of PIWI-interacting RNA in cancer. J Hematol Oncol 2021;14(1):93.
[24] Zhu J, He F, Hu S, Yu J. On the nature of human housekeeping genes. Trends Genet: TIG 2008;24(10):481–4.
[25] Ricketts CJ, De Cubas AA, Fan H, Smith CC, Lang M, Reznik E, et al. The cancer genome atlas comprehensive molecular characterization of renal cell carcinoma. Cell Rep 2018;23(1):313–26. e315.
[26] Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci USA 2005;102(43):15545–50.
[27] Bhattacharya S, Dunn P, Thomas CG, Smith B, Schaefer H, Chen J, et al. ImmPort, toward repurposing of open access immunological assay data for translational and clinical research. Sci data 2018;5:180015.
[28] Haber PK, Castet F, Torres-Martin M, Andreu-Oller C, Puigvehi M, Miho M, et al. Molecular markers of response to Anti-PD1 therapy in advanced hepatocellular carcinoma. Gastroenterology 2022.
[29] Cao L, Huang C, Cui Zhou D, Hu Y, Lih TM, Savage SR, et al. Proteogenomic characterization of pancreatic ductal adenocarcinoma. Cell 2021;184(19):5031–52. e5026.
[30] Sedano R, Cabrera D, Jimenez A, Ma C, Jairath V, Arrese M, et al. Immunotherapy for cancer: common gastrointestinal, liver, and pancreatic side effects and their management. Am J Gastroenterol 2022;117(12):1917–32.
[31] Osmani L, Askin F, Gabrielson E, Li QK. Current WHO guidelines and the critical role of immunohistochemical markers in the subclassification of non-small cell lung carcinoma (NSCLC): moving from targeted therapy to immunotherapy. Semin Cancer Biol 2018;52(Pt 1):103–9.
[32] Ino Y, Yamazaki-Itoh R, Shimada K, Iwasaki M, Kosuge T, Kanai Y, et al. Immune cell infiltration as an indicator of the immune microenvironment of pancreatic cancer. Br J Cancer 2013;108(4):914–23.
[33] Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, et al. Robust enumeration of cell subsets from tissue expression profiles. Nat Methods 2015;12(5):453–7.
[34] Racle J, de Jonge K, Baumgaertner P, Speiser DE, Gfeller D. Simultaneous enumeration of cancer and immune cell types from bulk tumor gene expression data. eLife 2017:**6**.
[35] Becht E, Giraldo NA, Lacroix L, Buttard B, Elarouci N, Petitprez F, et al. Estimating the population abundance of tissue-infiltrating immune and stromal cell populations using gene expression. Genome Biol 2016;17(1):218.
[36] Finotello F, Mayer C, Plattner C, Laschober G, Rieder D, Hackl H, et al. Molecular and pharmacological modulators of the tumor immune contexture revealed by deconvolution of RNA-seq data. Genome Med 2019;11(1):34.
[37] Li T, Fan J, Wang B, Traugh N, Chen Q, Liu JS, et al. TIMER: a web server for comprehensive analysis of tumor-infiltrating immune cells. Cancer Res 2017;77(21):e108–10.
[38] Aran D, Hu Z, Butte AJ. xCell: digitally portraying the tissue cellular heterogeneity landscape. Genome Biol 2017;18(1):220.
[39] Miao YR, Zhang Q, Lei Q, Luo M, Xie GY, Wang H, et al. ImmuCellAI: a unique method for comprehensive T-cell subsets abundance prediction and its application in cancer immunotherapy. Adv Sci 2020;7(7):1902880.
[40] Juzenas S, Venkatesh G, Hubenthal M, Hoeppner MP, Du ZG, Paulsen M, et al. A comprehensive, cell specific microRNA catalogue of human peripheral blood. Nucleic Acids Res 2017;45(16):9290–301.
[41] Rajkomar A, Dean J, Kohane I. Machine Learning in Medicine. N Engl J Med 2019;380(14):1347–58.
[42] Quail DF, Joyce JA. Microenvironmental regulation of tumor progression and metastasis. Nat Med 2013;19(11):1423–37.
[43] Bhattacharya S, Andorf S, Gomes L, Dunn P, Schaefer H, Pontius J, et al. ImmPort: disseminating data to the public for the future of immunology. Immunol Res 2014;58(2–3):234–9.
[44] Wang G, Xie Z, Su J, Chen M, Du Y, Gao Q, et al. Characterization of Immune-Related Alternative Polyadenylation Events in Cancer Immunotherapy. Cancer Res 2022;82(19):3474–85.
[45] Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinforma (Oxf, Engl) 2010;26(6):841–2.
[46] Wang J, Shi Y, Zhou H, Zhang P, Song T, Ying Z, et al. piRBase: integrating piRNA annotation in all aspects. Nucleic Acids Res 2022;50(D1):D265–72.
[47] Liao Y, Smyth GK, Shi W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinforma (Oxf, Engl) 2014;30(7):923–30.
[48] Xin J, Du M, Jiang X, Wu Y, Ben S, Zheng R, et al. Systematic evaluation of the effects of genetic variants on PIWI-interacting RNA expression across 33 cancer types. Nucleic Acids Res 2021;49(1):90–7.
[49] Tweedie S, Braschi B, Gray K, Jones TEM, Seal RL, Yates B, et al. Genenames.org: the HGNC and VGNC resources in 2021. Nucleic Acids Res 2021;49(D1):D939–46.
[50] Li T, Fu J, Zeng Z, Cohen D, Li J, Chen Q, et al. TIMER2.0 for analysis of tumor-infiltrating immune cells. Nucleic Acids Res 2020;48(W1):W509–14.
[51] Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 2014;15(12):550.

[52] Li G, Xu W, Zhang L, Liu T, Jin G, Song J, et al. Development and validation of a CIMP-associated prognostic model for hepatocellular carcinoma. EBioMedicine 2019;47:128–41.
[53] Liu J, Lichtenberg T, Hoadley KA, Poisson LM, Lazar AJ, Cherniack AD, et al. An integrated TCGA pan-cancer clinical data resource to drive high-quality survival outcome analytics. Cell 2018;173(2):400–16. e411.
[54] Thorsson V, Gibbs DL, Brown SD, Wolf D, Bortone DS, Ou Yang TH, et al. The immune landscape of cancer. Immunity 2019;51(2):411–2.
[55] Brown SD, Raeburn LA, Holt RA. Profiling tissue-resident T cell repertoires by RNA sequencing. Genome Med 2015;7:125.
[56] Rooney MS, Shukla SA, Wu CJ, Getz G, Hacohen N. Molecular and genetic properties of tumors associated with local immune cytolytic activity. Cell 2015;160 (1–2):48–61.