Scientific Reports
https://doi.org/10.1038/s41598-026-44997-1
Article in Press

# Immunogenomic classification reveals prognostic immune signatures in pediatric solid and hematological tumors

Received: 22 August 2024
Accepted: 16 March 2026
Published online: 31 March 2026

Zhouqi Xia, Qiaoli Hua, Jianqin Qian, Jinhu Wang, Jiangtao Liang, Qiaojun He, Shaoqing Ni & Ting Tao

**Cite this article as:** Xia Z., Hua Q., Qian J. *et al.* Immunogenomic classification reveals prognostic immune signatures in pediatric solid and hematological tumors. *Sci Rep* (2026). https://doi.org/10.1038/s41598-026-44997-1

We are providing an unedited version of this manuscript to give early access to its findings. Before final publication, the manuscript will undergo further editing. Please note there may be errors present which affect the content, and all legal disclaimers apply.

If this paper is publishing under a Transparent Peer Review model then Peer Review reports will publish with the final article.

ARTICLE IN PRESS

© The Author(s) 2026. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by-nc-nd/4.0/.

ARTICLE IN PRESS

# Immunogenomic Classification Reveals Prognostic Immune Signatures in Pediatric Solid and Hematological Tumors

**Zhouqi Xia[1,2], Qiaoli Hua[1,2], Jianqin Qian[1], Jinhu Wang[3,4,5,6], Jiangtao Liang[7], Qiaojun He[2,8,9*], Shaoqing Ni[1,2*], Ting Tao[3,4,5,6*]**

[1]National Clinical Trial Institute, Children's Hospital, Zhejiang University School of Medicine, National Clinical Research Center for Children and Adolescents' Health and Diseases, Hangzhou, Zhejiang 310052, P.R. China.

[2]Research Center for Clinical Pharmacy, College of Pharmaceutical Sciences, Zhejiang University, Hangzhou, Zhejiang 310058, P.R. China.

[3]Pediatric Cancer Research Center, Children's Hospital, Zhejiang University School of Medicine, National Clinical Research Center for Children and Adolescents' Health and Diseases, Hangzhou, Zhejiang 310052, P.R. China.

[4]Department of Surgical Oncology, Children's Hospital, Zhejiang University School of Medicine, National Clinical Research Center for Children and Adolescents' Health and Diseases, Hangzhou, Zhejiang 310052, P.R. China.

[5]Zhejiang Key Laboratory of Neonatal Diseases, Hangzhou, Zhejiang 310052, P.R. China.

[6]Cancer Center, Zhejiang University, Hangzhou, Zhejiang 310052, P.R. China.

[7]Hangzhou Universal Medical Imaging Diagnostic Center, Hangzhou, Zhejiang 310009, P.R. China.

[8]Center for Drug Safety Evaluation and Research, College of Pharmaceutical Sciences, Zhejiang University, Hangzhou, Zhejiang 310058, P.R. China.

ARTICLE IN PRESS

ARTICLE IN PRESS

[9]Hangzhou Institute of Innovative Medicine, College of Pharmaceutical Sciences, Zhejiang University, Hangzhou, Zhejiang 310058, P.R. China.

*Correspondence: Ting Tao, Pediatric Cancer Research Center, Children's Hospital, Zhejiang University School of Medicine, National Clinical Research Center for Children and Adolescents' Health and Diseases, Hangzhou, 310052, China.

Email: taot@zju.edu.cn.

Shaoqing Ni, National Clinical Trial Institute, Children's Hospital, Zhejiang University School of Medicine, National Clinical Research Center for Children and Adolescents' Health and Diseases, Hangzhou, 310052, China.

Email: chgcp@zju.edu.cn.

Qiaojun He, Research Center for Clinical Pharmacy, College of Pharmaceutical Sciences, Children's Hospital School of Medicine, Zhejiang University, Hangzhou, 310058, China.

Email: qiaojunhe@zju.edu.cn.

**Keywords**: Pediatric solid tumor; Pediatric hematological tumor; Immune features; Prognosis

**Running title**: Pediatric Tumor Classification with Immune Prognosis

ARTICLE IN PRESS

ARTICLE IN PRESS

**Abstract**

The immune features in pediatric tumor are poorly explored. To characterize immune features of pediatric cancer, we performed an immunogenomic analysis of public database (TARGET) for pediatric solid tumor (PST) (n=423) and pediatric hematological tumor (PHT) (n=2302). We clustered PST and PHT samples into 5 subtypes (S1-S5) and 4 subtypes (H1-H4), respectively, based on immune features. In the PST cohort, cluster S1 with elevated expression of Wound_CSR (fibroblast core serum response in wound healing) and B cell exhibited the worst overall survival. Conversely, cluster S4 (HR = 0.378, 95% CI: 0.24–0.59, *P*-value < 0.001) with down-regulated expression of these features were associated with prolonged survival. We also validated the prognostic significance of the S4 immune subtype in an independent neuroblastoma cohort from the ZJUCH (n = 127), which demonstrated favorable patient outcomes. In the PHT cohort, we observed that the relationships between immune clusters and prognosis differed between *FLT3*-ITD mutation-positive AML (AML-1) and *FLT3*-ITD mutation-negative AML (AML-2). In AML-1, cluster H2 featured upregulated infiltration of neutrophils, monocytes and antigen processing signatures, possibly leading to worst overall survival. While in AML-2, cluster H2 exhibited a favorable outcome. The study highlights the potential of immune features as biomarkers for prognosis and treatment planning in pediatric cancers and provides novel insights into their immunological landscape.

ARTICLE IN PRESS

ARTICLE IN PRESS

## Introduction

It is estimated that approximately 400,000 children worldwide are diagnosed with cancer each year[1]. Despite decreased mortality over the past 40 years, cancer remains the leading cause of childhood death[2]. While over 80% of children with cancer in high-income countries can be cured, 5-year survival rates for pediatric cancer in lower-middle-income regions remain less than 30%[3]. The most common categories of pediatric cancer include brain tumors, hematological tumors and solid tumors[4]. Unfortunately, many high-grade and relapsed tumors within these categories still exhibit a poor prognosis, such as high-risk neuroblastoma (NB)[5] and *TP53*-aneuploidy acute myeloid leukemia (AML)[6].

Pediatric tumors tend to have lower mutation burdens, presenting unique challenges due to the complexity of genetic dependencies[7]. Risk-based and response-adapted treatments have extended survival in pediatric cancers, yet the late effects of cytotoxic chemotherapy in high-risk subsets remain life-threatening and unavoidable[8]. Hence, there is an urgent need to develop advanced approaches that can either replace or complement the current relatively robust strategies. Immunotherapy[9] and targeted small molecules[10] have revolutionized the treatment of adult cancers. However, there has been limited progress in their effectiveness for pediatric counterparts[11]. The obstacles mainly come from less supporting evidences of adult drugs to treat pediatric cancer and a shortage of agents specially designed for childhood tumors.

ARTICLE IN PRESS

Immune checkpoint inhibitors have achieved effective response in multiple types of cancer, and the major mechanisms of responsiveness include the diversity of tumor

ARTICLE IN PRESS

immune microenvironment (TIME)[12]. TIME involves tumor-infiltrating cells associated with tumor-antagonizing or tumor-promoting functions[13]. These cells provide opportunities for cancer treatment, such as utilizing anti-PD1 antibodies and CAR-T cell therapies. In pediatric cancers, the immunogenomic features of tumors are less well understood. However, recent studies have highlighted promising potential in immunological signatures, such as the immunologic constant of rejection, and the presence of immune cell lineages that might correlate with prognosis, especially in neuroblastoma[14]. Additionally, mesenchymal and adrenergic cell lineages have drawn great attention in NB, and mesenchymal lineage is thought to possess a favorable immune state[15]. Although the immunotherapy has not been widely applied to pediatric cancers due to their relatively lower mutational burdens[11], the newly developed strategies, such as anti-GD2 antibodies[16] and CD19 CAR T-cell therapy[17], have shown hopeful results. For instance, a study demonstrated that Kymriah (tisagenlecleucel), a CD19-targeting CAR-T therapy, achieved a 75% complete remission rate in pediatric patients with relapsed or refractory B-cell acute lymphoblastic leukemia (ALL), further supporting the exploration of CAR-T therapies in pediatric cancers[18].

In this study, we aim to harness computational approaches to analyze immune infiltration patterns in pediatric solid tumors (PST) and hematological tumors (PHT). Specifically, we focus on understanding the prognostic significance of immune infiltration and identifying immune-related signaling pathways and genes that may serve as therapeutic targets. The key research questions guiding this study are: 1) What are the immune infiltration patterns in PST and PHT, and how do these patterns

ARTICLE IN PRESS

ARTICLE IN PRESS

correlate with patient prognosis? 2) Which immune-related signaling pathways and immune genes are associated with prognosis in pediatric cancer patients?

To address these questions, we collected RNA-seq data from pediatric solid and hematological tumors in the TARGET dataset. Using 184 immune cell signatures[19], we classified PST and PHT into distinct immune subtypes and investigated the prognostic significance of these immune subtypes. We also explored underlying prognostic biomarkers and validated our findings using neuroblastoma samples, where high gene scores (S4 gene score) of 44 upregulated genes in the S4 cluster were associated with a favorable prognosis. This study provides a comprehensive understanding of the immune landscape in pediatric tumors and identifies promising immune-related biomarkers and therapeutic targets. The findings could inform the development of novel immune-based therapies, offering new avenues for improving clinical outcomes in pediatric cancer treatment.

ARTICLE IN PRESS

## Methods

### Samples collection

Patients aged 0-19 years in the TARGET pediatric dataset (up until September 2023) diagnosed with NB, osteosarcoma (OS), Wilms tumor (WT), rhabdoid tumor (RT), clear cell sarcoma of kidney (CCSK), AML and ALL were included, detailed information regarding the patient demographics and clinical characteristics could be found on the TARGET website. For ALL, we included patients in TARGET ALL expansion effort (ALL Phase II, ALL-P2) and the latest TARGET ALL study (ALL Phase III, TARGET-ALL-P3).

ARTICLE IN PRESS

In contrast, we excluded the participants in TARGET ALL Pilot project (ALL Phase I, ALL-P1) due to its low proportion of pediatric patients and relatively small sample size. ALL-P1 was designed primarily as a feasibility study to test data collection and sequencing methods and included a higher proportion of adult patients, which made the dataset less suitable for our focused analysis of pediatric ALL. All metastatic and recurrent tumor samples were excluded, and primary tumor samples were retained. RNA-seq data in FASTQ format files were downloaded from database of Genotypes and Phenotypes (dbGaP) under the accession number phs000218.v26.p8 with approval by National Institutes of Health (NIH)[20]. The RNA-seq gene expression data consisting of gene counts, TPM (transcripts per kilobase of exon per million mapped reads), FPKM (fragments per kilobase of exon per million mapped fragments) in STAR-processed format files[21] and the clinical data were acquired from the GDC portal website.

In addition, NB samples (n=127) for validation were obtained from Department of Surgical Oncology, Children's Hospital Zhejiang University School of Medicine (refer to as ZJUCH cohort), and subjected to RNA-seq analysis. To eliminate batch effects and normalize the data, ComBat from the sva package (version 3.42.0) in R was applied, using the TARGET-NB samples as a reference for the adjustment across datasets. The patients enrolled in the study were all under the age of 18, had received first-line chemotherapy, and had undergone surgical resection or biopsy at the Children's Hospital of Zhejiang University School of Medicine. This study was approved by the Ethics Committee of Children’s Hospital, Zhejiang University School

ARTICLE IN PRESS

ARTICLE IN PRESS

of Medicine (2020-IRB-049). All methods were performed in accordance with the relevant guidelines and regulations. All data were acquired with informed consent from patients, their parents or guardians. All results were reported with arbitrary sample ID numbers without linked identifiers.

**RNA sequencing (RNA-seq) and data analysis**

Total RNA of tumor tissues from ZJUCH cohort were extracted by the TRIzol reagent (Thermo Fisher Scientific, USA). mRNA was purified from total RNA using Dynabeads Oligo (dT) (Thermo Fisher, USA) with two rounds of purification, and subjected to library preparation for Illumina NovaSeq 6000 S4 sequencing with a 150 bp paired-end sequencing strategy at Zhejiang Biosan Biochemical Technologies Co., Ltd. RNA-seq data were aligned to the GRCh38 reference genome with STAR[21], followed by duplicates marking with Picard (https://broadinstitute.github.io/picard/), and reads mapped to each gene were counted by featureCounts[22].

ARTICLE IN PRESS

**Immune signatures and immune subtypes**

The gene transcriptome data were utilized to perform single-sample gene set enrichment analysis (ssGSEA) using the R package GSVA[23]. The 184 immune cell signatures[19] were served as references to calculate gene set enrichment scores (GSEs), representing the expression levels of tumor-infiltrating immune cell signatures. With the coxph function from R packages survminer (v.0.4.9) and survival (v.3.5.7), the relationship between GSEs of 184 immune cell signatures and overall survival time

was analyzed through univariate Cox regression, and signatures of PST or PHT with *P*-values less than 0.25 or 0.05 were selected for further analysis. The relatively higher *P*-value threshold of 0.25 for PST was chosen because, in our dataset, the number of immune signatures with *P*-values below 0.05 was limited (Supplementary Table S1), and raising the threshold allowed for the inclusion of potentially informative immune signatures that showed a trend toward association with survival, reflecting the more subtle role of immune features in these tumors. Spearman correlation between the GSEs of these signatures was performed, and signatures with relatively larger values of coefficient were clustered into the same modules by R package corrplot (v.0.92). Finally, ssGSEA were operated to calculate the GSEs of these modules, which were utilized to draw heatmap with R package pheatmap (v.1.0.12) and cluster PST and PHT into S1-S5 and H1-H4 subtypes with clustering method "ward.D2", respectively. And univariate COX analysis was performed to search the relationship between these GSEs of immune modules and overall survival time (*P*-values were adjusted by Benjamini-Hochberg correction), and results were visualized by R package forestplot (v.3.1.3) and ggolot2 (v.3.4.4). The survival curves of immune subtypes were plotted by the survdiff function from R package survminer, and the *P*-values were calculated by log rank test. A flowchart outlining the overall bioinformatic pipeline is shown as Fig. 1.

ARTICLE IN PRESS

**Differentially expressed immune genes and pathways of immune subtypes**

To explore the differentially expressed immune genes of selected immune signatures

ARTICLE IN PRESS

among immune subtypes, R package DESeq2 (v.1.42.0)[24] was utilized to find out the differential gene expression between specific subgroup and others. Benjamini–Hochberg correction was applied to adjust *P*-values (denoted as q-values). Up-regulated ($log_2$-fold change > 1, q value < 0.05) and down-regulated ($log_2$-fold change < -1, q value < 0.05) genes were selected. Volcano maps of differentially expressed genes between specific immune subgroups and others were drawn by R package ggVolcano (v.0.0.2). The relationship between the expression levels of these genes and the overall survival were further processed with univariate Cox regression analysis (HR < 0.95 or > 1.05, Benjamini-Hochberg adjusted *P*-value < 0.05).

In order to find prognostic or functional pathways of immune clusters, 50 gene sets for hallmark pathways were obtained from GSEA Human Molecular Signatures Database (MSigDB) Collections. And GSEs of these pathways were calculated by GSVA, which were then leveraged to perform univariate COX analysis (*P*-values were adjusted by Benjamini-Hochberg correction). Quantitative set analysis for gene expression (QuSAGE) (v.2.36.0)[25] was conducted to illustrate the differential expression of prognostic hallmark pathways among immune clusters.

**Validation for the prognostic significance of S4 gene scores**

To identify immune genes with prognostic significance in NB, up-regulated genes in the S4 cluster with a $log_2$-fold change greater than 0.5 and q values less than 0.05 were utilized to perform univariate COX analysis on NB samples from the TARGET cohort. Forty-four genes (*P*-value < 0.05) were identified and used to construct gene



ARTICLE IN PRESS

ARTICLE IN PRESS

score (S4 gene score). The S4 gene score was created as follows: within each sample, we ranked all immune genes used in our study according to their expression levels. For each of the 44 genes, we calculated its percentile relative to the full list of immune genes. The average of the percentiles for the 44 genes was used as the S4 gene score. These values can be readily compared across samples. We calculated the S4 gene scores for NB samples from both the TARGET and ZJUCH cohorts and performed Kaplan-Meier analyses. Additionally, tumor purity for each sample was calculated using the ESTIMATE method (v.1.0.13), and a scatter plot was generated to visualize the relationship between the S4 gene score and tumor purity. R package ClusterProfiler (v.4.10.0)[26] was utilized to perform Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis for the 44 genes. The formula was as follows: $\text{S4 score} = \frac{1}{|G|}\sum_{g\in G}(1-\frac{\text{rank}_g-1}{\text{N}})$ (g, one of the 44 genes; $\text{rank}_g$, its descending expression rank among all N immune genes; |G| = 44). All figures were optimized for interpretation by individuals with common forms of color-vision deficiency. The underlying numerical data for all figures were provided in the Supplementary Tables to ensure accurate grayscale reproduction.

## Results

### The prognostic immune signatures of PST and PHT

Samples with transcriptome information were utilized to calculate enrichment scores of immune features. TARGET-NBL (n=150), TARGET-OS (n=73), TARGET-WT (n=124), TARGET-RT (n=63) and TARGET-CCSK (n=13) were included to depict the

immune signatures of the PST cohort, while TAEGET-AML (n=1780), TARGET-ALL-P2 (n=440) and TARGET-ALL-P3 (n=82) were included in the PHT cohort (Fig. 2a, b). As *FLT3*-ITD mutation is correlated with poor prognosis of AML and drives stem cell transformation through ERK signaling, PI3 kinase signaling and STAT5 signaling[27], we separated AML into *FLT3*-ITD mutation positive (AML-1) and negative (AML-2) subtype in the following analysis. Firstly, $log_2$(TPM+1) of the transcriptome data were utilized to perform ssGSEA of 184 immune signatures and GSEs were calculated. To filter out prognosis-related immune signatures, univariate Cox regression analysis was performed to find relationship between the GSEs of the 184 immune signatures and overall survival time. Seventy-three immune signatures with *P*-values less than 0.25 in the PST cohort and 115 immune signatures with *P*-values less than 0.05 in the PHT cohort were selected for further studies ( Supplementary Tables S1 and S2). The GSEs of these signatures were subjected to Spearman correlation analysis ( Supplementary Tables S3 and S4), and highly correlated signatures were clustered into immune modules (Fig. 2c, d). Those immune signatures with similar function were chosen to construct immune modules, while some known immune signatures were also utilized individually as immune modules. The selected immune signatures are representative and were chosen to cover the major immune cell types and immune responses. Eight immune modules were selected in each of the PST and PHT cohort ( Tables 1 and 2, Supplementary Tables S5 and S6). The GSEs of these immune modules were calculated through ssGSEA, which were used to classify samples into 5 and 4 immune clusters for the PST and PHT cohort, respectively (Fig. 2e, f). The univariate Cox

ARTICLE IN PRESS

ARTICLE IN PRESS

regression results for the immune modules, including the HR, *P*-values, and Benjamini-Hochberg adjusted *P*-values, are provided by Supplementary Tables S7 and S8.

In the PST cohort, cluster S1 (n=86, 20.3%) was featured as abundance of Wound_CSR (fibroblast core serum response in wound healing) and B cell signatures while scoring less for neutrophils, NK cell and angiogenesis. For cluster S2 (n=44, 10.4%), Wound_CSR, angiogenesis and TGFβ signatures were enriched. Cluster S3 (n = 86, 20.3%) had similar immune characteristics to S2 but with lower angiogenesis scores. Clusters S4 (n=116, 27.4%) and S5 (n=91, 21.5%) have different scores of B cell, T cell and angiogenesis, while the scores of other modules were similar (Fig. 2e). The clusters S1 and S4 exhibited completely different expressions of immune signatures, which was consistent with the later depicted different overall survival (Fig. 3c). Tumor cases were unevenly distributed across the 5 clusters: S1 contained 14% NB (n=21), 60.3% RT (n=38), and 21% WT (n=26); S2 and S3 were primarily composed of WT cases (S2: 26.6%, n=33; S3: 44.4%, n=55). S4 included all CCSK, 71.2% OS (n=52), and 28.7% NB (n=43), while 48% of NB cases (n=72) were in S5 (Fig. 2g).

For the PHT cohort, cluster H1 owned higher expressions of neutrophils, monocytes and TGFβ while harboring lower infiltration of antigen processing signature. Cluster H2 was featured as up-regulated levels of neutrophils, monocytes and antigen processing signatures. Cluster H3 scored highly for antigen processing and interferon signatures, but the infiltration of neutrophils and monocytes was down-regulated. Except for T cell and TGFβ, cluster H4 owned lower scores of all other immune modules (Fig. 2f).

ARTICLE IN PRESS

Interestingly, the levels of all immune features showed remarkable difference between the H2 and H4 clusters, and these two clusters showed distinct overall survival in pan-cancer or type-specific analysis of the PHT cohort (Fig. 3d). While the cases of ALL-P2 were mainly distributed in the H3 (B-precusor cell dominant) and H4 (T-cell dominant) clusters, the tumor cases in each of other types were all distributed in the 4 clusters (Fig. 2h).

The associations between tumor subtypes and immune clusters were also studied (Fig. 2i). *MYCN*-amplified NB had a higher proportion of cluster S1 (*P*-value < 0.001) and lower proportion of S5 (*P*-value < 0.001) compared to *MYCN* non-amplified NB. For WT, DAWT (diffuse anaplastic Wilms tumor) showed a higher proportion of S1 cases (*P*-value = 0.011) and fewer S2 (*P*-value = 0.018) and S4 (*P*-value = 0.05) cases than FHWT (favorable histology Wilms tumor). In ALL-P2, 95.3% of B-precursor and 90.9% of T-cell ALL cases were clustered in H3 and H4, respectively, with all *ETV6-RUNX1* fusion cases in H3. For AML-1, 50% of *CEBPA* mutation cases and 80% of *DEK-NUP214* fusion cases were in H3 and H1, respectively, while 63.6% of *NUP98-NSD1* fusion and 64.5% of *WT1* mutation cases were in H2. For AML-2, 97.8% of *CBFB-MYH11* fusion cases and 62.5% of *KMT2A* fusion cases were in H1/H2 and H2, respectively.

ARTICLE IN PRESS

**Immune clusters exhibit distinct overall survival both in PST and PHT**

We next sought to investigate the clinical features of patients in each of the immune clusters. Most of the RT cases with younger age of diagnosis belonged to cluster S1

and the majority of OS cases with elder age of diagnosis were classified as cluster S4 (Figs. 2g and Supplementary Fig. S1a), which resulted in significantly difference of the age (two-tailed *t*-test, *P*-value = 0.00021) between clusters S1 and S4 in the PST cohort (Fig. 3a). In the PHT cohort, AML-1 cases exhibited elder age of diagnosis than each of the other types of hematological tumors (Kruskal-Wallis test, *P*-value < 2.2e-16) (Supplementary Fig. S1b), while significantly older cases were found in cluster H1 than others (Kruskal-Wallis test, *P*-value = 8e-6) (Fig. 3b). However, no associations were identified between immune clusters and gender or race in both PST and PHT cohorts (Fig. 3a, b). Multivariate COX analysis showed immune clusters were prognosis indicators independent of age, gender and race in both PST and PHT cohorts (Supplementary Tables S9 and S10).

Survival analysis was performed to elucidate the associations between immune clusters and patient outcomes. In the PST cohort, S1 and S5 behaved as poor prognostic features while clusters S2 and S4 exhibited relatively prolonged survival time (log-rank test, *P*-value < 0.001) (Fig. 3c). To validate if the prognostic values of immune clusters still remained within tumor entities, we analyzed the overall survival of NB, OS, WT and RT across immune clusters (Fig. 3c). The analysis of CCSK was not available because all the CCSK cases belonged to cluster S4. In NB, clusters S1 and S3 exhibited poor overall survival while significant difference was found in survival of S4 (log-rank test, *P*-value = 0.008). For WT, S1 was still in the worst situation, while the survival rates of S2 and S3 were similar as shown in the whole PST cohort (log-rank test, p = 0.08). The difference of survival rate was not significant in RT (*P*-value =

ARTICLE IN PRESS

ARTICLE IN PRESS

0.67) and OS (*P*-value = 0.21), possibly due to the small sample size and uneven distribution of these cases in immune clusters.

For the PHT cohort (Fig. 3d), cluster H4 possessed highest survival rate than other clusters (log-rank test, *P*-value = 0.002). In ALL-P2, H3 and H4 owned completely different survival rate (log-rank test, *P*-value < 0.001). In this study, we divided AML into FLT-ITD mutation positive and negative subtype, and found they had completely different survival curves of immune clusters. Cluster H2 had the lowest survival rate in AML-1 (log-rank test, *P*-value = 0.015). While for AML-2, H2 had the highest survival rate, and H4 owned the lowest survival rate (log-rank test, *P*-value = 0.019).

Considering that tumors with small sample sizes, including ALL-P3, CCSK, RT, and OS, could potentially skew the overall survival curves, we performed additional analyses excluding these tumors. In the PST cohort, survival curves were recalculated after excluding RT (Supplementary Fig. S2a), CCSK (Supplementary Fig. S2b), and OS (Supplementary Fig. S2c), yielding *P*-values of 0.002, <0.001, and <0.001, respectively. In the PHT cohort, excluding ALL-P3 resulted in a *P*-value of 0.001(Supplementary Fig. S2d). Importantly, the overall trends of survival across different immune clusters remained consistent after excluding these small-sample tumors, indicating that their impact on our study conclusions was minimal.

ARTICLE IN PRESS

**Differentially expressed genes and pathways of immune clusters**

To figure out the underlying functional genes of immune pathways, DESeq2 was utilized to find out the differential gene expressions between specific subgroup and

ARTICLE IN PRESS

others. In the PST cohort, we found top 10 immune-signature-sepecifically up-rugulated genes (based on $log_2$-fold change) including genes encoding ribosome proteins (*RPL29*, *RPS14*, *RPL39*, *RPS15* and *RPS29*) in cluster S1, TGFβ (*COL1A2*, *MMP14* and *ITGB3*) in cluster S4 and B cell (*CR2* and immunogloblin genes) in cluster S5. On the contrary, immunogloblin genes (*IGKJ4*, *IGHJ4*, *IGHJ6*, *IGHV7-4-1* and *IGHJ5*) were significantly down-regulated ($log_2$-fold change < -5) in clusters S2 and S3 (Fig. 4a). Additionally, genes significantly up-regulated ($log_2$-fold change > 1, q value < 0.05) in cluster S4 were enriched for angiogenesis, NK cells and neutrophils, which may contribute to the better survival of cluster S4 (Figs. 2e and 3c and Supplementary Table S11).

For the PHT cohort (Fig. 4b), we found monocytes or neutrophils genes were enriched in H1 (*CLC*, *MNDA*, *TREM2*, *TNFAIP6*, *S100A12*, *HK3*, *CLEC10A*, *S100A9* and *CHI3L1*) and H2 (*MARCO*, *C5AR1*, *AQP9*, *FPR1*, *CXCR1*, *S100A12* and *VNN3*) clusters, while genes *CLEC10A*, *MARCO*, *S100A9* and *S100A12* were down-regulated in H3 and H4 clusters. And possibly because of B-precursor dominant ALL-P2 being mostly distributed in H3 cluster (Fig. 2i), B cell genes (*TCL1A*, *VPREB3*, *EBF1*, *PAX5*, *CD19*, *BLNK*, *CD79A*, *CORO2B*, *CD22* and *POU2AF1*) were highly expressed in H3 cluster.

In order to obtain immune-related genes with potentail prognostic values, we performed univariate COX analysis on differentialy expressed genes ($log_2$-fold change > 1 or < -1, q value < 0.05), and 11 and 9 genes (HR > 1.05 or < 0.95, adjusted *P*-value < 0.05) were acquired in the PST (Fig. 5a) and PHT (Fig. 5b) cohorts, repectively.

ARTICLE IN PRESS

ARTICLE IN PRESS

Additionally, genes with HR > 1 or < 1 and *P*-value < 0.05 are provided in Supplementary Tables S12 and S13 for further reference. In the PST cohort, the neutrophil-related gene *ALPK1* and the angiogenesis genes (*PTPRB*, *PALMD* and *HECW2*) were up-regulated in cluster S4, while the T cell gene *KLRK1* and the B cell genes (*IGLV5-48*, *IGKV3-7*, *IGHV3-35* and *PIK3CG*) showed elevated expression in S5. In the PHT cohort, the monocyte marker *CD68* was up-regulated in clusters H1 and H2, whereas the monocyte gene *CCL13*, NK-cell gene *SHC4*, TGFβ gene *ITGA3* and T-cell genes *RTKN2*, *LIME1*, *SPEG* and *JAKMIP1* were all enriched in cluster H4. The statistical significance of these differential expression results, including *P*-values and Benjamini-Hochberg adjusted *P*-values, is detailed in Supplementary Tables S14 and S15.

ARTICLE IN PRESS

**Heterogeneity of hallmark pathways in immune clusters**

The hallmark pathways are part of pathway modules in GSEA, which contain important and well-defined oncology pathways. The GSEs of these hallmark pathways were calculated and utilized to perform univariate COX regression analysis, and 18 and 31 prognostic pathways (univariate COX analysis, *P*-value < 0.05) were acquired in PST (Fig. 5c, Supplementary Table S16) and PHT (Fig. 5d, Supplementary Table S17), respectively. In the PST cohort, the prognostic hallmark pathways showed distinct expression among clusters, especially between clusters S1 and S4, which was consistent with the overall survival of these clusters (Figs. 3c and 5c). For the PHT cohort, cluster H2 and H4 showed distinct expressions of hallmark pathways.

ARTICLE IN PRESS

**NB samples with higher S4 gene scores exhibited prolonged survival time**

As the S4 cluster possessed favorable prognosis in the PST cohort, especially for NB, up-regulated genes in the S4 cluster ($log_2$-fold change > 0.5, q value < 0.05) were utilized to perform univariate COX analysis on NB samples from the TARGET cohort. Forty-four genes (*P*-value < 0.05) were identified and used to construct S4 gene score ( Supplementary Table S18). We first calculated the S4 gene scores for the neuroblastoma samples in all of the clusters from the TARGET cohort (Fig. 6a). The S4 cluster exhibited the highest score across all clusters, demonstrating that the S4 gene scores effectively represent the S4 cluster. Before calculating the S4 scores for the ZJUCH-NB cohort, we performed batch effect correction on the ZJUCH-NB expression data, using TARGET-NB as the reference cohort to eliminate any potential technical bias (see PCA plots before and after correction in Supplementary Fig. S3a, b). Following this, we found that high S4 gene scores were correlated with prolonged survival time for NB in both the TARGET and ZJUCH cohorts (*P*-value < 0.0001 and *P*-value = 0.0014) (Fig. 6b, c). GO and KEGG analyses showed that the genes used in the S4 gene score were mostly enriched in membrane signaling and oncology pathways (Fig. 6d, e), consistent with the properties of immune genes in neuroblastoma. Specifically, the S4 gene scores were mainly enriched in the KEGG pathways of “Aldosterone-regulated sodium reabsorption,” “Thyroid hormone signaling pathway,” “Carbohydrate digestion and absorption,” “FoxO signaling pathway,” “PI3K-Akt signaling pathway,” and “Phospholipase D signaling pathway”[28-30].



ARTICLE IN PRESS

ARTICLE IN PRESS

Tumor purity is known to potentially influence tumor immune infiltration. To investigate this, we calculated tumor purity for samples from both the TARGET database and ZJUCH cohort (Supplementary Tables S19). The cohorts with the highest tumor purity (>0.7) include TARGET CCSK (100%), TARGET ALL-P2 (83.6%), TARGET RT (82.5%), and TARGET NB (78.7%). Other cohorts, including TARGET AML, TARGET OS, TARGET ALL-P3 and ZJUCH NB, exhibited moderate tumor purity. Although tumor purity of ZJUCH NB was generally lower compared to the TARGET cohort, we found that the relationship between the calculated S4 gene scores and tumor purity was minimal (Supplementary Fig. S4), indicating that tumor purity had little impact on our analyses.

ARTICLE IN PRESS

**Discussion**

By utilizing immune signatures collected in a previous study[19], we described immune signatures in pediatric solid and hematological tumors from the TARGET cohort, and divided them into immune clusters with distinctly differential expression of immune cell infiltrating features. These immune clusters showed prognostic significance in pan-cancer and within some specific cancer types. In the PST cohort, Neutrophils and B cell behaved as the most favorable and poor prognosis indicators both in NB and WT samples. For NB, clusters S4 and S5 with highly expressed neutrophils signature owned prolonged survival time, aligning with previous findings indicating favorable outcomes in myeloid-predominant pediatric nervous system tumors[31]. In our work, WT samples in cluster S1 with abundance of B cell were in a worse situation compared to

ARTICLE IN PRESS

clusters S2 and S3, based on the transcriptome data. However, a study based on DNA methylation data stated that B cell enrichment was higher in the low-risk group than in the high-risk group of WT[32], possibly due to the tumor heterogeneity and other immune factors involved. In the PHT cohort, ALL-P2 samples were mainly divided into clusters H3 and H4. H3 was mainly consisted of B cell precursor ALL cases, while H4 was mainly consisted of T cell ALL cases. In contrast, AML samples were distributed in all the four groups. For AML, cases in H2 and H4 clusters, with (AML-1) or without (AML-2) *FLT3*-ITD mutation, exhibited completely different survival curves. This disparity may be attributed to the roles of *FLT3*-ITD mutation in the JAK/STAT signaling pathway, which acts as an important factor in inflammation response [27]. However, the detailed associations between *FLT3*-ITD mutation and immune cell infiltrating still need further investigation. Future work should experimentally validate the impact of FLT3-ITD mutations on the immune microenvironment in hematological cancers.

ARTICLE IN PRESS

To dig out the underlying mechanism contributing to the prognostic significance of immune clusters, we analyzed differentially expressed genes of immune modules and hallmark pathways across immune clusters. A series of immune genes were identified, and the expression levels and prognostic significance of these genes were consistent with relationship between immune infiltrating cell signatures and survival rate across immune clusters. Upregulated expression of *PIK3CG* is associated with a positive outcome of lung adenocarcinoma[33], but a poor outcome of immune-inflamed colorectal cancer[34]. Our data showed the prolonged survival was related to high expression of *PIK3CG* in the PST cohort. In the PST cohort, we also observed that immunoglobulin

ARTICLE IN PRESS

genes were highly expressed in the B cell dominant cluster S5, providing valuable information for the development of new therapies focused on these tumors. In addition, the gene scores of 44 up-regulated genes in the S4 cluster (S4 gene score) were correlated with favorable prognosis of NB in the ZJUCH cohort, and this result validated that our results of prognosis grouping were reasonable and highly reproducible.

For translational applications, our findings suggest that immune clusters could serve as potential biomarkers for prognosis in pediatric tumors. In particular, clusters enriched in immune signatures like neutrophils or B cells could guide treatment decisions or stratification in clinical trials. One potential next step is to validate the utility of immune cluster classification in prospective clinical trials, evaluating whether these clusters can predict treatment response or identify high-risk patients in need of more aggressive therapies.

ARTICLE IN PRESS

This study contributes to advancing several Sustainable Development Goals (SDGs), particularly those related to health and well-being, by improving cancer treatment outcomes and fostering the development of precision medicine. By identifying immune-related prognostic factors, we aim to enhance survival rates and quality of life for children with these challenging diseases, aligning with the global vision for equitable healthcare and better access to effective treatments. Future work will focus on refining these immune profiles through more detailed molecular approaches, including single-cell RNA sequencing, spatial transcriptomics, and proteomics. Additionally, the clinical applicability of these profiles should be tested across diverse patient cohorts to assess

ARTICLE IN PRESS

their relevance and effectiveness.

There are still some limitations in our study. For example, the sample sizes for CCSK, OS, RT, and ALL-P3 were relatively small and unevenly distributed across the immune clusters, which hindered the construction of a robust immune profile. Specifically, 60.3% of RT cases were concentrated in S1, 71.2% of OS cases in S4, and all CCSK samples were assigned to S4, resulting in insufficient representation in the other clusters for reliable comparisons. Another concern is that the analysis is based on bulk RNA-seq, which may introduce potential bias and lead to overestimation of gene expression levels in different immune clusters. Moreover, our study was conducted within the tumor patient populations and did not include comparisons with normal tissues. This limits our ability to delineate immune features that are uniquely altered in the tumor microenvironment versus those that may also be present in normal tissue. Further studies should include more patients with tumors and diverse racial backgrounds, especially those with small sample sizes, to depict immune features in more detail and breadth.

ARTICLE IN PRESS

## Conclusion

In summary, this study addressed key research questions regarding the immune infiltration patterns in pediatric solid and hematological tumors and their prognostic significance. We classified PST and PHT into distinct immune subtypes, revealing immune clusters that correlate with different overall survival rates. Specifically, clusters S2 and S4 of the PST cohort, as well as clusters H2 and H4 of the PHT cohort,

ARTICLE IN PRESS

exhibited unique immune features and survival outcomes, answering our first research question by identifying immune subtypes associated with patient prognosis.

Furthermore, we identified immune-related signaling pathways and genes that are prognostically relevant to pediatric cancers. In particular, 44 upregulated genes in the S4 cluster of the PST cohort were associated with a favorable prognosis in neuroblastoma samples, which addresses our second research question. These findings highlight potential immune biomarkers that could serve as prognostic indicators for pediatric cancer patients.

In conclusion, our study provides a comprehensive understanding of the immune landscape in pediatric tumors, identifies immune features that are significant for prognosis, and proposes potential biomarkers for future therapeutic development. These findings could inform the design of more precise immunotherapy strategies to improve clinical outcomes for pediatric cancer patients.

ARTICLE IN PRESS

**Data availability**

High-throughput sequencing and clinical data are acquired from TARGET database through Genomic Data Commons (GDC) data portal (https://gdc.cancer.gov/). The neuroblastoma validation cohort (refer to as ZJUCH cohort) were obtained from Children's Hospital Zhejiang University School of Medicine. The RNA-seq data of the ZJUCH cohort were available from the Genome Sequence Archive in National Genomics Data Center, China National Center for Bioinformation / Beijing Institute of Genomics, Chinese Academy of Sciences under accession numbers HRA006359 and

ARTICLE IN PRESS

HRA007828.

**Abbreviations**

| | |
|---|---|
| PST | Pediatric solid tumor |
| PHT | Pediatric hematological tumor |
| NB | Neuroblastoma |
| AML | Acute myeloid leukemia |
| TIME | Tumor immune microenvironment |
| ALL | Acute lymphoblastic leukemia |
| OS | Osteosarcoma |
| WT | Wilms tumor |
| RT | Rhabdoid tumor |
| CCSK | Clear cell sarcoma of kidney |
| dbGaP | Database of Genotypes and Phenotypes |
| NIH | National Institutes of Health |
| TPM | Transcripts per kilobase of exon per million mapped reads |
| FPKM | Fragments per kilobase of exon per million mapped fragments |
| ssGSEA | Single-sample gene set enrichment analysis |
| GSEs | Gene set enrichment scores |
| QuSAGE | Quantitative set analysis for gene expression |
| GO | Gene Ontology |
| KEGG | Kyoto Encyclopedia of Genes and Genomes |
| DAWT | Diffuse anaplastic Wilms tumor |
| FHWT | Favorable histology Wilms tumor |

ARTICLE IN PRESS

ARTICLE IN PRESS

## References


1. Steliarova-Foucher, E. *et al.* International incidence of childhood cancer, 2001-10: a population-based registry study. *Lancet Oncol* **18**, 719-731 (2017). https://doi.org/10.1016/S1470-2045(17)30186-9
2. Collaborators, G. B. D. C. C. The global burden of childhood and adolescent cancer in 2017: an analysis of the Global Burden of Disease Study 2017. *Lancet Oncol* **20**, 1211-1225 (2019). https://doi.org/10.1016/S1470-2045(19)30339-0
3. Lam, C. G., Howard, S. C., Bouffet, E. & Pritchard-Jones, K. Science and health for all children with cancer. *Science* **363**, 1182-1186 (2019). https://doi.org/10.1126/science.aaw4892
4. (2021), W. H. O. *CureAll framework: WHO global initiative for childhood cancer: increasing access, advancing quality, saving lives*. https://apps.who.int/iris/handle/10665/347370
5. Irwin, M. S. *et al.* Revised Neuroblastoma Risk Classification System: A Report From the Children's Oncology Group. *J Clin Oncol* **39**, 3229-3241 (2021). https://doi.org/10.1200/JCO.21.00278
6. Papaemmanuil, E. *et al.* Genomic Classification and Prognosis in Acute Myeloid Leukemia. *N Engl J Med* **374**, 2209-2221 (2016). https://doi.org/10.1056/NEJMoa1516192
7. Dharia, N. V. *et al.* A first-generation pediatric cancer dependency map. *Nat Genet* **53**, 529-538 (2021). https://doi.org/10.1038/s41588-021-00819-w
8. Laetsch, T. W., DuBois, S. G., Bender, J. G., Macy, M. E. & Moreno, L. Opportunities and Challenges in Drug Development for Pediatric Cancers. *Cancer Discov* **11**, 545-559 (2021). https://doi.org/10.1158/2159-8290.CD-20-0779
9. Bagchi, S., Yuan, R. & Engleman, E. G. Immune Checkpoint Inhibitors for the Treatment of Cancer: Clinical Impact and Mechanisms of Response and Resistance. *Annu Rev Pathol* **16**, 223-249 (2021). https://doi.org/10.1146/annurev-pathol-042020-042741
10. Bedard, P. L., Hyman, D. M., Davids, M. S. & Siu, L. L. Small molecules, big impact: 20 years of targeted therapy in oncology. *Lancet* **395**, 1078-1088 (2020). https://doi.org/10.1016/S0140-6736(20)30164-1
11. Butler, E. *et al.* Recent progress in the treatment of cancer in children. *CA Cancer J Clin* **71**, 315-332 (2021). https://doi.org/10.3322/caac.21665
12. Binnewies, M. *et al.* Understanding the tumor immune microenvironment (TIME) for effective therapy. *Nat Med* **24**, 541-550 (2018). https://doi.org/10.1038/s41591-018-0014-x
13. Lei, X. *et al.* Immune cells within the tumor microenvironment: Biological functions and roles in cancer immunotherapy. *Cancer Lett* **470**, 126-133 (2020). https://doi.org/10.1016/j.canlet.2019.11.009
14. Sherif, S. *et al.* The immune landscape of solid pediatric tumors. *J Exp Clin Cancer Res* **41**, 199 (2022). https://doi.org/10.1186/s13046-022-02397-z
15. Sengupta, S. *et al.* Mesenchymal and adrenergic cell lineage states in neuroblastoma possess distinct immunogenic phenotypes. *Nat Cancer* **3**, 1228-1246 (2022). https://doi.org/10.1038/s43018-022-00427-5
16. Yu, A. L. *et al.* Anti-GD2 antibody with GM-CSF, interleukin-2, and isotretinoin for neuroblastoma. *N Engl J Med* **363**, 1324-1334 (2010).


ARTICLE IN PRESS

https://doi.org/10.1056/NEJMoa0911123

17. Gardner, R. A. *et al.* Intent-to-treat leukemia remission by CD19 CAR T cells of defined formulation and dose in children and young adults. *Blood* **129**, 3322-3331 (2017). https://doi.org/10.1182/blood-2017-02-769208
18. Maude, S. L. *et al.* Tisagenlecleucel in Children and Young Adults with B-Cell Lymphoblastic Leukemia. *N Engl J Med* **378**, 439-448 (2018). https://doi/org/10.1056/NEJMoa1709866
19. Wang, S. *et al.* Clinical significance and immunogenomic landscape analyses of the immune cell signature based prognostic model for patients with breast cancer. *Brief Bioinform* **22** (2021). https://doi.org/10.1093/bib/bbaa311
20. Wong, K. M. *et al.* The dbGaP data browser: a new tool for browsing dbGaP controlled-access genomic data. *Nucleic Acids Res* **45**, D819-D826 (2017). https://doi.org/10.1093/nar/gkw1139
21. Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15-21 (2013). https://doi.org/10.1093/bioinformatics/bts635
22. Liao, Y., Smyth, G. K. & Shi, W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923-930 (2014). https://doi.org/10.1093/bioinformatics/btt656
23. Hanzelmann, S., Castelo, R. & Guinney, J. GSVA: gene set variation analysis for microarray and RNA-seq data. *BMC Bioinformatics* **14**, 7 (2013). https://doi.org/10.1186/1471-2105-14-7
24. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**, 550 (2014). https://doi.org/10.1186/s13059-014-0550-8
25. Meng, H., Yaari, G., Bolen, C. R., Avey, S. & Kleinstein, S. H. Gene set meta-analysis with Quantitative Set Analysis for Gene Expression (QuSAGE). *PLoS Comput Biol* **15**, e1006899 (2019). https://doi.org/10.1371/journal.pcbi.1006899
26. Yu, G., Wang, L. G., Han, Y. & He, Q. Y. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* **16**, 284-287 (2012). https://doi.org/10.1089/omi.2011.0118
27. Burchert, A. Maintenance therapy for FLT3-ITD-mutated acute myeloid leukemia. *Haematologica* **106**, 664-670 (2021). https://doi.org/10.3324/haematol.2019.240747
28. Kanehisa, M., Furumichi, M., Sato, Y., Matsuura, Y. & Ishiguro-Watanabe, M. KEGG: biological systems database as a model of the real world. *Nucleic Acids Res* **53**, D672-D677 (2025). https://doi.org/10.1093/nar/gkae909
29. Kanehisa, M. Toward understanding the origin and evolution of cellular organisms. *Protein Sci* **28**, 1947-1951 (2019). https://doi.org/10.1002/pro.3715
30. Kanehisa, M. & Goto, S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res* **28**, 27-30 (2000). https://doi.org/10.1093/nar/28.1.27
31. Nabbi, A. *et al.* Transcriptional immunogenomic analysis reveals distinct immunological clusters in paediatric nervous system tumours. *Genome Med* **15**, 67 (2023). https://doi.org/10.1186/s13073-023-01219-x
32. Tang, F. *et al.* DNA Methylation Data-Based Classification and Identification of Prognostic Signature of Children With Wilms Tumor. *Front Cell Dev Biol* **9**, 683242 (2021). https://doi.org/10.3389/fcell.2021.683242

ARTICLE IN PRESS

ARTICLE IN PRESS

33. Liu, Y. *et al.* A novel anoikis-related prognostic signature associated with prognosis and immune infiltration landscape in lung adenocarcinoma. *J Gene Med* **26**, e3610 (2024). https://doi.org/10.1002/jgm.3610
34. Liu, Y. *et al.* Exploring the potential mechanisms of Yi-Yi-Fu-Zi-Bai-Jiang-San therapy on the immune-inflamed phenotype of colorectal cancer via combined network pharmacology and bioinformatics analyses. *Comput Biol Med* **166**, 107432 (2023). https://doi.org/10.1016/j.compbiomed.2023.107432

## Author Contributions

T.T., S.N. and J.W. provided resources and obtained funding. T.T., S.N. and Q.H. conceived and supervised the study. T.T., Z.X. and J.Q. generated the data. T.T., Z.X., Q.H. and J.L. performed analysis. T.T. and Z.X. interpreted the data and wrote the main manuscript text. Z.X. prepared figures and tables. All authors reviewed and approved the final manuscript.

ARTICLE IN PRESS

## Funding

This work was supported by grants (No. 82373971 and No. 32270853) from National Natural Science Foundation of China, a National Key R&D Program of China (2023YFC2706100), grants (No. 2025C01106 and 2024C03181) from "Pioneer" and "Leading Goose" R&D Program of Zhejiang Province, grants (NO. 2022KY1047 and NO. 2024KY1428) from Medical Science and Technology Project of Zhejiang Province, a startup fund from Children's Hospital, Zhejiang University School of Medicine, and a research fund from Cancer Center, Zhejiang University.

## Declarations

ARTICLE IN PRESS

## Competing interests

The authors declare no competing interests.

## Figure legends

**Fig. 1.** Flowchart for the bioinformatic pipeline: A total of 127 neuroblastoma samples from the ZJUCH cohort were subjected to RNA-seq analysis, including RNA extraction, library preparation and sequencing, RNA-seq data alignment, duplicate marking, and gene counting. Additionally, 2,725 childhood cancer samples were collected: TARGET-NBL (n = 150), TARGET-OS (n = 73), TARGET-WT (n = 124), TARGET-RT (n = 63), and TARGET-CCSK (n = 13) were included to characterize the immune signatures of the PST cohort, while TARGET-AML (n = 1,780), TARGET-ALL-P2 (n = 440), and TARGET-ALL-P3 (n = 82) were included in the PHT cohort. Immune subtypes were constructed based on the distribution of immune signatures. Differentially expressed genes and pathways were then identified across these immune subtypes. Furthermore, up-regulated genes in the S4 cluster were used to construct a prognostic S4 gene score, which was subsequently validated in the ZJUCH cohort.

ARTICLE IN PRESS

**Fig. 2.** Immune clusters of 423 pediatric solid tumors (PST) and 2302 pediatric hematological tumors (PHT): **a, b** Barplots showing sample composition in the PST and PHT cohorts. CCSK: clear cell sarcoma of kidney; NB: neuroblastoma; OS: osteosarcoma; RT: rhabdoid tumor; WT: Wilms tumor; ALL: acute lymphoblastic leukemia; AML: acute myelogenous leukemia. **c, d** Heatmap acquired from Corrplot revealing spearman's correlation of 73 and 115 prognostic immune signatures in the

ARTICLE IN PRESS

PST (**c**) or PHT cohort (**d**). Black rectangles containing highly correlated signatures were utilized to construct immune modules. Wound_CSR: fibroblast core serum response in wound healing. **e, f** Enrichment heatmaps, forest plots and bubble plots of selected immune modules in the PST (**e**) and PHT cohorts (**f**). Enrichment heatmaps cluster patients into 5 (S1-S5) and 4 (H1-H4) immune clusters in the PST and PHT cohort, respectively. Forest plots show hazard ratios (HRs) of immune modules in the whole PST or PHT cohort. Bubble plots reveal HRs and Benjamini-Hochberg adjusted *P*-values in two cohorts and within specific tumors, the color of the circles depicts the HR and the size of the circles represents the $-\log_{10}(P\text{-value})$. **g, h** Stacked barplots showing the distribution of the tumor samples in immune clusters in the PST (**g**) and PHT cohort (**h**). **I** Stacked barplots revealing associations of tumor subtypes with immune clusters. *MYCN* amp: *MYCN*-amplified neuroblastoma; *MYCN* non-amp: *MYCN* non-amplified neuroblastoma; DAWT: diffuse anaplastic Wilms tumor; FHWT: favorable histology Wilms tumor. The most commonly *KMT2A* fusion includes *KMT2A-ELL*, *KMT2A-MLLT10*, *KMT2A-MLLT3* and *KMT2A-MLLT4*.

ARTICLE IN PRESS

**Fig. 3.** Clinical parameters and overall survival curves of immune clusters in PST and PHT: **a, b** Associations of clinical parameters with immune clusters in the PST (**a**) and PHT (**b**) cohorts. Boxplots reveal the distribution of diagnostic age among immune clusters. The boxes show the median (represented by a horizontal line) and interquartile range (IQR), while the whiskers indicate the largest and smallest values within 1.5 times the IQR. The pairwise comparisons between immune clusters were

ARTICLE IN PRESS

done by the two-tailed unpaired *t*-test, and Kruskal-Wallis test was used for all the groups. Stacked barplots showing fractions of gender and race in immune clusters. **c, d** Survival curves of immune clusters in the PST (**c**) and PHT (**d**) cohorts. AML-1: AML with *FLT3*-ITD mutation; AML-2: AML without *FLT3*-ITD mutation. Log-rank test *P*-values are denoted.

**Fig. 4.** Enrichment analysis of immune genes by DEseq2: **a, b** Volcano plots showing differential expression of immune genes in each immune cluster compared to others in the PST (**a**) or PHT cohort (**b**). Dashed lines represent the thresholds of $\log_2$-fold change ($\log_2$FC >1 or < -1) and q value (q < 0.05). Top 10 down-regulated and up-regulated immune genes are annotated. Black rectangle represents differentially expressed genes with *P*-values less than 1e-250.

ARTICLE IN PRESS

**Fig. 5.** Differentially expressed immune genes and hallmark pathways across immune clusters: **a, b** Differentially expressed immune genes acquired from DESeq2 with hazard ratios (HRs) less than 0.95 or more than 1.05 and Benjamini-Hochberg adjusted *P*-values less than 0.05 in the PST (**a**) and PHT cohorts (**b**). Heatmap depicting $\log_2$-fold change ($\text{Log}_2$FC) of genes in each immune cluster compared to others. Forest plots revealing the HRs and *P*-values in the whole PST or PHT cohort. **c, d** Heatmaps, forest plots and bubble plots of differentially expressed and prognostic hallmark pathways. Heatmaps revealing $\log_2$-fold change ($\text{Log}_2$FC) of hallmark pathways across immune clusters. Forest plots showing HRs in the whole

ARTICLE IN PRESS

PST (**c**) or PHT (**d**) cohort. Bubble plots showing HRs and Benjamini-Hochberg adjusted *P*-values in two cohorts and within specific cancer types.

**Fig. 6.** Construction of immune gene scores to evaluate prognosis of neuroblastoma: **a** The S4 gene scores of neuroblastoma samples in all of the clusters from the TARGET cohort. **b, c** The relationship between S4 gene scores and prognosis of neuroblastoma samples in the TARGET cohort (**a**) and ZJUCH cohort(**b**). D, E Gene Ontology (GO) analysis (**d**) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis (**e**) of 44 genes utilized to construct S4 gene score.

ARTICLE IN PRESS

**Table 1 Selected immune modules for the PST cohort.**

| Immune modules | Immune signatures | HR | CI (95%) | *P-value* | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | | PST | NB | OS | WT | RT |
| B cell | IGG_Cluster_21214954, IGG_Cluster, Bcell_mg_IGJ, Plasma_cells_ImSig, BCRSignalingPathway, TNBC_B_Cell, B_Cell_60gene, Bcell_21978456, IgG_19272155 | 2.06 | 1.39-3.06 | <0.01 | 0.02 | 0.24 | 0.01 | 0.72 |
| Neutrophils | Neutrophils_ImSig | 0.38 | 0.21-0.71 | <0.01 | <0.01 | 0.05 | 0.01 | 0.18 |
| Wound_CSR | CHANG_CORE_SERUM_RESPONSE_UP, LYMPHS_PCA_16704732, Translation_ImSig | 1.52 | 1.03-2.25 | 0.03 | 0.09 | 0.76 | 0.65 | 0.93 |
| Dendritic cell | Activated_dendritic_cell | 1.75 | 0.78-3.92 | 0.17 | 0.32 | 0.51 | 0.46 | 0.08 |
| Angiogenesis | Angiogenesis, Endothelial_cells_MCPcounter | 0.63 | 0.43-0.92 | 0.02 | 0.33 | 0.53 | 0.73 | 0.84 |
| NK cell | Natural_killer_cell | 0.26 | 0.11-0.60 | <0.01 | 0.21 | 0.05 | 0.29 | 0.40 |
| TGFb | TGFB_score_21050467, TGFb_Family_Member | 0.54 | 0.33-0.89 | 0.02 | 0.35 | 0.95 | 0.21 | 0.26 |
| T cell | T_cell_PCA_16704732, T_Cell | 0.22 | 0.08-0.64 | <0.01 | 0.03 | 0.70 | 0.66 | 0.19 |

ARTICLE IN PRESS

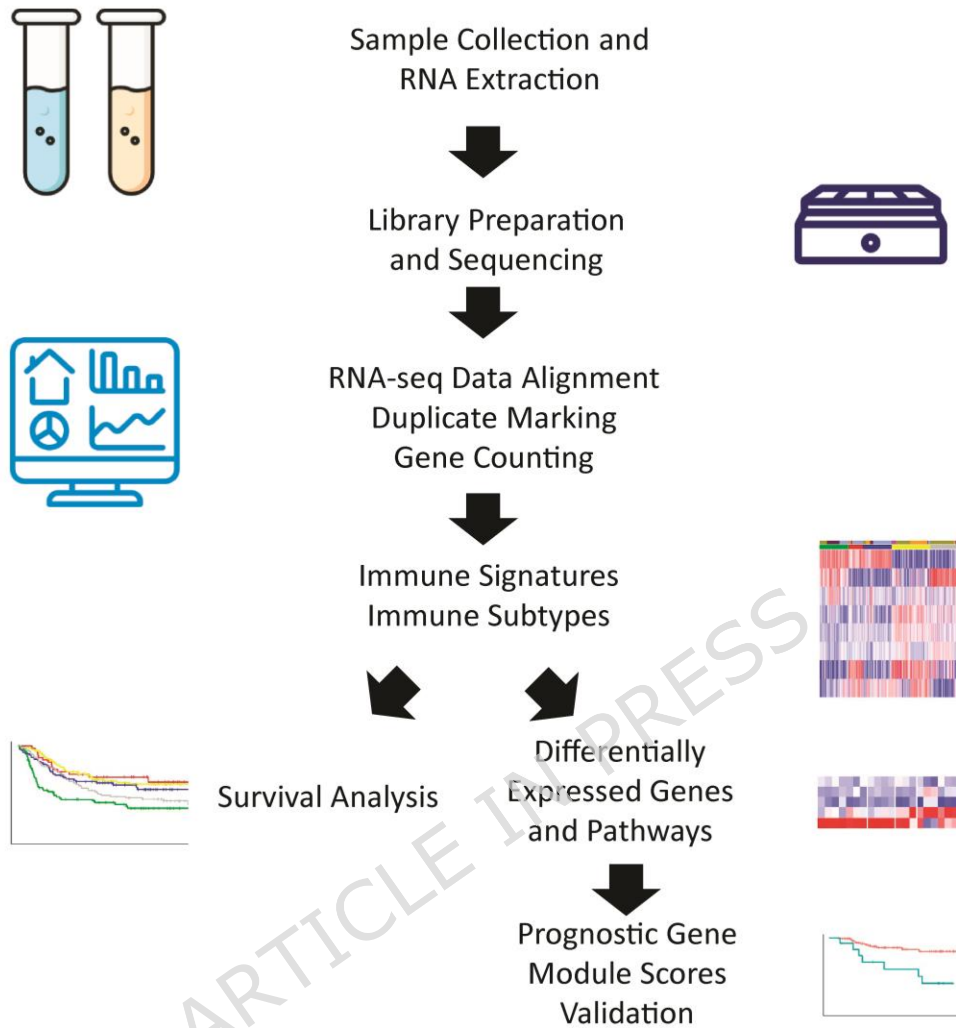
ARTICLE IN PRESS

**Table 2 Selected immune modules for the PHT cohort.**

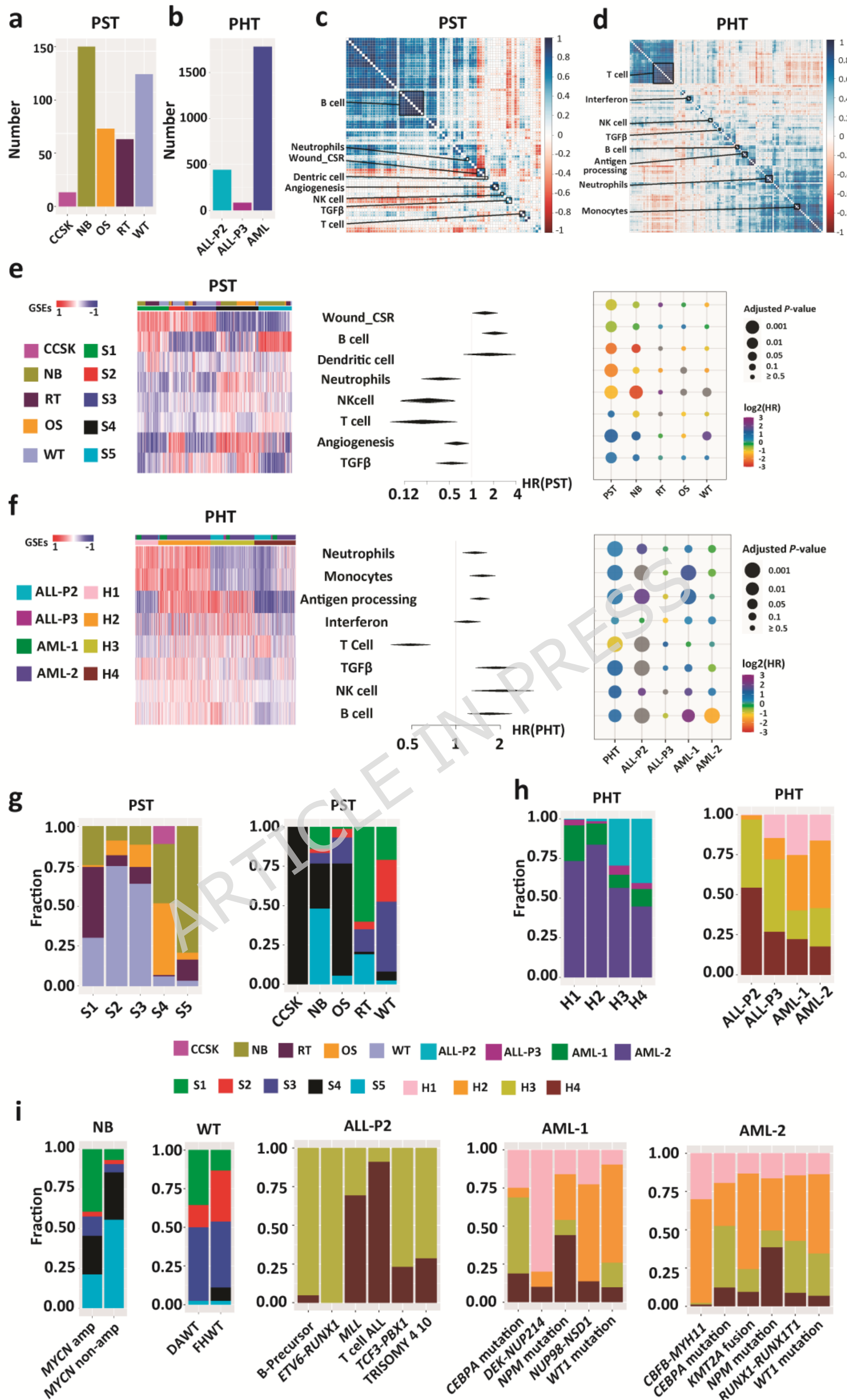
| Immune modules | Immune signatures | HR | CI (95%) | *P-value* | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | | PHT | ALL-P2 | ALL-P3 | AML-1 | AML-2 |
| T cell | T_cell_PCA_16704732, T_Cell, Tcell_receptors_score, T_cells_MCPcounter, T_cells_CD4_memory_activated, CD8_cluster, T_Cell_cluster, T_cells_CD4_naive, T_cells_follicular_helper, T_cells_regulatory__Tregs, T_cells_CD8, T_cells_CD4_memory_resting | 1.09 | 0.74-1.60 | 0.66 | <0.01 | 0.51 | 0.81 | 0.66 |
| Interferon | Minterferon_Cluster_21214954, Interferon_ImSig | 1.32 | 1.03-1.69 | 0.03 | 0.73 | 0.13 | 0.37 | 0.03 |
| NK cell | Natural_killer_T_cell, NaturalKiller_Cell_Cytotoxicity | 1.42 | 0.79-2.54 | 0.24 | 0.08 | 0.07 | 0.05 | 0.24 |
| TGFb | TGFB_score_21050467 | 0.7 | 0.46-1.06 | 0.10 | <0.01 | 0.69 | 0.24 | 0.10 |
| B cell | B_cell_PCA_16704732, B_Cell | 0.35 | 0.21-0.60 | <0.01 | <0.01 | 0.56 | <0.01 | <0.01 |
| Antigen processing | APM1, MHC_I_19272155, MHC1_21978456 | 1.01 | 0.83-1.22 | 0.95 | <0.01 | 0.20 | <0.01 | 0.95 |
| Neutrophils | Neutrophils_MCPcounter, Neutrophils_ImSig, Neutrophil, Neutrophils | 0.78 | 0.61-1.00 | 0.05 | 0.02 | 0.93 | 0.16 | 0.05 |
| Monocytes | Monocytes, Macrophages_M2, Macrophages_M0 | 0.74 | 0.57-0.97 | 0.03 | <0.01 | 0.63 | <0.01 | 0.03 |

ARTICLE IN PRESS

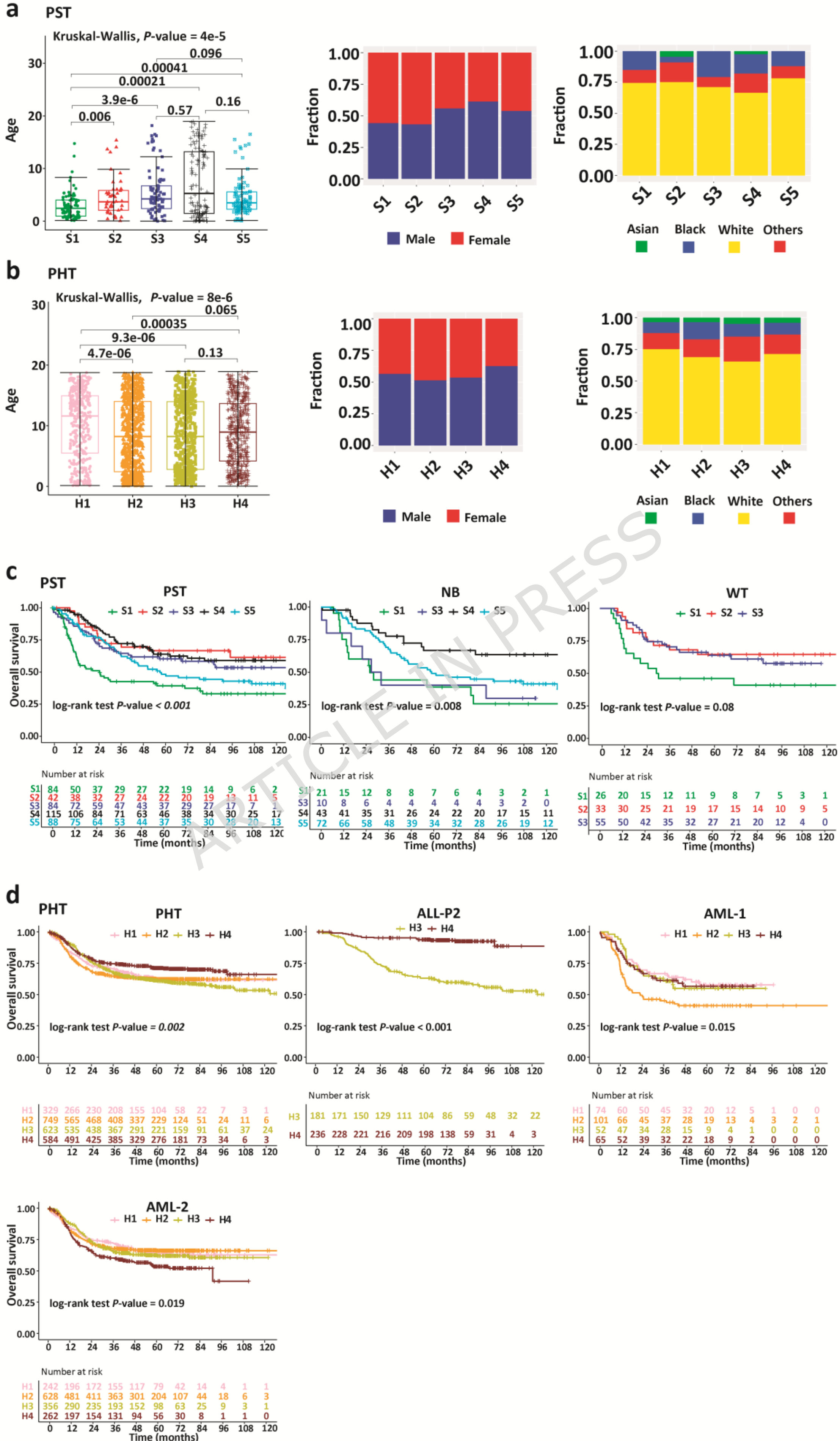
ARTICLE IN PRESS




ARTICLE IN PRESS

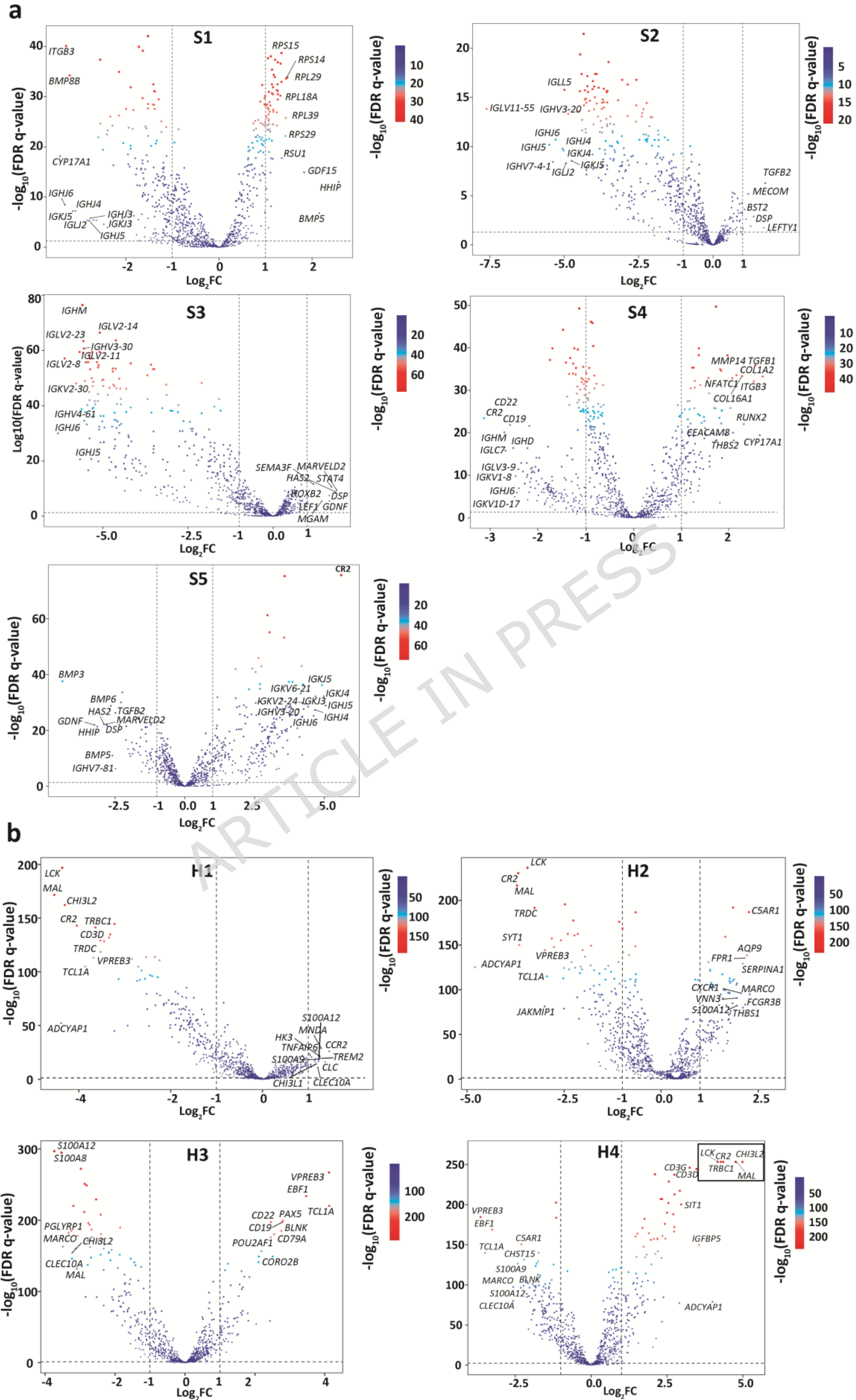



ARTICLE IN PRESS

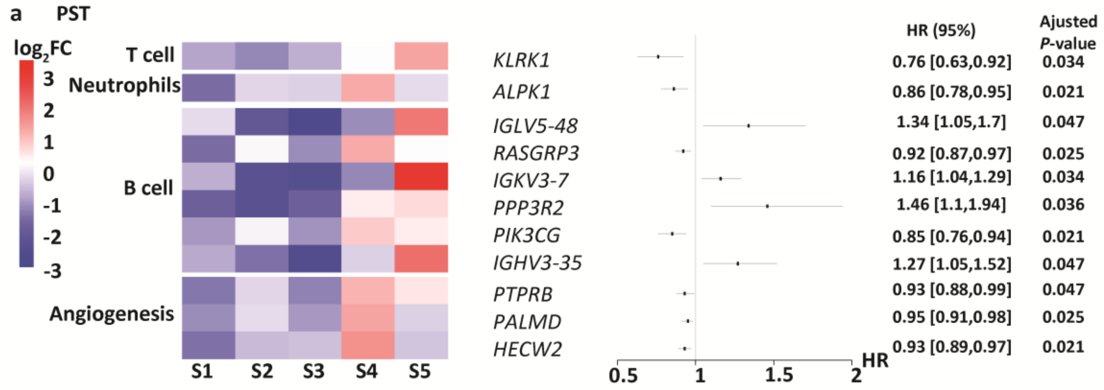


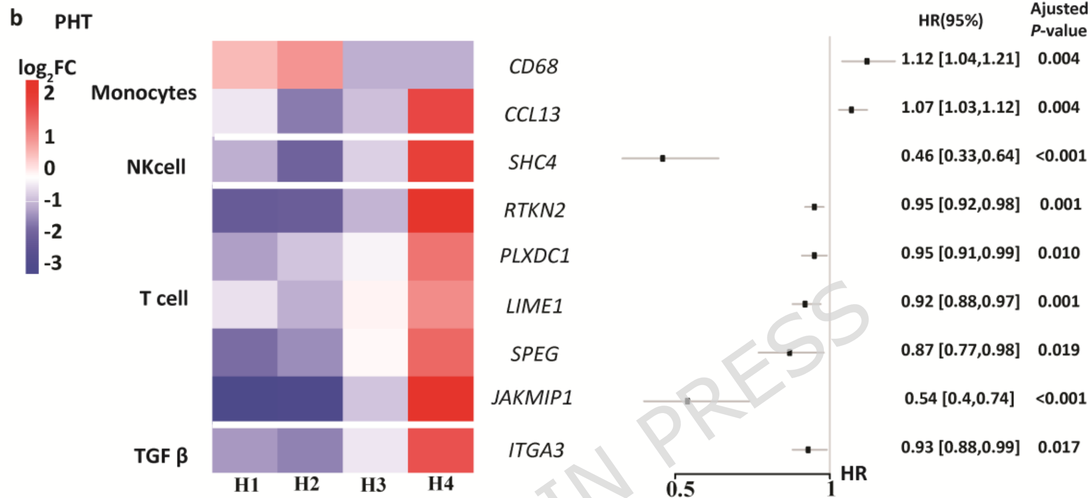

ARTICLE IN PRESS

ARTICLE IN PRESS

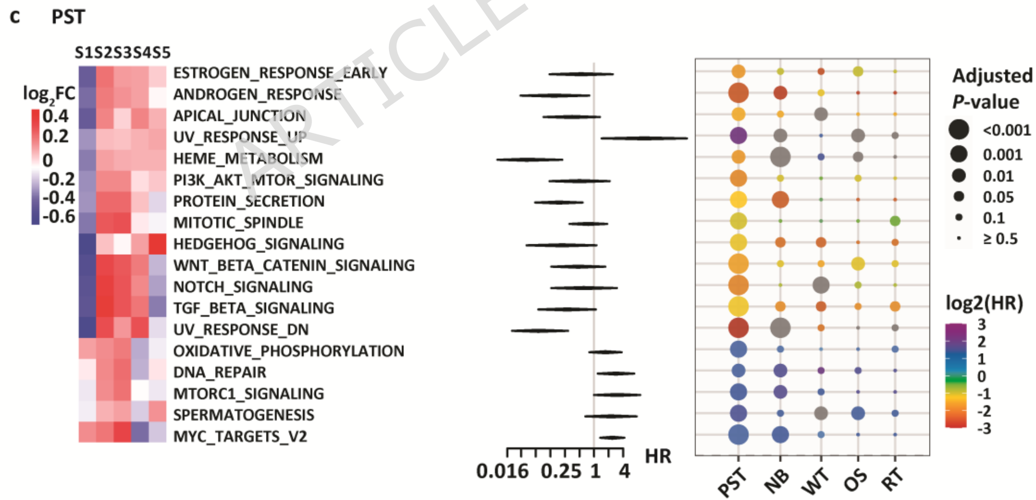


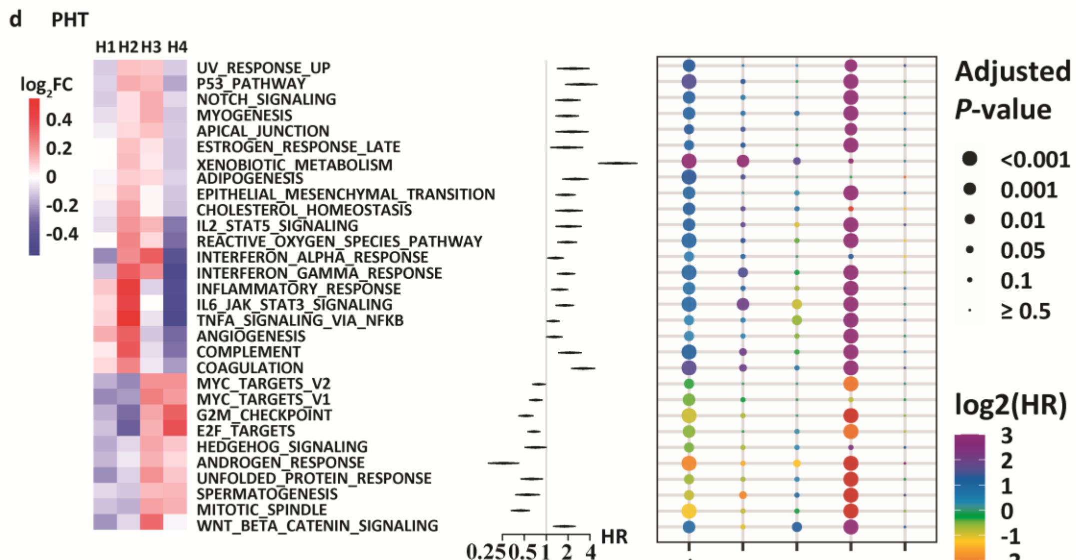
ARTICLE IN PRESS



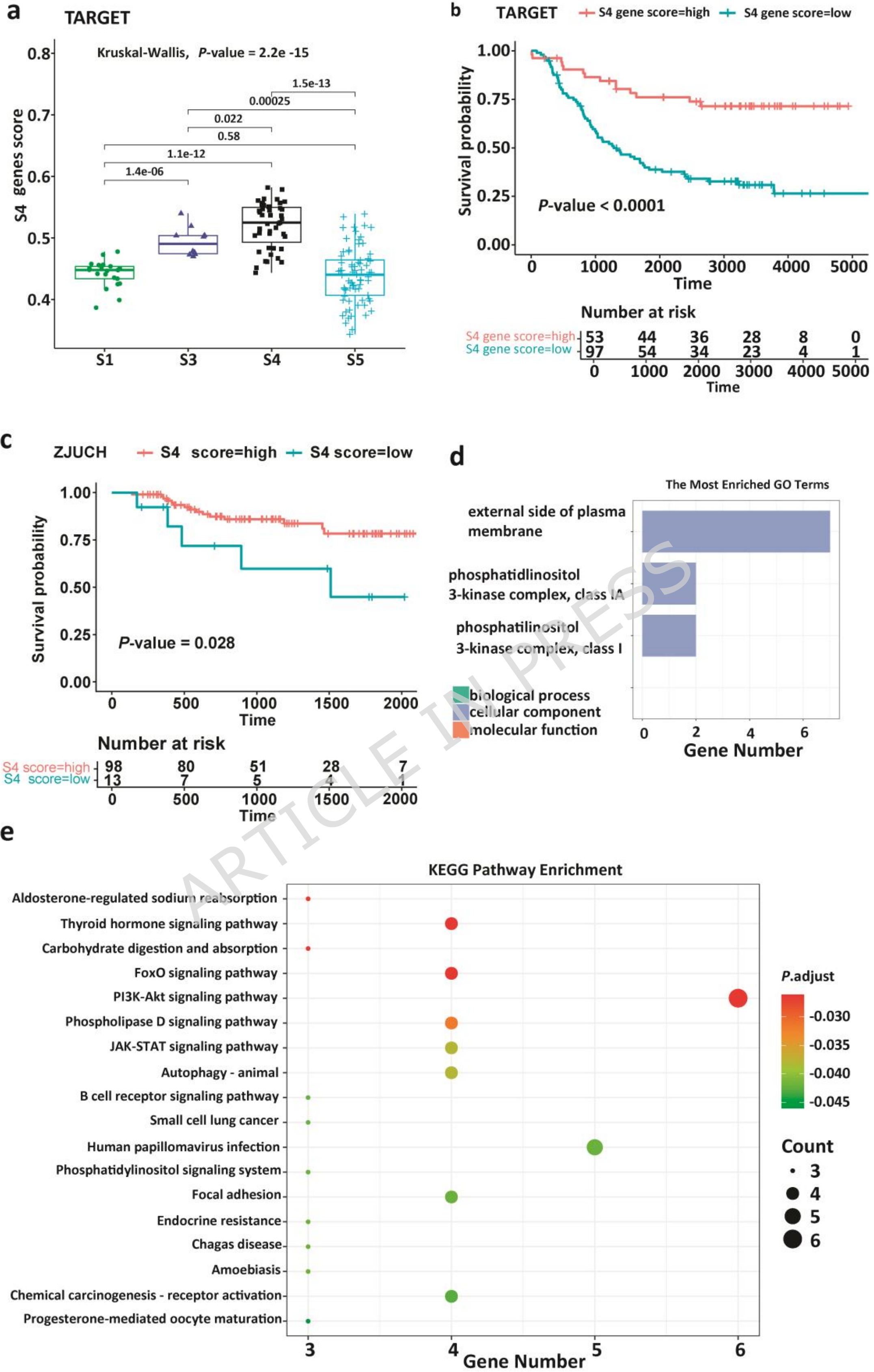




ARTICLE IN PRESS



ARTICLE IN PRESS




ARTICLE IN PRESS