pathogens

MDPI

Article

Cross-Cohort Integration of Blood DNA Methylation and Sepsis Transcriptomes Prioritizes TP53INP1 as a Candidate Associated with B-Cell Transcriptomic Patterns in Pediatric HAdV-7-Associated Sepsis

Peidan Hu †, Bolun Huang †, Wenmin Yang, Chunmin Zhang, Jinda Huang, Yiyu Yang * and Feiyan Chen *

Department of Pediatric Intensive Care Unit, Guangzhou Women and Children's Medical Centre, Guangzhou Medical University, Guangzhou 510623, China
* Correspondence: yangyiyu2019@126.com (Y.Y.); kelly20200407@outlook.com (F.C.)
† These authors contributed equally to this work.

Abstract

Human adenovirus type 7 (HAdV-7) can cause severe pneumonia and sepsis in children, but the associated host epigenetic and immune alterations remain poorly defined. We profiled peripheral-blood DNA methylation in 11 children with HAdV-7-associated sepsis (6 survivors and 5 non-survivors) and 5 pediatric controls without clinically detectable HAdV-7 infection using reduced representation bisulfite sequencing. Differentially methylated genes were integrated with pediatric septic shock transcriptomic modules from GSE26440, and their cellular distribution and predicted regulatory effects were explored using the single-cell dataset GSE167363 and scTenifoldKnk. We identified 106 genes showing differential methylation across all three pairwise comparisons. Cross-cohort integration prioritized 19 candidate genes, among which TP53INP1 showed preferential expression in B cells in the external sepsis single-cell dataset. In silico perturbation of TP53INP1 was associated with changes in interferon-stimulated and antiviral genes, including TRIM22, IFI44L, and PARP14. These findings identify TP53INP1 as a hypothesis-generating candidate linking HAdV-7-associated whole-blood methylation changes with B-cell regulatory networks observed in an external sepsis dataset. Given the small discovery cohort and absence of experimental validation, the findings are exploratory and require confirmation in independent cohorts.

Keywords: HAdV-7; pediatric sepsis; DNA methylation; host–pathogen interaction; B cells; TP53INP1

Check for updates

Academic Editors: Hongyan Guo and Heather S. Koehler

Received: 23 July 2026
Revised: 26 August 2026
Accepted: 29 August 2026
Published: 17 September 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

Human adenoviruses (HAdVs) are non-enveloped viruses belonging to the family Adenoviridae and contain a linear double-stranded DNA genome [1]. HAdVs are important respiratory pathogens in children and can cause a broad spectrum of disease ranging from self-limited respiratory infection to severe pneumonia and critical illness. Among the circulating HAdV types, HAdV-7 has been repeatedly associated with more severe lower respiratory tract disease in pediatric populations [2–4]. Molecular epidemiological surveillance in Guangzhou showed a shift in the predominant HAdV subtype from HAdV-3 to HAdV-7 during 2018–2019, with HAdV-7 infection associated with greater disease severity [2]. Studies from other regions of China have similarly reported an association

between HAdV-7 and severe pediatric pneumonia and respiratory failure [3,4]. However, the host molecular alterations associated with severe HAdV-7 infection and sepsis-related organ dysfunction remain incompletely understood.

The rapid clinical deterioration observed in septic patients is fundamentally driven by systemic immune dysregulation [5,6]. Current evidence indicates that tissue damage resulting from HAdV-7 infection is associated with the aberrant activation of the host's immune system, with a particular emphasis historically placed on the innate immune response [7]. However, the pathophysiology of sepsis extends far beyond initial hyperinflammation; it is characterized by a broad immune dysfunction that involves quantitative and functional alterations in both innate and adaptive immune cells [8]. Elucidating the deeper regulatory networks that govern this broad-spectrum immune exhaustion may help inform future therapeutic strategies [9].

Epigenetic modifications act as important regulators of gene expression during disease progression, with DNA methylation representing a widely studied epigenetic modification relevant to disease-associated molecular changes [10]. Previous studies have shown that DNA methylation is associated with immune activation and immune tolerance [10,11]. Furthermore, examining methylation profiles from peripheral blood offers an accessible approach to explore the pathogenesis of systemic diseases and identify biomarkers associated with disease severity [10–12]. Nevertheless, to date, epigenetic profiling has rarely been applied to explore the molecular landscape of children with HAdV-7-associated sepsis.

To explore potential molecular links between HAdV-7-associated methylation alterations and sepsis-related immune dysregulation, we combined RRBS-derived methylation data from the pediatric HAdV-7 cohort with independent public bulk and single-cell transcriptomic datasets. The external datasets were used for cross-cohort candidate prioritization and cellular-context exploration rather than direct validation in the same patients. This exploratory framework was further complemented by in silico network perturbation to generate hypotheses regarding TP53INP1-associated B-cell regulatory programs.

2. Materials and Methods

2.1. Subjects

This retrospective observational study analyzed archived whole-blood samples and clinical data from pediatric patients admitted to the pediatric intensive care unit (PICU) of Guangzhou Women and Children's Medical Centre, Guangzhou Medical University, between April 2018 and April 2020. The archived samples had been collected and stored under a broader approved research protocol on severe viral infections (Approval No. 202054301). Children aged >30 days and <18 years were included if they had severe pneumonia, laboratory-confirmed HAdV-7 infection, and infection-associated acute organ dysfunction requiring intensive care. HAdV infection was detected by PCR using throat-swab or lower-respiratory-tract specimens, and HAdV-7 typing was based on the institutional molecular diagnostic workflow described previously [2].

The term "HAdV-7-associated sepsis" was used because HAdV-7 could not be established as the sole cause of systemic organ dysfunction. The diagnosis was interpreted according to established pediatric sepsis frameworks based on severe infection accompanied by acute organ dysfunction [13,14].

Eleven children with HAdV-7-associated sepsis with available archived whole-blood samples were included in the RRBS analysis and classified as survivors (n = 6) or non-survivors (n = 5) according to in-hospital outcome. Blood samples were collected at PICU admission and stored at −80 °C. Five pediatric controls undergoing routine physical examination at the same hospital were also included. Because complete archived information on screening numbers, quantitative viral load, coinfections, and some pre-admission

treatments was unavailable, these factors could not be systematically analyzed and are acknowledged as study limitations.

The broader research protocol under which the biospecimens were collected was approved by the Ethics Committee of Guangzhou Women and Children's Medical Centre (Approval No. 202054301), with an approved study period from 1 April 2018 to 31 March 2021. The present study represents a retrospective analysis of archived biospecimens and clinical data obtained within this research framework. Written informed consent for participation and research use of biospecimens was obtained from the parents or legal guardians of all participants in accordance with the approved protocol. The study was conducted in accordance with the Declaration of Helsinki and relevant institutional guidelines.

2.2. RRBS Library Construction and Bioinformatics Processing

Genomic DNA was extracted from peripheral whole-blood samples using the QIAamp DNA Blood Mini Kit (Qiagen, Hilden, Germany). RRBS libraries were prepared according to established procedures [15]. Briefly, 1 μg of genomic DNA was digested with 100 U of MspI restriction enzyme (New England Biolabs, Ipswich, MA, USA) at 37 °C for 16 h. Following end repair, dATP addition, and methylated-adapter ligation, DNA fragments of 160–350 bp were size-selected from 2% agarose gels. Bisulfite conversion was performed using the EZ DNA Methylation-Gold Kit (Zymo Research, Irvine, CA, USA) according to the manufacturer's instructions. The converted DNA was amplified using JumpStart Taq DNA Polymerase (Sigma-Aldrich, St. Louis, MO, USA) for 13 cycles (94 °C for 30 s, 58 °C for 30 s, and 72 °C for 30 s). Library quality and concentration were assessed using an Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA) and quantitative PCR, respectively. The final RRBS libraries were sequenced on an Illumina HiSeq X Ten platform (Illumina, San Diego, CA, USA) using 150-bp paired-end reads.

Raw sequencing reads containing >30% ambiguous bases or low-quality bases ($Q < 20$) in >10% of the sequence were removed. Adapter sequences were trimmed using Cutadapt version 1.9 [16]. Clean reads were aligned to the human reference genome GRCh38 (UCSC hg38) using BSMAP version 2.73 [17], and only uniquely mapped reads containing the expected terminal MspI restriction sites were retained. CpG sites with a sequencing depth $\geq 5\times$ were included in downstream methylation analysis. The methylation level at each CpG site was calculated as the proportion of methylated cytosine reads relative to the total number of methylated and unmethylated reads.

Differentially methylated regions (DMRs) were identified using Metilene version 0.2-7 [18], which uses recursive segmentation of intergroup methylation differences and a two-dimensional Kolmogorov–Smirnov test to assess the statistical significance of candidate regions. DMRs were required to contain at least five CpG sites, with an inter-CpG distance ≤ 300 bp, an absolute methylation difference > 0.10, and a Benjamini–Hochberg-adjusted Q-value < 0.05. Gene promoters were defined as regions extending from 2 kb upstream to 0.5 kb downstream of the transcription start site. Genes were designated as differentially methylated genes (DMGs) when at least one qualifying DMR showed >50% sequence overlap with the corresponding promoter or gene body. GO enrichment analysis of DMGs was performed using a hypergeometric test, with a Benjamini–Hochberg false discovery rate (FDR) < 0.05 considered statistically significant.

Because complete leukocyte-subset data were unavailable for the archived cohort, formal adjustment for blood-cell composition was not performed. Accordingly, differential methylation findings were interpreted as whole-blood-associated signals rather than evidence of cell-intrinsic epigenetic alterations.

2.3. WGCNA and Candidate Gene Prioritization

To prioritize methylation-associated candidates using an independent transcriptomic context, we performed cross-cohort analysis using the pediatric septic shock dataset GSE26440. The pediatric septic shock transcriptomic dataset GSE26440 was downloaded from the Gene Expression Omnibus (GEO) database [19–21]. The dataset contains whole-blood gene-expression profiles from 98 children with septic shock and 32 healthy controls generated using the Affymetrix Human Genome U133 Plus 2.0 Array platform (GPL570; Affymetrix, Santa Clara, CA, USA). The publicly available expression data had been processed using robust multi-array average (RMA) normalization, with patient samples subsequently normalized to the median values of the healthy controls. The healthy controls were used for normalization and were not included in the mortality-associated WGCNA analysis.

WGCNA was performed on the 98 septic-shock samples using WGCNA version 1.72-1 in R version 4.2.2. A signed co-expression network was constructed using Pearson correlations, with a soft-thresholding power of $\beta = 6$, a minimum module size of 30 genes, and a module-merging threshold of 0.25. In-hospital outcome was coded as 0 for survivors and 1 for non-survivors. Module eigengenes were correlated with mortality using Pearson correlation.

The turquoise module was the largest co-expression module in the original analysis and contained 5451 genes. Because its correlation with mortality was weak and non-significant, it was not interpreted as a mortality-associated module in the revised analysis. Instead, it was retained as a broad exploratory co-expression context for cross-cohort candidate prioritization. Intersection with the 106 RRBS-derived DMGs yielded 19 candidate genes for subsequent single-cell contextualization. This overlap was used to refine the candidate set across independent molecular datasets and was not interpreted as a formal enrichment test. No formal module-preservation analysis was performed, and the WGCNA-based prioritization was therefore considered exploratory.

2.4. Single-Cell RNA Sequencing (scRNA-seq) Data Preprocessing and Clustering

To delineate the cellular specificities of the identified candidate genes, the public scRNA-seq dataset GSE167363 was downloaded and analyzed [22]. The GSE167363 dataset comprised samples from 7 individual donors, including 2 healthy controls and 5 patients with sepsis. Because the sepsis patients were sampled longitudinally, the dataset contained multiple samples from some donors. Single-cell data were processed using Seurat version 4.3.0 in R version 4.2.2 [23]. Cells expressing 200–6000 genes and containing less than 15% mitochondrial transcripts were retained. Gene counts were log-normalized with a scale factor of 10,000, and the top 2000 highly variable genes were selected. After sample integration and data scaling, principal component analysis was performed. The first 20 principal components were used for neighbor construction, clustering at a resolution of 0.15, and uniform manifold approximation and projection visualization.

2.5. Cell-Type Annotation and Exploratory Cell-Level Expression Analysis

Cluster marker genes were identified using the FindAllMarkers function with the Wilcoxon rank-sum test. Marker genes were required to be expressed in at least 10% of cells, with an average log2 fold change greater than 0.25 and an adjusted *p*-value below 0.05. Cell types were annotated using canonical markers and the CellMarker 2.0 database, including CD3D and CD3E for T cells, NKG7 and GNLY for cytotoxic T/NK cells, CD79A and MS4A1 for B cells, LST1 and LYZ for monocytes, and PPBP and PF4 for platelets [24]. The expression of the 19 prioritized candidate genes was examined across the annotated cell populations. All 19 intersecting genes were carried forward without a prespecified composite ranking score. TP53INP1 was selected for focused downstream analysis be-

cause it showed a prominent B-cell-enriched expression pattern in the external single-cell dataset. This selection was exploratory and was not based on the highest WGCNA gene significance or methylation effect size. TP53INP1 expression was predominantly enriched in B-cell clusters; therefore, B cells were subsetted for further analysis. Exploratory cell-level expression differences between sepsis and healthy-control B cells were assessed using FindMarkers with the Wilcoxon rank-sum test, using an absolute average log2 fold change >0.25 and a Benjamini–Hochberg-adjusted p value < 0.05 for visualization and functional exploration. Because cells from the same donor are not statistically independent, these results were not interpreted as donor-level differential-expression inference. Functional enrichment analysis was performed using clusterProfiler version 4.6.2, with Benjamini–Hochberg-adjusted p-values below 0.05 considered significant [25].

2.6. Computational Network Perturbation Analysis Using scTenifoldKnk

The predicted regulatory role of TP53INP1 in sepsis-associated B cells was evaluated using scTenifoldKnk version 1.0.1 [26]. The raw count matrix of B cells from the external sepsis dataset was used as input, and genes expressed in fewer than 25 cells were excluded. Gene regulatory networks were constructed using 10 networks, with 500 cells used for each network and three principal components retained. Tensor decomposition was performed with a rank of 3, and the original and TP53INP1-perturbed networks were compared using two-dimensional manifold alignment.

Genes were ranked according to their manifold-alignment perturbation distance. The archived analysis used nominal p-values for exploratory visualization rather than a multiple-testing-adjusted significance threshold. Accordingly, the perturbation results were interpreted as hypothesis-generating computational predictions and were not considered evidence of experimentally validated gene regulation or biological causality. Functional enrichment of the predicted perturbed gene set was performed using clusterProfiler.

2.7. Statistical Analysis

Statistical analyses were performed in R version 4.2.2 unless otherwise specified. Multiple testing was controlled using the Benjamini–Hochberg procedure where applicable. WGCNA module–trait associations were assessed using Pearson correlation. Single-cell marker and exploratory cell-level expression analyses were performed using Wilcoxon rank-sum tests with Benjamini–Hochberg adjustment. scTenifoldKnk results were interpreted using nominal p values as exploratory outputs. Given the small size of the clinical cohort, clinical characteristics were summarized descriptively without formal between-group hypothesis testing.

3. Results

3.1. Clinical Characteristics of HAdV-7-Associated Sepsis Patients

This study included 11 children with HAdV-7-associated sepsis, including 6 survivors and 5 non-survivors. Detailed individual-level clinical characteristics and treatment information are presented in Table 1, while clinical characteristics stratified by survival status are summarized in Table 2. All patients developed ARDS and required invasive mechanical ventilation and vasoactive support. ECMO was used in 3 of 6 survivors (50.0%) and 4 of 5 non-survivors (80.0%). The median PaO_2/FiO_2 ratio was 93.5 (IQR, 68.3–96.3) among survivors and 66.0 (IQR, 36.0–72.0) among non-survivors. Blood samples used for RRBS were collected at PICU admission. Because of the small sample size, clinical characteristics are presented descriptively without formal between-group hypothesis testing.

Table 1. Clinical characteristics and treatment information of children with HAdV-7-associated sepsis.

ID	Sex	Age (Months)	Admission Weight (kg)	Pre-Admission Illness Duration (Days)	Hospital Stay (Days)	PICU Length of Stay (Days)	Invasive Ventilation (Days)	HFV	ECMO	Outcome	PaO_2/FiO_2 Ratio	WBC Count ($\times 10^9$/L)	Neutrophil Count ($\times 10^9$/L)
1	Male	6	8	5	39	13	11	Y	Y	Survivor	93	13.3	7.76
2	Male	96	17	15	28	21	18	Y	Y	Survivor	43	5.6	4.07
3	Female	14	10	7	38	38	38	Y	Y	Survivor	60	6.2	5.1
4	Female	12	10	7	27	12	8	N	N	Survivor	97	2	1.29
5	Female	20	12	14	27	18	15	N	N	Survivor	94	1.6	0.67
6	Male	19	10	7	23	11	11	N	N	Survivor	138	2.7	1.48
7	Female	17	10	15	25	24	25	Y	N	Non-survivor	72	11.5	10.38
8	Male	7	10	10	7	7	7	N	Y	Non-survivor	36	7.1	4.5
9	Male	36	16	7	57	57	57	Y	Y	Non-survivor	66	2.7	1.35
10	Male	8	8	25	5	5	5	N	Y	Non-survivor	36	9.8	7.97
11	Male	10	9	7	24	22	22	Y	Y	Non-survivor	81	6.3	3

Abbreviations: PICU, pediatric intensive care unit; ECMO, extracorporeal membrane oxygenation; PaO_2/FiO_2 ratio, ratio of arterial oxygen partial pressure to the fraction of inspired oxygen; WBC, white blood cell; Neutrophil count, absolute neutrophil count. Y, yes; N, no. HFV, high-frequency ventilation.

Table 2. Clinical characteristics of patients with HAdV-7-associated sepsis stratified by survival status.

Characteristic	Survivors (n = 6)	Non-Survivors (n = 5)
Age, months	16.5 (12.5–19.8)	10.0 (8.0–17.0)
Male sex, n (%)	3 (50.0)	4 (80.0)
Admission weight, kg	10.0 (10.0–11.5)	10.0 (9.0–10.0)
Pre-admission illness duration, days	7.0 (7.0–12.3)	10.0 (7.0–15.0)
ARDS, n (%)	6 (100)	5 (100)
Invasive mechanical ventilation, n (%)	6 (100)	5 (100)
Invasive ventilation duration, days	13.0 (11.0–17.3)	22.0 (7.0–25.0)
HFV, n (%)	3 (50.0)	3 (60.0)
ECMO, n (%)	3 (50.0)	4 (80.0)
Vasoactive support, n (%)	6 (100)	5 (100)
PaO_2/FiO_2 ratio	93.5 (68.3–96.3)	66.0 (36.0–72.0)
WBC count, $\times 10^9$/L	4.15 (2.17–6.05)	7.10 (6.30–9.80)
Neutrophil count, $\times 10^9$/L	2.78 (1.34–4.84)	4.50 (3.00–7.97)
Sampling time	PICU admission	PICU admission

3.2. Identification of Differentially Methylated Genes

To characterize whole-blood methylation differences associated with clinical outcome, RRBS methylation profiles were compared among the control, survivor, and non-survivor groups. PCA showed partial separation among the three groups, although substantial overlap remained between the survivor and non-survivor groups (Figure 1). PC1 and PC2 explained 56.5% and 20.7% of the total variance, respectively. No formal clustering-performance metric was calculated; therefore, the PCA was interpreted descriptively rather than as evidence of distinct molecular separation. Unsupervised hierarchical clustering further illustrated group-associated methylation patterns (Figure 2A–C). DMGs identified in the Survivor vs. Control, Non-survivor vs. Control, and Non-survivor vs. Survivor comparisons were intersected, yielding 106 genes shared across all three comparisons (Figure 2D). The complete list of the 106 shared DMGs is provided in Supplementary Data S1. These genes were subsequently used for cross-cohort candidate-gene prioritization. Because statistical significance across all three pairwise comparisons does not necessarily imply concordant methylation direction, the 106-gene intersection was used as a candidate-prioritization set rather than interpreted as a uniformly directional methylation signature.

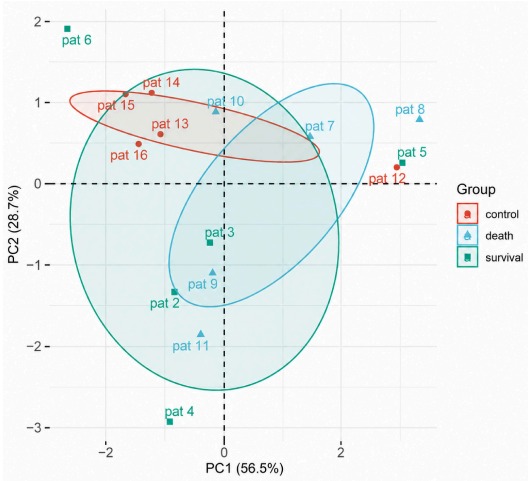


Figure 1. Principal component analysis (PCA) of whole-blood DNA methylation profiles from controls (n = 5), survivors (n = 6), and non-survivors (n = 5). PC1 and PC2 explained 56.5% and 28.7% of the variance, respectively. Each point represents one participant. PCA was interpreted descriptively.

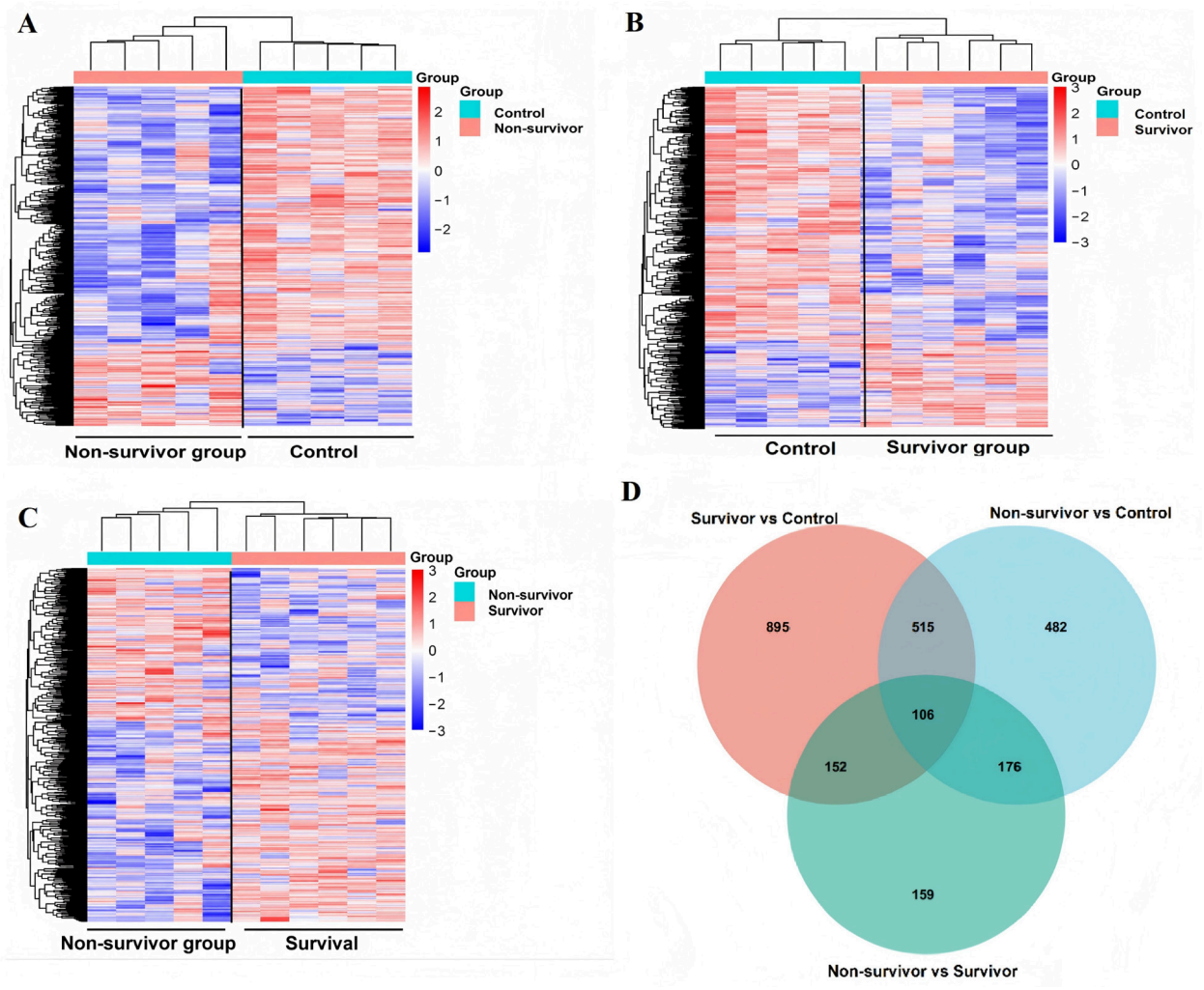


Figure 2. Unsupervised clustering and intersection of differential methylation signals. (**A**) Non-survivor (n = 5) vs. Control (n = 5). (**B**) Survivor (n = 6) vs. Control (n = 5). (**C**) Non-survivor (n = 5) vs. Survivor (n = 6). Red and blue indicate relatively higher and lower scaled methylation levels, respectively. (**D**) Venn diagram showing the overlap of DMGs across the three comparisons, yielding 106 shared DMGs.

3.3. Prioritization of Candidate Genes via WGCNA and Methylation Intersection

WGCNA of the GSE26440 septic-shock cohort identified multiple co-expression modules (Figure 3A–C). The turquoise module was the largest module, containing 5451 genes, but showed only a weak and non-significant correlation with in-hospital mortality (r = 0.072, p = 0.40). Accordingly, this module was not interpreted as a mortality-associated module. Instead, it was used as an exploratory co-expression context for cross-cohort candidate prioritization. Intersection of the 5451 turquoise-module genes with the 106 RRBS-derived DMGs yielded 19 candidate genes, including TP53INP1 (Figure 3D,E). This cross-cohort overlap provided an additional layer of candidate refinement before evaluation in the external single-cell dataset. The overlap was used for exploratory prioritization rather than interpreted as a formal statistical enrichment test. All 19 genes were subsequently examined in the external single-cell dataset to assess their cell-type expression patterns. The complete list of the 19 intersecting candidate genes is provided in Supplementary Data S2. TP53INP1 was not the highest-ranked gene by WGCNA gene significance; its subsequent prioritization was based on its B-cell-enriched expression pattern in the external single-cell analysis.

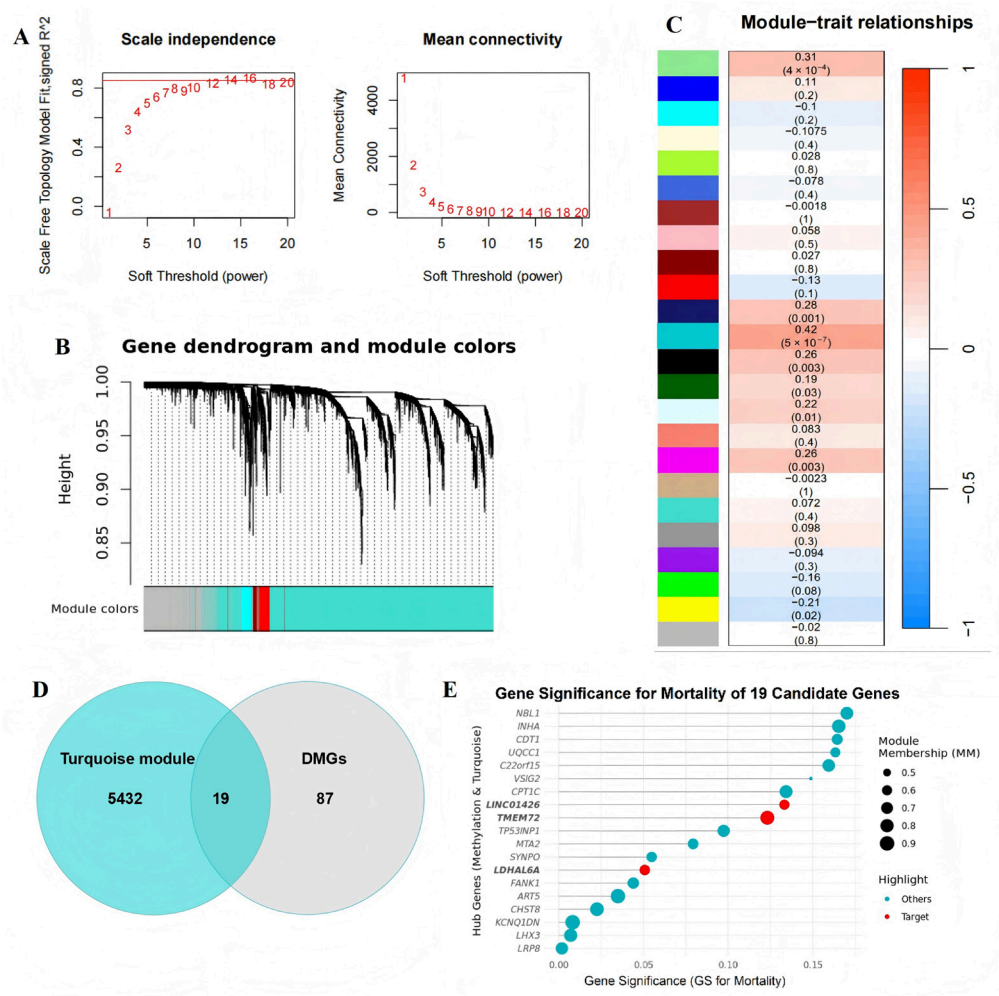


Figure 3. WGCNA-based co-expression analysis and candidate-gene prioritization in GSE26440. WGCNA was performed using 98 pediatric septic-shock samples. (**A**) Soft-threshold selection. (**B**) Gene dendrogram and module assignment. (**C**) Module–mortality relationships based on Pearson correlation. (**D**) Intersection of 5451 turquoise-module genes with 106 RRBS-derived DMGs, yielding 19 candidate genes. (**E**) Gene significance for mortality among the 19 candidates; point size represents module membership (MM).

3.4. Single-Cell Transcriptomic Characterization of the External Sepsis Dataset

UMAP visualization was used to display the cellular landscape of PBMCs in the healthy-control and sepsis samples from the external GSE167363 dataset (Figure 4A,B). The apparent distribution of cell populations in the UMAP embeddings was considered descriptive and was not interpreted as evidence of altered cell-type composition, because donor-level cell-proportion analysis was not performed.

Exploratory cell-level expression analysis revealed genes associated with broad immune and stress-response programs. GO-BP enrichment highlighted pathways related to immune and antiviral responses (Figure 4C), whereas KEGG enrichment identified several inflammatory and immune-associated pathways (Figure 4D). The distribution of cell-level expression differences was visualized using a volcano plot (Figure 4E). The expression patterns of the 19 cross-cohort candidate genes were subsequently examined across conditions (Figure 4F). These analyses were used for cellular contextualization and hypothesis generation rather than donor-level differential-expression inference.

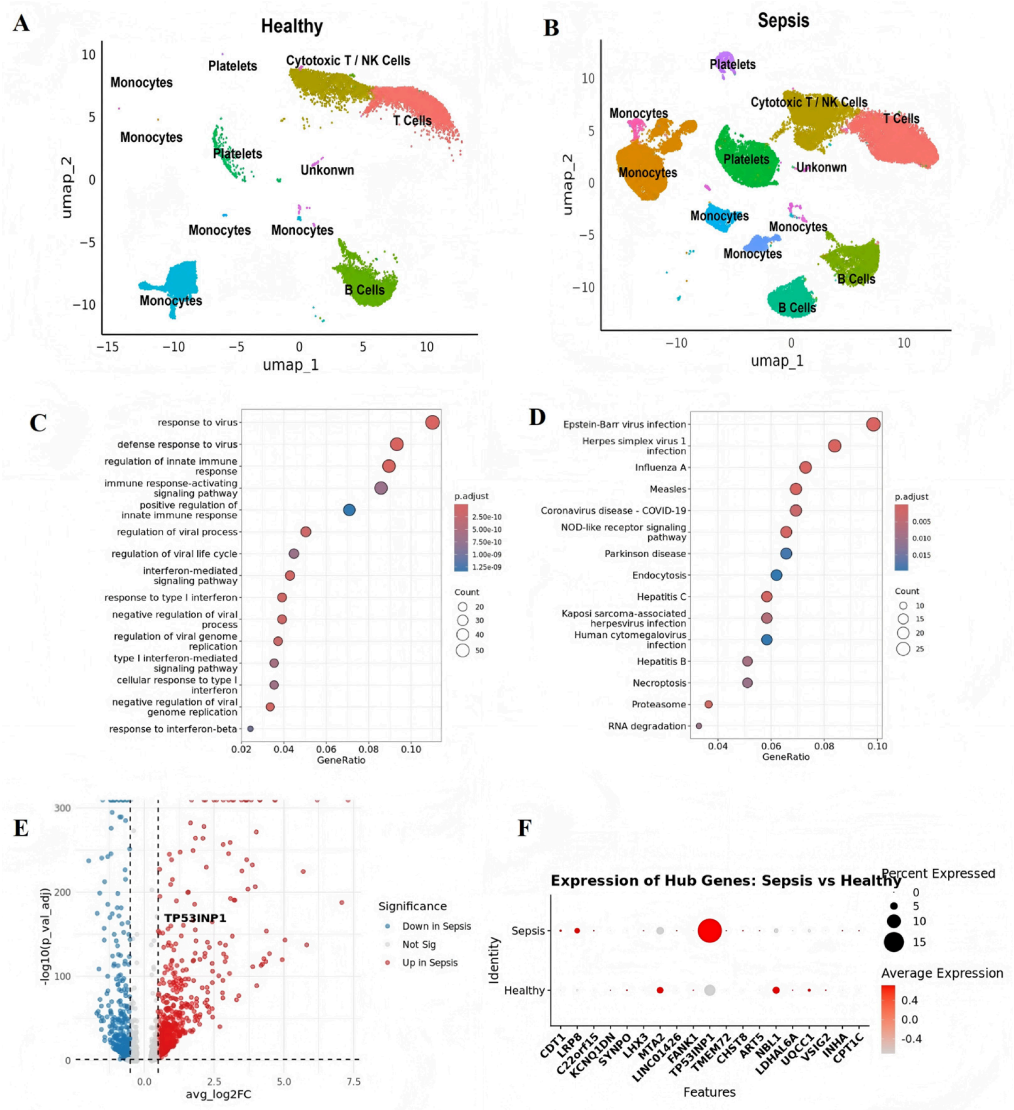


Figure 4. Exploratory single-cell transcriptomic analysis of GSE167363 comprising 7 donors (2 healthy controls and 5 patients with sepsis). (**A**,**B**) UMAP visualization of healthy and sepsis cells; colors indicate annotated cell populations. (**C**,**D**) GO-BP and KEGG enrichment of exploratory cell-level expression differences; point size represents gene count and color represents Benjamini–Hochberg-adjusted *p* value. (**E**) Volcano plot of exploratory cell-level expression differences identified using the Wilcoxon rank-sum test with Benjamini–Hochberg correction. (**F**) Expression patterns of the 19 candidate genes; point size represents the percentage of expressing cells and color represents average expression.

3.5. Cell-Type Annotation and B-Cell-Enriched Expression of TP53INP1

The expression patterns of the candidate genes were examined across the annotated cell populations (Figure 5A). Cell identities were assigned on the basis of canonical lineage-marker expression (Figure 5B). TP53INP1 showed a prominent B-cell-enriched expression pattern in the external single-cell dataset. Feature and violin plots were further used to descriptively visualize TP53INP1 expression in B cells from sepsis and healthy-control samples (Figure 5C–E). Because these observations were not evaluated using a donor-aware statistical model, no inferential conclusion regarding differential TP53INP1 expression between the two groups was drawn.

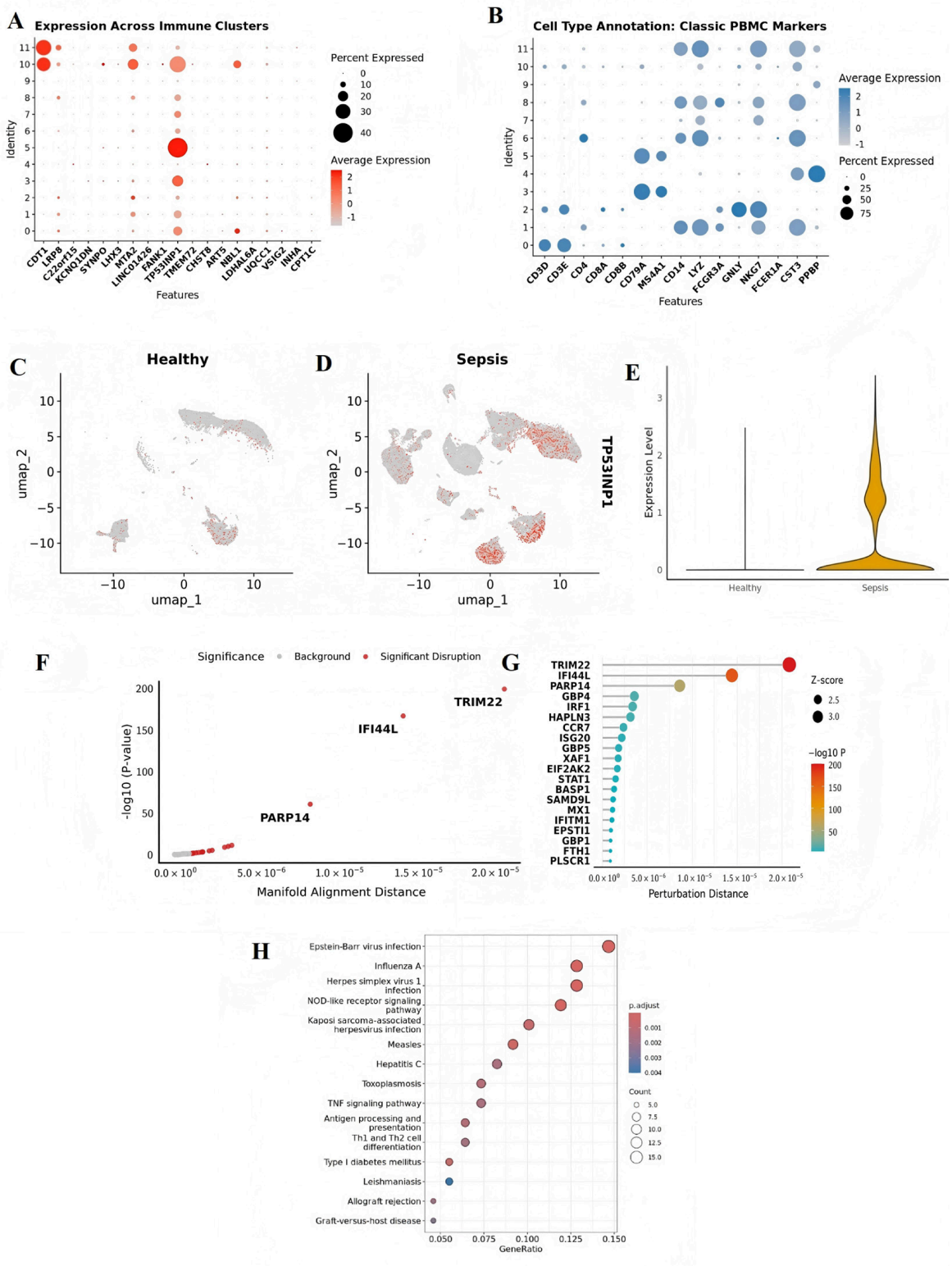


Figure 5. Cell-level TP53INP1 expression and computational network-perturbation analysis in the external GSE167363 sepsis dataset. (**A**,**B**) Candidate-gene and canonical marker expression across cell clusters; point size indicates percent expressed and color indicates average expression. (**C**,**D**) TP53INP1 expression in healthy and sepsis cells. (**E**) Cell-level TP53INP1 expression in B cells. (**F**,**G**) Predicted TP53INP1-associated network perturbations; nominal *p* values were used without multiple-testing correction. (**H**) KEGG enrichment of predicted perturbed genes; point size indicates gene count and color indicates adjusted *p* value.

3.6. Computational Network Perturbation Identifies TP53INP1-Associated B-Cell Gene Programs

Computational perturbation of TP53INP1 using scTenifoldKnk identified a set of genes with relatively large predicted network-perturbation distances (Figure 5F,G). Among the higher-ranking genes were several interferon-stimulated or antiviral-response genes, including TRIM22, IFI44L, and PARP14. Functional enrichment of the predicted perturbed gene set included pathways annotated as Epstein–Barr virus infection, influenza A, herpes simplex virus 1 infection, and NOD-like receptor signaling (Figure 5H). These pathway annotations reflect shared antiviral and inflammatory gene programs and should not be interpreted as evidence of coinfection or HAdV-7-specific regulatory effects. Overall, the analysis generates a computational hypothesis that TP53INP1 may be associated with B-cell antiviral-response networks in sepsis; it does not demonstrate that experimental alteration of TP53INP1 would produce these transcriptional changes.

4. Discussion

Severe HAdV-7 infection can be associated with critical illness and sepsis in children, but the host molecular alterations accompanying severe disease remain incompletely understood [7]. In the present study, we used an exploratory cross-cohort framework to examine these alterations at different levels of evidence. Whole-blood RRBS data from children with HAdV-7-associated sepsis provided the primary disease-specific methylation observations. An independent pediatric septic-shock transcriptomic dataset was subsequently used for candidate-gene prioritization, whereas an external sepsis single-cell dataset was used to examine the cellular expression context of the prioritized genes. Finally, computational network perturbation was applied to generate hypotheses regarding TP53INP1-associated gene programs. Because these molecular layers were obtained from different cohorts, the findings represent cross-cohort contextual evidence rather than direct multi-omic validation of an HAdV-7-specific mechanism. Because GSE26440 and GSE167363 differ from the present HAdV-7 cohort in disease etiology, patient characteristics, sample type, and profiling platform, the TP53INP1-associated signals identified across these datasets cannot be considered HAdV-7-specific. Instead, they may reflect broader molecular features of severe sepsis that warrant validation in disease-matched HAdV-7 cohorts.

The RRBS analysis identified group-associated whole-blood methylation differences, with 106 DMGs shared across the Survivor vs. Control, Non-survivor vs. Control, and Non-survivor vs. Survivor comparisons. These shared genes were used as a candidate-prioritization set and were not interpreted as a uniformly directional methylation signature. In addition, because methylation was measured in whole blood and leukocyte composition could not be formally adjusted, the observed differences cannot be attributed to specific immune-cell populations. Cross-cohort intersection with the GSE26440 turquoise module reduced the RRBS-derived candidate set to 19 genes. Importantly, although the turquoise module itself was not significantly associated with mortality, its overlap with the RRBS-derived DMGs provided an additional cross-cohort layer for candidate prioritization. The 19-gene overlap was therefore interpreted as an exploratory cross-cohort convergence that guided subsequent single-cell contextualization, rather than as a formal test of enrichment or mortality association.

Among the 19 candidate genes, TP53INP1 was selected for further exploration because it showed a prominent B-cell-enriched expression pattern in the external single-cell dataset. Alterations in B-cell number, phenotype, and function have been described in sepsis and septic shock [27–29], providing a biological context in which the observed expression pattern may be relevant. In the present analysis, TP53INP1 expression in sepsis and healthy-control B cells was visualized descriptively at the cell level. Because the single-cell dataset was derived from a clinically and etiologically distinct sepsis cohort and the comparison was

not evaluated using a donor-aware statistical model, no donor-level differential-expression inference was made. Consequently, these observations do not establish that TP53INP1 is specifically dysregulated in B cells from children with HAdV-7-associated sepsis.

Mechanistically, TP53INP1 functions as a stress-responsive protein orchestrating p53-mediated cell cycle arrest and apoptosis [30]. Beyond classical apoptotic pathways, emerging evidence indicates that TP53INP1 is a regulator associated with oxidative stress responses and a modulator of autophagy under severe pathogenic burdens [31]. p53 also participates in interferon-mediated antiviral immunity by enhancing interferon signaling and the transcription of interferon-stimulated genes [32]. Pathway enrichment data suggested that sepsis-derived B cells were subjected to antiviral-response signaling and type I interferon signaling. In this hyper-inflammatory microenvironment, TP53INP1 is a stress-responsive regulator associated with p53-dependent cell-cycle control and apoptosis. In the external sepsis single-cell dataset, TP53INP1 expression was enriched in B-cell clusters. Network perturbation analysis predicted that alterations in TP53INP1 could affect regulatory programs involving TRIM22, IFI44L, PARP14, antigen presentation, TNF signaling, and antiviral responses. These results provide a testable hypothesis linking TP53INP1 to B-cell immune dysfunction; however, they do not demonstrate that TP53INP1 directly induces B-cell apoptosis or depletion in HAdV-7 sepsis. Experimental knockdown or overexpression studies are required to establish causality.

This study has several limitations. First, the RRBS discovery cohort was small, comprising 11 patients and 5 controls, which limited statistical power and the ability to adjust comprehensively for potential confounders. Although multiple-testing correction was applied, the small sample size may still increase the instability and false-positive risk of the methylation findings. Given the small discovery cohort, the stability and reproducibility of the identified methylation signals remain uncertain, and the 106 shared DMGs should be considered exploratory candidates requiring validation in larger independent cohorts. Second, whole-blood DNA methylation profiles are influenced by leukocyte composition, and complete leukocyte-subset data were unavailable; consequently, the observed methylation differences cannot be interpreted as cell-intrinsic epigenetic alterations. Because cell-composition correction could not be performed, we could not determine whether the identified methylation differences would persist after adjustment for leukocyte composition. Third, the bulk and single-cell transcriptomic datasets were derived from independent sepsis cohorts with clinical and etiological characteristics that differed from those of the pediatric HAdV-7 cohort, and no within-subject methylation–expression relationships were available. Fourth, the single-cell expression comparisons were conducted at the cell level and did not account for the non-independence of cells derived from the same donor, introducing a potential pseudoreplication bias. Fifth, the WGCNA-based candidate prioritization was exploratory, and TP53INP1 was selected for focused analysis based on its B-cell-enriched expression pattern rather than a prespecified composite ranking score. Finally, the scTenifoldKnk analysis was computational and did not constitute an experimental gene knockout; detailed information regarding cell subsampling, random seeds, and robustness across alternative network parameters was not retained in the archived analysis records. Because TP53INP1 was prioritized entirely through cross-cohort computational analyses, its expression, methylation status, and functional role were not independently validated in the present study. The proposed TP53INP1-associated B-cell hypothesis therefore requires validation in larger disease-matched cohorts using donor-aware transcriptomic analyses and direct functional experiments.

5. Conclusions

In summary, cross-cohort integration of whole-blood RRBS data from pediatric HAdV-7-associated sepsis with independent sepsis transcriptomic datasets prioritized TP53INP1 as a candidate gene with B-cell-enriched expression in an external single-cell dataset. Computational network perturbation further generated hypotheses linking TP53INP1 with antiviral and inflammatory B-cell gene programs. These findings do not establish an HAdV-7-specific TP53INP1 mechanism and should be regarded as hypothesis-generating observations requiring validation in larger disease-matched cohorts and functional experimental models.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/pathogens15090988/s1, Supplementary Data S1: Complete list of the 106 shared differentially methylated genes identified across the three pairwise comparisons; Supplementary Data S2: Complete list of the 19 candidate genes obtained by intersection of the RRBS-derived DMGs and the WGCNA turquoise module.

Author Contributions: Conceptualization, Y.Y. and F.C.; methodology, P.H. and B.H.; formal analysis, P.H., B.H. and W.Y.; investigation, P.H., B.H., C.Z. and J.H.; resources, Y.Y. and F.C.; data curation, P.H., B.H. and W.Y.; writing—original draft preparation, P.H. and B.H.; writing—review and editing, Y.Y. and F.C.; visualization, P.H. and B.H.; supervision, Y.Y. and F.C.; project administration, Y.Y. and F.C.; funding acquisition, F.C. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the National Natural Science Foundation of China (Grant No. 82302432) and the Medical Science Research Foundation of Guangdong Province (Grant No. B2024274). The funders had no role in the study design; data collection, analysis, or interpretation; preparation of the manuscript; or the decision to submit the manuscript for publication.

Institutional Review Board Statement: The broader research protocol under which the biospecimens analyzed in this study were collected was approved by the Ethics Committee of Guangzhou Women and Children's Medical Centre (Approval No. 202054301), with an approved study period from 1 April 2018 to 31 March 2021. The present study represents a retrospective analysis of archived biospecimens and clinical data collected within this approved research framework.

Informed Consent Statement: Informed consent was obtained from the parents or legal guardians of all participants involved in the study.

Data Availability Statement: The raw RRBS sequencing data are not publicly deposited because of ethical and patient-privacy restrictions associated with the original consent and study approval. De-identified derived data used for candidate prioritization, including the complete lists of 106 shared DMGs and 19 intersecting candidate genes, are provided as Supplementary Data S1 and S2. Additional de-identified analytical materials may be requested from the corresponding authors, subject to institutional and ethical requirements. The public transcriptomic datasets are available from GEO under accession numbers GSE26440 and GSE167363.

Conflicts of Interest: The authors declare that they have no competing interests.

Abbreviations

ARDS, acute respiratory distress syndrome; DMGs, differentially methylated genes; DMRs, differentially methylated regions; DNA, deoxyribonucleic acid; ECMO, extracorporeal membrane oxygenation; FDR, false discovery rate; GO, gene ontology; GO-BP, gene ontology biological processes; GS, gene significance; HAdV, human adenovirus; KEGG, Kyoto encyclopedia of genes and genomes; PaO_2/FiO_2 ratio, ratio of arterial oxygen partial pressure to the fraction of inspired oxygen; PBMCs, peripheral blood mononuclear cells; PCA, principal component analysis; PCR, polymerase chain reaction; PICU, pediatric intensive care unit; RRBS, reduced representation bisulfite sequencing;

scRNA-seq, single-cell RNA sequencing; UMAP, uniform manifold approximation and projection; WGCNA, weighted gene co-expression network analysis.

References

1. Benkő, M.; Aoki, K.; Arnberg, N.; Davison, A.J.; Echavarría, M.; Heß, M.; Jones, M.S.; Kaján, G.L.; Kajon, A.E.; Mittal, S.K.; et al. ICTV Virus Taxonomy Profile: Adenoviridae 2022. *J. Gen. Virol.* **2022**, *103*, 001721. [CrossRef] [PubMed]
2. Chen, Y.; Lin, T.; Wang, C.; Liang, W.; Lian, G.; Zanin, M.; Wong, S.; Tian, X.; Zhong, J.; Zhang, Y.; et al. Human adenovirus (HAdV) infection in children with acute respiratory tract infections in Guangzhou, China, 2010-2021: A molecular epidemiology study. *World J. Pediatr.* **2022**, *18*, 545–552. [CrossRef] [PubMed]
3. Wo, Y.; Lu, Q.B.; Huang, D.; Li, X.; Guo, C.; Wang, H.; Zhang, X.A.; Liu, W.; Cao, W.C. Epidemical features of HAdV-3 and HAdV-7 in pediatric pneumonia in Chongqing, China. *Arch. Virol.* **2014**, *160*, 633–638. [CrossRef] [PubMed]
4. Huang, X.; Yi, Y.; Chen, X.; Wang, B.; Long, Y.; Chen, J.; Rongkavilit, C. Clinical Characteristics of 204 Children with Human Adenovirus Type 7 Pneumonia Identified by Whole Genome Sequencing in Liuzhou, China. *Pediatr. Infect. Dis. J.* **2020**, *40*, 91–95. [CrossRef] [PubMed]
5. van der Poll, T.; van de Veerdonk, F.L.; Scicluna, B.P.; Netea, M.G. The immunopathology of sepsis and potential therapeutic targets. *Nat. Rev. Immunol.* **2017**, *17*, 407–420. [CrossRef] [PubMed]
6. Singer, M.; Deutschman, C.S.; Seymour, C.W.; Shankar-Hari, M.; Annane, D.; Bauer, M.; Bellomo, R.; Bernard, G.R.; Chiche, J.; Coopersmith, C.M.; et al. The Third International Consensus Definitions for Sepsis and Septic Shock (Sepsis-3). *JAMA* **2016**, *315*, 801–810. [CrossRef] [PubMed]
7. Liang, Y.; Wei, J.; Shen, J.; Liang, Z.; Ma, X.; Du, Y.; Qian, W.; Dong, H.; Huang, P.; Chen, A.; et al. Immunological pathogenesis and treatment progress of adenovirus pneumonia in children. *Ital. J. Pediatr.* **2025**, *51*, 4. [CrossRef] [PubMed]
8. Hotchkiss, R.S.; Monneret, G.; Payen, D. Sepsis-induced immunosuppression: From cellular dysfunctions to immunotherapy. *Nat. Rev. Immunol.* **2013**, *13*, 862–874. [CrossRef] [PubMed]
9. Torres, L.K.; Pickkers, P.; van der Poll, T. Sepsis-Induced Immunosuppression. *Annu. Rev. Physiol.* **2022**, *84*, 157–181. [CrossRef] [PubMed]
10. Binnie, A.; Walsh, C.J.; Hu, P.; Dwivedi, D.J.; Robichaud, A.F.; Liaw, P.C.; Tsang, J.; Batt, J.; Carrasqueiro, G.; Gupta, S.; et al. Epigenetic Profiling in Severe Sepsis: A Pilot Study of DNA Methylation Profiles in Critical Illness. *Crit. Care Med.* **2019**, *48*, 142–150. [CrossRef] [PubMed]
11. Binnie, A.; Tsang, J.; Hu, P.; Carrasqueiro, G.; Branco, P.C.; Santos, C.C.D. Epigenetics of Sepsis. *Crit. Care Med.* **2020**, *48*, 745–756. [CrossRef] [PubMed]
12. Sorolla, C.L.; García-Gómez, A.; Moll, F.C.; Toledano, V.; Ciudad, L.; Ortiz, J.A.; Eid, C.M.; Quirós, A.M.; Gallo, M.M.; Ruiz-Sanmartín, A.; et al. Inflammatory cytokines and organ dysfunction associate with the aberrant DNA methylome of monocytes in sepsis. *Genome Med.* **2019**, *11*, 66. [CrossRef] [PubMed]
13. Goldstein, B.; Giroir, B.P.; Randolph, A.G. International pediatric sepsis consensus conference: Definitions for sepsis and organ dysfunction in pediatrics. *Pediatr. Crit. Care Med.* **2005**, *6*, 2–8. [CrossRef] [PubMed]
14. Weiss, S.L.; Peters, M.; Alhazzani, W.; Agus, M.S.D.; Flori, H.R.; Inwald, D.; Nadel, S.; Schlapbach, L.J.; Tasker, R.C.; Argent, A.C.; et al. Surviving sepsis campaign international guidelines for the management of septic shock and sepsis-associated organ dysfunction in children. *Intensive Care Med.* **2020**, *46*, 10–67. [CrossRef] [PubMed]
15. Gu, H.; Smith, Z.D.; Bock, C.; Boyle, P.; Gnirke, A.; Meissner, A. Preparation of reduced representation bisulfite sequencing libraries for genome-scale DNA methylation profiling. *Nat. Protoc.* **2011**, *6*, 468–481. [CrossRef] [PubMed]
16. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J.* **2011**, *17*, 10. [CrossRef]
17. Xi, Y.; Li, W. BSMAP: Whole genome bisulfite sequence MAPping program. *BMC Bioinform.* **2009**, *10*, 232. [CrossRef] [PubMed]
18. Jühling, F.; Kretzmer, H.; Wolf, S.; Otto, C.; Stadler, P.F.; Hoffmann, S. metilene: Fast and sensitive calling of differentially methylated regions from bisulfite sequencing data. *Genome Res.* **2015**, *26*, 256–262. [CrossRef] [PubMed]
19. Langfelder, P.; Horvath, S. WGCNA: An R package for weighted correlation network analysis. *BMC Bioinform.* **2008**, *9*, 559. [CrossRef] [PubMed]
20. Wong, H.R.; Cvijanovich, N.; Lin, R.; Allen, G.L.; Thomas, N.J.; Willson, D.F.; Freishtat, R.J.; Anas, N.; Meyer, K.; Checchia, P.A.; et al. Identification of pediatric septic shock subclasses based on genome-wide expression profiling. *BMC Med.* **2009**, *7*, 34. [CrossRef] [PubMed]
21. Wynn, J.L.; Cvijanovich, N.Z.; Allen, G.L.; Thomas, N.J.; Freishtat, R.J.; Anas, N.; Meyer, K.; Checchia, P.A.; Lin, R.; Shanley, T.P.; et al. The influence of developmental age on the early transcriptomic response of children with septic shock. *Mol. Med.* **2011**, *17*, 1146–1156. [CrossRef] [PubMed]
22. Qiu, X.; Li, J.; Bonenfant, J.; Jaroszewski, L.; Mittal, A.; Klein, W.; Godzik, A.; Nair, M.G. Dynamic changes in human single-cell transcriptional signatures during fatal sepsis. *J. Leukoc. Biol.* **2021**, *110*, 1253–1268. [CrossRef] [PubMed]

23. Hao, Y.; Hao, S.; Andersen-Nissen, E.; Mauck, W.M.R.; Zheng, S.; Butler, A.; Lee, M.J.; Wilk, A.J.; Darby, C.; Zager, M.; et al. Integrated analysis of multimodal single-cell data. *Cell* **2021**, *184*, 3573–3587. [CrossRef] [PubMed]
24. Hu, C.; Li, T.; Xu, Y.; Zhang, X.; Li, F.; Bai, J.; Chen, J.; Jiang, W.; Yang, K.; Ou, Q.; et al. CellMarker 2.0: An updated database of manually curated cell markers in human/mouse and web tools based on scRNA-seq data. *Nucleic Acids Res.* **2023**, *51*, D870–D876. [CrossRef] [PubMed]
25. Wu, T.; Hu, E.; Xu, S.; Chen, M.; Guo, P.; Dai, Z.; Feng, T.; Zhou, L.; Tang, W.; Zhan, L.; et al. clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. *Innovation* **2021**, *2*, 100141. [CrossRef] [PubMed]
26. Osorio, D.; Zhong, Y.; Li, G.; Xu, Q.; Yang, Y.; Tian, Y.; Chapkin, R.S.; Huang, J.Z.; Cai, J.J. scTenifoldKnk: An efficient virtual knockout tool for gene function predictions via single-cell gene regulatory network perturbation. *Patterns* **2022**, *3*, 100434. [CrossRef] [PubMed]
27. Venet, F.; Monneret, G. Advances in the understanding and treatment of sepsis-induced immunosuppression. *Nat. Rev. Nephrol.* **2017**, *14*, 121–137. [CrossRef] [PubMed]
28. Kelly-Scumpia, K.M.; Scumpia, P.O.; Weinstein, J.S.; Delano, M.J.; Cuenca, A.G.; Nacionales, D.C.; Wynn, J.L.; Lee, P.Y.; Kumagai, Y.; Efron, P.A.; et al. B cells enhance early innate immune responses during bacterial sepsis. *J. Exp. Med.* **2011**, *208*, 1673–1682. [CrossRef] [PubMed]
29. Gustave, C.; Gossez, M.; Demaret, J.; Rimmele, T.; Lepape, A.; Malcus, C.; Poitevin-Later, F.; Jallades, L.; Textoris, J.; Monneret, G.; et al. Septic Shock Shapes B Cell Response toward an Exhausted-like/Immunoregulatory Profile in Patients. *J. Immunol.* **2018**, *200*, 2418–2425. [CrossRef] [PubMed]
30. Tomasini, R.; Samir, A.A.; Carrier, A.; Isnardon, D.; Cecchinelli, B.; Soddu, S.; Malissen, B.; Dagorn, J.C.; Iovanna, J.; Dusetti, N. TP53INP1s and homeodomain-interacting protein kinase-2 (HIPK2) are partners in regulating p53 activity. *J. Biol. Chem.* **2003**, *278*, 37722–37729. [CrossRef] [PubMed]
31. Saadi, H.; Seillier, M.; Carrier, A. The stress protein TP53INP1 plays a tumor suppressive role by regulating metabolic homeostasis. *Biochimie* **2015**, *118*, 44–50. [CrossRef] [PubMed]
32. Fontela, C.M.; Macip, S.; Sobrido, L.M.; Brown, L.M.; Ashour, J.; Sastre, A.G.; Lee, S.W.; Aaronson, S.A. Transcriptional role of p53 in interferon-mediated antiviral immunity. *J. Exp. Med.* **2008**, *205*, 1929–1938. [CrossRef] [PubMed]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.