Age-Stratified Transcriptomic Profiling Reveals Biologically Distinct Molecular Phenotypes Across Pediatric, Adolescent, and Adult Osteosarcoma

Xingyu Liu[1], Feiyang Qi[1], Boyang Wang[1], Haijie Liang[1], Siyi Huang[1], Wei Guo[1], Yi Yang[1], Jichuan Wang[1]

1 Musculoskeletal Tumor Center, Beijing Key Laboratory for Musculoskeletal Tumors, Peking University People's Hospital, Beijing, China

Significance & Hypothesis

Osteosarcoma (OS) has a bimodal age distribution with age-dependent clinical behavior, yet treatment remains largely age-agnostic. Understanding age-associated transcriptomic programs may inform risk stratification and guide age-adapted therapeutic strategies. We aimed to define age-associated transcriptomic programs and develop an age-informed prognostic tool.

Innovation

This study provides the first systematic age-stratified transcriptomic characterization of osteosarcoma across pediatric, adolescent, and adult groups, identifying distinct molecular phenotypes for each age stratum. It further introduces a pediatric-specific 10-gene prognostic score with external evaluation, offering a candidate biomarker for age-adapted risk stratification.

Approach (Materials & Methods)

Bulk RNA sequencing was performed on 70 primary OS specimens from a single center, stratified as pediatric (≤14 years, n=37), adolescent (15–18 years, n=22), and adult (≥19 years, n=11) (median follow-up 22.1 months) (Figure 1). Baseline clinicopathologic variables were comparable among groups (all $p>0.05$). Differential expression (DESeq2), pathway enrichment (Hallmark/GO GSEA and ssGSEA), and immune-cell deconvolution (CIBERSORT) were conducted. A pediatric-specific 10-gene score (MT1G, CTXN1, CENPV, CDCA7, HPDL, BNIP3, SIGMAR1, UCHL1, NPW, TAC3) was defined as the mean z-score across the 10 genes and evaluated using Kaplan–Meier and Cox regression, with external evaluation in TARGET-OS (n=87).

Results

Transcriptomic divergence increased with age distance (pediatric vs adult: 328 differentially expressed genes [DEGs]; pediatric vs adolescent: 154; adolescent vs adult: 162). Unsupervised clustering positioned adolescents between pediatric and adult clusters, suggesting a transitional molecular state. Pediatric tumors displayed a hyperproliferative program enriched for E2F targets, G2M checkpoint, and DNA replication pathways. Adolescent tumors showed the strongest immune–inflammatory activation (inflammatory response, TNF-alpha–NF-kB, B-cell receptor signaling) and higher monocyte infiltration by CIBERSORT ($p=0.05$), whereas pediatric and adult tumors showed relatively "cold" immune features. Adult tumors preferentially activated osteogenic differentiation programs (e.g., BMP/TGF-beta signaling). The pediatric 10-gene score declined stepwise with increasing age and was independently associated with inferior overall survival (OS: HR 5.6, 95% CI 1.2–26.2, $p=0.01$) and progression-free survival (PFS: HR 2.1, 95% CI 1.1–4.2, $p=0.03$) (Figure 2). Directionally consistent trends were observed in TARGET-OS (OS: HR 1.8, $p=0.08$; PFS: HR 1.6, $p=0.09$).

Discussion

OS comprises age-linked transcriptional phenotypes—hyperproliferative pediatric, immune-active adolescent, and differentiation-oriented adult—supporting age-adapted risk stratification and trial design. The pediatric 10-gene score is a candidate prognostic biomarker warranting prospective multi-center validation before clinical implementation.

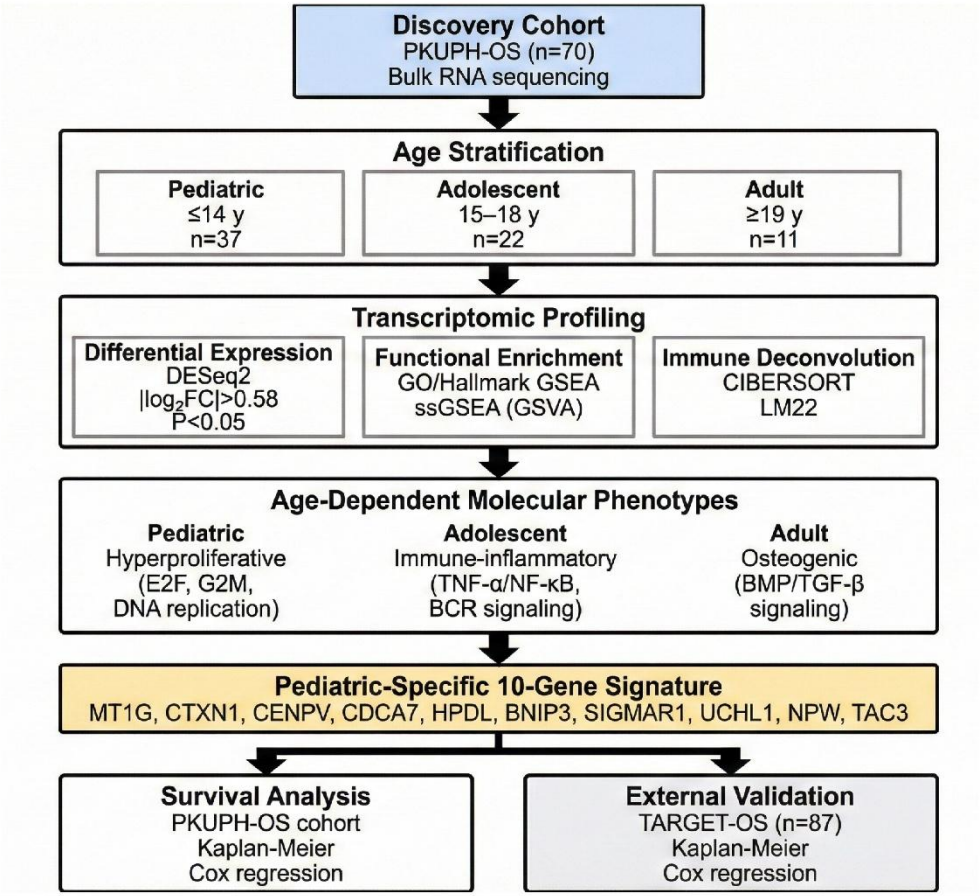

Figure 1. Schematic overview of the study design. Overview of the study workflow, in-cluding patient cohort stratification into pediatric (≤14 years), adolescent (15–18 years), and adult (≥19 years) groups. The diagram illustrates the integrated pipeline of bulk RNA sequencing, differential expression analysis, immune microenvironment characterization, and the development and validation of the pediatric-specific 10-gene prognostic signature.

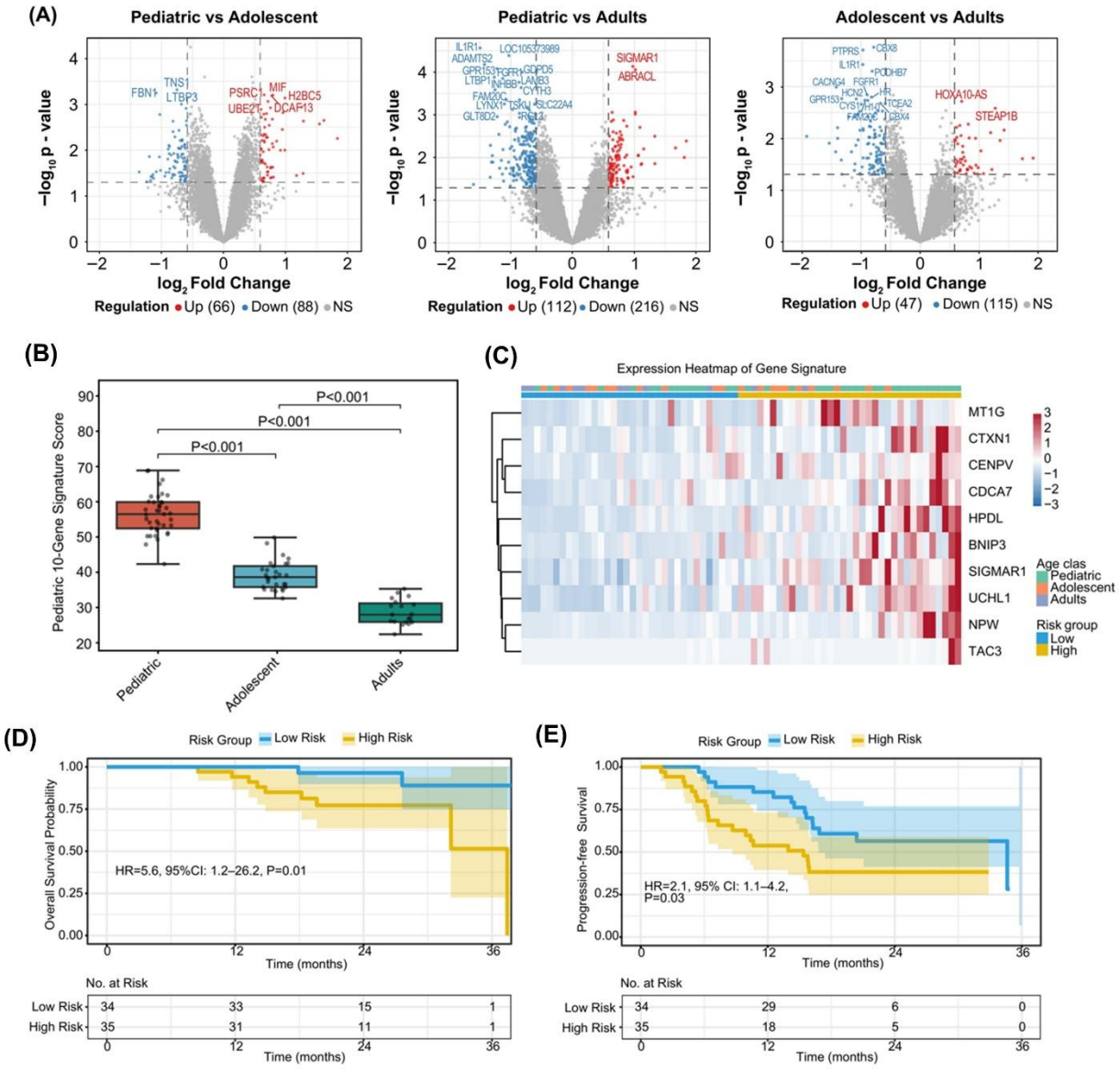

Figure 2. (A) Volcano plots showing differentially expressed genes (DEGs). Red dots indicate significantly upregulated genes, while blue dots indicate downregulated genes, defined by |log2FC| > 0.58 and adjusted p < 0.05. (B) Distribution of risk scores calculated by the 10-gene signature across the age groups. Pediatric patients exhibit significantly higher risk scores compared to adolescent and adult patients (p < 0.001). (C) Heatmap showing the expression of the 10 signature genes (MT1G, CTXN1, CENPV, CDCA7, UCHL1, HPDL, BNIP3, SIGMAR1, NPW, TAC3) in high-risk vs. low-risk groups. (D) Kaplan–Meier curve for overall survival (OS) in the discovery cohort. Patients are stratified into high-risk and low-risk groups based on the median risk score. Hazard ratio (HR) and log-rank p-value are shown. (E) Kaplan–Meier curve for progression-free survival (PFS) in the discovery cohort, demonstrating the prognostic utility of the signature in predicting tumor progression.