REGULAR ARTICLE

blood advances

Immune profiling after allogeneic hematopoietic cell transplantation in pediatric acute myeloid leukemia

Sanam Shahid,[1,*] Nicholas Ceglia,[2,*] Jean-Benoît Le Luduec,[3] Andrew McPherson,[4] Barbara Spitzer,[1] Theodota Kontopoulos,[3] Viktoria Bojilova,[2] M. Kazim Panjwani,[3] Mikhail Roshal,[5] Sohrab P. Shah,[4] Omar Abdel-Wahab,[6-8] Benjamin Greenbaum,[4] and Katharine C. Hsu[3,6,7,9]

[1]Department of Pediatrics, [2]Memorial Hospital Research, [3]Immunology Program, Sloan Kettering Institute, [4]Department of Epidemiology-Biostatistics, and [5]Department of Pathology, Memorial Sloan Kettering Cancer Center, New York, NY; [6]Department of Medicine, New York Presbyterian Hospital Weill Cornell Medical Center, New York, NY; and [7]Department of Medicine, [8]Center for Hematologic Malignancies, and [9]Human Oncology and Pathogenesis Program, Memorial Sloan Kettering Cancer Center, New York, NY

Key Points

- Downregulation of major histocompatibility complex class II expression is observed at posttransplant pediatric AML relapse.
- Natural killer cells and CD8$^+$ T-cell subsets are dysfunctional at posttransplant pediatric AML relapse.

Although allogeneic hematopoietic cell transplant (allo-HCT) is curative for high-risk pediatric acute myeloid leukemia (AML), disease relapse remains the primary cause of posttransplant mortality. To identify pressures imposed by allo-HCT on AML cells that escape the graft-versus-leukemia effect, we evaluated immune signatures at diagnosis and posttransplant relapse in bone marrow samples from 4 pediatric patients using a multimodal single-cell proteogenomic approach. Downregulation of major histocompatibility complex class II expression was most profound in progenitor-like blasts and accompanied by correlative changes in transcriptional regulation. Dysfunction of activated natural killer cells and CD8$^+$ T-cell subsets at relapse was evidenced by the loss of response to interferon gamma, tumor necrosis factor α signaling via NF-κB, and interleukin-2/STAT5 signaling. Clonotype analysis of posttransplant relapse samples revealed an expansion of dysfunctional T cells and enrichment of T-regulatory and T-helper cells. Using novel computational methods, our results illustrate a diverse immune-related transcriptional signature in posttransplant relapses not previously reported in pediatric AML.

Introduction

Pediatric acute myeloid leukemia (AML) is an aggressive cancer characterized by a rapid accumulation of immature cells of the myeloid lineage, representing from 15% to 20% of all pediatric acute leukemias.[1] Although allogeneic hematopoietic cell transplantation (allo-HCT) is curative for high-risk AML in children, posttransplantation relapse occurs in ~20% to 50% patients, comprising the primary cause of treatment failure and mortality.[2] The efficacy of allo-HCT in curing hematological malignancies partly relies on transferring an immune system from donor to patient that is capable of eliminating residual tumor cells via a mechanism known as the graft-versus-leukemia effect.[3] Computational advances permit the study of tumor evolution and relapsed disease in the context of immune-mediated selective pressures imposed by allo-HCT.

Submitted 5 December 2022; accepted 24 May 2023; prepublished online on *Blood Advances* First Edition 16 June 2023; final version published online 28 August 2023. https://doi.org/10.1182/bloodadvances.2022009468.

*S.S. and N.C. contributed equally to this study.

Presented in abstract form at the 64th annual meeting of the American Society of Hematology, New Orleans, LA, 10-13 December 2022.

Sequencing data are deposited in the Gene Expression Omnibus database (accession number GSE207104).

Data are available on request from the corresponding author, Katharine C. Hsu (hsuk@mskcc.org).

The full-text version of this article contains a data supplement.

Current evidence in adult patients with AML suggests that relapse after allo-HCT is mediated by a variety of immune escape mechanisms, including loss or altered expression of major histocompatibility complex (MHC) molecules, aberrant cytokine signaling, and expression of immune checkpoints that suppress immune response against leukemic cells.[4,5] Specifically, AML cells in adult patients with post-HCT relapse demonstrate downregulation of MHC class II (MHC-II) molecules, downregulation of activating markers, such as CD11A and LFA-1, and upregulation of inhibitory markers, including PD-L1 and B7-H3.[4] T cells from the same patients exhibit increases in exhaustion patterns, including upregulation of PD-1.[4] Meanwhile, in natural killer (NK) cells, target cell–induced interferon gamma (IFN-γ) production appears markedly diminished in multiple allo-HCT settings (T-cell–replete, T-cell–depleted, and umbilical cord blood transplantation).[6] These studies have exclusively addressed adult patients with AML.

Major differences in the disease spectrum between adult and pediatric patients with AML can be attributed to dissimilar environmental and genetic factors.[7] The homeostatic environment and immune system of a fully mature host differ significantly from that of a developing host, in whom a susceptible stem cell population experiences reduced protection from a less mature innate and adaptive immune system.[7,8] In addition, pediatric AML is typified by molecular drivers, including structural chromosomal rearrangements leading to expressed fusion genes, whereas adult AML largely results from acquisition of sequential somatic mutations over time.[9-11] These major differences may result in distinctively dysregulated biological processes within the leukemia cell, affecting cell differentiation, proliferation, apoptosis, and immune interaction.

Elucidating the factors that contribute to the ability of leukemia cells in children to escape immune recognition in allo-HCT is vital to the success of transplantation and adoptive cellular therapies for relapsed disease. We evaluated bone marrow aspirate (BMA) samples obtained at diagnosis and at posttransplant relapse for 4 pediatric patients with AML at Memorial Sloan Kettering Cancer Center (MSKCC). Using cellular indexing of transcriptomes and epitopes by sequencing (CITE-seq)[12] and T-cell receptor (TCR) sequencing, we identify several transcriptional and phenotypic alterations that may permit pediatric AML cells to escape the graft-versus-leukemia effect.

Methods

Patient samples

To identify patients for inclusion in this study, we performed a retrospective review of pediatric patients with AML who relapsed after receiving an allogeneic HCT between 2008 and 2020 at MSKCC. Cryopreserved paired samples of BMAs performed at the time of diagnosis and relapse after allo-HCT were obtained from MSKCC's Hematologic Oncology Tissue Bank and New York Presbyterian Hospital Weill Cornell Medical Center Hematopathology Department. All participants or their parents provided written consent for banking and use of these specimens for future research following research sample procurement and genomic profiling protocols, IRB#06-107 and 12-245, respectively, in accordance with the regulations of the MSKCC Institutional Review Board. In all cases, a retrospective review of clinical parameterswas performed, and cytogenetic, molecular genetic, and diagnostic flow cytometry results were obtained. Clinical follow-up information was obtained via retrospective review of the medical record under IRB#16-1564. Frozen primary human whole bone marrow mononuclear cells (from a 20-year-old female) were obtained from STEMCELL Technologies and was used as a healthy donor control.

Frozen bone marrow mononuclear cell preparation

Frozen human bone marrow specimens were thawed and transferred into 5 mL polypropylene fluorescence-activated cell sorting (FACS) tubes and resuspended in 4 mL phosphate-buffered saline (PBS) + 2 mM EDTA + 2% fetal bovine serum (FBS) (staining buffer). Cell suspensions were centrifuged at 500 *g* for 5 minutes at 4°C, and the supernatant was discarded (all subsequent wash steps were performed the same way). Cell pellets were resuspended in staining buffer and counted. For each sample with $>10^6$ live cells, 70% was used for T-cell sorting (described in "T-cell sorting"). The remaining 30% of the samples, or the entire sample for those with $<10^6$ live cells, were washed and then subjected to dead cell removal (Miltenyi Biotec). Dead cells were depleted per the manufacturer's instructions, with the exception of the incubation period shortened to 5 minutes and column washes reduced to 3 times. The live cell fraction was washed, resuspended in staining buffer, and counted. Samples were resuspended in PBS + 2% FBS (labeling buffer) and Human FcR Blocking Reagent (1:50, Miltenyi Biotec) and incubated on ice for 5 minutes. Samples were tagged with cell hashing (Stoeckius et al, 2018) oligo-tagged antibodies (TotalSeq-C0251 anti-human hashtag 1 antibody and TotalSeq-C0252 anti-human hashtag 2 antibody; BioLegend) per the manufacturer's instructions. After tagging, samples were washed 3 times in labeling buffer, and then each patient's samples were pooled together and washed. Samples were resuspended in labeling buffer and Human FcR Blocking Reagent (1:50, Miltenyi Biotec) and incubated on ice for 5 minutes. CITE-Seq antibodies (TotalSeq-C Human Universal Cocktail, BioLegend) were stained per the manufacturer's instructions, using the aforementioned labeling buffer for all steps. After staining, samples were washed 3 times in labeling buffer and resuspended at a concentration of 10^3 cells per μL in cold PBS + 0.04% bovine serum albumin (sequencing buffer).

T-cell sorting

Samples were incubated with Human FcR Blocking Reagent (1:50; Miltenyi Biotec) in staining buffer on ice for 10 minutes. Phycoerythrin-conjugated anti-human CD45 (clone HI30; to a final concentration of 1:50 in staining buffer; BD Biosciences), BV650-conjugated anti-human CD3 (clone UCHT1; to a final concentration of 1:100 in staining buffer; BD Biosciences), and fluorescein isothiocyanate–conjugated anti-human CD14 (clone M5E2; to a final concentration of 1:50 in staining buffer; BD Biosciences) were added to the samples and incubated on ice for 20 minutes. After incubation, cell suspensions were washed twice with staining buffer, resuspended at a concentration of 10^6 cells per 100 μL of staining buffer containing 4′,6-diamidino-2-phenylindole (1 μg/mL) for viability staining, and then passed through a 35 μm nylon mesh filter to remove cell clumps. Sorting was performed using FACSAria Cell Sorter (BD Biosciences). T cells were identified and sorted into 5 mL polypropylene FACS tubes coated with 350 μL

FBS on the basis of forward scatter and side scatter, singlets, viability (4′,6-diamidino-2-phenylindole^{low}), CD45^{+}, CD14^{-}, and CD3^{+}. After cell sorting, individual samples were centrifuged at 700 *g* at 4°C for 7 minutes, counted, tagged for cell hashing as described earlier, and resuspended at a concentration of 10^3 cells per μL in sequencing buffer.

Sequencing

Single-cell transcriptome sequencing. Fresh cells were stained with Trypan blue, and Countess II Automated Cell Counter (Thermo Fisher Scientific) was used to assess both cell number and viability. After quality control, the single-cell suspension was loaded onto a Chromium Next GEM Chip K (10X Genomics PN 1000286). GEM generation, complementary DNA (cDNA) synthesis, cDNA amplification, and library preparation of ~2000 to 14 000 cells was done using the Chromium Next GEM Single Cell 5′ Kit version 2 (10X Genomics PN 1000263) in accordance with the manufacturer's protocol. cDNA amplification included 13 or 14 cycles, and between 38 and 50 ng of the material was used to prepare sequencing libraries with 14 cycles of polymerase chain reaction (PCR). Indexed libraries were pooled equimolar and sequenced on a NovaSeq 6000 in a PE28/91 run using the NovaSeq 6000 S1 or S2 Reagent Kit (100 or 200 cycles) (Illumina). An average of 263 million paired reads were generated per sample.

Single-cell V(D)J analysis from RNA. An aliquot of cDNA generated using the methods described above was used to enrich for V(D)J regions using the Chromium Single Cell Human TCR Amplification Kit (10X Genomics PN 1000252) per the manufacturer's protocol, with 10 cycles of PCR during enrichment and 9 cycles during library preparation. Indexed libraries were pooled equimolar and sequenced on a NovaSeq 6000 in a PE28/91 run using the NovaSeq 6000 S2 Reagent Kit (100 cycles; Illumina). An average of 44 million reads were generated per sample.

Cell surface protein feature barcode analysis. Amplification products generated using the previously described methods included both cDNA and feature barcodes tagged with cell barcodes and unique molecular identifiers (UMIs). Smaller feature barcode fragments were separated from longer amplified cDNA using a 0.6X cleanup using aMPure XP beads (Beckman Coulter catalog # A63882). Libraries were constructed using the 5′ Feature Barcode Kit (10X Genomics PN 1000256) per the manufacturer's protocol, with 9 cycles of PCR. Indexed libraries were pooled equimolar and sequenced on a NovaSeq 6000 in a PE28/91 run using the NovaSeq 6000 S2 Reagent Kit (100 cycles; Illumina). An average of 36 million reads were generated per sample.

Data preprocessing

Twenty-two sequencing samples from 5 individuals (including a healthy donor) were registered in the Isabl platform[13] (a platform for the integration, management, and processing of individual-centric multimodal data). Metadata, such as experimental techniques (single-cell RNA sequencing [RNA-seq] VDJ, single-cell CITE-seq, and single-cell RNA-seq), sample category (tumor/germ line), sample ID, cohort ID, and disease information were recorded in a secured relational database.

Single-cell analysis

CellRanger software (version 6.0.1) was used to perform read alignment, barcode filtering, and UMI quantification using the 10X GRCh38 transcriptome (version 3.0.0) for gene expression. Paired TCR sequences and protein expression levels were generated using 10X CellRanger VDJ (version 4.0.0) and CellRanger Feature Barcoding (3.1.0), respectively. Filtered matrixes were processed using the Scanpy python package (version 3.0.1).[14] The resulting gene-by-cell matrix was log normalized and merged for each patient. Protein expression from CITE-seq was normalized using the centered log ratio method.[15] Cells with <300 total UMIs and genes found in <3 cells were excluded from the analysis.

Individual cells were classified to major cell types using a set of marker genes (supplemental Table 1) using GeneVector.[16,17] The AML blast cells were further divided into subtypes using annotation transfer within the ingest pipeline available in Scanpy python library[14] with a single-cell RNA-seq, AML-specific reference data set of 6 blast subtypes.[18] T-cell specific phenotypes were probabilistically assigned using a collection of referenced marker genes,[19] and NK activation and exhaustion signature were manually curated. Cell-type annotated matrixes for individual patients across timepoints were integrated using Harmony (version 0.1) into a single batch-corrected matrix. Dimensionality reduction and visualization as a uniform manifold approximation and projection (UMAP) embedding was performed using the Scanpy python package.[14] Differentially expressed genes (DEGs; $P < .001$) were computed using the Wilcoxon test both at diagnosis and post-transplantation relapse for each phenotype and among patients.

Pathway enrichment and gene set enrichment analysis plots were generated using the Gseapy python package,[20] with gene lists collected from Enrichr libraries on DEGs in our data set and from 2 additional articles based on adult patients using the reported DEGs.[21-23] Doublets were calculated using the Scrublet python package, with a maximum score of 0.3.[24] Clonotype annotations were assigned to barcodes using the Scirpy python library.[20] Transcription factors with binding sites found 10 kb upstream and 1 kb downstream of transcription start sites for MHC-II genes were identified using Motifmap[25] and annotated as activators or repressors using the TRRUST database.[26] Transcriptional activators, repressors, and MHC-II genes were assigned scored values for the 6 AML blast subtypes using the Seurat gene scoring module as implemented in Scanpy.[27]

Individual TCR β (TRB) clonotypes were defined from the TRB CDR3 amino acid sequence and processed using Scirpy,[27] defined by a unique amino acid sequence. Public clonotypes were defined as any clonotype found in >1 patient. Generation probabilities for individual TRB sequences was computed using OLGA python package.[28]

Results

Clinical information for the 4 pediatric patients with AML (UPN1-4) is shown in supplemental Table 2. UPN1-3 achieved sustained full donor chimerism before relapse, whereas UPN4 did not. A total of 79 334 single-cell transcriptomes were profiled from 9 BMA samples obtained from the 4 patients at diagnosis (n = 45 345) and posttransplantation relapse (n = 15 412) and from 1 healthy 20-year-old female donor (n = 18 577). For BMA samples with

$\geq 10^6$ viable cells present, a fraction of the sample was flow cytometrically sorted for T cells. Detailed profiling of the leukemia-associated immune compartment was performed, including TCR-seq (supplemental Figures 1 and 2).

Downregulation of MHC-II gene expression in pediatric AML blasts at posttransplantation relapse compared with that at disease diagnosis

We performed an initial probabilistic cell-type assignment using a set of marker genes (supplemental Table 1) for the major cell types (T cells, NK cells, normal myeloid cells, B cells, and AML blast cells) (Figure 1A). The distribution of cell types was diverse across patients and between diagnosis and relapse samples per patient. Major cell–type assignments were validated based on relative expression levels of canonical markers encompassing lineage-specific gene and surface markers (supplemental Figure 3). Although known AML blast transcriptional markers are derived from adult samples, we found that those cells labeled as AML blasts in our samples from pediatric patients also expressed high levels of myeloperoxidase (MPO), elastase, neutrophil expressed (ELANE), and azurocidin 1 (AZU1). The healthy myeloid and AML blast clusters in Figure 1A appear to cluster differently for each patient (supplemental Figure 4A), rather than at different time points (supplemental Figure 4B). Each myeloid cluster (supplemental Figure 4C) contains known myeloid/AML markers (supplemental Figure 4D-E). Although there is limited availability of single-cell transcriptional data for pediatric AML blasts, we validated our analysis by comparing genes upregulated in pediatric leukemic blasts vs healthy stem cells in our data with that previously generated from pediatric AML microarray data.[29] We found that of the 50 genes reported by Depreter et al, 5 genes were significantly upregulated in our pediatric AML blast analysis: ATF3, CFD, EMILIN2, FOSB, and TUBB6 ($P = .02$).

Pathway enrichment analysis identified several significantly downregulated immune-related processes in AML blasts at disease relapse after allo-HCT compared with that at initial disease diagnosis, including antigen processing and presentation, cytokine-mediated signaling, and IFN-γ mediated signaling (Figure 1B). Interestingly, the MHC-II antigen presentation pathways were common to pediatric and adult post-HCT relapse[4,5] (Figure 1C). Overall downregulation of MHC-II genes HLA-DR, -DP, -DQ, -DM, and -DO and MHC-II regulators CIITA and CD74 was observed on comparing mean expression at diagnosis with that at posttransplantation relapse (Figure 1D-E). However, there was variability among patients regarding the specific MHC-II genes downregulated (supplemental Figure 5). Although this downregulation was not significant in the CITE-seq protein expression data (supplemental Figure 6), it was readily apparent via flow cytometry (supplemental Figure 7).

It has been reported in adult AML that the absence of NKG2D ligands defines leukemia stem cells and mediates their immune evasion,[30] so we evaluated the change in expression level of NKG2D ligands (MICA, MICB, ULBP1, ULBP2, and ULBP3) on AML blasts between diagnosis and posttransplantation relapse. Although there was a decrease in gene expression level of all 5 ligands at posttransplantation relapse, only MICA was statistically significantly downregulated ($P = .01$) in our data (supplemental Figure 8).

Downregulation of MHC-II gene expression in pediatric AML blast subtypes

Because of known intratumoral heterogeneity in AML,[31,32] we sought to evaluate whether a specific subpopulation of AML blasts was driving the downregulation of MHC-II at relapse. Using annotation transfer on a published AML single-cell RNA-seq data,[18] previously identified blast cells from patients with AML (Figure 1A) were reprocessed and stratified into 6 subtypes along the myeloid differentiation axis (hematopoietic stem cell–like, progenitor-like, granulocyte-monocyte progenitor–like, promonocyte-like, monocyte-like, or cDC-like; Figure 2A-B). Mean MHC-II expression was computed to ascertain overall MHC-II gene activation for each subtype in all 4 patients (Figure 2C).[27]

Comparative analysis revealed that progenitor-like AML blasts exhibited the lowest MHC-II gene expression at diagnosis, which remained constant in posttransplantation relapse samples. Interestingly, in 3 of 4 patients, there was an increase in the proportion of progenitor-like blasts at posttransplantation relapse compared with that at diagnosis (Figure 2B). The statistical significance for each AML blast subtype change from diagnosis to posttransplantation relapse was evaluated using Fisher exact test (supplemental Table 3). Furthermore, we found that the 3 representative markers of progenitor-like blasts (ELANE, MPO, and AZU1) were relatively increased in frequency and/or expression level at posttransplantation relapse in UPN2, UPN3, and UPN4 (Figure 2D). We found that this change was significant in UPN2 and UPN4, as compared with that inUPN3, who had a comparatively lower number of AML blasts recovered (supplemental Figure 9). There was a statistically significant increase in ELANE, MPO, and AZU1 in AML blasts as compared with healthy myeloid cells (supplemental Figure 10).

Previously published transcription factors targeting ≥ 1 MHC-II pathway genes were then annotated as activators or repressors in our data set[25,26] (Figure 2E). The regulatory network constructed for upstream MHC-II activators (MYC, RELA, TP53, CREBBP, EZH2, IRF1, and CIITA) and repressors (MYCN and HDAC1) highlighted significant differences in regulatory expression between diagnosis and posttransplantation relapse stages within progenitor-like AML blasts (supplemental Figure 11).

Bone marrow environment immune dysfunction at posttransplantation relapse in pediatric patients with AML

To elucidate the immune pressures leading to the MHC-II changes in AML blasts, we characterized the bone marrow immune microenvironment using single-cell CITE-seq analysis (Figure 3A). Bone marrow from a healthy donor served as a comparison (supplemental Figure 12). T cells from BMA samples were separated into $CD8^+$ and $CD4^+$ populations, alongside NK cells ($CD56^+$), using CITE-seq cell surface protein expression. Individual phenotypes for T and NK cells were characterized by probabilistic assignment, using sets of known representative genes (supplemental Table 4).[19] After categorization into naive, memory, activated, and dysfunctional $CD8^+$ phenotypes, distinct subpopulations of $CD8^+$ T cells associated with high MHC-II expression or IFN-stimulated genes were then identified within activated or dysfunctional phenotypes. When compared to BMA samples

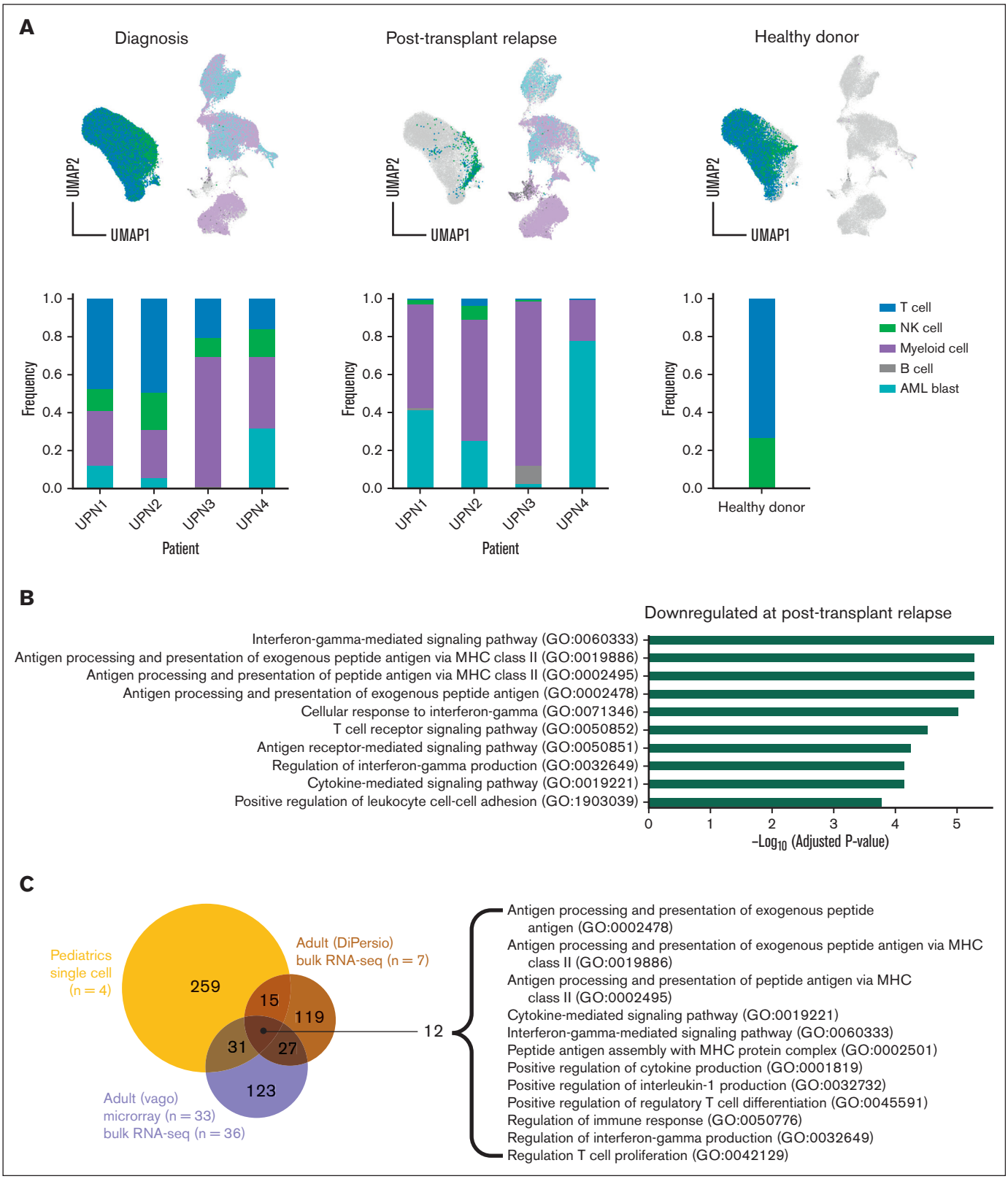

Figure 1. Downregulation of MHC-II gene expression in pediatric AML blasts at posttransplantation relapse compared with that at disease diagnosis. (A) UMAP plots and bar graphs of 5 transcriptionally distinct cell clusters/broad hematopoietic cell types identified in BMA samples in pediatric patients with AML at the time of diagnosis and posttransplantation relapse (n = 4) and in healthy donor control (n = 1). (B) Histogram outlining the top 10 most significantly downregulated gene ontology (GO) biological processes identified from the pairwise comparison of AML blasts from patients at disease diagnosis and at relapse after allo-HCT. (C) Venn diagram comparing GO pathways

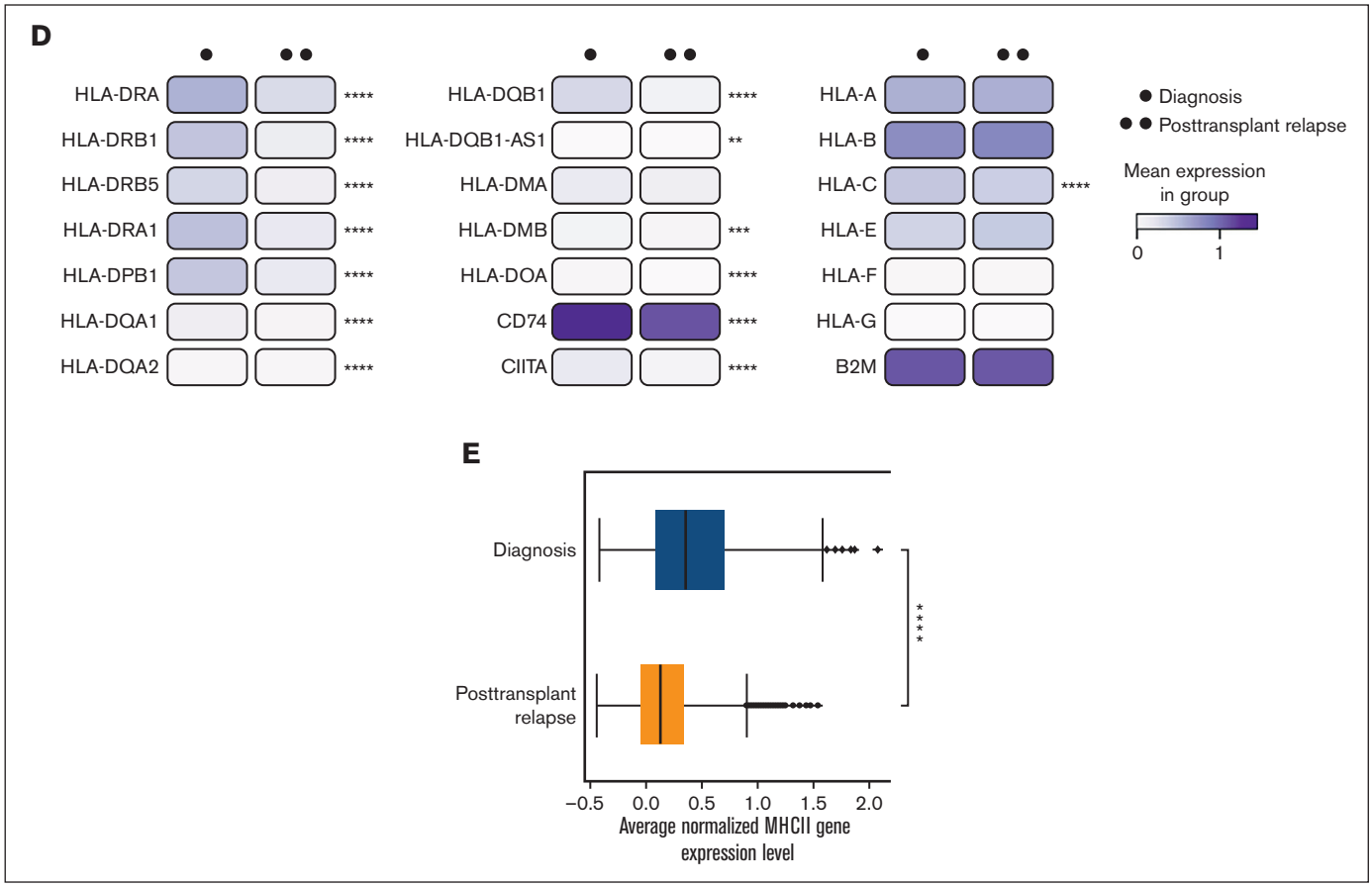

Figure 1 (continued) downregulated in 2 data sets of adult patients with AML with that in pediatric patients with AML at posttransplantation relapse. Twelve shared GO pathways among all 3 data sets are shown. (D) Heatmap representing MHC class I MHC-II gene expression at diagnosis and posttransplantation relapse in all pediatric AML blasts. (E) Box plot showing decreased average normalized MHC-II gene expression level in all pediatric AML blasts at posttransplantation relapse as compared with that at diagnosis. *$P < .05$; **$P < .01$; and *** $P < .001$.

obtained at diagnosis, samples obtained from patients who relapsed after transplantation revealed a relative expansion of exhausted CD8$^+$ cells (Figure 3B), suggesting recruitment of cytotoxic T cells to the bone marrow environment and lower effector functionality at the time of relapse after transplantation.

CD4$^+$ T cells were accordingly assigned into naive, regulatory, and helper phenotypes. In comparison to diagnosis samples, samples obtained from patients who relapsed after transplantation revealed an over-representation of regulatory T cells with a relative decrease in naive CD4$^+$ cells (Figure 3C), findings consistent with a relatively immunosuppressive tumor microenvironment. Although the ratio of activated to exhausted NK cells was similar at both timepoints, the frequency of NK cells was enriched at relapse after transplantation (Figure 3A, D; supplemental Figure 13). Pathway enrichment using a combined enrichment score[21] validated the overall dysfunctional pattern in T cells and NK cells at relapse after transplantation with lower observed response to IFN-γ, tumor necrosis factor α signaling via NF-κB, and interleukin-2 [IL-2]/STAT5 signaling in activated NK cells and lower observed response to IFN-γ and IL-2/STAT5 in activated T cells (Figure 3E).

Evolution of TCR clonotypes at diagnosis and relapse after transplantation in pediatric patients with AML

Phenotype assignment was coupled with TCR clonotype in all 4 patients at diagnosis and in 2 at relapse after transplantation. We observed a significant number of clonotypes present across >1 patient (Figure 4A), suggesting a conserved function and further highlighted by high generation probabilities (supplemental Figure 14).[28] Notably, the top 10 TCR clones by size across all 4 patients primarily were activated MHC-II$^+$ CD8$^+$ T cells (Figure 4B).[33]

Several TCR clonotypes were identified both at diagnosis and at relapse after transplantation within a single patient, UPN4 (Figure 4C-D; supplemental Table 5), demonstrating a persistence of host T cells despite ablative therapy. The evolution of each clonotype in UPN4 could be tracked from diagnosis to relapse after transplantation by defining the empirical probability of each phenotype at each timepoint. Persistent clones shifted from an activated to dysfunctional phenotype within MHC-II$^+$ CD8$^+$ T cells and from a memory or naive phenotype to an activation signature

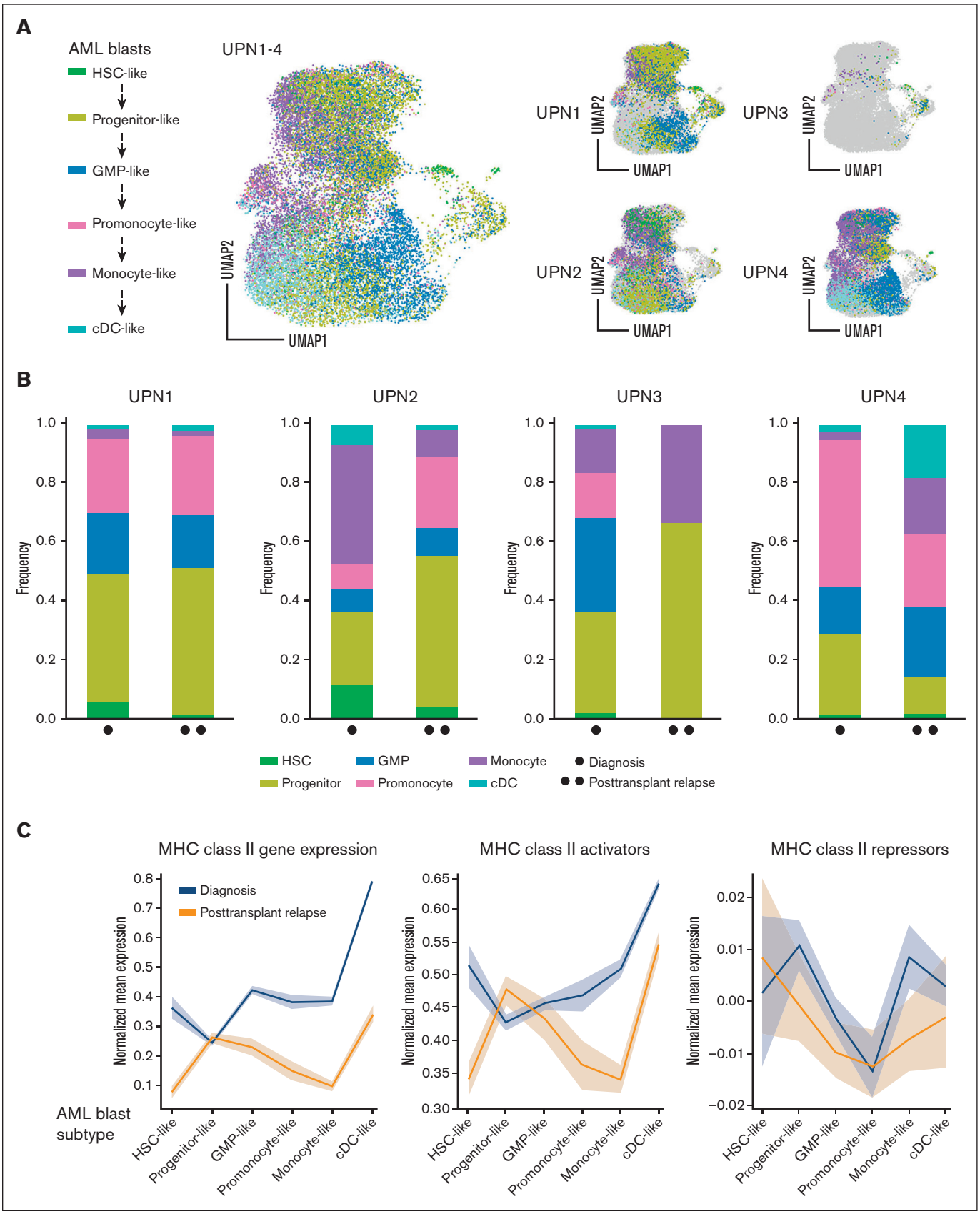

Figure 2. Downregulation of MHC-II gene expression in pediatric AML blast subtypes. (A) UMAP plots showing malignant myeloid cells stratified into 6 subtypes along the myeloid differentiation axis for UPN1-4. (B) Bar graphs showing relative composition of 6 malignant AML blast subtypes within each patient, UPN1-4, at diagnosis, and at

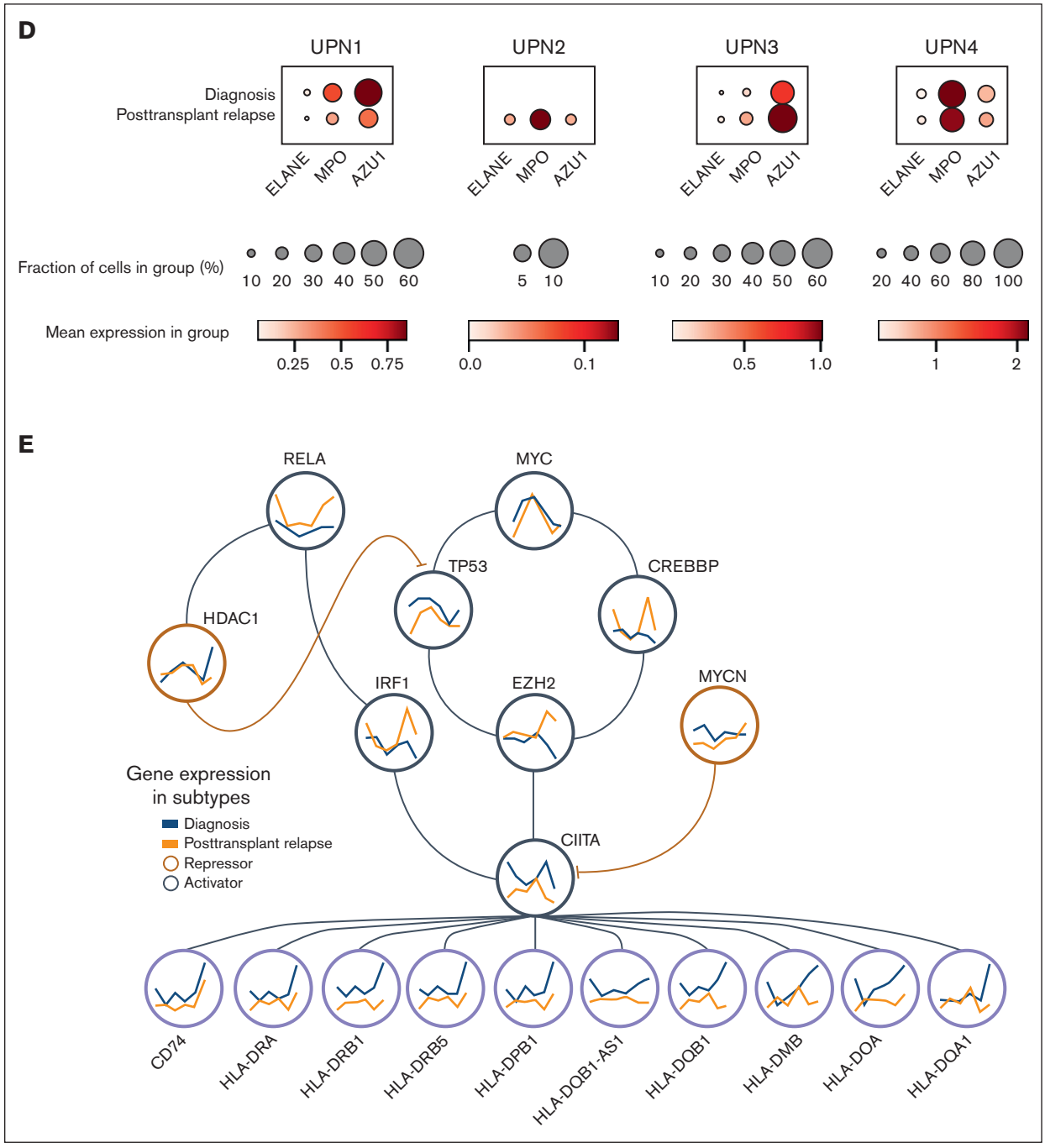

Figure 2 (continued) posttransplantation relapse. (C) Normalized mean gene expression profiles for MHC-II genes and upstream MHC-II activators/repressors across 6 malignant AML blast subtypes at diagnosis (blue) and posttransplantation relapse (orange). (D) Dot plots showing 3 representative markers of progenitor-like AML blasts (ELANE, MPO, and AZU1), with relative frequency and expression levels at diagnosis vs that at posttransplantation relapse for UPN1-4. (E) Transcriptional regulation of MHC-II genes showing MHC-II activators (gray circles) and repressors (orange circles) highlighting gene expression changes across 6 malignant AML blast subtypes at time of diagnosis (blue) and posttransplantation relapse (orange). AZU, azurocidin; cDC, classical dendritic cell; ELANE, elastase, neutrophil expressed; GMP, guanosine monophosphate; HSC, hematopoietic stem cell; MPO, myeloperoxidase.

specific to IFN-stimulated genes (Figure 4E). Phenotypic evaluation of clones shared between diagnosis and posttransplantation relapse in UPN4 revealed that the majority were MHC-II$^+$, CD8$^+$-exhausted T cells (Figure 4F). Taken together, the data reveal dynamic changes among MHC-II$^+$ CD8$^+$ cells in the tumor microenvironment, consistent with activation-induced exhaustion.

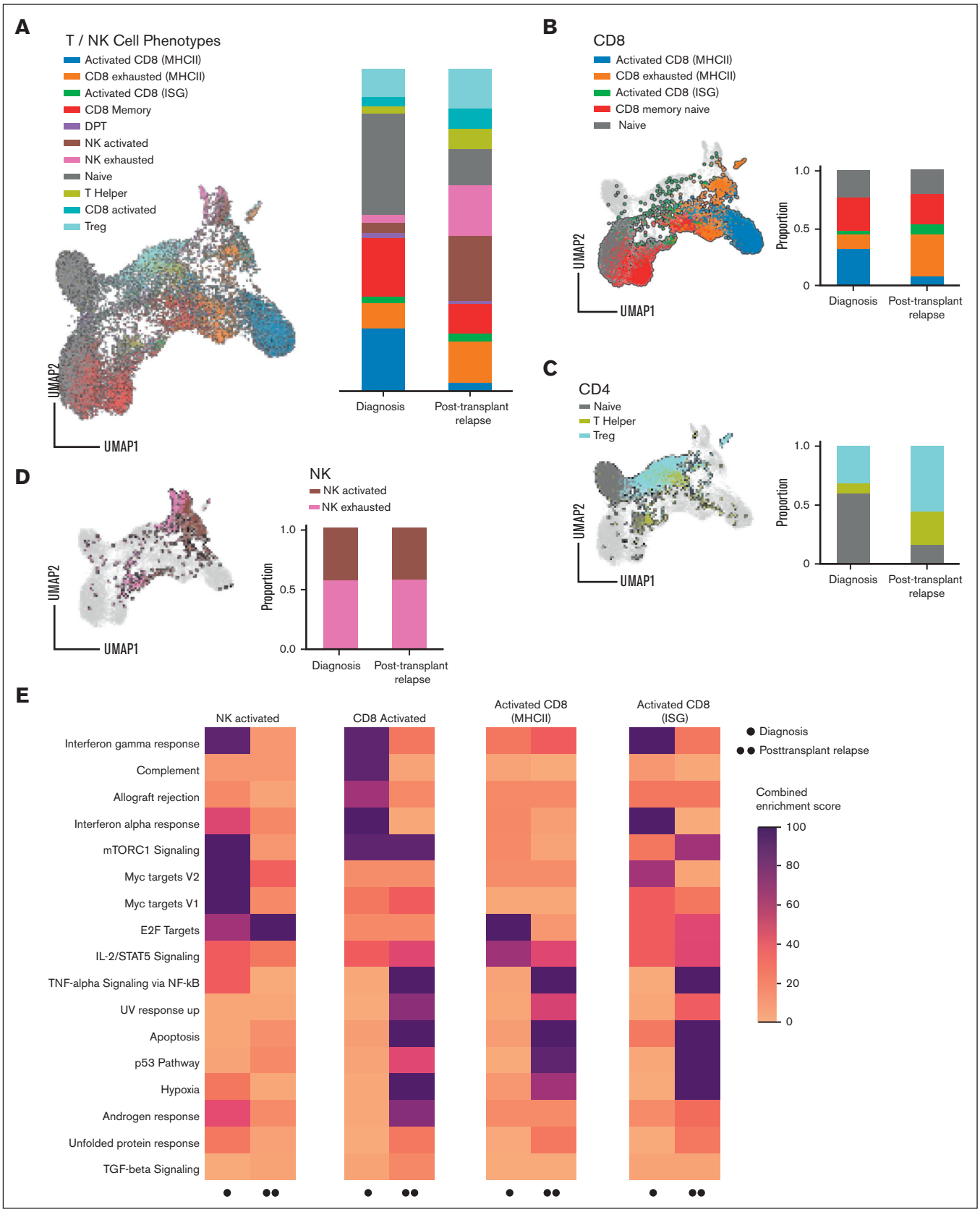

Figure 3. Bone marrow environment immune dysfunction at posttransplantation relapse in pediatric patients with AML. (A) Single-cell bone marrow immune microenvironment of (B) CD8$^+$ T cells, (C) CD4$^+$ T cells, and (D) NK cells for UPN1-4 at diagnosis and posttransplantation relapse. (E) Top differentially expressed Hallmark pathways between diagnosis and posttransplantation relapse in T and NK cells. Scale from 0 to 100 is a combined enrichment score $c = \log(p) \times z$, in which c is the combined

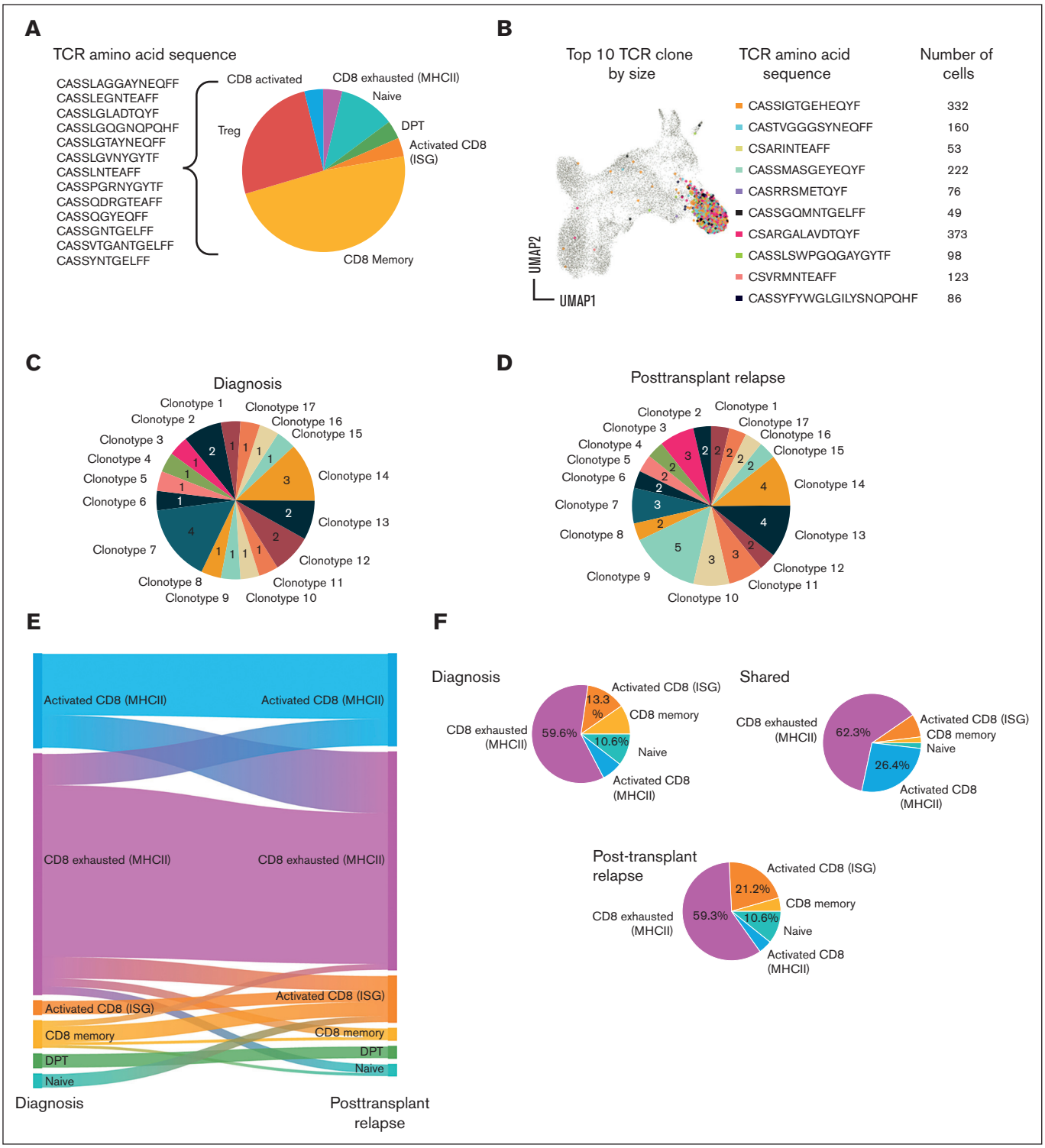

Figure 4. Evolution of TCR clonotypes at diagnosis and posttransplantation relapse in pediatric patients with AML. (A) Pie chart showing public TCR clonotypes found across >1 patient. (B) UMAP, TRB amino acid sequence, and cell counts for top 10 TCR clones based on size across all 4 patients. (C-D0 Pie charts showing number of each TCR clonotype identified at diagnosis (C) and at posttransplantation relapse (D) within a single patient, UPN4. (E) Alluvial description of T-cell phenotype probabilities within clonotypes from diagnosis to posttransplantation relapse for UPN4. (F) Phenotype of TCR clonotypes identified in diagnosis and/or posttransplantation relapse for UPN-4. ISG, IFN-stimulated gene.

Figure 3 (continued) score, *p* is the *P* value computed using the Fisher exact test, and *z* is the z-score computed by assessing the deviation from the expected rank. DPT, double positive T cell; TGF, transforming growth factor; TNF, tumor necrosis factor; Treg, regulatory T cell.

Discussion

Although limited by small sample size and heterogeneity among patients, this study uses a novel multimodal computational approach to demonstrate that pediatric AML has immune-mediated drivers of tumor escape related to MHC-II expression. These novel methods include mapping AML blasts in pediatrics to subtypes, evaluating trends in expression of pathways over the differentiation stages of AML blasts, and demonstrating phenotypic changes of AML clonotypes. Although not available for this pilot study, larger sample sizes as well as remission samples are necessary to further characterize the frequency of these transcriptional and phenotypic changes, and how they will influence treatment options. However, the emerging computational methods used in this study serve as a strong foundation to guide future efforts to define the functions and mechanisms of immunomodulatory AML, evaluate their relationship to the bone marrow microenvironment, and modify their activities for therapeutic benefit using future single-cell studies.

In our patient cohort, downregulation of MHC-II genes in AML blasts at posttransplantation relapse correlates with changes in upstream activating (MYC, RELA, TP53, CREBBP, EZH2, IRF1, and CIITA) and repressing (MYCN and HDAC1) transcription factors. A dampened $CD4^+$ T-cell response, suggested by the reduced conventional $CD4^+$ T-cell population and diminished IFN-γ, tumor necrosis factor α, and IL-2 response signature in NK and $CD8^+$ T cells, could be a possible mechanism for posttransplantation relapse driven by downregulation of AML MHC-II expression. We further propose that the progenitor-like AML blast subset may be driving this phenomenon based on the relative expansion of progenitor-like AML clones at posttransplantation relapse. However, it remains unclear whether AML blasts with intact MHC-II expression were surveyed and destroyed by the host's immune system before the time of diagnosis of posttransplantation relapse or whether the low-expressing MHC-II, progenitor-like AML blasts dominated immune escape after transplantation. Finally, consistent with published results that the absence of NKG2D ligands defines leukemia stem cells and mediates their immune evasion,[30] we observed a decrease in gene expression level of NKG2D ligands (MICA, MICB, ULBP1, ULBP2, and ULBP3) on AML blasts between diagnosis and posttransplantation relapse.

In summary, our findings suggest potential mechanisms for therapeutic epigenetic targeting to modify MHC-II gene expression and prevent immune escape, particularly among the progenitor-like AML blast stage, leading to improvements in prevention and therapy of pediatric AML.

Acknowledgments

The authors thank the patients and families whose samples were analyzed in this study. The authors thank Joseph Olechnowicz (Department of Pediatrics, Memorial Sloan Kettering Cancer Center [MSKCC]), for editorial assistance. The authors acknowledge Brian Shaffer (MSKCC), Colleen Lau (MSKCC), Eric Wang (MSKCC), and Audrey Lasry-Sandler (New York University Langone Medical Center) for helpful feedback and suggestions, and Hao Li (MSKCC) for assistance with data management.

The authors acknowledge the use of the Integrated Genomics Operation Core, funded by the National Cancer Institute Cancer Center Support grant P30 CA08748, Cycle for Survival, and the Marie-Josée and Henry R. Kravis Center for Molecular Oncology.

Authorship

Contribution: S.S. designed study, performed experiments, analyzed data, and wrote the manuscript; N.C. performed bioinformatics, analyzed data, and contributed to manuscript writing; J.-B.L.L., A.M., B.S., S.P.S., and B.G. contributed to data analysis; T.K. and V.B. contributed to bioinformatics and data visualization; M.K.P. contributed to study design and performed experiments; M.R. contributed clinical hematopathology data; O.A-W. contributed to study design and data analysis; K.C.H. contributed to study design, data analysis, and manuscript writing; and all authors reviewed the manuscript.

Conflict-of-interest disclosure: S.S. receives support from the MSK Kids Tow Center for Developmental Oncology. J.-B.L.L., M.K.P., and K.C.H. are listed as inventors on a provisional patent for the manufacture and use of $NKG2C^+CD8^+$ T cells (US provisional patent application number 63/188 435). M.K.P. is listed as an inventor on a patent for targeting CD20 in canine malignancies (patent number 11116795 from US Patent and Trademark Office). M.K.P and K.C.H. are listed as inventors on a provisional patent for the design and use of HLA-E:peptide chimeric molecules (US provisional patent application number 63/173 966). M.R. reports funding from Roche, NGM, Beat AML, and Auron Pharma and board membership in Auron Pharma. S.P.S is a shareholder in Imagia Canexia Health Inc and received consulting fees from AstraZeneca outside the scope of this work. O.A.-W. has served as a consultant for H3B Biomedicine, Foundation Medicine Inc, Merck, Prelude Therapeutics, and Janssen; is on the scientific advisory board of Envisagenics Inc, AIChemy, Harmonic Discovery Inc, and Pfizer Boulder; and has received prior research funding from H3B Biomedicine, Loxo Oncology, and Nurix Therapeutics. B.G. has received honoraria for speaking engagements from Merck, Bristol Myers Squibb, and Chugai Pharmaceuticals; has received research funding from Bristol Myers Squibb, Merck, and ROME Therapeutics; has been a compensated consultant for Darwin Health, Merck, PMV Pharma, Shennon Biotechnologies, and ROME Therapeutics, of which he is a cofounder; and is listed as an inventor on a patent application related to work on neoantigen quality modeling (63/303 500). K.C.H. serves on the scientific advisory board for WuGen. The remaining authors declare no competing financial interests.

ORCID profiles: S.S., 0000-0003-4281-4910; N.C., 0000-0001-9387-103X; J.-B.L.L., 0000-0002-1046-2941; A.M., 0000-0002-5654-5101; B.S., 0000-0003-1603-1381; T.K., 0000-0003-2054-4112; M.K.P., 0000-0002-2305-1164; S.P.S., 0000-0001-6402-523X; O.A.-W., 0000-0002-3907-6171; B.G., 0000-0001-6153-8793; K.C.H., 0000-0003-2827-5324.

Correspondence: Katharine C. Hsu, Memorial Sloan Kettering Cancer Center, 1275 York Ave, New York, NY 10065; email: hsuk@mskcc.org.

References

1. Reinhardt D, Antoniou E, Waack K. Pediatric acute myeloid leukemia-past, present, and future. *J Clin Med*. 2022;11(3):504.
2. Keating AK, Langenhorst J, Wagner JE, et al. The influence of stem cell source on transplant outcomes for pediatric patients with acute myeloid leukemia. *Blood Adv*. 2019;3(7):1118-1128.
3. Sweeney C, Vyas P. The graft-versus-leukemia effect in AML. *Front Oncol*. 2019;9:1217.
4. Toffalori C, Zito L, Gambacorta V, et al. Immune signature drives leukemia escape and relapse after hematopoietic cell transplantation. *Nat Med*. 2019; 25(4):603-611.
5. Christopher MJ, Petti AA, Rettig MP, et al. Immune escape of relapsed AML cells after allogeneic transplantation. *N Engl J Med*. 2018;379(24): 2330-2341.
6. Foley B, Cooley S, Verneris MR, et al. NK cell education after allogeneic transplantation: dissociation between recovery of cytokine-producing and cytotoxic functions. *Blood*. 2011;118(10):2784-2792.
7. Downing JR, Shannon KM. Acute leukemia: a pediatric perspective. *Cancer Cell*. 2002;2(6):437-445.
8. Simon AK, Hollander GA, McMichael A. Evolution of the immune system in humans from infancy to old age. *Proc Biol Sci*. 2015;282(1821): 20143085-20143085.
9. Bolouri H, Farrar JE, Triche T Jr, et al. The molecular landscape of pediatric acute myeloid leukemia reveals recurrent structural alterations and age-specific mutational interactions. *Nat Med*. 2018;24(1):103-112.
10. Kentsis A. Why do young people get cancer? *Pediatr Blood Cancer*. 2020;67(7):e28335-e28335.
11. Welch JS, Ley TJ, Link DC, et al. The origin and evolution of mutations in acute myeloid leukemia. *Cell*. 2012;150(2):264-278.
12. Stoeckius M, Hafemeister C, Stephenson W, et al. Simultaneous epitope and transcriptome measurement in single cells. *Nat Methods*. 2017;14(9): 865-868.
13. Medina-Martínez JS, Arango-Ossa JE, Levine MF, et al. Isabl Platform, a digital biobank for processing multimodal patient data. *BMC Bioinformatics*. 2020;21(1):549.
14. Wolf FA, Angerer P, Theis FJ. SCANPY: large-scale single-cell gene expression data analysis. *Genome Biol*. 2018;19(1):15.
15. Mulè MP, Martins AJ, Tsang JS. Normalizing and denoising protein expression data from droplet-based single cell profiling. *bioRxiv*. 2021.
16. Ceglia N, Sethna Z, Uhlitz F, et al. GeneVector: Identification of transcriptional programs using dense vector representations defined by mutual information. *bioRxiv*. 2022.
17. Lasry A, Nadorp B, Fornerod M, et al. An inflammatory state remodels the immune microenvironment and improves risk stratification in acute myeloid leukemia. *Nat Cancer*. 2023;4(1):27-42.
18. van Galen P, Hovestadt V, Wadsworth Ii MH, et al. Single-cell RNA-seq reveals AML hierarchies relevant to disease progression and immunity. *Cell*. 2019;176(6):1265-1281.e24.
19. van der Leun AM, Thommen DS, Schumacher TN. CD8(+) T cell states in human cancer: insights from single-cell analysis. *Nat Rev Cancer*. 2020; 20(4):218-232.
20. Sturm G, Szabo T, Fotakis G, et al. Scirpy: a Scanpy extension for analyzing single-cell T-cell receptor-sequencing data. *Bioinformatics*. 2020;36(18): 4817-4818.
21. Chen EY, Tan CM, Kou Y, et al. Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC Bioinformatics*. 2013;14(1):128.
22. Kuleshov MV, Jones MR, Rouillard AD, et al. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res*. 2016; 44(W1):W90-97.
23. Xie Z, Bailey A, Kuleshov MV, et al. Gene set knowledge discovery with enrichr. *Curr Protoc*. 2021;1(3):e90.
24. Wolock SL, Lopez R, Klein AM. Scrublet: computational identification of cell doublets in single-cell transcriptomic data. *Cell Syst*. 2019;8(4):281-291. e9.
25. Daily K, Patel VR, Rigor P, Xie X, Baldi P. MotifMap: integrative genome-wide maps of regulatory motif sites for model species. *BMC Bioinformatics*. 2011;12(1):495.
26. Han H, Cho JW, Lee S, et al. TRRUST v2: an expanded reference database of human and mouse transcriptional regulatory interactions. *Nucleic Acids Res*. 2018;46(D1):D380-D386.
27. Satija R, Farrell JA, Gennert D, Schier AF, Regev A. Spatial reconstruction of single-cell gene expression data. *Nat Biotechnol*. 2015;33(5):495-502.
28. Sethna Z, Elhanati Y, Callan CG Jr, Walczak AM, Mora T. OLGA: fast computation of generation probabilities of B- and T-cell receptor amino acid sequences and motifs. *Bioinformatics*. 2019;35(17):2974-2981.
29. Depreter B, De Moerloose B, Vandepoele K, et al. Deciphering molecular heterogeneity in pediatric AML using a cancer vs. normal transcriptomic approach. *Pediatr Res*. 2021;89(7):1695-1705.
30. Paczulla AM, Rothfelder K, Raffel S, et al. Absence of NKG2D ligands defines leukaemia stem cells and mediates their immune evasion. *Nature*. 2019; 572(7768):254-259.

31. Depreter B, Hofmans M, Terras E, et al. Deciphering molecular heterogeneity in pediatric aml using a cancer vs normal transcriptomic approach. *Blood*. 2019;134(suppl 1):1457-1457.

32. Fornerod M, Ma J, Noort S, et al. Integrative genomic analysis of pediatric myeloid-related acute leukemias identifies novel subtypes and prognostic indicators. *Blood Cancer Discov*. 2021;2(6):586-599.

33. Caushi JX, Zhang J, Ji Z, et al. Transcriptional programs of neoantigen-specific TIL in anti-PD-1-treated lung cancers. *Nature*. 2021;596(7870):126-132.