UBTF tandem duplications in pediatric myelodysplastic syndrome and acute myeloid leukemia: implications for clinical screening and diagnosis

Juan M. Barajas,[1]* Masayuki Umeda,[1]* Lisett Contreras,[1] Mahsa Khanlari,[1] Tamara Westover,[1] Michael P. Walsh,[1] Emily Xiong,[1] Chenchen Yang,[2] Brittney Otero,[2] Marc Arribas-Layton,[2] Sherif Abdelhamed,[1] Guangchun Song,[1] Xiaotu Ma,[3] Melvin E. Thomas 3rd,[1] Jing Ma[1] and Jeffery M. Klco[1]

[1]Department of Pathology, St. Jude Children's Research Hospital, Memphis, TN; [2]Mission Bio, South San Francisco, CA and [3]Department of Computational Biology, St. Jude Children's Research Hospital, Memphis, TN, USA

*JMB and MU contributed equally as first authors.

Abstract

Recent genomic studies in adult and pediatric acute myeloid leukemia (AML) demonstrated recurrent in-frame tandem duplications (TD) in exon 13 of upstream binding transcription factor (*UBTF*). These alterations, which account for approximately 4.3% of AML in childhood and about 3% in adult AML aged <60 years of age, are subtype-defining and associated with poor outcomes. Here, we provide a comprehensive investigation into the clinicopathological features of *UBTF*-TD myeloid neoplasms in childhood, including 89 unique pediatric AML and 6 myelodysplastic syndrome (MDS) cases harboring a tandem duplication in exon 13 of *UBTF*. We demonstrate that *UBTF*-TD myeloid tumors are associated with dysplastic features, low bone marrow blast infiltration, and low white blood cell count. Furthermore, using bulk and single-cell analyses, we confirm that *UBTF*-TD is an early and clonal event associated with a distinct transcriptional profile, whereas the acquisition of *FLT3* or *WT1* mutations is associated with more stem cell-like programs. Lastly, we report rare duplications within exon 9 of *UBTF* that phenocopy exon 13 duplications, expanding the spectrum of *UBTF* alterations in pediatric myeloid tumors. Collectively, we comprehensively characterize pediatric AML and MDS with *UBTF*-TD, and highlight key clinical and pathologic features that distinguish this new entity from other molecular subtypes of AML.

Introduction

Pediatric acute myeloid leukemia (AML) and myelodysplastic syndrome (MDS) are characterized by unique genetic backgrounds when compared to those in adults.[1-3] Recurrent tandem duplications (TD) of exon 13 of upstream binding transcription factor (*UBTF*) were only recently identified as potential initiating events in pediatric AML,[4–7] accounting for about 4% of newly diagnosed pediatric AML. PCR-based screening covering exon 13 of *UBTF* also identified *UBTF*-TD alterations in large adult AML cohorts.[8,9] These studies significantly contributed to the accumulation of evidence of *UBTF*-TD alterations in adult AML. However, PCR-based methods potentially underestimate partial tandem duplications (PTD) extending outside the regions covered by amplicons or possible alterations not involving exon 13.[9] Also, data on *UBTF* alterations in pediatric AML is limited to screening of relatively small cohorts,[4,5] and further efforts are needed to accumulate more knowledge about the bi-

ology and clinicopathologic features of this disease entity. *UBTF* encodes for the UBTF/UBF protein that regulates ribosomal RNA (rRNA) transcription and nucleolar formation.[10,11] We previously reported that expression of exon 13 UBTF-TD in cord blood CD34+ (cbCD34+) cells is sufficient to induce cellular proliferation, increase clonogenic activity, and it establishes a transcriptional signature that recapitulates what is observed in *UBTF*-TD AML patient samples.[4] Our previous analyses also demonstrated that *UBTF*-TD do not occur with other canonical alterations in pediatric AML, but that *UBTF*-TD AML often harbor additional somatic mutations, such as internal tandem duplications in *FLT3* (*FLT3*-ITD) and frameshift mutations in *WT1*. The acquisition of these co-operating mutations can likely contribute to the stepwise progression of the disease and clonal evolution. However, our understanding of how these co-operating mutations contribute to the cellular and disease status remains to be elucidated.

To bridge these knowledge gaps, we present an extended

pediatric and young adult cohort of 89 AML and 6 myelo-dysplastic syndrome (MDS) samples with exon 13 *UBTF*-TD, showing that *UBTF*-TD neoplasms are strongly associated with dysplastic features and unique patterns of co-op-erating mutations. By leveraging bulk RNA-sequencing (RNA-seq) and single cell proteogenomics, we show that the co-occurrence of *FLT3*-ITD and *WT1* mutations is as-sociated with stem cell-like phenotypes. Furthermore, we identified tandem duplications within exon 9 of UBTF in 2 cases transcriptionally resembling exon 13 *UBTF*-TD AML. Exon 9 *UBTF*-TD contain analogous hydrophobic leucine-rich sequences as the exon 13 duplications and similarly induce leukemic phenotypes in cbCD34+ cells, suggesting that they likely have a shared mechanism and should be classified as the same molecular entity. These findings offer valu-able insights to inform future diagnostic strategies and understanding of the molecular basis of *UBTF*-TD myeloid neoplasms.

Methods

Single-cell DNA sequencing and analysis

Single-cell targeted sequencing was performed using Tap-estri System from Mission Bio (missionbio.com). A panel consisting of 162 PCR amplicons targeting genes commonly mutated in pediatric AML with an average size of 260bp was designed using Tapestri Designer from Mission Bio (designer.missionbio.com), as well as a manually designed amplicon targeting the *UBTF* exon13 TD region (*Online Supplementary Table S1*). Cryopreserved *UBTF*-TD primary AML samples were thawed and subjected to dead cell removal using the EasySep™ Dead Cell Removal Kit (STEMCELL Technologies, cat# 17899). Live cells were then subjected to the Mission Bio DNA+Protein (TotalSeq™-D Heme Oncology Protein Panel) protocol per manufacturer's instructions (missionbio.com). Libraries were sequenced on the Novaseq platform (100M read pairs for DNA libraries and 225M read pairs for protein libraries). BAM files, loom files, h5 files, and QC metrics were produced via a customization of the Tapestri pipeline developed by Mission Bio (support.missionbio.com/hc/en-us). Analysis of the samples was completed using the Mosaic package v.3.0.1 (missionbio.github.io/mosaic/). Reads for the UBTF-TD and FLT3-ITD calls were isolated from bam files using the pysam python package v.0.21.0 (github.com/pysam-developers/pysam). Reads were then realigned to ITD contigs reported in previous studies[4] using the BWA aligner v.0.7.15-r1140.[12,13] When necessary, mutation variant allele frequencies (VAF) were adjusted using pysam.

Transcriptome analysis

Gene expression analysis was performed as previously described.[1] Briefly, an RNA-seq cohort was established by integrating *UBTF*-TD cases with RNA-seq data in this study (N=96) and AML in other categories (N=837) and cbCD34+ cells (N=5).[1] Reads from aligned RNA-seq BAM files were assigned to genes and counted using HTSeq (v.0.11.2)[14] with the GENCODE human release 19 gene annotation. The count data were transformed to log2-counts per million (log2CPM) using Voom available from R package Limma (v.3.50.3).[15] The top variable genes were selected using the "vst" method in Seurat package.[16] The expression data were then scaled, and PCA (Principal Component Analysis) was performed on the scaled data using the top 265 variable genes. Dimension reduction was performed using UMAP (Uniform Manifold Approximation and Projection)[17] with the top 100 principal components. Differential gene expression analysis was performed by Limma between groups as in-dicated in each figure, and we set Log2 CPM = -1 if it is < -1 based on the Log2 CPM data distribution. *P* values were adjusted by the Benjamini-Hochberg method to calculate the false discovery rate (FDR) using R function p.adjust. Genes with absolute fold change > 2 and FDR <0.05 were regarded as significantly differentially expressed. Gene Set Enrichment Analysis (GSEA) was performed by GSEA (v.4.2.3) using MSigDB gene sets c2.all (v.7.5.1).[18] Permutations were done 1,000 times among gene sets with sizes between 15 and 1,500 genes.

Institutional Review Board approval and ethics committee

Samples from patients with MDS or AML from St. Jude Children's Research Hospital tissue resource core facility were obtained with written informed consent using a pro-tocol approved by the St. Jude Children's Research Hospital Institutional Review Board (IRB). Studies were conducted in accordance with the International Ethical Guidelines for Biomedical Research Involving Human Subjects.

Statistical analysis

Details about statistical comparisons are provided in each figure legend. All the computations were performed using R or GraphPad Prism, and all *P* values are two-sided.

Results

UBTF-TD in pediatric myeloid neoplasms

Our previous study described the molecular features of 27 pediatric AML cases with exon 13 *UBFT*-TD.[4] To expand the cohort and better understand the biology, we screened RNA-seq data from pediatric and young adult MDS and AML samples and identified an additional 68 cases from available datasets and previously published studies,[2,19] and routine clinical service at St. Jude Children's Research Hospital. All 95 cases (median age = 14.0 years, range = 2.4-27.4) (*Online Supplementary Figure S1A*) possessed exon 13 *UBTF*-TD en-coding a consensus hydrophobic leucine-rich ELTRLLARM amino acid motif within the duplications (Figure 1A, *Online Supplementary Table S2*). The duplications resulted in an

increased size of exon 13 (median size = 60 bp, range = 45-339 bp) (Figure 1B). Consistent with previous findings, *UBTF*-TD did not co-occur with other subtype-defining alterations and showed high VAF (median = 36.3%, range = 13.5-78.1%), further supporting our previous assertion that *UBTF*-TD alterations are early clonal events[1,4] (Figure 1C). We further investigated the mutational background of this *UBTF*-TD cohort, confirming a strong association with a normal karyotype or trisomy 8 (*Online Supplementary Figure S1B, C),* as well as with *FLT3*-ITD (N=55, 57.9%) and mutations in *WT1* (N=39, 41.1%), which are highly co-occurring

(*Online Supplementary Figure S1D*) (P=0.011, Fisher's exact test) with 30.1% (N=29) cases harboring both alterations. In addition, 26.6% (N=25) of cases also had at least one mutation in Ras-MAPK pathway genes: *NRAS* (N=17, 17.9%), *PTPN11* (N=5, 5.3%), *RIT1* (N=5, 5.3%), *NF1* (N=4, 4.2%), *CBL* (N=2, 2.1%), and *KRAS* (N=2, 2.1%). Other recurrent mutations in myeloid malignancies were rarely observed, including *IDH1/ IDH2* (N=3, 3.2%), *BCOR* (N=2, 2.1%), and *RUNX1* (N=1, 1.1%). We also screened MDS cases and identified *UBTF*-TD in 6 cases of pediatric MDS, all with normal karyotype; these include 3 primary pediatric MDS cases from our previously published
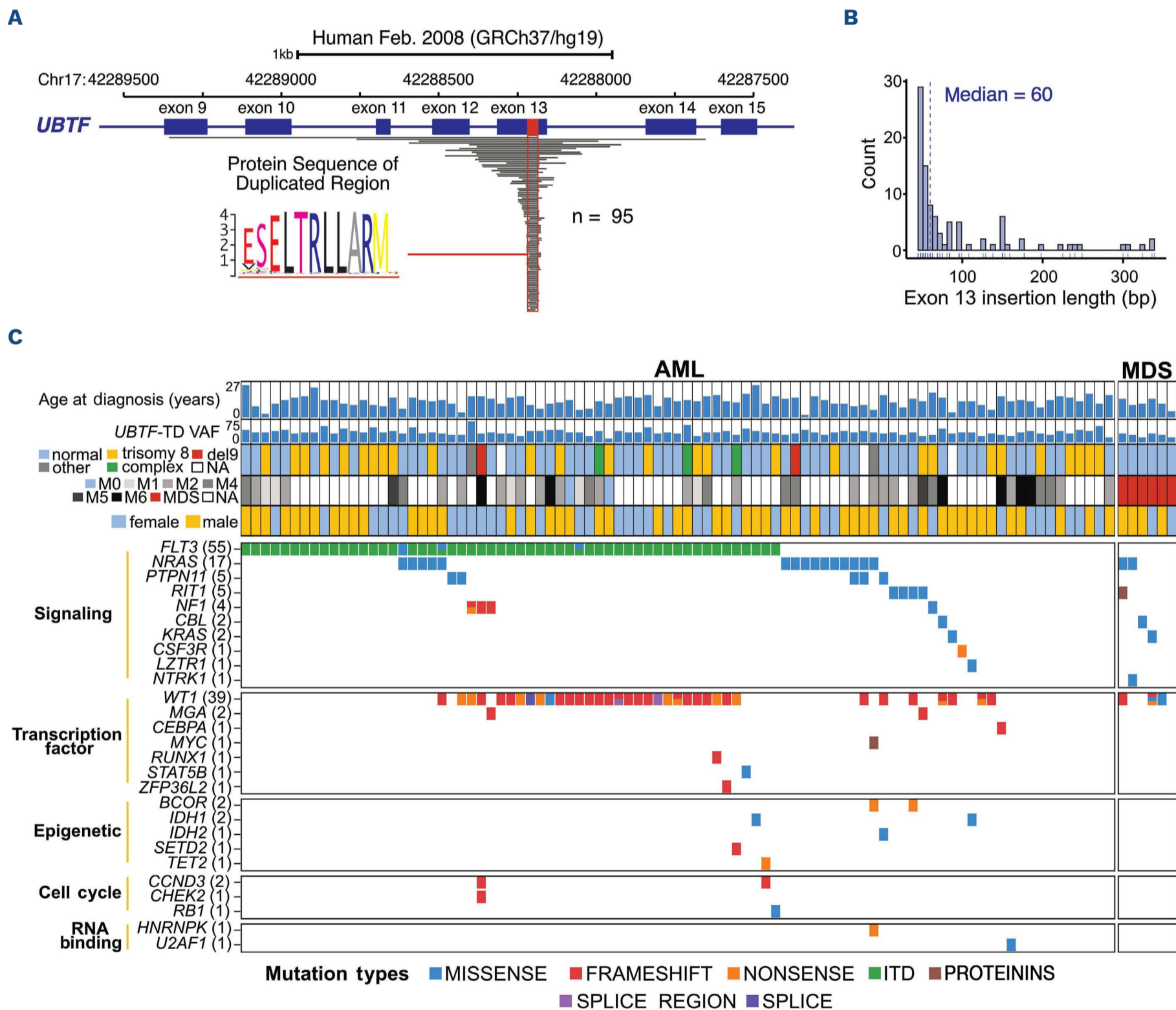

Figure 1. Characterization of tandem duplications in exon 13 of *UBTF*. (A) Genomic representation of duplicated regions (gray) from 95 cases with *UBTF* exon 13 duplications. A common duplicated region is outlined in red with the alignment of the resulting amino acid sequence. (B) Size assessment of exon 13 resulting from the duplicated regions and associated insertions / deletions. For partial tandem duplications (N=13) extending beyond exon 13, sizes of duplicated exon 13 with insertions or deletions are included. (C) Genomic landscape of *UBTF*-tandem duplication (TD) cases. Each column represents an individual patient. For each case, variant allele frequency (VAF), diagnosis, karyotype, and sex are shown.

cohort,[2] classified as childhood MDS with increased blasts according to the current WHO classification.[20] These 3 cases lacked a known germline predisposition (e.g., *GATA2, SAMD9, SAMD9L*) and had normal BM cytogenetics. Collectively, *UBTF*-TD was present in 3/38 (7.9%) of the primary MDS cases without a known germline predisposition, as well as 3/22 (13.6%) of all cases of childhood MDS with increased blasts. No *FLT3*-ITD mutations were detected in the 6 MDS cases, while *WT1* mutations were present in 3 out of the 6 MDS cases, suggesting clonal evolutionary patterns initiating with *UBTF*-TD alterations.

To address clonal evolution in *UBTF*-TD myeloid neoplasms, we utilized a droplet-based single cell proteogenomic platform from MissionBio (Figure 2).[21] This platform en-

ables the concurrent detection of *UBTF*-TD alterations and somatic mutations by a custom-targeted DNA panel and cell classification using DNA-oligo conjugated antibodies targeting cell surface markers at the single cell level. In a single timepoint AML case, we found a clonal *UBTF*-TD alteration. A small fraction of the cells only contained the *UBTF*-TD alteration, whereas the major population also contained a *WT1* (p.V359fs) mutation (Figure 2A, B, *Online Supplementary Table S3*). Distinct *UBTF*-TD+*WT1*+ minor subclones defined by *NRAS* (p.G12D) or *FLT3* (p.V592D) were also present, representing branched evolution of the *UBTF*-TD+*WT1*+ clone. We also found that cells with *WT1*+*FLT3*+ mutations were associated with high stem cell marker protein expression (CD34, CD117, or CD123) compared with
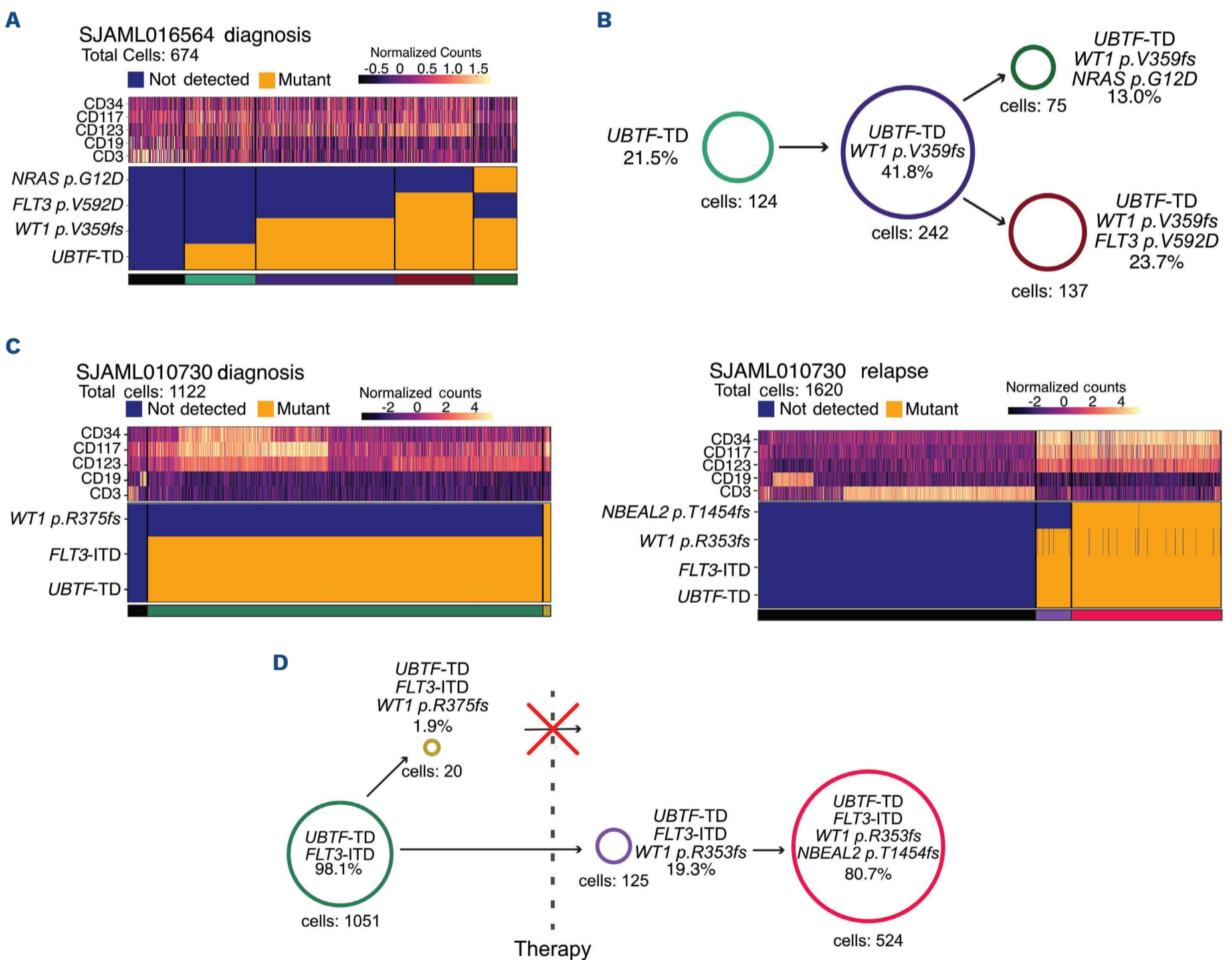

Figure 2. Clonal dynamics of *UBTF* tandem duplication leukemias. (A) Single-cell DNA sequencing coupled with surface marker expression of *UBTF* tandem duplication (TD) case at diagnosis. Heatmap depicting the presence of mutant allele (bottom) and relative protein abundance (top). (B) Schematic of clonal structure in *UBTF*-TD case from (A). Percentages are calculated as a proportion of total cells with a somatic mutation. (C) Single-cell DNA sequencing of a *UBTF*-TD case with diagnosis (left) and relapse (right) paired samples. Heatmap depicts the presence of a mutant allele and protein expression. (D) Schematic of clonal dynamics in diagnosis/relapse case from (C). Percentages are calculated as a proportion of total cells with a somatic mutation.
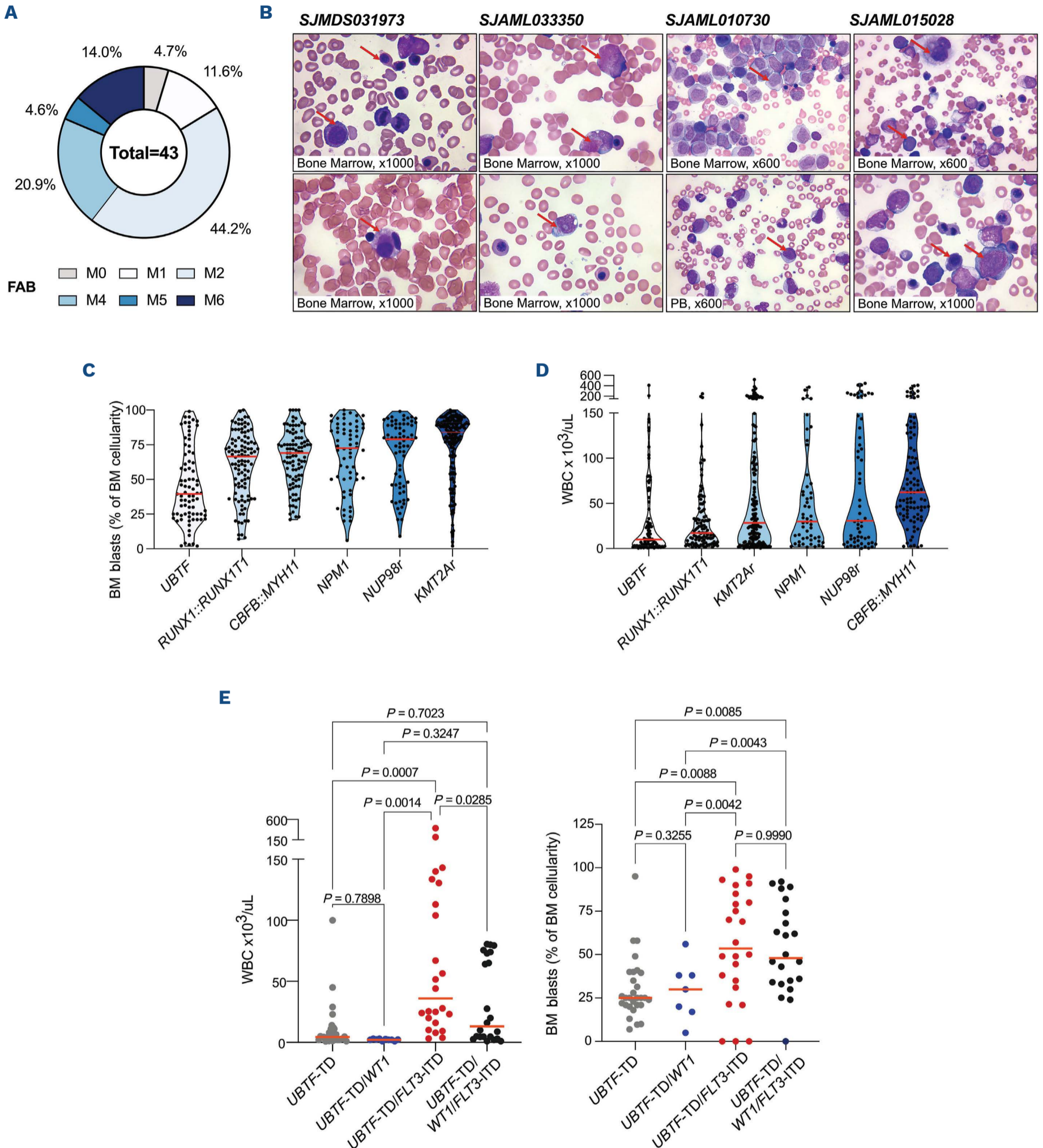
Figure 3. Morphological and biological assessment of *UBTF* tandem dupliacation myeloid neoplasms. (A) Available French-American-British (FAB) classification of *UBTF* tandem duplication (TD) acute myeloid leukemia (AML) (N=43). (B) Representative Wright-Giemsa staining of bone marrow (BM) aspirates and peripheral blood (PB) smears from 4 unique *UBTF*-TD cases. Case IDs are labeled above. SJMDS031973-childhood myelodysplastic syndromes (MDS) with increased blasts: increased erythroid cells with dysplastic morphologic features (top), dysplastic myeloid cells with salmon-colored granules, and dysplastic small megakaryocytes (bottom). SJAML033350-AML with other defined genetic alterations: characteristic blasts and myeloid precursor cells with salmon-colored granules (top) and blasts with small Auer rods (bottom). SJAML010730-AML with other defined genetic alterations: blasts with both myeloid and monocytic / monoblastic morphologic features. PB in this patient is one of the few cases in our cohort showing hyperleukocytosis with many circulating blasts (bottom). SJAML015028-AML with other defined genet-

ic alterations: dysplastic megakaryocytes, increased erythroid cells with dysplastic morphologic features, and blasts, many with monocytic / monoblastic morphologic features. (C) BM blast percentage for a pediatric AML cohort stratified by oncogenic driver subtypes. (D) White blood cell (WBC) counts for a pediatric AML cohort stratified by oncogenic driver subtypes. (E) WBC count and BM blast percentage among *UBTF*-TD cases with different *WT1* and *FLT3* mutation status.

the *UBTF*-TD-only population, whereas the *WT1⁺NRAS⁺* population was characterized by low expression of these markers. In a diagnosis and relapse-paired case, we found that the identical *UBTF*-TD was retained through disease progression along with an *FLT3*-ITD alteration (Figure 2C, D). Interestingly, a minor *WT1⁺* (p.R375fs) subclone at diagnosis was eradicated after chemotherapy, whereas a different *WT1⁺* (p.R353fs) clone became dominant at relapse, showing high expression of CD34, CD117, and CD123. These data collectively confirm that *UBTF*-TD is an early initiating event, while somatic mutations are subclonal to *UBTF*-TD, possibly contributing to disease progression toward subclones with unique expression profiles.

Clinical features of *UBTF*-TD pediatric myeloid neoplasms
UBTF-TD AML showed a variety of morphologic features associated with cellular differentiation, as evidenced by variable French-American-British (FAB) classifications (Figure 3A, *Online Supplementary Table S4*). Although AML with maturation (FAB M2) was the most common (19/43, 44.2%), cases with erythroid features, including FAB M6 (6/43, 14.0%), were also observed, as was also recently described for *UBTF*-TD AML in adults.[9] Morphologic assessments also revealed that *UBTF*-TD cases often displayed pleomorphic blasts (Figure 3B), accompanied by background multilineage dysplasia and increased erythroid precursors. *UBTF*-TD AML showed lower white blood cell count (median = 10.0×10^9/L, range = $0.6-409.4 \times 10^9$/L) and BM blast percentage (median = 39.5%, range = 2-97%) when compared to other AML, including those with similar transcriptional profiles like AML with *NUP98*-rearrangements or *NPM1* mutations[1,4] (Figure 3C, D). *FLT3*-ITD, but not *WT1* mutations, were associated with a higher white blood cell count (WBC) count and BM blasts in *UBTF*-TD AMLs (Figure 3E). Despite the presence of dysplastic features, cytogenetic studies commonly found either a normal karyotype (58/92, 63.0%) or trisomy 8 (27/92 29.3%). Myelodysplasia-related chromosomal changes or myelodysplasia-related mutations were overall rare, suggesting that *UBTF*-TD itself contributes to dysplastic features (*Online Supplementary Table S4*). Considering these overall features and other recent publications,[5,8,9,22,23] the majority of *UBTF*-TD AML (83/89, 93.3%) are best classified as "Acute myeloid leukemia with other defined genetic alterations" in the current WHO classification[20] (*Online Supplementary Table S4*).

Transcriptional features of *UBTF*-TD myeloid neoplasms
We and others have previously shown that AML with *UBTF*-TD is characterized by high *HOXA* and *HOXB* cluster gene expression, similar to *NPM1*-mutated or *NUP98::NSD1* AML.[1,8] To further define the unique expression profiles of *UBTF*-TD, we established an RNA-seq cohort consisting of various AML subtypes1 (N=837), cord blood CD34⁺ samples from healthy donors (N=5), and *UBTF*-TD AML and MDS samples (N=94: 1 *UBTF*-TD case did not have RNA-seq data available) (Figure 4A). Consistent with previous data, *UBTF*-TD cases clustered with *NPM1*-mutated and *NUP98::NSD1* AML. However, *UBTF*-TD samples displayed a significantly higher expression of a subset of HOXB cluster genes (e.g., HOXB8, HOXB9) compared with *NUP98::NSD1* AML (*Online Supplementary Figure S2A*). We also observed uniquely high expression of histone genes (e.g., *HIST1H4F* and *HIST1H1D*) compared to *NPM1*-mutated AML (*Online Supplementary Figure S2B*), suggesting transcriptional mechanisms unique to *UBTF*-TD AML. Within *UBTF*-TD samples, those with *FLT3*-ITD and *WT1* mutations showed unique distribution on the UMAP cluster (*Online Supplementary Figure S2C*), and each mutation group demonstrated differential gene expression against *UBTF*-TD samples without either mutation (Figure 4B). Co-occurrence of *WT1* and *FLT3*-ITD was associated with stemness-related genes (e.g., *CD34* and *DNTT*, *Online Supplementary Figure S2D*), and GSEA confirmed enrichment of stemness or cell cycle-related gene expression in *WT1⁺FLT3*-ITD⁺ *UBTF*-TD samples (Figure 4C). These results show the unique expression profile of *UBTF*-TD AML and the specific influence of additional co-operating mutations, which can likely impact patterns of clonal evolution.

Exon 9 tandem duplications in *UBTF*
Given the recurrent *UBTF* exon 13 alterations duplicating specific hydrophobic amino acid sequences, we hypothesized that *UBTF* alterations outside exon 13 resulting in similar amino acid sequences could be found in cases without defining alterations but with a similar expression signature. By close inspection of the *UBTF* gene using RNA-seq data, we found 2 pediatric AML cases without exon 13 *UBTF*-TD or other driver alterations that instead have in-frame tandem duplications (lengths of 78 and 153bp) in exon 9 of *UBTF* (TD-exon9), encoding short hydrophobic amino acid sequences (Figure 5A, B). These cases express *HOXA/B* cluster genes comparably to exon 13 *UBTF*-TD (*Online Supplementary Figure S3*), and one had a WT1 mutation (*Online Supplementary Table S5*). To test whether *UBTF*-TD-exon9 alterations could lead to cellular transformation, we expressed both UBTF-TD-exon9 in cbCD34⁺ cells using lentiviral vectors and assessed their impact on cell proliferation, clonogenic potential, and cellular morphology in comparison with control conditions and exon 13 UBTF-TD (Figure 5C, D). Colony-forming unit (CFU)

assay revealed that the expression of both UBTF-TD-exon9 increased the total colony number (Figure 5E). After the second round of replating, cells with UBTF-TD-exon9 showed an immature morphology along with erythroid features, similar to exon 13 UBTF-TD expressing cells in contrast to control conditions which displayed myeloid differentiation (Figure 5F). Furthermore, cells expressing UBTF-TD-exon9 as well as UBTF-TD-exon13 showed a proliferative advantage compared to UBTF-WT and vector controls (Figure 5G). Collectively, these data highlight a tandem duplication in *UBTF* exon 9 as a defining alteration functionally equivalent to exon 13 tandem duplications.

Discussion

In this study we extended our cohort of pediatric myeloid malignancies with exon 13 *UBTF*-TD, which now includes 95 pediatric and young adult cases. Similar to studies in adults,[9] pediatric myeloid tumors with *UBTF*-TD have a lower BM blast infiltration and lower peripheral WBC count, suggesting a continuum of a common entity across the age spectrum. We observed that *UBTF*-TD neoplasms showed variable cellular morphology and differentiation, including FAB M2 (44.2%), but also cases with erythroid features. These morphological features align with findings in adult cases, which showed a high prevalence of FAB M6 and M2 cases.[9] Furthermore, the presence of dysplastic features, and the observation that *UBTF*-TD occurs in cases of pediatric MDS, suggest that MDS and AML could be part of a continuum driven by *UBTF*-TD. These findings are consistent with the recognition of *UBTF*-TD alterations in nearly a third of pediatric patients with high-grade MDS[24] and recent studies have also identified the persistence of *UBTF*-TD in remission samples in patients with *UBTF*-TD myeloid neoplasms.[22] These collective data support the
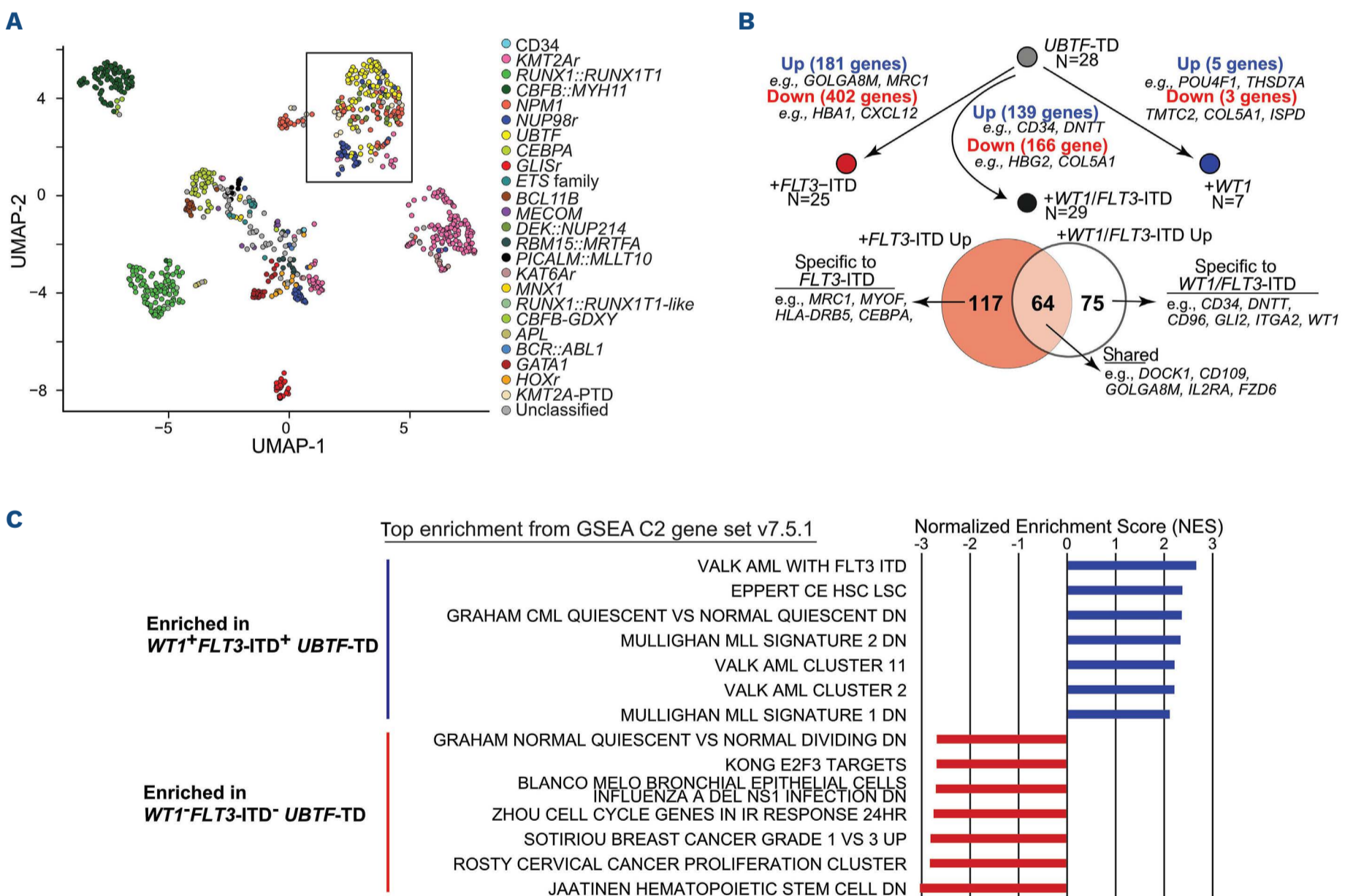

Figure 4. Transcriptional characterization of *UBTF* tandem duplication leukemias. (A) Uniform Manifold Approximation and Projection (UMAP) of expression profiles across a cohort of *UBTF* tandem duplication (TD) (N=94) and pediatric AML (N=837) acute myeloid leukemia (AML) cases, adapted from a previous study.[1,4] Each dot is colored by subtype-defining alterations. Black box outlines a cluster of cases with *HOXA/HOXB* dysregulation, which includes all *UBTF*-TD cases. (B) (Top) Schematic depicting the transcriptional comparison of *UBTF*-TD cases by mutational status. (Bottom) Venn diagram showing overlap of differentially expressed genes in *FLT3*-ITD only cases and *FLT3*-ITD+/*WT1*+ cases compared with cases without these mutations. (C) Gene Set Enrichment Analysis of *UBTF*-TD cases based on mutational status.

conclusion that UBTF-TD alterations can lead to both MDS and AML, and that UBTF-TD myeloid neoplasms should be recognized as a separate entity.

Our data also suggest that progression to AML may be promoted by the acquisition of co-operating mutations, such as *FLT3*-ITD. This is supported by the finding that
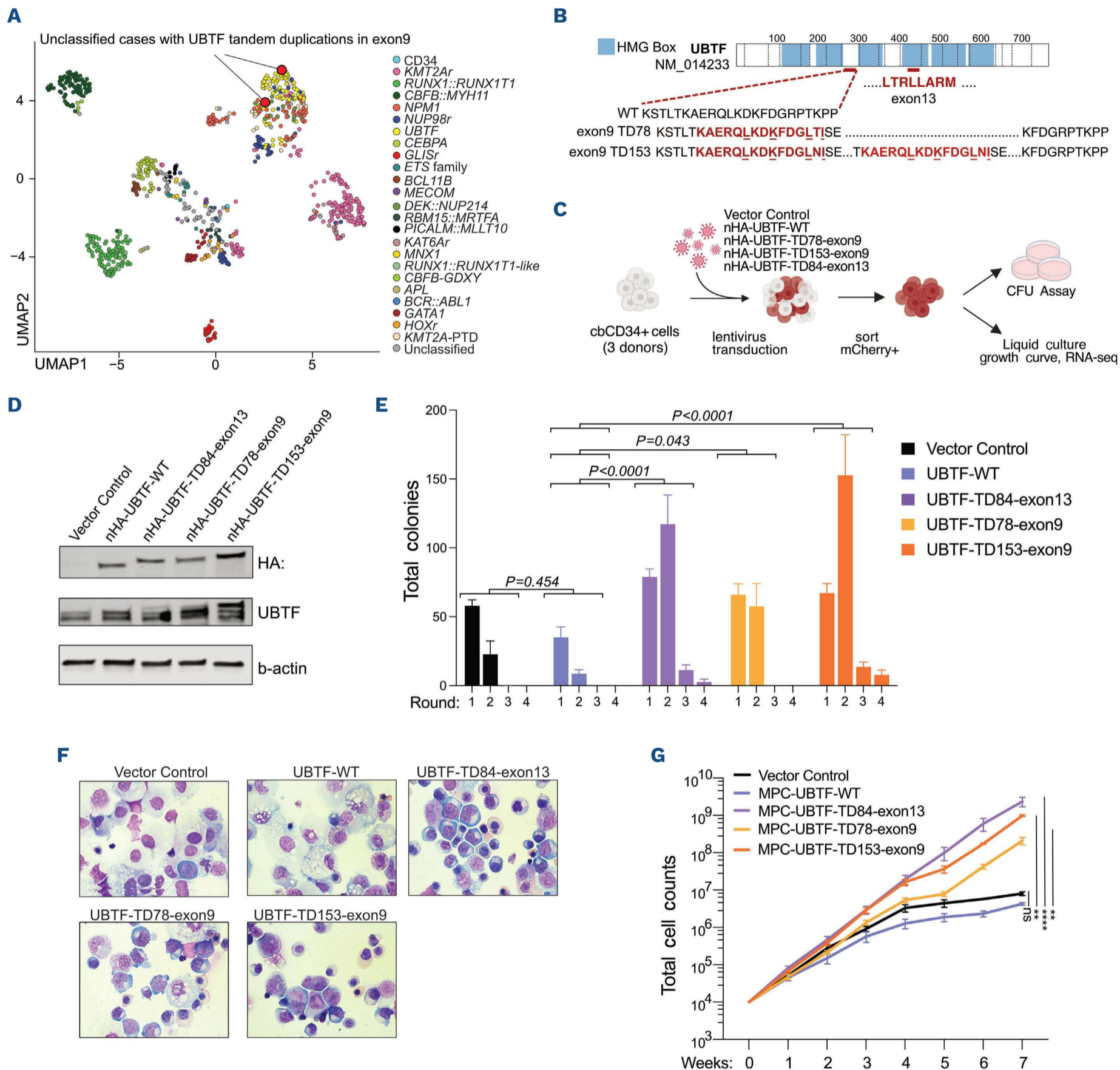

Figure 5. Exon 9 *UBTF* tandem duplications in pediatric acute myeloid leukemia. (A) Identification of *UBTF* tandem duplication (TD) cases with tandem duplications in exon 9 of *UBTF*. *UBTF*-TD-exon9 on the UMAP plot of the pediatric AML cohort (red circles), adapted from a previous study.[1,4] (B) Schematic of UBTF protein, highlighting amino acid sequences encoded in exon 9 of *UBTF*-WT, *UBTF*-TD78-exon9, *UBTF*-TD153-exon9. Duplications resulting in short hydrophobic sequences are labeled in red. Hydrophobic residues are underlined. (C) Experimental design to evaluate the transforming potential of UBTF-TD-exon9 mutants. (D) Immunoblot of lysates from cbCD34+ cells transduced with UBTF-TD-exon9 vectors and corresponding controls. Antibodies against HA-tag, UBTF, and β-actin were used. (E) Colony forming unit assay. Two-way ANOVA with two-sided Dunnett's test was applied with *UBTF* wild-type (WT) as control. (F) Cytospins isolated from colony forming unit (CFU) assay at round 2 of replating (Wright-Giemsa Staining, 400X magnification). (G) *In vitro* growth curves of cbCD34+ cells expressing controls or UBTF-TD-e9 mutants. Two-way ANOVA with Dunnett's test was applied with *UBTF*-WT as control.

none of the MDS cases in this cohort harbored a *FLT3*-ITD alteration, although this may be limited by the small size of the MDS cohort (N=6) in this study. However, the lack of *FLT3* or RAS pathway variants in a subset of AML cases suggest that UBTF-TD could be sufficient for leukemic transformation in some cases. Our analyses, including single cell studies, also showed that the co-occurrence of *FLT3*-ITD and *WT1* is strongly associated with progressive stem cell features as represented by CD34 expression or quiescent states. Thus, the patterns of mutational co-operativity likey will influence disease phenotypes.

Given that exon 13 *UBTF*-TD has been underappreciated in AML studies using standard computational pipelines, we investigated other possible *UBTF* alterations in pediatric AML cases without a defined driver event and demonstrated that in-frame tandem duplications in exon 9 of *UBTF* are additional possible driver alterations. Although rare within cases with *UBTF* alterations (2/97 UBTF-TD cases, 2.1%), cases with tandem duplications in exon 9 show similar transcriptional profiles to exon 13 *UBTF*-TD AML. We further show that exon 9 alterations can induce leukemic changes, including hematopoietic cell growth and increased clonogenicity in cbCD34+ cells similar to exon 13 alterations. At the amino acid level, exon 9 tandem duplications contain hydrophobic amino acid sequences (LKDKFDGL) that are similar to the sequences in exon 13 tandem duplications (LTRLLARM), suggesting a shared mechanism. Importantly, PCR-based exon 13 screening could potentially underestimate PTD involving exon 13 and flanking regions or these exon 9 alterations,[9] and we propose unbiased sequencing-based strategies to diagnose this entity accurately. The findings presented here will help build on our understanding of *UBTF*-TD myeloid neoplasms and further support its recognition as a distinct entity in future classification systems.

Disclosures

CY, BO and MA are employed by Mission Bio, Inc. All the other authors have no conflicts of interest to disclose.

Contributions

JMB is responsible for methodology, software validation, formal analysis, investigation, writing the original draft, vi-sualization, project administration and funding acquisition. MU is responsible for formal analysis, investigation, writing the original draft, visualization and project administration. LC is responsible for methodology, software validation and investigation. MK is responsible for formal analysis and investigation. TW is responsible for data curation and project administration. MPW, MET and XM are responsible for investigation. EX is responsible for methodology and software validation. BO and MA-L are responsible for formal analysis. GS is responsible for data curation. JMK is responsible for study concept, project administration and supervision, and funding acquisition. SA is responsible for project administration. CY, BO and MA-L are responsible for formal analysis. JM is responsible for formal analysis and investigation. All authors reviewed and edited the manuscript.

Data-sharing statement

The expression data newly generated in this study (RNA-sequencing: 3) and scDNA + protein sequencing (3) have been deposited in the European Genome-Phenome Archive (EGA) which is hosted by the European Bioinformatics Institute (EBI), under accession N. EGAS00001005760. The remaining RNA-sequencing data are available via EGA, St. Jude Cloud or TARGET/GDC as defined in Online Supplementary Table S6. Information about TARGET can be found at http://ocg.cancer.gov/programs/target. Other data generated in this study are available in the Online Supplementary Tables or upon request to the corresponding author.

References

1. Umeda M, Ma J, Westover T, et al. A new genomic framework to categorize pediatric acute myeloid leukemia. Nat Genet. 2024;56(2):281-293.
2. Schwartz JR, Ma J, Lamprecht T, et al. The genomic landscape of pediatric myelodysplastic syndromes. Nat Commun. 2017;8(1):1557.
3. 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.
4. Umeda M, Ma J, Huang BJ, et al. Integrated genomic analysis identifies UBTF tandem duplications as a recurrent lesion in pediatric acute myeloid leukemia. Blood Cancer Discov. 2022;3(3):194-207.
5. Kaburagi T, Shiba N, Yamato G, et al. UBTF-Internal tandem duplication as a novel poor prognostic factor in pediatric acute myeloid leukemia. Genes Chromosomes Cancer. 2022;62(4):202-209.
6. Stratmann S, Yones SA, Mayrhofer M, et al. Genomic characterization of relapsed acute myeloid leukemia reveals

novel putative therapeutic targets. Blood Adv. 2021;5(3):900-912.

7. Ma X, Liu Y, Liu Y, et al. Pan-cancer genome and transcriptome analyses of 1,699 paediatric leukaemias and solid tumours. Nature. 2018;555(7696):371-376.

8. Georgi JA, Stasik S, Eckardt JN, et al. UBTF tandem duplications are rare but recurrent alterations in adult AML and associated with younger age, myelodysplasia, and inferior outcome. Blood Cancer J. 2023;13(1):88.

9. Duployez N, Vasseur L, Kim R, et al. UBTF tandem duplications define a distinct subtype of adult de novo acute myeloid leukemia. Leukemia. 2023;37(6):1245-1253.

10. Sanij E, Hannan RD. The role of UBF in regulating the structure and dynamics of transcriptionally active rDNA chromatin. Epigenetics. 2009;4(6):374-382.

11. Moss T, Mars JC, Tremblay MG, et al. The chromatin landscape of the ribosomal RNA genes in mouse and human. Chromosome Res. 2019;27(1-2):31-40.

12. Li H, Durbin R. Fast and accurate long-read alignment with Burrows-Wheeler transform. Bioinformatics. 2010;26(5):589-595.

13. Li H, Handsaker B, Wysoker A, et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics. 2009;25(16):2078-2079.

14. Anders S, Pyl PT, Huber W. HTSeq--a Python framework to work with high-throughput sequencing data. Bioinformatics. 2015;31(2):166-169.

15. Ritchie ME, Phipson B, Wu D, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. 2015;43(7):e47.

16. Satija R, Farrell JA, Gennert D, et al. Spatial reconstruction of single-cell gene expression data. Nat Biotechnol. 2015;33(5):495-502.

17. Becht E, McInnes L, Healy J, et al. Dimensionality reduction for visualizing single-cell data using UMAP. Nat Biotechnol. 2019;37:38-44.

18. Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A. 2005;102(43):15545-15550.

19. 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.

20. Khoury JD, Solary E, Abla O, et al. The 5th edition of the World Health Organization Classification of Haematolymphoid Tumours: Myeloid and Histiocytic/Dendritic Neoplasms. Leukemia. 2022;36(7):1703-1719.

21. Miles LA, Bowman RL, Merlinsky TR, et al. Single-cell mutation analysis of clonal evolution in myeloid malignancies. Nature. 2020;587(7834):477-482.

22. Harrop S, Nguyen PC, Byrne D, et al. Persistence of UBTF tandem duplications in remission in acute myeloid leukaemia. EJHaem. 2023;4(4):1105-1109.

23. Barajas JM, Rasouli M, Umeda M, et al. Acute myeloid leukemias with UBTF tandem duplications are sensitive to Menin inhibitors. Blood. 2024;143(7):619-630.

24. Erlacher M, Stasik S, Yoshimi A, et al. UBTF tandem duplications account for a third of advanced pediatric MDS without genetic predisposition to myeloid neoplasia. Blood. 2022;140(Suppl 1):1355-1356.