



Article

<https://doi.org/10.1038/s43018-024-00863-5>

A multiomic atlas identifies a treatment-resistant, bone marrow progenitor-like cell population in T cell acute lymphoblastic leukemia

Received: 19 February 2024

Accepted: 17 October 2024

Published online: 25 November 2024

Check for updates

A list of authors and their affiliations appears at the end of the paper

Refractoriness to initial chemotherapy and relapse after remission are the main obstacles to curing T cell acute lymphoblastic leukemia (T-ALL). While tumor heterogeneity has been implicated in treatment failure, the cellular and genetic factors contributing to resistance and relapse remain unknown. Here we linked tumor subpopulations with clinical outcome, created an atlas of healthy pediatric hematopoiesis and applied single-cell multiomic analysis to a diverse cohort of 40 T-ALL cases. We identified a bone marrow progenitor (BMP)-like leukemia subpopulation associated with treatment failure and poor overall survival. The single-cell-derived molecular signature of BMP-like blasts predicted poor outcome across multiple subtypes of T-ALL and revealed that *NOTCH1* mutations additively drive T-ALL blasts away from the BMP-like state. Through *in silico* and *in vitro* drug screenings, we identified a therapeutic vulnerability of BMP-like blasts to apoptosis-inducing agents including venetoclax. Collectively, our study establishes multiomic signatures for rapid risk stratification and targeted treatment of high-risk T-ALL.

Acute lymphoblastic leukemia (ALL) is the most common pediatric cancer and leading cause of mortality¹. Outcomes in B cell ALL (B-ALL) have improved drastically because of optimization of chemotherapy², development of targeted therapies^{3–7} and genetically guided risk stratification⁸. In contrast, while outcomes have improved in T cell ALL (T-ALL), most persons who relapse are considered with low or favorable risk at diagnosis and few targeted therapies have successfully translated into the clinic⁹. Outcomes for persons with relapsed T-ALL are dismal as there are no effective salvage options. Accordingly, T-ALL is a disease where the goal is to use the most effective therapy at diagnosis. There is an urgent need to identify biologic risk factors to inform the development of targeted therapeutics and enable early identification of high-risk persons who need alternative treatment strategies¹⁰.

Systematic T-ALL classification used by the World Health Organization¹¹ and International Consensus Classification¹² relies on immunophenotype to characterize the maturation stage of T-ALL blasts and

has not reproducibly been associated with clinical importance. Recent T-ALL genomic initiatives have refined the transcriptomic classification of T-ALL, providing a unifying framework to link genetic alterations with outcome¹³. Among the emerging biomarkers for high-risk disease are the high prevalence of tumor subclones that could contribute to relapse^{13–15} and clonal selection^{16,17} that occurs with treatment failure. Here, we used single-cell multiomics to map the tumor landscape of >595,000 T-ALL blasts to the full hierarchy of pediatric hematopoiesis. We identify and characterize a chemotherapy-resistant and steroid-resistant bone marrow progenitor-like (BMP-like) tumor population shared between high-risk persons across the immunophenotypic spectrum of T-ALL. We used single-cell multiomics, large-cohort bulk genomics and primary patient-derived xenograft (PDX) models to establish multiomic signatures for rapid risk assessment and test the sensitivities of BMP-like blasts to currently available targeted therapeutics.

e-mail: tank1@chop.edu; teacheyd@chop.edu

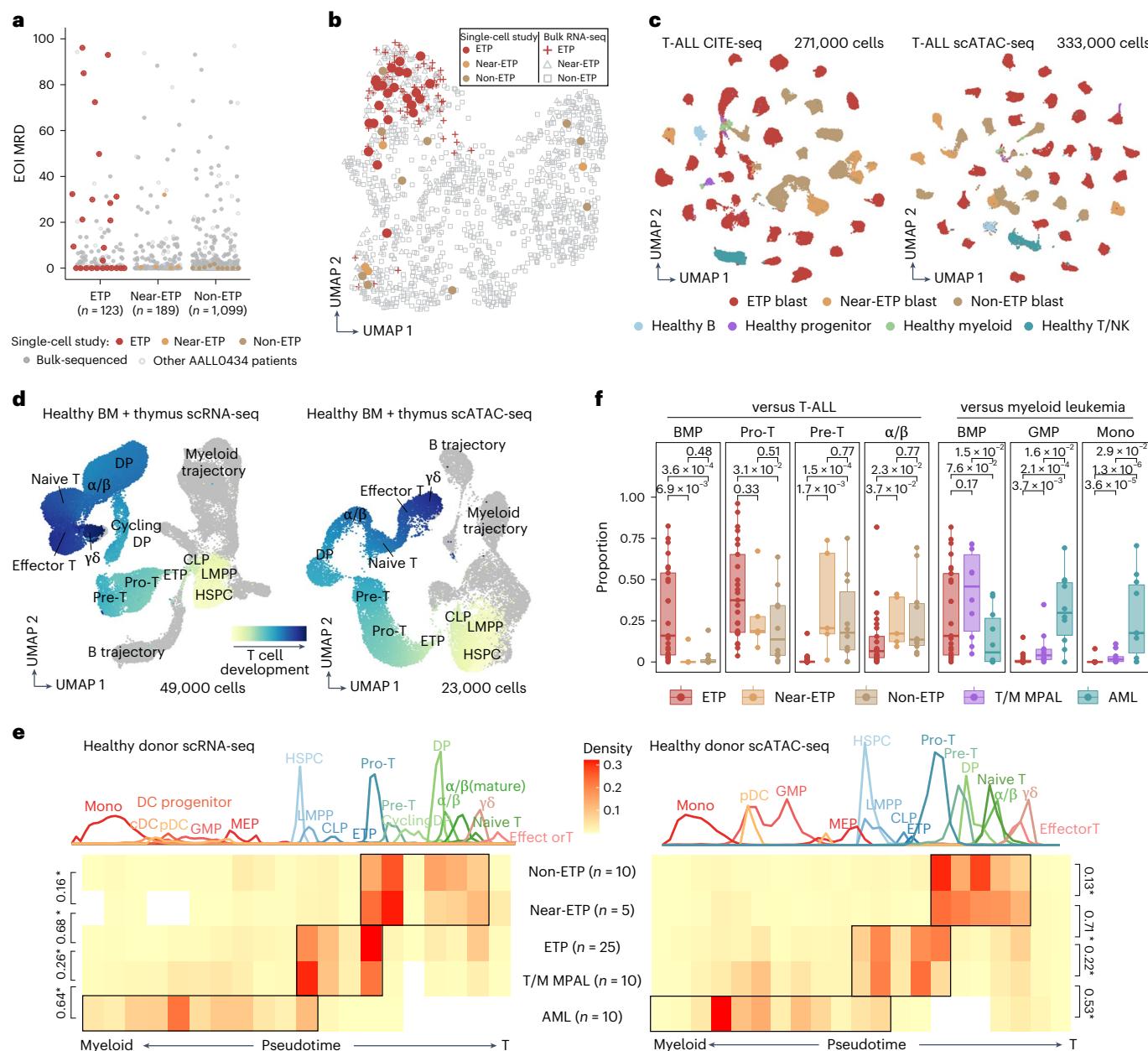


Fig. 1 | Arrest states of T-ALL subtypes in reference to human hematopoiesis. **a**, Selection of $n = 25$ participants with ETP-ALL, $n = 5$ participants with near-ETP-ALL and $n = 10$ participants with non-ETP-ALL from the COG AALL0434 cohort ($n = 1,411$) based on response to induction therapy (day 29 MRD). **b**, **c**, UMAP representation of bulk RNA-seq data from $n = 1,335$ diagnostic T-ALL samples from COG AALL0434. Each point represents the bulk RNA-seq transcriptome for one participant. Participants selected for single-cell study are indicated by circular points. All participants with ETP are colored red. **d**, UMAP representation of healthy human hematopoiesis development reference trajectories, based on scRNA-seq (left; $n = 49,623$ cells) and scATAC-seq (right; $n = 23,618$ cells) data. The key stages of T cell development implicated are labeled. α/β , alpha-beta; $\gamma\delta$, gamma-delta. **e**, Arrest states of leukemic cells from 40 participants with T-ALL based on projection to healthy

scRNA-seq (left) and scATAC-seq (right) reference. The D value from a two-sample Kolmogorov-Smirnov test is indicated to the side of brackets (* $P < 2.2 \times 10^{-16}$). Ten participants with T/M MPAL and ten participants with AML sequenced using identical assays are included as comparator groups ($n = 60$: ETP-ALL, 25; near-ETP-ALL, 5; non-ETP-ALL, 10; T/M MPAL, 10; AML, 10). **f**, Left: proportion of ETP blasts in four key T cell developmental stages, as compared to other participants with T-ALL. Right: proportion of ETP blasts in three key myeloid developmental stages, as compared to participants with T/M MPAL and AML. P values are based on a two-sided Mann-Whitney test. Results based on scRNA-seq projection are shown. The BMP stage encapsulates multipotent progenitors: HSPC, LMPP, CLP or ETP. The α/β stage encapsulates all cells that have moved past T cell commitment: DP, α/β , α/β (mature) or naive T. The box includes the median, hinges mark the 25th and 75th percentiles and whiskers extend 1.5 times the interquartile range ($n = 60$: ETP-ALL, 25; near-ETP-ALL, 5; non-ETP-ALL, 10; T/M MPAL, 10; AML, 10). Mono, monocyte; NK, natural killer.

Results

Multomic developmental atlases define T-ALL cellular arrest

We studied 25 participants with early T cell precursor (ETP)-ALL, 5 participants with near-ETP-ALL and 10 participants with non-ETP-ALL with

varied clinical response to treatment with AALL0434 (NCT00408005), an international phase 3 randomized Children's Oncology Group (COG) trial that reported the best published outcomes for children and young adults with T-ALL¹⁰ (Fig. 1a,b). We selected participants who quickly

went into minimal residual disease (MRD)-negative remission and were cured ($n = 16$), participants who had intrinsic chemotherapy resistance (refractory, enriched for cases with induction failure; $n = 10$) and participants who relapsed ($n = 14$) (Supplementary Tables 1 and 2). We had nine genetic subtypes represented, with the most common being the 'ETP-like' transcriptional subtype¹³ ($n = 24$), followed by T cell leukemia homeobox 3 ($n = 5$) and T-ALL protein 1αβ-like ($n = 3$). We performed cellular indexing of transcriptomes and epitopes by sequencing (CITE-seq) and single-cell assay for transposase-accessible chromatin sequencing (scATAC-seq) on live cells sorted from cryopreserved diagnostic bone marrow aspirate ($n = 32$) or peripheral blood mononuclear cells (PBMCs; $n = 8$), capturing over 600,000 high-quality cells across both modalities (Fig. 1c,d, Extended Data Fig. 1a–j and Supplementary Tables 3 and 4). To robustly phenotype our single-cell dataset within the context of normal hematopoietic development (Fig. 1e,f), we assembled a multiomic reference map of healthy pediatric hematopoiesis, using normal thymus and bone marrow tissues collected from children (Fig. 2a–h and Supplementary Tables 5 and 6).

We mapped T-ALL blasts to the hierarchy of human hematopoietic development, overcoming the limitations of using references restricted by bone marrow¹⁸, thymus¹⁹ or murine thymocyte signatures¹⁵. To assess the integrity of our reference mapping method, we additionally projected single-cell RNA sequencing (scRNA-seq) and scATAC-seq data from ten acute myeloid leukemia (AML) and ten T-myeloid mixed-phenotype acute leukemia (T/M MPAL) samples onto our reference, finding that AML blasts projected to the monocytic lineage and T/M MPAL blasts projected to both monocytic and T cell lineages, as expected (Fig. 1e). Notably, all subtypes of leukemia showed a spectrum of developmental arrest states with notable overlap at the subpopulation level. Arrest in a multipotent progenitor-like (hematopoietic stem and progenitor cell (HSPC) or lymphomyeloid primed progenitor (LMPP)) state represented a shared cell state in ETP-ALL, T/M MPAL and AML. Pro-T cell-like arrested blasts were shared between T/M MPAL and all three subtypes of T-ALL, while pre-T cell-like arrested blasts were shared between near-ETP-ALL and non-ETP-ALL.

A primary hypothesis for ETP-ALL treatment resistance has been that ETP-ALL retains myeloid populations that confer resistance to ALL-directed therapy^{20,21}. We found the median frequency of myeloid-projecting blasts (granulocyte macrophage progenitor (GMP), dendritic cell (DC) progenitor, plasmacytoid DC, conventional DC or monocyte) to be 0.167% in participants with ETP-ALL, in contrast to 16% and 82.5% in participants with T/M MPAL and AML, respectively (Fig. 1f). GMP-like and monocytic-like populations comprised <1% of blasts in 18/25 participants with ETP-ALL and were not detected in 5/25 participants, strongly supporting a lymphoid progenitor origin of ETP-ALL blasts and use of ALL-directed therapy.

We hypothesized that the developmental arrest spectrum of near-ETP-ALL would be similar to that of ETP-ALL. Therefore, we enriched our cohort with near-ETP cases from participants who relapsed (5/5 participants). The developmental arrest spectrum of near-ETP-ALL was closer to non-ETP-ALL, an unexpected finding given that near-ETP-ALL is defined by the ETP immunophenotype with the exception of high (>75%) CD5 expression^{11,12}. To assess whether the divergence of ETP and near-ETP developmental arrest spectra could explain the diverging clinical responses of near-ETP and ETP cases observed within AALL0434 (Extended Data Fig. 2a), we performed differential expression (DE) and differential chromatin accessibility (DA) analyses (Extended Data Fig. 2b,c).

Near-ETP blasts had downregulation of stem and myeloid markers (*SPINK2*, *C1QTNF4* and *HLA-DRA*) and upregulation of T cell receptor (TCR)-related machinery (*LAT*, *CD3E*, *CD28*, *LCK* and *PTCRA*) as compared to ETP-ALL, in line with commitment to the T cell lineage^{22,23}. Near-ETP blasts also had increased expression and motif accessibility of *TCF7* and *LEF1*, two transcription factors (TFs) central to T cell lineage commitment in healthy thymocytes²⁴. Within healthy thymocytes,

we observed the expression and accessibility of *TCF7* and *LEF1* to peak at the pre-T cell stage (Extended Data Fig. 2d). To ask whether *TCF7* and *LEF1* have an analogous T cell lineage-specific function in T-ALL blasts, we computationally inferred the regulators and targets of *TCF7* and *LEF1* in ETP-ALL, near-ETP-ALL and non-ETP-ALL by constructing subtype-specific transcriptional regulatory networks (Extended Data Fig. 2e–g). Our analysis predicted the repression of *TCF7* and *LEF1* expression by stem-related TFs^{25,26} (*MEF2C*, *IRF1* and *LYL1*) in ETP-ALL and activation of *TCF7* and *LEF1* by core TFs of T cell commitment²⁷ (*BCL11B*, *SIX6* and *TCF7L2*) in near-ETP-ALL and non-ETP-ALL. Our analysis also predicted the direct regulation of *CD5* expression and TCR-related genes (*LAT*, *DNTT*, *MAL* and *CD3E*) by *TCF7* and *LEF1* in near-ETP blasts, connecting *TCF7* and *LEF1* regulation to the CD5-bright phenotype observed within diagnostic flow cytometry. Elevated expression of our predicted *TCF7* and *LEF1* regulon was observed in bulk RNA-seq data of $n = 110$ participants with ETP and $n = 168$ participants with near-ETP (Extended Data Fig. 2h). Participants with ETP and near-ETP with higher expression of the *TCF7* and *LEF1* signature had more favorable outcomes within AALL0434 (Extended Data Fig. 2i–k), with the *TCF7* and *LEF1* signature having prognostic significance independent of MRD and central nervous system (CNS) status in participants with ETP-ALL (92.7% versus 79.3% 5-year overall survival (OS), $P = 0.024$; Extended Data Fig. 2j). Taken together, our results highlight functionally significant transcriptional regulatory circuits that underlie minute immunophenotypical differences.

BMP-like blasts are highly enriched in refractory ETP-ALL

High rates of treatment resistance, rather than eventual relapse, contribute to poor outcome in ETP-ALL²⁸. Within AALL0434, participants with ETP-ALL were 7.1-fold less likely to achieve remission (<5% bone marrow blasts by morphology) after the first month of chemotherapy (day 29 end of induction (EOI)) and >2-fold more likely to have flow-detectable MRD compared to participants without ETP.

We enriched our single-cell (SC) cohort for treatment-refractory ETP cases to identify tumor cell states associated with initial treatment resistance. We first compared the initial developmental arrest state of ten participants with high-EOI-MRD (>20%) ETP to ten participants with EOI-MRD-negative (0.1%) ETP (Fig. 3a,b) and asked whether response to chemotherapy was correlated with the fraction of actively cycling tumor cells. Although we observed a small enrichment of cycling cells in MRD-negative participants (20% versus 16%; Fig. 3c), we observed all treatment-refractory participants to have significant proportions of cycling cells, prompting us to investigate whether treatment response could be explained by differences in cell arrest state. Within scRNA-seq and scATAC-seq, we observed that participants with high-EOI-MRD ETP had an enrichment of blasts at the HSPC, LMPP, common lymphoid progenitor (CLP) or ETP developmental stages (Fig. 3d). Multi-lineage potency is retained^{29,30} in these states; we termed this cell state 'BMP-like'. In contrast, participants with MRD-negative ETP had an enrichment of blasts in the pro-T cell and pre-T cell stages. These states represent specification to the T cell lineage^{31,32}; we henceforth refer to them as 'T-specified'. We observed the proportion of tumor blasts in BMP-like and T-specified developmental stages to associate with day 29 MRD (Fig. 3d,e), event-free survival (EFS; Fig. 3f) and OS (Fig. 3g–h) and these populations were inversely correlated (Extended Data Fig. 3a–e) in single-cell-sequenced participants with ETP-ALL.

To test whether the molecular signatures of BMP-like and T-specified blasts could be used to stratify other participants with ETP within AALL0434, we performed DE and DA analyses to generate cell-type-specific molecular signatures (Supplementary Tables 7–9). We found that BMP-like blasts from non-responding participants with ETP had a high surface protein expression of myeloid and stem cell markers, including CD33, CD123, HLA-DR and CD34, and a low expression of T cell lineage surface protein markers, including CD3, CD4, CD2 and CD10 (Fig. 3i). The top DE genes (DEGs) for BMP-like blasts

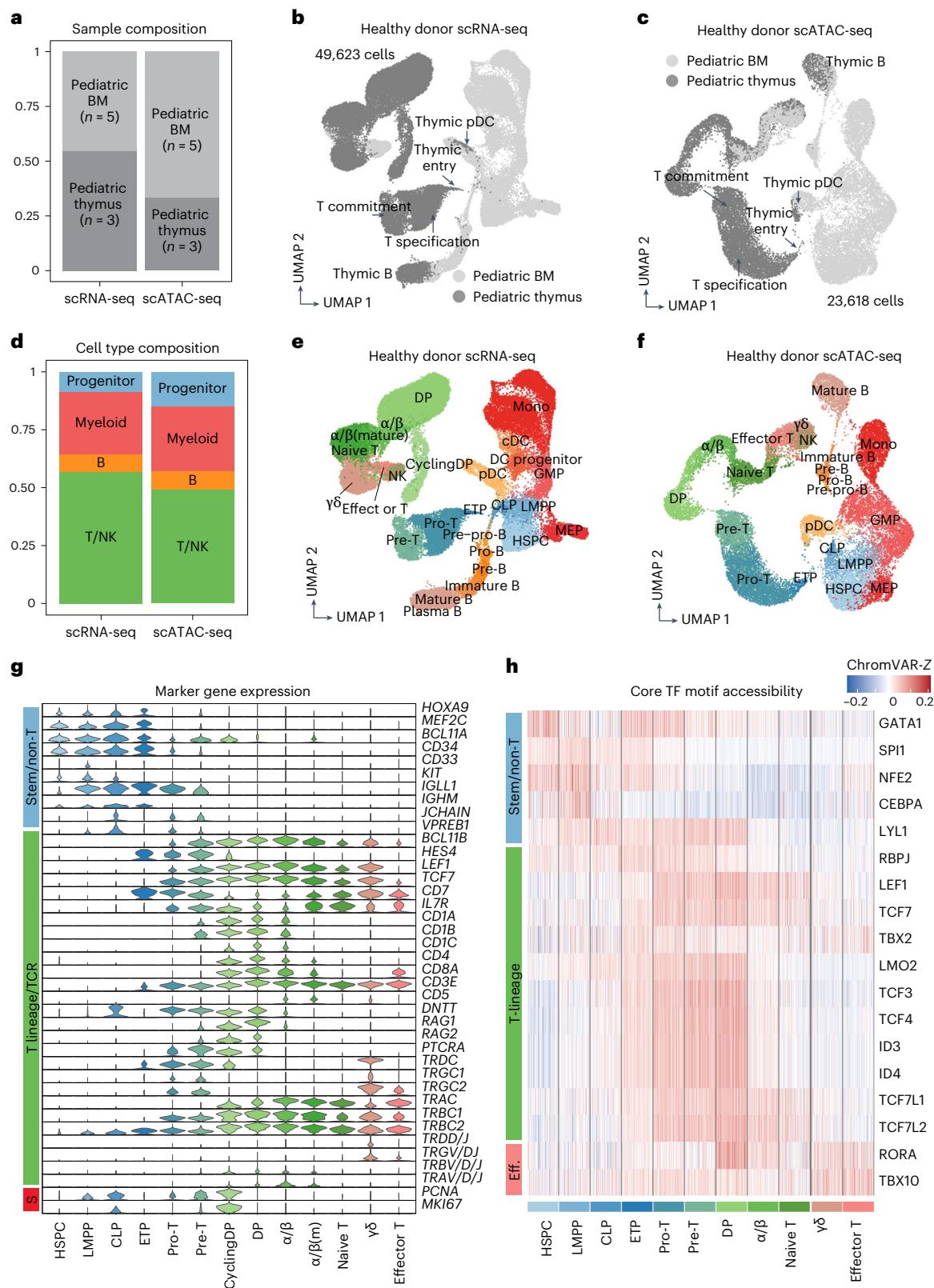


Fig. 2 | Multiomic developmental atlases to define cellular arrest state of leukemic blasts. **a**, Sample composition of scRNA-seq ($n = 49,623$ cells) and scATAC-seq ($n = 23,618$ cells) reference maps. **b,c**, UMAP representation of developmental reference trajectories, based on scRNA-seq (**b**; $n = 49,623$ cells) and scATAC-seq (**c**; $n = 23,618$ cells) data. The key stages of T cell development implicated in T-ALL and low-frequency thymic populations are labeled. **d**, Cell type composition of scRNA-seq ($n = 49,623$ cells) and scATAC-seq ($n = 23,618$ cells) reference maps. **e,f**, UMAP representation of scRNA-seq ($n = 49,623$ cells)

cells) and scATAC-seq ($n = 23,618$ cells) developmental reference trajectories. A total of 26 distinct cell populations defined by clustering and marker gene expression are labeled. **g**, Stem, T and S-phase marker gene expression (scRNA-seq, log-normalized counts) within developing T cell populations ($n = 49,623$ cells). **h**, Stem, T cell lineage and effector TF motif accessibility (scATAC-seq and chromVAR-Z) within developing T cell populations. Cells were randomly downsampled to $n = 100$ cells per group.

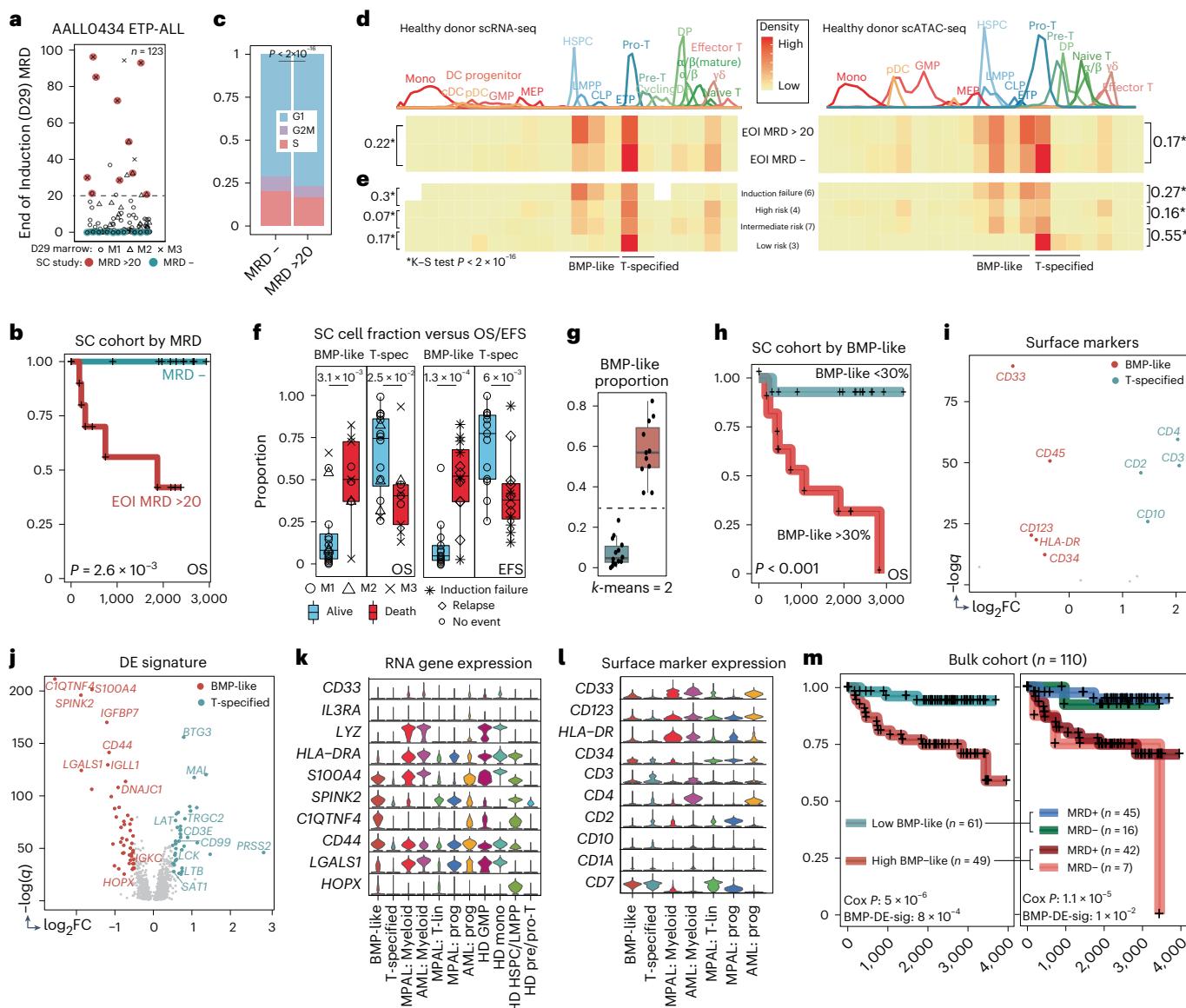


Fig. 3 | Treatment resistance in ETP-ALL is associated with a BMP-like population. **a**, Selection of ten high-MRD and ten MRD-negative (control) participants from $n = 123$ participants with ETP-ALL diagnosed within COG AALLO434. **b**, OS of $n = 10$ high-MRD versus $n = 10$ MRD-negative participants profiled using single-cell genomics. The P value for the log-likelihood statistic of a Cox proportional hazard test with day 29 MRD as a covariate is shown. **c**, Proportion of non-cycling (G1: MRD-negative, $n = 23,913$ cells; high-MRD, $n = 25,727$ cells), cycling (S: MRD-negative, $n = 2,862$ cells; high-MRD, $n = 2,274$ cells) and dividing (G2M: MRD-negative, $n = 6,125$ cells; high-MRD, $n = 5,499$ cells) cells in $n = 10$ high-MRD versus $n = 10$ MRD-negative participants. The chi-squared test statistic and P value were computed by comparing the proportion of cells in the G1 versus non-G1 phase in each group ($n = 3,500$ cells per participant, $33,500$ cells per group). **d,e**, Arrest states of leukemic cells from $n = 10$ high-MRD and $n = 10$ MRD-negative participants with ETP-ALL based on projection to the healthy scRNA-seq (left) and scATAC-seq (right) reference trajectory: proportion ranges from 0 to 0.3 (**d**) and from 0 to 0.5 (**e**). The D values from a two-sample Kolmogorov–Smirnov (K–S) test are indicated by the brackets ($*P < 2.2 \times 10^{-16}$; $n = 25$ participants with ETP-ALL: $n = 6$ induction failure, $n = 4$ high risk, $n = 7$ intermediate risk and $n = 3$ low risk). **f**, Proportion of ETP blasts in BMP-like and T-specified (T-spec) developmental stages in $n = 25$ single-cell-sequenced participants with ETP-ALL. The P values from two-sided t -tests are shown above the brackets. Alive indicates participants ($n = 16/25$) who were alive at last known follow-up (mean = 2,091 days). No event indicates participants ($n = 13/25$) who

had no event at last known follow-up (mean = 2,108 days). **g**, Proportions of BMP-like blasts in $n = 25$ single-cell-sequenced ETP-ALL blasts were stratified into high ($n = 11$ participants) and low ($n = 14$ participants) using k -means clustering. **h**, Stratification of $n = 25$ single-cell-sequenced participants with ETP-ALL by BMP-like proportion (high: >30%, $n = 11$; low: <30%, $n = 14$) determined through k -means clustering ($k = 2$). **i**, DE surface markers between BMP-like blasts from non-responding participants and T-specified blasts from responding participants. The input matrix to DE analysis was a matrix of G1-phase ETP-ALL blasts with an equal number of cells per participant ($n = 1,711$ cells per participant and $n = 42,775$ cells in total). **j**, DEGs between BMP-like blasts from non-responding participants and T-specified blasts from responding participants. The input matrix to DE analysis was a matrix of G1-phase ETP-ALL blasts with an equal number of cells per participant ($n = 1,711$ cells per participant and $n = 42,775$ cells in total). **k,l**, Normalized gene (**k**) and surface marker (**l**) expression for DE BMP-like genes across cell subpopulations in T-ALL, AML, MPAL and healthy donors (HD). Cells were randomly downsampled to $n = 200$ in each comparison group. **m**, Stratification of $n = 113$ participants with ETP from AALL0434 using 119 DEGs between BMP-like and T-specified blasts obtained in **d**. The prognostic value of the BMP-like signature (BMP-DE-sig) in multivariate analysis (with day 29 MRD, CNS status, WBC count and age at diagnosis) is shown below the Cox proportional hazard log-likelihood P value with day 29 MRD as a covariate. Left: stratification with signature alone. Right: stratification with signature and EOI MRD status.

included stem cell^{31,33,34} (*C1QTNF4*, *CD44*, *LGALS1* and *HOPX*), B cell lineage (*IGLL1*, *IGKC* and *IGHM*) and myeloid (*S100A4* and *SPINK2*) lineage markers (Fig. 3j). Although relatively enriched in BMP-like blasts, the myeloid lineage genes and surface markers were expressed at lower levels than myeloid blasts and healthy myeloid cells (Fig. 3k–l). The top DE TFs in BMP-like blasts included TFs associated with self-renewal (*MEF2C*, *HOXA3*–*HOXA6*, *HOXA9*–*HOXA11*, *MEIS1*, *HHEX* and *SPII*)^{26,35,36} and recovery from genotoxic stress³⁷ (*BCL11A*) (Extended Data Fig. 3f). Several of the BMP-like genes involved in self-renewal (*MEF2C*, *HOXA9* and *FLI1*) or T cell developmental block^{38,39} (*SPII*) also had increased motif accessibility (Extended Data Fig. 3g,h). In contrast, the top DEGs for T-specified blasts included T cell differentiation proteins (*MAL*), TCR components (*LAT*, *CD3E*, *LCK* and *TRGC2*) and thymic honing molecules (*CD99*). The top DE TFs in T-specified blasts included NOTCH targets⁴⁰ (*HES4* and *HES1*) and core T cell commitment TFs^{22,23} (*BCL11B*, *TCF7* and *RUNX1*). Interestingly, the top T-specified DEG by fold change was *PRSS2*, a trypsin gene lying adjacent to the TCRβ locus and becomes highly expressed as the TCRβ locus prepares for rearrangement⁴¹. We also assessed how the analogous populations differ between responding and non-responding participants through DEG and DE TF motif analyses (Extended Data Fig. 4a–d) and gene set enrichment analysis (GSEA) (Extended Data Fig. 5a–e), revealing that BMP-like non-responding blasts exhibited a similar differential phenotype compared to both T-specified responding and T-specified non-responding blasts.

Previous single-cell studies have also suggested that a portion of thymic seeding progenitors (TSPs) are more stem-like than others, retaining similar markers of multipotency seen in BMP-like ETP blasts^{19,29,42}. We transcriptomically matched BMP-like ETP-ALL blasts to the most stem-like subset of TSPs (Extended Data Fig. 6a–c), revealing two putative mechanisms of treatment resistance. We found corticosteroid receptor (*NR3C1*) expression to be directly correlated with T cell differentiation state and *NR3C1* expression highly predictive of EOI MRD (Extended Data Fig. 6d,e). BMP-like blasts had a significantly lower expression of *NR3C1*, rendering them highly resistant to prednisolone (>80-fold increase in half-maximal inhibitory concentration (IC₅₀)) in vitro (Extended Data Fig. 6f). We also predicted that BMP-like blasts would have high self-renewal capacity, much like TSPs⁴³. We observed the upregulation of leukemic stem-cell-related transcriptional programs in BMP-like blasts, associated with >100-fold resistance in vitro to T-ALL induction agents (Extended Data Fig. 6g,h).

Next, we sought to understand the transcriptional regulation of the BMP-like and T-specified states. We examined the TF motifs that were enriched in regions correlated with promoter accessibility of the BMP-like and T-specified DEGs (Supplementary Table 10). Motifs for TFs involved in the maintenance of progenitors (*SPII* and *GABPA*) were

enriched in the BMP-like state and TFs involved in T cell specification (*TCFL2* and *LEF1*) were enriched in the T-specified state (Extended Data Fig. 7a). We constructed transcriptional regulatory networks using SCENIC+ (ref. 44), which nominated *MEF2C* as the top activator of the BMP-like state. We found that *TCF7L2* and *BCL11B* were the most specific activators of the T-specified state but also served repressive roles in the BMP-like state, suggesting reciprocal regulation by these TFs (Extended Data Fig. 7b,c). Transcriptional activation of these regulons in the bulk RNA-seq cohort demonstrated that they are associated with OS across subtypes (Extended Data Fig. 7d–f).

Lastly, we applied the BMP-like DEG signature to stratify $n = 110$ bulk-sequenced participants with AALL0434 ETP-ALL. Our 119-gene BMP-like signature was predictive of OS (66.4% versus 94.3% 5-year OS, $P = 5 \times 10^{-6}$) and EFS (68.2% versus 94.3% 5-year EFS, $P = 2 \times 10^{-9}$) independent of EOI MRD and CNS status (Fig. 3m, left). Our BMP-like DEG signature provided robust stratification when MRD status was considered as a binary variable (Fig. 3m, right), identifying a subset of EOI-MRD-negative participants with inferior survival (high BMP-like and MRD-negative, 75.0% 5-year OS) and a group of EOI-MRD-positive participants with favorable survival outcomes (low BMP-like and MRD-positive, 94.9% 5-year OS).

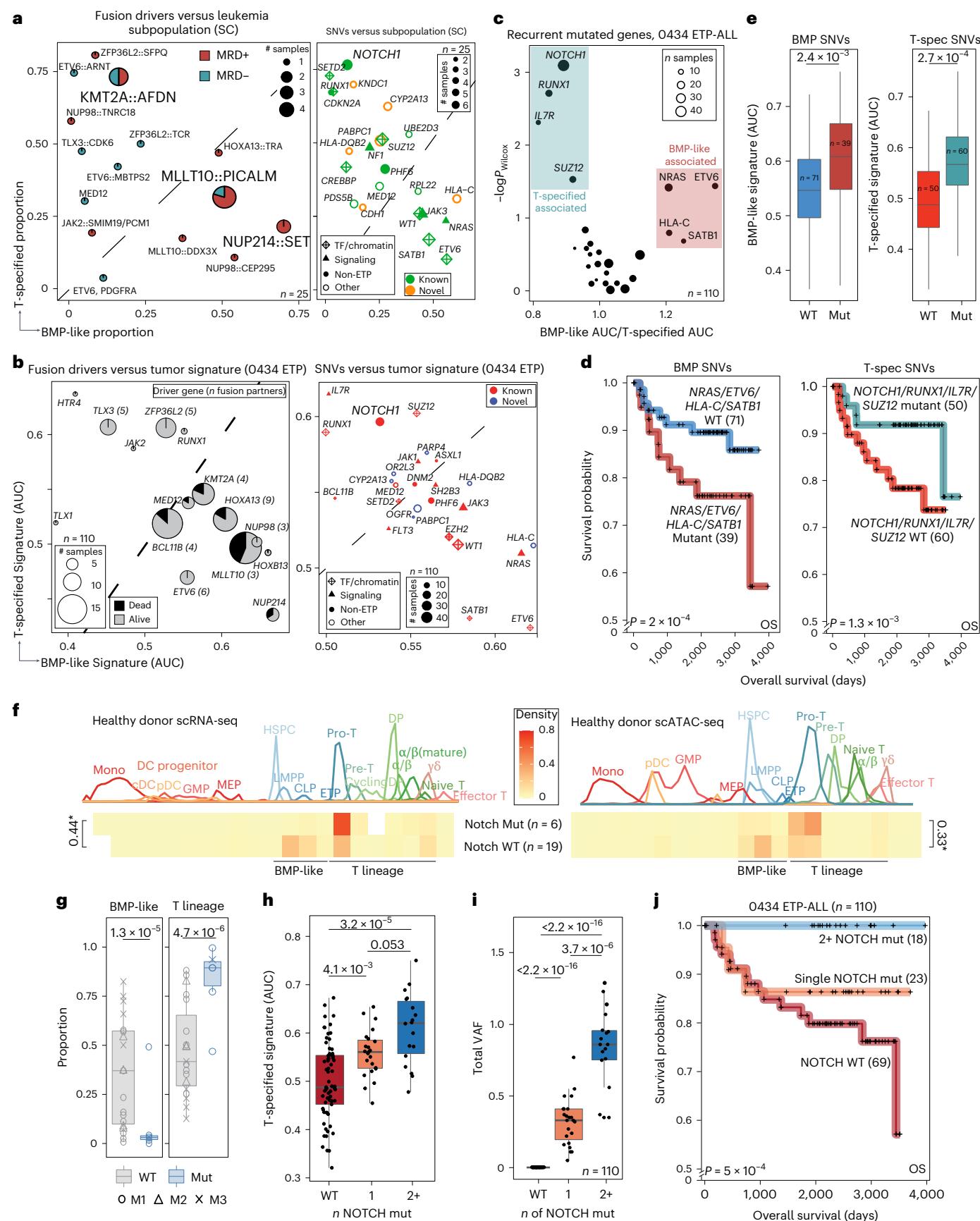
Distinct mutations underlie BMP-like and T-specified states

We next hypothesized that BMP-like and T-specified ETP blasts would harbor distinct mutational spectra. We leveraged the intersection of single-cell derived phenotypes with structural variant and single-nucleotide variant (SNV) calls (Supplementary Table 11) to identify associated drivers of these cell states (Fig. 4a,b and Extended Data Fig. 8a–c). BMP-like leukemias harbored fusion products known to drive high *HOXA* cluster expression, including *MLLT10*, *KMT2A*, *NUP214* and direct *HOXA*:*TCR* fusions. T-specified leukemias had *ZFP36L2* fusions (involved in cell-cycle control during T cell β selection), *TLX1* and *TLX3*. Notably, all participants with ETP who died from disease on AALL0434 had BMP-like associated fusions (Fig. 4b). On the SNV level, BMP-like high participants had recurrent mutations in TF and signaling pathways (that is, *JAK3*, *NRAS*, *WT1*, *ETV6* and *SATB1*) while T-specified high participants had mutations in T cell lineage regulators (that is, *IL7R*, *NOTCH1* and *RUNX1*). Top BMP-like associated SNVs were associated with inferior outcome, while top T-specified SNVs showed the opposite trend (Fig. 4c–e and Extended Data Fig. 8d,e).

In line with its essential role in T cell lineage commitment, the most recurrently altered gene associated with either cell state was *NOTCH1*. Within our single-cell cohort, we observed that *NOTCH1*-mutant participants ($n = 6$) had divergent arrest spectra compared to participants with wild-type (WT) *NOTCH1* ($n = 19$) (Fig. 4f), with a notable depletion of BMP-like blasts (Fig. 4g). We found that T-specified blasts had DE

Fig. 4 | Divergent mutational spectra associated with T-specified and BMP-like state. **a**, Recurrently mutated genes (left) and fusion drivers (right) among single-cell-sequenced participants with ETP-ALL versus tumor BMP-like and T-specified proportions among 25 participants with ETP-ALL. **b**, Recurrent driver fusions ($n = 16$, recurrent in >2 samples; left) and recurrently mutated genes ($n = 26$, recurrent in >5 samples; right) among 113 bulk-sequenced participants with ETP-ALL¹³ and associated BMP-like and T-specified signature scores. **c**, Top recurrently mutated genes associated with BMP-like and T-specified cell states in $n = 110$ ETP-ALL samples from AALL0434. Significance was assessed using a two-sided Wilcoxon rank-sum test. **d,e**, OS (d) and single-cell signature scores (e) among $n = 110$ bulk-sequenced participants with ETP-ALL grouped by mutation status within BMP-like (*ETV6*, *NRAS*, *HLA-C* and *SATB1*; mutant (mut), $n = 39$; WT, $n = 71$) and T-specified (*NOTCH1*, *IL7R*, *RUNX1* and *SUZ12*; mut, $n = 60$; WT, $n = 50$) associated genes. The P value for the log-likelihood statistic of a Cox proportional hazard test with day 29 MRD as a covariate is shown at the bottom left of the Kaplan–Meier curves. The box includes the median, hinges mark the 25th and 75th percentiles and whiskers extend 1.5 times the interquartile range. **f**, Arrest states of leukemic cells from *NOTCH1*-WT ($n = 19$) and *NOTCH1*-activated ($n = 6$) leukemias

based on scRNA-seq and scATAC-seq developmental trajectories. The D value from a two-sample Kolmogorov–Smirnov test is indicated to the side of brackets ($*P < 2.2 \times 10^{-16}$). **g**, Proportion of leukemic cells in BMP-like and T cell lineage (pro-T cell to αβ cell) states in *NOTCH1*-WT and *NOTCH1*-mut in single-cell-profiled participants ($n = 19$ *NOTCH1*-WT and $n = 6$ *NOTCH1*-mut). Significance was assessed using a two-sided Wilcoxon rank-sum test. **h**, T-specified signature score among 110 bulk-sequenced participants with ETP-ALL from AALL0434. Participants are divided into three groups by *NOTCH1* mutation status. P values from a two-sided Mann–Whitney test are shown. The box includes the median, hinges mark the 25th and 75th percentiles and whiskers extend 1.5× the interquartile range ($n = 69$ *NOTCH1*-WT, $n = 23$ *NOTCH1*-mut (single) and $n = 18$ *NOTCH1*-mut (two or more)). **i**, Summed VAF of *NOTCH1* mutations in participants with AALL0434 ETP with 0 (WT, $n = 69$), 1 ($n = 23$) and two or more *NOTCH1* ($n = 18$) mutations. P values from a two-sided Mann–Whitney test are indicated. The box includes the median, hinges mark the 25th and 75th percentiles and whiskers extend 1.5× the interquartile range. **j**, OS of $n = 110$ participants with AALL0434 ETP by *NOTCH1* mutation status. The P value for the log-likelihood statistic of a Cox proportional hazard test with day 29 MRD as a covariate is shown to the bottom left.



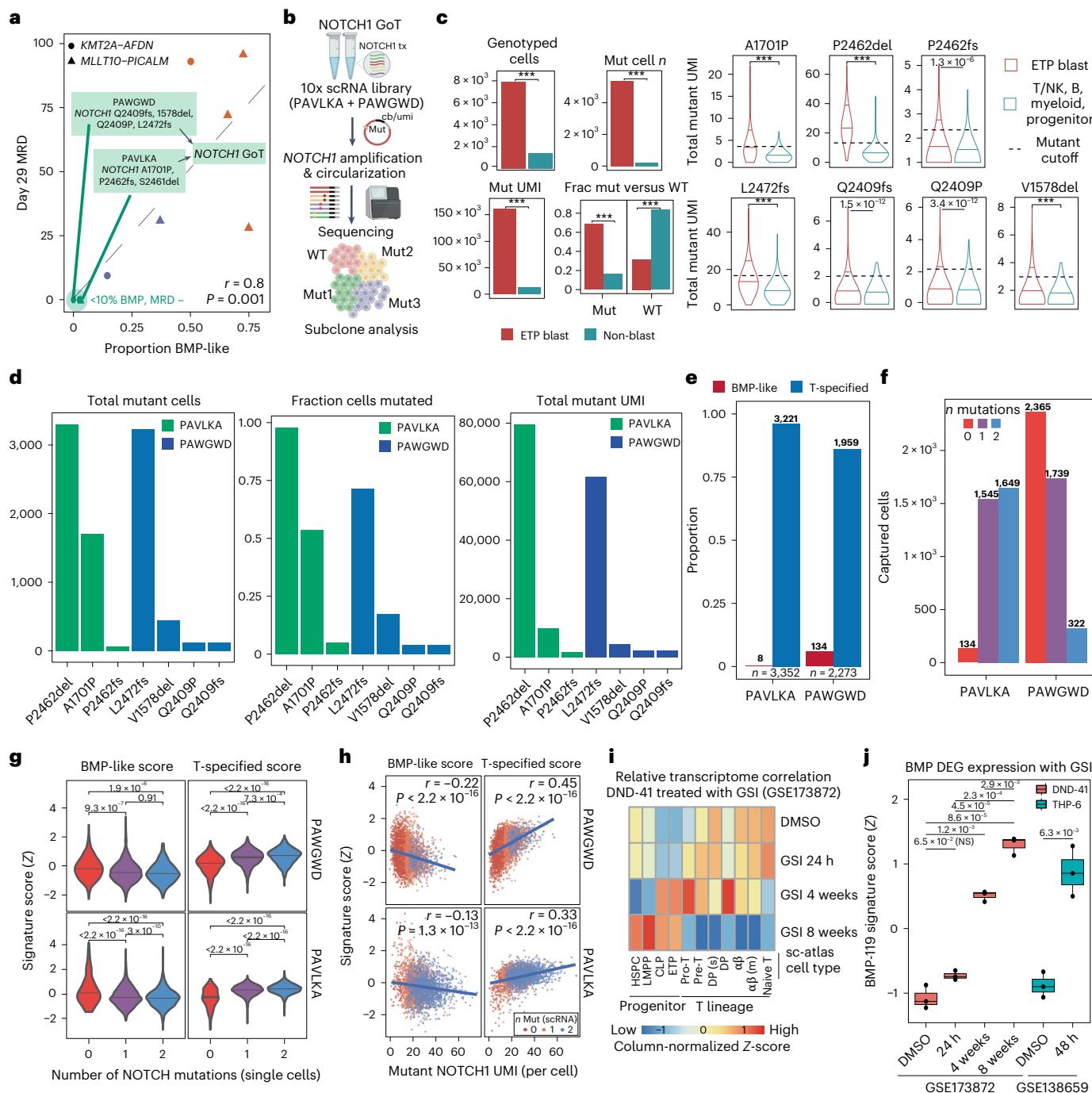


Fig. 5 | Subclonal NOTCH1 mutations additively contribute to differentiation toward the T-specified state. **a**, Identification of $n = 7$ subclonal NOTCH1 mutations in $n = 3$ MRD-negative participants with BMP-like associated fusion drivers. **b**, Left: experimental workflow for detection of NOTCH1-mut leukemic blasts in scRNA-seq libraries through GoT. cb, cell bar code; umi, unique molecular identifier. **c**, Summary statistics from $n = 9,314$ genotyped cells from seven independent GoT experiments. P values from a chi-square test are shown ($***P < 2.2 \times 10^{-16}$). Right: expression of NOTCH1-mut UMIs amongst ETP blast and non-ETP blast populations in scRNA-seq data. P values from a two-sided Mann-Whitney test are shown ($***P < 2.2 \times 10^{-16}$). The median of both populations and 90th percentile of UMI reads in non-blast populations are indicated within each violin plot. frac, fraction. **d**, GoT detection of 7/7 mut transcripts determined from bulk WES and WGS ($n = 7,754$ cells; $n = 5,625$ NOTCH1-mut: PAVLKA, $n = 3,352$; PAWGWD, $n = 2,273$). **e**, Fraction of BMP-like and T-specified cells within NOTCH1-mut cells ($n = 5,625$ cells: PAVLKA, $n = 3,352$; PAWGWD, $n = 2,273$). **f**, Detection of

$n = 1,971$ cells harboring two unique NOTCH1 mutations within leukemic blasts from PAVLKA ($n = 1,649$) and PAWGWD ($n = 322$). **g**, Association of BMP-like and T-specified signature scores with NOTCH1 mutation dosage in single cells. Cells are binarized into zero (WT; PAVLKA, $n = 134$; PAWGWD, $n = 2,365$), one (PAVLKA, $n = 1,545$; PAWGWD, $n = 1,739$) and two (PAVLKA, $n = 1,649$; PAWGWD, $n = 322$) mutations. P values from a two-sided Mann-Whitney test are shown. **h**, Association of NOTCH1-mut transcript expression and BMP-like and T-specified signatures ($n = 7,754$ genotyped single cells from PAWGWD and PAVLKA are plotted). The Pearson correlation coefficient and P value for Pearson's product moment are indicated to the top right. **i**, Transcriptome similarity of GSI-treated DND-41 T-ALL cells ($n = 3$ per condition) with T cell developmental stages identified in scRNA-seq data. **j**, Upregulation of BMP-like transcriptomic signature in GSI-treated DND-41 (GSE173872, $n = 3$ per condition) and THP-6 (near-ETP; GSE138659, $n = 3$ per condition) T-ALL cell lines. Cell lines were scored using the 119 BMP-like signatures established in Fig. 2. P values from a two-sided t-test are shown.

NOTCH activation pathways compared to BMP-like blasts (Extended Data Fig. 8f) and *NOTCH1* activation represented a universal marker of treatment sensitivity within the larger bulk-sequenced cohort. Of 110 bulk-sequenced participants with ETP-ALL, 41 harbored *NOTCH1* mutations, with 18 having two or more mutations (range: 2–5). Participants with two or more *NOTCH1* mutations had higher T-specified signature scores (Fig. 4h), aligning with an elevated *NOTCH1* variant allele frequency (VAF; >50% versus <50%) (Fig. 4i). Remarkably, all 18 participants with two or more *NOTCH1* mutations in the bulk-sequenced cohort were alive at last known follow-up, outperforming participants with single *NOTCH1* mutations and WT *NOTCH1* (Fig. 4j).

Interestingly, in EOI-MRD-negative participants with BMP-like associated drivers (*KMT2A* and *MLLT10* fusions), two of three harbored tumors with multiple activating *NOTCH1* mutations, suggesting that pan-tumor *NOTCH1* activation can drive differentiation away from the BMP-like state (Fig. 5a). To study *NOTCH1*-mutant subclones at the single-cell level, we performed genotyping of transcriptomes (GoT) on two participants harboring a total of seven unique activating *NOTCH1* mutations (Fig. 5b). We successfully detected 7/7 *NOTCH1* mutations in scRNA-seq libraries, corroborating bulk-derived variant calls (Fig. 5c,d). We found that *NOTCH1*-mutant cells were predominantly in the T-specified state and identified hundreds of leukemic blasts carrying two distinct *NOTCH1* mutations, likely resulting from selection for *NOTCH1* mutation in separate alleles (Fig. 5e,f). We found a direct connection between *NOTCH1* mutation dosage and the T-specified cell state, whereby blasts with two unique mutations had the highest expression of T-specified genes and lowest expression of BMP-like genes (Fig. 5g,h). Lastly, we analyzed the RNA-seq data of DND-41 and THP-6 cell lines treated with NOTCH pathway γ-secretase inhibitors^{45,46} (GSIs), finding that NOTCH pathway inhibition induced exposure-dependent transcriptomic shifts toward the BMP-like cell state (Fig. 5i,j). Taken together, these data offer high-resolution insight into how *NOTCH1* mutations alter T-ALL developmental arrest and cement *NOTCH1* mutation status as a critical biomarker for response to conventional therapy.

BMP-like genetic signatures risk-stratify non-ETP-ALL

Relapsed and refractory T-ALL is nearly universally fatal. Given that BMP-like blasts are highly resistant to conventional T-ALL therapy, we wondered whether an analogous, less differentiated subpopulation could be responsible for treatment resistance and relapse in non-ETP-ALL. Analysis of bulk RNA-seq data from the AALL0434 non-ETP cohort supported this hypothesis, with DE between 355 MRD-positive and 714 MRD-negative non-ETP-ALL cases revealing gross differences in differentiation state (Supplementary Table 12). Within this analysis, MRD-negative participants overexpressed markers of the α/β stage (*CD1B*, *CD1E*, *MAL*, *CD8A*, *PTCRA*, *RAG1* and *RAG2*), while MRD-positive participants expressed immature forms of the TCR

(*TRGC1* and *TRGC2*) and stem-related TFs (*HHEX* and *LYL1*). We used single-cell multiomics data to determine whether these differences were mediated by cell state or cell proportion differences.

We stratified our ten participants without ETP into two groups: six EOI-MRD-negative participants with complete response (CR) and four EOI-MRD-positive participants (EOI MRD > 0.1%). We observed that non-CR participants had an enrichment of cell states before T cell commitment (pro-T cell, CLP, LMPP, megakaryocyte-erythroid progenitor (MEP) and HSPC), while CR participants had an enrichment of cells in postcommitment states (CD4⁺CD8⁺ double-positive (DP) or α/β) (Fig. 6a,b). Because precommitted blasts represent a continuum of cell states from BMP-like to pro-T cell-like, we compared the distribution of precommitment blasts in CR participants and non-CR participants. We found that MRD-positive participants harbored a strong enrichment of BMP-like blasts, which were nearly absent in CR participants (Fig. 6c). DE analysis generated signatures for non-ETP precommitted and BMP-like blasts (Supplementary Tables 13 and 14), which both stratified participants without AALL0434 ETP independent of MRD, with the BMP-like signature having slightly better stratification (Fig. 6d).

We next intersected the molecular signature of BMP-like blasts obtained from participants with ETP-ALL and non-ETP-ALL, revealing a shared BMP-like gene set ('BMP-17') composed of 17 marker genes typically expressed in stem, myeloid and B cell progenitors (Fig. 6e,f and Supplementary Table 15). We applied BMP-17 to five different clinical scenarios for risk stratification, observing robust risk stratification in all instances (Fig. 6g,h). BMP-17 powered risk stratification within the smaller, partially sequenced AALL1231 cohort and was prognostic independent of EOI MRD and CNS status within the fully sequenced AALL0434 cohort, including participants stratified by ETP status. To further test whether BMP-like subpopulations were enriched at relapse time points, we performed RNA-seq on 27 relapsed T-ALL cases within AALL0434. We identified a consistent enrichment of BMP-like gene signatures and decrease in T-specified signatures within relapse cases, further supporting the use of BMP-like signatures for risk stratification (Fig. 6i,j).

We next determined the surface immunophenotype of BMP-like blasts using CITE-seq data, revealing a nine-marker phenotype ('BMP-surface-9') that reflected similar lineage aberrancy to BMP-17 (Fig. 7a–c and Supplementary Table 16). We found that RNA expression of BMP-surface-9 genes robustly stratified AALL0434 and AALL1231 T-ALL cohorts (Fig. 7d–h) and further validated BMP-surface-9 in AALL0434 diagnostic flow cytometry cases ($n = 99$), finding a robust correlation of clinically used surface markers with the BMP-17 gene signature (Fig. 7i,j).

Because non-pediatric T-ALL is enriched for both treatment-refractory cases and ETP phenotype, we next applied BMP-17 to young adult (age ≥ 18) cases on AALL0434. This analysis isolated a subset of high-BMP cases (Extended Data Fig. 9a–c) with high rates of EOI

Fig. 6 | A consensus 17-gene BMP-like signature predicts OS across all subtypes of T-ALL. a, Arrest states of leukemic cells from CR ($n = 6$) and MRD-positive ($n = 4$) participants with non-ETP-ALL based on projection to healthy scRNA-seq (left) and scATAC-seq (right) reference trajectory. **b**, Proportion of non-ETP blasts in precommitment (pre-commit, all cells before the pre-T cell stage) and postcommitment (post-commit, after the pre-T cell stage) developmental arrest state. P values from a two-sided Mann-Whitney test are indicated ($n = 68,801$ cells; total cells: CR, 52,971; MRD-positive, 15,830; precommitted blasts: CR, 7,152; MRD-positive, 11,047). **c**, Arrest states of precommitted non-ETP blasts in CR ($n = 6$) and MRD-positive ($n = 4$) participants. BMP-like encapsulates all cells that possess multipotent potential (HSPC, LMPP, CLP or ETP). **d**, Kaplan-Meier plot showing OS of participants with non-ETP-ALL in AALL0434 when binarized using signatures (sig) derived from precommitted non-ETP blasts and BMP-like non-ETP blasts. The P value for the log-likelihood statistic of a Cox proportional hazard test with day 29 MRD as a covariate is shown to the bottom left. **e**, Overlap of ETP BMP-like and non-ETP BMP-like DEGs to

create consensus signature for risk stratification in AALL0434 (fully sequenced) and AALL1231 (partially sequenced). BMP-like DEGs were filtered for mean log₂ fold change (FC) > 0.9 between ETP and non-ETP comparisons. **f**, Expression score of BMP-17 signature score and BMP-17 marker genes within bone marrow and thymus scRNA-seq reference ($n = 49,623$ cells). Multipotent BMP populations with high BMP-17 expression are circled. **g,h**, Kaplan-Meier plot showing OS of bulk RNA-seq participants with T-ALL in AALL0434 ($n = 1,335$ participants) and AALL1231 ($n = 75$ participants) binarized using the BMP-17 signature. The prognostic value of the BMP-17 signature in multivariate analysis (with day 29 MRD, CNS status, WBC count and age at diagnosis) is shown below the Cox proportional hazard log-likelihood P value with day 29 MRD as covariate. **i,j**, Change in BMP-like and T-specified signature scores in AALL0434 diagnosis and relapse sample pairs ($n = 27$; near-ETP, 4; non-ETP, 23). P values from a two-sided paired t-test are shown. The box includes the median, hinges mark the 25th and 75th percentiles and whiskers extend 1.5× the interquartile range.



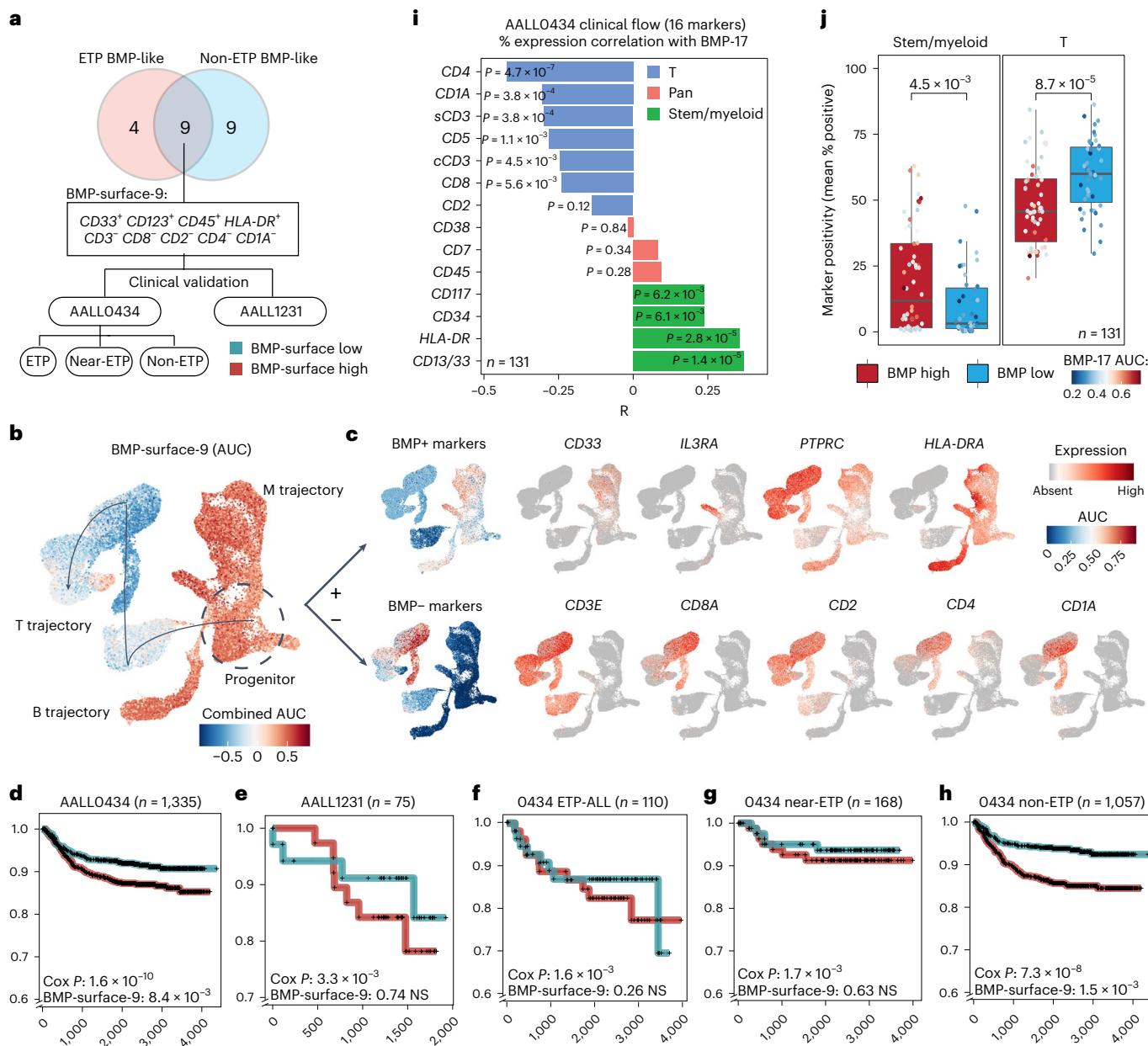


Fig. 7 | Clinical utility of the BMP-like surface immunophenotype in risk stratifying participants with T-ALL. **a**, Overlap of ETP BMP-like and non-ETP BMP-like DE surface markers to create consensus surface marker signature for risk stratification in AALL0434 (fully sequenced) and AALL1231 (partially sequenced). Positive surface markers were filtered for $\log_2 FC > 0.5$ and adjusted P value < 0.001 . Negative surface markers were filtered for $\log_2 FC < -0.5$ and adjusted P value < 0.001 . **b**, Aggregate signature score of BMP-surface-9 signature (AUC of positive markers – AUC of negative markers) calculated using AUCell in scRNA-seq reference ($n = 49,623$ cells). The T cell developmental trajectory is indicated with an arrow. The progenitor populations are circled. **c**, RNA expression of BMP-surface-9 marker genes within scRNA-seq reference of normal hematopoiesis ($n = 49,623$ cells). Positive marker genes are shown in the top row; negative marker genes are shown in the bottom row. Left: the AUC of positive and negative surface marker genes within healthy hematopoiesis. **d–h**, Kaplan–Meier plot showing the OS of bulk RNA-seq participants with T-ALL in AALL0434 and

AALL1231. Participants in each analysis were binarized using RNA-seq-derived expression of BMP-surface-9 signature genes. Participants are grouped by subtype, with non-subtyped participants ($n = 194$) grouped with participants without ETP. The prognostic value of the BMP-surface-9 signature in multivariate analysis (with day 29 MRD, CNS status, WBC count and age at diagnosis) is shown below the Cox proportional hazard log-likelihood P value with day 29 MRD as a covariate. **i**, Pearson correlation of percentage expression of T, pan, stem and myeloid flow cytometric markers among $n = 99$ participants without ETP. P values from a correlation test are shown. **j**, Average percentage expression of stem or myeloid (CD117, CD34, HLA-DR, CD13 and CD33) and T (CD4, CD1A, CD3, CD5, CD8 and CD2) in BMP-high versus BMP-low non-ETP cases. BMP-high and BMP-low were defined using bulk RNA-seq and matched with flow cytometric values by unique sequence index. P values from a two-sided Mann–Whitney test are shown. The box includes the median, hinges mark the 25th and 75th percentiles and whiskers extend 1.5× the interquartile range ($n = 56$ BMP high; $n = 43$ BMP low).

MRD and induction failure (Extended Data Fig. 9d,e) and reduced OS and EFS (Extended Data Fig. 9f). Lastly, we sought to determine whether we could make these signatures more parsimonious using machine learning regression, narrowing the BMP-17, BMP-surface-9 and BMP-119 signatures

to 6, 3 and 20 genes, respectively, while maintaining or improving prognostic stratification (Extended Data Fig. 9g–i). Together, our results support a treatment paradigm for T-ALL involving early genetic screening for the chemotherapy-refractory BMP-like phenotype.

BMP-like blasts are uniquely sensitive to BCL-2 inhibition

The universal existence of BMP-like populations across treatment-refractory T-ALL cases prompted us to develop a pipeline for the discovery of BMP-like directed targeted therapy. To support the modeling of BMP-like therapy response, we first expanded blasts from 22 single-cell-sequenced participants in NOD *scid* gamma (NSG) mice (Fig. 8a). scRNA-seq on engrafted blasts from 16 participants indicated strong retention of specific features (Fig. 8b), with BMP-high and BMP-low participants maintaining their respective phenotypes after engraftment (Fig. 8c–f).

We performed computational screening for targets specific to BMP-like blasts (Fig. 8g). We queried 552 BMP-like genes against three drug target databases (TTD, DrugIDB and OpenTargets), one transcriptomic-based compound screening database (LINCS1000) and one cancer gene vulnerability database (DepMap). The consensus results nominated four druggable surface proteins (CD44, LGALS1, ITGA4 and CD74), three homeostatic enzymes (S100A4, BCL-2 and Hsp90), two signal transduction molecules (Syk and Btk) and one TF (BCL-11A) (Extended Data Fig. 10a–d and Supplementary Table 17).

To test these computational predictions, we first performed in vitro drug screening using an established panel of 40 leukemia active drugs (Fig. 8h and Supplementary Table 18). PDX-expanded blasts from five of the 16 participants were screened using a stromal cell coculture system and dose–response curves were generated for each compound (Fig. 8i). Of 40 compounds, nine were active in all five PDX models and eight were active in at least one PDX model (Extended Data Fig. 10e,f). These compounds had different activity across BMP-high participants ($n = 3$) and BMP-low participants ($n = 2$) (Fig. 8i and Supplementary Table 19). After adding an additional five PDX models ($n = 5$ BMP-high and $n = 5$ BMP-low), we confirmed that BMP-high participants had increased sensitivity to venetoclax and navitoclax, while BMP-low participants were more sensitive to conventional cytotoxics (prednisolone, mercaptopurine and daunorubicin) (Fig. 8j). These associations were observed using bulk transcriptomic signatures across an integrated cohort⁴⁷ (Fig. 8k and Extended Data Fig. 10g). Although BTK returned as a four-database hit, comparable ibrutinib sensitivity was seen in BMP-like and T-specified blasts, likely resulting from DE of Tec kinases⁴⁸ between immature and mature T cells (Extended Data Fig. 10h).

The strong in vitro activity of BCL-2 inhibitors against BMP-like blasts prompted us to initiate in vivo efficacy studies in BMP-high (PATTDP, $n = 6$) and BMP-low (PAUNDK, $n = 9$) PDX models (Extended Data Fig. 10i–m). In the peripheral blood, venetoclax treatment resulted in the halting of disease progression in BMP-low models compared to control (Extended Data Fig. 10j). However, after the conclusion of the study, BMP-low PDX models still harbored notable residual disease within the bone marrow (>38% blasts) and spleen (>4% blasts) (Extended Data Fig. 10k). In contrast, venetoclax treatment resulted

in the robust clearance of disease in BMP-high PDX (Extended Data Fig. 10k–l), with a reduction in blasts beyond our limit of detection (<0.01%) in the majority of PATTDP PDX models of the bone marrow and spleen (Extended Data Fig. 10m and Supplementary Table 21). Our results support further clinical testing of BCL-2 inhibitors in refractory BMP-like T-ALL.

Discussion

Our study reports a comprehensive mapping of T-ALL to healthy human hematopoiesis. We report the surprising discovery that T cell leukemias differing drastically by bulk immunophenotype are linked at the sub-population level. Our integrated analysis identifies a shared BMP-like population tightly associated with treatment failure in ETP-ALL, near-ETP-ALL and non-ETP-ALL. This subpopulation can represent <5% of blasts at diagnosis, illustrating the limitations of current bulk-level tumor classification schemes. The cell of origin of ETP-ALL remains unknown and it is widely hypothesized that the transformation occurs in a T cell progenitor that maintains its immature cell state⁴⁹. However, experimental work in mice demonstrated that dedifferentiation of T cell-committed blasts can also lead to murine ETP-ALL⁵⁰. Interestingly, BMP-like blasts are also found in myeloid and mixed-phenotype leukemias, raising the possibility of one common cell of origin, which could be inferred through lineage-tracing approaches. Supporting evidence indicating a shared progenitor cell of origin includes an enrichment of TF and signaling gene mutations within BMP-like T-ALL blasts (similar to myeloid leukemias), non-T cell lineage marker expression and shared drug sensitivity profiles with myeloid leukemia stem cells. The unique opportunity to intersect our single-cell data with large-cohort bulk whole-exome sequencing (WES) and whole-genome sequencing (WGS) data allowed us to associate BMP-like phenotype and genotype. Within AALL0434, the NOTCH pathway was mutated in >70% of cases. Most *NOTCH1*-activating mutations (indels, SNVs and stop, frameshift and splice variants) were associated with superior 5-year OS and EFS; however, intronic *NOTCH1* SNVs and intragenic deletions were associated with worse outcomes. Our data isolate transcriptomic signatures from leukemic blasts carrying two co-occurring *NOTCH1* mutations, revealing a direct connection between *NOTCH1* mutation dosage and T cell lineage differentiation and refining our understanding of the functional transcriptomic changes that occur with *NOTCH1* mutation.

Although our study was heavily focused on pediatric T-ALL, our findings are perhaps even more relevant in the adult setting, where the ETP phenotype represents up to 52% of cases and 5-year survival rates are <50%. Detection of high-risk cases within young adults treated on AALL0434 using BMP-like signatures supports the hypothesis that BMP-like blast-mediated treatment failure extends beyond pediatric T-ALL cases. Multiple studies of pediatric⁵¹ (ALL-BFM 2000) and adult⁵² (LALA-94 and GRAALL-2003) trials found that NOTCH pathway activation was associated with favorable outcomes, further supporting a

Fig. 8 | Nomination and preclinical validation of targeted therapy against BMP-like blasts.

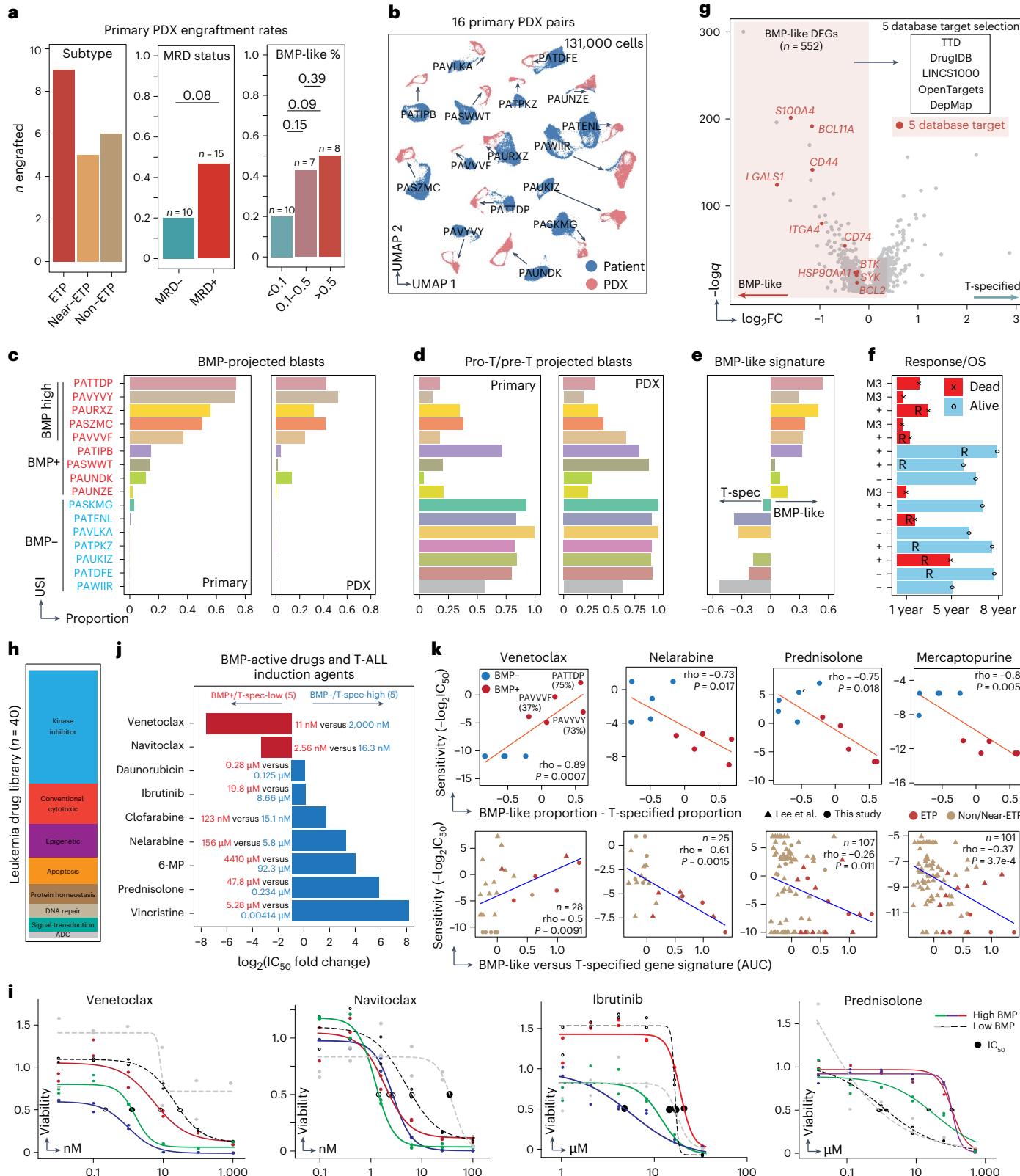
a, Total engrafted PDX by subtype, ETP-ALL PDX by MRD status and BMP-like proportion. *P* values from a two-sided proportion test are shown ($n = 22$ PDX models: ETP, 9; near-ETP, 5; non-ETP, 6). **b**, UMAP representation of $n = 16$ primary participant samples and $n = 16$ corresponding PDX models profiled using scRNA-seq ($n = 131,168$ cells: primary, 93,458; PDX, 37,710). PDX engrafted blasts are connected to their primary sample by arrows. **c**, Proportion of BMP-projected blasts (HSPC, LMPP or CLP) in $n = 16$ participant–PDX pairs. Left: samples are ordered by the proportion of BMP-projected blasts in the primary sample. Participants with detectable blasts (>1%) are considered BMP-positive, while participants with >25% blasts are considered BMP high. USI, unique study identifier. **d**, Proportion of T-specified projected blasts (pro-T cell and pre-T cell) in each participant–PDX pair ($n = 16$). **e**, Difference between BMP-like and T-specified signatures scored using AUCell on bulk RNA-seq samples from $n = 16$ participants. **f**, OS and EFS swimmer plot of $n = 16$ participants with paired PDX models. Events are labeled. Relapse is indicated by R. **g**, Computational screening

approach used to identify targetable genes within BMP-like blasts. A total of 552 BMP-like blast-specific DEGs (FDR < 0.05) were overlapped with drug target and dependency databases and ranked on the basis of the number of database hits and DE scores. The top ten targets by aggregate score are highlighted in red.

h, A panel of 40 drugs was tested on PDX engrafted blasts from $n = 5$ BMP-high and $n = 5$ BMP-low participants ($n = 10$ participants: ETP, 4; near-ETP, 2; non-ETP, 4). **i**, Representative dose–response curves for $n = 4$ nominated therapeutics that showed differential activity in BMP-high ($n = 3$) versus BMP-low ($n = 2$) leukemias. **j**, Relative activity of drugs active in BMP-high versus BMP-low leukemias ($n = 5$ each). Mean values are shown. **k**, Correlations between drug sensitivity ($-\log_{10}IC_{50}$) and the scRNA-seq derived BMP-like percentage (top) and the BMP-like signature score computed using 119 DEGs on bulk RNA-seq data (bottom). Bottom: the bulk RNA-seq correlations include the data from this study ($n = 10$) and data from Lee et al.⁴⁷ (venetoclax, $n = 28$; nelarabine, $n = 25$; prednisolone, $n = 107$; mercaptopurine, $n = 101$). Spearman's correlations and significance are shown.

common mechanism of treatment resistance mediated by *NOTCH1*-WT BMP-like subpopulations. Lastly, our single-cell multiomic reference maps of human hematopoiesis present a valuable resource to further delineate the impact of developmental heterogeneity in human leukemia by enabling a higher-resolution dissection of rare phenotypes beyond what is possible through the deconvolution of bulk transcriptomic data. Use of these reference maps within five subtypes of acute

leukemia underlined a BMP-like arrest state shared among lymphoid, myeloid and mixed-phenotype leukemic disease, highlighting an opportunity for further study. For instance, these reference maps and gene signatures could be used to study tumor evolution in the context of relapsed tumors and serial samples. Likewise, additional profiling is needed to identify the characteristics of blast states among the recently established transcriptionally defined T-ALL subtypes such



as the broader category of ETP-like leukemia¹³ in relation to the ETP, near-ETP and non-ETP immunophenotypic classification.

Collectively, our study identifies a rare but clinically important BMP-like subpopulation, which represents a promising therapeutic target for relapsed and refractory persons with T-ALL. Single-cell approaches on a carefully selected cohort were uniquely powered to isolate the BMP-like gene signature for risk stratification and therapeutic targeting, illustrating how high-resolution single-cell analyses are needed to supplement high-throughput bulk genomic approaches for understanding clinically relevant tumor biology. Further studies are needed to demonstrate the clinical utility of the prognostic gene signatures across a range of populations and mutational subtypes. Lastly, future mechanistic studies are expected to clarify the actions of the drugs identified to understand their specificity and the subsequent alteration of leukemic phenotypes to demonstrate their translational applicability.

Methods

AALL0434 participant identification and clinical annotation

COG studies AALL0434 ([NCT00408005](#)) and AALL1231 ([NCT02112916](#)) were approved by the National Cancer Institute Cancer Evaluation and Therapeutic Program, the US Food and Drug Administration, the Pediatric Central Institutional Review Board (IRB) and local IRBs at all participating centers. Written informed consent was obtained from study participants and, when appropriate, their legal guardians, in accordance with the Declaration of Helsinki. All participant data were deidentified and written informed consent was obtained to publish the indirect identifiers in the present manuscript. Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Secondary genomic studies were approved by the Children's Hospital of Philadelphia (CHOP) IRB. In total, 40 cases from AALL0434 (Supplementary Table 1) and eight healthy thymus and bone marrow controls (Supplementary Tables 5 and 6) were selected for single-cell study. The healthy thymus and bone marrow used for this work were residual tissues after collection for clinical care. Leukemia samples were bone marrow or blood samples collected and banked for COG trials. Within AALL0434, ETP status was centrally assessed in diagnostic bone marrow or peripheral blood samples using 8–9-color multiparameter flow cytometry²⁶. ETP was defined as having lymphoblasts that were CD8⁺ CD1a⁻ (<5% positive), weakly expressed CD5 (either <75% positive or median intensity more than one log less than mature T cells) and expressed one or more myeloid or stem cell markers (>25% positive) including CD13, CD33, CD34, CD117 and HLA-DR (ref. 13). Subjects meeting the ETP immunophenotypic criteria but with stronger expression of CD5 were classified as near-ETP. Subjects with neither ETP nor near-ETP were defined as non-ETP. MRD was assessed using 8–9-color flow cytometry and was performed using established methods at a COG flow cytometry reference laboratory (University of Washington or Johns Hopkins University).

Processing of T-ALL diagnosis samples

Peripheral blood or bone marrow aspirate samples were thawed at 37 °C, treated with 1:10 (v/v) 1 mg ml⁻¹ DNase I (Sigma-Aldrich, D4513) for 90 s at 37 °C, resuspended in 10 ml of Iscove's modified Dulbecco's medium (IMDM) + 2% FBS and centrifuged (160g for 5 min). Samples were retreated with DNase I and resuspended in fluorescence-activated cell sorting (FACS) buffer (Ca²⁺-free and Mg²⁺-free PBS + 1% BSA). Cell number and viability were recorded using a Countess II cell counter (Invitrogen). More than 1 million live cells were aliquoted for tail-vein injection into NSG mice, with the remaining stained with DAPI (Invitrogen, D1306) and subjected to FACS sorting (FACSAria Fusion, BD).

scRNA-seq and CITE-seq library preparation

FACS-sorted DAPI-negative live cells were centrifuged and resuspended in cell staining buffer (BioLegend, 420201) at 45 µl per million cells.

Cells were blocked with Human TruStain FcX (BioLegend, 422301) at 5 µl per million cells (4 °C, 15 min). After blocking, cells were stained with a TotalSeq-A antibody cocktail (30 min, 4 °C). Cells were washed three times using cell staining buffer (BioLegend, 420201) and resuspended in PBS + 0.04% BSA. Cells were counted using a Countess II cell counter. A total of 17,000 cells per sample were then loaded onto 10x Genomics Chromium controller and processed with the Chromium NEXT GEM single-cell 3' reagent kit (version 3.1). GEX libraries were constructed using the 10x Genomics library preparation kit following the instructions. Antibody-derived tag (ADT) libraries were constructed using the KAPA HiFi HotStart ReadyMix kit (Kapa Biosystems, KK2601). The following program was used for ADT library PCR: 98 °C for 2 min, 14–15 cycles of 98 °C for 20 s, 60 °C for 30 s and 72 °C for 20 s, followed by 72 °C for 5 min and a hold at 4 °C. Library quality was checked using the Agilent high-sensitivity DNA kit (Agilent, 5067-4626) and Bioanalyzer 2100. Libraries were quantified using the dsDNA high-sensitivity assay kit (Invitrogen, Q33231) on a Qubit fluorometer and quantified using the qPCR-based KAPA quantification kit (Kapa Biosystems, KK4844). Libraries were sequenced on an Illumina NovaSeq 6000 with 28:8:0:87 paired-end format.

scATAC-seq library preparation

DAPI-negative live cells were centrifuged at 300g (5 min at 4 °C), mixed in 45 µl of lysis buffer and incubated (3 min on ice). Next, 50 µl of prechilled wash buffer was added without mixing and centrifuged immediately at 300g (5 min at 4 °C). Then, 95 µl of supernatant was discarded, 45 µl of diluted nuclei buffer (10x Genomics) was added and the sample was centrifuged (300g; 5 min at 4 °C). The nuclear pellet was then resuspended in 7 µl of prechilled diluted nuclei buffer and the nuclear concentration was determined using a Countess II cell counter. A total of 7,000–20,000 nuclei were used for the transposition reaction in bulk, loaded onto the 10x Genomics Chromium controller and processed with the Chromium NEXT GEM scATAC reagent kit (version 1.1). Library quality was checked using the Agilent high-sensitivity DNA kit and Bioanalyzer 2100. Libraries were quantified using the dsDNA high-sensitivity assay kit on a Qubit fluorometer and quantified using the qPCR-based KAPA quantification kit. Libraries were sequenced on an Illumina NovaSeq 6000 with 49:8:16:49 paired-end format.

Expansion and profiling of T-ALL blasts in PDX

NSG mice (RRID:IMSR_JAX:005557) were used for all experiments. For the development of PDX models, we injected ~10⁶ blasts from viably frozen participant samples (bone marrow or blood) per mouse to develop primagrafts (Supplemental Table 22). PDX-expanded blasts were isolated from the spleen or bone marrow. Frozen samples were thawed (37 °C), resuspended in IMDM + 2% FBS and treated with DNase I twice. Cells were washed twice with RPMI medium, resuspended in flow buffer, stained with DAPI and anti-human CD45 antibody (BD Pharmingen, 555485) and subjected to FACS sorting (FACSAria Fusion, BD). DAPI-negative hCD45⁺ sorted cells were stained with 10x Genomics 3' CellPlex multiplexing solution, washed three times and immediately processed using the 10x Genomics Chromium controller and the Chromium NEXT GEM single-cell 3' reagent kit (version 3.1). The 3' GEX libraries were constructed using the 10x Genomics library preparation kit. CellPlex libraries were constructed using the 10x Genomics 3' CellPlex kit. Library quality was checked using the Agilent high-sensitivity DNA kit and Bioanalyzer 2100. Libraries were quantified using the dsDNA high-sensitivity assay kit on a Qubit fluorometer and quantified using the qPCR-based KAPA quantification kit. Libraries were sequenced on an Illumina NovaSeq 6000 with 28:8:0:87 paired-end format.

CD34⁺ progenitor isolation from infant or pediatric thymi

Pediatric thymi were obtained and used according to and with the approval of the CHOP IRB. Thymus tissue was mechanically disrupted and treated with liberase (0.2 mg ml⁻¹, 30 min at 37 °C; Roche) with

intermittent shaking, as previously described¹⁹. Thymocytes were resuspended into flow buffer, sorted into DAPI-negative lineage-negative CD34⁺CD1A⁻ fractions and subjected to scRNA-seq and scATAC-seq.

Projection onto healthy reference trajectory

Participant-derived cells were projected onto the healthy reference trajectory using the MapQuery function in Seurat 4.0.5. For scRNA-seq data, participant and healthy control data were coembedded into a low-dimensional space using the default anchor-based canonical correlation analysis (CCA) method in Seurat 4.0.5 (30 dimensions, 2,000 anchor features) and cell type label transfer was performed on a sample-by-sample basis using the TransferData function. For scATAC-seq data, peaks from participant and healthy reference data were merged using the mergePeaks module from scATAC-pro⁵³ and peak × cell matrices with merged peaks were reconstructed for each participant with the scATAC-pro reConstMtx module. This allowed for participant and healthy control data to be coembedded into a low-dimensional space analogous to the scRNA-seq data.

AALL0434 ETP-ALL stratification using BMP-119 signature

BMP-like and T-specified DEGs were stringently filtered using cutoffs of false discovery rate (FDR) < 0.001 and average log₂ fold change (log₂FC) > 0.5, leaving 66 BMP-like DEGs and 53 T-specified DEGs. The z-score-based signature scoring was performed on 110 bulk-sequenced participants with ETP-ALL with BMP-like DEGs as positive features and T-specified DEGs as negative features. For each participant, the mean T-specified feature z-score was subtracted from the mean BMP-like feature z-score, with a score of >0 being interpreted as more BMP-like than T-specified. This cutoff was selected to compare participants on the basis of a relative enrichment of either phenotype. Participants were binarized by BMP-like signature score (BMP-like > T-specified versus T-specified > BMP-like) and OS and EFS were compared using the Cox proportional hazard model with day 29 MRD and CNS status taken as covariates using the survfit function from survival 3.2-13 ('survfit(Surv(time.survival, status.survival) ~ high.BMP + D29.MRD + D29.CNS.status').

Integration of single-cell signatures with mutation calls

Bulk RNA-seq data for $n = 110$ ETP samples with corresponding WES and WGS mutation calls were scored using 66 BMP-like DEGs and 53 T-specified DEGs using AUCell 1.12.0. For 1,490 mutant genes in 110 ETP samples, the number of samples carrying mutations was quantified and the mean BMP-like area under the curve (AUC) and T-specified AUC were calculated. Mutant genes observed in ≥ 5 samples with mean VAF > 0.05 were plotted for visualization. Classification of genes was derived from a previous bulk genomics study on ETP-ALL. For fusion drivers, the mean BMP-like AUC, the mean T-specified AUC, the percentage of participants with positive EOI MRD, the percentage of participants that died during the trial and the number of unique fusion partners were calculated.

Identification of a consensus BMP-like gene signature

BMP-like DEGs from participants with ETP-ALL ($n = 56$ BMP-like versus T-specified) and participants with non-ETP-ALL ($n = 445$ BMP-like versus postcommitment) were overlapped and the average log₂FC was calculated. A total of 17 genes with average log₂FC > 0.9 were retained as a consensus BMP-17 signature. We performed AUC-based signature scoring using AUCell 1.12.0 (with the top 25% of expressed genes) on bulk RNA-seq diagnostic T-ALL samples from two independent COG trials using BMP-17 DEGs. We then binarized participants on the basis of AUCell signature score and used the Cox proportional hazard model with EOI MRD and CNS status taken as covariates using the survfit function from survival 3.2-13 ('survfit(Surv(time.OS, status.OS) ~ BMP-17 + high.BMP-17 + D29.MRD + CNS.status)'). In each case, the top half of participants was compared to the bottom half of participants.

In silico drug screening against BMP-like blasts

Drug-target data from two independent drug target databases (TTD⁵⁴ and DrugIDB⁵⁵) and a third database (OpenTargets⁵⁶) that focuses on next-generation targets were overlapped with BMP-like DEGs (log₂FC > 0.2; adjusted $P < 0.01$). Targetable gene products were given a score of 1 for each database in which a resulting hit was obtained. To search for drugs that could specifically modify the BMP-like state, we inputted top BMP-like DEGs and TFs ($n = 56$) and top T-specified DEGs and TFs ($n = 62$) into the LINCS1000 (ref. 57) database under default parameters. Perturbation results were filtered in R to filter compound-mediated perturbations for compounds with defined targets, statistical significance (log₁₀FDR > 1), effect size (normalized connectivity score > 0.8), specificity to BMP-like state (raw connectivity score > 0) and activity in two or more leukemia cell lines. Non-compound perturbations were filtered for statistical significance (log₁₀FDR > 1) and effect size (normalized connectivity score > 0.8) and further separated into gene overexpression and gene knockdown (including short hairpin RNA knockdown, clustered regularly interspaced short palindromic repeats knockout and ligand-based perturbation) classes. BMP-like DEGs targeted by top compound perturbations and/or genes with overexpression or knockdown were given a score of 1. Lastly, we identified BMP-like DEGs that showed increased dependency in leukemia cell lines ($n = 59$) compared to non-leukemia and non-lymphoma cell lines ($n = 1,052$) in the cancer dependency map⁵⁸ (DepMap) portal. Genes with negative dependency scores in leukemia cell lines (mean dependency score < -0.1), dependency FC > 2 and >25% expression in BMP-like blasts were assigned a score of 1. Next, BMP-like DEGs with log₂FC > 1 were assigned a score of 1 and genes with log₂FC between 0.5 and 1 were assigned a score of 0.5. Beyond high expression change, we prioritized BMP-like DEGs with high percentage expression in BMP-like blasts; genes with >80% expression were assigned a value of 1, whereas genes with 50–80% expression were assigned a value of 0.5. Finally, we ranked genes with high statistical significance (adjusted $P < 1 \times 10^{-100}$ was given a score of 1; adjusted $P < 1 \times 10^{-50}$ was given a score of 0.5). The sum of DE evidence (ranging from 0 to 3) and database evidence (ranging from 0 to 5) was taken to rank BMP-like DEGs for follow-up experimental studies.

In vitro drug screening with leukemia active drug panel

Human leukemia blasts were collected from mouse spleen and enriched using an immunomagnetic isolation kit (StemCell Technologies, 19849) and screened with a panel of 40 leukemia active drugs (Supplementary Table 18) using a previously described imaging-based assay with a stromal cell coculture system⁵⁹.

Nomination of BMP-like specific drugs from drug screening

PDX-expanded blasts from five participants were screened in a stromal coculture system and dose-response curves were generated for each compound with the primary readout being cell viability (as a percentage of control). We defined ETP active drugs as compounds with IC₅₀ < 1,000 nM and categorized each compound as not active (active in 0/4 participants with ETP), partially active (active in 1–3 participants with ETP) or active (active in 4/4 participants with ETP). We then compared IC₅₀ values for ETP active compounds and used three comparisons to nominate drugs that were differentially active in high-BMP participants: BMP-high and MRD-positive ($n = 3$) versus BMP-low and MRD-negative ($n = 2$: one ETP and one non-ETP); BMP-high and MRD-positive ($n = 2$) versus BMP-low ($n = 1$); BMP-high and MRD-positive ($n = 3$) versus BMP-low ($n = 1$). Drugs with differential activity in BMP-high participants in all three comparisons were nominated as BMP-specific drugs. The sensitivity of these drugs was confirmed using PDX-expanded blasts from five additional participants, for a total of five BMP-high models and five BMP-low models.

scRNA-seq and CITE-seq data processing

Demultiplexing and alignment of RNA and ADT sequences were performed with Cell Ranger 3.1.0. Low-quality cells and red blood cells were then filtered by retaining only cells with between 300 and 2,500 genes in the scRNA-seq data, greater than 1,500 RNA counts, less than 10% mitochondrial RNA and fewer than three unique molecular identifiers (UMIs) mapping to hemoglobin B. To remove cell doublets in scRNA-seq data, DoubletFinder 2.0.3 was run with 5% of the expected rate of doublets. Participant cell × gene and cell × ADT count matrices were individually saved and subsequently concatenated using Seurat 4.0.5 for downstream analyses. For some analyses as specified, cell × gene and cell × ADT matrices for each participant were subset for G1 cells (representing the phenotype most resistant to conventional therapy) and were downsampled to match the lowest value in the cohort. After log-normalization, the FindVariableFeatures function in Seurat 4.0.5 was used to identify the top 5,000 features with greater than expected variance. Variable features with expression in >1% of cells were kept as the input to principal component analysis (PCA), with subsequent visualization being performed using uniform manifold approximation and projection (UMAP) of the top 50 principal components (PCs), 30 neighbors and two PCs. For visualization, we used the IntegrateLayers function in Seurat 5.0.3 with the RPCAIntegration method and default parameters.

scATAC-seq data processing

Demultiplexing of scATAC-seq reads was performed with Cell Ranger-ATAC 1.1.0 (alignment to hg38) and peak calling was performed with BWA and MACS2 using the scATAC-pro pipeline⁵³ with default parameters. Low-quality cells were filtered for those cells with <3,000 (low quality) and >50,000 unique fragments (doublets), <40% reads in peaks (fraction of reads in peaks < 0.4) and >20% reads mapping to mitochondria. To construct a common peak set, the top 100,000 peaks (defined by MACS2 MapQ score) were selected for downstream merging, alongside 1,500 randomly selected cell barcodes from each participant. We defined two sample sets for merging peaks: one with 25 participants with ETP-ALL and one with 40 participants with T-ALL. For each sample set, peaks were merged with the scATAC-pro mergePeaks module and peak × cell matrices with merged peaks and downsampled cell barcodes were reconstructed with the scATAC-pro reConstMtx module.

Bulk RNA-seq analysis and visualization

Sequencing read adaptors were removed using Trim Galore 0.4.4 (https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/) with parameters ‘-q 20 --phred33 -paired’. Reads were aligned to the human genome GRCh38 using STAR. The resulting gene counts for each sample were estimated by RSEM, and combined as gene count matrix. RSEM expected counts were processed and filtered. First, samples were required to exhibit expression of over one count per million in ≥5 samples. Second, batch correction was performed using the sva R package 3.46 function ComBat_seq. Batches were defined on the basis of library type as stranded or unstranded and on the basis of cohort as TARGET or X01 sequenced samples. Third, the DESeq2 R package 1.38.3 vst function was used for data normalization. Limma 3.46.0 was used for DE analysis. For visualization, the raw counts were transformed into transcripts per million and visualized using UMAP with the top ten PCs and $k = 30$ neighbors.

Construction of healthy reference trajectory

Construction of the healthy reference trajectory began with sample-by-sample cell annotation followed by consensus clustering and annotation. Annotations from previously published bone marrow samples were kept⁶⁰. Cell × gene matrices from each thymus sample were log-normalized and subjected to dimensionality reduction. Cells were clustered at multiple resolutions ($k = 1, 2$ and 3) and clusters were given preliminary labels on the basis of marker gene

expression. Cell × gene matrices from all thymus donors ($n = 3$) were then concatenated, log-normalized and subjected to dimensionality reduction. Cells were reclustered at high resolutions ($k = 3$) and clusters were relabeled on the basis of marker gene expression and prior labels.

scRNA-seq

Cell × gene matrices from healthy thymus donors were then concatenated with cell × gene matrices from healthy bone marrow donors ($n = 5$), log-normalized and subjected to dimensionality reduction using the top 25 PCs. The FindVariableFeatures function in Seurat was used to identify top 2,000 variable genes. These 2,000 genes were then filtered in two iterations on the basis of the Gini coefficient⁶¹. Briefly, a shared nearest neighbor graph was constructed using 50 and 20 PCs; cells were clustered at $k = 0.1$ resolution and the Gini coefficient was calculated for each variable gene. Genes with a low Gini coefficient (bottom 10% percentile) and cluster level expression < 10% were removed in each iteration. The 134 cell-cycle-related genes previously described¹⁹ were removed. The remaining 931 variable features were used as input to PCA and UMAP dimension reductions (25 PCs). Trajectory analysis was performed using Slingshot 1.8.0 with HSPC as the start cluster and effector T, mature B and monocyte as the end clusters for T, B and myeloid trajectories, respectively. Principal curves were selected for T cell and myeloid cell trajectories and values were scaled to a maximum of 1 in each curve. Pseudotime values of shared cell states that occurred in both myeloid and T cell development (multipotent progenitors: HSPC and LMPP) were then averaged. Pseudotime values were scaled for myeloid development (0 to -1) and T cell development (0 to 1). Statistical comparisons in the overall arrest state were made using a two-sample Kolmogorov–Smirnov test, as previously described⁶².

scATAC-seq

Gene–activity matrices for scATAC-seq were constructed by summing counts within the gene body and 2 kb upstream, as previously described⁶⁰. Integration of scATAC-seq samples was performed using gene–activity matrices and Seurat 4.0.5 using the default anchor-based CCA method with 30 dimensions, 2,000 anchor features and k.filter = 100. To learn labels for scATAC-seq data from scRNA-seq data, transfer anchors were computed using CCA with scRNA-seq as the reference and cell type label transfer was performed on a sample-by-sample basis using the TransferData function. Cell × peak matrices from all thymus donors were then concatenated. Cells were reclustered at high resolution ($k = 3$) and clusters were reannotated according to consensus labels. Dimensionality reduction was performed using UMAP of the top ten PCs of the concatenated scATAC-seq data and trajectory analysis was performed as described above.

Level 1 annotation of CITE-seq data

To distinguish malignant blasts from non-malignant cells, we first used a cluster-based statistic, Shannon entropy, to identify clusters of cells at multiple clustering resolutions ($k = 1, 2$ and 3) to identify four cell populations that had contribution from every participant. The Shannon entropy statistic was calculated using the formula $-\sum p(x) \times \log p(x)$, where $p(x)$ is the frequency of cells arising from any one participant in any one cluster, ranging from 0 to 1. Second, we concatenated and clustered participant-derived single-cell data with healthy bone marrow and thymus controls. Third, we calculated a similarity score to healthy controls across all participant-derived cells. Participant data and healthy control data were coembedded into a low-dimensional space using the default anchor-based CCA method in Seurat 4.0.5 (30 dimensions and 2,000 anchor features) and a $k = 30$ mutual nearest neighbor score was assigned for each cell using the TransferData function. Copy number profiles were analyzed using InferCNV 1.6.0 on a randomly downsampled (1:10) subset of participant data. We then compared blast percentages calculated in scRNA-seq to pathology reports of blast percentage obtained from diagnostic aspirate (mean

absolute deviation = 8.9%; non-significant difference according to paired two-tailed *t*-test).

Level 1 annotation of single-nucleus (sn)ATAC-seq data

Firstly, annotated scRNA-seq data were used as a reference to annotate participant scATAC-seq data on a paired, sample-by-sample basis. For each participant, gene–activity matrices for scATAC-seq were constructed by summing counts within the gene body and 2 kb upstream, as previously described⁶³. Integration of scATAC-seq samples with scRNA-seq data was performed using gene–activity matrices and Seurat 4.2.0 using the default anchor-based CCA method with 30 dimensions, 2,000 anchor features and *k.filter* = 100 using the TransferData function on a sample-by-sample basis. Then, participant data and healthy control data were coembedded into a low-dimensional space using the default anchor-based CCA method in Seurat 4.0.5 (30 dimensions, 2,000 anchor features and *k.filter* = 100) and a *k* = 30 mutual nearest neighbor score was assigned using the TransferData function to assess their similarity. Lastly, blast percentages calculated in scATAC-seq were compared to blast percentages calculated in scRNA-seq, showing high concordance (median deviation = 1.2%; non-significant difference according to paired two-tailed *t*-test).

Differential activity analyses

For TF motif enrichment analysis, cell × deviation score matrices were generated using the addGCBias, matchMotifs, getBackgroundPeaks and computeDeviations functions in chromVAR 1.12.0 with hg38 as the reference genome. Differential activity analysis was performed using the Wilcoxon rank-sum test with Benjamini–Hochberg multiple-testing correction with downsampling. For each motif in any particular comparison, we calculated the Δ median chromVAR deviation score, Δ mean chromVAR deviation score, adjusted *P* value, percentage expression of corresponding TF in paired scRNA-seq data and the ratio of median and mean chromVAR deviation score. DA TF motifs were defined by Δ median chromVAR deviation score > 0.0025, FDR < 0.001, >20% cell expression of corresponding TF and a ratio of median and mean chromVAR deviation score between 0.7 and 1.3, unless otherwise specified.

Subtype-specific transcriptional regulatory analysis

As we previously described⁶⁰, for each cell in the scRNA-seq dataset, an scRNA-seq and scATAC-seq ‘metacell’ was defined by pooling counts for each gene or peak from the 30 nearest neighbors in the PC space by cosine distance. Metacell counts were log-normalized and scaled. For a gene of interest, we ran a linear regression model using metacell gene expression as the dependent variable and putative enhancer peaks within 500 kb of the transcription start site as regressors. Bonferroni-adjusted *P* values < 0.01 with a positive coefficient were considered significant. Top induced targets of *TCF7* and *LEF1* were defined by high-confidence EP regression (regression coefficient > 0.3) and $\log_2\text{FC}$ > 0.5 for z-scoring on bulk RNA-seq data.

Promoter–enhancer coaccessibility networks (CCANs)

Cicero⁶⁴, which identifies coaccessible pairs of DNA elements, was implemented in Signac⁶⁵ through the make_cicero_cds function followed by the run_cicero function with the following parameters: sample_num = 100, window = 500,000. These links were aggregated into *cis*-coaccessible networks using the generate_ccans function with default parameters. The BMP-like and T-specified CCANs were isolated by identifying links to regions within 2,000 bp of the transcription start sites for the 66 and 53 DEGs for each state, respectively. Any peaks that overlapped regions within that coaccessibility group were then subset as potential regulators. This yielded 660 peaks in the BMP-like CCAN and 1,011 peaks in the T-specified CCAN. Those peaks were then used as input to HOMER with parameters ‘-size 200 -mask’ to identify motifs enriched in coaccessible regions. Motifs with a *q* value < 0.05 were considered significant.

Transcriptional regulatory analysis of developmental states

An integrated enhancer-driven transcriptional regulatory analysis was conducted using SCENIC+ 1.0a1 (ref. 44) following the standard vignettes with minor modifications. BMP-like and T-specified states from the scRNA-seq data and scATAC-seq data were extracted and 35 topics were empirically selected. The SCENIC+ pipeline was then run in non-multiome mode, using five cells per metacell. The search space was defined as 0–500 kb. Regulons were filtered with the following parameters: rho_threshold = 0.03, min_regions_per_gene = 0 and min_target_genes = 10. All other parameters were maintained as the defaults. Region-based and gene-based specificity scores for the BMP-like and T-specified states were calculated using the regulon_specificity_scores function.

AUCcell pathway analysis and GSEA

Pathway analysis was conducted using two methods. First, pathway enrichment scores for gene signatures were defined from our single-cell analysis, including the gene sets defined through our transcriptional regulatory analyses or downloaded from the Molecular Signatures Database⁶⁶, and were determined using AUCcell 1.12.0 with the top 5% of genes. Additional gene sets for NOTCH activation were previously published^{67,68}. GSEA was conducted to compare DE pathways between blast populations. A full gene list was constructed using the FindMarkers function in Seurat 5.0.3 with the following parameters: min.pct = 0.001, logfc.threshold = 0, only.pos = FALSE, max.cells.per.ident = 1,500. This gene list was sorted by $\log_2\text{FC}$ to use as input to preranked GSEA using the fgsea package⁶⁹.

Cell-cycle analyses in single-cell-sequenced participants with ETP-ALL

Cell-cycle signature scoring and phase classification was performed on ten high-MRD and ten MRD-negative participants with ETP using the CellCycleScoring function in Seurat 4.0.5 with default parameters. A total of 43 S-phase and 54 G2M-phase signature genes⁷⁰ were provided as input. Cells were then randomly downsampled so that each participant would be represented by an equal cell number (3,350 per participant and 33,500 per group).

BMP-like DE analyses

BMP-like and T-specified DEGs were computed using the FindMarkers function in Seurat 4.0.5 with the following parameters: assay = RNA, logfc.threshold = 0, ident.1 = T-specified-R (T-specified blasts from ten MRD-negative participants), ident.2 = BMP-like-NR (BMP-like blasts from 15 MRD-positive participants) and max.cells.per.ident = 1,500. The input matrix to DE analysis was a matrix of G1-phase ETP-ALL blasts with an equal number of cells per participant (1,711 per participant and 42,775 cells total). To identify DE TFs and DE surface markers, the same process was repeated with using genes encoding human TFs^{71,72} (feature = TFs) and a change of assay to normalized ADT count matrix (assay = ADT).

Intersection of DE TF and DA motifs

DE TFs from scRNA-seq data, defined by average $\log_2\text{FC}$ > 0.15 and FDR < 0.001, were intersected with DA TF motifs from scATAC-seq data. DA TF motifs were defined by Δ median chromVAR deviation score > 0.0025, FDR < 0.001, >20% cell expression of corresponding TF and a ratio of median and mean chromVAR deviation score between 0.7 and 1.3.

Identification of NOTCH1 mutations in scRNA-seq

Samples were demultiplexed into FASTQ files using bcl2fastq. FASTQ files were then processed using IronThrone 2.1 with the default parameters and inputs for 10x version 3.1 scRNA-seq data. Specifically, for each variant, IronThrone was run in circularization mode (–run = circ) with UMI length 12 (–umilen 12) and cell barcodes from each sample’s

Cell Ranger output (--whitelist sample_specific.barcodes.tsv), following the configuration set within IronThrone 2.1 documentation (<https://github.com/dan-landau/IronThrone-GoT>).

Identification of BMP-like blasts in participants without ETP

Precommitted blasts in ten participants without ETP (six EOI-MRD-negative participants, 7,152 precommitted blasts; four EOI-MRD-positive participants, 11,047 precommitted blasts) were subsetted (total of 52,971 blasts and 15,830 blasts, respectively) and mean proportions for corresponding cell fractions (BMP-like, MEP-like and pro-T cell-like) were quantified in each participant. The mean proportion of each cell type for participants of each group was plotted, with the proportion of BMP-like blasts in MRD-negative versus MRD-positive participants being compared using the prop.test function.

Single-cell signature-based stratification of non-ETP cases

Precommitment and postcommitment DEGs were computed as described above. The input matrix to DE analysis was a matrix of G1-phase non-ETP-ALL blasts with a maximum of $n = 1,500$ cells per participant (34,384 cells in total). DEGs located on the X and Y chromosomes were filtered out to retain the core biology of both cell fractions. A z-score-based signature scoring was performed on 1,051 bulk-sequenced diagnostic participants with non-ETP-ALL with BMP-like DEGs as positive features and T-specified DEGs as negative features. Human TFs were previously curated^{71,72}. Survival was analyzed using Cox proportional hazards as described above.

Identification of a BMP-like surface marker signature

DE ADTs from ETP-ALL BMP-like blasts and non-ETP BMP-like blasts were overlapped and the average log₂FC was calculated. Nine genes with |log₂FC| > 0.5 (five with positive expression and four with negative expression) were retained as a consensus BMP-surface-9 signature. Gene signature scoring and survival analysis were conducted as described above.

LASSO (least absolute shrinkage and selection operator) optimization of prognostic gene signatures

Refined gene sets were found using LASSO penalized regression. The Cox proportional hazards model was used with gene z-scores as features (glmnet and survival R packages). The model was adjusted for participant age, white blood cell (WBC) count at diagnosis, CNS status and treatment protocol, by including these as covariates on which no penalty was applied. The penalty was only applied to the gene features but the range of the predicted coefficients was bounded such that genes enriched in the BMP signature were given positive coefficients (increased hazard) and genes enriched in the T-specified signature were given negative coefficients (decreased hazard). The models were fit to the entire AALLO434 RNA-seq dataset and ETP status was used to stratify the survival allowing for different baseline hazards, followed by 100-fold cross-validation to determine the penalty parameter with the lowest root-mean-squared error.

Identification of BMP-like signature expression patterns

To contextualize BMP-17 and BMP-surface-9 marker genes within normal hematopoiesis, we performed AUC-based signature scoring of healthy donor scRNA-seq reference maps using AUCell with the top 10% of expressed genes considered for computational efficiency (auc-MaxRank = 0.1). BMP-surface-9 marker genes were divided into positive DEGs ($n = 4$) and negative DEGs ($n = 5$) and AUCs were calculated for each gene set. Overall enrichment (that is, the aggregate AUC) of the BMP-surface-9 signature was calculated by taking the difference in AUC between positive and negative surface markers.

Integration of AALLO434 and Lee et al.'s bulk RNA-seq data

In vitro drug sensitivity data were integrated from the current study (ten PDX models) with previously published drug screening data⁴⁷. To

generate a consensus BMP-like signature across data, we first converted the AALLO434 bulk transcriptomic data to fragments per kilobase of transcript per million mapped reads format using the convertCounts function in the DGEobj.utils package. A BMP-like gene signature was scored using the 119 BMP-like and T-specified DEGs with a robust z-score. The negative features (T-specified DEGs) were subtracted from the positive features (BMP-like DEGs).

Statistics and reproducibility

No statistical method was used to predetermine sample size. All data meeting the quality control threshold were included. The investigators were not blinded to allocation during genomic profiling and assessment of participant data. No data points were excluded from analyses related to single-cell and bulk genomics. No animal models were excluded from PDX-related analyses. Randomization and blinding were used for all in vitro and in vivo experiments. Statistical comparisons were made using a two-sided Wilcoxon rank-sum test unless otherwise specified in the figure legend. As the Wilcoxon test is non-parametric, we did not formally test for normality of the data. The chi-squared test was used to compare cell type proportions. The Cox proportional hazards model was used for the assessment of survival outcomes.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

All primary sample and PDX sequencing data generated for this study are available through dbGaP under accession number [phs003432](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1000032/) as part of the Childhood Cancer Data Initiative. This includes raw sequencing data of all primary participants with T-ALL (CITE-seq and scATAC-seq), healthy thymus (CITE-seq and scATAC-seq) and T-ALL PDX models (CITE-seq), along with the corresponding count matrices and Seurat objects. Healthy bone marrow samples included in the analysis were previously published⁶⁰. AML and MPAL samples included for comparative analysis can be accessed at <https://scpc.alexlemonade.org/projects/SCPCP000003#samples>. All other data supporting the findings of this study are available from the corresponding author on reasonable request. Requests will be processed within 30 days. Source data are provided with this paper.

Code availability

All code generated for processing of the CITE-seq and scATAC-seq data, as well as for all key analyses in this work are available from GitHub (https://github.com/tanlabcode/SC_TALL). Any other code involved in conducting the analysis and generating the figures will be made available upon reasonable request to the corresponding author.

References

1. Hunger, S. P. & Mullighan, C. G. Acute lymphoblastic leukemia in children. *N. Engl. J. Med.* **373**, 1541–1552 (2015).
2. Teachey, D. T., Hunger, S. P. & Loh, M. L. Optimizing therapy in the modern age: differences in length of maintenance therapy in acute lymphoblastic leukemia. *Blood* **137**, 168–177 (2021).
3. Salvaris, R. & Fedele, P. L. Targeted therapy in acute lymphoblastic leukaemia. *J. Pers. Med.* **11**, 715 (2021).
4. Foà, R. et al. Dasatinib–blinatumomab for Ph-positive acute lymphoblastic leukemia in adults. *N. Engl. J. Med.* **383**, 1613–1623 (2020).
5. Kantarjian, H. et al. Blinatumomab versus chemotherapy for advanced acute lymphoblastic leukemia. *N. Engl. J. Med.* **376**, 836–847 (2017).
6. Kantarjian, H. et al. Inotuzumab ozogamicin versus standard care for acute lymphoblastic leukemia. *N. Engl. J. Med.* **375**, 740–753 (2016).

7. Park, J. H. et al. Long-term follow-up of CD19 CAR therapy in acute lymphoblastic leukemia. *N. Engl. J. Med.* **378**, 449–459 (2018).
8. Den Boer, M. L. et al. A subtype of childhood acute lymphoblastic leukaemia with poor treatment outcome: a genome-wide classification study. *Lancet Oncol.* **10**, 125–134 (2009).
9. Dt, T. & Ch, P. Comparative features and outcomes between paediatric T-cell and B-cell acute lymphoblastic leukaemia. *Lancet Oncol.* **20**, e142–e154 (2019).
10. Wood, B. L. et al. T-lymphoblastic leukemia (T-ALL) shows excellent outcome, lack of significance of the early thymic precursor (ETP) immunophenotype, and validation of the prognostic value of end-induction minimal residual disease (MRD) in Children's Oncology Group (COG) study AALL0434. *Blood* **124**, 1 (2014).
11. Swerdlow, S. H. et al. The 2016 revision of the World Health Organization classification of lymphoid neoplasms. *Blood* **127**, 2375–2390 (2016).
12. Duffield, A. S., Mullighan, C. G. & Borowitz, M. J. International Consensus Classification of acute lymphoblastic leukemia/lymphoma. *Virchows Arch.* **482**, 11–26 (2023).
13. Pöllönen, P. et al. The genomic basis of childhood T-lineage acute lymphoblastic leukaemia. *Nature* **632**, 1082–1091 (2024).
14. Liu, Y. et al. The genomic landscape of pediatric and young adult T-lineage acute lymphoblastic leukemia. *Nat. Genet.* **49**, 1211–1218 (2017).
15. Anand, P. et al. Single-cell RNA-seq reveals developmental plasticity with coexisting oncogenic and immune evasion programs in ETP-ALL. *Blood* **137**, 2463–2480 (2021).
16. De Bie, J. et al. Single-cell sequencing reveals the origin and the order of mutation acquisition in T-cell acute lymphoblastic leukemia. *Leukemia* **32**, 1358–1369 (2018).
17. Albertí-Servera, L. et al. Single-cell DNA amplicon sequencing reveals clonal heterogeneity and evolution in T-cell acute lymphoblastic leukemia. *Blood* **137**, 801–811 (2021).
18. Granja, J. M. et al. Single-cell multiomic analysis identifies regulatory programs in mixed-phenotype acute leukemia. *Nat. Biotechnol.* **37**, 1458–1465 (2019).
19. Park, J.-E. et al. A cell atlas of human thymic development defines T-cell repertoire formation. *Science* **367**eaay3224 (2020).
20. Bell, J. J. & Bhandoola, A. The earliest thymic progenitors for T cells possess myeloid lineage potential. *Nature* **452**, 764–767 (2008).
21. Zhang, J. et al. The genetic basis of early T-cell precursor acute lymphoblastic leukaemia. *Nature* **481**, 157–163 (2012).
22. Rothenberg, E. V., Moore, J. E. & Yui, M. A. Launching the T-cell-lineage developmental programme. *Nat. Rev. Immunol.* **8**, 9–21 (2008).
23. Rothenberg, E. V. Programming for T-lymphocyte fates: modularity and mechanisms. *Genes Dev.* **33**, 1117–1135 (2019).
24. Harly, C. et al. The transcription factor TCF-1 enforces commitment to the innate lymphoid cell lineage. *Nat. Immunol.* **20**, 1150–1160 (2019).
25. Ferrando, A. A. et al. Gene expression signatures define novel oncogenic pathways in T cell acute lymphoblastic leukemia. *Cancer Cell* **1**, 75–87 (2002).
26. Canté-Barrett, K. et al. MEF2C opposes Notch in lymphoid lineage decision and drives leukemia in the thymus. *JCI Insight* **7**, e150363 (2022).
27. Roels, J. et al. Distinct and temporary-restricted epigenetic mechanisms regulate human $\gamma\delta$ T cell development. *Nat. Immunol.* **21**, 1280–1292 (2020).
28. Wood, B. L. et al. Prognostic significance of ETP phenotype and minimal residual disease in T-ALL: a Children's Oncology Group study. *Blood* **142**, 2069–2078 (2023).
29. Weerkamp, F. et al. Human thymus contains multipotent progenitors with T/B lymphoid, myeloid, and erythroid lineage potential. *Blood* **107**, 3131–3137 (2006).
30. García-Peydró, M., de Yébenes, V. G. & Toribio, M. L. Notch1 and IL-7 receptor interplay maintains proliferation of human thymic progenitors while suppressing non-T cell fates. *J. Immunol.* **177**, 3711–3720 (2006).
31. Canté-Barrett, K. et al. Loss of CD44^{dim} expression from early progenitor cells marks T-cell lineage commitment in the human thymus. *Front. Immunol.* **8**, 32 (2017).
32. Hao, Q.-L. et al. Human intrathymic lineage commitment is marked by differential CD7 expression: identification of CD7⁺ lympho-myeloid thymic progenitors. *Blood* **111**, 1318–1326 (2008).
33. Vester, S. K. et al. Nucleolin acts as the receptor for C1QTNF4 and supports C1QTNF4-mediated innate immunity modulation. *J. Biol. Chem.* **296**, 100513 (2021).
34. Ruvolo, P. P. et al. LGALS1 acts as a pro-survival molecule in AML. *Biochim. Biophys. Acta Mol. Cell Res.* **1867**, 118785 (2020).
35. Huang, Y. et al. Identification and characterization of Hoxa9 binding sites in hematopoietic cells. *Blood* **119**, 388–398 (2012).
36. Wang, W. et al. MEF2C protects bone marrow B-lymphoid progenitors during stress haematopoiesis. *Nat. Commun.* **7**, 12376 (2016).
37. Luc, S. et al. Bcl11a deficiency leads to hematopoietic stem cell defects with an aging-like phenotype. *Cell Rep.* **16**, 3181–3194 (2016).
38. Rothenberg, E. V., Hosokawa, H. & Ungerbäck, J. Mechanisms of action of hematopoietic transcription factor PU.1 in initiation of T-cell development. *Front. Immunol.* **10**, 228 (2019).
39. Ungerbäck, J. et al. Pioneering, chromatin remodeling, and epigenetic constraint in early T-cell gene regulation by SPI1 (PU.1). *Genome Res.* **28**, 1508–1519 (2018).
40. Decker, M. D. et al. HES1 and HES4 have non-redundant roles downstream of Notch during early human T-cell development. *Haematologica* **106**, 130–141 (2021).
41. L, R., Bf, K. & L, H. The complete 685-kilobase DNA sequence of the human β T cell receptor locus. *Science* **272**, 1755–1762 (1996).
42. Lavaert, M. et al. Integrated scRNA-seq identifies human postnatal thymus seeding progenitors and regulatory dynamics of differentiating immature thymocytes. *Immunity* **52**, 1088–1104.e6 (2020).
43. Six, E. M. et al. A human postnatal lymphoid progenitor capable of circulating and seeding the thymus. *J. Exp. Med.* **204**, 3085–3093 (2007).
44. Bravo González-Blas, C. et al. SCENIC+: single-cell multiomic inference of enhancers and gene regulatory networks. *Nat. Methods* **20**, 1355–1367 (2023).
45. McCarter, A. C. et al. Combinatorial ETS1-dependent control of oncogenic NOTCH1 enhancers in T-cell leukemia. *Blood Cancer Discov.* **1**, 178–197 (2020).
46. Zhou, Y. et al. EBF1 nuclear repositioning instructs chromatin refolding to promote therapy resistance in T leukemic cells. *Mol. Cell* **82**, 1003–1020.e15 (2022).
47. Lee, S. H. R. et al. Pharmacotypes across the genomic landscape of pediatric acute lymphoblastic leukemia and impact on treatment response. *Nat. Med.* **29**, 170–179 (2023).
48. Liang, C. et al. The development of Bruton's tyrosine kinase (BTK) inhibitors from 2012 to 2017: a mini-review. *Eur. J. Med. Chem.* **151**, 315–326 (2018).
49. Tsunemine, H. & Takahashi, T. Early T-cell precursor acute lymphoblastic leukemia—a characteristic neoplasm presenting the phenotype of common hematopoietic progenitors for both myeloid and lymphoid lineages. In *Leukemias—Updates and New Insights* (eds Guenova, M. & Balatzenko, G.) (IntechOpen, 2015).

50. Berquam-Vrieze, K. E. et al. Cell of origin strongly influences genetic selection in a mouse model of T-ALL. *Blood* **118**, 4646–4656 (2011).
51. Kox, C. et al. The favorable effect of activating NOTCH1 receptor mutations on long-term outcome in T-ALL patients treated on the ALL-BFM 2000 protocol can be separated from FBXW7 loss of function. *Leukemia* **24**, 2005–2013 (2010).
52. Asnafi, V. et al. NOTCH1/FBXW7 mutation identifies a large subgroup with favorable outcome in adult T-cell acute lymphoblastic leukemia (T-ALL): a Group for Research on Adult Acute Lymphoblastic Leukemia (GRAALL) study. *Blood* **113**, 3918–3924 (2009).
53. Yu, W., Uzun, Y., Zhu, Q., Chen, C. & Tan, K. scATAC-pro: a comprehensive workbench for single-cell chromatin accessibility sequencing data. *Genome Biol.* **21**, 94 (2020).
54. Chen, X., Ji, Z. L. & Chen, Y. Z. TTD: Therapeutic Target Database. *Nucleic Acids Res.* **30**, 412–415 (2002).
55. Freshour, S. L. et al. Integration of the Drug–Gene Interaction Database (DGIdb 4.0) with open crowdsource efforts. *Nucleic Acids Res.* **49**, D1144–D1151 (2021).
56. Ochoa, D. et al. Open Targets Platform: supporting systematic drug–target identification and prioritisation. *Nucleic Acids Res.* **49**, D1302–D1310 (2021).
57. Subramanian, A. et al. A next generation connectivity map: L1000 platform and the first 1,000,000 profiles. *Cell* **171**, 1437–1452 (2017).
58. Tsherniak, A. et al. Defining a cancer dependency map. *Cell* **170**, 564–576 (2017).
59. Rowland, L. et al. Ex vivo drug sensitivity imaging-based platform for primary acute lymphoblastic leukemia cells. *Bio Protoc.* **13**, e4731 (2023).
60. Chen, C. et al. Single-cell multiomics reveals increased plasticity, resistant populations, and stem-cell-like blasts in KMT2A-rearranged leukemia. *Blood* **139**, 2198–2211 (2022).
61. Zhu, Q. et al. Developmental trajectory of prehematopoietic stem cell formation from endothelium. *Blood* **136**, 845–856 (2020).
62. Street, K. et al. Slingshot: cell lineage and pseudotime inference for single-cell transcriptomics. *BMC Genomics* **19**, 477 (2018).
63. Hao, Y. et al. Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573–3587 (2021).
64. Pliner, H. A. et al. Cicero predicts cis-regulatory DNA interactions from single-cell chromatin accessibility data. *Mol. Cell* **71**, 858–871 (2018).
65. Stuart, T., Srivastava, A., Madad, S., Lareau, C. A. & Satija, R. Single-cell chromatin state analysis with Signac. *Nat. Methods* **18**, 1333–1341 (2021).
66. Liberzon, A. et al. The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst.* **1**, 417–425 (2015).
67. Wilkens, A. B. et al. NOTCH1 signaling during CD4⁺ T-cell activation alters transcription factor networks and enhances antigen responsiveness. *Blood* **140**, 2261–2275 (2022).
68. Wang, H. et al. NOTCH1–RBPJ complexes drive target gene expression through dynamic interactions with superenhancers. *Proc. Natl Acad. Sci. USA* **111**, 705–710 (2014).
69. Korotkevich, G. et al. Fast gene set enrichment analysis. Preprint at bioRxiv <https://doi.org/10.1101/060012> (2024).
70. Tirosh, I. et al. Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. *Science* **352**, 189–196 (2016).
71. Lambert, S. A. et al. The human transcription factors. *Cell* **172**, 650–665 (2018).
72. Weirauch, M. T. et al. Determination and inference of eukaryotic transcription factor sequence specificity. *Cell* **158**, 1431–1443 (2014).

Acknowledgements

We acknowledge the contributions of the CHOP Research Information Services for providing computing support, the Flow Cytometry Core for flow cytometry equipment maintenance and the Animal Facility for providing mouse-related equipment. This AALLO434 clinical trial was supported by Novartis, Inc. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health (NIH). Funding for this study was obtained from Alex's Lemonade Stand Foundation (to D.T.T., K.T. and S.P.H.), Gabriella Miller Kids First X01HD100702 (to D.T.T., C.G.M., P.P., M.L.L., S.P.H., S.W., E.A.R., B.L.W., M.D., S.P.B., K.P.D. and J.J.Y.), the Leukemia and Lymphoma Society (to D.T.T.), Hyundai Hope of Wheels (to D.T.T., K.T. and R.S.), R03CA256550 (to D.T.T., C.G.M., P.P., M.L.L., S.P.H., S.W., E.A.R., B.L.W., M.D., S.P.B., K.P.D. and J.J.Y.), R01CA193776 (to D.T.T., B.W., K.T., C.G.M., S.P.H., J.J.Y., R.S. and M.D.), U10CA180886 (to D.T.T. and M.L.L.), R01CA264837 (to D.T.T., J.J.Y., C.G.M., K.T., B.W. and R.S.), U10CA18099 (to M.D.), U24CA114766 (to D.T.T. and M.L.L.), U24CA196173 (to D.T.T.), St. Baldrick's Foundation (to D.T.T. and M.L.L.), Pennsylvania Department of Health (to D.T.T. and K.T.), the Harrison Willing Memorial Research Fund (to D.T.T.), the Invisible Prince Foundation (to D.T.T.), the Aiden Everett Davies Innovation Fund (to D.T.T.), U2CCA233285 and U54HL165442 (to K.T.), NIH grant F30-CA-268782 (to J.X.) and the NIH Medical Scientist Training Program T32 GM07170 and NIH grant F30-CA-277965 (to S.B.).

Author contributions

J.X., C.C., J.H.S., S.Y., T.V., P.P., J.H., S.B., O.E., W.Y., J.T., C.-H.C., E.Y.L., C.D., R.S., H.N., L.U., A.L., G.M.C., D.W.W., Y.-Y.D., J.A.X., D.K., T.L., M.H., A.T., K.J.A., C.-Y.W., J.P., Y.S., A.W., R.M., D.F., L.M., B.W., J.J.Y., C.G.M., K.T. and D.T.T. performed experiments or analyzed the data. M.L.L., E.A.R., Z.C., M.D., K.P.D., S.S.W., T.-C.C., B.L.W., G.W., S.B.P., N.R.Z., W.C., S.P.H. and K.B. provided resources. J.X. and C.C. contributed equally. J.H.S., S.Y. and T.V. contributed equally. K.T. and D.T.T. jointly supervised this work. All authors helped design the experiments and contribute to data interpretation. J.X., C.C., K.T. and D.T.T. wrote the paper.

Competing interests

D.T.T. received research funding from BEAM Therapeutics and Neolimmune Tech and serves on advisory boards for BEAM Therapeutics, Janssen, Servier, Sobi and Jazz. D.T.T. has multiple patents pending on chimeric antigen receptor T cell therapy. The other authors declare no competing interests.

Additional information

Extended data is available for this paper at <https://doi.org/10.1038/s43018-024-00863-5>.

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s43018-024-00863-5>.

Correspondence and requests for materials should be addressed to Kai Tan or David T. Teachey.

Peer review information *Nature Cancer* thanks Simon Haas, Shai Izraeli and Seishi Ogawa for their contribution to the peer review of this work.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the

article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

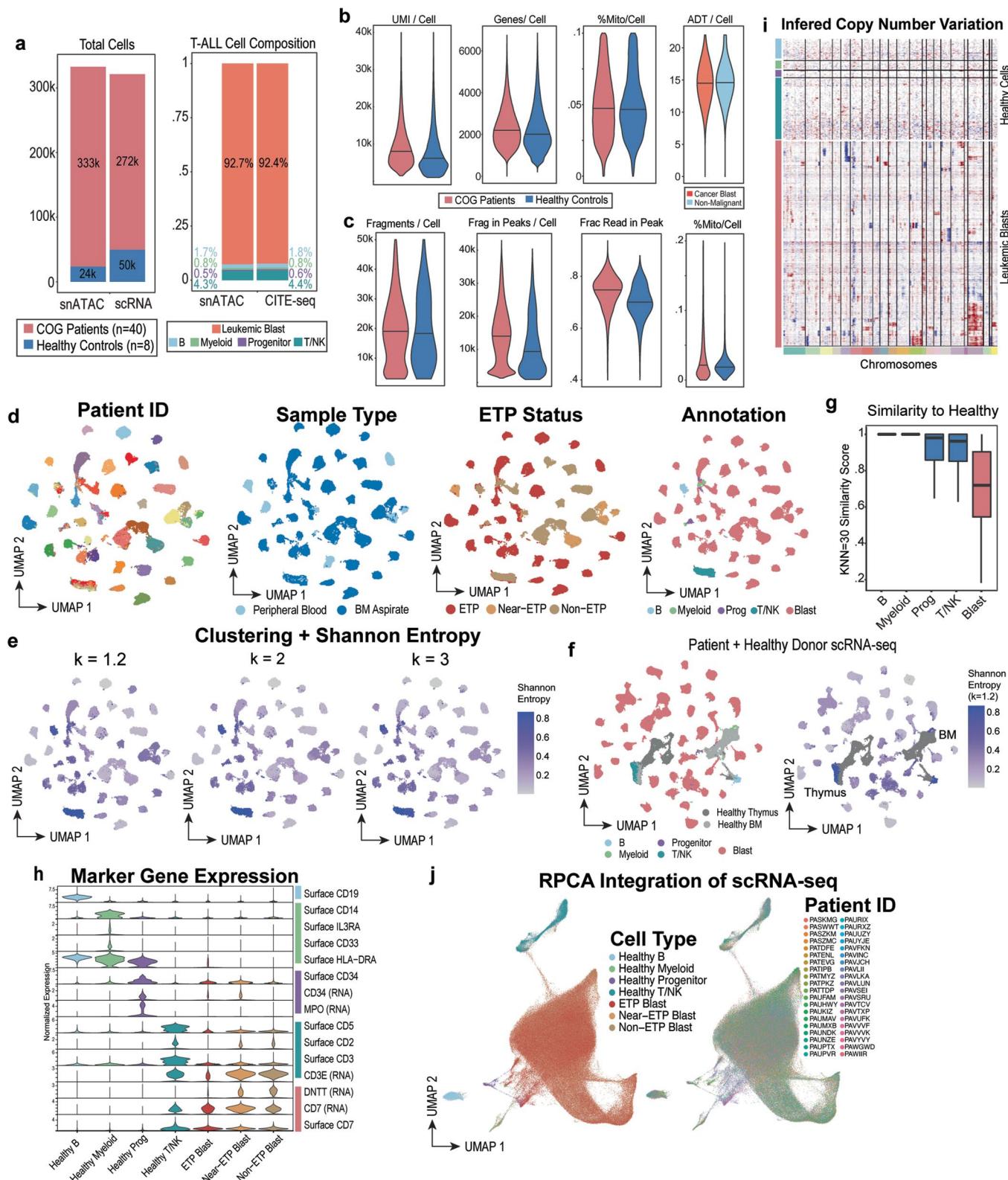
© The Author(s) 2024

Jason Xu^{1,2,23}, **Changya Chen**^{3,22,23}, **Jonathan H. Sussman**^{①,2}, **Satoshi Yoshimura**⁴, **Tiffaney Vincent**³, **Petri Pöllönen**^⑤, **Jianzhong Hu**⁴, **Shovik Bandyopadhyay**^{2,6}, **Omar Elghawy**³, **Wenbao Yu**³, **Joseph Tumulty**^③, **Chia-hui Chen**^③, **Elizabeth Y. Li**^{3,7}, **Caroline Diorio**³, **Rawan Shraim**³, **Haley Newman**³, **Lahari Uppuluri**^③, **Alexander Li**³, **Gregory M. Chen**^{1,2}, **David W. Wu**^{③,1,2}, **Yang-yang Ding**^③, **Jessica A. Xu**³, **Damjan Karanfilovski**³, **Tristan Lim**², **Miles Hsu**², **Anusha Thadi**^③, **Kyung Jin Ahn**³, **Chi-Yun Wu**¹, **Jacqueline Peng**¹, **Yusha Sun**^{③,2,7}, **Alice Wang**^{③,1}, **Rushabh Mehta**⁷, **David Frank**⁸, **Lauren Meyer**^{9,10}, **Mignon L. Loh**^{③,9,10}, **Elizabeth A. Raetz**¹¹, **Zhiguo Chen**¹², **Brent L. Wood**¹³, **Meenakshi Devidas**^{③,14}, **Kimberly P. Dunsmore**¹⁵, **Stuart S. Winter**^{③,16}, **Ti-Cheng Chang**^{③,17}, **Gang Wu**^{③,17}, **Stanley B. Pounds**^{③,18}, **Nancy R. Zhang**^{③,19}, **William Carroll**¹¹, **Stephen P. Hunger**^{3,20}, **Kathrin Bernt**^{③,20}, **Jun J. Yang**^{③,4}, **Charles G. Mullighan**^{③,5}, **Kai Tan**^{③,20,21,24}✉ & **David T. Teachey**^{③,20,24}✉

¹Graduate Group in Genomics and Computational Biology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA. ²Medical Scientist Training Program, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA. ³Division of Oncology and Center for Childhood Cancer Research, Children's Hospital of Philadelphia, Philadelphia, PA, USA. ⁴Department of Pharmacy and Pharmaceutical Sciences, St. Jude Children's Research Hospital, Memphis, TN, USA. ⁵Department of Pathology, St. Jude Children's Research Hospital, Memphis, TN, USA. ⁶Graduate Group in Cell & Molecular Biology, University of Pennsylvania, Philadelphia, PA, USA. ⁷Yale School of Medicine, New Haven, CT, USA. ⁸Division of Cardiology, Children's Hospital of Philadelphia, Philadelphia, PA, USA. ⁹The Ben Town Center for Childhood Cancer Research, Seattle Children's Hospital, Seattle, WA, USA. ¹⁰Department of Pediatric Hematology Oncology, Seattle Children's Hospital, Seattle, WA, USA. ¹¹Department of Pediatrics and Perlmutter Cancer Center, New York University Langone Health, New York, NY, USA. ¹²Department of Biostatistics, University of Florida, Gainesville, FL, USA. ¹³Department of Pathology and Laboratory Medicine, Children's Hospital Los Angeles, Los Angeles, CA, USA. ¹⁴Department of Global Pediatric Medicine, St Jude Children's Research Hospital, Memphis, TN, USA. ¹⁵Division of Oncology, University of Virginia Children's Hospital, Charlottesville, VA, USA. ¹⁶Children's Minnesota Hospital, Minneapolis, MN, USA. ¹⁷Center for Applied Bioinformatics, St. Jude Children's Research Hospital, Memphis, TN, USA. ¹⁸Department of Biostatistics, St. Jude Children's Research Hospital, Memphis, TN, USA. ¹⁹Department of Statistics, University of Pennsylvania, Philadelphia, PA, USA. ²⁰Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA. ²¹Center for Single Cell Biology, Children's Hospital of Philadelphia, Philadelphia, PA, USA. ²²Present address: State Key Laboratory of Experimental Hematology, National Clinical Research Center for Blood Diseases, Institute of Hematology & Blood Diseases Hospital, Chinese Academy of Medical Sciences & Peking Union Medical College, Tianjin, China.

²³These authors contributed equally: Jason Xu, Changya Chen. ²⁴These authors jointly supervised this work: Kai Tan, David T. Teachey.

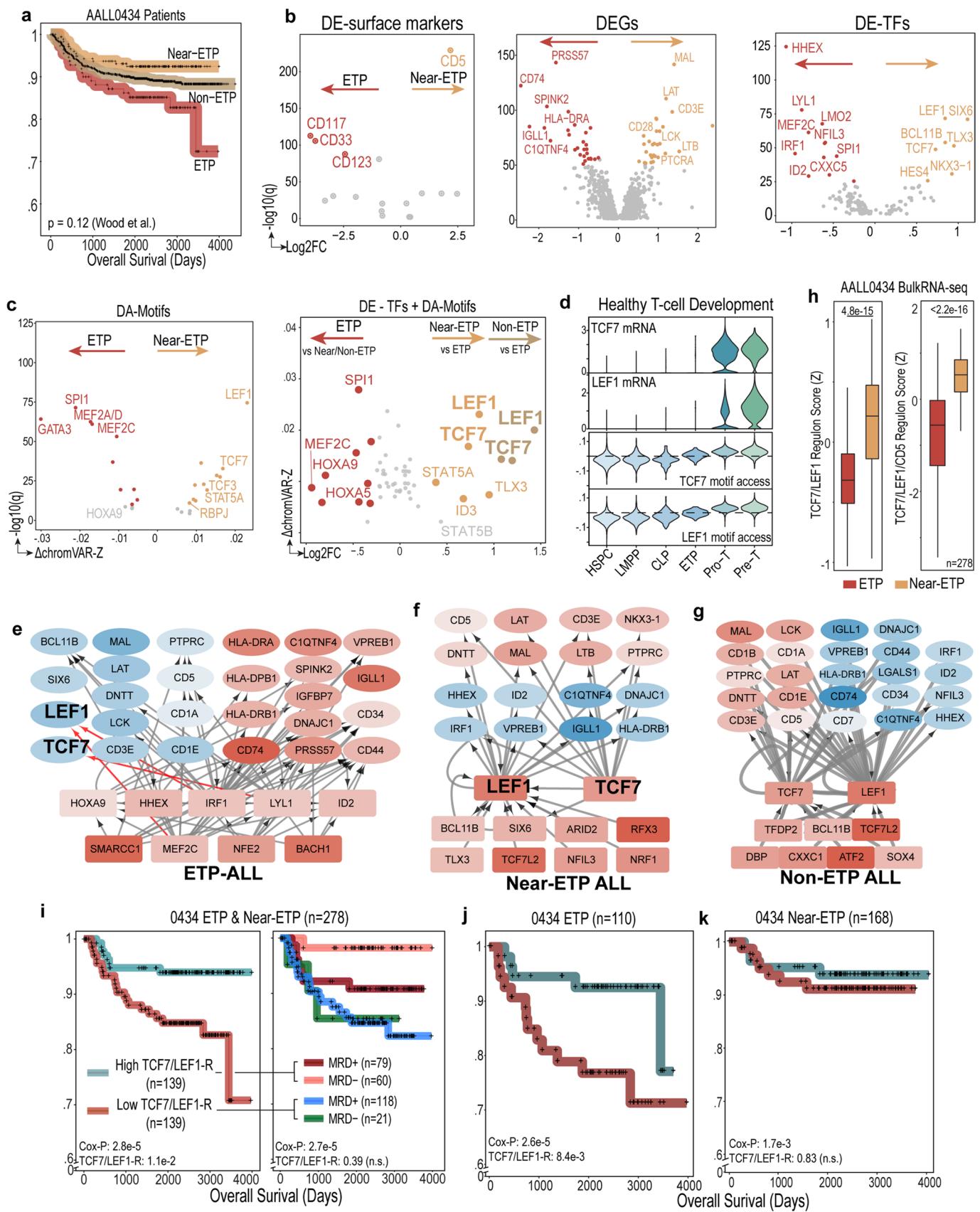
✉e-mail: tank1@chop.edu; teacheyd@chop.edu



Extended Data Fig. 1 | See next page for caption.

Extended Data Fig. 1 | Single cell multiomics to identify high-risk cell states in T-ALL. (a) size (left) and cell type composition for scRNA-seq and scATAC-seq dataset from n = 40 T-ALL cases and n = 8 healthy thymus/BM controls. COG: Children's Oncology Group. (b) Quality of scRNA-seq dataset after filtering (n = 328,820 cells; COG patients: n = 271,603 cells; Healthy Control: n = 49,623 cells). (c) Quality of scATAC-seq dataset after filtering (COG patients: n = 333,490 cells; Healthy Control: n = 23,618 cells). (d) UMAP representation of scRNA-seq dataset (n = 328,820 cells) colored by patient ID, sample type, ETP status, and cell type annotation. (e) Shannon Entropy (1 = equal contribution from each sample; 0 = contribution from only 1 sample) of cell clusters at k = 1.2, k = 2, and k = 3 clustering resolutions. (f) Clustering of T-ALL patient-derived data with n = 8

healthy bone marrow/thymus controls (n = 328,820 cells). Left: colored by cell type annotation; right: colored by Shannon entropy. (g) k = 30 nearest neighbor similarity score to known healthy controls. Patient derived cells (n = 271,603) were mapped to known healthy controls using the RPCA method in Seurat 4.0. The average similarity score to 30 nearest healthy control neighbors in principal component space is shown for each cell. (h) Marker gene expression of annotated cell types. (i) inferCNV results for annotated patient derived cells within scRNA-seq data. Cells are ordered with hierarchical clustering within each category. (j) RPCA-integrated UMAP of scRNA-seq dataset (n = 271,603 cells) colored by annotation (left) and patient ID (right).

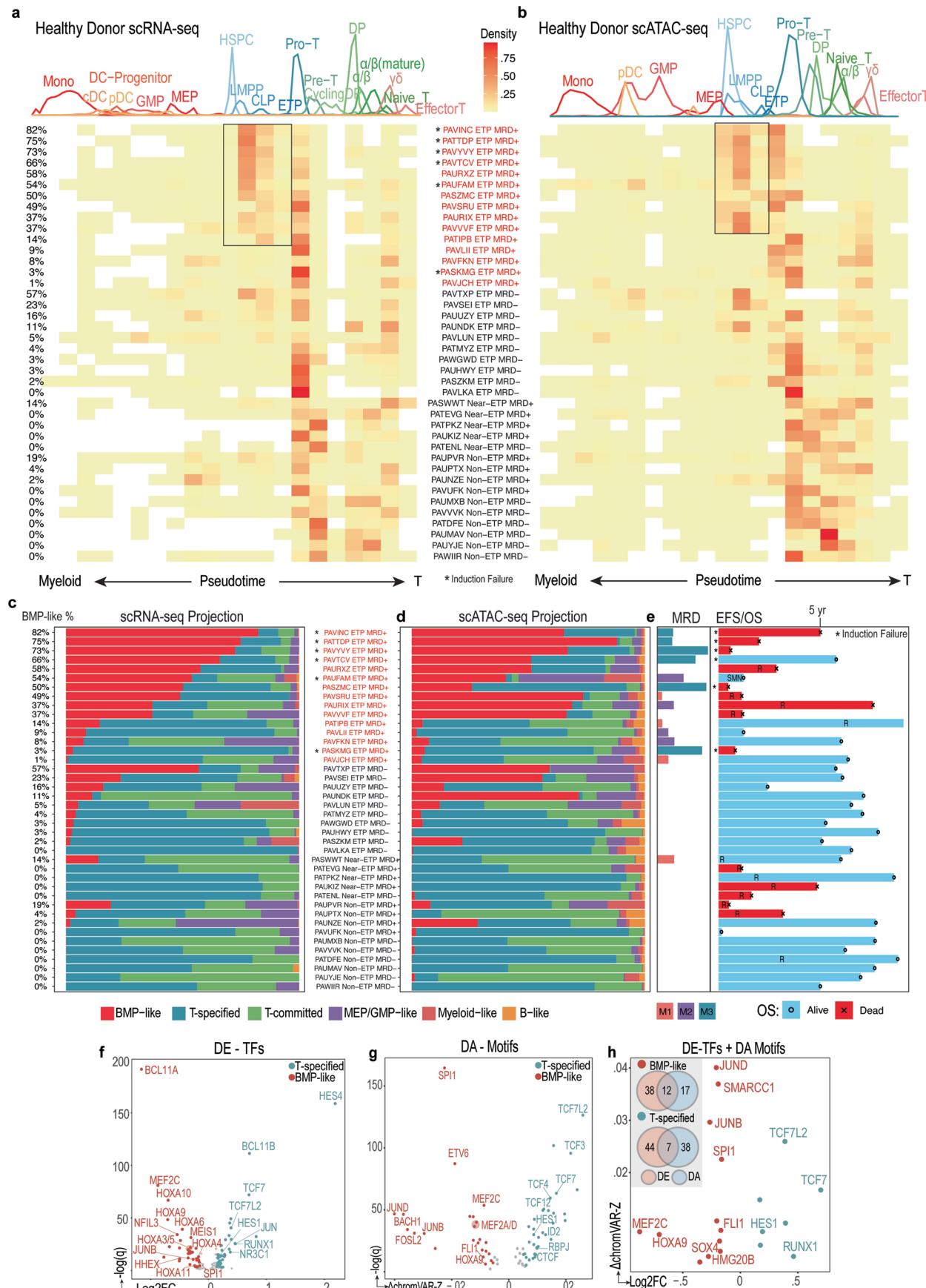


Extended Data Fig. 2 | See next page for caption.

Extended Data Fig. 2 | TCF7/LEF1 activation underlies CD5 expression in**Near-ETP T-ALL and contributes to positive outcome in ETP-ALL Patients.**

(a) Overall survival of ETP, Near-ETP and Non-ETP patients from Children's Oncology Group AALL0434 cohort ($n = 1411$). (b) Differentially expressed surface markers, genes, and transcription factors in single-cell sequenced ETP ($n = 25$) and Near-ETP ($n = 5$) patients. (c) Intersection of differentially expressed transcription factors and differentially accessible motifs in single-cell sequenced ETP ($n = 25$) and Near-ETP ($n = 5$) patients. (d) Expression of *TCF7* and *LEF1* and accessibility of *TCF7* and *LEF1* motifs in healthy T-cell development references. $n = 1200$ cells; $n = 100$ randomly downsampled cells per group. (e-g) Subset of transcriptional regulatory network constructed using integrated scRNA and scATAC data from (e) ETP-ALL ($n = 25$), (f) Near-ETP ALL ($n = 5$), and (g) Non-ETP ALL ($n = 10$) patients. Transcription factors are represented as squares, gene targets as ovals. In (e), color is proportional to expression fold change in comparison to Near-ETP and Non-ETP blasts. In (f-g), color is proportional to expression fold change in

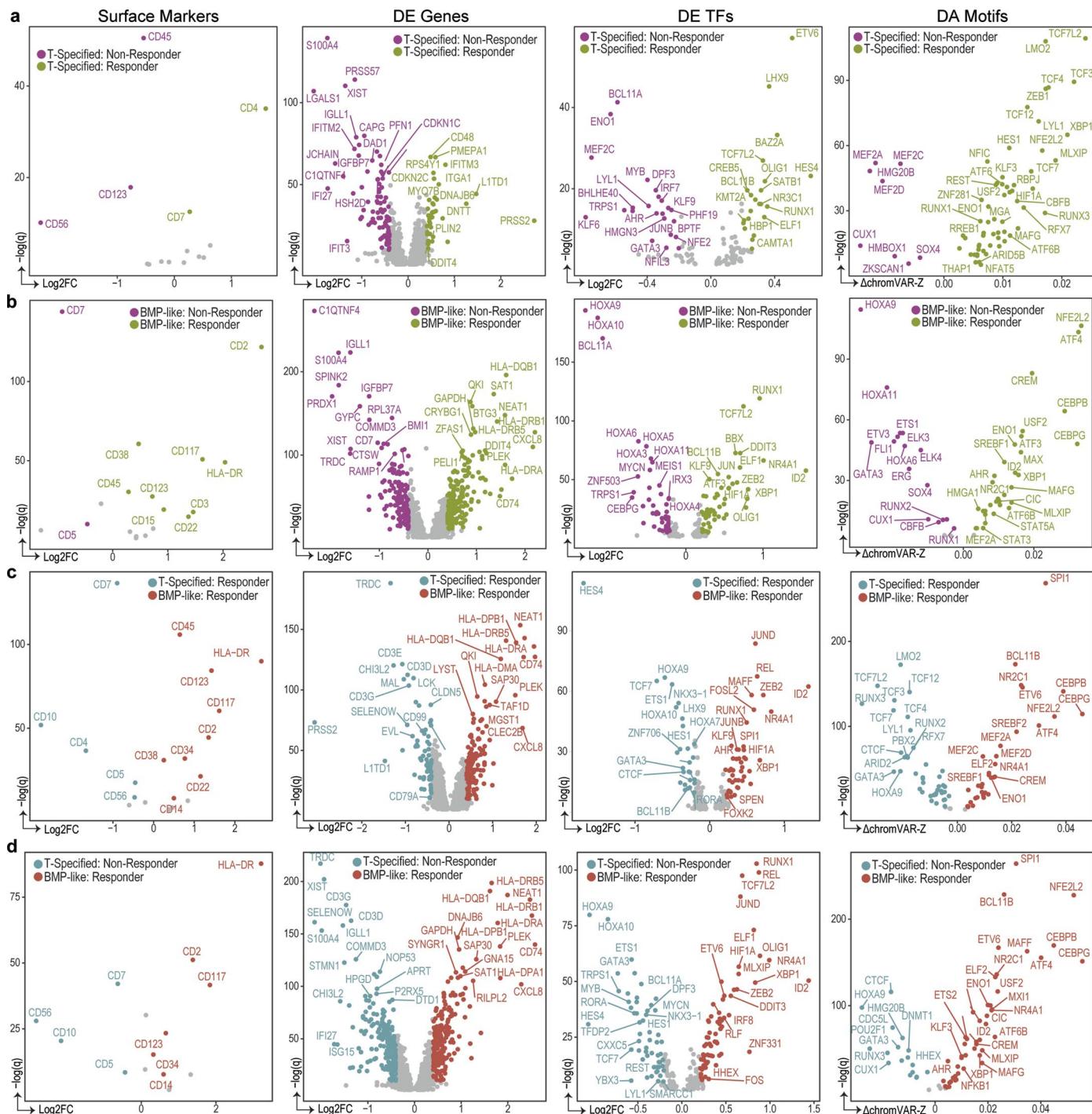
comparison to ETP blasts blue is downregulated, red upregulated. In (e), edges with >100 edge score are shown, with edge score representing the sum of $-\log(p\text{-value})$ of all predicted EP interactions. In (f-g), Edges contacting *TCF7* and *LEF1* with regression coefficient > 0.3 are shown. Predicted regulators of *TCF7* and *LEF1* are highlighted. (h) Signature score of top 28 target genes of the TCF7/LEF1 regulon and top 11 predicted TF regulators of TCF7/LEF1/CD5 in bulk-sequenced ETP ($n = 110$) and Near-ETP ($n = 168$) T-ALL patients from COG AALL0434 trial. The box includes the median, hinges mark the 25th and 75th percentiles, and whiskers extend 1.5 times the interquartile range. (i-k) Kaplan-Meier plot showing overall survival of bulk-RNA-sequenced ETP-ALL ($n = 110$) and Near-ETP ($n = 168$) patients in AALL0434 binarized using the TCF7-LEF1 targets and activator signature. Prognostic value of the TCF7/LEF1 signature in multivariate analysis (with Day 29 MRD, CNS status, WBC count, and age at diagnosis) is shown below the Cox-proportional hazard log-likelihood p-value with Day 29 MRD as covariate.



Extended Data Fig. 3 | See next page for caption.

Extended Data Fig. 3 | Sample-specific developmental arrest state of > 500,000 T-ALL blasts from CITE-seq/scATAC-seq data and clinical response correlates. (a-b) Arrest state of T-ALL blasts over T and myeloid development based on projection to a healthy reference using (a) scRNA-seq data and (b) scATAC-seq data. BMP-like proportion is shown on the left. ETP-ALL patients with D29 residual disease are highlighted in red and those with >10% BMP-like are boxed; patients with induction failure (D29 M3 bone marrow morphology) are marked with an asterisk. n = 40 patients: 25 ETP-ALL, 5 Near-ETP, 10 Non-ETP. (c-d) Fraction of key cell states in (c) scRNA-seq data and (d) scATAC-seq data of T-ALL blasts from 40 AALL0434 patients. Rows are in the same order as panel a. BMP-like (HSPC/LMPP/CLP/ETP projected), T-specified (Pro-T/Pre-T), T-committed (DP to Naïve T). n = 40 patients: 25 ETP-ALL, 5 Near-ETP, 10 Non-ETP. (e) Left: MRD (range, 0-100) and day 29 marrow status. Right: Overall survival (OS)/event free survival (EFS) swimmer plot of 40 AALL0434 T-ALL patient cohort; 5 year timepoint is marked at the top. Events

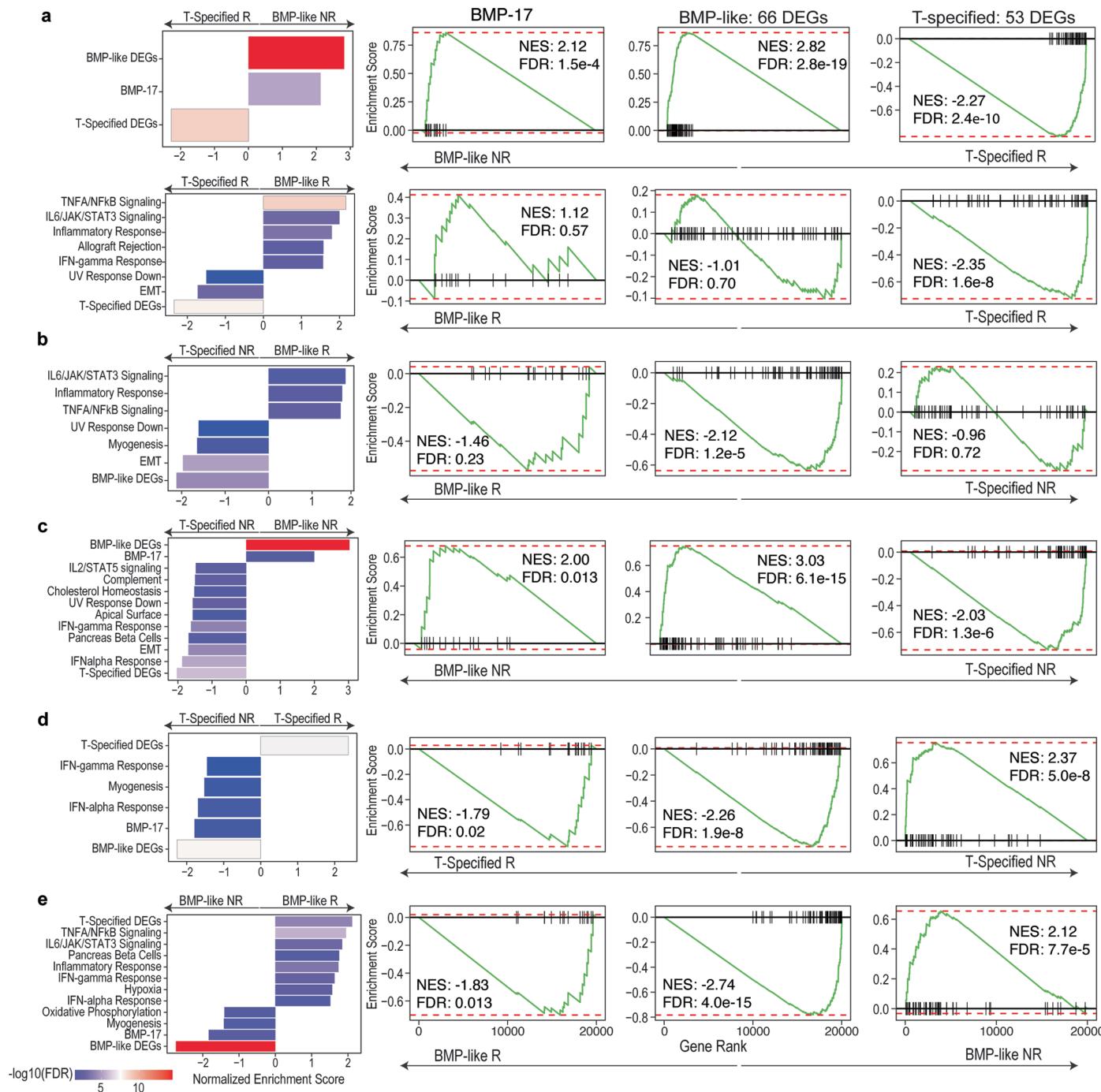
are labeled: induction failure is indicated by an asterisk; relapse is indicated by R; second malignant neoplasm is indicated by "SMN". n = 40 patients: 25 ETP-ALL, 5 Near-ETP, 10 Non-ETP. (f-h) Differentially expressed transcription factors (f), differentially accessible transcription factor motifs, and intersection of differentially expressed transcription factors and differentially accessible motifs between BMP-like blasts from non-responding patients (n = 15) and T-specified blasts from responding patients (n = 10). Differentially expressed transcription factors were defined by Log2FC > 0.15, adjusted p-value < 0.01; differentially accessible motifs were defined by Δ median chromVAR deviation > 0.005, adjusted p-value < 0.01. The input matrix to differential expression was a matrix of G1-phase ETP-ALL blasts with equal number of cells per patient (n = 1,711 per patient and 42,775 cells total). Differential expression was performed using n = 1500 randomly downsampled cells per cluster. Differential accessibility was performed using n = 1500 randomly downsampled cells per cluster.



Extended Data Fig. 4 | Transcriptional and epigenetic characteristics of BMP-like and T-specified cell states between responders and non-responders.

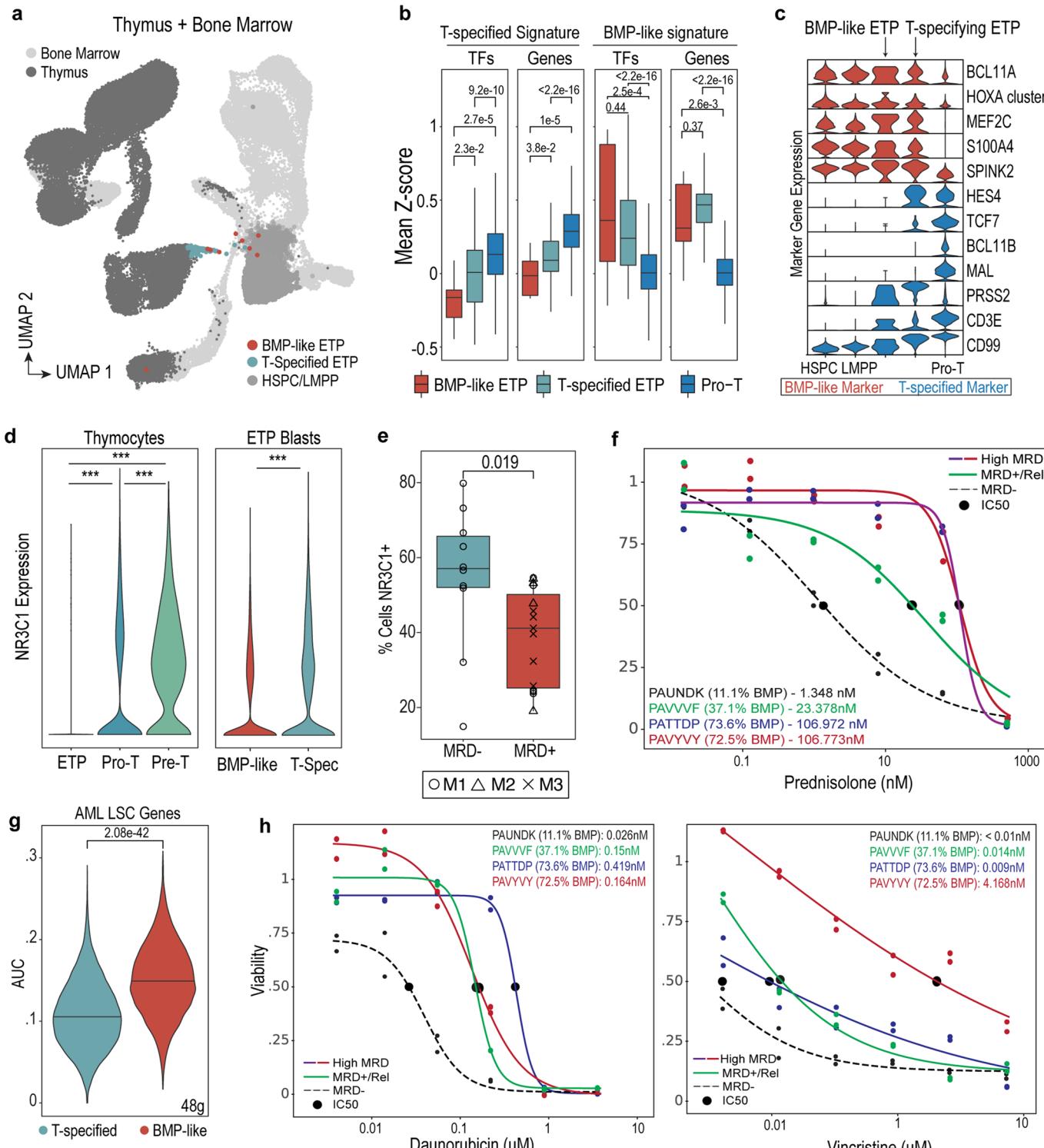
(a-d) Differentially expressed surface markers, genes, and transcription factors as well as differentially accessible transcription factor motifs in (a) T-specified blasts between responding ($n = 10$) and non-responding ($n = 15$) patients

(b) BMP-like blasts between responding and non-responding patients,
 (c) BMP-like and T-specified blasts from responding patients, (d) BMP-like blasts
 from responding patients and T-specified blasts from non-responding patients.
 Differential expression was performed using $n = 1500$ randomly downsampled
 cells per cluster.



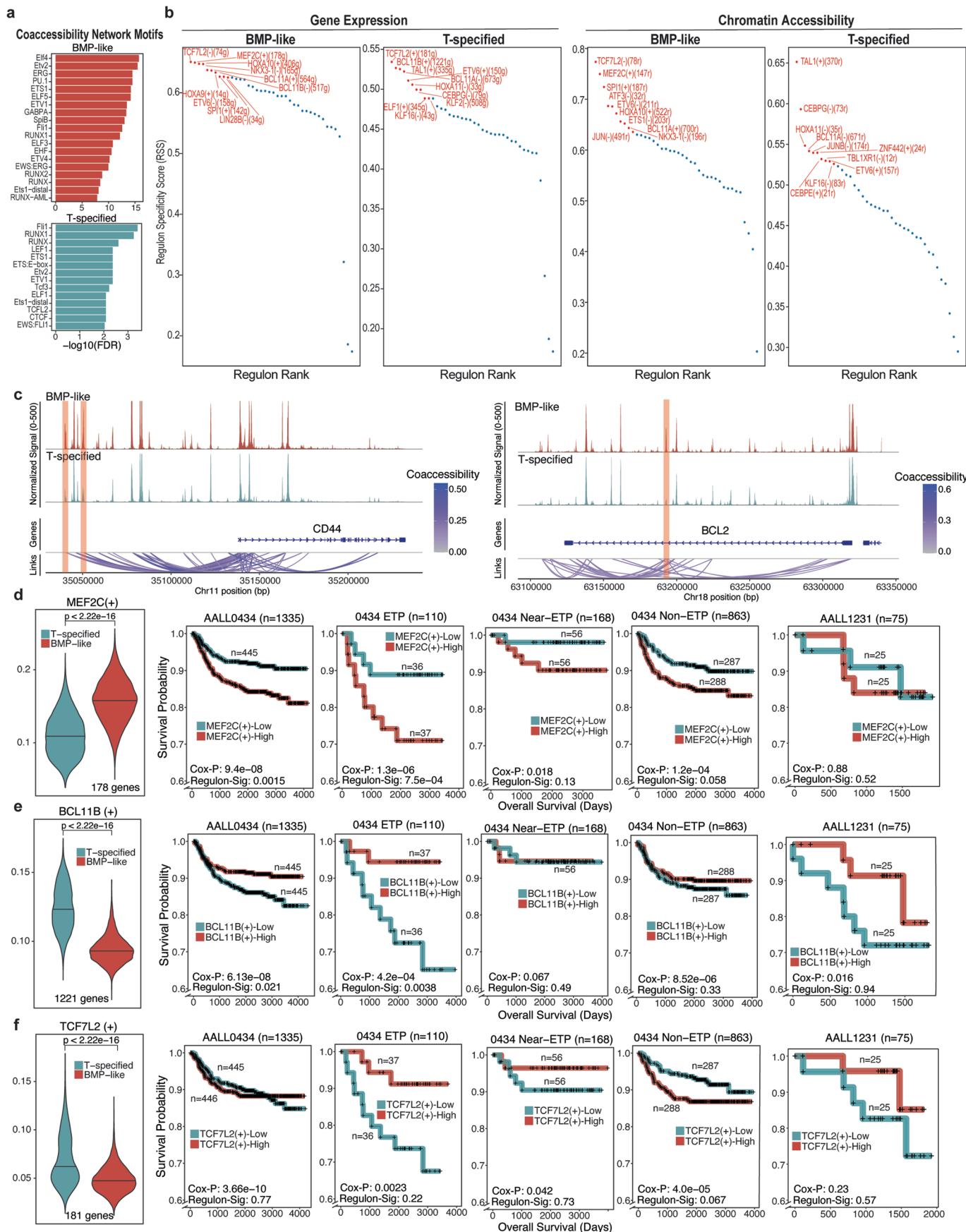
Extended Data Fig. 5 | Pathway analysis of BMP-like and T-specified blasts from responding and non-responding patients. (a-e) Gene Set Enrichment Analysis (GSEA) was performed based on differential gene expression of BMP-like and T-specified blasts from responding ($n = 10$) and non-responding ($n = 15$) patients. Left, pathway enrichment is shown for the Hallmark gene sets combined

with the BMP-17 genes and differentially expressed genes (DEGs) from BMP-like and T-specified blasts (as shown in Fig. 2). Right, Enrichment plots are shown for the BMP-17 genes, BMP-like DEGs, and T-specified DEGs. FDR, false discovery rate; NES, normalized enrichment score.



Extended Data Fig. 6 | Non-Malignant counterparts to BMP-like blasts and associated resistance to frontline ALL therapy. (a) Healthy reference trajectory with BMP-like ETP and T-specified ETP highlighted. Thymus, bone marrow, and subset of BM progenitors (HSPC/LMPP) are colored in different shades of gray ($n = 49,623$ cells). (b) Average Z-score of BMP-like and T-specified-like gene signatures derived from non-responding and responding ETP-ALL patients were computed for BMP-like ETP, T-specifying ETP, and Pro-T cells. $N = 13$ BMP-like ETP, 101 T-specified ETP, 5,141 Pro-T. (c) Expression of BMP-like and T-specified TFs and marker genes across multipotent BM progenitors, BMP-like ETP, T-specifying ETP, and Pro-T cells. HOXA cluster denotes sum of expression across HOXA cluster genes and MEIS1, the HOX co-factor. (d) Expression of NR3C1 during thymic entry, T-specification, and T-commitment. P-value was calculated based on a two-sided

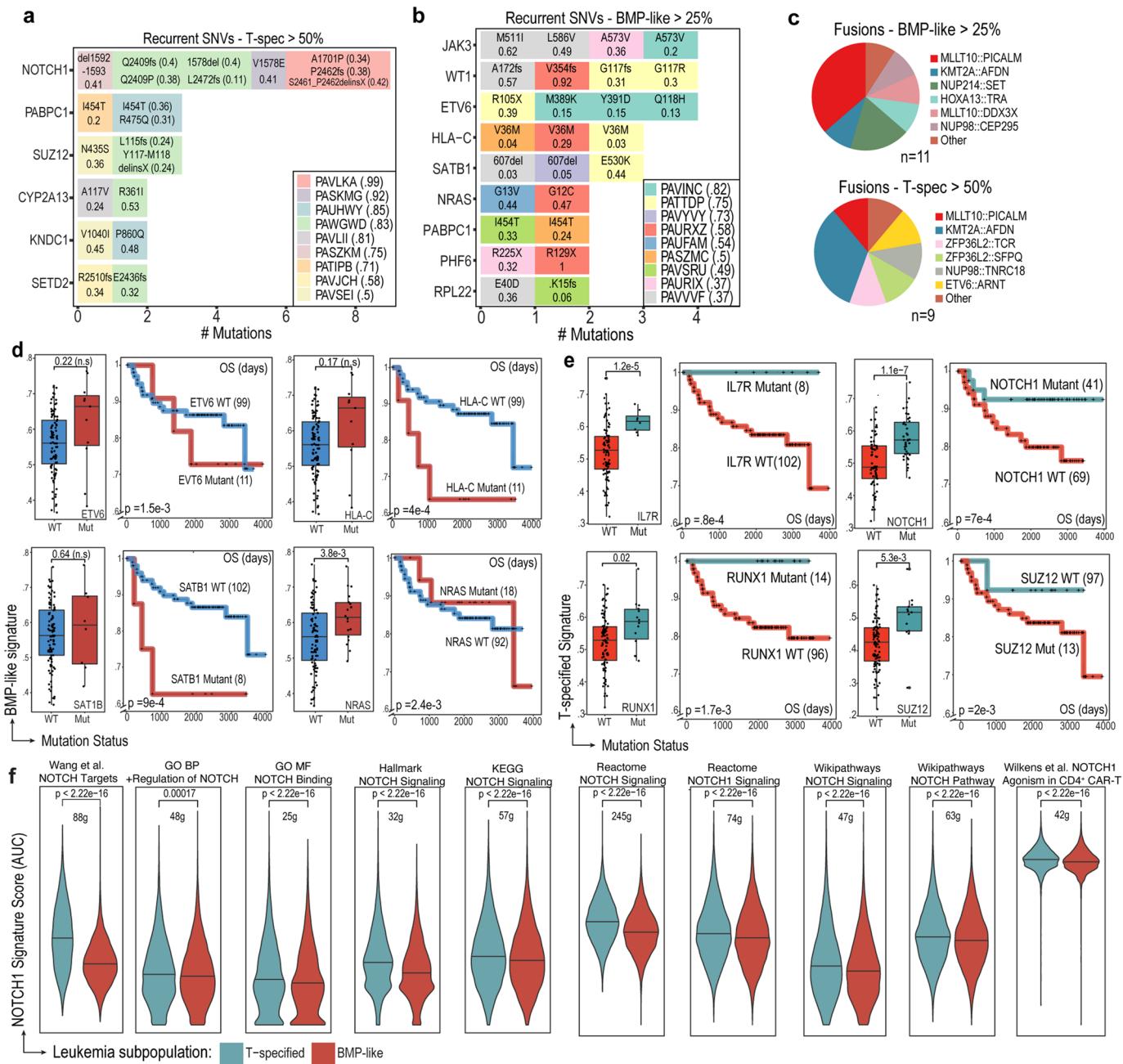
Mann Whitney test on log normalized data. *** $p < 0.001$. (e) Expression of NR3C1 in BMP-like ETP blasts from non-responding patients and T-specified ETP blasts from responding patients. P-value was calculated based on a two-sided Mann Whitney test on log normalized data. (b, e) The box includes the median, hinges mark the 25th and 75th percentiles, and whiskers extend 1.5 times the interquartile range. $n = 15$ MRD+, 10 MRD-. (f) Response of $n = 3$ High BMP and $n = 1$ MRD Negative ETP patient to prednisolone. High MRD patients both had > 30% BMP-like blasts; MRD Negative patient and Non-ETP Patients had > 50% T-specified blast. (g) AUCell signature score for $n = 48$ LSC-related genes (Ng et al.)³⁷ in T-specified and BMP-like ETP blasts. (h) Response of $n = 3$ High BMP-like and $n = 1$ MRD Negative ETP patient to daunorubicin and vincristine. High MRD patients all had > 30% BMP-like blasts; MRD Negative patient had > 50% T-specified blasts.



Extended Data Fig. 7 | See next page for caption.

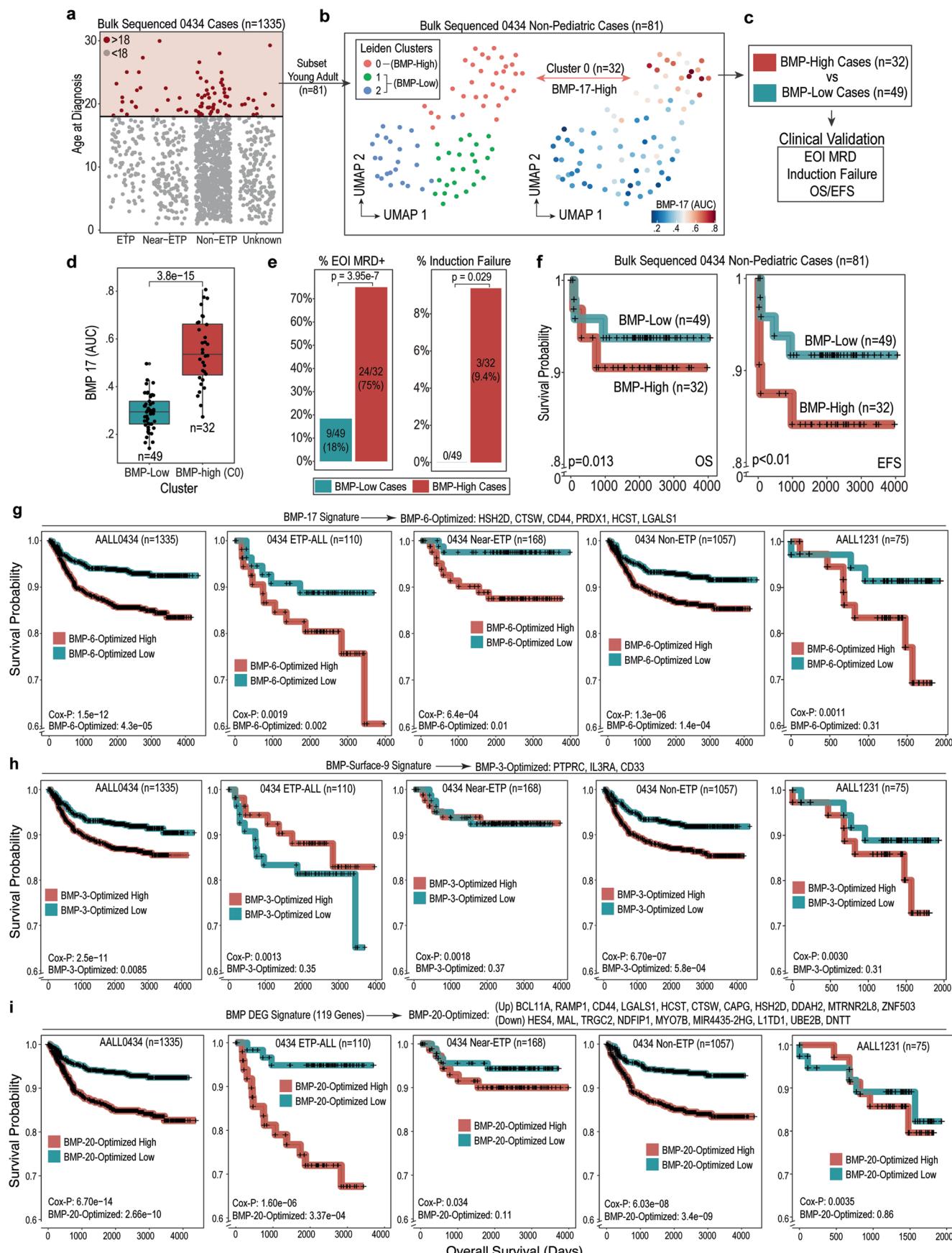
Extended Data Fig. 7 | Transcriptional regulation of BMP-like and T-specified blast states. (a) Top significantly enriched motifs in the BMP-like and T-specified co-accessibility networks involving the promoter regions of the $n = 119$ differentially-expressed genes shown in Fig. 3j. (b) Regulon specificity scores based on gene expression (left) and chromatin accessibility (right). The top 10 regulons are listed, and the number of genes or regions contained in the regulon is indicated in parentheses. (c) Representative scATAC-seq signal tracks highlighting regulatory regions that are differentially accessible between BMP-like and T-specified populations. The number of region coaccessibility links were reduced and the range of the normalized signal track was truncated to 500

for visualization purposes. Links are colored by Cicero coaccessibility score. (d-f) Regulon gene signatures for (d) MEF2C(+), (e) BCL11B(+), and (f) TCF7L2(+) were scored between T-specified and BMP-like populations using AUCell (left), and Kaplan-Meir plots showing overall survival of bulk-RNA-sequenced T-ALL patients in AALL0434 and AALL1232 stratified by upper and lower third using the regulon signatures (right). The regression coefficient of the regulon signatures in multivariate analysis (with Day 29 MRD, CNS status, WBC, and age at diagnosis) is shown below the Cox proportional-hazards log-likelihood p-value controlling for D29 MRD. Non-subtyped samples were not included in this analysis.



Extended Data Fig. 8 | Integration of bulk-derived mutation calls with single-cell-derived tumor phenotype. (a) Recurrently mutated genes seen in low-risk T-specified >50% patients. (b) Recurrently mutated genes seen in high-risk BMP-like > 25% patients. (c) Driver fusion profile of high BMP-like patients (high risk, left) and high T-specified patients (low risk, right). (d-e) Prognostic value of individual mutated genes associated with BMP-like (d) and T-specified states

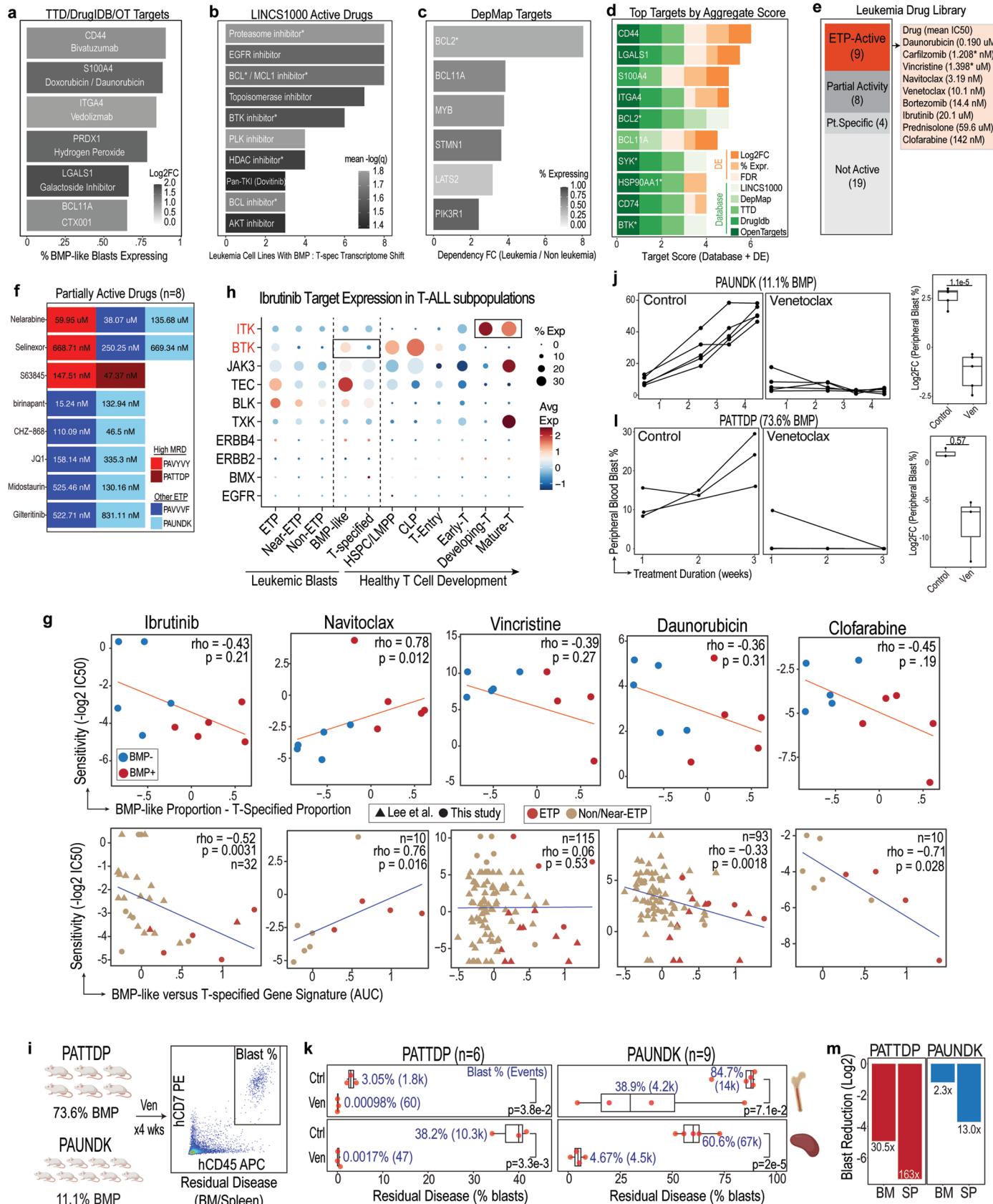
(e). The p-value for two-sided t-test is shown above each boxplot; the p-value for Log-likelihood statistic of Cox-proportional hazard test run with Day 29 MRD as a healthy donor curves. (f) NOTCH1 signature scores for T-specified and BMP-like blasts. Pathways were obtained from the Molecular Signatures Database or as published by Wang et al. and Wilkins et al. Significance was assessed using a two-sided Wilcoxon rank-sum test.



Extended Data Fig. 9 | See next page for caption.

Extended Data Fig. 9 | Clinical utility of the BMP-like-17 in risk stratifying non-pediatric and relapsed T-ALL patients. (a) Selection of n = 81 young adult (age at diagnosis >=18 years) T-ALL cases from bulk-sequenced AALL0434 cohort (n = 1335). (b) Identification of BMP-17-high cluster within AALL0434 young adult cases via Leiden clustering. n = 81 patients; n = 32 BMP-high, n = 49 BMP-low. (c) Clinical outcome comparison strategy between BMP-17-High and BMP-17-Low cases. n = 81 patients; n = 32 BMP-high, n = 49 BMP-low. (d) BMP-17 signature score in cluster 0 (BMP-high) vs clusters 1-2 (BMP-low) cases. n = 81 patients; n = 32 BMP-high, n = 49 BMP-low. (e) enrichment of EOI MRD positive and induction failure cases within BMP-High cases. n = 81 patients; n = 32 BMP-high, n = 49 BMP-low. (f) Overall (left) and event free (right) survival outcomes in BMP-High and BMP-Low cases. Cox-proportional hazard log-likelihood p-value is shown in the bottom left. n = 81 patients; n = 32 BMP-high, n = 49 BMP-low.

(g-i) LASSO regression model was used to narrow the (g) BMP-17 signature, (h) BMP-surface-9 signature, and (i) BMP-119 DEG signature to the genes that were most predictive of overall survival when stratified by ETP subtype. Kaplan-Meier plots show overall survival of bulk-RNA-sequenced T-ALL patients in AALL0434 (n = 1335) and AALL1231 (n = 75) binarized using the optimized signatures. The BMP-Optimized-6 and BMP-Optimized-3 signatures were calculated using AUCell and binarized at the 50th percentile, as in Fig. 6g,h to be directly comparable. The BMP-Optimized-20 signature was computed using a z-score to positively or negatively weight genes and binarized at a z-score=0, thus directly comparable to Fig. 3m. The prognostic value of the signatures in multivariate analysis (with Day 29 MRD, CNS status, age at diagnosis, and WBC count) are shown below the Cox-proportional hazard log-likelihood p-value with Day 29 MRD as the covariate.



Extended Data Fig. 10 | See next page for caption.

Extended Data Fig. 10 | Computational and *in vitro* drug screening results against BMP-like blasts. (a) Top predicted drugs from LINCS1000 ($n = 10$). BMP-like DEGs (High Risk) and T-specified DEGs (Low Risk) were inputted into LINCS1000. Drug treated leukemia cell lines were filtered for statistical significance ($FDR < 0.1$) and connectivity score ($NCS > 0.8$). Drugs are ranked by number of leukemia cell lines with favorable transcriptomic shift after treatment (downregulation of BMP-like DEG, upregulation of T-specified DEG). Each drug is colored by the mean -log(FDR). (b) Top leukemia specific targets ($n = 6$) predicted from DepMap screening. Dependency scores in leukemic ($n = 59$) and non-leukemic cell lines ($n = 1,052$) were calculated for all BMP-like DEGs and ranked by fold change in dependency (mean dependency in leukemia / mean dependency in non-leukemia cell lines). The top druggable (with score 1+ from other drug databases) targets are shown. (c) Top druggable targets ($n = 6$) from TTD/DrugIDB drug database screening. Targets are ranked by percentage expression and selected based on $\text{Log2FC} > 1$. An example of drug is listed below the target. (d) Top 10 targets by aggregate database (1-5) and DE (1-3) score. (e) Drugs active in $n = 4/4$ ETP patients tested with mean IC₅₀. Drugs with marked asterisk had IC₅₀ below lowest tested dose in 1 sample. $n = 40$ drugs, $n = 9$ ETP-active. (f) Drugs active in some, but not all ETP patients. High MRD patients are colored in red. $n = 40$ drugs, $n = 8$ partially active. (g) Correlations between drug sensitivity (-log₂ of the IC₅₀ concentration) and the scRNA-seq derived BMP-like percentage (top) and the BMP-like signature score computed using $n = 119$ differentially expressed genes on bulk RNA-sequenced data (bottom).

The bulk RNA-seq correlations (bottom) include the data from this study ($n = 10$) and data by Lee et al. Total number of data points for each drug is indicated in the figure. Spearman's correlations and significance are shown. (h) Gene expression of ibrutinib targets across ETP subtypes, BMP-like/T-specified blast phenotypes, and stages of healthy T cell development. Dot size indicates percent of cells with gene expression detected, and color indicates normalized average expression. ($n = 328,820$ cells; T-ALL patients: $n = 271,603$ cells; Healthy Control: $n = 49,623$ cells). (i) Representative flow gating for quantification of hCD7 + hCD45+ leukemic blasts during venetoclax or control treatment. (j) peripheral blast percentage (left) and log₂ fold change (right) of peripheral blast % over study period for PAUNDK (BMP-low, $n = 8$: $n = 4$ control, $n = 4$ venetoclax) PDX model during control or venetoclax treatment. P-value from two-sided t-test is shown. (k) Bone Marrow (BM, top) and spleen (bottom) leukemic burden in High-BMP (left, $n = 6$: $n = 3$ control, $n = 3$ venetoclax) and low-BMP (right, $n = 8$: $n = 4$ control, $n = 4$ venetoclax) PDX models after 1 month of venetoclax or vehicle (ctrl) treatment. P-value from two-sided t-test is shown. The box includes the median, hinges mark the 25th and 75th percentiles, and whiskers extend 1.5 times the interquartile range. (l) Peripheral blast percentage (left) and log₂ fold change (right) of peripheral blast % over study period for PATTDP (BMP-high, $n = 6$: $n = 3$ control, $n = 3$ venetoclax) PDX model during control or venetoclax treatment. P-value from two-sided t-test is shown. (m) Fold-reduction of leukemic burden in BM and spleen with venetoclax treatment in BMP-high ($n = 6$) and BMP-low ($n = 8$) PDX models.

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
- A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided
Only common tests should be described solely by name; describe more complex techniques in the Methods section.
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g. F , t , r) with confidence intervals, effect sizes, degrees of freedom and P value noted
Give P values as exact values whenever suitable.
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen's d , Pearson's r), indicating how they were calculated

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

Data collection	No software was used
Data analysis	<p>Publicly available tools were used for data analysis, statistics and visualization, with specific tools and versions described in the supplementary information methods section, and listed below. GRIN2 Tools and libraries were sourced from: https://raw.githubusercontent.com/stjude/TALL-example/main/GRIN2.0.ALEX.library.09.29.2022.R.</p> <p>Software: R (v. 4.2.2)</p> <p>R packages:</p> <ul style="list-style-type: none"> survminer (v0.4.9) sva (v3.46) DESeq2 (v1.38.3) uwot (v0.1.14) igraph (v1.3.5) scran (v1.26.2) edgeR (v3.4) limma (v3.54.1) biomaRt (v104) DNDscv (github: im3sanger/dndscv) glmnet (v.4.1.6)

cmprsk (v2.2-11)
tidy cmprsk (v0.2.0)
ranger (v.0.14.1)
survival (v. 3.2-3.13)
rpart (v4.1.16)
intsurv (v0.2.2)
Seurat (v4.0.53)

Sequence analysis:
BWA-MEM (v0.7.12)
Trim Galore (v0.4.4)
SAMtools (v1.10)
deepTools (v3.4.0)
STAR v(2.7.11)
RSEM v(1.3.0)
Mutect2(v4.1.2.0)
SomaticSniper(v1.0.5.0)
VarScan2(v2.4.3)
MuSE(v1.0rc)
Strelka2(v2.9.10).
Annovar (v1.0)
CONCERTING (v1.0)
Delly (v0.8.2)
Manta (v1.5.0)
GRIDSS (v2.5.0)
LUMPY(v0.3.0)
SURVIVOR (v1.0.7)
SVtyper (v0.7.1)
Arriba (v2.4.0)
Fusioncatcher (v1.0)
Pizzly (v1.0)
SQUID (v1.0)
STAR-fusion (v1.2)

Other:
IGV (v2.14)
Minimap2 (v1.0)
MiXCR (v3.0.13)
Cis-X (v1.0)
Gistic2 (v1.0)
MAPS (v2.0)
Juicer (v1.0)

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our [policy](#)

All primary sample and PDX sequencing data generated for this study are available through dbGaP at the accession number: phs003432, as part of the Childhood Cancer Data Initiative. This includes raw sequencing data of all primary T-ALL patients (CITE-seq and scATAC-seq), healthy thymus (CITE-seq and scATAC-seq), and T-ALL PDX models (CITE-seq), along with the corresponding count matrices and Seurat objects. Healthy bone marrow samples included in the analysis are available as previously published. AML and MPAL samples included for comparative analysis were previously reported can be accessed at <https://scpc.alexlemonade.org/projects/SCPCP000003#samples>. All other source data supporting the findings of this study are provided with this paper.

Research involving human participants, their data, or biological material

Policy information about studies with [human participants or human data](#). See also policy information about [sex, gender \(identity/presentation\), and sexual orientation](#) and [race, ethnicity and racism](#).

Reporting on sex and gender

All patients in the cohort were male or female according to biological sex, and no information on gender identity was collected.

Six of the 16 PDX were derived from female patients with T-ALL. 10 of the PDX were derived from male patients.

Reporting on race, ethnicity, or

Racial breakdown (self-reported) of the patients whose samples were in PDX single cell profiling: 4 Black/African American, 1

other socially relevant groupings

Asian, 1 not reported, and 10 White. 12 of the 16 patients whose samples were used for single cell analysis of PDX were <16 years of age at diagnosis.

Blasts from harvested spleens from 10 PDX cases underwent ex vivo drug profiling. These cases were selected by having sufficient cells for ex vivo drug profiling, which requires ~50 million leukemic blasts. These 10 PDX included samples from 1 female patient, 1 Black patient, and 1 Asian patient. 7 of the 10 patients whose samples were sent for ex vivo drug profiling were <16 years of age.

All other details about the racial characteristics of our cohort are reported in the supplemental materials.

Population characteristics

AALL0434 (NCT04408005) is a Children's Oncology Group (COG) phase 3 international clinical trial for patients with newly diagnosed T-cell acute lymphoblastic leukemia (T-ALL) and lymphoblastic lymphoma (T-LL) aged 1-30 years.

Recruitment

Subjects were enrolled from 01/22/07 until 07/25/14 at 214 centers in the United States, Canada, Australia, New Zealand, and Switzerland. All subjects with T-ALL were required to enroll on a companion classification study that was used for sample banking and risk stratification, AALL03B1 (NCT00482352) from 1/22/07 until 08/08/10 or AALL08B1 (NCT01142427) from 08/09/10 until 07/25/14.

Ethics oversight

AALL0434, AALL03B1, and AALL08B1 were approved by the Pediatric Central Institutional Review Board (IRB), local IRBs at all participating centers, and NCI Cancer Evaluation and Therapeutic Program (CTEP). Written informed consent/assent was obtained from all study participants and/or their legally authorized representative in accordance with the Declaration of Helsinki. Genomic studies performed for this work were approved by COG, CTEP, and the local IRBs at the Children's Hospital of Philadelphia and St Jude Children's Research Hospital. Samples were decoded and assigned a unique study identifier (USI). Samples were banked at the COG biorepository at Nationwide Children's Hospital in Chicago, IL. NIH limits on maximum blood/marrow collection for research were followed and there were no deviations: no more than 5cc/kg may be drawn for research purposes in a single day and no more than 9.5cc/kg may be drawn for research purposes over any 8-week period.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

Life sciences

Behavioural & social sciences

Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size

Power calculations were based on anticipated sample size of 1287 patients based on available materials and estimation of usable data. The objective was to have sufficient statistical power to identify genomic alterations (Frequency >5%, adjusting for End of induction MRD), that have a significant impact on EFS.

Data exclusions

1309 subjects had successful complete sequencing defined as WGS (whole genome sequencing), WES (whole exome sequencing), and WTS (whole transcriptome sequencing) of tumor and WGS of normal matched control. An additional 53 subjects had successful WTS sequencing without either WGS/WES of tumor or normal matched control and were included in transcriptome only analyses. A comparison of important clinical and demographic features as well as outcomes between the eligible and evaluable T-ALL patients and the sequenced cohort are provided in supplemental Table 2, demonstrating the sequenced cohort was representative of the overall trial cohort.

Replication

No replication could be performed given that a single sample was collected from each patient, but large sample size (N=1309) ensured statistical robustness.

Randomization

Randomization was not applicable because this is retrospective study for clinical and genomic analyses. Therefore, patients were already determined to have particular disease subtypes and/or treatments and could not be randomized into groups solely for the purpose of this study.

Blinding

Blinding was not applicable because this is retrospective study for clinical and genomic analyses. Therefore, patient metadata was already available and known by researchers a priori in order to select appropriate samples for our cohort design. All sequencing and sample preparations were blinded and randomized.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
<input type="checkbox"/>	<input checked="" type="checkbox"/> Antibodies
<input checked="" type="checkbox"/>	<input type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology and archaeology
<input type="checkbox"/>	<input checked="" type="checkbox"/> Animals and other organisms
<input type="checkbox"/>	<input checked="" type="checkbox"/> Clinical data
<input checked="" type="checkbox"/>	<input type="checkbox"/> Dual use research of concern
<input checked="" type="checkbox"/>	<input type="checkbox"/> Plants

Methods

n/a	Involved in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input type="checkbox"/>	<input checked="" type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging

Antibodies

Antibodies used

CD1a HI149 300133 GATCGTGTGTTA
 CD2 TS1/8 309229 TACGATTGTCAGGG
 CD3 UCHT1 300475 CTCATTGTAACCTCCT
 CD4 RPA-T4 300563 TGTTCCGCTCAACT
 CD5 UCHT2 300635 CATTAAACGGGATGCC
 CD7 CD7-6B7 343123 TGGATTCCGGACTT
 CD8 RPA-T8 301067 GCTGCGCTTCCATT
 CD10 HI10a 312231 CAGCCATTCAATTAGG
 CD14 63D3 367131 CAATCAGACCTATGA
 CD19 HIB19 302259 CTGGGCAATTACTCG
 CD22 S-HCL-1 363514 GGTTGTTGCTTTG
 CD33 P67.6 366629 TAACTCAGGGCCTAT
 CD34 581 343537 GCAGAAATCTCCCTT
 CD38 HB7 356635 CCTATTCCGATTCCG
 CD45 HI30 304064 TGCAATTACCCGGAT
 CD56 QA17A16 392421 TTGCGCGATTGAGT
 CD94 DX22 305521 CTTTCCGGTCCTACA
 CD117 104D2 313241 AGACTAATAGCTGAC
 CD123 6H6 306037 CTTCACTCTGTCAGG
 HLA-DR L243 307659 AATAGCGAGCAAGTA
 CD15 W6D3 323046 TCACCACTACCTAGT
 CD66b 6/40c 392905 AGCTGTAAGTTTCGG

anti-CD7 BioLegend 343106 PE
 anti-CD38 eBioscience 12-0388-42 PE
 anti-hCD45 BD Biosciences 561864 APC
 anti-CD34 Biolegend 561 343607 B273991 APC
 anti-CD1A Biolegend HI149 300105 B263720 PE

anti-CD235A Biolegend HI264 349103 B277415 FITC
 anti-CD56 Biolegend 5.1H11 362545 B263356 FITC
 anti-CD4 Biolegend RPA-T4 300505 B283934 FITC
 anti-CD16 Biolegend 3G8 302005 B309216 FITC
 anti-CD8A Biolegend RPA-T8 301005 B275512 FITC
 anti-CD19 Biolegend HIB19 302205 B265558 FITC
 anti-CD3 Biolegend HIT3a 300306 B274310 FITC

All antibodies were used at the concentrations recommended by the respective manufacturer

Validation

All antibodies are validated for detecting human and mouse proteins by the manufacturer, confirmed for each specific application using cells of known origin and differentiation state, and compared to isotype controls and cells that are known to express or lack the antigen. Each manufacturer above provides further validation information on their respective websites.

Animals and other research organisms

Policy information about [studies involving animals; ARRIVE guidelines](#) recommended for reporting animal research, and [Sex and Gender in Research](#)

Laboratory animals

NOD.Cg-Prkdcscid Il2rgtm1Wjl/SzJ (NSG, global reference ID RRID:IMSR_JSX:005557) mice were used for all experiments.

For all experiments, mice were maintained at three to four mice per cage under humidity and temperature-controlled conditions with a light/dark cycle that is set at 12-hours. Animals were maintained under microisolator tops in a HEPA-filtered rack. Animals were fed autoclaved Purina mouse chow and water ad libitum. Handling was performed with universal sterile precautions and experienced personnel will perform all procedures. Mice were sacrificed, and tissues harvested, consistent with the recommendation of the Panel of Euthanasia of the American Veterinary Medical Association. Animal sacrifice is performed by administration of

isoflurane or CO₂ sedation followed by cervical dislocation. This methodology is consistent with the recommendations of the Panel on Euthanasia of the American Veterinary Medical Association. Please refer to the "AALL0434 patient identification and clinical annotation" section of methods for details on patient enrollment and IRB approval for the work.

For the mouse colony, 1 male and 1-2 females were maintained in each cage. Mice were bred from approximately 8 weeks of age until 12 months of age, depending on health of the animals. For the development of patient derived xenograft (PDX) models, we injected ~106 blasts from viably frozen patient samples (bone marrow or blood) per mouse to develop primagrafts. We injected 1-5 mice for each patient sample. Mice for each sample for injection were housed in a single cage. Ages of the mice were 2 to 6.5 months at injection.

Wild animals

Provide details on animals observed in or captured in the field; report species and age where possible. Describe how animals were caught and transported and what happened to captive animals after the study (if killed, explain why and describe method; if released, say where and when) OR state that the study did not involve wild animals.

Reporting on sex

Male and female mice were used for PDX generation, but as male and female mice cannot be housed together and with limited cell numbers, each patient sample was injected either into a cage of male mice or a cage of female mice. The use of male or female mice was randomly selected for each patient sample. PDX were successfully made from 24 unique patients.

For the in vivo PDX experiments, male mice from 2 PDX were used (6 mice for sample PATTDP and 9 mice for sample PAUNDK).

Field-collected samples

For laboratory work with field-collected samples, describe all relevant parameters such as housing, maintenance, temperature, photoperiod and end-of-experiment protocol OR state that the study did not involve samples collected from the field.

Ethics oversight

Mice were purchased from the Jackson Laboratory to establish a breeding colony at the Children's Hospital of Philadelphia (CHOP) and mouse bred in the colony were used for experiments. All animals were housed in the laboratory animal facility (LAF) vivarium in the Colket Translational Research Building (CTR) at CHOP. The LAF is accredited by the American Association for Accreditation of Laboratory Animal Care (AAALAC), registered with the USDA and complies with the Public Health Service Policy on Humane care and Use of Laboratory Animals (Section: A3442-01). Additionally, the LAFs activities involving animals comply with the Guide for the Care and use of Laboratory Animals. Animal work in our laboratory was reviewed and approved by our Institutional Animal Care and Utilization Committee (21-000232, most recent re-approval 5/07/2024).

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Clinical data

Policy information about [clinical studies](#)

All manuscripts should comply with the ICMJE [guidelines for publication of clinical research](#) and a completed [CONSORT checklist](#) must be included with all submissions.

Clinical trial registration

No study protocol was provided for the current study because this is retrospective study for clinical and genomic analyses of the AALL0434 clinical trial cohort (NCT00408005).

Study protocol

No study protocol was provided for the current study because this is retrospective study for clinical and genomic analyses of the AALL0434 clinical trial cohort (NCT00408005).

Data collection

Randomization was not applicable because this is retrospective study for clinical and genomic analyses. No intervention was performed.

Outcomes

No pre-determined primary and secondary outcomes because this is retrospective study for clinical and genomic analyses.

Flow Cytometry

Plots

Confirm that:

- The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).
- The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).
- All plots are contour plots with outliers or pseudocolor plots.
- A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation

Pediatric thymus tissue was mechanically disrupted and treated with liberase (0.2mg/mL Roche; 30 min at 37°C) with intermittent shaking, as previously described²⁴. Thymocytes were resuspended into flow buffer, sorted into DAPI-Lin-CD34+CD1A- fractions and subject to scRNA-seq and scATAC-seq.

PDX expanded blasts were harvested from spleen or bone marrow. Frozen samples were thawed (37°C) and resuspended in IMDM + 2% FBS and treated with DNase I twice. Cells were gently washed 2x with RPMI, resuspended in flow buffer, stained with DAPI and anti-human CD45 antibody (BD Pharmingen, Cat #: 555485) and subject to FACS sorting (FACSAria Fusion, BD).

Instrument

Flow cytometry was carried out by the Flow Cytometry Core at the Children's Hospital of Philadelphia. FACS sorting was conducted on the FACSAria Fusion, BD

Software

FlowJo was used for flow analysis

Cell population abundance

The purity of populations sorted for single-cell sequencing was confirmed and further purified in silico through computational analysis.

Gating strategy

For in vivo experiments, gates were set using CD38-FITC/CD45-APC FMO controls or CD7-PE/CD45-APC FMO controls. The percentage of hCD38+CD45+ blasts or hCD45+CD7+ blasts amongst all live singlets were plotted

Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.