haematologica
Journal of the Ferrata Storti Foundation

Distinct clinical and molecular profiles of acute myeloid leukemia defined by mutations in *IDH1*, *IDH2*, and *TET2*

by Fanny Asbaghi, Erik Håkansson, Jakob Hofvander, Yuhong Ning, Yuan Xue, Tierney Philips, Vladimir Lazarevic, Christer Nilsson, Sören Lehmann, Samuel Levy, Gunnar Juliusson, Ravindra Majeti and Niklas Landberg

Received: May 27, 2026.
Accepted: October 29, 2026.

Citation: Fanny Asbaghi, Erik Håkansson, Jakob Hofvander, Yuhong Ning, Yuan Xue, Tierney Philips, Vladimir Lazarevic, Christer Nilsson, Sören Lehmann, Samuel Levy, Gunnar Juliusson, Ravindra Majeti and Niklas Landberg. Distinct clinical and molecular profiles of acute myeloid leukemia defined by mutations in IDH1, IDH2, and TET2.
Haematologica. 2026 Nov 5. doi: 10.3324/haematol.2026.301319 [Epub ahead of print]

Publisher's Disclaimer.
E-publishing ahead of print is increasingly important for the rapid dissemination of science. Haematologica is, therefore, E-publishing PDF files of an early version of manuscripts that have completed a regular peer review and have been accepted for publication.
E-publishing of this PDF file has been approved by the authors.
After having E-published Ahead of Print, manuscripts will then undergo technical and English editing, typesetting, proof correction and be presented for the authors' final approval, the final version of the manuscript will then appear in a regular issue of the journal.
All legal disclaimers that apply to the journal also pertain to this production process.

Distinct clinical and molecular profiles of acute myeloid leukemia defined by mutations in *IDH1*, *IDH2*, and *TET2*

Fanny Asbaghi[1], Erik Håkansson[1,2], Jakob Hofvander[1], Yuhong Ning[3], Yuan Xue[3], Tierney Philips[3], Vladimir Lazarevic[2], Christer Nilsson[4], Sören Lehmann[5], Samuel Levy[3], Gunnar Juliusson[2], Ravindra Majeti[6], Niklas Landberg[1,2]

1 Division of Clinical Genetics, Department of Laboratory Medicine, Lund University, Sweden

2 Department of Hematology, Oncology and Radiation Physics, Skåne University Hospital, Lund, Sweden

3 ClearNote Health, San Diego, California, USA

4 Division of Hematology, Department of Medicine, Karolinska University Hospital, Huddinge, Stockholm, Sweden

5 Department of Medicine Huddinge, Center for Hematology and Regenerative Medicine, Karolinska Institutet, Stockholm, Sweden

6 Institute for Stem Cell Biology and Regenerative Medicine, Stanford University School of Medicine, Stanford, California, USA

Authors' contributions: F.A. and N.L. designed the study; G.J., C.N., V.L., Sö.L., E.H., R.M., and N.L. collected patient material and clinical data; T.P. and N.L. performed the experiments; F.A., E.H., and N.L. analyzed the data, with Y.N. and Y.X. contributing to the data analysis; F.A. made the figures; F.A. and N.L. wrote the manuscript; all authors provided intellectual input and reviewed the manuscript.

Running title: AML subgroups defined by *IDH1*, *IDH2*, and *TET2* mutations

Correspondence:

Fanny Asbaghi, MSE, Division of Clinical Genetics, Department of Laboratory Medicine, BMC D14, Lund University, SE-221 84 Lund, Sweden.

fanny.asbaghi@med.lu.se

Niklas Landberg, MD, PhD, Division of Clinical Genetics, Department of Laboratory Medicine, BMC D14, Lund University, SE-221 84 Lund, Sweden; Department of Hematology, Oncology and Radiation Physics, Skåne University Hospital, SE-221 85 Lund, Sweden

niklas.landberg@med.lu.se

Data-sharing-statement: Sequencing data are available through dbGaP accession phs003853.

Acknowledgements: We wish to thank all the patients for providing AML samples for this study. We also thank Thomas Koehnke at Stanford University School of Medicine for his assistance with data management.

Funding: This work was supported by the Swedish Research Council, the Crafoord Foundation, the Swedish Society of Medicine, the Royal Physiographic Society in Lund, the Maggie Stephens Foundation, the John Persson Foundation for Hematological Research, the Lundgren Foundation, and the Knut and Alice Wallenberg Foundation.

Artificial intelligence disclosure: Artificial intelligence-based tools were used for limited language assistance and code debugging. No advanced artificial intelligence (aAI) was used in the preparation of this manuscript.

Conflict-of-interest disclosure: N.L. is cofounder of Lead Biologics International and has received honoraria from Servier; V.L has received honoraria from Servier; R.M. is on the Advisory Boards of Kodikaz Therapeutic Solutions, Orbital Therapeutics, Pheast Therapeutics, 858 Therapeutics, Prelude Therapeutics, Mubadala Capital, and Aculeus Therapeutics; R.M. is a cofounder and equity holder of Pheast Therapeutics, MyeloGene, Orbital Therapeutics and Sequentify; Y.N., Y.X., T.P., and Sa.L. are employed by ClearNote Health; G.J. reports institutional funding from Servier; the rest of the authors declare no conflict of interest.

ABSTRACT

Acute myeloid leukemia (AML) is a heterogeneous disease driven by recurring somatic mutations, frequently affecting genes involved in epigenetic regulation. Mutations in *IDH1*, *IDH2*, and *TET2* are largely mutually exclusive and thought to converge on a shared pathway by impairing TET-mediated DNA demethylation, either directly through loss of TET2 activity or indirectly through the accumulation of the oncometabolite (R)-2-hydroxyglutarate (2-HG) in *IDH1*- and *IDH2*-mutant AML. Whether this biochemical convergence translates into similar disease biology remains unclear. Here, we analyzed AML defined by mutations in *IDH1*, *IDH2*, or *TET2* using population-based data from the Swedish AML Registry, complemented by transcriptomic and 5-hydroxymethylcytosine (5hmC) profiling of purified blasts from a highly selected subset of primary AML samples. *TET2*-mutant AML occurred at an older age and was associated with a more pronounced thrombocytopenia, whereas *IDH1*- and *IDH2*-mutant AML showed deeper neutropenia and more favorable overall survival, including after adjustment for clinical covariates. Transcriptomic and 5hmC alterations showed only modest concordance, with 5hmC profiles providing clearer separation of the mutational subgroups. *TET2*-mutant samples showed significant enrichment of the oxidative phosphorylation pathway compared with *IDH1*- and *IDH2*-mutant AML. The mutational groups also showed distinct closest normal counterparts among healthy hematopoietic cell populations. Together, these findings demonstrate that *IDH1*, *IDH2*, and *TET2* mutations define clinically and molecularly distinct subgroups of AML, highlighting the relevance of mutation-specific clinical stratification.

INTRODUCTION

Acute myeloid leukemia (AML) is a genetically and epigenetically heterogeneous hematologic malignancy arising through clonal expansion and a stepwise acquisition of somatic mutations, with early mutations frequently affecting genes involved in epigenetic regulation.[1-3] Frequently recurring aberrations include heterozygous mutations in *isocitrate dehydrogenase 1* (*IDH1*), *isocitrate dehydrogenase 2* (*IDH2*), and loss-of-function mutations in the epigenetic regulator *TET methylcytosine dioxygenase 2* (*TET2*).[4] These mutations are largely mutually exclusive, and are thought to converge on the same epigenetic regulatory pathway.[5-7] Mutant IDH1 and IDH2 enzymes acquire the neomorphic function to convert α-ketoglutarate (α-KG) into the oncometabolite (R)-2-hydroxyglutarate (2-HG), which can inhibit α-KG–dependent dioxygenases including the TET family of 5-methylcytosine (5mC) hydroxylases. This leads to reduced active DNA demethylation and the resulting hypermethylation, generally believed to occur in *IDH1*-, *IDH2*-, and *TET2*-mutant AML.[8, 9] More recent data, however, indicate that 2-HG can also inhibit the KDM5 family of histone lysine demethylases, thereby altering histone methylation and promoting accumulation of active histone marks such as H3K4me3 and H3K4me2, promoting oncogenic transcription and cellular transformation in *IDH1*- and *IDH2*-mutant AML.[10] Despite this apparent mechanistic convergence, biological and clinical differences have been described between *IDH1*- and *IDH2*-mutant AML, as well as between *IDH1*-, *IDH2*-, and *TET2*-mutant AML, although the mechanisms underlying these differences remain incompletely understood.[11, 12] Neither *IDH1*, *IDH2*, nor *TET2* is included as an independent risk-defining factor in the ELN 2022 risk classification, although *TET2* mutations have been associated with shorter overall survival.[13] While *IDH1* and *IDH2* mutations have historically not been associated with outcome, the introduction of targeted therapies has substantially improved outcomes in these subgroups.[14] Combining azacitidine with ivosidenib for *IDH1*-mutant AML and venetoclax for *IDH2*-mutant AML significantly increases the overall survival of these subgroups, reflected in the ELN 2024 risk classification for less intensively treated AML.[15-17]

Here, we compare AML defined by mutations in *IDH1*, *IDH2*, or *TET2* using population-based data from the Swedish AML Registry and integrate these findings with transcriptomic and 5-hydroxymethylcytosine (5hmC) profiling of genetically matched primary AML samples. We

further examine whether molecular differences between these mutational subgroups are associated with distinct developmental states by comparison with normal hematopoietic reference populations. Despite the assumed convergence on disrupted TET function, *IDH1*-, *IDH2*-, and *TET2*-mutant AML show clear differences in clinical characteristics and outcome, 5hmC levels, and, to some degree, transcriptomic programs. These findings indicate that shared perturbation of DNA demethylation does not translate into uniform disease biology across these mutational subgroups.

METHODS

Patient samples and registry data

Primary AML bone marrow and peripheral blood samples were collected after written informed consent according to Stanford University guidelines [Stanford University Institutional Review Board (IRB) No. 6453]. Mononuclear cells were isolated by Ficoll separation and viably frozen. For the transcriptomic and epigenetic analyses, samples were matched for additional genetic aberrations beyond the group-defining alteration in *IDH1*, *IDH2*, or *TET2*. The "Other" group was defined as wild-type for *IDH1*, *IDH2*, and *TET2* but matched for other mutations. Patient data were obtained from the population-based Swedish AML Registry and included individuals diagnosed with AML between 2007 and 2024. Comprehensive mutational data have been systematically recorded in the registry since 2020. Additional mutational data for approximately 800 patients diagnosed between 2013 and 2018 were retrieved from the Swedish National Acute Leukemia Biobank.[18, 19] Analyses were approved by the Swedish Ethical Review Authority (Dnr 2023-07681-02).

RNA sequencing

Samples were shipped to the Novogene commercial sequencing facility where mRNA was enriched using poly-T oligo-attached magnetic beads and first-strand cDNA was synthesized using random hexamer primers. Second-strand synthesis, end repair, A-tailing, and adapter ligation were then performed, followed by size selection, amplification, and purification. After Bioanalyzer quality control, Illumina-based massively parallel sequencing was performed. The mean RNA integrity number (RIN) was 7.3 (range 5.3-9.2), and sequencing resulted in >92% of bases with Phred scores ≥30, indicating excellent base-calling quality. Adapter trimming and

quality filtering were performed using TrimGalore within the nf-core/rnaseq Nextflow pipeline. Reads were aligned to the GRCh38.p13 human reference genome using STAR in two-pass mode, and transcript quantification was performed using Salmon. Downstream analyses of gene counts were restricted to protein-coding genes filtered using Bioconductor/org.Hs.eg.db (v.3.21.0). Differential expression analyses were performed using DESeq2 (v1.49.4).[20]

5hmC sequencing

To study epigenetic differences, FACS-isolated AML blasts were analyzed using click chemistry enrichment for 5-hydroxymethylcytosine (5hmC)-containing DNA fragments coupled with sequencing as previously described.[21] Sequencing libraries were aligned to GRCh38 and deduplicated. Fragments were collapsed into gene-body counts based on fragment overlap and counts per gene were obtained. Low 5hmC counts were filtered using edgeR (v4.7.5).[22] Genes were retained if they had counts per million (CPM) >1 in at least 15% of samples, with a minimum of two samples. The counts were normalized using trimmed mean of M-values (TMM) normalization. LogCPM values were computed using the edgeR::cpm(log = TRUE, prior.count = 1).

Additional methods are provided in the Supplementary Appendix.

RESULTS

Population-level differences in clinical characteristics and outcome between *IDH1-, IDH2-*, and *TET2*-mutant AML

Out of the patients included in the Swedish AML Registry during 2007–2024, 2,515 had molecular genetic information available. Of these, 2,471 were classified into four mutually exclusive mutational subgroups: *IDH1* mutation (n=204, 8.1%), *IDH2* mutation (n=392, 15.6%), *TET2* mutation (n=548, 21.8%), and *Other*, comprising samples lacking these mutations (n=1,327, 52.8%). Acute promyelocytic leukemia was excluded from downstream analyses. An additional 44 patients (1.7%) harbored mutations in more than one of these genes and were categorized separately and excluded from downstream analyses (Supplementary Figure 1A). Clinical characteristics are summarized in Table 1. *TET2*-mutant AML patients were significantly older, with a median age of 73 years, compared with 71 years for *IDH2*, 68 years for *IDH1*, and 67 years for *Other* AML (Supplementary Figure 1B). There were no significant age differences

between *IDH1*- and *IDH2*-mutant patients. Notable differences in baseline clinical characteristics included lower platelet counts in *TET2*-mutant AML (median $57x10^9$/L) compared with *IDH1* ($93x10^9$/L), *IDH2* ($93x10^9$/L), and *Other* ($64x10^9$/L), and a more pronounced neutropenia in *IDH1* ($0.3x10^9$/L) and *IDH2* ($0.7x10^9$/L) compared with *TET2* ($2.4x10^9$/L) and *Other* AML ($1.1x10^9$/L, Supplementary Figure 1C).

When comparing overall survival between the mutational groups, Kaplan-Meier analysis indicated significant differences between the groups with the longest median overall survival observed for *IDH1*-mutant patients (2.5 years), followed by *IDH2* (1.9 years), compared with *TET2* (0.8 years; log-rank p<0.0001, Figure 1A). A multivariable Cox proportional hazards model adjusted for mutation group, age, sex, treatment intensity (low-intensity/HMA-based or intensive chemotherapy), and allogeneic stem cell transplantation (SCT) modeled as a time-dependent covariate, confirmed that both *IDH1*- and *IDH2*-mutant AML were associated with lower hazards of death (HR=0.73 and 0.68, respectively) compared with *Other* AML, as was intensive chemotherapy (HR=0.65) compared with low-intensity/HMA-based therapy (Figure 1B). Older age was associated with inferior overall survival (HR=1.20 per 5 years). With similar survival seen in the *IDH1*- and *IDH2*-mutant AML groups, we next combined them into an *IDH1*/*IDH2* group (n=594) and compared survival with *TET2*-mutant AML (n=546). *TET2*-mutant AML showed a worse outcome with a median overall survival of 0.8 years compared to 2.0 years for *IDH1*/*IDH2*-mutant AML (Supplementary Figure 1D). *TET2* mutation remained associated with worse prognosis after adjusting for age, treatment intensity, sex, and stem cell transplantation in a multivariable Cox regression model (Supplementary Figure 1E).

We next explored the effect of adding venetoclax to low-intensity/HMA-based therapies in *IDH1*- and *IDH2*-mutant AML. Addition of venetoclax was associated with a significant increase in overall survival for *IDH2*-mutant AML (median OS 1.78 versus 0.87 years, HR=0.42, 95% CI 0.23-0.78, p=0.005), with a similar, non-significant trend observed in *IDH1*-mutant AML (median OS 1.91 versus 0.43 years, HR=0.44, 95% CI 0.18-1.08, p=0.072, Supplementary Figure 2A-B). In intensively treated patients, *NPM1* co-mutations were associated with longer overall survival in *IDH2*-mutant AML (median OS not reached versus 2.25 years, HR=0.30, 95% CI 0.13-0.67, p=0.003), *TET2*-mutant AML (median OS not reached versus 0.82 years, HR=0.40, 95% CI 0.24-0.66, p<0.001), and *Other* AML (median OS not reached versus 2.72 years, HR=0.50, 95%

CI 0.33-0.75, p<0.001). A similar non-significant trend was observed in *IDH1*-mutant AML with a co-occurring *NPM1* mutation (median OS not reached versus 2.47 years, HR=0.38, 95% CI 0.14-1.08, p=0.07, Supplementary Figure 2C-F).

Because *TET2* allelic state and mutational burden have been implicated in disease biology and outcome, we compared survival in patients with <45% *TET2* VAF and >55% VAF, with high VAF suggestive of biallelic *TET2* inactivation.[23] No significant difference in overall survival was observed between the groups, although allelic state cannot be definitively established from VAF alone (Supplementary Figure 2G-H).

Co-mutational patterns differ between *IDH1*-, *IDH2*-, and *TET2*-mutant AML

To characterize the molecular profiles of each mutational group, we compared co-occurring genetic alterations (Figure 1C). *DNMT3A* was the most common mutation, occurring in 44.6% of *IDH1*- and 39.5% of *IDH2*-mutant AML, with lower frequencies in *TET2*-mutant (26.8%) and *Other* AML (20.7%, Bonferroni-adjusted pairwise p<0.001 for *IDH1* and *IDH2* versus *TET2* and *Other*). Mutations in *NPM1* were most prevalent in *IDH1*- (38.2%) and *TET2*-mutant (34.1%), with significantly lower prevalence in *IDH2*-mutant AML (26%; adjusted pairwise p=0.012 compared to *IDH1*). *SRSF2* mutations were most frequent in *IDH2* (34.2%) and were also frequently observed in both *IDH1* (19.1%) and *TET2* (26.6%) compared with *Other* AML (8.5%; adjusted pairwise p<0.001 for all three comparisons). Mutations in *TP53* were underrepresented in the *IDH1*-, *IDH2*-, and *TET2*-mutant compared with *Other* AML (adjusted pairwise p<0.001). Additional smaller (<10% absolute differences) yet significant differences were seen for *ASXL1* and *NRAS*, while *RUNX1*, *FLT3-ITD*, and *FLT3-TKD* showed no significant differences in prevalence between groups. Interestingly, no individual cytogenetic aberration differed significantly between the *IDH1*-, *IDH2*-, and *TET2*-mutant groups (Figure 1D). Instead, all three groups were more frequently associated with a normal karyotype compared with *Other* AML (62.8%, 61.8%, and 53.0% compared to 35.7%, respectively; adjusted pairwise p<0.001). The *Other* group showed higher frequencies of 5q/-5, 7q/-7, and 17p/-17 associated with adverse risk, as well as more favorable alterations inv(16) and t(8;21).

Transcriptomic differences between epigenetic subgroups of AML

To compare gene expression across *IDH1*-, *IDH2*-, *TET2*-mutant AML, and *Other* AML, we selected 12 primary AML samples closely matched for additional genetic aberrations and FACS-isolated AML blasts for mRNA sequencing (Table 2, Supplementary Figure 3A). Post-sort analyses confirmed purity ≥99% for all samples (Supplementary Figure 3B). PCA based on variance-stabilized RNA-seq expression across all measured protein-coding genes was used to assess global transcriptional variation and the separation of the predefined mutation groups (Figure 2A). Given the limited sample size and inter-patient variability, group separation and within-group heterogeneity were assessed using Euclidean distances and silhouette scores in the PC1-PC2 space. *IDH2*-mutant AML samples showed the most distinct group structure (mean silhouette score = 0.551), whereas *IDH1*, *TET2*, and particularly *Other* samples showed greater within-group heterogeneity (mean silhouette scores = −0.108, −0.082, and −0.602, respectively). Consistent with the heterogeneity of the *Other* AML group, exclusion of these samples increased the overall silhouette score from −0.060 to 0.190, while the mean silhouette scores were −0.075 for *IDH1*, 0.551 for *IDH2*, and 0.093 for *TET2*. Notably, all three *IDH1*-mutant samples had *IDH2* as their nearest alternative group, while all three *IDH2*-mutant samples had *IDH1* as their nearest alternative group. Thus, despite greater heterogeneity within *IDH1*, the *IDH1*- and *IDH2*-mutant samples were consistently more similar to each other than to *TET2*-mutant AML in the PC1-PC2 space.

To explore shared expression patterns, *IDH1*, *IDH2*, and *TET2* were combined and compared with *Other*, resulting in 11 significantly differentially expressed genes (Supplementary Figure 4A). These included relative upregulation of *HOXB8* and *HOXB9*, although expression levels varied considerably between samples (Supplementary Figure 4B). We next compared *IDH1*/*IDH2*-mutant with *TET2*-mutant AML, identifying 434 upregulated and 762 downregulated genes in the *IDH1*/*IDH2*-mutant group (Figure 2B-C). GSEA identified multiple Hallmark pathways enriched in the *TET2*-mutant samples relative to *IDH1*/*IDH2*-mutant samples, including pathways related to inflammatory signaling, metabolism, and oxidative phosphorylation, while no pathways were significantly enriched in *IDH1*/*IDH2*-mutant samples (Figure 2D). Comparison of *IDH1*/*IDH2*-mutant with *Other* AML identified 97 differentially expressed genes, while comparison of *TET2*-mutant AML with *Other* AML identified 17 differentially expressed genes. The relatively small number of differences may in part reflect the high variability within the *Other* group, reducing the robustness of these comparisons (Supplementary

Figure 4C-D). Overall, transcriptomic heterogeneity was pronounced across samples, but pathway-level differences were observed between *IDH1*/*IDH2*- and *TET2*-mutant AML, with enrichment of inflammatory and metabolic transcriptional programs in *TET2*-mutant AML.

Distinct 5-hydroxymethylcytosine patterns in *IDH1*-, *IDH2*-, and *TET2*-mutant AML

To complement the transcriptional data, we subjected AML blasts to 5hmC pull-down sequencing. PCA of gene-body 5hmC levels was used to visualize the separation of the predefined mutation groups (Figure 3A). Similar to the RNA-seq data, PCA-based silhouette analysis showed the most distinct group structure for *IDH2*-mutant samples (mean silhouette score = 0.759), whereas the *IDH1*, *TET2*, and *Other* AML groups showed greater within-group heterogeneity (−0.116, −0.100, and −0.281, respectively). After exclusion of the heterogeneous *Other* AML samples, the overall silhouette score increased from 0.066 to 0.306, with mean silhouette scores of −0.063 for *IDH1*, 0.759 for *IDH2*, and 0.221 for *TET2*. All three *IDH1*-mutant samples had *IDH2* as their nearest alternative group and all three *IDH2*-mutant samples had *IDH1* as their nearest alternative group, indicating that the *IDH1*- and *IDH2*-mutant AML samples were consistently more similar to each other than to *TET2*-mutant AML in the PC1-PC2 space. Overall, global gene-body 5hmC levels were reduced in *IDH1*, *IDH2*, and *TET2* compared with *Other* AML (Figure 3B). Using a multi-group ANOVA, subgroup-specific 5hmC patterns emerged, with *IDH1*- and *IDH2*-mutant samples showing similar patterns (Figure 3C). We therefore combined *IDH1* and *IDH2* and compared them with *TET2*-mutant samples, identifying 708 genes with higher and 409 genes with lower 5hmC gene-body levels (Figure 3D). GSEA identified 12 significantly enriched Hallmark pathways, with apoptosis and TNF-α signaling via NF-κB enriched in *IDH1*/*IDH2*-mutant samples and oxidative phosphorylation enriched in the *TET2*-mutant samples (Figure 3E). Compared with *Other* AML, *IDH1*/*IDH2* showed 237 genes with higher and 98 genes with lower gene-body 5hmC levels (Supplementary Figure 5A). Comparison of *TET2*-mutant AML with *Other* AML identified only one significantly different gene-body 5hmC level (Supplementary Figure 5B).

Finally, integration of transcriptomic and 5hmC data revealed a modest correlation between the two data modalities (Spearman $\rho=0.27$, $p<0.001$, Supplementary Figure 5C). A total of 150 genes displayed concordant differential regulation across RNA and 5hmC, 81 genes with upregulation in *IDH1*/*IDH2*-mutant AML and 69 genes upregulated in *TET2*-mutant AML (Figure

4A). However, the majority (1604 genes) instead showed significantly different levels in only the RNA or 5hmC modality, indicating substantial uncoupling between 5-hydroxymethylation and transcription. To explore pathway-level differences between the groups across both modalities, RNA and 5hmC log2 fold changes were independently standardized to z-scores and combined into an integrated score used for preranked GSEA. This identified 11 significantly enriched Hallmark pathways. Eight pathways were enriched in *TET2*-mutant cases, including oxidative phosphorylation, coagulation, and complement, whereas three gene sets were enriched in the *IDH1*/*IDH2*-mutant group (Figure 4B). We next evaluated these pathways in publicly available TCGA AML transcriptomic data.[24] All 11 gene sets were significantly enriched in *TET2*-mutant AML compared with *IDH1*/*IDH2*-mutant AML (FDR<0.05), including oxidative phosphorylation, coagulation, and complement (Figure 4C), as well as the three pathways enriched in the *IDH1*/*IDH2*-mutant group in the integrated analysis. This discordance suggests that 5hmC does not directly predict transcriptional output but may capture additional information on molecular differences between the mutational subgroups.

Molecular resemblance to normal hematopoietic populations

To determine whether the observed molecular differences could reflect differences in the developmental state of the AML blasts, we next compared the AML samples with healthy hematopoietic populations. Publicly available single-cell RNA-seq data from healthy bone marrow were aggregated into donor-level pseudobulk profiles to construct a reference landscape of normal myeloid differentiation.[25] Projection of the AML samples onto this reference revealed a highly consistent pattern among the *IDH*-mutant samples: all *IDH1*- and *IDH2*-mutant AML samples were most similar to neutrophil-committed granulocyte-monocyte progenitor cell populations (GMP-Neut; Spearman ρ=0.547–0.636). In contrast, *TET2*-mutant AML displayed greater heterogeneity, with individual samples showing greatest similarity to GMP-Neut, promonocyte, or monocyte populations (ρ=0.560, 0.734, and 0.660, respectively; Figure 4D, Supplementary Figure 5D). This pattern was supported using independent bulk RNA-seq data from normal hematopoietic populations where, again, *IDH1*- and *IDH2*-mutant samples were most strongly correlated with GMPs (ρ=0.433–0.509), whereas two of the three *TET2*-mutant samples were most similar to monocytes (ρ=0.634 and 0.735), and the remaining sample to GMPs (ρ=0.519; Supplementary Figure 5E–F).[26]

A corresponding analysis of normal hematopoietic 5hmC profiles showed that *IDH1*- and *IDH2*-mutant AML were most similar to immature progenitor populations, spanning GMP, lymphoid-primed multipotent progenitors (LMPP), and common myeloid progenitors (CMP), with spearman ρ=0.803–0.832. In contrast, two of the three *TET2*-mutant samples were most similar to monocytes (ρ=0.792 and 0.851), while the third was most similar to GMP (ρ=0.835; Figure 4E, Supplementary Figure 5G).[21] Together, these analyses suggest that the transcriptional and epigenetic differences between *IDH1*/*IDH2*-mutant and *TET2*-mutant AML are associated, at least in part, with differences in their resemblance to normal stages of myeloid differentiation.

DISCUSSION

The genetic landscape of AML has been extensively characterized, revealing that early mutations frequently occur in genes involved in epigenetic regulation.[1, 4] Among these recurrent aberrations, mutations in *IDH1*, *IDH2*, and *TET2* are thought to converge on impaired TET-mediated DNA demethylation either via direct loss of TET2 function or indirectly via accumulation of the oncometabolite 2-HG.[6, 8] Despite this shared biochemical effect, AML defined by these mutations has been shown to be biologically and clinically heterogeneous.[6, 11, 12] Here, we combined nationwide registry data with transcriptomic and epigenetic profiling of primary samples to demonstrate that AML with *IDH1*, *IDH2*, and *TET2* mutations comprises clinically and molecularly distinct disease subgroups.

Overall survival remains poor in AML and one of the strongest predictors for outcome is age at diagnosis.[27, 28] The ELN 2022 risk classification used for younger patients intended for intensive chemotherapy has recently been complemented with the ELN 2024 response prediction model for patients receiving less-intensive therapy.[17, 29] Mutations in *IDH1* stratify AML patients into a favorable risk category only in the context of combining hypomethylating agents with the *IDH1* inhibitor ivosidenib in patients lacking a *TP53* mutation.[16, 30] *IDH2* mutations have also been shown to confer a better prognosis.[31] Our population-based analyses revealed significant differences in overall survival between mutational subgroups, with both *IDH1*- and *IDH2*-mutant AML exhibiting more favorable outcomes than *TET2*-mutant and *Other* AML, even after adjustment for clinical covariates, including age, treatment intensity, sex, and stem cell transplantation. This is of particular interest because the observed survival differences are

unlikely to be explained by routine use of IDH inhibitors, as ivosidenib and enasidenib had not yet been implemented in clinical practice in Sweden as of early 2026.

IDH1, *IDH2*, and *TET2* mutations have been described as preleukemic events and are present in clonal hematopoiesis as well as myelodysplastic syndrome.[32-35] These are early events and most often mutually exclusive.[6, 7] In our data, we observed different co-mutational patterns between these groups, suggesting potentially distinct evolutionary trajectories. *IDH1*- and *IDH2*-mutant AML frequently co-occurred with *DNMT3A* mutations, while *IDH1* was more commonly associated with *NPM1* mutations compared with *IDH2*-mutant AML. This is of particular interest considering recent findings from the HARMONY study, which demonstrated that the prognostic impact of *IDH1* mutations varies depending on their co-mutational context.[36] *SRSF2* mutations were frequently observed in all three groups but significantly enriched in *IDH2*-mutant AML, in line with previous reports linking *IDH2* and spliceosome-associated aberrations.[37] *TP53* mutations were rare in all three groups compared with *Other* AML. Another interesting distinction was the higher neutrophil counts seen in *TET2*-mutant AML compared to all other samples. This could potentially reflect a direct effect of reduced TET2 activity, which has been shown to increase production of dysfunctional neutrophils in xenograft models of *TET2*-mutant clonal hematopoiesis.[38]

Transcriptome-based classification of AML has proven less effective compared to diseases such as acute lymphoblastic leukemia (ALL), where gene expression profiling can delineate subtypes with high precision.[39] In AML, the most distinct transcriptomic signatures are typically associated with specific chromosomal translocations.[24, 40, 41] To increase the fidelity of our transcriptomic analyses, we matched primary AML samples for additional mutational patterns to the extent possible, selecting samples with normal karyotypes, *NPM1* mutations, and when possible, a signaling mutation. Despite selecting cases from the Stanford AML biobank containing more than 1000 AML samples, our transcriptomic analyses showed high variability within groups and, with our limited sample size, this precluded robust analyses of subgroup-specific expression patterns. The *Other* AML samples contributed substantially to the limited separation between groups, while inter-patient heterogeneity remained evident within the genetically defined *IDH1*-, *IDH2*-, and *TET2*-mutant AML subgroups. This underscores a key limitation of our study: the small number of genetically highly matched primary AML samples

available for integrative analyses. Instead, we focused on shared features of *IDH1*- and *IDH2*-mutant AML compared with *TET2*-mutant AML and identified *TET2*-specific enrichment of gene sets related to oxidative phosphorylation and inflammatory signaling. Both oxidative phosphorylation and inflammatory response gene sets were enriched in *TET2*-mutant AML samples in the independent TCGA AML transcriptomic cohort. Although no functional data are provided in this paper, the findings are consistent with previous reports of increased mitochondrial oxidative metabolism in *TET2*-deficient hematopoietic cells as well as altered mitochondrial metabolism in *IDH1*- and *IDH2*-mutant AML and preleukemic cells.[7, 11, 42-44] Targeting oxidative phosphorylation in AML has been suggested and explored in clinical trials but early results from the phase I trial of the complex I inhibitor IACS-010759 showed elevated lactate levels and dose-limiting toxicities including neurotoxicity.[45] The potential use of inhibitors of oxidative phosphorylation would most likely require better-tolerated compounds as well as better patient stratification to predict responders. Our analyses raise the possibility that oxidative phosphorylation may represent a relevant metabolic vulnerability in *TET2*-mutant AML, although pathway enrichment alone cannot establish sensitivity to oxidative phosphorylation inhibition. Conversely, *IDH1*/*IDH2*-mutant AML may retain dependencies on mitochondrial oxidative metabolism that are not reflected by higher transcriptional pathway enrichment. These hypotheses require functional validation. Our findings of increased inflammatory signaling pathways in *TET2*-mutant AML are consistent with previous mechanistic studies linking *TET2* loss to enhanced IL-1β and IL-6 signaling, suggesting that inflammatory pathways may represent a therapeutic vulnerability.[46, 47] In line with these findings, the CANTOS trial showed that individuals with *TET2*-mutant clonal hematopoiesis had a reduced risk of cardiovascular events during IL-1β blockade with canakinumab.[48]

Compared with transcriptomic profiling, the 5hmC data revealed a clearer overall subgroup structure among *IDH1*-, *IDH2*-, and *TET2*-mutant AML samples, as supported by higher silhouette scores, although substantial within-group heterogeneity was observed. Integrated analysis of RNA expression and 5hmC data showed only partial concordance, suggesting that altered 5hmC alone is insufficient to universally drive transcriptional change. This decoupling, previously reported but not fully understood, may reflect the context-dependent role of 5hmC and the contribution of additional regulatory mechanisms, including chromatin accessibility

and transcription factor activity, highlighting the need for further investigation into the mechanisms linking epigenetic modification and gene expression.[21, 49] The discordance further suggests that 5hmC may capture information on molecular differences between the mutational subgroups not reflected in transcriptional data alone.

Mapping to normal bone marrow reference populations revealed distinct closest normal counterparts across the mutational groups. *IDH1*/*IDH2*-mutant AML showed greatest similarity to GMP-like or other immature progenitor populations, whereas *TET2*-mutant AML more frequently resembled more mature monocytic populations in both the RNA and 5hmC analyses. These findings suggest that part of the molecular distinction between the groups may reflect differences in developmental state.

Together, our findings demonstrate that AML with mutations in *IDH1*, *IDH2*, or *TET2* represents biologically distinct subgroups with distinct clinical trajectories and epigenetic profiles, underscoring the need for mutation-informed disease stratification.

REFERENCES

1. Corces-Zimmerman MR, Hong WJ, Weissman IL, Medeiros BC, Majeti R. Preleukemic mutations in human acute myeloid leukemia affect epigenetic regulators and persist in remission. Proc Natl Acad Sci U S A. 2014;111(7):2548-2553.
2. Shlush LI, Zandi S, Mitchell A, et al. Identification of pre-leukaemic haematopoietic stem cells in acute leukaemia. Nature. 2014;506(7488):328-333.
3. Dohner H, Weisdorf DJ, Bloomfield CD. Acute myeloid leukemia. N Engl J Med. 2015;373(12):1136-1152.
4. Papaemmanuil E, Gerstung M, Bullinger L, et al. Genomic classification and prognosis in acute myeloid leukemia. N Engl J Med. 2016;374(23):2209-2221.
5. Abdel-Wahab O, Mullally A, Hedvat C, et al. Genetic characterization of TET1, TET2, and TET3 alterations in myeloid malignancies. Blood. 2009;114(1):144-147.
6. Figueroa ME, Abdel-Wahab O, Lu C, et al. Leukemic IDH1 and IDH2 mutations result in a hypermethylation phenotype, disrupt TET2 function, and impair hematopoietic differentiation. Cancer Cell. 2010;18(6):553-567.
7. Bassal MA, Samaraweera SE, Lim K, et al. Germline mutations in mitochondrial complex I reveal genetic and targetable vulnerability in IDH1-mutant acute myeloid leukaemia. Nat Commun. 2022;13(1):2614.
8. Xu W, Yang H, Liu Y, et al. Oncometabolite 2-hydroxyglutarate is a competitive inhibitor of alpha-ketoglutarate-dependent dioxygenases. Cancer Cell. 2011;19(1):17-30.
9. Wilson ER, Helton NM, Heath SE, et al. Focal disruption of DNA methylation dynamics at enhancers in IDH-mutant AML cells. Leukemia. 2022;36(4):935-945.
10. Gunn K, Myllykoski M, Cao JZ, et al. (R)-2-hydroxyglutarate inhibits KDM5 histone lysine demethylases to drive transformation in IDH-mutant cancers. Cancer Discov. 2023;13(6):1478-1497.
11. Landberg N, Kohnke T, Feng Y, et al. IDH1-mutant preleukemic hematopoietic stem cells can be eliminated by inhibition of oxidative phosphorylation. Blood Cancer Discov. 2024;5(2):114-131.
12. Fortin J, Chiang MF, Meydan C, et al. Distinct and opposite effects of leukemogenic Idh and Tet2 mutations in hematopoietic stem and progenitor cells. Proc Natl Acad Sci U S A. 2023;120(4):e2208176120.
13. Wang R, Gao X, Yu L. The prognostic impact of tet oncogene family member 2 mutations in patients with acute myeloid leukemia: a systematic-review and meta-analysis. BMC Cancer. 2019;19(1):389.
14. Zarnegar-Lumley S, Alonzo TA, Gerbing RB, et al. Characteristics and prognostic impact of IDH mutations in AML: a COG, SWOG, and ECOG analysis. Blood Adv. 2023;7(19):5941-5953.
15. DiNardo CD, Jonas BA, Pullarkat V, et al. Azacitidine and venetoclax in previously untreated acute myeloid leukemia. N Engl J Med. 2020;383(7):617-629.
16. Montesinos P, Recher C, Vives S, et al. Ivosidenib and azacitidine in IDH1-mutated acute myeloid leukemia. N Engl J Med. 2022;386(16):1519-1531.
17. Dohner H, DiNardo CD, Appelbaum FR, et al. Genetic risk classification for adults with AML receiving less-intensive therapies: the 2024 ELN recommendations. Blood. 2024;144(21):2169-2173.
18. Juliusson G, Lazarevic V, Horstedt AS, Hagberg O, Hoglund M, Swedish Acute Leukemia Registry G. Acute myeloid leukemia in the real world: why population-based registries are needed. Blood. 2012;119(17):3890-3899.

19. Juliusson G, Jadersten M, Deneberg S, et al. The prognostic impact of FLT3-ITD and NPM1 mutation in adult AML is age-dependent in the population-based setting. Blood Adv. 2020;4(6):1094-1101.
20. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15(12):550.
21. Nakauchi Y, Azizi A, Thomas D, et al. The cell type-specific 5hmC landscape and dynamics of healthy human hematopoiesis and TET2-mutant preleukemia. Blood Cancer Discov. 2022;3(4):346-367.
22. Chen Y, Chen L, Lun ATL, Baldoni PL, Smyth GK. edgeR v4: powerful differential analysis of sequencing data with expanded functionality and improved support for small counts and larger datasets. Nucleic Acids Res. 2025;53(2):gkaf018.
23. Awada H, Nagata Y, Goyal A, et al. Invariant phenotype and molecular association of biallelic TET2 mutant myeloid neoplasia. Blood Adv. 2019;3(3):339-349.
24. Cancer Genome Atlas Research Network. Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. N Engl J Med. 2013;368(22):2059-2074.
25. Zeng AGX, Iacobucci I, Shah S, et al. Single-cell transcriptional atlas of human hematopoiesis reveals genetic and hierarchy-based determinants of aberrant AML differentiation. Blood Cancer Discov. 2025;6(4):307-324.
26. Corces MR, Buenrostro JD, Wu B, et al. Lineage-specific and single-cell chromatin accessibility charts human hematopoiesis and leukemia evolution. Nat Genet. 2016;48(10):1193-1203.
27. Appelbaum FR, Gundacker H, Head DR, et al. Age and acute myeloid leukemia. Blood. 2006;107(9):3481-3485.
28. Hemminki K, Zitricky F, Forsti A, et al. Age-specific survival in acute myeloid leukemia in the Nordic countries through a half century. Blood Cancer J. 2024;14(1):44.
29. Dohner H, Wei AH, Appelbaum FR, et al. Diagnosis and management of AML in adults: 2022 recommendations from an international expert panel on behalf of the ELN. Blood. 2022;140(12):1345-1377.
30. Montesinos P, Marchione DM, Recher C, et al. Long-term results from the AGILE study of azacitidine plus ivosidenib vs placebo in newly diagnosed IDH1-mutated AML. Blood Adv. 2025;9(20):5177-5189.
31. Hoff FW, Huang Y, Welkie RL, et al. IDH2 mutation is associated with favorable outcome among older adults with newly diagnosed acute myeloid leukemia treated with hypomethylating agent-based therapy. Haematologica. 2026;111(5):1625-1633.
32. Jaiswal S, Fontanillas P, Flannick J, et al. Age-related clonal hematopoiesis associated with adverse outcomes. N Engl J Med. 2014;371(26):2488-2498.
33. Genovese G, Kahler AK, Handsaker RE, et al. Clonal hematopoiesis and blood-cancer risk inferred from blood DNA sequence. N Engl J Med. 2014;371(26):2477-2487.
34. DiNardo CD, Jabbour E, Ravandi F, et al. IDH1 and IDH2 mutations in myelodysplastic syndromes and role in disease progression. Leukemia. 2016;30(4):980-984.
35. Jung HA, Jung CW, Jang JH. Mutations in genes affecting DNA methylation enhances responses to decitabine in patients with myelodysplastic syndrome. Korean J Intern Med. 2021;36(2):413-423.

36. Hernandez-Sanchez A, Villaverde Ramiro A, Strang E, et al. Unravelling co-mutational patterns with prognostic implications in NPM1 mutated adult acute myeloid leukemia - a HARMONY study. Leukemia. 2026;40(2):418-428.
37. Telonis AG, Stanley RF, Adelman ER, et al. Synergistic intragenic epigenetic deregulation by IDH2 and SRSF2 mutations causes mis-splicing of key transcriptional regulators. Sci Adv. 2026;12(1):eadu8292.
38. Huerga Encabo H, Aramburu IV, Garcia-Albornoz M, et al. Loss of TET2 in human hematopoietic stem cells alters the development and function of neutrophils. Cell Stem Cell. 2023;30(6):781-799.e9.
39. Li JF, Dai YT, Lilljebjorn H, et al. Transcriptional landscape of B cell precursor acute lymphoblastic leukemia based on an international study of 1,223 cases. Proc Natl Acad Sci U S A. 2018;115(50):E11711-E11720.
40. Tyner JW, Tognon CE, Bottomly D, et al. Functional genomic landscape of acute myeloid leukaemia. Nature. 2018;562(7728):526-531.
41. Lilljebjorn H, Pena-Martinez P, Thorsson H, et al. The AML cellular state space unveils NPM1 immune evasion subtypes with distinct clinical outcomes. Nat Commun. 2025;16(1):10592.
42. Chan SM, Thomas D, Corces-Zimmerman MR, et al. Isocitrate dehydrogenase 1 and 2 mutations induce BCL-2 dependence in acute myeloid leukemia. Nat Med. 2015;21(2):178-184.
43. Thomas D, Wu M, Nakauchi Y, et al. Dysregulated lipid synthesis by oncogenic IDH1 mutation is a targetable synthetic lethal vulnerability. Cancer Discov. 2023;13(2):496-515.
44. Loke J, Kim PG, Nguyen TTP, et al. An in vivo barcoded CRISPR-Cas9 screen identifies Ncoa4-mediated ferritinophagy as a dependence in Tet2-deficient hematopoiesis. Blood. 2025;146(10):1174-1186.
45. Yap TA, Daver N, Mahendra M, et al. Complex I inhibitor of oxidative phosphorylation in advanced solid tumors and acute myeloid leukemia: phase I trials. Nat Med. 2023;29(1):115-126.
46. Cai Z, Kotzin JJ, Ramdas B, et al. Inhibition of inflammatory signaling in Tet2 mutant preleukemic cells mitigates stress-induced abnormalities and clonal hematopoiesis. Cell Stem Cell. 2018;23(6):833-849.e5.
47. McClatchy J, Strogantsev R, Wolfe E, et al. Clonal hematopoiesis related TET2 loss-of-function impedes IL1beta-mediated epigenetic reprogramming in hematopoietic stem and progenitor cells. Nat Commun. 2023;14(1):8102.
48. Svensson EC, Madar A, Campbell CD, et al. TET2-driven clonal hematopoiesis and response to canakinumab: an exploratory analysis of the CANTOS randomized clinical trial. JAMA Cardiol. 2022;7(5):521-528.
49. Rasmussen KD, Berest I, Kebetaler S, et al. TET2 binding to enhancers facilitates transcription factor recruitment in hematopoietic cells. Genome Res. 2019;29(4):564-575.

TABLES

Table 1. Clinical characteristics of AML according to mutational group

Variable[1]	Mutation group				p-value[2]
	IDH1 N=204	IDH2 N=392	TET2 N=548	Other N=1,327	
Age (years)	68.0 (59.5, 74.0)	71.0 (62.5, 76.0)	73.0 (67.0, 78.0)	67.0 (53.0, 75.0)	<0.001
Sex					n.s.
Male	97 (47.5%)	222 (56.6%)	317 (57.8%)	740 (55.8%)	
Treatment intensity, n (%)[3]					
Intensive chemo	152 (74.5%)	265 (67.9%)	290 (53.1%)	874 (66.1%)	<0.001
Low intensity/HMA	42 (20.6%)	93 (23.8%)	181 (33.2%)	323 (24.4%)	<0.001
Palliative care	10 (4.9%)	32 (8.2%)	75 (13.7%)	126 (9.5%)	0.003
Hemoglobin (g/L)	93 (80, 105)	95 (81, 108)	92 (80, 107)	90 (77, 103)	<0.001
White blood cell count ($\times 10^9$/L)	3.1 (1.4, 27.7)	4.9 (1.9, 36.6)	19.2 (4.2, 64.1)	7.2 (2.5, 32.6)	<0.001
Blood blast count ($\times 10^9$/L)	0.7 (0.1, 20.1)	0.7 (0.0, 18.7)	3.0 (0.3, 24.6)	1.1 (0.1, 10.4)	<0.001
Lactate dehydrogenase (µkat/L)	5.2 (3.4, 8.2)	4.5 (3.3, 8.8)	6.7 (4.3, 11.7)	6.3 (4.0, 11.0)	<0.001
Bone marrow blasts (%)	59 (35, 76)	51.5 (30, 75)	45 (26, 71)	43 (26, 65)	<0.001
Albumin (g/L)	34 (29, 38)	34 (30, 38)	35 (30, 38)	34 (29, 38)	n.s.
Neutrophil count ($\times 10^9$/L)	0.3 (0.1, 0.8)	0.7 (0.2, 2.3)	2.4 (0.7, 8.0)	1.1 (0.4, 3.4)	<0.001
Platelet count ($\times 10^9$/L)	93 (52, 168)	93 (44, 167)	57 (30, 104)	64 (31, 116)	<0.001

[1] Median (Q1, Q3)

[2] Kruskal-Wallis rank sum test for continuous variables; Pearson's chi-squared test for categorical variables. Treatment intensity row-specific p-values were Bonferroni-adjusted.

[3] Intent at diagnosis.

Table 2. AML samples analyzed by RNA and 5hmC enrichment sequencing

Sample	Sex	Age	Tissue	Cytogenetics	Mutation	VAF (%)
A1_IDH1	M	59	Leukapheresis	46,XY[14]	IDH1-R132H	45.57
					NPM1-L287ins(TCTG)	48.95
					FLT3-ITD	41.18
A2_IDH1	M	48	Leukapheresis	46,XY[20]	IDH1-R132H	39.84
					NPM1-L287ins(TCTG)	28.15
					KIT-D816V	37.82
					TP53-R175H	7.61
A3_IDH1	F	54	PB	46,XX[20]	IDH1-R132C	50.62
					NPM1-L287ins(TCTG)	32.33
					FLT3-ITD	NA
					FLT3-V592D	10.33
B1_IDH2	M	70	BM	No analyzable metaphases	IDH2-R140Q	46.43
					NPM1-L287ins(TCTG)	51.14
					FLT3-I836del(GAT)	37.24
B2_IDH2	M	69	BM	46,XY[21]	IDH2-R140Q	47.66
					NPM1-L287ins(TCTG)	34.75
					GATA2-K378ins(T)	51.53
B3_IDH2	F	71	PB	46,XX[13]	IDH2-R140Q	44.29
					NPM1-L287ins(TCTG)	27.85
					FLT3-V592D	31.34
C1_TET2	F	81	PB	No analyzable metaphases	TET2-E100del(C)	41
					NPM1-W288ins(GCAA)	31.27
C2_TET2	M	82	BM	46,XY[20]	TET2-501*	45.07
					NPM1-L287ins(TCTG)	43.7
C3_TET2	F	66	PB	46,XX[21]	TET2-H1868Y	46.91
					NPM1-L287ins(TCTG)	35.48
					FLT3-ITD	NA
D1_Other	F	54	BM	46,XX	NPM1-L287ins(TCTG)	52.14
					RAD21-Q592*	41.40
					PTPN11-A72T	45.25
D2_Other	F	41	BM	46,XX[20]	NPM1-W288ins(GCCA)	20.49

					GATA2-A372T	29.21
D3_Other	M	54	PB	46,XY[7]	NPM1-W288ins(GTTC)	31.91
					FLT3-ITD	NA
					JAK3-V722I	47.71

FIGURE LEGENDS

Figure 1: Distinct clinical characteristics of AML with mutations in *IDH1*, *IDH2*, or *TET2*.

A. Kaplan–Meier estimates of overall survival according to mutation group. Patients without mutations in *IDH1*, *IDH2*, or *TET2* were classified as *Other*. Significance is indicated with log-rank p<0.0001.

B. Multivariable Cox proportional hazards regression showing hazard ratios (HR) and 95% confidence interval (CI) for overall survival. The model included mutation group, treatment intensity, age (5-year increments), sex, and allogeneic stem cell transplantation (SCT), with SCT modeled as a time-dependent covariate.

C. Frequencies of the most common co-mutations within the designated mutation groups. Significant p-values from Pearson's chi-squared tests are indicated for pairwise comparisons with an absolute frequency difference ≥10% after Bonferroni correction (* p<0.05, ** p<0.01, *** p<0.001).

D. Frequencies of normal karyotype and selected cytogenetic aberrations across mutation groups. Significant pairwise Fisher's exact test p-values are indicated after Bonferroni correction (* p<0.05, ** p<0.01, *** p<0.001).

Figure 2: Transcriptomic profiling of AML shows similarities between epigenetic subgroups.

A. Principal component analysis (PCA) of RNA-seq profiles from *IDH1*-, *IDH2*-, *TET2*-mutant, and *Other* AML samples (n=12) matched for additional genetic aberrations.

B. Volcano plot of differential gene expression between *IDH1*/*IDH2*- and *TET2*-mutant AML. For visualization, absolute log2 fold-change values exceeding 12 were capped at 12.

C. Hierarchically clustered heatmap of the 50 most differentially expressed genes between *IDH1*/*IDH2*- and *TET2*-mutant AML. Expression is shown as row-wise z-scores.

D. Gene set enrichment analysis (GSEA) of MSigDB Hallmark pathways comparing *IDH1*/*IDH2*- and *TET2*-mutant AML. Negative normalized enrichment scores (NES) indicate enrichment in *TET2*-mutant AML. FDR-adjusted p-values are indicated by color.

Figure 3: Distinct 5-hydroxymethylcytosine patterns in *IDH1*-, *IDH2*-, and *TET2*-mutant AML.

A. Principal component analysis (PCA) of gene-body 5hmC profiles from *IDH1*-, *IDH2*-, *TET2*-mutant, and *Other* AML (n=12) matched for additional genetic aberrations.

B. Global 5hmC mass fraction across *IDH1*-, *IDH2*-, *TET2*-mutant, and *Other* AML samples. Kruskal-Wallis p-value is indicated.

C. Hierarchically clustered heatmap of genes with significantly different gene-body 5hmC levels across mutation groups, identified by multi-group analysis of variance (ANOVA) with a false discovery rate (FDR) threshold of <0.10.

D. Volcano plot of differential gene-body 5hmC levels between *IDH1*/*IDH2*- and *TET2*-mutant AML.

E. Gene set enrichment analysis (GSEA) of MSigDB Hallmark pathways based on differential gene-body 5hmC levels between *IDH1*/*IDH2*- versus *TET2*-mutant AML. Positive normalized enrichment scores (NES) indicate enrichment in *IDH1*/*IDH2*-mutant AML, whereas negative NES indicate enrichment in *TET2*-mutant AML. FDR-adjusted p-values are indicated by color.

Figure 4: Integrated transcriptomic and 5hmC profiling of AML.

A. Scatter plot showing the relationship between differential RNA expression and differential gene-body 5hmC levels between *IDH1*/*IDH2*- and *TET2*-mutant AML. Each point represents one gene (n=14,219). Dashed lines indicate the absolute log2 fold-change threshold of 1. Points are colored according to significance and concordance between data modalities: yellow, significantly higher RNA expression and gene-body 5hmC levels in *IDH1*/*IDH2*-mutant AML; red, significantly higher RNA expression and gene-body 5hmC levels in *TET2*-mutant AML; green, significant differential RNA expression only; blue, significantly different gene-body 5hmC levels only; and grey, no significant difference in either modality.

B. Gene set enrichment analysis (GSEA) of MSigDB Hallmark pathways based on integrated RNA and 5hmC data comparing *IDH1*/*IDH2*- and *TET2*-mutant AML. Positive normalized enrichment scores (NES) indicates enrichment in *IDH1*/*IDH2*-mutant AML, whereas negative NES indicates enrichment in *TET2*-mutant AML. FDR-adjusted p-values are indicated by color.

C. External validation of selected Hallmark pathways identified in the integrated RNA/5hmC analysis using the TCGA AML cohort. Genes were ranked by a Welch t-statistic comparing *IDH1*/*IDH2*- and *TET2*-mutant AML. Positive normalized enrichment scores indicate enrichment in *IDH1*/*IDH2*-mutant AML, whereas negative scores indicate enrichment in *TET2*-mutant AML. Significance was defined as FDR<0.05.

D. Projection of AML RNA-seq samples onto a normal bone marrow reference map constructed from publicly available single-cell RNA-seq data aggregated into pseudobulk profiles.

Normal reference populations included hematopoietic stem cells (HSC), hematopoietic multipotent progenitors (MPP), lymphoid-primed multipotent progenitors (LMPP), megakaryocyte-erythroid progenitors (MEP), granulocyte-monocyte progenitors (GMP), monocyte-committed GMP (GMP-Mono), neutrophil-committed GMP (GMP-Neut), promonocytes, and monocytes.

E. Projection of AML 5hmC samples onto publicly available 5hmC sequencing data from normal hematopoietic populations. Normal reference populations included HSC, MPP, LMPP, common myeloid progenitors (CMP), GMP, MEP, and monocytes.

Figure 1

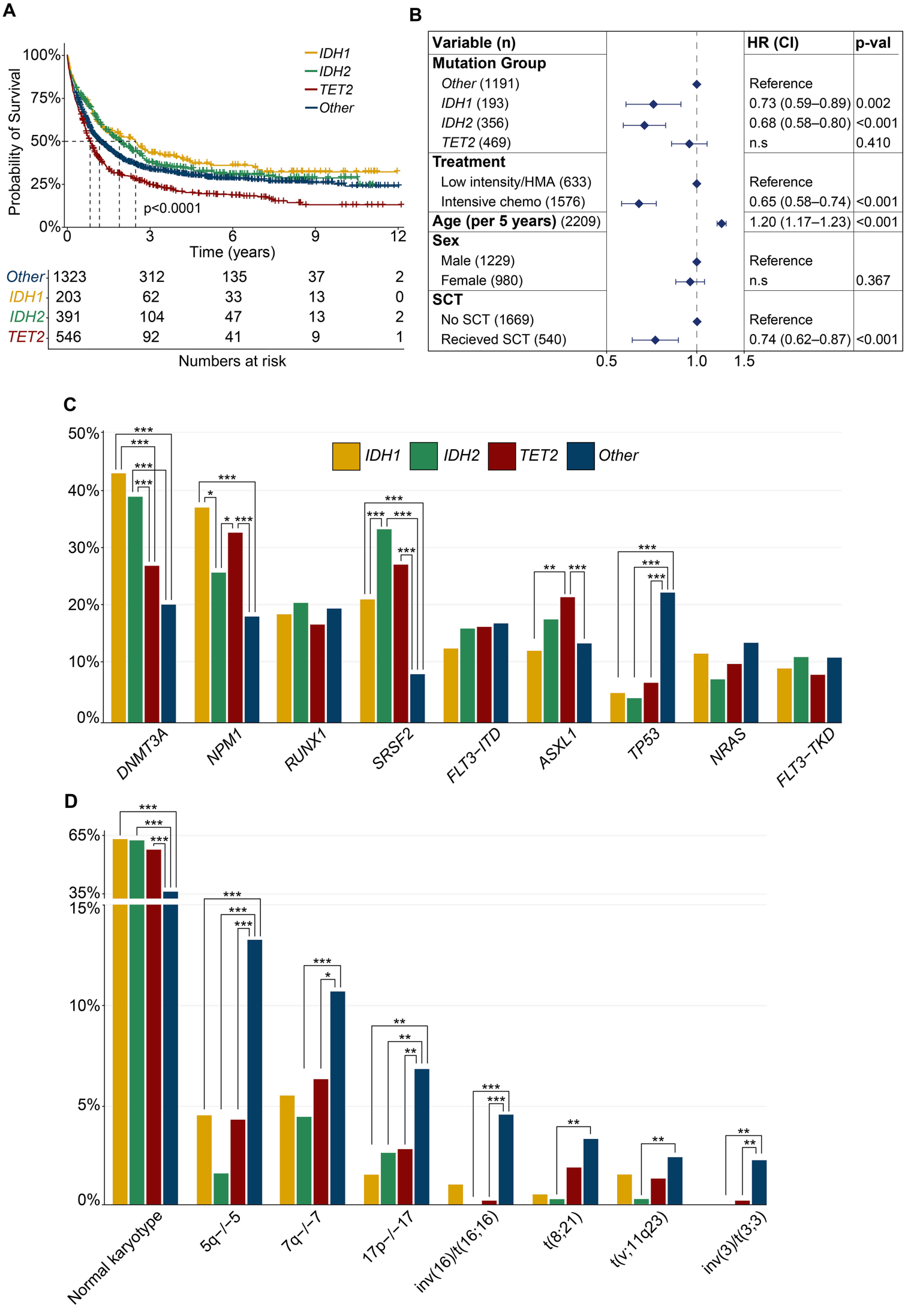

A
100%
75%
50%
25%
0%
Probability of Survival
IDH1
IDH2
TET2
Other
p<0.0001
0
3
6
9
12
Time (years)
Other 1323 312 135 37 2
IDH1 203 62 33 13 0
IDH2 391 104 47 13 2
TET2 546 92 41 9 1
Numbers at risk
B
Variable (n)
HR (CI)
p-val
Mutation Group
Other (1191) Reference
IDH1 (193) 0.73 (0.59–0.89) 0.002
IDH2 (356) 0.68 (0.58–0.80) <0.001
TET2 (469) n.s 0.410
Treatment
Low intensity/HMA (633) Reference
Intensive chemo (1576) 0.65 (0.58–0.74) <0.001
Age (per 5 years) (2209) 1.20 (1.17–1.23) <0.001
Sex
Male (1229) Reference
Female (980) n.s 0.367
SCT
No SCT (1669) Reference
Recieved SCT (540) 0.74 (0.62–0.87) <0.001
0.5
1.0
1.5
C
50%
40%
30%
20%
10%
0%
IDH1
IDH2
TET2
Other
DNMT3A
NPM1
RUNX1
SRSF2
FLT3−ITD
ASXL1
TP53
NRAS
FLT3−TKD
D
65%
35%
15%
10%
5%
0%
Normal karyotype
5q−/−5
7q−/−7
17p−/−17
inv(16)/t(16;16)
t(8;21)
t(v;11q23)
inv(3)/t(3;3)

Figure 2

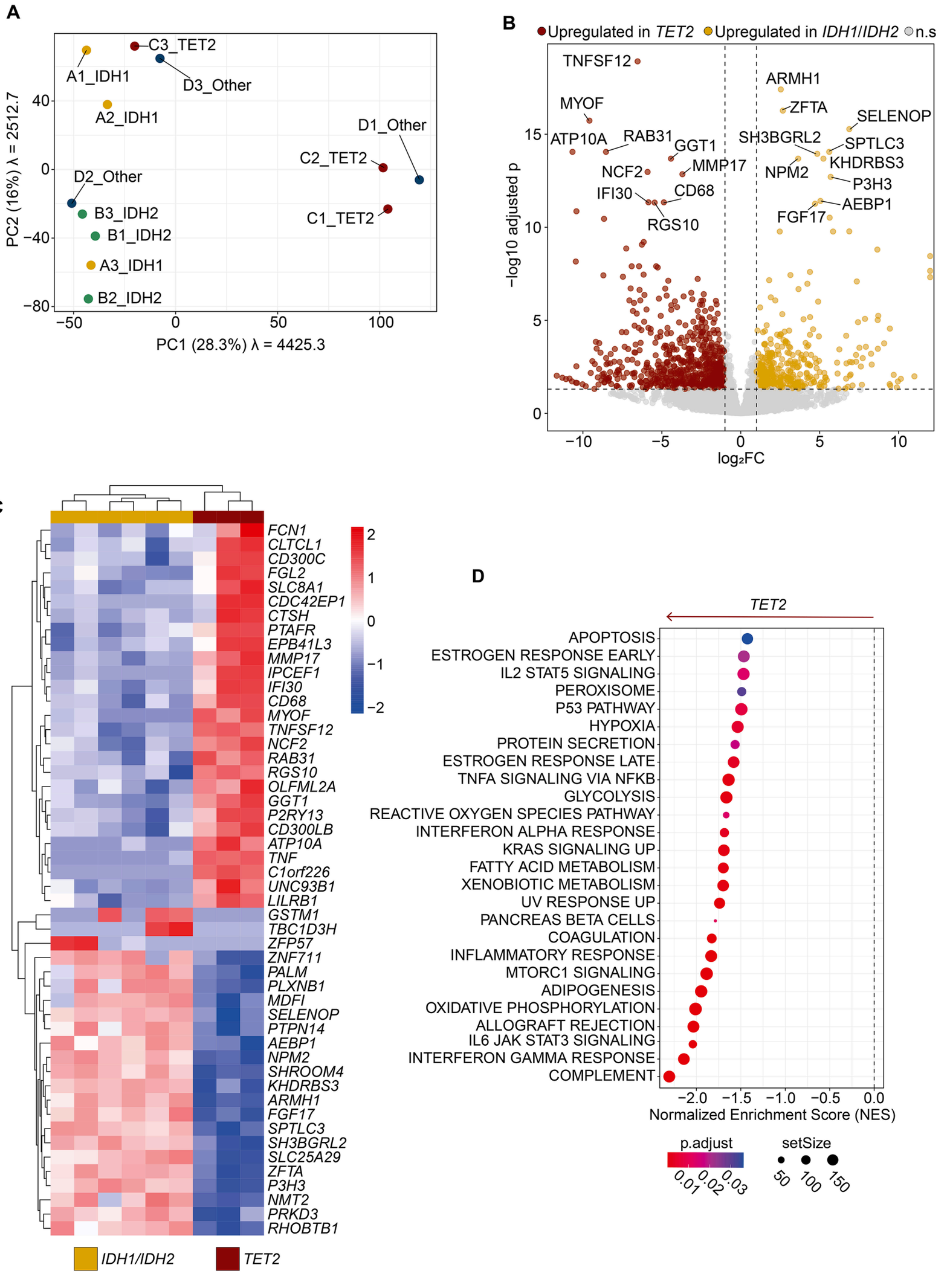
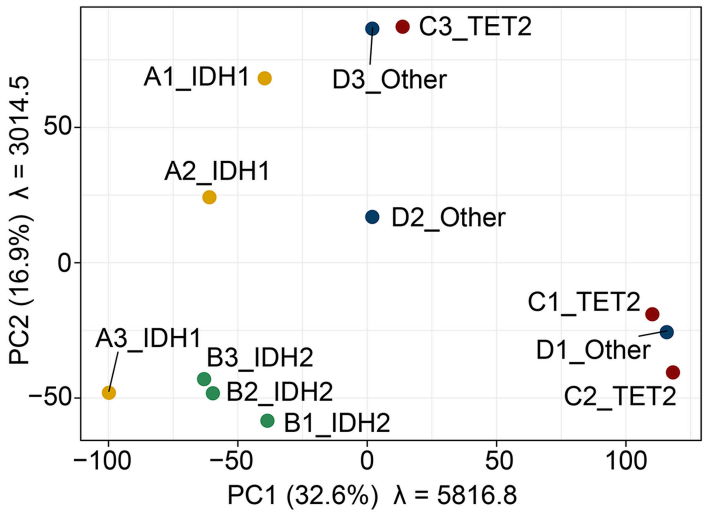
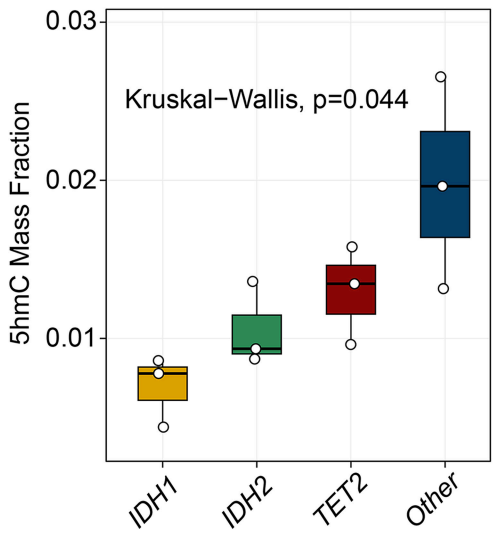
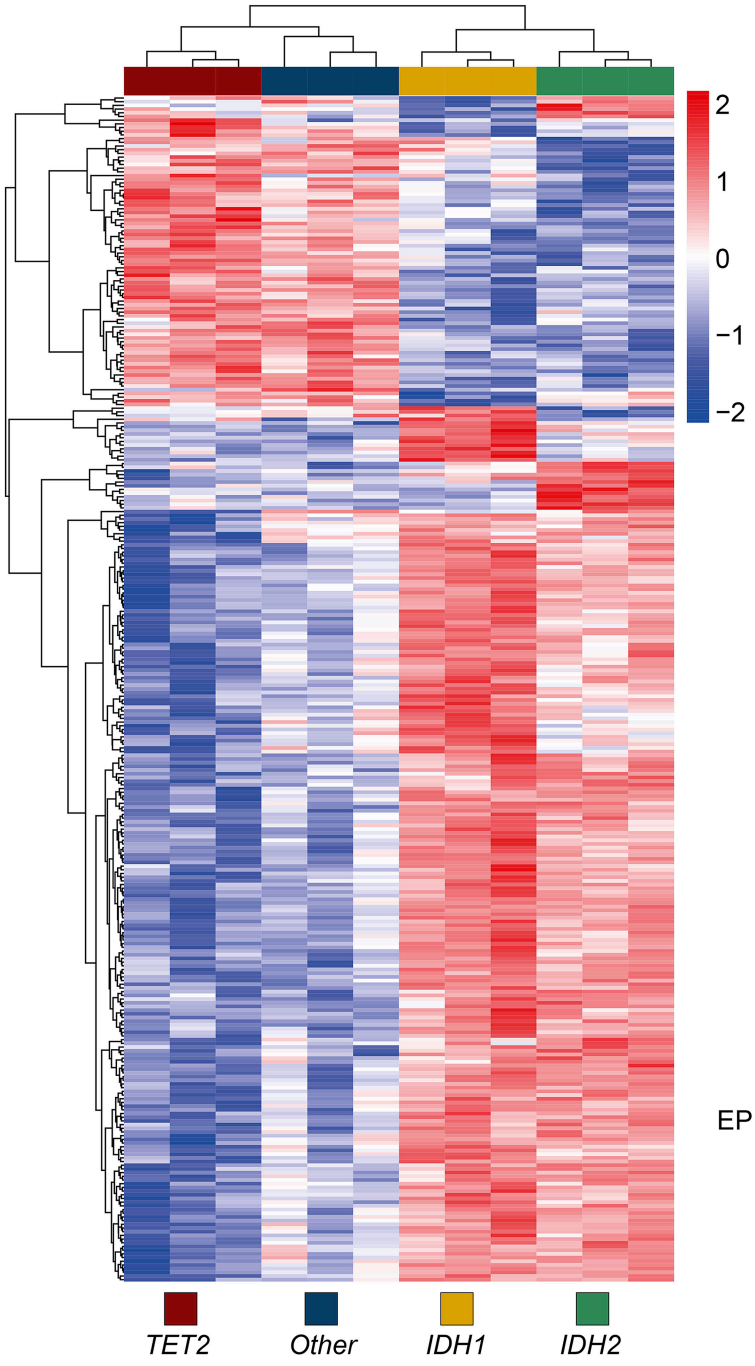
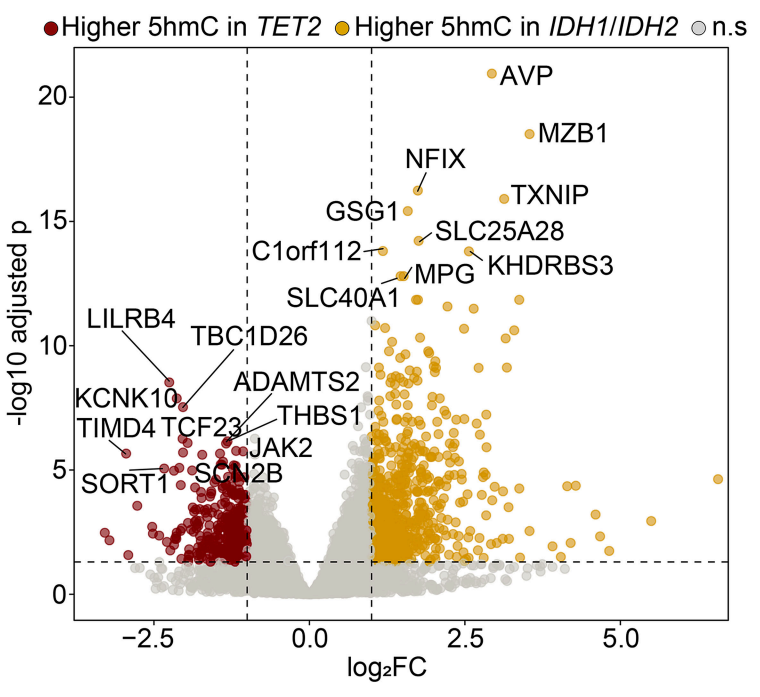

A
C3_TET2
A1_IDH1
D3_Other
A2_IDH1
D1_Other
C2_TET2
D2_Other
B3_IDH2
B1_IDH2
C1_TET2
A3_IDH1
B2_IDH2
PC2 (16%) λ = 2512.7
PC1 (28.3%) λ = 4425.3
B
Upregulated in TET2
Upregulated in IDH1/IDH2
n.s
TNFSF12
ARMH1
MYOF
ZFTA
SELENOP
ATP10A
RAB31
GGT1
SH3BGRL2
SPTLC3
NCF2
MMP17
NPM2
KHDRBS3
IFI30
CD68
P3H3
AEBP1
RGS10
FGF17
−log10 adjusted p
log₂FC
C
FCN1
CLTCL1
CD300C
FGL2
SLC8A1
CDC42EP1
CTSH
PTAFR
EPB41L3
MMP17
IPCEF1
IFI30
CD68
MYOF
TNFSF12
NCF2
RAB31
RGS10
OLFML2A
GGT1
P2RY13
CD300LB
ATP10A
TNF
C1orf226
UNC93B1
LILRB1
GSTM1
TBC1D3H
ZFP57
ZNF711
PALM
PLXNB1
MDFI
SELENOP
PTPN14
AEBP1
NPM2
SHROOM4
KHDRBS3
ARMH1
FGF17
SPTLC3
SH3BGRL2
SLC25A29
ZFTA
P3H3
NMT2
PRKD3
RHOBTB1
IDH1/IDH2
TET2
D
TET2
APOPTOSIS
ESTROGEN RESPONSE EARLY
IL2 STAT5 SIGNALING
PEROXISOME
P53 PATHWAY
HYPOXIA
PROTEIN SECRETION
ESTROGEN RESPONSE LATE
TNFA SIGNALING VIA NFKB
GLYCOLYSIS
REACTIVE OXYGEN SPECIES PATHWAY
INTERFERON ALPHA RESPONSE
KRAS SIGNALING UP
FATTY ACID METABOLISM
XENOBIOTIC METABOLISM
UV RESPONSE UP
PANCREAS BETA CELLS
COAGULATION
INFLAMMATORY RESPONSE
MTORC1 SIGNALING
ADIPOGENESIS
OXIDATIVE PHOSPHORYLATION
ALLOGRAFT REJECTION
IL6 JAK STAT3 SIGNALING
INTERFERON GAMMA RESPONSE
COMPLEMENT
Normalized Enrichment Score (NES)
p.adjust
setSize

Figure 3


A
C3_TET2
A1_IDH1
D3_Other
A2_IDH1
D2_Other
C1_TET2
A3_IDH1
B3_IDH2
D1_Other
B2_IDH2
C2_TET2
B1_IDH2
PC2 (16.9%) λ = 3014.5
PC1 (32.6%) λ = 5816.8


B
Kruskal−Wallis, p=0.044
5hmC Mass Fraction
IDH1
IDH2
TET2
Other


C
TET2
Other
IDH1
IDH2


D
Higher 5hmC in TET2
Higher 5hmC in IDH1/IDH2
n.s
AVP
MZB1
NFIX
TXNIP
GSG1
SLC25A28
C1orf112
KHDRBS3
MPG
SLC40A1
LILRB4
TBC1D26
KCNK10
ADAMTS2
THBS1
TIMD4
TCF23
JAK2
SORT1
SCN2B
-log10 adjusted p
log₂FC

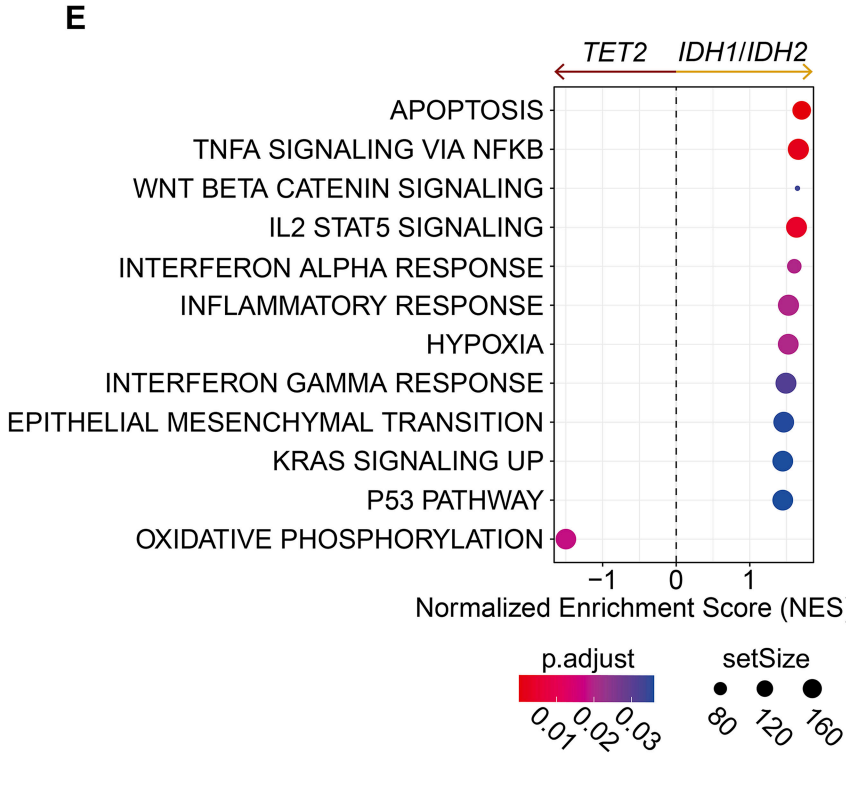
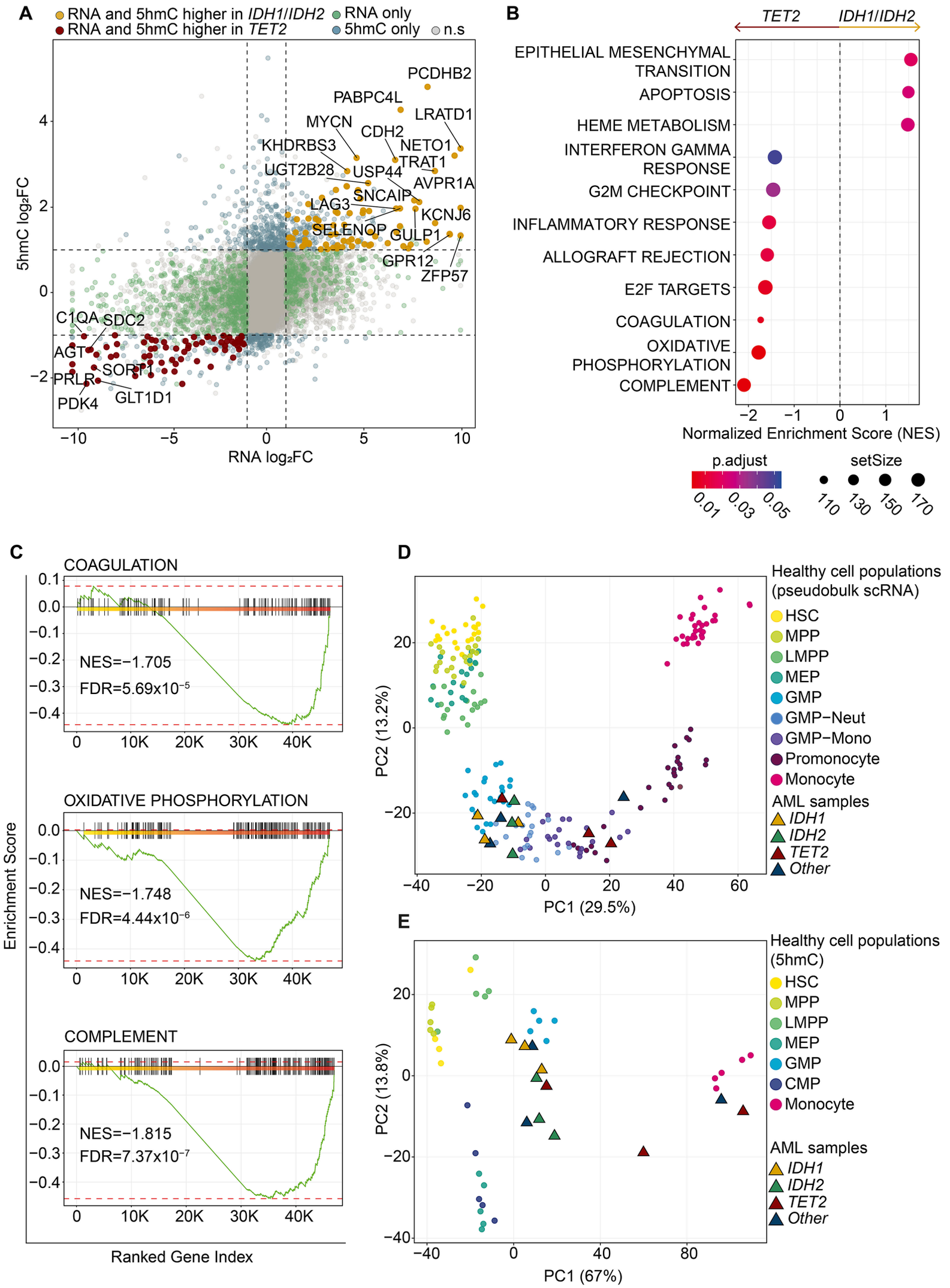

E
TET2
IDH1/IDH2
APOPTOSIS
TNFA SIGNALING VIA NFKB
WNT BETA CATENIN SIGNALING
IL2 STAT5 SIGNALING
INTERFERON ALPHA RESPONSE
INFLAMMATORY RESPONSE
HYPOXIA
INTERFERON GAMMA RESPONSE
EPITHELIAL MESENCHYMAL TRANSITION
KRAS SIGNALING UP
P53 PATHWAY
OXIDATIVE PHOSPHORYLATION
Normalized Enrichment Score (NES)
p.adjust
setSize

Figure 4


A
RNA and 5hmC higher in IDH1/IDH2
RNA and 5hmC higher in TET2
RNA only
5hmC only
n.s
PCDHB2
PABPC4L
LRATD1
MYCN
CDH2
KHDRBS3
NETO1
TRAT1
UGT2B28
USP44
AVPR1A
SNCAIP
LAG3
KCNJ6
SELENOP
GULP1
GPR12
ZFP57
C1QA
SDC2
AGT
SORT1
PRLR
GLT1D1
PDK4
5hmC log₂FC
RNA log₂FC
B
TET2
IDH1/IDH2
EPITHELIAL MESENCHYMAL TRANSITION
APOPTOSIS
HEME METABOLISM
INTERFERON GAMMA RESPONSE
G2M CHECKPOINT
INFLAMMATORY RESPONSE
ALLOGRAFT REJECTION
E2F TARGETS
COAGULATION
OXIDATIVE PHOSPHORYLATION
COMPLEMENT
Normalized Enrichment Score (NES)
p.adjust
setSize
C
COAGULATION
NES=−1.705
FDR=5.69x10⁻⁵
OXIDATIVE PHOSPHORYLATION
NES=−1.748
FDR=4.44x10⁻⁶
COMPLEMENT
NES=−1.815
FDR=7.37x10⁻⁷
Enrichment Score
Ranked Gene Index
D
Healthy cell populations (pseudobulk scRNA)
HSC
MPP
LMPP
MEP
GMP
GMP−Neut
GMP−Mono
Promonocyte
Monocyte
AML samples
IDH1
IDH2
TET2
Other
PC1 (29.5%)
PC2 (13.2%)
E
Healthy cell populations (5hmC)
HSC
MPP
LMPP
MEP
GMP
CMP
Monocyte
AML samples
IDH1
IDH2
TET2
Other
PC1 (67%)
PC2 (13.8%)

SUPPLEMENTARY APPENDIX - Distinct clinical and molecular profiles of acute myeloid leukemia defined by mutations in *IDH1*, *IDH2*, and *TET2*

SUPPLEMENTARY MATERIALS AND METHODS

Fluorescence-Activated Cell Sorting

For florescence-activated cell sorting (FACS) of AML blasts, cells were thawed and stained in PBS buffer containing 2% FBS and 2 mM EDTA, using propidium iodide (PI, Thermo Fisher Scientific, USA, Cat#P3566) as a viability dye and the following antibodies: V450-conjugated anti-CD45 antibody (BD Bioscience, USA, Cat# 560367), PE-Cy5 conjugated anti-CD19 antibody (BD Bioscience, Cat# 555414), PE-Cy5 conjugated anti-CD20 antibody (BD Bioscience, USA, Cat# 555624), APC-Cy7 conjugated anti-CD3 antibody (BD Bioscience, Cat# 557832). Cells were kept in the dark for 20 minutes at 4°C, washed and resuspended in buffer containing 0.5 µg/ml propidium iodide (PI). For each sample, 400 000 viable AML blasts were defined as $PI^-CD3^-CD19^-CD20^-CD45^{dim}$ and sorted into two Eppendorf tubes and post-sort purity was analyzed by flow cytometry. Cells were washed in PBS and resuspended in DNA/RNA Shield storage buffer (Zymo, USA, Cat# R1100-250) and snap-frozen on dry ice and stored at -80°C until analysis.

RNA and 5hmC downstream analyses

For each differential comparison, the DESeq2 design was specified according to the contrast being tested. Significance was defined as a Benjamini-Hochberg (BH)-adjusted p-value <0.05 and an absolute log2 fold change ($|\log_2 FC| > 1$. Volcano plots were created using ggplot2 (v.4.0.0). Heatmaps were created with pheatmap (v.1.0.13)[1] using normalized DESeq2 counts with row-wise z-score scaling. Principal component analysis (PCA) was performed on variance-stabilized expression matrices after removal of zero-variance features. Silhouette scores were calculated based on Euclidean distances in the two-dimensional PC1-PC2 space. For each sample, the nearest alternative mutation group was defined as the group with the smallest mean Euclidean distance to that sample in the PC1-PC2 space, excluding its own mutation group. Global 5hmC mass fractions were compared across mutation groups using the Kruskal-Wallis test. Multi-group analysis of variance (ANOVA) was performed to identify genes with significantly different 5hmC levels across mutation groups. For each gene, an ANOVA F-statistic and corresponding p-value were calculated. Multiple testing correction was applied using the BH

false discovery rate (FDR) method, with an FDR ≤0.10 considered significant. Preranked Gene Set Enrichment Analysis (GSEA) was performed using clusterProfiler (v.4.17.0)[2] with MSigDB Hallmark gene sets. Genes were ranked according to sign(log_2FC) * (-log_{10}(p-value)), using unadjusted DESeq2 gene-level p-values. Gene identifiers were mapped to Entrez IDs where required, and duplicated identifiers were collapsed by their mean rank. Multiple testing was controlled using BH adjustment, with FDR <0.05 considered significant.

To integrate transcriptional and epigenetic data, differential results from the *IDH1*/*IDH2*- versus *TET2*-mutant AML comparisons were matched by gene symbols. Concordance between RNA and 5hmC effect sizes was assessed using Spearman rank correlation (Spearman ρ) across genes represented in both modalities. Genes were considered concordant if they were significant in both modalities (FDR <0.05) and showed log2FC values in the same direction. For integrated pathway analysis, RNA and 5hmC log2FC values were standardized separately to z-scores. An integrated gene-ranking score was calculated as z(log2FC(RNA)) + z(log2FC(5hmC)) and used for preranked GSEA of MSigDB Hallmark gene sets.

External datasets

Publicly available single-cell RNA-seq data from normal hematopoietic populations as described in Zeng *et al*. 2025[3], comprising datasets from Oetjen *et al*. 2018[4], Human Cell Atlas (HCA) Bone Marrow dataset, Mende *et al.* 2022[5], Setty *et al*. 2019[6], Ainciburu *et al.* 2022[7] and Granja *et al*. 2019[8], were retrieved from the CZ CELLxGENE collection "A (Balanced) Bone Marrow Reference Map of Hematopoietic Development." The single-cell RNA-seq data were restricted to normal bone marrow cells belonging to predefined hematopoietic populations. Cell-type annotations were harmonized into healthy hematopoietic reference groups, including hematopoietic stem cells (HSC), multipotent progenitors (MPP), lymphoid-primed multipotent progenitors (LMPP), megakaryocyte-erythroid progenitors (MEP), granulocyte-monocyte progenitors (GMP), monocyte-committed GMP (GMP-Mono), neutrophil-committed GMP (GMP-Neut), promonocytes, and monocytes. Pseudobulk profiles were generated separately for each donor and hematopoietic population by summing raw gene counts across cells. Donor cell-type combinations represented by fewer than 50 cells were excluded. Duplicate gene symbols were collapsed by summing counts, and pseudobulk libraries were normalized to counts per million and log2-transformed as log2(CPM + 1) prior to reference projection. For

the single-cell RNA-seq reference, study-associated variation was adjusted using limma (v.3.66.0) prior to PCA while preserving differences between hematopoietic reference populations. Single-cell RNA-seq data were processed in Python (v.3.13.5) using Scanpy (v.1.12.1)[9] for handling AnnData objects, pandas (v.2.3.3) for data manipulation, NumPy (v.2.4.6)[10] for numerical operations, and SciPy (v.1.18.0)[11] for sparse matrix handling.

Unstranded bulk RNA-seq data from normal hematopoietic populations previously generated by Corces *et al*. 2016[12] were downloaded from the NCBI Gene Expression Omnibus (GEO) under accession number GSE74246. Publicly available 5hmC sequencing data from normal hematopoietic populations described in Nakauchi *et al*. 2022[13] were obtained from GEO under accession number GSE198908. Reference populations included HSC, MPP, LMPP, common myeloid progenitors (CMP), GMP, MEP, and monocytes. Normal reference counts were TMM-normalized and transformed to logCPM values. For each reference analysis, genes shared between the normal reference and the corresponding AML dataset were identified, and the 3,000 genes with the highest variance in the normal reference were selected. PCA was performed on the normal reference only, with features centered and scaled, and AML samples were subsequently projected into the PCA space defined by the normal reference. Mean profiles were calculated for each normal hematopoietic population, and similarity between individual AML samples and the reference populations was quantified using Spearman rank correlation across the selected genes. The nearest normal population was defined as the population with the highest correlation.

Selected Hallmark pathways identified in the integrated RNA/5hmC analysis were externally evaluated using transcriptomic data from The Cancer Genome Atlas (TCGA) AML cohort.[14] Samples with mutations in *IDH1* or *IDH2* in the absence of *TET2* mutation were classified as *IDH1*/*IDH2*-mutant, whereas samples with *TET2* mutation in the absence of *IDH1*/*IDH2* mutations were classified as *TET2*-mutant. Fusion-positive samples were excluded from the primary analysis. TCGA gene expression data were obtained as TPM values and log2-transformed as log2(TPM + 1). Genes were mapped to gene symbols with duplicated symbols collapsed. Genes were ranked using a Welch t-statistic comparing the *IDH1*/*IDH2*- and TET2-mutant groups, and preranked GSEA was performed using fgseaMultilevel for the prespecified MSigDB Hallmark pathways. FDR <0.05 was considered significant.

Swedish AML registry statistical analyses

Clinical, molecular, and cytogenetic characteristics were summarized by mutation group (*IDH1*, *IDH2*, *TET2*, and *Other*). Continuous variables were reported as medians with interquartile range (Q1, Q3) and compared using the Kruskal-Wallis rank-sum test. Where applicable, post hoc pairwise comparisons were performed using two-sided Wilcoxon rank-sum test (Mann-Whitney U test). Categorical variables were compared using Pearson's chi-squared test. For treatment intensity (intensive chemotherapy, low intensity/HMA, or palliative care) separate row-specific chi-squared tests were performed. Co-mutation frequencies were compared both globally and pairwise using Pearson's chi-square test. Cytogenetic abnormalities were analyzed using Fisher's exact test for global comparisons, followed by two-sided Fisher's exact tests for pairwise comparisons. P-values from multiple pairwise or row-specific comparisons were adjusted using the Bonferroni method.

Overall survival was analyzed using multivariable Cox proportional hazards regression to estimate the associations between mutation groups, treatment and overall survival. For the main cohort analyses, overall survival was calculated from the date of diagnosis to death, with patients alive at last follow-up censored at that date. The model included mutation group (*IDH1*, *IDH2*, *TET2*, and *Other*), treatment (low intensity/HMA or intensive chemotherapy), age (per 5-year increments), sex, and allogeneic stem cell transplantation (SCT). SCT was modeled as a time-dependent covariate when transplantation dates were available. The proportional hazards assumption was assessed using Schoenfeld residuals. Univariable Cox proportional hazards regression was used for predefined subgroup analyses to assess the association of *NPM1* mutation status with overall survival among patients receiving intensive chemotherapy, and to compare HMA plus venetoclax with HMA monotherapy among patients receiving low-intensity treatment. For these treatment-specific subgroup analyses, survival time was calculated from treatment initiation to death, with patients alive at last follow-up censored at that date. Cox models were fitted using the Efron method for tied event times. Hazard ratios (HR) and 95% confidence intervals (CI) were obtained by exponentiating the model coefficients and corresponding confidence limits. Patients with missing data for variables required in a given analysis were excluded from that analysis. All statistical analyses were performed using R (v.4.5.1) with the following packages: dplyr (v.1.1.4)[15] for data manipulation, rstatix (v.0.7.2)[16] for statistical analyses, and ggplot2 (v.4.0.0)[17] for visualization.

REFERENCES

1. Kolde R. pheatmap: Pretty Heatmaps. 1.0.13 ed, 2025.
2. Xu S, Hu E, Cai Y, et al. Using clusterProfiler to characterize multiomics data. Nat Protoc. 2024;19(11):3292–3320.
3. Zeng AGX, Iacobucci I, Shah S, et al. Single-cell Transcriptional Atlas of Human Hematopoiesis Reveals Genetic and Hierarchy-Based Determinants of Aberrant AML Differentiation. Blood Cancer Discov. 2025;6(4):307–324.
4. Oetjen KA, Lindblad KE, Goswami M, et al. Human bone marrow assessment by single-cell RNA sequencing, mass cytometry, and flow cytometry. JCI Insight. 2018;3(23):e124928.
5. Mende N, Bastos HP, Santoro A, et al. Unique molecular and functional features of extramedullary hematopoietic stem and progenitor cell reservoirs in humans. Blood. 2022;139(23):3387–3401.
6. Setty M, Kiseliovas V, Levine J, Gayoso A, Mazutis L, Pe'er D. Characterization of cell fate probabilities in single-cell data with Palantir. Nat Biotechnol. 2019;37(4):451–460.
7. Ainciburu M, Ezponda T, Berastegui N, et al. Uncovering perturbations in human hematopoiesis associated with healthy aging and myeloid malignancies at single-cell resolution. Elife. 20230111 ed, 2023.
8. Granja JM, Klemm S, McGinnis LM, et al. Single-cell multiomic analysis identifies regulatory programs in mixed-phenotype acute leukemia. Nat Biotechnol. 2019;37(12):1458–1465.
9. Wolf FA, Angerer P, Theis FJ. SCANPY: large-scale single-cell gene expression data analysis. Genome Biol. 2018;19(1):15.
10. Harris CR, Millman KJ, van der Walt SJ, et al. Array programming with NumPy. Nature. 2020;585(7825):357–362.
11. Virtanen P, Gommers R, Oliphant TE, et al. SciPy 1.0: fundamental algorithms for scientific computing in Python. Nat Methods. 2020;17(3):261–272.
12. Corces MR, Buenrostro JD, Wu B, et al. Lineage-specific and single-cell chromatin accessibility charts human hematopoiesis and leukemia evolution. Nat Genet. 2016;48(10):1193–1203.
13. Nakauchi Y, Azizi A, Thomas D, et al. The Cell Type-Specific 5hmC Landscape and Dynamics of Healthy Human Hematopoiesis and TET2-Mutant Preleukemia. Blood Cancer Discov. 2022;3(4):346–367.
14. Cancer Genome Atlas Research N, Ley TJ, Miller C, et al. Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. N Engl J Med. 2013;368(22):2059–2074.
15. Wickham H, François R, Henry L, Müller K, Vaughan D. dplyr: A Grammar of Data Manipulation. 2023.
16. Kassambara A. rstatix: Pipe-Friendly Framework for Basic Statistical Tests. 2023.
17. Wickham H. ggplot2: Elegant Graphics for Data Analysis: Springer-Verlag New York, 2016.

SUPPLEMENTARY FIGURE LEGENDS

Supplementary Figure 1: Clinical characteristics of epigenetic subgroups of AML.

A. Distribution of AML patients in the Swedish AML registry. Co-mutated* (≥2) indicates co-occurring mutations in two or more of *IDH1*, *IDH2*, and *TET2*.

B. Age distributions across mutation groups. Violin plot shows full distribution with embedded boxplots indicating median and quartiles. Pairwise Wilcoxon rank-sum p-values are indicated after Bonferroni correction (*** p<0.001).

C. Distribution of key clinical parameters across mutation groups. Median values are indicated by dotted lines, and healthy reference ranges are shown in grey.

D. Kaplan-Meier estimates of overall survival comparing *IDH1*/*IDH2*-mutant and *TET2*-mutant AML. Significance is indicated by log-rank p<0.0001.

E. Multivariable Cox proportional hazards regression showing hazard ratios (HR) with 95% confidence interval, adjusting for mutation group (*IDH1*/*IDH2*-mutant as one group compared to *TET2*-mutant), treatment intensity, age (5-year increments), sex, and stem cell transplantation (SCT), with SCT modeled as a time-dependent covariate.

Supplementary Figure 2: Survival analyses in mutational subgroups.

A-B. Kaplan-Meier estimates of overall survival of AML patients receiving low-intensity treatment consisting of HMA with or without venetoclax, results shown separately for *IDH2*-mutant (A) and *IDH1*-mutant (B) AML. P-values are from univariable Cox proportional hazards regression.

C-F. Kaplan-Meier estimates of overall survival according to *NPM1* mutation status among patients receiving intensive chemotherapy, shown separately for *IDH1*-mutant (C), *IDH2*-mutant (D), *TET2*-mutant (E), and *Other* AML (F) groups. P-values are from univariable Cox proportional hazards regression.

G. Distribution of *TET2* variant allele frequency (VAF) among patients with *TET2*-mutant AML. The median VAF is indicated as a dotted line.

H. Kaplan-Meier estimates of overall survival according to *TET2* VAF, comparing patients with VAF <45% and VAF >55%. Patients with VAF between 45% and 55% were excluded. The log-rank p-value is indicated.

Supplementary Figure 3: FACS isolation of AML blast cells for RNA sequencing and gene-body 5hmC enrichment sequencing.

A. Representative FACS gating strategy used to isolate AML blasts.

B. Post-sort purity analyses of FACS-isolated AML blasts from all samples, showing blast purity ≥99%.

Supplementary Figure 4: Differential gene expression analyses.

A. Volcano plot of differential gene expression between *IDH1*/*IDH2*/*TET2*-mutant and *Other* AML samples.

B. Hierarchically clustered heatmap of the 11 differentially expressed genes between *IDH1*/*IDH2*/*TET2*-mutant and *Other* AML samples. Expression is shown as row-wise z-scores.

C. Volcano plot of differential gene expression comparing *IDH1*/*IDH2*-mutant and *Other* AML samples.

D. Volcano plot of differential gene expression comparing *Other* and *TET2*-mutant AML samples.

Supplementary Figure 5: Differential gene-body 5hmC patterns and comparison with normal hematopoietic reference populations.

A. Volcano plot of differential gene-body 5hmC levels between *IDH1*/*IDH2*-mutant and *Other* AML samples.

B. Volcano plot of differential gene-body 5hmC levels between *Other* and *TET2*-mutant AML samples.

C. Scatter plot showing gene-wise RNA and 5hmC log2 fold changes for genes represented in both datasets (n=14,219). The fitted linear trend is shown in blue, and the Spearman correlation coefficient and p-value are indicated.

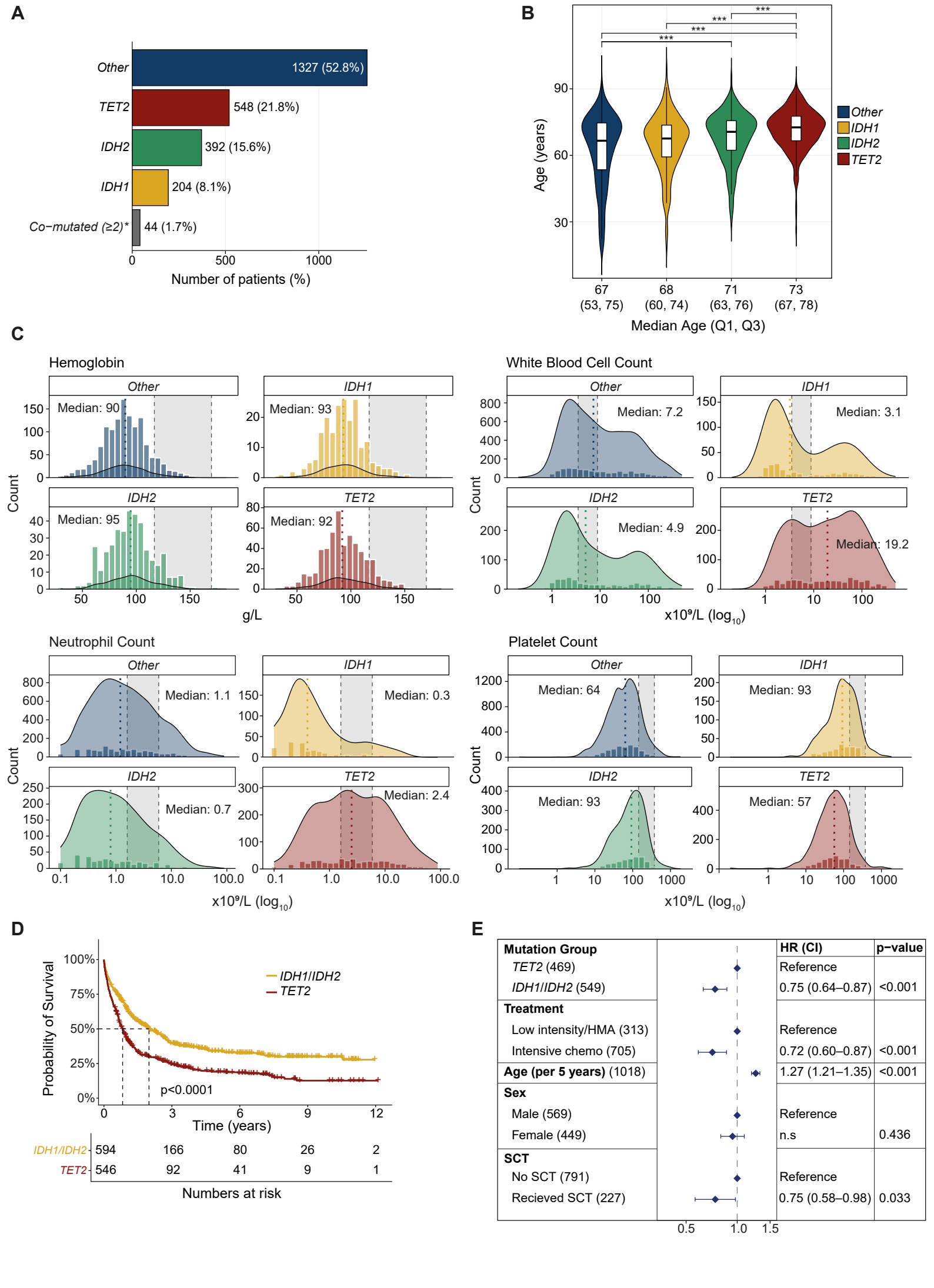
D. Heatmap of Spearman correlations between AML RNA-seq samples and normal bone marrow single-cell pseudobulk reference populations. Rows represent AML samples and columns represent normal hematopoietic populations.

E. Projection of AML RNA-seq samples onto a publicly available normal hematopoietic bulk RNA-seq reference. Normal reference populations included HSC, MPP, LMPP, CMP, GMP, MEP, and monocytes.

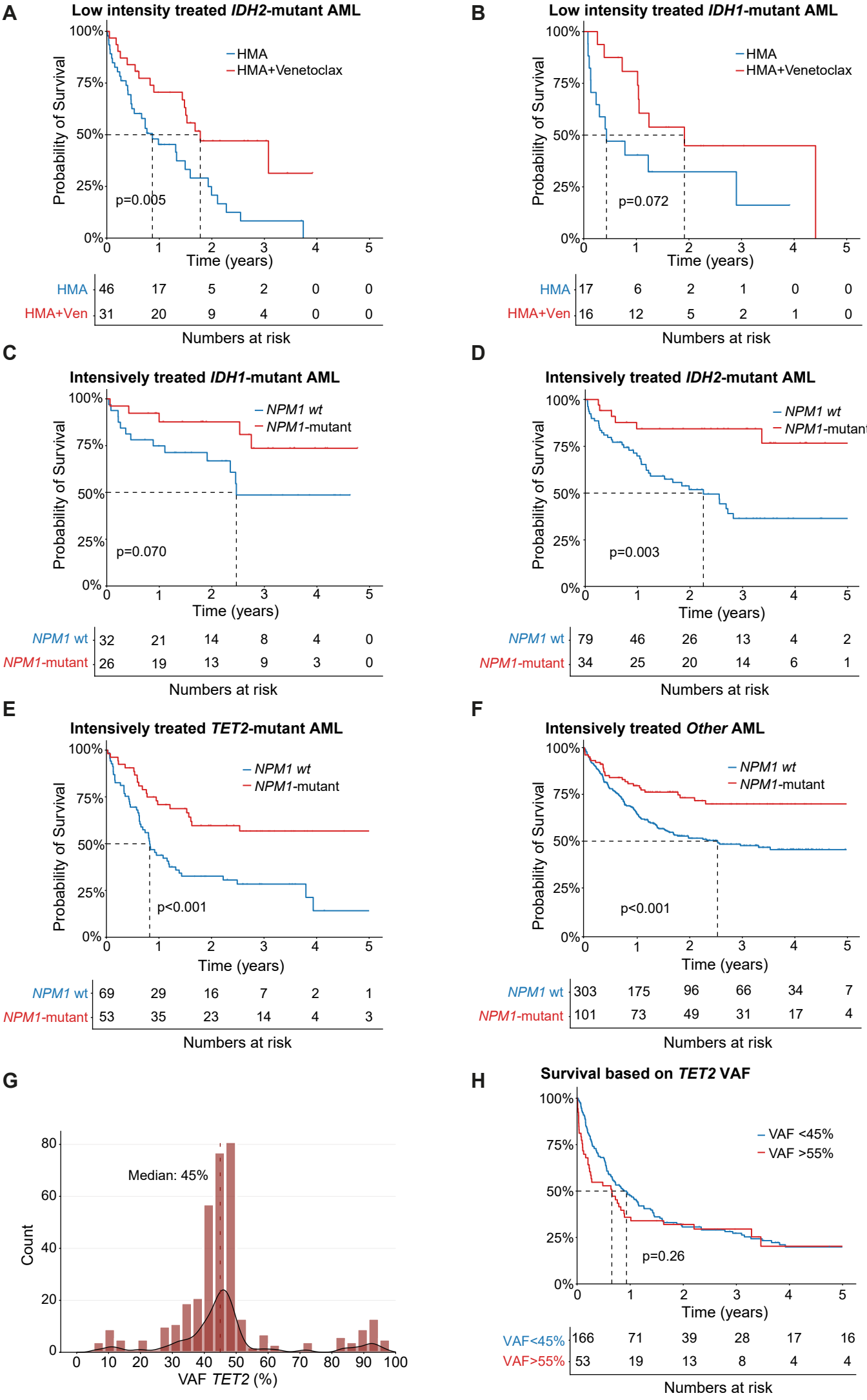
F. Heatmap of Spearman correlations between AML RNA-seq samples and normal hematopoietic bulk RNA-seq reference populations. Rows represent AML samples and columns represent normal hematopoietic populations.

G. Heatmap of Spearman correlations between AML 5hmC samples and normal hematopoietic 5hmC reference populations. Rows represent AML samples and columns represent normal hematopoietic populations.

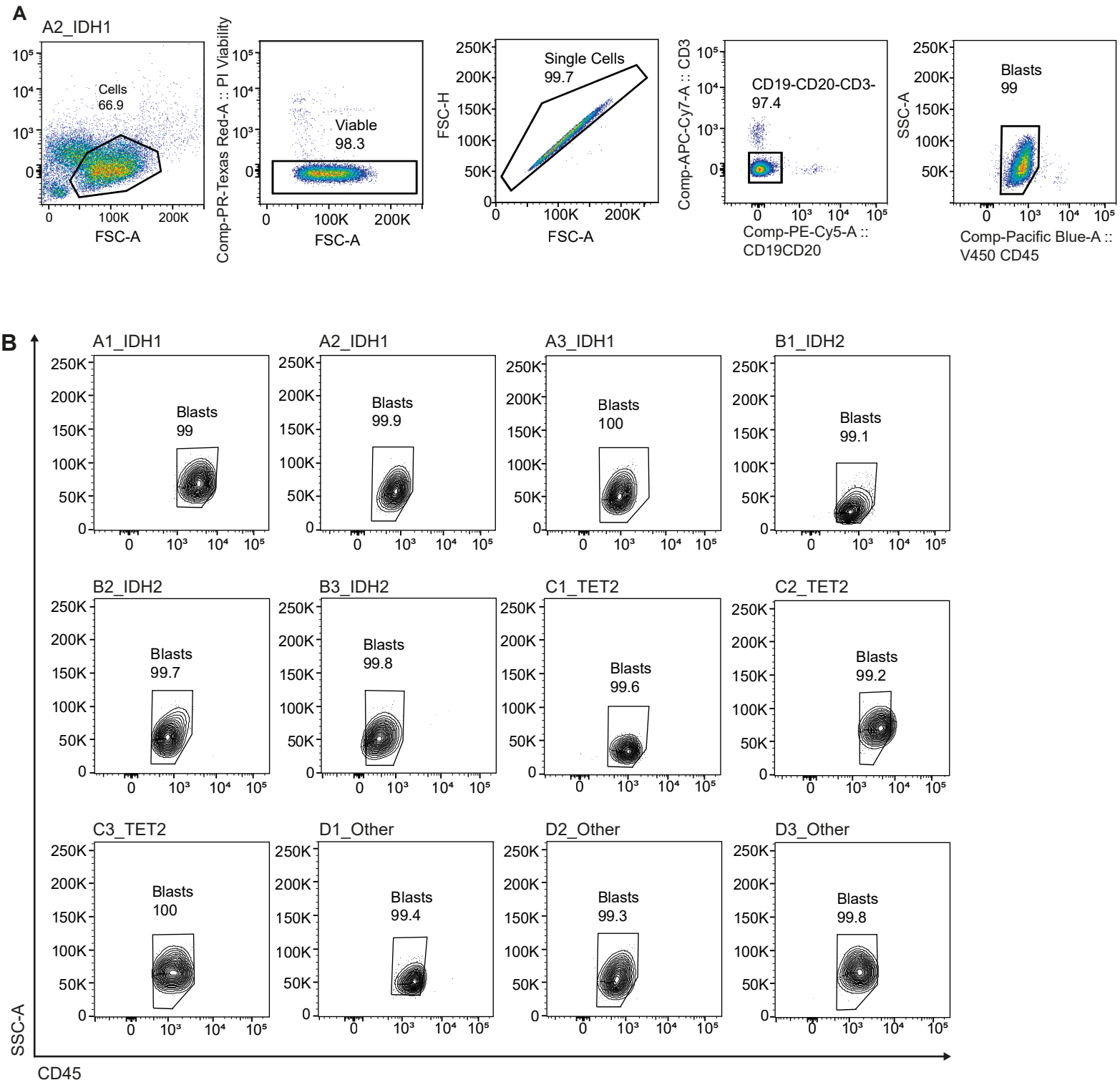
Supplementary Figure 1



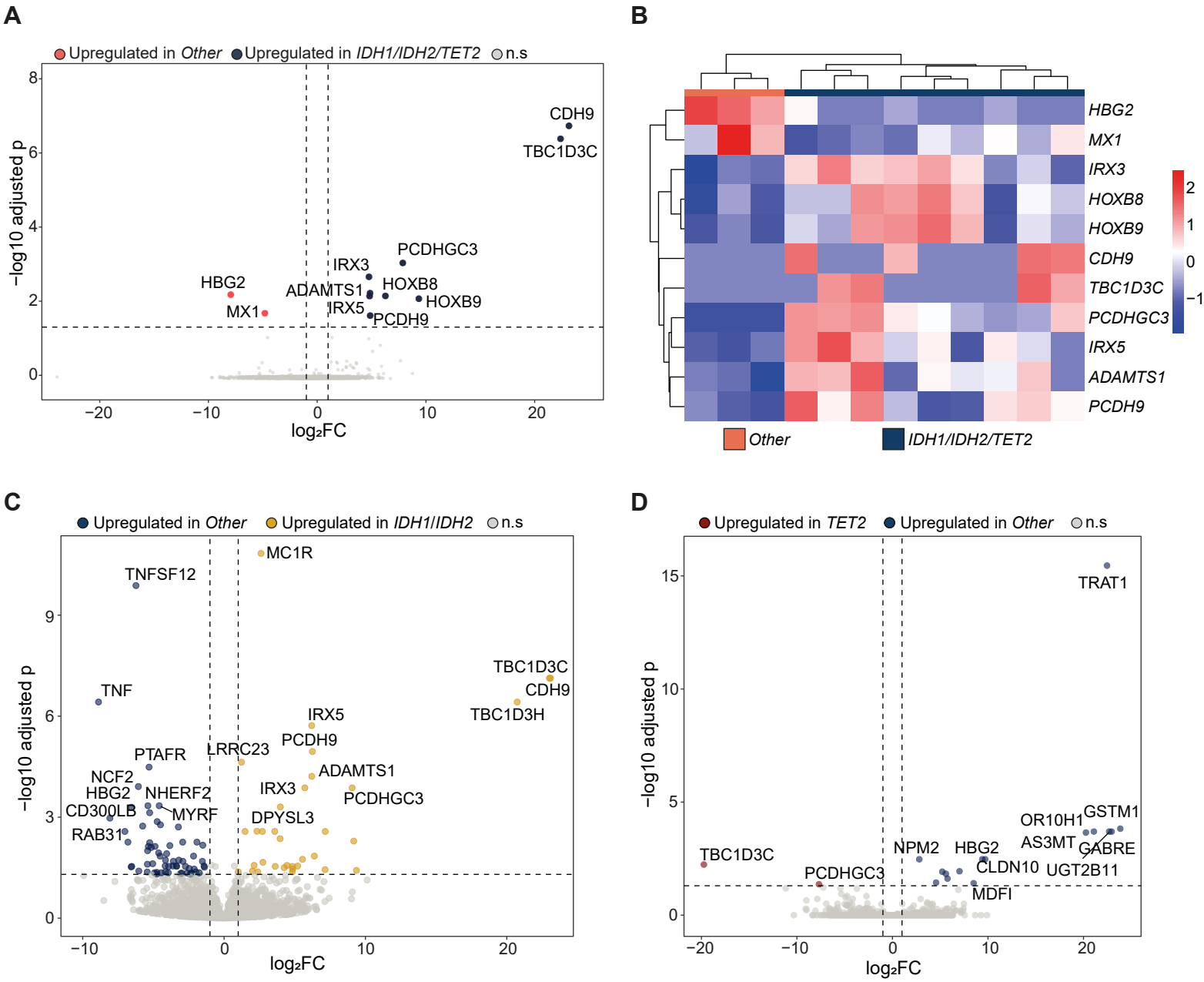
Supplementary Figure 2



Supplementary Figure 3



Supplementary Figure 4



Supplementary Figure 5

