Single-cell analysis of transcription factor regulatory networks reveals molecular basis for subtype-specific dysregulation in acute myeloid leukemia

Ruixia Sun[a,b], Lina Sun[a,b], Xiaowei Xie[a,b], Xuan Li[a,b], Peng Wu[a,b,*], Lu Wang[a,b,*], Ping Zhu[a,b,*]

[a]State Key Laboratory of Experimental Hematology, National Clinical Research Center for Blood Diseases, Haihe Laboratory of Cell Ecosystem, Institute of Hematology & Blood Diseases Hospital, Chinese Academy of Medical Sciences & Peking Union Medical College, Tianjin, China; [b]Center for Stem Cell Medicine and Department of Stem Cell & Regenerative Medicine, Chinese Academy of Medical Sciences & Peking Union Medical College, Tianjin, China

Abstract

Highly heterogeneous acute myeloid leukemia (AML) exhibits dysregulated transcriptional programs. Transcription factor (TF) regulatory networks underlying AML subtypes have not been elucidated at single-cell resolution. Here, we comprehensively mapped malignancy-related TFs activated in different AML subtypes by analyzing single-cell RNA sequencing data from AMLs and healthy donors. We first identified six modules of regulatory networks which were prevalently dysregulated in all AML patients. AML subtypes featured with different malignant cellular composition possessed subtype-specific regulatory TFs associated with differentiation suppression or immune modulation. At last, we validated that ERF was crucial for the development of hematopoietic stem/progenitor cells by performing loss- and gain-of-function experiments in zebrafish embryos. Collectively, our work thoroughly documents an abnormal spectrum of transcriptional regulatory networks in AML and reveals subtype-specific dysregulation basis, which provides a prospective view to AML pathogenesis and potential targets for both diagnosis and therapy.

Keywords: Acute myeloid leukaemia, Co-expression analysis, Single-cell RNA-sequencing, Transcription factor, Transcriptional regulatory network

1. INTRODUCTION

Acute myeloid leukemia (AML) is an aggressive hematological malignancy with increasing incidence and poor prognosis in the elderly.[1,2] AML is highly heterogeneous and harbors recurrent genetic abnormalities, which involve in the dysregulation of transcriptional regulatory networks (TRNs) that may directly change the expression of transcription factors (TFs) or affect the binding of TFs to regulatory regions, resulting in farraginous clonal proliferation of blasts in bone marrow.[3] TFs such as CEPBα, GATA2, and RUNX1 are essential in hematopoietic regulation and cell fate decision, and their frequent genetic alterations would give rise to abnormal hematopoietic differentiation in AML.[3–7] A deep survey of AML at the regulatory network level will help shed light on the pathogenesis of AML.

By classifying AML according to different genetic mutations, Assi et al defined distinctive TF networks for each AML subtype, providing new insights for finding therapeutic targets.[8] However, blast cells mingled with nonmalignant cells in bulk analysis might decrease the sensitivity of identifying aberrant TRNs. In recent years, single-cell sequencing technologies have contributed to dissecting the molecular and cellular heterogeneity in physiological and pathological hematopoiesis.[9] Previous studies showed that AML cells were stagnated at different stages of hematopoietic differentiation, including primitive cells with the features of self-renewal, quiescence, and resistance to apoptosis, and differentiated cells lacking self-renewal ability that affect hematopoietic function.[10,11] Nevertheless, it remains unclear about the characteristics and mechanisms of TRNs underlie AML subtypes with different malignant cellular composition at single-cell level.

Here, we comprehensively mapped malignancy-related TF regulatory networks by analyzing scRNA-seq data in AML patients, which were classified by their cellular composition of primitive and differentiated blasts.[12] AML subtype-specific

* *Address correspondence: Ping Zhu, Lu Wang, and Peng Wu, PhD, State Key Laboratory of Experimental Hematology, National Clinical Research Center for Blood Diseases, Haihe Laboratory of Cell Ecosystem, Institute of Hematology & Blood Diseases Hospital, Chinese Academy of Medical Sciences & Peking Union Medical College, 288 Nanjing Road, Heping District, Tianjin 300020, China. E-mail address: zhuping@ihcams.ac.cn (P. Zhu), wanglu1@ihcams.ac.cn (L. Wang), and wupeng1@ihcams.ac.cn (P. Wu).*

This work was supported by grants from the National Key Research and Development Program of China (2018YFA0107804), the National Natural Science Foundation of China (81900117, 82131430173), and the CAMS Initiative for Innovative Medicine (2021-I2M-1–040).

Blood Science, (2022) 4, 65-75

Received March 10, 2022; Accepted March 30, 2022.

http://dx.doi.org/10.1097/BS9.0000000000000113

regulatory aberrations were identified and significantly enriched in cell cycle arrest, differentiation suppression, and immune modulation, which were closely relevant to the varied malignant cell blocking in AMLs. Moreover, we validated the effects of ERF dysfunction on stem/progenitor cells by loss- and gain-of-function in zebrafish embryos. Collectively, our work unveiled the regulatory aberrations and diversities of AML in view of TFs, therefore, provides a valuable perspective for understanding the pathogenesis of AML and exploring potential therapy targets.

2. RESULTS

2.1. Prevalent and common network dysregulations in AML patients

To investigate the aberrant regulations at network level in AML, we analyzed scRNA-seq data of 30,712 AML (n = 16) derived cells and 4677 donor (n = 4) derived cells and subjected to pySCENIC to define the TF regulatory networks (hereafter named regulons) (Fig. 1A).[12] In total, we obtained 196 regulons representing the activation status of TFs and their target genes.

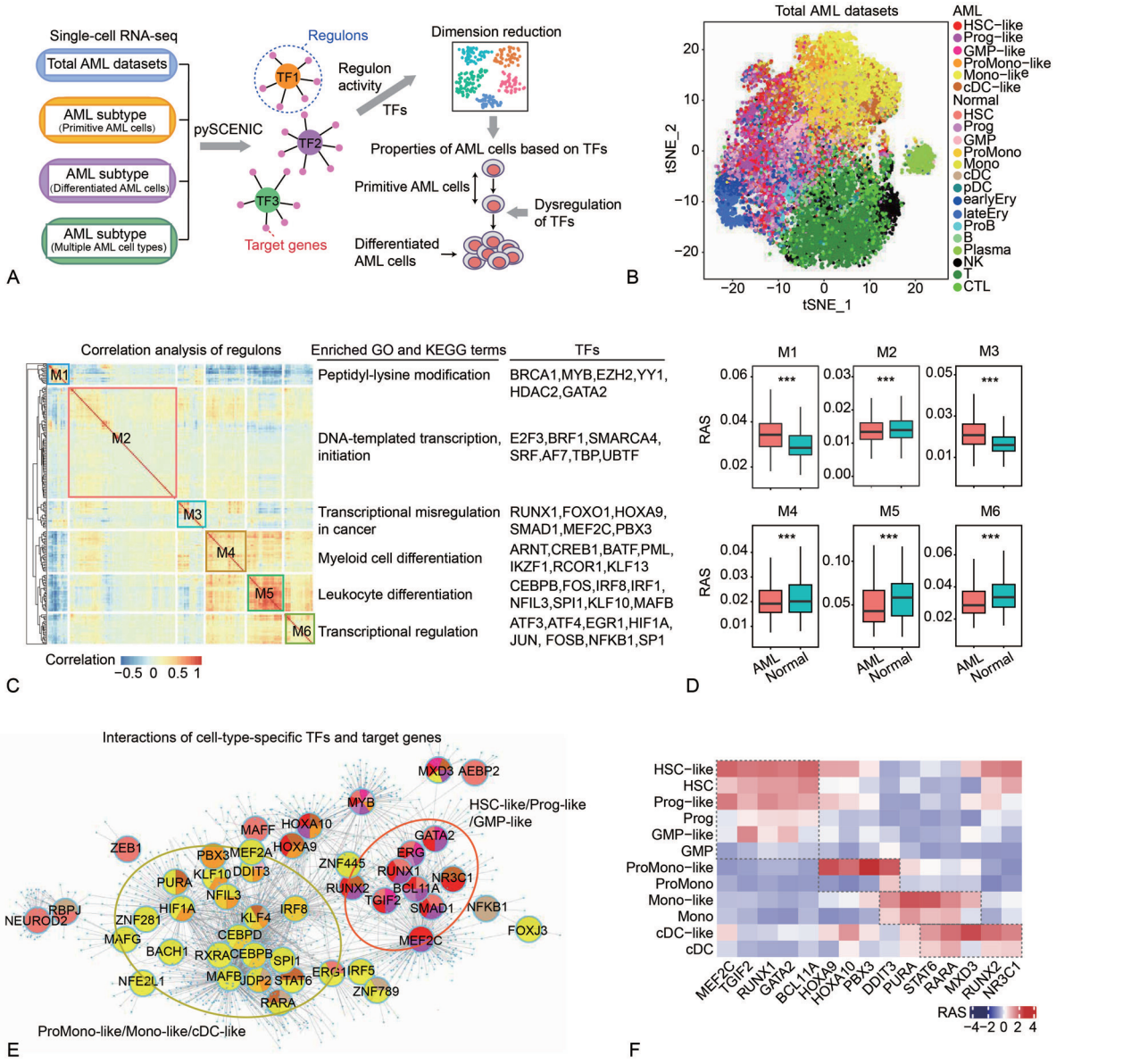

Figure 1. Differences of transcriptional regulatory networks between normal and malignant hematopoiesis based on scRNA-seq data. (A) Schematic overview of the scRNA-seq datasets and computational procedures used in this study. (B) t-SNE plot based on RAS of total normal and malignant hematopoietic cells. Cells were labeled by cell-type annotations from the reference.[12] (C) Correlation analysis of regulons based on RAS of total AML and healthy donor datasets. GO and KEGG enrichment analysis of TFs in each module. Top one enriched term and related TFs were displayed. (D) Comparison of RAS between AML and corresponding normal cells of the 6 modules in (C) ($P < .001$, Wilcoxon test). (E) Transcriptional regulatory network (TRN) constructed by TFs and their targets in AML and corresponding normal cells. Circles indicate TFs; colors indicate cell clusters; dots around a circle mean target genes. (F) Heatmap of the RAS of tumor-specific TFs in AML cells and corresponding normal cells. AML = acute myeloid leukemia, B = mature B cell, cDC = conventional dendritic cell, CTL = cytotoxic T Lymphocyte, early Ery = early erythroid progenitor, GMP = granulocyte-macrophage progenitor, GO = gene ontology, HSC = hematopoietic stem cell, late Ery = late erythroid progenitor, Mono = monocyte, NK = natural killer cell, pDC = plasmacytoid dendritic cell, Plasma = plasma cell, Pro B = progenitor B cell, Prog = progenitor, ProMono = promonocyte, RAS = regulon activity scores, T = naïve T cell.

As expected, regulon activity scores (RAS) could recapitulate the hematopoietic hierarchy and AML derived cells were aligned to the myeloid compartment, including 6 AML cell types (HSC-like, Prog-like, GMP-like, Promono-like, Mono-like, and cDC-like malignant cells) (Fig. 1B and Supplemental Figure 1D, http://links.lww.com/BS/A44). Cell clusters were dominated by their cell types instead of donors, indicating that RAS were capable to remove batch or individual effects (Supplemental Figure 1A-B, http://links.lww.com/BS/A44).

To systematically characterize the regulatory patterns of AML and normal cells and further evaluate the regulatory differences between them, we first performed unsupervised clustering of 196 regulons based on RAS by integrating cells from both AMLs and healthy donors. Six modules each containing regulons with potential co-regulation were determined (Fig. 1C and Supplemental Table 1, http://links.lww.com/BS/A45). Remarkably, we found that regulons in M3 were highly activated in malignant cells and strongly associated with transcriptional misregulation in cancer, involving AML-related TFs such as RUNX1, HOXA9, and PBX3, suggesting that M3 might play important roles in AML progression. Meanwhile, the activity scores of regulons in M4 and M5 were downregulated and these regulons were related to leukocyte differentiation and myeloid cell differentiation (Fig. 1C–D), indicating a common suppress of myeloid differentiation in all AML patients. In addition, other modules related to peptidyl-lysine modification in M1, DNA-templated transcription in M2, and transcriptional regulation in M6 were differentially activated ($P < .001$, Wilcoxon test; Fig. 1D). These results revealed a prevalent dysregulation at TF level shared by all AML patients, conferring the molecular basis of myeloid suppression and AML pathogenesis.

To further identify regulatory differences between AML and corresponding normal cells, we constructed the TRNs based on regulation weight scores and revealed the cell-type-specific TFs with high regulatory activity in malignant cells (Fig. 1E; see Methods for details). Concretely, the RAS of TFs such as RUNX1, GATA2, and MEF2C in primitive AML cells were higher than that of corresponding normal cells. PBX3 and HOXA9 were activated and upregulated in ProMono-like cells. Highly activated regulons in Mono-like cells include STAT6 and PURA, while RARA was upregulated in cDC-like cells. These TFs were reported to be associated with the pathogenesis of AML,[13–15] and we found that their dysregulation occurred in specific AML cell types. Notably, stem/progenitor AML cells exhibited a relatively similar transcriptional regulatory pattern, in contrast, differentiated AML cells featured with their own specific TFs, which further inspired us to study AML subtypes according to different malignant cellular composition (Fig. 1F). Therefore, we performed unsupervised clustering of AML patients based on tumor cellular composition and defined three distinct AML subtypes, which were represented by AML707B with primitive AML cells, AML556 with differentiated AML cells, and AML921A with multiple AML cell types. These patients were selected to representatively explore AML subtype-specific dysregulations at TF level (Supplemental Figure 1C, http://links.lww.com/BS/A44).

Collectively, these results indicated an overall and common dysregulation in AML at transcriptional regulatory level. Meanwhile, malignant cell-type-specific aberrations suggested the demands for in-depth analysis of AML subtypes classified based on cellular composition.

2.2. Inhibition of cell differentiation in AML subtype with malignant primitive cells

Next, we focused on AML subtypes to reveal the regulatory networks involving different hematopoietic blockades. We first generated regulons and estimated the RAS for each TF in cells obtained from 4 healthy donors and AML707B featured with malignant stem/progenitor cells (Supplemental Figure 1C, http://links.lww.com/BS/A44). Eight cell clusters were well-separated and patient derived malignant cells were grouped together as a unique cluster, comprised of malignant stem/progenitor cells, which was adjacent to its corresponding normal population C3 (Fig. 2A). Cell type annotations were further verified by independent gene expression analysis (Supplemental Figure 2A, http://links.lww.com/BS/A44).

We subsequently investigated the regulatory dynamics of stem/progenitor blockade by comparing C2 and C3 cells (Supplemental Table 2, http://links.lww.com/BS/A46). Notably, the highly activated TFs in malignant cells were enriched for positive regulation of cell cycle arrest, negative regulation of cell differentiation, and transcriptional misregulation in cancer, which might explain the regulatory basis of differentiation blockade at primitive stage. Besides, TFs including MEF2C, SOX4, and IRX3, have been reported to be closely associated with the occurrence, development, and poor prognosis of AML (Fig. 2B–C).[16,17]. To further verify the reliability of the disrupted cell cycle and cell differentiation in primitive AML cells, we scored the transcriptional activity of C2 and C3 using the gene sets generated from AmiGO.[18] We found that scores of C2 were lower than that of C3 in terms of cell cycle arrest and cell differentiation inhibition, while G2M scores of C2 were higher than that of C3 ($P < .001$, Wilcoxon test; Fig. 2D and Supplemental Figure 2B, http://links.lww.com/BS/A44). To make the results more reliable, we also performed the parallel transcriptional regulatory analysis for other patients in this AML subtypes and got the same conclusions (Supplemental Figure 2C–F, http://links.lww.com/BS/A44). Taken together, these results indicated that AML subtype predominated by malignant stem/progenitor cells exhibited cell cycle and cell differentiation inhibition at regulatory network level.

To further identify potential key factors in AML, we obtained the top differentially activated TFs including IRX3, ERF, and FOXD2 in C2 (Fig. 2E–F). Previous studies showed that elevated expression of IRX3 could inhibit terminal differentiation of AML cells,[19] while dysregulation of FOXD2 would result in abnormalities in epigenetic programming of leukemia cells and was associated with poor prognosis.[20] However, the roles of ERF in AML progression have not been investigated. Notably, the regulatory networks composed by differentially activated TFs and their target genes in C2 suggested that ERF might be closely related to SOX4. In addition, both ERF and SOX4 regulated *RUNX1T1*, which could pathologically fuse with *RUNX1* to affect RNA transcription, block cell differentiation, and promote AML progression (Fig. 2G).[21] Hence, these results suggested that ERF might play an important role in malignant transformation of stem/progenitor cell-dominated AMLs.

2.3. Aberrations of immunoregulation and relevant TFs in AML subtype with malignant differentiated cells

By contrast to the typical stem/progenitor-malignant AML subtype described above, malignant differentiated cells were significantly enriched in AML556 (Supplemental Figure 1C, http://links.lww.com/BS/A44). To figure out the regulatory dynamics in this AML subtype, a total of 9659 single cells from 4 healthy donors and AML556 were grouped into 8 clusters. Likewise, the malignant population C2 composed of differentiated AML cells were separated from majority normal cells and

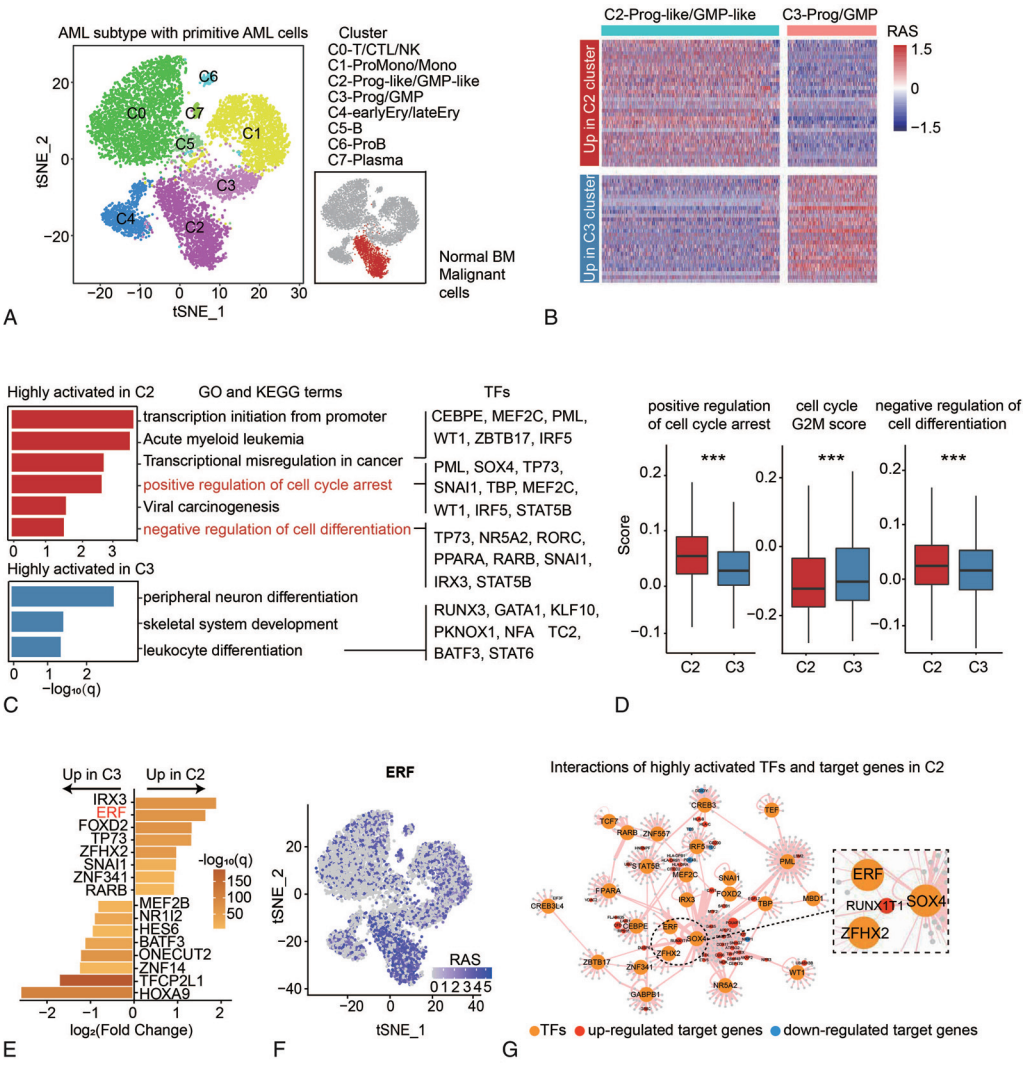

Figure 2. AML subtype featured with malignant stem/progenitor cells exhibited cell differentiation related dysregulation. (A) t-SNE plot based on RAS in AML subtype (AML707B) featured with malignant stem and progenitor cells. Points are color-coded by cell-type annotations. t-SNE plot located on the lower right emphasizes the distribution of AML cells (colored in red). (B) Heatmap of differentially activated TFs in Prog-like/GMP-like population C2 and its corresponding normal population C3. (C) Enrichment analysis of differentially activated TFs in (B). Related signaling pathways and corresponding TFs are listed on the right. (D) Transcriptional activity of cell cycle and differentiation related genes in C2 and C3 using the gene sets generated from AmiGO corresponding to the GO terms marked in red in (C) ($P < .001$, Wilcoxon test). (E) Differentially activated TFs between C2 and C3. The colors represent logarithmic transformed adjusted *P* values (Benjamini-Hochberg correction). (F) High regulon activity scores of ERF in C2 cells. (G) TRN constructed by differentially activated TFs and their target genes in C2. Circles colored in orange indicate TFs; dots around a circle indicate target genes; dots colored in red represent highly expressed target genes; dots colored in blue mean target genes with low expression; dot sizes indicate the average $\log_2$(fold change); line thickness represents the regulatory weight score. The enlarged section of TRN located on the right illustrates that ERF, SOX4, and ZFHX2 co-regulate the target gene *RUNX1T1*. AML = acute myeloid leukemia, GO = gene ontology, RAS = regulon activity scores, TF = transcription factor, TRN = transcriptional regulatory network.

close to normal population C1 (Fig. 3A and Supplemental Figure 3A, http://links.lww.com/BS/A44).

To explore the differences between malignant C2 and corresponding normal C1, and their relationship with stem/progenitor population C3, we leveraged Monocle to construct the differentiation trajectory.[22] We found a single branch point from progenitors to two termini corresponding to Prog/GMP and ProMono-like/Mono-like fates, revealing the progressive differentiation of Prog/GMP cells into ProMono-like/Mono-like and ProMono/Mono populations, respectively. Notably, the trajectory diverged and terminated at C1 and C2, indicating the significant differences between AML and corresponding normal cells (Fig. 3B and Supplemental Figure 3B, http://links.lww.com/BS/A44). Moreover, to identify branch-dependent gene expression and potential drivers of cell fate specification, we carried out branch expression analysis modeling (BEAM) analysis between C2 and C1. One thousand two hundred ninety-one branch-dependent genes ($P < 1e\text{-}6$) were differentially expressed and 4 kinetic modules were generated by unsupervised clustering (Supplementary Table 3, http://links.lww.com/BS/A47). Concretely, genes highly expressed in differentiated AML cells including *GRN*, *BTG1*, and *CFD* were enriched for interferon signaling, oxidative stress and immune-related processes, indicating that the immunomodulatory functions of differentiated malignant cells might contribute to the pathogenesis of AML as previously reported[12] (Fig. 3C and Supplemental Figure 3C, http://links.lww.com/BS/A44). We further investigated the molecular basis of C1 and C2

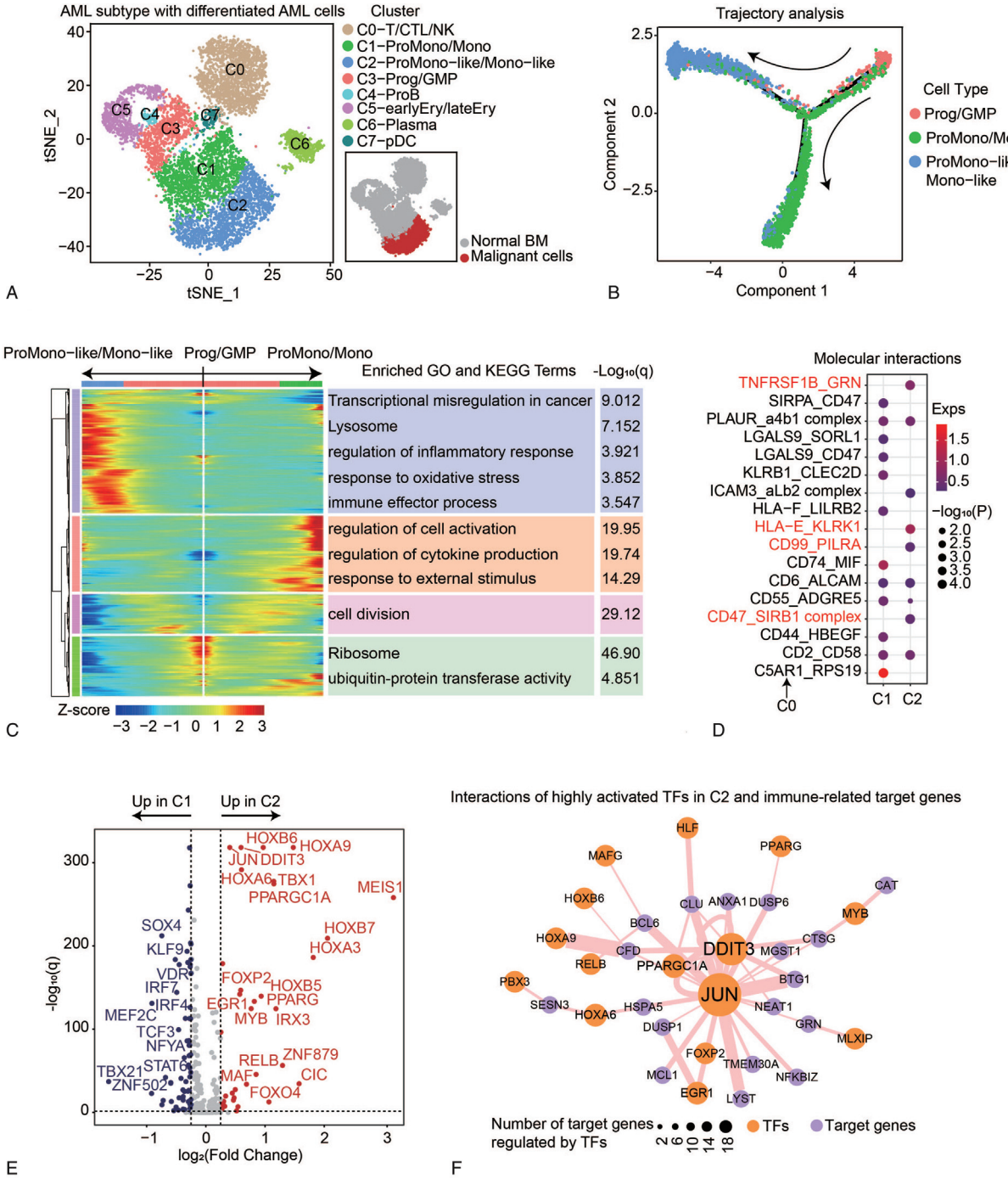

Figure 3. Transcriptional regulatory aberrations underline AML subtype with differentiated malignant cells. (A) t-SNE plot based on RAS in AML subtype (AML556) with differentiated malignant cells. Points are color-coded by cell-type annotations. t-SNE plot located on the lower right emphasized the distribution of AML cells (colored in red). (B) Trajectory analysis of Prog/GMP, ProMono/Mono, and ProMono-like/Mono-like cells based on gene expression. (C) Heatmap shows branch-dependent genes and unsupervised clustering of four kinetic modules with corresponding enrichment analysis. Related enriched GO terms and -$\log_{10}q$ value were listed on the right. (D) Spectrum of ligand-receptor pairs (rows) between C1/C2 and T/CTL/NK cells C0 (columns). Dot sizes represent logarithmic transformed *P* values (permutation test) and dot colors mean the expression of interacting molecules in corresponding cell subsets. (E) Volcano plot shows the differentially activated TFs in C1 (dots colored in blue) and C2 (dots colored in red). (F) TRN constructed by differentially activated TFs in C2 and target genes enriched in immune-related GO terms in (E). Circles colored in orange represent TFs; dots colored in purple around a circle indicate target genes; circle sizes indicate the number of regulated targeted genes; line thickness represents the regulatory weight scores. AML = acute myeloid leukemia, GO = gene ontology, RAS = regulon activity scores, TF = transcription factor, TRN = transcriptional regulatory network.

crosstalk with C0 (T/CTL/NK population) by assessing the expression of paired ligand receptor. We detected 12 actively expressed cell-type-specific ligand-receptor pairs in C1 and 9 in C2. In detail, immune-related receptor ligand pairs TNFRSF1B_GRN, HLA-E_KLRK1, CD99_PILRA, and CD47_SIRB1 complex appeared only in C2 (Fig. 3D), supporting the aberrant immunoregulatory role of differentiated AML cells.

Next, to investigate the dysregulated TRNs related to immunoregulation in C2 and find TFs that regulated the immunomodulatory genes mentioned above, we identified the differentially activated TFs between C2 and corresponding normal population C1. Notably, the elevated activity of MEIS1 and HOX-related TFs in C2 was associated with *NPM1* mutations in AML556 (Fig. 3E).[23] We constructed TRNs derived from highly activated TFs and target genes enriched in immune-related processes in C2. We observed that JUN and DDIT3, both of which regulated lots of immune-related genes (Fig. 3F). As the first oncogenic TF discovered, JUN could increase proliferation and prevents differentiation in AML by inhibiting C/EBPα DNA binding.[24] Besides, JUN targeted immune-regulatory molecule *GRN*,[25] suggesting its role in immune regulation during AML progression. DDIT3 is a member of C/EBP family, which functions in endoplasmic reticulum stress response, thereby inducing cell cycle arrest.[26] In addition, we also performed the parallel transcriptional regulatory analysis on patient AML419, and the results showed that differentiated AML cells were associated with immune-related signaling pathways such as oxidative stress and cytokine response compared with corresponding normal cells, which verified the rationality of our previous analysis (Supplemental Figure 3D–G, http://links.lww.com/BS/A44). Collectively, these results indicated the abnormal immunomodulatory functions of malignant differentiated cells and uncovered the underlying dynamics of TFs relevant to immune functions in AMLs.

2.4. Dysregulated transcriptional regulatory patterns and enhanced cell differentiation capacity of primitive malignant cells in AML subtype with multiple cell types

Unlike the two AML subtypes mentioned above, the third subtype represented by AML921A had multiple malignant cell types along the HSC to myeloid axis (Supplemental Figure 1C, http://links.lww.com/BS/A44). We distinguished 9 clusters based on RAS of TFs from healthy donors and AML921A, among which C0 and C2 were mainly composed of malignant cells (Fig. 4A and Supplemental Figure 4A, http://links.lww.com/BS/A44). As recent studies showed that AML cells exhibited abnormal gene expression patterns, we further investigated whether the transcriptional regulatory patterns of AML cells were dysregulated compared with normal cells. To relate the differentiation states of AML and normal cells at transcriptional regulatory level, we first generated TF signatures in normal cells including HSC/Prog, GMP and Myeloid. As expected, these TFs could clearly distinguish the normal cell types. However, when we applied the same TFs to AML cells, we found that HSC/Prog-like and GMP-like cells followed similar regulatory patterns that majority TFs turned down their activities as compared to normal cells. By contrast, differentiated myeloid-like cells maintained the normal regulatory networks. Therefore, in addition to the co-expression of genes, these results showed the aberrations of transcriptional regulatory patterns of primitive malignant cells in AML (Fig. 4B and Supplemental Figure 4B, http://links.lww.com/BS/A44).

Next, we identified the highly activated TFs in primitive AML cells to examine the transcriptional regulatory differences between AML921A and AML707B, representing patients belong to subtypes dominated by multiple cell types and primitive cells, respectively. Gene ontology (GO) enrichment analysis revealed that the former was active in leukocyte/myeloid differentiation and cell fate commitment, while the latter was enriched for negative regulation of cell differentiation, suggesting an enhanced differentiation capacity of primitive cells in AML921A, in concordant with the multiple malignant cellular composition in this type of patients (Fig. 4C and Supplemental Table 4, http://links.lww.com/BS/A48). Besides, dysregulated TFs such as HOXB6, NKX2-1, RELA, STAT6, and their target genes were associated with cell differentiation, transcriptional misregulation in cancer, and immune signals, which further supported a distinct phenotype of AML patients with multiple malignant cell types (Fig. 4D–E and Supplemental Figure 4E, http://links.lww.com/BS/A44). Parallel transcriptional regulatory analysis was also performed on patient AML475 whose primitive AML cells were associated with cell differentiation (Supplemental Figure 4F–I, http://links.lww.com/BS/A44). Notably, among the top ranked TFs based on regulon specificity score (RSS), HOXB6, a structural protein with many characteristics of carcinogenesis, had been implicated in the development of AML (Supplemental Figure 4C, http://links.lww.com/BS/A44).[27] Indeed, we found that patients with higher HOXB6 signals had significantly worse outcomes (Fig. 4F). In addition to previously reported AML associated TFs, we also identified new TFs with potential roles in AML. For instance, NKX2-1 was reported as a potential oncogene in T cell acute lymphoblastic leukemia,[28] but its regulatory mechanism in AML had not been further studied yet. Additional survival analysis showed poor prognosis of patients with NKX2-1 and other highly activated TFs (Fig. 4F and Supplemental Figure 4D, http://links.lww.com/BS/A44). Collectively, these results revealed an abnormal co-regulatory network related to cell differentiation capacity that might confer the formation of cellular composition by multiple malignant cell types in this AML subtype.

2.5. ERF is required for stem/progenitor cell development

To verify the functions of our defined key TFs in hematopoiesis, we selected ERF as an illustration of the potential role in stem/progenitor cells. The zebrafish has been known as an excellent vertebrate model system to study hematopoiesis due to several appealing advantages, including external fertilization, genetic accessibility, and optical transparency for live imaging.[29] More importantly, the developmental process and molecular mechanism involved in zebrafish hematopoiesis are highly conserved with mammals.[30,31] Therefore, to investigate whether ERF is essential for stem/progenitor cells, we first examined the expression pattern of *erf* in zebrafish embryos by whole-mount in situ hybridization (WISH). Our results showed that the expression of *erf* was observed in aorta-gonad-mesonephric (AGM) region at 36 hours post-fertilization (hpf) (Supplemental Figure 5A, http://links.lww.com/BS/A44). We further performed double fluorescence in situ hybridization (FISH) and found that *erf* was co-expressed with *runx1*, which is as a master regulator of hematopoiesis,[32,33] indicating that *erf* was specifically expressed in HSPCs (Supplemental Figure 5B, http://links.lww.com/BS/A44).

Next, we performed knockdown experiments by injecting antisense morpholino (MO) into zebrafish embryos at the one-cell stage to block endogenous *erf* expression. The MO efficiency was validated by co-injection of *erf* MO and *erf*-EGFP construct with MO binding sites, and the results showed EGFP expression was blocked efficiently in co-injection embryos (Supplemental Figure 5C, http://links.lww.com/BS/A44). WISH data showed that the expression of *runx1*/*cmyb*, the HSPC markers, was absent in embryos upon *erf* knockdown (Fig. 5A). Consistent results were also achieved by the quantitative RT-PCR (qPCR). The expression of *runx1* and *cmyb* was significantly decreased in

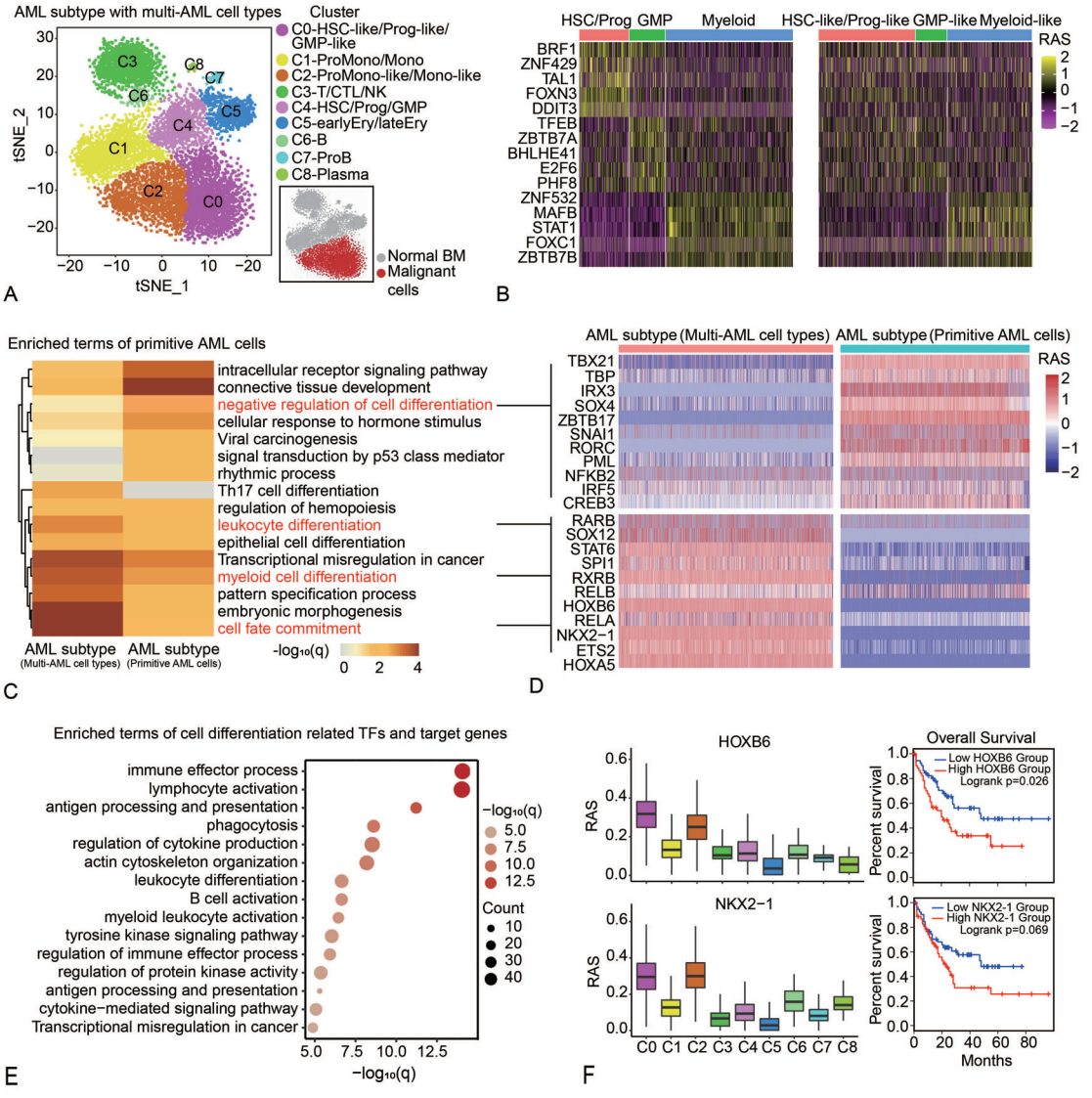

Figure 4. Aberrant regulation and cell differentiation of primitive malignant cells in AML subtype with multiple cell types. (A) t-SNE plot based on RAS of AML subtype (AML921A) with multiple malignant cell types along the HSC to myeloid axis. Points are color-coded by cell-type annotations. t-SNE plot located on the lower right emphasized the distribution of AML cells (colored in red). (B) Heatmaps show RAS of normal TF signatures for HSC/Prog, GMP, and differentiated myeloid cells in normal cells (left) and malignant AML cells (right). (C) GO and KEGG pathway analysis of highly activated TFs of primitive malignant cells in AML921A and AML707B. Color shows the -log$_{10}$(*q* value). (D) Heatmap shows the RAS of TFs related to the signaling pathways marked in red from (C) in the primitive malignant cells of AML921A and AML707. (E) Enrichment analysis of cell differentiation related TFs and their targets. Dot sizes represent number of genes enriched in each GO term and dot colors represent logarithmic transformed adjusted *P* values (Benjamini-Hochberg correction). (F) RAS of HOXB6 and NKX2-1 in all AML921A cell clusters (left) and their overall survival curves (right). AML = acute myeloid leukemia, GO = gene ontology, RAS = regulon activity scores, TF = transcription factor.

erf morphants, compared with the control group (Fig. 5B). To visualize the developmental dynamics of HSPCs in vivo, we injected *erf* MO into Tg (*kdrl*:mCherry/*cmyb*:EGFP) transgenic embryos, in which *kdrl*$^+$ *cmyb*$^+$ cells were indicated as hemogenic endothelial cells.[34] As expected, time-lapse confocal imaging showed that the number of *kdrl*$^+$*cmyb*$^+$ cells in the AGM was decreased at 36 hpf in *erf* morphants (Fig. 5C). And the number of *cmyb*$^+$ cells, which labeled HSPCs, in the CHT was also reduced at 50 hpf compared with control embryos (Fig. 5D). Moreover, overexpression of *erf* induced an increased expression of *runx1* and *cmyb* by qPCR and WISH (Supplemental Figure 5D–F, http://links.lww.com/BS/A44). Similarly, live imaging using Tg (*kdrl*:mCherry/*cmyb*:EGFP) double transgenic embryos showed that the number of *kdrl*$^+$ *cmyb*$^+$ cells was increased at 36 hpf in *erf* mRNA-injected embryos, as well as the number of *cmyb*$^+$ cells at 50 hpf (Fig. 5G–H). Taken together, these data indicated that *erf* is required for stem/progenitor cell development; overexpression of *erf* could enhance the expansion of stem/progenitor cells and might confer the AML phenotype dominated by primitive malignant cells in patients.

3. DISCUSSION

AML is characterized by the suppression of hematopoiesis which caused accumulation of immature myeloid cells. Here, we analyzed available scRNA-seq data of AMLs and healthy donors

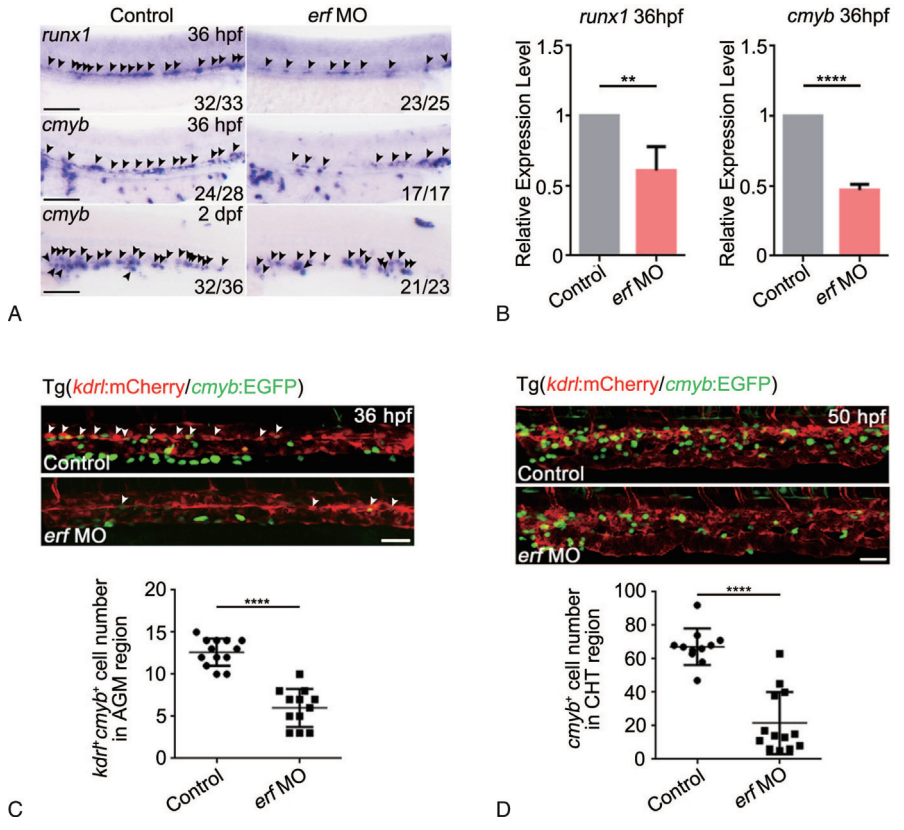

Figure 5. Functional validation of ERF in stem/progenitor cells. (A) Expression of *runx1* and *cmyb* in control and *erf* morphants at 36 hpf and 2 dpf, respectively. Black arrowheads denote runx1 and *cmyb* expression in AGM region. Scale bar, 100 μm. (B) qPCR analysis of *runx1* and *cmyb* in control and *erf* morphants at 36 hpf. The expression level of *runx1* and *cmyb* were normalized to *β-actin* (mean ± SD, Student *t* test; n = 3; $^{**}P < .01$, $^{****}P < .0001$). (C) Confocal microscopy of Tg (*kdrl*:mCherry/*cmyb*:EGFP) transgenic line at 36 hpf in control and *erf* morphants (upper panel). White arrowheads mark the hemogenic endothelial cells. The bottom panel showing quantification of *kdrl*$^+$ *cmyb*$^+$ cells in AGM region of control and *erf* morphants at 36 hpf (mean ± SD, Student *t* test, n = 12, $^{****}P < .0001$). Scale bar, 50 μm. (D) Visualization of *cmyb*$^+$ cells in CHT region of control and *erf* morphants at 50 hpf (upper panel) with quantification (bottom panel) (mean ± SD, Student *t* test, n = 10, $^{****}P < .0001$). White arrowheads mark HSPCs. Scale bar, 50 μm.

and focused on the dysregulated TRNs underline different AML subtypes. We found global differences between AMLs and healthy donors at transcription factor regulatory network level. Furthermore, we classified different types of AMLs by their malignant cellular composition and unveiled the subtype-specific aberrant regulatory networks and properties of distinct malignant cell types. Overall, our study illustrates an abnormal spectrum of TRNs in AML and provides insights for targeted therapy.

Our single-cell analysis precisely identified the aberrant regulation of malignant cells which were robust to confounding signals from nonmalignant cells. Based on the published bone marrow-derived AML and normal single-cell RNA-seq data, we have comprehensively revealed how abnormally activated TFs shape the distinct TRNs of AMLs. Remarkably, AML presented prevalent dysregulation as compared to normal cells at network level, which were strongly associated with tumoral transcriptional misregulation and suppressed cell differentiation. Though heterogeneity exists in AML patients, for instances, cellular compositions and gene expressions at the whole transcriptome level, the aberrations of six modules of TFs were shared among patients. This result may suggest a possible hint of common therapy strategy by interfering pathways enriched in these modules for AML patients.

We also analyzed specific TRNs in AML subtypes with different malignant cellular compositions. Patients predominated by malignant stem/progenitor cells showed disrupted cell cycle and cell differentiation at regulatory network level. Highly activated TFs including IRX3, ERF, and FOXD2 were identified in malignant stem/progenitor cells. To verify the malignancy-related functions of TFs dysregulated in this AML subtype, we evaluated the effects of knockdown and overexpression of ERF on hematopoiesis in zebrafish embryos, confirming its contribution to HSPC development. Other key TFs identified in this study need efforts to distinguish from passenger aberrations.

We further documented regulatory aberrations of immune functions and identified two TFs with high immune-regulatory activity (JUN and DDIT3) in AML subtypes dominated by malignant differentiated cells. Studies have shown that JUN is a positive regulator of myeloid differentiation and involved in directing monopoiesis during normal hematopoiesis, and DDIT3 emerges as a regulatory node during erythro-myeloid cell fate bifurcation.[35,36] As the abnormal expression of these two TFs was recognized as the mechanisms leading to leukemia progression, our analysis further elucidated their regulatory roles in immunity, which paved the way to study AML immune microenvironment.[26]

In patients with multiple malignant cell types, primitive cells exhibited a disordered stemness and myeloid regulatory pattern, and an enhanced cell differentiation capacity. We also identified cell differentiation-related TFs including HOXB6 and NKX2-1

in primitive AML cells, which were associated with worse clinical outcomes. The regulated pattern of HOXB6 for normal monocytopoiesis and granulopoiesis has been extensively characterized and abnormalities of the HOXB6 expression may contribute to the development of the leukemic phenotype, which is consistent with our findings.[27,37] Besides, NKX2-1 has been proved to promote leukemogenesis in T-cell progenitors, especially in T cell acute lymphoblastic leukemia.[38] Our findings further suggest its regulatory role in AML progression. Taken together, these key TFs we identified are potentially essential in physiological and pathological hematopoiesis and worthy of investigation.

In summary, we comprehensively explored the TRNs in AML, unveiled properties of malignant cell types, and identified key TFs in AML subtypes at single-cell resolution. Importantly, TRNs in AML subtypes dominated by different malignant cells revealed the specific regulatory characteristics and mechanisms for aberrant hematopoiesis. Our findings provide clues for future efforts to identify potential therapy targets for AML patients.

4. MATERIALS AND METHODS

4.1. Experimental model and subject details

Zebrafish strains including Tubingen and transgenic lines, *kdrl*:mCherry/*cmyb*:EGFP were raised in a standard circulating water system at 28°C. Zebrafish embryos were obtained by the natural mating and grown at 28.5°C in embryo medium. All zebrafish work was approved by the Ethical Review Committee of 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, China.

4.2. Source of AML Data

The UMI matrix of single cells for AML and normal samples and the corresponding cell type annotations were downloaded from the processed data previously published (https://www.ncbi.nlm.nih.gov/geo/GSE116256).[12]

4.3. Regulatory network analysis

Here we used pySCENIC (v0.10.4), a lightning-fast python implementation of the SCENIC pipeline (Single-Cell rEgulatory Network Inference and Clustering),[39] to construct the regulatory networks through three steps: (i) inferring the co-expression relationship between TFs and genes with GRNBoost2, (ii) create candidate regulons from TF-target gene interactions and prune indirect targets from candidate regulons based on motif discovery using cisTarget. We chose motifs-v9-nr.hgnc-m0.001-o0.0.tbl and hg19–500bp-upstream-10species.mc9nr.feather database for our cisTarget analysis. Each TF together with its potential direct targets is a regulon, and (iii) calculating the activity of each TF and then generate the Regulon Activity Matrix using AUCell. The TF list were merged from Animal TFDB,[40] JASPAR[41] and the supplementary data.[42] Thus, we obtained the activity scores of each TF in each cell based on the integrated UMI matrix from AML and normal bone marrows. In the following, Seurat package (v 2.3.2)[43] was applied to explore the regulatory clusters and detect their specifically activated TFs, on the basis of activity matrix. Noticeably, the activity matrix was submitted to Seurat without any filtering and all the enriched TFs were used to dimensionality reduction. Furthermore, the scale data was replaced by the raw activity matrix for subsequent t-SNE and clustering analysis. Finally, regulatory networks made up of specific TFs and their direct targets were displayed by Cytoscape (v3.5.1) software.[44]

4.4. Regulon module analysis

To evaluate the transcriptional regulatory patterns, we identified 6 regulon modules by unsupervised clustering of a correlation matrix, which was generated by calculating the Pearson correlation coefficient of single-cell RAS scores between every two TFs in total AML and normal cell datasets. GO and KEGG enrichment analysis of regulon sets in each module was performed by Metascape (http://metascape.org/gp/index.html#/main/step1) and we listed the enriched terms with the highest $-\log_{10}(P)$ value for each module.[45] Related boxplots were generated by comparing the RAS of TFs in the same module between AML and corresponding normal cells using the ggplot2 package (v3.3.5).

4.5. Interaction analysis of cell-type-specific TFs and target genes

Interactions between cell-type-specific TFs and target genes in Figure 1E was generated by following steps: (i) we first calculated the logarithm of regulatory weight score of every regulon pair generated by pySCENIC, and filtered out pairs with regulatory weight scores less than 1; (ii) we selected top 10 markers of 6 AML clusters and corresponding normal cells generated by Seurat; (iii) we took the intersection of the TFs in two datasets mentioned above, and finally obtained 49 TFs and 1195 target genes. Next, TF-target pairs were entered into Cytoscape.[44] Subsequent TRNs of different AML subtypes were also constructed based on Cytoscape software.

4.6. Generation of differentially activated TFs

Seurat package was applied to identify cell cluster-highly activated TFs based on RAS matrix. TFs with adjusted *P*-value lower than .05 and average fold change higher than 1.5 were defined as significantly differentially activated. Pheatmap package (v1.0.12) in R language (v3.4.3) was used to plot heatmaps. Highly activated TFs in heatmap were derived from the AML-related TFs circled in Figure 1E and the RAS of those TFs in malignant cells were 1.5 times higher than that in corresponding normal cells, as shown in Figure 1F. Besides, the cell-type-specific TF signatures in Figure 4B were generated by correlating RAS of TFs to cell type prediction scores from previously published in normal and malignant cells types along the HSC to myeloid axis.

4.7. Cell cycle and cell differentiation scoring

The AddModuleScore function in Seurat was used to score malignant C2 and normal C3 in AML707B at gene expression level. The gene sets related to signaling pathways of cell cycle and cell differentiation used for scoring were generated from AmiGO[18] (http://amigo.geneontology.org/amigo) and GSVA (v1.32.0).[46]

4.8. Trajectory and pseudotime analysis

We applied single-cell trajectory and pseudotime analysis to cells of normal C1, malignant C2, and stem/progenitor population C3 in AML556 using Monocle2 (v2.18.0)[22] with default parameters. Besides, we performed BEAM to identify the branch-dependent genes and chose *P* value <1e-6 as the threshold to filter genes generated by BEAM analysis. Finally, 1291 branch-dependent genes were identified and clustered into 4 kinetic modules for further functional enrichment analysis.

4.9. Cell-cell communication analysis

CellPhoneDB (v2.1.7)[47] was used with default parameters to estimate cell-cell communications of C1 and C2 crosstalk with T/CTL/NK population C0 in AML556.

4.10. Survival analysis

Overall survival time was calculated from diagnosis to first recurrence or death from disease. The survival curves were estimated according to the Kaplan-Meier algorithm by the GEPIA2 web-based tool (http://gepia2.cancer-pku.cn/), which calculated the *P* value through log-rank test from the coxph function[48]

4.11. Calculation of RSS

The cell cluster specificity of a TF was evaluated by RSS previously reported.[49,50] RSS was defined as:

$$RSS = 1 - JSD$$

where JSD is the Jensen-Shannon Divergence, and was used to quantify the difference between a TF RAS probability distribution (*R*) and a cluster-specific control RAS probability distribution (*C*). The calculation formula of JSD was:

$$JSD(R,C) = \sqrt[2]{H\left(\frac{R+C}{2}\right) - \frac{H(R)+H(C)}{2}}$$

where *H* was the Shannon entropy of a probability distribution, *R* was a probability distribution of a TF RAS value in all cells, and *C* was a control distribution in which the interesting cluster cells were 1 and others were 0. Both *R* and *C* were normalized by the sum.

Morpholinos, mRNA synthesis, and microinjection

Antisense MO was purchased from GeneTools and prepared as 1 mM stock solutions using ddH_2O. The full-length ERF mRNA was synthesized from a pCS2 plasmid using the mMessage mMachine SP6 kit (Ambion). For fish embryo injections, *erf* MO (0.8 ng) and *erf* mRNA (100 pg) was injected into one-cell stage embryos at the yolk/blastomere boundary. The sequence of *erf* MO used was described in Supplemental Table 5, http://links.lww.com/BS/A49.

4.12. Whole-mount and FISH

WISH for zebrafish embryos was performed as reported previously with RNA probes targeting the genes of *erf*, *runx1*, and *cmyb*.[51] Briefly, fixed embryos were permeabilized with Proteinase K (10 μg/ml) after rehydration, re-fixed in 4% paraformaldehyde, and then hybridized with antisense DIG labeled RNA probe at 65°C overnight. After washing with sodium citrate buffer and blocking with MAB, embryos were incubated in anti-DIG antibody (1:5000) at 4°C overnight. After removing antibody and washing embryos with MABT, embryos were stained with BM Purple. The embryos after staining were mounted in 90% glycerol and photographed with Nikon SMZ18 camera. Double FISH was performed as previously described,[52] with FLU-labeled *runx1* and DIG-labeled *erf* probes.

4.13. Quantitative Real-Time PCR

Total RNA used for qPCR was collected from dissected AGM regions of control and mRNA-injected embryos (50 embryos pooled for each sample) at 36 hpf, and extracted by TRNzol reagent. The cDNA templates were all reversely transcribed from total RNA by M-MLV Reverse Transcriptase and subsequently diluted 5 times. The qPCR experiment was performed using QuantStudio3 Real-Time PCR system. *P* values were determined using two-sided Student *t* test for three samples with equal variance. $^{*}P < .05$; $^{**}P < .01$; $^{****}P < .0001$. Primer sequences used were provided in Supplemental Table 5, http://links.lww.com/BS/A49.

4.14. Plasmids Construction

The *erf*-EGFP fusion construct was generated to validate the efficiency of *erf* MO using PEGFP-N1 basic vector. The construct contained a partial sequence of *erf* gene including the MO recognition site and fused EGFP in frame. The primers were listed in Supplemental Table 5, http://links.lww.com/BS/A49.

4.15. Confocal Microscopy

After anesthetized and mounted in 1.2% low melting agarose, transgenic zebrafish embryos were imaged by Dragonfly 200 spinning disk confocal microscope with 20× objective. And the imaging was edited by ImageJ.

4.16. Quantification and Statistical Analysis

All statistical analyses of qPCR and confocal imaging were performed at least three independent biological or experimental replicates. Student's two-tailed unpaired *t* test was used for statistical comparisons and data are shown as mean ± SD. *P* value was used for significance. $^{*}P < .05$; $^{**}P < .01$; $^{****}P < .0001$.

ACKNOWLEDGEMENTS

This work was supported by grants from the National Key Research and Development Program of China (2018YFA0107804), the National Natural Science Foundation of China (81900117, 82131430173), and the CAMS Initiative for Innovative Medicine (2021-I2M-1-040).

REFERENCES

[1] Shah A, Andersson TM, Rachet B, Bjorkholm M, Lambert PC. Survival and cure of acute myeloid leukaemia in England, 1971–2006: a population-based study. *Br J Haematol* 2013;162 (4):509–516. doi:10.1111/bjh.12425.

[2] De Kouchkovsky I, Abdul-Hay M. Acute myeloid leukemia: a comprehensive review and 2016 update. *Blood Cancer J* 2016;6 (7): e441. doi:10.1038/bcj.2016.50.

[3] Li S, Mason CE, Melnick A. Genetic and epigenetic heterogeneity in acute myeloid leukemia. *Curr Opin Genet Dev* 2016;36:100–106. doi:10.1016/j.gde.2016.03.011.

[4] Stavast CJ, Leenen PJM, Erkeland SJ. The interplay between critical transcription factors and microRNAs in the control of normal and malignant myelopoiesis. *Cancer Lett* 2018;427:28–37. doi:10.1016/j.canlet.2018.04.010.

[5] Iwasaki H, Mizuno S, Arinobu Y, et al. The order of expression of transcription factors directs hierarchical specification of hematopoietic lineages. *Genes Dev* 2006;20 (21):3010–3021. doi:10.1101/gad.1493506.

[6] Chen L, Kostadima M, Martens JHA, et al. Transcriptional diversity during lineage commitment of human blood progenitors. *Science* 2014;345 (6204):1251033. doi:10.1126/science.1251033.

[7] Churpek JE, Bresnick EH. Transcription factor mutations as a cause of familial myeloid neoplasms. *J Clin Invest* 2019;129 (2):476–488. doi:10.1172/JCI120854.

[8] Assi SA, Imperato MR, Coleman DJL, et al. Subtype-specific regulatory network rewiring in acute myeloid leukemia. *Nat Genet* 2019;51 (1):151–162. doi:10.1038/s41588-018-0270-1.
[9] Wu J, Xiao Y, Sun J, et al. A single-cell survey of cellular hierarchy in acute myeloid leukemia. *J Hematol Oncol* 2020;13 (1):128. doi:10.1186/s13045-020-00941-y.
[10] Karamitros D, Stoilova B, Aboukhalil Z, et al. Single-cell analysis reveals the continuum of human lympho-myeloid progenitor cells. *Nat Immunol* 2018;19 (1):85–97. doi:10.1038/s41590-017-0001-2.
[11] Velten L, Haas SF, Raffel S, et al. Human haematopoietic stem cell lineage commitment is a continuous process. *Nat Cell Biol* 2017;19 (4):271–281. doi:10.1038/ncb3493.
[12] van Galen P, Hovestadt V, Wadsworth Ii MH, et al. Single-Cell RNA-Seq reveals AML hierarchies relevant to disease progression and immunity. *Cell* 2019;176 (6):1265–1281e24. doi:10.1016/j.cell.2019.01.031.
[13] Sood R, Kamikubo Y, Liu P. Role of RUNX1 in hematological malignancies. *Blood* 2017;129 (15):2070–2082. doi:10.1182/blood-2016-10-687830.
[14] Laszlo GS, Alonzo TA, Gudgeon CJ, et al. High expression of myocyte enhancer factor 2C (MEF2C) is associated with adverse-risk features and poor outcome in pediatric acute myeloid leukemia: a report from the Children's Oncology Group. *J Hematol Oncol* 2015;8 (1):115. doi:10.1186/s13045-015-0215-4.
[15] Li Z, Zhang Z, Li Y, et al. PBX3 is an important cofactor of HOXA9 in leukemogenesis. *Blood* 2013;121 (8):1422–1431. doi:10.1182/blood-2012-07-442004.
[16] Puccetti E, Ruthardt M. Acute promyelocytic leukemia: PML/RARα and the leukemic stem cell. *Leukemia* 2004;18 (7):1169–1175. doi:10.1038/sj.leu.2403367.
[17] Rampal R, Figueroa ME. Wilms tumor 1 mutations in the pathogenesis of acute myeloid leukemia. *Haematologica* 2016;101 (6):672–679. doi:10.3324/haematol.2015.141796.
[18] Carbon S, Ireland A, Mungall CJ, et al. AmiGO: online access to ontology and annotation data. *Bioinformatics* 2008;25 (2):288–289. doi:10.1093/bioinformatics/btn615.
[19] Somerville TDD, Simeoni F, Chadwick JA, et al. Derepression of the Iroquois homeodomain transcription factor gene IRX3 confers differentiation block in acute leukemia. *Cell Rep* 2018;22 (3):638–652. doi:10.1016/j.celrep.2017.12.063.
[20] Li Y, Zhao H, Xu Q, et al. Detection of prognostic methylation markers by methylC-capture sequencing in acute myeloid leukemia. *Oncotarget* 2017;8 (66):110444–110459. doi:10.18632/oncotarget.22789.
[21] Al-Harbi S, Aljurf M, Mohty M, Almohareb F, Ahmed SOA. An update on the molecular pathogenesis and potential therapeutic targeting of AML with t(8;21)(q22;q22.1);RUNX1-RUNX1T1. *Blood Adv* 2020;4 (1):229–238. doi:10.1182/bloodadvances.2019000168.
[22] Trapnell C, Cacchiarelli D, Grimsby J, et al. The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nat Biotechnol* 2014;32 (4):381–386. doi:10.1038/nbt.2859.
[23] Nagy Á, Ősz Á, Budczies J, et al. Elevated HOX gene expression in acute myeloid leukemia is associated with NPM1 mutations and poor survival. *J Adv Res* 2019;20:105–116. doi:10.1016/j.jare.2019.05.006.
[24] Rangatia J, Vangala RK, Singh SM, et al. Elevated c-Jun expression in acute myeloid leukemias inhibits C/EBPα DNA binding via leucine zipper domain interaction. *Oncogene* 2003;22 (30):4760–4764. doi:10.1038/sj.onc.1206664.
[25] Jian J, Konopka J, Liu C. Insights into the role of progranulin in immunity, infection, and inflammation. *J Leukoc Biol* 2013;93 (2):199–208. doi:10.1189/jlb.0812429.
[26] Wang Y-l, Qian J, Lin J, et al. Methylation status of DDIT3 gene in chronic myeloid leukemia. *J Exp Clin Cancer Res* 2010;29 (1):54–154. doi:10.1186/1756-9966-29-54.
[27] Fischbach NA, Rozenfeld S, Shen W, et al. HOXB6 overexpression in murine bone marrow immortalizes a myelomonocytic precursor in vitro and causes hematopoietic stem cell expansion and acute myeloid leukemia in vivo. *Blood* 2005;105 (4):1456–1466. doi:10.1182/blood-2004-04-1583.
[28] Nagel S, Pommerenke C, Scherr M, et al. NKL homeobox gene activities in hematopoietic stem cells, T-cell development and T-cell leukemia. *PLoS One* 2017;12 (2):e0171164. doi:10.1371/journal.pone.0171164.
[29] Gore AV, Pillay LM, Venero Galanternik M, Weinstein BM. The zebrafish: a fintastic model for hematopoietic development and disease. *Wiley Interdiscip Rev Dev Biol* 2018;7 (3):e312. doi:10.1002/wdev.312.
[30] Ciau-Uitz A, Monteiro R, Kirmizitas A, Patient R. Developmental hematopoiesis: ontogeny, genetic programming and conservation. *Exp Hematol* 2014;42 (8):669–683. doi:10.1016/j.exphem.2014.06.001.
[31] Paik EJ, Zon LI. Hematopoietic development in the zebrafish. *Int J Dev Biol* 2010;54 (6–7):1127–1137. doi:10.1387/ijdb.093042ep.
[32] Lam EY, Chau JY, Kalev-Zylinska ML, et al. Zebrafish runx1 promoter-EGFP transgenics mark discrete sites of definitive blood progenitors. *Blood* 2009;113 (6):1241–1249. doi:10.1182/blood-2008-04-149898.
[33] North TE, de Bruijn MF, Stacy T, et al. Runx1 expression marks long-term repopulating hematopoietic stem cells in the midgestation mouse embryo. *Immunity* 2002;16 (5):661–672. doi:10.1016/s1074-7613(02)00296-0.
[34] Bertrand JY, Chi NC, Santoso B, Teng S, Stainier DY, Traver D. Haematopoietic stem cells derive directly from aortic endothelium during development. *Nature* 2010;464 (7285):108–111. doi:10.1038/nature08738.
[35] Friedman AD. C/EBPalpha induces PU.1 and interacts with AP-1 and NF-kappaB to regulate myeloid development. *Blood Cells Mol Dis* 2007;39 (3):340–343. doi:10.1016/j.bcmd.2007.06.010.
[36] Pina C, Teles J, Fugazza C, et al. Single-cell network analysis identifies DDIT3 as a nodal lineage regulator in hematopoiesis. *Cell Rep* 2015;11 (10):1503–1510. doi:10.1016/j.celrep.2015.05.016.
[37] Giampaolo A, Felli N, Diverio D, et al. Expression pattern of HOXB6 homeobox gene in myelomonocytic differentiation and acute myeloid leukemia. *Leukemia* 2002;16 (7):1293–1301. doi:10.1038/sj.leu.2402532.
[38] Homminga I, Pieters R, Meijerink JPP. NKL homeobox genes in leukemia. *Leukemia* 2012;26 (4):572–581. doi:10.1038/leu.2011.330.
[39] Van de Sande B, Flerin C, Davie K, et al. A scalable SCENIC workflow for single-cell gene regulatory network analysis. *Nat Protoc* 2020;15 (7):2247–2276. doi:10.1038/s41596-020-0336-2.
[40] Zhang HM, Chen H, Liu W, et al. AnimalTFDB: a comprehensive animal transcription factor database. *Nucleic Acids Res* 2012;40 (Database issue):D144–D149. doi:10.1093/nar/gkr965.
[41] Mathelier A, Fornes O, Arenillas DJ, et al. JASPAR 2016: a major expansion and update of the open-access database of transcription factor binding profiles. *Nucleic Acids Res* 2016;44 (D1):D110–D115. doi:10.1093/nar/gkv1176.
[42] Ravasi T, Suzuki H, Cannistraci CV, et al. An atlas of combinatorial transcriptional regulation in mouse and man. *Cell* 2010;140 (5):744–752. doi:10.1016/j.cell.2010.01.044.
[43] Butler A, Hoffman P, Smibert P, Papalexi E, Satija R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat Biotechnol* 2018;36 (5):411–420. doi:10.1038/nbt.4096.
[44] Shannon P, Markiel A, Ozier O, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* 2003;13 (11):2498–2504. doi:10.1101/gr.1239303.
[45] Zhou Y, Zhou B, Pache L, et al. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat Commun* 2019;10 (1):1523. doi:10.1038/s41467-019-09234-6.
[46] Hänzelmann S, Castelo R, Guinney J. GSVA: gene set variation analysis for microarray and RNA-Seq data. *BMC Bioinform* 2013;14 (1):7. doi:10.1186/1471-2105-14-7.
[47] Efremova M, Vento-Tormo M, Teichmann SA, Vento-Tormo R. CellPhoneDB: inferring cell-cell communication from combined expression of multi-subunit ligand-receptor complexes. *Nat Protoc* 2020;15 (4):1484–1506. doi:10.1038/s41596-020-0292-x.
[48] Tang Z, Kang B, Li C, Chen T, Zhang Z. GEPIA2: an enhanced web server for large-scale expression profiling and interactive analysis. *Nucleic Acids Res* 2019;47 (W1):W556–w560. doi:10.1093/nar/gkz430.
[49] Cabili MN, Trapnell C, Goff L, et al. Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. *Genes Dev* 2011;25 (18):1915–1927. doi:10.1101/gad.17446611.
[50] Suo S, Zhu Q, Saadatpour A, Fei L, Guo G, Yuan GC. Revealing the critical regulators of cell identity in the mouse cell atlas. *Cell Rep* 2018;25 (6):1436–1445e3. doi:10.1016/j.celrep.2018.10.045.
[51] Wang L, Zhang P, Wei Y, Gao Y, Patient R, Liu F. A blood flow-dependent klf2a-NO signaling cascade is required for stabilization of hematopoietic stem cell programming in zebrafish embryos. *Blood* 2011;118 (15):4102–4110.
[52] Xue Y, Liu D, Cui G, et al. A 3D atlas of hematopoietic stem and progenitor cell expansion by multi-dimensional RNA-seq analysis. *Cell Rep* 2019;27 (5). 1567–1578.e5.