Supplemental materials for

C1Q labels a highly aggressive macrophage-like leukemia population indicating extramedullary infiltration and relapse

Li-Xue Yang[1*], Cheng-Tao Zhang[2*], Meng-Ying Yang[1*], Xue-Hong Zhang[3*], Hong-Chen Liu[2], Chen-Hui Luo[1], Yue Jiang[2], Zhang-Man Wang[2], Zhong-Yin Yang[4], Zhao-Peng Shi[5], Yi-Ci Yang[2], Ruo-Qu Wei[1], Li Zhou[5], Jun Mi[5], Ai-Wu Zhou[5], Zhi-Rong Yao[1], Li Xia[5*#], Jin-Song Yan[2#], Ying Lu[1#]

Methods

Patient samples and cell cultures

Primary samples were collected from AML patients diagnosed according to French-American-British classification at Department of Hematology of the Second Hospital of Dalian Medical University. All patients provided written consent to protocols that were approved by the Institutional Review Board and Medical Science Ethic Committee of Dalian Medical University in accordance with the Declaration of Helsinki, which enabled de-identified research use of biospecimens.

Mononuclear cell layer from BM sample was separated by Ficoll Hypaque (1114547, Axis-Shield) density gradient separation. Human monocytic AML cell lines (U937, THP-1, MV4-11 and Molm13) and human keratinocyte cells HaCaT were purchased from ATCC. CNS fibroblast HEB cells were purchased from BNCC. Primary bone marrow (BM) cells from patients or healthy donors, Molm13, U937, THP-1 and MV4-11 leukemia cells were cultured in complete RPMI 1640 medium supplemented with 10% fetal bovine serum (FBS, Gibco) and maintained in a humidified atmosphere of 5% CO_2 at 37°C. HaCaT and HEB cells were cultured in complete DMEM supplemented with 10% FBS (Gibco). Primary fibroblasts were obtained from skin or gastrointestinal biopsy and cultured in DMEM supplemented with 10% FBS (Australia Origin, Gibco).

Reagents and antibodies

Human cells were assessed using anti-human CD45 (Invitrogen, 11045942), anti-human CD33 (Invitrogen, 48033742), anti-human CD14 (Invitrogen, 17-0168-41), anti-human CD16 (Beckman Coulter, IM0650U), anti-human C1Q (ab4223, Abcam), anti-human CD123 (Biolegend, 306013), anti-human CD4 (Invitrogen, 56-0049-41) and anti-human CD56 (Biolegend, 318305) by flow cytometry. Antibodies against the following proteins were used: C1QA (Abcam, ab4223; ab189922), gC1QR (Abcam, ab24733; Proteintech,

24474-1-AP), M60.11 (anti-gC1QR, Abcam, ab24733), TGF-β1 (Proteintech, 21898-1-AP), MAFB (santa cruz, sc-376387), Collagen I (Proteintech, 14695-1-AP), Flag (Sigma, F3165), β-actin (MBL, PM053-7), FSP1 (Proteintech, 66489-1-Ig), F-actin (Thermo scientific, R415), MPO (Proteintech, 22225-1-AP), KI67 (Proteintech, 27309-1-AP) and CD3 (Proteintech, 17617-1-AP). Secondary antibodies for immunofluorescence staining are Alexa Fluor 488 and Alexa Fluor 555 (Invitrogen Molecular Probes, A-21206 and A-31572). Puromycin was from Merck/Millipore (MERCK, 540411-100mg). Galunisertib was from Beyotime (Beyotime, SF7926-10mM). Elisa kit were purchased from X-Y Biotechnology (XY9H2727).

Bulk RNA-seq data processing

For the gene expression, we mapped the sequencing data to reference genome (hg38) using STAR[1], and defined transcript coordinates according to the gene annotation format file (GTF file) from GENCODE (Release 27, GRCh38). The Gene abundances are given as Reads Per Kilobase per Million mapped reads (RPKM) with "cuffnorm" command using the Cufflinks package [2].

Flow cytometry

BM or PB (peripheral blood) samples were processed using a whole blood lysis technique. The immunophenotyping was performed by multiparameter flow cytometry (EPICSTM XL-MCL, Beckman Coulter). The leukemic blasts were stained and analyzed with anti-human CD45 (Invitrogen, 11045942), anti-mouse CD45 (Invitrogen, 12045182), anti-human CD33 (Invitrogen, 48033742).

Immunofluorescence microscopy

Molm13 cells were harvested and spotted onto slides coated with poly-l-lysine. HaCaT and fibroblasts were harvested on borosilicate cover glass. Cells were fixed with 4% (w/v) paraformaldehyde for 15min and washed with PBS for 5min for 3 times. Then cells were

permeabilized using a 0.2% Triton-PBS solution for 15min and were blocked with 3% (w/v) BSA-PBS for 30min and stained with indicated primary and secondary antibodies. Cells were concurrently stained with DAPI. Fluorescence signals were detected on a Nikon A1R confocal laser microscope.

Immunofluorescent (IF) staining was performed to examine the expression of FSP1 and gC1QR in skin, peritoneum, CNS and lymph node tissues. Briefly, slides were deparaffinized, rehydrated and subjected to antigen retrieval by heating the sample with citrate buffer (pH 6.0) in a microwave oven. After blocking, sections were incubated with gC1QR and FSP1 primary antibodies at 4℃ overnight. The slides were then incubated with fluorochrome-conjugated secondary antibodies for 1 h at RT. Finally, DAPI was applied for nuclear staining, and images were viewed using confocal fluorescence microscope (Leica).

shRNA (short hairpin RNA) knockdown cells

C1QA-knockdown cell lines were generated by shRNA of C1QA. Briefly, the sequences targeting C1QA (shC1QA-1: TCTATGGTGACCGAGGACT; shC1QA-2: GGCAACGTGGTCATCTTCG) were cloned into pGIPZ lentiviral vector (Thermo Scientific). The pGIPZ lentiviral empty vector were used as negative control. The pGIPZ lentiviral vector containing shRNA, psPAX2, and pMD2G were transfected using Lipofectamine 2000 (Thermo Fisher Scientific) into HEK293T cells. Supernatant were harvested 48 hours after transfection and infected the indicated leukemia cells using polybrene (Sigma-Aldrich). The infected cells were selected in medium containing puromycin (MERCK, 540411-100mg), and single-cell clones were selected and were evaluated by western blot.

CRISPR–Cas9 knockout (KO) cells

The gC1QR and MAFB KO cell lines were generated by sgRNA. The targeting sequences

were as followed: sg-gC1QR-1: TTAGTGCGGAAAGTTGCCG; sg-gC1QR-2: GTGCTGGGCTCCTCCGTCGC; sg-MAFB-1: GACGCAGCTCATTCAGCAGG; sg-MAFB-2: ATGGCCGCGGAGCTGAGCAT. The empty vector lentiCRISPR v2 were used as negative control. The lentivirus preparation and infection were the same as above.

Cell migration and transwell assays

Molm13 cells were suspended with RPMI 1640 medium with serum free and then seeded (1×10^6 cells in 200 μL medium) in the top chamber of transwell plate (Jet biofill, PET Tissue Culture Plate Insert 3.0um, TCS012024PET). The RPMI 1640 medium with 10% FBS was added in underneath chamber. After 24h and 48h, the chambers were removed and cells in the plates were counted under microscope at three fields per filter.

Real-time RT-PCR (Reverse Transcription-Polymerase Chain Reaction)

Total RNA was extracted from Molm13 cells by TRIzol reagent (Invitrogen), and RT was performed by FastKing RT Kit (TIANGEN, KR116) following the manufacturer's instructions. The human *β-actin* housekeeping gene was used as an internal control (F: 5'-CATCCTCACCCTGAAGTACCC-3', R: 5'-AGCCTGGATAGCAACGTACATG-3'). Real-time RT-PCR was performed using SYBR Selected Master Mix (Thermo Fisher Scientific, 447308) on QuantStudio 5 instrument (Applied Biosystems) using a primer-specific standard curve with denaturation (95°C for 10 minutes), amplification repeated 40 times (95°C for 15 seconds, 60°C for 1min). For each sample, ddCt (crossing point) values were calculated as the Ct of the target gene minus the Ct of the actin gene.

Human *C1QA*-F: 5'-TCTGCACTGTACCCGGCTA-3', human *C1QA*-R: 5'-CCCTGGTAAATGTGACCCTTTT-3'.

Human *C1QB*-F: 5'-TTCTGTGACTATGCCTACAACAC-3', human *C1QB*-R: 5'-GCCCAGTAGTGAGTTCTTGTC-3'.

Human *C1QC*-F: 5'-GAATCCCAGCCATTCCCGGGA-3', human *C1QC*-R: 5'-

GCCCTCCTCACCTGGCTCTCC-3'.

Human *CXCL2*-F: 5'-CTGCTCCTGCTCCTGGTG-3', human *CXCL2*-R: 5'-AGGGTCTGCAAGCACTGG-3'.

Human *CXCL8*-F: 5'-CTGGCCGTGGCTCTCTTG-3', human *CXCL8*-R: 5'-CCTTGGCAAAACTGCACCTT-3'.

Human *MAFB*-F: 5'-TCAAGTTCGACGTGAAGAAGG-3', human *MAFB*-R: 5'-GTTCATCTGCTGGTAGTTGCT-3'.

Human *TGF-β1*-F: 5'-CTAATGGTGGAAACCCACAACG-3', human *TGF-β1*-R: 5'-TATCGCCAGGAATTGTTGCTG-3'.

Human *TGF-β2*-F: 5'-CCCCGGAGGTGATTTCCATC-3', human *TGF-β2*-R: 5'-GGGCGGCATGTCTATTTTGTAAA-3'.

Human *TGF-β3*-F: 5'-GGAAAACACCGAGTCGGAATAC-3', human *TGF-β3*-R: 5'-GCGGAAAACCTTGGAGGTAAT-3'.

FACS (fluorescence-activated cell sorting) of Molm13^{C1Q+} and Molm13^{C1Q-} cells and BMMC

Molm13 cells were stained with anti-human C1Q (Abcam, ab4223). Cells were then sorted using FACS-AriaII cell sorter and Diva Version 5.1 software (BD Biosciences). After initial scatter-based gating dead cells, cells were gated for C1Q^{+} and C1Q^{-} cells. Bone marrow mononuclear cells (BMMC) from P-S1022 were stained with anti-human CD45 (TONBO, 20-0459) and anti-human C1Q (Abcam, ab4223). After initial scatter-based gating to dead cells, the hCD45^{+} populations were gated followed by collection of C1Q^{-} population.

Cell line derived (CDX) xenograft models

FACS-sorted Molm13^{C1Q+} and Molm13^{C1Q-} cells (5×10^6 cells per mouse) were injected intravenously into NOD.Cg-Prkdc^{scid}Il2rg^{tm1Wjl}/SzJ NSG mice (The Jackson Laboratory).

Mice were weighed every 3 days. Mice were euthanized depending on weight loss, disruption of locomotor coordination, hunching, lack of grooming, and lethargy.

Growth curve assay

FACS-sorted Molm13^{C1Q+} and Molm13^{C1Q-} cells or C1QA-knockdown Molm13 were seeded (1×10^5/ml) in 24 well plate. Each kind of cells had 3 repeat wells. Cell number was counted under microscope every day.

scRNA-seq integrated analysis and unsupervised clustering

We carried out Seurat anchor-based integration analysis to eliminate biological and technical batch effect of the samples. The total cells from patient samples (BM and cutis) were processed. First, the count data was normalized by a scale factor (10,000), and then natural-log transformed. The 2,000 highly variable genes (HVGs) detected by ‘‘FindVariableGenes’’ function were selected to identify the integration anchors based on the first 30 dimensions followed by the “IntegrateData” function. Further, we performed the principal component analysis (PCA) with “RunPCA” function to project the cells in two-dimensional space. We performed the graph-based Louvain clustering and partitioned clusters using “FindClusters” function with 30 nearest neighbors and resolution parameter equal to 0.9. Finally, we visualized gene expression and clustering results on a *t*-Distributed Stochastic Neighbor Embedding (*t*SNE) using RunTSNE function.

Mouse tissues

BM cells were prepared by washing femurs and tibias with injection needles in PBS and filtered by 70 μm filter. Spleen cells were prepared by grinding spleens in PBS and filtering cells by 70 μm filter. Red cell lysis was performed with lysis buffer before flow cytometric analysis. For histology, samples were fixed in 4% paraformaldehyde overnight before standard processing for paraffin-embedded tissues. Then, 10-μm sections were stained with hematoxylin-eosin/immunohistochemistry. The paraffin sections of mice

tissues were dewaxed, antigen retrieval and blocked with 5% BSA as per standard protocols. Sections were stained using the poly-clonal antibodies hCD45 (Invitrogen, 11045942) and hC1Q (ab4223, Abcam).

Generation of antibody against globular C1Q receptor (gC1QR)

Human gC1QR 76-100 peptide were synthesis which locates in the C1Q binding domain of gC1QR[3]. Balb/c mice were immunized by subcutaneous injection of peptide for 3 times. The mouse spleen cells and myeloma cell line sp2/0 were fused with PEG to produce hybridoma, and subcloned by limiting dilution to identify a high-producing subclone. Hybridoma supernatant was tested by ELISA to select positive subclone. The selected hybridoma cells were injected into mice intraperitoneally. The antibodies were purified from mice.

IC_{50} (half-maximal inhibitory concentration) assay

FACS-sorted Molm13^{C1Q+} and Molm13^{C1Q-} cells were seeded (1×10^5/ml) in 24 well plate. Cells were treated with etoposide, daunorubicin (DNR) and cytarabine at concentrations of 0.1, 1, 10, 100 and 1000 nM. Cell number and viability were counted under microscope after trypan blue staining. After nonlinear regression (curve fit), IC_{50} was calculated by Graphpad Prizm 8.0.1. (GraphPad Software, Inc.).

Cell counting and apoptosis assays

Cells were seeded at 1×10^6 cells/ml before treatment and counted by trypan blue (Sigma-Aldrich) exclusion. Apoptosis was quantified by staining with AnnexinV–FITC and PI (BD Biosciences, 556547).

Western blot

Protein extracts were equally loaded on 10% SDS-polyacrylamide gel, and transferred to nitrocellulose membrane (Amersham Bioscience). After blocking with 3% (w/v) BSA-PBS, the membranes were incubated with indicated antibodies, followed by horseradish

perioxidase (HRP)-linked secondary antibodies (Cell Signaling). Detection was performed by chemiluminescence phototope-HRP kit (Cell Signaling).

Chromatin Immunoprecipitation (ChIP) and quantitative PCR (Q-PCR)[4,5]

Flag-MAFB-expressing vector pCMV-3Tag-1-MAFB was transfected into Molm13 cells for 72h using Lipofectamine 2000. ChIP assay was conducted using ChIP-IT High Sensitivity kit (ACTIVE MOTIF, 53040). Briefly, 1×10^7 cells were fixed with fixation buffer for 15 min at room temperature. After washing cells with PBS buffer, resuspend cells with chromatin prep buffer. Centrifuge cells at 1250g for 3 min at 4 ℃ and resuspend cells with 500 µl ChIP buffer (5 µl PIC, 5 µl 150 mM PMSF). Cells were sonicated on ice with amplitude of 25 for 6 min. Cells were centrifuged at 16,000g for 5 min at 4 ℃ and the supernatant was collected. After the "input" samples were separated, 200 µl sample was incubated with 5 µl PIC, 4 µg anti-Flag antibody (Sigma, F3165), and 5 µL blocker at 4 ℃ overnight. Then add 30 µl agarose beads into mixture and incubate the mixture at 4 ℃ for 3h. Filter the mixture through the filter column and eluate ChIP DNA with AM4. Add 2 µl protease K into ChIP DNA and incubate mixture at 55 ℃ for 30 min and 80 ℃ for 2h. After adding 500 µl DNA purification buffer and 5 µl 3M sodium acetate, add the mixture into the purification column and centrifuge the column at 14,000 rpm for 1min. At last, add 750 µl DNA purification washing buffer into the column and collect ChIP DNA after centrifuge the column at 14,000 rpm for 2min. Quantitative PCR was conducted following method of Real-time RT-PCR.

Mass spectrometry (MS) analysis

The eluted peptides were lyophilized using a SpeedVac and resuspended in 10µl 1% formic acid/5% acetonitrile. All mass spectrometric experiments are performed on a Orbitrap Fusion LUMOS mass spectrometer (Thermo Fisher Scientific) connected to an Easy-nLC 1200 via an Easy Spray (Thermo Fisher Scientific). The peptides mixture was

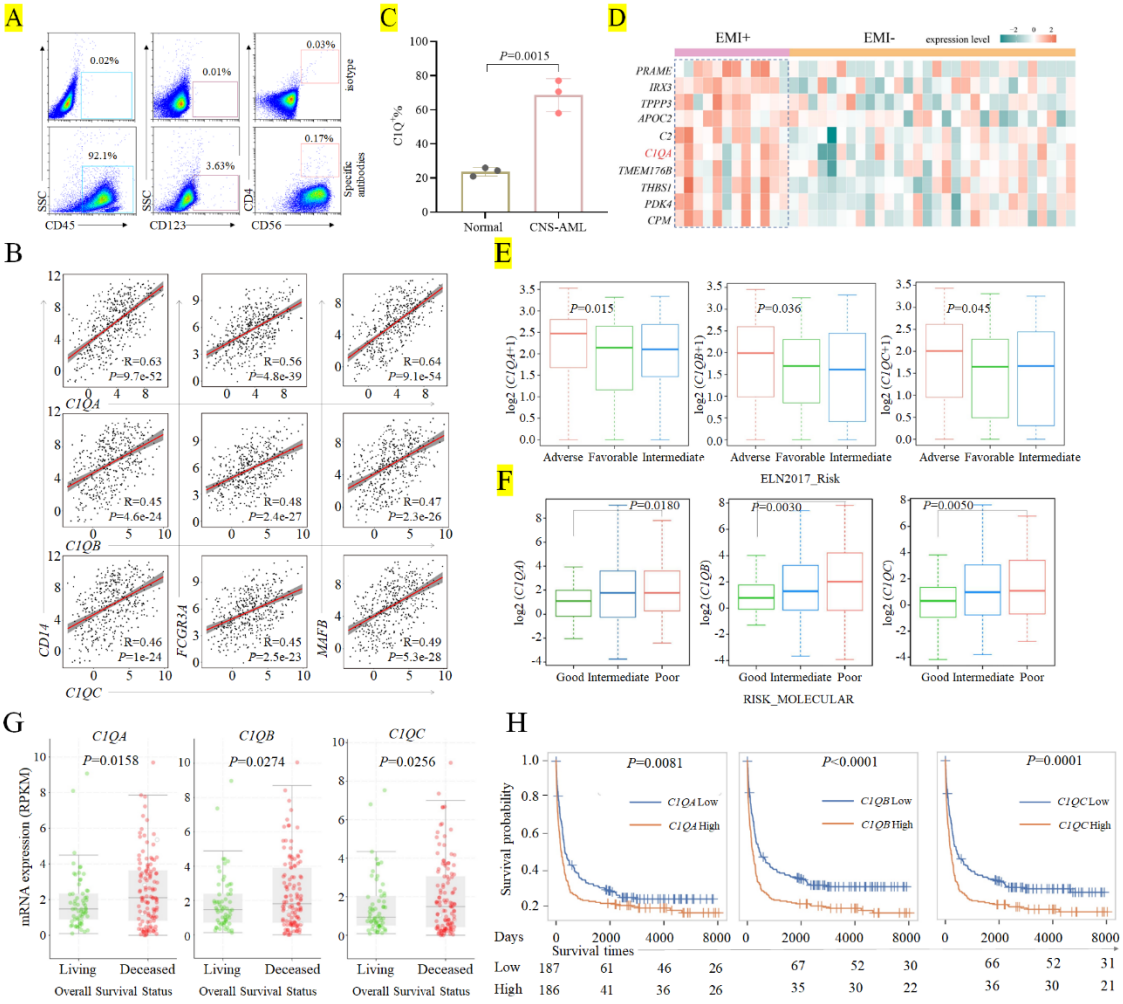
loaded onto a self-packed analytical PicoFrit column with integrated spray tip (New Objective) (75μm x 20cm length) packed with 130A C18 1.7 μm (waters) and separated within a 60-minute linear gradient from 95% solvent A (0.1% formic acid/2% acetonitrile/98% water) to 28% solvent B (0.1% formic acid/80% acetonitrile) at a flow rate of 300 nl/min at 50°C. The spray voltage was set to 2.1KV and the temperature of ion transfer capillary was 275°C, and RF lens was 40%. The mass spectrometer was operated in positive ion mode and employed in the data-dependent acquisition (DDA) mode within the specialized cycle time (3s) to automatically switch between MS and MS/MS. One full MS scan from 350 to 1500 m/z was acquired at high resolution R=120,000 (defined at m/z=400); MS/MS scans were performed at a resolution of 30,000 with an isolation window of 4 Da and higher energy collisional dissociation (HCD) fragmentation with collision energy of 30% +/- 5. Dynamic exclusion was set to 30s.

MS data processing

All MS/MS ion spectra were analyzed using PEAKS X (Bioinformatics Solutions) for processing, *de novo* sequencing, database searching and label-free quantification. Resulting sequences were searched against the UniProt Human Proteome database (downloaded 5 May 2018) with mass error tolerances of 10 ppm and 0.02 Da for parent and fragment, respectively, the digestion enzyme semiTrypsin allowed for two missed tryptic cleavages, Carbamidomethyl of cysteine specified as a fixed modification, and Oxidation of methionine, acetyl of the N-terminus and phosphorylation of tyrosine, serine and threonine as variable modifications. FDR estimation was enabled. Peptides were filtered for $-10\log P \geq 15$, and proteins were filtered for $-10\log P \geq 15$ and one unique peptide. For all experiments, this gave an FDR of <1% at the peptide-spectrum match level. Proteins sharing significant peptide evidence were grouped.

Survival analysis

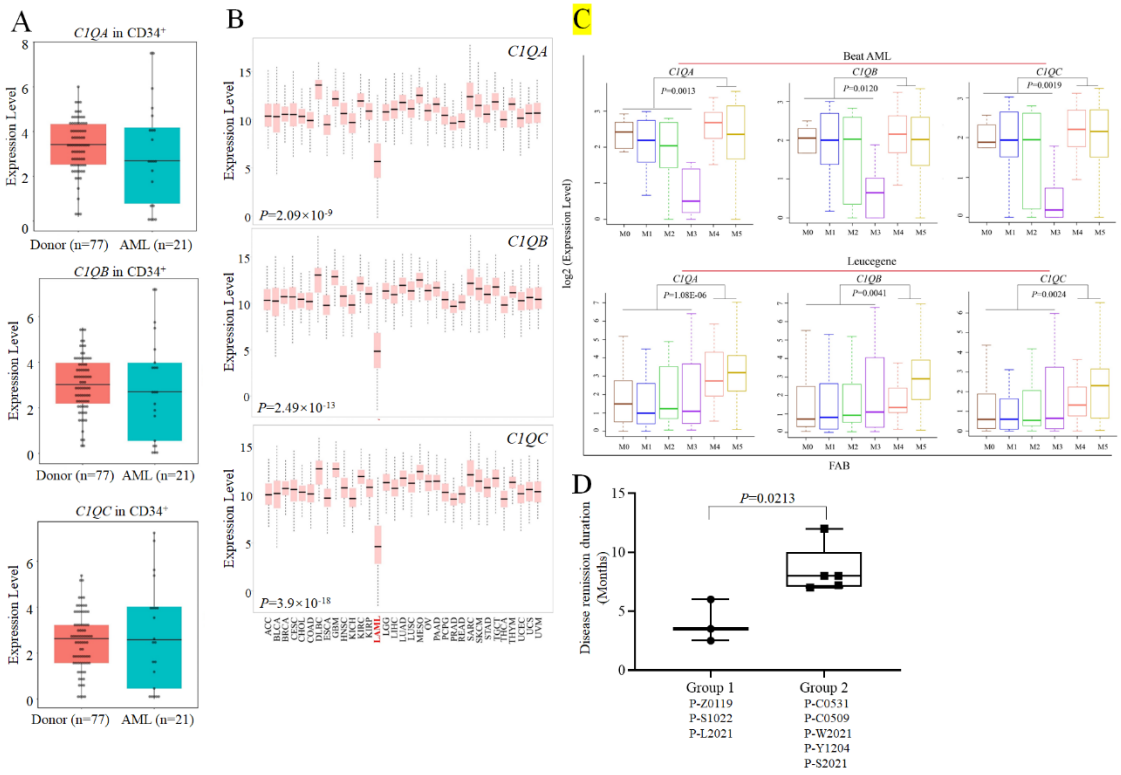
We performed the survival analysis (*C1QA, C1QB, C1QC, CXCL2*) on the website (https://data.leucegene.iric.ca/survival) based on the Leucegene project cohort consisting of 373 diagnostic AML samples (excluding acute promyelocytic leukemia). Then, we defined the cluster 5 feature as mean log2 (RPKM+1) normalized expression of the following genes (*C1QA*, *C1QB*, *C1QC*, *MAFB*, *CD14* and *FCGR3A*). Further, we carried out the survival analysis on the bulk RNA cohorts from the Cancer Genome Atlas (TCGA) and BEATAML1.0[6]. The gene expression and survival data were downloaded by R function "GDCdownload" in TCGAbiolinks package (v2.18.0)[6]. We divided AML samples into the high and low expression groups by the median value. Kaplan-Meier survival curve with the events table showing differences in survival time was plotted by survminer package (v0.4.9). The survival risk and statistical significance were determined according to the Hazard ratio (HR) and log-rank *P* values reported by R package survival (v3.2-11).



Supplemental Figure 1. Expression level of C1Q is related to extramedullary infiltration.

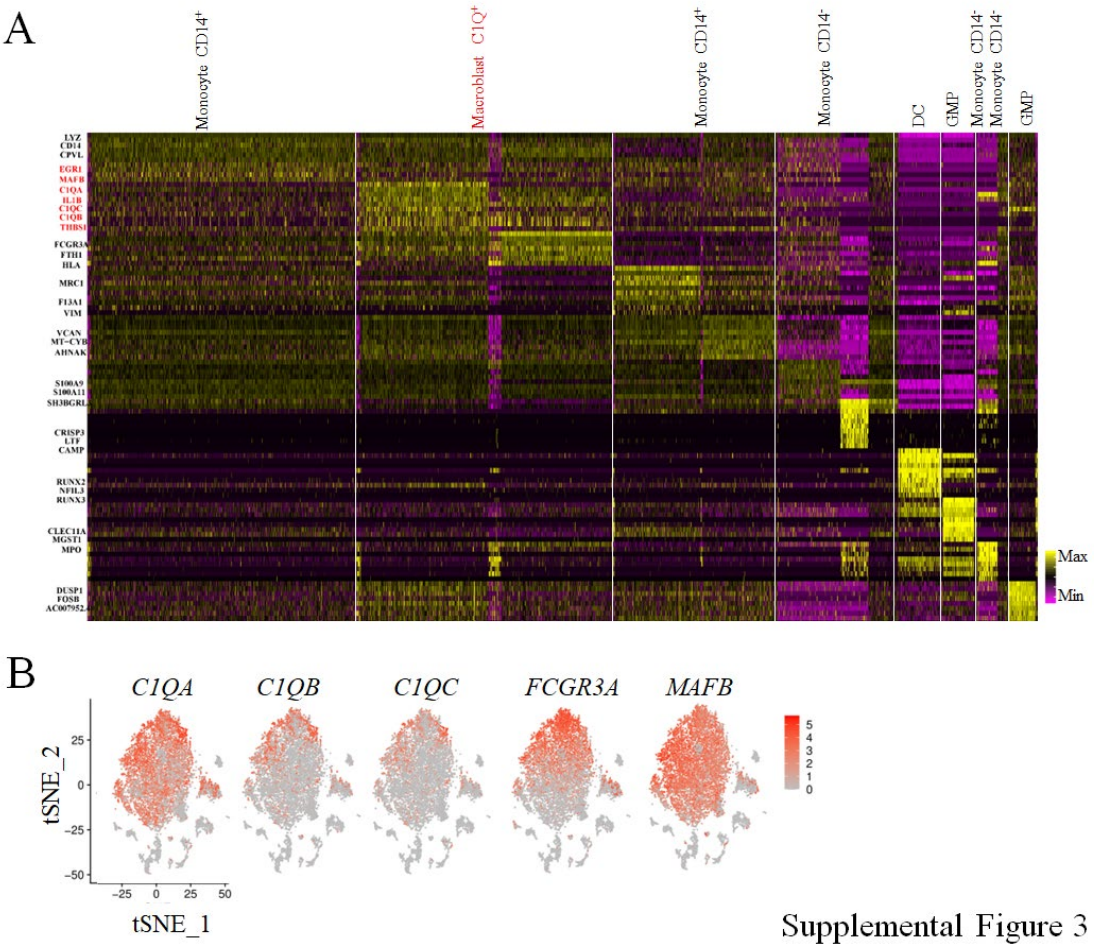
(A) Expression levels of marker genes of blastic plasmacytic dendritic cell neoplasm (CD123, CD4 and CD56) in bone marrow mononuclear cell (BMMC) of P-S1022 were analyzed by flow cytometry. (B) Matched scatterplots of *C1QA*, *C1QB*, *C1QC* and other signature genes of C1Q^{+} macroblast including *CD14*, *FCGR3A* and *MAFB*. R was calculated by Pearson correlation coefficient. (C) The percentages of C1Q^{+} cells in cerebrospinal fluids (CSF) from AML patients with (CNS-AML) or without (Normal) central nervous system (CNS) infiltration were evaluated by flow cytometry. (D) Heatmap of the top 10 genes upregulated in EMI+ *vs* EMI- patients. (E-F) Box plots for the

expression of *C1QA*, *C1QB* and *C1QC* grouped by (E) ELN2017_Risk and (F) RISK_MELOCULAR in AML patients from BeatAML dataset (n = 441) and TCGA dataset (n = 173). *P* values were calculated by two-sample one-tailed t-test. (G) Overall survival status based on mRNA expression level of *C1QA*, *C1QB* and *C1QC* in AML patients from the Cancer Genome Atlas (TCGA). RPKM: reads per kilobase million. *P* value was calculated by two-sample t-test. (H) Overall survival of patients from the Leucegene database, performed in Xena browser (https://xena.ucsc.edu). *P* values were calculated by Kaplan-Meier log-rank test.



Supplemental Figure 2. Characterization of C1Q expression in AML patients and multiple cancer types.

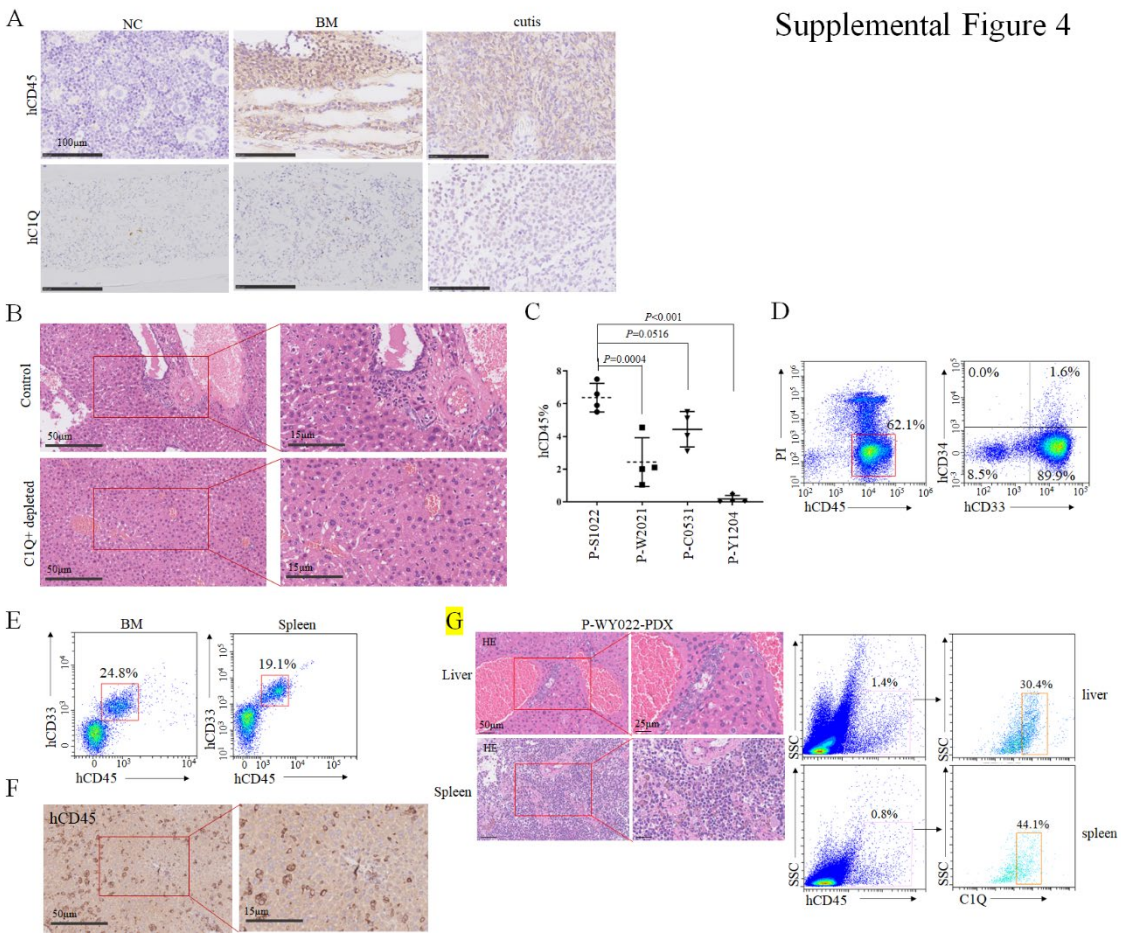
(A) RNA-seq analysis of sorted $CD34^+$ cells samples collected from *de novo* AML patients (AML, n=21) and healthy donors (HD, n=77). The mRNA expression level of *C1QA*, *C1QB* and *C1QC* are shown. (B) *C1QA*, *C1QB* and *C1QC* level across all the cancer types in TCGA are shown. (C) *C1QA*, *C1QB* and *C1QC* level across all the FAB subtypes of AML in BeatAML (upper) and TCGA (lower), respectively. (D) Disease remission duration to chemotherapy of AML individuals with high (Group 1) or low (Group 2) level of C1Q. *P* value was calculated by two-sample t-test.



Supplemental Figure 3. Featured genes of clusters identified through scRNA-seq in P-Z0119.

(A) Heatmap of the top 10 genes marking indicated clusters (from Figure 1E). (B) *t*-SNE projection of representative marker genes of relapse-associated clusters identified in P-Z0119.

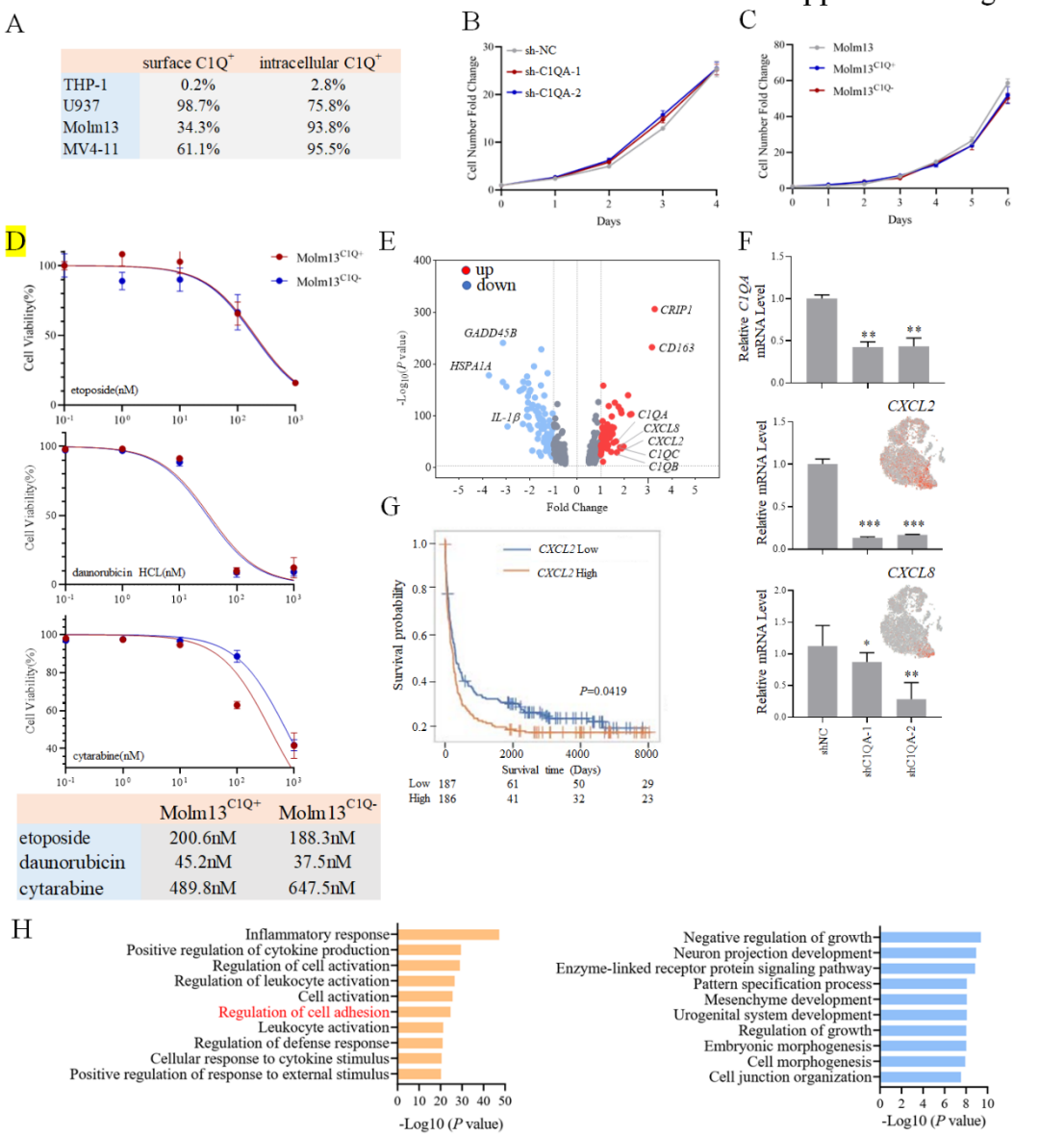
Supplemental Figure 4



Supplemental Figure 4. C1Q^{+} population is tumorigenic.

(A) The leukemic invasions in BM and cutis were analyzed by immunohistochemistry (IHC) staining of human CD45 (hCD45) and hC1Q. NC: healthy control mice. (B) The leukemic invasions in liver were analyzed by hematoxylin and eosin (H&E) staining. (C) Flow cytometry analysis of hCD45 of Passage 1 recipient mice of indicated patients at eight weeks post transplantation. *P* value was calculated by two-sample t-test. (D-E) Flow cytometry analysis of hCD45, hCD33 and hCD34 expression on cells recovered from PDX recipient mice constructed using cutis cells of P-S1022 (D) or cutis of Passage 1 PDX mice (E). (F) The leukemic invasions in liver of Passage 2 PDX mice were analyzed by IHC staining of hCD45. (G) The leukemic invasions in liver and spleen of Passage 1 recipient mice of P-WY022 were analyzed by H&E staining. Flow cytometry analysis of human CD45 (hCD45) and C1Q expression on cells were shown on the right.

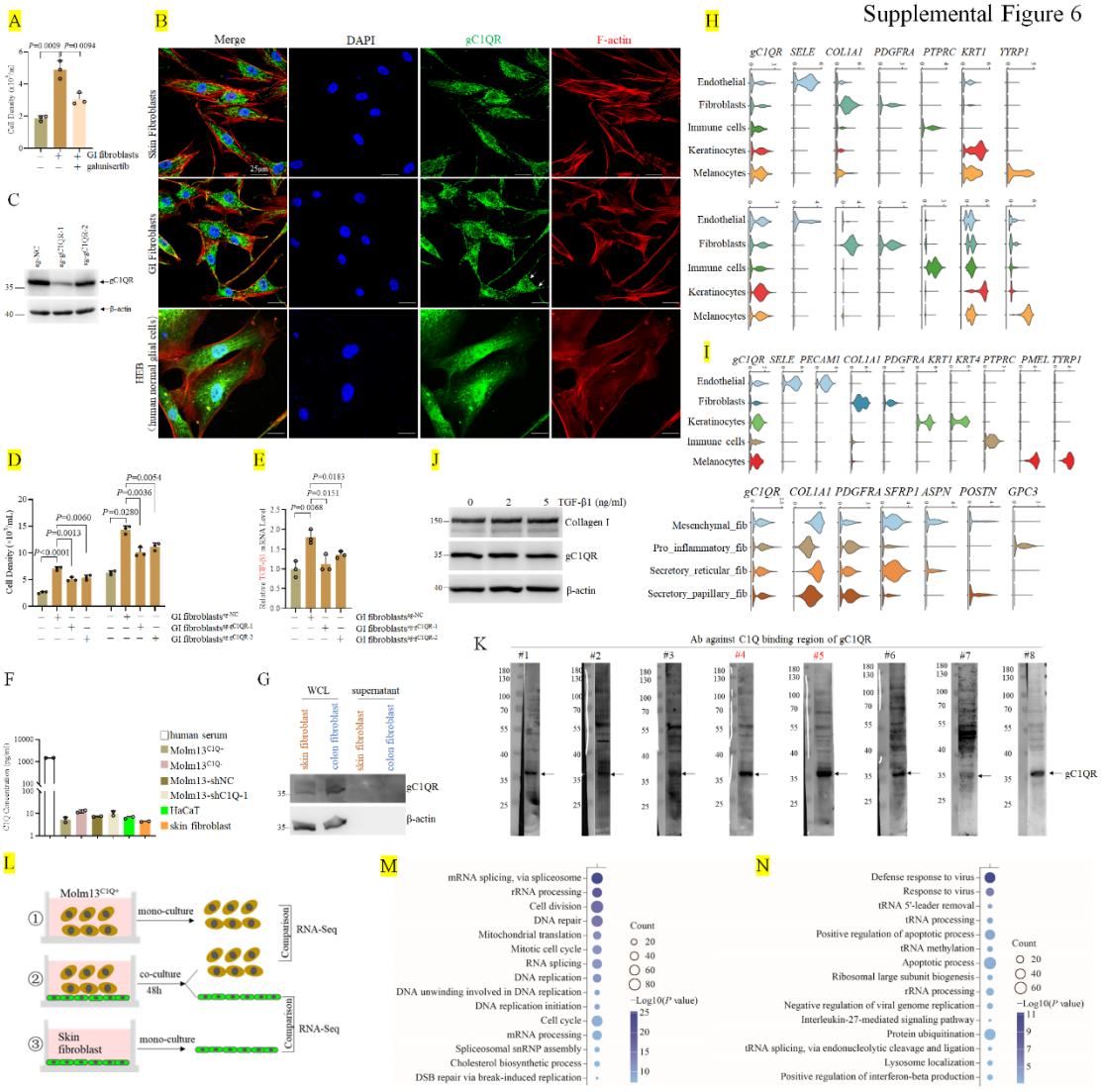
Supplemental Figure 5



Supplemental Figure 5. Molm13^{C1Q+} cells display migration-associated features in contrast to Molm13^{C1Q-}.

(A) Percentages of membrane and intracellular C1Q^{+} cells in human monocytic leukemia cell lines were analyzed by flow cytometry. (B) Growth curve of Molm13 cells upon *C1QA* deletion. sh-NC: Molm13 transfected with empty vector (pGIPZ) using for shRNA knockdown. Cell number was counted by trypan blue. (C) Growth curve of Molm13^{C1Q+}, Molm13^{C1Q-} and parental unsorted Molm13. Cell number was counted by trypan blue. (D) IC_{50} of etoposide, daunorubicin and cytarabine in Molm13^{C1Q+} and Molm13^{C1Q-} cells. Cell

number and viability were counted by trypan blue. (E) Volcano plot of DEGs that are upregulated (red) or downregulated (blue) in $C1Q^{+}$ macroblast of cutis vs BM. Relevant DEGs identified in the pathways are labeled. *P* value was derived by Wilcoxon rank-sum test. (F) Q-PCR analysis of *C1QA*, *CXCL2* and *CXCL8* in Molm13 cells upon *C1QA* depletion. Expression of *CXCL2* and *CXCL8* were projected onto the *t*-SNE plot (top). *$P < 0.05$; **$P < 0.01$; ***$P < 0.001$. (G) Overall survival status based on mRNA expression level of *CXCL2* in AML patients from Leucegene group. The *P* value was from Kaplan-Meier log-rank test. (H) RNA-seq were performed between $Molm13^{C1Q+}$ and $Molm13^{C1Q-}$ cells (n=3) followed by GO (Gene Oncology) analysis. Orange: top 10 enriched pathways based on upregulated genes in $Molm13^{C1Q+}$ *vs* $Molm13^{C1Q-}$. Blue: top 10 enriched pathways based on downregulated genes in $Molm13^{C1Q+}$ *vs* $Molm13^{C1Q-}$.

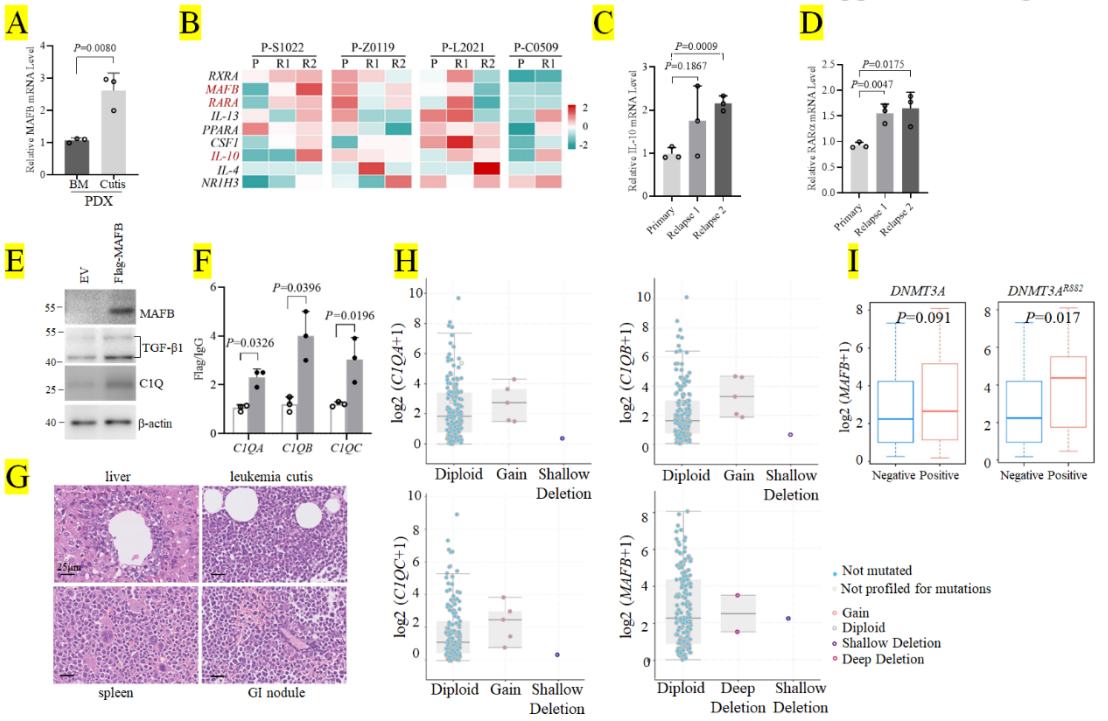


Supplemental Figure 6. C1Q^{+} leukemia cells communicate with fibroblasts via surface C1Q-gC1QR signaling.

(A) Transwell assay of Molm13^{C1Q+} cells upon GI fibroblast co-culture and TGF-β1 inhibitor galunisertib treatment. GI: gastrointestinal. (B) Immunofluorescent staining of gC1QR and F-actin in permeabilized skin fibroblasts, colon fibroblasts and HEB. (C) Western blot analysis of skin fibroblasts upon gC1QR depletion. (D) Transwell assay of Molm13^{C1Q+} cells co-cultured with GI fibroblasts with or without gC1QR deletion for 24h (left) and 48h (right). (E) Q-PCR analysis of *TGF-β1* in Molm13^{C1Q+} cells co-cultured with or without GI fibroblasts (with or without gC1QR deletion). (F) C1Q concentration

was measured in human serum and supernatant of indicated cells by ELISA. (G) Western blot analysis of gC1QR using whole cell lysate (WCL) and supernatant fractions prepared from skin fibroblast and colon fibroblast. (H-I) Expression levels (*x* axis) of marker genes of indicated cell types were analyzed with three publicly available scRNA-seq datasets. Violin plots show the distribution of normalized expression levels of genes and are color-coded on the basis of cell types. (J) Western blot analysis of Collagen I and gC1QR of skin fibroblasts upon TGF-β1 treatment. (K) Generation of antibodies against C1Q binding sequence of gC1QR. Western blots showed the specificity of these antibodies targeting gC1QR. (L) Flow chart of RNA-Seq performed on mono- and co-cultured Molm13^{C1Q+} cells and skin fibroblasts. Briefly, Molm13^{C1Q+} cells and skin fibroblasts were mono- or co-cultured for 48h. RNA of Molm13^{C1Q+} cells and skin fibroblasts were isolated and subjected to RNA-Seq. Comparison was performed between mono- and co-cultured Molm13^{C1Q+} cells, as well as mono- and co-cultured skin fibroblasts. (M-N) Top 15 biological processes for downregulated DEGs in co-cultured Molm13^{C1Q+} (M) and skin fibroblasts (N) were shown.

Supplemental Figure 7



Supplemental Figure 7. MAFB modulates C1Q expression and contributes to EMI.

(A) *MAFB* mRNA level in BM and leukemia cutis from PDX mice was measured by Q-PCR. *P* value was calculated by two-sample t-test. (B) Relative expression heatmap of genes that regulate expression of MAFB on longitudinally collected samples of indicated AML patients. P: primary; R: relapse. (C-D) The mRNA levels of *IL-10* (C) and *RARα* (D) were measured by Q-PCR in leukemia blasts of P-S1022 at the time point of primary, relapse 1, and relapse 2, respectively. *P* value was calculated by two-sample t-test. (E) Western blot analysis of TGF-β1 and C1Q upon Flag-MAFB overexpression in Molm13 cells. (F) ChIP assay was conducted using Molm13 cells expressing Flag-MAFB fusion protein. Gene promoter sequences were amplified and analyzed by Q-PCR. *P* value was calculated by two-sample t-test. (G) The leukemic invasions in liver, spleen, cutis and GI nodule of Molm13^{C1Q+} cells recipient mice were analyzed by H&E staining. (H) Expression levels of *C1QA*, *C1QB*, *C1QC* and *MAFB* of AML patients from indicated cytogenetic and molecular subsets. (I) Expression levels of *MAFB* in AML patients with

or without *DNMT3A* mutation (left) or *DNMT3A*^{R882} mutation (right).

References

1. Dobin A, Davis CA, Schlesinger F, et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics.* Jan 1 2013;29(1):15-21. doi: 10.1093/bioinformatics/bts635.
2. Trapnell C, Williams BA, Pertea G, et al. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat Biotechnol.* May 2010;28(5):511-515. doi: 10.1038/nbt.1621.
3. Peerschke EI, Ghebrehiwet B. cC1qR/CR and gC1qR/p33: observations in cancer. *Mol Immunol.* Oct 2014;61(2):100-109. doi: 10.1016/j.molimm.2014.06.011.
4. Hamada M, Nakamura M, Tran MT, et al. MafB promotes atherosclerosis by inhibiting foam-cell apoptosis. *Nat Commun.* 2014;5:3147. doi: 10.1038/ncomms4147.
5. Tran MTN, Hamada M, Jeon H, et al. MafB is a critical regulator of complement component C1q. *Nat Commun.* Nov 22 2017;8(1):1700. doi: 10.1038/s41467-017-01711-0.
6. Tyner JW, Tognon CE, Bottomly D, et al. Functional genomic landscape of acute myeloid leukaemia. *Nature.* Oct 2018;562(7728):526-531. doi: 10.1038/s41586-018-0623-z.

Supplemental Table 4. Multivariate Cox regression analysis of DFS in TCGA cohort.
CBF_Fusion: Positive for the RUNX1-RUNX1T1 or CBFB-MYH11 fusion.

Prognostic factors		β	HR	95% CI		*P*
				Lower	Upper	
***C1QA* expression** (*Reference: Low*)	High *C1QA* expression	**1.8**	**6.05**	**1.47**	**24.85**	**.013**
Age (year) (*Reference: < 60*)	> 60 years	0.25	1.29	0.71	2.32	.401
Sex (*Reference: Female*)	Male	-0.38	0.68	0.41	1.13	.134
WBC (10^9) (*Reference: < 30.5*)	$> 30.5 \times 10^9$	**1.18**	**3.25**	**1.76**	**5.98**	**<.001**
FAB (*Reference: M0*)	M1	0.29	1.34	0.48	3.77	.577
	M2	0.16	1.17	0.42	3.27	.766
	M3	-0.01	0.99	0.18	5.57	.991
	M4	0.03	1.03	0.36	2.93	.995
	M5	-0.99	0.37	0.09	1.61	.186
Risk_Molecular (*Reference: Good*)	Intermediate	**1.33**	**3.77**	**1.25**	**11.36**	**.018**
	Poor	0.7	2.02	0.72	5.65	.179
CBF_Fusion (*Reference: Negative*)	Positive	-0.36	0.7	0.3	1.64	.41
FLT3 (*Reference: Negative*)	Positive	**0.7**	**2.02**	**1.11**	**3.66**	**.021**
NPM1 (*Reference: Negative*)	Positive	-0.66	0.52	0.24	1.12	.095
CEBPA (*Reference: Negative*)	Positive	-0.27	0.76	0.26	2.2	.615
TP53 (*Reference: Negative*)	Positive	0.32	1.37	0.23	8.25	.73
RUNX1 (*Reference: Negative*)	Positive	-0.16	0.86	0.32	2.26	.754
DNMT3A (*Reference: Negative*)	Positive	0.08	1.09	0.58	2.04	.8
KIT (*Reference: Negative*)	Positive	0.27	1.31	0.38	4.52	.67

Multivariate Cox analysis of DFS in *de novo* AML patients in TCGA dataset (n=173) according to *C1QA* expression, age, sex, WBC counts, FAB classification, risk_molecular, CBF_Fusion and genetic characteristics known for their prognostic importance. Patient number and percentage, regression coefficient (β), hazard ratio (HR) and *P*-values are shown for each parameter.