Supplementary Information: The landscape of KMT2A-PTD AML: Concurrent mutations, gene expression signatures, and clinical outcome.

Adil S.A. Al Hinai, Marta Pratcorona, Tim Grob, François G. Kavelaars, Elena Bussaglia, Mathijs A. Sanders, Josep F. Nomdedeu, Peter J.M. Valk

# Materials and method:

## Patient material, RNA isolation cDNA synthesis, and DNA isolation

Bone marrow aspirates or peripheral blood samples of cohorts of patients with AML were collected after a written informed consent in accordance with the Declaration of Helsinki. These procedures were performed as previously described [1]. DNA was isolated as per the standard molecular techniques.

## Screening for *MLL*-PTD by conventional PCR and Sanger sequencing

The primer set used for the detection of partial tandem duplication (PTD) of the *MLL* gene are flanking exon 8 and exon 4 (Supplementary figure S1). The primer set used to detect the common type PTD of the *MLL* gene (ex9/ex3, ex10/ex3, and ex11/ex3) were adopted from Caligiuri *et al*., 1998 [2]. The sequence of the forward primer MLL-ex8: 5’-GGAAGTCAAGCAAGCAGGTC and reverse primer MLL-ex4: 5’-AGGAGAGAGTTTACCTGCTC-3’ were used at 200nM for PCR in a total reaction volume of 50µL using 3µL cDNA, 1×PCR buffer, 1.5mM MgCl2, 200nM dNTPs (Thermo Fisher, Walthan, MA) and 200nM of each primer. The thermocycling condition was 94°C for 3min, followed by 35 cycles of 94°C for 15sec, 63°C for 1min, 72°C for 1min, ending with a final extension for 10min at 72°C. The PCR product was then visualized on a 2% agarose gel. Serial dilution of the EOL-1 cell line was taking along each PCR as a positive control. The PCR products of the positive cases were purified using MultiScreen plates (Merck Millipore, Amsterdam, the Netherlands) and Sanger sequenced on the ABI PRISM 3130*xl* genetic analyzer (Thermo Fisher, Waltham, MA).

## Next-Generation Sequencing

Genomic DNA of 50ng used for generating libraries using multiplexed oligonucleotide probes as per the protocol described by the manufacturer of TruSight Myeloid Sequencing Panel (Illumina, San Diego, CA). Subsequently, sample-specific indexes were ligated to each library. The libraries were then pooled, and paired-end sequenced (2× 221-bp) on the Illumina HiSeq 2500 in Rapid Mode according to manufacturer’s recommendation (Illumina, San Diego, CA). Bioinformatics analysis was performed as previously described [3].

## EVI1 Expression assay

EVI1 expression was assessed by RQ-PCR assay as previously described [4].

## Gene Expression Profiling

Samples were analyzed using Affymetrix Human Genome U133Plus2.0 GeneChips (Affymetrix, Santa Clara, CA, USA) as previously described [5].

## Statistical analysis

The Fisher’s exact test and the Mann-Whitney U test were used for categorical and continuous variables, respectively. Survival time was calculated from the date of diagnosis until the date of death from any cause or censoring. Event-free survival was defined as the date of diagnosis to relapse or death from any cause or censoring, whichever came first. Survival estimates were graphically represented with the Kaplan-Meier method. The log-rank test was used to compare the survival distributions of the groups and the Cox proportional hazards model for multivariable analysis. P-values are two-sided and p-values <0.05 were considered statistically significant. Statistical analysis was performed in STATA statistical software, Release 15.1 (Stata, College Station, TX, USA).

# References

1 Valk PJ, Verhaak RG, Beijen MA, Erpelinck CA, Barjesteh van Waalwijk van Doorn-Khosrovani S, Boer JM, Beverloo HB, Moorhouse MJ, van der Spek PJ, Lowenberg B, Delwel R: Prognostically useful gene-expression profiles in acute myeloid leukemia. N Engl J Med 2004;350:1617-1628.

2 Caligiuri MA, Strout MP, Lawrence D, Arthur DC, Baer MR, Yu F, Knuutila S, Mrozek K, Oberkircher AR, Marcucci G, de la Chapelle A, Elonen E, Block AW, Rao PN, Herzig GP, Powell BL, Ruutu T, Schiffer CA, Bloomfield CD: Rearrangement of all1 (mll) in acute myeloid leukemia with normal cytogenetics. Cancer Res 1998;58:55-59.

3 Jongen-Lavrencic M, Grob T, Hanekamp D, Kavelaars FG, Al Hinai A, Zeilemaker A, Erpelinck-Verschueren CAJ, Gradowska PL, Meijer R, Cloos J, Biemond BJ, Graux C, van Marwijk Kooy M, Manz MG, Pabst T, Passweg JR, Havelange V, Ossenkoppele GJ, Sanders MA, Schuurhuis GJ, Lowenberg B, Valk PJM: Molecular minimal residual disease in acute myeloid leukemia. N Engl J Med 2018;378:1189-1199.

4 Groschel S, Lugthart S, Schlenk RF, Valk PJ, Eiwen K, Goudswaard C, van Putten WJ, Kayser S, Verdonck LF, Lubbert M, Ossenkoppele GJ, Germing U, Schmidt-Wolf I, Schlegelberger B, Krauter J, Ganser A, Dohner H, Lowenberg B, Dohner K, Delwel R: High evi1 expression predicts outcome in younger adult patients with acute myeloid leukemia and is associated with distinct cytogenetic abnormalities. J Clin Oncol 2010;28:2101-2107.

5 Verhaak RG, Wouters BJ, Erpelinck CA, Abbas S, Beverloo HB, Lugthart S, Lowenberg B, Delwel R, Valk PJ: Prediction of molecular subtypes in acute myeloid leukemia based on gene expression profiling. Haematologica 2009;94:131-134.