A gene expression signature for defining aggressive phenotype and prognosis in intermediate-risk acute myeloid leukemia

Ailing Deng[1]

[1] Medicine College of Nankai university, Tianjin, China

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.1833v1 | CC-BY 4.0 Open Access | rec: 7 Mar 2016, publ: 7 Mar 2016

Abstract

Intermediate-risk AML is a group of heterogeneous disease, complete and accurate understanding of their molecular pattern are urgent. We downloaded gene expression profile from TCGA, and performed the differential gene expression analysis of 98 intermediate-risk AML patients according to the different clinical outcome. By studying the overlap of differential genes based on OS and EFS classification, we identified 6 distinct genes (SPANXC, ADAMTS15, C8B, FAM183A, FLJ42875 and PXDN). Only FLJ42875 and PXDN were relatively widely expressed in intermediate-risk AML. Moreover, Q1 (0-25%) of PXDN was associated significantly shortened OS and EFS (P=0.0001, P=0.0032, respectively), while OS and EFS of Q2 (25-50%) of FLJ42875 expression group were significantly longer than that of low FLJ42875 expression patients (P=0.0157, P=0.0074, respectively). In the multivariable model of OS, PXDN Q1 group had a shorter OS and EFS ($P < 0.001$, P =0.009; respectively). While FLJ42875 Q2 group had a tendency of prolonged OS and EFS (P=0.050, P=0.023; respectively). Our results suggest that FLJ42875 and PXDN may play a role in the leukemogenesis of intermediate-risk AML.

Key words: FLJ42875, PXDN, intermediate-risk AML

Introduction

Cytogenetic lesions is used as diagnostic and prognostic markers and also to elucidate the molecular pathogenesis of acute myeloid leukemia (AML) for more than three decades (Mrozek et al. 2004; Rowley 2008). However, nearly half of the AML patients don't have apparent cytogenetic structural abnormalities, they are intermediate-cytogenetic risk (Martelli et al. 2013). However, their clinical outcomes are greatly heterogeneous and lacks effective prognostic indicators at the same time. Some of these patients respond well to chemotherapy, while others not. Therefore, recent studies have focused on improving classification of intermediate-risk AML patients. Mutations, aberrantly expressed genes, microRNAs and changes in DNA methylation are potential prognostic markers. For example, mutations of CEBPA and NPM1 are associated with favorable outcomes (Ahn et al. 2016; Han et al. 2013; Li et al. 2015), whereas mutations of FLT3-ITD, WT1, MLL, TET2 and DNMT3A are associated with an unfavorable clinical outcome (Ley et al. 2010; Mikulasova et al. 2010; Ribeiro et al. 2012; Stirewalt & Radich 2003). High expression of WT1, ERG, and DNMT3B as well as low expression of LEF1 have also been shown to be unfavorable prognostic factor (Metzeler et al. 2009; Metzeler et al. 2012; Schulze et al. 2016; Yi-Ning et al. 2015). However, none of the present classification strategies conform to clinical needs entirely, which suggests that a more complete and accurate understanding of the genetic

and epigenetic alterations associated with the pathogenesis of intermediate-risk AML is crucial.

In order to further identify molecular determinants associated with aggressive behavior of intermediate-risk AML, we downloaded gene expression profile from TCGA, and we compared the gene expression pattern characterized by the distinct clinical outcome.

Methods

Patients and treatment

98 patients with previously untreated intermediate-risk AML (median age, 58 years; range, 21–88 years) were studied, all of them were diagnosed and received treatment according to NCCN guidelines between November of 2001 and March of 2010. 51 patients (52.0%) were aged <60 years (younger patients) and 47 patients (48.0%) were ≥ 60 years (older patients). The diagnosis and risk stratification of AML were based on NCCN Guidelines. All patients were assessed for somtatic mutations such as IDH1, NPM1, FLT3 and gene expression. Clinical, gene expression and somatic mutations profiles of all primary AML cases could be publicly downloaded from the TCGA project via the data portal on 01/10/2015. This research was approved by the Washington University Human Studies Committee (WU HSC #01 -1014), and informed consent from all patients was obtained in accordance with the Declaration of Helsinki(Cancer Genome Atlas Research 2013).

RNA Sequencing analyses

Gene expression data have been previously published. Briefly, gene expression data were obtained using RNA Sequencing. All the design and quality control for this experiment were based on the standard protocols. Patients with OS and EFS values above the median of all patients were subdivided into the OS^{longer} and EFS^{longer} group while others were into $OS^{shorter}$ and $EFS^{shorter}$ group. Gene expression levels were also determined from the normalized data.

Statistical analyses

The time from date of diagnosis to removal from study due to absence of complete remission, relapse or death defined EFS, and the time from date of diagnosis to death due to any cause defined OS. Firstly, we classified the 98 AML patients into two groups based on OS and EFS value respectively. Log-rank Test was used to estimate the association between gene expressions and EFS and OS of the patients, which were further validated using Gehan-Breslow-Wilcoxon test. Student's t-test and multiple hypothesis correction (False Discovery Rate, FDR) was used to identify differences in gene expression profiles. The statistical cutoff values were a fold-change (FC) ≥ 1.5 and an adjusted P-value < 0.05. All analyses were performed using the R3.2.2 software packages and GraphPad Prism 5.0.

Results

Clinical characteristics of primary intermediate-risk AML patients

In the cohort of 98 primary intermediate-risk AML patients, 6 were M0, 25 were M1, 24 were M2, 1 was M3, 24 were M4, 16 were M5, 1 was M7, 1 was not classified. The median age was 58 years, ranging from 21 to 88 years. 51 patients (52.0%) were aged <60 years (younger patients) and 47 patients (48.0%) were ≥ 60 years (older patients).49(50%) patients were male and 49(50%) patients were female. Pre-treatment hemoglobin was 9.629±0.1621g/L, WBC counts was (45.63±5.348)*10^9/L, and PLT counts was (72.60±6.057)*10^9/L. The median OS was 16.35(0-100.5) months, while median EFS was 9.4(0-100.5) months (Table 1).

Table 1. Clinical characteristics of intermediate-risk AML patients.

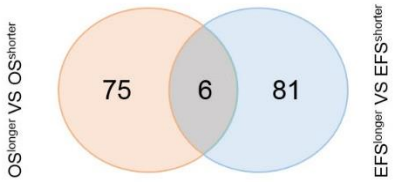
Variable	Intermediate-risk AML patients (n=98)
Age, y(range)	58(21-88)
≥ 60 y, no (%)	47(48.0)
Female	49(50%)
Hemoglobin(g/L)	9.629±0.1621
WBC(10^9/L)	45.63±5.348
PLT(10^9/L)	72.60±6.057
Blast percentage in BM (%)	80.26±2.129
FAB subtype, no (%)	
M0	6(6.1)
M1	25(25.5)
M2	24(24.5)
M3	1(1.0)
M4	24(24.5)
M5	16(16.3)
M7	1(1.0)
Not classified	1(1.0)
Mutation, no (%)	
FLT3	33(33.7)
NPM1	42(42.9)
FLT3+NPM1	21(21.4)
IDH1	10(10.2)

IDH2	12(12.2)
NRAS	9(9.2)
KRAS	4(4.1)

Differential gene expression in primary intermediate-risk AML samples

OS^{longer}, $OS^{shorter}$, EFS^{longer}, $EFS^{shorter}$ groups were subdivided as described above. Normalized RNA Sequencing data was downloaded from TCGA database and differential expression analysis was performed to obtain expression of AML patients. 75 differential genes were detected after the comparison of OS^{longer} and $OS^{shorter}$ groups, while 81 differential genes were discovered in that of EFS^{longer} and $EFS^{shorter}$ groups (Supplemental Table S1 and S2). The overlaps were six genes including SPANXC, ADAMTS15, C8B, FAM183A, FLJ42875 and PXDN (Figure 2). No strong association among them was found.

A



B

Gene	$Log_2FC:OS^{longer}/OS^{shorter}$	P-value	$Log_2FC:EFS^{longer}/EFS^{shorter}$	P-value
ADAMTS15	-1.64824	0.021	-1.6474	0.026
C8B	-2.15338	0.016	-2.52137	0.025
FAM183A	-2.62104	0.047	-2.62104	0.047
FLJ42875	1.530341	0.029	2.195344	0.001
PXDN	-2.97264	<0.001	-1.69152	0.034
SPANXC	3.410173	0.015	3.410173	0.015

Figure 1. Differential genes in intermediate-risk AML patients. (A)Overlap of OS and EFS differential genes. (B)List of differential genes.

Expression of differential genes in AML samples

Next, we analyzed the RNA Sequencing data to obtain the expression of six differential genes mentioned above in intermediate-risk AML patients. As indicated in the Figure 2, SPANXC, ADAMTS15, C8B,

FAM183A were expressed narrowly at a low level in intermediate-risk AML, they could hardly be detected in most patients. However, FLJ42875 and PXDN could be detected widely in intermediate-risk AML patients.

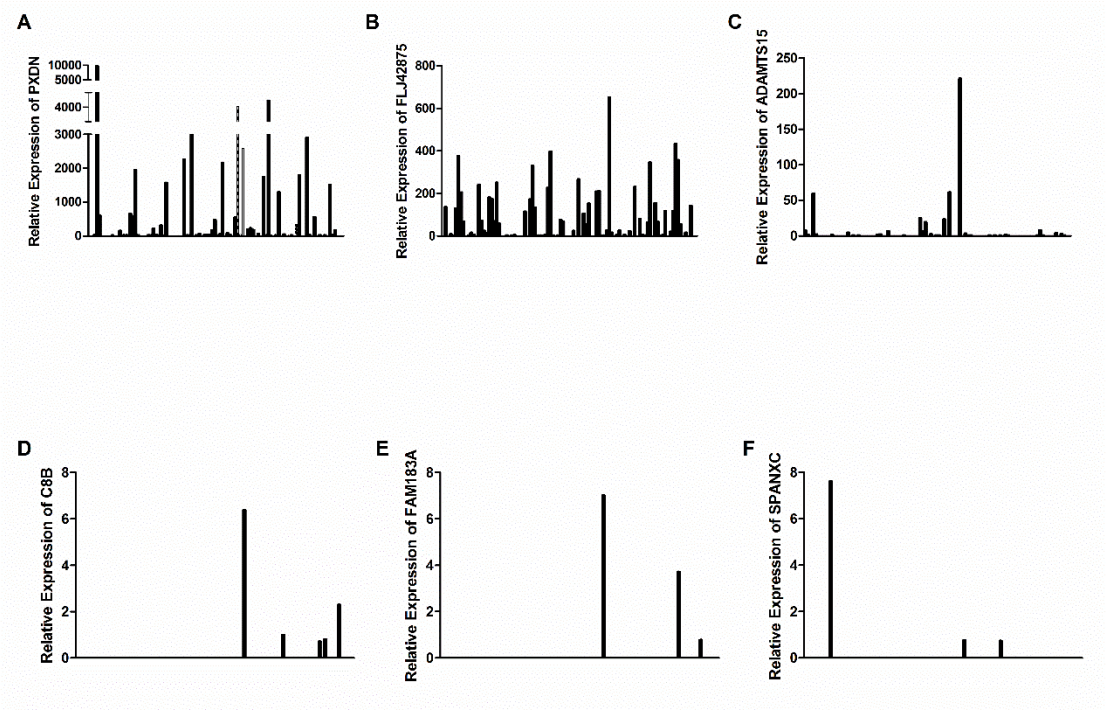


Figure 2. Expression of six differential genes in intermediate-risk AML patients.

Relationship between genes' expression and survival in primary intermediate-risk AML patients

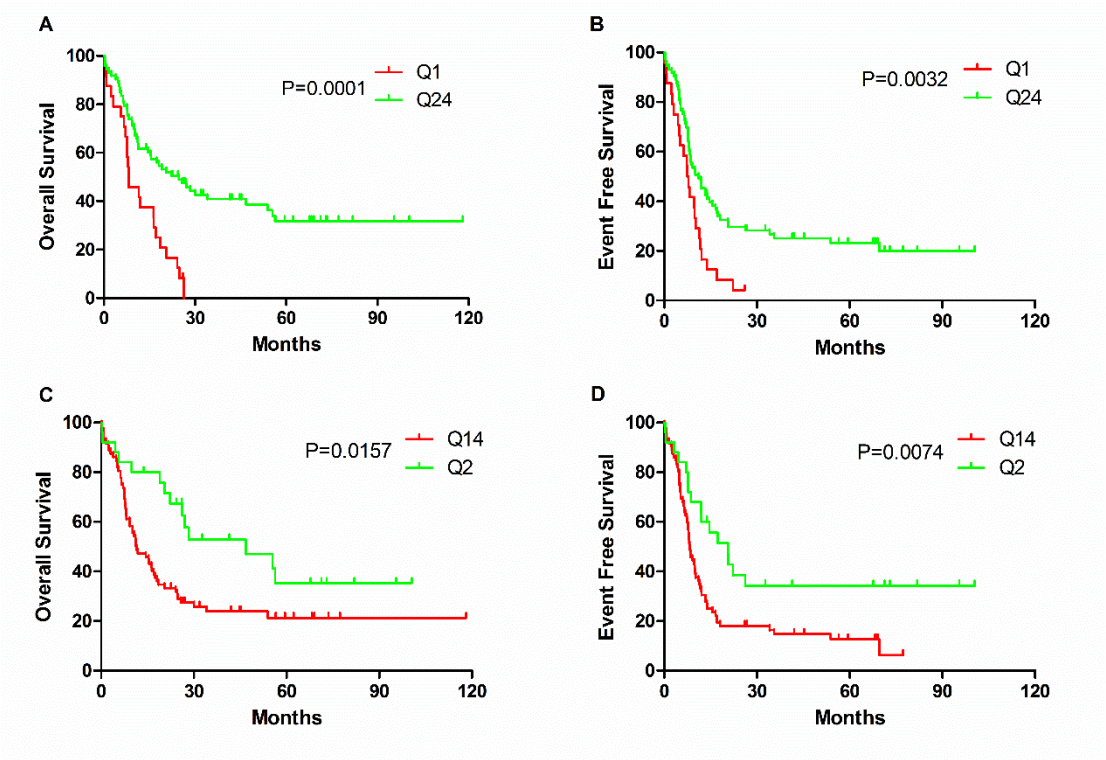


Figure 3. Relationship between gene expression and survival of intermediate-risk AML patients. (A) and (**B**) Kaplan-Meier curves showing the relationship of PXDN expression with Overall Survival and Event Free Survival of intermediate-risk AML patients; (C) and (**D**) Kaplan-Meier curves showing the relationship of FLJ42875 expression with Overall Survival and Event Free Survival of intermediate-risk AML patients.

We performed the survival analysis based on the expression of two widely expressed genes, FLJ42875 and PXDN. Firstly, we subdivided 98 AML patients into four quartiles (Q1: <25%, Q2: 25~50%, Q3: 50~75%, Q4: >75%) based on their expression value. The survival curves showed that PXDN Q1 expression was associated significantly shortened OS and EFS (P=0.0001, P=0.0032, respectively; Figure 3A and Figure 3B). However, OS and EFS of FLJ42875 Q2 expression group were

significantly prolonged than others (*P*=0.0157, *P*=0.0074, respectively; Figure 3C and Figure 3D), the difference was still significant by revalidation of Gehan-Breslow-Wilcoxon test. After adjusting for the impact of several known risk factors, we performed multivariable analyses to determine the prognostic significance of FLJ42875 and PXDN expression. In the multivariable model of OS, PXDN Q1 group had a shorter OS ($P < 0.001$, Table 2). The other factors associated with shorter OS was age more than or equal to 60 years old. While in the multivariable model of EFS, PXDN Q1 group had a shorter EFS (P =0.009, Table 2). FLJ42875 Q2 group had a prolonged OS and EFS (*P*=0.050, *P*=0.023; respectively).

Table 2: Multivariable analysis of PXDN expression with OS and EFS for intermediate-risk AML patients

Variable	OS, n=98		EFS, n=98	
	HR (95% CI)	P value	HR (95% CI)	P value
PXDN(Q1)	2.713(1.589-4.631)	**<0.001**	1.995(1.189-3.345)	**0.009**
Age, ≥60 year	2.655(1.629-4.327)	**<0.001**	2.237(1.419-3.524)	**0.001**
Female	1.155(0.717-1.859)	0.553	1.138(0.730-1.773)	0.569
FLT3 mutation	1.333(0.806-2.204)	0.262	1.602(0.998-2.570)	0.288

Table 3: Multivariable analysis of FLJ42875 expression with OS and EFS for intermediate-risk AML patients

Variable	OS, n=98		EFS, n=98	
	HR (95% CI)	P value	HR (95% CI)	P value
FLJ42875 (Q2)	0.548(0.300-1.000)	**0.050**	0.521(0.297-0.912)	**0.023**
Age, ≥60 year	2.341(1.441-3.806)	**0.001**	2.015(1.288-3.151)	**0.002**
Female	1.186(0.737-1.910)	0.483	1.191(0.764-1.858)	0.569
FLT3 mutation	1.353(0.820-2.230)	0.237	1.685(1.057-2.684)	0.028

Discussion

Nearly half of the AML patients have an intermediate-cytogenetic risk, and their clinical outcome is greatly heterogeneous and lacks effective prognostic indicators at the same time. Some of these patients respond well to chemotherapy, while others not. For this reason, we conducted this analysis to focus on establishing new indicators for complete classification of intermediate-risk.

First, we performed the differential analysis of gene expression profile of 98 intermediate-risk AML patients by using the RNA Sequencing data from TCGA based on different OS and EFS. We identified 6 distinct genes, SPANXC, ADAMTS15, C8B, FAM183A, FLJ42875 and PXDN. Then we detected the overall expression pattern in intermediate-risk AML patients. Only FLJ42875 and PXDN were relatively widely expressed in intermediate-risk AML. At last, we focused on the association of their expression pattern and survival of intermediate-risk AML patients. We found that OS and EFS of high FLJ42875 expression group were significantly shorter than that of low FLJ42875 expression patients (P=0.0409, P=0.0136, respectively), while high PXDN expression was associated significantly prolonged EFS and OS on the contrary (P=0.0185, P=0.0469, respectively). In the multivariable model of OS, PXDN Q1 group had a shorter OS and EFS ($P < 0.001$, P =0.009; respectively). While FLJ42875 Q2 group had a prolonged OS and EFS (P=0.050, P=0.023;

respectively).

PXDN, a member of the larger peroxidase gene family, is an adhesion molecule involved in ECM formation (Peterfi et al. 2014; Tauber et al. 2010). It is one of the target genes of miR-29a. Ectopic overexpression of miR-29a promoted self-renewal of myeloid progenitors and led to AML ultimately in mouse model (Oliveira et al. 2015). What's more, PXDN were diminished in AML (Desmond et al. 2007). In this article, high PXDN expression was associated with favorable prognosis in intermediate-risk AML, it was hypothesized that PXDN downregulation might play a role in the leukemogenesis of intermediate-risk AML.

Researches on FLJ42875 were limited. Only in a case report, upregulation of FLJ42875 by t(1;2)(p36;p21) translocation during evolution of CMML was reported with no further insight into mechanism of this gene(Storlazzi et al. 2008). No biological function especially in hematology malignancies were elucidated. In this article, Q2 of FLJ42875 was associated with favorable prognosis in intermediate-risk AML, it could be that FLJ42875 expression at particular level might predict the prognosis of intermediate-risk AML patients. More researches on this gene are required in the future.

In summary, our results detected two differential genes of which expression were related to the prognosis of primary intermediate AML patients, FLJ42875 and PXDN might be new markers of risk stratification

especially for intermediate-risk AML patients and new targets for clinical therapy. QPCR should be performed to revalidate the RNA Sequencing analysis and making the results more credible. What' more, further studies is needed to elucidate the mechanism behind it.

CONFLICT OF INTEREST

The authors declare no conflicts of interest.

ACKNOWLEDGMENTS

References

Ahn JS, Kim JY, Kim HJ, Kim YK, Lee SS, Jung SH, Yang DH, Lee JJ, Kim NY, Choi SH, Minden MD, Jung CW, Jang JH, Kim HJ, Moon JH, Sohn SK, Won JH, Kim SH, and Kim DD. 2016. Normal karyotype acute myeloid leukemia patients with CEBPA double mutation have a favorable prognosis but no survival benefit from allogeneic stem cell transplant. *Ann Hematol* 95:301-310. 10.1007/s00277-015-2540-7

Cancer Genome Atlas Research N. 2013. Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. *N Engl J Med* 368:2059-2074. 10.1056/NEJMoa1301689

Desmond JC, Raynaud S, Tung E, Hofmann WK, Haferlach T, and Koeffler HP. 2007. Discovery of epigenetically silenced genes in acute myeloid leukemias. *Leukemia* 21:1026-1034. 10.1038/sj.leu.2404611

Han C, Lin D, Ai XF, Wang F, Sun HY, Wang M, Mi YC, Wang JX, and Ru K. 2013. [CEBPA gene mutation analysis in acute myeloid leukemia]. *Zhonghua Xue Ye Xue Za Zhi* 34:566-571. 10.3760/cma.j.issn.0253-2727.2013.07.002

Ley TJ, Ding L, Walter MJ, McLellan MD, Lamprecht T, Larson DE, Kandoth C, Payton JE, Baty J, Welch J, Harris CC, Lichti CF, Townsend RR, Fulton RS, Dooling DJ, Koboldt DC, Schmidt H, Zhang Q, Osborne JR, Lin L, O'Laughlin M, McMichael JF, Delehaunty KD, McGrath SD, Fulton LA, Magrini VJ, Vickery TL, Hundal J, Cook LL, Conyers JJ, Swift GW, Reed JP, Alldredge PA, Wylie T, Walker J, Kalicki J, Watson MA, Heath S, Shannon WD, Varghese N, Nagarajan R, Westervelt P, Tomasson MH, Link DC, Graubert TA, DiPersio JF, Mardis ER, and Wilson RK. 2010. DNMT3A mutations in acute myeloid leukemia. *N Engl J Med* 363:2424-2433. 10.1056/NEJMoa1005143

Li HY, Deng DH, Huang Y, Ye FH, Huang LL, Xiao Q, Zhang B, Ye BB, Lai YR, Mo ZN, and Liu ZF. 2015. Favorable prognosis of biallelic CEBPA gene mutations in acute myeloid leukemia patients: a meta-analysis. *Eur J Haematol* 94:439-448. 10.1111/ejh.12450

Martelli MP, Sportoletti P, Tiacci E, Martelli MF, and Falini B. 2013. Mutational landscape of AML with

normal cytogenetics: biological and clinical implications. *Blood Rev* 27:13-22. 10.1016/j.blre.2012.11.001

Metzeler KH, Dufour A, Benthaus T, Hummel M, Sauerland MC, Heinecke A, Berdel WE, Buchner T, Wormann B, Mansmann U, Braess J, Spiekermann K, Hiddemann W, Buske C, and Bohlander SK. 2009. ERG expression is an independent prognostic factor and allows refined risk stratification in cytogenetically normal acute myeloid leukemia: a comprehensive analysis of ERG, MN1, and BAALC transcript levels using oligonucleotide microarrays. *J Clin Oncol* 27:5031-5038. 10.1200/JCO.2008.20.5328

Metzeler KH, Heilmeier B, Edmaier KE, Rawat VP, Dufour A, Dohner K, Feuring-Buske M, Braess J, Spiekermann K, Buchner T, Sauerland MC, Dohner H, Hiddemann W, Bohlander SK, Schlenk RF, Bullinger L, and Buske C. 2012. High expression of lymphoid enhancer-binding factor-1 (LEF1) is a novel favorable prognostic factor in cytogenetically normal acute myeloid leukemia. *Blood* 120:2118-2126. 10.1182/blood-2012-02-411827

Mikulasova Z, Ilencikova D, Slamka T, and Durovcikova D. 2010. [Acute myeloblastic leukaemia with alternations of MLL proto-oncogene protein (11q23/MLL+ AML)]. *Klin Onkol* 23:401-407.

Mrozek K, Heerema NA, and Bloomfield CD. 2004. Cytogenetics in acute leukemia. *Blood Rev* 18:115-136. 10.1016/S0268-960X(03)00040-7

Oliveira LH, Schiavinato JL, Fraguas MS, Lucena-Araujo AR, Haddad R, Araujo AG, Dalmazzo LF, Rego EM, Covas DT, Zago MA, and Panepucci RA. 2015. Potential roles of microRNA-29a in the molecular pathophysiology of T-cell acute lymphoblastic leukemia. *Cancer Sci* 106:1264-1277. 10.1111/cas.12766

Peterfi Z, Toth ZE, Kovacs HA, Lazar E, Sum A, Donko A, Sirokmany G, Shah AM, and Geiszt M. 2014. Peroxidasin-like protein: a novel peroxidase homologue in the human heart. *Cardiovasc Res* 101:393-399. 10.1093/cvr/cvt256

Ribeiro AF, Pratcorona M, Erpelinck-Verschueren C, Rockova V, Sanders M, Abbas S, Figueroa ME, Zeilemaker A, Melnick A, Lowenberg B, Valk PJ, and Delwel R. 2012. Mutant DNMT3A: a marker of poor prognosis in acute myeloid leukemia. *Blood* 119:5824-5831. 10.1182/blood-2011-07-367961

Rowley JD. 2008. Chromosomal translocations: revisited yet again. *Blood* 112:2183-2189. 10.1182/blood-2008-04-097931

Schulze I, Rohde C, Scheller-Wendorff M, Krause A, Herbst F, Riemke P, Hebestreit K, Tschanter P, Lin Q, Linhart H, Godley LA, Glimm H, Dugas M, Wagner W, Berdel WE, Rosenbauer F, and Muller-Tidow C. 2016. Increased DNA methylation of Dnmt3b-targets impairs leukemogenesis. *Blood*. 10.1182/blood-2015-07-655928

Stirewalt DL, and Radich JP. 2003. The role of FLT3 in haematopoietic malignancies. *Nat Rev Cancer* 3:650-665. 10.1038/nrc1169

Storlazzi CT, Albano F, Guastadisegni MC, Impera L, Muhlematter D, Meyer-Monard S, Wuillemin W, Rocchi M, and Jotterand M. 2008. Upregulation of MEL1 and FLJ42875 genes by position effect resulting from a t(1;2)(p36;p21) occurring during evolution of chronic myelomonocytic leukemia. *Blood Cells Mol Dis* 40:452-455. 10.1016/j.bcmd.2007.11.004

Tauber S, Jais A, Jeitler M, Haider S, Husa J, Lindroos J, Knofler M, Mayerhofer M, Pehamberger H, Wagner O, and Bilban M. 2010. Transcriptome analysis of human cancer reveals a functional role of heme oxygenase-1 in tumor cell adhesion. *Mol Cancer* 9:200. 10.1186/1476-4598-9-200

Yi-Ning Y, Xiao-rui W, Chu-xian Z, Chun W, and You-wen Q. 2015. Prognostic significance of diagnosed WT1 level in acute myeloid leukemia: a meta-analysis. *Ann Hematol* 94:929-938. 10.1007/s00277-014-2295-6