

Mutational spectrum of acute myeloid leukemia patients with double *CEBPA* mutations based on next-generation sequencing and its prognostic significance

Long Su^{1,*}, YeHui Tan^{1,*}, Hai Lin¹, XiaoLiang Liu¹, Li Yu², YanPing Yang¹, ShanShan Liu¹, Ou Bai¹, Yan Yang¹, FengYan Jin¹, JingNan Sun¹, ChunShui Liu¹, QiuJu Liu¹, SuJun Gao¹ and Wei Li¹

¹Department of Hematology, The First Hospital, Jilin University, Changchun, China

²Department of Hematology, Chinese PLA General Hospital, Peking, China

*These authors have contributed equally to this work

Correspondence to: SuJun Gao, email: gaosujunjdyy@163.com

Keywords: acute myeloid leukemia; *CEBPA* mutations; next generation sequencing; prognoses; Chinese population

Received: March 16, 2017 **Accepted:** December 27, 2017 **Epub:** January 03, 2018 **Published:** May 18, 2018

Copyright: Su et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License 3.0 (CC BY 3.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

ABSTRACT

The aim of this study was to profile the spectrum of genetic mutations in acute myeloid leukemia (AML) patients co-occurring with *CEBPA* double mutation (*CEBPA*^{dm}). Between January 1, 2012, and June 30, 2017, 553 consecutive patients with *de novo* AML were screened for *CEBPA* mutations. Out of these, 81 patients classified as *CEBPA*^{dm} were analyzed further by a sensitive next-generation sequencing assay for mutations in 112 candidate genes. Within the *CEBPA* gene itself, we found 164 mutations. The most common mutated sites were c.936_937insGAG (n = 11/164, 6.71%) and c.939_940insAAG (n = 11/164, 6.71%), followed by c.68dupC (n = 10/164, 6.10%). The most common co-occurring mutations were found in the *CSF3R* (n = 16/81, 19.75%), *WT1* (n = 15/81, 18.52%), and *GATA2* (n = 13/81, 16.05%) genes. Patients with *CSF3R* mutations had an inferior four-year relapse-free survival (RFS) than those with the wild-type gene (15.3% versus 46.8%, respectively; *P* = 0.021). Patients with *WT1* mutations had an inferior five-year RFS compared with those without such mutations (0% versus 26.6%, respectively, *P* = 0.003). However, *GATA2*, *CSF3R*, *WT1* mutations had no significant influence on the overall survival. There were some differences in the location of mutational hotspots within the *CEBPA* gene, as well as hotspots of other co-occurring genetic mutations, between AML patients from Chinese and Caucasian populations. Some co-occurring mutations may be potential candidates for refining the prognoses of AML patients with *CEBPA*^{dm} in the Chinese population.

INTRODUCTION

Mutations in the *CCAAT/enhancer binding protein α* (*CEBPA*) gene occur in 7%–15% of all acute myeloid leukemia (AML) cases. The subgroup of biallelic *CEBPA* mutations in AML patients has now been acknowledged in ‘The 2016 revision to the World Health Organization classification of myeloid neoplasms and acute leukemia’

as a definite entity, given its distinct biological and clinical features, as well as its prognostic significance [1]. *CEBPA* belongs to the basic-leucine zipper (b-ZIP) family of transcription factors whose C-terminal regions contain two highly conserved motifs: a DNA-binding motif rich in basic amino acids and a leucine zipper dimerization motif. They also contain two less conserved N-terminal transactivation domains (TADs) [2]. *CEBPA* mutations

Table 1: The characteristics of 81 AML patients with *CEBPA*^{dm}

Characteristics	Number of patients	Percentage
Age (years), median (range)	44 (9 ~ 67)	
Gender		
Male	45	55.56%
Female	36	44.44%
FAB classification		
M1	2	2.47%
M2	47	58.02%
M4	24	29.63%
M5	5	6.17%
M6	3	3.70%
Cytogenetics		
Normal karyotype	60	92.31%
Abnormal karyotypes	5	7.69%
Peripheral blood		
White blood cells ($\times 10^9/L$)	17.10 [8.84, 62.64]	
Hemoglobin (g/L)	97.56 \pm 27.85	
Platelets ($\times 10^9/L$)	23.00 [12.00, 41.00]	
Marrow blasts (%)	61.00 \pm 17.38	

can occur across the whole gene, but cluster in two main hotspots: N-terminal frame-shift insertions/deletions—these cause translation of a 30 kDa protein from an internal ATG start site that lacks transactivation domain 1 and has a dominant negative effect over the full-length p42 protein; C-terminal mutations—these are generally in-frame insertions/deletions, in the DNA-binding or leucine zipper domains, that disrupt binding to DNA or dimerization [3].

AML patients with double *CEBPA* mutations (*CEBPA*^{dm}) show a favorable outcome, which was also observed in our previous study [4]. Both others' and our studies suggest that the frequency of *CEBPA* mutations (17.1%–21.6%) may be higher in Chinese patients with AML than what has been reported for populations of Western countries [4-5]. We also noticed some genetic differences between patients with AML from China and Western countries [4, 6-7]. Although the genetic profiling of AML patients with *CEBPA*^{dm} has been reported in previous studies [8-9], there is no data available for Chinese patients. Furthermore, the prognostic significance of co-occurring mutations remains unclear in patients with

CEBPA^{dm}. In this study, we screened 553 patients with *de novo* AML and profiled genetic mutations in those with *CEBPA*^{dm} (n = 81) by a sensitive next-generation sequencing assay. The prognostic significance of the top three co-occurring genetic mutations was also evaluated.

RESULTS

Patients' characteristics

Of the 553 consecutive patients with *de novo* AML, *CEBPA* mutations were detected in 105 patients (18.99%), with 81 cases (14.65%) harboring double mutations and 24 cases (4.34%) harboring single mutations. Characteristics of the patients are summarized in Table 1. Most (60.49%) of the patients were morphological M1 and M2 subtypes according to the French–American–British (FAB) classification system. Of the 65 patients who underwent successful cytogenetic analysis, 60 cases (92.31%) presented with normal karyotypes.

CEBPA mutation screening based on next-generation sequencing

Among the 81 *CEBPA*^{dm} patients, 164 genetic mutations, classifiable into 91 different kinds, were detected in the *CEBPA* gene. The median mutation load was 45.3% (range: 4.5%–58.2%). The most common mutated sites were c.936_937insGAG (n = 11/164, 6.71%) and c.939_940insAAG (n = 11/164, 6.71%), followed by c.68dupC (n = 10/164, 6.10%), c.247delC (n = 7/164, 4.27%), c.275dupA (n = 7/164, 4.27%) (Figure 1). The majority of *CEBPA*^{dm} comprised frame-shift insertions or deletions (n = 83/164, 50.61%). The next most common were in-frame insertions or deletions (n = 72/164, 43.90%). The least common were missense mutations (n = 5/164, 3.05%), and stop-gain mutations (n = 4/164, 2.44%). A majority of *CEBPA*^{dm} patients (n = 64/81, 79.01%) showed a combination of an N-terminal frame-shift and a C-terminal in-frame mutation.

When the DNA sequences were translated into the corresponding amino acid sequences, the most common mutation site was p.Pro23fs (n = 17/164, 10.37%; 13 frame-shift insertions and four frame-shift deletions), followed by p.Gln312_Lys313insGln (n = 12/164, 7.32%), and p.Lys313_Val314insLys (n = 11/164, 6.71%; Figure 2).

Correlation of the *CEBPA*^{dm} status to other molecular mutations

Twenty-seven types of other molecular mutations were detected in patients with *CEBPA*^{dm}. Seventeen patients (20.99%) had no additional molecular mutation, 23 (28.40%) had one, 20 (24.69%) had two, 12 (14.81%) had three, six had (7.41%) four, and three (3.70%) patients had five additional mutations (Supplementary Figure 1).

CSF3R (n = 16), *WT1* (n = 15), and *GATA2* (n = 13), were the most common co-occurring mutations, with frequencies of 19.75%, 18.52%, and 16.05%, respectively (Figure 3).

Next, we analyzed the clinical characteristics of patients with mutations in other genes which co-occurred with frequencies exceeding 10%. These included mutations in the *CSF3R*, *WT1*, *GATA2*, *NRAS*, and *TET2*

genes. *CSF3R* mutation was associated with a lower platelet ($18.50 [11.25, 32.75] \times 10^9/L$ versus $23.00 [13.50, 47.00] \times 10^9/L$; $u = 2.873, P = 0.005$) and higher leukocyte ($53.57 [28.76, 73.39] \times 10^9/L$ versus $14.00 [7.33, 36.44] \times 10^9/L$; $u = 3.001, P = 0.030$) counts as compared to the wild-type. *WT1* mutation was associated with a higher white blood cell count as compared to the wild-type ($36.22 [13.22, 121.31] \times 10^9/L$ versus $16.36 [7.00, 50.00] \times 10^9/L$;

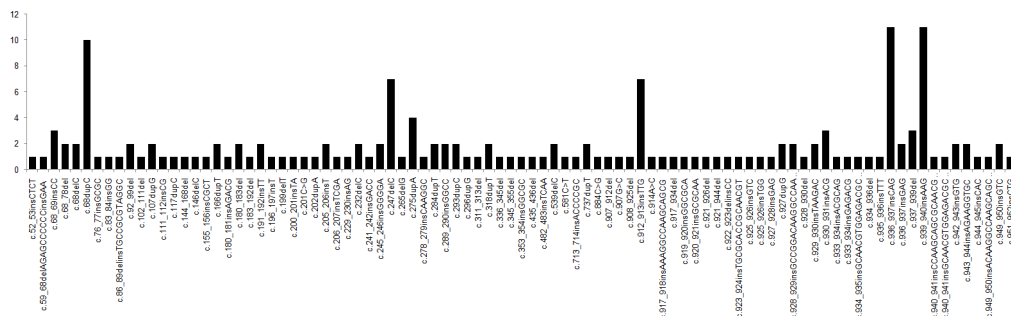


Figure 1: The mutation locations of *CEBPA* gene in AML patients with *CEBPA*^{dm}.

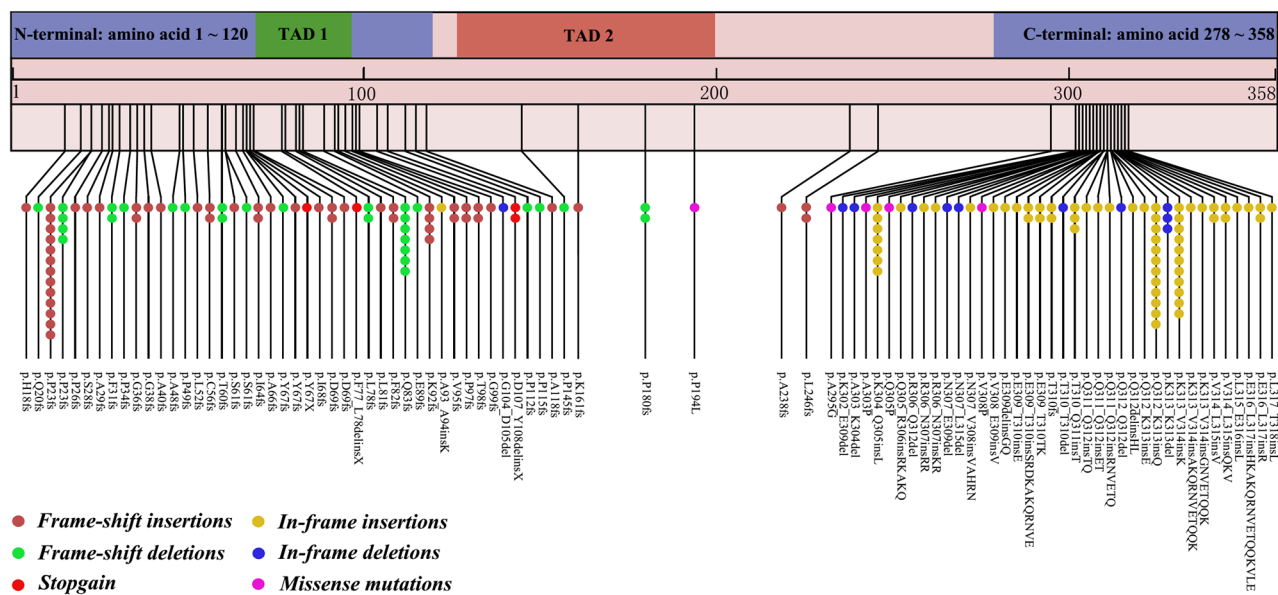


Figure 2: Amino acid alterations in *CEBPA* protein in AML patients with *CEBPA*^{dm}.

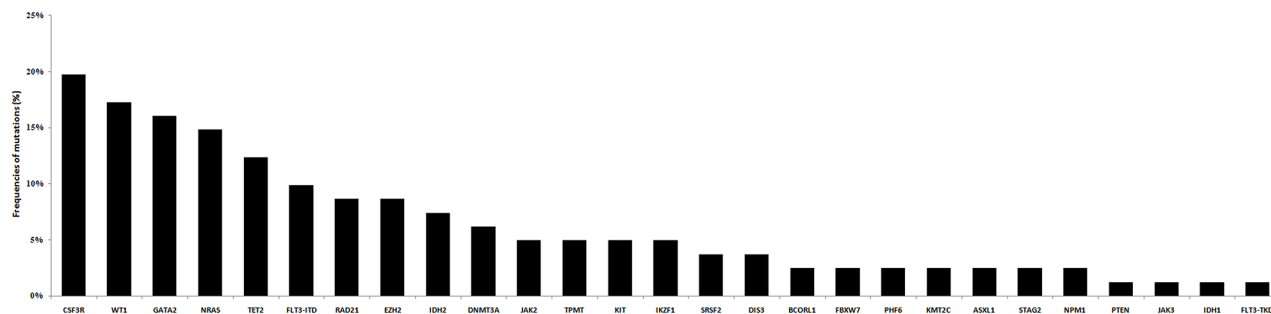


Figure 3: The distribution of co-occurring mutations in AML patients with *CEBPA*^{dm}.

Table 2: The characteristics of AML patients with different molecular mutations

	<i>CSF3R</i>		<i>WT1</i>		<i>GATA2</i>	
	Mutations (n = 16)	Wide-type (n = 65)	Mutations (n = 14)	Wide-type (n = 67)	Mutations (n = 13)	Wide-type (n = 68)
Age (years)	38.88 ± 13.54	40.35 ± 14.52	28.14 ± 12.13	42.56 ± 13.46*	37.38 ± 14.75	40.57 ± 14.22
Gender						
Male	10 (62.50%)	35 (53.85%)	10 (71.43%)	35 (52.24%)	8 (61.54%)	37 (54.41%)
Female	6 (37.50%)	30 (46.15%)	4 (28.57%)	32 (47.76%)	5 (38.46%)	31 (45.59%)
FAB classification						
M1	0 (0.00%)	2 (3.08%)	1 (7.14%)	1 (1.49%)	0 (0.00%)	2 (2.94%)
M2	6 (37.50%)	33 (50.77%)	6 (42.86%)	33 (49.25%)	10 (76.92%)	29 (42.65%)
M4	9 (56.25%)	23 (35.38%)	6 (42.86%)	26 (38.81%)	2 (15.38%)	30 (44.12%)
M5	1 (6.25%)	4 (6.15%)	1 (7.14%)	4 (5.97%)	1 (7.69%)	4 (5.88%)
M6	0 (0.00%)	3 (4.62%)	0 (0.00%)	3 (4.48%)	0 (0.00%)	3 (4.41%)
Cytogenetics						
Normal karyotype	12 (92.31%)	48 (92.31%)	9 (100.00%)	51 (91.07%)	11 (100.00%)	49 (90.74%)
Abnormal karyotypes	1 (7.69%)	4 (7.69%)	0 (0.00%)	5 (8.93%)	0 (0.00%)	5 (9.26%)
Peripheral blood						
White blood cells (×10 ⁹ /L)	53.57 [28.76, 73.39]	14.00 [7.33, 36.44]*	36.22 [13.22, 121.31]	16.36 [7.00, 50.00] *	16.45 [11.44, 67.81]	18.06 [8.48, 62.64]
Hemoglobin (g/L)	96.00 ± 25.10	97.94 ± 28.65	84.71 ± 27.94	100.24 ± 27.27	86.15 ± 28.57	99.74 ± 27.38
Platelets (×10 ⁹ /L)	18.50 [11.25, 32.75]	23.00 [13.50, 47.00]*	23.00 [10.75, 46.50]	23.00 [12.00, 40.00]	25.00 [9.50, 43.50]	23.00 [12.00, 39.75]
Marrow blasts (%)	53.63 ± 18.55	63.00 ± 16.66	62.07 ± 20.14	60.75 ± 16.86	59.23 ± 12.45	61.37 ± 18.31

	<i>NRAS</i>		<i>TET2</i>	
	Mutations (n = 12)	Wide-type (n = 69)	Mutations (n = 10)	Wide-type (n = 71)
Age (years)	37.38 ± 14.75	40.57 ± 14.22	44.10 ± 11.20	39.49 ± 14.61
Gender				
Male	6 (50.00%)	39 (56.52%)	3 (30.00%)	42 (59.15%)
Female	6 (50.00%)	30 (43.48%)	7 (70.00%)	29 (40.85%)
FAB classification				
M1	1 (8.33%)	1 (1.45%)	1 (10.00%)	1 (1.41%)
M2	6 (50.00%)	33 (47.83%)	6 (60.00%)	33 (46.48%)
M4	2 (16.67%)	30 (43.38%)	3 (30.00%)	29 (40.85%)

(Continued)

	<i>NRAS</i>		<i>TET2</i>	
	Mutations (n = 12)	Wide-type (n = 69)	Mutations (n = 10)	Wide-type (n = 71)
M5	3 (25.00%)	2 (2.90%)	0 (0.00%)	5 (7.04%)
M6	0 (0.00%)	3 (4.35%)	0 (0.00%)	3 (4.23%)
Cytogenetics				
Normal karyotype	6 (75.00%)	54 (94.74%)	9 (90.00%)	51 (92.73%)
Abnormal karyotypes	2 (25.00%)	3 (5.26%)	1 (10.00%)	4 (7.27%)
Peripheral blood				
White blood cells ($\times 10^9/L$)	25.70 [9.72, 93.26]	17.10 [8.43, 62.64]	34.72 [21.35, 72.37]	16.36 [7.66, 50.00]
Hemoglobin (g/L)	86.15 \pm 28.57	99.74 \pm 27.38	87.80 \pm 27.35	98.93 \pm 27.83
Platelets ($\times 10^9/L$)	22.00 [14.00, 23.00]	23.00 [12.00, 42.00]	23.00 [12.50, 24.50]	23.00 [12.00, 45.00]
Marrow blasts (%)	59.23 \pm 12.45	61.37 \pm 18.31	69.49 \pm 12.60	59.69 \pm 17.72

*Compared with patients with mutations, $P < 0.05$.

$u = 2.024$, $P = 0.043$). The average age of patients with a *WT1* mutation was less than the average age of those without one (28.14 ± 12.13 versus 42.56 ± 13.46 ; $t = 3.702$, $P < 0.001$; Table 2).

Co-occurring mutations were categorized as falling into various pathways and gene families: tyrosine kinase pathway, transcription factor gene, tumor suppressor gene, DNA methylation gene, chromatin-modifier gene, cohesion molecule gene, spliceosome complex gene, and others. The most frequent mutation involved genes affecting the tyrosine kinase pathway (33.33%), followed by DNA methylation (15.94%), and tumor suppressor (13.77%) gene families (Supplementary Figure 2).

Treatment response and long-term outcome

For 67 patients received induction therapy, 50 patients achieved complete remission (CR), 14 achieved partial remission (PR), and the remaining three cases were classified as non-remission (NR) after one course of chemotherapy. *CSF3R*, *WT1*, and *GATA2* mutations had no influence on the CR rate ($P = 0.320$, $P = 0.130$, and $P = 0.158$ respectively). Finally, 66 cases who achieved CR entered long-term follow-up. The follow-up time ranged from two to 66 months (median: 8 months). In total, 18 patients relapsed, and 13 patients died. Five-year relapse-free survival (RFS) (Figure 4A) and overall survival (OS) (Figure 4B) rates were 20.7% and 47.0%, respectively.

We also evaluated the prognostic significance of *CSF3R*, *WT1*, *GATA2* mutations in patients with *CEBPA*^{dm}. The four-year RFS of patients with *CSF3R* mutations was 15.3%, which was lower than those with wild-type *CSF3R* (46.8%) ($P = 0.021$). The median RFS of patients with mutated and wild-type *CSF3R* were 10 and 43 months,

respectively. Patients with *WT1* mutations had an inferior five-year RFS compared with those without the mutations (0% versus 26.6%, $P = 0.003$). The median RFS of patients with mutated and wild-type *WT1* were 10 and 64 months, respectively. The five-year RFS rates were 38.1% and 46.4% in patients with the mutated and wild-type *GATA2*, respectively ($P = 0.641$). *GATA2*, *CSF3R*, *WT1* mutations had no significant influence on OS in this study (Figure 5).

DISCUSSION

AML is a heterogeneous disease, and DNA sequencing can offer clues to its etiology and predict prognoses of patients with AML. With the development of new sequencing technology, more and more genetic mutations are being identified in AML patients [10]. In our previous studies, we observed some differences in genetic alterations between AML patients from China and Western countries [4, 6-7]. The frequencies of *NPM1* (15.4%) and *FLT3*-ITD mutations (14%) are lower in AML patients from China [7], whereas the frequency of *CEBPA* mutations is higher (17.1%) [4]. These results accord with the literature published by others from China [5, 11-12]. It is known that AML with *CEBPA*^{dm} indicates a favorable outcome, which was also confirmed in our cohort of patients [4]. However, it is unknown whether there are genetic differences among the geographic or ethnic subgroups of AML patients with *CEBPA*^{dm}.

The present subset of AML patients was derived from 553 consecutive patients with *de novo* diagnoses, which avoided selection bias. The majority (60.49%) of patients presented with M1 and M2 subtypes, according to the FAB classification system. A normal karyotype

was present in 92.31% of the patients, while aberrant karyotypes included del(9q) and +8 trisomy. Only two patients with *NPM1* mutation were detected in this study. All these features are consistent with previous reports [3, 8].

We found that the most common *CEBPA* mutation types were frame-shift insertions or deletions, followed by in-frame insertions or deletions, which is in accord with previous studies [3, 8]. A combination of an N-terminal frame-shift and a C-terminal in-frame mutation was present in the majority of patients in this study, which was also reported previously [8-9]. Fasen *et al.* reported that the most frequent mutation site was p.Lys313del, followed by p.His24A1afs, and p.Gln312del [8]. However, we

observed a different result. The most common mutation site in the present study was p.Pro23fs, followed by p.Gln312_Lys313insGln, and p.Lys313_Val314insLys. We profiled for genetic mutations co-occurring in *CEBPA*^{dm} AML patients. Interestingly, we observed that the percentage of patients with three or more co-occurring molecular mutations was higher in this study than in previous studies (25.93% versus 2.88%, respectively, $\chi^2 = 21.412$, $P < 0.001$; [8]). We hypothesize that these differences between AML patients from Chinese and Caucasian populations may be due to their differing ethnic backgrounds.

Among AML patients with *CEBPA*^{dm}, Grossmann *et al.* from Germany reported that *TET2* was found to be

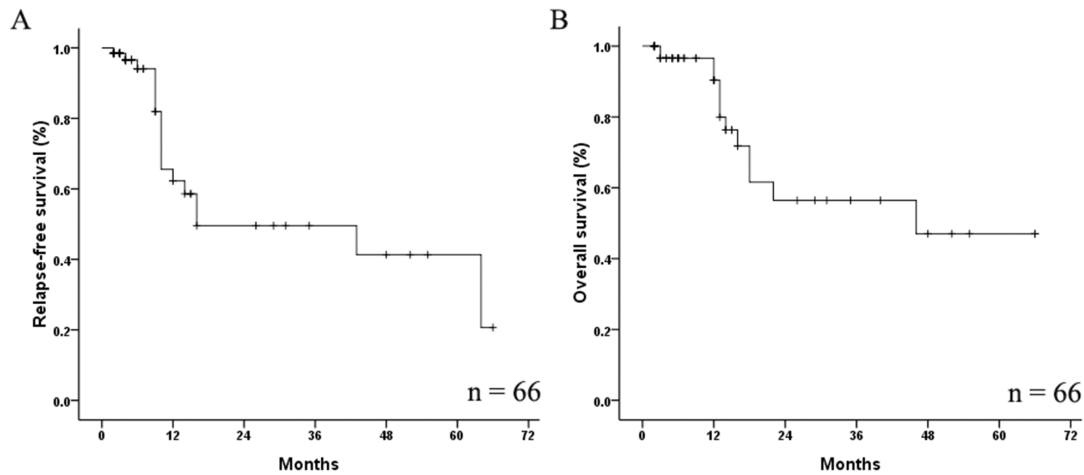


Figure 4: Relapse-free survival (A) and overall survival (B) in AML patients with *CEBPA*^{dm}.

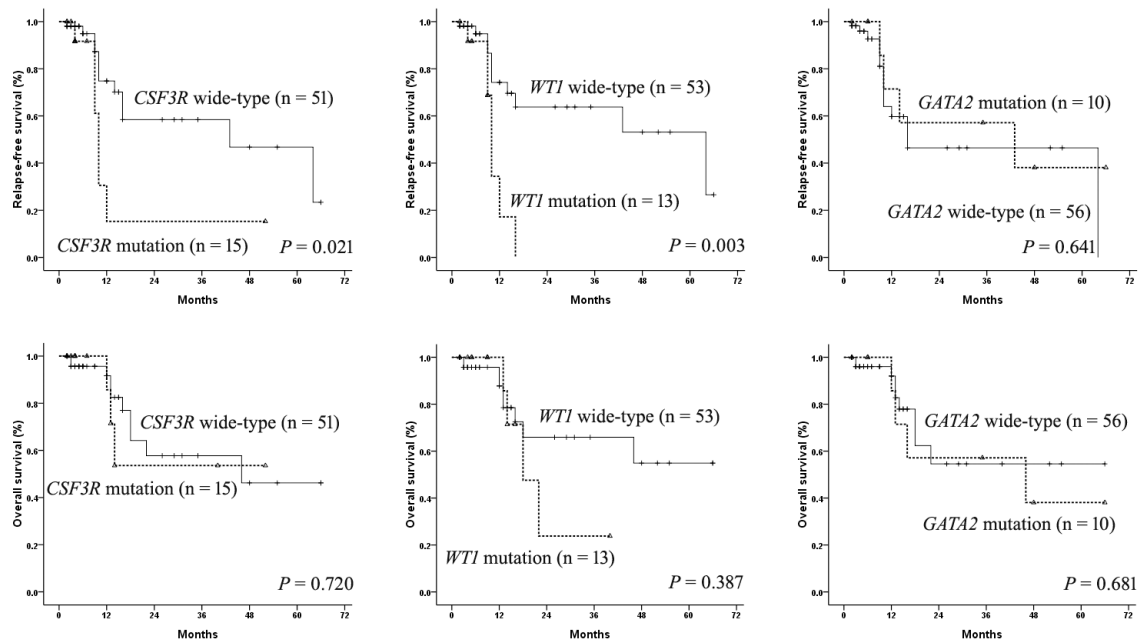


Figure 5: The influence of *CSF3R*, *WT1*, and *GATA2* mutations on outcomes in AML patients with *CEBPA*^{dm}.

Table 3: Mutations of 112 genes analyzed in this study

Number	Gene	Number	Gene	Number	Gene	Number	Gene
1	<i>CEBPA</i>	29	<i>CALR</i>	57	<i>CCND1</i>	85	<i>KMT3A</i>
2	<i>NPM1</i>	30	<i>CSF3R</i>	58	<i>CD79B</i>	86	<i>MAP2K4</i>
3	<i>FLT3-ITD</i>	31	<i>SH2B3</i>	59	<i>CDA</i>	87	<i>MAP3K7</i>
4	<i>FLT3-TKD</i>	32	<i>IKZF1</i>	60	<i>CREBBP</i>	88	<i>MDM2</i>
5	<i>KIT</i>	33	<i>ABL</i>	61	<i>CRLF2</i>	89	<i>MEF2B</i>
6	<i>DNMT3A</i>	34	<i>NOTCH1</i>	62	<i>CSF1R</i>	90	<i>MLH1</i>
7	<i>IDH1</i>	35	<i>FBXW7</i>	63	<i>CTLA4</i>	91	<i>MTHFR</i>
8	<i>IDH2</i>	36	<i>TPMT</i>	64	<i>CUX1</i>	92	<i>NF2</i>
9	<i>MLL</i>	37	<i>CDKN2A</i>	65	<i>CYP2C19</i>	93	<i>NOTCH2</i>
10	<i>TET2</i>	38	<i>ATM</i>	66	<i>CYP3A4</i>	94	<i>NQO1</i>
11	<i>WT1</i>	39	<i>HRAS</i>	67	<i>DIS3</i>	95	<i>NT5C2</i>
12	<i>RUNX1</i>	40	<i>RB1</i>	68	<i>DNAH9</i>	96	<i>NTRK1</i>
13	<i>KRAS</i>	41	<i>MYD88</i>	69	<i>E2A</i>	97	<i>NTRK2</i>
14	<i>NRAS</i>	42	<i>ABCB1</i>	70	<i>EGFR</i>	98	<i>PDGFRA</i>
15	<i>ASXL1</i>	43	<i>ABCC3</i>	71	<i>ERCC1</i>	99	<i>PIGA</i>
16	<i>PHF6</i>	44	<i>AKT2</i>	72	<i>ERG</i>	100	<i>PIK3CA</i>
17	<i>TP53</i>	45	<i>AKT3</i>	73	<i>FAM46C</i>	101	<i>PTEN</i>
18	<i>SF3B1</i>	46	<i>AMER1</i>	74	<i>GATA1</i>	102	<i>RAD21</i>
19	<i>SRSF2</i>	47	<i>APC</i>	75	<i>GATA2</i>	103	<i>SMAD4</i>
20	<i>U2AF1</i>	48	<i>ATRX</i>	76	<i>GNAS</i>	104	<i>SMC1A</i>
21	<i>ZRSR2</i>	49	<i>BCL2</i>	77	<i>GSTM1</i>	105	<i>SMC3</i>
22	<i>EZH2</i>	50	<i>BCOR</i>	78	<i>GSTP1</i>	106	<i>STAG2</i>
23	<i>CBL</i>	51	<i>BCORL1</i>	79	<i>ID3</i>	107	<i>STAT5A</i>
24	<i>JAK2</i>	52	<i>BRAF</i>	80	<i>IL17R</i>	108	<i>STAT5B</i>
25	<i>SETBP1</i>	53	<i>CACNA1E</i>	81	<i>JAK1</i>	109	<i>SYK</i>
26	<i>ETV6</i>	54	<i>CARD11</i>	82	<i>JAK3</i>	110	<i>TERC</i>
27	<i>PTPN11</i>	55	<i>CBLB</i>	83	<i>KDM6A</i>	111	<i>TRAF3</i>
28	<i>MPL</i>	56	<i>CBLC</i>	84	<i>KMT2C</i>	112	<i>XRCC1</i>

the most frequently mutated gene (34.0%), followed by *GATA2* (21.0%), and *WT1* (13.7%) genes [9]. Another German group reported that the top-three mutated genes were *TET2* (15.7%), *ASXL1* (13.7%), and *WT1* (13.6%) [8]. The top-three mutated genes identified in this study were *CSF3R* (19.75%), *WT1* (18.52%), and *GATA2* (16.05%).

The frequency of *GATA2* mutations in *CEBPA*^{dm} patients in this study (16.05%) was lower than that reported in previous studies [9, 13]. There are still some controversies regarding the prognostic significance of *GATA2* mutations in patients with *CEBPA*^{dm} [9, 13-15].

Grossmann *et al.* reported that *GATA2*-mutated patients show a longer OS than *GATA2* wild-type cases (n = 95; [9]). Hou and colleagues observed that among patients with *CEBPA*^{dm}, those with *GATA2* mutations had a trend of better OS and RFS than those without (n = 62; [13]). In univariate analysis, *GATA2* mutations were associated with better event-free survival (EFS) and OS ($P = 0.03$ and $P = 0.041$, respectively; n = 98; [14]). However, no significant difference in CR rate, RFS, and OS was also observed in *CEBPA*^{dm} patients with and without *GATA2* mutations (n = 113; [15]). In the present study, we found that *GATA2* mutations had no influence on CR, RFS and OS. Due to

the relatively small number of patients in these studies, further research is still needed to evaluate the prognostic significance of *GATA2* mutation in patients with *CEBPA*^{dm}. Furthermore, we argue that ethnicity should also be taken into account when conducting analyses.

Recently, Lavallée *et al.* from Canada reported that *CSF3R* mutations were the most frequent mutations (29%) in AML patients with *CEBPA*^{dm} [16]. Maxson and colleagues confirmed those findings in a cohort of pediatric patients with AML. They found a significant enrichment of *CSF3R* mutations (46%) among the *CEBPA*-mutated AML patients in America [17]. A high frequency of *CSF3R* mutations was also observed in our cohort of AML patients. In accordance with a previous study (n = 11/19, 57.89%) [17], we also found the majority of *CSF3R* mutations (n = 11/16, 68.75%) were p.T618I. Collectively, these findings suggest that *CEBPA*^{dm} AML patients may benefit from treatment with Janus kinase inhibitors.

Although AML with *CEBPA*^{dm} indicates a favorable outcome, recent data show that more than 50% of the patients finally relapsed when consolidated with chemotherapy alone [18]. Hence, a new marker is needed to stratify patients with *CEBPA*^{dm}. Patients with *CSF3R* and *WT1* mutations showed inferior RFS compared with those with the wild-type genes. As a result, *WT1* and *CSF3R* mutations may be adopted as potential markers to stratify patients with *CEBPA*^{dm} in the Chinese population.

Consistent with a previous study [19], we also found that the most frequent mutations in patients with *CEBPA*^{dm} occurred in the tyrosine kinase signaling pathway. Exploration or evaluation of drugs targeting these pathways, and translational research integrating these molecular findings, may improve the treatment of patients with *CEBPA*^{dm}.

In summary, we found that there were some differences in hotspots of *CEBPA* mutations, and in hotspots of co-occurring genetic mutations, between AML patients from Chinese and Caucasian populations. Some of the co-occurring mutations may even be potential candidates, for treating patients with *CEBPA*^{dm}, specific to the Chinese population. The continuation of such studies may uncover more mutational differences based on ethnicity, which may similarly reveal information pertinent to research into the etiology of AML and treatment of AML patients with *CEBPA*^{dm}.

MATERIALS AND METHODS

Patients and treatment

From January 1, 2012, to June 30, 2017, 553 consecutive patients with *de novo* AML were screened for *CEBPA* mutations from our center and Chinese People's Liberation Army (PLA) General Hospital. They were categorized into FAB subtypes (M0–M7) based on morphological diagnoses [20] (Supplementary S3).

Patients in this study were treated with the standard '3+7' regimen (darubicin/idarubicin + cytarabine) or CAG (acliarubicin + cytarabine + G-CSF) regimen (for some elderly patients) for induction therapy. The response was assessed by bone marrow aspiration performed on days 14 and 28. The first consolidation therapy was the same as that generally used to achieve CR. Three to four courses of scheduled, high-dose cytarabine, at 2–3 g/m², were administered for consolidation therapy. Five patients with *CEBPA*^{dm} received allo-HSCT. All of the participating patients gave informed consent prior to enrolment in the study. This study was approved by the ethics committee of Jilin University and Chinese PLA General Hospital, and conducted in accordance with the Declaration of Helsinki.

Cytogenetic analysis

Standard culturing and chromosome-banding techniques were used to analyze the karyotypes. Their clonal abnormalities were defined and described according to the International System for Human Cytogenetic Nomenclature [21].

Molecular mutations screening by next-generation sequencing

Eighty-one patients with *CEBPA*^{dm} were analyzed by a sensitive next-generation sequencing (NGS) assay for 112 genes (see Table 3). The NGS assay was performed as previously described [22], covering 654 coding regions, and approximately 2610000 base pairs. A NimbleGen SeqCap EZ Choice kit was used according to the manufacturer's protocol with some modifications. Multiplexed libraries were sequenced using 75-bp paired-end runs on an Illumina NextSeq 550AR system. Reads were aligned using the Burrows-Wheeler alignment (BWA) tool (version 0.7.5a) against human genomic reference sequences (HG19, NCBI build 37). To identify single nucleotide polymorphisms (SNPs) and short insertions and deletions (INDELs), MuTect2 operation was performed with recommended parameters. All mutations were annotated by the ANNOVAR software. A subset of somatic mutations was randomly selected for validation using Sanger sequencing. Cell line dilutions were prepared for evaluation of sensitivity and specificity. For AML patients in this study, the SCARF file was converted to the FASTQ format by the CASAVA software (version 1.8, Illumina). Raw sequence reads were filtered with an indigenous program. Reads with more than 5% N bases or in which at least 50% bases had Q ≤ 5 were eliminated. The remaining reads were aligned using the BWA tool to the human genomic reference sequences (HG19, NCBI build 37) with certain parameters (mem -t 10 -k 32 -M). To decrease PCR duplication bias, the resulting Bam files were processed with Sam tools. Only unique reads were delivered for analyses. For

identification of SNP and indel, MuTect2 operation was performed with recommended parameters. All mutations were annotated by the ANNOVAR software using the following resources: all annotated transcripts in RefSeq Gene; known constitutional polymorphisms as reported in human variation databases, such as 1000 Genomes (release date 20130308), the Exome Aggregation Consortium (ExAC release date 20151129) and dbSNP (version 135) were download from ANNOVAR; known somatic variations in myeloid and other malignancies as reported in COSMIC (version 70). To identify high-confidence somatic variants in AML samples in the absence of matched control samples, the following criteria were used: removal of all variants within intronic, UTR and intergenic regions, and retention of only nonsynonymous, frame-shift and stop-gain mutations in exonic regions; removal of all variants present in at least one of 81 healthy individuals; removal of all variants with one of the following features in MuTect results: mutation depth of less than four, Phred-scaled *p*-value using Fisher's exact test to detect strand bias of more than 60, mapping quality lower than 40. Because we lacked matched normal samples, somatic mutations could not be selected by comparing a tumor with a matched, normal sample. Thus, a series of steps were used to remove germline mutations and harmless mutations. Mutations were removed unless they satisfied all of the following conditions: the mutation depth was more than four; the mutation occurred in an exonic region; the mutation function was not "synonymous SNV"; the annotation from ClinVar was not "benign" or the mutation did not appear in a dbSNP135 or the 1000 Genomes Project (2012 Feb) database.

Statistics

Statistics Package for Social Sciences (SPSS) software (Version 17.0, SPSS Inc., Chicago, IL, USA) was used to calculate the statistical difference. For categorical variables, the *Chi-square* test or *Fisher's exact* test was used to assess the statistical significance of differences between groups. Independent-samples *t*-test or Mann-Whitney *U*-test was used to compare differences between groups for continuous variables. Kaplan-Meier method was employed for survival analysis, and the log-rank test was used to compare differences between groups. $P < 0.05$ was considered significant in all tests.

Author contributions

SL, TYH, GSJ and YL conceived the study. SL, TYH, GSJ, LXL, and LW designed the study and analyzed the data. SL, LH, LXL, SJN, JFY, BO, YY, YL and LCS recruited the patients. LSS, YYP, YL and LQJ performed the next-generation sequence analysis. SL, TYH, and GSJ wrote this manuscript. All authors discussed and revised the manuscript before submission.

ACKNOWLEDGMENTS AND FUNDINGS

We thank Department of Hematology of the First Hospital, Bethune Medical College of Jilin University, and Department of Hematology, Chinese PLA General Hospital, for their assistance in this work. This work was supported by the Program from Finance Department of Jilin Province (No. 2016swszx004), Clinical Research Foundation of First Hospital of Jilin University (No. LCFYJJ2017005), and the Fifth Youth Developmental Foundation of First Hospital of Jilin University (No. JDYY52014004).

CONFLICTS OF INTEREST

The authors declare no conflicts of interest.

REFERENCES

1. Arber DA, Orazi A, Hasserjian R, Thiele J, Borowitz MJ, Le Beau MM, Bloomfield CD, Cazzola M, Vardiman JW. The 2016 revision to the world health organization classification of myeloid neoplasms and acute leukemia. *Blood*. 2016; 127:2391–2405.
2. Tenen DG, Hromas R, Licht JD, Zhang DE. Transcription factors, normal myeloid development, and leukemia. *Blood*. 1997; 90:489–519.
3. Dufour A, Schneider F, Metzeler KH, Hoster E, Schneider S, Zellmeier E, Benthaus T, Sauerland MC, Berdel WE, Büchner T, Wörmann B, Braess J, Hiddemann W, et al. Acute myeloid leukemia with biallelic CEBPA gene mutations and normal karyotype represents a distinct genetic entity associated with a favorable clinical outcome. *J Clin Oncol*. 2010; 28:570–577.
4. Su L, Gao SJ, Liu XL, Tan YH, Wang L, Li W. CEBPA mutations in patients with de novo acute myeloid leukemia: data analysis in a Chinese population. *Onco Targets Ther*. 2016; 9:3399–403.
5. Ruan GR, Jiang B, Niu JH, Li LD, Li JL, Li N, Leng X, Lai YY, Shi HX, Xu LP, Liu YR, Chen SS, Huang XJ. Detection of CEBPA gene mutations in acute myeloid leukemia and its clinic significance. *Clin J Clin*. 2011; 5:4987–4991.
6. Su L, Gao SJ, Li W, Tan YH, Yang L, Liu ZL, Bai O, Yang Y, Yao C, Song YQ, Wang GJ. Age-specific distributions of cytogenetic subgroups of acute myeloid leukemia: data analysis in a Chinese population. *Acta Haematol*. 2013; 129:175–81.
7. Su L, Gao SJ, Li W, Tan YH, Cui JW, Hu RP. NPM1, FLT3-ITD, CEBPA, and c-kit mutations in 312 Chinese patients with de novo acute myeloid leukemia. *Hematology*. 2014; 19:324–328.
8. Fasan A, Haferlach C, Alpermann T, Jeromin S, Grossmann V, Eder C, Weissmann S, Dicker F, Kohlmann A, Schindela S, Kern W, Haferlach T, Schnittger S. The role of different

- genetic subtypes of CEBPA mutated AML. *Leukemia*. 2014; 28:794–803.
9. Grossmann V, Haferlach C, Nadarajah N, Fasan A, Weissmann S, Roller A, Eder C, Stopp E, Kern W, Haferlach T, Kohlmann A, Schnittger S. CEBPA double-mutated acute myeloid leukaemia harbours concomitant molecular mutations in 76.8% of cases with TET2 and GATA2 alterations impacting prognosis. *Br J Haematol*. 2013; 161:649–658.
 10. Papaemmanuil E, Gerstung M, Bullinger L, Gaidzik VI, Paschka P, Roberts ND, Potter NE, Heuser M, Thol F, Bolli N, Gundem G, Van Loo P, Martincorena I, et al. Genomic classification and prognosis in acute myeloid leukemia. *N Engl J Med*. 2016; 374:2209–2221.
 11. Shen Y, Zhu YM, Fan X, Shi JY, Wang QR, Yan XJ, Gu ZH, Wang YY, Chen B, Jiang CL, Yan H, Chen FF, Chen HM, et al. Gene mutation patterns and their prognostic impact in a cohort of 1185 patients with acute myeloid leukemia. *Blood*. 2011; 118:5593–5603.
 12. Gou H, Zhou J, Ye Y, Hu X, Shang M, Zhang J, Zhao Z, Peng W, Zhou Y, Zhou Y, Song X, Lu X, Ying B. The prevalence and clinical profiles of FLT3-ITD, FLT3-TKD, NPM1, C-KIT, DNMT3A, and CEBPA mutations in a cohort of patients with de novo acute myeloid leukemia from southwest china. *Tumour Biol*. 2016; 37:7357–7370.
 13. Hou HA, Lin YC, Kuo YY, Chou WC, Lin CC, Liu CY, Chen CY, Lin LI, Tseng MH, Huang CF, Chiang YC, Liu MC, Liu CW, et al. GATA2 mutations in patients with acute myeloid leukemia-paired samples analyses show that the mutation is unstable during disease evolution. *Ann Hematol*. 2015; 94:211–221.
 14. Fasan A, Eder C, Haferlach C, Grossmann V, Kohlmann A, Dicker F, Kern W, Haferlach T, Schnittger S. GATA2 mutations are frequent in intermediate-risk karyotype AML with biallelic CEBPA mutations and are associated with favorable prognosis. *Leukemia*. 2013; 27:482–485.
 15. Theis F, Corbacioglu A, Gaidzik VI, Paschka P, Weber D, Bullinger L, Heuser M, Ganser A, Thol F, Schlegelberger B, Göhring G, Köhne CH, Germing U, et al. Clinical impact of GATA2 mutations in acute myeloid leukemia patients harboring CEBPA mutations: a study of the AML study group. *Leukemia*. 2016; 30:2248–2250.
 16. Lavallée VP, Kros J, Lemieux S, Boucher G, Gendron P, Pabst C, Boivin I, Marinier A, Guidos CJ, Meloche S, Hébert J, Sauvageau G. Chemo-genomic interrogation of CEBPA mutated AML reveals recurrent CSF3R mutations and subgroup sensitivity to JAK inhibitors. *Blood*. 2016; 127:3054–61.
 17. Maxson JE, Ries RE, Wang YC, Gerbing RB, Kolb EA, Thompson SL, Guidry Auvil JM, Marra MA, Ma Y, Zong Z, Mungall AJ, Moore R, Long W, et al. CSF3R mutations have a high degree of overlap with CEBPA mutations in pediatric AML. *Blood*. 2016; 127:3094–3098.
 18. Schlenk RF, Taskesen E, van Norden Y, Krauter J, Ganser A, Bullinger L, Gaidzik VI, Paschka P, Corbacioglu A, Göhring G, Kündgen A, Held G, Götze K, et al. The value of allogeneic and autologous hematopoietic stem cell transplantation in prognostically favorable acute myeloid leukemia with double mutant CEBPA. *Blood*. 2013; 122:1576–1582.
 19. Duployez N, Marceau-Renaut A, Boissel N, Petit A, Bucci M, Geffroy S, Lapillonne H, Renneville A, Ragu C, Figeac M, Celli-Lebras K, Lacombe C, Micol JB, et al. Comprehensive mutational profiling of core binding factor acute myeloid leukemia. *Blood*. 2016; 127:2451–2459.
 20. Bennett JM, Catovsky D, Daniel MT, Flandrin G, Galton DA, Gralnick HR, Sultan C. Proposals for the classification of the acute leukaemias. French-American-British (FAB) co-operative group. *Br J Haematol*. 1976; 33:451–458.
 21. Shaffer LG, Slovak ML, Campbell LJ, editors. *ISCN 2009: an international system for human cytogenetic nomenclature (2009)*. Basel: Karger; 2009.
 22. Wang B, Liu Y, Hou G, Wang L, Lv N, Xu Y, Xu Y, Wang X, Xuan Z, Jing Y, Li H, Jin X, Deng A, et al. Mutational spectrum and risk stratification of intermediate-risk acute myeloid leukemia patients based on next-generation sequencing. *Oncotarget*. 2016; 7:32065–32078. <https://doi.org/10.18632/oncotarget.7028>.