

ARTICLE

Open Access

GATA2 zinc finger 1 mutations are associated with distinct clinico-biological features and outcomes different from *GATA2* zinc finger 2 mutations in adult acute myeloid leukemia

Feng-Ming Tien^{1,2,3}, Hsin-An Hou¹, Cheng-Hong Tsai^{1,3}, Jih-Luh Tang^{1,3}, Yu-Chiao Chiu⁴, Chien-Yuan Chen¹, Yuan-Yeh Kuo⁵, Mei-Hsuan Tseng¹, Yen-Ling Peng¹, Ming-Chih Liu⁶, Chia-Wen Liu⁶, Xiu-Wen Liao³, Liang-In Lin⁷, Chien-Ting Lin^{1,3}, Shang-Ju Wu¹, Bor-Sheng Ko¹, Szu-Chun Hsu⁸, Shang-Yi Huang¹, Ming Yao¹, Wen-Chien Chou^{1,8} and Hwei-Fang Tien¹

Abstract

Mutations of the GATA binding protein 2 (*GATA2*) gene in myeloid malignancies usually cluster in the zinc finger 1 (ZF1) and the ZF2 domains. Mutations in different locations of *GATA2* may have distinct impact on clinico-biological features and outcomes in AML patients, but little is known in this aspect. In this study, we explored *GATA2* mutations in 693 *de novo* non-M3 AML patients and identified 44 *GATA2* mutations in 43 (6.2%) patients, including 31 in ZF1, 10 in ZF2, and three outside the two domains. Different from *GATA2* ZF2 mutations, ZF1 mutations were closely associated with French-American-British (FAB) M1 subtype, *CEBPA* double mutations (*CEBPA*^{double-mut}), but inversely correlated with FAB M4 subtype, *NPM1* mutations, and *FLT3*-ITD. ZF1-mutated AML patients had a significantly longer overall survival (OS) than *GATA2*-wild patients and ZF2-mutated patients in total cohort as well as in those with intermediate-risk cytogenetics and normal karyotype. ZF1 mutations also predicted better disease-free survival and a trend of better OS in *CEBPA*^{double-mut} patients. Sequential analysis showed *GATA2* mutations could be acquired at relapse. In conclusion, *GATA2* ZF1 mutations are associated with distinct clinico-biological features and predict better prognosis, different from ZF2 mutations, in AML patients.

Introduction

GATA binding protein 2 (*GATA2*) belongs to the GATA family of transcription factors which regulate hematopoietic stem cell proliferation and differentiation^{1,2}. *GATA2* mutations have been reported in acute

myeloid transformation of chronic myeloid leukemia (CML)³, familial myelodysplastic syndrome-related acute myeloid leukemia (MDS/AML), pediatric MDS^{4,5}, Emberger syndrome⁶, and monocytopenia and mycobacterial infection (MonoMAC) syndrome^{7,8}. Mutations of *GATA2* are also identified in AML patients, with an incidence varied from 3.6% in patients with French-American-British (FAB) M5 subtype⁴ to 8.1–14.4% in non-selected AML patients^{9–11}.

Somatic *GATA2* mutations mainly cluster in the two zinc finger (ZF) domains, which can occupy GATA DNA motif

Correspondence: H-A. Hou (hsinanhou@ntu.edu.tw) or H-F. Tien (hftien@ntu.edu.tw)

¹Division of Hematology, Department of Internal Medicine, National Taiwan University Hospital, Taipei, Taiwan

²Graduate Institute of Clinical Medicine, College of Medicine, National Taiwan University, Taipei, Taiwan

Full list of author information is available at the end of the article.

© The Author(s) 2018



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

in thousands of genes⁹. The patterns of somatic *GATA2* mutations differ among myeloid diseases. ZF1 mutations predominate in AML, and ZF2 mutations are frequently identified in CML blastic phase³. *GATA2* mutations are strongly associated with *CEBPA* double mutations (*CEBPA*^{double-mut})^{9,10,12}. However, discrepancies exist among different reports regarding prognostic impact of *GATA2* mutations in AML patients^{10,13}. We hypothesize that mutations in different domains of *GATA2* may have distinct impact on clinico-biological features and outcomes in AML patients, like *IDH2* mutations in which *IDH2* R172 is associated with gene mutations and clinical outcomes different from other *IDH* mutations¹⁴. However, little is known about this issue till now.

In this study, we investigated the clinical and prognostic relevance of mutations in different *GATA2* domains in a large cohort of 693 unselected *de novo* non-M3 AML patients. To our knowledge, this is the first study to show *GATA2* ZF1 mutations are associated with distinct clinical features, gene mutations, and outcomes different from ZF2 mutations. Longitudinal follow-ups were also performed in 419 samples from 124 patients to evaluate the dynamic changes of the mutations. Furthermore, we analyzed the global gene expression profiles in 328 patients to interrogate the possible molecular pathways associated with mutations in different *GATA2* domains.

Methods and materials

Subjects

We consecutively enrolled 693 newly diagnosed *de novo* non-M3 AML patients at the National Taiwan University Hospital (NTUH) from 1994 to 2011. Diagnosis and classification of AML were made according to the FAB Cooperative Group Criteria and the 2016 WHO classification¹⁵. To focus on a more homogeneous group of patients with *de novo* AML, those with antecedent hematological diseases, history of cytopenia, and family history of myeloid neoplasms or therapy-related AML were excluded¹⁶. Survival analyses were performed in 469 (67.7%) patients who received standard chemotherapy. This study was approved by the Institutional Review Board of the NTUH, and written informed consents were obtained from all participants in accordance with the Declaration of Helsinki.

Cytogenetics

Chromosomal analyses were performed as described previously¹⁷. Karyotypes were classified using Medical Research Council (MRC) risk groups¹⁸.

Mutation analysis

Mutation analysis of *GATA2* exons 2–6¹² and 20 other genes, including *FLT3*-ITD¹⁹, *FLT3*-TKD¹⁹, *NRAS*¹⁹,

*KRAS*¹⁹, *KIT*¹⁹, *PTPN11*²⁰, *CEBPA*²¹, *RUNX1*²², *MLL-PTD*²³, *ASXL1*²⁴, *IDH1*²⁵, *IDH2*²⁵, *TET2*²⁶, *DNMT3A*¹⁶, *SF3B1*²⁷, *SRSF2*²⁷, *U2AF1*²⁷, *NPM1*²⁸, *WT1*²⁹, *TP53*³⁰, and *ETV6*³¹ were performed by Sanger sequencing as previously described for patients ($n = 455$) diagnosed from 1994 to 2007. For patients ($n = 238$) diagnosed after 2008, Ion torrent next-generation sequencing (NGS) (Thermo Fisher Scientific, MA, USA) was performed³². Serial analyses of mutations at diagnosis, complete remission (CR), and relapse were performed in 419 samples from 124 patients by targeted NGS using TruSight Myeloid Panel (Illumina, San Diego, CA, USA). HiSeq platform (Illumina) was used for sequencing with a median reading depth of 12,000 \times ³².

Functional annotation analysis of *GATA2* mutation-regulated genes

We analyzed the differentially expression genes associated with *GATA2* mutations by the knowledge-based Ingenuity Pathway Analysis (IPA) (Qiagen, Redwood City, CA) software for associated functions. We also used Gene Set Enrichment Analysis (GSEA) software to investigate systematic enrichments of *GATA2* mutation-governed expressional profile in biological functions³³. Statistical significance of the degree of enrichment was assessed by a 1000-time random permutation test.

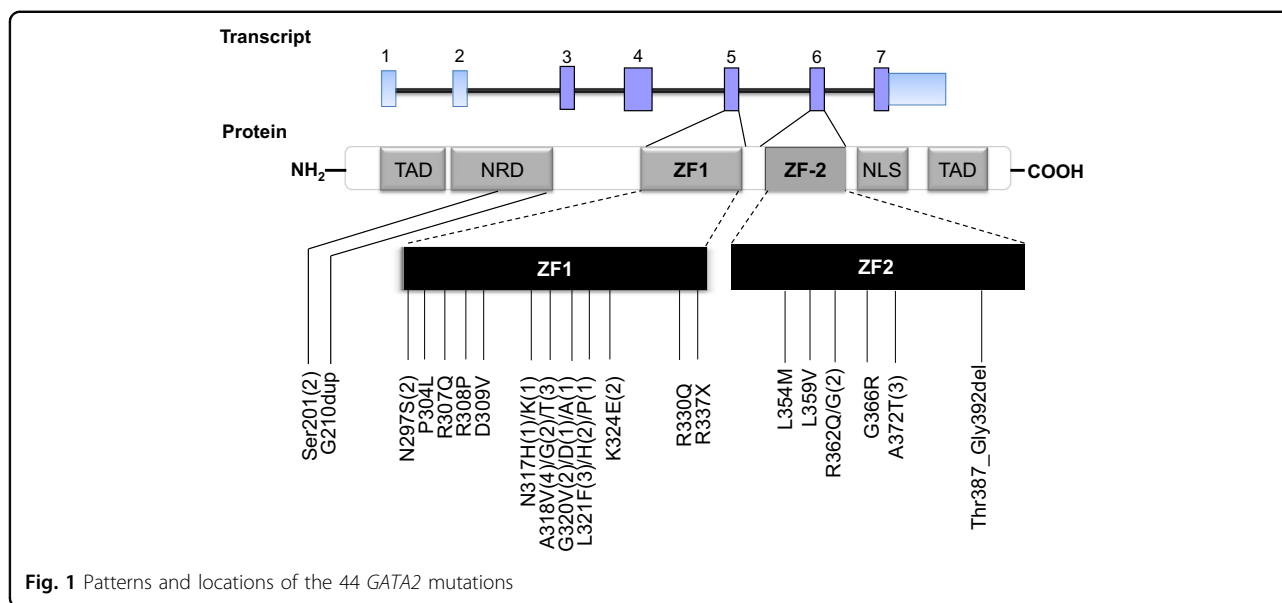
Statistical analysis

The discrete variables were compared using the χ^2 tests, but if the expected values of contingency tables were <5 , Fisher's exact test was used. Mann–Whitney U tests were used to compare continuous variables and medians of distributions. Overall survival (OS) was measured from the date of first diagnosis to the date of last follow-up or death from any cause. Disease-free survival (DFS) was measured from the date of diagnosis until treatment failure, relapse from CR, or death from any cause, whichever occurred first. To ameliorate the influence of hematopoietic stem cell transplantation (HSCT) on survival, DFS and OS were censored at the time of HSCT in patients receiving the treatment³⁴. Multivariate Cox proportional hazard regression analysis was used to investigate independent prognostic factors for OS and DFS. A P value <0.05 was considered statistically significant. All statistical analyses were performed with the SPSS 18 (SPSS Inc., Chicago, IL, USA) and StatsDirect (Cheshire, England, UK).

Results

GATA2 mutations in patients with AML

Excluding two single-nucleotide polymorphisms (A164T, M400T)³⁵ and eight missense mutations (N114T, M223I, P250A, A256V, L315P, C319E, V369A, S429T)



with unknown biologic significance (because they were not reported previously and could not be verified because of lack of matched bone marrow samples in CR), we identified 44 distinct *GATA2* mutations in 43 (6.2%) of 693 patients (Fig. 1). Forty *GATA2* mutations were missense mutations. The other four were in-frame deletion or duplication: p.Ser201*(c.598_599insG) in two, p.Thr387_Gly392del (c.1160_1177delCCATGAAGAAGGAAGGGA) and G210dup (c.631_632insGCG) in one each. With regard to the functional sites, 31 mutations were clustered in the highly conserved N-terminal ZF domain (ZF1 domain), and other 10 mutations were within C-terminal ZF domain (ZF2 domain). The remaining three mutations scattered outside of the ZF domains. The most common mutations were A318V ($n=4$), followed by L321F and A318T ($n=3$ each). p.Ser201*(c.598_599insG), N297S, A318G, G320V, L321H, and K324E occurred in two patients each. All other mutations were detected in only one patient each (Table 1). Only one patient had two *GATA2* mutations (patient no. 20). All mutations were heterozygous. The mutant burden ranged from 4.89 to 52% with a median of 39.07% in ZF1 mutations, and from 10.74 to 50.26% with a median of 36.16% in ZF2 mutations.

Correlation of *GATA2* mutations with clinical and laboratory features

Table 2 depicted the clinical characteristics of patients with and without *GATA2* mutations. ZF1-mutated patients were younger (median, 39 years vs. 55 years, $P=0.004$), and had higher incidence of FAB M1 subtype (56.7% vs. 22.1%, $P<0.0001$), but lower incidence of FAB M4 subtype (3.3% vs. 28.1%, $P=0.003$) than *GATA2*-wild patients. ZF1-mutated patients also had a higher

incidence of FAB M1 subtype than ZF2-mutated patients ($P=0.044$). The patients with ZF2 mutations showed similar clinical features to the *GATA2*-wild group, including peripheral white blood cell counts (median, 47.3 vs. 18.7 k/ μ L), incidences of FAB M1 subtype (20% vs. 22.1%), and M4 subtype (20% vs. 28.1%).

Association of *GATA2* mutations with cytogenetics abnormalities

Chromosome data were available in 669 patients at diagnosis, including 43 *GATA2*-mutated and 626 *GATA2*-wild patients (Supplementary Table 1). Totally, *GATA2* mutations were closely associated with intermediate-risk cytogenetics. Compared to *GATA2*-wild patients, ZF1-mutated patients had more intermediate-risk cytogenetics (100% vs. 70.9%, $P<0.0001$), normal karyotype (73.3% vs. 46.5%, $P=0.004$), and t(3;3) (6.7% vs. 1.0%, $P=0.048$), but less favorable-risk (0% vs. 13.6%, $P=0.024$) or unfavorable-risk cytogenetics (0% vs. 15.5%, $P=0.014$). There was no association of ZF1 mutations with other chromosomal abnormalities, including +8, +11, +13, and +21.

Association of *GATA2* mutations with other molecular alterations

To investigate the interaction of *GATA2* ZF1 and ZF2 mutations with other genetic alterations in the pathogenesis of adult AML, a complete mutational screening of 20 other genes was performed. Only ZF1-mutated patients had a significantly higher frequency of *CEBPA*^{double-mut} (66.7% vs. 6.7%, $P<0.0001$) than wild-type patients, but not ZF2-mutated patients (Table 3). ZF1-mutated patients had lower frequencies of *NPM1* mutations (0% vs. 22%, $P=0.004$) and *FLT3*-ITD (4% vs.

Table 1 The mutation patterns in 43 patients with GATA2 mutations at diagnosis

UPN	Age/sex	Karyotype	Location	GATA2 mutations		Mutant burden (%)	Protein change	Other mutations
				DNA change				
1	29F	CN	ZF1	c.953C>T		52	A318V	<i>CEBPA</i> ^{dm} , <i>FLT3</i> -ITD, <i>NRAS</i>
2	40M	CN	ZF1	c.961C>T		49.37	L321F	<i>CEBPA</i> ^{dm} , <i>NRAS</i>
3	65F	t(3;3)(q21;q26),del(12)(p11p13)	ZF1	c.890A>G		49.04	N297S	<i>NRAS</i> , <i>ASXL1</i>
4	36M	CN	ZF1	c.961C>T		47.42	L321F	<i>CEBPA</i> ^{dm}
5	37M	−Y	ZF1	c.959G>T		47.19	G320V	<i>CEBPA</i> ^{dm}
6	36M	CN	ZF1	c.970A>G		46.14	K324E	<i>CEBPA</i> ^{dm} , <i>NRAS</i>
7	27M	CN	ZF1	c.953C>G		45.45	A318G	<i>CEBPA</i> ^{dm}
8	78F	+8	ZF1	c.1009C>T		45.3	R337X	<i>FLT3</i> -ITD, <i>NRAS</i> , <i>IDH2</i> , <i>SRSF2</i>
9	42M	t(3;3)(q21;q26)/46, idem,add(17)(p13)	ZF1	c.959G>A		44.62	G320D	<i>ASXL1</i> , <i>U2AF1</i>
10	34M	CN	ZF1	c.962T>A		43.97	L321H	<i>CEBPA</i> sm , <i>NRAS</i> , <i>KIT</i> , <i>IDH2</i> , <i>DNMT3A</i>
11	20F	CN	ZF1	c.989G>A		43.14	R330Q	<i>CEBPA</i> ^{dm} , <i>ASXL1</i>
12	32M	CN	ZF1	c.952G>A		42.99	A318T	<i>CEBPA</i> ^{dm} , <i>KIT</i>
13	39M	CN	ZF1	c.911C>T		42.41	P304L	<i>MLL</i> , <i>TET2</i>
14	43M	CN	ZF1	c.923G>C		41.21	R308P	<i>CEBPA</i> ^{dm} , <i>NRAS</i>
15	18M	del(9)(q22q34)	ZF1	c.926A>G		39.07	D309V	<i>CEBPA</i> sm
16	36F	CN	ZF1	c.920G>A		39.06	R307Q	<i>CEBPA</i> ^{dm} , <i>NRAS</i>
17	31F	CN	ZF1	c.970A>G		37.98	K324E	<i>CEBPA</i> ^{dm}
18	55M	CN	ZF1	c.952G>A		32.72	A318T	<i>CEBPA</i> ^{dm}
19	69F	CN	ZF1	c.953C>T		30.26	A318V	<i>CEBPA</i> ^{dm}
20	57M	CN	ZF1	c.962T>A c.949A>C		23.94	L321H, N317H	<i>CEBPA</i> ^{dm} , <i>TET2</i>
21	51M	+21	ZF1	c.953C>T		23.48	A318V	<i>CEBPA</i> ^{dm} , <i>RUNX1</i>
22	39F	46,XX,der(3)t(3;17)(q26;q21),t(16;17)(p11;q11)	ZF1	c.962T>C		20.48	L321P	<i>SF3B1</i>
23	82M	CN	ZF1	c.951T>A		20.46	N317K	<i>RUNX1</i> , <i>SF3B1</i>
24	19F	CN	ZF1	c.959G>C		18.41	G320A	<i>CEBPA</i> ^{dm} , <i>FLT3</i> -TKD
25	59M	CN	ZF1	c.953C>T		18.15	A318V	<i>CEBPA</i> ^{dm} , <i>NRAS</i>
26	29M	CN	ZF1	c.961C>T		17.58	L321F	<i>CEBPA</i> sm
27	50M	CN	ZF1	c.959G>T		13.81	G320V	<i>CEBPA</i> ^{dm} , <i>U2AF1</i>
28	54M	CN	ZF1	c.953C>G		10.81	A318G	<i>CEBPA</i> ^{dm}
29	22F	del(9q)	ZF1	c.952G>A		6.02	A318T	<i>CEBPA</i> ^{dm}
30	78M	CN	ZF1	c.890A>G		4.89	N297S	<i>PTPN11</i> , <i>DNMT3A</i>
31	76M	CN	ZF2	c.1075T>G		50.26	L359V	<i>RUNX1</i>

Table 1 continued

UPN	Age/sex	Karyotype	Location	GATA2 mutations		Mutant burden (%)	Protein change	Other mutations
				DNA change				
32	53F	CN	ZF2	c.1085G>A		48.68	R362Q	ASXL1, IDH2, DNMT3A
33	28F	CN	ZF2	c.1114G>A		46.82	A372T	NPM1
34	69F	CN	ZF2	c.1096G>A		46.33	G366R	NPM1
35	18F	CN	ZF2	c.1114G>A		39.8	A372T	NPM1, PTPN11
36	20M	CN	ZF2	c.1084C>G		23.05	R362G	CEBPA ^{dm} , ASXL1
37	40F	t(7;11)	ZF2	c.1114G>A		21.36	A372T	FLT3-ITD, NRAS
38	60M	-Y	ZF2	c.1084C>G		32.52	R362G	-
39	32F	+10	ZF2	c.1160_1177delCCATGAAGAAGGAAGGGA		17.59	Thr387_Gly392del	CEBPA ^{dm} , NRAS
40	80F	CN	ZF2	c.1061C>T		10.74	T354M	CEBPA ^{dm} , FLT3-ITD
41	71M	del(12)(p12p13), -7		c.598_599insG		35.31	Ser201	PTPN11, RUNX1, ASXL1
42	68F	CN		c.598_599insG		34.1	Ser201	FLT3-ITD, RUNX1, MLL
43	76M	CN		c.631_632insGCG		40.36	G210dup	TP53

UPN unique patient number, CEBPA^{dm} CEBPA double mutation, CN cytogenetically normal, ZF zinc finger

19.9%, $P = 0.024$) than wild-type patients. In contrast, ZF2-mutated patients had similar frequencies of *NPM1* mutations (30%) and *FLT3*-ITD (20%) to those with wild type of *GATA2*. Both ZF1 and ZF2 mutations were mutually exclusive with *KRAS*, *WT1*, *IDH1*, *TP53*, and *ETV6* mutations (Table 3).

Impact of different *GATA2* domains mutations on treatment response and clinical outcomes

Of the 469 AML patients, including 27 *GATA2* ZF1-mutated and nine *GATA2* ZF2-mutated patients, undergoing conventional intensive induction chemotherapy, 352 (75.1%) patients achieved a CR. The CR rate was 85.2% in ZF1-mutated patients and 60% in ZF2-mutated patients (Table 2). The relapse rate was similar between the two groups.

With a median follow-up time of 78.6 months (ranges, 0.1–236 months), patients with *GATA2* mutations as a whole had a trend of longer OS (5-year survival rate, 56% vs. 43%, $P = 0.078$) and DFS (median, 32.9 vs. 8.8 months, $P = 0.091$) than those without *GATA2* mutations (Supplementary Figure 1). Focusing on the prognostic implication of mutation sites, patients with *GATA2* ZF1 mutations had a significantly better OS (5-year survival rate, 72% vs. 43%, $P = 0.003$) and DFS than *GATA2*-wild patients (median, 91.2 vs. 8.8 months, $P = 0.022$) (Fig. 2). In contrast, patients

with *GATA2* ZF2 mutations had similar OS (5-year survival rate, 31%, $P = 0.297$) and DFS (median, 4.4 months, $P = 0.882$) as the *GATA2*-wild group. Intriguingly, ZF1 mutations were also associated with better OS compared with ZF2 mutations ($P = 0.001$) (Fig. 2). In intermediate-risk cytogenetics group, ZF1-mutated patients had significantly superior OS (5-year survival rate, 72% vs. 39%, $P = 0.009$) and DFS (median, 91.2 vs. 7.8 months, $P = 0.006$) than *GATA2*-wild patients, and a longer OS (5-year survival rate, 72% vs. 31%, $P = 0.007$) and a trend toward longer DFS (median, 91.2 vs. 4.4 months, $P = 0.133$) than ZF2-mutated patients (Fig. 3). The finding also held true in normal karyotype subgroup (Supplementary Figure 2). Multivariate analysis demonstrated that ZF1 mutation was an independent favorable prognostic factor for OS (HR 0.207, 95% CI 0.066–0.652, $P = 0.007$) and DFS (HR 0.529, 95% CI 0.295–0.948, $P = 0.032$) irrespective of age, white blood cell counts, cytogenetics, *NPM1*, and *FLT3*-ITD status. However, the prognostic independence of ZF1 mutation was lost if we included *CEBPA*^{double-mut} as a covariable (Supplementary Table 2). We could not find the survival difference stratified by the degree of mutational burden in either ZF1 or ZF2-mutated patients (data not shown). Allo-HSCT in CR1 for ZF1-mutated patients did not offer survival benefit compared to postremission chemotherapy alone (data not shown).

Table 2 Comparison of clinical and laboratory features between AML patients with *GATA2* ZF1 domain and ZF2 domain mutations

Variables	<i>GATA2</i> -wild (<i>n</i> = 650)	<i>GATA2</i> mutations (<i>n</i> = 43)	<i>P</i> value ^a	ZF1 domain mutations (<i>n</i> = 30)	<i>P</i> value ^b	ZF2 domain mutations (<i>n</i> = 10)	<i>P</i> value ^c
Sex ^d			0.876		0.291		0.112
Male	370 (56.9)	25 (58.1)		20 (66.7)		3 (30)	
Female	280 (43.1)	18 (41.9)		10 (33.3)		7 (70)	
Age (year) ^e	55 (15–94)	40 (18–82)	0.017	39 (18–82)	0.004	47 (18–80)	0.365
Lab data ^e							
WBC (k/ μ L)	18.7 (0.12–423)	21.2 (1.23–627.8)	0.200	23.4 (1.33–627.8)	0.195	47.3 (1.23–212.7)	0.494
Hb (g/dL)	8.1 (2.9–16.2)	8.1 (4.2–13.2)	0.704	8.1 (4.4–12.5)	0.436	7.4 (4.2–13.2)	0.311
Platelet (k/ μ L)	47 (3–802)	45 (6–1017)	0.565	47 (6–1017)	0.937	47 (11–119)	0.606
PB Blast(k/ μ L)	7.33 (0–371.9)	9.09 (0–456.7)	0.077	11.3 (0.06–456.7)	0.067	29.9 (0–140.7)	0.358
LDH (U/L)	859 (206–15000)	917 (299–4220)	0.575	970 (327–4220)	0.385	1029 (394–2970)	0.629
FAB ^d							
M0	16 (2.5)	2 (4.7)	0.309	2 (6.7)	0.186	0 (0)	>0.999
M1	144 (22.1)	21 (48.8)	<0.0001	17 (56.7)	<0.0001	2 (20)	>0.999
M2	239 (36.8)	17 (39.5)	0.716	10 (33.3)	0.703	6 (60)	0.186
M4	183 (28.1)	3 (7.0)	0.002	1 (3.3)	0.003	2 (20)	0.734
M5	31 (4.8)	0 (0)	0.248	0 (0)	0.633	0 (0)	>0.999
M6	27 (4.2)	0 (0)	0.403	0 (0)	0.625	0 (0)	>0.999
Unclassified	10 (1.5)	0 (0)	>0.999	0 (0)	>0.999	0 (0)	>0.999
2016 WHO classification ^d							
t(8;21)	57 (8.7)	0 (0)	0.041	0 (0)	0.165	0 (0)	>0.999
Inv(16)	27 (4.2)	0 (0)	0.403	0 (0)	0.625	0 (0)	>0.999
t(9;11)	9 (1.4)	0 (0)	>0.999	0 (0)	>0.999	0 (0)	>0.999
t(6;9)	3 (0.5)	0 (0)	>0.999	0 (0)	>0.999	0 (0)	>0.999
Inv(3)	1 (0.2)	2 (4.6)	0.011	2 (6.7)	0.005	0 (0)	>0.999
t(1;22)	0 (0)	0 (0)	–	0 (0)	–	0 (0)	–
<i>CEBPA</i> ^{dm}	43 (6.6)	22 (51.2)	<0.0001	20 (66.7)	<0.0001	2 (20)	0.144
<i>NPM1</i>	139 (21.3)	3 (7.0)	0.023	0 (0)	0.005	3 (30)	0.455
<i>RUNX1</i>	73 (11.2)	4 (9.3)	>0.999	1 (3.3)	0.237	1 (10)	>0.999
BCR-ABL	1 (0.2)	0 (0)	>0.9999	0 (0)	>0.999	0 (0)	>0.999
MRC	93 (14.3)	0 (0)	0.008	0 (0)	0.025	0 (0)	0.372
AML, NOS	204 (31.4)	12 (27.9)	0.633	7 (23.3)	0.351	4 (40)	0.516
Induction response ^f	431	38		27		9	
Complete remission	323 (74.9)	29 (76.3)	0.851	23 (85.2)	0.230	5 (60)	0.241
Induction death	32 (7.4)	1 (2.6)	0.503	0 (0)	0.243	1 (10)	0.508
Relapse	161 (49.8)	9 (31)	0.052	8 (34.8)	0.163	1 (16.7)	0.371

*CEBPA*sm *CEBPA* single mutation, *CEBPA*^{dm} *CEBPA* double mutation, *MRC* myelodysplasia-related change, *NOS* not otherwise specified, *PB* peripheral blood

^a*GATA2*-mutated patients vs. *GATA2* wild-type patients

^b*GATA2* ZF1-mutated patients vs. *GATA2* wild-type patients

^c*GATA2* ZF2-mutated patients vs. *GATA2* wild-type patients

^dNumber of patients (%)

^eMedian (range)

^fOnly the 469 patients, including 27 with *GATA2* ZF1 domain mutations, nine with *GATA2* ZF2 domain mutations, and 431 without, who received conventional intensive induction chemotherapy and then consolidation chemotherapy if CR was achieved, as mentioned in the text, were included in the analysis

In *CEBPA*^{double-mut} subgroup, *GATA2* ZF1-mutated patients had a trend of longer OS (5-year survival rate, 76% vs. 68%, $P = 0.075$) and a significantly longer DFS (median, 91.2 vs. 14.0 months, $P = 0.034$) than *GATA2*-wild patients (Fig. 4). ZF1 mutations allowed further refinement of the clinical outcome of *CEBPA*^{double-mut} patients. The small number of ZF2-mutated patients ($n = 3$) in this group did not allow statistically meaningful correlations.

Sequential studies of *GATA2* mutations in AML patients

GATA2 mutations were serially studied in 419 samples from 124 patients who had ever obtained a CR and had available samples for study, including 19 patients with and 105 patients without *GATA2* mutations at diagnosis (Table 4). Among the 19 *GATA2*-mutated patients who had paired samples, all lost the original *GATA2* mutations at remission. Five of the six patients regained the original *GATA2* mutations at first relapse, but one (no. 27) lost the

mutation. In the former five patients, the mutation burden, compared to that at diagnosis, was increased in one patient (no. 25), decreased in two (nos. 13 and 16), and stable in the remaining two (nos. 5 and 9). One patient (no. 9) retained the co-occurring *ASXL1* mutations at CR status. Among the 105 patients who had no *GATA2* mutations at diagnosis, four patients (nos. 44, 45, 46, and 47) acquired novel *GATA2* mutations at relapse (Table 4).

***GATA2* expression and biological functions associated with *GATA2* mutations**

We analyzed the microarray dataset of 328 patients studied to assess the impact of *GATA2* mutations on gene expression and biological functions. By comparing the mRNA expression profiles between patients with and without *GATA2* mutations, we found *GATA2* expression levels were higher in those with *GATA2* mutations ($P = 0.003$). More specifically, both ZF1 and ZF2 mutations

correlated with higher *GATA2* expression level compared to *GATA2* wild-type. *GATA2* mutations were associated with significant differential expression of 159 probes (t -test, $P < 0.05$ and >2 -fold change). IPA analysis revealed different molecular networks between the *GATA2* ZF1 and ZF2-mutated group (Supplementary Figure 3). We also performed the GSEA analysis to identify biological functions associated with genes significantly enriched in *GATA2*-mutated AML, compared with *GATA2*-wild AML. Three-hundred and thirteen patients with wild-type *GATA2*, 12 patients with *GATA2* ZF1 mutations, and three patients with *GATA2* ZF2 mutations were analyzed. We identified significant underrepresentation of genes hyper-methylated in AML ($P = 0.006$; normalized enrichment score (NES) = -1.49 ; Supplementary Figure 4A) and genes related to apoptosis ($P = 0.042$; NES = -1.33) in the ZF1-mutated patients compared to *GATA2* wild-type patients. ZF2-mutations were associated with

Table 3 Comparison of other genetic alterations between AML patients according to *GATA2* mutation domain

Mutation	Total pts examined	Pts with the other gene mutations (%)					<i>P</i> value ^a	<i>P</i> value ^b	<i>P</i> value ^c
		Whole cohort	<i>GATA2</i> wt pts	<i>GATA2</i> mutated pts	ZF1	ZF2			
<i>FLT3</i> -ITD	685	19.3	19.9	9.3	4.0	20	0.087	0.024	>0.999
<i>FLT3</i> -TKD	690	8.8	9.2	2.4	3.3	0	0.248	0.508	>0.999
<i>NRAS</i>	691	15.5	14.8	26.8	30	25	0.038	0.035	0.340
<i>KRAS</i>	688	3.6	3.9	0	0	0	0.391	0.620	>0.999
<i>PTPN11</i>	658	5	4.8	7.7	3.6	12.5	0.436	>0.999	0.335
<i>KIT</i>	690	4.8	4.8	4.9	6.7	0	>0.999	0.652	>0.999
<i>WT1</i>	688	6.8	7.3	0	0	0	0.103	0.257	>0.999
<i>NPM1</i>	693	21.1	22	7	0	30	0.019	0.004	0.467
<i>CEBPA</i>	689	14.2	11.1	60.5	76.7	30	<0.0001	<0.0001	0.095
<i>CEBPA</i> ^{dm}	689	9.4	6.7	51.2	66.7	20	<0.0001	<0.0001	0.146
<i>RUNX1</i>	684	14	14.2	11.9	6.7	11.1	0.682	0.413	>0.999
<i>MLL</i> /PTD	636	5.7	5.7	5.1	3.6	0	>0.999	>0.999	>0.999
<i>ASXL1</i>	691	14	14	14	10	20	0.987	0.786	0.640
<i>IDH1</i>	690	6.4	6.8	0	0	0	0.101	0.250	>0.999
<i>IDH2</i>	691	12.7	13.1	7.1	6.7	11.1	0.262	0.410	>0.999
<i>TET2</i>	670	11.9	12.4	4.9	6.9	0	0.212	0.562	0.610
<i>DNMT3A</i>	685	17.4	18	7.3	6.9	11.1	0.080	0.124	>0.999
<i>TP53</i>	685	7.7	8.1	2.4	0	0	0.241	0.158	>0.999
<i>ETV6</i>	649	0.9	0.9	0	0	0	>0.999	>0.999	>0.999
SF	653	11.8	11.7	12.5	17.9	0	0.802	0.366	0.608

Pts patients, *CEBPA*^{dm} *CEBPA*^{double-mutation}, SF splicing factors, including *SF3B1*, *SRSF2*, and *U2AF1*

^a*GATA2*-mutated patients vs. *GATA2* wild-type patients

^b*GATA2* ZF1-mutated patients vs. *GATA2* wild-type patients

^c*GATA2* ZF2-mutated patients vs. *GATA2* wild-type patients

the Gene Oncology term of myeloid leukocyte differentiation ($P = 0.03$; NES = -1.46) (Supplementary Figure 4B). Comparing with ZF2-mutated AML, we identified significant overrepresentation of genes related to myeloid leukocyte differentiation ($P = 0.042$; NES = 1.36) and underrepresentation of genes hyper-methylated in AML ($P = 0.029$; NES = -1.37) in the ZF1-mutated AML.

Discussion

To the best of our knowledge, this is the first study to explore differences in clinical and biological implications between the *GATA2* ZF1 and ZF2 mutations in AML patients. We found that mutations in different domains were associated with distinct clinical features, co-occurring mutations and outcomes (Supplementary Table 3).

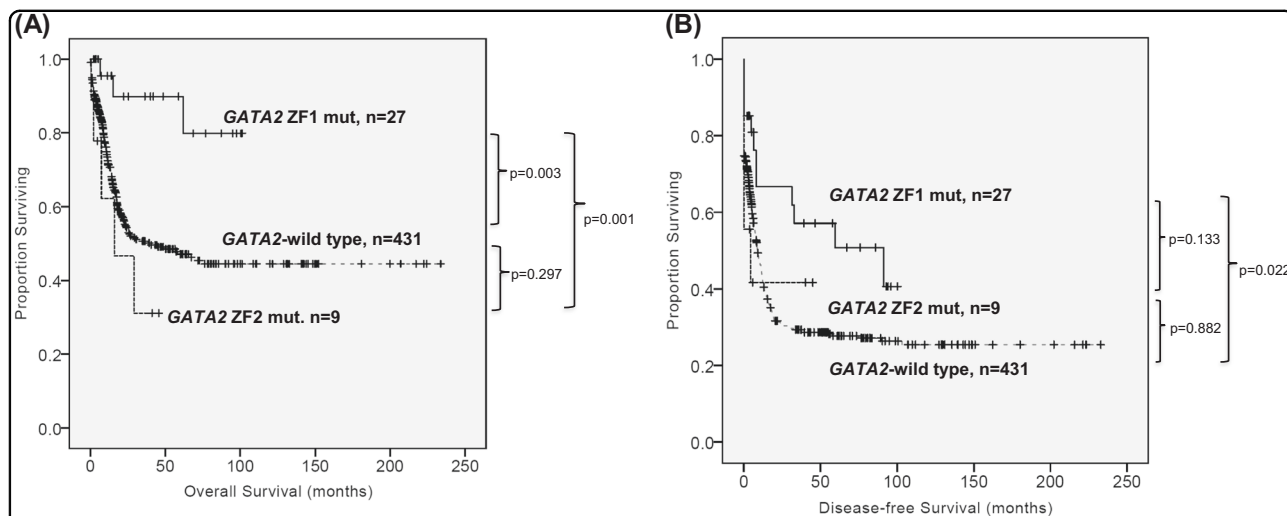


Fig. 2 Kaplan–Meier survival curves for OS (a) and DFS (b) stratified by the *GATA2* mutation status and the sites of mutations in 467 AML patients who received standard intensive chemotherapy. Patients with *GATA2* ZF1 mutations had a significantly better OS (5-year survival rate, 72% vs. 43%, $P = 0.003$) and DFS than *GATA2*-wild patients (median, 91.2 vs. 8.8 months, $P = 0.022$). Patients with *GATA2* ZF2 mutations had similar OS (5-year survival rate, 31%, $P = 0.297$) and DFS (median, 4.4 months, $P = 0.882$) as the wild-type group. ZF1 mutations were also associated with better OS compared with ZF2 mutations ($P = 0.001$)

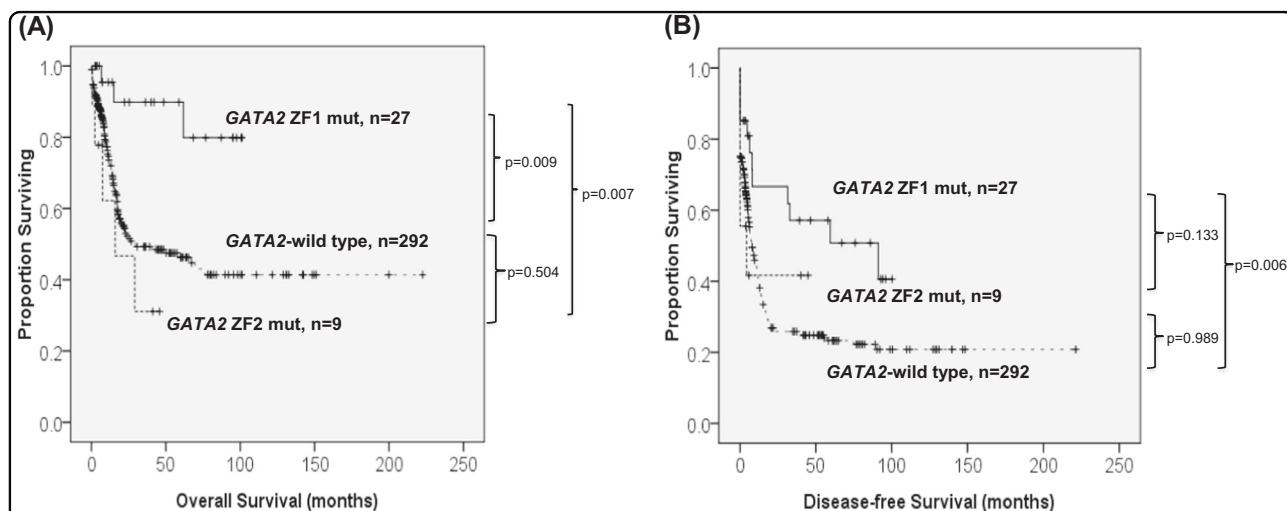
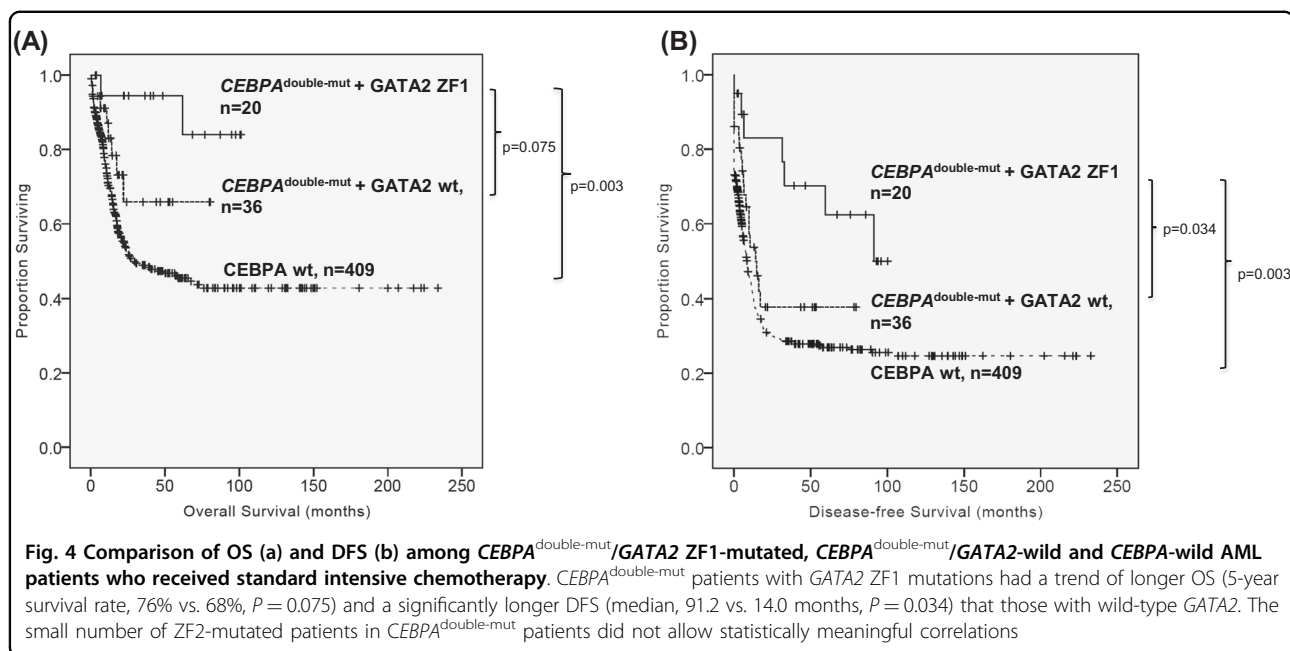


Fig. 3 Kaplan–Meier survival curves for OS (a) and DFS (b) stratified by the *GATA2* mutation status and the sites of mutations in 328 intermediate-risk cytogenetics patients who received standard intensive chemotherapy. Patients with *GATA2* ZF1 mutations had a significantly better OS (5-year survival rate, 72% vs. 39%, $P = 0.009$) and DFS (median, 91.2 vs. 7.8 months, $P = 0.006$) than *GATA2*-wild patients. Patients with *GATA2* ZF2 mutations had similar OS and DFS as the wild-type group ($P = 0.504$, $P = 0.989$, respectively). ZF1 mutations were also associated with a longer OS (5-year survival rate, 72% vs. 31%, $P = 0.007$) and a trend toward longer DFS (median, 91.2 vs. 4.4 months, $P = 0.133$) compared with ZF2 mutations



The *GATA2* mutation landscape in adult *de novo* AML differs from that in blastic crisis of CML³, familial MDS/AML⁴, and pediatric AML⁵. In adult AML, ZF1 mutations predominate, while ZF2 mutations are reported sporadically^{10,36,37}. In concordance with the findings, two-thirds of the 44 distinct *GATA2* mutations in our study were located in the ZF1 domain. We also reported two novel missense mutations in ZF2 domain (L359V and G366R) that had not been reported before in adult *de novo* AML patients, but ever identified in blastic crisis of CML.

AML with *CEBPA*^{double-mut} has been included as a definite entity in the 2016 WHO Classification of Myeloid Neoplasms¹⁵. It is well established that *GATA2* mutations frequently co-occur with *CEBPA*^{double-mut} with an incidence of 18–41%^{9,10,12} and the two proteins show direct protein–protein interaction³⁸. Further study revealed *GATA2* ZF1 mutants, but not the ZF2 L359V that is commonly seen at the progression of CML to blast crisis, had reduced capacity to enhance *CEBPA*-dependent activation of transcription⁹. Based on this functional study and the frequent co-occurrence of *CEBPA*^{double-mut} and ZF1 mutations, but not ZF2 mutations, in AML patients, it is possible that *GATA2* ZF1 mutations and *CEBPA*^{double-mut} interact together to induce leukemogenesis. In addition, we found ZF1 mutations were associated with lower incidences of *NPM1* mutations and *FLT3*-ITD than wild-type *GATA2*, different from ZF2 mutations as ZF2-mutated patients had similar incidences of these two mutations to those in *GATA2*-wild patients. *GATA2* ZF1 and ZF2 mutations may induce AML through different oncogenic mechanisms and have

distinct impact on clinical outcomes. Truly, in this study, we demonstrated that patients with *GATA2* ZF1 mutations had a significantly longer OS than ZF2-mutated patients in total cohort, as well as in patients with intermediate-risk cytogenetics and normal karyotype.

The prognostic impact of *GATA2* mutations in *CEBPA*^{double-mut} patients was conflicting^{12,13,37,39}. Greif et al. and Theis et al. found that *GATA2* mutations did not impact clinical outcome in *CEBPA*^{double-mut} patients. On the contrary, *GATA2* mutations correlated with improved survival among *CEBPA*^{double-mut} patients in other reports^{12,13}. In a study of Theis et al., 31 (74%) of *GATA2* mutations were detected in ZF1 domain, and 11 (26%) in ZF2 domain. They did not show different clinical outcomes with respect to *GATA2* ZF1 and ZF2 mutations in a cohort with both *CEBPA*^{double-mut} and *CEBPA*^{single-mut} patients³⁹. We were the first to investigate the prognostic implication of *GATA2* ZF1 mutations in *CEBPA*^{double-mut} patients and showed its association with a better DFS and a trend of longer OS than wild-type *GATA2* among the *CEBPA*^{double-mut} subgroup.

The poor prognostic impact of *GATA2* ZF2 mutations was also witnessed in blast crisis CML patients as in *de novo* AML patients shown in this study⁴. The reason that ZF1 and ZF2 mutations had different survival impacts on *de novo* AML patients might be partially explained by their difference in association with *CEBPA*^{double-mut}, and by different oncogenic mechanisms. Further studies are warranted to explore the underlying mechanisms of the differences.

The study also recruited the largest number of *de novo* AML patients for sequential analyses of *GATA2*

Table 4 Sequential studies in the AML patients with *GATA2* mutations^a

UPN	Interval ^b (months)	Status	<i>GATA2</i> mutations	Allele burden	Other mutations
1		Initial	Ala318Val	52	<i>CEBPA, FLT3-ITD, NRAS</i>
	0.9	CR1	—	0	—
4		Initial	Leu321Phe	47.42	<i>CEBPA</i>
	1.3	CR1	—	0	—
5		Initial	Gly320Val	47.19	<i>CEBPA</i>
	6.6	CR1	—	0	—
	27.1	Relapse1	Gly320Val	43.1	<i>CEBPA</i>
	1.0	CR2	—	0	—
6		Initial	Lys324Glu	46.14	<i>CEBPA, NRAS</i>
	0.9	CR1	—	0	—
7		Initial	Ala318Gly	45.45	<i>CEBPA</i>
	0.9	CR1	—	0	—
9		Initial	Gly320Asp	44.62	<i>ASXL1, U2AF1</i>
	3.2	CR1	—	0	<i>ASXL1</i>
	6.5	Relapse1	Gly320Asp	43.2	<i>ASXL1</i>
12		Initial	Ala318Thr	42.99	<i>CEBPA, KIT</i>
	0.9	CR1	—	0	—
13		Initial	Pro304Leu	42.41	<i>MLL, TET2</i>
	3.5	CR1	—	0	—
	6.3	Relapse1	Pro304Leu	3.2	—
14		Initial	Arg308Pro	41.21	<i>CEBPA, NRAS</i>
	1.4	CR1	—	0	—
16		Initial	Arg307Gln	39.06	<i>CEBPA, NRAS</i>
	3.0	CR1	—	0	—
	34.7	Relapse1	Arg307Gln	27	<i>CEBPA</i>
18		Initial	Ala318Thr	32.72	<i>CEBPA</i>
	2.1	CR1	—	0	—
20		Initial	Leu321His, Asn317His	11.3	<i>CEBPA, TET2</i>
				23.94	
	1.4	CR1	—	0, 0	—
21		Initial	Ala318Val	23.48	<i>CEBPA, RUNX1</i>
	1.0	CR1	—	0	—
24		Initial	Gly320Ala	18.41	<i>CEBPA, FLT3-TKD</i>
	0.9	CR1	—	0	—
25		Initial	Ala318Val	18.15	<i>NRAS, CEBPA</i>
	1.2	CR1	—	0	—
	12.0	Relapse1	Ala318Val	43.5	<i>CEBPA</i>
27		Initial	Gly320Val	13.81	<i>CEBPA, U2AF1</i>
	1.0	CR1	—	0	—

Table 4 continued

UPN	Interval ^b (months)	Status	GATA2 mutations	Allele burden	Other mutations
	3.5	Relapse1	—	0	<i>CEBPA</i>
	11.7	CR2	—	0	—
	5.9	Relapse2	—	0	<i>CEBPA</i>
29		Initial	Ala318Thr	6.02	<i>CEBPA</i>
	1.0	CR1	—	0	
39		Initial	Thr387_Gly392del	17.59	<i>CEBPA, NRAS</i>
	1.0	CR1	—	0	—
41		Initial	Ser201	35.31	<i>PTPN11, RUNX1, ASXL1</i>
	0.8	CR1	—	0	—
44		Initial	—	0	<i>CEBPA, DNMT3A</i>
	4.5	CR1	—	0	<i>DNMT3A</i>
	2.9	Relapse1	Glu180LysfsTer38	7.1	<i>DNMT3A</i>
	1.1	CR2	—	0	—
	6.0	Relapse2	—	0	<i>DNMT3A</i>
	2.0	CR3	—	0	—
45		Initial	—	0	<i>DNMT3A, NPM1, NRAS, PTPN11</i>
	7.3	CR1	—	0	<i>DNMT3A</i>
	12.5	Relapse1	Arg307Leu	5.6	<i>DNMT3A, NPM1</i>
	1.2	CR2	—	0	<i>DNMT3A</i>
	13.6	Relapse2	—	0	<i>DNMT3A, NPM1</i>
46		Initial	—	0	<i>CEBPA</i>
	2.9	CR1	—	0	—
	14.2	Relapse1	Leu321Pro	26	<i>CEBPA</i>
47		Initial	—	0	—
	1.0	CR1	—	0	—
	15.6	Relapse1	Gly320Asp	15.9	—
			Leu321His	15.1	—
	3.6	CR2	—	0	—
	11.8	Relapse2	Leu321His	39.4	—
	4.8	CR3	—	0	—

UPN unique patient number, CR complete remission, ND not done, “—” negative

^aThe results of serial studies in 101 patients without GATA2 mutation at both diagnosis and relapse were not shown in this table

^bInterval between the two successive statuses

mutations by NGS during clinical follow-ups. The original mutations in all 19 *GATA2*-mutated patients were lost at remission status, confirming them to be truly somatic mutations. We showed *GATA2* mutation was not stable during disease evolution. One (no. 27) of the six patients with *GATA2* mutations at diagnosis lost the mutation at relapse. Among the 105 patients who had no *GATA2* mutations at diagnosis, four (nos. 44, 45, 46, 47) acquired

novel *GATA2* mutations at relapse. The four mutations were all ZF1 mutations.

In conclusion, *GATA2* ZF1 mutations, but not ZF2 mutations, are closely associated with *CEBPA*^{double-mut}, and inversely correlated with *NPM1* mutations and *FLT3*-ITD. The two *GATA2* ZF domain mutations have different impacts on OS in AML patients. *GATA2* ZF1 mutations also affect clinical outcome in *CEBPA*^{double-mut}

patients. Incorporation of *GATA2* ZF1, not ZF2 mutations, allows further refinement of the WHO Classification in the specific entity of AML with *CEBPA*^{double-mut.}

Acknowledgements

We would like to acknowledge the service provided by the DNA Sequencing Core of the First Core Laboratory, National Taiwan University College of Medicine. This work was partially sponsored by grants MOST 104-2314-B-002-128-MY4 and 106-2314-B-002-226-MY3 from the Ministry of Science and Technology (Taiwan), and MOHW 107-TDU-B-211-114009 from the Ministry of Health and Welfare (Taiwan), NTUH 102P06, from the Department of Medical Research, National Taiwan University Hospital, and Taiwan Health Foundation.

Authors' contributions

F.-M.T. was responsible for literature collection, data management and interpretation, statistical analysis, and manuscript writing; H.-A.H. was responsible for study design, plan and coordination, literature collection, data management and interpretation, statistical analysis, and manuscript writing; C.-H. T. and Y.-C.C. were responsible for bioinformatics analysis; L.-I.L. and Y.-Y.K. were responsible for mutation analysis and interpretation; C.-T.L., C.-Y.C., W.-C. C., M.-Y., S.-Y.H., J.-L.T., B.-S.K., S.-C.H., S.-J.W., X.-W.L., and W.T. contributed patient samples and clinical data; Y.-L.P., M.-H.T., C.-W.L., and M.-C.L. performed the gene mutation and chromosomal studies; and H.-F.T. planned, designed, coordinated the study over the entire period, and wrote the manuscript.

Author details

¹Division of Hematology, Department of Internal Medicine, National Taiwan University Hospital, Taipei, Taiwan. ²Graduate Institute of Clinical Medicine, College of Medicine, National Taiwan University, Taipei, Taiwan. ³Tai-Cheng Stem Cell Therapy Center, National Taiwan University, Taipei, Taiwan. ⁴Greehey Children's Cancer Research Institute, University of Texas Health Science Center at San Antonio, San Antonio, TX, USA. ⁵Graduate Institute of Oncology, College of Medicine, National Taiwan University, Taipei, Taiwan. ⁶Department of Pathology, National Taiwan University Hospital, Taipei, Taiwan. ⁷Department of Clinical Laboratory Sciences and Medical Biotechnology, College of Medicine, National Taiwan University, Taipei, Taiwan. ⁸Department of Laboratory Medicine, National Taiwan University Hospital, Taipei, Taiwan

Conflict of interest

The authors declare that they have no conflict of interest.

Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Supplementary Information accompanies this paper at (<https://doi.org/10.1038/s41408-018-0123-2>).

Received: 4 June 2018 Revised: 20 July 2018 Accepted: 1 August 2018

Published online: 31 August 2018

References

- Tsai, F. Y. et al. An early haematopoietic defect in mice lacking the transcription factor *GATA-2*. *Nature* **371**, 221–226 (1994).
- Tsai, F. Y. & Orkin, S. H. Transcription factor *GATA-2* is required for proliferation/survival of early hematopoietic cells and mast cell formation, but not for erythroid and myeloid terminal differentiation. *Blood* **89**, 3636–3643 (1997).
- Zhang, S. J. et al. Gain-of-function mutation of *GATA-2* in acute myeloid transformation of chronic myeloid leukemia. *Proc. Natl Acad. Sci. USA* **105**, 2076–2081 (2008).
- Hahn, C. N. et al. Heritable *GATA2* mutations associated with familial myelodysplastic syndrome and acute myeloid leukemia. *Nat. Genet.* **43**, 1012–1017 (2011).
- Wlodarski, M. W. et al. Prevalence, clinical characteristics, and prognosis of *GATA2*-related myelodysplastic syndromes in children and adolescents. *Blood* **127**, 1387–1397 (2016).
- Ostergaard, P. et al. Mutations in *GATA2* cause primary lymphedema associated with a predisposition to acute myeloid leukemia (Emberger syndrome). *Nat. Genet.* **43**, 929–931 (2011).
- Hsu, A. P. et al. Mutations in *GATA2* are associated with the autosomal dominant and sporadic monocytopenia and mycobacterial infection (MonoMAC) syndrome. *Blood* **118**, 2653–2655 (2011).
- Kazenwadel, J. et al. Loss-of-function germline *GATA2* mutations in patients with MDS/AML or MonoMAC syndrome and primary lymphedema reveal a key role for *GATA2* in the lymphatic vasculature. *Blood* **119**, 1283–1291 (2012).
- Greif, P. A. et al. *GATA2* zinc finger 1 mutations associated with biallelic *CEBPA* mutations define a unique genetic entity of acute myeloid leukemia. *Blood* **120**, 395–403 (2012).
- Fasan, A. et al. *GATA2* mutations are frequent in intermediate-risk karyotype AML with biallelic *CEBPA* mutations and are associated with favorable prognosis. *Leukemia* **27**, 482–485 (2013).
- Green, C. L. et al. *GATA2* mutations in sporadic and familial acute myeloid leukaemia patients with *CEBPA* mutations. *Br. J. Haematol.* **161**, 701–705 (2013).
- Hou, H. A. et al. *GATA2* mutations in patients with acute myeloid leukemia-paired samples analyses show that the mutation is unstable during disease evolution. *Ann. Hematol.* **94**, 211–221 (2015).
- Grossmann, V. et al. *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.* **161**, 649–658 (2013).
- Papaemmanuil, E. et al. Genomic classification and prognosis in acute myeloid leukemia. *N. Engl. J. Med.* **374**, 2209–2221 (2016).
- Arber, D. A. et al. The 2016 revision to the World Health Organization classification of myeloid neoplasms and acute leukemia. *Blood* **127**, 2391–2405 (2016).
- Hou, H. A. et al. DNMT3A mutations in acute myeloid leukemia: stability during disease evolution and clinical implications. *Blood* **119**, 559–568 (2012).
- Hou, H. A. et al. Incorporation of mutations in five genes in the revised International Prognostic Scoring System can improve risk stratification in the patients with myelodysplastic syndrome. *Blood Cancer J.* **8**, 39 (2018).
- Grimwade, D. et al. The importance of diagnostic cytogenetics on outcome in AML: analysis of 1,612 patients entered into the MRC AML 10 trial. The Medical Research Council Adult and Children's Leukaemia Working Parties. *Blood* **92**, 2322–2333 (1998).
- Hou, H. A. et al. Clinical implications of the SETBP1 mutation in patients with primary myelodysplastic syndrome and its stability during disease progression. *Am. J. Hematol.* **89**, 181–186 (2014).
- Hou, H. A. et al. Characterization of acute myeloid leukemia with PTPN11 mutation: the mutation is closely associated with NPM1 mutation but inversely related to FLT3/ITD. *Leukemia* **22**, 1075–1078 (2008).
- Lin, I. I. et al. Characterization of *CEBPA* mutations in acute myeloid leukemia: most patients with *CEBPA* mutations have biallelic mutations and show a distinct immunophenotype of the leukemic cells. *Clin. Cancer Res.* **11**, 1372–1379 (2005).
- Tang, J. L. et al. AML1/RUNX1 mutations in 470 adult patients with de novo acute myeloid leukemia: prognostic implication and interaction with other gene alterations. *Blood* **114**, 5352–5361 (2009).
- Shiah, H. S. et al. Clinical and biological implications of partial tandem duplication of the MLL gene in acute myeloid leukemia without chromosomal abnormalities at 11q23. *Leukemia* **16**, 196–202 (2002).
- Chen, T. C. et al. Dynamics of ASXL1 mutation and other associated genetic alterations during disease progression in patients with primary myelodysplastic syndrome. *Blood Cancer J.* **4**, e177 (2014).
- Lin, C. C. et al. IDH mutations are closely associated with mutations of DNMT3A, ASXL1 and SRSF2 in patients with myelodysplastic syndromes and are stable during disease evolution. *Am. J. Hematol.* **89**, 137–144 (2014).
- Chou, W. C. et al. TET2 mutation is an unfavorable prognostic factor in acute myeloid leukemia patients with intermediate-risk cytogenetics. *Blood* **118**, 3803–3810 (2011).
- Hou, H. A. et al. Splicing factor mutations predict poor prognosis in patients with de novo acute myeloid leukemia. *Oncotarget* **7**, 9084–9101 (2016).

28. Falini, B. et al. Cytoplasmic nucleophosmin in acute myelogenous leukemia with a normal karyotype. *N. Engl. J. Med.* **352**, 254–266 (2005).
29. Hou, H. A. et al. WT1 mutation in 470 adult patients with acute myeloid leukemia: stability during disease evolution and implication of its incorporation into a survival scoring system. *Blood* **115**, 5222–5231 (2010).
30. Hou, H. A. et al. TP53 mutations in de novo acute myeloid leukemia patients: longitudinal follow-ups show the mutation is stable during disease evolution. *Blood Cancer J.* **5**, e331 (2015).
31. Silva, F. P. et al. ETV6 mutations and loss in AML-M0. *Leukemia* **22**, 1639–1643 (2008).
32. Tsai, C. H. et al. Prognostic impacts and dynamic changes of cohesin complex gene mutations in de novo acute myeloid leukemia. *Blood Cancer J.* **7**, 663 (2017).
33. Subramanian, A. et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl Acad. Sci. USA* **102**, 15545–15550 (2005).
34. Hou, H. A. et al. Integration of cytogenetic and molecular alterations in risk stratification of 318 patients with de novo non-M3 acute myeloid leukemia. *Leukemia* **28**, 50–58 (2014).
35. Richards, S. et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet. Med.* **17**, 405–424 (2015).
36. Niimi, K. et al. GATA2 zinc finger 2 mutation found in acute myeloid leukemia impairs myeloid differentiation. *Leuk. Res. Rep.* **2**, 21–25 (2013).
37. Marceau-Renaut, A. et al. Classification of CEBPA mutated acute myeloid leukemia by GATA2 mutations. *Am. J. Hematol.* **90**, E93–E94 (2015).
38. Tong, Q., Tsai, J., Tan, G., Dalgin, G. & Hotamisligil, G. S. Interaction between GATA and the C/EBP family of transcription factors is critical in GATA-mediated suppression of adipocyte differentiation. *Mol. Cell. Biol.* **25**, 706–715 (2005).
39. Theis, F. et al. Clinical impact of GATA2 mutations in acute myeloid leukemia patients harboring CEBPA mutations: a study of the AML study group. *Leukemia* **30**, 2248–2250 (2016).