

blood

2012 120: 395-403
Prepublished online May 30, 2012;
doi:10.1182/blood-2012-01-403220

GATA2 zinc finger 1 mutations associated with biallelic *CEBPA* mutations define a unique genetic entity of acute myeloid leukemia

Philipp A. Greif, Annika Dufour, Nikola P. Konstandin, Bianka Ksienzyk, Evelyn Zellmeier, Belay Tizazu, Jutta Sturm, Tobias Benthous, Tobias Herold, Marjan Yaghmaie, Petra Dörge, Karl-Peter Hopfner, Andreas Hauser, Alexander Graf, Stefan Krebs, Helmut Blum, Purvi M. Kakadia, Stephanie Schneider, Eva Hoster, Friederike Schneider, Martin Stanulla, Jan Braess, Maria Cristina Sauerland, Wolfgang E. Berdel, Thomas Büchner, Bernhard J. Woermann, Wolfgang Hiddemann, Karsten Spiekermann and Stefan K. Bohlander

Updated information and services can be found at:
<http://bloodjournal.hematologylibrary.org/content/120/2/395.full.html>

Articles on similar topics can be found in the following Blood collections
[Myeloid Neoplasia](#) (794 articles)

Information about reproducing this article in parts or in its entirety may be found online at:
http://bloodjournal.hematologylibrary.org/site/misc/rights.xhtml#repub_requests

Information about ordering reprints may be found online at:
<http://bloodjournal.hematologylibrary.org/site/misc/rights.xhtml#reprints>

Information about subscriptions and ASH membership may be found online at:
<http://bloodjournal.hematologylibrary.org/site/subscriptions/index.xhtml>



GATA2 zinc finger 1 mutations associated with biallelic CEBPA mutations define a unique genetic entity of acute myeloid leukemia

*Philipp A. Greif,^{1,2} *Annika Dufour,¹ Nikola P. Konstandin,^{1,2} Bianka Ksienzyk,¹ Evelyn Zellmeier,^{1,2} Belay Tizazu,^{1,2} Jutta Sturm,^{1,2} Tobias Benthous,¹ Tobias Herold,¹ Marjan Yaghmaie,^{1,2} Petra Dörge,³ Karl-Peter Hopfner,⁴ Andreas Hauser,⁵ Alexander Graf,⁶ Stefan Krebs,⁶ Helmut Blum,⁶ Purvi M. Kakadia,¹ Stephanie Schneider,¹ Eva Hoster,¹ Friederike Schneider,¹ Martin Stanulla,³ Jan Braess,⁷ Maria Cristina Sauerland,⁸ Wolfgang E. Berdel,⁹ Thomas Büchner,⁹ Bernhard J. Woermann,¹⁰ Wolfgang Hiddemann,^{1,2} †Karsten Spiekermann,^{1,2} and †Stefan K. Bohlander^{1,2,11}

¹Department of Internal Medicine 3, Ludwig-Maximilians-Universität (LMU) Munich, Munich, Germany; ²Clinical Cooperative Group Leukemia, Helmholtz Zentrum München, German Research Center for Environmental Health, Munich, Germany; ³Department of Pediatrics, University Hospital Schleswig-Holstein, Campus Kiel, Kiel, Germany; Departments of ⁴Biochemistry and ⁵Scientific Computing, and ⁶Laboratory for Functional Genome Analysis at the Gene Center, LMU Munich, Munich, Germany; ⁷Oncology and Hematology, St John-of-God Hospital, Regensburg, Germany; ⁸Institute of Biostatistics and Clinical Research and ⁹Department of Medicine A–Hematology, Oncology and Pneumology, Universität Münster, Münster, Germany; ¹⁰German Society of Hematology and Oncology, Berlin, Germany; and ¹¹Center for Human Genetics, Philipps Universität, Marburg, Germany

Cytogenetically normal acute myeloid leukemia (CN-AML) with biallelic CEBPA gene mutations (biCEBPA) represents a distinct disease entity with a favorable clinical outcome. So far, it is not known whether other genetic alterations cooperate with biCEBPA mutations during leukemogenesis. To identify additional mutations, we performed whole exome sequencing of 5 biCEBPA patients and detected somatic GATA2 zinc finger 1 (ZF1) mutations in 2 of 5 cases. Both GATA2 and CEBPA are transcription factors crucial for

hematopoietic development. Inherited or acquired mutations in both genes have been associated with leukemogenesis. Further mutational screening detected novel GATA2 ZF1 mutations in 13 of 33 biCEBPA-positive CN-AML patients (13/33, 39.4%). No GATA2 mutations were found in 38 CN-AML patients with a monoallelic CEBPA mutation and in 89 CN-AML patients with wild-type CEBPA status. The presence of additional GATA2 mutations (n=10) did not significantly influence the clinical outcome of 26 biCEBPA-positive patients. In reporter

gene assays, all tested GATA2 ZF1 mutants showed reduced capacity to enhance CEBPA-mediated activation of transcription, suggesting that the GATA2 ZF1 mutations may collaborate with biCEBPA mutations to deregulate target genes during malignant transformation. We thus provide evidence for a genetically distinct subgroup of CN-AML. The German AML cooperative group trials 1999 and 2008 are registered with the identifiers NCT00266136 and NCT01382147 at www.clinicaltrials.gov. (Blood. 2012;120(2):395-403)

Introduction

In 2008, the World Health Organization classification of Tumors of the Hematopoietic and Lymphoid Tissues included “AML with mutated CEBPA” as a provisional favorable prognostic entity, and a routine CEBPA mutation screening was recommended for all acute myeloid leukemia (AML) patients with no detectable chromosomal abnormalities.^{1,2} However, only patients with mutations in both CEBPA alleles (biCEBPA-mutated AML) have a prognostically favorable outcome.³⁻⁷ The biCEBPA patients typically have a combination of an N-terminal frameshift mutation leading to a 30-kDa dominant-negative CEBPA isoform and a C-terminal in-frame mutation in the bZIP region, disrupting dimerization and DNA-binding activities of CEBPA. In murine bone marrow transplantation models, distinct but collaborative roles of both types of CEBPA mutations have been detected.⁸ Furthermore, a combination of an N- and C-terminal disruption of the CEBPA gene synergistically resulted in fast and efficient development of leukemia in mice.⁸⁻¹⁰ These experiments suggest that a biallelic disruption of CEBPA might be responsible for both the differentiation

block and the enhanced proliferation of progenitor cells and thus be sufficient for leukemogenesis. Indeed, biCEBPA mutant AML represents as a very homogeneous AML subgroup with a distinct immunophenotype¹¹ and a characteristic gene expression profile.^{3,6,7,12} Furthermore, biCEBPA mutant AML patients rarely have mutations in genes that are frequently mutated in CN-AML, such as the nucleophosmin (NPM1) gene, internal tandem duplications (ITDs) of the FLT3 gene, partial tandem duplications (PTDs) of the MLL gene (MLL-PTD),³⁻⁵ or mutations in the TET2, isocitrate dehydrogenase (IDH)–1 and 2 or DNMT3A genes.¹³⁻¹⁵ To identify potentially cooperating mutations associated with biCEBPA-mutated AML, we performed whole exome sequencing. Thereby, we discovered a high frequency of N-terminal zinc finger (ZF1) mutations of GATA2 in patients with biCEBPA-mutated AML. GATA2 is a zinc finger transcription factor important for hematopoietic stem cell proliferation¹⁶⁻¹⁹ and normal megakaryocytic development.²⁰ Somatic mutations affecting the C-terminal zinc finger (ZF2) of GATA2 are associated with the progression of chronic myeloid leukemia (CML),²¹

Submitted January 12, 2012; accepted May 15, 2012. Prepublished online as Blood First Edition paper, May 30, 2012; DOI 10.1182/blood-2012-01-403220.

*P.A.G. and A.D. contributed equally to this work.

†K.S. and S.K.B. jointly supervised this work.

The online version of the article contains a data supplement.

Presented in abstract form at the 53rd Annual Meeting of the American Society of Hematology, San Diego, CA, December 11, 2011 (Abstract 72).

The publication costs of this article were defrayed in part by page charge payment. Therefore, and solely to indicate this fact, this article is hereby marked “advertisement” in accordance with 18 USC section 1734.

© 2012 by The American Society of Hematology

whereas hereditary *GATA2* ZF2 mutations predispose to AML and myelodysplastic syndrome.²² Mutations targeting either of the 2 ZFs were described as a rare event in the M5 subtype of AML.²³ Here, we report a unique association of *GATA2* ZF1 mutations and *biCEBPA* mutations in AML.

Methods

Patients

In this analysis, we included diagnostic bone marrow or peripheral blood samples from 160 adult AML patients with a normal karyotype of which 146 AML patients were enrolled in the German AML cooperative group (AMLCG) 1999 multicenter treatment trial. These patients were also investigated in a previous publication with a different objective.³ Four patients with *biCEBPA* mutations were treated within the AMLCG 2008 study. Sample selection for exome sequencing was dependent on availability of remission samples. Available clinical characteristics were age, sex, de novo versus secondary AML, French-American-British (FAB) subtype, white blood cell and platelet counts, the percentage of bone marrow blasts, and hemoglobin and lactate dehydrogenase levels. All patients treated within the AMLCG 1999 trial received intensive induction therapy with either thioguanine, cytarabine, and daunorubicin as standard therapy or high-dose cytarabine and mitoxantrone (HAM) in the experimental arm followed by 1 course of HAM and consolidation therapy. In patients with an age of ≥ 60 years, a second induction was only administered in case of inadequate response to initial intensive induction treatment. Details of the trial protocols have been published previously.²⁴ Patients treated within the AMLCG 2008 trial received dose-dense sequential HAM versus age-dependent standard double induction and thioguanine, cytarabine, and daunorubicin-9 consolidation.²⁵ The study protocols were approved by the ethics committees of the participating centers, and all patients provided written informed consent to the scientific use of surplus samples (eg, bone marrow, peripheral blood) in accordance with the Declaration of Helsinki.

Sample preparation and high-throughput sequencing

Genomic DNA was extracted from patients' bone marrow or peripheral blood samples using QIAcube technology (QIAGEN). Using the Bioruptor sonicator (Diagenode), 5 μ g of genomic DNA was fragmented to an average size of 150 bp. Ultrasound was applied during 3 cycles of 15 minutes each at low power as described previously.^{26,27} Sequencing libraries were prepared using DNA sample prep reagent set 1 (NEBNext). In brief, library preparation included end repair, adapter ligation, and PCR enrichment. Exon-coding sequences were then captured using SureSelect human all exon 50Mb kit (Agilent) according to the manufacturer's instructions. Exome libraries were then sequenced by performing 76-bp paired end reads on the Genome Analyzer IIx platform (Illumina).

Sequence alignment

Short read alignment to the human genome assembly (build NCBI36/hg18) was performed using the BWA sequence alignment program²⁸ with the default parameters. For each sample, we generated at least 90 000 000 paired end reads of 76 bp (7 Gbp), of which at least 84 000 000 (6.4 Gbp) could be aligned to the reference sequence (supplemental Table 1, available on the *Blood* Web site; see the Supplemental Materials link at the top of the online article). Read mapping, subsequent assembly, and variant calling were performed using the resequencing software packages BWA and SAMtools.^{28,29} During alignment, apparently duplicated reads were removed. Global exon coverage and histograms (supplemental Table 1; supplemental Figure 1) were calculated using the BEDtools software.³⁰

Variant detection

Variant calling was performed using the VarScan software.³¹ For the AML samples, we excluded all positions with a read depth of less than 10 and required (1) each putative variant to have a median quality value of the

variant bases of at least 15, (2) that at least 20% of all reads covering the position show the variant allele with an absolute minimum number of 5 variant reads, and (3) that reads showing the variant allele are from both strands. For the remission samples, we included all potential variants with a minimum frequency of 5% regardless of read depth, base quality, absolute number of variant reads, or number of affected strands. Analysis of the effect of the variants on the protein level was performed with VarScan scripts using datasets from Agilent, Ensembl, and the Santa Cruz Genome Browser as described previously.^{32,33} Variants were annotated using a Variant Effect Predictor perl script.³⁴ Alignment inspection was performed using the Integrative Genomics Viewer.³⁵

Sanger sequencing

Candidate nonsynonymous somatic variants (present in the AML but not in the remission sample) were verified by bidirectional DNA sequencing using ABI 3100-Avant technology (Applied Biosystems) of the corresponding gene after PCR amplification using exon-spanning primers.

High-resolution melting curve analysis

To verify the frequency of selected gene mutations in a larger cohort of *biCEBPA*-mutated patients ($n = 33$), we scanned all coding exons 1 to 5 of the *GATA2* gene and all coding exons 1 to 7 of the *IKZF1* gene and exons 15 to 23 of the *DNMT3A* gene for sequence variations using high-resolution melting analysis. Because of limited sample availability, we confined the *DNMT3A* mutation screening to exons 15 to 23 in which most mutations have been detected in AML so far.^{13,36,37} Exon-spanning primers were designed to generate amplicons of up to 300 bp. Approximately 15 to 30 ng of genomic DNA was amplified using the LightCycler High Resolution Melting master (Roche Diagnostics) following the manufacturer's instructions. PCR products were directly examined on a 1.5% Gel Red (Biotium, distributed by Biotrend) agarose gel. Differences in the high-resolution melting profile in comparison with the wild-type genes in HL-60 cells were confirmed by bidirectional DNA sequencing.

MLPA analysis

Multiplex ligation-dependent probe amplification (MLPA) analysis was performed using the MLPA SALSA kit P335-A4 ALL-*IKZF1* (MRC-Holland) according to the manufacturer's protocol. The assay includes probes for each of the 8 exons of the *IKZF1* gene and is therefore able to detect deletions of the whole gene as well as all types of focal intragenic deletions in exons. Separation of MLPA products were performed on a 3730 DNA Analyzer with GeneScan 500 LIZ as internal size standard (both Applied Biosystems). Data were analyzed using Peak Scanner Version 1.0 and Coffalyser Version 9.4 software. Relative copy number was calculated after intrasample normalization against control fragments and intersample normalization against control samples (healthy blood donor DNA). A ratio of 1 ± 0.3 represents a normal copy number of 2, ratios < 0.7 and < 0.3 represent heterozygous and homozygous deletions, respectively, whereas ratios > 1.3 indicate amplifications.

Plasmids

The pFLAG-CMV2-GATA2 construct was purchased from Addgene. For the following *GATA2* point mutants (A318T, G320D, L321V, R293Q, V296G, and L359V), a 560-bp fragment was synthesized by GeneArt (Regensburg) and inserted into pFLAG-CMV2-GATA2 using restriction sites SacII and HpaI. The pcDNA3-CEBPA expression plasmids (wild type and p30) and the 4xCAAT-binding-site-TK-LUC reporter are a gift from Daniel G. Tenen (Harvard Medical School, Boston, MA). The *GATA-LUC* reporter was purchased from Panomics.

Immunoblotting analysis

Cellular lysates from human embryonic kidney (HEK) 293T cells were electrophoresed on 10% to 12% SDS-PAGE and transferred to polyvinylidene difluoride membrane (Hybond P; Amersham Pharmacia Biotech). The membranes were blocked for 30 minutes with 5% nonfat dried milk at

Table 1. Confirmed tumor-specific somatic mutations identified by exome sequencing in 5 patients with biCEBPA-mutated AML

Patient	Gene	Genomic position (hg18)	Reference genotype	Variant genotype	Amino acid	Ensembl transcript	Read depth in AML	Variant frequency, %
1	<i>RBM39</i>	chr20:33765551	G/G	G/A	R356C	ENST00000361162	92	40.22
	<i>IKZF1</i>	chr7:50417782	G/G	G/A	G158S	ENST00000359197	37	43.24
	<i>FBLN7</i>	chr2:112639148	G/G	G/A	R112Q	ENST00000331203	36	25
	<i>GATA2</i>	chr3:129685449	G/G	G/C	L321V	ENST00000341105	22	40.91
	<i>KRAS</i>	chr12:25289548	C/C	C/T	G13D	ENST00000256078	28	21.43
2	<i>GALNT10</i>	chr5:153769409	G/G	G/A	R427H	ENST00000297107	85	40
	<i>STAG2</i>	chrX:123006890	-/A	-/-C	L221Tfs	ENST00000371145	85	83.5
3	<i>GABRE</i>	chrX:150874947	-/G	-/A	R276W	ENST00000370328	57	96.49
	<i>ZNF318</i>	chr6:43431833	G/G	G/A	S406F	ENST00000318149	44	40.91
	<i>STK19</i>	chr6_cox_hap1:3386347	A/A	A/T	T76S	ENST00000375333	13	46.15
	<i>DNMT3A</i>	chr2:25323985	G/G	*/+A	G332Rfs	ENST00000321117	51	27.45
4	<i>TRPM8</i>	chr2:234543699	C/C	C/T	L749F	ENST00000324695	121	48.76
	<i>AFTPH</i>	chr2:64633197	C/C	C/T	T362I	ENST00000238855	110	45.45
	<i>PSD3</i>	chr8:18703129	G/G	G/A	R118X	ENST00000286485	76	50
	<i>DSC1</i>	chr18:26979591	G/G	G/A	T307I	ENST00000257198	83	44.58
	<i>ASAP3</i>	chr1:23632600	G/G	G/A	Q678X	ENST00000336689	78	34.62
	<i>B4GALNT1</i>	chr12:56306985	G/G	G/A	L471F	ENST00000341156	24	54.17
	<i>FAM117A</i>	chr17:45152254	G/G	G/A	A192V	ENST00000240364	12	50
	<i>DNMT3A</i>	chr2:25310746	C/C	C/T	R882H	ENST00000321117	32	46.88
5	<i>FLJ35848</i>	chr17:40099629	G/G	G/C	G275A	ENST00000409122	177	49.72
	<i>PCDHA13</i>	chr5:140243097	T/T	T/C	S354P	ENST00000289272	92	40.22
	<i>NEUROD4</i>	chr12:53706752	C/C	C/T	R88X	ENST00000242994	78	42.31
	<i>GATA2</i>	chr3:129685421	C/C	C/T	R330Q	ENST00000341105	17	58.82

* indicates wild-type allele.

room temperature and probed with polyclonal rabbit GATA2 antibodies (H116; sc-9008, Santa Cruz Biotechnology) or mouse monoclonal β -actin antibodies (clone AC-15; Sigma-Aldrich), followed by secondary antibodies conjugated to horseradish peroxidase. Proteins were detected with enhanced chemiluminescence (ECL; GE Healthcare). Quantification of Western blot signals was performed using ImageJ Version 1.45 software.

Immunostaining and confocal laser scanning fluorescence microscopy

For intracellular localization studies, HEK 293T cells were grown on coverslips and cotransfected with FLAG-GATA wt or FLAG-GATA2 mutants. After 24 hours, cells were fixed with 2% paraformaldehyde in PBS for 10 minutes, permeabilized with 0.1% Triton X in PBS for 10 minutes, and blocked with 10% FCS in PBS for 1 hour. Coverslips were incubated with monoclonal mouse FLAG antibodies (M2; Sigma-Aldrich) overnight at 4°C. After extensive washing with PBS, Alexa 555-conjugated secondary antibodies were added for 1 hour at room temperature. After further washing steps, cells were stained with 4,6-diamidino-2-phenylindole and mounted using fluorescent mounting medium (Dako Deutschland). Finally, immunostained specimen were analyzed with a confocal fluorescence laser scanning system (TCS-SP2 scanning system and DM IRB inverted microscope; Leica).

Reporter gene assays

HEK 293T cells were seeded in 24-well plates and cotransfected with 100 ng of luciferase reporter plasmid, 10 ng of pRL coreporter vector (Promega), and variable amounts of pFLAG-GATA2 wt, pFLAG-GATA2 mutants, or pcDNA3-CEBPA. The total amount of transfected DNA in each well was kept constant by adding pcDNA3 empty vector. After 24 hours, cells were harvested and assayed for firefly and *Renilla* luciferase activities using a Dual-Luciferase reporter assay system (Promega). Measurements of *Renilla* luciferase activity were used for normalization. Experiments were performed in triplicates and repeated at least 3 times. All results are reported as mean $\pm$ standard error of the mean. Pairwise comparisons were performed using a Student *t* test.

Statistical analyses

Comparisons of different categoric parameters according to the *GATA2* mutation status were performed using the χ^2 test. Fisher exact test was used instead of the χ^2 test when at least 1 cell had an expected frequency of < 5 . For continuous variables, we used the Mann-Whitney *U* test. Outcome variables were overall survival (OS) and event-free survival (EFS). OS was calculated from randomization to death from any cause or to the latest follow-up date. EFS was defined as the period from the start of therapy until lack of a complete remission, relapse after complete remission or death without relapse. Patients treated within the AMLCG 2008 study were excluded from the survival analysis because of differences in treatment protocols and a short follow-up time. Survival curves were calculated according to the Kaplan-Meier method and compared with the log rank test. All tests were 2-tailed, and *P* values $< .05$ were considered statistically significant.

Statistical analyses were performed with SPSS Version 16.0 and the R 2.7.2 software package (R Foundation for Statistical Computing).

Structural model of GATA2

The model is based on the crystal structure of the highly homologous GATA3³⁸ (PDB accession 3DVV). The backbone was left unaltered. All displayed side chains that harbor mutations except one are conserved between GATA2 and GATA3 and were left unaltered. The remaining side chain was modeled with COOT³⁹ using the most common rotamer.

Results

Exome sequencing of 5 biCEBPA AML cases

To identify collaborating mutations, we sequenced the exomes (protein coding regions) of 5 biCEBPA-mutated CN-AML leukemia samples and the corresponding remission samples representing the patients' germlines. We generated a minimum of 7 Gbp of sequence

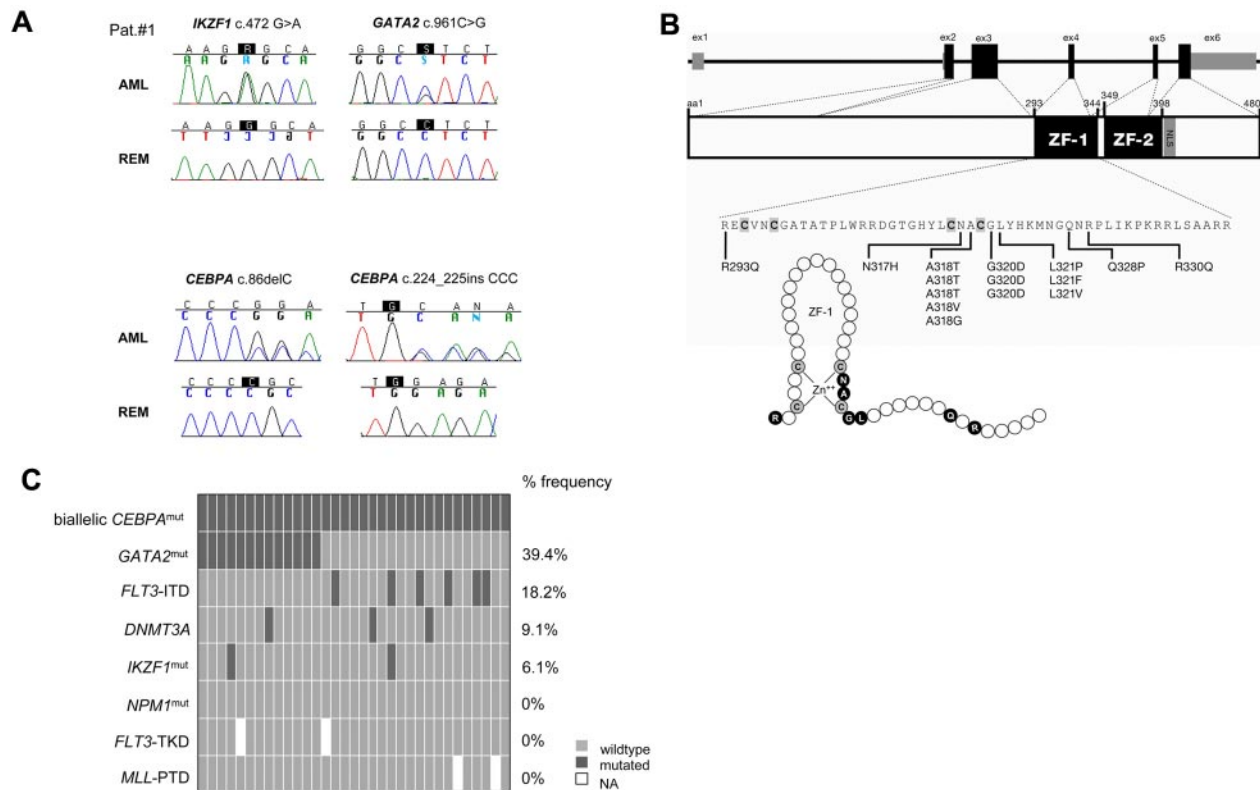


Figure 1. Additional genetic aberrations identified in *biCEBPA* mutant CN-AML. (A) Somatic mutations affected 3 key regulators of hematopoiesis, namely, *CEBPA*, *GATA2*, and *Ikaros* (*IKZF1*) in 1 CN-AML patient analyzed by exome sequencing (also see Table 1). Chromatograms are shown for both diagnostic AML samples and corresponding follow-up samples at complete remission (REM) from the same patients. (B) Structure of the human *GATA2* gene with 5 coding exons is shown in black, and noncoding exons are shown in gray (top). The *GATA2* protein structure includes the N- and C-terminal zinc finger domains (ZF1, ZF2) and the nuclear localization signal (NLS; middle). The amino acid sequence of the N-terminal ZF1 domain of *GATA2* is detailed with cysteine residues that are responsible for zinc binding are shown in bold (bottom). Fifteen heterozygous missense mutations were detected in coding exon 4 of *GATA2* in 13 of 33 *biCEBPA* patients (overall frequency, 39.4%). All *GATA2* mutations clustered in the highly conserved N-terminal zinc finger domain (ZF domain) of *GATA2*. Positions A318, G320, and G321 were recurrently affected. The ZF1 also is illustrated as polypeptide chain with the amino acids (circles) targeted by mutations in black (bottom). (C) Frequency distribution of additional genetic aberrations in 33 *biCEBPA* patients. Each box indicates 1 patient. Dark gray boxes are indicative for patients who are positive for the respective mutation; light gray boxes indicate wild-type status. Missing information is shown as a white space. All 33 *biCEBPA*-mutated patients were *NPM1* wild type and did not carry an additional *FLT3*-TKD or *MLL*-PTD. *FLT3*-ITD indicates internal tandem duplications in the *FLT3* gene; *FLT3*-TKD, *FLT3* mutations in the tyrosine kinase domain; *MLL*-PTD, partial tandem duplications in the *MLL* gene; and *NPM1*, nucleophosmin.

for each exome. This allowed us to cover at least 80% of RefSeq coding exon positions with a minimum read depth of 10 (supplemental Table 1; supplemental Figure 1). Comparison of the AML exome sequence with the remission exome sequence and exclusion of annotated polymorphisms led to the identification of leukemia-specific variants. Between 2 and 8 nonsynonymous coding somatic mutations per patient were confirmed using Sanger sequencing (Table 1). Thus, we detected tumor-specific mutations (nonsense and missense) in a total of 21 genes. *DNMT3A* and *GATA2* were found to be mutated in 2 of the 5 *biCEBPA* CN-AML samples. *DNMT3A* encodes a DNA-methyltransferase with an overall mutation frequency of ~20% in AML.

The *IKZF1* G158S mutation in patient 1 has been identified previously in acute lymphoblastic leukemia and was functionally characterized as dominant negative.^{40,41} Deletions of *IKZF1* are frequent events in acute lymphocytic leukemia and also occur at the progression of chronic myeloid leukemia to lymphoid blast crisis.^{41,42} Thus, 3 key regulators of hematopoiesis were targeted by somatic mutations in patient 1, namely, *CEBPA*, *GATA2*, and *IKZF1* (*Ikaros*; Figure 1A; Table 1). In patient 2 we found a mutation of *STAG2* (Table 1) that is part of the cohesin complex. Alterations of the cohesin genes were reported previously in

myeloid diseases.⁴³ *KRAS* mutations (patient 1) occur in 5% of AML patients and are mainly associated with FAB M4.⁴⁴

High frequency of *GATA2* ZF1 mutations in *biCEBPA* AML

Based on the exome sequencing results, *GATA2*, *DNMT3A*, and *IKZF1* were selected for mutational screening in 33 *biCEBPA*-mutated AML patients. Strikingly, we detected 15 heterozygous missense mutations in coding exon 4 of *GATA2* in 13 of 33 *biCEBPA* patients (39.4%). Two patients were found to carry 2 different mutations in *GATA2*. The *GATA2* mutations affected 7 different amino acid positions (Figure 1B). All mutations were in the highly conserved N-terminal zinc finger domain (ZF1 domain) of *GATA2* (Figure 1B). A mutational hotspot within the ZF1 domain between amino acids alanine 317 and glycine 321 surrounding cysteine 319 accounted for 12 of the 15 mutations. This is a highly conserved region in different *GATA* factors and among different species (supplemental Figure 2). The missense mutations Ala318Thr and Gly320Asp were detected in 6 of 13 *GATA2* mutated *biCEBPA* patients (3 with Ala318Thr and 3 with Gly320Asp). Four of 13 *biCEBPA* patients with *GATA2* mutations also were analyzed during remission. All 4 patients had lost the *GATA2* mutation at this point (Figure 1A; data not shown). Furthermore, no

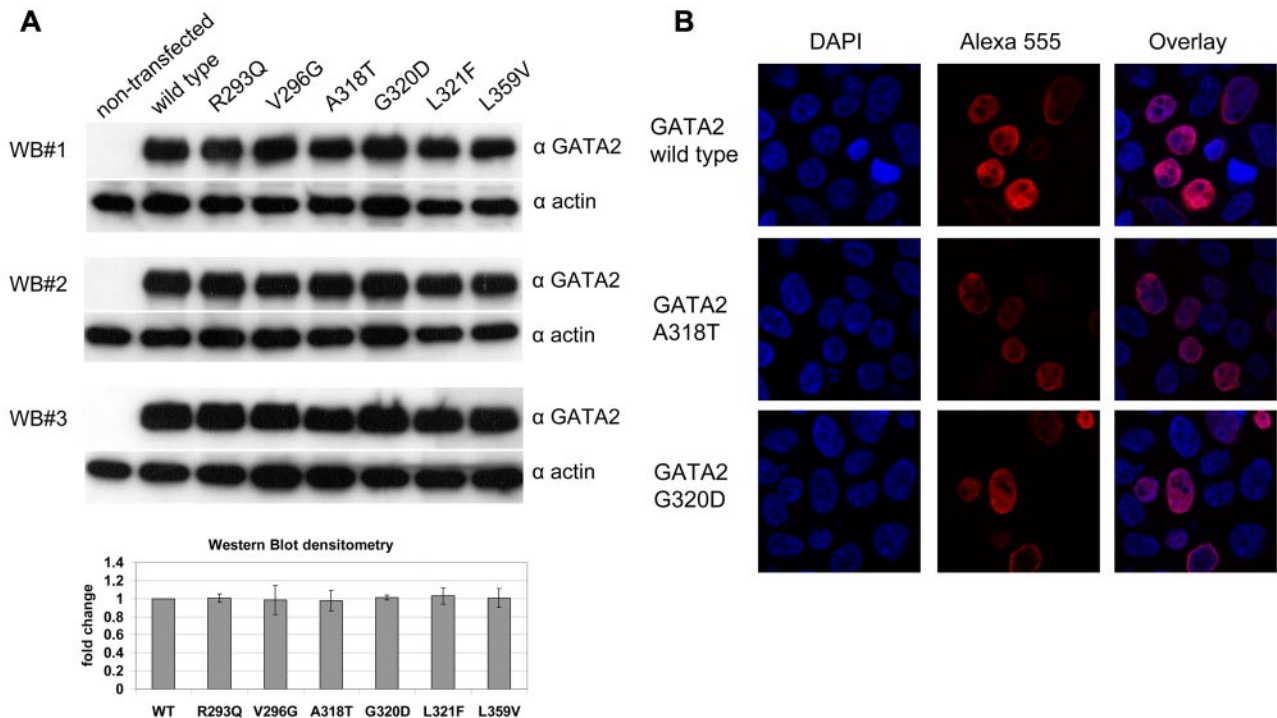


Figure 2. Expression of GATA2-mutants in HEK 293T cells. (A) Expression of FLAG-tagged GATA2-mutants confirmed by 3 independent Western blot (WB) experiments using anti-GATA2 antibodies (top) and anti- β -actin antibodies (bottom). Graph shows mean values of the GATA2 signals normalized to β -actin signals as fold change compared with wild-type GATA2. Error bars indicate standard error of the mean. (B) Immunofluorescence of GATA2 using anti-FLAG antibodies and Alexa 555-labeled secondary antibodies. DNA was counter stained using 4,6-diamidino-2-phenylindole. Confocal laser scans show nuclear localization for GATA2 wild type (top) and the recurring GATA2-mutants A318T (middle) or G320D (bottom). Similar results were obtained for all GATA2-mutants tested (see supplemental Figure 3).

GATA2 mutations were found in 38 patients with a monoallelic *CEBPA* mutation and in 89 patients with wild-type *CEBPA*. The association between GATA2 ZF1 mutations and bi*CEBPA* mutations in a total cohort of 160 screened CN-AML patients was highly significant ($P < .001$; χ^2 test).

High-resolution melting curve analysis of all exons of *IKZF1* detected only 1 additional patient with a nonsynonymous substitution in exon 4 (patient 21, supplemental Table 2). Unfortunately, no remission sample from this patient was available. No *IKZF1* sequence variations were identified in 38 monoallelic *CEBPA* mutation patients. Whole gene and intragenic deletions of *IKZF1* were ruled out by performing MLPA in 30 of 33 bi*CEBPA* patient samples. We detected an additional mutation in *DNMT3A* (Table 1), resulting in an overall *DNMT3A* mutation frequency of 9.1% (3/33) in bi*CEBPA* patients. The frequency distribution of additional mutations in bi*CEBPA* mutant AML is shown in Figure 1C. Interestingly, *FLT3*-ITD (6/33) was mutually exclusive with GATA2 mutations in bi*CEBPA* mutant AML ($P = .029$, Fisher exact test).

GATA2 ZF1 mutants show reduced capacity to enhance CEBPA-dependent activation of transcription

To study the functional consequences of the GATA2 ZF1 mutations, we performed reporter gene assays in HEK 293T cells. Western blotting confirmed that the GATA2 mutants were expressed at similar levels (Figure 2A). All mutants showed nuclear localization like the wild-type GATA2 protein (Figure 2B and supplemental Figure 3). On cotransfection of a GATA-responsive reporter together with expression plasmids for GATA2 wild-type or GATA2 mutants, a reduced activation was observed for mutants Ala318Thr and Leu321Phe. In contrast, an increased activation was observed for the Gly320Asp mutant (Figure 3A). As a control, we included the Leu359Val mutant of GATA2, located in ZF2, which is found in

~ 10% of patients at the progression of CML to blast crisis.²¹ The L359V mutant did not show a significant difference in transcriptional activation compared with wild-type GATA2. GATA2 and CEBPA are protein-protein interactors.⁴⁵ In coimmunoprecipitation experiments, the GATA2 mutants Ala318Thr, Gly320Asp, Leu321Phe, and Leu359Val were still able to interact with wild-type CEBPA (supplemental Figure 5). To further analyze the role of GATA2 mutations in the context of this interaction, we used a CEBPA-responsive luciferase reporter (4xCAAT-binding-site-TK-LUC) together with expression plasmids for wild-type CEBPA or the truncated CEBPA p30 mutant and GATA2 wild-type or GATA2 mutants. CEBPA-dependent activation was enhanced by coexpression of GATA2 wild type. Interestingly, this enhancement was significantly reduced for all the GATA2 ZF1 mutants identified in bi*CEBPA* CN-AML but not for the CML-associated Leu359Val mutant (Figure 3B). Residual activation of the reporter by the p30 CEBPA mutant also could be enhanced by coexpression of wild-type GATA2, but all GATA2 mutants tested showed a markedly reduced enhancement (Figure 3B).

GATA2 ZF1 mutations and clinical characteristics

There were no significant differences in the clinical parameters (eg, white blood cells, age, sex) between bi*CEBPA*-mutated patients with ($n = 13$) and without ($n = 20$) mutations in GATA2 (supplemental Table 3).

In Kaplan-Meier survival analysis, the presence of GATA2 mutations ($n = 10$) did not negatively impact on the favorable OS and EFS of 26 bi*CEBPA* patients (OS, $P = .274$; EFS, $P = .185$; Figure 4).

In a 3-group comparison, there was a trend toward a better OS ($P = .012$) and EFS ($P = .172$) for bi*CEBPA* patients with GATA2 mutations ($n = 10$) compared with bi*CEBPA* patients without a

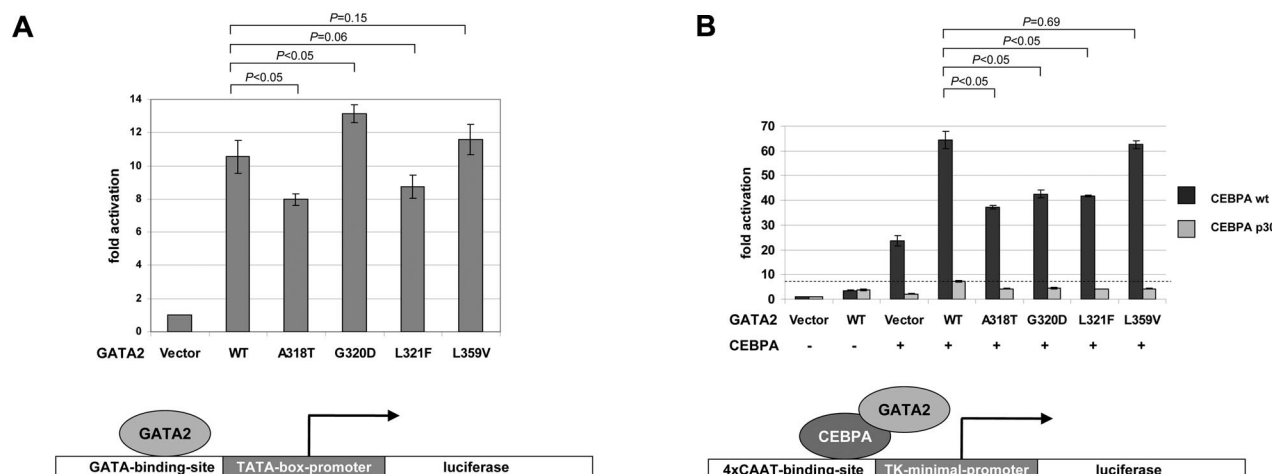


Figure 3. Analysis of GATA2 mutants by transcription assays in HEK 293T cells. Bars indicate fold activation of the luciferase reporters in at least 3 independent experiments, each performed in triplicates. Errors bars represent standard error of the mean. *P* values are indicated for pairwise comparisons using Student *t* test. (A) A GATA2-responsive reporter (GATA-LUC) containing GATA-binding elements derived from the TCR δ enhancer was cotransfected with expression plasmids for either GATA2 wild-type or GATA2 mutants. A reduced activation was observed for mutants A318T and L321F compared with wild-type GATA2. In contrast, an increased activation was observed for the G320D mutant. The L359V mutant of GATA2, located in ZF2, which is found in ~ 10% of patients at the progression of CML to AML,²¹ did not show a significant difference in transcriptional activation compared with wild-type GATA2. (B) A CEBPA-responsive reporter (4xCAAT-BS-TK-LUC) containing CEBPA-binding elements was cotransfected with expression plasmids for wild-type CEBPA or a p30 CEBPA truncation mutant in combination with either GATA2 wild type or GATA2 mutants. Enhancement of CEBPA-dependent transcriptional activation by wild-type GATA2 was significantly reduced for all the ZF1-mutants but not for L359V ZF2 mutant. Residual activation of the reporter by the p30 CEBPA truncation mutant was enhanced by coexpression of wild-type GATA2 (dashed line), but this enhancement was markedly reduced for all of the GATA2 mutants tested.

GATA2 mutation (*n* = 16) and 117 patients with wild-type GATA2 status that included 89 CN-AML patients with wild-type CEBPA and 28 patients with monoallelic CEBPA mutations (Figure 4).

Structural modeling of the GATA2 ZF1 mutants

To gain further insight into the consequences of the GATA ZF1 mutations on the protein level, we performed structural modeling. Based on the homology model (Figure 5), amino acids asparagine 317, alanine 318, lysine 321, and arginine 330 are directly implicated in DNA binding and mutations in these residues likely alter DNA affinity. Glycine 320 is important for proper attachment of an adjacent β -hairpin loop that provides additional DNA binding contacts. Glutamine 328 is not directly involved in DNA binding. However, the Gln328Pro mutation could either affect the backbone fold of ZF1 and indirectly alter DNA binding by perturbing the adjacent Arg330, or alternatively be involved in interactions with other domains such as ZF2.

CEBPA target genes are enriched in biCEBPA-mutated patients with concurrent GATA2 mutations

To assess the impact of GATA2 mutations on gene expression, we performed microarray analysis. Nine biCEBPA-mutated CN-AML patient samples were analyzed. Samples were selected based on availability. Four patients had an additional GATA2 mutation (44.4%). No patient had an additional NPM1 mutation or an MLL-PTD. One patient without GATA2 mutation had an additional FLT3-ITD (11.1%). Gene set enrichment analysis was performed to assess significant changes in gene expression levels. We were able to identify 18 pathways significantly enriched in the GATA2-mutated subgroup (false discovery rate < 25% and nominal *P* value < .05). These belonged, for example, to the cell cycle, ERBB, MTOR, p53 signaling, and apoptosis pathway (supplemental Table 4). We then focused on validated CEBPA and GATA2 target genes (supplemental Table 5). Very interestingly, CEBPA target genes were significantly enriched in the GATA2-mutated subgroup (nominal *P* value = .003; false discovery rate = 1.7%;

Normalized enrichment score = 1.62; supplemental Figure 4). Genes enriched in the GATA2-mutated subgroup were, for example, MYC, JUNB, FOS, and CEBPB (supplemental Table 6; supplemental Figure 4). GATA2 target genes showed no significant enrichment.

Discussion

Ours is the first report of recurrent GATA2 mutations associated exclusively with biCEBPA mutations in AML.

The specific occurrence of GATA2 mutations in the biCEBPA mutant AML subgroup, which accounts for ~ 4% of AML, might explain why GATA2 mutations remained undetected in previous full-length GATA2 mutation screens of unselected AML patients.^{22,46} Another study limited GATA2 mutational screening to specific GATA2 mutations, for example, Leu359Val in GATA2 ZF2,²¹ and thus would have missed GATA2 ZF1 mutations. GATA2 ZF1 and -2 mutations were recently detected at a low frequency of 3.6% (4/112) in the M5 subtype of AML.²³ However, the CEBPA status was not reported in these studies. We found that biCEBPA mutations are significantly associated with AML M1 or M2 subtype and rarely occur in the AML FAB M5 subtype.³

Notably, all mutations found in the present study affected exon 4 encoding the N-terminal zinc finger domain (ZF1) of GATA2, whereas recently GATA2 germ line mutations were reported in the C-terminal zinc finger (ZF2) in familial cases of myelodysplastic syndrome and AML.^{22,47-49} The ZF2 domain of GATA2 also is targeted by somatic mutations in blast crisis of CML.^{21,46} Although the GATA2 ZF2 mutations in CML are associated with poor prognosis,²¹ the GATA2 ZF1 mutants found in the present study do not seem to have a negative impact on clinical outcome (Figure 4). Analysis of a larger cohort of biCEBPA-mutated patients will elucidate whether concurrent GATA2 and biCEBPA mutations define a prognostically distinct subgroup of AML.

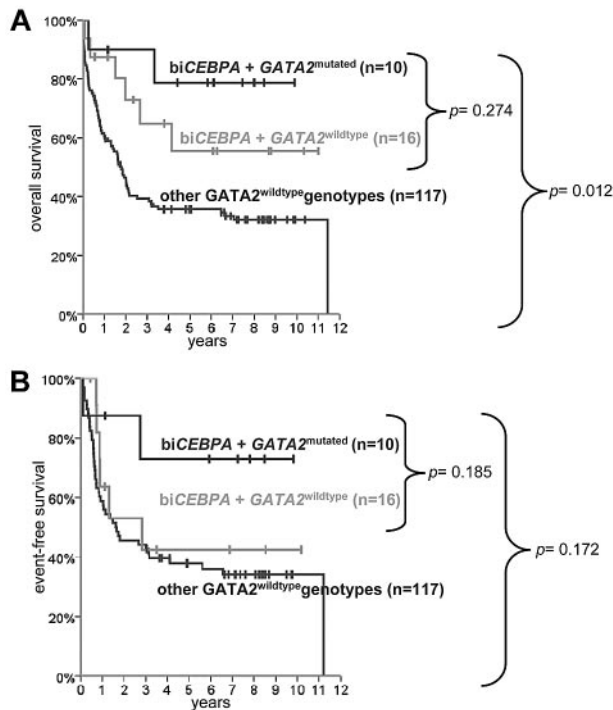
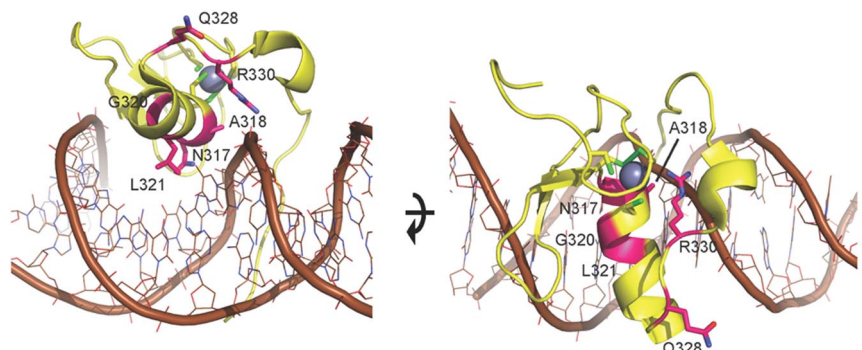


Figure 4. Survival according to GATA2 mutational status in 143 CN-AML patients. OS (A) and EFS (B) within biCEBPA patients according to the GATA2 mutational status and compared with other GATA2 wild-type CN-AML patients (P, log rank test).

The mutual exclusiveness of GATA2 mutations and FLT3-ITD within biCEBPA-mutated patients (Figure 1C) suggests alternative mechanisms of leukemogenesis in these genetic subgroups. A potential pathogenetic relevance of FLT3-ITD in CEBPA mutated AML has been suggested in a bone marrow transplantation model in which a C-terminal CEBPA mutant collaborated with a FLT3-ITD in the process of leukemogenesis.⁸ Considering the direct protein-protein interaction between GATA2 and CEBPA,⁴⁵ and their prominent expression in early hematopoietic progenitor cells,^{9,16,17,50} mutations of both genes may affect the same protein complex in a synergistic manner during hematopoietic differentiation, leading to leukemogenesis. We show that the GATA2 ZF1 mutants found in our patients have a reduced capacity to cooperate with wild-type CEBPA to activate transcription (Figure 3B). This is in contrast to the behavior of the GATA2 ZF2 mutant Leu359Val found in CML blast crisis; this mutant exhibits gain of function properties.²¹

Figure 5. Model of N-terminal zinc finger of GATA2.

Model of ZF1 (yellow ribbon with brown DNA and gray zinc ion), based on the crystal structure of the DNA complex of the highly homologous zinc finger of GATA3³⁸ (PDB accession 3DFV). Mutated residues (magenta) are annotated and displayed with side chains. The mutations cluster at the DNA binding side of ZF1, suggesting they perturb DNA binding. Based on the homology model, N317, A318, L321, and R330 are directly implicated in DNA binding, so mutations in these residues probably alter the affinity to DNA or prevent DNA binding. G320 is important for proper attachment of an adjacent β -hairpin loop that provides additional DNA binding contacts. Q328 is not directly involved in DNA binding. However, Q328P could either affect the backbone fold of ZF1 and indirectly alter DNA binding by perturbing the adjacent R330, or alternatively is involved in interaction with other domains such as ZF2. All in all, the location of the mutations suggests that they influence DNA binding of ZF1.



The ZF1 domain of GATA2 is known to contribute to the stabilization and specificity of DNA binding and mediates the interaction with the transcriptional cofactor Friend of GATA2 (FOG1),⁵¹ whereas the interaction with CEBPA was mapped to the C terminus of GATA2.⁴⁵ Thus, the novel GATA2 ZF1 mutants probably influence DNA-binding properties (Figure 5) or protein-protein interactions of GATA2. In coimmunoprecipitation experiments, all GATA2 mutants tested were able to interact with CEBPA (supplemental Figure 5). However, we cannot exclude that the GATA2 mutations may result in subtle differences in the binding affinity between GATA2 and CEBPA that might be very challenging to quantify. The structural model (Figure 5) may provide an explanation for the differences observed in the activation of the GATA-LUC reporter by GATA ZF1 mutants (Figure 3A): the amino acids affected by those mutants that show reduced activation (Ala318Thr and Leu321Phe) seem to be more directly involved in DNA binding, whereas the mutant Gly320Asp showing increased activation might alter the conformation of ZF1 in a way which is favorable for DNA binding or transactivation. Additional experiments are necessary to analyze the impact of GATA2 ZF1 mutants on DNA binding.

The gene expression profiling results have to be interpreted carefully because of the limited sample size. Considering the high degree of homogeneity within the analyzed groups (9 patients with normal karyotype and biCEBPA mutations including 4 patients with additional GATA2 mutations and only 1 patient with an additional FLT3-ITD), the results show distinct differences between the GATA2-mutated and GATA2 wild-type groups. These differences affect many well known oncogenic pathways in AML. Very interestingly, CEBPA target genes showed significant enrichment in gene set enrichment analysis, whereas GATA2 target genes showed no significant enrichment. Interpreting our gene expression profiling analysis in the context of our reporter gene assays, it seems that GATA2 mutants are able to modulate the effect of the mutated CEBPA transcription factor, rather than affecting GATA2 function and GATA2 target genes.

It is tempting to speculate that the ZF1 mutations of GATA2 may further reduce the residual activity of CEBPA p30 (Figure 3B) that results from the N-terminal nonsense mutations in biCEBPA AML. This would support the recent hypothesis that a “dosage window” of nuclear factors is critical for the process of hematopoietic differentiation and leukemogenesis. For example, the absence of PU.1 does not lead to the development of leukemia in knockout mice, whereas lowering the levels of PU.1 to 20% of normal levels will result in leukemia.⁵²

In summary, we describe the specific association of biallelic *CEBPA* mutations in cytogenetically normal AML with novel mutations in the N-terminal zinc finger of GATA2. In contrast to the results of most other high-throughput sequencing studies in AML that have painted a picture of increasing genetic complexity, our results suggest that there are indeed striking associations of defined mutations in subgroups of AML. The specific association of mutations affecting 2 interacting regulators of hematopoiesis introduces a novel concept for leukemogenesis: the simultaneous mutational targeting of 2 transcription factors that function in the same differentiation pathway in AML. Studies like ours might help to eventually define the various pathways that are critical to AML leukemogenesis.

Acknowledgments

The authors thank the participating centers of the AMLCG clinical trial and Klaus H. Metzeler for support with the gene expression analysis.

This work was supported by Deutsche Krebshilfe grant 109031 (P.A.G., S.K.B.) and Bundesministerium für Bildung und Forschung (BMBF) grant Nationales Genomforschungsnetzwerk (NGFN) Plus, PKL-01-GS0876-6 (S.K.B.). P.A.G., K.P.H., K.S., and S.K.B. acknowledge support by the Deutsche Forschungsgemeinschaft (SFB 684 Molecular Mechanisms of Normal and Malignant Hematopoiesis).

References

1. Swerdlow, Harris. *WHO classification on Tumours of the Haematopoietic and Lymphoid Tissues*. Geneva, Switzerland: World Health Organization Press; 2008.
2. Döhner H, Estey EH, Amadori S, et al. Diagnosis and management of acute myeloid leukemia in adults: recommendations from an international expert panel, on behalf of the European LeukemiaNet. *Blood*. 2010;115(3):453-474.
3. Dufour A, Schneider F, Metzeler KH, 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(4):570-577.
4. Green CL, Koo KK, Hills RK, Burnett AK, Linch DC, Gale RE. Prognostic significance of *CEBPA* mutations in a large cohort of younger adult patients with acute myeloid leukemia: impact of double *CEBPA* mutations and the interaction with FLT3 and NPM1 mutations. *J Clin Oncol*. 2010;28(16):2739-2747.
5. Pabst T, Eyholzer M, Fos J, Mueller BU. Heterogeneity within AML with *CEBPA* mutations; only *CEBPA* double mutations, but not single *CEBPA* mutations are associated with favourable prognosis. *Br J Cancer*. 2009;100(8):1343-1346.
6. Taskesen E, Bullinger L, Corbacioglu A, et al. Prognostic impact, concurrent genetic mutations and gene expression features of AML with *CEBPA* mutations in a cohort of 1182 cytogenetically normal AML: further evidence for *CEBPA* double mutant AML as a distinctive disease entity. *Blood*. 2011;117(8):2469-2475.
7. Wouters BJ, Lowenberg B, Erpelinck-Verschueren CA, et al. Double *CEBPA* mutations, but not single *CEBPA* mutations, define a subgroup of acute myeloid leukemia with a distinctive gene expression profile that is uniquely associated with a favorable outcome. *Blood*. 2009;113(13):3088-3091.
8. Kato N, Kitaura J, Doki N, et al. Two types of C/EBP[alpha] mutations play distinct but collaborative roles in leukemogenesis: lessons from clinical data and BMT models. *Blood*. 2011;117(1):221-233.
9. Bereshchenko O, Mancini E, Moore S, et al. Hematopoietic stem cell expansion precedes the generation of committed myeloid leukemia-initiating cells in C/EBPalpha mutant AML. *Cancer Cell*. 2009;16(5):390-400.
10. Kirstetter P, Schuster MB, Bereshchenko O, et al. Modeling of C/EBPalpha mutant acute myeloid leukemia reveals a common expression signature of committed myeloid leukemia-initiating cells. *Cancer Cell*. 2008;13(4):299-310.
11. Lin LI, Chen CY, Lin DT, 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*. 2005;11(4):1372-1379.
12. Verhaak RG, Wouters BJ, Erpelinck CA, et al. Prediction of molecular subtypes in acute myeloid leukemia based on gene expression profiling. *Haematologica*. 2009;94(1):131-134.
13. Thol F, Damm F, Ludeking A, et al. Incidence and prognostic influence of DNMT3A mutations in acute myeloid leukemia. *J Clin Oncol*. 2011;29(21):2889-2896.
14. Masuda S. DNMT3A mutations in acute myeloid leukemia: impact on low-risk patients with *CEBPA* mutations. *J Clin Oncol*. 2011;29(34):4592-4593.
15. Marcucci G, Haferlach T, Döhner H. Molecular genetics of adult acute myeloid leukemia: prognostic and therapeutic implications. *J Clin Oncol*. 2011;29(5):475-486.
16. Tsai FY, Keller G, Kuo FC, et al. An early hematopoietic defect in mice lacking the transcription factor GATA-2. *Nature*. 1994;371(6494):221-226.
17. Tsai FY, Orkin SH. 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*. 1997;89(10):3636-3643.
18. Persons DA, Allay JA, Allay ER, et al. Enforced expression of the GATA-2 transcription factor blocks normal hematopoiesis. *Blood*. 1999;93(2):488-499.
19. Rodrigues NP, Boyd AS, Fugazza C, et al. GATA-2 regulates granulocyte-macrophage progenitor cell function. *Blood*. 2008;112(13):4862-4873.
20. Huang Z, Dore LC, Li Z, et al. GATA-2 reinforces megakaryocyte development in the absence of GATA-1. *Mol Cell Biol*. 2009;29(18):5168-5180.
21. Zhang SJ, Ma LY, Huang QH, et al. Gain-of-function mutation of GATA-2 in acute myeloid transformation of chronic myeloid leukemia. *Proc Natl Acad Sci U S A*. 2008;105(6):2076-2081.
22. Hahn CN, Chong CE, Carmichael CL, et al. Heritable GATA2 mutations associated with familial myelodysplastic syndrome and acute myeloid leukemia. *Nat Genet*. 2011;43(10):1012-1017.
23. Yan XJ, Xu J, Gu ZH, et al. Exome sequencing identifies somatic mutations of DNA methyltransferase gene DNMT3A in acute monocytic leukemia. *Nat Genet*. 2011;43(4):309-315.
24. Büchner T, Berdel WE, Schoch C, et al. Double induction containing either two courses or one course of high-dose cytarabine plus mitoxantrone and postremission therapy by either autologous stem-cell transplantation or by prolonged maintenance for acute myeloid leukemia. *J Clin Oncol*. 2006;24(16):2480-2489.
25. Braess J, Spiekermann K, Staib P, et al. Dose-dense induction with sequential high-dose cytarabine and mitoxantrone (S-HAM) and pegfilgrastim results in a high efficacy and a short duration of critical neutropenia in de novo acute myeloid leukemia: a pilot study of the AMLCG. *Blood*. 2009;113(17):3903-3910.
26. Gregory BD, O'Malley RC, Lister R, et al. A link between RNA metabolism and silencing affecting Arabidopsis development. *Dev Cell*. 2008;14(6):854-866.
27. Lister R, O'Malley RC, Tonti-Filippini J, et al.

Authorship

Contribution: P.A.G., A.D., K.S., and S.K.B. conceived and designed the experiments; P.A.G., A.D., N.P.K., B.K., E.Z., B.T., J.S., T. Benthous, M.Y., P.D., and S.K. performed experiments; P.A.G., A.D., and T.H. analyzed data; K.-P.H. performed structural modeling; A.H. and A.G. provided bioinformatics support; H.B. managed the Genome Analyzer IIx platform; A.D., E.Z., T. Benthous, P.M.K., S.S., J.B., and K.S. characterized patient samples; H.B. and M.S. provided administrative support; J.B., M.C.S., W.E.B., T. Büchner, B.J.W., and W.H. coordinated the AMLCG clinical trial; A.D., E.H., and F.S. performed statistical analysis; K.S. and S.K.B. jointly supervised the project; and P.A.G., A.D., K.S., and S.K.B. wrote the manuscript.

Conflict-of-interest disclosure: P.A.G. and S.K. received honoraria from Illumina. The remaining authors declare no competing financial interests.

The current affiliation for P.M.K. is Center for Human Genetics, Philipps Universität, Marburg, Germany. The current affiliation for M.Y. is Hematology-Oncology and Stem Cell Transplantation Research Center, Tehran University of Medical Sciences, Tehran, Iran.

Correspondence: Stefan K. Bohlander, Clinical Cooperative Group Leukemia, Helmholtz Zentrum München, Marchioninistr 25, 81377 Munich, Germany; e-mail: bohlander@helmholtz-muenchen.de.

- Highly integrated single-base resolution maps of the epigenome in Arabidopsis. *Cell*. 2008;133(3):523-536.
28. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*. 2009;25(14):1754-1760.
29. Li H, Handsaker B, Wysoker A, et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics*. 2009;25(16):2078-2079.
30. Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics*. 2010;26(6):841-842.
31. Koboldt DC, Chen K, Wylie T, et al. VarScan: variant detection in massively parallel sequencing of individual and pooled samples. *Bioinformatics*. 2009;25(17):2283-2285.
32. Greif PA, Eck SH, Konstantin NP, et al. Identification of recurring tumor-specific somatic mutations in acute myeloid leukemia by transcriptome sequencing. *Leukemia*. 2011;25(5):821-827.
33. Greif PA, Yaghmaie M, Konstantin NP, et al. Somatic mutations in acute promyelocytic leukemia (APL) identified by exome sequencing. *Leukemia*. 2011;25:1519-1522.
34. McLaren W, Pritchard B, Rios D, Chen Y, Flicek P, Cunningham F. Deriving the consequences of genomic variants with the Ensembl API and SNP Effect Predictor. *Bioinformatics*. 2010;26(16):2069-2070.
35. Robinson JT, Thorvaldsdottir H, Winckler W, et al. Integrative genomics viewer. *Nat Biotechnol*. 2011;29(1):24-26.
36. Ley TJ, Ding L, Walter MJ, et al. DNMT3A mutations in acute myeloid leukemia. *N Engl J Med*. 2010;363(25):2424-2433.
37. Marková J, Michková P, Burčková K, et al. Prognostic impact of DNMT3A mutations in patients with intermediate cytogenetic risk profile acute myeloid leukemia. *Eur J Haematol*. 2012;88(2):128-135.
38. Bates DL, Chen Y, Kim G, Guo L, Chen L. Crystal structures of multiple GATA zinc fingers bound to DNA reveal new insights into DNA recognition and self-association by GATA. *J Mol Biol*. 2008;381(5):1292-1306.
39. Emsley P, Lohkamp B, Scott WG, Cowtan K. Features and development of Coot. *Acta Crystallogr D Biol Crystallogr*. 2010;66(Pt 4):486-501.
40. Cobb BS, Morales-Alcayá S, Kleiger G, Brown KE, Fisher AG, Smale ST. Targeting of Ikaros to pericentromeric heterochromatin by direct DNA binding. *Genes Dev*. 2000;14(17):2146-2160.
41. Mullighan CG, Su X, Zhang J, et al. Deletion of IKZF1 and prognosis in acute lymphoblastic leukemia. *N Engl J Med*. 2009;360(5):470-480.
42. Mullighan CG, Miller CB, Radtke I, et al. BCR-ABL1 lymphoblastic leukaemia is characterized by the deletion of Ikaros. *Nature*. 2008;453(7191):110-114.
43. Rocquain J, Gelsi-Boyer V, Adelaide J, et al. Alteration of cohesin genes in myeloid diseases. *Am J Hematol*. 2010;85(9):717-719.
44. Bowen DT, Frew ME, Hills R, et al. RAS mutation in acute myeloid leukemia is associated with distinct cytogenetic subgroups but does not influence outcome in patients younger than 60 years. *Blood*. 2005;106(6):2113-2119.
45. Tong Q, Tsai J, Tan G, Dalgin G, Hotamisligil GS. Interaction between GATA and the C/EBP family of transcription factors is critical in GATA-mediated suppression of adipocyte differentiation. *Mol Cell Biol*. 2005;25(2):706-715.
46. Zhang SJ, Shi JY, Li JY. GATA-2 L359 V mutation is exclusively associated with CML progression but not other hematological malignancies and GATA-2 P250A is a novel single nucleotide polymorphism. *Leuk Res*. 2009;33(8):1141-1143.
47. Dickinson RE, Griffin H, Bigley V, et al. Exome sequencing identifies GATA-2 mutation as the cause of dendritic cell, monocyte, B and NK lymphoid deficiency. *Blood*. 2011;118(10):2656-2658.
48. Ostergaard P, Simpson MA, Connell FC, et al. Mutations in GATA2 cause primary lymphedema associated with a predisposition to acute myeloid leukemia (Emberger syndrome). *Nat Genet*. 2011;43(10):929-931.
49. Hsu AP, Sampaio EP, Khan J, et al. Mutations in GATA2 are associated with the autosomal dominant and sporadic monocytopenia and mycobacterial infection (MonoMAC) syndrome. *Blood*. 2011;118(10):2653-2655.
50. Wölfler A, Danen-van Oorschot AA, Haanstra JR, et al. Lineage-instructive function of C/EBPalpha in multipotent hematopoietic cells and early thymic progenitors. *Blood*. 2010;116(20):4116-4125.
51. Trainor CD, Ghirlando R, Simpson MA. GATA zinc finger interactions modulate DNA binding and transactivation. *J Biol Chem*. 2000;275(36):28157-28166.
52. Rosenbauer F, Wagner K, Kutok JL, et al. Acute myeloid leukemia induced by graded reduction of a lineage-specific transcription factor, PU. 1. *Nat Genet*. 2004;36(6):624-630.