11. Bacher U, Schnittger S, Haferlach T. Molecular genetics in acute myeloid leukemia. Curr Opin Oncol. 2010;22:646–55.
12. Papaemmanuil E, Cazzola M, Boultwood J, Malcovati L, Vyas P, Bowen D, et al. Somatic SF3B1 mutation in myelodysplasia with ring sideroblasts. New Engl J Med. 2011;365:1384–95.
13. Yoshida K, Sanada M, Shiraishi Y, Nowak D, Nagata Y, Yamamoto R, et al. Frequent pathway mutations of splicing machinery in myelodysplasia. Nature. 2011;478:64–9.
14. Graubert TA, Shen D, Ding L, Okeyo-Owuor T, Lunn CL, Shao J, et al. Recurrent mutations in the U2AF1 splicing factor in myelodysplastic syndromes. Nat Genet. 2012;44:53–7.
15. Visconte V, Makishima H, Jankowska A, Szpurka H, Traina F, Jerez A, et al. SF3B1, a splicing factor is frequently mutated in refractory anemia with ring sideroblasts. Leukemia Off J Leuk Soc Am Leuk Res Fund UK. 2012;26:542–5.
16. Cui R, Gale RP, Xu Z, et al. Clinical importance of SF3B1 mutations in Chinese with myelodysplastic syndromes with ring sideroblasts. Leuk Res. 2012;36:1428–33.
17. Damm F, Kosmider O, Gelsi-Boyer V, Renneville A, Carbuccia N, Hidalgo-Curtis C, et al. Mutations affecting mRNA splicing define distinct clinical phenotypes and correlate with patient outcome in myelodysplastic syndromes. Blood. 2012;119:3211–8.
18. Damm F, Thol F, Kosmider O, Kade S, Loffeld P, Dreyfus F, et al. SF3B1 mutations in myelodysplastic syndromes: clinical associations and prognostic implications. Leukemia Off J Leuk Soc Am Leuk Res Fund UK. 2012;26:1137–40.
19. Jeromin S, Haferlach T, Grossmann V, et al. High frequencies of SF3B1 and JAK2 mutations in refractory anemia with ring sideroblasts associated with marked thrombocytosis strengthen the assignment to the category of myelodysplastic/myeloproliferative neoplasms. Haematologica. 2012. doi:10.3324/haematol.2012.072538.
20. Lasho TL, Finke CM, Hanson CA, Jimma T, Knudson RA, Ketterling RP, et al. SF3B1 mutations in primary myelofibrosis: clinical, histopathology and genetic correlates among 155 patients. Leukemia Off J Leuk Soc Am Leuk Res Fund UK. 2012;26:1135–7.
21. Meggendorfer M, Roller A, Haferlach T, et al. SRSF2 mutations in 275 cases with chronic myelomonocytic leukemia (CMML). Blood. 2012. doi:10.1182/blood-2012-01-404863.
22. Patnaik MM, Lasho TL, Hodnefield JM, Knudson RA, Ketterling RP, Garcia-Manero G, et al. SF3B1 mutations are prevalent in myelodysplastic syndromes with ring sideroblasts but do not hold independent prognostic value. Blood. 2012;119:569–72.
23. Thol F, Kade S, Schlarmann C, Loffeld P, Morgan M, Krauter J, et al. Frequency and prognostic impact of mutations in SRSF2, U2AF1, and ZRSR2 in patients with myelodysplastic syndromes. Blood. 2012.
24. Visconte V, Rogers HJ, Singh J, et al. SF3B1 haploinsufficiency leads to formation of ring sideroblasts in myelodysplastic syndromes. Blood. 2012. doi:10.1182/blood-2012-05-430876.
25. Wu SJ, Kuo YY, Hou HA, et al. The clinical implication of SRSF2 mutation in patients with myelodysplastic syndrome and its stability during disease evolution. Blood. 2012. doi:10.1182/blood-2012-02-412296.
26. Zhang SJ, Rampal R, Manshouri T, Patel J, Mensah N, Kayserian A, et al. Genetic analysis of patients with leukemic transformation of myeloproliferative neoplasms shows recurrent SRSF2 mutations that are associated with adverse outcome. Blood. 2012;119:4480–5.
27. Bejar R, Stevenson KE, Caughey BA, et al. Validation of a prognostic model and the impact of mutations in patients with lower-risk myelodysplastic syndromes. J Clin Oncol Off J Am Soc Clin Oncol. 2012;30(27):3376–82.
28. Hirabayashi S, Flotho C, Moetter J, et al. Spliceosomal gene aberrations are rare, coexist with oncogenic mutations, and are unlikely to exert a driver effect in childhood MDS and JMML. Blood. 2012;119(11):e96–99.
29. Takita J, Yoshida K, Sanada M, et al. Novel splicing-factor mutations in juvenile myelomonocytic leukemia. Leukemia Off J Leuk Soc Am Leuk Res Fund UK. 2012;26(8):1879–81.
30. Wahl MC, Will CL, Luhrmann R. The spliceosome: design principles of a dynamic RNP machine. Cell. 2009;136:701–18.
31. Chen M, Manley JL. Mechanisms of alternative splicing regulation: insights from molecular and genomics approaches. Nat Rev Mol Cell Biol. 2009;10:741–54.
32. Tronchere H, Wang J, Fu XD. A protein related to splicing factor U2AF35 that interacts with U2AF65 and SR proteins in splicing of pre-mRNA. Nature. 1997;388:397–400.
33. Malcovati L, Papaemmanuil E, Bowen DT, Boultwood J, Della Porta MG, Pascutto C, et al. Clinical significance of SF3B1 mutations in myelodysplastic syndromes and myelodysplastic/myeloproliferative neoplasms. Blood. 2011;118:6239–46.
34. Rossi D, Bruscaggin A, Spina V, Rasi S, Khiabanian H, Messina M, et al. Mutations of the SF3B1 splicing factor in chronic lymphocytic leukemia: association with progression and fludarabine-refractoriness. Blood. 2011;118:6904–8.
35. Wang L, Lawrence MS, Wan Y, Stojanov P, Sougnez C, Stevenson K, et al. SF3B1 and other novel cancer genes in chronic lymphocytic leukemia. New Engl J Med. 2011;365:2497–506.
36. Damm F, Nguyen-Khac F, Fontenay M, Bernard OA. Spliceosome and other novel mutations in chronic lymphocytic leukemia, and myeloid malignancies. Leukemia Off J Leuk Soc Am Leuk Res Fund UK. 2012;26(9):2027–31.
37. Quesada V, Conde L, Villamor N, Ordonez GR, Jares P, Bassaganyas L, et al. Exome sequencing identifies recurrent mutations of the splicing factor SF3B1 gene in chronic lymphocytic leukemia. Nat Genet. 2012;44:47–52.
38. Rossi D, Rasi S, Spina V, et al. Different impact of NOTCH1 and SF3B1 mutations on the risk of chronic lymphocytic leukemia transformation to Richter syndrome. Br J Haematol. 2012;158(3):426–29.
39. Ellis MJ, Ding L, Shen D, Luo J, Suman VJ, Wallis JW, et al. Whole-genome analysis informs breast cancer response to aromatase inhibition. Nature. 2012;486:353–60.
40. Isono K, Mizutani-Koseki Y, Komori T, Schmidt-Zachmann MS, Koseki H. Mammalian polycomb-mediated repression of Hox genes requires the essential spliceosomal protein Sf3b1. Genes Dev. 2005;19:536–41.
41. Xiao R, Sun Y, Ding JH, Lin S, Rose DW, Rosenfeld MG, et al. Splicing regulator SC35 is essential for genomic stability and cell proliferation during mammalian organogenesis. Mol Cell Biol. 2007;27:5393–402.
42. Makishima H, Visconte V, Sakaguchi H, et al. Mutations in the spliceosome machinery, a novel and ubiquitous pathway in leukemogenesis. Blood. 2012;119(14):3203–10.
43. Tanaka N, Ishihara M, Kitagawa M, Harada H, Kimura T, Matsuyama T, et al. Cellular commitment to oncogene-induced transformation or apoptosis is dependent on the transcription factor IRF-1. Cell. 1994;77:829–39.
44. Serrano M, Lee H, Chin L, Cordon-Cardo C, Beach D, DePinho RA. Role of the INK4a locus in tumor suppression and cell mortality. Cell. 1996;85:27–37.
45. Kaida D, Motoyoshi H, Tashiro E, Nojima T, Hagiwara M, Ishigami K, et al. Spliceostatin A targets SF3b and inhibits both splicing and nuclear retention of pre-mRNA. Nat Chem Biol. 2007;3:576–83.
46. Kotake Y, Sagane K, Owa T, Mimori-Kiyosue Y, Shimizu H, Uesugi M, et al. Splicing factor SF3b as a target of the antitumor natural product pladienolide. Nat Chem Biol. 2007;3:570–5.
47. Webb TR, Joyner AS, Potter PM. The development and application of small molecule modulators of SF3b as therapeutic agents for cancer. Drug Discov Today. 2012 [Epub ahead of print].

SRSF2 mutations in 275 cases with chronic myelomonocytic leukemia (CMML)

Manja Meggendorfer,[1] Andreas Roller,[1] Torsten Haferlach,[1] Christiane Eder,[1] Frank Dicker,[1] Vera Grossmann,[1] Alexander Kohlmann,[1] Tamara Alpermann,[1] Kenichi Yoshida,[2] Seishi Ogawa,[2] H. Phillip Koeffler,[3,4] Wolfgang Kern,[1] Claudia Haferlach,[1] and Susanne Schnittger[1]

[1]MLL Munich Leukemia Laboratory, Munich, Germany; [2]Cancer Genomics Project, Graduate School of Medicine, University of Tokyo, Tokyo, Japan; [3]UCLA School of Medicine, Cedars-Sinai Medical Center, Los Angeles, CA; [4]National University Cancer Institute Singapore, National University Hospital, Singapore

We analyzed the mutational hotspot region of *SRSF2* (Pro95) in 275 cases with chronic myelomonocytic leukemia (CMML). In addition, *ASXL1*, *CBL*, *EZH2*, *JAK2*V617F, *KRAS*, *NRAS*, *RUNX1*, and *TET2* mutations were investigated in subcohorts. Mutations in *SRSF2* (*SRSF2*mut) were detected in 47% (129 of 275) of all cases. In detail, 120 cases had a missense mutation at Pro95, leading to a change to Pro95His, Pro95Leu, Pro95Arg, Pro95Ala, or Pro95Thr. In 9 cases, 3 new in/del mutations were observed: 7 cases with a 24-bp deletion, 1 case with a 3-bp duplication, and 1 case with a 24-bp duplication. In silico analyses predicted a damaging character for the protein structure of SRSF2 for all mutations. *SRSF2*mut was correlated with higher age, less pronounced anemia, and normal karyotype. *SRSF2*mut and *EZH2*mut were mutually exclusive, but *SRSF2*mut was associated with *TET2*mut. In the total cohort, no effect of *SRSF2*mut on survival was observed. However, in the *RUNX1*mut subcohort, *SRSF2* Pro95His had a favorable effect on overall survival. This comprehensive mutation analysis found that 93% of all patients with CMML carried at least 1 somatic mutation in 9 recurrently mutated genes. In conclusion, these data show the importance of *SRSF2*mut as new diagnostic marker in CMML. (*Blood*. 2012;120(15):3080-3088)

Introduction

Chronic myelomonocytic leukemia (CMML) is a clonal hematopoietic malignancy that can be characterized by features of both a myelodysplastic syndrome (MDS) and a myeloproliferative neoplasm (MPN). Therefore, the World Health Organization classification of 2008 assigned CMML to the mixed category MDS/MPN.[1] A further characteristic feature is the wide heterogeneity of clinical presentations and course, leading to variable prognosis. Beside cytologic criteria for diagnosis, the only genetic criterion, until recently, was the absence of the *BCR-ABL1* fusion transcript. The number of blasts in the peripheral blood (PB) and bone marrow (BM) is a prognostic factor dividing CMML cases into 2 morphologic categories: CMML-1 with fewer than 5% blasts in PB or 10% in BM, and CMML-2 with 5%-19% blasts in PB or 10%-19% in BM.[1] Median overall survival (OS) is approximately 20 months in CMML-1 and 15 months in CMML-2, but wide variations exist.[2] In approximately 15%-30% of patients with CMML, the disease evolves into acute myeloid leukemia (AML).[1,2] On the basis of patient characteristics of 213 patients Onida et al defined a scoring system for CMML, named M.D. Anderson (MDA) prognostic score, stratifying patients with CMML in the 4 subgroups: low, intermediate-1, intermediate-2, and high risk. The level of risk is defined by 4 scores assigned by the following variables: hemoglobin levels below 12 g/dL, lymphocyte count higher than 2.5×10^9/L, presence of circulating immature myeloid cells, and bone marrow blasts 10% or more.[3]

Most patients show a normal karyotype in the CMML cells, and only 20%-40% show clonal cytogenetic abnormalities.[1] Such and coworkers investigated 414 patients with CMML to evaluate the prognostic effect of cytogenetic abnormalities and identified 3 risk categories.[4] A normal karyotype or loss of the Y-chromosome as a sole abnormality represent the low-risk group; trisomy 8, abnormalities of chromosome 7, or a complex karyotype (defined as 3 or more abnormalities) were related to the high-risk group. All other abnormalities were assigned to the intermediate-risk category.

In contrast to cytogenetic aberrations, several molecular gene mutations recently have been found to be frequent in CMML (resulting in overall mutation frequencies of $> 55\%$[5-8]); but, unfortunately, none of these alterations is specific for CMML. Gene mutations identified in CMML cases affect different cellular targets and processes, such as *RUNX1*[9] (transcriptional regulation); isocitrate dehydrogenases *IDH1/2*[10] (metabolism); or *KRAS*, *NRAS*,[11,12] *CBL*,[13] and *JAK2*[14] (tyrosine-signaling pathways). *TET2*,[15] *DNMT3A*,[16] *ASXL1*,[17] *UTX*,[18] and *EZH2*[19] contribute in the broadest sense to epigenetic regulatory mechanisms. All of the cytogenetic changes and molecular mutations have been associated with the pathogenesis of CMML but do not fully explain leukemogenesis.

Thus far, mutations in several of these genes already show prognostic relevance. To date, *EZH2* is the best molecularly analyzed gene in CMML and implies an unfavorable prognosis.[5] Mutation of *ASXL1* correlates with evolution to AML and a shorter OS.[20] The effect of *TET2* mutations remains controversial; in patients with MDS it is associated with a favorable outcome,[1,21] and in CMML different studies found favorable to adverse clinical courses for it.[6,22] Mutations in *RUNX1* clearly correlate with a poor outcome in patients with MDS and patients with AML.[23,24]

Submitted January 16, 2012; accepted August 9, 2012. Prepublished online as *Blood* First Edition paper, August 23, 2012; DOI 10.1182/blood-2012-01-404863.

The online version of this article contains a data supplement.

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.

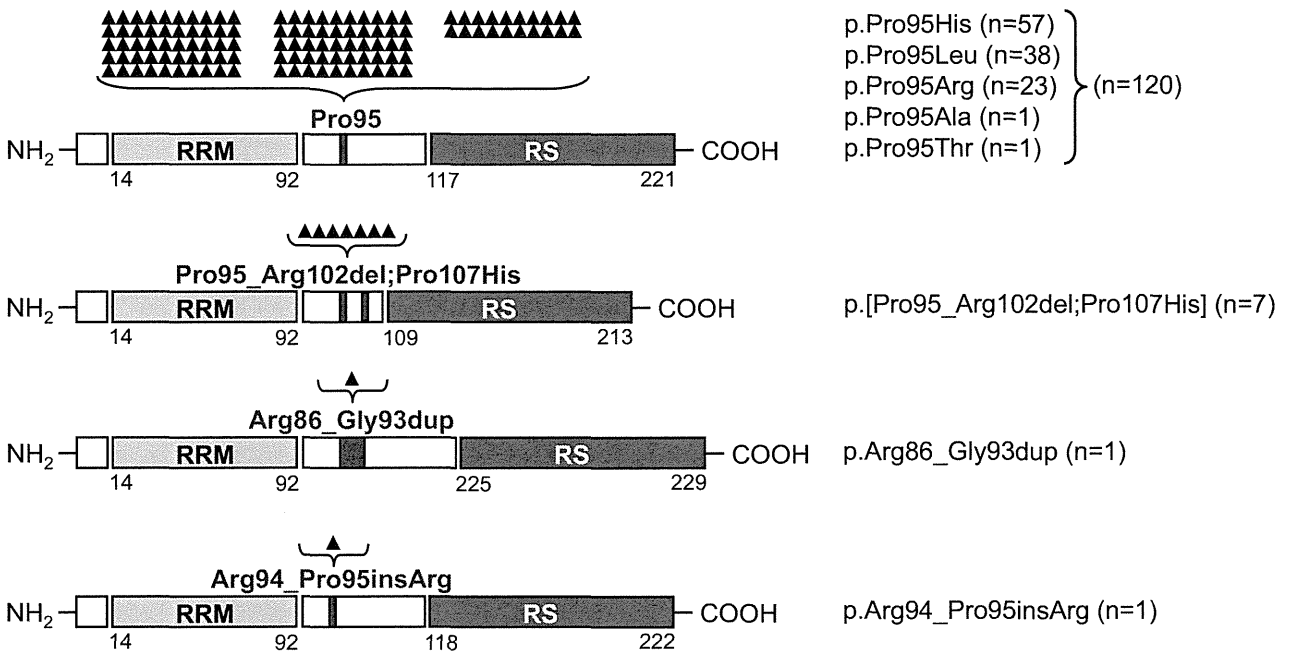


Figure 1. Schematic overview of SRSF2 protein organization, mutation type, and mutation frequency. SRSF2 consists of an RRM (AA 14-92; light gray box), a linker region (white box), and an RS-rich domain (AA 117-221; dark gray box). The 4 different mutation types and their mutation localization are indicated in red. The mutation frequency is listed and also shown as black triangles above each mutation type. Each triangle is representing 1 mutated case. From top to bottom, Pro95 missense mutations, p.Pro95_Arg102del;Pro107His, p.Arg86_Gly93dup, and p.Arg94_Pro95insArg alterations are depicted.

Previously, we have investigated 81 CMML cases and analyzed the mutation frequency of a number of genes that were found to be recurrently mutated in CMML. These comprehensive studies resulted in an overall mutation frequency of 82%,[5,6] indicating that there is a certain percentage of patients with unknown molecular alterations.

More recently, an additional cellular process was found to be altered in MDS. A whole-exome sequencing approach of 29 MDS specimens and their normal controls detected mutations in several components of the splicing machinery (ie, spliceosome; such as *SF3B1* and *U2AF1*), mostly involved in 3′-splice site recognition. In this context a new candidate gene, *SRSF2* (serine/arginine-rich splicing factor 2, also known as *SC35*, a classic member of the SR-protein family), was identified in close cooperation with our laboratory.[7] Members of the SR-protein family function in constitutive and alternative splicing. They contain a RNA recognition motif (RRM) for binding to RNA and a arginine/serine-rich (RS) domain for interaction with other SR-proteins (Figure 1). As a component of the spliceosome, SRSF2 binds to exonic splicing enhancers, preventing exon skipping and ensuring the correct linear order of exons in spliced mRNA.[25,26] In our recent study, mutations within the *SRSF2* sequence occurred exclusively at position 95 (Pro95), located in a linker sequence between the 2 functional RRM and RS domains. *SRSF2* was found to be most frequently mutated in CMML (28%), less frequently in MDS without increased ring sideroblasts (12%), and to some extent in refractory anemia with ring sideroblasts (6%) and AML/MDS (7%). It was rarely seen to be mutated in MPN (2%) or de novo AML (1%).[7]

To characterize further the genetic defects of CMML, we analyzed the frequency of *SRSF2* mutations, their coincidence with other mutations, and their prognostic relevance in a large cohort of 275 cases.

Methods

Patient cohort

In total, 275 cases with CMML were analyzed. All cases were validated on peripheral blood and/or bone marrow smears according to standards of the World Health Organization[1] and included in all cases May-Grünwald-Giemsa staining, as well as myeloperoxidase, nonspecific esterase, and iron stains.[27] The cohort comprised 189 men and 86 women with a median age of 72.8 years (range, 21.9-93.3 years). Eighty-one patients who have been published previously by our group except for *SRSF2* entered the cohort.[5,6] There is no overlap with the CMML cohort analyzed in Yoshida et al.[7] Cytogenetic analyses were performed after short-term culture. Karyotypes were analyzed after G-banding and were described according to the International System for Human Cytogenetic Nomenclature (1995 guidelines).[28] Further parameters are given in Table 1. All patients gave their consent for genetic analyses and the use of laboratory results for research purposes. The study design adhered to the tenets of the Declaration of Helsinki and was approved by our institutional review board before its initiation.

Sequencing analyses

Isolation of mononuclear cells, DNA and mRNA extraction, and random primed cDNA synthesis were performed as described previously.[29] A 187-bp fragment, containing the mutational hotspot region of *SRSF2* around Pro95, was amplified with the GC-RICH PCR system (Roche Applied Science) from either genomic DNA (n = 201) or cDNA templates (n = 74), using the following primers: SRSF2-for, TTCGCCTT-CGTTCGCTTT; SRSF2-rev, TCCGGCGTCCGTAGCCA. The single amplicon was analyzed by Sanger sequencing in all cases with the use of BigDye Term v1.1 cycle sequencing chemistry (Applied Biosystems). Estimation of the mutational load was based on the electropherograms of the forward and reverse reactions. In addition, in 10 cases the mutational load was confirmed by next-generation sequencing, showing the correlation of both methods (supplemental Figure 2B, available on the *Blood* Web site; see the Supplemental Materials link at the top of the online article). Additional mutational data obtained by Sanger sequencing, next-generation deep amplicon sequencing,[30] or melting curve analyses were available in subcohorts and are described methodically elsewhere: *ASXL1* exon

Table 1. Clinical characteristics, cytogenetics, and molecular mutations of 275 patients with CMML

	Total cohort (n = 275)	*SRSF2*mut (n = 129; 47%)	*SRSF2*wt (n = 146; 53%)	*P*
Clinical characteristics				
Male/female (ratio)	189/86 (2.2)	91/38 (2.4)	98/48 (2.0)	NS
CMML-1, n (%)	193 (70)	91 (47)	102 (53)	NS
CMML-2, n (%)	82 (30)	38 (46)	44 (54)	NS
Median age, y (range)	72.8 (21.9-93.3)	73.6 (49.9-89.5)	71.5 (21.9-93.3)	.011
Median WBC count, $\times 10^3/\mu L$, (range; n = 247)	15.3 (0.9-160.0)	17.4 (2.2-113.2)	12.9 (0.9-160.0)	NS
Median platelet count, $\times 10^3/\mu L$ (range; n = 227)	90.0 (3.0-1385)	80.5 (3.0-1119)	105.0 (5.0-1385)	NS
Median Hb level, g/dL (range; n = 226)	11.0 (4.0-18.2)	11.3 (6.0-15.5)	10.2 (4.0-18.2)	.006
Cytogenetics (n = 269)				
Normal karyotype, n (%)	190 (71)	101 (53)	89 (47)	.001
Aberrant karyotype, n (%)	79 (29)	24 (30)	55 (70)	
Trisomy 8, n (%)	26 (33)	9 (35)	17 (65)	NS
Y-chromosome, n (%)	13 (17)	2 (15)	11 (85)	NS
Chromosome 7 aberration, n (%)	9 (11)	4 (44)	5 (56)	NS
Complex, n (%)	4 (5)	1 (25)	3 (75)	
All other, n (%)	27 (34)	8 (30)	19 (70)	
Molecular mutations				
ASXL1; (n = 261)				NS
Mutated, n (%)	115 (44)	56 (49)	59 (51)	
Wild-type, n (%)	146 (56)	68 (47)	78 (53)	
CBL (n = 274)				NS
Mutated, n (%)	51 (19)	27 (53)	24 (47)	
Wild-type, n (%)	223 (81)	101 (45)	122 (55)	
EZH2 (n = 208)				< .001
Mutated, n (%)	20 (10)	1 (5)	19 (95)	
Wild-type, n (%)	188 (90)	106 (56)	82 (44)	
*JAK2*V617F (n = 275)				NS
Mutated, n (%)	18 (7)	9 (50)	9 (50)	
Wild-type, n (%)	257 (93)	120 (47)	137 (53)	
KRAS (n = 266)				NS
Mutated, n (%)	28 (11)	10 (36)	18 (64)	
Wild-type, n (%)	238 (89)	117 (49)	121 (51)	
NRAS (n = 273)				NS
Mutated, n (%)	43 (16)	17 (40)	26 (60)	
Wild-type, n (%)	230 (84)	111 (48)	119 (52)	
RUNX1 (n = 274)				NS
Mutated, n (%)	61 (22)	34 (56)	27 (44)	
Wild-type, n (%)	213 (78)	94 (44)	119 (56)	
TET2 (n = 160)				.001
Mutated, n (%)	97 (61)	60 (62)	37 (38)	
Wild-type, n (%)	63 (39)	22 (35)	41 (65)	

P values are given for significant differences.
Hb indicates hemoglobin.

12 (n = 261),[31] *CBL* (n = 274),[6,32] *EZH2* (n = 208),[5] *IDH1/2* (n = 82),[5] *JAK2*V617F (n = 275),[33] *KRAS* codons 12/13 and 61 (n = 266),[6] *NRAS* codons 12/13 and 61 (n = 273),[6,34] *RUNX1* (n = 274),[6] and *TET2* (n = 160).[6] The coding sequence of *SF3B1* (n = 171) was analyzed by Sanger sequencing. *U2AF1* Ser34 and Gln157 (n = 265) were analyzed by melting curve analyses.

In silico analyses

For protein structure prediction, we used the Robetta prediction server (http://robetta.bakerlab.org).[35] In first iteration, we applied Robetta to predict models for the known RRM domain (2KN4.pdb) of the SRSF2 wild-type (wt) protein. On the basis of the resulting model, the 3-dimensional full model option was applied to obtain a complete model of SRSF2. Next, we repeated these steps to generate full models for our detected novel mutations. The altered protein sequences were submitted to Robetta, and resulting full models were compared with the SRSF2 wild-type model. For each submitted sequence, we selected the best model based on a manual validation process of the RRM domain. Finally, to analyze the differences between the best resulting models we calculated the Cα-Cα distances.[36] For a more detailed report, see supplemental Methods.

Statistical analyses

Statistical analyses were performed with SPSS version 19.0.0 (SPSS Inc); the reported *P* values are 2-sided.

Survival curves were calculated for OS according to Kaplan-Meier and compared with the 2-sided log-rank test. OS was the time from diagnosis to death or last follow-up. Follow-up data were available in 180 cases, which were included in survival analyses. Results were considered significant at $P < .05$. Adjustment for multiple testing was not done. Dichotomous variables were compared between different groups with the use of the chi-square test and continuous variables by Student *t* test.

Results

Characterization of 275 patients with CMML

According to the classification of the World Health Organization, the 275 patients were categorized as 193 cases of CMML-1 and 82 cases of CMML-2. Morphologic features of monocytes and

monoblasts and erythroid dysplastic changes are given in supplemental Methods. On the basis of biologic parameters 61 patients were categorized to the MDA score,[3] with 10 patients belonging to the low-risk group, 15 to the intermediate-1 group, 25 to the intermediate-2 group, and 11 to the high-risk group.

Cytogenetic analyses were performed in 269 of 275 cases (in 6 cases, no metaphases were available). As typical in CMML a majority of patients had a normal karyotype (71%; 190 of 269), yet 29% (79 of 269) showed an aberrant karyotype. Within the aberrant karyotype group of 79 patients, a loss of the Y-chromosome (n = 13), chromosome 7 aberrations (n = 9), and a trisomy 8 (n = 26) were the most frequent abnormalities (for further parameters, see Table 1). Therefore, 203 cases belong to the low-risk category, whereas 27 belong to the intermediate- and 39 to the high-risk categories, defined by Such et al.[4]

Characterization and frequency of *SRSF2* mutations

To analyze the mutation frequency of *SRSF2* in our CMML cohort of 275 patients (Table 1), we investigated the sequence of an amplicon covering the mutation hotspot codon Pro95. Alterations of Pro95 or adjacent sequences were detected in 47% (129 of 275) of all cases. Mutation frequencies were similar in CMML-1 (47%; 91 of 193) and CMML-2 (46%; 38 of 82). In detail, 119 cases had a missense mutation leading to a change of Pro95 to 1 of the following 5 residues: p.Pro95His (n = 56), p.Pro95Leu (n = 38), p.Pro95Arg (n = 23), p.Pro95Ala (n = 1), and p.Pro95Thr (n = 1). In all cases, an estimated mutation load of 30%-50% in accordance with a heterozygous mutation status was detected. One additional case showed 2 different mutations p.[Arg94Pro;Pro95His] in a subset of 50% each. Next-generation sequencing validated it as a mono-allelic mutation.

Interestingly, beyond the previously described missense mutations leading to alterations of Pro95,[7] 3 new in/del mutations were observed, affecting the immediate neighboring amino acids (AA) of Pro95. In 7 cases a deletion of 24 bp with a start in the codon of Pro95 resulted in deletion of 8 AAs, ranging from Pro95 to Arg102. All of these 7 cases showed an additional missense mutation at Pro107 (p.[Pro95_Arg102del;Pro107His]). Furthermore, 1 single case showed a 24-bp duplication of the AA Arg86 to Gly93 (p.Arg86_Gly93dup), and another sample had a 3-bp duplication that resulted in an insertion of arginine between Arg94 and Pro95 (p.Arg94_Pro95insArg). None of these mutations led to a frameshift. Buccal swab controls of 2 patients, carrying the p.[Pro95_Arg102del;Pro107His] mutation, were *SRSF2* wild-type. Furthermore, 1 patient obtained this mutation during disease course. The National Institutes of Health dbSNP databases[37] as well as the National Heart, Lung, and Blood Institute Exome variant server both report no missense single nucleotide polymorphisms for the analyzed region (AA 86-107), indicating that these novel mutations are somatic mutations and no germline polymorphisms. Figure 1 gives a schematic overview of the protein organization (based on information given by UniProtKB Q01130) and the mutation type, localization, and frequency. *SRSF2*-mutated sequences are shown in supplemental Figure 2A.

In silico analyses

To estimate the damaging character of these specific missense mutations at Pro95, we used SIFT (http://sift.jcvi.org), PolyPhen-2 (http://genetics.bwh.harvard.edu/pph2/index.shtml), and MutationTaster (www.mutationtaster.org) online analysis tools. The straightforward physical and comparative considerations showed predominantly a damaging character for all missense mutations leading to AA exchanges at position Pro95 (see supplemental Methods).

To gain insights into the extent by which other *SRSF2* mutations might alter the protein folding and therefore the protein function, we generated and compared structural models of SRSF2wt and SRSF2mut. A crystal structure of the SRSF2 protein is only available for the RRM domain, and a complete structure for any of the SR proteins has not yet been achieved. To evaluate any altering character of the 3 novel mutation types p.[Pro95_Arg102del; Pro107His], p.Arg86_Gly93dup, and p.Arg94_Pro95insArg on the protein structure, we used the Robetta server (http://robetta.bakerlab.org)[35] to calculate a complete structural model of the wild-type SRSF2 protein and the different mutant SRSF2 proteins (Figure 2A).

The differences between these models were determined by calculating the Cα-Cα distances between the 2 corresponding amino acids of SRSF2wt and mutant SRSF2 for AA 88-99. This area covers the mutation hotspot Pro95 and represents the linker sequence (AA 92-117) and, therefore, reflects the proper folding of the 2 functional domains (RRM and RS) relative to each other. The 3 analyzed novel mutations all found different distances relative to SRSF2wt, summarized in a table in Figure 2B. The 3-bp duplication showed the smallest divergence to the reference model with a distance range of 0.4-6.3 Å. The 24-bp deletion and the 24-bp insertion models show greater differences with distances ranging from 0.2 to 20.1 Å and 0.5 to 22.7 Å, respectively. Because only 1 AA is changed by the missense mutations, the models for the missense mutations show only slight divergences, being congruent with the wt SRSF2 model (Cα-Cα distances and a more detailed report about the whole procedure are given in the supplemental Methods). These data show that all calculated models very well fit the known crystal structure of the RRM domain up to AA 92, and larger changes appear within the mutated linker sequences.

Taken together, the in silico analyses indicate that the linker sequence, particularly AA 95, probably has a relevant effect on protein structure.

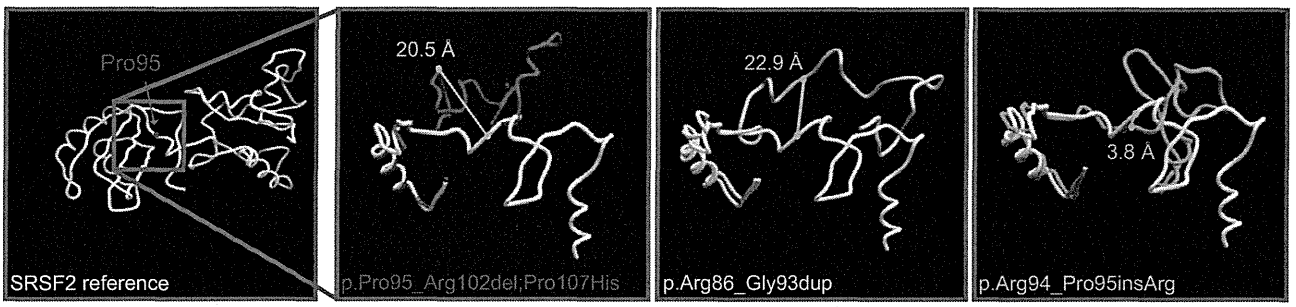
Correlation of *SRSF2* with karyotype

As shown before, the majority of patients had a normal karyotype (71%; 190 of 269), whereas 29% (79 of 269) showed an aberrant karyotype. Within the aberrant karyotype group, the most frequent abnormalities were a loss of the Y-chromosome, aberrations of chromosome 7, and trisomy 8. Therefore, we correlated *SRSF2*mut with both a normal or aberrant karyotype and with subgroups exhibiting the respective chromosomal changes. These analyses found a normal karyotype in 81% of the *SRSF2*mut cases. Stated differently, in the group with a normal karyotype, 53% (101 of 190) had a *SRSF2* mutation, whereas only 30% (24 of 79) were *SRSF2*mut in the aberrant karyotype group ($P = .001$; Table 1; Figure 3). Therefore, *SRSF2*mut correlated significantly with the low-risk group (composed of normal karyotype and loss of the Y-chromosome)[4] compared with the intermediate-risk group (103 of 203, 51% vs 8/27, 30%; $P = .043$). No correlation of *SRSF2*mut was noted for the subcohorts with either loss of the Y-chromosome, chromosome 7 aberrations, or trisomy 8.

Correlation of *SRSF2* with biologic parameters

Mutations in *SRSF2* correlated with higher age (73.6 years vs 71.5 years in the *SRSF2*wt cases; $P = .011$) and higher hemoglobin levels (11.3 vs 10.2 g/dL in the *SRSF2*wt cases; $P = .006$), whereas white blood cell (WBC) and platelet counts were not different. No correlations were observed between cases with *SRSF2*mut and the

A



B

Calculated mutation model	Cα - Cα distance of reference to mutation model for AA 88 to 99 in Å											
	88	89	90	91	92	93	94	95	96	97	98	99
	RRM					linker						
p.Pro95_Arg102del;Pro107His	0.2	0.3	2.3	4.3	7.0	13.8	17.1	20.5	16.7	19.7	20.0	16.3
p.Arg86_Gly93dup	0.7	0.5	0.6	1.2	4.0	7.9	19.6	22.9	19.1	19.6	22.7	22.3
p.Arg94_Pro95insArg	0.4	0.6	2.4	3.6	5.9	6.3	6.3	3.8	3.8	6.6	7.8	3.5

Figure 2. In silico analyses of the structural models of the *SRSF2* in/del mutations. The structural changes of the in/del mutations were calculated with the Robetta server algorithm (http://robetta.bakerlab.org).[35] The calculated model for the complete SRSF2wt protein (white structure) is depicted (A), the mutational hotspot Pro95 is marked in red. In addition, the enlargement shows the structure of AA 61-129 of the calculated models: p.[Pro95_Arg102del;Pro107His]; p.Arg86_Gly93dup; and p.Arg94_Pro95insArg. The Cα-Cα distance measurement of the corresponding AA of SRSF2wt to SRSF2mut is shown exemplarily for Pro95. All Å values of the Cα-Cα distance measurements for AA 88-99 are given (B).

CMML categories 1 and 2 or sex (Table 1). There was also no significant correlation of *SRSF2*mut with other morphologic features (supplemental Methods), any MDA risk category, or proliferative CMML (WBC counts > 13 000/μL) and dysplastic CMML (WBC counts < 13 000/μL).

Coincidence of *SRSF2* with other mutations

We further investigated our CMML cohort for mutations in genes that have been described to be relevant in CMML. *ASXL1*, *CBL*, *EZH2*, *KRAS*, *NRAS*, *IDH1/2*, *JAK2*V617F, *RUNX1*, *SF3B1*, *TET2*, and *U2AF1* were analyzed in large fractions of the 275 cases (Table 1). Comparison of the mutation frequencies of these genes showed that *SRSF2* is the second most frequently mutated gene in this cohort (47%; 129 of 275) after *TET2* (61%; 97 of 160), followed by *ASXL1* (44%; 115 of 261), *RUNX1* (22%; 61 of 274), *CBL* (19%; 51 of 274), *NRAS* (16%; 43 of 273), *KRAS* (11%; 28 of 266), *EZH2* (10%; 20 of 208), and *JAK2* (7%; 18 of 275). The mutation frequencies and associations are shown in Table 1 and Figure 3, respectively. Mutations in *IDH1/2*, *U2AF1*, and *SF3B1* occurred in ≤ 5% of patients and are therefore depicted in supplemental Figure 3.

Analyses of coincidences showed that *SRSF2* mutations were nearly mutually exclusive of *EZH2* mutations. Of the 20 cases with an *EZH2* mutation only 1 had a *SRSF2* mutation. In counterdistinction, in the 208 cases with wt *EZH2*, *SRSF2* was mutated in 106 samples (56%; $P < .001$). In contrast, a high coincidence of *SRSF2* mutations occurred with *TET2* mutations as 62% (60 of 97) of the samples with *TET2*mut had a *SRSF2* mutation; whereas in the *TET2*wt group, only 35% (22 or 63) also carried a mutation in *SRSF2* ($P = .001$). For associations with all the other genes, no specific associations were observed (Figure 3B). In a further analysis the coincidences of *SRSF2*mut with any other gene mutation were analyzed separately for CMML-1 and CMML-2 cases. Both groups reflect the same associations as observed in the total cohort (supplemental Figure 4).

Comprehensive analysis of gene mutations

In a subset of 148 cases of the cohort, the mutational status data of 9 genes were available (*SRSF2*, *ASXL1*, *CBL*, *EZH2*, *JAK2*, *KRAS*, *NRAS*, *RUNX1*, and *TET2*). Overall, 93% (137 of 148) of the samples had at least 1 mutation in any of these genes, whereas only 7% (11 of 148) showed no molecular mutation. Eight of these 11 patients without mutation had a normal karyotype; 3 patients carried an aberrant karyotype. This consequently leads to a combined detection rate of alterations in 140 of 148 patients with CMML (95%) having cytogenetic and/or molecular genetic aberrations. Twelve percent (18 of 148) showed mutations in 1 gene; but in none of these 18 cases did a sole mutation of either *SRSF2* or *RUNX1* occur. Most of the cases had simultaneous mutations in 2 (33%; 49 of 148) or 3 (28%; 42 of 148) genes. In cases with mutations involving 2 genes, 1 of the 2 mutated genes was *SRSF2* in 49% (24 of 49) of the samples. In these cases the mutational load of *SRSF2*mut was equal or beneath the mutational load of the second mutated gene. Four mutations occurred in 22 of 148 cases (15%). In only 5 patients mutations in 5 genes were observed (5 of 148; 3%), 1 patient carried mutations in 7 genes (1 of 148; 1%).

Effect of *SRSF2* mutation on clinical outcome

Follow-up data were available in 180 cases (median follow-up, 12 months; median OS, 29.6 months). This cohort comprised 117 CMML-1 (65%) and 63 CMML-2 (35%) cases, and 93 patients had mutations in *SRSF2* (52%). Calculation of the OS for prognostic relevance of *ASXL1*, *EZH2*, *TET2*, and *RUNX1* mutations in the

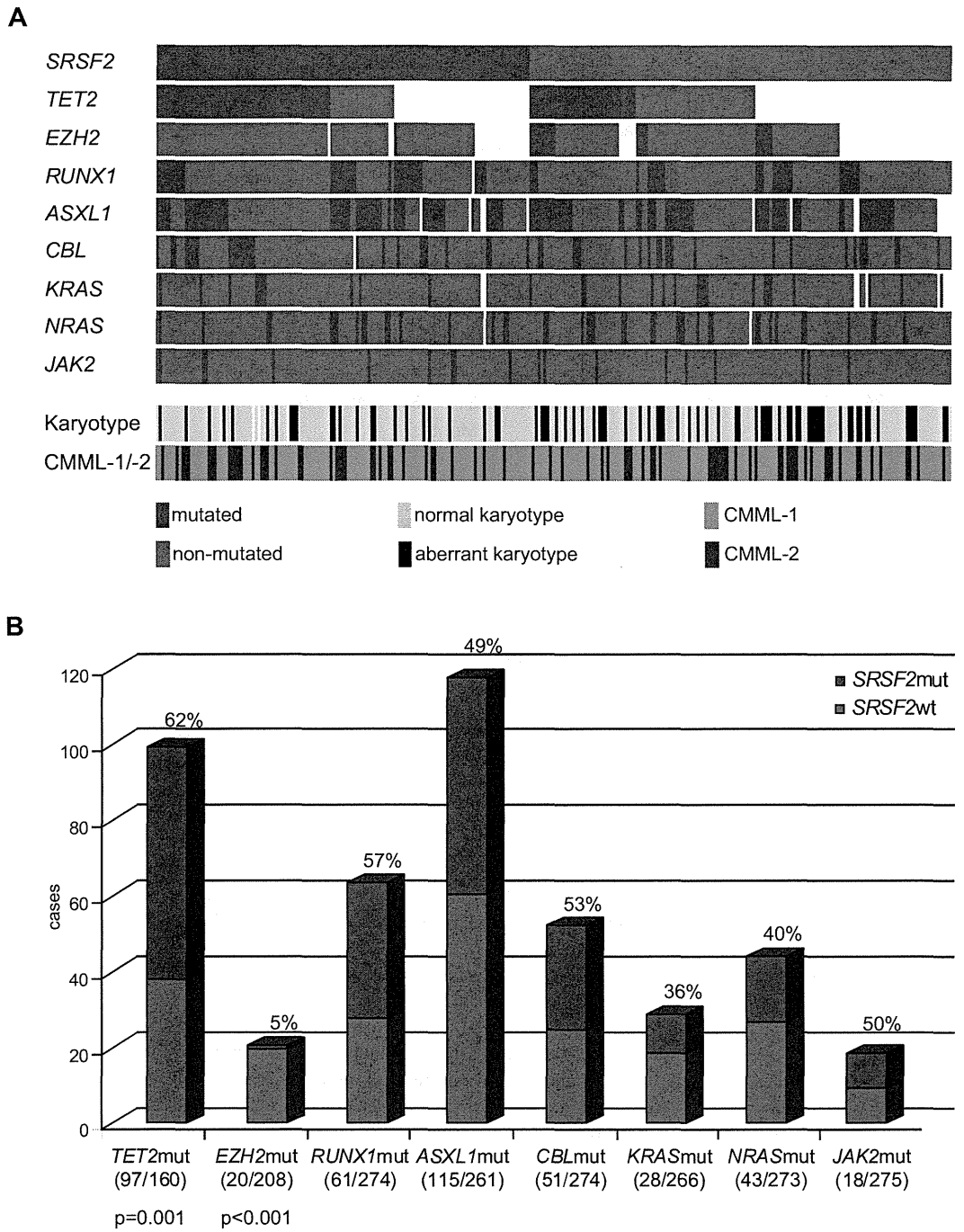


Figure 3. Alignment of gene mutations, karyotype information, and CMML category for 275 patients. (A) Each column represents 1 of the 275 analyzed samples. Analyses of 9 investigated genes, the karyotype, and CMML category-1 or -2 are depicted by colored bars. Red bar indicates mutated gene; dark gray bar, nonmutated gene; white bar, no data available; light-gray bar, normal karyotype; black bar, aberrant karyotype; gray bar, CMML-1; and anthracite bar, CMML-2. (B) Concomitant events of *SRSF2* with other mutations are also shown as a bar chart. The gray part represents *SRSF2*wt, the red one *SRSF2*mut within the analyzed subcohorts. *SRSF2*mut frequencies and significances (*P* values) are denoted; numbers of mutated/analyzed cases of the subcohorts are given in parentheses below the bars.

total CMML cohort found an adverse effect of *ASXL1*mut compared with *ASXL1*wt (median OS, 17.3 months vs not reached; $P = .001$) and a slightly adverse effect of *EZH2*mut relative to *EZH2*wt (median OS, 18.3 vs 29.3 months; $P = .073$). *TET2* and *RUNX1* mutations showed no effect on OS (supplemental Figure 5).

Finally, the influence of *SRSF2* mutation on survival was analyzed. In the total cohort, no effect of *SRSF2* mutations on OS was observed (Figure 4A). Because of the high coincidence of *SRSF2* mutations with *TET2* mutations and the prognostic relevance of *RUNX1* and *ASXL1* alterations in MDS and CMML, respectively, we additionally analyzed these specific subcohorts, resulting in no statistically significant differences. Further, the 3 most frequently appearing missense mutations (Pro95His, Pro95Leu, and Pro95Arg) were analyzed separately. The OS curve of Pro95His-mutated cases showed a slightly better course compared with the wt *SRSF2* cases, whereas the OS of Pro95Leu and Pro95Arg was slightly shorter than of the wt (see supplemental

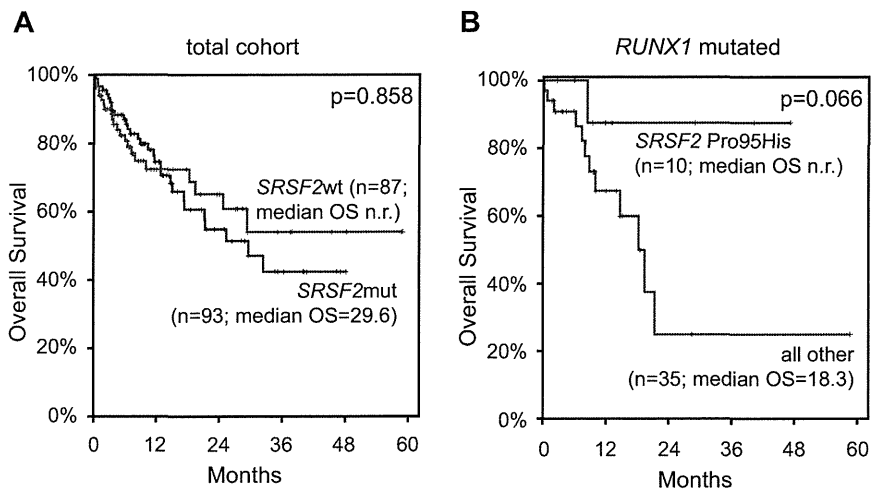


Figure 4. OS by Kaplan-Meier analyses of patients with CMML according to *SRSF2* mutations. (A) OS of patients with *SRSF2*mut did not significantly differ from patients with *SRSF2*wt. (B) OS of patients positive for *SRSF2* Pro95His compared with patients with all other *SRSF2* mutations and *SRSF2*wt (= all other) within the *RUNX1*-mutated subcohort showed a favorable trend. OS is indicated in months and was compared with the 2-sided log-rank test. *P* values are denoted in each graph, respectively.

Figure 6). On the basis of these finding, we calculated the prognostic relevance of Pro95His separately in the above-mentioned subcohorts: *TET2*mut, *RUNX1*mut, and *ASXL1*mut. Pro95His tends to have a favorable effect on OS in the *RUNX1*mut group compared with other *SRSF2*mut or *SRSF2*wt cases (median OS, not reached vs 18.3 months; $P = .066$; Figure 4B). *SRSF2*mut had no influence on OS within any of the cytogenetic risk categories or MDA risk groups.

Discussion

A number of molecular targets have been identified that are frequently mutated in MDS or MDS/MPN. Thereby, some cellular pathways became apparent that are affected by mutations of several genes, including tyrosine kinase signaling and epigenetic regulation.[5,6,8,20] Recently, components of the splicing machinery were found to be frequently mutated in MDS, including mutations in *U2AF1*, *ZRSR2*, *SF3B1*, and *SRSF2*.[7] All of these factors are involved in 3′-splice site recognition of pre-mRNA, inducing abnormal RNA splicing.

In the present study, we analyzed 275 patients with CMML for mutations in *SRSF2* and found a high frequency of mutations (47%). This frequency is even higher than the 28% that is described in the primary publication of Yoshida et al.[7] This difference in frequencies may be caused by ethnic differences of the 2 cohorts, more stringent patient selection and diagnostic procedures (using in all cases nonspecific esterase for calculation of monocytes), or partially by methodologic differences. For example, the next-generation short read sequencing platform that was used in the previous study may have missed the in/del mutations. *SRSF2*, therefore, belongs to the most frequently mutated genes in CMML together with *TET2* and *ASXL1*, with incidences of 61% and 44%, respectively, which is comparable with the frequencies of 44%-50%[6,22] and 49%[20] in previous studies.

Of note, all other results of our mutational screening were in line with published data. *RUNX1* was mutated in 22% of the cases, which is in the range of findings reported by Kohlmann et al,[6] and Gelsi-Boyer et al,[38] with frequencies of 9%, and 30%, respectively. Likewise, *CBL* was mutated in 19% of the cases, and therefore in the range of 13% and 22% reported by Grand et al[32] and Kohlmann et al,[6] respectively. In this article with our enlarged cohort we also confirmed the mutation frequency of *RAS* gene mutations of 30%, also observed by Kohlmann et al[6] (16% for *NRAS* and 11% for *KRAS*). Grossmann et al found a mutation frequency of 11% for *EZH2*,[5] which was confirmed with 10%. Levine et al noted a mutation frequency of 8% for *JAK2*,[39] which is in line with the 7% mutated cases observed in this study. *IDH1/2* showed a mutation frequency of 5%,[5] being in line with 4% presented in Jankowska et al.[8]

The cytogenetic risk stratification suggested by Such et al could not be confirmed in this cohort by Kaplan-Meier analysis,[4] which may be because of small case numbers for the intermediate- (n = 18) and low-risk (n = 27) categories. The median OS for the low- and intermediate-risk groups were not reached, and the median OS for the high-risk group was 21.1 months, but there was no statistically significant difference between the 3 cohorts. This is also true for the MDA risk stratification, whereby the case numbers were even smaller (low, n = 7; intermediate-1, n = 8; intermediate-2, n = 17; high, n = 8). The median OS for the low and intermediate-1 risk groups were not reached and was 11.6 months for the intermediate-2 and 17.3 months for the high-risk groups.

For functional insights of the *SRSF2* mutations various computational analyses were performed. Because a crystal structure of SR proteins is not available, bioinformatic tools were used to predict the character of the missense mutations and to generate SRSF2-structural models that were based on the amino acid sequence. All missense mutations of Pro95 in this study were predicted to be damaging. Recently, Daubner et al analyzed the RNA binding mode of SRSF2 and indicated that Pro95 forms extensive contact with RNA.[40] In addition, the 3 newly described mutations with deletions and insertions are suggestive of being even more deleterious. Comparison of the calculated models indicated that the mutations affected the linker sequence. Therefore, the topography of the 2 domains (RRM and RS) might have changed as a result of an altered number or structure of the amino acid. Considering the fact that no frameshift or nonsense mutations occurred, the protein probably retains both structural integrity and any other modified function.

SRSF2 belongs to the SR protein family and is therefore a splicing factor involved in alternative splicing (reviewed in Long and Caceres[25] and Shepard and Hertel[26]). Alternative splicing is an essential process by which eukaryotes generate high protein diversity from a single gene through the selective joining of different exons. More than 60% of human genes have been estimated to be alternatively spliced,[41] indicating that regulation of

alternative splicing is an important event. Mutations in both the nucleotide sequence of splicing regulatory elements and the components of the cellular splicing machinery can result in aberrant splicing. In addition, aberrant splicing has been found to be associated with various diseases, including cancer.[42,43] Many cancer-related genes are regulated by alternative splicing, and changes in the splicing pattern appear to be unique to the malignant state.[44,45] Daubner et al report that mutations of *SRSF2* affecting the RRM also affect the function, showing a decreased splicing activity of the protein.[40] More recently, Makishima et al showed that *SRSF2*mut leads to defective splicing of the *RUNX1* gene.[46] Moreover, miss-expression of SR proteins changes the alternative splicing pattern and is associated with the development of cancer. Increased expression of SR proteins correlates with cancer progression, as was shown for SRSF2 in ovarian cancers.[47] However, depletion of SRSF2 in the thymus of a mouse model changed the alternative splicing of CD45, causing a defect in T-cell maturation.[48] Lareau et al reported that SRSF2 directs the splicing of its own transcripts and autoregulates its own expression by coupling alternative splicing with RNA decay.[49] Recent reports indicate further functions of *SRSF2* in transcription, promoting RNA Pol II elongation, genome stability, and cell-cycle progression (reviewed in Long and Caceres[25] and Zhong et al[50]). Taken together, mutations in *SRSF2*, although occurring in a region without any obvious functional domain, may cause changes in protein function or expression levels, both possibly contributing to a change of alternative splicing patterns, leading to developmental defects and the onset of cancer.

SRSF2 mutations frequently overlapped with other mutations in our cohort of 275 patients with CMML. Only mutations of *EZH2* did not overlap, pointing to their mutual exclusiveness. One may speculate that this occurs because either no advantageous cooperating effect results from both proteins being altered or concomitant mutations of both proteins is lethal for the cell. Overall, in 18 cases only 1 mutation was detected and this was never *SRSF2*. Thus, *SRSF2* never occurs as a sole mutation, either indicating that *SRSF2* mutations are not early events in the pathogenesis of CMML or that a sole mutation in *SRSF2* results in no clinical manifestation. This is further supported because the mutational load of *SRSF2*mut was always equal or below the mutational load of the second mutated gene in cases with only 2 mutations. By contrast, *SRSF2* is frequently mutated in cases with either 2 or 3 mutations. *SRSF2* mutations may result in a dysfunction of the protein that affects transcriptional elongation and therefore genome stability. Depletion of SRSF2 has been reported to trigger overwhelming double-strand breaks (reviewed in Zhong et al[50]).

Mutations in *SRSF2* were highly associated with *TET2* mutations, a protein converting 5-methyl-cytosin to 5-hydroxymethyl-cytosin. Depletion of *TET2* in bone marrow progenitor cells promotes an expansion of monocyte/macrophage cells,[51] indicating that loss of function can promote clonal expansion of mutant cells. Addressing the WBC count in cases with *TET2*wt + *SRSF2*wt (n = 32) showed a mean of 16 036 cells/μL, whereas *TET2*mut + *SRSF2*wt cases (n = 33) showed 31 112 WBC/μL ($P = .044$). *SRSF2*mut seems to antagonize this leukocytosis, mostly by monocytosis, because the mean WBC count was 16 864/μL in cases with *TET2*mut + *SRSF2*mut (n = 57; *TET2*mut + *SRSF2*mut vs *TET2*wt + *SRSF2*wt; $P = .047$). As mentioned earlier, SRSF2 depletion has been reported to cause genome instability by triggering double-strand breaks, which induced the S phase checkpoint and ended in cell cycle arrest or apoptosis (reviewed in Zhong et al[50]). Furthermore, *SRSF2* mutation is correlated with higher hemoglobin levels. Thus, patients with *SRSF2*mut show a less pronounced leuko/monocytosis in the presence of a concomitant *TET2* mutation and have a less pronounced anemia, both indicating a better state of health.

The median OS of our cohort is 29.6 months, indicating that the outcome of our cohort is somehow better than in other datasets published (eg, Onida et al[3]). This may be because 60% of the patients sent for diagnosis to our institution are referred from outpatient units and hematologist practices at first suspect of CMML and thus are diagnosed very early. Many of our patients were not treated upfront but followed a watch-and-wait-strategy. This may in part explain the differences in the survival curves in comparison with other studies published from centers to which the patients were referred to receive treatment, including enrollment into clinical trials. The mutational status of *SRSF2* did not affect OS, although the median OS was not reached in *SRSF2*wt cases in contrast to 29.6 months in cases with a mutated *SRSF2* ($P = .858$). In *RUNX1*-mutated cases the addition of a *SRSF2* mutation prolonged the OS. Analyzing the most frequently occurring *SRSF2* missense mutations (Pro95His, Pro95Leu, and Pro95Arg) separately indicated that CMML with a Pro95His showed a better outcome than the other 2 frequent mutations as well as the wild-type *SRSF2*, in the *RUNX1*, *TET2*, and *ASXL1* mutated groups. This goes in line with the idea that a *SRSF2* mutation, especially Pro95His, affects protein function; this may result in a favorable effect in cases with concomitant (adverse) mutation, possibly because of inhibition of cell cycle progression.

In summary, *SRSF2* mutations are common in CMML and seem to have a deleterious effect on protein structure and function. This may on the one hand result in promoting further gene mutations and therefore disease progression. On the other hand, it could have a favorable effect on the OS of patients with an additional (adverse) mutation. *SRSF2*mut further correlated with a normal karyotype and confined the cytogenetic categories low and intermediate. *SRSF2*, therefore, represents a novel molecular marker that is helpful for diagnosis of CMML or suspected CMML and for further genetic characterization of this disease. A possible positive prognostic effect in cases with other, partially adverse mutations (*RUNX1*, *TET2*, and *ASXL1*), that was suggested based on our results has to be validated in further independent studies. Of note, based on this data, overall 93% of patients with CMML in the present cohort carried at least 1 mutated gene. However, cases are still found without any detectable genetic defect, warranting further efforts to identify new genetic aberrations that are essential to better understand the molecular pathology of this disease.

Authorship

Contribution: M.M. investigated the molecular mutations of *SRSF2* and *ASXL1*, analyzed the data, and wrote the manuscript; A.R. made the bioinformatic analyses; T.H. was responsible for cytomorphologic analysis and was involved in the collection of clinical data; C.E., F.D., V.G., and A.K. contributed to molecular analyses of the *ASXL1*, *CBL*, *EZH2*, *JAK2*V617F, *KRAS*, *NRAS*, *RUNX1*, and *TET2* mutations; T.A. collected and documented clinical data and compiled statistical analyses; K.Y., S.O., and H.P.K. originally detected *SRSF2* gene mutations and shared unpublished data; W.K. was responsible for immunophenotyping and was involved in statistical analyses; C.H. was responsible for chromosome banding

analysis; S.S. was the principle investigator of the study and wrote the manuscript. All authors read and contributed to the final version of the manuscript.

Conflict-of-interest disclosure: Several of the authors (T.H., W.K., C.H., and S.S.) are part owners of the MLL Munich Leukemia Laboratory. Several of the authors (M.M., A.R., C.E., F.D., A.K., V.G., and T.A.) are employed by the MLL Munich Leukemia Laboratory. K.Y., S.O., and H.P.K. declare no competing financial interests.

Correspondence: Susanne Schnittger, MLL Munich Leukemia Laboratory, Max-Lebsche-Platz 31, 81377 Munich, Germany; e-mail: susanne.schnittger@mll.com.

References

1. Swerdlow SH, Campo E, Harris NL, et al. *WHO Classification of Tumours of Haematopoietic and Lymphoid Tissues*. Lyon, France: International Agency for Research on Cancer (IARC); 2008.
2. Germing U, Kundgen A, Gattermann N. Risk assessment in chronic myelomonocytic leukemia (CMML). *Leuk Lymphoma*. 2004;45(7):1311-1318.
3. Onida F, Kantarjian HM, Smith TL, et al. Prognostic factors and scoring systems in chronic myelomonocytic leukemia: a retrospective analysis of 213 patients. *Blood*. 2002;99(3):840-849.
4. Such E, Cervera J, Costa D, et al. Cytogenetic risk stratification in chronic myelomonocytic leukemia. *Haematologica*. 2011;96(3):375-383.
5. Grossmann V, Kohlmann A, Eder C, et al. Molecular profiling of chronic myelomonocytic leukemia reveals diverse mutations in >80% of patients with TET2 and EZH2 being of high prognostic relevance. *Leukemia*. 2011;25(5):877-879.
6. Kohlmann A, Grossmann V, Klein HU, et al. Next-generation sequencing technology reveals a characteristic pattern of molecular mutations in 72.8% of chronic myelomonocytic leukemia by detecting frequent alterations in TET2, CBL, RAS, and RUNX1. *J Clin Oncol*. 2010;28(24):3858-3865.
7. Yoshida K, Sanada M, Shiraishi Y, et al. Frequent pathway mutations of splicing machinery in myelodysplasia. *Nature*. 2011;478(7367):64-69.
8. Jankowska AM, Makishima H, Tiu RV, et al. Mutational spectrum analysis of chronic myelomonocytic leukemia includes genes associated with epigenetic regulation: UTX, EZH2, and DNMT3A. *Blood*. 2011;118(14):3932-3941.
9. Asou N. The role of a Runt domain transcription factor AML1/RUNX1 in leukemogenesis and its clinical implications. *Crit Rev Oncol Hematol*. 2003;45(2):129-150.
10. Ward PS, Patel J, Wise DR, et al. The common feature of leukemia-associated IDH1 and IDH2 mutations is a neomorphic enzyme activity converting alpha-ketoglutarate to 2-hydroxyglutarate. *Cancer Cell*. 2010;17(3):225-234.
11. Delgado MD, Vaque JP, Arozarena I, et al. H-, K- and N-Ras inhibit myeloid leukemia cell proliferation by a p21WAF1-dependent mechanism. *Oncogene*. 2000;19(6):783-790.
12. Crespo P, Leon J. Ras proteins in the control of the cell cycle and cell differentiation. *Cell Mol Life Sci*. 2000;57(11):1613-1636.
13. Reindl C, Quentmeier H, Petropoulos K, et al. CBL exon 8/9 mutants activate the FLT3 pathway and cluster in core binding factor/11q deletion acute myeloid leukemia/myelodysplastic syndrome subtypes. *Clin Cancer Res*. 2009;15(7):2238-2247.
14. Kralovics R, Passamonti F, Buser AS, et al. A gain-of-function mutation of JAK2 in myeloproliferative disorders. *N Engl J Med*. 2005;352(17):1779-1790.
15. Mullighan CG. TET2 mutations in myelodysplasia and myeloid malignancies. *Nat Genet*. 2009;41(7):766-767.
16. Goll MG, Bestor TH. Eukaryotic cytosine methyltransferases. *Annu Rev Biochem*. 2005;74(481-514.
17. Fisher CL, Randazzo F, Humphries RK, Brock HW. Characterization of Asxl1, a murine homolog of Additional sex combs, and analysis of the Asx-like gene family. *Gene*. 2006;369(109-118.
18. Swigut T, Wysocka J. H3K27 demethylases, at long last. *Cell*. 2007;131(1):29-32.
19. Ernst T, Chase AJ, Score J, et al. Inactivating mutations of the histone methyltransferase gene EZH2 in myeloid disorders. *Nat Genet*. 2010;42(8):722-726.
20. Gelsi-Boyer V, Trouplin V, Roquain J, et al. ASXL1 mutation is associated with poor prognosis and acute transformation in chronic myelomonocytic leukaemia. *Br J Haematol*. 2010;151(4):365-375.
21. Kosmider O, Gelsi-Boyer V, Cheok M, et al. TET2 mutation is an independent favorable prognostic factor in myelodysplastic syndromes (MDSs). *Blood*. 2009;114(15):3285-3291.
22. Kosmider O, Gelsi-Boyer V, Ciudad M, et al. TET2 gene mutation is a frequent and adverse event in chronic myelomonocytic leukemia. *Haematologica*. 2009;94(12):1676-1681.
23. Dicker F, Haferlach C, Sundermann J, et al. Mutation analysis for RUNX1, MLL-PTD, FLT3-ITD, NPM1 and NRAS in 269 patients with MDS or secondary AML. *Leukemia*. 2010;24(8):1528-1532.
24. Schnittger S, Dicker F, Kern W, et al. RUNX1 mutations are frequent in de novo AML with noncomplex karyotype and confer an unfavorable prognosis. *Blood*. 2011;117(8):2348-2357.
25. Long JC, Caceres JF. The SR protein family of splicing factors: master regulators of gene expression. *Biochem J*. 2009;417(1):15-27.
26. Shepard PJ, Hertel KJ. The SR protein family. *Genome Biol*. 2009;10(10):242.
27. Löffler H, Rastetter J, Haferlach T. *Atlas of Clinical Hematology*. Heidelberg, Germany: Springer; 2005.
28. ISCN. In: Mitelman F, ed. *ISCN 1995, Guidelines for Cancer Cytogenetics, Supplement to: An International System for Human Cytogenetic Nomenclature*. Basel, Switzerland: S. Karger; 1995:1-110.
29. Schnittger S, Schoch C, Dugas M, et al. Analysis of FLT3 length mutations in 1003 patients with acute myeloid leukemia: correlation to cytogenetics, FAB subtype, and prognosis in the AMLCG study and usefulness as a marker for the detection of minimal residual disease. *Blood*. 2002;100(1):59-66.
30. Kohlmann A, Klein HU, Weissmann S, et al. The Interlaboratory RObustness of Next-generation sequencing (IRON) study: a deep sequencing investigation of TET2, CBL and KRAS mutations by an international consortium involving 10 laboratories. *Leukemia*. 2011;25(12):1840-1848.
31. Gelsi-Boyer V, Trouplin V, Adelaide J, et al. Mutations of polycomb-associated gene ASXL1 in myelodysplastic syndromes and chronic myelomonocytic leukaemia. *Br J Haematol*. 2009;145(6):788-800.
32. Grand FH, Hidalgo-Curtis CE, Ernst T, et al. Frequent CBL mutations associated with 11q acquired uniparental disomy in myeloproliferative neoplasms. *Blood*. 2009;113(24):6182-6192.
33. Schnittger S, Bacher U, Kern W, et al. Report on two novel nucleotide exchanges in the JAK2 pseudokinase domain: D620E and E627E. *Leukemia*. 2006;20(12):2195-2197.
34. Nakao M, Janssen JW, Seriu T, Bartram CR. Rapid and reliable detection of N-ras mutations in acute lymphoblastic leukemia by melting curve analysis using LightCycler technology. *Leukemia*. 2000;14(2):312-315.
35. Kim DE, Chivian D, Baker D. Protein structure prediction and analysis using the Robetta server. *Nucleic Acids Res*. 2004;32(Web Server issue):W526-W531.
36. Holm L, Sander C. Protein structure comparison by alignment of distance matrices. *J Mol Biol*. 1993;233(1):123-138.
37. Database of Single Nucleotide Polymorphisms (dbSNP). Bethesda (MD): National Center for Biotechnology Information, National Library of Medicine. (dbSNP Build ID: 137). Available from: http://www.ncbi.nlm.nih.gov/SNP/.
38. Gelsi-Boyer V, Trouplin V, Adelaide J, et al. Genome profiling of chronic myelomonocytic leukemia: frequent alterations of RAS and RUNX1 genes. *BMC Cancer*. 2008;8:299.
39. Levine RL, Wadleigh M, Cools J, et al. Activating mutation in the tyrosine kinase JAK2 in polycythemia vera, essential thrombocythemia, and myeloid metaplasia with myelofibrosis. *Cancer Cell*. 2005;7(4):387-397.
40. Daubner GM, Clery A, Jayne S, Stevenin J, Allain FH. A syn-anti conformational difference allows SRSF2 to recognize guanines and cytosines equally well. *EMBO J*. 2012;31(1):162-174.
41. Lander ES, Linton LM, Birren B, et al. Initial sequencing and analysis of the human genome. *Nature*. 2001;409(6822):860-921.
42. Ghigna C, Valacca C, Biamonti G. Alternative splicing and tumor progression. *Curr Genomics*. 2008;9(8):556-570.
43. Grosso AR, Martins S, Carmo-Fonseca M. The emerging role of splicing factors in cancer. *EMBO Rep*. 2008;9(11):1087-1093.
44. Wang Z, Lo HS, Yang H, et al. Computational analysis and experimental validation of tumor-associated alternative RNA splicing in human cancer. *Cancer Res*. 2003;63(3):655-657.
45. Xu Q, Lee C. Discovery of novel splice forms and functional analysis of cancer-specific alternative splicing in human expressed sequences. *Nucleic Acids Res*. 2003;31(19):5635-5643.
46. Makishima H, Visconte V, Sakaguchi H, et al. Mutations in the spliceosome machinery, a novel and ubiquitous pathway in leukemogenesis. *Blood*. 2012;119(14):3203-3210.
47. Fischer DC, Noack K, Runnebaum IB, et al. Expression of splicing factors in human ovarian cancer. *Oncol Rep*. 2004;11(5):1085-1090.
48. Wang HY, Xu X, Ding JH, Bermingham JR Jr, Fu XD. SC35 plays a role in T cell development and alternative splicing of CD45. *Mol Cell*. 2001;7(2):331-342.
49. Lareau LF, Inada M, Green RE, Wengrod JC, Brenner SE. Unproductive splicing of SR genes associated with highly conserved and ultraconserved DNA elements. *Nature*. 2007;446(7138):926-929.
50. Zhong XY, Wang P, Han J, Rosenfeld MG, Fu XD. SR proteins in vertical integration of gene expression from transcription to RNA processing to translation. *Mol Cell*. 2009;35(1):1-10.
51. Ko M, Huang Y, Jankowska AM, et al. Impaired hydroxylation of 5-methylcytosine in myeloid cancers with mutant TET2. *Nature*. 2010;468(7325):839-843.

AUTHORSHIP CONTRIBUTION

LLS, AJW, JH, and Y-YY performed the experiments. LLS, AJJ, and JCB analyzed the results and made the figures. XZ performed the statistical analysis. JF, JJ, MRG, and JCB provided clinical samples. AJJ, RB, MRG and JCB designed the research and wrote the paper.

AJ Johnson[1,2], Y-Y Yeh[1], LL Smith [1], AJ Wagner [1], J Hessler[1], S Gupta[3], J Flynn[1], J Jones[1], X Zhang[4], R Bannerji[5], MR Grever[1] and JC Byrd[1,2]
[1]*Department of Internal Medicine, Division of Hematology, Comprehensive Cancer Center, The Ohio State University, Columbus, OH, USA;*
[2]*Division of Medicinal Chemistry, College of Pharmacy, The Ohio State University, Columbus, OH, USA;*
[3]*Division of Pharmaceutics, College of Pharmacy, The Ohio State University, Columbus, OH, USA;*
[4]*Center for Biostatistics, The Ohio State University, Columbus, OH, USA and*
[5]*Merck and Company, Oncology Clinical Research, Kenilworth, NJ, USA*
Correspondence: E-mail: amy.johnson@osumc.edu

REFERENCES

1 Grever MR, Lucas DM, Johnson AJ, Byrd JC. Novel agents and strategies for treatment of p53-defective chronic lymphocytic leukemia. *Best Pract Res Clin Haematol* 2007; **20**: 545–556.
2 Byrd JC, Shinn C, Waselenko JK, Fuchs EJ, Lehman TA, Nguyen PL *et al.* Flavopiridol induces apoptosis in chronic lymphocytic leukemia cells via activation of caspase-3 without evidence of bcl-2 modulation or dependence on functional p53. *Blood* 1998; **92**: 3804–3816.
3 Konig A, Schwartz GK, Mohammad RM, Al-Katib A, Gabrilove JL. The novel cyclin-dependent kinase inhibitor flavopiridol downregulates Bcl-2 and induces growth arrest and apoptosis in chronic B-cell leukemia lines. *Blood* 1997; **90**: 4307–4312.
4 Kitada S, Zapata JM, Andreeff M, Reed JC. Protein kinase inhibitors flavopiridol and 7-hydroxy-staurosporine down-regulate antiapoptosis proteins in B-cell chronic lymphocytic leukemia. *Blood* 2000; **96**: 393–397.
5 Phelps MA, Lin TS, Johnson AJ, Hurh E, Rozewski DM, Farley KL *et al.* Clinical response and pharmacokinetics from a phase 1 study of an active dosing schedule of flavopiridol in relapsed chronic lymphocytic leukemia. *Blood* 2009; **113**: 2637–2645.
6 Lin TS, Ruppert AS, Johnson AJ, Fischer B, Heerema NA, Andritsos LA *et al.* Phase II study of flavopiridol in relapsed chronic lymphocytic leukemia demonstrating high response rates in genetically high-risk disease. *J Clin Oncol* 2009; **27**: 6012–6018.
7 Paruch K, Dwyer MP, Alvarez C, Brown C, Chan T-Y, Doll RJ *et al.* Discovery of dinaciclib (SCH 727965): a potent and selective inhibitor of cyclin-dependent kinases. *ACS Med Chem Lett* 2010; **1**: 204–208.
8 Parry D, Guzi T, Shanahan F, Davis N, Prabhavalkar D, Wiswell D *et al.* Dinaciclib (SCH 727965), a novel and potent cyclin-dependent kinase inhibitor. *Mol Cancer Ther* 2010; **9**: 2344–2353.
9 Niedermeier M, Hennessy BT, Knight ZA, Henneberg M, Hu J, Kurtova AV *et al.* Isoform-selective phosphoinositide 3′-kinase inhibitors inhibit CXCR4 signaling and overcome stromal cell-mediated drug resistance in chronic lymphocytic leukemia: a novel therapeutic approach. *Blood* 2009; **113**: 5549–5557.
10 Hussain SR, Lucas DM, Johnson AJ, Lin TS, Bakaletz AP, Dang VX *et al.* Flavopiridol causes early mitochondrial damage in chronic lymphocytic leukemia cells with impaired oxygen consumption and mobilization of intracellular calcium. *Blood* 2008; **111**: 3190–3199.
11 Woyach JA, Lozanski G, Ruppert AS, Lozanski A, Blum KA, Jones JA *et al.* Outcome of patients with relapsed or refractory chronic lymphocytic leukemia (CLL) treated with flavopiridol: Impact of Genomic Features. *Leukemia* 2012; **26**: 1442–1444.

Supplementary Information accompanies the paper on the Leukemia website (http://www.nature.com/leu)

EED mutants impair polycomb repressive complex 2 in myelodysplastic syndrome and related neoplasms

Leukemia (2012) **26,** 2557–2560; doi:10.1038/leu.2012.146

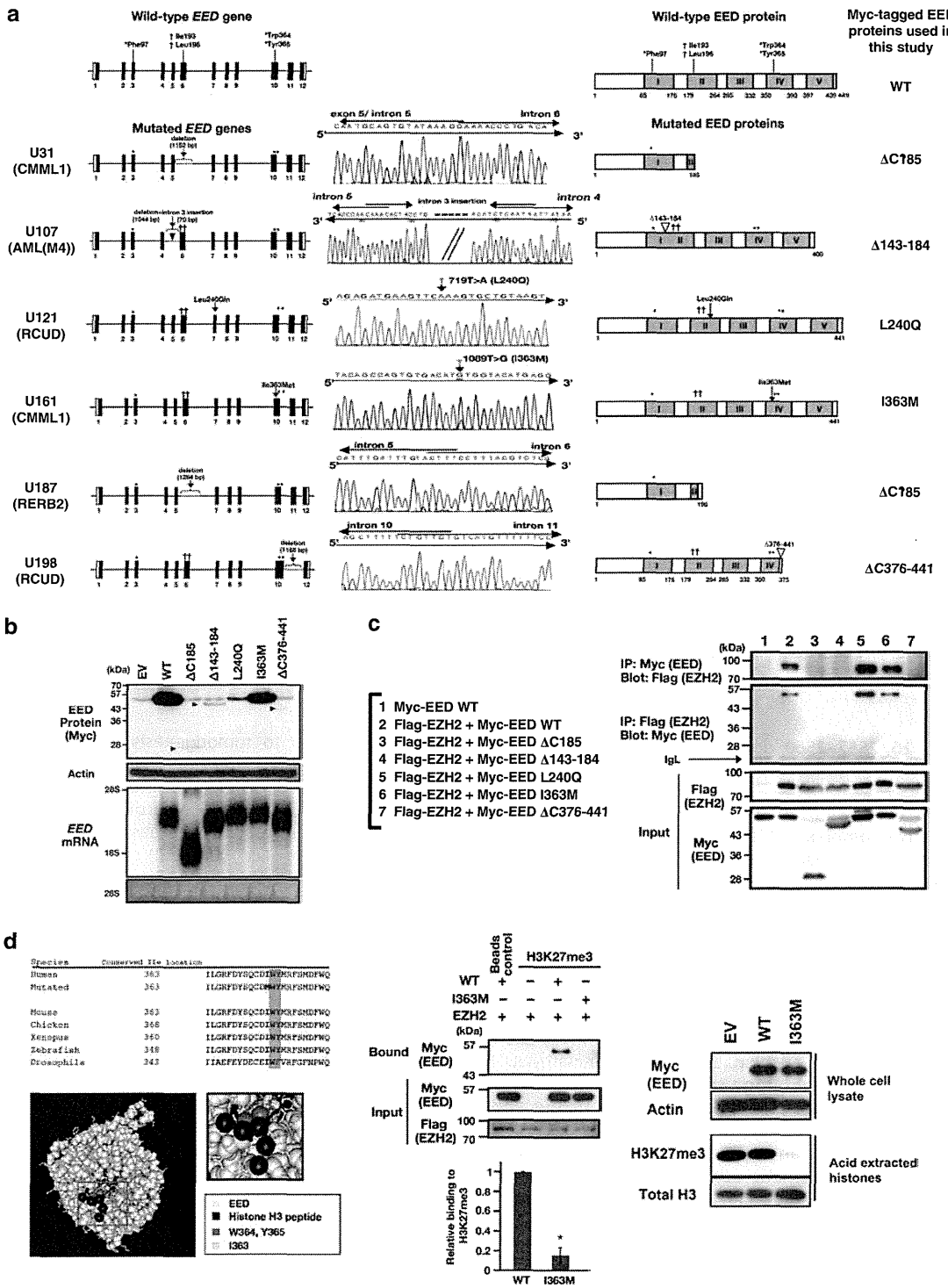
Polycomb recessive complex 2 (PRC2) is an epigenetic regulator that marks repressive chromatin domain through trimethylation of histone H3 lysine 27 (H3K27).[1,2] Recently, inactivating mutations of EZH2, the catalytic subunit of PRC2, have been identified in subsets of myeloid disorders including myelodysplastic syndrome (MDS), and are predicted to inactivate PRC2 function.[3–5] PRC2 comprises four core components, EZH2, EED, SUZ12 and RBBP4 (also known as RbAp48). Although EZH2 possesses the methyltransferase activity of PRC2, EZH2 is inactive on its own, and direct interaction of EED to EZH2 is required for EZH2 to fully exert its enzymatic activity.[1,2,6] In addition, EED binds to trimethylated H3K27 (H3K27me3) through the so-called 'aromatic cage' composed of three aromatic amino acids (Phe97, Trp364 and Tyr365) to activate PRC2.[1,7] A previous report demonstrated that PRC2 complexes possessing an aromatic-cage mutation in EED show severely reduced enzymatic activities in the presence of H3K27me3 peptides.[7] EED haploinsufficiency (Leu196Pro) in mice leads spontaneously to a myeloproliferative disorder,[8] and exposure of hypomorphic (Ile193Asn) homozygotes to genotoxic stresses gives rise to tumorigenesis.[8,9] These findings strongly suggest that dysfunction of EED might be involved in the pathogenesis of myeloid disorders. Here we searched for EED mutations in MDS and related diseases.

The genomic regions encompassing the EED gene (NCBI accession number NM_003797.2) were sequenced in bone marrow samples obtained from 192 patients with MDS or related diseases (Figure 1a). Detailed clinical information of individual samples was given in our previous report.[10] We identified EED mutations in six cases (Figure 1a and Supplementary Table 1), which were confirmed by repeated amplification and sequencing. None of these mutations are reported in the 1000 genomes database (a deep catalog of human genetic variation)[11] and the Ensemble gene and the transcript sequences currently available. We were unable to evaluate whether the mutations arose from germline or somatic tissues or to examine mRNA expression patterns, due to limited sample availability. Of note, three of the six mutations occurred in patients with chronic myelomonocytic leukemia (CMML) or acute myelogenous leukemia possibly preceded by CMML (Figure 1a). This is consistent with the finding that EZH2 abnormalities are most common in CMML and MDS/myelodysplastic-myeloproliferative neoplasms (MDS/MPN).[4] In addition, we observed no compound mutations of the EED and

Accepted article preview online 1 June 2012; advance online publication, 26 June 2012

EZH2 genes (Figure 2a), strongly suggesting that these genes may independently affect PRC2 function.

Two distinct mutations lacking a part of intron 5 and the entire exon 6 were found in individuals with CMML1 and refractory anemia with excess blasts 2 (RAEB2) (U31 and U187, respectively) (Figure 1a). In these patients, exon 5 should be either spliced to exon 7 or followed by intron 5, and in each case consequently creates a new stop codon; therefore the mutations are predicted to produce proteins lacking part of the WD40 II motif and all of the WD40 motifs III–V (hereafter referred to as the ΔC185 mutant) (Figure 1a). A genomic region containing exon 5 was absent in an acute myelogenous leukemia (M4) patient with myelodysplasia-related changes arising from CMML (U107) (Figure 1a). In U107 subject, exon 4 should be followed by exon 6 in frame; therefore, the corresponding protein product (hereafter referred to as the Δ143-184 mutant) lacks part of the first and second



WD40 repeats (Figure 1a). One refractory cytopenia with unilineage dysplasia subject (U121) possessed a missense mutation (T719A; CTA > CAA), producing a Leu363Gln amino-acid substitution (hereafter referred to as the L240Q mutant) (Figure 1a). In addition, another CMML1 subject (U161) possessed a missense mutation (T1089G; ATT > ATG), producing an Ile363Met amino-acid substitution (hereafter referred to as the I363M mutant) (Figure 1a). The exon 11 deletion in another refractory cytopenia with unilineage dysplasia subject (U198) is predicted to produce an EED protein that lacks one and a half of the WD40 repeat motifs located at the C terminus (hereafter referred to as the ΔC376-441 mutant), as exon 10 should be followed by exon 12 with premature stop

a

Patient ID[a]	WHO classification	Age	Sex	IPSS[b]	Cytogenetics	SNP array[a]	*EED* mutation	Predicted EED protein	Expression	Binding EZH2	Binding H3K27me3	*EZH2* status
U31	CMML1	64	F	INT-1	46,XX	+Xp	Δexon 6	ΔC185	low	-	ND	WT
U107	AML(M4)	50	M	NA	45,XY,-7	-7	Δexon 5	Δ143-184	low	-	ND	WT
U121	RCUD	32	F	INT-1	46,XX	Normal	T719A	L240Q	low	+	ND	WT
U161	CMML1	35	M	INT-1	46,XY	11qUPD, 18qUPD	T1089G	I363M	normal	+	↓	WT
U187	RAEB2	59	M	HIGH	46,XY,-7,+13	-7,+13	Δexon 6	ΔC185	low	-	ND	WT
U198	RCUD	58	F	INT-1	46,XX	Normal	Δexon 11	ΔC376-441	low	-	ND	WT

b

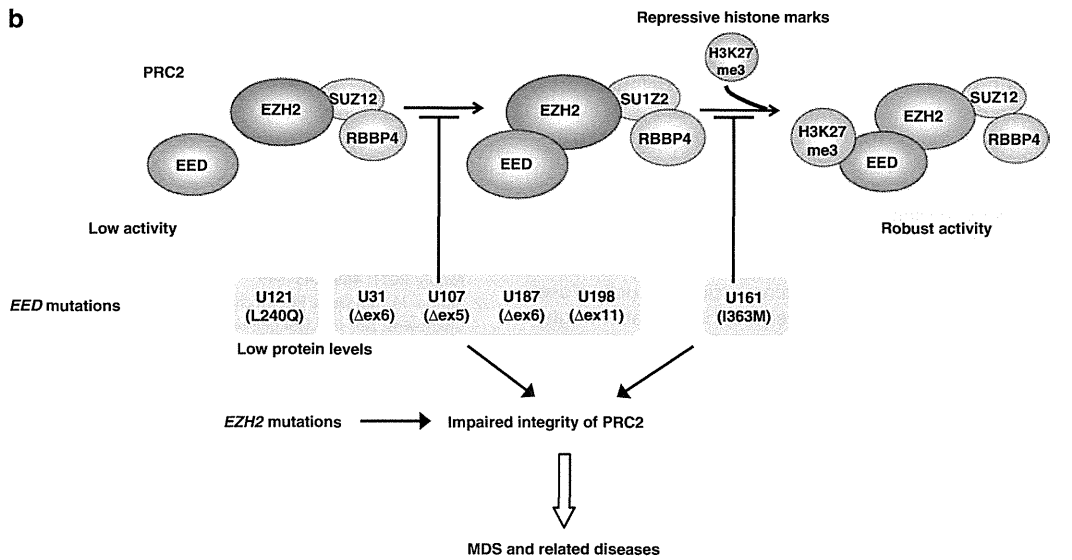


Figure 2. *EED* mutations in MDS and related diseases. (**a**) Characterization of samples with *EED* mutation. [a]Patient ID and SNP array are as previously described.[10] [b]Classification is according to the International Prognostic Scoring System. ND, not determined. (**b**) Impaired integrity of PRC2 promotes the pathogenesis of MDS and related diseases. *EED* exon deletions and L240Q mutation inhibit the PRC2 function through lower expression of EED proteins with or without impaired binding to EZH2. The I363M mutation interrupts the active interaction between H3K27me3 and EED.

Figure 1. Functionally defective mutations of *EED* in MDS and related diseases. (**a**) *EED* mutations detected in individuals with MDS and related diseases. The exon/intron structure, nucleotide sequence chromatogram, predicted protein structure and the Myc-tagged EED proteins used in this study are shown. I-V, WD40 repeat motifs. Dagger symbols indicate amino acids whose mutations were linked to myeloproliferative disorder in mice. Asterisks indicate amino acids necessary for binding to H3K27me3. In exon-deleted subjects (U31, U107, U187 and U198), PCR spanning the deletion point resulted in two distinct bands. Sequence chromatograms show the results from direct sequencing of the purified smaller bands, whereas the larger bands represented the corresponding wild-type *EED* sequences (data not shown). (**b**) EED protein and mRNA levels following forced expression in 293T cells. Cells were harvested with 2% SDS-sample buffer for western blot analysis (top) or Trizol reagent for northern blot analysis (bottom). Arrowheads indicate poorly expressed proteins of exon deletion mutants. EV, empty vector. (**c**) Interaction of EZH2 with EEDs. 293T cells were transfected with plasmids expressing the proteins indicated. The amounts of plasmids were adjusted to achieve similar protein expression levels at the stage of transfection. Cell lysates were immunoprecipitated and analyzed by western blot using anti-Flag (EZH2) and anti-Myc (EED) antibodies. Input represents 0.5% or 2% of cell lysate used for IP (Myc or Flag, respectively). IgL, immunoglobulin light chain. (**d**) I363M mutant decreases global H3K27me3 levels through impaired interaction to H3K27me3. Left: The alignment of amino-acid sequences surrounding Ile363. Blue-shaded boxes indicate well-conserved aromatic cage residues. The yellow-shaded box indicates the Ile363Met mutation found in subject U161 CMML. (Protein DataBank ID code: 3IIW). The boxed area is magnified on the right side. Middle: Myc-tagged wild-type EED (WT) or I363M mutant was co-expressed with Flag-tagged EZH2 in 293T cells. Cell lysates were analyzed in a pull-down assay using H3K27me3 peptide by western blot using an anti-Myc antibody. Input represents 0.5% of cell lysate used for the pull-down assay. Relative binding efficiencies (WT = 1) were estimated by normalizing the densitometry values representing the bound EED against those from the 'Input EED'. Bar graph: mean binding ± s.e. from three independent experiments. *$P = 0.0028$ (Student's *t*-test). Right: Myc-tagged wild-type EED (WT) or I363M mutant was retrovirally transduced in NIH3T3 cells. Acid-extracted histones were analyzed by western blot using an anti-H3K27me3 or an anti-total H3.

codon (Figure 1a). In subjects with a deletion (U31, U107, U187 and U198), the deletion points demonstrated the joining of the 5′ part and 3′ part with 1- to 8-base-pair microhomology (Figure 1a), which suggests the possibility that these deletions resulted from recombination with unequal crossover.[12]

We first evaluated the cellular expression levels of the mutant proteins following forced expression in 293T cells. The result showed that exon deletions (ΔC185, Δ143-184, L240Q and ΔC376-441) and L240Q mutant expressed significantly less protein than did either the wild-type or I363M (Figure 1b, top, indicated by arrowheads), despite a similar level of expression of the corresponding mRNAs (Figure 1b, bottom). These results indicate that the exon deletions and L240Q impaired translation and/or stability of the EED proteins, and suggest that these mutated forms are hypomorphic, and thus are functionally defective. Interaction of EED with EZH2 was demonstrated to be necessary for the full activity of PRC2;[2] hence we next examined the binding ability of the mutant proteins to EZH2. As expected from the results of the previous studies showing that WD repeats are important for EED–EZH2 interaction,[6,13] the ΔC185, Δ143-184 and ΔC376-441 failed to bind to EZH2 (Figure 1c). L240Q, in contrast, showed comparable binding ability to EZH2 as compared with wild-type EED. Thus, these results suggest that exon deletions and an L240Q mutation disrupt the functional integrity of PRC2, owing to poor protein expression coupled with or without their inability to bind EZH2.

On the other hand, the I363M mutant retained the ability to bind EZH2 and was expressed at a level similar to that of the wild-type EED (Figures 1b and c), suggesting that this mutant could incorporate into the PRC2 complex comparably with the wild type. However, substitution of an amino acid in such close proximity to the cage residues raised the possibility that it might affect the EED-H3K27me3 interaction, and therefore PCR2 function (Figure 1d, left).[7] We compared the binding of wild-type and I363M mutant EED in a pull-down assay that employed a synthetic H3K27me3 peptide ligand. Intriguingly, the I363M substitution significantly inhibited the EED-H3K27me3 interaction when co-expressed in the presence of EZH2 (Figure 1d, middle). In addition, global H3K27me3 levels were severely decreased in cells stably overexpressing the I363M mutant (Figure 1d, right), suggesting that the PRC2 complex incorporating the I363M mutant is functionally compromised, possibly through impaired structural integrity of the aromatic cage.

In summary, all the six mutated forms of EED displayed functional defects involving changes: (i) protein stability, (ii) interaction with EZH2 and/or (iii) binding to H3K27me3, thereby impairing PRC2 function (Figure 2b). We suggest that, in addition to inactivating mutations of catalytic EZH2,[3–5] non-catalytic EED mutations exclusively perturb PRC2-mediated epigenetic regulation and substantially contribute to the pathogenesis of MDS and related diseases (Figure 2b). Recently, Score *et al.*[14] reported a set of defective gene mutations of PRC2 constituents, including an EED point mutation, Gly255Asp, in 148 MDS/MPN cases. Our data suggest that various types of defective EED mutations contribute to the MDS pathogenesis. Analysis of more samples would clarify the clinical features of patients with EED mutation(s) in MDS and related diseases. Our findings highlight that recurrent mutations in PRC2 may constitute a new molecular-based disease category of myeloid malignancies.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

We thank Yu Iwai and Megumi Nakamura for technical assistance, and Dragon Genomics Center, Takara Bio Inc. for sequencing of the PCR products. This work was supported by a Grant-in-Aid from the Ministry of Education, Science and Culture of Japan.

T Ueda[1,7], M Sanada[2,7], H Matsui[3], N Yamasaki[1], Z-i Honda[4], L-Y Shih[5], H Mori[6], T Inaba[3], S Ogawa[2] and H Honda[1]

[1]*Department of Disease Model, Research Institute for Radiation Biology and Medicine, Hiroshima University, Hiroshima, Japan;*

[2]*Cancer Genomics Project, The University of Tokyo, Tokyo, Japan;*

[3]*Department of Molecular Oncology, Research Institute for Radiation Biology and Medicine, Hiroshima University, Hiroshima, Japan;*

[4]*Health Care Center of Humanities and Sciences, Ochanomizu University, Tokyo, Japan;*

[5]*Division of Hematology-Oncology, Department of Internal Medicine, Chang Gung Memorial Hospital, Chang Gung University, Taipei, Taiwan and*

[6]*Division of Hematology, Internal Medicine, Showa University Fujigaoka Hospital, Kanagawa, Japan*

E-mail: sogawa-tky@umin.ac.jp (or) hhonda@hiroshima-u.ac.jp

[7]*These authors contributed equally to this work.*

REFERENCES

1 Margueron R, Reinberg D. The Polycomb complex PRC2 and its mark in life. *Nature* 2011; **469**: 343–349.
2 Simon JA, Kingston RE. Mechanisms of polycomb gene silencing: knowns and unknowns. *Nat Rev Mol Cell Biol* 2009; **10**: 697–708.
3 Nikoloski G, Langemeijer SM, Kuiper RP, Knops R, Massop M, Tönnissen ER *et al.* Somatic mutations of the histone methyltransferase gene EZH2 in myelodysplastic syndromes. *Nat Genet* 2010; **42**: 665–667.
4 Ernst T, Chase AJ, Score J, Hidalgo-Curtis CE, Bryant C, Jones AV *et al.* Inactivating mutations of the histone methyltransferase gene EZH2 in myeloid disorders. *Nat Genet* 2010; **42**: 722–726.
5 Makishima H, Jankowska AM, Tiu RV, Szpurka H, Sugimoto Y, Hu Z *et al.* Novel homo- and hemizygous mutations in EZH2 in myeloid malignancies. *Leukemia* 2010; **24**: 1799–1804.
6 Denisenko O, Shnyreva M, Suzuki H, Bomsztyk K. Point mutations in the WD40 domain of Eed block its interaction with Ezh2. *Mol Cell Biol* 1998; **18**: 5634–5642.
7 Margueron R, Justin N, Ohno K, Sharpe ML, Son J, Drury V *et al.* Role of the polycomb protein EED in the propagation of repressive histone marks. *Nature* 2009; **461**: 762–767.
8 Lessard J, Schumacher A, Thorsteinsdottir U, van Lohuizen M, Magnuson T, Sauvageau G. Functional antagonism of the Polycomb-Group genes eed and Bmi1 in hemopoietic cell proliferation. *Genes Dev* 1999; **13**: 2691–2703.
9 Sauvageau M, Miller M, Lemieux S, Lessard J, Hébert J, Sauvageau G. Quantitative expression profiling guided by common retroviral insertion sites reveals novel and cell type specific cancer genes in leukemia. *Blood* 2008; **111**: 790–799.
10 Sanada M, Suzuki T, Shih LY, Otsu M, Kato M, Yamazaki S *et al.* Gain-of-function of mutated C-CBL tumour suppressor in myeloid neoplasms. *Nature* 2009; **460**: 904–908.
11 1000 Genomes Project Consortium. A map of human genome variation from population-scale sequencing. *Nature* 2010; **467**: 1061–1073.
12 Maruyama H, Morino H, Ito H, Izumi Y, Kato H, Watanabe Y *et al.* Mutations of optineurin in amyotrophic lateral sclerosis. *Nature* 2010; **465**: 223–226.
13 Montgomery ND, Yee D, Montgomery SA, Magnuson T. Molecular and functional mapping of EED motifs required for PRC2-dependent histone methylation. *J Mol Biol* 2007; **374**: 1145–1157.
14 Score J, Hidalgo-Curtis C, Jones AV, Winkelmann N, Skinner A, Ward D *et al.* Inactivation of polycomb repressive complex 2 components in myeloproliferative and myelodysplastic/myeloproliferative neoplasms. *Blood* 2011; **119**: 1208–1213.

Supplementary Information accompanies the paper on the Leukemia website (http://www.nature.com/leu)

Open

Leukemia (2012) **26,** 1879–1898

www.nature.com/leu

LETTERS TO THE EDITOR

Novel splicing-factor mutations in juvenile myelomonocytic leukemia

Leukemia (2012) **26,** 1879–1881; doi:10.1038/leu.2012.45

Myelodysplastic syndromes (MDS) and myelodysplastic/myeloproliferative neoplasms (MDS/MPN) are heterogeneous groups of chronic myeloid neoplasms characterized by clonal hematopoiesis, varying degrees of cytopenia or myeloproliferative features with evidence of myelodysplasia and a propensity to acute myeloid leukemia (AML).[1] In recent years, a number of novel gene mutations, involving *TET2*, *ASXL1*, *DNMT3A*, *EZH2*, *IDH1/2*, and *c-CBL*, have been identified in adult cases of chronic myeloid neoplasms, which have contributed to our understanding of disease pathogenesis.[2–7] However, these mutations are rare in pediatric cases, with the exception of germline or somatic *c-CBL* mutations found in 10–15% of chronic myelomonocytic leukemia (CMML) and juvenile myelomonocytic leukemia (JMML),[8] highlighting the distinct pathogenesis of adult and pediatric neoplasms.[9]

Recently, we reported high frequencies of mutations, involving the RNA splicing machinery, that are largely specific to myeloid neoplasms, showing evidence of myeloid dysplasia in adult.[10] Affecting a total of eight components of the RNA splicing machinery (U2AF35, U2AF65, SF3A1, SF3B1, SRSF2, ZRSR2, SF1 and PRPF40B) commonly involved in the 3′ splice-site (3′SS) recognition, these pathway mutations are now implicated in the pathogenesis of myelodysplasia.[10] To investigate the role of the splicing-pathway mutations in the pathogenesis of pediatric myeloid malignancies, we have examined 165 pediatric cases with AML, MDS, chronic myeloid leukemia (CML) and JMML for mutations in the four major splicing factors, *U2AF35*, *ZRSR2*, *SRSF2*, and *SF3B1*, commonly mutated in adult cases.

Bone marrow or peripheral blood tumor specimens were obtained from 165 pediatric patients with various myeloid malignancies, including *de novo* AML ($n=93$), MDS ($n=28$), CML ($n=17$) and JMML ($n=27$), and the genomic DNA (gDNA) was subjected to mutation analysis (Supplementary Table 1). The status of the RAS pathway mutations for the current JMML series has been reported previously (Supplementary Table 2).[11,12] Nineteen leukemia cell lines derived from AML (YNH-1, ML-1, KASUMI-3, KG-1, HL60, inv-3, SN-1, NB4 and HEL), acute monocytic leukemia (THP-1, SCC-3, J-111, CTS, P31/FUJ, MOLM-13, IMS/MI and KOCL-48) and acute megakaryoblastic leukemia (CMS and CMY) were also analyzed for mutations. Peripheral blood gDNA from 60 healthy adult volunteers was used as controls. Informed consent was obtained from the patients and/or their parents and from the healthy volunteers. We previously showed that for *U2AF35*, *SRSF2* and *SF3B1*, most of the mutations in adult cases were observed in exons 2 and 7, exon 1, and exons 14 and 15, respectively.[10] Therefore, we confirmed mutation screening to these 'hot-spot' exons. In contrast, all the coding exons were examined for *ZRSR2*, because no mutational hot spots have been detected. Briefly, the relevant exons were amplified using PCR and mutations were examined by Sanger sequencing, as previously described.[10] The Fisher's exact test was used to evaluate the statistical significance of frequencies of mutations for *U2AF35*, *SF3B1*, *ZRSR2* or *SRSF2* in adult cases and pediatric cases. This study was approved by the Ethics Committee of the University of Tokyo (Approval number 948-7).

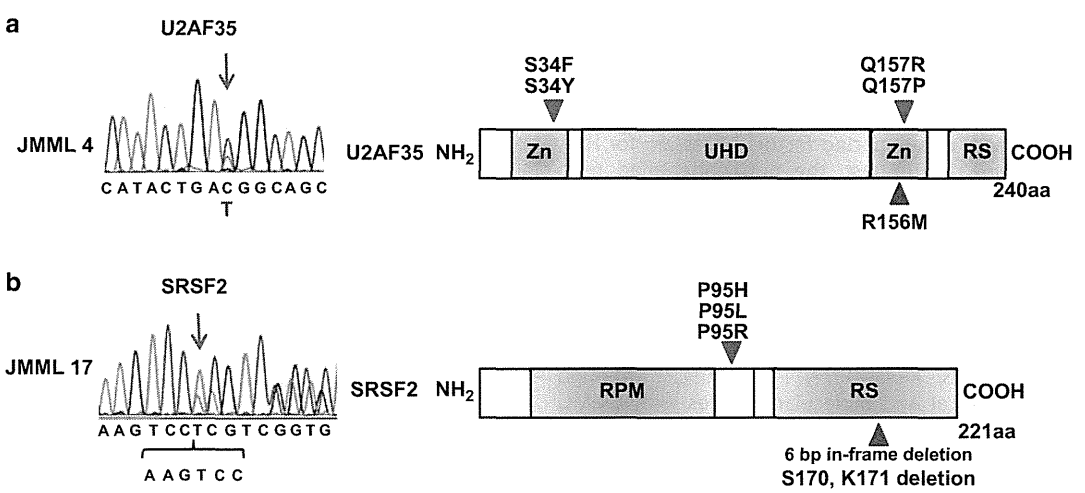


Figure 1. Novel *U2AF35*and *SRSF2* mutations detected in JMML cases. (**a**) Left panel: sequence chromatogram of a heterozygous mutation at R156 in N-terminal zinc-finger motifs of *U2AF35* detected in a JMML case (JMML 4) is shown. Mutated nucleotides are indicated by arrows. Right panel: illustration of functional domains and mutations of U2AF35. Red arrow heads indicate hot-spot mutations at S34 and Q157 detected in the adult cases.[10] Blue arrow head indicates the missense mutation at R156. (**b**) Left panel: sequence chromatogram of a 6-bp in-frame deletion (c.518-523delAAGTCC) in *SRSF2* detected in JMML 17 is shown. Mutated nucleotides are indicated by arrows. Right panel: illustration of functional domains and mutations of SRSF2. Red arrow head indicates hot-spot mutation at P95 frequently detected in the adult cases.[10] Blue arrow head indicates a 6-bp in-frame deletion leading to deletion of S170 and K171.

Accepted article preview online 20 February 2012; advance online publication, 2 March 2012

No mutations were identified in the 28 cases with pediatric MDS, which included 13 cases with refractory anemia with excess blasts, 5 with refractory cytopenia of childhood, 2 with Down syndrome-related MDS, 2 with Fanconi anemia-related MDS, 2 with secondary MDS and 4 with unclassified MDS. Similarly, no mutations were detected in 93 cases with *de novo* AML or in 17 with CML, as well as 19 leukemia-derived cell lines. Our previous study in adult patients showed the frequency of mutations in *U2AF35*, *SF3B1*, *ZRSR2* or *SRSF2* to be 60/155 cases with MDS without increased ring sideroblasts and 8/151 *de novo* AML patients, emphasizing the rarity of these mutations in pediatric MDS ($P<5.0\times10^{-6}$) and AML ($P<0.02$) compared with adult cases. We found mutations in two JMML cases, JMML 4 and JMML 17. JMML 4 carried a heterozygous *U2AF35* mutation (R156M), whereas JMML 17 had a 6-bp in-frame deletion (c.518-523delAAGTCC) in *SRSF2* that resulted in deletion of amino acids S170 and K171 (Figure 1). Both nucleotide changes found in *U2AF35* and *SRSF2* were neither identified in the 60 healthy volunteers nor registered in the dbSNP database (http://www.ncbi.nlm.nih.gov/projects/SNP/) or in the 1000 genomes project, indicating that they represent novel spliceosome mutations in pediatric cases.

U2AF35 is the small subunit of the U2 auxiliary factor (U2AF), which binds an AG dinucleotide at the 3′SS, and has an essential role in RNA splicing.[13] With the exception of a single A26V mutation found in a case of refractory cytopenia with multilinage dysplasia, all the U2AF35 mutations reported in adult myeloid malignancies involved one of the two hot spots within the two zinc-finger domains, S34 and Q157, which are highly conserved across species, suggesting the gain-of-function mutations.[10] In JMML 4, the R156M U2AF35 mutation affects a conserved amino acid adjacent to Q157, suggesting it may also be a gain-of-function mutation, leading to aberrant pre-mRNA splicing possibly in a dominant fashion.

SRSF2, better known as SC35, is a member of the serine/arginine-rich (SR) family of proteins.[14] SRSF2 binds to a splicing-enhancer element in pre-mRNA and has a crucial role not only in constitutive and alternative pre-mRNA splicing but also in transcription elongation and genomic stability.[14] All mutations thus far identified in adult cases exclusively involved P95 within the intervening sequence between the N-terminal RNA-binding domain and the C-terminal RS domain.[10] This region interacts with other SR proteins, again suggesting that the P95 mutation may result in gain-of-function.[10] This proline residue is thought to determine the relative orientation of the two flanking domains of SRSF2, and a substitution at this position could compromise critical interactions with other splicing factors necessary for RNA splicing to take place. In contrast, the newly identified 6-bp in-frame deletion in JMLL 17 results in two conserved amino acids, S170 and K171, within the RS domain. Although it may affect protein–protein interactions, the functional significance of this deletion remains elusive.

JMML is a unique form of pediatric MDS/MPN characterized by activation of the RAS/mitogen-activated protein kinase signaling pathway; in 90% of cases, there are germ line and/or somatic mutations of *NF1*, *NRAS*, *KRAS*, *PTPN11* and *CBL*.[8] Although JMML shares some clinical and molecular features with CMML, its spectrum of gene mutations suggests that it is a neoplasm distinct from CMML.[15] This was also confirmed by the current results that the splicing-pathway mutations are rare in JMML, whereas they are extremely frequent (~60%) in CMML.[10] Although the two JMML cases carrying the splicing-pathway mutations had no known RAS-pathway mutations, both the pathway mutations frequently coexisted in CMML.[8]

To summarize, no mutations of *SF3B1*, *U2AF35*, *ZRSR2* or *SRSF2* are found in pediatric MDS and AML. In our study, except for *ZRSR2*, mutations were examined focusing on the reported hot spots in adult studies, raising a possibility that we may have missed some mutations occurring in other regions. However, these hot spots represent evolutionally conserved amino acids and have functional relevance, it is unlikely that the distribution of hot spots in children significantly differs from adult cases and as such, we could safely conclude that mutations of *SF3B1*, *U2AF35*, *ZRSR2* and *SRSF2* are rare in myeloid neoplasms in children. Finally, mutations of *U2AF35* and *SRSF2* may have some role in the pathogenesis of JMML, although further evaluations are required.

CONFLICT OF INTEREST

The authors declare no conflict interest.

ACKNOWLEDGEMENTS

This work was supported by Research on Measures for Intractable Diseases, Health and Labor Sciences Research Grants, Ministry of Health, Labor and Welfare, by Research on Health Sciences focusing on Drug Innovation, and the Japan Health Sciences Foundation. We would like to thank M Matsumura, M Yin, N Hoshino and S Saito for their excellent technical assistance.

J Takita[1,2], K Yoshida[3], M Sanada[3], R Nishimura[1], J Okubo[1], A Motomura[1], M Hiwatari[1], K Oki[1], T Igarashi[1], Y Hayashi[4] and S Ogawa[3]
[1]*Department of Pediatrics, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan;*
[2]*Department of Cell Therapy and Transplantation Medicine, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan;*
[3]*Cancer Genomics Project, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan and*
[4]*Gunma Children's Medical Center, Gunma, Japan*
E-mail: sogawa-tky@umin.ac.jp

REFERENCES

1 Garcia-Manero G. Myelodysplastic syndromes: 2011 update on diagnosis, risk-stratification, and management. *Am J Hematol* 2011; **86**: 490–498.
2 Delhommeau F, Dupont S, Della Valle V, James C, Trannoy S, Masse A *et al.* Mutation in TET2 in myeloid cancers. *N Engl J Med* 2009; **360**: 2289–2301.
3 Thol F, Friesen I, Damm F, Yun H, Weissinger EM, Krauter J *et al.* Prognostic significance of ASXL1 mutations in patients with myelodysplastic syndromes. *J Clin Oncol* 2011; **29**: 2499–2506.
4 Ley TJ, Ding L, Walter MJ, McLellan MD, Lamprecht T, Larson DE *et al.* DNMT3A mutations in acute myeloid leukemia. *N Engl J Med* 2010; **363**: 2424–2433.
5 Nikoloski G, Langemeijer SM, Kuiper RP, Knops R, Massop M, Tonnissen ER *et al.* Somatic mutations of the histone methyltransferase gene EZH2 in myelodysplastic syndromes. *Nature Genet* 2010; **42**: 665–667.
6 Green A, Beer P. Somatic mutations of IDH1 and IDH2 in the leukemic transformation of myeloproliferative neoplasms. *N Engl J Med* 2010; **362**: 369–370.
7 Sanada M, Suzuki T, Shih LY, Otsu M, Kato M, Yamazaki S *et al.* Gain-of-function of mutated C-CBL tumour suppressor in myeloid neoplasms. *Nature* 2009; **460**: 904–908.
8 Perez B, Kosmider O, Cassinat B, Renneville A, Lachenaud J, Kaltenbach S *et al.* Genetic typing of CBL, ASXL1, RUNX1, TET2 and JAK2 in juvenile myelomonocytic leukaemia reveals a genetic profile distinct from chronic myelomonocytic leukaemia. *Br J Haematol* 2010; **151**: 460–468.
9 Oki K, Takita J, Hiwatari M, Nishimura R, Sanada M, Okubo J *et al.* IDH1 and IDH2 mutations are rare in pediatric myeloid malignancies. *Leukemia* 2011; **25**: 382–384.
10 Yoshida K, Sanada M, Shiraishi Y, Nowak D, Nagata Y, Yamamoto R *et al.* Frequent pathway mutations of splicing machinery in myelodysplasia. *Nature* 2011; **478**: 64–69.
11 Chen Y, Takita J, Hiwatari M, Igarashi T, Hanada R, Kikuchi A *et al.* Mutations of the PTPN11 and RAS genes in rhabdomyosarcoma and pediatric hematological malignancies. *Genes Chromosomes Cancer* 2006; **45**: 583–591.
12 Shiba N, Kato M, Park MJ, Sanada M, Ito E, Fukushima K *et al.* CBL mutations in juvenile myelomonocytic leukemia and pediatric myelodysplastic syndrome. *Leukemia* 2010; **24**: 1090–1092.

13 Zhang M, Zamore PD, Carmo-Fonseca M, Lamond AI, Green MR. Cloning and intracellular localization of the U2 small nuclear ribonucleoprotein auxiliary factor small subunit. *Proc Natl Acad Sci USA* 1992; **89**: 8769–8773.

14 Edmond V, Brambilla C, Brambilla E, Gazzeri S, Eymin B. SRSF2 is required for sodium butyrate-mediated p21(WAF1) induction and premature senescence in human lung carcinoma cell lines. *Cell Cycle* 2011; **10**: 1968–1977.

15 Emanuel PD. Juvenile myelomonocytic leukemia and chronic myelomonocytic leukemia. *Leukemia* 2008; **22**: 1335–1342.

Supplementary Information accompanies the paper on the Leukemia website (http://www.nature.com/leu)

Sequencing histone-modifying enzymes identifies UTX mutations in acute lymphoblastic leukemia

Leukemia (2012) **26,** 1881–1883; doi:10.1038/leu.2012.56

Mutations affecting epigenetic regulators have long been known to have a crucial role in cancer and, in particular, hematological malignancies.[1,2] One of the earliest epigenetic factors described altered in leukemia was the mixed lineage leukemia (*MLL*) protein which is found translocated in 10% of adult acute myeloid leukemia (AML), 30% of secondary AML and >75% of infants with both AML and acute lymphocytic leukemia (ALL). *MLL* is a SET domain-containing protein, which is recruited to many promoters and mediates histone 3 lysine 4 (H3K4) methyltransferase activity, thought to promote gene expression.[3]

In addition to *MLL* fusions, recently, somatic mutations of *UTX* (also known as *KDM6A*), encoding an H3K27 demethylase, were described in multiple hematological malignancies, including multiple myeloma and many types of leukemia cell lines.[4,5] H3K27 methylation is generally thought to cause gene repression. Complimentary to *UTX*, mutations of *EZH2*, a H3K27 methyltransferase, have been reported in both lymphoid and myeloid tumors (Figure 1).[6,7] These mutations lead to altered EZH2 activity and influence H3K27 in tumor cells. Mutations in EZH2, EED and SUZ12, which all cooperate in Polycomb repressive complex 2 have been recently described in early T-cell precursor ALL.[8] Similarly, point mutations affecting the functional jumonji C (jmjC) domain of UTX inactivates its H3K27 demethylase activity. In addition, UTX associates with MLL2 in a multiprotein complex, which promotes H3K4 methylation, and recently *MLL2* has also been found mutated in cancer, further pointing to a common and complex epigenetic deregulation in cancer.[9] In line with the growing evidence for epigenetic regulators as important in tumorigenesis, additional mutations affecting epigenetic regulators such as SETD2, a H3K36 methyltransferase, KDM3B, a H3K9 demethylase, and KDM5C, a H3K4 demethylase, have been reported and are associated with distinct gene expression patterns (Figure 1).[4]

Though the clinical significance of these findings remains to be explored, it is evident that epigenetic deregulation is having an important role in both lymphoid and myeloid leukemogenesis. Furthermore, with novel drugs at hand, such as histone deacetylase inhibitors or demethylating agents that can target and reverse epigenetic alterations, understanding the underlying molecular aberrations is of growing interest.[10] We therefore undertook an effort to examine the prevalence of somatic mutations in genes encoding histone-modifying proteins, in particular, *KDM3B, KDM5C, UTX, MLL2, EZH2* and *SETD2*, which previously were reported mutated in cancer.[4,5]

For an initial screen, we analyzed banked diagnostic primary leukemia samples from 44 childhood B-cell ALL and 50 adult AML patients, and, where available, used bone marrow samples obtained in complete remission to validate the somatic nature of the mutations. Samples had been collected with patient/parental informed consent from patients enrolled on Dana–Farber Cancer Institute protocols for childhood ALL (DFCI 00-001 (NCT00165178), DFCI 05-001 (NCT00400946)) or AML treatment protocols of the German-Austrian AML Study Group (AMLSG) for younger adults (AMLSG-HD98A (NCT00146120), AMLSG 07-04 (NCT00151242)), and the study was approved by the IRB of the participating centers.

Using conventional Sanger sequencing of primary leukemia sample-derived genomic DNA, we first screened all coding exons in which mutations have been reported previously.[4,5] Initially, we analyzed a total of 36 of 174 exons (*KDM3B* (2/24), *KDM5C* (9/26), *UTX* (7/29), *MLL2* (8/54), *EZH2* (1/20) and *SETD2* (9/21)) and found 7 non-synonymous tumor-specific aberrations. In AML, we found one *EZH2* mutation (p.G648E) in a t(8;21)-positive, and two *MLL2* missense mutations (p.R5153Q and p.Y5216S; Table 1) and one

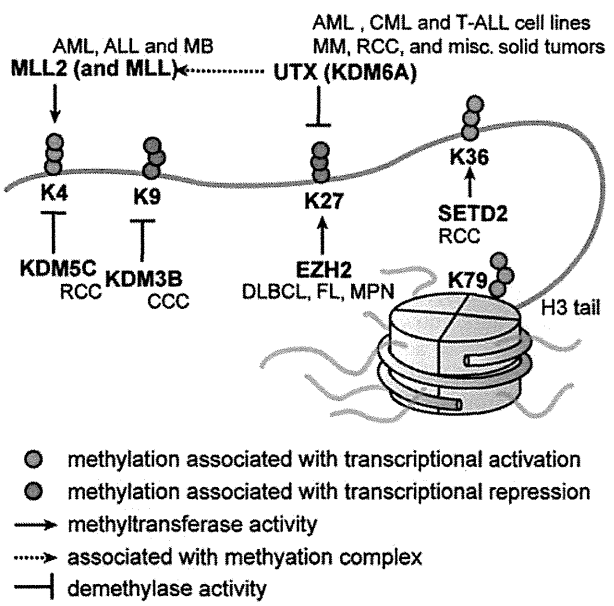


Figure 1. Histone 3 methylation and selected histone demethylases and methyltransferases. Cancers are shown in italics next to the mutated protein they are associated with. *MM*, multiple myeloma; FL, follicular lymphoma; DLBCL, diffuse large B-cell lymphoma; RCC, renal cell carcinoma; CCC clear cell carcinoma; MPN, myeloproliferative neoplasm; MB, medulloblastoma.

Accepted article preview online 1 March 2012; advance online publication, 3 April 2012

Biochemical and Biophysical Research Communications 419 (2012) 662–669

Contents lists available at SciVerse ScienceDirect

Biochemical and Biophysical Research Communications

journal homepage: www.elsevier.com/locate/ybbrc

Using peripheral blood circulating DNAs to detect CpG global methylation status and genetic mutations in patients with myelodysplastic syndrome

Chisako Iriyama [a], Akihiro Tomita [a,*], Hideaki Hoshino [a], Mizuho Adachi-Shirahata [a], Yoko Furukawa-Hibi [b], Kiyofumi Yamada [b], Hitoshi Kiyoi [a], Tomoki Naoe [a]

[a] Department of Hematology and Oncology, Nagoya University Graduate School of Medicine, Nagoya, Japan
[b] Department of Neuropsychopharmacology and Hospital Pharmacy, Nagoya University School of Medicine, Nagoya, Japan

ARTICLE INFO

Article history:
Received 7 February 2012
Available online 20 February 2012

Keywords:
MDS
Circulating DNA
Genetic mutations
Epigenetics
LINE-1
TET2

ABSTRACT

Myelodysplastic syndrome (MDS) is a hematopoietic stem cell disorder. Several genetic/epigenetic abnormalities are deeply associated with the pathogenesis of MDS. Although bone marrow (BM) aspiration is a common strategy to obtain MDS cells for evaluating their genetic/epigenetic abnormalities, BM aspiration is difficult to perform repeatedly to obtain serial samples because of pain and safety concerns. Here, we report that circulating cell-free DNAs from plasma and serum of patients with MDS can be used to detect genetic/epigenetic abnormalities. The plasma DNA concentration was found to be relatively high in patients with higher blast cell counts in BM, and accumulation of DNA fragments from mono-/di-nucleosomes was confirmed. Using serial peripheral blood (PB) samples from patients treated with hypomethylating agents, global methylation analysis using bisulfite pyrosequencing was performed at the specific CpG sites of the *LINE-1* promoter. The results confirmed a decrease of the methylation percentage after treatment with azacitidine (days 3–9) using DNAs from plasma, serum, and PB mononuclear cells (PBMNC). Plasma DNA tends to show more rapid change at days 3 and 6 compared with serum DNA and PBMNC. Furthermore, the *TET2* gene mutation in DNAs from plasma, serum, and BM cells was quantitated by pyrosequencing analysis. The existence ratio of mutated genes in plasma and serum DNA showed almost equivalent level with that in the CD34+/38- stem cell population in BM. These data suggest that genetic/epigenetic analyses using PB circulating DNA can be a safer and painless alternative to using BM cells.

1. Introduction

Myelodysplastic syndrome (MDS) is one of the hematopoietic stem cell disorders, showing the features of dysplasia and a high rate of progression to acute myeloid leukemia (AML). Recent reports using next generation DNA sequencing techniques and single nuclear polymorphism (SNP) array analyses have suggested that specific gene mutations, resulting in aberrant DNA methylation [1], histone modification [2], and RNA splicing [3], may contribute to the pathogenesis of MDS. Furthermore, it is also speculated that an aberrant methylation status in some specific gene promoters [4,5] also contributes to pathogenesis and disease progression [6].

DNA hypomethylating reagents such as azacitidine and decitabine, known to be DNA methyltransferase inhibitors (DNMTi) [7], have recently come to be considered as standard therapeutics for patients with MDS [8]. Although DNMTi provide improvement of cytopenia and reduction of blast counts for certain patients, biomarkers, such as genetic mutations and methylation status of specific promoters, that predict the effectiveness of DNMTi before and/or during treatment, are still unclear. So far, cytogenetic and molecular analyses using bone marrow (BM) cells are the standard strategies for confirming the disease status of MDS; however, a disadvantage is that patients hesitate the performance of repeated BM aspiration because the procedure is more painful than peripheral blood (PB) aspiration. It is therefore highly desirable to find alternative strategies for detecting the serial genetic/epigenetic alterations those occur in BM cells.

Recently, circulating cell-free nucleic acids in the plasma and serum of PB, such as genomic DNA, mRNA, and microRNA, are recognized as useful materials for the detection of genetic/epigenetic abnormalities in malignant cells especially in patients who have solid tumors [9]. Previous reports suggest that a higher concentration of these nucleic acids correlates with disease progression or a higher tumor burden of solid tumors [10]. Furthermore, specific genetic mutations and epigenetic abnormalities including DNA methylation in solid tumors are also detectable by using circulating nucleic acids [9]. Circulating nucleic acids are expected to be

* Corresponding author. Address: Department of Hematology and Oncology, Nagoya University Graduate School of Medicine, Tsurumai-cho 65, Showa-ku, Nagoya 466-8550, Japan. Fax: +81 52 744 2161.
E-mail address: atomita@med.nagoya-u.ac.jp (A. Tomita).

0006-291X/$ - see front matter © 2012 Elsevier Inc. All rights reserved.
doi:10.1016/j.bbrc.2012.02.071

good materials for determining tumor status without performing re-operations or re-biopsies.

This report aims to show the usefulness of PB circulating DNAs for genetic/epigenetic analyses in patients with MDS. Plasma and serum circulating DNAs were obtained repeatedly from patients' PB after the administration of DNMTi. Serial changes of global DNA methylation status were successfully confirmed by using bisulfite pyrosequencing analysis. We propose that analysis using circulating DNAs can be a safer and painless alternate strategy compared to repeated BM aspiration for determining the genetic and epigenetic events in BM cells.

2. Materials and methods

2.1. Patients

Five patients with MDS in Nagoya University Hospital were enrolled into this analysis after obtaining appropriate informed consent. Patient UPN1 was a 74-year-old male who was diagnosed with MDS refractory anemia with excess blasts (RAEB)-1 by the World Health Organization (WHO) classification. He died after three courses of therapy with a demethylating agent because of disease progression to acute myeloid leukemia (AML). UPN2 was a 75-year-old male who was diagnosed as chronic myelomonocytic leukemia. He died at day 9 after the first course of azacitidine treatment (5 days) because of severe pulmonary bleeding originating from a background of autoimmunity. UPN3 was 74-year-old male who was diagnosed with MDS RAEB-2, and received azacitidine for 5 days, every 4 weeks. UPN4 was 65-year-old male diagnosed with MDS RAEB-t/AML, and received azacitidine for 7 days, every 4 weeks. UPN5 was 77-year-old female diagnosed with MDS refractory cytopenia with multilineage dysplasia, who received therapy with a demethylating agent and achieved complete remission.

2.2. Preparation of serum, plasma, and MNC from PB and BM of MDS patients

PB was drawn and placed into plain tubes with a separating agent for serum, and tubes with sodium ethylenediaminetetraacetate or heparin for plasma. Plasma and serum were aliquoted into 1.5 mL tubes after centrifugation at 415*g* or 1660*g* for 10 min at room temperature, and stored at −80 °C until genetic analysis. PBMNC and BM cells were collected using Ficoll paque [11].

2.3. DNA extraction

Genomic DNAs from PBMNC and BM cells were extracted using the QIAamp DNA Blood Mini Kit (QIAGEN, Valencia, CA). Circulating DNAs in plasma and serum (450 μL each) were extracted using MinElute Virus Vacuum Kit (QIAGEN) according to the manufacturer's instructions.

2.4. DNA methylation analysis with the bisulfite pyrosequencing strategy

Bisulfite conversion of genomic or circulating DNAs was performed using MethylEasy Xceed Rapid DNA Bisulphite Modification Kit (Takara, Ohtsu, Japan). DNA (15–300 ng) was utilized for one assay of conversion. For the polymerase chain reaction (PCR) for the analysis of *long interspersed nuclear elements-1* (*LINE-1*) (GenBank; X58075), the following primers were used, as indicated previously [12]: LINE-1-F; 5′-TTTTGAGTTAGGTGTGGGATATA-3′, and LINE-1-R; 5′-AAAATCAAAAAATTCCCTTTC-3′ with 5′ end biotinylation. PCR conditions were 95 °C for 30 s, 54 °C for 30 s, and 72 °C for 40 s for a total of 40 cycles. The PCR products were purified and single stranded with the PyroMark Vacuum Workstation according to the manufacturer's protocol and analyzed by the PSQ96MA Pyrosequencing System (QIAGEN) [12]. The primer for the pyrosequencing of the *Line-1* PCR product was as follows: LINE-1-pyro; 5′-AGTTAGGTGTGGGATATAGT-3′ [12]. The ratio of methylation was calculated from the existence percentage of unmethylated and methylated cytosines, indicated as thymine and cytosine. The methylation status for each sample was analyzed at least twice, and the results were statistically evaluated.

2.5. Single nuclear polymorphism (SNP) array analysis

The SNP array was analyzed with BM DNA using the 250k_Nsp GeneChip-SNP (Affymetrix), as previously described [13].

2.6. Pyrosequencing analysis for TET2 mutation

The PCR primers for *TET2* exon 6 were as follows: TET2-Ex6-F; 5′-GGCTGCAGTGATTGTGATTC-3′, and TET2-Ex6-R; 5′-TTGGGCTTTCCTATCAGTGG-3′ with 5′ end biotinylation, with the following conditions: 95 °C for 15 s, 56 °C for 20 s, and 72 °C for 30 s for a total of 50 cycles. Sequencing analysis was performed by the ABI 310 genetic analyzer (Applied Biosystems, Foster City, CA) with the following sequencing primer; TET2-Ex6-F2; 5′-GTCTCTGGCTGACAAACTCT-3′. The existence percentage of the *TET2* mutation was measured by pyrosequencing analysis using the primer as follows; TET2-Ex6-pyro; 5′-AGTTAGGTGTGGGATATAGT-3′.

2.7. Cell sorting

BM cells were sorted into CD34(+)/CD38(−), CD34(+)/CD38(+), and CD34(−) subpopulations with the BD FACS Aria (Becton Dickinson, Franklin Lakes, NJ) using the anti-CD34 APC antibody (Becton Dickinson) and the anti-CD38 PE-Cy7 antibody (Becton Dickinson), as shown previously [14].

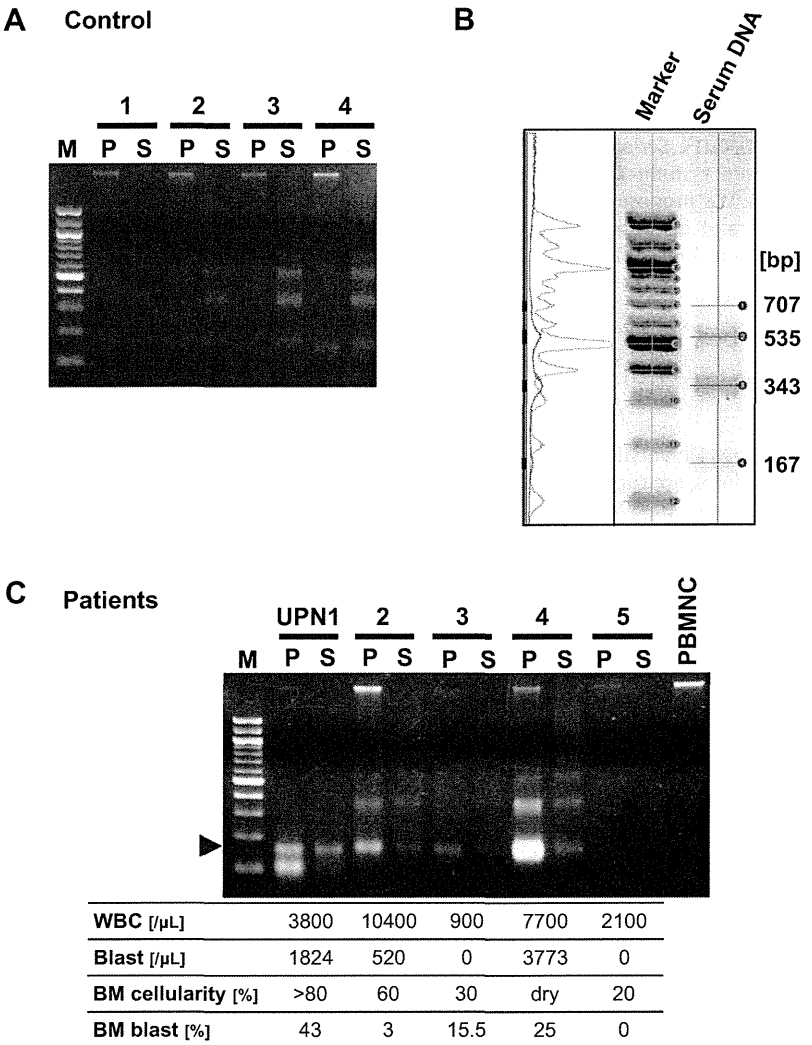
2.8. Statistical analysis

Using Prism version 5 software (Graph Pad Software, Inc., La Jolla, CA), differences in methylation and the mutation percentage were analyzed with two way repeated measure ANOVA and one-way factorial ANOVA, respectively. The *p* values were 2-tailed, and a *p*-value of less than .05 was considered statistically significant.

3. Results

3.1. Peripheral blood circulating DNAs in MDS patients

Peripheral blood circulating DNA from plasma and serum obtained from healthy volunteer donors (Fig. 1A, lanes 1–4, and B) and MDS patients (Fig. 1C, Patient UPN1 to UPN5) were visualized via agarose gel electrophoresis. These DNAs were all prepared from 450 μL of serum or plasma, and suspended in 30 μL of distilled water after DNA extraction. Twenty-five percent of the total obtained DNA was applied for each one lane of the electrophoresis. The DNA from healthy volunteers was confirmed in all lanes; the DNA concentration in serum was much higher than that in plasma (Fig. 1A, lanes S versus P). The circulating DNA showed the ladder pattern suggesting the fragmentation of genomic DNA by the effect of deoxyribonuclease against chromatinized genomic DNA [15], as reported previously [9]. The size of the DNA fragments, analyzed by BioMax 1D software (Kodak), supported this phenomenon showing the multiple numbers of 160–180 base pairs from mono-, di-nucleosomes, and so on.



WBC [/μL]	3800	10400	900	7700	2100
Blast [/μL]	1824	520	0	3773	0
BM cellularity [%]	>80	60	30	dry	20
BM blast [%]	43	3	15.5	25	0

Fig. 1. Detection of peripheral blood circulating DNA from normal and MDS patients. (A) Circulating DNA from plasma (P) and serum (S) harvested from four healthy volunteer donors was visualized via agarose gel electrophoresis. DNA ladders were confirmed, especially with respect to serum DNA. Weak bands were confirmed also in plasma DNA after a longer exposure (data not shown). (B) DNA ladders were measured by DNA analysis software, and the size was calculated. The estimated band size of the serum DNA was indicated at the right side of the gel image. (C) Plasma and serum circulating DNAs were harvested from five MDS patients (UPN1–5) and visualized via agarose gel electrophoresis. The laboratory data for the white blood cells (WBC) and the blast cell count in PB and BM are also indicated in the bottom panel. Note that the DNA concentration of the plasma DNA was higher in some patients who had higher blast counts in their bone marrow cells. M: 100 bp DNA ladder marker.

When using the samples from the patients with MDS, the DNA concentration in the plasma was relatively higher than that in the serum; and the concentration seemed to parallel the amount of tumor (blast) cells in the BM and peripheral blood (Fig. 1C, Patient UPN1, 2, and 4). In these patients, DNA fragments from mononucleosome were relatively accumulated (the black triangle in Fig. 1C), and further digestion of the DNA was confirmed in one patient (UPN1, P). The total amount of the DNA from 1 mL of plasma or serum varied from 1.40 to 141 μg. The genomic DNA was obtained from PBMNC and also loaded into the gel electrophoresis.

These data suggest that the circulating DNAs from plasma and serum are also confirmed in MDS patients. The concentration of the plasma DNA may tend to reflect the blast cell amounts in the BM cells.

3.2. *Analysis of* LINE-1 *methylation in plasma and serum circulating DNAs*

LINE-1 are repeated sequences that exist in the amount of about 85,000 copies in normal cells. *LINE-1* are moderately rich in CpG sites. Most methylated CpGs are located in the 5′ region of the sequence that can function as an internal promoter [9]. Recent reports indicate that the methylation status of the CpG sites reflects the global methylation status, and that aberrant hypomethylation of these sites may correlate with malignant tumor biology [12,16].

We performed bisulfite pyrosequencing analysis for four CpG sites of the *LINE-1* promoter (Fig. 2A) using plasma and serum circulating DNA to confirm the usefulness of those DNAs to determine the global methylation status in MDS patients. Plasma, serum and MNC were obtained from UPN2 at days 1, 4 and 6 after starting the first course of azacitidine treatment, and the *LINE-1* methylation percentage was confirmed. The methylation percentage of each CpG site generally decreased after started the azacitidine treatment (Fig. 2B). Next, the average methylation ratio of all four CpG sites was confirmed. The methylation percentage at day 1 (untreated) was adjusted as 1, and the relative methylation rates at days 4 and 6 were calculated (Fig. 2C). In this assay, the demethylation effect produced by azacitidine on the *LINE-1* element was confirmed in PBMNC and plasma at day 6 and also confirmed in