Ⅳ．研究成果の刊行物・別刷

ARTICLE

doi:10.1038/nature10496

Frequent pathway mutations of splicing machinery in myelodysplasia

Kenichi Yoshida[1]*, Masashi Sanada[1]*, Yuichi Shiraishi[2]*, Daniel Nowak[3]*, Yasunobu Nagata[1]*, Ryo Yamamoto[4], Yusuke Sato[1], Aiko Sato-Otsubo[1], Ayana Kon[1], Masao Nagasaki[5], George Chalkidis[6], Yutaka Suzuki[7], Masashi Shiosaka[1], Ryoichiro Kawahata[1], Tomoyuki Yamaguchi[8], Makoto Otsu[4], Naoshi Obara[9], Mamiko Sakata-Yanagimoto[9], Ken Ishiyama[10], Hiraku Mori[11], Florian Nolte[3], Wolf-Karsten Hofmann[3], Shuichi Miyawaki[10], Sumio Sugano[7], Claudia Haferlach[12], H. Phillip Koeffler[13,14], Lee-Yung Shih[15], Torsten Haferlach[12], Shigeru Chiba[9], Hiromitsu Nakauchi[4,8], Satoru Miyano[2,6] & Seishi Ogawa[1]

Myelodysplastic syndromes and related disorders (myelodysplasia) are a heterogeneous group of myeloid neoplasms showing deregulated blood cell production with evidence of myeloid dysplasia and a predisposition to acute myeloid leukaemia, whose pathogenesis is only incompletely understood. Here we report whole-exome sequencing of 29 myelodysplasia specimens, which unexpectedly revealed novel pathway mutations involving multiple components of the RNA splicing machinery, including *U2AF35*, *ZRSR2*, *SRSF2* and *SF3B1*. In a large series analysis, these splicing pathway mutations were frequent (~45 to ~85%) in, and highly specific to, myeloid neoplasms showing features of myelodysplasia. Conspicuously, most of the mutations, which occurred in a mutually exclusive manner, affected genes involved in the 3′-splice site recognition during pre-mRNA processing, inducing abnormal RNA splicing and compromised haematopoiesis. Our results provide the first evidence indicating that genetic alterations of the major splicing components could be involved in human pathogenesis, also implicating a novel therapeutic possibility for myelodysplasia.

Myelodysplastic syndromes (MDS) and related disorders (myelodysplasia) comprise a group of myeloid neoplasms characterized by deregulated, dysplastic blood cell production and a predisposition to acute myeloid leukaemia (AML)[1]. Although the prevalence of MDS has not been determined precisely, more than 10,000 people are estimated to develop myelodysplasia annually in the United States[2]. Their indolent clinical course before leukaemic transformation and ineffective haematopoiesis with evidence of myeloid dysplasia indicate a pathogenesis distinct from that involved in *de novo* AML. Currently, a number of gene mutations and cytogenetic changes have been implicated in the pathogenesis of MDS, including mutations of *RAS*, *TP53* and *RUNX1*, and more recently *ASXL1*, *c-CBL*, *DNMT3A*, *IDH1/2*, *TET2* and *EZH2* (ref. 3). Nevertheless, mutations of this set of genes do not fully explain the pathogenesis of MDS because they are also commonly found in other myeloid malignancies and roughly 20% of MDS cases have no known genetic changes (ref. 4 and unpublished data). In particular, the genetic alterations responsible for the dysplastic phenotypes and ineffective haematopoiesis of myelodysplasia are poorly understood. Meanwhile, the recent development of massively parallel sequencing technologies has provided an expanded opportunity to discover genetic changes across the entire genomes or protein-coding sequences in human cancers at a single-nucleotide level[5–10], which could be successfully applied to the genetic analysis of myelodysplasia to obtain a better understanding of its pathogenesis.

Overview of genetic alterations

In this study, we performed whole-exome sequencing of paired tumour/control DNA from 29 patients with myelodysplasia (Supplementary Table 1). Although incapable of detecting non-coding mutations and gene rearrangements, the whole-exome approach is a well-established strategy for obtaining comprehensive registries of protein-coding mutations at low cost and high performance. With a mean coverage of 133.8, 80.4% of the target sequences were analysed at more than ×20 depth on average (Supplementary Fig. 1). All the candidates for somatic mutations ($N = 497$) generated through our data analysis pipeline were subjected to validation using Sanger sequencing (Supplementary Methods I and Supplementary Fig. 2). Finally, 268 non-synonymous somatic mutations were confirmed with an overall true positive rate of 53.9% (Supplementary Fig. 3), including 206 missense, 25 nonsense, and 10 splice site mutations, and 27 frameshift-causing insertions/deletions (indels) (Supplementary Fig. 4). The mutation rate of 9.2 (0–21) per sample was significantly lower than that in solid tumours (16.2–302)[7,11,12] and multiple myeloma (32.4)[6], but was comparable to that in AML (7.3–13)[13–15] and chronic lymphocytic leukaemia (11.5)[16]. Combined with the genomic copy number profile obtained by single nucleotide polymorphism (SNP) array karyotyping, this array of somatic mutations provided a landscape of myelodysplasia genomes (Supplementary Fig. 5)[17,18].

[1]Cancer Genomics Project, Graduate School of Medicine, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8655, Japan. [2]Laboratory of DNA Information Analysis, Human Genome Center, Institute of Medical Science, The University of Tokyo, 4-6-1 Shirokanedai, Minato-ku, Tokyo 108-8639, Japan. [3]Department of Hematology and Oncology, Medical Faculty Manheim of the University of Heidelberg, 1–3 Theodor-Kutzer-Ufer, Mannheim 68167, Germany. [4]Division of Stem Cell Therapy, Center for Stem Cell Biology and Regenerative Medicine, Institute of Medical Science, The University of Tokyo, 4-6-1 Shirokanedai, Minato-ku, Tokyo 108-8639, Japan. [5]Laboratory of Functional Genomics, Human Genome Center, Institute of Medical Science, The University of Tokyo, 4-6-1 Shirokanedai, Minato-ku, Tokyo 108-8639, Japan. [6]Laboratory of Sequence Data Analysis, Human Genome Center, Institute of Medical Science, The University of Tokyo, 4-6-1 Shirokanedai, Minato-ku, Tokyo 108-8639, Japan. [7]Division of Systems Biomedical Technology, Institute of Medical Science, The University of Tokyo, 4-6-1 Shirokanedai, Minato-ku, Tokyo 108-8639, Japan. [8]Nakauchi Stem Cell and Organ Regeneration Project, Exploratory Research for Advanced Technology, Japan Science and Technology Agency, 4-6-1 Shirokanedai, Minato-ku, Tokyo 108-8639, Japan. [9]Department of Hematology, Institute of Clinical Medicine, University of Tsukuba, 1-1-1 Tennodai, Tsukuba-shi, Ibaraki, 305-8571, Japan. [10]Division of Hematology, Tokyo Metropolitan Ohtsuka Hospital, 2-8-1 Minami-Ohtsuka, Toshima-ku, Tokyo 170-0005, Japan. [11]Division of Hematology, Internal Medicine, Showa University Fujigaoka Hospital, 1-30 Fujigaoka, Aoba-ku, Yokohama, Kanagawa 227-8501, Japan. [12]Munich Leukemia Laboratory, Max-Lebsche-Platz 31, Munich 81377, Germany. [13]Hematology/Oncology, Cedars-Sinai Medical Center, 8700 Beverly Blvd, Los Angeles, California 90048, USA. [14]National University of Singapore, Cancer Science Institute of Singapore, 28 Medical Drive, Singapore 117456, Singapore. [15]Division of Hematology-Oncology, Department of Internal Medicine, Chang Gung Memorial Hospital, Chang Gung University, 199 Tung Hwa North Rd, Taipei 105, Taiwan.
*These authors contributed equally to this work.

Novel gene targets in myelodysplasia

The list of the somatic mutations (Supplementary Table 2) included most of the known gene targets in myelodysplasia with similar mutation frequencies to those previously reported, indicating an acceptable sensitivity of the current study. The mutations of the known gene targets, however, accounted for only 12.3% of all detected mutations ($N = 33$), and the remaining 235 mutations involved previously unreported genes. Among these, recurrently mutated genes in multiple cases are candidate targets of particular interest, for which high mutation rates are expected in general populations. In fact, 8 of the 12 recurrently mutated genes were among the well-described gene targets in myelodysplasia (Supplementary Table 3). However, what immediately drew our attention were the recurrent mutations involving *U2AF35* (also known as *U2AF1*), *ZRSR2* and *SRSF2* (*SC35*), because they belong to the common pathway known as RNA splicing. Including an additional three genes mutated in single cases (*SF3A1*, *SF3B1* and *PRPF40B*), six components of the splicing machinery were mutated in 16 out of the 29 cases (55.2%) in a mutually exclusive manner (Fig. 1, Supplementary Fig. 6 and Supplementary Table 2).

Frequent mutations in splicing machinery

RNA splicing is accomplished by a well-ordered recruitment, rearrangement and/or disengagement of a set of small nuclear ribonucleoprotein (snRNP) complexes (U1, U2, and either U4/5/6 or U11/12), as well as many other protein components onto the pre-mRNAs. Notably, the mutated components of the spliceosome were all engaged in the initial steps of RNA splicing, except for PRPF40B, whose functions in RNA splicing are poorly defined. Making physical interactions with SF1 and a serine/arginine-rich (SR) protein, such as SRSF1 or SRSF2, the U2 auxiliary factor (U2AF) that consists of the U2AF65 (U2AF2)–U2AF35 heterodimer, is involved in the recognition of the 3′ splice site (3′SS) and its nearby polypyrimidine tract, which is thought to be required for the subsequent recruitment of the U2 snRNP, containing SF3A1 as well as SF3B1, to establish the splicing A complex (Fig. 1)[19]. ZRSR2 (or Urp), is another essential component of the splicing machinery. Showing a close structural similarity to U2AF35, ZRSR2 physically interacts with U2AF65, as well as SRSF1 and SRSF2, with a distinct function from its homologue, U2AF35 (ref. 20).

To confirm and extend the initial findings in the whole-exome sequencing, we studied mutations of the above six genes together with three additional spliceosome-related genes, including *U2AF65*, *SF1* and *SRSF1*, in a large series of myeloid neoplasms ($N = 582$) using a high-throughput mutation screen of pooled DNA followed by confirmation/identification of candidate mutations (refs 21 and 22 and Supplementary Methods II).

In total, 219 mutations were identified in 209 out of the 582 specimens of myeloid neoplasms through validating 313 provisional positive events in the pooled DNA screen (Supplementary Tables 4 and 5). The mutations among four genes, *U2AF35* ($N = 37$), *SRSF2* ($N = 56$), *ZRSR2* ($N = 23$) and *SF3B1* ($N = 79$), explained most of the mutations with much lower mutational rates for *SF3A1* ($N = 8$), *PRPF40B* ($N = 7$), *U2AF65* ($N = 4$) and *SF1* ($N = 5$) (Fig. 2). Mutations of the splicing machinery were highly specific to diseases showing myelodysplastic features, including MDS either with (84.9%) or without (43.9%) increased ring sideroblasts, chronic myelomonocytic leukaemia (CMML) (54.5%), and therapy-related AML or AML with myelodysplasia-related changes (25.8%), but were rare in *de novo* AML (6.6%) and myeloproliferative neoplasms (MPN) (9.4%) (Fig. 3a). The mutually exclusive pattern of the mutations in these splicing pathway genes was confirmed in this large case series, suggesting a common impact of these mutations on RNA splicing and the pathogenesis of myelodysplasia (Fig. 3b). The frequencies of mutations showed significant differences across disease types. Surprisingly, *SF3B1* mutations were found in the majority of the cases with MDS characterized by increased ring sideroblasts, that is, refractory anaemia with ring sideroblasts (RARS) (19/23 or 82.6%) and refractory cytopenia with multilineage dysplasia with ≥ 15% ring sideroblasts (RCMD-RS) (38/50 or 76%) with much lower mutation frequencies in other myeloid neoplasms. RARS and RCMD-RS account

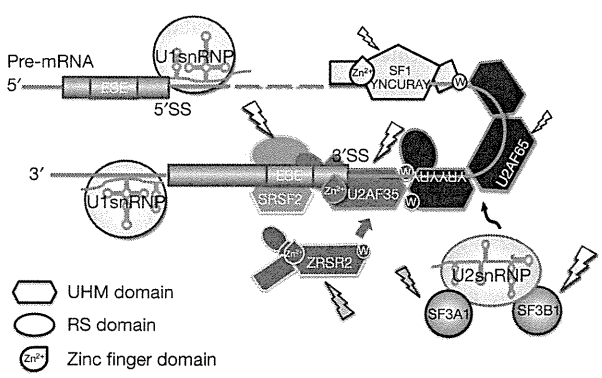


Figure 1 | Components of the splicing E/A complex mutated in myelodysplasia. RNA splicing is initiated by the recruitment of U1 snRNP to the 5′SS. SF1 and the larger subunit of the U2 auxiliary factor (U2AF), U2AF65, bind the branch point sequence (BPS) and its downstream polypyrimidine tract, respectively. The smaller subunit of U2AF (U2AF35) binds to the AG dinucleotide of the 3′SS, interacting with both U2AF65 and a SR protein, such as SRSF2, through its UHM and RS domain, comprising the earliest splicing complex (E complex). ZRSR2 also interacts with U2AF and SR proteins to perform essential functions in RNA splicing. After the recognition of the 3′SS, U2 snRNP, together with SF3A1 and SF3B1, is recruited to the 3′SS to generate the splicing complex A. The mutated components in myelodysplasia are indicated by arrows.

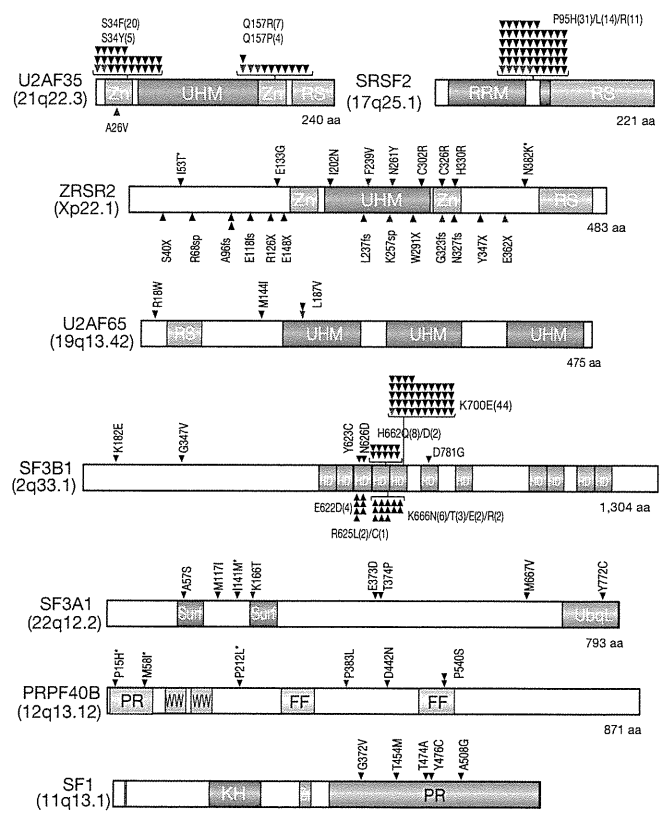


Figure 2 | Mutations of multiple components of the splicing machinery. Each mutation in the eight spliceosome components is shown with an arrowhead. Confirmed somatic mutations are discriminated by red arrows. Known domain structures are shown in coloured boxes as indicated. Mutations predicted as SNPs by MutationTaster (http://www.mutationtaster.org/) are indicated by asterisks. The number of each mutation is indicated in parenthesis. ZRSR2 mutations in females are shown in blue.

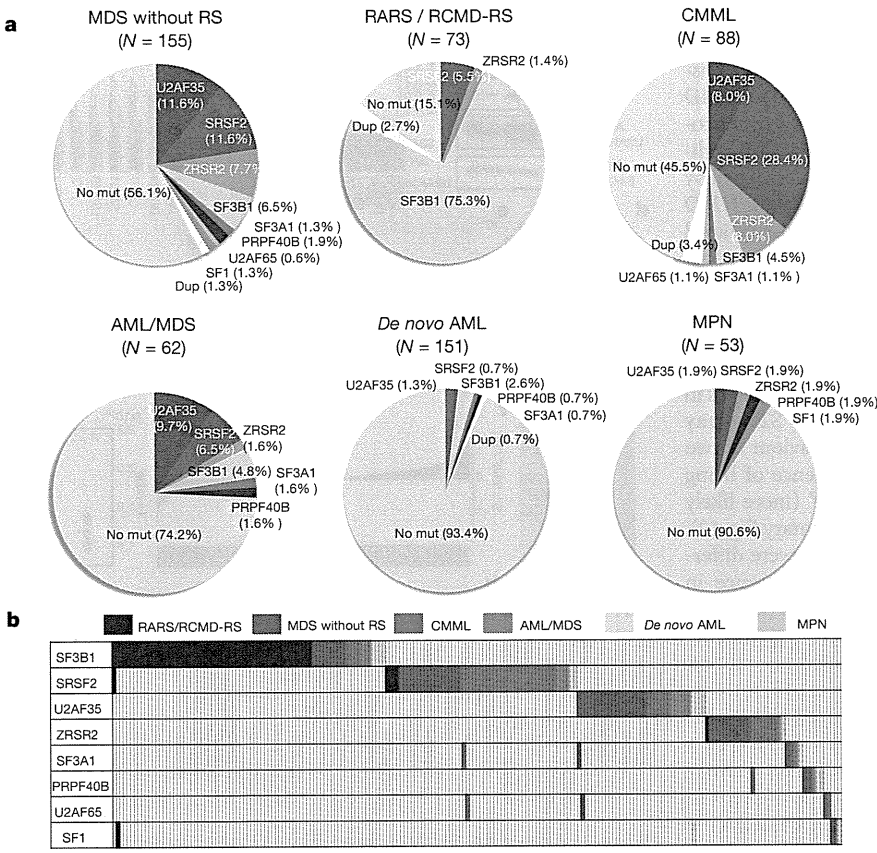


Figure 3 | Frequencies and distribution of spliceosome pathway gene mutations in myeloid neoplasms. a, Frequencies of spliceosome pathway mutations among 582 cases with various myeloid neoplasms. **b**, Distribution of mutations in eight spliceosome genes, where diagnosis of each sample is shown by indicated colours.

for 4.3% and 12.9% of MDS cases, respectively, where deregulated iron metabolism has been implicated in the development of refractory anaemia[23]. With such high mutation frequencies and specificity, the *SF3B1* mutations were thought to be almost pathognomonic to these MDS subtypes characterized by increased ring sideroblasts, and strongly implicated in the pathogenesis of MDS in these categories. Less conspicuously but significantly, *SRSF2* mutations were more frequent in CMML cases (Fig. 3 and Supplementary Table 4). Thus, although commonly involving the E/A splicing complexes, different mutations may still have different impacts on cell functions, contributing to the determination of discrete disease phenotypes. For example, studies have demonstrated that *SRSF2* was also involved in the regulation of DNA stability and that depletion of SRSF2 can lead to genomic instability[24]. Of interest in this context, regardless of disease subtypes, samples with *SRSF2* mutations were shown to have significantly more mutations of other genes compared with *U2AF35* mutations ($P = 0.001$, multiple regression analysis) (Supplementary Table 6 and Supplementary Fig. 7).

Notably, with a rare exception of A26V in a single case, the mutations of U2AF35 exclusively involved two highly conserved amino acid positions (S34 or Q157) within the amino- and the carboxyl-terminal zinc finger motifs flanking the U2AF homology motif (UHM) domain. SRSF2 mutations exclusively occurred at P95 within an intervening sequence between the RNA recognition motif (RRM) and arginine/serine-rich (RS) domains (Fig. 2 and Supplementary Figs 8 and 9). Similarly, SF3B1 mutations predominantly involved K700 and, to a lesser extent, K666, H662 and E622, which are also conserved across species (Fig. 2 and Supplementary Fig. 10). The involvement of recurrent amino acid positions in these spliceosome genes strongly indicated a gain-of-function nature of these mutations, which has been a well-documented scenario in other oncogenic mutations[25]. On the other hand, the 23 mutations in *ZRSR2* (Xp22.1) were widely distributed along the entire coding region (Fig. 2). Among these, 14 mutations were nonsense or frameshift changes, or involved splicing donor/acceptor sites that caused either a premature truncation or a large structural change of the protein, leading to loss-of-function. Combined with their strong male preference for the mutation (14/14 cases), *ZRSR2* most likely acts as a tumour suppressor gene with an X-linked recessive mode of genetic action. The remaining nine *ZRSR2* mutations were missense changes and found in both males (six cases) and females (three cases), whose somatic origin was only confirmed in two cases. However, neither the dbSNP database (build131 and 132) nor the 1000 Genomes database (May 2011 snp calls) contained these missense nucleotides, suggesting that many, if not all, of these missense changes are likely to represent functional somatic changes, especially those found in males. Interrogation of these hot spots for mutations in *U2AF35* and *SRSF2* found no mutations among lymphoid neoplasms, including acute lymphoblastic leukaemia ($N = 24$) or non-Hodgkin's lymphoma ($N = 87$) (data not shown).

RNA splicing and spliceosome mutations

Because the splicing pathway mutations in myelodysplasia widely and specifically affect the major components of the splicing complexes E/A in a mutually exclusive manner, the common consequence of these mutations is logically the impaired recognition of 3′SSs that would lead to the production of aberrantly spliced mRNA species. To appreciate this and also to gain an insight into the biological/biochemical impact of these splicing mutations, we expressed the wild-type and the mutant (S34F) U2AF35 in HeLa cells using retrovirus-mediated gene transfer with enhanced green fluorescent protein (EGFP) marking (Fig. 4a and Supplementary Methods III) and examined their effects on gene expression in these cells using GeneChip Human genome U133 plus 2.0 arrays (Affymetrix), followed by gene set enrichment analysis (GSEA) (Supplementary Methods IV)[26]. Intriguingly, the GSEA disclosed a significant enrichment of the genes on the nonsense-mediated mRNA decay (NMD) pathway among the significantly upregulated genes in mutant *U2AF35*-transduced HeLa cells (Fig. 4b, Supplementary Fig. 11a and Supplementary Table 7), which was

confirmed by quantitative polymerase chain reactions (qPCR) (Fig. 4c and Supplementary Methods 5V). A similar result was also observed for the gene expression profile of an MDS-derived cell line (TF-1) transduced with the S34F mutant (Supplementary Figs 11b, c). The NMD activation by the mutant U2AF35 was suppressed significantly by the co-overexpression of the wild-type protein (Supplementary Fig. 11d), indicating that the effect of the mutant protein was likely to be mediated by inhibition of the functions of the wild-type protein. Given that the NMD pathway, known as mRNA surveillance, provides a post-transcriptional mechanism for recognizing and eliminating abnormal transcripts that prematurely terminate translation[27], the result of the GSEA analyses indicated that the mutant U2AF35 induced abnormal RNA splicing in HeLa and TF-1 cells, leading to the generation of unspliced RNA species having a premature stop codon and induction of the NMD activity.

To confirm this, we next performed whole transcriptome analysis in these cells using the GeneChip Human exon 1.0 ST Array (Affymetrix), in which we differentially tracked the behaviour of two discrete sets of probes showing different level of evidence of being exons, that is, 'Core' (authentic exons) and 'non-Core' (more likely introns) sets (Supplementary Methods IV and Supplementary Fig. 12). As shown in Fig. 4d, the Core and non-Core set probes were differentially enriched among probes showing significant difference in expression between wild-type and mutant-transduced cells (false discovery rate (FDR) = 0.01). The Core set probes were significantly enriched in those probes significantly downregulated in mutant *U2AF35*-transduced cells compared with wild-type *U2AF35*-transduced cells, whereas the non-Core set probes were enriched in those probes significantly upregulated in mutant *U2AF35*-transduced cells (Fig. 4e). The significant differential enrichment was also demonstrated, even when all probe sets were included (Fig. 4f). Moreover, the significantly differentially expressed Core set probes tended to be up- and downregulated in wild-type and mutant *U2AF35*-transduced cells compared with mock-transduced cells, respectively, and vice versa for the differentially expressed non-Core set probes (Fig. 4e). Combined, these exon array results indicated that the wild-type U2AF35 correctly promoted authentic RNA splicing, whereas the mutant U2AF35 inhibited this processes, rendering non-Core and therefore, more likely intronic sequences to remain unspliced.

The abnormal splicing in mutant *U2AF35*-transduced cells was more directly demonstrated by sequencing mRNAs extracted from HeLa cells, in which expression of the wild-type and mutant (S34F) U2AF35 were induced by doxycycline. First, after adjusting by the total number of mapped reads, the wild-type *U2AF35*-transduced cells showed an increased read counts in the exon fraction, but reduced counts in other fractions, compared with mutant *U2AF35*-transduced cells (Fig. 4g). The reads from the mutant-transduced cells were mapped to broader genomic regions compared with those from the wild-type *U2AF35*-transduced cells, which were largely explained by non-exon reads (Fig. 4h). Finally, the number of those reads that encompassed the authentic exon/intron junctions was significantly increased in mutant *U2AF35*-transduced cells compared with wild-type *U2AF35*-transduced cells (Fig. 4i and Supplementary Methods VI). These results clearly demonstrated that failure of splicing ubiquitously occurred in mutant *U2AF35*-transduced cells. A typical example of abnormal splicing in mutant-transduced cells and the list of significantly unspliced exons are shown in Supplementary Fig. 13 and Supplementary Table 8, respectively.

Biological consequence of U2AF35 mutations

Finally, we examined the biological effects of compromised functions of the E/A splicing complexes. First, TF-1 and HeLa cells were transduced with lentivirus constructs expressing either the S34F U2AF35 mutant or wild-type U2AF35 under a tetracycline-inducible promoter (Fig. 5a and Supplementary Figs 14a and 15a), and cell proliferation was examined after the induction of their expression. Unexpectedly, after the induction of gene expression with doxycycline, the mutant *U2AF35*-transduced cells, but not the wild-type *U2AF35*-transduced cells, showed reduced cell proliferation (Fig. 5b and Supplementary Fig. 15b) with a marked increase in the G2/M fraction (G2/M arrest) together with enhanced apoptosis as

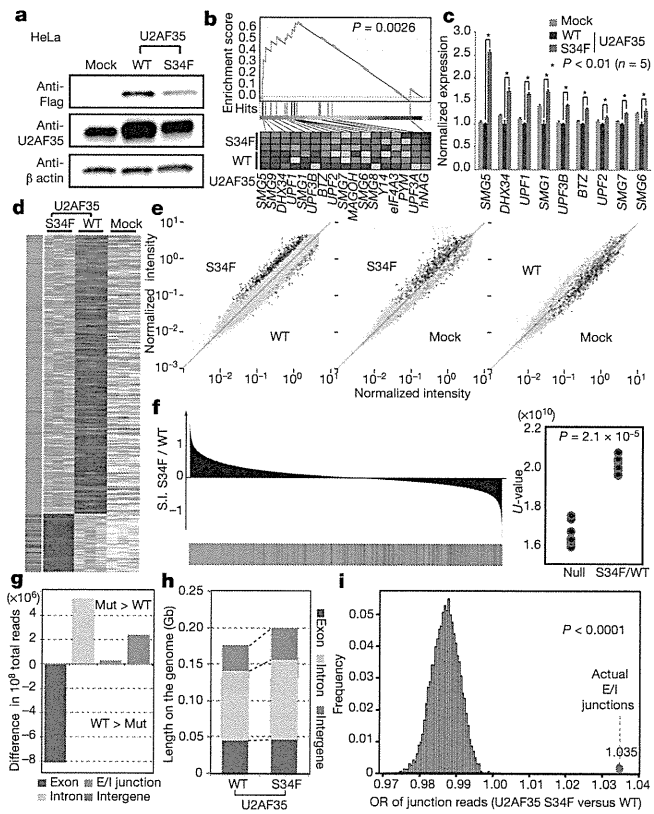


Figure 4 | Altered RNA splicing caused by a U2AF35 mutant. **a**, Western blot analyses showing expression of transduced wild-type or mutant (S34F) U2AF35 in HeLa cells used for the analyses of expression and exon microarrays. **b**, The GSEA demonstrating a significant enrichment of the set of 17 NMD pathway genes among significantly differentially expressed genes between wild-type and mutant *U2AF35*-transduced HeLa cells. The significance of the gene set was empirically determined by 1,000 gene-set permutations. **c**, The confirmation of the microarray analysis for the expression of nine genes that contributed to the core enrichment in the NMD gene set. Means ± s.e. are provided for the indicated NMD genes. *P* values were determined by the Mann–Whitney *U* test. **d**, Significantly upregulated and downregulated probe sets (FDR = 0.01) in mutant *U2AF35*-transduced cells compared with wild-type *U2AF35*-transduced cells in triplicate exon array experiments are shown in a heat map. The origin of each probe set is depicted in the left lane, where red and green bars indicate the Core and non-Core sets, respectively. **e**, Pair-wise scatter plots of the normalized intensities of entire probe sets (grey) across different experiments. The Core and non-Core set probes that were significantly differentially expressed between the wild-type and mutant *U2AF35*-transduced cells are plotted in red and green, respectively. **f**, Distribution of the Core (red) and non-Core (green) probe sets within the entire probe sets ordered by splicing index (S.I.; Supplementary Methods IV), calculated between wild-type and mutant *U2AF35*-transduced cells. In the right panel, the differential enrichment of both probe sets was confirmed by Mann–Whitney *U* test. **g**, Difference in read counts for the indicated fractions per 10^8 total reads in RNA sequencing between wild-type and mutant U2AF35-expressing HeLa cells analysis. Increased/decreased read counts in mutant U2AF35-expressing cells are plotted upward/downward, respectively. **h**, Comparison of the genome coverage by the indicated fractions in wild-type- and mutant-U2AF35-expressing cells. The genome coverage was calculated for each fraction within the 10^8 reads randomly selected from the total reads and averaged for ten independent selections. **i**, The odds ratio of the junction reads within the total mapped reads was calculated between the two experiments (red circle), which was evaluated against the 10,000 simulated values under the null hypothesis (histogram in blue).

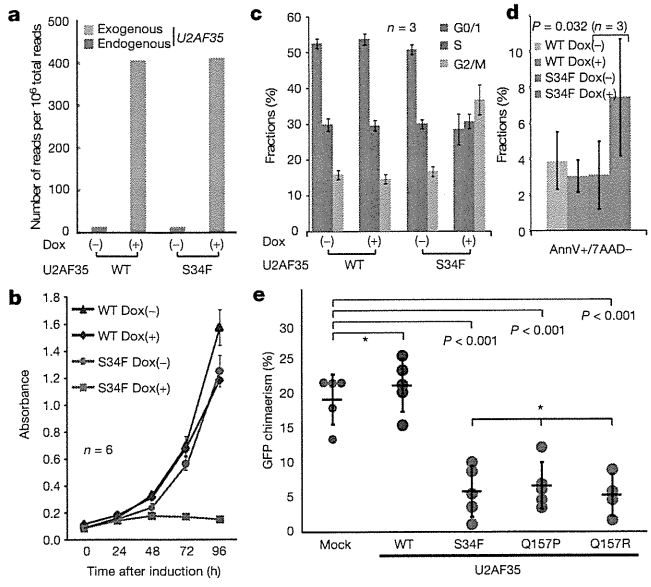


Figure 5 | Functional analysis of mutant U2AF35. a, Expression of endogenous and exogenous *U2AF35* transcripts in HeLa cells before and after induction determined by RNA sequencing. *U2AF35* transcripts were differentially enumerated for endogenous and exogenous species, which were discriminated by the Flag sequence. **b**, Cell proliferation assays of *U2AF35*-transduced HeLa cells, where cell numbers were measured using cell-counting apparatus and are plotted as mean absorbance ± s.d. **c**, The flow cytometry analysis of propidium iodide (PI)-stained HeLa cells transduced with the different U2AF35 constructs. Mean fractions ± s.d. in G0/G1, S and G2/M populations after the induction of U2AF35 expression are plotted. **d**, Fractions of the annexin V-positive (AnnV+) populations among the 7-amino-actinomycin D (7AAD)-negative population before and after the induction of U2AF35 expression are plotted as mean ± s.d. for indicated samples. The significance of difference was determined by paired *t*-test. **e**, Competitive reconstitution assays for CD34-negative KSL cells transduced with indicated U2AF35 mutants. Chimaerism in the peripheral blood 6 weeks after transplantation are plotted as mean %EGFP-positive Ly5.1 cells ± s.d., where outliers were excluded from the analysis. The significance of differences was evaluated by the Grubbs test with Bonferroni's correction for multiple testing. *not significant.

indicated by the increased sub-G1 fraction and annexin V-positive cells (Fig. 5c, d, Supplementary Fig. 14b and Supplementary Methods VI). To confirm the growth-suppressive effect of U2AF35 mutants *in vitro*, a highly purified haematopoietic stem cell population ($CD34^-$c-$Kit^+Scal^+Lin^-$, $CD34^-$KSL) prepared from C57BL/6 (B6)-Ly5.1 mouse bone marrow[28] was retrovirally transduced with either the mutant (S34F, Q157P and Q157R) or wild-type *U2AF35*, or the mock constructs, each harbouring the EGFP marker gene (Supplementary Fig. 16). The ability of these transduced cells to reconstitute the haematopoietic system was tested in a competitive reconstitution assay. The transduced cells were mixed with whole bone marrow cells from B6-Ly5.1/5.2 F1 mice, transplanted into lethally irradiated B6-Ly5.2 recipients, and peripheral blood chimaerism derived from EGFP-positive cells was assessed 6 weeks after transplantation by flow cytometry. We confirmed that each recipient mouse received comparable numbers of EGFP-positive cells among the different retrovirus groups by estimating the percentage of EGFP-positive cells and overall proliferation in transduced cells by *ex vivo* tracking. Also no significant difference was observed in their homing capacity to bone marrow as assessed by transwell migration assays (Supplementary Fig. 17). As shown in Fig. 5e, the wild-type *U2AF35*-transduced cells showed a slightly higher reconstitution capacity than the mock-transduced cells. On the other hand, the recipients of the cells transduced with the various *U2AF35* mutants showed significantly lower EGFP-positive cell chimaerism than those of either the mock- or the wild-type *U2AF35*-transduced cells, indicating a compromised reconstitution capacity of the haematopoietic stem/progenitor cells expressing the U2AF35 mutants. In summary, these mutants lead to loss-of-function of U2AF35 most probably by acting in a dominant-negative fashion to the wild-type protein.

Discussion

Our whole-exome sequencing study unexpectedly unmasked a complexity of novel pathway mutations found in approximately 45% to 85% of myelodysplasia patients depending on the disease subtypes, which affected multiple but distinctive components of the splicing machinery and, as such, demonstrated the unquestionable power of massively parallel sequencing technologies in cancer research.

The RNA splicing system comprises essential cellular machinery, through which eukaryotes can achieve successful transcription and guarantee the functional diversity of their protein species using alternative splicing in the face of a limited number of genes[29]. Accordingly, the meticulous regulation of this machinery should be indispensable for the maintenance of cellular homeostasis[30], deregulation of which causes severe developmental abnormalities[31,32]. The current discovery of frequent mutations of the splicing pathway in myelodysplasia, therefore, represents another remarkable example that illustrates how cancer develops by targeting critical cellular functions. It also provides an intriguing insight into the mechanism of 'cancer specific' alternative splicing, which have long been implicated in the development of cancer, including MDS and other haematopoietic neoplasms[33,34].

In myelodysplasia, the major targets of spliceosome mutations seemed to be largely confined to the components of the E/A splicing complex, among others to *SF3B1*, *SRSF2*, *U2AF35* and *ZRSR2*, and to a lesser extent, to *SF3A1*, *SF1*, *U2AF65* and *PRPF40B*. The broad coverage of the wide spectrum of spliceosome components in our exome sequencing was likely to preclude frequent involvement of other components on this pathway (Supplementary Fig. 18). The surprising frequency and specificity of these mutations in this complex, together with the mutually exclusive manner they occurred, unequivocally indicate that the compromised function of the E/A complex is a hallmark of this unique category of myeloid neoplasms, playing a central role in the pathogenesis of myelodysplasia. The close relationship between the mutation types and unique disease subtypes also support their pivotal roles in MDS.

Given the critical functions of the E/A splicing complex on the precise 3′SS recognition, the logical consequence of these relevant mutations would be the impaired splicing involving diverse RNA species. In fact, when expressed in HeLa cells, the mutant U2AF35 induced global abnormalities of RNA splicing, leading to increased production of transcripts having unspliced intronic sequences. On the other hand, the functional link between the abnormal splicing of RNA species and the phenotype of myelodysplasia is still unclear. Mutant U2AF35 seemed to suppress cell growth/proliferation and induce apoptosis rather than confer a growth advantage or promote clonal selection. *ZRSR2* knockdown in HeLa cells has been reported to also result in reduced viability, arguing for the common consequence of these pathway mutations[35]. These observations suggested that the oncogenic actions of these splicing pathway mutations are distinct from what is expected for classical oncogenes, such as mutated kinases and signal transducers, but could be more related to cell differentiation. Of note in this regard, the commonest clinical presentation of MDS is severe cytopenia in multiple cell lineages due to ineffective haematopoiesis with increased apoptosis rather than unlimited cell proliferation[1]. In this regard, lessons may be learned from the recent findings on the pathogenesis of the 5q− syndrome, where haploinsufficiency of *RPS14* leads to increased apoptosis of erythroid progenitors, but not myeloproliferation[36,37].

A lot of issues remain to be answered, however, to establish the functional link between these splicing pathway mutations and the

pathogenesis of MDS, where the broad spectrum of RNA species affected by impaired splicing hampers identification of responsible gene targets. Moreover, the mutated components of the splicing machinery have distinct function of their own other than direct regulation of RNA splicing, involved in elongation and DNA stability, which may be important to determine specific disease phenotypes. Clearly, more studies are required to answer these questions through understanding of the molecular basis of their oncogenic actions.

METHODS SUMMARY

Whole-exome sequencing of paired tumour/normal DNA samples from the 29 patients was performed after informed consent was obtained. SNP array-based copy number analysis was performed as previously described[17,18]. Mutation analysis of the splicing pathway genes in a set of 582 myeloid neoplasms were performed by first screening mutations in PCR-amplified pooled targets from 12 individuals, followed by validation/identification of the candidate mutations within the corresponding 12 individuals by Sanger sequencing. Flag-tagged cDNAs of the wild-type and mutant *U2AF35* were generated by *in vitro* mutagenesis, constructed into a murine stem cell virus-based retroviral vector as well as a tetracycline-inducible lentivirus-based expression vector, and used for gene transfer to $CD34^{-}$KSL cells and cultured cell lines, with EGFP marking, respectively. Total RNA was extracted from wild-type or mutant *U2AF35*-transduced HeLa and TF-1 cells, and analysed on microarrays. RNA sequencing was performed according to the manufacturer's instructions (Illumina). Cell proliferation assays (MTT assays) on HeLa and TF-1 cells stably transduced with lentivirus U2AF35 constructs were performed in the presence or absence of doxycycline. For competitive reconstitution assays, $CD34^{-}$KSL cells collected from C57BL/6 (B6)-Ly5.1 mice were retrovirally transduced with various *U2AF35* constructs with EGFP marking, and transplanted with competitor cells (B6-Ly5.1/5.2 F1 mouse origin) into lethally irradiated B6-Ly5.2 mice 48 h after gene transduction. Frequency of EGFP-positive cells was assessed in peripheral blood by flow cytometry 6 weeks after the transplantation (Supplementary Methods VII). The primer sets used for validation of gene mutations and qPCR of NMD gene expression are listed in Supplementary Tables 9–11. A complete description of the materials and methods is provided in the Supplementary Information. This study was approved by the ethics boards of the University of Tokyo, Munich Leukaemia Laboratory, University Hospital Mannheim, University of Tsukuba, Tokyo Metropolitan Ohtsuka Hospital and Chang Gung Memorial Hospital. Animal experiments were performed with approval of the Animal Experiment Committee of the University of Tokyo.

Received 7 June; accepted 24 August 2011.
Published online 11 September 2011.

1. Corey, S. J. *et al.* Myelodysplastic syndromes: the complexity of stem-cell diseases. *Nature Rev. Cancer* **7,** 118–129 (2007).
2. Ma, X., Does, M., Raza, A. & Mayne, S. T. Myelodysplastic syndromes: incidence and survival in the United States. *Cancer* **109,** 1536–1542 (2007).
3. Bejar, R., Levine, R. & Ebert, B. L. Unraveling the molecular pathophysiology of myelodysplastic syndromes. *J. Clin. Oncol.* **29,** 504–515 (2011).
4. Sanada, M. *et al.* Gain-of-function of mutated *C-CBL* tumour suppressor in myeloid neoplasms. *Nature* **460,** 904–908 (2009).
5. Campbell, P. J. *et al.* Identification of somatically acquired rearrangements in cancer using genome-wide massively parallel paired-end sequencing. *Nature Genet.* **40,** 722–729 (2008).
6. Chapman, M. A. *et al.* Initial genome sequencing and analysis of multiple myeloma. *Nature* **471,** 467–472 (2011).
7. Lee, W. *et al.* The mutation spectrum revealed by paired genome sequences from a lung cancer patient. *Nature* **465,** 473–477 (2010).
8. Ley, T. J. *et al.* DNA sequencing of a cytogenetically normal acute myeloid leukaemia genome. *Nature* **456,** 66–72 (2008).
9. Metzker, M. L. Sequencing technologies — the next generation. *Nature Rev. Genet.* **11,** 31–46 (2010).
10. Shendure, J. & Ji, H. Next-generation DNA sequencing. *Nature Biotechnol.* **26,** 1135–1145 (2008).
11. Shah, S. P. *et al.* Mutational evolution in a lobular breast tumour profiled at single nucleotide resolution. *Nature* **461,** 809–813 (2009).
12. Varela, I. *et al.* Exome sequencing identifies frequent mutation of the SWI/SNF complex gene *PBRM1* in renal carcinoma. *Nature* **469,** 539–542 (2011).
13. Ley, T. J. *et al. DNMT3A* mutations in acute myeloid leukemia. *N. Engl. J. Med.* **363,** 2424–2433 (2010).
14. Mardis, E. R. *et al.* Recurring mutations found by sequencing an acute myeloid leukemia genome. *N. Engl. J. Med.* **361,** 1058–1066 (2009).
15. Yan, X. J. *et al.* Exome sequencing identifies somatic mutations of DNA methyltransferase gene *DNMT3A* in acute monocytic leukemia. *Nature Genet.* **43,** 309–315 (2011).
16. Puente, X. S. *et al.* Whole-genome sequencing identifies recurrent mutations in chronic lymphocytic leukaemia. *Nature* **475,** 101–105 (2011).
17. Nannya, Y. *et al.* A robust algorithm for copy number detection using high-density oligonucleotide single nucleotide polymorphism genotyping arrays. *Cancer Res.* **65,** 6071–6079 (2005).
18. Yamamoto, G. *et al.* Highly sensitive method for genomewide detection of allelic composition in nonpaired, primary tumor specimens by use of Affymetrix single-nucleotide-polymorphism genotyping microarrays. *Am. J. Hum. Genet.* **81,** 114–126 (2007).
19. Wahl, M. C., Will, C. L. & Luhrmann, R. The spliceosome: design principles of a dynamic RNP machine. *Cell* **136,** 701–718 (2009).
20. Tronchère, H., Wang, J. & Fu, X. D. A protein related to splicing factor $U2AF^{35}$ that interacts with $U2AF^{65}$ and SR proteins in splicing of pre-mRNA. *Nature* **388,** 397–400 (1997).
21. Bevilacqua, L. *et al.* A population-specific *HTR2B* stop codon predisposes to severe impulsivity. *Nature* **468,** 1061–1066 (2010).
22. Calvo, S. E. *et al.* High-throughput, pooled sequencing identifies mutations in *NUBPL* and *FOXRED1* in human complex I deficiency. *Nature Genet.* **42,** 851–858 (2010).
23. Haase, D. *et al.* New insights into the prognostic impact of the karyotype in MDS and correlation with subtypes: evidence from a core dataset of 2124 patients. *Blood* **110,** 4385–4395 (2007).
24. Xiao, R. *et al.* Splicing regulator SC35 is essential for genomic stability and cell proliferation during mammalian organogenesis. *Mol. Cell. Biol.* **27,** 5393–5402 (2007).
25. Morin, R. D. *et al.* Somatic mutations altering EZH2 (Tyr641) in follicular and diffuse large B-cell lymphomas of germinal-center origin. *Nature Genet.* **42,** 181–185 (2010).
26. Subramanian, A. *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl Acad. Sci. USA* **102,** 15545–15550 (2005).
27. Maquat, L. E. Nonsense-mediated mRNA decay: splicing, translation and mRNP dynamics. *Nature Rev. Mol. Cell Biol.* **5,** 89–99 (2004).
28. Ema, H. *et al.* Adult mouse hematopoietic stem cells: purification and single-cell assays. *Nature Protocols* **1,** 2979–2987 (2007).
29. Chen, M. & Manley, J. L. Mechanisms of alternative splicing regulation: insights from molecular and genomics approaches. *Nature Rev. Mol. Cell Biol.* **10,** 741–754 (2009).
30. Ni, J. Z. *et al.* Ultraconserved elements are associated with homeostatic control of splicing regulators by alternative splicing and nonsense-mediated decay. *Genes Dev.* **21,** 708–718 (2007).
31. He, H. *et al.* Mutations in U4atac snRNA, a component of the minor spliceosome, in the developmental disorder MOPD I. *Science* **332,** 238–240 (2011).
32. Edery, P. *et al.* Association of TALS developmental disorder with defect in minor splicing component *U4atac* snRNA. *Science* **332,** 240–243 (2011).
33. David, C. J. & Manley, J. L. Alternative pre-mRNA splicing regulation in cancer: pathways and programs unhinged. *Genes Dev.* **24,** 2343–2364 (2010).
34. Pajares, M. J. *et al.* Alternative splicing: an emerging topic in molecular and clinical oncology. *Lancet Oncol.* **8,** 349–357 (2007).
35. Shen, H., Zheng, X., Luecke, S. & Green, M. R. The U2AF35-related protein Urp contacts the 3′ splice site to promote U12-type intron splicing and the second step of U2-type intron splicing. *Genes Dev.* **24,** 2389–2394 (2010).
36. Barlow, J. L. *et al.* A p53-dependent mechanism underlies macrocytic anemia in a mouse model of human 5q− syndrome. *Nature Med.* **16,** 59–66 (2010).
37. Ebert, B. L. *et al.* Identification of *RPS14* as a $5q^{-}$ syndrome gene by RNA interference screen. *Nature* **451,** 335–339 (2008).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements This work was supported by Grant-in-Aids from the Ministry of Health, Labor and Welfare of Japan and from the Ministry of Education, Culture, Sports, Science and Technology, and also by the Japan Society for the Promotion of Science (JSPS) through the 'Funding Program for World-Leading Innovative R&D on Science and Technology (FIRST Program)', initiated by the Council for Science and Technology Policy (CSTP). pGCDNsamIRESEGFP vector was a gift from M. Onodera. We thank Y. Mori, O. Hagiwara, M. Nakamura and N. Mizota for their technical assistance. We are also grateful to K. Ikeuchi and M. Ueda for their continuous encouragement throughout the study.

Author Contributions Y.Sh., Y.Sa., A.S.-O., Y.N., M.N., G.C., R.K. and S.Miyano were committed to bioinformatics analyses of resequencing data. M.Sa., A.S.-O. and Y.Sa. performed microarray experiments and their analyses. R.Y., T.Y., M.O., M.Sa., A.K., M.Sh. and H.N. were involved in the functional analyses of U2AF35 mutants. N.O., M.S.-Y., K.I., H.M., W.-K.H., F.N., D.N., T.H., C.H., S.Miyawaki, S.C., H.P.K. and L.-Y.S. collected specimens and were also involved in planning the project. K.Y., Y.N., Y.Su., A.S.-O. and S.S. processed and analysed genetic materials, library preparation and sequencing. K.Y., M.Sa., Y.Sh., A.S.-O., Y. Sa. and S.O. generated figures and tables. S.O. led the entire project and wrote the manuscript. All authors participated in the discussion and interpretation of the data and the results.

Author Information Sequence data have been deposited in the DDBJ repository under accession number DRA000433. Microarray data have been deposited in the GEO database under accession numbers GSE31174 (for SNP arrays), GSE31171 (for exon arrays) and GSE31172 (for expression arrays). Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to S.O. (sogawa-tky@umin.ac.jp).

Open

Leukemia (2012) 1 - 3

www.nature.com/leu

LETTER TO THE EDITOR

Novel splicing-factor mutations in juvenile myelomonocytic leukemia

Leukemia advance online publication, 2 March 2012;
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).

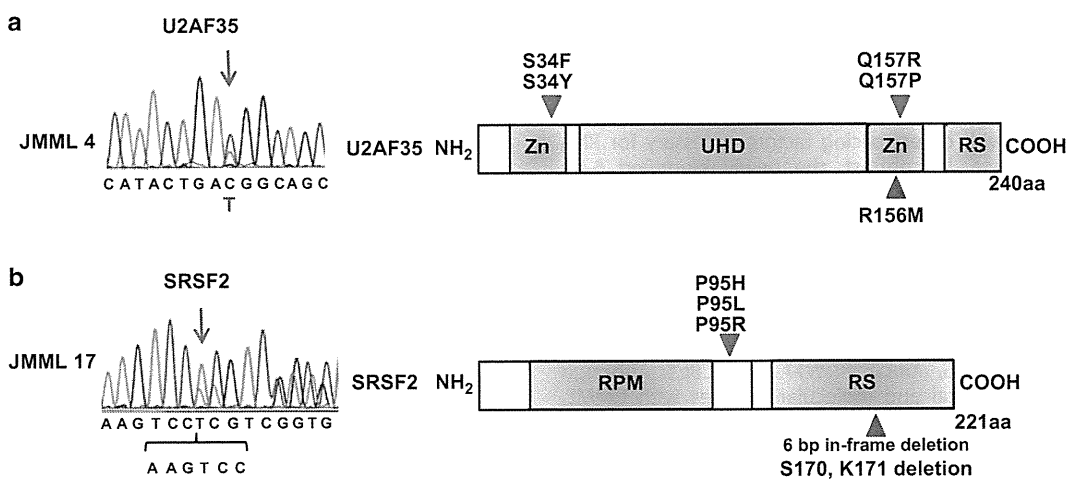


Figure 1. Novel *U2AF35*and *SRSF2* mutations detected in JMML cases. (**a**) Left panel: sequence chromatogram of a heterozygous mutation at R156 in N-terminal zinc-finger motifs of *U2AF35* detected in a JMML case (JMML 4) is shown. Mutated nucleotides are indicated by arrows. Right panel: illustration 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

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)

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].

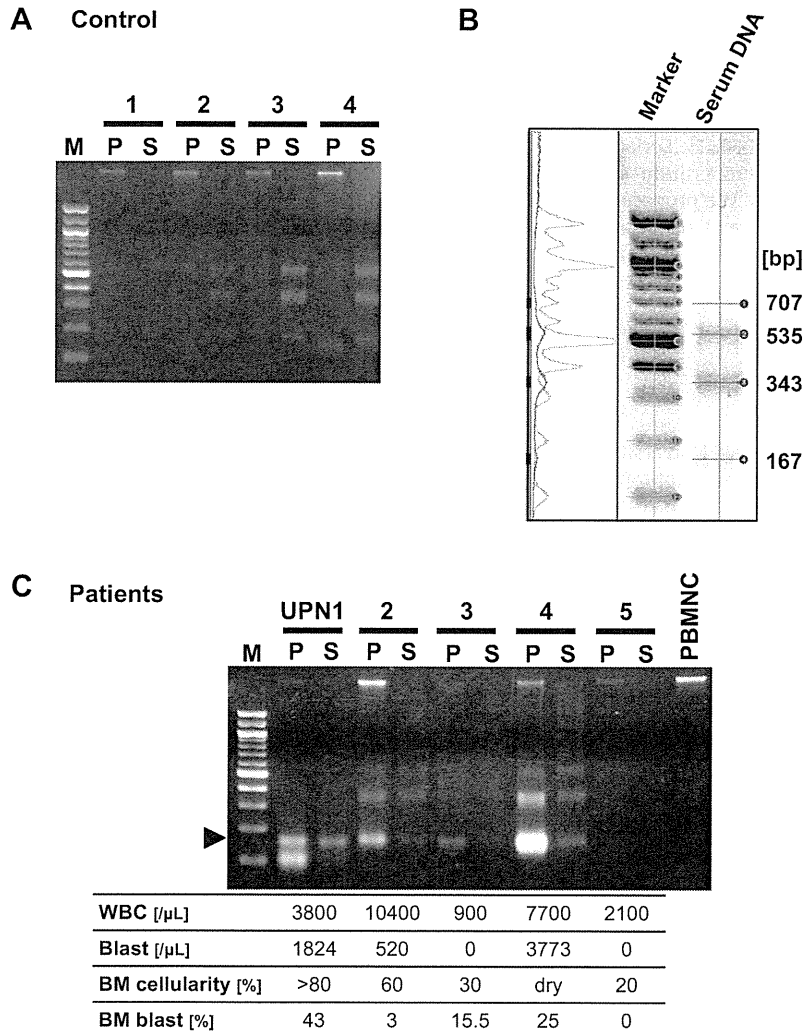
2.8. Statistical analysis

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

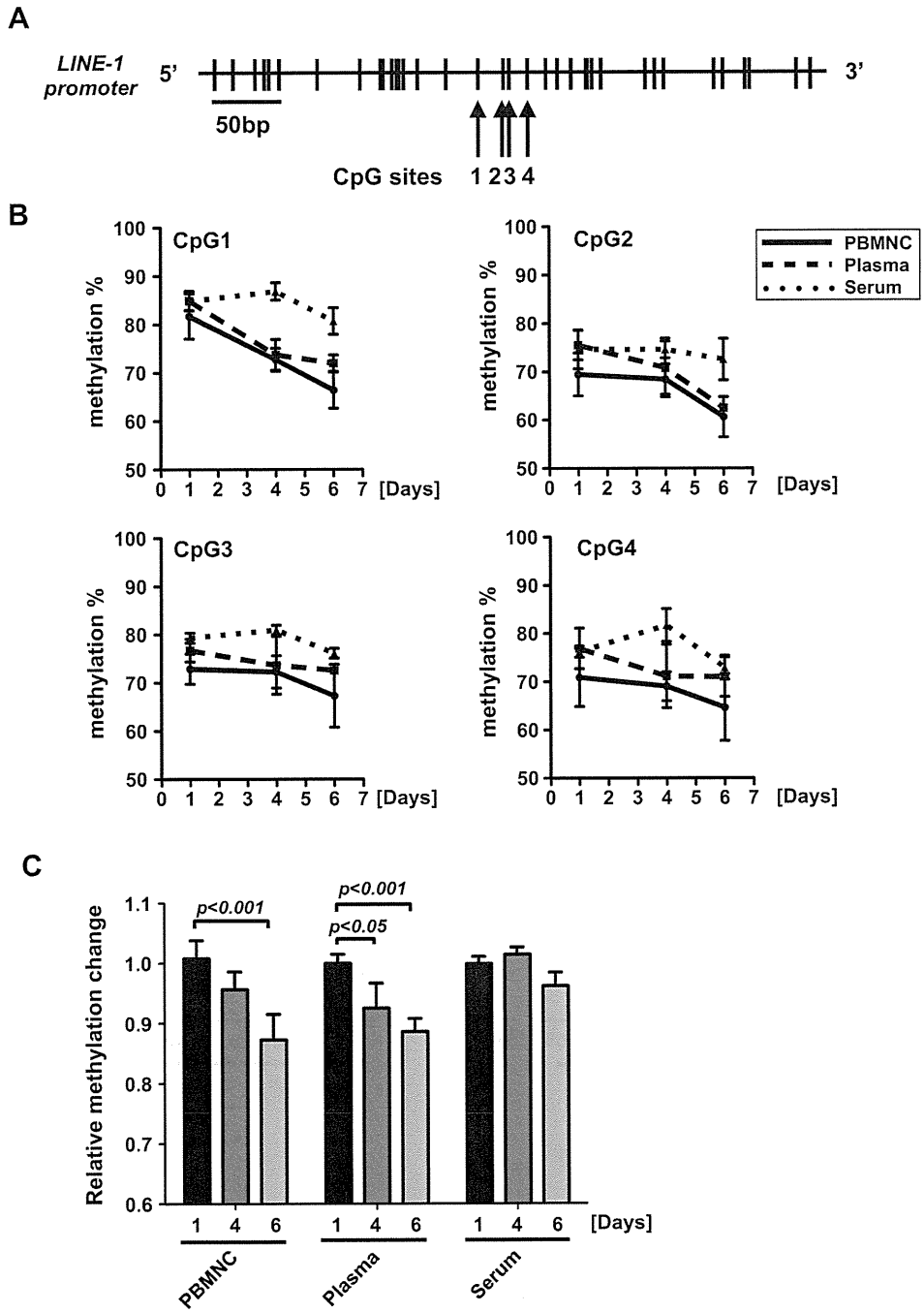


Fig. 2. *LINE-1* promoter CpG methylation status in a MDS patient treated with hypomethylating agent can be measured by using circulating DNAs. (A) Schematic representation of the *LINE-1* promoter CpG sites (GenBank; X58075, nucleotide position 108–520, complementary strand). Four CpG sites (1–4) were selected to measure cytosine methylation status by bisulfite pyrosequencing. (B) Peripheral blood was obtained from MDS UPN2 during the first azacitidine treatment cycle at days 1, 4, and 6. DNAs were prepared from plasma, serum, and PBMNC, and the methylation percentage of the four sites (CpG1–4) was quantitated by pyrosequencing analysis. The percentage of the methylation generally decreased after treatment with azacitidine. (C) The changing ratio of the methylation of all four sites was calculated using DNAs from PBMNC, plasma and serum, and the mean changing ratio was indicated in the bar graphs with the mean standard deviation (S.D.). The *p*-value was also indicated if the differences were significant.

plasma at day 4. In PBMNC, the tendency toward a decrease of the methylation ratio can be seen also at day 4, but was not significant. These data suggest that circulating DNA from plasma can be used for global methylation analysis with analysis of the *LINE-1* promoter as an alternative strategy using MNC in peripheral blood and/or BM.

3.3. Longer follow up of the global methylation status after treatment with a demethylating agent using plasma and serum circulating DNAs

Almost the same analysis was performed using peripheral blood samples from UPN4. PB was harvested 7 and 5 times during azacitidine treatment cycles 1 and 2 (Fig. 3A) and analyzed. The global

methylation ratio at the four CpG sites of the *LINE-1* elements generally decreased after starting treatment until day 9, and then increased again until the next treatment was started (days 12–28). To determine which samples of DNA were suitable for detection of the changing ratio after administration of azacitidine, the average methylation percentage of those four CpG sites was calculated using the data from days 1 to 9 in the first azacitidine course (Fig. 3B). In this assay, all DNA samples from PBMNC, plasma, and serum could detect the significant change of methylation status after azacitidine treatment. In particular, in this assay, plasma circulating DNA could detect the change much earlier at day 3 than DNA from PBMNC and serum. These data suggest that repeated sample collection from MDS patients can provide more accurate information about methylation status, and the methylation ratio of the *LINE-1* elements can be measured by using DNA circulating in the peripheral blood. Furthermore, taken together with the data of DNA concentration as shown in Fig. 1C, it is suggested that the DNA in plasma may be much more sensitive for detecting the change of methylation status in the blast cells in MDS patients.

3.4. Detection of genetic mutation using peripheral blood circulating DNAs

To test whether circulating DNAs can also be used for the detection of genetic mutations in MDS cells, mutation analysis was performed using plasma and serum circulating DNAs. BM cells from UPN1, showing 4q uni-parental disomy (UPD) by a SNP array (Fig. 4A), were utilized for genetic mutation analysis for the *TET2* gene. A non-sense point mutation of the *TET2* gene in exon 6, resulting in C-terminus truncation at the cysteine rich domain (CRD), was confirmed (Fig. 4B and C). Next, we performed pyrosequencing analysis to show the existence ratio of the *TET2* mutation

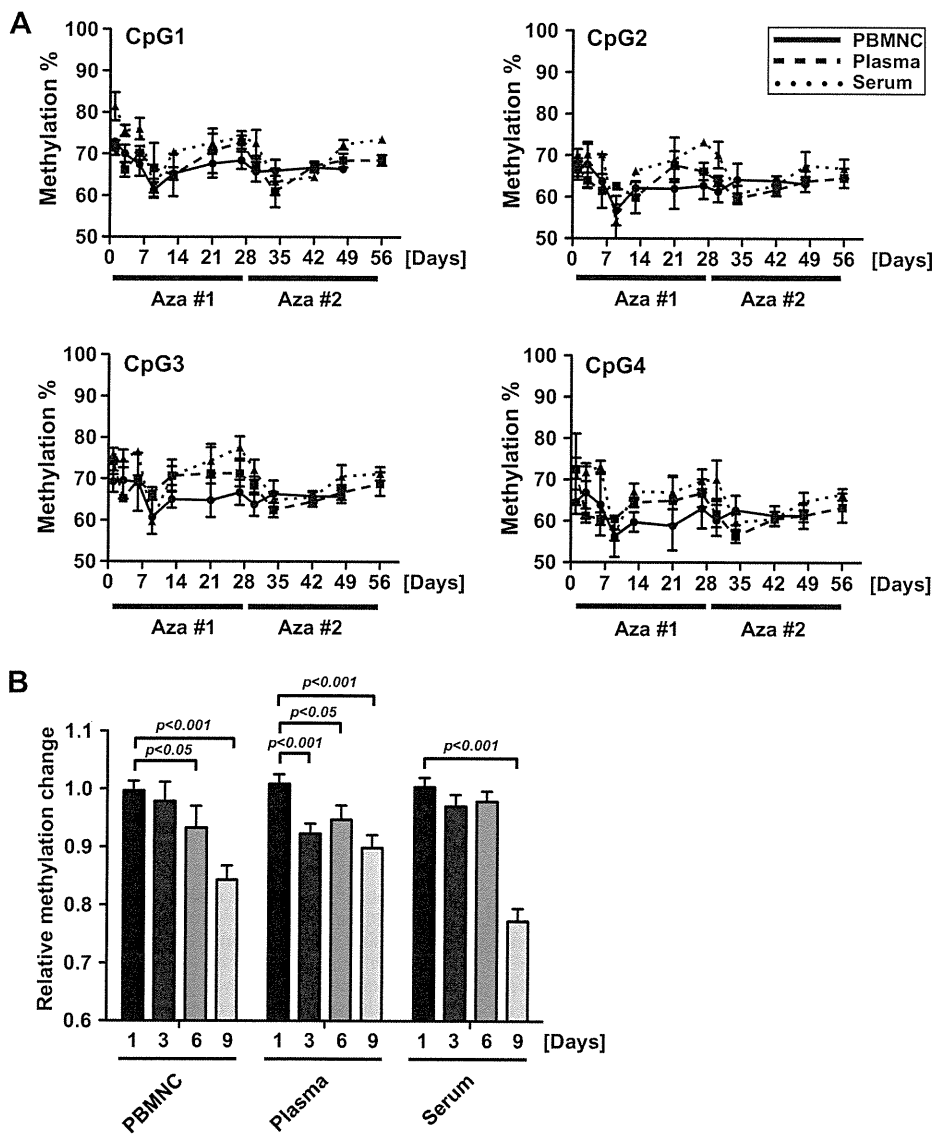


Fig. 3. A longer follow up of the *LINE-1* promoter CpG methylation status in an MDS patient treated with a hypomethylating agent using circulating DNAs. (A) Peripheral blood was obtained from MDS UPN4 during the first (#1) and the second (#2) azacitidine treatment cycle as indicated. DNAs were prepared from plasma, serum, and PBMNC, and the methylation percentage of the four sites (CpG1–4) was quantitated as analyzed in Fig. 2. Note that the methylation percentage was generally decreased until day 9, and the ratio was reversed until the next azacitidine treatment was started. (B) The mean changing ratio of the methylation of all four sites was calculated using DNAs from PBMNC, plasma, and serum, as analyzed in Fig. 2C.

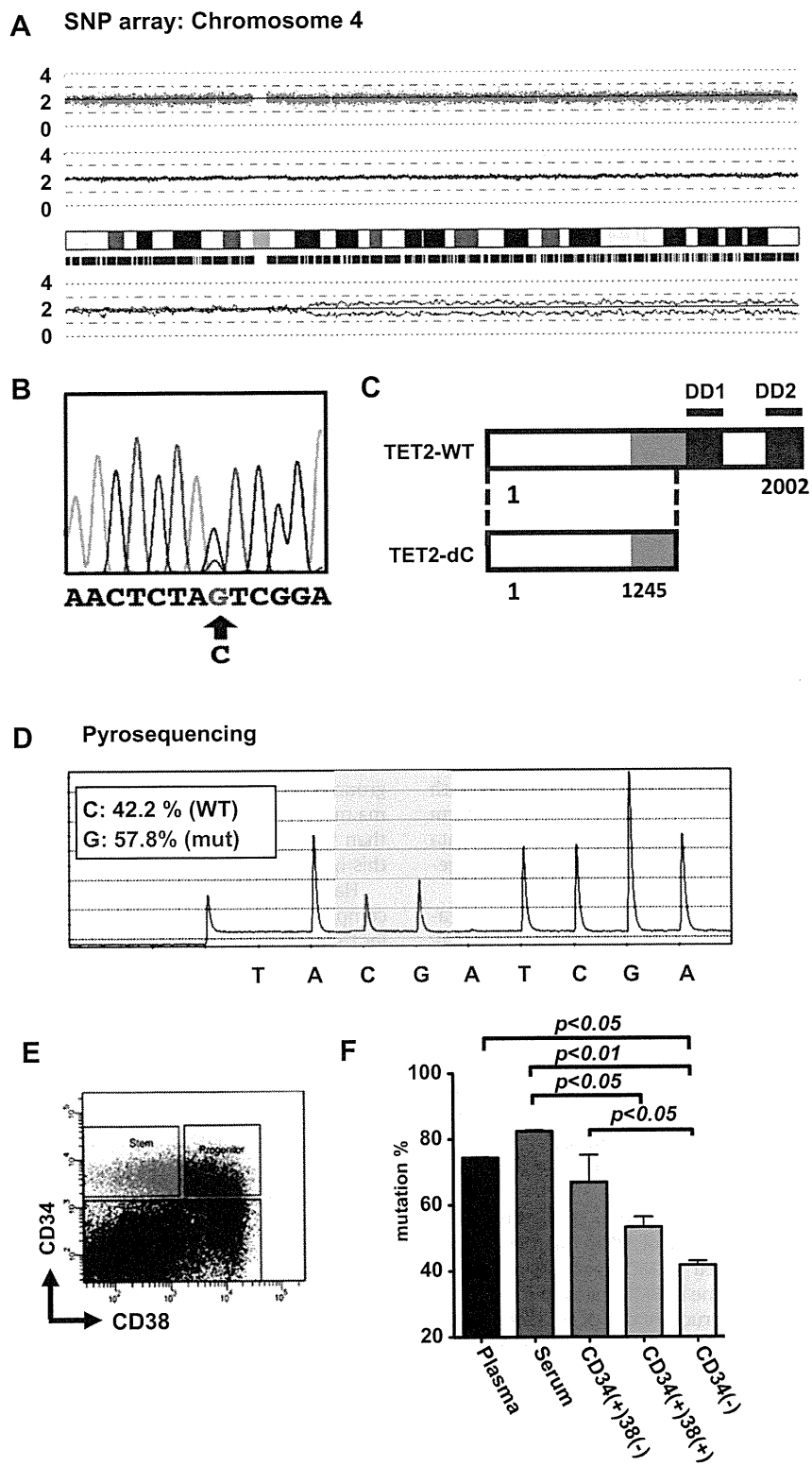


Fig. 4. Detection of the genetic mutations in an MDS patient using peripheral blood circulating DNAs. (A) SNP array analysis of UPN1. The DNA copy numbers were indicated on the green and red lines. UPD was observed in chromosome 4q. (B) Non-sense point mutation in the *TET2* gene exon 6, from TAC (tyrosine) to TAG (stop), was confirmed by the dye-terminator method. (C) Functional domains of wild-type TET2 and the putative C-terminally truncated TET2 protein (TET2-dC) in UPN1. CRD: cysteine rich domain, DD: dioxygenase domain. (D) Pyrosequencing analysis for the mutated region of the *TET2* gene using BM cells. The existence percentage of the wild-type (blue) and mutated (red) nucleotide was calculated by analyzing software. (E) BM cells of UPN1 were sorted into three subpopulations: CD34+/CD38− (stem), CD34+/CD38+(progenitor), and CD34− (others). (F) The existence ratio of the *TET2* mutation in each DNA prepared from circulating DNA (plasma and serum) and BM cell subpopulations (CD34+/CD38−, CD34+/CD38+, and CD34−).

using genomic DNA from whole BM cells of this patient. This mutation was detected by this assay indicating 57.8% of mutated clones. Since the mutation can be observed in both alleles in one cell (UPD), it was suggested that the wild type sequence came from the normal component of the patient.

To confirm whether the *TET2* mutation can also be detected in circulating DNA, and if so, whether the existing ratio of the mutation in circulating DNAs is much higher than that in components of BM cells, especially in stem cell population, three sorted BM cell populations (stem: CD34+/38−, progenitors: CD34+/38+, and others: CD34−) were utilized for the *TET2* mutation analysis using the pyrosequencing strategy. BM and peripheral blood aspiration from this patient were carried out in the same day. DNAs from plasma, serum, and the three populations of BM cells (Fig. 4E) were prepared, and pyrosequencing analysis was performed to confirm the existence ratio of the mutated *TET2* gene. The result was that the existence ratio of the mutants was significantly higher in serum and plasma than that in CD34-negative BM cells (Fig. 4F). These results indicated that circulating DNAs from plasma and serum can be used for the detection of genetic mutations in MDS cells. Furthermore, these data also suggest that plasma and serum DNA may likely come from the MDS clones distributed in these three populations, especially in CD34+/38− population, which might be relatively more fragile than normal cells. Further investigation is required to prove this hypothesis.

4. Discussion

The aim of this study was to check whether peripheral blood circulating DNA is useful for the detection of disease specific genetic and epigenetic changes in patients with MDS. Because MDS usually shows relatively slow disease progression compared with acute leukemia, patients and clinicians tend to hesitate to perform BM aspiration repeatedly to estimate the disease status. Our data suggest that plasma and serum DNAs are able to be used for the genetic and epigenetic assays instead of BM cells.

Interestingly, our data suggest that the plasma DNA concentration tends to reflect BM blast counts (Fig. 1C). This finding may indicate that the PB circulating DNA may come from relatively fragile MDS cells rather than normal cells. If so, there is a possibility that abnormal DNAs from the MDS clone may be much more enriched in the circulating DNA compared with DNAs from BM or PBMNC. This hypothesis may also explain two of our results: (1) the methylation status can be decreased more rapidly in plasma DNA than in PBMNC, as shown in Figs. 2C and 3B, (2) the existence ratio of the genetic mutation in plasma and serum DNA is relatively higher than that in the BM cells. A large number of patients are needed to confirm this phenomenon.

As shown in Fig. 1C, a higher concentration of digested DNA fragments, especially the accumulation of DNA fragments from mono-nucleosomes, can be observed in MDS patients who may have a relatively higher tumor volume in BM. Since the fragmentation of the DNA in the chromatin structure can occur through the enzymatic activity of deoxyribo-nuclease I (DNase I), a higher degree of enzymatic activity in plasma may serve as a biomarker that shows disease activity in MDS. Interestingly, further digestion of DNA fragments from one nucleosome structure (~200 bp) could be observed in UPN1 (Fig. 1C). This fragmentation could not be observed when we used the plasma sample after several freeze and thaw treatments (data not shown). This may indicate that some other DNase activity in specific patients correlates with this phenomenon. Further investigation is needed.

In our experiments, the plasma circulating DNA concentration of MDS patients was much higher than that in serum. An explanation for this phenomenon may be that many DNA fragments can be absorbed into blood clots after collection into sample tubes. The difference between the DNA concentrations in plasma and serum may correlate with the phenomenon that plasma DNA, compared with serum DNA, tends to more effectively detect the epigenetic changes in MDS cells (Figs. 2 and 3). These findings indicate that circulating DNA from plasma is better suited for epigenetic analysis than that from serum in MDS patients.

For detection of the CpG methylation status of specific gene promoters, we analyzed the methylation status of the $p15^{INK4B}$ promoter, known to be one of the hypermethylated genes in MDS patients [4,17]. Unfortunately, the result did not seem to be accurate with relatively varied data among the repeated experiments, even when using plasma DNA and nested PCR technique (data not shown). One possibility is that the DNA concentration was not sufficient for the bisulfite sequencing analysis. In contrast, the result of the *LINE-1* methylation (Figs. 2 and 3) was much more reliable compared with the analysis for the $p15^{INK4B}$ gene, because of the sufficient number of DNA copies for the *LINE-1* promoter: ~85,000 copies for *LINE-1* versus only two copies for the $p15^{INK4B}$ gene in one cell. From this standpoint, DNA methylation analysis focusing on the *LINE-1* promoter using plasma circulating DNA may be a good strategy to confirm the global methylation status in MDS patients.

Interestingly, an experiment for UPN4 indicated that the decreasing ratio of DNA methylation of the *LINE-1* promoter after treatment with azacitidine was detected much earlier when using plasma DNA (Figs. 2C, day 4, and 3B, day 3) than when using PBMNC (Figs. 2C and 3B, day 6). Since DNMTi are incorporated into DNA (and RNA) during cell division, MDS/AML cells that have a more rapid cell division cycle compared with intact cells may be affected more quickly by DNMTi in BM. Furthermore, MDS clones have a more fragile character resulting from "dysplastic" backgrounds. Taking this into consideration, circulating DNA from plasma may better reflect the DNA that comes from MDS clones in BM than from intact cells. Further investigation is required to prove this hypothesis.

Harvesting PB is painless, and is a safer and easier procedure compared with obtaining BM cells. Preparing plasma is much easier than preparing PBMNC, and plasma can very easily be stored in tubes in the freezer. Utilization of plasma circulating DNA for MDS patients may provide a new way to analyze the serial genetic/epigenetic changes that are integral to an understanding of MDS pathogenesis and their disease condition.

Acknowledgments

This work was supported by Grants-in-Aid from the Ministry of Health, Labor and Welfare, and the Ministry of Education, Culture, Sports, Science and Technology, Japan. We thank Dr. Seishi Ogawa and Dr. Masashi Sanada for performing the SNP array analysis. We thank Chika Wakamatsu, Eriko Ushida, Yukie Konishi, Mari Otsuka, Manami Kira, Mirei Okamoto, and Rie Kojima for valuable laboratory assistance.

References

[1] M.J. Walter, L. Ding, D. Shen, J. Shao, M. Grillot, M. McLellan, R. Fulton, H. Schmidt, J. Kalicki-Veizer, M. O'Laughlin, C. Kandoth, J. Baty, P. Westervelt, J.F. DiPersio, E.R. Mardis, R.K. Wilson, T.J. Ley, T.A. Graubert, Recurrent DNMT3A mutations in patients with myelodysplastic syndromes, Leukemia 25 (2011) 1153–1158.

[2] T. Ernst, A.J. Chase, J. Score, C.E. Hidalgo-Curtis, et al., Inactivating mutations of the histone methyltransferase gene EZH2 in myeloid disorders, Nat. Genet. 42 (2010) 722–726.

[3] K. Yoshida, M. Sanada, Y. Shiraishi, et al., Frequent pathway mutations of splicing machinery in myelodysplasia, Nature 478 (2011) 64–69.

[4] M. Kim, B. Oh, S.Y. Kim, et al., P15INK4b methylation correlates with thrombocytopenia, blast percentage, and survival in myelodysplastic

syndromes in a dose dependent manner: quantitation using pyrosequencing study, Leuk. Res. 34 (2010) 718–722.
[5] J. Lin, Y.L. Wang, J. Qian, et al., Aberrant methylation of DNA-damage-inducible transcript 3 promoter is a common event in patients with myelodysplastic syndrome, Leuk. Res. 34 (2010) 991–994.
[6] L. Shen, H. Kantarjian, Y. Guo, et al., DNA methylation predicts survival and response to therapy in patients with myelodysplastic syndromes, J. Clin. Oncol. 28 (2010) 605–613.
[7] G. Egger, G. Liang, A. Aparicio, P.A. Jones, Epigenetics in human disease and prospects for epigenetic therapy, Nature 429 (2004) 457–463.
[8] G. Garcia-Manero, P. Fenaux, Hypomethylating agents and other novel strategies in myelodysplastic syndromes, J. Clin. Oncol. 29 (2011) 516–523.
[9] H. Schwarzenbach, D.S. Hoon, K. Pantel, Cell-free nucleic acids as biomarkers in cancer patients, Nat. Rev. Cancer 11 (2011) 426–437.
[10] D. Allen, A. Butt, D. Cahill, M. Wheeler, R. Popert, R. Swaminathan, Role of cell-free plasma DNA as a diagnostic marker for prostate cancer, Ann. NY Acad. Sci. 1022 (2004) 76–80.
[11] E. Goto, A. Tomita, F. Hayakawa, A. Atsumi, H. Kiyoi, T. Naoe, Missense mutations in PML-RARA critical for the lack of responsiveness to arsenic trioxide treatment, Blood 118 (2011) 1600–1609.
[12] F. Ohka, A. Natsume, K. Motomura, et al., The global DNA methylation surrogate LINE-1 methylation is correlated with MGMT promoter methylation and is a better prognostic factor for glioma, PLoS ONE 6 (2011) e23332.
[13] M. Sanada, T. Suzuki, L.Y. Shih, et al., Gain-of-function of mutated C-CBL tumour suppressor in myeloid neoplasms, Nature 460 (2009) 904–908.
[14] A. Abe, Y. Minami, F. Hayakawa, et al., Retention but significant reduction of BCR-ABL transcript in hematopoietic stem cells in chronic myelogenous leukemia after imatinib therapy, Int. J. Hematol. 88 (2008) 471–475.
[15] A.P. Wolffe, Transcriptional activation. Switched-on chromatin, Curr. Biol. 4 (1994) 525–528.
[16] K. Balassiano, S. Lima, M. Jenab, et al., Aberrant DNA methylation of cancer-associated genes in gastric cancer in the European Prospective Investigation into Cancer and Nutrition (EPIC-EURGAST), Cancer Lett. 311 (2011) 85–95.
[17] T. Uchida, T. Kinoshita, H. Nagai, et al., Hypermethylation of the p15INK4B gene in myelodysplastic syndromes, Blood 90 (1997) 1403–1409.

blood

2011 118: 1600-1609
Prepublished online May 25, 2011;
doi:10.1182/blood-2011-01-329433

Missense mutations in *PML-RARA* are critical for the lack of responsiveness to arsenic trioxide treatment

Emi Goto, Akihiro Tomita, Fumihiko Hayakawa, Akihide Atsumi, Hitoshi Kiyoi and Tomoki Naoe

Blood (print ISSN 0006-4971, online ISSN 1528-0020), is published weekly by the American Society of Hematology, 2021 L St, NW, Suite 900, Washington DC 20036.

Missense mutations in *PML-RARA* are critical for the lack of responsiveness to arsenic trioxide treatment

Emi Goto,[1] Akihiro Tomita,[1] Fumihiko Hayakawa,[1] Akihide Atsumi,[1] Hitoshi Kiyoi,[1] and Tomoki Naoe[1]

[1]Department of Hematology and Oncology, Nagoya University Graduate School of Medicine, Nagoya, Japan

Arsenic trioxide (As_2O_3) is a highly effective treatment for patients with refractory/relapsed acute promyelocytic leukemia (APL), but resistance to As_2O_3 has recently been seen. In the present study, we report the findings that 2 of 15 patients with refractory/relapsed APL treated with As_2O_3 were clinically As_2O_3 resistant. Leukemia cells from these 2 patients harbored missense mutations in promyelocytic leukemia gene–retinoic acid receptor-α gene (*PML-RARA*) transcripts, resulting in amino acid substitutions of A216V and L218P in the PML B2 domain. When wild-type or mutated PML-RARA (PR-WT and PR-B/L-mut, respectively) were overexpressed in HeLa cells, immunoblotting showed SUMOylated and/or oligomerized protein bands in PR-WT but not in PR-B/L-mut after As_2O_3 treatment. Protein-localization analysis indicated that PR-WT in the soluble fraction was transferred to the insoluble fraction after treatment with As_2O_3, but PR-B/L-mut was stably detected in fractions both with and without As_2O_3. Immunofluorescent microscopy analysis showed PR-WT localization as a microgranular pattern in the cytoplasm without As_2O_3 and as a macrogranular pattern with As_2O_3. PR-B/L-mut was diffusely observed in the cytoplasm with and without As_2O_3. Nearly identical localization patterns were observed in patients' primary cells. Therefore, B2 domain mutations may play an important role in aberrant molecular responses to As_2O_3 and may be critical for As_2O_3 resistance in APL. (*Blood.* 2011;118(6):1600-1609)

Introduction

Acute promyelocytic leukemia (APL) is characterized by the reciprocal chromosomal translocation t(15;17)(q22;q21), leading to fusion of the promyelocytic leukemia gene (*PML*) on chromosome 15 and the retinoic acid receptor-α gene (*RARA*) on chromosome 17.[1] *PML-RARA* fusions are detectable in > 95% of patients with APL. In 1985, *all-trans* retinoic acid (ATRA) was introduced for the treatment of APL as a differentiation therapy, and a dramatic improvement in the overall survival of patients with APL has been obtained.[2-4] However, approximately 10%-30% of patients eventually relapse after treatment with combination chemotherapies with ATRA.[5-7]

Arsenic trioxide (As_2O_3) is a critical drug for the treatment of APL and is clinically effective even in ATRA-resistant patients.[8] As_2O_3 is a natural substance that has been used medically for over 2400 years. In the 1970s, a group in China identified As_2O_3 as a component of an anticancer reagent.[9] Over the last 18 years, clinical trials conducted worldwide have demonstrated the efficacy of As_2O_3 for the treatment of relapsed patients with APL.[10,11] Recently, it was also reported that As_2O_3 improves event-free survival and overall survival of adult APL when As_2O_3 is used as a consolidation treatment after obtaining the first remission.[12] Currently, the role of As_2O_3 in frontline therapy is under investigation.[10,13]

Rapid degradation of PML-RARA via targeting of PML has been reported as a molecular mechanism for the effectiveness of As_2O_3.[14] Furthermore, As_2O_3 induces posttranslational modifications of PML-RARA with small ubiquitin-related modifier (SUMO) and ubiquitin, resulting in the transfer of PML-RARA from the soluble fraction to the insoluble nuclear matrix[14] and the degradation of both PML and PML-RARA.[14-17] In addition to the significant clinical effectiveness of As_2O_3 for patients with APL, acquired resistance to As_2O_3 therapy has been recognized in clinical practice.[18] Several studies have indicated that arsenic-resistant NB4 cells in vitro show higher glutathione levels than in parental cells.[19-21] However, the detailed molecular mechanisms of resistance to As_2O_3 remain unclear.

Very recently, 2 studies reported that As_2O_3 binds directly to cysteine residues in zinc fingers located within the RBCC motif that contains 3 cysteine-rich zinc-binding domains, a RING-finger (R), 2 B-box motifs (B1 and B2), and a coiled-coil (CC) domain,[22,23] in PML-RARA and PML.[24,25] An intriguing hypothesis is that impairment of As_2O_3 binding to PML-RARA due to conformational changes may result from genetic mutations and/or abnormal posttranslational modifications. These events may be related to resistance to As_2O_3 therapy.

We report the clinical significance and frequency of As_2O_3 resistance in patients with APL. Fifteen patients with APL were treated with As_2O_3 after combination chemotherapy with ATRA, and 2 patients showed clinical As_2O_3 resistance. Interestingly, in both of these As_2O_3-resistant patients, missense genetic mutations in the *PML-RARA* fusion transcript were observed in the leukemia cells. We demonstrated that the mutations, which were located in the PML RBCC region, were critical for PML localization and As_2O_3 responsiveness in vitro. Our observations suggest that acquired genetic mutations in the *PML-RARA* transcript may be a critical molecular mechanism of resistance to As_2O_3 therapy.

Methods

Patients

From January 2000 to December 2008 at Nagoya University Hospital, Japan, 15 patients with APL who showed relapse or disease progression

Submitted January 6, 2011; accepted May 9, 2011. Prepublished online as *Blood* First Edition paper, May 25, 2011; DOI 10.1182/blood-2011-01-329433.

An Inside *Blood* analysis of this article appears at the front of this issue.

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.