Review

Genetic basis of myelodysplastic syndromes

By Seishi OGAWA*1,†

(Edited by Takao SEKIYA, M.J.A.)

Abstract: During the past decade, substantial progress has been made in the field of the genetics of myelodysplastic syndromes (MDS). These comprise a group of chronic myeloid neoplasms with abnormal cell morphology and progression to acute myeloid leukemia (AML), where revolutionary sequencing technologies have played a major role. Through extensive sequencing of a large number of MDS genomes, a comprehensive registry of driver mutations involved in the pathogenesis of MDS has been revealed, along with their impacts on clinical phenotype and prognosis. The most frequently affected molecules are involved in DNA methylations, chromatin modification, RNA splicing, transcription, signal transduction, cohesin regulation, and DNA repair. These mutations show strong positive and negative correlations with each other, suggesting the presence of functional interactions between mutations, which dictate disease progression. Because these mutations are associated with disease phenotype, drug response, and clinical outcomes, it is essential to be familiar with MDS genetics not only for better understanding of MDS pathogenesis but also for management of patients.

Keywords: myelodysplastic syndromes (MDS), splicing factors, cohesin, secondary acute myeloid leukemia (sAML), next-generation sequencing

Introduction

Myelodysplastic syndromes (MDS) are a heterogeneous group of chronic myeloid neoplasms, which are characterized by refractory cytopenia with abnormal bone marrow (BM) morphology and frequent progression to acute myeloid leukemia (AML) (secondary AML [sAML]).1,2) Although AML progression, which is found in one-third of these patients, is a strong predictor of poor outcomes, many of the remaining patients also die from BM failure, which leads to fatal bleeding and infection.3) Recently, several new drugs, such as lenalidomide1) and hypomethylating agents,5) have been introduced into clinics, contributing improved outcomes for a subset of MDS patients. However, no curative therapy is available for MDS patients, except for allogeneic hematopoietic stem cell transplantation, which unfortunately is applicable only to younger patients, even though the vast majority of patients are elderly and not eligible to transplantation. Thus, there is an urgent need to understand the molecular pathogenesis of MDS to develop effective but less toxic therapies.

Little was known about relevant gene mutations that are implicated in MDS pathogenesis until early 2000s. This was in contrast to AML, whose molecular pathogenesis was intensively studied during the 1990s through the analysis of highly recurrent chromosomal translocations. In fact, only a handful of genes, including NRAS, TP53, ATRX, and RUNX1, were known to be mutated in MDS.6–10) The first breakthrough was brought about by the development of single nucleotide polymorphism (SNP) array-based genomic copy number analysis (or SNP array karyotyping), using millions of SNP probes, which enabled the comprehensive detection of genomic copy number abnormalities (CNAs) at an unprecedented resolution (~several kilobase pairs) across the entire cancer genome.11,12) Thus, using SNP-array karyotyping, we and other groups identified a number of recurrent deletions, amplifications,
and other allelic imbalances, such as copy neutral loss-of-heterozygosity (LOH) or uniparental disomies (UPDs), from which novel mutational/functional target genes were identified, including TET2, CBL, EZH2, CUX1, and ASXL1, which are targeted by common copy number lesions in 4q (UPD/deletion), 11q (UPD), 7q (UPD/deletion), and 20q (deletion), respectively. Other gene targets of common CNAs include NRAS, FLT3, TP53, and RUNX1, whose function is not well-established, all SFs mutated in 29 cases are included in the U2 RNP complex and involved in one of the initial steps of mRNA splicing, i.e., 5′ splice site recognition (Fig. 1a). To confirm these initial results from whole exome sequencing in the 29 cases, we performed extended mutation analysis in a large cohort of 582 cases with a wide variety of myeloid neoplasms, including not only different subtypes of MDS and sAML, but also de novo AML as well as classical myeloproliferative neoplasms, focusing on SFs involved in 5′ splice site recognition. Through this large-scale analysis, we confirmed that SF mutations were found in all subtypes of MDS, such as refractory anemia (RA), RA with excess blasts (RAEB), refractory cytopenia with multilineage dysplasia (RCMD) as well as RA with ringed sideroblasts (RARS) and refractory cytopenia with multilineage dysplasia and ringed sideroblasts (RCMD-RS), and also chronic myelomonocytic leukemia (CMML) at very high frequencies ranging from 45% to 85%, depending on subtype (Fig. 1b). Conspicuously, in two unique subtypes, RARS and RCMD-RS, which were characterized by elevated numbers of ring sideroblasts (also known as sideroblastic anemia), >70% of the patients had SF3B1 mutations, making these mutations almost pathognomonic to these MDS subtypes, which was also confirmed in other studies. Moreover, SF mutations were also found in other myeloid neoplasms, including de novo and secondary AML as well as classical myeloproliferative neoplasm (MPN) cases, although mutation frequencies were largely <10%, except in sAML progressed from MDS, where an intermediate mutation frequency (~25%) was observed. These findings strongly suggest that SF mutations are cardinal features of MDS and to a lesser extent, other myeloid neoplasms.

SF mutations were found in at least 8 SFs, of which by far the most frequently affected were SF3B1, SRSF2, U2AF1, and ZRSR2. Among these, mutations in SF3B1, SRSF2, and U2AF1 showed prominent mutational hotspots (Fig. 1c); in SRSF2 most mutations affected the P95 residue between two functional domains, whereas mutations in U2AF1 exclusively affected two highly conserved

Identification of splicing factor mutations in MDS using whole exome sequencing

To investigate novel mutational targets in MDS in an unbiased manner, we performed whole exome sequencing analysis of paired tumor/normal DNA from 29 cases with MDS and related myeloid neoplasms, including chronic myelomonocytic leukemia and sAML. On average, we identified 265 mutations (9 somatic mutations/sample). Most of the mutated genes were affected in a single sample, except for 13 genes shown in Table 1 (in bold), which were recurrently mutated in multiple samples, suggesting their role in MDS pathogenesis. Although 8 of these genes were well-known mutational targets in MDS, the remaining 4 were not reported previously, of which three genes (U2AF1, SRSF2, and ZRSR2) encode RNA splicing factors (SFs). Notably, also including an additional three genes (SF3A1, SF3B1, and PRPF40B), 16 out of 29 MDS cases analyzed had a mutated SF, suggesting that deregulated RNA splicing plays a central role in MDS pathogenesis.

RNA splicing is a basic cellular apparatus ubiquitously found in all eukaryotes, which generates a wide variety of different splicing forms, enabling high-degrees of protein species expressed in eukaryotic organisms. The process begins with recognition of exon/intron boundaries, which is accomplished by U1 or U2 snRNP complex, followed by two trans-ester reactions and excision of the intronic sequence as a lariat structure. Of interest, except for PRPF40B, whose function is not well-established, all SFs mutated in 29 cases have been found in the U2 RNP complex and involved in one of the initial steps of mRNA splicing, i.e., 5′ splice site recognition (Fig. 1a). To confirm these initial results from whole exome sequencing in the 29 cases, we performed extended mutation analysis in a large cohort of 582 cases with a wide variety of myeloid neoplasms, including not only different subtypes of MDS and sAML, but also de novo AML as well as classical myeloproliferative neoplasms, focusing on SFs involved in 5′ splice site recognition. Through this large-scale analysis, we confirmed that SF mutations were found in all subtypes of MDS, such as refractory anemia (RA), RA with excess blasts (RAEB), refractory cytopenia with multilineage dysplasia (RCMD) as well as RA with ringed sideroblasts (RARS) and refractory cytopenia with multilineage dysplasia and ringed sideroblasts (RCMD-RS), and also chronic myelomonocytic leukemia (CMML) at very high frequencies ranging from 45% to 85%, depending on subtype (Fig. 1b). Conspicuously, in two unique subtypes, RARS and RCMD-RS, which were characterized by elevated numbers of ring sideroblasts (also known as sideroblastic anemia), >70% of the patients had SF3B1 mutations, making these mutations almost pathognomonic to these MDS subtypes, which was also confirmed in other studies. Moreover, SF mutations were also found in other myeloid neoplasms, including de novo and secondary AML as well as classical myeloproliferative neoplasm (MPN) cases, although mutation frequencies were largely <10%, except in sAML progressed from MDS, where an intermediate mutation frequency (~25%) was observed. These findings strongly suggest that SF mutations are cardinal features of MDS and to a lesser extent, other myeloid neoplasms.

SF mutations were found in at least 8 SFs, of which by far the most frequently affected were SF3B1, SRSF2, U2AF1, and ZRSR2. Among these, mutations in SF3B1, SRSF2, and U2AF1 showed prominent mutational hotspots (Fig. 1c); in SRSF2 most mutations affected the P95 residue between two functional domains, whereas mutations in U2AF1 exclusively affected two highly conserved
Table 1. Genes mutated in multiple samples

| Gene    | Type       | Chr | Mutation          | A.A          | Sample ID |
|---------|------------|-----|-------------------|--------------|-----------|
| ASXL1   | Nonsense   | Chr20 | C > A            | S852X        | MDS-19    |
| ASXL1   | Nonsense   | Chr20 | C > T            | R693X        | CMMML-01 |
| BCOR    | Frameshift | ChrX  | */+C             | V1351fs      | MDS-19    |
| BCOR    | Nonsense   | ChrX  | C > T            | Q1187X       | CMMML-02 |
| DNMT3A  | Missense   | Chr2  | G > T            | G543C        | MDS-19    |
| DNMT3A  | Missense   | Chr2  | G > A            | R882H        | CMMML-02 |
| EZH2    | Missense   | Chr7  | A > C            | D664A        | tAML-01   |
| EZH2    | Frameshift | Chr7  | */-AA            | K718fs       | CMMML-01 |
| KRAS    | Missense   | Chr12 | G > A            | G12D         | tAML-06   |
| KRAS    | Missense   | Chr12 | A > G            | K5E          | CMMML-01 |
| NRAS    | Missense   | Chr1  | G > A            | G13D         | MDS-19    |
| NRAS    | Missense   | Chr1  | G > A            | G13D         | MDS-20    |
| NRAS    | Missense   | Chr1  | G > A            | G12S         | MDS-20    |
| NRAS    | Missense   | Chr1  | G > A            | G12S         | tAML-05   |
| RUNX1   | Frameshift | Chr21 | */+G             | L294fs       | MDS-19    |
| RUNX1   | Missense   | Chr21 | G > A            | D171N        | tAML-02   |
| SRSF2   | Missense   | Chr17 | C > T            | P95L         | MDS-09    |
| SRSF2   | Missense   | Chr17 | C > A            | P95H         | MDS-18    |
| SRSF2   | Missense   | Chr17 | C > A            | P95H         | CMMML-04 |
| SRSF2   | Deletion   | Chr17 | —                | P955HR103>R  | tAML-07   |
| STAG2   | Nonsense   | ChrX  | C > T            | R1012X       | MDS-12    |
| STAG2   | Splice site| ChrX  | G > A            | Q399_spl     | MDS-19    |
| TET2    | Frameshift | Chr4  | */-C             | I340fs       | MDS-06    |
| TET2    | Frameshift | Chr4  | */+C             | Y1255Fs      | MDS-16    |
| TET2    | Nonsense   | Chr4  | C > T            | Q323X        | MDS-16    |
| TET2    | Nonsense   | Chr4  | C > G            | S825X        | MDS-19    |
| TET2    | Frameshift | Chr4  | */+A             | L1046fs      | MDS-19    |
| TET2    | Nonsense   | Chr4  | C > T            | R556X        | tAML-01   |
| TET2    | Frameshift | Chr4  | */-C             | H994fs       | tAML-01   |
| TET2    | Nonsense   | Chr4  | C > T            | Q1680X       | tAML-07   |
| TET2    | Frameshift | Chr4  | */+A             | S402Fs       | tAML-07   |
| TET2    | Frameshift | Chr4  | */-AAAT          | Q749fs       | CMMML-01 |
| TP53    | Missense   | Chr17 | C > T            | R116W        | MDS-09    |
| TP53    | Frameshift | Chr17 | */-T             | M1fs         | tAML-03   |
| U2AF1   | Missense   | Chr21 | A > G            | Q157R        | MDS-03    |
| U2AF1   | Missense   | Chr21 | C > T            | A26V         | MDS-12    |
| U2AF1   | Missense   | Chr21 | A > C            | Q157P        | MDS-15    |
| U2AF1   | Missense   | Chr21 | A > G            | Q157R        | CMMML-01 |
| U2AF1   | Missense   | Chr21 | C > T            | S34F         | CMMML-02 |
| ZRSR2   | Splice site| ChrX  | G > A            | K257splice   | MDS-06    |
| ZRSR2   | Frameshift | ChrX  | */-G             | G323fs       | MDS-08    |
| ZRSR2   | Nonsense   | ChrX  | G > T            | E362X        | MDS-16    |
| ZRSR2   | Missense   | ChrX  | A > T            | N261Y        | tAML-05   |

amino-acids within Zinc-finger domains flanking the RNA binding domain (S34 and Q157).26),31),32) In addition, in SF3B1 nearly half of the mutations affected K700 and surrounding amino-acids.26) The presence of these prominent mutational hot spots strongly suggested that these cause neomorphic or gain-of-function mutations. By contrast, mutations in ZRSR2 were distributed along the entire protein,
most of which resulted in a premature truncation, and thus were thought to cause loss-of-function. Additional important features of SF mutations in MDS and other myeloid neoplasms are that these mutations occurred in a mutually exclusive manner and that except for ZRSR2 mutations, all SF mutations were heterozygous with no homozygous mutations (Fig. 1d). These features suggest that multiple or homozygous SF mutations were lethal and not compatible with cell viability and positive selection of cells having SF mutations in such a manner.

**Functional consequence of SF mutations**

Several studies, including ours, addressed the effects of SF mutations on RNA splicing and biological phenotypes. To investigate the effects of SF mutations on RNA splicing, we performed RNA sequencing of CD34+ cells from a total of 100 MDS patients, including 32 cases with SF3B1 mutations and 23 with SRSF2 mutations. RNA from whole BM cells were also sequenced for 165 patients, which consisted of 54 SF3B1-mutated and 32 SRSF2-mutated, and 14 U2AF1-mutated cases. As expected, SF-mutated samples showed extensive changes in splicing compared with samples without any SF mutations (n = 41 for CD34+ cells and n = 66 for whole BM cells) (Fig. 2a). However, the patterns of splicing changes differed substantially depending on the mutated SF. For example, SF3B1-mutated samples showed significantly increased alternative 3' splice sites, which were 5' upstream of a canonical branch point, suggesting that they were caused by the recognition of an alternative 5' branch point by the mutated SF3B1. SF3B1-mutated samples also exhibited a dramatically reduced number of retained introns (Fig. 2a). Of note many of these alternative splicing forms in SF3B1-mutated samples cause nonsense-mediated decay, leading to significantly reduced gene expression (Fig. 2b). A notable example is 3' alternative splicing in ABCB7 encoding an iron transporter in mitochondria, haploinsufficiency of which is responsible for congenital sideroblastic anemia (Fig. 2c). By contrast, mutations in SRSF2, which binds to exonic splicing enhancer (ESEs) via
consensus sequences (SSNG) and is implicated in the regulation of splicing efficiency, were associated with enhanced or suppressed cassette (or alternative) exons, where alternative exons containing the CCNG consensus have higher affinity to mutant SRSF2 tended to be more frequently included, whereas those carrying the consensus with less affinity were more often excluded (Fig. 2d).36)40) Splicing changes in U2AF1-mutated samples were dominated by alterations in cassette exons as seen in SRSF2-mutated samples. In contrast to a 3’ alternative splice site, alternative cassette exons seen in SRSF2 or U2AF1 mutants were less likely to influence gene expression. Nevertheless, some genes with mutant SRSF2- or U2AF1-associated alternative splicing resulted in reduced gene expression (Fig. 2e). An example is EZH2, a catalytic component of PRC2 frequently mutated in MDS, whose alternative exons in introns...
10 and 11 are more frequently included or excluded caused a premature termination in SRSF2-mutated cells, of which the cassette exon in intron 10 is also included in the U2AF1 S34F mutants (Fig. 2f). Alternative splicing affecting EZH2 provides an example in which different SF mutants affected splicing of a common target, which was rarely seen despite large numbers of gene targets of SF mutant-associated alternative splicing. However, in general there were a few genes commonly mutated in MDS whose alternative splicing was caused by different SF mutants.

We and other groups also investigated the biological consequences of mutant SFs using an engineered mouse model with conditional mutant SF alleles. However, even though targets of abnormal splicing of SF mutations overlapped with known mutational targets, the molecular pathogenesis of mutated SFs has not fully been elucidated. Indeed all these mouse models showed abnormal splicing of expected patterns, as seen in human samples, and developed dysplastic cell morphology as well as reduced production of mature blood cells. However, BM cells from these mice did not show any evidence of clonal advantage in competitive repopulation analysis; when equivalent numbers of BM cells from mutant and wild-type mice were transplanted into lethally irradiated recipients. When chimerism between wild-type and mutant SFs was evaluated, mutant cells unexpectedly showed a reduced rather than elevated chimerism, suggesting that SF-mutant BM cells have compromised repopulation capacity. One of the explanations might be an altered BM microenvironment favoring SF-mutant cells. Future studies are clearly warranted to clarify the exact mechanism of positive selection of SF-mutant cells in MDS patients.

**Cohesin mutations**

An additional mutational targets mutated in MDS more frequently than expected only by chance (i.e., driver genes) were identified through exome sequencing of an extended cohort. Cohesin is a multimeric protein complex comprising RAD21, SMC1, SMC3, and a STAG protein (STAG1-3) (Fig. 3a). Taking a ring-like structure, these molecules are recruited on chromatin in concert with cohesin-associated molecules, such as NIPBL and ESCO proteins, to prevent sister chromatids from premature separation during cell division. However, these also participate in post-replicative DNA repair and long-range regulation of gene expression through the formation of large-scale chromatin structures (Fig. 3b). In MDS and other myeloid malignancies, approximately 10–15% of cases harbor mutations in cohesin and related molecules. Among these cohesin components, STAG2 was most frequently mutated, in which most mutations were truncating, leading to loss-of-function, and so were RAD21 mutations (Fig. 3c). As is the case with SF mutations, cohesin mutations are largely mutually exclusive (Fig. 3d), suggesting synthetic lethality of mutations in multiple cohesin components. Cohesin mutations were also reported in several solid tumors, including glioblastoma, melanoma, Ewing sarcoma, and colon cancers, where aneuploidy caused by loss of cohesin functions has been implicated in tumorigenesis. However, in myeloid malignancies, many of cohesin-mutated tumors had normal or near normal karyotypes (Fig. 3d), arguing against the role of aneuploid but supporting altered gene regulation as the mechanism of cohesin-mutated myeloid neoplasms. Recent studies showed that cohesin-deficient cells had altered chromatin structures that allowed for accessibility to a number of transcription factors, including RUNX1, GATA2, and ERG, suggesting functional interactions of cohesin mutations and these transcription factors.

**Landscape of gene mutations in MDS**

Other newly identified driver genes in MDS include BCOR/BCORL1, BRCC3, PHF6, ETV6, SETBP1, PRPF8, and RIT1. Thus, to elucidate the landscape of mutations in MDS, we conducted a comprehensive analysis of mutations in these major driver or candidate driver genes in a total of 944 MDS patients using targeted-capture sequencing. In total 47 genes were mutated at frequencies expected only by chance, where one or more of the 47 driver genes were mutated in as many as 864 (91.5%) of the 944 patients (Fig. 4). These driver genes were grouped into several discrete functional pathways, including RNA splicing, DNA methylation (TET2, DNMT3A, IDH1 and IDH2), chromatin modification (ASXL1, EZH2, and KDM6A), transcription (RUNX1, ETV6, BCOR/BCORL1, GATA2, CEBPA, and NPM1), receptor/tyrosine kinase, cohesin, RAS signaling (KRAS, NRAS, PTPN11, and CBL), and DNA repair (ATM, BRCC3, and FANCL). With a median of 3 per sample, the number of driver mutations increased with WHO risk. Among these, only 6 driver genes were mutated at 10% ≥ frequencies, followed by many low-frequency mutated genes. Most of the
driver genes commonly mutated in MDS were also
mutated in AML, but their frequency is substantially
different between these malignancies (Fig. 5),
suggesting that distinct pathogenesis between MDS
and AML despite a substantial overlap in mutated
drivers. In particular, a number of driver genes,
including NRAS, FLT3, KRAS, PTPN11, IDH1,
IDH2, WT1, RAD21, KIT, and CEBPA, were less
frequently mutated in MDS compared with in
primary AML (Fig. 5). This was also in line with
the differences in the cytogenetic or copy number
alteration profiles between AML and MDS. AML is
characterized by frequent disease-specific reciprocal
translocations, such as t(8;21)(q22;q22), inv(16)
(p13q22), t(3;21)(q26;q22), t(15;17)(q22;q21) and
other MLL(11q23)-affecting translocations, leading
to RUNX1/RUNX1T1, RUNX1/EVII, PML/RARA,
and a variety of MLL fusions, respectively. By
contrast, most of the chromosomal lesions in MDS
are unbalanced changes, which are caused by deletion
and amplification of one or more chromosomal
segments, including del(5q)/-5, del(7q)/-7, 12pLOH,
del(20q), and 17pLOH (Fig. 5), most of which are seen in combinations as complex karyotypes in association with biallelic TP53 involvement. t(3;21)(q26;q22), inv(3)(q21q26), t(3;3)(q21;q26), t(1;3)(q36;q21), as well as t(6;9)(q22;q34) are typically seen in borderline cases between AML and MDS, predicting poor prognosis. Other abnormalities, such as D8, D21, are also seen in both AML and MDS at similar frequencies.

Co-occurring and mutually exclusive mutations

Among these driver mutations and copy number alterations, there exist strong positive and negative correlations. By exhaustively interrogating pairwise combination of mutations among the 944 patients, we detected a total of 82 significant (q < 0.01) combinations of mutations (Fig. 6), suggesting strong functional interactions between these mutations. Most of these significant combinations of mutations showed excellent concordance with those detected by analysis of independent MDS cohorts (Table 2), confirming the validity of these mutational combinations. For example, SF mutations are almost mutually exclusive with each other and so are mutations in components of the cohesin complex, likely due to synthetic lethality of multiple mutations in these pathways and complexes. By contrast, STAG2 mutations significantly cooccurred with mutations in RUNX1, SRSF2, ASXL1, EZH2, BCOR, and IDH2.55,64) Mutations and copy number alterations are largely mutually exclusive with some exceptions, of which significant correlations between copy number alterations and TP53 mutations. TP53 mutations, particularly when biallelic, are tightly associated with complex karyotypes with aneuploidy, which uniformly shows very poor prognosis resistant even to allogeneic stem cell transplantation, whereas TP53 mutations tend to be mutually exclusive with other mutations (Fig. 6).63)

Moreover, not only combinations of mutations but also their order of acquisition is not totally random, but there is a strong trend that some
mutations tend to precede others, frequently associated with progression from low-risk to high-risk diseases and to sAML. To examine this, we interrogated significant enrichment of mutations between low- and high-risk MDS and also between high-risk MDS and sAML in a cohort of 2,250 MDS cases.\(^6\) As shown in Fig. 7a, there was a significant enrichment of mutations in SF3B1 in low-risk vs. high-risk MDS cases, whereas mutations in TP53, GATA2, RUNX1, KRAS, STAG2, ASXL1, ZRSR2, and TET2 (Type-2 mutations) were significantly enriched in high-risk compared with low-risk MDS. Between high-risk and sAML cases, there was a significant enrichment of mutations in FLT3, PTPN11, WT1, IDH1, NPM1, IDH2, and NRAS (Type-1 mutations) (Fig. 7a). These results suggested that Type-2 and Type-1 mutations are associated with progression from low-risk to high-risk MDS and high-risk MDS to sAML, respectively. To see this, we investigated the impact of Type-1 and Type-2 mutations on progression to sAML. We found that patients carrying one or more of Type-1 mutations at the time of diagnosis tended to progress to sAML earlier than those cases with Type-2 mutations alone (Fig. 7b). Patients with SF3B1 mutations but not Type-1 or Type-2 mutations rarely progressed to sAML and those without Type-1 or Type-2 mutations displayed an intermediate risk of AML progression. Notably, all Type-1 mutations were also strongly enriched in primary

---

Fig. 5. Different mutation frequency between MDS, sAML, and primary AML. Frequency of major driver genes mutated in common in different myeloid malignancies, including MDS, sAML, and de novo primary AML. Risks of MDS are indicated by color. Mutations commonly seen in primary AML but only rarely in MDS are shown in dark blue in the AML panel. Cited from Ogawa, Blood, 2019.\(^6\)
AML, compared with MDS. Of further interest, upon progression to sAML, acquisition of one or more Type-1 mutations was confirmed in some patients with Type-2 mutations alone. These Type-1 and Type-2 mutations demonstrated significant co-occurrence between each other, suggesting that mutations are associated with disease progression. Recent studies more closely described the dynamics of clones during progression from MDS to AML using longitudinal samples for deep sequencing and single-cell sequencing. The presence of multiple clones in MDS, particularly in immature stem cell fractions, from which a clone evolves to sAML, while swiping other clones.

**Clinical impact of mutations**

Finally, we evaluated the impact of mutations on survivals of 944 MDS patients. To examine the
impact of gene mutations, we generated two Cox proportional hazard models, in which the model was constructed using mutation alone (Model-II) or both clinical factors and mutations (Model-I), including complete blood counts, BM blast%, and cytogenetics, which were compared with IPSS and IPSS-R conventionally used for prognostication of MDS patients, using LASSO for selection of variables. In univariate analysis, mutations in 25 genes were significantly associated with overall survival (OS), of which only SF3B1 mutations predicted a better OS. The model combining clinical and gene mutations significantly outperformed traditional IPSS and IPSS-R models, which was in accordance with other studies, but in accordance of an earlier studies evaluating the clinical effects of mutations in MDS.

Conclusions
During the past two decades, our knowledge about the genetics of MDS has improved dramatically, with the revolution in genomics, particularly next-generation sequencing, providing a comprehensive registry of driver mutations in MDS, their frequency, and mutual correlations, as well as their impacts on disease phenotypes and prognosis. Now, allied efforts are in progress by the international working group for the prognosis of MDS (IWG-PM), in which thousands of MDS patients are being investigated for major driver mutations to confirm and further extend these initial findings. These studies provide a solid basis for novel molecular classifications of MDS, which help accurate diagnosis and prognostication, as well as better choice of therapies to ultimately improve patients’ clinical outcomes. Finally, despite a large number of newly revealed driver mutations, their functional roles in MDS pathogenesis remain to be elucidated for many drivers. For example, as seen above, SF mutations unequivocally play critical roles in MDS pathogenesis, whereas no functional evidence has been demonstrated for how SF-mutated hematopoietic cells achieve a clonal dominance; SF-mutated stem cells show a compromised competitive repopulation potential over normal stem cells in mouse models. The importance of understanding of the functional/molecular mechanisms of these driver mutations identified through genetic analyses cannot be underestimated.

Acknowledgement
This work is supported by Grants-in-Aid for Scientific Research (MEXT/JSPS KAKENHI
JP26221308 and JP26253060), a Grant-in-Aid from the Japan Agency for Medical Research and Development (Project for Development of Innovative Research on Cancer Therapeutics [16cm0106501h0001, and Practical Research for Innovative Cancer Control Project [15ck0106073h0002 and 16ck0106073h0003]), and grants from the Takeda Science Foundation.

References

1) Nimer, S.D. (2008) Myelodysplastic syndromes. Blood 111, 4841–4851.
2) Telfer, A. and Vardiman, J.W. (2009) Myelodysplastic syndromes. N. Engl. J. Med. 361, 1872–1885.
3) Corey, S.J., Minden, M.D., Barber, D.L., Kantarjian, H., Wang, J.C. and Schimmer, A.D. (2007) Myelodysplastic syndromes: The complexity of stem-cell diseases. Nat. Rev. Cancer 7, 118–129.
4) List, A., Kurtin, S., Roe, D.J., Buresh, A., Mahadevan, D., Fuchs, D. et al. (2005) Efficacy of lenalidomide in myelodysplastic syndromes. N. Engl. J. Med. 352, 549–557.
5) Fenaux, P., Mufti, G.J., Hellstrom-Lindberg, E., Santimi, V., Finelli, C., Giagounidou, A. et al. (2009) Efficacy of azacitidine compared with that of conventional care regimens in the treatment of higher-risk myelodysplastic syndromes: A randomised, open-label, phase III study. Lancet Oncol. 10, 223–232.
6) Hirai, H., Kobayashi, Y., Mano, H., Hagiwara, K., Maru, Y., Omine, M. et al. (1987) A point mutation at codon 13 of the N-ras oncogene in myelodysplastic syndrome. Nature 327, 430–432.
7) Sugimoto, K., Hirano, N., Toyoshima, H., Chiba, S., Mano, H., Takaku, F. et al. (1993) Mutations of the p53 gene in myelodysplastic syndrome (MDS) and MDS-derived leukemia. Blood 81, 3022–3026.
8) Osato, M., Asou, N., Abdalla, E., Hoshino, K., Yamasaki, H., Okubo, T. et al. (1999) Biallelic and heterozygous point mutations in the runt domain of the AML1/PEBP2αB gene associated with myeloblastic leukemias. Blood 93, 1817–1824.
9) Imai, Y., Kurokawa, M., Iizutsu, K., Hagaishi, A., Takeuchi, K., Maki, K. et al. (2000) Mutations of the AML1 gene in myelodysplastic syndrome and their functional implications in leukemogenesis. Blood 96, 3154–3160.
10) Steensma, D.P., Higgs, D.R., Fisher, C.A. and
Gibbons, R.J. (2004) Acquired somatic ATRX mutations in myelodysplastic syndrome associated with α-thalassemia (ATMDS) convey a more severe hematologic phenotype than germline ATRX mutations. Blood 103, 2019–2026.

11) Nannya, Y., Sanada, M., Nakazaki, K., Hosoya, N., Wang, L., Hangaiishi, A. et al. (2005) A robust algorithm for copy number detection using high-density oligonucleotide single nucleotide polymorphism genotyping arrays. Cancer Res. 65, 6071–6079.

12) Yamamoto, G., Nannya, Y., Kato, M., Sanada, M., Levine, R.L., Kawamata, N. et al. (2007) 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.

13) Sato-Otsubo, A., Sanada, M. and Ogawa, S. (2012) Single-nucleotide polymorphism array karyotyping in clinical practice: Where, when, and how? Semin. Oncol. 39, 13–25.

14) Delhiommeau, F., Dupont, S., Della Valle, V., James, C., Trannoy, S., Masse, A. et al. (2009) Mutation in TET2 in myeloid cancers. N. Engl. J. Med. 360, 2289–2301.

15) Langemeijer, S.M., Kuiper, R.P., Berends, M., Knops, R., Aslanyan, M.G., Massop, M. et al. (2009) Acquired mutations in TET2 are common in myelodysplastic syndromes. Nat. Genet. 41, 838–842.

16) Sanada, M., Suzuki, T., Shihi, L.Y., Otsu, M., Kato, M., Yamazaki, S. et al. (2009) Gain-of-function of mutated C-CBL tumour suppressor in myeloid neoplasms. Nature 460, 904–908.

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

18) Nikologi, G., Langemeijer, S.M., Kuiper, R.P., Knops, R., Massop, M., Tonnissen, E.R. et al. (2010) Somatic mutations of the histone methyltransferase gene EZH2 in myeloid syndromes. Nat. Genet. 42, 665–667.

19) Thoenissen, N.H., Lasho, T., Thoenissen, G.B., Ogawa, S., Tefferi, A. and Koeffler, H.P. (2011) Novel CUX1 missense mutation in association with 7q− at leukemic transformation of MPN. Am. J. Hematol. 86, 703–705.

20) Gelsi-Boyer, V., Trouplin, V., Adaléa, J., Bonans, J., Cervera, N., Carbuchia, N. et al. (2009) Mutations of polycomb-associated gene ASXL1 in myelodysplastic syndromes and chronic myelomonocytic leukaemia. Br. J. Haematol. 145, 788–800.

21) Ley, T.J., Mardis, E.R., Ding, L., Fulton, B., McLellan, M.D., Chen, K. et al. (2008) DNA sequencing of a cytogenetically normal acute myeloid leukaemia genome. Nature 456, 66–72.

22) Meyerson, M., Gabriel, S. and Getz, G. (2010) Advances in understanding cancer genomes through second-generation sequencing. Nat. Rev. Genet. 11, 685–696.

23) Vogelstein, B., Papadopoulos, N., Velculescu, V.E., Zhou, S., Diaz, L.A. Jr. and Kinzler, K.W. (2013) Cancer genome landscapes. Science 339, 1546–1558.

24) Mardis, E.R., Ding, L., Dooling, D.J., Larson, D.E., McLellan, M.D., Chen, K. et al. (2009) Recurring mutations found by sequencing an acute myeloid leukemia genome. N. Engl. J. Med. 361, 1058–1066.

25) Ley, T.J., Ding, L., Walter, M.J., McLellan, M.D., Lamprecht, T., Larson, D.E. et al. (2010) DNMT3A mutations in acute myeloid leukemia. N. Engl. J. Med. 363, 2424–2433.

26) Yoshiida, K., Sanada, M., Shiraishi, Y., Nowak, D., Nagata, Y., Yamamoto, R. et al. (2011) Frequent pathway mutations of splicing machinery in myelodysplasia. Nature 478, 64–69.

27) Wahl, M.C., Will, C.L. and Lührmann, R. (2009) The spliceosome: Design principles of a dynamic RNP machine. Cell 136, 701–718.

28) Papaemmanuil, E., Cazzola, M., Boulwood, J., Malcovati, L., Vyas, P., Bowen, D. et al. (2011) Somatic SF3B1 mutation in myelodysplasia with ring sideroblasts. N. Engl. J. Med. 365, 1384–1395.

29) Malcovati, L., Karimi, M., Papaemmanuil, E., Ambaglio, I., Jädersten, M., Jansson, M. et al. (2015) SF3B1 mutation identifies a distinct subset of myelodysplastic syndrome with ring sideroblasts. Blood 126, 233–241.

30) Yoshiida, K. and Ogawa, S. (2014) Splicing factor mutations and cancer. Wiley Interdiscip. Rev. RNA 5, 445–459.

31) Meegendorfer, M., Roller, A., Haferlach, T., Eder, C., Dicker, F., Grossmann, V. et al. (2012) SRSF2 mutations in 275 cases with chronic myelomonocytic leukaemia (CMMI). Blood 120, 3080–3088.

32) Graubert, T.A., Shen, D., Ding, L., Okeyo-Owner, T., Lunn, C.L., Shao, J. et al. (2012) Recurrent mutations in the U2A/F1 splicing factor in myelodysplastic syndromes. Nat. Genet. 44, 53–57.

33) Shao, C., Yang, B., Wei, J., Huang, J., Tang, P., Zhou, Y. et al. (2014) Mechanisms for U2AF to define 3′ splice sites and regulate alternative splicing in the human genome. Nat. Struct. Mol. Biol. 21, 997–1005.

34) Darman, R.B., Seiler, M., Agrawal, A.A., Lim, K.H., Peng, S., Aird, D. et al. (2015) Cancer-associated SF3B1 hotspot mutations induce cryptic 3′ splice site selection through use of a different branch point. Cell Rep. 13, 1033–1045.

35) Ilagan, J.O., Ramakrishnan, A., Hayes, B., Murphy, M.E., Zebari, A.S., Bradley, P. et al. (2015) U2AF1 mutations alter splice site recognition in hematological malignancies. Genome Res. 25, 14–26.

36) Kim, E., Ilagan, J.O., Liang, Y., Daubner, G.M., Lee, S.C., Ramakrishnan, A. et al. (2015) SRSF2 mutations contribute to myelodysplasia by mutant-specific effects on exon recognition. Cancer Cell 27, 617–630.
49) Phillips-Cremins, J.E., Sauria, M.E., Sanyal, A., Gerasimova, T.I., Lajoie, B.R., Bell, J.S. et al. (2013) Architectural protein subclasses shape 3D organization of genomes during lineage commitment. Cell 153, 1281–1295.

50) Welch, J.S., Ley, T.J., Link, D.C., Miller, C.A., Larson, D.E., Koboldt, D.C. et al. (2012) The origin and evolution of mutations in acute myeloid leukemia. Cell 150, 264–278.

51) Kon, A., Shih, L.Y., Minamino, M., Sanada, M., Shiraishi, Y., Nagata, Y. et al. (2013) Recurrent mutations in multiple components of the cohesin complex in myeloid neoplasms. Nat. Genet. 45, 1232–1237.

52) Solomon, D.A., Kim, T., Diaz-Martinez, L.A., Fair, J., Eklahloum, A.G., Harris, B.T. et al. (2011) Mutational inactivation of STAG2 causes aneuploidy in human cancer. Science 333, 1039–1043.

53) Noutson, M., Li, J., Ling, J., Jones, J., Wang, Y., Chen, Y. et al. (2017) The cohesin complex is necessary for epidermal progenitor cell function through maintenance of self-renewal genes. Cell Rep. 20, 3005–3013.

54) Damm, F., Chesnais, V., Nagata, Y., Yoshida, K., Scourlic, L., Okuno, Y. et al. (2013) BCOR and BCORL1 mutations in myelodysplastic syndromes and related disorders. Blood 122, 3169–3177.

55) Haferlach, T., Nagata, Y., Grossmann, V., Okuno, Y., Bacher, U., Nagae, G. et al. (2014) Landscape of genetic lesions in 944 patients with myelodysplastic syndromes. Leukemia 28, 241–247.

56) Huang, D., Nagata, Y., Grossmann, V., Radivojevitch, T., Okuno, Y., Nagae, G. et al. (2015) BRCC3 mutations in myeloid neoplasms. Haematologica 100, 1051–1057.

57) Mori, T., Nagata, Y., Makishima, H., Sanada, M., Shiozawa, Y., Kon, A. et al. (2016) Somatic PHF6 mutations in 1760 cases with various myeloid neoplasms. Leukemia 30, 2270–2273.

58) Makishima, H., Yoshida, K., Nguyen, N., Przychoden, B., Sanada, M., Okuno, Y. et al. (2013) Somatic SETBP1 mutations in 941 patients with myelodysplastic syndromes. Leukemia 27, 621–635.

59) Muro, A., Seiler, M., Sathiaseelan, V., Pance, A., Yang, Y., Agrawal, A.A. et al. (2017) Hemopoietic-specific Sf3b1K700E causes impaired erythropoiesis, aberrant splicing, and sensitivity to therapeutic spliceosome modulation. Cancer Cell 30, 404–417.

60) Kon, A., Yamazaki, S., Nannya, Y., Kataoka, K., Ota, Y., Nakagawa, M.M. et al. (2018) Physiologic expression of Sf3b1K700E causes impaired erythropoiesis, aberrant splicing, and sensitivity to therapeutic spliceosome modulation. Cancer Cell 30, 404–417.

61) Dolatshad, H., Pellagatti, A., Liberante, F.G., Llorian, M., Repapi, E., Steepeles, V. et al. (2016) Cryptic splicing events in the iron transporter ABCB7 and other key target genes in Sf3b1 mutant myelodysplastic syndromes. Leukemia 30, 2322–2331.

62) Gruber, S., Arumugam, P., Katou, Y., Kuglitsch, D., Helmhart, W., Shirahige, K. et al. (2006) Evidence that loading of cohesin onto chromosomes involves opening of its SMC hinge. Cell 127, 523–537.

63) Nasmyth, K. and Haering, C.H. (2000) Cohesin: its roles and mechanisms. Annu. Rev. Genet. 43, 525–558.

64) Phillips-Cremins, J.E., Sauria, M.E., Sanyal, A., Gerasimova, T.I., Lajoie, B.R., Bell, J.S. et al. (2013) Architectural protein subclasses shape 3D organization of genomes during lineage commitment. Cell 153, 1281–1295.

65) Welch, J.S., Ley, T.J., Link, D.C., Miller, C.A., Larson, D.E., Koboldt, D.C. et al. (2012) The origin and evolution of mutations in acute myeloid leukemia. Cell 150, 264–278.

66) Kon, A., Shih, L.Y., Minamino, M., Sanada, M., Shiraishi, Y., Nagata, Y. et al. (2013) Recurrent mutations in multiple components of the cohesin complex in myeloid neoplasms. Nat. Genet. 45, 1232–1237.

67) Solomon, D.A., Kim, T., Diaz-Martinez, L.A., Fair, J., Eklahloum, A.G., Harris, B.T. et al. (2011) Mutational inactivation of STAG2 causes aneuploidy in human cancer. Science 333, 1039–1043.

68) Noutson, M., Li, J., Ling, J., Jones, J., Wang, Y., Chen, Y. et al. (2017) The cohesin complex is necessary for epidermal progenitor cell function through maintenance of self-renewal genes. Cell Rep. 20, 3005–3013.

69) Damm, F., Chesnais, V., Nagata, Y., Yoshida, K., Scourlic, L., Okuno, Y. et al. (2013) BCOR and BCORL1 mutations in myelodysplastic syndromes and related disorders. Blood 122, 3169–3177.

70) Haferlach, T., Nagata, Y., Grossmann, V., Okuno, Y., Bacher, U., Nagae, G. et al. (2014) Landscape of genetic lesions in 944 patients with myelodysplastic syndromes. Leukemia 28, 241–247.

71) Huang, D., Nagata, Y., Grossmann, V., Radivojevitch, T., Okuno, Y., Nagae, G. et al. (2015) BRCC3 mutations in myeloid neoplasms. Haematologica 100, 1051–1057.

72) Mori, T., Nagata, Y., Makishima, H., Sanada, M., Shiozawa, Y., Kon, A. et al. (2016) Somatic PHF6 mutations in 1760 cases with various myeloid neoplasms. Leukemia 30, 2270–2273.

73) Makishima, H., Yoshida, K., Nguyen, N., Przychoden, B., Sanada, M., Okuno, Y. et al. (2013) Somatic SETBP1 mutations in 941 patients with myelodysplastic syndromes. Leukemia 27, 621–635.

74) Muro, A., Seiler, M., Sathiaseelan, V., Pance, A., Yang, Y., Agrawal, A.A. et al. (2017) Hemopoietic-specific Sf3b1K700E causes impaired erythropoiesis, aberrant splicing, and sensitivity to therapeutic spliceosome modulation. Cancer Cell 30, 404–417.

75) Kon, A., Yamazaki, S., Nannya, Y., Kataoka, K., Ota, Y., Nakagawa, M.M. et al. (2018) Physiologic expression of Sf3b1K700E causes impaired erythropoiesis, aberrant splicing, and sensitivity to therapeutic spliceosome modulation. Cancer Cell 30, 404–417.

76) Dolatshad, H., Pellagatti, A., Liberante, F.G., Llorian, M., Repapi, E., Steepeles, V. et al. (2016) Cryptic splicing events in the iron transporter ABCB7 and other key target genes in Sf3b1 mutant myelodysplastic syndromes. Leukemia 30, 2322–2331.

77) Gruber, S., Arumugam, P., Katou, Y., Kuglitsch, D., Helmhart, W., Shirahige, K. et al. (2006) Evidence that loading of cohesin onto chromosomes involves opening of its SMC hinge. Cell 127, 523–537.

78) Nasmyth, K. and Haering, C.H. (2000) Cohesin: its roles and mechanisms. Annu. Rev. Genet. 43, 525–558.
Tauro, S., Gundem, G., Van Loo, P. et al. (2013) Clinical and biological implications of driver mutations in myelodysplastic syndromes. Blood 122, 3616–3627.

65) Makishima, H., Yoshizato, T., Yoshida, K., Sekeres, M.A., Radivoyevitch, T., Suzuki, H. et al. (2017) Dynamics of clonal evolution in myelodysplastic syndromes. Nat. Genet. 49, 204–212.

66) da Silva-Coelho, P., Kroeze, L.I., Yoshida, K., Koorenhof-Scheele, T.N., Knops, R., van de Locht, L.T. et al. (2017) Clonal evolution in myelodysplastic syndromes. Nat. Commun. 8, 15099.

Chen, J., Kao, Y., Sun, D., Todorova, T., Reynolds, D., Narayanagari, S.-R. et al. (2019) Myelodysplastic syndrome progression to acute myeloid leukemia at the stem cell level. Nat. Med. 25, 103–110.

Bejar, R., Stevenson, K., Abdel-Wahab, O., Galli, N., Nilsson, B., Garcia-Manero, G. et al. (2011) Clinical effect of point mutations in myelodysplastic syndromes. N. Engl. J. Med. 364, 2496–2506.

(Received Aug. 16, 2019; accepted Jan. 15, 2020)

Profile

Seishi Ogawa was born in Okayama Prefecture in 1962 and graduated from Medical School at the University of Tokyo in 1988 and then from the Graduate School of Medicine, the University of Tokyo in 1993. He became Professor at Kyoto University in 2013. He has been the Head of the Department of Pathology and Tumor Biology since 2013 developing work on genome analysis of human cancers. Since 2017, he has been a guest professor at the Department of Medicine, Center for Hematology and Regenerative Medicine, Karolinska Institute, Stockholm, Sweden. He is a member of the Institute for the Advanced Study of Human Biology (WPI-ASHBi), Kyoto University since 2018. His scientific contributions have been central to understanding the pathogenesis of myelodysplastic syndromes, clear cell renal carcinoma, neuroblastoma, and other hematological cancers through the identification of key genetic alterations and mutations. He has exploited various approaches from the development of genome analysis tools, including the use of massively parallel sequencing technologies to comprehensively explore cancer genomes, giving rise to a series of publications (over 400 peer-reviewed papers). His work has also addressed the implication of these findings for cancer biology. He has received numerous prestigious awards, most recently, Education, Culture, Sports, Science and Technology Minister’s Commendation (2017), Uehara Prize (2017), Princes Takamatsu Cancer Research Fund Prize (2017), Takamine Prize (2018), Medal with Purple Ribbon (2018), and Erwin von Baelz Prize (2018). His work has received continuous funding support, including Gant-in-aids from the Ministry of Education, Culture, Sports, Science and Technology, Japan.