Abstract. At least one mutation is present in 70-80% of patients with myelodysplastic syndrome (MDS). Genetic alterations and other molecular biological markers have been included in the diagnostic and treatment guidelines for MDS. The aim of the present study was to analyze the association between genetic alterations and clinicopathological features among 47 Chinese patients with a novel diagnosis of MDS using a next-generation sequencing approach. The results indicated that from the 47 patients, 66.0% had genetic alterations. Furthermore, seven genes, U2 small nuclear RNA auxiliary factor 1 (23.4%), splicing factor 3b subunit (12.8%), ASXL transcriptional regulator 1 (10.6%), tet methylcytosine dioxygenase 2 (8.5%), BCL6 corepressor (8.5%), TP53 (8.5%) and DNA methyltransferase 3α (6.4%), indicated a higher prevalence of alterations in >5% of patients. Among the 16 (51.6%) patients with ≥2 mutations, 12 (75%) had mutations in different genetic functional groups. Variant allele frequencies in signaling pathways were generally low, suggesting that mutations in the corresponding genes were acquired relatively late during the evolution of the leukemic clones. The mutation prevalence rates of Janus kinase 2 and SH2B adaptor protein 3 were significantly higher in the MDS unclassified group and in the very high-risk groups with a karyotype as a prognostic indicator, respectively (both P<0.05). The mutation prevalence rates of SET binding protein 1 and enhancer of zeste 2 polycomb repressive complex 2 subunit were significantly higher in the high-risk group (both P<0.05). In summary, 66.0% of the 47 patients with a novel MDS diagnosis had a genetic mutation as detected by 127-target gene next-generation sequencing. The results for the genetic alterations in the present study will supplement the database of patients with MDS in China.

Introduction

Myelodysplastic syndrome (MDS) is a group of acquired clonal disorders that originate in the hematopoietic stem/progenitor cells, and are characterized by ineffective erythropoiesis of the bone marrow, long-term progressive refractory cytopenia, and high risk of conversion to acute leukemia (1-3).

With the continuous progress in the field of life sciences, researchers have begun to examine the pathogenesis of MDS at the gene level and have reported that more and more gene abnormalities are associated with MDS pathogenesis (4-8). Studies on MDS genetic alterations have revealed that 70-80% of the patients with MDS have at least one mutation (4-6,9). With the development of sequencing technology and its broad applications, gene mutations and other molecular biological markers have been included in the guidelines for the diagnosis and treatment of MDS (9,10). In November 2016, the National Comprehensive Cancer Network released the Clinical Practice Guidelines in Oncology: Myelodysplastic Syndromes (version 2.2017), which proposed that frequent mutations in MDS-associated genes may be suggestive of the presence of clonal hematopoiesis (11). A previous study reported that there are ~60 MDS-affected genes, which are subdivided into RNA splicing, DNA methylation, chromatin remodeling, transcription, receptors/kinases, cohesion, RAS pathway and DNA repair (9). The completion of whole-genome sequencing and targeted gene sequencing in patients with MDS has preliminarily revealed the molecular mechanism underlying the pathogenesis of MDS (12,13). In the present study, the mutant genes of patients with a novel MDS diagnosis were determined by the next-generation sequencing technology to analyze the association between the mutant genes and clinicopathological features of the patients.

Materials and methods

Diagnosis and classification criteria. The diagnostic criteria, classification criteria, and international prognostic scoring system for MDS were based on the 2007 Vienna standards...
for the diagnosis of MDS (14), the 2008 World Health Organization (WHO) classification criteria for MDS (15), and the Revised International Prognostic Scoring System (IPSS-R) (16), respectively.

Sample collection. The subjects were 47 patients with a novel MDS diagnosis in the Department of Hematology of Xiyuan Hospital of China Academy of Chinese Medical Sciences (Beijing, China) between July 15th, 2017 and December 31st, 2017. The sample included 26 females (55.3%) and 21 males (44.7%) with median age of 56 years (range, 19-82 years). The median peripheral white blood cell count, hemoglobin level, platelet count, neutrophil count and bone marrow blast percentage were 2.52 (1.9-8.2x10\(^9\)/l), 77 (35-168 g/l), 44 (3-540x10\(^9\)/l), 1.12 (0-8.48x10\(^9\)/l) and 2% (0-17.2%), respectively. According to the 2008 WHO classification criteria, 1 case of refractory anemia (RA), 27 cases of refractory cytopenia with multilineage dysplasia (RCMD), 12 cases of type 1 RA with excess blasts (RAEB-1), 5 cases of type 2 RAEB (RAEB-2), and 2 cases of MDS-unclassified (MDS-U) were included. According to the cytogenetic risk classification, 36 cases of good-prognosis karyotype, 10 cases of intermediate-risk, 10 high-risk, and 5 very high-risk cases were noted (Table I). The study protocol was approved by the Clinical Research Ethics Committee of Xiyuan Hospital, China Academy of Chinese Medical Sciences (Beijing, China). All patients provided written informed consent to participate in the study.

Next-generation sequencing. The genomic DNA (gDNA) was extracted following bone marrow or peripheral blood sample collection from patients. The concentration of gDNA was >10 ng/l, with optical density (OD)\(_{260}/OD\(_{280}\)=1.7-1.9, and the total amount was >1,000 ng. If quality inspection yielded good results, the gDNA was subsequently used for the construction of an Illumina standard library (Illumina, Inc.), and the Roche NimbleGen liquid phase hybrid capture chip was employed to perform 127-target gene sequencing (Table SI). The captured exon library was sequenced on the Illumina NextSeq 550AR platform (Illumina, Inc.), and each sample was required to have an average effective depth ≥1,000x in the target area. Using the Burrows-Wheeler Alignment algorithm version 0.7.12 (17) to compare the sequence data with the human genome (version: GRCh37), Picard version 1.115 (https://github.com/broadinstitute/picard) was used to mark the polymerase chain reaction duplicates, and the quality value of the sequence alignment results was corrected by means of BaseRecalibrator in Genome Analysis Toolkit version 3.5 (18). The MuTect2 version 3.5 software (18) was employed for the detection frequency among the 47 patients with MDS, and the threshold was a statistically significant difference.

Statistical analysis. Data analysis was performed in the Python 3.5.2 statistical software (https://www.python.org/) using the \(\chi^2\) test or Fisher's exact test. The raw values (Fig. 1), the median (minimum to maximum), if appropriate (Table I), or the maximum, upper quartile, median, lower quartile, and minimum values (Fig. 4) are presented. The odds ratio was calculated as the ratio of mutation frequency between two different groups. \(P<0.05\) was considered to indicate a statistically significant difference.

Results

Analysis of mutant genes in patients with MDS. Among the 47 patients with a novel MDS diagnosis, 31 patients had a gene mutation(s), and the overall rate of mutation prevalence was 66.0% (31/47). A total of 23 mutant genes of clinical significance were detected. According to the descending order of the detection frequency among the 47 patients with MDS, there were 11 cases with a U2 small nuclear RNA auxiliary factor 1 (U2AF1) mutation (23.4%); 6 cases with a splicing factor 3b subunit 1 (SF3B1) mutation (12.8%); 5 cases with an ASXL1 transcriptional regulator 1 (ASXL1) mutation (10.6%); 4 cases each with a mutation in tet methylcytosine dioxygenase 2 (TET2), BCL6 corepressor (BCOR) or TP53 (8.5%); 3 cases with a DNA methyltransferase 3a (DNMT3A) mutation (6.8%); 2 cases each with a mutation in serine and arginine rich splicing factor 2 (SRSF2), enhancer of zeste 2 polycomb repressive complex 2 subunit (EZH2), PHD finger protein 6
and 2 cases were respectively associated with the following categories among the 47 patients with a novel MDS diagnosis: RNA splicing (40.4%), chromatin remodeling (23.4%), DNA methylation (23.4%), a signaling pathway (10.6%), a tumor suppressor (8.5%), a transcription factor (6.4%), and the cohesin complex (4.26%) (Fig. 2).

Of the 31 patients with mutations, 16 patients had ≥2 mutations (51.6%). Among them, 7 patients had two genetic alterations, 6 patients had three genetic alterations, and 3 patients had four genetic alterations (Fig. 2). Of the 16 patients with ≥2 mutations, 4 (25%) had a synergistic mutation within the same functional group and the remaining 12 (75%) had mutations in different functional groups. The prevalence of synergistic mutations in different functional groups was significantly higher compared with that in the single functional group (P=0.036; Fig. 2).

Association analysis of the mutant genes with MDS. The results indicated that the mutations in genes EZH2 and ASXL1 (P=0.009), IDH2 and KRAS (P=0.021), IDH2 and STAG2 (P=0.043), IDH2 and SRSF2 (P=0.043), KRAS and STAG2 (P=0.043), RUNX1 and PHF6 (P=0.043), NPM1 and NRAS (P=0.043), and EZH2 and ZRSR2 (P=0.043) co-occurred and these associations were statistically significant (Fig. 3).

VAF analysis. In the present study, 23 mutant genes were detected, and the median VAFs were compared and sorted in descending order. The results revealed that the four genes associated with ‘signaling pathway’, JAK2, KRAS, NRAS and SH2B3, had a low VAF, which suggested that the corresponding mutations were acquired relatively late during the evolution of the leukemic clones (Fig. 4).

Analysis of the association between genetic alterations and clinicopathological features of the patients.

Mutant genes and MDS subtypes. The mutation prevalence rates of the JAK2 gene in subtypes RA (0/1), RCMD (0/27), RAEB-1 (0/12), and RAEB-2 (0/5) were all 0%, and in the MDS-U subtype, the mutation prevalence was 50% (1/2); the mutation prevalence rate of the JAK2 gene was significantly higher in the MDS-U subtype compared with non-MDS-U subtypes (P=0.043; Fig. 5).

Mutant genes and karyotype. The mutation prevalence rates of the SH2B3 gene in the patients with good, intermediate and very poor prognosis karyotypes were 0 (0/36), 0 (0/10), and 100% (1/1), respectively; mutation prevalence was significantly higher in the patients with the very poor prognosis karyotype (P=0.021). The mutation prevalence rates of the U2AF1 gene in the patients with good, intermediate and very poor prognosis karyotypes were 16.7 (6/36), 50.0 (5/10) and 0 (0/1), respectively; mutation prevalence tended to be highest in the patients with the intermediate prognosis karyotype (P=0.07; Fig. 6).

Mutant genes and IPSS-R. The mutation prevalence rates of the SETBP1 gene among low-risk, intermediate-risk, high-risk, and very high-risk patients were 0 (0/8), 0 (0/24), 20 (2/10), and 0 (0/5), respectively; and mutation prevalence was significantly higher in the high-risk group as defined by IPSS-R (P=0.042).
The mutation rates of \textit{EZH2} among low-risk, intermediate-risk, high-risk, and very high-risk patients were 0 (0/8), 0 (0/24), 20 (2/10), and 0 (0/5), respectively; and mutation prevalence was significantly the highest in the high-risk group on the basis of the IPSS-R score (P=0.042; Fig. 7).

**Discussion**

The positive gene mutation detection rates in the study by Haferlach \textit{et al} (9) in a 104-target gene panel, Xu \textit{et al} (20) in a 28-target gene panel, and the present study in a 127-target gene panel were 89.5, 84.0, and 66.0%, respectively, suggesting that mutations in patients with newly diagnosed MDS are relatively common. Twelve genes, \textit{TET2} (33.3%), \textit{SF3B1} (32.9%), \textit{ASXL1} (23.4%), \textit{SRSF2} (17.5%), \textit{DNMT3A} (13.1%), \textit{RUNX1} (10.6%), \textit{U2AF1} (7.7%), \textit{ZRSR2} (7.6%), \textit{STAG2} (7.5%), \textit{TP53} (6.4%), \textit{EZH2} (5.5%) and Cbl proto-oncogene (5.1%), with a mutation frequency prevalence >5% in the MDS population have been previously reported (9). In the present study, seven genes with a mutation prevalence >5% were detected, including \textit{U2AF1} (23.4%), \textit{SF3B1} (12.8%), \textit{ASXL1} (10.6%), \textit{TET2} (8.5%), \textit{BCOR} (8.5%), \textit{TP53} (8.5%) and \textit{DNMT3A} (6.8%).

The pathogenesis of MDS is associated with genetic alterations. Previous studies from China reported that the
U2AF1 mutation has one of highest prevalence rates among other mutations in the Chinese MDS population (20-22). The mutation prevalence of U2AF1 according to different Chinese research groups was 16.8% among 511 patients (21), 9.4%
among 320 patients (22) and 8.0% among 125 patients (20). Furthermore, \(U2AF1\) mutations are more common among patients with trisomy 8 (21-23). The results of the present study are in accordance with the aforementioned results, since \(U2AF1\) mutations had the highest prevalence (23.4%) among the 47 Chinese patients with MDS and tended to occur in the patients with the intermediate-prognosis karyotype. In addition, the genetic alterations with clear clinical significance and poor prognosis were prone to be accompanied by poor prognostic clinical (objective) indicators. Furthermore, in the present study population, it was also indicated that the prevalence of \(SETBP1\) mutations (4.3%) was relatively low, similar to the result (4.7%) obtained by Xu et al (22); however, a mutation in this gene has not been reported in patients with MDS in western countries (9). The prevalence of \(SRSF2\) mutations (17.5%) is reported to be higher in patients with MDS in western countries (9). The results of the present study revealed that the prevalence of \(SRSF2\) mutations (4.3%) was lower in Chinese patients with MDS, which is in accordance with the findings (3.4%) of Xu et al (22) in
China. In the present study, the prevalence rates of mutations in genes, including \textit{IDH2}, \textit{TP53}, \textit{BCOR} and \textit{EZH2}, resemble those reported by the other two groups in China (22) and in western country (9).

The most common mutant genes in these 47 patients were the splicing genes, followed by the methylation genes; consistent with previous literature data (9). Among the 16 patients with ≥2 mutations, the mutations that co-occurred were detected in gene pairs \textit{IDH2-SRSF2}, \textit{IDH2-STAG2} and \textit{EZH2-ASXL1}, which were distributed among different functional groups. This result suggested that genes in different functional groups may undergo synergistic mutations. Among the patients with newly diagnosed MDS, the median VAFs of the genes associated with ‘signaling pathway’ were relatively low, suggesting that mutations in signaling-pathway-associated genes appeared later in the clonal evolution of MDS.

In this study, a 31-year-old patient had a synergistic interaction of mutations in \textit{SETBP1}, \textit{ASXL1}, \textit{EZH2} and \textit{U2AF1}; according to the RAEB-1 subtype with a blast percentage of 6%, this patient belonged to the high-risk group on the basis of IPSS-R and had agranulocytosis status. The percentages of \textit{WT1} and \textit{PRAME} quantitative gene detection were 12.8 and 387.8, respectively, which were all associated with acute leukemia. This result is consistent with the findings of Inoue \textit{et al} (24) who concluded that the \textit{SETBP1} gene mutation can trigger the \textit{ASXL1} mutation in patients with MDS and the conversion of MDS to leukemia. Since the patient with synergistic genetic alterations had a worse prognosis, synergistic genetic alterations were targeted by therapeutic interventions in the present study. In the 2016 WHO classification system, \textit{JAK2} gene mutation was the main indicator of chronic myeloproliferative neoplasms (25). The mutation prevalence of \textit{JAK2} was significantly higher in the MDS-U subtype compared with non MDS-U subtypes in this study, suggesting that \textit{JAK2} could help with the differential diagnosis of this disease.

In myeloid neoplasms, \textit{EZH2} gene mutations often occur in MDS and myeloproliferative neoplasms, and have been associated with a poor prognosis (15). A knockout mouse model revealed that after \textit{EZH2} undergoes inactivating mutations, the number of modifications of H3k27me3 sharply diminishes, and the transcription of oncogenes, including target genes \textit{Hmga2}, \textit{Pbx3}, \textit{Lmo1} and \textit{Myc}, is inhibited, leading to MDS or myeloproliferative neoplasm-like phenotypes (26). In the present study, the \textit{EZH2} gene mutation mostly occurred in the high-risk group on the basis of the IPSS-R score, suggesting that \textit{EZH2} mutations are associated with a poor prognosis among patients with MDS, a finding that is in accordance with previous literature (15).

In summary, 66.0% of 47 Chinese patients with a novel MDS diagnosis were indicated to have a genetic mutation, as detected by the highly promising next-generation sequencing technology. The results for gene mutations in this study will supplement the database of patients with MDS in China. Due to the small sample size, the results concerning the association between genetic alterations and clinicopathological features of patients with MDS in this study require further confirmation with a larger cohort.

\section*{Acknowledgements}
Not applicable.

\section*{Funding}
This study was supported by the National Natural Science Foundation of China (grant no. 81673821), the Beijing Municipal Science & Technology Commission (grant no. Z141100006014003) and the Special Research Foundation of Central Level Public Scientific Research Institutes (grant no. ZZ10-016).
Availability of data and materials

All data and materials analyzed during the current study are included in this published article.

Authors’ contributions

XH contributed to the study design; PZ, JQ and XH wrote the manuscript; XH, PZ, WL, RQ, HK, CL, LL, YL, QZ, HW and XG conducted the clinical research; JQ and JW performed the next-generation sequencing; PZ and JQ performed the data processing and statistical analysis. All authors have read and agreed to the final version of the manuscript.

Ethics approval and consent to participate

The study protocol was approved by the Clinical Research Ethics Committee of Xiyuan Hospital, China Academy of Chinese Medical Sciences (approval no. 2017XLA019-2). All patients provided written informed consent to participate in the study.

Patient consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

References

12. ZHAO et al: GENETIC ALTERATIONS IN MYELODYSPLASTIC SYNDROME