Molecular genetic methods in the diagnosis of myelodysplastic syndromes. 
A review
Renata Lukackovaa, Maria Gerykova Bujalkovaa, Lubica Majerova, Beata Mladosieovicova

Background. Myelodysplastic syndromes (MDS) represent a heterogeneous group of premalignant hematological disorders characterized by ineffective hematopoiesis, peripheral blood cytopenias and increased risk of progression to acute leukemia. Cytogenetic analysis still plays a central role in the diagnosis of MDS, as clonal chromosomal abnormalities are observed in 30-50% of MDS patients. Despite their technical limitations, standard karyotyping and fluorescence in situ hybridization (FISH) are routinely used for identifying recurrent chromosomal rearrangements. However, using this approach means that submicroscopic and not targeted chromosomal aberrations, as well as somatic mutations and epigenetic changes remain largely undetected.

Methods and Results. Introduction of methods for the analysis of copy-number variations (CNV), including array-based technologies and Multiplex ligation-dependent probe amplification (MLPA) has provided novel insights into the molecular pathogenesis of MDS and considerably extended possibilities for genetic laboratory testing. Several novel molecular markers have been discovered and used for diagnosis and prognostic evaluation of patients with MDS. At present, mutational analysis is not routinely performed, as the clinical significance of somatic mutations in MDS has only begun to emerge. However, recently introduced Next-generation sequencing (NGS) technologies could help to elucidate the relationship between chromosomal and molecular aberrations in MDS and lead to further improvement in its diagnosis.

Conclusion. This review focuses on the advantages, limitations, clinical applications and future perspectives of three molecular methods (array-based analysis, MLPA and NGS) currently used in genetic testing and/or translational research of MDS. In conclusion, a brief summary for clinicians from the routine diagnostic point of view is given.

Key words: myelodysplastic syndromes, aCGH, SNP array, multiplex ligation-dependent probe amplification, CNV (copy-number variations), NGS, somatic mutations, prognostic markers, genetic testing

INTRODUCTION

Myelodysplastic syndromes (MDS) comprise a heterogeneous spectrum of premalignant disorders of hematopoietic stem cells characterized by hypercellular bone marrow that exhibits defective maturation of all blood cell lineages. MDS are clinically associated with ineffective hemopoiesis, peripheral blood cytopenia and increased risk of progression to acute myeloid leukemia (AML) in approximately one-third of patients. MDS are most frequently diagnosed in elderly individuals and have highly variable clinical course and prognosis. In many patients, MDS are asymptomatic and appear as an abnormality on a routine blood cell count. Symptoms develop as the production of normal blood cells is more and more compromised. These are usually non-specific and depend on the cell type affected (e.g. typically anemia). About 10% of MDS (ref.7) cases are secondary, most often due to radiation treatment or chemotherapy (particularly with alkylating agents) for primary cancer. Secondary MDS have a worse prognosis than de novo cases. Several classifications and prognostic systems are used for selection of the most appropriate therapy. The classifications of myelodysplastic syndromes are largely based on well-established morphological criteria. The widely accepted revised WHO classification (2008) takes into consideration also cytogenetic characteristics (1. number of peripheral cytopenias, 2. percentage of blasts in peripheral blood and bone marrow, 3. proportion of ring sideroblasts, 4. presence of Auer rods, 5. detection of a cytogenetic abnormality - isolated del(5q)). Thus the following MDS subtypes are defined: refractory cytopenia with unilineage or multilineage dysplasia, MDS with isolated del(5q) and unclassifiable MDS (ref.3).

The WHO-based prognostic scoring system is based on morphological classification (WHO category), cytogenetic findings (good: diploidy, loss of Y chromosome, del(5q), del(20q); poor: complex, i.e. more than three aberrations or chromosome 7 alterations; intermediate: other results) and the frequency of blood transfusions. Resulting score designates lower-risk or higher-risk prognostic category, as well as treatment options. In addition to supportive therapy (blood transfusions), in lower risk
disease, erythropoiesis-stimulating or immunosuppressive agents and lenalidomide for MDS with isolated del(5q) are the treatment of choice. DNA hypomethylating methyltransferase inhibitors (azacitidine, decitabine) are used in higher risk cases, where cytotoxic chemotherapy (similar to AML treatment), hematopoietic stem cell transplantation or experimental treatments in clinical trials are also offered. Molecular alterations with possible prognostic value in MDS identified by array-based and/or NGS analyses (e.g. c-CBL, ASXL1, and TET2, see also below) have not yet been incorporated into the prognostic models. Discovery of the relationship between the presence of del5q and response to lenalidomide resulted in the targeted treatment in these patients and discovery of RPS14 as a critical gene in the del(5)(q31.5) region. However, there is no clear molecular alteration to explain the lenalidomide response. Research of predictive biomarkers for the response to hypomethylating agents is one of the most important in the MDS field, however no presumed association between specific methylation patterns and treatment response was found.

Genetic evolution of MDS is a dynamic process characterized by multiple cycles of acquisition of genetic abnormalities and subsequent clonal selection. The frequency of clonal chromosomal aberrations is remarkably high: they occur in 30% of de novo and up to 50% of therapy-related MDS (ref.4). The karyotype of bone marrow cells remains one of the most important prognostic markers for MDS. Numerical aberrations (monosomy 5 or 7, trisomy 8, loss of the Y-chromosome) and structural abnormalities or copy number changes (deletions or amplifications) of chromosomes 5, 7, 8, and 20 are among the most common cytogenetic abnormalities observed in MDS (ref.5). The spectrum of acquired genetic lesions in de novo and secondary MDS is similar, although their frequencies are different. For example, balanced translocations are infrequent in de novo MDS, while unbalanced loss of material from chromosomes 5 and 7, a frequent event in de novo MDS, is even more common in alkylator-associated MDS (ref.6). Technology development has been a major driver of recent discoveries in MDS genetics as the field has moved from “low resolution genome scans” using conventional cytogenetics and candidate gene resequencing to array-based and NGS technology.

Genome-wide analysis tools have led to the discovery of submicroscopic genomic structural variations, so called DNA copy-number variations (CNV). Twelve percent of the human genome can be classified as structurally variant, containing deletions, duplications, inversions, or large tandem repetitions. Common (benign) CNV identified in healthy controls have been published in public databases, such as the Database of genomic variants (DGV, http://projects.tcgca.ca/variation) (ref.7). The discovery of CNV has dramatically changed our perspective of the relationship between DNA structural variations and diseases. Detection of CNV, most commonly deletions, amplifications/duplications and regions of copy-neutral loss of heterozygosity using array-based platforms has led to the identification of genes that are often mutated in myeloid malignancies. For diagnostic purposes with clinical implications, CNV should include at least the critical DNA region of the syndrome and, if known, the causative genes. The greatest interpretation challenge is discrimination between CNV that are likely to be disease causing and CNV that represent only a benign DNA polymorphism. Clinical interpretation of the data is simple and explicit, when the probe sets assess only known aberrations (MLPA analysis) in comparison with genome-wide screening by the aCGH method that may detect chromosomal abnormalities that have not been described before and their clinical interpretation is unclear. At present, several different approaches are used for detection of copy number changes including standard chromosome analysis, fluorescent in situ hybridization (FISH), array-based techniques and MLPA (ref.8).

Recent Next-generation sequencing-based study reported whole-exome sequencing of 29 MDS specimens and revealed novel somatic mutations involving genes coding for several components of the RNA splicing machinery (U2AF35, ZRSR2, SRSF2 and SF3B1) (ref.9). Further investigations using NGS technology have extended the implications of mRNA splicing abnormalities in MDS, demonstrating mutations in SF3B1 in 20% (ref.10) of 354 patients, but with a higher degree of specificity for MDS subtype with ring sideroblasts. Targeted deep-sequencing studies are performed by amplicon resequencing of mutational hotspot regions of established disease markers, e.g. RAS pathway alterations in myeloproliferative disorders, CEBPA alterations in AML or TP53 mutations in low-risk MDS (ref.11). Testing of a gene panel begins to be state-of-the-art in getting as much information as possible to define the molecular profile of an individual MDS patient. Only NGS platforms can provide this information today and in the near future. The portfolio of genes to be investigated and potentially routinely tested in MDS may be following: TET2, KRAS, CBL, ETV6, EZH2, ASXL1, SF3B1, U2AF1, UTX, WT1, SRSF2, RUNX1, FLT3, MLL-PTD and NRAS (ref.12). At present, patient-specific RUNX1 mutations have been proposed to represent clinically useful biomarkers for disease progression from MDS to secondary AML.

Epigenetic mechanisms, such as DNA methylation and histone modifications, drive clonal changes in gene expression and may mediate pathway dysfunction in neoplastic transformation. Although ethiopathogenesis of MDS is considered highly heterogeneous, subtypes of the disease can be largely explained by disordered stem cell epigenetics. The epigenetic nature of MDS may account for the observation that it is the disease most responsive to DNA methylation inhibitors: other epigenetic-acting drugs are being explored in MDS as well. DNA hypermethylation of a variety of genes may be found in patients with early-stage MDS, affecting essential cell functions: CDKN2A (cell cycle), DAPK1, RIL, CDH1 and CDH13 (adhesion and motility). Methylation of CpG dinucleotides is mediated by DNA methyltransferases, including DNMT1, DNMT3A, and DNMT3B. Heterozygous mutations of DNMT3A were identified in 8% (ref.13) of patients with MDS. These patients have poorer overall survival and more rapid progression to AML compared with patients.
without such mutations. These findings suggest a link to the treatment response to hypomethylating agents in MDS (ref.21), with azacitidine being the first approved and used22. Progression in MDS is characterized by further acquisition of epigenetic defects, as well as somatic mutations in growth-controlling genes that result in the development of AML.

**Array-based Comparative genomic hybridization (aCGH)**

Originally, comparative genomic hybridization (CGH) was a method for determining copy number gains and losses between two samples of DNA, by completely hybridizing these two differentially labeled DNAs to metaphase chromosomes23. More recently, microarray-based formats (array-CGH) using inserted genomic clones or oligonucleotides have replaced metaphase chromosomes24. aCGH enables genome-wide detection of copy-number variations at high resolution and is currently recommended as the first-line genetic test in MDS instead of standard karyotyping. Commercially available array platforms have approximately 50-fold higher resolution than cytogenetic analysis and detect chromosomal abnormalities in 15 to 20% of samples in comparison to 4% using classical karyotyping in the same patient population. The main advantage of aCGH over FISH is the ability to detect DNA copy changes simultaneously at multiple loci in a genome, i.e. to analyze hundreds to thousands of genes present on microarray in one experiment (Table 1) (ref.25,26). DNA is extracted both from patient and control samples, labeled with different fluorescent dyes and cohybridized to the chip containing arrayed genomic clones or oligonucleotide probes27. Following approximately 24-hour hybridization and washing steps, the arrays are scanned using fluorescent imaging scanner (Fig. 1) (ref.28). The fluorescence ratio of the tested vs. reference sample hybridization signals is determined at different positions along the genome and provides information about the relative DNA copy number in the tested genome compared with a normal diploid genome29.

A major advantage of aCGH in the MDS genetic testing is that it does not require proliferating cells for successful analysis. However, the proportion of malignant cells in the sample has to reach the detection limit of about 25-35% (ref.30). In general, balanced rearrangements, low-level mosaicism and polyploidy are not detectable by aCGH (ref.31). The standard resolution varies between 1-5 Mb and can be increased up to approximately 40 kb. Part of the aCGH procedure is semi-automated and results are analysed using a biostatistical algorithm32,33.

A number of different array-based approaches have been developed, aiming to improve resolution of aCGH in particular34. Currently, there are commercially available array-based DNA platforms (bacterial artificial chromosome - BAC, oligonucleotide, single-nucleotide polymorphism - SNP), synonymous names being chromosomal microarray or molecular karyotyping31,32. The genomic resolution of the different aCGH platforms is determined by spacing and length of the DNA probes (e.g. BACs 75-200kb; smaller insert clones, cosmids 30-40 kb; fosmids 40-50 kb and oligonucleotides 25-85mers) (ref.35,36). In oncohematology, microarrays can be used to refine diagnosis, identify novel disease subtypes, predict response to treatment and identify genes and pathways linked to pathogenesis, thus defining targets to rational therapy37.

**Table 1. Novel methods using array-based technology in comparison with classical techniques.**

<table>
<thead>
<tr>
<th>Method</th>
<th>Resolution</th>
<th>Sensitivity</th>
<th>Screening for new aberrations</th>
<th>Balanced aberrations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cytogenetics</td>
<td>Low</td>
<td>10%</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>FISH</td>
<td>Low</td>
<td>&gt;50-80%</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>aCGH, SNP array</td>
<td>High</td>
<td>2-30%</td>
<td>Yes</td>
<td>No</td>
</tr>
</tbody>
</table>

**Fig. 1.** Schematic overview of aCGH. Test and control DNA are labeled with green and red fluorochromes. Images of the fluorescent signals are captured and analyzed. Red spots indicate a loss of tested DNA, green spots indicate a gain of tested DNA, and yellow spots indicate the presence of equal amounts of tested and control DNA. These results can be translated into a high-resolution overview of copy number changes throughout the whole genome (modified according to Oostlander et al. 2004) (ref.33).
microarray, novel CNV in 47% (14/30) patients. Cryptic 344-kb RUNX1 deletions were found in three patients at the time of AML transformation. Other detected CNV involved 3q26.2/EVI1, 5q22/APC, 5q32/TCERG1, 12p13.1/EMP1, 12q21.3/KITLG, and 17q11.2/NFI.

SNP-based oligonucleotide arrays (array SNP) have been used for detection of genomic copy number changes and allelic imbalances in a variety of hematologic malignancies\textsuperscript{38-39}. High resolution, genome-wide coverage and minimal DNA amount requirements are advantageous for use in a clinical setting\textsuperscript{40}. Diagnostic laboratories have begun to implement array SNP to eliminate the need for multiple FISH tests, to provide high resolution analysis of technologically challenging samples and generate CNV information simultaneously with genotyping\textsuperscript{41}. MDS offers a compelling model for SNP array investigation to uncover novel clonal molecular changes with implications for pathogenesis, evaluation of their effects on disease progression and response to treatment\textsuperscript{40}.

Multiplex ligation-dependent probe amplification (MLPA)

MLPA assay is a recently developed technique for targeted CNV identification in many human genes simultaneously. The development of this analysis for multiple genetic loci has provided a new approach to routine diagnostic testing of CNV. So far, over 300 probe sets specific for a very large range of genetic disorders are commercially available. MLPA is a multiplex polymerase chain reaction (PCR)-based technique that can quantify up to 50 different genomic targets simultaneously in a single experiment through amplification of specific hybridizing probes\textsuperscript{42}. As the sequence recognized by an MLPA probe is only 50-70 nucleotides long, this assay is very useful for detection of deletions or amplifications of single exons\textsuperscript{44}. One of the major advantages is the high specificity of MLPA, because this method is able to distinguish sequences differing in length by only one nucleotide. Another advantage is the low amount of input DNA (minimum of 20-50 ng) required for a successful MLPA reaction\textsuperscript{44}. Assay shares several limitations with aCGH: balanced rearrangements, polyploidy and low proportion of tested CNV are not identified by MLPA. The detection limit is approximately 20% (ref.\textsuperscript{43}) proportion of a tested cell clone in the sample. In the case of low-quality DNA, data analysis can be difficult and the findings have to be confirmed by an alternative method, preferably aCGH (ref.\textsuperscript{39}). Up to 40-50 small specific probes are directed at DNA regions of interest and to reference regions not associated with the disease, providing a resolution greater than FISH or BAC – based aCGH (ref.\textsuperscript{39}). Each probe consists of two oligonucleotides (5’ and 3’ end-probes), that hybridize to adjacent sites of the target sequence. The short oligonucleotide contains a target-specific sequence and a universal PCR primer X. The long probe consists of a target-specific sequence, a universal PCR primer Y and a stuffer sequence of variable length in between (19-370 nucleotides) to generate the size differences necessary for electrophoretic resolution\textsuperscript{44,45}. The MLPA reaction comprises five steps (Fig. 2) and requires a thermocycler and automated genetic analyzer for fragment analysis by capillary electrophoresis. A crucial step of the MLPA assay, especially when used in the diagnostic setting (i.e. gene deletion or duplication/amplification detection) is the data analysis and interpretation of results. The most widely used is the Coffalyser software, an Excel-based program able to perform data normalization steps and necessary corrections\textsuperscript{46,47}. The results interpretation is based on the mathematical comparison between relative quantities of target DNA amplified from a tested (patient) sample vs. those of a normal (control) sample\textsuperscript{43}. In general, signals are interpreted as aberrant at cut-off values below 0.7 (deletion) and above 1.3 (duplication/amplification) (ref.\textsuperscript{39}). The utility of MLPA assay in the testing of acute leukemias and myelodysplastic syndromes has been analyzed, confirming excellent accuracy and specificity of MLPA as compared to FISH (ref.\textsuperscript{39}). The SALSA MLPA MDS kits (P144 and P145, MRC Holland) are used to detect chromosomal abnormalities commonly associated with MDS. The probe mix contains probes for chromosomes 5, 7, 8, 11, 12 (ETV6), 17 (TP53), 20 and 21 (RUNX1). MDS MLPA assay is sufficiently specific and reproducible to be
used in routine diagnostic settings as the first-line genetic screening tool\(^4\). Recently, several modifications of the original MLPA technique have been implemented. Novel applications of MLPA include e.g. methylation status determination and expression profiling\(^5\). RT-MLPA (Reverse Transcriptase-MLPA) can be used for mRNA profiling and is an alternative method to real-time PCR or expression microarray. The technique enables simultaneous detection and relative quantification of the expression of 40 genes in a single PCR reaction, studies of expression profiles of a pre-defined set of genes in a small cell populations such as those present in minimal residual disease (MRD) conditions\(^6\). In methylation-specific MLPA (MS-MLPA), the detection of CNV is combined with methylation profiling and assay is widely used for simultaneous detection of aberrant methylation patterns of CpG islands and CNV of target genes\(^7\). MS-MLPA is considered to be an equivalent alternative to other methylation assays (bisulfite sequencing, pyrosequencing and methylation-specific PCR). The most recent application combines MLPA with array-based hybridization and uses up to 200 different probes. Array-MLPA significantly increases the potential for high multiplexing and represents a valuable tool in routine diagnostics.

**Next-generation Sequencing (NGS)**

NGS (or massive parallel sequencing) approaches provide an accurate and comprehensive tool for the detection of somatic mutations in heterogeneous tumor specimens. In pivotal studies investigating hematological malignancies, NGS enabled an unbiased approach for discovering cancer-initiating mutations. Somatic mutations may influence the clinical phenotype of MDS, but are not included in current scoring systems. The principal commercially available NGS platforms (Roche 454 GS Junior, Illumina HiSeq/ MiSeq, Life Technologies Ion Proton/ PGM) vary substantially in terms of template preparation, sequencing chemistry, imaging and data analysis\(^8\). While each instrument is based on fundamentally different sequencing chemistry, the modes of parallel amplification and detection are similar, e.g. by using universal codeable adapters to facilitate the capture and subsequent amplification of fragments from mixed samples\(^9\). At present, most of the major cancer types have been already characterized by whole-genome (WGS) and whole-exome (WES) sequencing studies. To enable unbiased biological interpretation, the variants obtained by sequencing data analysis must be compared against reference data, e.g. from public SNP databases to rule out germline variants and to identify true somatic alterations. Using 116 samples from 25 patients (18 AML and 7 MDS), Kohlmann et al. demonstrated that amplicon-based NGS is highly-sensitive method for accurate detection and quantification of various RUNX1 aberrations with subsequent individualized monitoring of disease progression and treatment efficacy\(^10\). Godley et al. predict that the future diagnostic testing of hematological malignancies will consist of various PCR assays, microarray analysis and NGS (ref.\(^\text{11}\)). Jointly, these tests are capable of gene expression profiling and potential determination of causal cell lineage, as well as of identification of common chromosomal rearrangements and gene mutations\(^12\). Molecular screening by NGS and its combination with other approaches needs to be prospectively tested in large cohorts of MDS patients to define the clinical utility in disease diagnosis, prognosis (including MRD measurements) and guiding of targeted therapy. However, limitations of the NGS have to be considered; currently, the method is highly time- and cost-consuming, encompasses complicated bioinformatic analysis and is facing the problem of complex data interpretation.

**CONCLUSION**

In 2012, the International Standard Cytogenomic Array Consortium (www.iscaconsortium.org) has examined routine clinical applications and performance of aCGH techniques with the aim to replace the standard karyotyping as a first-line genetic test . Subsequently, aCGH (and accordingly MLPA) were proposed as an important step in the diagnostic algorithm of MDS (Table 2). Higher resolution, reproducibility and throughput, as well as precise mapping of aberrations are the most significant advantages of both aCGH and MLPA. Both techniques are suitable for large scale testing of chromosome aneuploidies and CNV in routine laboratory diagnosis. aCGH / MLPA have been also increasingly used as a methods of choice for diagnosis of MDS patients with unexplained genetic aberrations. The choice of assays is mostly dependent on 1) clinical vs. research utility, 2) spectrum of assessed aberrations (targeted vs. genome wide, small vs. large), 2) desired throughput level, 4) laboratory resources (aCGH equipment availability, cost- and time-effectiveness). From the laboratory analyst point of view, clinicians naturally prefer consistent laboratory assays leading to largely unequivocal results helpful in subsequent patient-related decision making. MLPA and microarrays offer well-established and time-effective analytical platform with rather uncomplicated data interpretation. The powerful combination of the MLPA with array technology should make them a useful test in screening for multiple somatic aberrations and ultimately lead to successful first attempts at tailored therapy for MDS patients. Though at present NGS techniques (and in lesser extent also epigenetic analyses) only emerge as novel tools in translational research settings, their role in
the personalized medicine of hematological malignancies is yet to come.

ABBREVIATIONS

aCGH, array-based Comparative genomic hybridization; CNV, copy-number variations; MDS, Myelodysplastic syndromes; MLPA, Multiplex ligation-dependent probe amplification; NGS, Next-generation sequencing.

ACKNOWLEDGEMENTS

Authorship contributions: RL, LM: literature search; RL, MGB: manuscript writing; RL: study design; RL: figures; RL, MGB, BM: final approval.

Conflict of interest statement: The authors state that there are no conflicts of interest regarding the publication of this article.

REFERENCES

tic test for individuals with developmental disabilities or congenital anomalies. Am J Hum Genet 2010;86:749-64.


