The genetic basis of myelodysplasia and its clinical relevance

Mario Cazzola,1,2 Matteo G Della Porta,1,3 and Luca Malcovati 1,2

1Department of Hematology Oncology, Fondazione Istituto di Ricerca e Cura a Carattere Scientifico (IRCCS) Policlinico San Matteo; 2Department of Molecular Medicine and 3Department of Internal Medicine, University of Pavia, Pavia, Italy

Running title: Molecular basis of myelodysplasia

Correspondence: Mario Cazzola, MD, Department of Hematology Oncology, Fondazione IRCCS Policlinico San Matteo, 27100 Pavia, Italy; e-mail: mario.cazzola@unipv.it
Abstract

Myelodysplasia is a diagnostic feature of myelodysplastic syndromes (MDS) but is found also in other myeloid neoplasms. Its molecular basis has been recently elucidated by means of massive parallel sequencing studies. About 90% of MDS patients carry one or more oncogenic mutations, and two thirds of them are found in individuals with normal karyotype. Driver mutant genes include those of RNA splicing (SF3B1, SRSF2, U2AF1, ZRSR2), DNA methylation (TET2, DNMT3A, IDH1/2), chromatin modification (ASXL1, EZH2), transcription regulation (RUNX1), DNA repair (TP53), signal transduction (CBL, NRAS, KRAS), and cohesin complex (STAG2). Only 4-6 genes are consistently mutated in 10% or more MDS patients, while a long tail of about 50 genes are mutated less frequently. At presentation, most patients typically have 2 or 3 driver oncogenic mutations and hundreds of background mutations. MDS driver genes are frequently mutated also in other myeloid neoplasms. Reliable genotype/phenotype relationships include the association of SF3B1 mutation with refractory anemia with ring sideroblasts, TET2/SRSF2 co-mutation with chronic myelomonocytic leukemia, and activating CSF3R mutation with chronic neutrophilic leukemia. Although both founding and subclonal driver mutations have been shown to have prognostic significance, prospective clinical trials that include the molecular characterization of the patient’s genome are now needed.
**Introduction**

Myelodysplasia is a term used in pathology for describing morphologic abnormalities, or dysplasia, in one or more of the major myeloid cell lines of hematopoiesis, and is a typical feature of myelodysplastic syndromes (MDS).\(^1\) In the World Health Organization (WHO) classification of tumors of hematopoietic and lymphoid tissues, MDS are defined as clonal hematopoietic stem cell disorders characterized by cytopenia, myelodysplasia, ineffective hematopoiesis, and increased risk of progression to acute myeloid leukemia (AML).\(^1\) Representative examples of morphologic abnormalities of myelodysplasia are reported in Figure 1.

Myelodysplasia is not restricted to MDS but may be found also in other myeloid neoplasms of the WHO classification (Table 1). Although the different subtypes of myeloid neoplasms have distinctive characteristics, they may share morphologic abnormalities. The paradigmatic example is refractory anemia with ring sideroblasts associated with marked thrombocytosis (RARS-T), which has both the myelodysplastic features of RARS and the myeloproliferative characteristics of essential thrombocythemia. This suggests that the myelodysplastic features of various myeloid neoplasms may reflect common underlying genetic lesions, and that these latter contribute to determining clinical phenotypes.

In this article, we will review the most recent advances in our understanding of the genetic basis of myelodysplasia, and will discuss its clinical relevance. The Chronic Myeloid Disorders Working Group of the International Cancer Genome Consortium has just completed a study of targeted gene sequencing in a large cohort of patients with MDS and closely related neoplasms.\(^2\) For additional information on the genomic characterization of myeloid neoplasms, the reader is referred to recent landmark studies of genomic and epigenomic landscapes of AML,\(^3,4\) and a review article in this journal.\(^5\) A very detailed analysis of what has been learned about cancer genomes in the last few years has been published recently by Vogelstein et al.\(^6\) Definitions for basic terms used in studies of the genetic basis of myeloid neoplasms are reported in Table 2.
Clonal expansion of myelodysplastic stem cells, ineffective hematopoiesis, and leukemic transformation: a working model of the pathophysiology of myelodysplasia

For a better understanding of the pathophysiology of myelodysplasia, we have summarized the current concepts in the model reported in Figure 2. Some steps of this model are still working hypotheses, which will be hopefully verified in the near future.

An essential component of the WHO definition of MDS is the clonal nature of myelodysplastic hematopoiesis. Although various approaches can be used in order to prove the existence of a clonal population of hematopoietic cells, the more straightforward one is the use of chromosomal abnormalities or discrete gene rearrangements in order to show the uniform presence of these markers in purified hematopoietic cell populations. Using this approach, chromosomal aberrations were found to be restricted to committed myeloid progenitor cells in MDS patients, suggesting that the genetic lesion occurred in a hematopoietic cell with the capacity to differentiate into mature myeloid cells.

In a landmark study, Walter et al used whole genome sequencing to identify somatic mutations in bone marrow samples from patients with AML developing from MDS, and then genotyped in each patient a bone marrow sample obtained during the antecedent MDS phase. About 85-90% of unfractionated bone marrow cells were found to be clonal in these patients, both in the MDS and in AML phase, irrespective of the number of blasts. This study formally proved that almost all cells of the bone marrow myeloid cell lines (i.e., immature red cells, granulocytic/monocytic precursors, and megakaryocytes) are clonally derived in MDS patients at any stage of the disease, not only after AML transformation.

The clonal architecture of MDS has been elegantly studied by Delhommeau et al following the identification of somatic TET2 mutations. CD34+ cells from MDS patients were fractionated into immature CD34+CD38− and mature CD34+CD38+ progenitors. While TET2 mutations were detected in only a small fraction of CD34+CD38− cells, they were present in a high proportion of more mature progenitors. This suggests that the initial somatic TET2 mutation occurred in a CD34+CD38− cell, and was then transmitted to its CD34+CD38+ progeny. A similar clonal architecture has been more recently observed also in patients with CMML.
The occurrence in an immature hematopoietic stem cell of a somatic mutation that provides survival and growth advantage (for instance, lower propensity to apoptosis) leads to formation of a local clone (step #1 Figure 2). For this clone to become fully dominant in the whole body, the mutated stem cells must have additional advantages. In adulthood, migration and trafficking of hematopoietic stem cells are of crucial importance in maintaining homeostasis of the hematopoietic system.\textsuperscript{13,14} Despite several investigations, the mechanisms by which neoplastic hematopoietic cells leave the primary site and migrate to other bone marrow districts remain largely unclear.\textsuperscript{13} Ultimately, however, mutated hematopoietic stem cells achieve full clonal dominance in the bone marrow, and the vast majority of circulating mature cells derive from the dominant clone (step #2 Figure 2).

Once the myelodysplastic clone has become fully dominant in the bone marrow, the disease may or may not become clinically apparent. For instance, a somatic \textit{SF3B1} mutation appears to be able to cause a clinical phenotype per se,\textsuperscript{15,16} while a driver \textit{TET2} mutation can determine clonal hematopoiesis without hematologic manifestations,\textsuperscript{17} suggesting that cooperating mutant genes might be required for phenotypic expression. Myelodysplastic hematopoiesis is characterized by excessive apoptosis of hematopoietic precursors, at least in patients with low-risk disease.\textsuperscript{18} Ineffective hematopoiesis, i.e., the premature intramedullary death of erythroblasts, immature granulocytes/monocytes and megakaryocytes, is primarily responsible for the defective production of mature blood cells and peripheral blood cytopenia. We must therefore assume that the somatic mutation responsible for gain-of-function at the stem cell level, involves loss-of-function at the hematopoietic precursor level (step #3 in Figure 2). RARS associated with \textit{SF3B1} mutation represents an illustrative example of gain-of-function at hematopoietic stem cell level combined with loss-of-function (excessive apoptosis of immature red cells) at hematopoietic precursor level.\textsuperscript{16} In CMML, the early clonal dominance of \textit{TET2} mutations has been shown to lead to granulo-monocytic differentiation skewing at the expense of erythroid and megakaryocytic differentiation.\textsuperscript{12}

During the natural course of the disease, patients with MDS are at high risk of progressing to AML.\textsuperscript{1} The most likely interpretation is that the acquisition of additional driver mutations leads to formation of subclones of hematopoietic cells with further impaired differentiation and/or maturation capacity. The proportion of blast cells progressively increases over time, and overt AML eventually develops (step #4 in Figure 2). This has been demonstrated by Walter et al\textsuperscript{10} in their study of the clonal architecture of secondary AML. In each of the seven
patients studied, in fact, progression to AML was characterized by the persistence of the antecedent founding myelodysplastic clone and the emergence of one or more subclones harboring new somatic mutations. Thus, a secondary AML developing from MDS is not monoclonal in the strict sense, but is instead a mosaic of several clones/ genomes with different sets of somatic mutations, illustrating the concept of intradonal or intratumoral heterogeneity.¹⁰

**Recurrent chromosomal abnormalities are mostly secondary genetic events with established clinical relevance in MDS**

Recurrent chromosomal abnormalities have been very important so far for diagnosis and prognostication of MDS. As regards diagnosis, the detection of a cytogenetic aberration in a patient with peripheral cytopenia and bone marrow dysplasia provides an important marker of clonal proliferation. By contrast, the diagnosis of MDS may be difficult in patients with a normal karyotype or non-informative cytogenetics.¹⁹

Recurrent chromosomal abnormalities are detected in about half of patients with MDS,²⁰ and the most common single cytogenetic aberrations include del(5q), trisomy 8, del(20q), and monosomy 7 or del(7q).²⁰⁻²² These are likely secondary genetic events, deriving from the genome instability caused by the founding genetic mutation.⁵ The only exception to the rule known so far is isolated del(5q), which characterizes the 5q- syndrome: in fact, haploinsufficiency for RPS14 and miR-145, mapping to the common deleted region, represents the pathophysiological basis of this MDS subtype.²³⁻²⁵

With respect to the prognostic relevance of recurrent chromosomal abnormalities, in a recent collaborative study aimed to develop the revised International Prognostic Scoring System (IPSS-R) for MDS, patient databases from international institutions were coalesced to assemble a combined database including 7012 patients.²⁶ Cytogenetic abnormalities were categorized into five prognostic subgroups that were shown to have significant prognostic relevance with different median survival and risk of evolution into AML (Table 3).

The 5-group cytogenetic risk classification reported in Table 3 was recently found to predict the outcome of allogeneic hematopoietic stem cell transplantation in MDS patients.²⁷ In
particular, patients with complex karyotype (very poor cytogenetic subgroup) had a very poor outcome after transplantation. This was also true for monosomal karyotype, defined as the karyotype of patients who had two (or more) autosomal monosomies or one monosomy in combination with other structural abnormalities.27

Thus, chromosomal aberrations will likely continue to have clinical relevance in MDS even in the era of genomic medicine. Since they basically consist in copy number changes, their detection will likely be improved by array-based karyotyping and/or by massive parallel sequencing itself.2

**Somatic gene mutations in myelodysplasia: from studies of candidate genes to whole genome sequencing**

Our understanding of the molecular basis of MDS has improved dramatically in the last 4 years. The first major breakthrough was the identification of somatic mutations of \( TET2 \) in patients with rearrangements of chromosome 4q24, where the gene maps.11,29 Subsequently, Bejar et al.30 used next-generation sequencing and mass spectrometry-based genotyping to screen mutational hotspots in 111 genes in 439 patients with MDS. In the last two years, studies of whole exome or whole genome sequencing led to the discovery of mutations in the RNA splicing machinery,15,31,32 of \( SETBP1 \) mutations in aCML,33 and of \( CSF3R \) mutations in chronic neutrophilic leukemia.34

Recent studies have performed a systematic analysis of panels of known or putative genes relevant in myelodysplasia by coupling high-throughput sample handling with massive parallel sequencing2,35 or by combining deep sequencing with array-based genomic hybridization (Seishi Ogawa, Kyoto University, email, August 9, 2013). The panels employed included from 94 to 111 candidate genes, and in the two largest studies enrolling more than 700 hundred patients, oncogenic mutations were identified with confidence in about half of the genes sequenced. Accounting for both gene mutations and chromosomal aberrations, the overall frequency of genetic lesions range from 78 to 90%. A list of the most common recurrently mutated genes in patients with MDS or MDS/MPN, based on studies published so far, is reported in Table 4.
**Spliceosome mutations: their unexpected promotion of clonal proliferation and the concept of genetic predestination**

Pre-mRNA splicing is catalyzed by the spliceosome, a macromolecule composed of 5 small nuclear RNAs associated with proteins to form particles termed small nuclear ribonucleoproteins (snRNP). More than 50% of patients with myelodysplasia carry somatic mutations in spliceosome genes encoding proteins involved in the 3’ splice site recognition and U2 snRNP function. Spliceosome mutations are rarely found in childhood myeloid neoplasms, suggesting that they are typically acquired in the elderly.

Mutations of the RNA splicing machinery are largely mutually exclusive, and are most often founding events. In fact, the mutant allele burden is typically between 40 and 50%, indicating a dominant bone marrow clone that is heterozygous for the mutation. Mutation hotspots have been shown in the three most frequently mutated genes, i.e., *SF3B1*, *SRSF2*, and *U2AF1*, and almost all described mutations are missense, with no evidence of nonsense or frameshift changes. Different patterns of missplicing associated with the above mutant genes have been described. Altogether, the available evidence suggests that spliceosome mutations affecting the 3’ splice site recognition and U2 snRNP function, are likely to create novel protein isoforms that can drive clonal dominance of mutated hematopoietic stem cells. This conclusion was totally unpredictable since, as emphasized by Vogelstein et al, a spliceosome mutation was expected to lead to a plethora of nonspecific cellular abnormalities rather than to promote clonal proliferation.

The different spliceosome mutations are associated with different phenotypes and different clinical outcomes (Table 4). Somatic *SF3B1* mutations are found almost exclusively in patients with refractory anemia with ring sideroblasts without or with thrombocytosis (RARS and RARS-T, respectively), and this clearly suggests a causal relationship between mutation and ring sideroblast formation. In addition, the vast majority of patients with *SF3B1* mutation have a good clinical outcome with a low propensity to AML transformation. *SRSF2* mutations are found mainly in patients with multilineage dysplasia and/or excess blasts and, at variance with *SF3B1* mutations, predict for high risk of leukemic evolution and poor survival. *SRSF2* mutations have been detected in about one fifth of cases of AML transformed from MPN, and in particular have been reported in about 40 to 50% of patients with CMML, where they are frequently associated with *TET2* mutations.
mutations of \textit{U2AF1} have been reported in various MDS subtypes, and found to be predictive of high risk of leukemic evolution and shorter survival.\textsuperscript{32,41}

The observation that spliceosome mutations are mainly founding mutations associated with different clinical outcomes has lead Papaemmanuil et al\textsuperscript{2} to hypothesize that they give rise to initiating clones with different genetic predestination. More specifically, the initial driver mutation would shape the future trajectory of clonal evolution through constraints on the repertoire of co-operating genetic lesions. The molecular mechanisms underlying genetic predestination remain to be defined.

\textbf{Somatic mutations in genes encoding epigenetic regulators}

In a recent review article in this journal, Issa\textsuperscript{45} has described cellular differentiation as an epigenetic process that requires specific and highly ordered DNA methylation and chromatin modification programs. The disordered differentiation of MDS is often associated with somatic mutations in genes that control DNA methylation (\textit{TET2, DNMT3A, IDH1/IDH2}) or regulate chromatin modification (\textit{ASXL1, EZH2}).\textsuperscript{5,45}

Somatic \textit{TET2} mutations were first described in patients with myeloid neoplasms in 2009.\textsuperscript{11,29} \textit{TET2} is mutated in 20-25\% of patients with MDS,\textsuperscript{2,30} and in 50 to 60\% of patients with CMML.\textsuperscript{46} In an elegant study, Busque et al\textsuperscript{47} found recurrent somatic \textit{TET2} mutations in elderly females who had clonal hematopoiesis demonstrated by X-chromosome inactivation skewing but no hematologic phenotype. This and other observations support the notion that \textit{TET2} mutation can lead per se to increased hematopoietic stem cell self-renewal and clonal myeloid proliferation.\textsuperscript{48,49} \textit{TET2} mutations are frequently found in patients with normal karyotype, and therefore represent a useful marker of clonality in these subjects.\textsuperscript{30} In addition, co-occurrence of \textit{TET2} and \textit{SRSF2} mutation is typically found in CMML.\textsuperscript{2} So far, no prognostic relevance in terms of overall survival has been clearly defined,\textsuperscript{30,50} but recent studies suggests that \textit{TET2} mutation may predict response to hypomethylating agents.\textsuperscript{51,52} Interestingly, \textit{TET2} mutations were found to be associated with reduced overall survival among patients with intermediate-risk AML.\textsuperscript{53}
In a study based on massively parallel DNA sequencing, Ley et al.\textsuperscript{54} found that DNMT3A mutations are highly recurrent in patients with de novo AML and are associated with a poor outcome. Somatic DNMT3A mutations have been later detected in about 10-15\% of patients with different subtypes of MDS.\textsuperscript{35,55,56} They are associated with unfavorable clinical outcome and more rapid progression to AML in patients with RCMD or RAEB,\textsuperscript{55} but not in those with RARS, likely because the co-occurrence of the SF3B1 mutation mitigates the negative effect of DNMT3A mutation.\textsuperscript{56} As suggested by Papaemmanuil et al.,\textsuperscript{2} this observation may indicate that some genes may only be transforming in specific genomic contexts.

Recurrent mutations in the isocitrate dehydrogenase genes IDH1 and IDH2 are found in AML and MDS.\textsuperscript{5,57} In AML, co-occurrence of NPM1 and IDH1 or IDH2 mutations is associated with a good clinical outcome.\textsuperscript{53} By contrast, IDH1 mutation has been found to be associated with a short leukemia-free survival in MDS.\textsuperscript{38}

Two genes involved in chromatin modification and regulation are recurrently mutated in MDS: ASXL1, which interacts with the polycomb-group repressive complex 1 and 2 (PRC1, PRC2),\textsuperscript{59,60} and EZH2, which belongs to PRC2.\textsuperscript{2,30,61-65} In cellular and animal models, ASXL1 mutations have been shown to promote myeloid transformation through loss of PRC2-mediated gene repression.\textsuperscript{60} ASXL1 mutations are common not only in MDS, but also in AML, CMML, and PMF, and are generally associated with poor clinical outcome in all myeloid neoplasms.\textsuperscript{46,61,62,66} Of note, ASXL1 mutation has been recently incorporated into a prognostic scoring system for CMML as a negative prognostic factor.\textsuperscript{46} Similarly, EZH2 mutations were found to be significantly associated with a shorter overall in lower risk MDS, and their incorporation in a prognostic model improved risk stratification in these patients.\textsuperscript{56}

Several studies, recently reviewed in this journal by Issa,\textsuperscript{45} have shown that MDS patients not only carry epigenetic effector mutations, but also an abnormal epigenome. Some of these patients respond to epigenetic therapy, including hypomethylating drugs (azacitidine and decitabine) and drugs that inhibit multiple histone deacetylases.\textsuperscript{45} However, although the efficacy of some of these treatments has been demonstrated in prospective clinical trials,\textsuperscript{67} we still lack reliable predictors of response to epigenetic therapy.
Somatic mutations in other cellular pathways

Acquired mutations of transcription factors have been described not only in AML but also in MDS.\textsuperscript{4,30} Somatic mutations of \textit{RUNX1} are found in about 7-8\% of all patients with MDS, and are generally associated with advanced disease, severe thrombocytopenia, and poor clinical outcome.\textsuperscript{2,30,56}

The gene \textit{TP53}, located on chromosome 17p13.1, encodes p53, which coordinates transcription programs contributing to tumor suppression, and mutant p53 proteins have been identified in various cancers.\textsuperscript{6,68} \textit{TP53} mutations are found in about 5\% of patients with MDS, mainly in subjects with advanced disease, complex karyotype, abnormalities of chromosome 17, or deletions of chromosome 5 and 7.\textsuperscript{30,55} MDS patients carrying \textit{TP53} mutation have an unfavorable clinical outcome and a high risk of leukemic evolution,\textsuperscript{2,30} and the same is true for patients with MPN.\textsuperscript{69} In particular, \textit{TP53} mutated subclones may occur at an early disease stage in MDS with del(5q), where they are associated with a lower response to lenalidomide and an increased risk of progression to AML.\textsuperscript{70}

The Ras superfamily includes small GTP-binding proteins involved in intracellular signal transduction. Several genes of this superfamily have been found to be mutated in patients with myelodysplasia, including \textit{NRAS}, \textit{KRAS}, \textit{NF1}, \textit{PTPN11}, and \textit{CBL}.\textsuperscript{2} Somatic or germline mutations of RAS pathway gene are present in about 90\% of patients with juvenile chronic myelomonocytic leukemia,\textsuperscript{71} an MDS/MPN in which secondary mutations of \textit{SETBP1} and \textit{JAK3} may cause disease progression.\textsuperscript{72}

\textit{SETBP1} encodes a protein that binds the \textit{SET} nuclear oncogene involved in DNA replication. While heterozygous de novo germline mutations in \textit{SETBP1} have been shown to be associated with the Schirzel-Giedion midface retraction syndrome,\textsuperscript{73} somatic mutations have been recently detected in patients with myeloid malignancies.\textsuperscript{33,74} In particular, \textit{SETBP1} mutations are found in about 25-30\% of patients with aCML.\textsuperscript{33,75} \textit{SETBP1} has a direct role in the transcriptional regulation of other genes,\textsuperscript{76} and that \textit{SETBP1} mutations are more often genetic events associated with disease progression in MDS.\textsuperscript{74,75,77}

\textit{CSF3R} encodes the receptor for colony stimulating factor 3. The acquisition of nonsense mutations in this gene, resulting in the expression of truncated CSF3R protein, has been
previously found to be associated with progression to MDS/AML in patients with severe congenital neutropenia.\textsuperscript{78} Activating somatic mutations in \textit{CSF3R} have been recently detected in about 90\% of patients with chronic neutrophilic leukemia, and in about 40\% of those with aCML.\textsuperscript{34} This study also showed that distinguishing between these two disorders may be difficult using the WHO criteria, while follow-up investigations demonstrated that \textit{CSF3R} and \textit{SETBP1} mutations are not mutually exclusive.\textsuperscript{34} As pointed out by Gotlib et al.,\textsuperscript{79} chronic neutrophilic leukemia and aCML are likely overlapping neoplasms: while the pathogenesis of the former is mainly characterized by \textit{CSF3R} mutation, that of aCML is likely more multifactorial.

Cohesin is a highly conserved four-subunit ring structure that encircles sister chromatids during metaphase, allowing their cohesion, and also plays critical roles in transcriptional regulation and post-replicative DNA repair.\textsuperscript{80} Somatic mutations in \textit{STAG2}, a gene of the cohesin complex, have been found in about 6\% of MDS patients.\textsuperscript{10} In a recent work, Kon et al.\textsuperscript{81} detected mutations and deletions involving various genes of the cohesin complex (\textit{STAG2}, \textit{RAD21}, \textit{SMC1A}, and \textit{SMC3}) in 8\% of patients with MDS, 10\% of those with CMML, and 12\% of those with AML. A similar frequency was previously reported in AML patients,\textsuperscript{3} suggesting that altered cohesin function plays a role in myeloid leukemogenesis.

The gene \textit{BCOR}, located on chromosome Xp11.4, encodes a corepressor of BCL6, a POZ/zinc finger transcription repressor that is required for germinal center formation and may influence apoptosis. Germline mutations of this gene are associated with oculofaciocardiodental and Lenz microphthalmia syndromes.\textsuperscript{82} Inactivating somatic mutations of \textit{BCOR} have been described in AML with normal karyotype,\textsuperscript{83} and more recently in a subset of patients with MDS.\textsuperscript{2,84} These are typical subclonal driver mutations, associated with a poor clinical outcome.\textsuperscript{84}

\textbf{Familial myelodysplastic syndromes and genetic predisposition to acquisition of somatic mutations associated with myelodysplasia: Germline \textit{GATA2} mutations}

For a deeper analysis of this subject, the reader is referred to a comprehensive review article by Liew and Owen.\textsuperscript{85} Familial syndromes predisposing to MDS or AML include bone marrow failure inherited disorders (Diamond-Blackfan, dyskeratosis congenita, severe congenital
neutropenia), DNA repair deficiency syndromes, Noonan syndrome and neurofibromatosis I, Down syndrome, and familial platelet disorder with propensity to myeloid malignancy (associated with germline mutations of \textit{RUNX1} or \textit{CEBPA}).

More recently, germline mutations of \textit{GATA2} have been described in familial syndromes characterized by predisposition to MDS and AML.\textsuperscript{86} In 3 families, subjects carrying the \textit{GATA2} (C1061T) mutation had macrocytic anemia and developed MDS/AML between the second to fifth decade of life. Various germline mutations of \textit{GATA2} have been detected in patients with the Emberger syndrome, characterized by primary lymphedema associated with a predisposition to AML.\textsuperscript{87} Finally, several germline mutations in \textit{GATA2} have been reported to be associated with the autosomal dominant and sporadic monocytopenia and mycobacterial infection (MonoMAC) syndrome, which predisposes to myeloid malignancy.\textsuperscript{88-90} Typically these patients have a combination of severe monocytopenia with mild neutropenia and marginally reduced hemoglobin level, and progression to MDS/AML typically occurs in the second or third decade of life.\textsuperscript{90} It should be noted that while germline mutations of both \textit{RUNX1} or \textit{GATA2} may predispose to MDS,\textsuperscript{91} somatic mutations of the same genes may represent mechanisms of disease progression in myeloid neoplasms.\textsuperscript{2}

\textbf{Several driver genes may cause myelodysplasia, and genotype/phenotype relationships have been defined}

In their recent article, Vogelstein et al.\textsuperscript{6} concluded that there are about 140 genes whose intragenic mutations contribute cancer. Based on currently available data, the number of potential myelodysplasia driver genes is likely somewhat lower (50 to 60). However, reliable estimates can be provided only by whole genome sequencing studies.\textsuperscript{10,35} Whole exome sequencing and candidate gene sequencing are clearly less informative, and may not allow the identification of more rarely mutated genes.

The patient cohorts studied so far are heterogeneous and all studies are basically retrospective. Taking into account these limitations, only 4 to 6 genes (\textit{SF3B1}, \textit{TET2}, \textit{SRSF2}, \textit{ASXL1}, \textit{DNMT3A}, \textit{RUNX1}) have been found to be consistently mutated in more than or about 10\% of MDS patients, while a long tail of about 40 to 50 additional genes are mutated in smaller subsets.\textsuperscript{2} In the study by Walter et al.,\textsuperscript{35} the two most frequently mutated genes were
TP53 and U2AF1: this likely reflects the fact that their patient cohort included a low proportion of low risk, and a high proportion of high risk MDS subtypes.

At presentation, most MDS patients have 2 or 3 oncogenic driver mutations and hundreds of background or passenger mutations. Considering the variant allele frequency, some mutant genes, typically those of RNA splicing and DNA methylation, appear to be mainly associated with the initial clonal proliferation, while others are mainly involved in subclonal evolution (Table 4). However, the temporal order of acquisition of driver mutations is not fixed and varies from subject to subject. Thus, the same mutant gene, e.g., TET2, may be an early driver in some patients and a subclonal driver in others. Walter et al. observed that mutations in driver genes belonging to the same biologic pathway tended not to co-occur, suggesting that a second mutation in the same pathway provides no additional growth advantage, or is not even tolerated.

According to Vogelstein et al., most human cancers are caused by 2 to 8 sequential genetic lesions that develop over the course of 20 to 30 years. In their studies of whole genome sequencing, Welch et al. estimated that as few as two key somatic mutations are needed to cause the malignant founding clone and clinically manifest AML. The available evidence suggests that this may apply also to MDS: at variance with AML, however, the genetic lesions responsible for MDS likely occur sequentially over years, rather than over months or weeks, at least in low risk subtypes with long natural history of disease, as is typically RARS.

Unfortunately only few functional studies linking the various mutant genes to a cellular or disease phenotype have been performed. In an animal model of conditional Tet2 loss, Tet2 haploinsufficiency was shown to lead to a disorder resembling human CMML. Another study reported findings suggesting that SF3B1 or Sf3b1 haploinsufficiency leads to ring sideroblast formation in human cells or heterozygous knockout mice, respectively. The CSF3R (T618I) mutation has been shown to drive a lethal myeloproliferative disorder in a murine model.

Genotype/phenotype relationships have anyhow been defined in MDS, MDS/MPN and related myeloid neoplasms. A tentative schematic representation of our current knowledge of these relationships is reported in Figure 3. Although reliable conclusions will be made possible only by prospective studies, this scheme provides a proof of concept of the potential feasibility of a molecular classification of MDS and related myeloid neoplasms. More specifically, SF3B1
mutation appears to be strictly associated with refractory anemia with ring sideroblasts with or without marked thrombocytosis, the combination of SRSF2 and TET2 mutation with chronic myelomonocytic leukemia, and activating CSF3R mutation with chronic neutrophilic leukemia. Refractory anemia has no peculiar genotype so far, raising the question as to whether it should be considered as a separate entity. Refractory cytopenia with multilineage dysplasia and refractory anemia with blast excess can be associated with different combinations of founding mutations, primarily involving genes of RNA splicing (with the only exclusion of SF3B1) or epigenetic regulation, and subclonal driver mutations.

Interestingly, as shown in Table 5, a molecular classification appears to be already feasible in MDS/MPN, a group of myeloid neoplasms so far defined on the basis of cumbersome clinical, hematologic and morphologic criteria. It is also apparent that specific driver genes are responsible for the myeloproliferative component of the different MDS/MPN, like JAK2 or MPL in RARS-T, SETBP1 in aCML, and CSF3R in chronic neutrophilic leukemia. We have previously shown that RARS-T develops from RARS through the occurrence of a subclonal driver mutation in JAK2 or MPL in the initial SF3B1 mutated clone. A similar evolutionary process likely operates in CMML and aCML, which may develop from a pre-existing myelodysplastic syndrome through the acquisition of subclonal driver mutations that cause monocytosis and granulocytic leukocytosis, respectively.

**Myelodysplasia driver genes are recurrently mutated also in other myeloid neoplasms**

The driver genes whose mutations are responsible for MDS (Table 4) are frequently mutated also in other myeloid neoplasms listed in Table 1. This is primarily true for AML, although, as underlined by Walter et al, some genes are overrepresented in MDS compared with AML and vice versa. Indeed, in the recent study by the Cancer Genome Atlas Research Network, the 20 most recurrently mutated genes in AML were FLT3, NPM1, DNMT3A, IDH1/2, TET2, RUNX1, TP53, N/KRAS, CEBPA, WT1, PTPN11, KIT, U2AF1, SMC1A, SMC3, PHF6, STAG2, and RAD21. Only half of these genes are within the 20 most recurrently mutated ones in MDS (Table 4). Although the molecular pathophysiology of MDS is different from that of AML, some AML driver genes might behave as subclonal drivers in MDS and thereby drive leukemic transformation.
Somatic mutations of \textit{CBL}, \textit{TET2}, \textit{ASXL1}, and \textit{IDH/IDH2} have been detected in advanced phase of chronic myeloid leukemia.\textsuperscript{96} Several of the genes reported in Table 4 can be mutated also in primary myelofibrosis, in combination with \textit{JAK2} (V617F) or \textit{MPL} exon 10 mutations, and co-mutation has been shown to have a negative impact on the clinical course of this\textsuperscript{66} myeloproliferative neoplasm.\textsuperscript{66} Recently, somatic mutations of \textit{TET2}, \textit{SRSF2}, \textit{ASXL1}, \textit{CBL}, and \textit{RUNX1} have been detected in about 90\% of patients with advanced mastocytosis, and overall survival was found to be significantly shorter in patients with additional mutations than in those carrying \textit{KIT} (D816V) only.\textsuperscript{97}

Myelodysplasia driver genes may also interact with somatic mutations involving lymphoid cell lines, thus giving rise to peculiar phenotypes. T-cell large granular lymphocytic (LGL) leukemia is characterized by clonal expansion of CD3+ cytotoxic T lymphocytes and may be associated with autoimmune disorders and immune-mediated cytopenias.\textsuperscript{98} The expansion of clonal T cells has been shown to be caused by somatic mutations of \textit{STAT3} or \textit{STAT5b}.\textsuperscript{99,100} Autoimmune processes have been shown to contribute to cytopenia in a subset of MDS patients, and these patients may benefit from immunosuppressive treatment.\textsuperscript{101} Interestingly, a recent study describes the occurrence of \textit{STAT3} mutated T cell clones in a subset of patients with MDS, suggesting that this may represent a molecular mechanism for the occurrence of autoimmune phenomena.\textsuperscript{102}

\textbf{Relevance of gene mutations in the diagnostic approach to myelodysplasia: Towards a molecular classification of myeloid neoplasms}

The current diagnostic approach to MDS includes peripheral blood and bone marrow morphology to evaluate abnormalities of peripheral blood cells and hematopoietic precursors (Figure 1), bone marrow biopsy to assess marrow cellularity, fibrosis and topography,\textsuperscript{103} and cytogenetics to identify non-random chromosomal abnormalities (Table 3).\textsuperscript{104}

Massive parallel sequencing has the potential of dramatically improving our approach to diagnosis of MDS as illustrated in Figure 5. The price of whole genome sequencing is expected to drop below $1,000 in the next years, while that of targeted gene sequencing\textsuperscript{2} should be definitely lower. Deep sequencing may allow the simultaneous detection of both somatic gene mutations and copy number changes, the cytogenetic abnormalities typical of MDS, in a single
assay. While whole genome sequencing is clearly more informative, massive parallel sequencing of a panel of myeloid genes is more feasible in a clinical laboratory. The genes to be sequenced may include about 50-60 myelodysplasia driver genes, genes associated with inherited disorders that predispose to MDS, and a reasonable number of germline SNPs that allow reliable detection of copy number changes from sequencing data.

Prognostic and predictive significance of driver mutations, and development of molecular models for clinical decision making

As shown in Figure 6, the current WHO classification of MDS has valuable prognostic relevance, with blast percentage and multilineage dysplasia representing the most relevant parameters from this point of view. However, the reproducibility of these latter parameters is far from optimal, and there is a need for more robust prognostic factors. The IPSS-R clearly represents a step forward but does take into account only cytogenetic abnormalities, which are secondary genetic events and not the driver lesions.

The definition of founding and subclonal driver mutations might considerably improve prognostication of MDS and more generally clinical decision-making in this field. First, the identification of the mutant gene responsible for the initial clone is relevant to clinical outcome. For instance, ring sideroblasts may be found not only in patients with a founding mutation in SF3B1, but also in those with an initiating oncogenic lesion in SRSF2. However, the median leukemia-free survival is greater than 10 years in the former versus less than 2 years in the latter. Second, the detection of subclonal driver mutations associated with small clones may allow early diagnosis of disease progression, including evolution into AML. In chronic lymphocytic leukemia, the presence of subclonal driver mutations adversely impacts clinical outcome. Third, the number of driver mutations per patient represents an important prognostic factor per se. In the recent study by Papaemmanuil et al, the median leukemia-free survival was greater than 3 years in patients with 1 or 2 driver mutations, versus less than 2 years in patients with 3 or more driver mutations.

A few studies have already suggested that incorporation of somatic mutations into prognostic scoring systems can improve prognostication of MDS. Solary and coworkers have proposed a new prognostic score for CMML that includes not only age and hematologic
parameters, but also ASXL1 mutations status. Under the aegis of the MDS Foundation, the International Working Group for Prognosis in MDS (IWG-PM) has started a research project aimed to develop an IPSS-Mol that includes clinical, hematologic and molecular parameters. Our preliminary evidence suggests that parameters such as hemoglobin level, blast count and high-risk cytogenetic abnormalities will continue to retain strong independent prognostic value (M.C, L.M. and M.G.D.P., unpublished data, August 8, 2013). Additional parameters that may significantly contribute to a refined risk assessment of MDS include gene expression profiling-based signatures.107

Finally, characterization of the patient’s genome may guide therapeutic programs, and its inclusion in prospective clinical trials is therefore of crucial importance. TET2 mutations might be associated with response to hypomethylating agents,52 while U2AF1 mutations would independently predict for poor outcome after allogeneic stem cell transplantation.108 There is considerable therapeutic potential for epigenetic-targeted therapies in AML,109 and this may be also true in MDS. Several drugs that target the spliceosome are being investigated for potential use in various malignancies,110 while drugs targeting oncogenic Ras signaling might be useful in many myeloid neoplasms. Case reports suggest that ruxolitinib may be effective in chronic neutrophilic leukemia associated with CSF3R mutation,34 and in chronic eosinophilic leukemia associated with a PCM1-JAK2 fusion gene.112 More generally, the identification of the biologic pathways that are activated by mutation might allow personalized treatment in the individual patient with myelodysplasia. It should also be considered that characterization of the patient’s genome before and after treatment may allow a correct assessment of response, in particular, the impact of treatment on clonal pattern of hematopoiesis.

For the above expectations to be realized, functional studies of the mutant genes and prospective clinical trials that include the molecular characterization of the patient’s genome are now needed.
Acknowledgments

The studies on the genetic basis of myeloid neoplasms conducted at the Department of Hematology Oncology, Fondazione IRCCS Policlinico San Matteo, and the Department of Molecular Medicine, University of Pavia, were supported by grants from Associazione Italiana per la Ricerca sul Cancro (AIRC), Fondazione Cariplo, Fondo per gli investimenti della ricerca di base (FIRB, project no. RBAP11CZLK), and Ministero dell’Istruzione, dell’Università e della Ricerca (MIUR, PRIN 2010-2011) to MC; and from Fondazione Berlucchi to MGDP. In particular, MC acknowledges funding from the AIRC Special Program “Molecular Clinical Oncology 5 per mille” (project #1005).

Authorship

Contribution: MC conceived this review article, analyzed the literature, wrote the manuscript and prepared the illustrations; MGDP and LM analyzed the literature and contributed to the manuscript preparation.

Conflict-of-interest disclosure: The authors declare no competing financial interests.

The authors wish to thank the many colleagues of the Chronic Myeloid Disorders Working Group of the International Cancer Genome Consortium, and those of the AIRC Gruppo Italiano Malattie Mieloproliferative (AGIMM).

Correspondence: Mario Cazzola, MD, Department of Hematology Oncology, Fondazione IRCCS Policlinico San Matteo, 27100 Pavia, Italy; e-mail: mario.cazzola@unipv.it
References


75. Meggendorfer M, Bacher U, Alpermann T, et al. SETBP1 mutations occur in 9% of MDS/MPN and in 4% of MPN cases and are strongly associated with atypical CML, monosomy 7, isochromosome i(17)(q10), ASXL1 and CBL mutations. Leukemia. 2013;27(9):1852-1860.


### Table 1. WHO classification of myeloid neoplasms.*†

<table>
<thead>
<tr>
<th>Major subgroups</th>
<th>Conditions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Myelodysplastic syndromes (MDS)</td>
<td>Refractory cytopenia with unilineage dysplasia (RCUD), refractory anemia with ring sideroblasts (RARS), refractory cytopenia with multilineage dysplasia (RCMD), refractory cytopenia with excess blasts (RAEB, type 1 and 2), myelodysplastic syndrome with isolated del(5q)</td>
</tr>
<tr>
<td>Myeloproliferative neoplasms (MPN)</td>
<td>Chronic myeloid leukemia (CML, BCR-ABL1 positive), chronic neutrophilic leukemia (CNL), polycythemia vera (PV), essential thrombocytopenia (ET), primary myelofibrosis (PMF), chronic eosinophilic leukemia (CEL), mastocytosis</td>
</tr>
<tr>
<td>Myelodysplastic/myeloproliferative neoplasms (MDS/MPN)</td>
<td>Chronic myelomonocytic leukemia (CMML), atypical chronic myeloid leukemia (aCML, BCR-ABL1 negative), juvenile myelomonocytic leukemia (JMML), refractory anemia with ring sideroblasts associated with marked thrombocytosis (RARS-RT)</td>
</tr>
<tr>
<td>Acute myeloid leukemia (AML)</td>
<td>Various subtypes based on blast morphology and/or underlying genetic lesion(s)</td>
</tr>
</tbody>
</table>

* Information is from Swerdlow et al.1
† The WHO classification of myeloid neoplasms includes also myeloid and lymphoid neoplasms with eosinophilia and abnormalities of PDGFRα, PDGFRβ or FGFR1.
**Table 2. Definitions for basic terms used in studies of the genetic basis of myeloid neoplasms.**

**Clone**: A group of cells that derive from a common parent cell and share its genome.

- **Malignant clone**: A clone generated by the occurrence in the parent cell of a somatic mutation that alters the cell biology and function.
- **Founding or initiating clone**: A malignant clone generated by a founding somatic mutation.
- **Subclone**: A clone generated by the occurrence of an additional driver mutation in a cell of an already established clone.
- **Clonal architecture**: Clonal composition of a neoplasm based on its clonal origin and subclonal evolution.
- **Clonal dominance**: A condition in which most cells of a tissue belong to a clone. In myelodysplastic syndromes, about 85-90% of bone marrow cells are clonally derived.
- **Subclonal evolution or expansion**: The process by which the founding malignant clone generates subclones through the acquisition of additional driver mutations. In some patients, the order of genetic changes, i.e., the temporal order of acquisition of driver mutations, can be inferred by means of massive parallel sequencing though calculation of the mutant allele burden or variant allele frequency.
- **Intraclonal (intratumoral) heterogeneity**: A founding malignant clone that has undergone subclonal evolution is no longer monoclonal in the strict sense, but is instead a mosaic of several clones/genomes with different sets of somatic mutations.

**Mutation**: A change of the nucleotide sequence of the genome.

- **Germline mutation**: A mutation inherited though a germ cell (oocyte or spermatozoan) at the time of conception, and is therefore present in all cells of a developed body.
- **Somatic mutation**: A mutation that occurs in a non-germ cell of a body after conception (the ancient Greek somatos means "of the body").
- **Nonsynonymous mutation**: A nucleotide mutation that alters the amino acid sequence of a protein.
- **Driver mutation**: A mutation that causes a selective advantage in a cell with capacity for self-renewal, leading to formation of a clone of mutated cells.
  - **Founding or initiating driver mutation**: A driver mutation that gives rise to the initial clone of a malignancy.
  - **Subclonal or cooperating driver mutation**: A driver mutation that occurs in a cell of an already established clone and generates a subclone carrying both the founding and the newly acquired mutation.
- **Background or passenger mutation**: A mutation that occurs a tissue before neoplastic transformation and has no pathophysiological significance. In most tissues, including hematopoietic stem cells, the number of passenger mutations is a function of age. Whole genome sequencing studies have shown that in hematopoietic stem cells, their number may range from about 100 to 1000, depending on age. When an initiating driver mutation gives rise to a malignant clone (neoplastic transformation), all background or passenger mutations present at that time are captured and carried forward. Additional passengers mutations can be captured during subclonal evolution.

**Mutant allele burden or variant allele frequency (VAF)**: The relative proportion of a mutant or variant allele (i.e., the allele carrying a somatic mutation) in a tissue or tumor sample. The mutant allele burden can be estimated using various approaches: i) by means of allele-specific quantitative PCR [e.g., the procedure employed for estimating the proportion of JAK2 (V617F)-mutant alleles in granulocytes from a patient with polycythemia vera]; or ii) more directly by assessing variant and wild-type reads using next-generation sequencing. A mutant allele burden or VAF of about 50% in regions of diploid DNA content in a homogenous cell population (e.g., circulating granulocytes in myeloid neoplasms) suggest a fully clonal population of cells that are heterozygous for the mutation. In myeloid neoplasms, the mutant allele burden in a bone marrow sample is most commonly between 40 and 50% due to the presence of non-myeloid cells.

**Variant allele clusters**: Group of mutation with similar VAF, potentially belonging to the same clone or subclone.

**Whole exome sequencing**: A procedure that allows to sequence all the coding regions (exomes) of a genome.

**Whole genome sequencing**: A procedure that allows to sequence the coding and non-coding regions of a genome.

*For a deeper insight into these concepts the reader is referred to recent articles by Ley and coworkers*[^34] and Vogelstein et al.*[^6]
**Table 3.** MDS cytogenetic scoring system: prognostic relevance of cytogenetic abnormalities in 7012 patients included in the International Working Group for Prognosis in MDS (IWG-PM) database.*

<table>
<thead>
<tr>
<th>Prognostic cytogenetic subgroup</th>
<th>Cytogenetic abnormalities</th>
<th>Proportion of MDS patients, %</th>
<th>Median overall survival, years</th>
<th>Median time to 25% AML evolution, years</th>
</tr>
</thead>
<tbody>
<tr>
<td>Very good</td>
<td>~Y, del(11q)</td>
<td>4</td>
<td>5.4</td>
<td>Not reached</td>
</tr>
<tr>
<td>Good</td>
<td>Normal karyotype, del(5q), del(12p), del(20q), double abnormality including del(5q)</td>
<td>72</td>
<td>4.8</td>
<td>9.4</td>
</tr>
<tr>
<td>Intermediate</td>
<td>del(7q), +8, +19, i(17q), any other single or double independent clones</td>
<td>13</td>
<td>2.7</td>
<td>2.5</td>
</tr>
<tr>
<td>High</td>
<td>~7, inv(3)/t(3q)/del(3q), double including −7/del(7q), complex: 3 abnormalities</td>
<td>4</td>
<td>1.5</td>
<td>1.7</td>
</tr>
<tr>
<td>Very High</td>
<td>Complex karyotype: &gt; 3 abnormalities</td>
<td>7</td>
<td>0.7</td>
<td>0.7</td>
</tr>
</tbody>
</table>

* Reproduced from Greenberg et al.26
**Table 4. Most common driver genes in patients with MDS and MDS/MPN.**

<table>
<thead>
<tr>
<th>Biological pathways &amp; genes</th>
<th>Frequency, %*</th>
<th>Timing of mutation acquisition†</th>
<th>Relationship between mutant gene and clinical phenotype</th>
<th>Prognostic or predictive relevance of mutant gene</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>RNA splicing</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SF3B1</td>
<td>15-30%</td>
<td>More often a founding mutation</td>
<td>Strictly associated with ring sideroblasts phenotype (RARS, RARS-T)</td>
<td>Associated with good overall survival and low risk of leukemic evolution</td>
</tr>
<tr>
<td>SRPS2</td>
<td>10-20%</td>
<td>More often a founding mutation</td>
<td>Associated with RCMD or RAEB, co-mutated with TET2 in CMML</td>
<td>Associated with poor overall survival and high risk of leukemic evolution</td>
</tr>
<tr>
<td>U2AF1</td>
<td>&lt;10%</td>
<td>More often a founding mutation</td>
<td>Mainly associated with RCMD or RAEB</td>
<td>Associated with high risk of leukemic evolution</td>
</tr>
<tr>
<td>ZRS2</td>
<td>&lt;10%</td>
<td>More often a founding mutation</td>
<td>Not defined</td>
<td>Not defined</td>
</tr>
<tr>
<td><strong>DNA methylation</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TET2</td>
<td>20-30%</td>
<td>More often a founding mutation</td>
<td>Found in all MDS subtypes, high mutation frequency (50-60%) in CMML</td>
<td>No impact on overall survival, predicts response to hypomethylating agents</td>
</tr>
<tr>
<td>DNMT3A</td>
<td>~ 10%</td>
<td>More often a founding mutation</td>
<td>Found in all MDS subtypes, co-mutated with SF3B1 in RARS</td>
<td>Associated with unfavorable clinical outcome (negative prognostic relevance mitigated by SF3B1 co-mutation in RARS)</td>
</tr>
<tr>
<td>IDH1/IDH2</td>
<td>~ 5%</td>
<td>More often a founding mutation</td>
<td>Associated with RCMD or RAEB</td>
<td>Associated with unfavorable clinical outcome</td>
</tr>
<tr>
<td><strong>Chromatin modification</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ASXL1</td>
<td>15-20%</td>
<td>More often a subclonal mutation</td>
<td>Associated with RCMD or RAEB, high mutation frequency (40%) in CMML</td>
<td>Associated with unfavorable clinical outcome in all myeloid neoplasms (MDS, MDS/MPN, MPN)</td>
</tr>
<tr>
<td>EZH2</td>
<td>~ 9%</td>
<td>More often a subclonal mutation</td>
<td>Associated with RCMD or RAEB</td>
<td>Associated with unfavorable clinical outcome in all myeloid neoplasms</td>
</tr>
<tr>
<td><strong>Transcription</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RUNX1</td>
<td>~ 10%</td>
<td>Typical subclonal mutation</td>
<td>Associated with RCMD or RAEB</td>
<td>Associated with unfavorable clinical outcome</td>
</tr>
<tr>
<td><strong>BCOR</strong></td>
<td>&lt;5%</td>
<td>Typical subclonal mutation</td>
<td>Associated with RCMD or RAEB</td>
<td>Associated with unfavorable clinical outcome</td>
</tr>
<tr>
<td><strong>DNA repair control</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TP53</td>
<td>~ 5%</td>
<td>Typical subclonal mutation</td>
<td>Associated with advanced disease and complex karyotype, mutated in 20% of patients with MDS with del(5q)</td>
<td>Associated with poor overall survival and high risk of leukemic evolution, predicts poor response to lenalidomide in MDS with del(5q)</td>
</tr>
<tr>
<td>STAG2</td>
<td>&lt;10%</td>
<td>More often a subclonal mutation</td>
<td>Associated with RCMD or RAEB, Mutated in about 10% of patients with AML</td>
<td>Associated with unfavorable clinical outcome</td>
</tr>
<tr>
<td><strong>RAS pathway</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GBL</td>
<td>&lt;5%</td>
<td>More often a subclonal mutation</td>
<td>Found in different MDS subtypes, associated with JMML in children</td>
<td>Not defined in MDS</td>
</tr>
<tr>
<td>NRAS/KRAS</td>
<td>&lt;5%</td>
<td>More often a subclonal mutation</td>
<td>Found in different MDS subtypes, associated with JMML in children</td>
<td>Not defined in MDS</td>
</tr>
<tr>
<td><strong>DNA replication</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SETBP1</td>
<td>&lt;5%</td>
<td>More often a subclonal mutation</td>
<td>Found in 2.9% of patients with aCML and in subsets of patients with advanced MDS or CMML</td>
<td>Associated with poor overall survival and high risk of leukemic evolution</td>
</tr>
<tr>
<td><strong>Receptors</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CSF3B1</td>
<td>&lt;1%</td>
<td>More often a subclonal mutation</td>
<td>Strictly associated with CNL, found in a subset of patients with aCML</td>
<td>Mutation type may predict response to specific inhibitors</td>
</tr>
</tbody>
</table>

* Approximate proportion of patients with MDS carrying the mutant gene reported in studies published so far.
† Based on values for mutant allele burden or variant allele frequency.
Table 5. Somatic mutations that characterize the different types of MDS/MPN. Chronic neutrophilic leukemia has been included because of its overlapping features with atypical chronic myeloid leukemia.

<table>
<thead>
<tr>
<th>Myeloid neoplasm according to the 2008 WHO classification</th>
<th>Main diagnostic/clinical feature(s)</th>
<th>Most common mutant gene(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chronic myelomonocytic leukemia (CMML) [classified as MDS/MPN]</td>
<td>Persistent peripheral blood monocytosis (&gt;1 × 10⁹/L).</td>
<td>Co-occurrence of TET2 and SRSF2 mutations, or combinations of ASXL1 mutations with other mutant driver genes. ASXL1 mutation is associated with poor overall survival and high risk of progression to AML.</td>
</tr>
<tr>
<td>Atypical chronic myeloid leukemia (aCML) [classified as MDS/MPN]</td>
<td>Peripheral leukocytosis (≥13 × 10⁹/L) with dysgranulopoiesis and 10% or more circulating immature granulocytes.</td>
<td>Combinations of founding mutations in various genes and subclonal mutations of SETBP1 or ASXL1.</td>
</tr>
<tr>
<td>Chronic neutrophilic leukemia (CNL) [classified as MPN]</td>
<td>Neutrophilic leukocytosis (≥25 × 10⁹/L) with less than 10% circulating immature granulocytes.</td>
<td>Activating somatic mutations of CSF3R (encoding the receptor for G-CSF) in the vast majority of patients. Oncogenic lesions can be classified as truncation or membrane proximal mutations, involving different preferential downstream kinase signaling.</td>
</tr>
<tr>
<td>Juvenile myelomonocytic leukemia (JMML) [classified as MDS/MPN]</td>
<td>Persistent peripheral blood monocytosis (&gt;1 × 10⁹/L) in children.</td>
<td>Somatic mutations of the RAS pathway (NRAS, KRAS, NF1, PTPN11, and CBL). Heterozygous germline CBL mutations may predispose to JMML. Subclonal driver mutations of SETBP1 and JAK3 may cause disease progression.</td>
</tr>
<tr>
<td>Refractory anemia with ring sideroblasts associated with marked thrombocytosis (RARS-T) [classified as a provisional entity within MDS/MPN]</td>
<td>Macrocytic anemia, ring sideroblasts, and thrombocytosis.</td>
<td>Combinations of a founding somatic mutation of SF3B1 and subclonal driver mutations of JAK2 or MPL (and likely of other as-yet-unknown genes).</td>
</tr>
</tbody>
</table>
Legends to Figures

**Figure 1.** Representative examples of morphologic abnormalities of myelodysplasia. May Grünwald Giemsa staining in all cases with the only exception of ring sideroblasts (Perls staining). Magnification from 200 to 1000X, courtesy of Erica Travaglino.

**Figure 2.** Schematic representation of our current understanding of the pathophysiology of myelodysplasia. In this example, the founding driver mutation is assumed to occur in a hematopoietic cell located in the bone marrow of the right ilium, and the sternum is shown as an anatomically separated bone marrow district to illustrate the concept of mutated stem cell migration through peripheral blood. Bone marrow microphotograph s; magnification from 600X, courtesy of Erica Travaglino.

**Figure 3.** Schematic representation of our current knowledge of genotype/phenotype relationships in MDS, MDS/MPN and a related myeloproliferative neoplasm like chronic neutrophilic leukemia. *SF3B1* mutation is strictly associated with RARS, while the combination of *SF3B1* mutation with subclonal driver mutations in *JAK2* or *MPL* is associated with RARS-T. So far, no conclusive genotype/phenotype relationship has been defined within RCUD. Various combinations of founding and subclonal driver mutations can be found in RCMD and RAEB. CMML has a relatively well-defined molecular basis, involving primarily somatic mutations of *TET2* and *SRSF2*: co-mutation of these genes is almost invariably associated with CMML, while *ASXL1* mutation involves poor outcome. Within MDS/MPN, aCML is characterized by various founding mutations plus a subclonal mutation of *SETBP1*. Activating mutations of *CSF3R* are strictly associated with chronic neutrophilic leukemia.

**Figure 4.** Current approach to diagnosis and prognostication of MDS and MDS/MPN, and the hypothetical future one based on massive parallel sequencing.

**Figure 5.** Kaplan-Meier analysis of overall survival and leukemia-free survival of 1110 patients diagnosed with MDS at the Department of Hematology Oncology, Fondazione IRCCS Policlinico San Matteo, Pavia, Italy, between 1990 and 2012. MDS patients are here stratified according to the 2008 WHO classification categories. Multilineage dysplasia and excess of blasts have a considerable impact on outcomes.
<table>
<thead>
<tr>
<th>Erythroid lineage</th>
<th>Megakaryocyte lineage</th>
<th>Granulocytic lineage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Erythroid hyperplasia</td>
<td>Micromegakaryocyte</td>
<td>Pseudo-Pelger anomaly</td>
</tr>
<tr>
<td>Megaloblastoid changes</td>
<td>Multiple separated nuclei</td>
<td>Abnormal nuclear shape</td>
</tr>
<tr>
<td>Multinuclearity</td>
<td>Small binucleated cell</td>
<td>Hypo-degranulation</td>
</tr>
<tr>
<td>Nuclear pyknosis</td>
<td>Monolobar cell</td>
<td>Myeloblasts</td>
</tr>
</tbody>
</table>

Figure 1
2. Because of their clonal advantage, mutated hematopoietic stem cells actively follow migratory pathways and progressively settle in other bone marrow districts, and eventually achieve full clonal dominance in the body. Clonal cells carry also hundreds of background or passenger mutations that have been captured at the start of clonal proliferation. At the end, nearly all bone marrow cells are clonally derived and carry the founding driver mutation plus additional passenger mutations.

3. The development of clinically apparent disease may or may not require superimposing mutations, and is primarily caused by abnormal differentiation/mutation of clonal hematopoietic progenitors and precursors. Skewed proliferation and/or defective maturation and/or excessive apoptosis most often lead to dysplasia, ineffective hematopoiesis and peripheral blood cytopenia, but leukocytosis or thrombocytosis may be observed with some mutant driver genes.

4. The acquisition by clonal cells of subclonal driver mutations, typically involving genes of chromatin modification or transcription regulation or signal transduction, involves the formation of sublines of hematopoietic cells with further impaired differentiation/maturation capacity. As a result, the proportion of blasts only progressively increases until the development of overt AML. This condition is not monoclonal in the strict sense, but is instead a mosaic of clones/genomes characterized by different sets of somatic mutations (concept of intridual or intratumoral heterogeneity).

Figure 2
Figure 3
Adult/elderly patient with cytopenia or a combination of anemia with leukocytosis and/or thrombocytosis (potential MDS or MDS/MPN)

Current approach

Peripheral blood examination, bone marrow aspiration, bone marrow biopsy, cytogenetic analysis (a marker of clonality can be identified in about 50% of patients).

Diagnosis is mainly based on morphology. Risk stratification is based on the IPSS-R.

Therapeutic decision making is still based on the IPSS risk. Follow-up monitoring is based on hematologic parameters and cytogenetics.

Future approach

Peripheral blood examination and bone marrow aspiration (bone marrow biopsy in selected patients). Candidate gene and SNP sequencing using DNA from peripheral blood granulocytes (or bone marrow) and from control tissue to detect gene mutations and copy number changes. Clonal marker(s) can be detected in 90% of patients or more, and possible germline mutations associated with a predisposing familial condition can also be identified.

Diagnosis is based on both morphology and somatic/germline mutations. Risk stratification is performed using the IPSS-Mol, a prognostic scoring system based on a combination of clinical/hematologic parameters and type and number of founding and subclonal driver mutations.

Therapeutic decision making is based not only on the individual patient’s risk, but also on his/her set of genetic lesions. When available, targeted molecular therapy is performed. Evaluation of response and follow-up monitoring is based on sequential evaluation of somatic genetic lesions and their evolutionary development.

Figure 4
Figure 5
The genetic basis of myelodysplasia and its clinical relevance

Mario Cazzola, Matteo G. Della Porta and Luca Malcovati