Molecular similarity between myelodysplastic form of chronic myelomonocytic leukemia and refractory anemia with ring sideroblasts

Véronique Gelsi-Boyer,[1,2,3,*] Nathalie Cervera,[1,*] François Bertucci,[1,3] Mandy Brecqueville,[1] Pascal Finetti,[1] Anne Murati,[1,2] Christine Arnoulet,[2] Marie-Joelle Mozziconacci,[1,2] Ken I. Mills,[4] Nicholas C. P. Cross,[5] Norbert Vey,[3,6] and Daniel Birnbaum[1]

[1]Centre de Recherche en Cancérologie de Marseille; Laboratoire d'Oncologie Moléculaire; UMR1068 Inserm; Institut Paoli-Calmettes; Marseille, France; [2]Département de BioPathologie, Institut Paoli-Calmettes, Marseille, France; [3]Faculté de Médecine, Aix-Marseille Université, Marseille, France; [4]Centre for Cancer Research and Cell Biology, Queens University Belfast, UK; [5]Faculty of Medicine, University of Southampton, UK and Wessex Regional Genetics Laboratory, Salisbury, UK; and [6]Département d'Hématologie, Institut Paoli-Calmettes, Marseille, France

ABSTRACT

Chronic myelomonocytic leukemia is similar to but a separate entity from both myeloproliferative neoplasms and myelodysplastic syndromes, and shows either myeloproliferative or myelodysplastic features. We ask whether this distinction may have a molecular basis. We established the gene expression profiles of 39 samples of chronic myelomonocytic leukemia (including 12 CD34-positive) and 32 CD34-positive samples of myelodysplastic syndromes by using Affymetrix microarrays, and studied the status of 18 genes by Sanger sequencing and array-comparative genomic hybridization in 53 samples. Analysis of 12 mRNAS from chronic myelomonocytic leukemia established a gene expression signature of 122 probe sets differentially expressed between proliferative and dysplastic cases of chronic myelomonocytic leukemia. As compared to proliferative cases, dysplastic cases over-expressed genes involved in red blood cell biology. When applied to 32 myelodysplastic syndromes, this gene expression signature was able to discriminate refractory anemias with ring sideroblasts from refractory anemias with excess of blasts. By comparing mRNAS from these two forms of myelodysplastic syndromes we derived a second gene expression signature. This signature separated the myelodysplastic and myeloproliferative forms of chronic myelomonocytic leukemias. These results were validated using two independent gene expression data sets. We found that myelodysplastic chronic myelomonocytic leukemias are characterized by mutations in transcription/epigenetic regulators (*ASXL1, RUNX1, TET2*) and splicing genes (*SRSF2*) and the absence of mutations in signaling genes. Myelodysplastic chronic myelomonocytic leukemias and refractory anemias with ring sideroblasts share a common expression program suggesting they are part of a continuum, which is not totally explained by their similar but not, however, identical mutation spectrum.

Introduction

Chronic myelomonocytic leukemia (CMML) is a malignant hematologic disease of the elderly characterized by peripheral blood monocytosis, overproduction of bone marrow monocytes with dysplasia of one or more lineages, and less than 20% of blasts in the bone marrow. Its prognosis is poor with a median survival of 12-18 months and a 15-20% risk of transformation into acute myeloid leukemia (AML).[1,2] CMML is classified by the World Health Organization (WHO) into the myelodysplastic/myeloproliferative neoplasms and, based on the number of blasts, subclassified into CMML1 and CMML2 (5-9% and 10-19%, respectively).[3] Like myelodysplastic syndromes (MDS), CMML shows dysplastic features that reflect ineffective hematopoiesis; however, dysplasia is associated with bone marrow proliferation.[4,5] Because of this duality, CMML had been separated into a myeloproliferative form (MP-CMML) and a myelodysplastic form (MD-CMML) based on a semi-arbitrary threshold of 13 x 10^9/L for peripheral white blood cell (WBC).[6] However, due to its lack of impact on outcome, this separation is not included in the WHO classification.[3] Yet, the prognosis of MD-CMML but not MP-CMML may be evaluated by the international prognostic scoring system, underlining a similarity of MD-CMML with MDS. Moreover, even if, given the limited treatments currently available, MD and MP-CMMLs have similar outcome, this situation may change with the advent of new therapies, in which case they would each need to be recognized separately.

Because CMML has both dysplastic and proliferative features it is likely that the disease is heterogeneous. We wanted to determine whether these MD and MP features may have any relevant molecular basis that may help classify and understand CMML. To this aim, we established the gene expression profiles and the mutational status of CMML and compared them to those of MDS.

Design and Methods

Patients and samples

We selected 53 CMML and 32 MDS bone marrow (BM) samples previously studied by array-comparative genome hybridization

©2013 Ferrata Storti Foundation. This is an open-access paper. doi:10.3324/haematol.2012.071506
*VG-B and NC contributed equally to this work.
The online version of this article has a Supplementary Appendix.
Manuscript received on June 4, 2012. Manuscript accepted on September 21, 2012.
Correspondence: gelsiv@marseille.fnclcc.fr

(aCGH) and sequencing of candidate genes.[7,8] According to the French-American-British (FAB)[6] and WHO[3] classifications, the CMML series was made up of 31 MP and 22 MD cases (*Online Supplementary Table S1*) and the MDS panel 8 refractory anemia (RA) with ring sideroblasts (RARS), 13 RA with excess of blasts type 1 (RAEB1) and 11 RAEB type 2 (RAEB2). CMML and MDS cases selected for gene expression profiling were collected at the time of diagnosis or in therapeutic abstention; none had been treated. All patients signed an informed consent for research and the study was approved by our institutional review board ("Comité d'Orientation Scientifique" of the Institut Paoli-Calmettes).

CD34 enrichment

Samples were enriched in CD34-positive (CD34+) cells for 12 CMML and 32 MDS cases. Leukocytes were obtained after bone marrow red cell lysis and washing with PBS, and labeled with magnetic bead-conjugated anti-CD34 monoclonal antibody (AC34 MicroBead; Miltenyi Biotec, Auburn, CA, USA). CD34+ hematopoietic stem cell populations were then purified through a miniMACS magnetic cell separation column (Miltenyi Biotec).

RNA/DNA extraction

RNAs and DNAs were extracted from whole BM CMML samples. After BM aspiration, a red cell lysis was carried out, followed by rinses with PBS. Leukocytes were processed immediately or cryopreserved at –80°C at the sample bank of the Institute and processed later. DNA and RNA were extracted using Nucleobond RNA/DNA kit from Macherey-Nagel (Düren, Germany) as recommended by the supplier. RNA from CD34+ cells were similarly extracted using Nucleobond RNA/DNA kit from Macherey-Nagel.

Sequencing of 18 candidate genes

Mutations in *ASXL1* (exon 12), *CBL* (exons 8, 9), *DNMT3A* (exons 15-23), *EZH2* (all exons), *FLT3* (exons 14, 15, 20), *IDH1/2* (exons 4), *JAK2* (exon 14), *NF1* (exons 1-50), *N/KRAS* (exons 1, 2), *PTPN11* (exons 3, 11)), *RUNX1* (exons 1-8), *SF3B1* (exons 12-16), *SUZ12* (exons 14-16), *SRSF2* (exon 2), *TET2* (exons 3-11), *U2AF35/U2AF1* (exons 2, 6), and *ZRSR2* (exons 1-11) were analyzed using BM DNA as previously described.[7-11]

Gene expression profiling

Gene expression profiles of 39 CMML (out of the 53) and 32 MDS (all from CD34+ cells) mRNAs were established. Among the 39 CMML cases, 37 were studied as BM (10 of these were also studied as CD34+) and 2 as CD34+ RNAs. In other words, 10 CMML samples were studied as both CD34+ and whole BM RNAs, and 2 as CD34+ only (12 CD34+ in total).

RNA quality and purity were assessed with Agilent Bioanalyzer 2100 (Agilent Technologies, Palo Alto, CA, USA). Preparation of cRNA, hybridizations onto Affymetrix U133 Plus 2.0 human oligonucleotide microarrays, washes and detection were carried out as recommended by the supplier and as previously described.[12] Data were analyzed by the Robust Multichip Average (RMA) method in R using Bioconductor and associated package, as previously described.[12] Before analysis, a first filtering process removed from the data set the probe sets with low and poorly measured expression as defined by an expression value inferior to 100 units, thus retaining 19,730 probe sets in the 12 CMML CD34+ data set and 23,515 probe sets in the 32 MDS CD34+ data set.

Before hierarchical clustering, a second filter, based on the intensity of standard deviation (SD >0.5), was applied and retained 9179 probe sets in the 37 CMML from the whole BM data set, 12,660 probe sets in the 12 CMML CD34+ data set, and 11,623 probe sets in the 32 MDS CD34+ data set. Filtered data were then log2-transformed and submitted to the Cluster program using data median-centered on genes, Pearson's correlation as similarity metrics and centroid linkage clustering. Results were shown using the TreeView program.

Supervised analyses identified and ranked genes that discriminate two groups of samples. A discriminating score (DS) was calculated for each of the 19,730 probe sets for the 12 CMML and of the 23,515 probe sets for the 32 MDS.[13] A 'leave-one-out' (LOO) cross-validation procedure was applied to estimate the accuracy of prediction of the signature and the validity of the supervised analysis. Functional processes and pathways were identified using Ingenuity software (Ingenuity Systems, Redwood City, CA, USA).

To test the performance of our signature on independent panels, we analyzed publicly-available external data sets[14,15] collected from NCBI/Genbank GEO database (series entry GES4619 and entry GES15061). Gene set enrichment analysis (GSEA) was carried out as reported.[16] Fisher's exact test was used when appropriate. All statistical tests were two-sided at the 5% level of significance. All statistical analyses were carried out in R (2.8.0) and its associated packages.

Results

Gene expression analysis separates MD- from MP-CMML cases

We first determined the gene expression profiles of 37 BM CMML samples. Unsupervised analysis identified two clusters (S1 and S2) including 17 and 20 cases, respectively (*Online Supplementary Figure S1*). S1 and S2 cases did not correlate with clinical or hematologic data and were not separated according to MD/MP features. We next determined the gene expression profiles of 12 available RNAs from CD34+ CMML samples (5 MD and 7 MP). Hierarchical clustering separated the 12 samples into two clusters (*Online Supplementary Figure S2*). The two clusters differed (Fisher's exact test, P=0.04) in terms of leukocytosis and overlapped the MP/MD definition: the left cluster contained 4 of 5 MD-CMML cases whereas the right cluster comprised 6 of 7 MP-CMML cases (black boxes). This showed that the MD/MP distinction has a molecular basis at the transcriptional level on a whole-genome scale.

MD-CMML over-expresses genes involved in red blood cell biology as compared to MP-CMML

To understand this MD/MP difference, we compared the gene expression profiles from the 5 MD-CMML samples to those of the 7 MP-CMML samples in a supervised analysis. A total of 122 probe sets (corresponding to 96 unique genes and 6 ESTs; *Online Supplementary Table S2*) were differentially expressed between the two forms. The accuracy of prediction and validity of our procedure was cross validated by LOO with overall accuracy of 92% (Fisher's exact test, P=0.015) with high sensitivity and specificity (86% and 100%; only one MP-CMML was misplaced) and with a theoretical number of false positive of 30.

Among the 122 probe sets, 61 were up-regulated and 61 were down-regulated in the MP samples (the top 20 up-regulated genes are listed in the *Online Supplementary Table S3*). Inspection of the list (hereafter called MD/MP CMML gene expression signature or CMML GES) showed that up-regulated genes in MD-CMML belonged to pathways and cell processes found in red blood cells: they encoded

enzymes involved in heme synthesis (ALAS2, HMBS, FECH), glycophorins (GYPA, GYPB), globins (HBA1, HBB, HBM), and proteins associated with blood groups (RHD, RHCE) and erythrocyte differentiation (TRIM10). Ingenuity analysis of this GES confirmed that the most relevant over-expressed genes in MD-CMML cases were involved in erythropoiesis (*data not shown*). Down-regulated genes in MD-CMML included *ZCCHC11/TUT4, PHC1* and *BMI1*.

We applied this CMML GES to our 12 CD34+ CMML RNAs. As expected, the MD and MP samples were separated (Fisher's exact test, $P=1x10^{-3}$) (Figure 1A). We applied this GES to the 37 BM CMML RNAs (including 10 of the 10 CD34+ samples and 27 additional samples). Two clusters were observed: in the left cluster, 16 of the 19 samples were MP-CMML, whereas in the right cluster, 13 of the 18 samples were MD-CMML (Figure 1B), supporting the validity of our MD/MP CMML GES (Fisher's exact test;

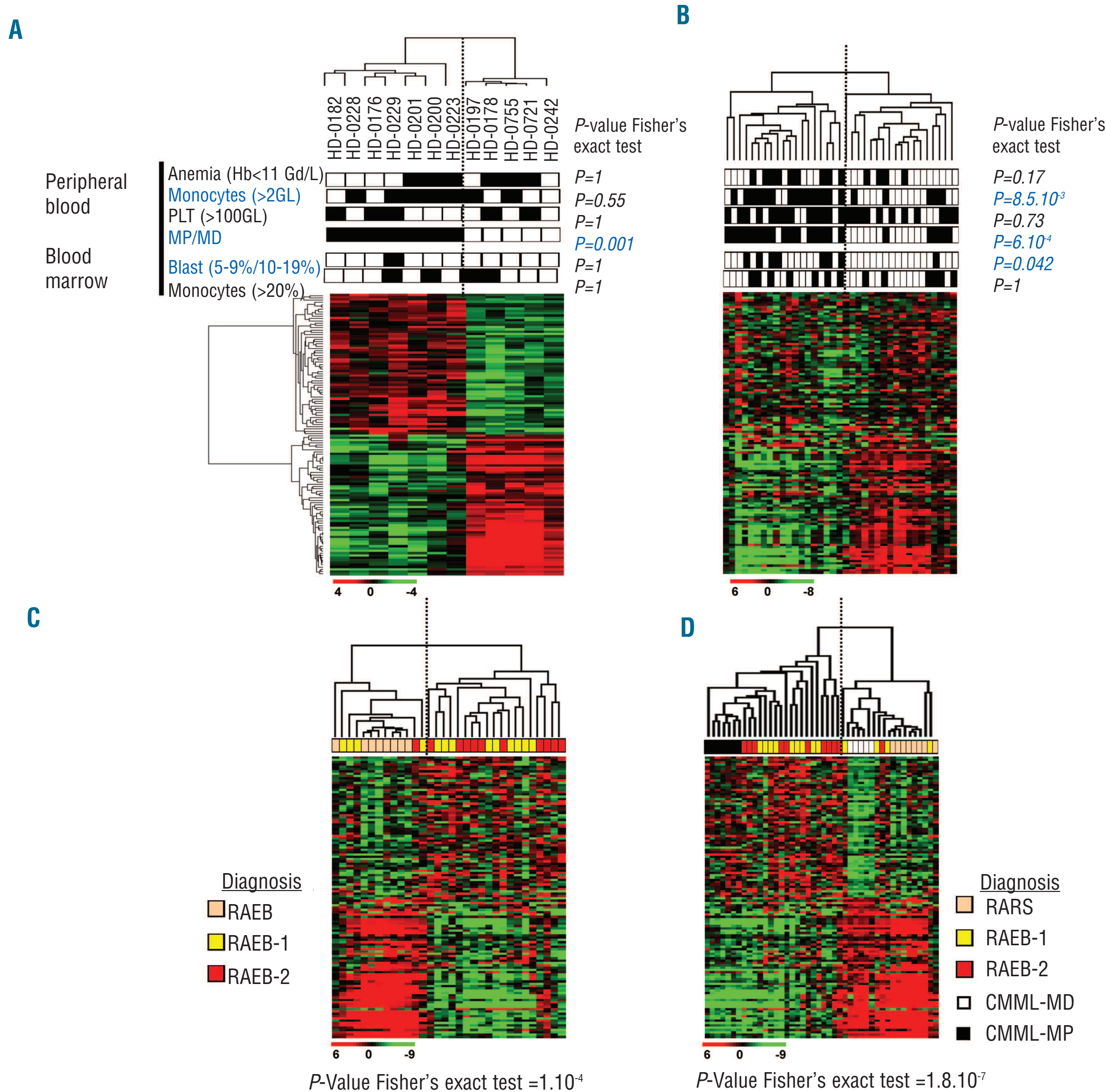

Figure 1. Classification of CMML (A, B) and MDS samples (D, E) using a CMML gene expression signature. (A) Hierarchical clustering of 12 CD34+ CMML RNAs with the 122-gene MD/MP CMML gene expression signature (GES). Each row of the data matrix represents a gene and each column represents a sample. Expression levels are depicted according to the color scale shown at the bottom. Red and green indicate expression levels respectively above and below the median. The magnitude of deviation from the median is represented by the color saturation. Below the horizontal sample dendrogram, are some histoclinical and molecular features of the samples (PLT: platelets), including the MD or MP phenotype (white for MP and black for MD). The GES (vertical dendrogram) classifies the samples into two clusters (black vertical line) associated with the MD/MP definition. The correlation of the two clusters with the histoclinical data is indicated using the *P* value of Fisher's exact test to the right of the stripes (significant *P*-values are in blue). (B) Similar to (A), but applied to 37 bone marrow CMML samples. The genes are not clustered and are ordered as in (A). (C) Hierarchical clustering of 32 CD34+ MDS samples using the 122-gene MD/MP CMML gene expression signature (GES). Legend as in (A). RARS and RAEB are indicated by colored boxes under the dendrogram. The GES significantly distinguishes the RARS from the RAEB samples. (D) As in (C), but applied to the 32 mRNAs from CD34-positive MDS samples and 12 mRNAs from CD34-positive CMML samples. The genes are not clustered and are ordered as in (A). The CMML GES separates the RARS and MD-CMML samples from the MP-CMML and RAEB samples. MD/MP forms are indicated by black and white boxes as indicated.

$P=6x10^{-4}$). Using GSEA, we confronted the gene expression profiles of our 37 BM CMML mRNAs to the 122 gene signature. We found a significant enrichment in the red cell genes of this signature in the MD-CMML samples (Enrichment Score=0.76; Normalized Enrichment Score=2.08; FDR q-value<0.01).

CMML gene expression signature classifies MDS samples

Overexpression of genes involved in red cell biology has been observed in previous gene expression analyses of RARS samples.[14,17] When applied to our 32 CD34+ MDS samples, the CMML GES perfectly separated RARS from RAEB samples (Fisher's exact test, $P=1x10^{-4}$, Figure 1C). When the CMML GES was applied to the pool of 12 CMML and 32 MDS CD34+ samples the MD-CMMLs clustered with the RARS samples and the MP-CMMLs with the RAEB samples (Figure 1D) (Fisher's exact test, $P=1.8x10^{-7}$). These results showed that MD-CMML and RARS share gene similar expression programs.

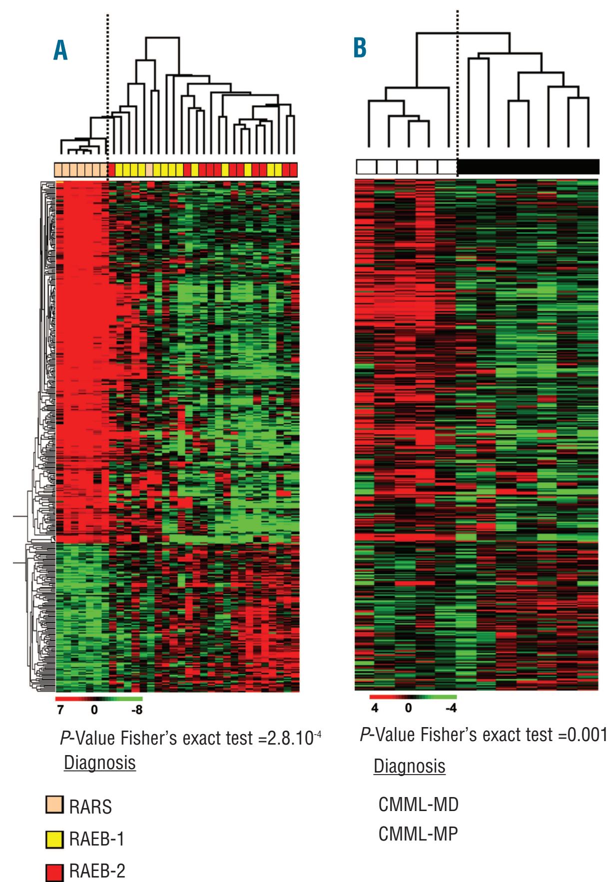

Figure 2. Classification of MDS samples using an MDS gene expression signature. (A) Hierarchical clustering of 32 CD34+ MDS samples using the 428-gene MDS gene expression signature. Legend as in Figure 1A. MDS classes are indicated by colored boxes in the stripe under the dendrogram. The GES significantly distinguishes the RARS from the RAEB samples. (B) As in (A), but applied to the 12 CD34-positive CMML samples. The genes are not clustered and are ordered as in (A). The GES significantly distinguishes the MD from the M¨-CMML samples. MD/MP forms are indicated by white and black boxes in the stripe under the dendrogram.

MDS gene expression signature classifies CMML samples

Then we derived an MDS GES by comparing the gene expression profiles of our 8 RARS to those of our 24 RAEB samples. A total of 428 probe sets (295 unique genes and 25 ESTs; *Online Supplementary Table S4*) were differentially expressed between RARS and RAEB (hereafter called MDS GES). The accuracy of prediction and validity of our procedure was cross-validated by LOO with overall accuracy of 78% (Fisher's exact test, $P=6x10^{-4}$) with high sensitivity and specificity (72% and 100%, respectively) and with a theoretical number of false positive of 5.

A total of 304 probe sets were up-regulated and 124 were down-regulated in the RARS cases (the top 20 up-regulated genes are listed in *Online Supplementary Table S3*). Up-regulated genes in RARS belonged to the pathway and cell processes found in red blood cells (*Online Supplementary Figure S3*): they encoded proteins involved in heme synthesis (ALAS2, ALAD, HMBS, UROD, CPOX, PPOX, FECH), iron-sulfur cluster biogenesis (SLC25A37, GLRX5), mitochondrial biology (TRAK2), antioxidant defense (HAGH), glycophorins (GYPA, GYPB, GYPE), globins (HBA1, HBA2, HBB, HBD, HBG1, HBM, HBQ1), proteins associated with blood groups (RHCE, RHD, DARC, KEL), red cell structure (ANK1, EPB42, EPB49, ERMAP, SPTA1, SLC4A1), differentiation (TRIM10) and regulation (GATA1, KLF1, TAL1, EPOR). Ingenuity analysis of this GES confirmed that the most relevant over-expressed genes in RARS were involved in erythropoiesis (*data not shown*). When applied to our 32 CD34+ MDS RNAs, the MDS GES perfectly separated the RARS from the RAEB samples (Figure 2A), as expected. More surprisingly, when applied to the 12 CD34+ (Figure 2B) or BM CMML mRNAs (Fisher's exact test, $P=8.4.10^{-4}$; *data not shown*) it perfectly separated the MD-CMML from the MP-CMML samples.

Thus, MD-CMML and RARS share expression of genes involved in erythropoiesis and red blood cell biology, with a total of 38 probe sets corresponding to 25 unique genes common to the MD/MP CMML and MDS GESs (*Online Supplementary Table S3*).

External validation on two independent data sets

To validate these unexpected results, we applied our two GESs (CMML GES and MDS GES) to two external published data sets (a CMML and an MDS data set)[14,15] and a GES obtained from one of these external data sets to our CMML and MDS samples.

First, we studied the external data sets with our two GESs. The first data set corresponded to the expression profiles of 25 CMML BM samples profiled using Affymetrix microarrays.[15] Unfortunately, information on MD/MP forms was not associated with the data. Our CMML GES separated these external CMML cases into two groups, one of which over-expressed red blood cell genes (*Online Supplementary Figure S4A*, stripe 1). All but one of the 25 CMML samples was similarly sorted with our MDS GES (*Online Supplementary Figure S4B*, stripe 2). We studied a second external data set, consisting in the expression profiles of 66 CD34+ MDS samples studied by Affymetrix microarrays.[14] Our MDS GES separated the RARS from the other MDS classes (Figure 3A), as did our CMML GES (Figure 3B).

Second, we derived a third GES (hereafter called MDS-ext GES) by comparing RARS and RAEB samples from the MDS external data set.[14] This MDS-ext GES contained 738

probe sets (597 unique genes and 28 ESTs). The genes found over-expressed in RARS are involved in red blood cell biology. This GES was able to separate our RARS from RAEB samples (*Online Supplementary Figure S4C*) and our MD-CMML from MP-CMML CD34-positive samples (Figure 3C). When applied to the CMML external data set,[15] it separated the 25 samples (stripe 3) in the same way as did our CMML and MDS GESs (*Online Supplementary Figure S4D*).

Thus, GESs obtained from comparison from either CMML or MDS were similarly able to distinguish CMML and MDS classes, showing that MD-CMML and RARS share common molecular features. The CMML GES, MDS GES and MDS-ext GES had 26 probe sets in common (corresponding to 16 genes and 1 EST)(Figure 3D), all overexpressed in MD-CMML and RARS and involved in red cell biology (*Online Supplementary Table S3*). This small core GES separated MD-CMML from MP-CMML (Figure 3E) and RARS from RAEB (*Online Supplementary Figure S5*) as efficiently as the three larger GESs.

Analysis of mutated genes in CMML and MDS

These results showed a molecular similarity between MD-CMML and RARS. Could this similarity be the result of gene mutations common and specific to the two diseases? We[7,8,18] and others[19-36] have previously studied several leukemogenic genes in CMML and RARS. However, several of those (e.g. *ASXL1, RUNX1, TET2*) are neither specific of MD-CMML nor of RARS and, therefore, can account neither for the similarity between the two diseases nor for the differences from the other myeloid malignancies.

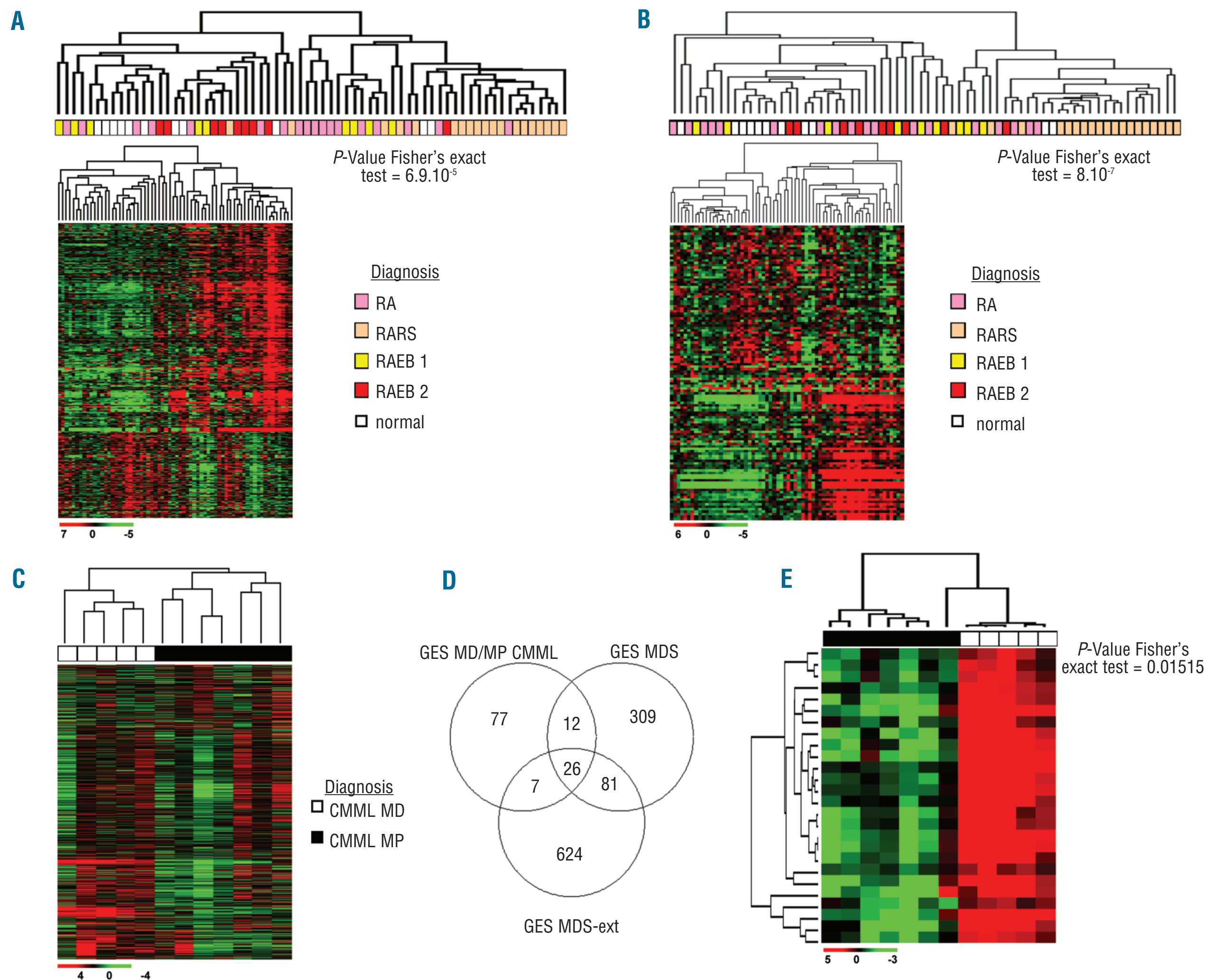

Figure 3. Independent validation of the gene expression signatures. (A) Hierarchical clustering of 66 Pellagatti's CD34+ MDS external samples[14] using our 428-gene MDS GES. Legend similar to Figure 1A. MDS classes are indicated by colored boxes in the stripe under the dendrogram. The GES significantly separates the RARS from the other MDS classes. (B) Similar to (A), but using our 122-gene MD/MP CMML. The GES significantly separates the RARS from the other MDS classes. (C) Hierarchical clustering of our 12 CD34-positive CMML samples using the MDS-ext GES. MD/MP forms are indicated by black and white boxes in the stripe under the dendrogram. The GES significantly separates the MD-CMML samples from the MP-CMML samples. (D) Venn diagram showing the overlap between the three GES: 26 genes overlap the three GES. (E) Hierarchical clustering of the 12 mRNAs from CD34+ CMML samples using these 26 overlapping genes. This 26-gene GES separates the MD-CMML from the MP-CMML samples.

We studied the sequence of 18 genes involved in the regulation of transcription (*ASXL1, DNMT3A, EZH2, IDH1, IDH2, RUNX1, SUZ12, TET2*), splicing (*SF3B1, SRSF2, U2AF1, ZRSR2*) and signaling (*CBL, FLT3, JAK2, KRAS, NRAS, PTPN11*) in our series of 53 CMMLs. Results are shown in Figure 4, *Online Supplementary Table S5* (see also *Online Supplementary Figure S1*). The samples had been studied by aCGH7,[18] and deletions of these genes were taken into account when appropriate. In CMML, splicing mutations have been recently described,[31] but how mutations in 4 splicing genes combine with alterations in the other 14 genes has not yet been reported. Only seven samples (3 MD and 4 MP) did not show any mutation in the genes studied (13%). Among transcription/epigenetic regulators, *ASXL1, RUNX1* and *TET2* were the most frequently mutated genes.

Mutations in these genes were found in both MP and MD cases. *DNMT3A* mutations were only found in MD-CMML and *EZH2* mutations in MP-CMML; however, the number of these events was low, confirming findings of previous studies.[27-30] MD-CMML showed only one CBL mutation and one NF1 deletion, whereas MP-CMML showed 14 mutations in signaling genes and one NF1 deletion. Thus, mutations in these selected signaling genes account for a first difference between MD-CMML and MP-CMML.

Except in one case, mutations in splicing genes were mutually exclusive, as previously described[31] and recently reviewed.[37] We found one mutation in *SF3B1*, 3 in *U2AF1*, 2 in *ZRSF2* and 24 in *SRSF2*. In total, 15 of 22 (68%) MD and 15 of 31 (45%) MP cases were mutated in the splicing genes studied. They were not differentially distributed between the two forms (P=0.17): the 22 MD cases showed no *SF3B1* and 13 *SRSF2* mutations (59%) whereas the 31 MP cases showed one *SF3B1* and 11 *SRSF2* mutations (38.7%). In addition, *U2AF1* and *ZRSR2* were mutated in one and one MD-CMML cases and in 2 and one MP-CMML cases, respectively.

To evaluate the possible relationships between gene mutations and gene expression we looked at the expression of the mutated genes (*Online Supplementary Figure S2B*). As expected, since they were not found in the GES, there was no difference in expression of the 18 genes between MD and MP-CMML. Gene mutations in our series of MDS cases have been described in previous studies[7,8,11] and will not be detailed here again except for the sake of comparison with CMML. We found 6 mutations in *SF3B1* and 3 in *SRSF2* in our series of 32 MDSs (*Online Supplementary Table S6*). They were differentially distributed between RARS and RAEB cases: the 8 RARS showed 5 *SF3B1* and one *SRSF2* mutations whereas the 24 RAEB cases showed one *SF3B1* and 2 *SRSF2* mutations. Three RAEB cases were also mutated in *U2AF1* and *ZRSR2*.

Thus, both MD-CMML and RARS display frequent mutations in genes encoding components of the RNA splicing machinery. However, MP-CMML cases also show alterations of this process. RARS and MD-CMML also show more mutations in *DNMT3A*.[38]

Discussion

We studied CMML by gene expression profiling and by sequencing analyses of 18 candidate genes. Unsupervised analysis of mRNAs from CD34-positive cells separated CMML into two molecular subtypes that overlapped with the MD and MP forms initially distinguished by the FAB classification. Supervised analysis established an MD/MP CMML GES characterized by the overexpression in MD-CMML of genes involved in red blood cells. The comparison of RARS and RAEB samples allowed the establishment of an MDS GES that was also characterized by the overexpression in RARS of the same genes and functions, as observed in previous studies.[14] The CMML GES recognized the RARS and also the MDS GES recognized the MD-CMML samples. Thus, MD-CMML and RARS share

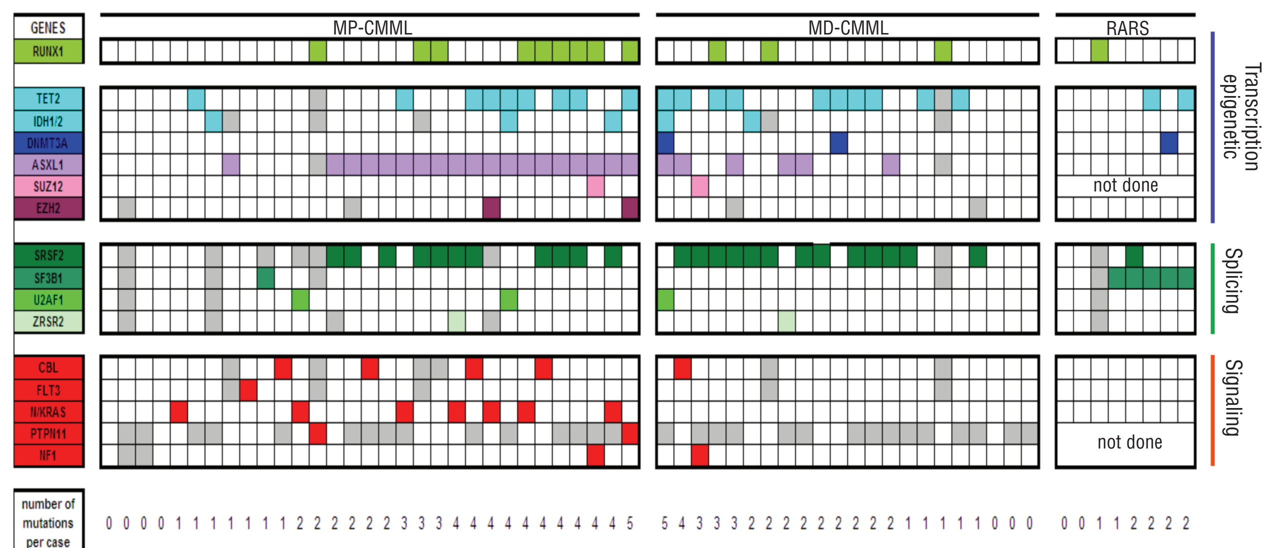

Figure 4. Gene mutations in CMML. Profiles of concomitant gene mutations of 18 genes in 53 CMML cases. Mutations in RARS (*Online Supplementary Table S6*) are shown for comparison. The number of gene mutations per case is presented below. Missing data are indicated by gray squares.

a similar transcriptional program involving red blood cell homeostasis These results were validated by using two external published data sets and a third GES (MDS-ext GES) suggesting the robustness of the results. In MD-CMML and RARS, several red blood cell functions were affected, including heme biosynthesis, iron-sulfur cluster biogenesis, calcium uptake, antioxidant defense, and transcription regulation, suggesting that MD-CMML and RARS share a common differentiation pathway characterized by an erythrocytic program.

In MD-CMML and RARS these results could reflect a mere abundance of erythroid progenitors, an overexpression of the program to compensate for abnormal red blood cell maturation, or a true molecular defect with pathological consequences. This defect in MD-CMML and RARS may be different or similar. In the latter case at least two possibilities should be considered. First, genes present in the signatures and/or involved in the identified functions could be directly affected by mutations. Many of these functions take place in the mitochondrion and it is possible that the defect affects primarily this organelle; the role of the iron transporter gene *ABCB7* had been suspected, but the gene is not mutated.[39] Actually, mutations in mitochondrion genes are associated with congenital rather than acquired anemias. A second hypothesis is perhaps the most likely. Mutations in splicing factors may indirectly affect red cell processes by leading to abnormal processing of mRNAs, including RNAs involved in red cell biology.[40] These frequently show alternative splicing[41] which may make them hypersensitive to splicing mutations. *SF3B1* mutations are prominent in MDS with ring sideroblasts (RS) (a hallmark of iron overload and dyserythropoiesis) such as RARS, and *SF3B1* haploinsufficiency or inhibition leads to the formation of RS.[42] In contrast, *SRSF2* mutations are frequent in CMML.[31] Splicing defects in *SRSF2* in CMML and in *SF3B1* in RARS may affect genes involved in erythropoiesis such as *ABCB7, FTL, GATA1*, or *HAMP* for example, or a master transcription factor of hematopoietic cell lineages such as TIF1γ, which controls erythroid cell fate and acts as a tumor suppressor in CMML.[43-45] However, *SRSF2* mutations are also found in MP-CMML which does not show a red cell program. MP-CMML is characterized by mutations in signaling genes, and this could modify the effect of *SRSF2*. Overexpression of components of the pathway may also compensate for the defect (e.g. *DICER1, CUGBP2/CELF2, ZCCHC11/TUT4* and *SYNCRIP* which are involved in miRNA and mRNA editing).

Current molecular findings are shedding new light on myeloid diseases.[46] Our study establishes a molecular bridge between CMML and MDS, suggesting that these diseases are part of a continuum of pathologies. Interestingly, we (*Online Supplementary Table S5*, HD-0376) and others[47] have observed rare cases of passage from MDS to CMML. This continuum is likely to include RARS associated with marked thrombocytosis (RARS-T, a disease characterized by mutations in *SF3B1* and *JAK2* or *MPL*), since RARS and RARS-T are characterized by common gene expression features[48] and the presence of RS. No such marker has been regularly described in CMML; nor has it been systematically investigated either. Perhaps this is due to a stronger defect in RARS than in MD-CMML because mutations in *SF3B1* have a stronger impact on erythropoiesis than *SRSF2*.

We did not find a difference in overall survival between our 37 MD and MP-CMML cases separated by the CMML GES (*P*=0.18, Fisher's exact test). Neither did *SRSF2* mutation have any impact on prognosis. If, given the current possibilities of treatment, the prognosis of MD and MP-CMML is similar, in the future these diseases may be treated differently and we will need to be able to identify the difference between them. Our discovery of a molecular similarity between MD-CMML and RARS could be useful by improving the classification of these diseases, providing pathophysiological clues and suggesting the possibility of using treatment approaches common to the two diseases but different from those for MP-CMML.

Acknowledgments

We are grateful to the patients who agreed to participate to the study.

Funding

This work was supported by Inserm, Institut Paoli-Calmettes and grants from the Fondation ARC pour la Recherche sur le Cancer (DB), Association Laurette Fugain (MJM 2010) and Leukaemia and Lymphoma Research (NCPC).

Authorship and Disclosures

Information on authorship, contributions, and financial & other disclosures was provided by the authors and is available with the online version of this article at www.haematologica.org.

References

1. Tefferi A, Hoagland HC, Therneau TM, Pierre RV. Chronic myelomonocytic leukemia: natural history and prognostic determinants. Mayo Clin Proc. 1989; 64(10):1246-54.
2. Onida F. Prognostic factors in chronic myelomonocytic leukemia: a retrospective analysis of 213 patients. Blood 2002;99(3):840-9.
3. Vardiman JW, Thiele J, Arber DA, Brunning RD, Borowitz MJ, Porwit A, et al. The 2008 revision of the World Health Organization (WHO) classification of myeloid neoplasms and acute leukemia: rationale and important changes. Blood. 2009;114(5):937-51.
4. Ramshaw HS, Bardy PG, Lee MA, Lopez AF. Chronic myelomonocytic leukemia requires granulocyte-macrophage colony-stimulating factor for growth in vitro and in vivo. Exp Hematol. 2002;30(10):1124-31.
5. Invernizzi R, Travaglino E, Benatti C, Malcovati L, Della Porta M, Cazzola M, et al. Survivin expression, apoptosis and proliferation in chronic myelomonocytic leukemia. Eur J Haematol. 2006;76(6):494-501.
6. Bennett JM, Catovsky D, Daniel MT, Flandrin G, Galton DA, Gralnick H, et al. The chronic myeloid leukaemias: guidelines for distinguishing chronic granulocytic, atypical chronic myeloid, and chronic myelomonocytic leukaemia. Proposals by the French-American-British Cooperative Leukaemia Group. Br J Haematol. 1994; 87(4):746-54.
7. Gelsi-Boyer V, Trouplin V, Roquain J, Adélaïde J, Carbuccia N, Esterni B, et al. ASXL1 mutation is associated with poor prognosis and acute transformation in chronic myelomonocytic leukaemia. Br J Haematol. 2010;151(4):365-75.
8. Rocquain J, Carbuccia N, Trouplin V, Raynaud S, Murati A, Nezri M, et al. Combined mutations of ASXL1, CBL, FLT3, IDH1, IDH2, JAK2, KRAS, NPM1, NRAS, RUNX1, TET2 and WT1 genes in myelodysplastic syndromes and acute myeloid leukemias. BMC Cancer. 2010; 10:401-17.
9. Ernst T, Chase AJ, Score J, Hidalgo-Curtis CE, Bryant C, Jones AV, et al. Inactivating mutations of the histone methyltransferase gene EZH2 in myeloid disorders. Nat Genet. 2010;42(8):722-6.
10. Brecqueville M, Rey J, Bertucci F, Coppin E, Finetti P, Carbuccia N, et al. Mutation analysis of ASXL1, CBL, DNMT3A, IDH1,

IDH2, JAK2, MPL, NF1, SF3B1, SUZ12 and TET2 in myeloproliferative neoplasms. Genes Chromosome Cancer. 2012; 51(8):743-55.
11. Damm F, Kosmider O, Gelsi-Boyer V, Renneville A, Carbuccia N, Hidalgo-Curtis C, et al. Mutations affecting mRNA splicing define distinct clinical phenotypes and correlate with patient outcome in myelodysplastic syndromes. Blood. 2012;119(14): 3211-8.
12. Bertucci F, Finetti P, Cervera N, Charafe-Jauffret E, Mamessier E, Adélaïde J, et al. Gene expression profiling shows medullary breast cancer is a subgroup of basal breast cancers. Cancer Res. 2006;66(9):4636-44.
13. Golub TR, Slonim DK, Tamayo P, Huard C, Gaasenbeek M, Mesirov JP, et al. Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. Science. 1999;286(5439):531-7.
14. Pellagatti A, Cazzola M, Giagounidis AA, Malcovati L, Porta MG, Killick S, et al. Gene expression profiles of CD34+ cells in myelodysplastic syndromes: involvement of interferon-stimulated genes and correlation to FAB subtype and karyotype. Blood. 2006;108(1):337-45.
15. Mills KI, Kohlmann A, Williams PM, Wieczorek L, Liu WM, Li R, et al. Microarray-based classifiers and prognosis models identify subgroups with distinct clinical outcomes and high risk of AML transformation of myelodysplastic syndrome. Blood. 2009;114(5):1063-72.
16. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci USA. 2005;102(43):15545-50.
17. Pellagatti A, Cazzola M, Giagounidis A, Perry J, Malcovati L, Della Porta MG, et al. Deregulated gene expression pathways in myelodysplastic syndrome hematopoietic stem cells. Leukemia. 2010;24(4):756-64.
18 Gelsi-Boyer V, Trouplin V, Adélaïde J, Bonansea J, Cervera N, Carbuccia N, et al. Mutations of polycomb-associated gene ASXL1 in myelodysplastic syndromes and chronic myelomonocytic leukaemia. Br J Haematol. 2009;145(6):788-800.
19. Reiter A, Invernizzi R, Cross NC, Cazzola M. Molecular basis of myelodysplastic/ myeloproliferative neoplasms. Haematologica. 2009;94(12):1634-8.
20. Bejar R, Stevenson K, Abdel-Wahab O, Galili N, Nilsson B, Garcia-Manero G, et al. Clinical effect of point mutations in myelodysplastic syndromes. N Engl J Med. 2011;364(26):2496-506.
21. Nikoloski G, van der Reijden BA, Jansen JH. Mutations in epigenetic regulators in myelodysplastic syndromes. Int J Hematol. 2012;95(1):8-16.
22. Kohlmann A, Grossmann V, Klein HU, Schindela S, Weiss T, Kazak B, et al. Next-generation sequencing technology reveals a characteristic pattern of molecular mutations in 72.8% of chronic myelomonocytic leukemia by detecting frequent alterations in TET2, CBL, RAS, and RUNX1. J Clin Oncol. 2010;28(24):3858-65.
23. Bacher U, Weissmann S, Kohlmann A, Schindela S, Alpermann T, Schnittger S, et al. TET2 deletions are a recurrent but rare phenomenon in myeloid malignancies and are frequently accompanied by TET2 mutations on the remaining allele. Br J Haematol. 2011;156(1):67-75.
24. Grossmann V, Kohlmann A, Zenger M, Schindela S, Eder C, Weissmann S, et al. A deep-sequencing study of chronic myeloid leukemia patients in blast crisis (BC-CML) detects mutations in 76.9% of cases. Leukemia. 2011;25(3):557-60.
25. Jankowska AM, Makishima H, Tiu RV, Szpurka H, Huang Y, Traina F, et al. Mutational spectrum analysis of chronic myelomonocytic leukemia includes genes associated with epigenetic regulation: UTX, EZH2, and DNMT3A. Blood. 2011; 118(4):3932-41.
26. Muramatsu H, Makishima H, Maciejewski JP. Chronic myelomonocytic leukemia and atypical chronic myeloid leukemia: novel pathogenetic lesions. Semin Oncol. 2012; 39(1):67-73.
27. Makishima H, Jankowska AM, Tiu RV, Szpurka H, Sugimoto Y, Hu Z, et al. Novel homo- and hemizygous mutations in EZH2 in myeloid malignancies. Leukemia. 2010; 24(10):1799-804.
28. Abdel-Wahab O, Pardanani A, Rampal R, Lasho TL, Levine RL, Tefferi A. DNMT3A mutational analysis in primary myelofibrosis, chronic myelomonocytic leukemia and advanced phases of myeloproliferative neoplasms. Leukemia. 2011;25(7):1219-20.
29. Abdel-Wahab O, Pardanani A, Patel J, Wadleigh M, Lasho T, Heguy A, et al. Concomitant analysis of EZH2 and ASXL1 mutations in myelofibrosis, chronic myelomonocytic leukemia and blast-phase myeloproliferative neoplasms. Leukemia. 2011;25(7):1200-2.
30. Walter MJ, Ding L, Shen D, Shao J, Grillot M, McLellan M, et al. Recurrent DNMT3A mutations in patients with myelodysplastic syndromes. Leukemia. 2011;25(7):1153-8.
31. Yoshida K, Sanada M, Shiraishi Y, Nowak D, Nagata Y, Yamamoto R, et al. Frequent pathway mutations of splicing machinery in myelodysplasia. Nature. 2011; 478(7367):64-9.
32. Visconte V, Makishima H, Jankowska A, Szpurka H, Traina F, Jerez A, et al. SF3B1, a splicing factor is frequently mutated in refractory anemia with ring sideroblasts. Leukemia 2012;26(3):542-5.
33. Papaemmanuil E, Cazzola M, Boultwood J, Malcovati L, Vyas P, Bowen D, et al. Somatic SF3B1 mutation in myelodysplasia with ring sideroblasts. N Engl J Med. 2011; 365(15):1384-95.
34. Abu Kar S, Jankowska AM, Makishima H, Visconte V, Jerez A, Sugimoto Y, et al. Spliceosomal gene mutations are frequent in the diverse mutational spectrum of chronic myelomonocytic leukemia but largely absent in juvenile myelomonocytic leukemia. Haematologica. 2013;98(1):107-13.
35. Malcovati L, Papaemmanuil E, Bowen DT, Boultwood J, Della Porta MG, Pascutto C, et al. Clinical significance of SF3B1 mutations in myelodysplastic syndromes and myeloproliferative/myelodysplastic neoplasms. Blood. 2011;118(24):6239-46.
36. Thol F, Kade S, Schlarmann C, Loffeld P, Morgan M, Krauter J, et al. Frequency and prognostic impact of mutations in SRSF2, U2AF1, and ZRSR2 in patients with myelodysplastic syndromes. Blood. 2012; 119(15):3578-84.
37. Visconte V, Makishima H, Maciejewski JP, Tiu RV. Emerging roles of the spliceosomal machinery in myelodysplastic syndromes and other hematological disorders. Leukemia. 2012;26(12):2447-54.
38. Brecqueville M, Cervera N, Gelsi-Boyer V, Murati A, Adélaïde J, Chaffanet M, et al. Rare mutations in DNMT3A in myeloproliferative neoplasms and myelodysplastic syndromes. Blood Cancer J. 2011;1(5):e18.
39. Boultwood J, Pellagatti A, Nikpour M, Pushkaran B, Fidler C, Cattan H, et al. The role of the iron transporter ABCB7 in refractory anemia with ring sideroblasts. PLoS One. 2008;3(4):e1970.
40. Makishima H, Visconte V, Sakaguchi H, Jankowska AM, Abu Kar S, Jerez A, et al. Mutations in the spliceosome machinery, a novel and ubiquitous pathway in leukemogenesis. Blood. 2012;119(14):3203-10.
41. Liu P, Barb J, Woodhouse K, Taylor JG 6th, Munson PJ, Raghavachari N. Transcriptome profiling and sequencing of differentiated human hematopoietic stem cells reveal lineage-specific expression and alternative splicing of genes. Physiol Genomics. 2011;43(20):1117-34.
42. Visconte V, Rogers HJ, Singh J, Barnard J, Bupathi M, Traina F, et al. SF3B1 haploinsufficiency leads to formation of ring sideroblasts in myelodysplastic syndromes. Blood. 2012;120(16):3173-86.
43.. Bai X, Kim J, Yang Z, Jurynec MJ, Akie TE, Lee J, et al. TIF1γ controls erythroid cell fate by regulating transcription elongation. Cell. 2010;142(1):133-43.
44. Monteiro R, Pouget C, Patient R. The gata1/pu.1 lineage fate paradigm varies between blood populations and is modulated by tif1gamma. EMBO J. 2011; 30(6):1093-103.
45. Aucagne R, Droin N, Paggetti J, Lagrange B, Largeot A, Hammann A, et al. Transcription intermediary factor 1γ is a tumor suppressor in mouse and human chronic myelomonocytic leukemia. J Clin Invest. 2011;21(6):2361-70.
46. Murati A, Brecqueville M, Devillier R, Mozziconacci MJ, Gelsi-Boyer V, Birnbaum D. Myeloid malignancies: mutations, models and management. BMC Cancer. 2012; 12(1):304.
47. Ben Salah N, Gouider E, Belakhal F, El Borgi W, Menif S, Ben Abid H, et al. [Evolution d'une cytopénie réfractaire avec dysplasie multilignage et sidéroblastes en couronne en leucémie myélomonocytaire chronique]. Tunis Med. 2011;89(10):806-7.
48. Malcovati L, Della Porta MG, Pietra D, Boveri E, Pellagatti A, Gallì A, et al. Molecular and clinical features of refractory anemia with ringed sideroblasts associated with marked thrombocytosis. Blood. 2009;114(17):3538-45.