Department of MEDICINE, HEMATOLOGY
CENTRO DE INVESTIGACIÓN DEL CÁNCER-IBMCC
(USAL-CSIC)

VNiVERSiDAD D SALAMANCA

CAMPUS OF INTERNATIONAL EXCELLENCE

Doctoral Dissertation

Molecular Characterization of Myelodysplastic Syndromes (MDS):

Analysis of genomic abnormalities in the development of MDS, progression to Acute Myeloblastic Leukemia and response to treatment with 5-Azacytidine

With the approval of Salamanca University, Department of Medicine, this thesis will be defended on 18th November 2015, in the Lecture Hall, Centro de Investigación del Cáncer-IBMCC, Salamanca.

Supervisors:

Prof. Dr. Jesús M. Hernández Rivas
Dra. M. del Rocío Benito Sánchez

María Abáigar Alvarado

2015

D. **Jesús María Hernández Rivas**, Doctor en Medicina, Profesor Titular del Departamento de Medicina de la Universidad de Salamanca, Médico Adjunto del Servicio de Hematología del Hospital Clínico Universitario de Salamanca e Investigador del Centro de Investigación del Cáncer de Salamanca,

D.ª **María del Rocío Benito Sánchez**, Doctora en Ciencias Biológicas e Investigadora del Centro de Investigación del Cáncer de Salamanca,

CERTIFICAN:

Que D.ª **María Abáigar Alvarado**, licenciada en Biología por la Universidad de Salamanca, ha realizado bajo nuestra dirección el trabajo de Tesis Doctoral titulado *"Molecular Characterization of Myelodysplastic Syndromes (MDS): Analysis of genomic abnormalities in the development of MDS, progression to Acute Myeloblastic Leukemia and response to treatment with 5-Azacytidine"*, y que éste reúne, a nuestro juicio, las condiciones de originalidad y calidad científica requeridas para su presentación y defensa ante el tribunal correspondiente para optar al grado de Doctor, con mención "Doctor Europeus", por la Universidad de Salamanca.

Y para que así conste a los efectos oportunos, firmamos el presente certificado en Salamanca, a 15 de Octubre de 2015.

Fdo:

Prof. Dr. Jesús M. Hernández Rivas

Dra. M. del Rocío Benito Sánchez

This Thesis was performed being María Abáigar Alvarado fully supported by a "*Junta para Ampliación de Estudios*" fellowship of the Spanish National Research Council (*Consejo Superior de Investigaciones Científicas*, CSIC), co-funded by the European Social Fund. (JAEPre_09_02402)

Acknowledgments:

- Red Temática de Investigación Cooperativa en Cáncer (RTICC), Instituto de Salud Carlos III (ISCIII), Spanish Ministry of Economy and Competitiveness & European Regional Development Fund (ERDF) "*Una manera de hacer Europa*" (Innocampus, CEI-2010-1-0010) (RD12/0036/0069; RD12/0036/0029; RD12/0036/0044)
- Celgene Corporation, Spain.
- Roche Diagnostics. Microarray innovations in LEukemia (MILE). 2006 – 2009
- The European Union COST Action. BMBS Action BM0801. "Translating genomic and epigenetic studies of MDS and AML" (EuGESMA). 2008 - 2012
- Consejería de Sanidad y Gerencia Regional de Salud de la Junta de Castilla y León (SACYL):
  (GRS 355/A/09). Estudio genómico y epigenético de los síndromes mielodisplásicos (SMD) mediante microarrays de alta densidad. Evaluación de la respuesta terapéutica. 2009 - 2011
  (GRS/874/A13). Estudio de los cambios genómicos y transcriptómicos asociados a la presencia de mutaciones en Síndromes Mielodisplásicos. 2013
  (BIO/SA47/13). Estudio mediante secuenciación masiva de las mutaciones de los genes implicados en síndromes mielodisplásicos. 2013
  (BIO/SA52/14). Estudio de las mutaciones presentes en las células progenitoras hematopoyéticas en los Síndromes Mielodisplásicos mediante secuenciación masiva. 2014
- Fundación Memoria de D. Samuel Solórzano Barruso. Estudio genómico y epigenético de los Síndromes Mielodisplásicos (SMD) mediante microarrays de alta densidad. Evaluación de la respuesta terapéutica. 2011
- Roche Diagnostics. The Interlaboratory Robustness of Next-generation sequencing (IRON) study II. Roche Molecular Diagnostics. 2012 – 2013
- The European Union Seventh Framework Programme [FP7/2007-2013] under Grant Agreement 306242-2, Next Generation Sequencing platform for targeted Personalized Therapy of Leukemia (NGSPTL). FP7-HEALTH-2012-INNOVATION-1. 2013-2015

"When we search for the treasure, we realize that the journey is the treasure itself"

Paulo Coelho

Table of Contents

List of Abbreviations

4D	Carboxy-terminal tetramerization domain
53bp1	Tumor protein p53 binding protein 1
AA	Amino acid
aCGH	Array-based comparative genomic hybridization
AKT	v-akt murine thymoma viral oncogene homolog 1
ALL	Acute lymphoblastic leukemia
AML	Acute myeloid leukemia
ANC	Absolute neutrophil nount
Arg	Arginine
Asp	Aspartic acid
AZA	5-azacytidine
Bak1	BCL2-Antagonist/Killer 1
BM	Bone marrow
bp	Base pairs
Brca1	Breast cancer 1
CBFB	Core binding factor beta
CC	Conventional cytogenetics
CDF	Chip definition file
CI	Confidence interval
CLL	Chronic lymphocytic leukemia
CMML	Chronic myelomonocytic leukemia
CN-LOH	Copy-neutral loss of heterozygosity
CNA	Copy number abnormalities
CNV	Copy number variation
CPSS	CMML-Specific Prognostic Scoring System
CR	Complete remission/response
CyR	Cytogenetic response
Cys	Cysteine
DAC	Decitabine
DAVID	Database for Annotation, Visualization and Integrated Discovery
DBA	Diamond–Blackfan anemia
DBD	Sequence-specific DNA-binding domain
DDR	DNA damage response
del	Deletion
der	Derivative chromosome
DNMT	DNA methyltransferase
dsDNA	Double-stranded DNA
EGFR	Epidermal growth factor receptor
ELN	European LeukemiaNet
EPO	Erythropoietin
ETS2	v-ets avian erythroblastosis virus E26 oncogene homolog 2
FAB	French-American-British
FDA	US Food and Drug Administration
FDR	False discovery rate
FEA	Functional Enrichment Analysis
FISH	Fluorescence in situ hybridization
FPSS	French Prognostic Scoring System
fRMA	Frozen RMA (Robust multichip average)
G-CSF	Granulocyte-colony stimulating factor
GATExplorer	Genomic and Transcriptomic Explorer
gDNA	Genomic DNA
GEO	Gene Expression Omnibus
GEP	Gene expression profiling
Gln	Glutamine
Gly	Glycine
GO-BP	Gene Ontology-Biological Process
GST P1-1	Glutathione S-transferase P1-1
HDAC	Histone deacetylase
HhaI	Restriction enzyme HhaI
HI	Hematological improvement
HIV	Human immunodeficiency virus
HMAs	Hypomethylating agents
HOX	Homeobox
HR	Hazard ratio
HR-MDS	High-risk MDS
HSC	Hematopoietic stem cell
i	Isochromosome
ICT	Intensive chemotherapy
Ile	Isoleucine

Int	Intermediate
inv	Inversion
IPSS	International Prognostic Scoring System
IPSS-R	Revised International Prognostic Scoring System
ISCN	International System for Human Cytogenetic Nomenclature
IWG	International Working Group
JNK	c-Jun N-terminal kinase JNK
Kb	Kilobase
KEGG	Kyoto Encyclopedia of Genes and Genomes
LDH	Lactate dehydrogenase
Leu	Leucine
LOH	Loss of heterozygosity
LR-MDS	Low-risk MDS
LR-PSS	Lower-Risk MDS Prognostic Scoring System
Lys	Lysine
MAPK	Mitogen activated protein kinase
Mb	Megabase
Mcl-1	Myeloid Cell Leukemia 1
MDA-CSS	MD Anderson Comprehensive Scoring System
Mdc1	Mediator of DNA-damage checkpoint 1
MDM2	MDM2 Proto-Oncogene
MDS	Myelodysplastic syndromes
MDS del(5q)	MDS associated with isolated deletion 5q
MDS-U	MDS unclassified
MDS/MPN	Myelodysplastic/myeloproliferative neoplasms
MIC	Morphologic, immunophenotypic and cytogenetic
MILE	Microarray Innovations in LEukemia
MM	Multiple myeloma
MPN	Myeloproliferative neoplasm
MS-MLPA	Methylation specific multiplex ligation-dependent probe amplification
MTase	Methyltransferase domain
mTOR	Mammalian target of rapamycin
MYC	v-myc avian myelocytomatosis viral oncogene homolog
Neg	Negative regulation domain
NF-kB	Nuclear factor kappa-B
NGS	Next-generation sequencing
NHEJ	Non-homologous end joining
NK	Natural killer
NLS	Nuclear localization signalling domain
NoL	Non-leukemic
NR1H3	Nuclear receptor subfamily 1, group H, member 3
ORR	Overall response rate
OS	Overall survival
PANTHER	Protein ANalysis THrough Evolutionary Relationships
PB	Peripheral blood
PCR	Polymerase chain reaction
Phe	Phenylalanine
PI3K	Phophatidylinositol 3-kinase
PR	Partial remission/response
Pro	Proline
PWWP	Proline-tryptophan-tryptophan-proline domain
RA	Refractory anemia
RAEB-1	Refractory anemia with excess of blasts type 1
RAEB-2	Refractory anemia with excess of blasts type 2
RARS	Refractory anemia with ringed sideroblasts
RARS-T	RARS and marked thrombocytosis
RBC	Red blood cell
RCMD	Refractory cytopenia with multilineage dysplasia
RCMD-RS	Refractory cytopenia with multilineage Dysplasia and ringed sideroblasts
RCUD	Refractory cytopenia with unilineage dysplasia
rDNA	Ribosomal DNA
RFS	Relapse-free survival
RN	Refractory neutropenia
RP	Ribosomal protein
RPL11	Ribosomal Protein L11

RPL23	Ribosomal Protein L23
RPL5	Ribosomal Protein L5
RPS13	Ribosomal Protein S13
RPS7	Ribosomal Protein S7
RT	Refractory thrombocytopenia
sAML	Secondary AML
SCT	Stem-cell transplantation
SD	Stable disease
Ser	Serine
SNP	Single nucleotide polymorphism
SNP-A	Single nucleotide polymorphism arrays
snRNP	Small nuclear ribonucleoprotein
SOM	Self Organizing Maps
t	Translocation
TAD1	Amino-terminal transactivation domain 1
TAD2	Amino-terminal transactivation domain 2
TF	Transcription factor
TFBS	Transcription factor binding sites
TGIF2	TGFB-induced factor homeobox 2
Tie-2	Protein receptor tyrosine kinase (epithelial-specific)
tMDS	Therapy-related MDS
Trp	Tryptophan
TSG	Tumor-suppressor gene
Tyr	Tyrosine
UPD	Uniparental disomy
VAF	Variant allele frequency
Val	Valine
WBC	White blood cells
WHO	World Health Organization
WPSS	WHO classification-based Prognostic Scoring System
XBP1	X-box binding protein 1
XPA	Xeroderma pigmentosum, complementation group A
ZNF	Zinc finger domain

General Introduction

"A journey of a thousand miles begins with a single step"

Lao Tse

1. Myelodysplastic syndromes

1.1. Disease overview

Myelodysplastic syndromes (MDS) are a heterogeneous group of hematopoietic stem cell disorders characterized by clonal and ineffective hematopoiesis, resulting in various degrees of peripheral blood (PB) cytopenias, bone marrow (BM) hypercellularity, and morphological dysplasia in one or more myeloid lineages.[1-3] MDS present a highly variable clinical course, ranging from indolent conditions with a near-normal life expectancy over many years, to an aggressive disease with a rapid progression to acute myeloid leukemia (AML).[4,5]

1.2. Incidence and causes

MDS are one of the most common hematological malignancies in Western countries and is generally considered to be a disease of the elderly, with a median age of 65–70 years at diagnosis, and with <10% of the patients younger than 50 years. The annual incidence of MDS is about 4 patients per 100,000 people, but the disease becomes more common with increasing age, such that the incidence rises to 30 per 100,000 per year for people over 70 years of age. In Asian populations MDS tend to appear at an earlier age.[6-8] MDS more frequently affect males, except for the 5q- syndrome, which is slightly more frequent in women.[3,6-11] MDS in children are considerably less common and have different characteristics from those of adults.[11]

Several risk factors have been implicated in the development of MDS. These include age, male gender, previous use of radiotherapy or chemotherapy, immunosuppressive agents, viral infections, exposure to ionizing radiation or to benzene, smoking tobacco, excess alcohol intake, and other environmental or occupational exposures.[6,7,12,13] In fact, the incidence of secondary MDS has been increasing in recent years, maybe in relation to the use of chemotherapy and radiotherapy in cancer.[11] An inherited predisposition was seen in a third of pediatric cases, including children with Bloom syndrome, neurofibromatosis, and

inherited bone marrow-failure syndromes such as Fanconi anemia, Schwachman-Diamond syndrome, severe congenital neutropenia, dyskeratosis congenital and Diamond-Blackfan anemia.[6,7,11,13]

1.3. Diagnosis: clinical features and diagnostic tests

The diagnosis of all hematological malignancies, including MDS, needs accurate clinical evaluation combined with precise pathological and genetic analyses. MDS are typically diagnosed on the basis of peripheral blood and bone marrow findings.[11]

Clinical presentation, signs and symptoms

The predominant clinical manifestations in MDS patients are fatigue, infections and bleeding as a result from the presence of peripheral blood cytopenias (anemia, neutropenia and thrombocytopenia), which are a consequence of bone marrow failure.[7,12] However, those symptoms are non-specific and could be related to other pathological processes including nutritional disorders (folate and vitamin B12 deficiencies), toxic exposures (alcohol, environmental toxins, certain drugs), infectious states (HIV, parvovirus B19, hepatitis viruses), autoimmune disease, liver disease, hypersplenism, rare forms of hereditary anemias, and other clonal hematopoietic conditions (aplastic anemia, paroxysmal nocturnal hemoglobinuria or myeloproliferative disorders). A careful study must be performed to exclude these diseases with similar characteristics to those of MDS. The diagnosis of MDS should be considered in any patient with unexplained cytopenias(s).[7,11-13]

Peripheral blood and bone marrow examination

The suspicion of a diagnosis of MDS is based on the presence of cytopenias in a routine analysis of the peripheral blood.[9] In general, 90% of MDS patients have anemia (hemoglobin levels <10 g/dL), while a third of patients show thrombocytopenia (platelet count $<50x10^9$/L) and/or neutropenia (neutrophil count $<1.5x10^9$/L). In addition, small numbers of circulating blasts (hematopoietic cell precursors) can be found, but at a level rarely exceeding 5%.[7]

A diagnosis is confirmed by performing a morphological examination of the peripheral blood smear and bone marrow aspirate.[7,9,11] The BM in MDS is usually hyper- or normocellular and shows dysplasia in one or several myeloid hematopoietic lineages (erythroid, granulocytic, megakaryocytic).[7,11,13,14] Dysplasia is considered when at least 10% of the cells, of at least one myeloid BM lineage, show unequivocal morphological changes.[11,14] The proportion of blasts in the BM must also be assessed to provide a correct classification.[7,13] By contrast, BM biopsy is not mandatory, but is important for identifying fibrotic MDS or hypocellular MDS.[7]

Cytogenetic findings

The key feature of MDS is the clonal nature of the dysplasia, and the cytogenetic evaluation of bone marrow samples from patients with MDS has become the most widely available and standardized method for establishing the clonality of these disorders and is needed to complete the laboratory evaluation of a patient with MDS.[9,11,13] Therefore, a cytogenetic analysis should be performed in any patient with a suspected MDS in whom BM examination is indicated, and at least 20 metaphases should be analyzed whenever possible, as a standard criterion for a reliable cytogenetic result.[11,13,15,16] Chromosomal abnormalities should be described according to the International System for Human Cytogenetic Nomenclature (ISCN) recommendations.[17]

At the time of diagnosis, clonal recurrent chromosomal abnormalities are detected in approximately half (40-60%) of primary MDS patients and 80% of therapy-related MDS (tMDS) by chromosome banding analysis.[2,3,7,13,16,18] These chromosomal aberrations appear as single abnormalities or combined with one or several other aberrations within a complex karyotype, defined as the presence of more than 3 abnormalities.[19] The most frequent single aberrations in MDS patients include del(5q), -7/del(7q), +8, del(20q), and –Y.[2,7,13,16,20] Other less frequent abnormalities have been identified as isolated aberrations or in the frame of complex karyotypes, and include del(11q), del(12p), i(17)(q10), inv(3)/t(3q)/del(3q), +19, +21, der(1,7), -X, -21, -13/del(13q), del(16q), t(11q23), -1/del(1p), t(5q), del(17p), +1q, or +11.[16,19,20] These karyotypic aberrations are not specific to MDS and are commonly found in other myeloid neoplasms such as chronic myelomonocytic leukemia (CMML) and acute myeloid leukemia (AML).[21] Many of these chromosomal abnormalities not only confirm the diagnosis of MDS but also are very important for defininig the outcome and the risk of

progression to AML, and, in some subsets of patients, to select the most effective therapy.[2,19,22] In addition, a significant fraction of MDS patients acquired additional cytogenetic changes over time. This evolution has been associated with an increased risk of leukemic transformation and worse survival.[19,23,24] In the absence of clonal cytogenetic abnormalities, a careful evaluation of other causes of bone marrow abnormalities, with pathological characteristics similar to MDS, must be assessed to define the MDS diagnosis.[11]

1.4. Classification of MDS

Given the clinical heterogeneity of MDS, several classification systems have been developed to identify groups of patients with similar morphological features, molecular etiology, prognosis, and likelihood of response to common therapies.[25,26]

Morphological classification

The first classification system of MDS was proposed by the French-American-British (FAB) cooperative group in 1982, and has been widely used.[27] Some years later, cytogenetics was incorporated in an attempt to include the 5q- syndrome (MIC classification).[28] More recently, the World Health Organization (WHO) 2001 and currently the WHO 2008 classification systems have also been used to classify MDS.[14,29]

1.4.1.1. WHO 2008 classification

The currently used system of classification of MDS is that of the World Health Organization (WHO) 2008. This classification divides MDS into several disease entities based on the morphological findings in peripheral blood and bone marrow such as the type and degree of dysplasia, the number of cytopenias, the proportion of blasts in the BM and PB, and the presence of a specific chromosomal abnormality (deletion of chromosome 5q).[11,14,25] The WHO classification scheme identified the following MDS subtypes: refractory cytopenia with unilineage dysplasia (RCUD), which includes refractory anemia (RA), refractory neutropenia (RN) and refractory thrombocytopenia (RT); refractory anemia with ringed sideroblasts

(RARS); refractory cytopenia with multilineage dysplasia (RCMD); refractory anemia with excess of blasts-type 1 (RAEB-1); refractory anemia with excess of blasts-type 2 (RAEB-2); MDS unclassified (MDS-U); MDS associated with isolated deletion 5q (MDS with del(5q)) (Table 1). In addition, in the WHO 2008 classification of myeloid neoplasms, chronic myelomonocytic leukemia (CMML), characterized by the presence of $\geq 1x10^9$/L monocytes in the PB or the BM, and refractory anemia with ringed sideroblasts and marked thrombocytosis (RARS-T), with a platelet count $>450x10^9$/L, ≥15% ringed sideroblasts and large megakaryocytes in the BM, are included in a new category of mixed myelodysplastic/myeloproliferative neoplasms (MDS/MPN). The presence of a ≥20% blasts in the PB or BM is considered as acute myeloid leukemia (AML) when it occurs *de novo,* or secondary AML (sAML) when it occurs in the setting of a previously diagnosed MDS, MDS/MPN, or myeloproliferative neoplasm (MPN).[14]

Table 1. **WHO 2008 criteria for classifying MDS patients**. (Adapted from *Vardiman et al. (2009)* [14]).

Category	Peripheral blood findings		Bone marrow findings			Cytogenetics
	Blasts (%)	Cytopenias	Blasts (%)	Myeloid lineages with dysplasia **	Ringed sideroblasts (%)	
RCUD (RA, RN, RT) †	< 1 ***	1 or 2 *	< 5	1	< 15	Possible abnormal
RARS	< 1	1 (Anemia)	< 5	1 (Erythroid)	≥ 15	Possible abnormal
RCMD	< 1 *** no Auer rods	≥ 1	< 5 no Auer rods	≥ 2	± 15	Possible abnormal
RAEB-1	< 5 †† no Auer rods	≥ 1	5 – 9 no Auer rods	≥ 1	Irrelevant	Possible abnormal
RAEB-2	5 – 19 †† ± Auer rods	≥ 1	10 - 19 †† ± Auer rods	≥ 1	Irrelevant	Possible abnormal
MDS-U	< 1 ***	≥ 1	< 5	≥ 1		Cytogenetic abnormality ‡
MDS with del(5q)	< 1	1 (Anemia) Normal/increased platelet count	< 5 no Auer rods	Normal/increased megakaryotyctes with hypolobated nuclei	Irrelevant	Isolated del(5q)

† RA, for those with anemia and only erythroid displasia (dyserythropoiesis); RN, for those with neutropenia and only dysgranulopoiesis; RT, for those with thrombocytopenia and only megakaryocytic dysplasia (dysmegakaryocytopoiesis).
‡ Cytogenetic abnormality considered as presumptive evidence for a diagnosis of MDS.
* Bicytopenia in the PB may occasionally be observed. Cases with pancytopenia should be classified as MDS-U.
** Myeloid lineages with dysplasia in ≥ 10% cells, neutrophil and/or erythroid precursors and/or megakaryocytes.
*** Cases with <5% myeloblasts in the BM but with 2-4% myeloblasts in the PB, should be classified as RAEB-1. Cases of RCUD and RCMD with 1% myeloblasts in the PB should be classified as MDS-U.
†† Cases with <10% myeloblasts in the BM and <5% myeloblasts in the PB but with Auer rods in the blood should be classified as RAEB-2. Although the finding of 5-19% blasts in the PB is, in itself, diagnostic of RAEB-2, cases of RAEB-2 may have <5% blasts in the PB if they have Auer rods or 10-19% myeloblasts in the BM or both. Similarly, cases of RAEB-2 may have <10% blasts in the BM but may be diagnosed by the other two findings, Auer rods+ and/or 5-19% blasts in the PB.

Risk stratification / Prognostic scoring systems

MDS patients have a highly variable clinical course, with large differences in overall survival (OS) and risk of transformation to AML, ranging from indolent conditions over many years to forms rapidly progressing to leukemia.[4,5,21] This clinical heterogeneity is relevant in clinical decision-making regarding therapeutic modalities and timing of intervention.[4] Thus, a prognostic classification of MDS is of great importance for risk stratification, prediction of survival, risk of evolution to AML, and also to help in the timing and choice of therapy.[9] Several studies have identified important prognostic factors, and in some instances, have developed prognostic scoring systems to predict the outcome of MDS patients. These include the International Prognostic Scoring System (IPSS), the WHO classification-based Prognostic Scoring System (WPSS), the Lower-Risk MDS Prognostic Scoring System (LR-PSS), the MD Anderson Comprehensive Scoring System (MDA-CSS), the French Prognostic Scoring System (FPSS), and more recently published the Revised International Prognostic Scoring System (IPSS-R).[5,22,25,30,31] However, some of these prognostic scoring systems have not been formally included in common clinical guidelines.[25] It is important to point out that none of these prognostic systems was designed to predict the response to any particular therapy in MDS.[26] However, they may help clinical decision-making. The MDS prognostic scoring systems most currently used are reviewed below.

1.4.1.2. IPSS-R (2012)

The IPSS-R is an MDS prognostic scoring system based on the proportion of blasts in the BM, cytogenetic abnormalities, and the severity of peripheral cytopenias, each considered individually (hemoglobin, platelet and neutrophil levels).[22,25] These five prognostic variables are weighted as shown in Table 2. Moreover, chromosomal abnormalities are stratified over 5 cytogenetic risk groups, according to the new MDS cytogenetic classification proposed by Schanz et al.,[19] and are summarized in Table 2. According to the score calculated by combining these five parameters, the IPSS-R assigned patients to 1 of 5 risk groups with significantly different OS and risk of progression to AML. In addition, the IPSS-R recognizes the role of age, performance status, serum ferritin, and lactate dehydrogenase (LDH) levels for overall survival, but not for AML evolution. An on-line calculator tool to help apply the

IPSS-R is available at http://www.ipss-r.com and http://www.mds-foundation.org/calculator/index.php.[22,25,31]

Table 2. **Revised International Prognostic Scoring System (IPSS-R)**. (Adapted from *Greenberg et al. (2012)* [22]).

	Score Value						
Prognostic Variables	**0**	**0.5**	**1**	**1.5**	**2**	**3**	**4**
Cytogenetics	Very good	_	Good	_	Intermediate	Poor	Very Poor
BM blasts %	≤ 2	_	> 2 - < 5%	_	5 - 10%	> 10%	_
Hemoglobin (g/dL)	≥ 10	_	8 - < 10	< 8	_	_	_
Platelets (x10^9/L)	≥ 100	50 - < 100	< 50	_	_	_	_
ANC (x10^9/L)	≥ 0.8	< 0.8	_	_	_	_	_
Cytogenetic risk groups							
Very good	del(11q), -Y						
Good	Normal, del(5q), del(12p), del(20q), double including del(5q)						
Intermediate	del(7q), +8, i(17q), +19, any other single or double independent clones						
Poor	inv(3)/t(3q)/del(3q), -7, double including -7/del(7q), complex: 3 abnormalities						
Very poor	Complex: >3 abnormalities						
IPSS-R Risk group	**Very low**	**Low**	**Intermediate**		**High**		**Very high**
Risk score	≤ 1.5	> 1.5 - 3	> 3 - 4.5		> 4.5 - 6		> 6
Median OS (years)	8.8	5.3	3.0		1.6		0.8
AML progression*	Not reached	10.8	3.2		1.4		0.7

*Median time to 25% AML evolution (years).
ANC, absolut neutrophil count.

1.4.1.3. Refined WPSS

The refined WPSS is an MDS prognostic scoring system mainly based on the WHO 2008 classification system.[25] However, other variables, such as the presence of severe anemia, have been shown to be of prognostic value in MDS, and are therefore included in the final refined WPSS model.[5,25,31,32] The presence of severe anemia, recently defined as hemoglobin levels <9g/dL in males and <8g/dL in females, has been correlated with poor clinical outcomes in MDS and has proved to be as effective as red blood cell (RBC) transfusion-dependency in the prognostic assessment of MDS.[5,22,31] These three disease variables: the

WHO 2008 MDS categories, cytogenetic abnormalities, and the presence of severe anemia are scored as shown in Table 3. Based on the scores of these three variables, the refined WPSS stratified patients into five risk groups with significantly different survival and probability of leukemic evolution[5,25,32] (Table 3). The refined WPSS is very simple to apply and does not require additional testing to implement. In addition, this prognostic system can be applied at any time during follow-up, enabling it to be used as a dynamic scoring system.[5,25]

Table 3. **Refined WHO Classification-based Prognostic Scoring System (WPSS)**. (Adapted from *Malcovati et al. (2011)*[5]).

	Score Value			
Prognostic Variables	**0**	**1**	**2**	**3**
WHO category	RCUD, RARS, MDS with isolated del(5q)	RCMD	RAEB-1	RAEB-2
Cytogenetics	Good	Intermediate	Poor	_
Severe anemia*	Absent	Present	_	_
Cytogenetic risk groups				
Good	Normal, del(5q), del(20q), -Y			
Intermediate	Other abnormalities			
Poor	Chromosome 7 abnormalities or complex (≥3 abnormalities)			
WPSS Risk group	**Very low**	**Low**	**Intermediate**	**High**	**Very high**
Risk score	0	1	2	3 - 4	5 - 6
Median OS (years)**	>10	8 - 9	4.5 - 5.5	1.8 - 2.5	0.5 - 1
AML progression** [a]	6%	24%	48%	63%	100%

*Severe anemia: Hemoglobin < 9 g/dL in males or < 8 g/dL in females.
**From *Malcovati et al. (2007)*[32] and *Estephan et al. (2014)*[26].
a) Cumulative probability of AML progression in 5 years.

2. Biological and molecular abnormalities of MDS

The natural history of MDS is highly variable. This probably reflects the large number of cytogenetic, genetic and epigenetic alterations that are associated with MDS.[11,33] Cytogenetic abnormalities are present in a high proportion of MDS patients, where they are often associated with specific clinical features.[33] However, more than 50% of patients have normal cytogenetics. These findings indicate that many underlying molecular lesions within the MDS bone marrow remain to be identified. There is increasing evidence that haploinsufficiency, epigenetic changes, abnormal apoptosis, immune system deregulation, abnormal signal transduction pathways, and the BM microenvironment contribute to the development and progression of MDS.[11,33] Several recent studies have demonstrated the presence of gene mutations in most MDS. The presence of a genetic lesion is not isolated, and a combination of gene mutations or cytogenetic abnormalities is usual.[2,34-36] Therefore, in MDS, as in other types of cancer, a single genetic event is not sufficient for a cell to develop into cancer. Instead, several gradually acquired DNA alterations are required, resulting in the development and progression of the disease.[37]

This section will summarize some of the biological mechanisms involved in the pathogenesis of MDS.

2.1. Cryptic genomic lesions

The study of structural variation in the human genome has become of great interest in cancer since the acquisition of genomic changes such as deletions, amplifications, translocations, and inversions that affect certain regions of the genome can drive the development of cancer through the activation or inactivation of genes that may promote tumor progression.[38-41]

High-resolution genome-wide scanning techniques such as array-based comparative genomic hybridization (aCGH) and single nucleotide polymorphism arrays (SNP-A) have been

applied in the study of several hematological malignancies, and have identified new chromosomal defects that are not detected by standard cytogenetics, suggesting their potential clinical utility in karyotyping analysis.[18,41-57] Array-CGH and SNP-A represent two whole-genome approaches for the identification of copy number abnormalities (CNAs) with a higher resolution than conventional cytogenetics (CC).[43,47,48] These CNAs consists of unbalanced chromosomal defects in the form of losses (deletions) or gains (duplications/amplifications) of genetic material.[41,53] In addition, SNP-A also enables detection of copy-neutral loss of heterozygosity (CN-LOH), a lesion that can occur due to uniparental disomy (UPD). UPD arises as a consequence of the loss of genetic material on one chromosome and subsequent duplication of the missing part from the remaining chromosome.[43,51] This abnormality cannot be detected by classic cytogenetic methods because there is no change in the quantity of genetic material.[43,47,48,51,53]

To date, several aCGH and SNP-A studies in MDS and other myeloid malignancies have been reported, in which copy number changes have been identified in cases with a normal karyotype, single chromosome aberrations, complex karyotypes, as well as with non-informative cytogenetic results due to no growth. These studies have demonstrated that array-based techniques can detect almost all the aberrations that can be identified by conventional cytogenetics at higher resolution, except for balanced translocations and inversions and when the abnormal clone is below the limits of aCGH/SNP-A sensitivity (<25-30% of abnormal cells).[18,43,46,48-50,52,53] Thus, array-based karyotyping allows the detection of clonal chromosomal aberrations in 50% of MDS cases with non-informative CC.[46-48] These include the most frequent abnormalities found in MDS such as del(5q), -7/del(7q) and trisomy 8, which occur at similar frequencies to those in patients with successful CC exam. However, large defects were also identified involving regions previously described in myeloid malignancies, such as chromosomes 1, 12, and 13.[48] Nevertheless, these high-resolution techniques have the advantage of revealing cryptic chromosomal defects that are not detected by standard cytogenetics due to their size (<5 Mb, below the resolution of conventional karyotyping) or the technical shortcomings of CC. Such cryptic lesions have been identified in approximately 70-90% cases of MDS and related myeloid neoplasms and in >50% of MDS patients with normal cytogenetics.[18,41-57] These submicroscopic abnormalities involved chromosomal regions containing genes of known significance in the pathogenesis of

these disorders, including *TET2*, *RUNX1*, *ETV6*, *TP53*, *NF1*, and *DNMT3A*.[42,44,49,50] SNP-A studies also showed that recurrent areas of CN-LOH are wide-spread lesions occurring in 10-20% of MDS, and are particularly frequent in CMML.[43,47,48] We would emphasize that these newly detected chromosomal aberrations, including CN-LOH and submicroscopic deletions, are often the most informative for implicating individual genes for detailed sequence analysis. However, these are also the most difficult to distinguish from nonpathological variations.[44,47,48]

In addition to the assessment of clonality in newly diagnosed cases of MDS, aCGH and SNP-A analyses have provided new insights into the molecular pathogenesis of these complex stem cell diseases. Indeed, the identification of cryptic clonal aberrations in MDS by high-resolution genome-wide scanning techniques is likely to be the best starting point for the discovery of new genetic mutations and signal transduction pathways involved in MDS, leading to the development of more effective targeted therapies.[44]

2.2. Genomic Instability: Chromothripsis

Cancer is driven by the progressive accumulation of genomic changes, such as somatically acquired point mutations and chromosomal rearrangements, leading to the inactivation of tumor suppressor genes, activation of oncogenes and/or the formation of fusion proteins with oncogenic potential. This wide variety of genomic changes has been well documented in tumor cells, indicating that chromosome instability is a central aspect of cancer cell biology.[39,40,58] As a result, cellular processes including cell cycle control, apoptosis, and DNA repair are impaired, conferring a growth advantage on cells and promoting tumor progression.[37,38]

Recent studies using next-generation DNA sequencing, array-based copy number profiling (aCGH, SNP-A) described an alternative mechanism of genomic instability, termed chromothripsis, (from the Greek; 'chromo' for chromosome; 'thripsis' for shattering into pieces), whereby tens to hundreds of clustered genomic rearrangements occur in an apparently one-step catastrophic event.[37-40,58-60] These multiple and complex

rearrangements usually criss-cross back and forth across the involved regions, resulting in a large number of copy number oscillations rapidly alternating between two, or occasionally three, copy number states.[37-39,58,60,61] For example, a normal region with two copies would be followed by a region with one copy, followed by another region with two copies, followed by another region with three copies[61] (Figure 1). Existing data have shown that this complex genomic abnormality usually affects an entire chromosome, a chromosome arm, or even a few megabases of a chromosome, and in some instances, several chromosomes.[37-40,58,60,61] Based on previous studies, chromothripsis could be inferred when at least 10 changes in segmental copy number involving two or three distinct copy-number states on a single chromosome are detected[59] (Figure 1).

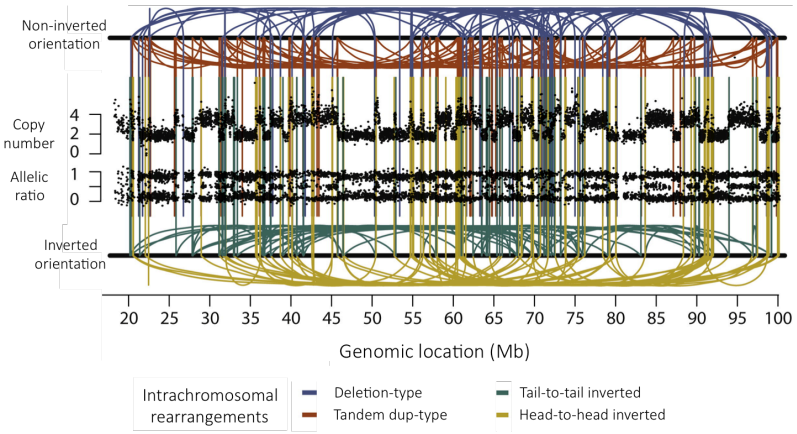


Figure 1. **Plot showing complex rearrangements on chromosome 8 resulting from chromothripsis in a small cell lung cancer cell line**. The illustration shows the copy number profile, allelic ratio, and rearrangements of chromosome 8, showing a complex pattern of alternating gains and losses. (Taken from *Stephens et.al. (2012)* [39]).

By far the simplest explanation for this genomic chaos is that, at some point during cancer development, distinct chromosomes or chromosomal regions are broken into many pieces, some, but not all of them being inaccurately stitched back together by the DNA repair machinery into a derivative chromosome.[38-40,61] Some of these chromosomal pieces, from diverse locations, are joined together in all possible orientations generating a highly mosaic

derivative chromosome, while other fragments from shattered chromosomes may not be incorporated into this derivative chromosome and are lost to the cell, or could become stitched together in a circular extra-chromosomal structure (so-called "double-minute" chromosome) that may then become amplified[40,59] (Figure 2). Under this scenario, those fragments that are retained in the eventual derivative chromosome or in the "double-minute" chromosome will have the higher copy number states, while those that are lost to the cell will be present in the lower copy number state.[39] Thus, chromothripsis generates a wide variety of genomic changes that lead to loss or disruption of tumor suppressor genes, activation of oncogenes and/or the formation of potentially oncogenic fusion genes, all occurring at the same time.[39,61] A cell suffering this genomic crisis would be expected to undergo apoptosis. Surprisingly, a cell can survive this catastrophic event and emerge with a highly aberrant genomic landscape that confers a significant selective growth advantage on that clone, thereby promoting cancer progression.[38,39]

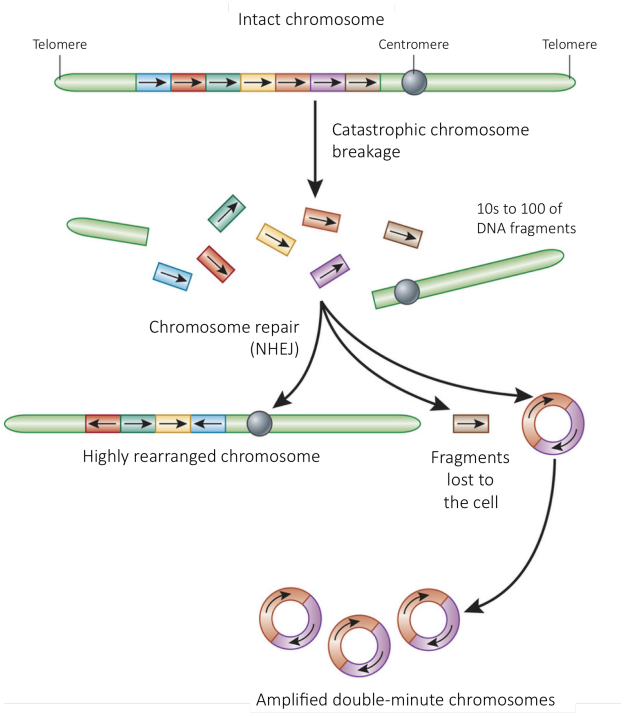


Figure 2. **Mechanism generating complex chromosomal rearrangements after chromosome shattering**. (Adapted from *Holland et.al. (2012)* [62]).

The mechanisms that could produce such massive and highly localized genomic rearrangements remain to be determined. However, the following mechanistic hypotheses have been suggested: high-energy ionizing radiation during mitosis, premature chromosome compaction, DNA replication stress, critical telomere shortening, abortive apoptosis and *TP53* mutations.[37,38,40,59,61,63]

Chromothripsis has been described in in various types of cancer, including hematological malignancies and solid tumors. In certain tumors, such as bone cancers, chromothripsis is particularly prevalent with an incidence of ~25%. In addition, some congenital disorders also display chromothripsis.[37-39,58-61,64-71] Furthermore, given the potential impact that a single catastrophic event may have on a cell, chromothripsis has been particularly associated with more aggressive tumors and forms of cancer known to gradually evolve into a more aggressive disease. Consequently, this phenomenon has been related to poor prognosis.[37,38,59,60] However a systematic analysis of the incidence of chromothripsis in MDS and the possible relation with the outcome of the patients has not been undertaken.

2.3. DNA Methylation

DNA methylation is an epigenetic process that involves the addition of a methyl group (CH_3) to the C5 position of the cytosine ring in the CpG dinucleotide, when the cytosine (C) is followed by a guanine (G).[72-75] The CpG dinucleotides are underrepresented in the human genome but they frequently cluster together in the so-called CpG-islands. These CpG-islands are often located in or near the promoter regions of the genes (promoter-associated CpG-islands).[72,74,75] Methylation of DNA is catalyzed by DNA methyltransferases (DNMTs) including DNMT1, DNMT3a, and DNMT3b, and is a key mechanism for controlling the stabilization of the genome, the remodeling of the chromatin and gene transcription.[72,74,76,77] In normal cells, most of the intergenic DNA is methylated, whereas promoter-associated CpG-islands are generally unmethylated, allowing for gene transcription.[72,74,75,78] DNA methylation patterns are tissue-specific and conserved through cell division, with the parental pattern as the template for the methylation of the newly synthesized DNA.[72,74] Methylation of

promoter-associated CpG-islands has been associated with gene silencing through inhibition of gene transcription.[72,75,79-81]

Aberrant DNA methylation is now recognized as as having an important role in carcinogenesis and, particularly, in the pathogenesis of myeloid neoplasms. In cancer cells, global DNA hypomethylation has been associated with genomic instabililty, whereas gene promoter hypermethylation has been related to gene silencing.[72,75,77,79-82] In recent years, several tumor-suppressor genes (TSGs) have been described as being transcriptionally inactivated by aberrant promoter hypermethylation in myeloid neoplasms. Some of these genes are tumor-suppressors involved in cell cycle regulation (*CDKN1B*, *CDKN2B*, *HIC1* and *FHIT*), while others are involved in cell adhesion (*IGSF4*, *CDH1* and *CDH13*), apoptosis regulation (*DAPK1*), cell growth (*ESR1*) and DNA damage repair (*MLH1*).[77,79,80,83-89] For instance, hypermethylation of the *CDKN2B* gene promoter, also known as $P15^{ink4b}$, has been found in 20-50% of MDS and AML patients and has been correlated with poor outcome.[78,79,83,84,86,88,90] In addition, hypermethylation of some other genes, such as *HIC1*, *CDH1*, *ESR1* and *FHIT*, has also been associated with poor prognosis and increased risk of progression to AML.[79,80,84,87,89,90] The exact mechanism that leads to aberrant DNA methylation is not fully understood. However, the recent discovery of mutations in regulators of the DNA methylation, such as *DNMT3A*, *IDH1*, *IDH2*, and *TET2* genes, may help to clarify some of these mechanisms.[74,77] Mutations in these epigenetic regulators will be reviewed below.

In contrast to genetic aberrations, silencing of genes by DNA methylation is a reversible process. Modification of this hypermethylation and consequently, gene reactivation, is now being used therapeutically for treating MDS and AML patients with hypomethylating agents such as 5-azacytidine and decitabine.[74,77,91]

2.4. Somatic gene mutations

Our understanding of the molecular pathogenesis of MDS has dramatically improved in the last 10 years with the application of next-generation sequencing techniques (whole exome,

whole genome sequencing and targeted gene set approaches).[2,31,36] More than 40 recurrently mutated genes have been identified in MDS and related myeloid neoplasms, with several more being discovered as genome sequencing technologies have continued to improve.[26,34-36,92]

Recent studies have demonstrated that more than 90% of MDS patients had at least one mutation, even those patients with a normal karyotype. In addition, approximately 40% of MDS patients had 2 or 3 gene mutations, while 10% of MDS show between 4 and 8 mutations.[35,36] These recurrently mutated genes are involved in signal transduction (*JAK2*, *KRAS*, *NRAS*, *CBL*), transcriptional regulation (*RUNX1*, *BCOR*, *BCORL1*, *ETV6*, *EVI1*, *GATA2*, *TP53*), DNA methylation (*DNMT3A*, *TET2*, *IDH1*, *IDH2*), chromatin modification (*EZH2*, *ASXL1*, *ATRX*, *KDM6A*), RNA-splicing machinery (*SF3B1*, *U2AF1*, *SRSF2*, *ZRSR2*, *PRPF8*), DNA replication (*SETBP1*) or cohesin complex (*STAG2*, *RAD21*, *SMC1A*, *SMC3*), among others.[2,34-36,93-101] Moreover, some of these mutations are related to morphological and clinical phenotypes.[34,35] Thus, mutations of splicing factors genes, especially *SF3B1*, are present in 12.4% of MDS patients, and are associated with the presence of ringed sideroblasts and better outcome.[26,100,102,103] By contrast, the presence of mutations in *RUNX1*, *EZH2*, *ASXL1* or *TP53* genes are associated with a worse outcome.[2,26,34-36,95,97,99] A summary of recurrent mutations found in *de novo* MDS patients in multiple studies and their prognostic impact is shown in Figure 3.

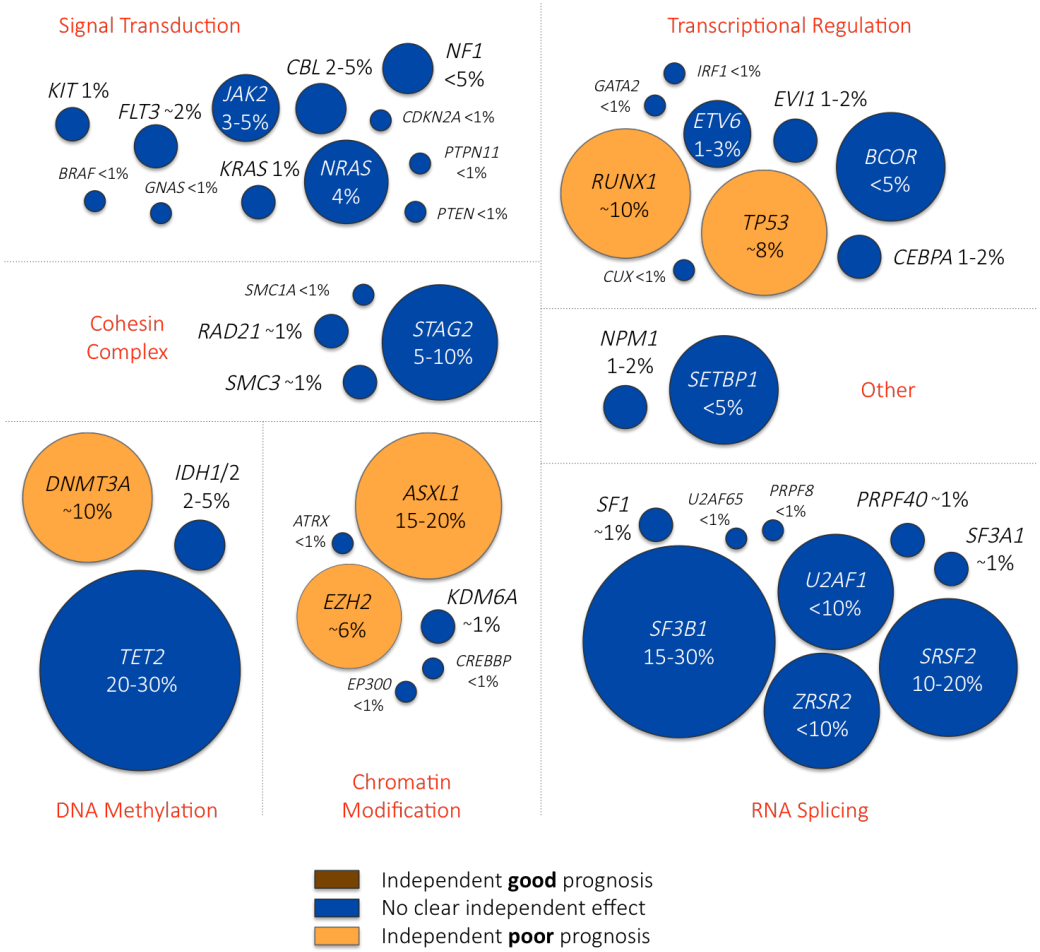


Figure 3. **Mutational landscape in MDS**. Summary of the current knowledge of somatic mutations in MDS and their effect on overall survival. (Adapted from *Steensma (2015)* [8]).

2.5. Gene expression deregulation in MDS

Microarray-based gene expression profiling (GEP) studies have proved to be a powerful tool in the study of transcriptional pathways involved in cancer. These studies have helped identify gene expression signatures associated with distinct cancer subtypes and contributed to better classify neoplasias. They have also enabled patients to be stratified into different risk classes and the deregulated genes and gene pathways implicated in cancer development to be identified.[104-108] GEP has also been used to analyze tumor progression, in order to

identify genes specifically activated or inactivated during the different stages of the tumor.[109-111] Furthermore, GEP studies have identified new prognostic markers related to patient outcomes and response to treatment, providing new information about potential target genes for therapeutic intervention.[106,107,111-113]

During the last decade, many microarray-based GEP studies have been directed towards the study of hematological malignancies, including lymphoma, leukemia, and MDS.[104,106,107,109-111,114-121] The first GEP report was able to distinguish AML and ALL patients without previous knowledge of the respective leukemia classes, and even to distinguish novel leukemia subclasses.[111,112] Subsequently, the international multi-institutional Microarray Innovations in LEukemia (MILE) study, centered on the European LeukemiaNet (ELN, www.leukemia-net.org), assessed the clinical utility of microarray-based GEP in the diagnosis and sub-classification of 16 clinically recognized subtypes of leukemia, as well as MDS.[106,107] Although GEP accurately classified leukemias, it failed to confirm the clinical diagnosis of MDS in half of the MDS samples.[106] However, GEP has shown itself capable of distinguishing clinical MDS subgroups of prognostic relevance, in terms of time to AML transformation.[106,108,110]

Several GEP studies of MDS have been undertaken, mainly focusing on the analysis of enriched CD34+ cell populations, and predominantly in MDS with del(5q). These studies have facilitated the discrimination between MDS and healthy individuals, between specific MDS subtypes, and between cytogenetic subclasses, providing valuable insights into the biology of these disorders.[106,109-111,115-121] Interestingly, patients with del(5q), -7/del(7q) or trisomy 8 have characteristically different gene expression signatures, reflecting the fact that gene expression is affected by gene dosage.[116,119-121] Thus, the most significantly deregulated gene pathways in MDS patients with del(5q) were related to ribosome, translation and chromatin assembly.[116] In fact, these patients showed reduced expression of multiple ribosomal genes, including *RPS14*, which is contained in the deleted region.[116,120,121] Similarly, patients with chromosome 7 abnormalities were characterized by the activation of the JNK pathway, and MDS with trisomy 8 featured deregulation of gene pathways associated with the immune response and leukocyte activation.[116,119] The study of specific MDS subtypes revealed that RARS patients were characterized by the up-regulation of mitochondrial and iron-related genes when compared to RA patients and non-leukemic controls.[117,118,122] GEP studies also

demonstrated that low-risk MDS were characterized by an increased apoptosis and deregulated immune response, while in high-risk MDS patients DNA damage response, cell cycle, and DNA repair were the most deregulated pathways.[109-111,114-116] Thus, microarray-based GEP has helped the identification of genes and gene pathways deregulated in patients with MDS, which are implicated in the pathogenesis of these disorders, providing novel targets for treatment and predicting prognosis and response to therapy.[111] However, the mechanisms underlying the progression from MDS to AML are still not well characterized and further investigation is required.

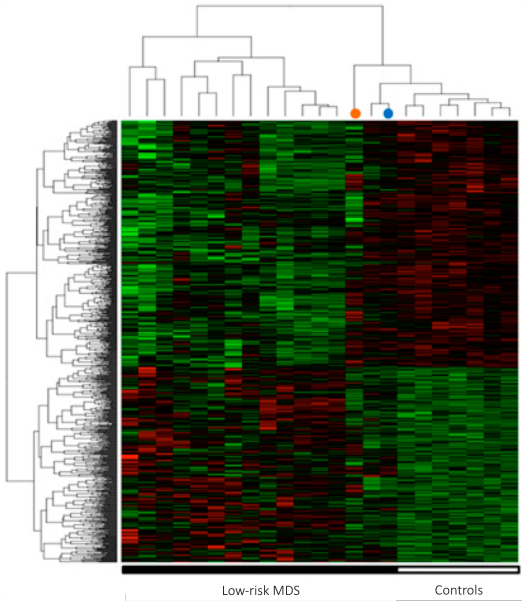


Figure 4. **Hierarchical clustering of differentially expressed genes between low-risk MDS and normal BM samples**. Green indicates up-regulation and red indicates down-regulation. Each row represents a single gene and each column corresponds to a single patient sample. (Adapted from *Del Rey et al (2013)* [115]).

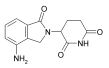
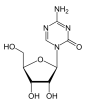
3. Treatment of MDS: 5-Azacytidine

Treatment options for MDS patients are highly variable and depend on the MDS subtype, disease severity, patient age and comorbidities.[26,123] Thus, the main goal of treatment is to ameliorate symptoms and side effects, and to improve the quality of life for low-risk MDS (LR-MDS) patients, whereas for high-risk MDS (HR-MDS), the main objective is to modify the natural course of the disease, delaying the progression to AML, and to extend overall survival.[7,26]

Patients with LR-MDS are either asymptomatic or suffer from anemia, infections and bleeding as a result of peripheral cytopenias.[6,7,26] Asymptomatic patients may only need observation and periodic PB count monitoring. In symptomatic LR-MDS patients, therapy is based on best supportive care consisting of red blood cell (RBC) or platelet transfusions, hematopoietic growth factors (EPO, darbepoietin alfa or G-CSF), antibiotics, thrombopoietin-receptor agonists (romiplostim, eltrombopag) and iron chelation agents (deferasirox, deferiprone).[7,26,123] In addition, lenalidomide is an immunomodulatory agent with proven clinical efficacy that has been approved by the US Food and Drug Administration (FDA) for the treatment of the 5q-syndrome (Table 4).[7,26]

In contrast, patients with HR-MDS have a very poor prognosis, with a survival of less than 1 year if they are not treated. Allogeneic stem-cell transplantation (SCT) remains the only potential curative treatment for high-risk MDS, with prolonged disease-free survival of 35–50%. As such, it should always be considered as a potential therapeutic option in all HR-MDS patients.[7,26,123] However, SCT plays a limited role due to the advanced age of many patients, associated comorbidities, the toxicity of this procedure and/or lack of a suitable donor. Thus, fewer than 5% of MDS patients can benefit from it.[7,26] For these reason, in HR-MDS patients other treatment options should be considered. These include intensive chemotherapy (ICT, combinations of anthracyclines, cytarabine and/or fludarabine) or hypomethylating agents (HMA) such as 5-azacytidine and decitabine.[7,26,123,124] These two hypomethylating agents have been approved for the treatment of all subtypes of MDS and are new promising therapies with proven clinical efficacy, especially in HR-MDS patients (Table 4).[26]

Table 4. **Approved drugs for the treatment of MDS**.

Agent	Chemical Structure	Mechanism of action	Population
Lenalidomide		Antiangiogenic and immunomodulatory agent	Low- / Int-1 risk MDS with del(5q)
Azacytidine		DNMTs inhibitors	Int-2 / high-risk MDS Primary treatment failure
Decitabine		DNMTs inhibitors	Int-2 / high-risk MDS Primary treatment failure

DNMTs, DNA methyltransferases; Int, intermediate; MDS, myelodysplastic syndromes.

These therapies can be effective for improving cytopenias, achieving cytogenetic remissions, and reducing the proportion of bone marrow blasts. However, increasing knowledge about the pathogenesis of MDS, through genetic, epigenetic, immunological, and other biological mechanisms, as well as the mechanisms of action and resistance to current therapies have facilitated the development of novel and targeted treatments and the refinement of existing therapies that can lead to improvements in the outcome of MDS.[26] These novel therapies are summarized in the Table 5.

Table 5. **Novel therapies under investigation for the treatment of MDS**. (Adapted from *Ridgeway et al. (2012)* [125]).

Agent	Target	Mechanism of action	Trial / Population	Response	Grade 3 / 4 adverse effects
ARRY-614	P38/Tie-2	P38-MAPK inhibitor, Antineoplastic, anti-infammatory, and antiangiogenic activity	Phase I Low- / Int-1 risk MDS	–	–
Entinostat (SNDX-275/ MS-275)	HDAC	HDAC-1 / HDAC-3 inhibitor	Combination with AZA Phase III High-risk MDS	HI and CyR did not differ between AZA/Placebo *vs* AZA/Entinostat	• Thrombocytopenia: 63% • Fatigue: 23%
Erlotinib	EGFR signaling	Tyrosine kinase inhibitor, Blocks EGFR signaling	Phase II Int-2 and high-risk MDS	ORR: 17%	• Diarrhea: 21% • Thrombocytopenia: 17% • Rash: 17%
Everolimus	mTOR	mTOR inhibitor, Induces G1 arrest	Phase II Low- / Int-1 risk MDS	–	–
Ezatiostat	GST P1-1	GST P1-1 inhibitor, Stimulates proliferation of myeloid precursors	Phase I Int-2 risk MDS	HI: 38%	• Neutropenia: 7%
Panobinostat (LBH589)	HDAC	HDACs inhibitor, Inhibits differentiation, induces apoptosis	Phase II Relapsed / refractory MDS	SD: 70%	• Thrombocytopenia: 80% • neutropenia: 70% • leukopenia: 60% • anemia: 50% • febrile neutropenia: 20%
Rigosertib (ON-01910; Estybon)	Polo-1 kinase, PI3K, AKT	Inhibits mitotic progression, induces apoptosis	Phase II Int-1 / Int-2 / high-risk MDS	>50% blast decrease: 27% (34% of 38% patients relapsed / refractory after HMAs) Median OS (weeks): responders, 51 / SD, 37; progressive disease, 15	Well tolerated without evidence of myelotoxicity

Tie-2; protein receptor tyrosine kinase (epithelial-specific); MAPK, mitogen activated protein kinase; Int-intermediate; MDS-myelodysplastic syndromes; HDAC, histone deacetylase; AZA, azacytidine; HI, hematological improvement; CyR, cytogenetic response; EGFR, epidermal growth factor receptor; ORR, overall response rate; mTOR, mammalian target of rapamycin; GST P1-1, glutathione S-transferase P1-1; SD, stable disease; PI3K, phophatidylinositol 3-kinase; AKT, protein kinase B; HMAs, hypomethylating agents; OS, overall survival.

The two hypomethylating agents, 5-azacytidine (AZA) and decitabine (DAC), have recently been approved for the treatment of patients with MDS. Although the mechanisms of action of these agents are not fully understood, they have been shown to induce gene and global hypomethylation in vivo [124] and to inhibit DNA methyltransferases (DNMTs), which are the enzymes responsible for DNA methylation. Incorporation of AZA or decitabine into DNA results in the dose- and time-dependent inhibition of DNA methyltransferase activity by irreversible binding to the enzymes, which prevents the methylation of newly synthesized DNA.[72]

As mentioned in section 2.3, aberrant DNA methylation of promoter-associated CpG-islands has been associated with gene silencing and has been related to the pathophisiology of MDS, especially in HR-MDS, who are associated with a higher number of methylated gene loci.[72,75,77,79,80] In addition, the increased number of methylated loci is associated with progression of the disease from LR-MDS to HR-MDS.[26,72] Thus, the use of hypomethylating agents results in a reversal of the epigenetically silenced genes, a progressive reduction of DNA methylation and gene expression reactivation in MDS patients, which restores the normal cell growth and differentiation of hematopoietic cells.[72,74,77,91] For these reasons, these drugs have become the first-line treatment for most HR-MDS patients and also those with AML with a low blast count.[7,126] In clinical practice, it is usually used in patients with an initial diagnosis of HR-MDS, and in LR-MDS patients who have failed previous and other treatments.[26]

In recent years an increasing number of patients with hematological malignancies are being treated with hypomethylating agents (azacytidine and decitabine). Several studies have shown that AZA induces 50-60% responses in MDS and AML patients, with 10-20% of complete responses (CR), 10-20% partial remissions (PR), and 30-40% hematological improvement (HI).[7,78,126,127] Treatment with AZA has shown to improve peripheral blood values, reduce transfusion requirements, and significantly improve the quality of life. Moreover, AZA has demonstrated to delay progression to AML and prolong overall survival compared with best supportive care regimens or traditional chemotherapy.[72,74,78,124,128-130] Clinical experience has demonstrated the effectiveness of this epigenetic therapy in MDS and AML patients. However, there are some unresponsive patients, and therefore it is important to better identify the predictive factors for response to these agents.

In summary, despite the advances in our knowledge of the mechanisms involved in the pathogenesis MDS in recent years, there are still questions to be resolved, such as the potential clinical utility of aCGH in the study of MDS cases with an insufficient number of mitosis, the assessment of chromothripsis, and the potential diagnostic and prognostic value of cryptic abnormalities detected by aCGH. In addition, the dynamics of gene expression changes in MDS patients has not been extensively studied, and thereby the analysis of the GEP could be a valuable tool for a better understanding of the biology of these diseases. Finally, since 5-azacytidine is one of the most widely used drugs in the treatment of MDS, the definition of new markers for predicting the response to this drug could be of great interest.

Hypothesis

Myelodysplastic syndromes comprise a highly heterogeneous disease due to its complex pathophysiology. An accurate diagnosis is essential to understand the clinical and biological behavior of MDS. Therefore, it is necessary to complement the clinical data, morphology and cytogenetic studies with other markers that may provide important insights into MDS biology, in order to predict more accurately the course of the disease and to explore novel and more effective targeted therapies.

Diagnosis and classification of MDS depends on the presence of morphological and cytogenetic abnormalities found in BM elements combined with biological parameters. Conventional metaphase cytogenetics (CC) remains the gold standard in the identification of chromosomal abnormalities occurring in MDS patients and still holds a central role in the prognostic assessment of these disorders. However, prognosis and risk stratification may be difficult especially in those patients with inconclusive morphological findings, low blast count, and without clonal cytogenetic aberrations, as is the case of normal or non-informative karyotypes. In fact, approximately 40-60% of MDS have a normal karyotype, and in 10-15% of MDS only poor-quality or few metaphases are available for karyotypic studies. In these patients some subtle chromosomal abnormalities can be undetectable or masked. Furthermore, the use of genome-wide scanning techniques (aCGH, SNP-A) enables the identification of novel abnormalities and a much improved characterization of unbalanced genetic changes in MDS. For this reason, these CC techniques may not be sufficient for a thorough study of these disorders.

In the era of high-resolution genomic technologies, mutations represent a new factor in the diagnosis and prognostic evaluation of MDS, since 70-90% of these patients carry at least one oncogenic mutation, even in patients with a normal karyotype. More than 40 recurrently mutated genes have been identified in patients with MDS, *SF3B1*, *TET2*, *SRSF2*, *ASXL1*, *DNMT3A* and *RUNX1* (>10%) being the most frequently mutated genes. Some of these mutations have been associated with specific MDS subtypes while others have been shown to have a clear impact on clinical outcome through their association with shorter survival and increased risk of progression to AML. Furthermore, some of these genes, such as *TET2*, *RUNX1*, and *TP53*, were also common targets of genomic deletions in the wild-type allele as revealed by CC, aCGH and SNP-A.

Thus, next-generation sequencing and genome-wide copy number studies have both identified several abnormalities in a number genes with a wide variety of functions in MDS and other hematological malignancies. The broad spectrum of these genetic defects highlights the clinical and pathophysiological heterogeneity observed in MDS. Therefore, high-throughput deep sequencing combined with genome-wide copy number analyses will enable us to have a large-scale comprehensive landscape of the genetic lesions occurring in MDS and increase our understanding of the biology of these disorders.

Gene expression profiling (GEP) studies using microarrays identified gene expression signatures distinguishing specific subgroups of MDS, as well as genes and biological pathways implicated in these disorders. Our group has shown that early stages of MDS (low-risk MDS) are characterized by an impairment of cell differentiation and intramedullary increased apoptosis, which would explain the ineffective hematopoiesis and PB cytopenias seen in MDS. In more advanced stages (high-risk MDS) further genetic and epigenetic events may occur, resulting in increased cell proliferation and decreased apoptosis, leading to the accumulation of blasts in the BM and progression to AML. Despite the advances in understanding the biology of MDS and AML through GEP and high-resolution techniques described above, the molecular mechanisms underlying the progression from MDS to AML are largely unknown. Therefore, the study of GEP in MDS and AML could help improve our knowledge of the mechanisms of leukemic transformation.

A large number of hypermethylated genes and gene mutations in key regulators of the DNA methylation have been described in MDS and AML patients. Thus, the use of hypomethylating agents such as 5-azacytidine appears to be a good therapeutic approach for the treatment of these disorders. In fact, it has become the first-line treatment for high-risk MDS patients and AML with low blast count, proving to be an effective treatment for these diseases. However, it remains unclear whether the number or type of hypermethylated genes is associated with clinical response to these agents. Therefore, the evaluation of the DNA methylation status in MDS and AML treated with AZA could select the group of patients in whom this drug could be effective. Furthermore, it has been reported that patients with noncomplex chromosome 7 abnormalities, who have a particularly poor outcome with traditional management strategies, had a longer survival with AZA with more than 40% of

responding patients. Therefore, we support the analysis of the methylation status before AZA treatment, complemented by genomic data, for identifying the predictive factors that may be related to the response to this drug.

Aims

General aim:

To identify new genetic markers that may contribute to a better diagnosis and prognostic evaluation of patients with MDS, and that could be related to the biological processes involved in the pathogenesis of MDS and in the progression from MDS to AML, as well as molecular markers of response to treatment with hypomethylating agents.

Specific aims:

- To explore the presence of novel genetic abnormalities in MDS patients through array-based comparative genomic hybridization (aCGH) analysis.

- To assess the application of genome-wide scanning techniques, such array-based karyotyping (aCGH) and next-generation sequencing (NGS) in the clinical diagnosis of MDS as a complementary tool to conventional cytogenetics (CC).

- To analyze the gene expression profile (GEP) of patients with MDS, AML and nonmalignant disorders in order to identify those genes and mechanisms underlying the development of MDS and their progression into AML.

- To evaluate the influence of the aberrant methylation status and cytogenetic abnormalities on survival and clinical response to 5-azacytidine.

Results

This section includes the experimental work performed on this thesis, including Material and Methods, Results and Discussion. This section has been divided into three chapters:

Chapter 1. M Abáigar, C Robledo, R Benito, F Ramos, M Díez-Campelo, L Hermosín, J Sánchez-del-Real, J María Alonso, R Cuello, M Megido, JN Rodríguez, G Martín-Núñez, C Aguilar, M Vargas, AA Martín, JL García, A Kohlmann, MC del Cañizo, JM Hernández-Rivas. **Chromothripsis is a recurrent genomic abnormality in high-risk myelodysplastic syndromes**. Genes Chromosomes Cancer. Second review.

Chapter 2. M Abáigar, S Aibar, R Benito, M Díez-Campelo, F Ramos, E Lumbreras, FJ Campos-Laborie, M Megido, I Recio, L Hermosín, J Sánchez-del-Real, C Olivier, R Cuello, L Zamora, K Mills, MC del Cañizo, J De Las Rivas, JM Hernández-Rivas. **Common and progressive gene expression changes in the progression of myelodysplastic syndromes to acute myeloid leukemia**. Haematologica. Submitted.

Chapter 3. M Abáigar, F Ramos, R Benito, M Díez-Campelo, J Sánchez-del-Real, L Hermosín, JN Rodríguez, C Aguilar, I Recio, JM Alonso, N de las Heras, M Megido, M Fuertes, MC del Cañizo, JM Hernández-Rivas. **Prognostic impact of the number of methylated genes in myelodysplastic syndromes and acute myeloid leukemias treated with azacytidine**. Ann Hematol. 2013 Nov;92(11):1543-52. doi: 10.1007/s00277-013-1799-9. PubMed PMID: 23740492.

All of them have been developed to accomplish the general aim of this work and give the title to this doctoral dissertation: “Molecular Characterization of Myelodysplastic Syndromes (MDS): Analysis of genomic abnormalities in the development of MDS, progression to Acute Myeloblastic Leukemia and response to treatment with 5-Azacytidine”.

A General Discussion, with additional data and which comprises all research, is addressed in a separate section of this thesis.

In addition, the supplementary material corresponding to each of the above chapters is collected at the end of the digital version of the thesis.

Chromothripsis is a recurrent genomic abnormality in high-risk myelodysplastic syndromes

1

María Abáigar[1], Cristina Robledo[1], Rocío Benito[1], Fernando Ramos[2,3], María Díez-Campelo[4], Lourdes Hermosín[5], Javier Sánchez-del-Real[3], Jose M Alonso[6], Rebeca Cuello[7], Marta Megido[8], Juan N Rodríguez[9], Guillermo Martín-Núñez[10], Carlos Aguilar[11], Manuel Vargas[12], Ana A Martín[4], Juan L García[1], Alexander Kohlmann[13], M Consuelo del Cañizo[4,14], Jesús M Hernández-Rivas[1,4,14]

[1]Unidad de Diagnóstico Molecular y Celular del Cáncer, Centro de Investigación del Cáncer-IBMCC (USAL-CSIC), Salamanca, Spain

[2]IBIOMED, Instituto de Biomedicina, Universidad de León, Spain

[3]Servicio de Hematología, Hospital Universitario de León, Spain

[4]Servicio de Hematología, Hospital Universitario de Salamanca, Spain

[5]Servicio de Hematología, Hospital Jerez de la Frontera, Spain

[6]Servicio de Hematología, Hospital Río Carrión, Palencia, Spain

[7]Servicio de Hematología, Hospital Clínico Universitario de Valladolid, Spain

[8]Servicio de Hematología, Hospital del Bierzo, Ponferrada, León, Spain

[9]Servicio de Hematología, Hospital Juan Ramón Jiménez, Huelva, Spain

[10]Servicio de Hematología, Hospital Virgen del Puerto, Plasencia, Spain

[11]Servicio de Hematología, Hospital General de Soria, Spain

[12]Servicio de Hematología, Hospital Comarcal de Jarrio, Jarrio-Coaña, Asturias, Spain

[13]AstraZeneca, Personalized Healthcare and Biomarkers, Innovative Medicines and Early Development, Cambridge, UK

[14]IBSAL, Instituto de Investigación Biomédica de Salamanca, Spain

Genes Chromosomes Cancer. (Second review)

Chromothripsis is a recurrent genomic abnormality in high-risk myelodysplastic syndromes

María Abáigar,[1] Cristina Robledo,[1] Rocío Benito,[1] Fernando Ramos,[2,3] María Díez-Campelo,[4] Lourdes Hermosín,[5] Javier Sánchez-del-Real,[3] Jose M Alonso,[6] Rebeca Cuello,[7] Marta Megido,[8] Juan N Rodríguez,[9] Guillermo Martín-Núñez,[10] Carlos Aguilar,[11] Manuel Vargas,[12] Ana A Martín,[4] Juan L García,[1] Alexander Kohlmann,[13] M Consuelo del Cañizo,[4,14] and Jesús M Hernández-Rivas[1,4,14]

[1]Unidad de Diagnóstico Molecular y Celular del Cáncer, Centro de Investigación del Cáncer-IBMCC (USAL-CSIC), Salamanca, Spain
[2]IBIOMED, Instituto de Biomedicina, Universidad de León, Spain
[3]Servicio de Hematología, Hospital Universitario de León, Spain
[4]Servicio de Hematología, Hospital Universitario de Salamanca, Spain
[5]Servicio de Hematología, Hospital Jerez de la Frontera, Spain
[6]Servicio de Hematología, Hospital Río Carrión, Palencia, Spain
[7]Servicio de Hematología, Hospital Clínico Universitario de Valladolid, Spain
[8]Servicio de Hematología, Hospital del Bierzo, Ponferrada, León, Spain
[9]Servicio de Hematología, Hospital Juan Ramón Jiménez, Huelva, Spain
[10]Servicio de Hematología, Hospital Virgen del Puerto, Plasencia, Spain
[11]Servicio de Hematología, Hospital General de Soria, Spain
[12]Servicio de Hematología, Hospital Comarcal de Jarrio, Jarrio-Coaña, Asturias, Spain
[13]AstraZeneca, Personalized Healthcare and Biomarkers, Innovative Medicines and Early Development, Cambridge, UK
[14]IBSAL, Instituto de Investigación Biomédica de Salamanca, Spain

To explore novel genetic abnormalities occurring in myelodysplastic syndromes (MDS) through an integrative study combining array-based comparative genomic hybridization (aCGH) and next-generation sequencing (NGS) in a series of MDS and MDS/MPN patients. 301 patients diagnosed with MDS (n=240) or MDS/MPN (n=61) were studied at the time of diagnosis. A genome-wide analysis of DNA copy number abnormalities was performed. In addition, a mutational analysis of *DNMT3A*, *TET2*, *RUNX1*, *TP53* and *BCOR* genes was performed by NGS in selected cases. 285 abnormalities were identified in 71 patients (23.6%). Three high-risk MDS cases (1.2%) displayed chromothripsis involving exclusively chromosome 13 and affecting some cancer genes: *FLT3*, *BRCA2* and *RB1*. All three cases carried *TP53* mutations as revealed by NGS. Moreover, in the whole series, the integrative analysis of aCGH and NGS enabled the identification of cryptic recurrent deletions in 2p23.3 (*DNMT3A*; n=2.8%), 4q24 (*TET2*; n=10%) 17p13 (*TP53*; n=8.5%), 21q22 (*RUNX1*; n=7%), and Xp11.4 (*BCOR*; n=2.8%), while mutations in the non-deleted allele where found only in *DNMT3A* (n=1), *TET2* (n=3), and *TP53* (n=4). These cryptic abnormalities were detected mainly in patients with normal (45%) or non-informative (15%) karyotype by conventional cytogenetics, except for those with *TP53* deletion and mutation (15%), which had a complex karyotype. In addition to well-known copy number defects, the presence of chromothripsis involving chromosome 13 was a novel recurrent change in high-risk MDS patients. Array CGH analysis revealed the presence of cryptic abnormalities in genomic regions where MDS-related genes, such as *TET2*, *DNMT3A*, *RUNX1* and *BCOR*, are located.

Keywords: Chromothripsis, array-based comparative genomic hybridization, amplicon-based next-generation sequencing, myelodysplastic syndromes

INTRODUCTION

The progressive accumulation of genetic aberrations such as copy number abnormalities, in the form of gains or losses of genetic material affecting certain regions of the genome, and particular gene mutations, provide the basis for cancer development (Suela et al., 2007; Forment et al., 2012). However, recent studies have revealed the presence of an alternative mechanism, termed chromothripsis, in which massive chromosome rearrangements occur in a one-step catastrophic event, indicating that

chromosome instability is a central aspect of cancer cell biology. A key feature of chromothripsis is the occurrence of tens to hundreds of clustered genomic rearrangements usually in one or, in some instances, several chromosomes. This complex abnormality can affect an entire chromosome, a chromosome arm, or be confined to a single region of a chromosome (Kloosterman et al., 2011; Magrangeas et al., 2011; Stephens et al., 2011; Forment et al., 2012; Maher and Wilson, 2012; Rausch et al., 2012; Korbel and Campbell, 2013; Kloosterman et al., 2014). These rearrangements usually appear crisscrossing the involved regions (Stephens et al., 2011), and chromosomes affected by chromothripsis show a characteristic pattern of copy number oscillations between two (or occasionally three) copy number states (Stephens et al., 2011; Korbel and Campbell, 2013). By far the simplest explanation of such rearrangements is that, at some stage during cancer development, distinct chromosomes or chromosomal regions are broken into many segments and then inaccurately stitched back together by DNA repair mechanisms (Meyerson and Pellman, 2011; Forment et al., 2012; Maher and Wilson, 2012). Chromothripsis was initially observed in chronic lymphocytic leukaemia (CLL), but it is present in a wide range of human cancers, including multiple myeloma (MM), acute lymphoblastic leukaemia (ALL), acute myeloid leukaemia (AML), Hodgkin lymphoma, bone cancers, medulloblastoma, neuroblastoma, colorectal cancer and melanoma. Some congenital disorders also show chromothripsis (Kloosterman et al., 2011; Magrangeas et al., 2011; Meyerson and Pellman, 2011; Stephens et al., 2011; Edelmann et al., 2012; Maher and Wilson, 2012; Rausch et al., 2012; Boeva et al., 2013; Hirsch et al., 2013; Kloosterman and Cuppen, 2013; Korbel and Campbell 2013;Mackinnon and Campbell, 2013; Nagel et al., 2013; Kloosterman et al., 2014; Li et al., 2014).

Myelodysplastic syndromes (MDS) are a clinically heterogeneous group of clonal hematopoietic stem cell disorders characterized by morphological dysplasia, ineffective haematopoiesis and peripheral blood cytopenias (Cazzola et al., 2013). MDS and chronic myelomonocytic leukaemia (CMML), an entity sharing features of myelodysplastic syndromes and chronic myeloproliferative disorders (Such et al., 2011), have a highly variable clinical course (Malcovati et al., 2005). The presence of chromosomal abnormalities is a recurrent hallmark of both MDS and CMML patients, with consequences for their diagnosis, risk stratification and prognosis (Greenberg et al., 2012; Cazzola et al., 2013; Such et al., 2013). In fact, these genetic changes are directly incorporated into the Revised International Prognostic Scoring System (IPSS-R) for MDS and the CMML-Specific Prognostic Scoring System (CPSS) (Greenberg et al., 2012; Such et al., 2013). In addition, gene mutations are also very frequent (80-90%) in MDS and related myeloid neoplasms (Papaemmanuil et al., 2013; Haferlach et al., 2014). These mutations affect transcription factors such as *RUNX1* and *BCOR*, epigenetic modulators such as *TET2*, *DNMT3A*, *IDH1/2*, *ASXL1* and *EZH2*, tumour suppressor genes such as *TP53*, several components of the RNA-splicing machinery such as *SF3B1*, *SRSF2*, *U2AF1*, and *ZRSR2*, genes involved in DNA replication such as *SETBP1*, and genes of the cohesin complex such as *STAG2*, *RAD21*, *SMC1A*, and *SMC3*. The list of genes carrying mutations involved in the pathogenesis of MDS is still growing (Bejar et al., 2011; Cazzola et al., 2013; Papaemmanuil et al., 2013; Haferlach et al., 2014).

Conventional metaphase cytogenetics (CC) is still the gold standard for karyotypic studies; however, diagnosis and prognostication may be difficult in the 10-15% of patients with non-informative cytogenetics, due to the absence of mitosis, or the 40-60% with a normal karyotype (Thiel et al., 2011; Merkerova et al., 2012). Additionally, the presence of complex karyotypes with three or more chromosomal abnormalities may hinder the identification of the chromosomes involved in these changes (Martínez-Ramírez et al., 2005). Therefore, these CC techniques are not sufficient for a thorough study of these myeloid malignancies. The use of molecular genome-wide scanning techniques allows the identification of cryptic abnormalities in patients with a normal karyotype and the better characterization of unbalanced genetic changes (Heinrichs et al., 2009; Kolquist et al., 2011; Thiel et al., 2011; Merkerova et al., 2012; Papaemmanuil et al., 2013; Haferlach et al., 2014). Thus, array-based karyotyping revealed MDS and related myeloid neoplasms with a normal karyotype to have one or more genomic abnormalities, including deletions of *TET2* and *RUNX1* genes (40%) (Heinrichs et al., 2009; Thiel et al., 2011; Merkerova et al., 2012). These studies suggest the potential clinical utility of such genome-wide scanning techniques for the management of MDS and related diseases.

To gain insight into the characterization of the molecular changes in MDS and to explore novel

genetic abnormalities occurring in these disorders, an integrative study combining array-based comparative genomic hybridization (aCGH) and next-generation sequencing (NGS) in a series of MDS and MDS/MPN patients was carried out. The results demonstrated the presence of infrequent but recurrent chromothripsis involving chromosome 13 in these diseases and also showed that aCGH could be used in a clinical setting as a complementary method to conventional cytogenetics for identifying copy number changes in MDS and CMML patients.

MATERIALS AND METHODS

Patients

A total of 301 patients diagnosed with MDS (n=240) or MDS/MPN (n=61) at the time of diagnosis were studied. The main clinical characteristics of the patients are summarised in Table 1. The median age was 77 years (range, 11-93 years), and 191 patients (63.5%) were male. Diagnoses were established according to the 2008 World Health Organization criteria (Vardiman et al., 2009) (Table 1). This study was performed in accordance with the Declaration of Helsinki guidelines, and was approved by the Local Ethical Committees. All patients provided written informed consent.

Array-based comparative genomic hybridization studies

Genome-wide DNA copy number abnormalities (CNAs) were analysed in all samples using the Human CGH 12x135K Whole-Genome Tiling v3.0 Array (Roche NimbleGen, Madison, WI). For sample preparation and hybridization the NimbleGen CGH array standard protocol was followed (Robledo et al., 2011) (See Supplementary Methods). All detected CNAs were carefully reviewed to identify regions overlapping those previously reported to be copy number variants (CNVs) in the Database of Genomic Variants (http://dgv.tcag.ca/); these were excluded from subsequent analysis. Genomic abnormalities were interpreted and reported in accordance with the International System for Human Cytogenetic Nomenclature (ISCN 2013) guidelines (Shaffer et al., 2013). All the array data discussed in this publication have been deposited in NCBI's Gene Expression Omnibus and are accessible through GEO Series accession number GSE67682.

Additionally, conventional metaphase cytogenetic results were available from all patients. Based on CC results, patients were divided into three groups:

TABLE 1. Main characteristics of the whole series of patients included in the study

Variables	Median [Range]
Number of patients	301
Gender (Male/Female)	191 / 110
Age (years)	77 [11 - 93]
Peripheral blood values	
Hemoglobin level (g/dl)	10 [4 - 17]
Neutrophil count ($x10^9$/L)	2.5 [0.1 - 90]
Platelet count ($x10^9$/L)	125 [5 - 1018]
Bone marrow blasts (%)	1.4 [0 - 20]
WHO 2008 classification	
MDS	**240**
RCUD	20
RARS	11
RCMD	147
RAEB-1	23
RAEB-2	23
MDS-U	9
MDS del(5q)	7
MDS/MPN	**61**
CMML-1	51
CMML-2	7
RARS-T	3
Conventional cytogenetics	
Normal	**216**
≤ 10 metaphases	14
11 - 19 metaphases	38
≥ 20 metaphases	164
Abnormal	**45**
-5/del(5q)	9
double including del(5q)	1
-7/del(7q)	1
+8	3
del(11q)	2
del(20q)	1
-Y	8
complex (≥3 abnormalities)	12
any other single abnormality	8
Non-informative	**40**

WHO, World Health Organization; MDS, myelodysplastic syndromes; RCUD, refractory cytopenia with unilineage dysplasia; RARS, refractory anemia with ringed sideroblasts; RCMD, refractory cytopenia with multilineage dysplasia; RAEB, refractory anemia with excess of blasts; MDS-U, MDS unclassified; MDS del(5q), MDS associated with isolated del(5q); MDS/MPN, myelodysplastic/myeloproliferative neoplasms; CMML, chronic myelomonocytic leukemia; RARS-T, RARS with thrombocytosis; del, deletion.

patients with non-informative cytogenetics (n=40, 13.3%) due to the absence of mitosis, patients with an abnormal karyotype (n=45), and patients with a normal karyotype (n=216, 71.7%). The latter group was further divided into three categories according to the number of good-quality metaphases evaluated: ≥20 metaphases (n=164), between 11 and 19 metaphases (n=38), and ≤10 metaphases (n=14). Detailed information about the cytogenetic groups is summarised in Table 1, and the cytogenetic abnormalities found in the whole series are listed in Supplementary Table S1.

All genomic changes found by aCGH but not detected by conventional metaphase cytogenetics were validated by interphase fluorescence in situ hybridization (FISH), in the case of large recurrent deletions and gains, or by using an independent genome-wide analysis of DNA copy number changes with the SurePrint G3 Human CGH Microarray (8x60k) (Agilent Technologies, Palo Alto, CA) for small recurrent and individual abnormalities.

Next-generation sequencing studies

Mutations in *DNMT3A*, *TET2*, *RUNX1*, *TP53* and *BCOR* genes were screened by amplicon-based next-generation sequencing (NGS) in selected cases using 454 Titanium amplicon chemistry (Roche Applied Science, Penzberg, Germany). Briefly, all coding exons of *TET2*, *RUNX1* and *BCOR*, exons 7-23 of *DNMT3A* and exons 4-11 of *TP53* were covered by 27, 7, 29, 16 and 8 amplicons, respectively (Supplementary Table S3). Amplicon libraries were prepared following the manufacturer's recommendations and previously described methods (Kohlmann et al., 2011) (See Supplementary Methods).

For the detection of variants, all amplicon reads were analysed with the Sequence Pilot software (v3.5.2; JSI medical systems, Kippenheim, Germany) and GS Amplicon Variant Analyzer Software (v2.9; Roche Applied Science). Single nucleotide polymorphisms (SNPs) (http://www.ncbi.nlm.nih.gov/projects/SNP/) and variants within introns were not scored.

RESULTS

Chromothripsis on chromosome 13 is a recurrent abnormality in high-risk MDS

The analysis of copy number profiles derived from aCGH data identified complex genomic rearrangements on the aCGH chromosome plots, showing multiple non-contiguous CNAs, with the hallmarks of chromothripsis. Based on previous studies, evidence of chromothripsis was defined as the presence of at least ten changes in segmental copy number between two or three copy number states on an individual chromosome (Rausch et al., 2012). Using these criteria, three MDS cases (1.2%) exhibited chromothripsis (Fig 1A and B), with more than 11 copy number changes involving exclusively chromosome 13. The copy number states rapidly alternate between one (deletion), two (normal) and three (gain) copies. The patterns of genomic alteration were different between the three high-risk MDS patients. However, it should be pointed out that involvement of a total of 91 genes mapping on chromosome 13 were common to the three patients. As examples *XPO4*, *FLT3* and *FLT1* were commonly amplified genes; *BRCA2* and *RB1* were commonly deleted genes (Supplementary Table S6). All these results were validated using an independent microarray (SurePrint G3 Human CGH Microarray, 8x60k, Agilent Technologies). The three patients with chromothripsis were diagnosed as high-risk MDS (RAEB 3/40; 7.5%), two of them were RAEB-1, with 6% and 8% of BM blasts, respectively, and the remaining patient was RAEB-2, with 12% of BM blasts. All three patients died within one year. All of them had a complex karyotype revealed by aCGH (3/17; 17.6%; patients #026, #027, #072; Fig 1A), with a median of 21 CNAs (range, 19-33) throughout the whole genome. The three patients showed genomic losses on 5q23.2-q35.3, two of them also carried losses on 7q22.3-q36.3 and 15q11.1-21.2 (Supplementary Table S1).

All cases showing chromothripsis carried *TP53* mutations as revealed by NGS (Fig 1C). Specifically, two missense mutations (p.Lys132Arg, p.Pro278Leu) and two nonsense mutations (p.Gln136*, p.Gln167*) were identified. The p.Gln167* and p.Lys132Arg mutations, located in exon 5, were observed in one patient each with a variant allele frequency (VAF) of 93.5% and 92.5%, respectively. However, the p.Gln136* and p.Pro278Leu mutations were observed in the same patient with VAFs of 46.5% and 43.5%, and were located on exon 5 and exon 8, respectively. Taken together, most of these mutations affected exon 5 of *TP53*, and all of them were located in the sequence-specific DNA binding domain, which plays a central role in transcriptional transactivation.

aCGH and NGS allow the identification of hidden recurrent genetic CNAs and gene mutations in MDS

A total of 285 abnormalities were identified (1-33 changes per patient) in 71 of the entire series of 301

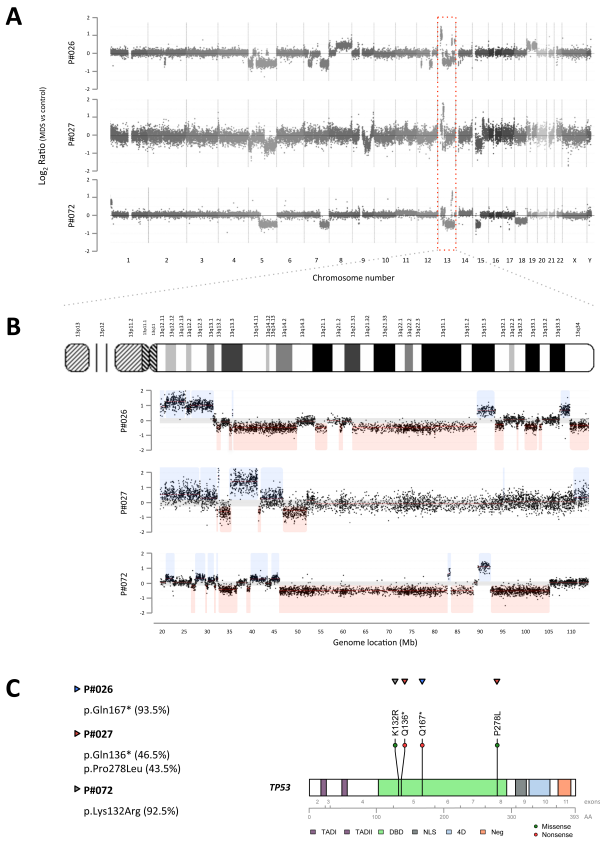


Figure 1. **Recurrent chromothripsis on chromosome 13 in high-risk MDS**. (**A**) Whole genome view ratio plots derived from aCGH data of MDS patients (#026, #027 and #072) showing chromothripsis on chromosome 13, indicated by the red-shaded box. The three patients had complex karyotypes: patient #026 had 33 aberrations and affecting seven chromosomes; patient #027 had 19 abnormalities affecting five chromosomes; patient #072 had 21 aberrations affecting six chromosomes. The Y-axis represents the $\log_2$ ratio values of MDS:control signal intensities for each probe. The X-axis illustrates all the probes in the array sorted by chromosome and physical mapping position. Chromosome numbers are indicated below the X-axis. (**B**) Detailed view of the whole chromosome 13 in patients #026, #027 and #072 showing a complex pattern of alternating copy number gains and losses. Boxes shaded in pale-red and pale-blue depict copy number losses and gains, respectively. Patient #026 had 18 alternating copy number changes, patient #027 had 11, and patient #072 had 16 changes along the whole chromosome 13. Copy number profiles differed between these patients. The Y-axis represents $\log_2$ ratios and the X-axis shows all probes of chromosome 13 sorted by chromosome position. Genomic location (Mb) is indicated below the X-axis. (**C**) Distribution of *TP53* mutations identified by amplicon-based deep sequencing in the three MDS patients with chromothripsis. All *TP53* mutations were located in the sequence-specific DNA binding domain. One patient had two mutations in heterozygosis, while the other two patients had one mutation each in homozygosis. The variant allele frequencies (VAFs) are represented in brackets. Each circle represents a mutation found in one patient. Green and red circles depict missense and nonsense mutations, respectively. Each patient is illustrated by a different-colour triangle. The complete coding region of *TP53* is illustrated and the respective exons and amino acid (AA) positions are indicated at the bottom. The following protein domains are shown: TAD1 and TAD2, amino-terminal transactivation domains 1 and 2; DBD, sequence-specific DNA-binding domain; NLS, nuclear localization signalling domain; 4D, carboxy-terminal tetramerization domain; Neg, negative regulation domain. Gene variants were represented using the R package "R453PlusToolbox" (Klein et al., 2011).

patients (23.6%): 61 of 244 (25%) were MDS patients, while 10 of 58 (17.3%) had a diagnosis of CMML. Copy number losses (72.6%) were more frequent than gains (27.4%). The detected CNAs were distributed amongst all chromosomes except for chromosomes 14 and 16.

Among the global series, CNAs were present in 9.3% of the normal karyotype patients, in 86.7% of cases with abnormal cytogenetics, and in 30% of patients with unsuccessful cytogenetic analyses (Supplementary Fig S1). The most frequent large recurrent aberrations were del(5q) (35.2%), del(20q) (18.3%), -Y (14.1%), trisomy 8 (14.1%), del(7q) (12.7%), +1/+1q (7%), -18/del(18p) (5.6%), del(17p) (5.6%), del(11q) (5.6%), del(4q) (4.2%), del(15q) (4.2%), and del(12q) (4.2%) (Supplementary Fig S2). The most frequent aberrations in CMML were –Y (5.2%) and gains on 1q (3.4%), while in RCMD losses on 5q (6.1%), 20q (5.4%), and –Y (3.4%) predominated. By contrast, in RAEB patients, losses involving 5q (21.7%), 7q (13%) and 17p (6.5%), and trisomy 8 (10.9%) were frequently observed (Supplementary Fig S2). In addition, in 38 of 301 cases (12.6%), aCGH revealed the presence of 81 small aberrations (≤5 Mb), which were below the detection limit of CC. In 21 of these cases only one cryptic CNA was detected: 17 deletions and four gains. In the other 17 cases, two or more cryptic aberrations were observed, consisting of 49 deletions and 11 gains. Notably, these cryptic CNAs involved regions such as 2p23.3, 4q24, 5q33.1, 7q22.1, 21q22.12, 21q22.3 and Xp14, where genes implicated in the pathogenesis of MDS and MDS/MPN are located, including *DNMT3A*, *TET2*, *SPARC*, *CUX1*, *RUNX1*, *U2AF1* and *BCOR*, respectively.

Furthermore, an in-depth analysis by NGS of the regions with genomic deletions of >100 kb by aCGH, where recurrently mutated genes in MDS and CMML are located, was carried out. Thus, aCGH analysis identified two cases with a deletion in 2p23.3 (*DNMT3A*)

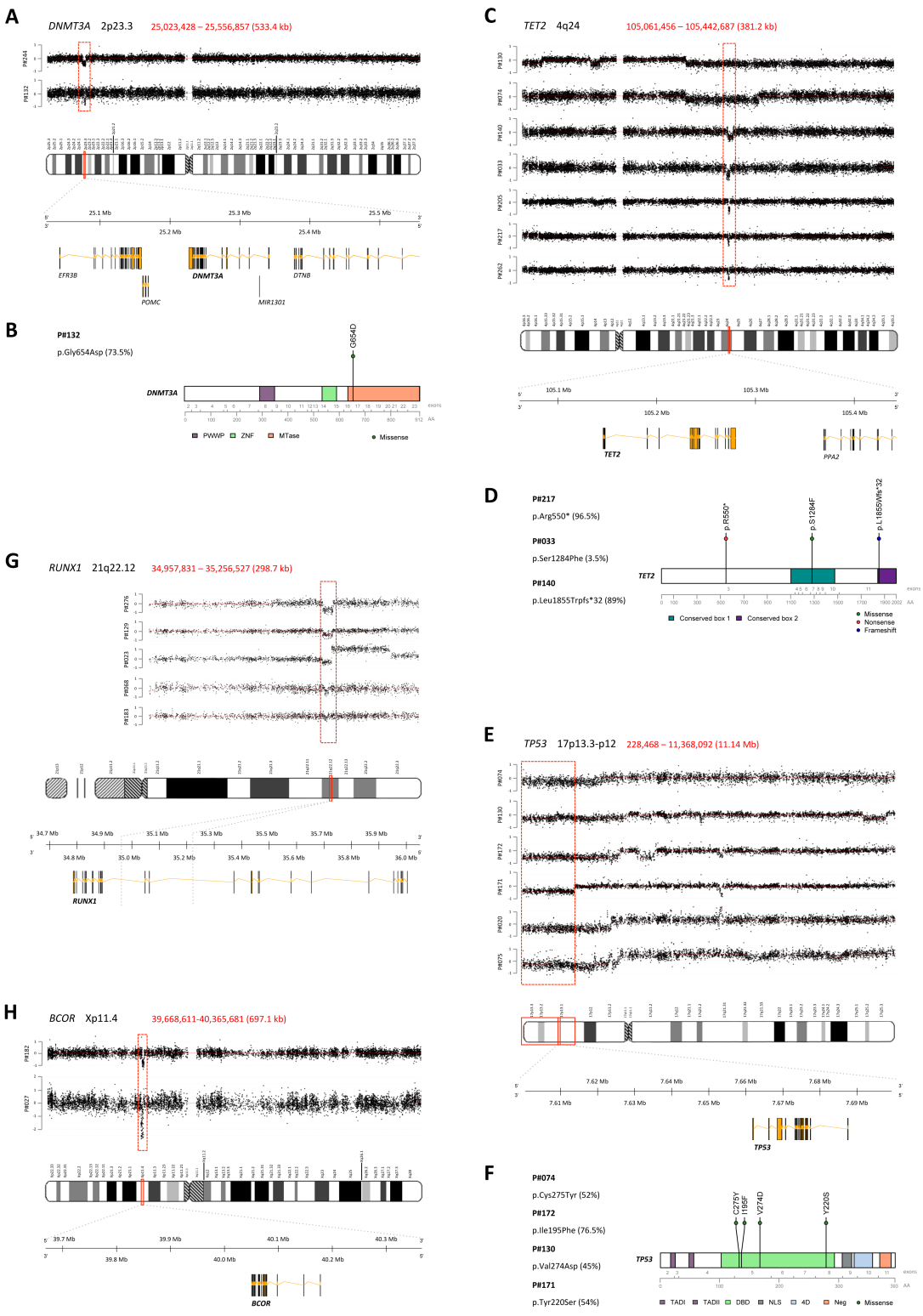

A
DNMT3A 2p23.3 25,023,428 – 25,556,857 (533.4 kb)
P#244
P#132
25.1 Mb
25.3 Mb
25.5 Mb
25.2 Mb
25.4 Mb
EFR3B
DNMT3A
DTNB
POMC
MIR1301
B
P#132
p.Gly654Asp (73.5%)
G654D
DNMT3A
exons
AA
PWWP
ZNF
MTase
Missense
C
TET2 4q24 105,061,456 – 105,442,687 (381.2 kb)
P#130
P#074
P#140
P#033
P#205
P#217
P#262
105.1 Mb
105.3 Mb
105.2 Mb
105.4 Mb
TET2
PPA2
D
P#217
p.Arg550* (96.5%)
P#033
p.Ser1284Phe (3.5%)
P#140
p.Leu1855Trpfs*32 (89%)
p.R550*
p.S1284F
p.L1855Wfs*32
TET2
exons
AA
Conserved box 1
Conserved box 2
Missense
Nonsense
Frameshift
E
TP53 17p13.3-p12 228,468 – 11,368,092 (11.14 Mb)
P#074
P#130
P#172
P#171
P#020
P#075
7.62 Mb
7.64 Mb
7.66 Mb
7.68 Mb
7.61 Mb
7.63 Mb
7.65 Mb
7.67 Mb
7.69 Mb
TP53
F
P#074
p.Cys275Tyr (52%)
P#172
p.Ile195Phe (76.5%)
P#130
p.Val274Asp (45%)
P#171
p.Tyr220Ser (54%)
C275Y
I195F
V274D
Y220S
TP53
exons
AA
TADI
TADII
DBD
NLS
4D
Neg
Missense
G
RUNX1 21q22.12 34,957,831 – 35,256,527 (298.7 kb)
P#276
P#125
P#023
P#068
P#183
34.7 Mb
34.9 Mb
35.1 Mb
35.3 Mb
35.5 Mb
35.7 Mb
35.9 Mb
34.8 Mb
35.0 Mb
35.2 Mb
35.4 Mb
35.6 Mb
35.8 Mb
36.0 Mb
RUNX1
H
BCOR Xp11.4 39,668,611-40,365,681 (697.1 kb)
P#182
P#027
39.7 Mb
39.9 Mb
40.1 Mb
40.3 Mb
39.8 Mb
40.0 Mb
40.2 Mb
BCOR

Figure 2. **Combination of aCGH and NGS analysis for regions with frequently mutated genes in MDS and MDS/MPN**. (**A**, **C**, **E**, **G**, **H**) Detailed view of the whole chromosomes 2, 4, 17, 21 and X, where recurrent regions of deletion where found by aCGH, indicated by the red-shaded box. A magnified view of the minimal deleted regions with a schematic diagram showing the genes included within the deletion. For all figures, genomic locations are indicated in Mb, and the chromosome position (bp) and size (kb) of the minimal deleted regions are indicated in the upper part of each chromosome view ratio plots. The Y-axis represents the $\log_2$ ratio values and all probes for each chromosome are sorted by genomic position along the X-axis. (**A**) A 533.4-kb deletion on 2p23.3 affecting the *DNMT3A* locus. (**C**) A 381.2-kb deletion on 4q24 affecting the *TET2* locus. (**E**) An 11.14-Mb deletion on 17p13.3-p12 affecting the *TP53* locus. (**G**) A 298.7-kb deletion on 21q22.12 affecting the *RUNX1* locus. (**H**) A 697.1-kb deletion on Xp11.4 affecting the *BCOR* locus. Genes were represented using the R package "GenomeGraphs". (**B**, **D**, **F**) Distribution of *DNMT3A*, *TET2* and *TP53* mutations identified by targeted amplicon-based deep sequencing. The variant allele frequencies (VAFs) are represented in brackets. The complete coding regions of *DNMT3A*, *TET2* and *TP53* are illustrated and the respective exons and amino acid (AA) positions are indicated below. Each circle represents a mutation found in a single patient. Green, red and blue circles depict missense, nonsense and frameshift mutations, respectively. (**B**) One patient carried a *DNMT3A* missense mutation located in the MTase domain. The following protein domains are shown: PWWP, proline-tryptophan-tryptophan-proline domain; ZNF, zinc finger domain; MTase, methyltransferase domain. (**D**) Three patients with a *TET2* deletion harbored one nonsense, one missense and one frameshift mutation each. The two evolutionarily conserved domains, boxes 1 and 2, are shown. (**F**) Four patients carried one *TP53* missense mutation each. The following protein domains are shown: TAD1 and TAD2, amino-terminal transactivation domains 1 and 2; DBD, sequence-specific DNA-binding domain; NLS, nuclear localization signalling domain; 4D, carboxy-terminal tetrame-rization domain; Neg, negative regulation domain. Gene variants were represented using the R package "R453PlusToolbox" (Klein et al., 2011).

(Fig 2A), seven cases with a deletion in 4q24 (*TET2*) (Fig 2C), six cases with a 17p13 deletion (*TP53*) (Fig 2E), five cases with a deletion in 21q22 (*RUNX1*) (Fig 2G) and two cases with an Xp11.4 deletion (*BCOR*) (Fig 2H). NGS studies detected that one patient with a *DNMT3A* deletion carried a missense mutation (p.Gly654Asp) in the other allele with a VAF of 73.5%. This mutation was located in the methyltransferase domain (Fig 2B). Three patients harbouring a *TET2* deletion carried one nonsense mutation (p.Arg550*), one missense mutation (p.Ser1284Phe) and one frameshift deletion (p.Leu1855Trpfs*32) each in the non-deleted allele. The variant allele frequencies were 96.5%, 3.5% and 89%, respectively. The latter two mutations affected the two evolutionarily conserved domains in the TET family proteins (Fig 2D). Of the six patients with a 17p13 deletion affecting the *TP53* locus, four carried one missense mutation each (p.Cys275Tyr, p.Ile195Phe, p.Val274Asp, p.Tyr220Ser) with VAFs of 52%, 76.5%, 45% and 54%, respectively. These *TP53* mutations were located in the sequence-specific DNA binding domain. None of the other studied genes showed mutations in those patients with losses in the regions of interest (Fig 2F).

Relationship between cytogenetic results and aCGH data

Array CGH results were compared with the cytogenetic data from each patient. 56 of 83 chromosomal imbalances previously identified by CC were detected by aCGH. A remarkably high correlation between CC and aCGH results was observed in this series. In addition, given that patients from different groups ascertained by CC were included in the study, we decided to analyse the concordance between CC and aCGH, but considering each cytogenetic group of patients separately: the non-informative cytogenetic, normal and abnormal karyotype groups.

Focusing on the 216 patients with a normal karyotype as determined by CC, the aCGH results were in excellent agreement with the cytogenetics of those patients with ≥20 and 11-19 metaphases studied (92.1 and 89.5%, respectively). However, only 78.6% of those patients with ≤10 successful metaphases and no changes by CC displayed no copy number changes by aCGH (Supplementary Fig S3). Thus, 20 patients (9.3%) with a normal karyotype as determined by CC showed at least one genomic abnormality by aCGH. Only one chromosome was affected in 16 of these patients. Considering those patients with an aberrant karyotype, aCGH revealed the same genomic abnormalities as previously identified by CC in 86.7% of cases. Indeed, only six of the 45 (13.3%) cases with an abnormal karyotype established by G-banding analysis showed no copy number abnormalities with aCGH. One of these patients had a balanced translocation, which was not detectable by aCGH, and four patients had chromosomal abnormalities in three metaphases of the analysed cells, clonal cell populations below the detection limit of the aCGH. Detailed information about these discordant cases is presented in Supplementary Table S1. With respect to the patients with unsuccessful cytogenetic analyses, 70% of cases displayed a normal aCGH profile, while 30% had at least one copy number aberration. Four patients (three high-risk MDS and one CMML) had a complex karyotype, defined by the presence of at least five copy number changes, as revealed by aCGH.

DISCUSSION

The presence of specific chromosomal abnormalities and genetic changes is a hallmark of MDS

(Haase et al., 2007; Bejar et al., 2011; Mallo et al., 2011; Cazzola et al., 2013; Papaemmanuil et al., 2013; Haferlach et al., 2014). In this study, we analysed a large cohort of MDS patients by integrating two genetic methodologies: array-based comparative genomic hybridization (aCGH) and amplicon-based deep sequencing (NGS). Our results demonstrated the presence of chromothripsis, inferred from aCGH profiles, as an infrequent but recurrent genomic abnormality in high-risk MDS. In addition, our results showed that the combination of these conventional and genome-wide scanning approaches enables a better characterization of MDS and related neoplasms, and provides new information that could improve the current diagnostic and treatment of these patients.

Chromothripsis is a genetic abnormality in which tens to hundreds of clustered genomic rearrangements occur in a one-step catastrophic event (Kloosterman et al., 2011; Magrangeas et al., 2011; Stephens et al., 2011; Forment et al., 2012; Maher and Wilson, 2012; Rausch et al., 2012; Korbel and Campbell, 2013; Kloosterman et al., 2014). In the present study, chromothripsis was observed in three cases of high-risk MDS (two cases of RAEB-1 and one patient with RAEB-2), two entities known to progress gradually to a more aggressive disease. The occurrence of chromothripsis in myeloid malignancies has been demonstrated in AMLs by SNP array-based copy number profiling, in which 8% of AML patients carried massive and complex rearrangements consistent with chromothripsis (Rausch et al., 2012; Zemanova et al., 2014). The molecular basis of this genomic chaos remains to be fully explained. However, several reports suggested that mechanisms such as ionizing radiation, premature chromosome compaction, DNA replication stress, telomere shortening, abortive apoptosis and *TP53* mutations could be involved in the origin of this complex genomic abnormality (Meyerson and Pellman, 2011; Crasta et al., 2012; Forment et al., 2012; Jones and Jallepalli, 2012; Maher and Wilson, 2012; Rausch et al., 2012; Korbel and Campbell, 2013). In addition, approximately 50% of AML patients carrying *TP53* mutations and approximately 40% of complex karyotype AML patients displayed chromothripsis, while only 1% of AML with wild-type *TP53* and no cases with non-complex karyotype showed this aberration (Rausch et al., 2012; Zemanova et al., 2014). Moreover, it was also demonstrated that almost all medulloblastomas showing evidence of chromothripsis had *TP53* mutations (Rausch et al., 2012). These findings reinforce the link between somatically acquired *TP53* mutations and the presence of complex karyotypes with chromothripsis. It is of particular interest that the three MDS cases with chromothripsis reported in this study had complex karyotypes as revealed by aCGH and carried *TP53* mutations, a previously described association (Rausch et al., 2012). Additionally, the outcome of the three cases was poor, a feature that is in accordance with high-risk MDS patients, the presence of a complex karyotype or *TP53* mutations (Greenberg et al., 2012; Cazzola et al., 2013). This complex genomic abnormality has not previously been comprehensively described in MDS patients, perhaps because the previous high-resolution copy number studies in MDS mainly concerned cases with a normal karyotype (Heinrichs et al., 2009; Thiel et al., 2011; Merkerova et al., 2012; Haferlach et al., 2014). Other copy number studies in MDS with abnormal and complex karyotypes have been reported. These studies aimed to analyse the relationship of copy number to the CC data and their prognostic impact, and they did not show the presence of chromothripsis (Martínez-Ramírez et al., 2005; O'Keefe et al., 2007; Slovak et al., 2010; Kolquist et al., 2011; Tiu et al., 2011). It should be noted that, in our study, chromothripsis was seen exclusively to affect the entire chromosome 13 in all three MDS patients. The presence of this genomic chaos restricted to a single chromosome has been described. Chromothripsis is typically found to affect different chromosomes at random, but was reported to involve chromosome 16p in 3/7 MM patients, and chromosome 21 in 5/9 iAMP21 ALL patients (Magrangeas et al., 2011; Li et al., 2014). However, the presence of chromothripsis involving chromosome 13 has not been previously reported in myeloid malignancies (Rausch et al., 2012; Cancer Genome Atlas Research Network, 2013; Mackinnon and Campbell, 2013; Zemanova et al., 2014). In this study, chromothripsis affected only chromosome 13, with patterns of genomic alteration differing between the three high-risk MDS patients. *FLT3* was commonly amplified, while *BRCA2* and *RB1* were commonly deleted in the cases with chromothripsis. *FLT3* is an oncogene that regulates hematopoietic stem cell differentiation, proliferation and survival. *FLT3*-activating mutations were recurrently described in myeloid malignancies, mainly in AML, and are associated with poor prognosis (Bains et al., 2011). *RB1* and *BRCA2* abnormalities play a role in the development of several cancers and are considered

to be tumour suppressor genes (Di Fiore et al., 2013; Lee, 2014). Our results suggest the involvement of the *FLT3* gene and *BRCA2* and *RB1* inactivation in the pathogenesis of some cases of MDS.

The present study revealed the presence of cryptic abnormalities that were not targeted by FISH and that were below the threshold of resolution by conventional cytogenetics. In fact, 12.6% of MDS patients showed cryptic changes. Some of these submicroscopic CNAs involved regions with genes of known significance in MDS pathogenesis and were deletions in 2p23.3 (*DNMT3A*), 4q24 (*TET2*), 5q33.1 (*SPARC*), 7q22.1 (*CUX1*), 21q22.12 (*RUNX1*) and Xp11.4 (*BCOR*), and gains in 21q22.3 (*U2AF1*), that were detected in 19 patients (Bejar et al., 2011; Kolquist et al., 2011; Thiel et al., 2011; Cazzola et al., 2013). These regions were equally likely to be involved in low- and high-risk MDS or CMML. In addition, the genes included in these recurrent cryptic deletions were further investigated by NGS to identify whether mutations occurred in the other allele. The sequencing results showed that only two of the seven cases with the *TET2* deletion carried *TET2* mutations, while only one of the two cases with losses in 2p23.3 showed *DNMT3A* mutations. Our results did not indicate any correlation between the presence of deletions and mutations in these MDS patients, which is in accordance with previously reported data (Haferlach et al., 2014). Therefore, our results support the idea that conventional cytogenetic and aCGH studies could be complemented by the sequencing of multi-gene panels, which have already been described for MDS and related myeloid neoplasms, instead of single genes, in routine workflows for the study of these group of patients.

The presence of cytogenetic changes is a keystone in the prognosis of MDS and CMML patients. In fact, conventional cytogenetics is essential in most prognostic systems to stratify these haematological disorders (Greenberg et al., 2012; Cazzola et al., 2013; Such et al., 2013). However, in some subsets of MDS and CMML, conventional cytogenetic techniques fail to provide any results due to a lack of cell growth during culturing, and consequently some aberrations may be missed, making the diagnosis and prognostic stratification very difficult. Our study showed that 30% of karyotype failures carried genomic abnormalities revealed by aCGH. Four patients had complex karyotypes, two had trisomy 8 and one showed a del(5q), while the others showed deletions in 9p, 12q, 17p, 21q22, and -Y, and gains in 1q and 15q (Supplementary Table S1). Therefore, identifying these clinically relevant lesions is significant in patients with failed CC results. The clinical utility of SNP-A as a karyotyping tool in a series of MDS patients with unsuccessful cytogenetics has been previously demonstrated (Arenillas et al., 2013). This method was also useful in those MDS patients with a normal karyotype when fewer than 20 good-quality metaphases are available for analysis. We demonstrated that in 10% of cases with 11-19 successful metaphases and in 21% of those with ≤10 harboured genomic aberrations aCGH will provide additional information that could redefine the prognostic risk of these patients, as previously suggested (Coleman et al., 2011). Thus, at least 20 metaphases need to be analysed for a karyotype to be considered normal (Haferlach et al., 2007). Therefore, the use of aCGH enabled the prognostic stratification according to the IPSS-R that could change the clinical management of this group of patients.

In summary, the present report describes the presence of a high incidence of genomic changes in MDS and CMML patients by the integrative analysis of several molecular genetic methodologies. In addition to well-known copy number defects, the presence of chromothripsis involving chromosome 13 was a novel recurrent change in high-risk MDS patients. Moreover, aCGH analysis revealed the presence of cryptic abnormalities in genomic regions where MDS-related genes, such as *TET2*, *DNMT3A*, *RUNX1* and *BCOR*, are located. Thus, the integrative analysis of conventional cytogenetics, aCGH and NGS in MDS will provide a better understanding of the molecular abnormalities occurring in these patients, and could improve the clinical management of MDS. The potential diagnostic and prognostic value of these new genomic abnormalities should be studied in prospective studies.

ACKNOWLEDGEMENTS

This work was partially supported by Grants from the Spanish Fondo de Investigaciones Sanitarias FIS (PI12/00281); Proyectos de Investigación del SACYL (BIO/SA47/13; BIO/SA52/14; GRS/874/A13; GRS 994/A/14); COST Action "EuGESMA" (BM0801); Red Temática de Investigación Cooperativa en Cáncer (RTICC), Instituto de Salud Carlos III (ISCIII), Spanish Ministry of Economy and Competitiveness & European Regional Development Fund (ERDF) "*Una manera de hacer Europa*" (Innocampus, CEI-2010-1-0010) (RD12/0036/0069; RD12/0036/0029; RD12/

0036/0044); and the European Union Seventh Framework Programme [FP7/2007-2013] under Grant Agreement n°306242-NGS-PTL. Mutation analyses were performed using oligonucleotide primer plates designed as part of the IRON-II collaborative network of haematological laboratories applying 454 amplicon next-generation deep-sequencing. MA was supported by a "*Junta para Ampliación de Estudios*" fellowship [09-02402] of the Spanish National Research Council (*Consejo Superior de Investigaciones Científicas*, CSIC), co-funded by the European Social Fund.

The authors would like to thank Irene Rodríguez, Sara González, Teresa Prieto, Mª Ángeles Ramos, Almudena Martín, Ana Díaz, Ana Simón, María del Pozo, Vanesa Gutiérrez and Sandra Pujante from the Centro de Investigación del Cáncer, Salamanca, Spain, for their technical assistance.

CONFLICTS-OF-INTEREST

AK: current employment by AstraZeneca, UK. The remaining authors declare no competing financial interests.

SUPPORTING INFORMATION

Supplementary Methods: Patients. Array-based comparative genomic hybridization studies. Next-generation sequencing studies.

Supplementary Table S1. All genomic copy number changes detected by aCGH and all conventional cytogenetic information for the whole series.

Supplementary Table S2. Ensembl gene and transcript IDs for the five genes selected for NGS.

Supplementary Table S3. PCR primer-pair sequences for all amplicons representing DNMT3A (16), RUNX1 (7), TET2 (27), TP53 (8) and BCOR (29) genes.

Supplementary Table S4. PCR amplification mixes for NGS.

Supplementary Table S5. PCR protocols for NGS.

Supplementary Table S6. Details of chromosome 13 rearrangements in the three high-risk MDS patients affected by chromothripsis.

Supplementary Figure S1. Proportion of the whole series of patients with normal and abnormal aCGH profiles.

Supplementary Figure S2. Frequency of large recurrent genomic abnormalities and frequency of cryptic recurrent CNAs involving genes of known significance in MDS and MDS/MPN patients only seen by aCGH.

Supplementary Figure S3. Relationship between aCGH and CC studies in the normal karyotype group.

REFERENCES

Arenillas L, Mallo M, Ramos F, Guinta K, Barragán E, Lumbreras E, Larráyoz MJ, De Paz R, Tormo M, Abáigar M, Pedro C, Cervera J, Such E, Calasanz MJ, Díez-Campelo M, Sanz GF, Hernández JM, Luño E, Saumell S, Maciejewski J, Florensa L, Solé F. 2013. Single nucleotide polymorphism array karyotyping: a diagnostic and prognostic tool in myelodysplastic syndromes with unsuccessful conventional cytogenetic testing. Genes Chromosomes Cancer 52:1167-1177.

Bains A, Luthra R, Medeiros LJ, Zuo, Z. 2011. FLT3 and NPM1 Mutations in Myelodysplastic Syndromes Frequency and Potential Value for Predicting Progression to Acute Myeloid Leukemia. Am J Clin Pathol 135:62-69.

Bejar R, Stevenson K, Abdel-Wahab O, Galili N, Nilsson B, Garcia-Manero G, Kantarjian H, Raza A, Levine RL, Neuberg D, Ebert BL. 2011. Clinical effect of point mutations in myelodysplastic syndromes. N Engl J Med 364:2496-2506.

Boeva V, Jouannet S, Daveau R, Combaret V, Pierre-Eugène C, Cazes A, Louis-Brennetot C, Schleiermacher G, Ferrand S, Pierron G, Lermine A, Rio Frio T, Raynal V, Vassal G, Barillot E, Delattre O, Janoueix-Lerosey I. 2013. Breakpoint features of genomic rearrangements in neuroblastoma with unbalanced translocations and chromothripsis. PLoS One 8:e72182.

Cancer Genome Atlas Research Network. 2013. Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. N Engl J Med 368:2059-2074.

Cazzola M, Della Porta MG, Malcovati L. 2013. The genetic basis of myelodysplasia and its clinical relevance. Blood 122:4021-4034.

Coleman JF, Theil KS, Tubbs RR, Cook JR. 2011. Diagnostic yield of bone marrow and peripheral blood FISH panel testing in clinically suspected myelodysplastic syndromes and/or acute myeloid leukemia: a prospective analysis of 433 cases. Am J Clin Pathol 135:915-920.

Crasta K, Ganem NJ, Dagher R, Lantermann AB, Ivanova EV, Pan Y, Nezi L, Protopopov A, Chowdhury D, Pellman D. 2012. DNA breaks and chromosome pulverization from errors in mitosis. Nature 482:53-58.

Di Fiore R, D'Anneo A, Tesoriere G, Vento R. 2013. RB1 in cancer: different mechanisms of RB1 inactivation and alterations of pRb pathway in tumorigenesis. J Cell Physiol 228:1676-1687.

Edelmann J, Holzmann K, Miller F, Winkler D, Bühler A, Zenz T, Bullinger L, Kühn MW, Gerhardinger A, Bloehdorn J, Radtke I, Su X, Ma J, Pounds S, Hallek M, Lichter P, Korbel J, Busch R, Mertens D, Downing JR, Stilgenbauer S, Döhner H. 2012. High-resolution genomic profiling of chronic lymphocytic leukemia reveals new recurrent genomic alterations. Blood 120:4783-4794.

Forment JV, Kaidi A, Jackson SP. 2012. Chromothripsis and cancer: causes and consequences of chromosome shattering. Nat Rev Cancer 12:663-670.

Greenberg PL, Tuechler H, Schanz J, Sanz G, Garcia-Manero G, Solé F, Bennett JM, Bowen D, Fenaux P, Dreyfus F, Kantarjian H, Kuendgen A, Levis A, Malcovati L, Cazzola M, Cermak J, Fonatsch C, Le Beau MM, Slovak ML, Krieger O, Luebbert M, Maciejewski J, Magalhaes SM, Miyazaki Y, Pfeilstöcker M, Sekeres M, Sperr WR, Stauder R, Tauro S, Valent P, Vallespi T, van de Loosdrecht AA, Germing U, Haase D. 2012. Revised international prognostic scoring system for myelodysplastic syndromes. Blood 120:2454-2465.

Haase D, Germing U, Schanz J, Pfeilstöcker M, Nösslinger T, Hildebrandt B, Kundgen A, Lübbert M, Kunzmann R, Giagounidis AA, Aul C, Trümper L, Krieger O, Stauder R, Müller TH, Wimazal F, Valent P, Fonatsch C, Steidl C. 2007. New insights into the prognostic impact of the karyotype in MDS and correlation with subtypes: evidence from a core dataset of 2124 patients. Blood 110:4385-4395.

Haferlach C, Rieder H, Lillington DM, Dastugue N, Hagemeijer A, Harbott J, Stilgenbauer S, Knuutila S, Johansson B, Fonatsch C. 2007. Proposals for standardized protocols for cytogenetic

analyses of acute leukemias, chronic lymphocytic leukemia, chronic myeloid leukemia, chronic myeloproliferative disorders, and myelodysplastic syndromes. Genes Chromosomes Cancer 46:494-499.

Haferlach T, Nagata Y, Grossmann V, Okuno Y, Bacher U, Nagae G, Schnittger S, Sanada M, Kon A, Alpermann T, Yoshida K, Roller A, Nadarajah N, Shiraishi Y, Shiozawa Y, Chiba K, Tanaka H, Koeffler HP, Klein HU, Dugas M, Aburatani H, Kohlmann A, Miyano S, Haferlach C, Kern W, Ogawa S. 2014. Landscape of genetic lesions in 944 patients with myelodysplastic syndromes. Leukemia 28:241-247.

Heinrichs S, Kulkarni RV, Bueso-Ramos CE, Levine RL, Loh ML, Li C, Neuberg D, Kornblau SM, Issa JP, Gilliland DG, Garcia-Manero G, Kantarjian HM, Estey EH, Look AT. 2009. Accurate detection of uniparental disomy and microdeletions by SNP array analysis in myelodysplastic syndromes with normal cytogenetics. Leukemia 23:1605-1613.

Hirsch D, Kemmerling R, Davis S, Camps J, Meltzer PS, Ried T, Gaiser T. 2013. Chromothripsis and focal copy number alterations determine poor outcome in malignant melanoma. Cancer Res 73:1454-1460.

Jones MJ, Jallepalli PV. 2012. Chromothripsis: chromosomes in crisis. Dev Cell 23:908-917.

Klein HU, Bartenhagen C, Kohlmann A, Grossmann V, Ruckert C, Haferlach T, Dugas M. 2011. R453Plus1Toolbox: an R/Bioconductor package for analyzing Roche 454 Sequencing data. Bioinformatics 27:1162-1163.

Kloosterman WP, Hoogstraat M, Paling O, Tavakoli-Yaraki M, Renkens I, Vermaat JS, van Roosmalen MJ, van Lieshout S, Nijman IJ, Roessingh W, van 't Slot R, van de Belt J, Guryev V, Koudijs M, Voest E, Cuppen E. 2011. Chromothripsis is a common mechanism driving genomic rearrangements in primary and metastatic colorectal cancer. Genome Biol 12:R103.

Kloosterman WP, Cuppen E. 2013. Chromothripsis in congenital disorders and cancer: similarities and differences. Curr Opin Cell Biol 25:341-348.

Kloosterman WP, Koster J, Molenaar JJ. 2014. Prevalence and clinical implications of chromothripsis in cancer genomes. Curr Opin Oncol 26:64-72.

Kohlmann A, Klein HU, Weissmann S, Bresolin S, Chaplin T, Cuppens H, Haschke-Becher E, Garicochea B, Grossmann V, Hanczaruk B, Hebestreit K, Gabriel C, Iacobucci I, Jansen JH, te Kronnie G, van de Locht L, Martinelli G, McGowan K, Schweiger MR, Timmermann B, Vandenberghe P, Young BD, Dugas M, Haferlach T. 2011. The Interlaboratory RObustness of Next-generation sequencing (IRON) study: a deep sequencing investigation of TET2, CBL and KRAS mutations by an international consortium involving 10 laboratories. Leukemia 25:1840-1848.

Kolquist KA, Schultz RA, Furrow A, Brown TC, Han JY, Campbell LJ, Wall M, Slovak ML, Shaffer LG, Ballif BC. 2011. Microarray-based comparative genomic hybridization of cancer targets reveals novel, recurrent genetic aberrations in the myelodysplastic syndromes. Cancer Genet 204:603-628.

Korbel JO, Campbell PJ. 2013. Criteria for inference of chromothripsis in cancer genomes. Cell 152:1226-1236.

Lee H. 2014. Cycling with BRCA2 from DNA repair to mitosis. Exp Cell Res 329:78-84.

Li Y, Schwab C, Ryan SL, Papaemmanuil E, Robinson HM, Jacobs P, Moorman AV, Dyer S, Borrow J, Griffiths M, Heerema NA, Carroll AJ, Talley P, Bown N, Telford N, Ross FM, Gaunt L, McNally RJ, Young BD, Sinclair P, Rand V, Teixeira MR, Joseph O, Robinson B, Maddison M, Dastugue N, Vandenberghe P, Haferlach C, Stephens PJ, Cheng J, Van Loo P, Stratton MR, Campbell PJ, Harrison CJ. 2014. Constitutional and somatic rearrangement of chromosome 21 in acute lymphoblastic leukaemia. Nature 508:98-102.

Mackinnon RN, Campbell LJ. 2013. Chromothripsis under the microscope: a cytogenetic perspective of two cases of AML with catastrophic chromosome rearrangement. Cancer Genet 206:238-251.

Magrangeas F, Avet-Loiseau H, Munshi NC, Minvielle S. 2011. Chromothripsis identifies a rare and aggressive entity among newly diagnosed multiple myeloma patients. Blood 118:675-678.

Maher CA, Wilson RK. 2012. Chromothripsis and human disease: piecing together the shattering process. Cell 148:29-32.

Malcovati L, Porta MG, Pascutto C, Invernizzi R, Boni M, Travaglino E, Passamonti F, Arcaini L, Maffioli M, Bernasconi P, Lazzarino M, Cazzola M. 2005. Prognostic factors and life expectancy in myelodysplastic syndromes classified according to WHO criteria: a basis for clinical decision making. J Clin Oncol 23:7594-7603.

Mallo M, Cervera J, Schanz J, Such E, García-Manero G, Luño E, Steidl C, Espinet B, Vallespí T, Germing U, Blum S, Ohyashiki K, Grau J, Pfeilstöcker M, Hernández JM, Noesslinger T, Giagounidis A, Aul C, Calasanz MJ, Martín ML, Valent P, Collado R, Haferlach C, Fonatsch C, Lübbert M, Stauder R, Hildebrandt B, Krieger O, Pedro C, Arenillas L, Sanz MÁ, Valencia A, Florensa L, Sanz GF, Haase D, Solé F. 2011. Impact of adjunct cytogenetic abnormalities for prognostic stratification in patients with myelodysplastic syndrome and deletion 5q. Leukemia 25:110-120.

Martínez-Ramírez A, Urioste M, Melchor L, Blesa D, Valle L, de Andrés SA, Kok K, Calasanz MJ, Cigudosa JC, Benítez J. 2005. Analysis of myelodysplastic syndromes with complex karyotypes by high-resolution comparative genomic hybridization and subtelomeric CGH array. Genes Chromosomes Cancer 42:287-298.

Merkerova MD, Bystricka D, Belickova M, Krejcik Z, Zemanova Z, Polak J, Hajkova H, Brezinova J, Michalova K, Cermak J. 2012. From cryptic chromosomal lesions to pathologically relevant genes: integration of SNP-array with gene expression profiling in myelodysplastic syndrome with normal karyotype. Genes Chromosomes Cancer 51:419-428.

Meyerson M, Pellman D. 2011. Cancer genomes evolve by pulverizing single chromosomes. Cell 144:9-10.

Nagel S, Meyer C, Quentmeier H, Kaufmann M, Drexler HG, MacLeod RA. 2013. Chromothripsis in Hodgkin lymphoma. Genes Chromosomes Cancer 52:741-747.

O'Keefe CL, Tiu R, Gondek LP, Powers J, Theil KS, Kalaycio M, Lichtin A, Sekeres MA, Maciejewski JP. 2007. High-resolution genomic arrays facilitate detection of novel cryptic chromosomal lesions in myelodysplastic syndromes. Exp Hematol 35:240-251.

Papaemmanuil E, Gerstung M, Malcovati L, Tauro S, Gundem G, Van Loo P, Yoon CJ, Ellis P, Wedge DC, Pellagatti A, Shlien A, Groves MJ, Forbes SA, Raine K, Hinton J, Mudie LJ, McLaren S, Hardy C, Latimer C, Della Porta MG, O'Meara S, Ambaglio I, Galli A, Butler AP, Walldin G, Teague JW, Quek L, Sternberg A, Gambacorti-Passerini C, Cross NC, Green AR, Boultwood J, Vyas P, Hellstrom-Lindberg E, Bowen D, Cazzola M, Stratton MR, Campbell PJ, Chronic Myeloid Disorders Working Group of the International Cancer Genome Consortium. 2013. Clinical and biological implications of driver mutations in myelodysplastic syndromes. Blood 122:3616-3627; quiz 3699.

Rausch T, Jones DT, Zapatka M, Stütz AM, Zichner T, Weischenfeldt J, Jäger N, Remke M, Shih D, Northcott PA, Pfaff E, Tica J, Wang Q, Massimi L, Witt H, Bender S, Pleier S, Cin H, Hawkins C, Beck C, von Deimling A, Hans V, Brors B, Eils R, Scheurlen W, Blake J, Benes V, Kulozik AE, Witt O, Martin D, Zhang C, Porat R, Merino DM, Wasserman J, Jabado N, Fontebasso A, Bullinger L, Rücker FG, Döhner K, Döhner H, Koster J, Molenaar JJ, Versteeg R, Kool M, Tabori U, Malkin D, Korshunov A, Taylor MD, Lichter P, Pfister SM, Korbel JO. 2012. Genome sequencing of pediatric medulloblastoma links catastrophic DNA rearrangements with TP53 mutations. Cell 148:59-71.

Robledo C, García JL, Benito R, Flores T, Mollejo M, Martínez-Climent JÁ, García E, Gutiérrez NC, Piris MA, Hernández JM. 2011. Molecular characterization of the region 7q221 in splenic marginal zone lymphomas. PLoS One 6:e24939.

Shaffer LG, McGowan-Jordan J, Schmid M (eds). 2013. ISCN 2013: An International System for Human Cytogenetic Nomenclature: Recommendation of the International Standing Committee on Human Cytogenetic Nomenclature. S Karger, Basel.

Slovak ML, Smith DD, Bedell V, Hsu YH, O'Donnell M, Forman SJ, Gaal K, McDaniel L, Schultz R, Ballif BC, Shaffer LG. 2010. Assessing karyotype precision by microarray-based comparative genomic hybridization in the myelodysplastic/myeloproliferative syndromes. Mol Cytogenet 3:23.

Stephens PJ, Greenman CD, Fu B, Yang F, Bignell GR, Mudie LJ, Pleasance ED, Lau KW, Beare D, Stebbings LA, McLaren S, Lin ML, McBride DJ, Varela I, Nik-Zainal S, Leroy C, Jia M, Menzies A, Butler AP, Teague JW, Quail MA, Burton J, Swerdlow H, Carter NP, Morsberger LA, Iacobuzio-Donahue C, Follows GA, Green AR, Flanagan AM, Stratton MR, Futreal PA, Campbell PJ. 2011. Massive genomic rearrangement acquired in a single catastrophic event during cancer development. Cell 144:27-40.

Such E, Cervera J, Costa D, Solé F, Vallespí T, Luño E, Collado R, Calasanz MJ, Hernández-Rivas JM, Cigudosa JC, Nomdedeu B, Mallo M, Carbonell F, Bueno J, Ardanaz MT, Ramos F, Tormo M, Sancho-Tello R, del Cañizo C, Gómez V, Marco V, Xicoy B, Bonanad S, Pedro C, Bernal T, Sanz GF. 2011. Cytogenetic risk stratification in chronic myelomonocytic leukemia. Haematologica 96:375-383.

Such E, Germing U, Malcovati L, Cervera J, Kuendgen A, Della Porta MG, Nomdedeu B, Arenillas L, Luño E, Xicoy B, Amigo ML, Valcarcel D, Nachtkamp K, Ambaglio I, Hildebrandt B, Lorenzo I, Cazzola M, Sanz G. 2013. Development and validation of a prognostic scoring system for patients with chronic myelomonocytic leukemia. Blood 121:3005-3015.

Suela J, Alvarez S, Cigudosa JC. 2007. DNA profiling by arrayCGH in acute myeloid leukemia and myelodysplastic syndromes. Cytogenet Genome Res 118:304-309.

Thiel A, Beier M, Ingenhag D, Servan K, Hein M, Moeller V, Betz B, Hildebrandt B, Evers C, Germing U, Royer-Pokora B. 2011. Comprehensive array CGH of normal karyotype myelodysplastic syndromes reveals hidden recurrent and individual genomic copy number alterations with prognostic relevance. Leukemia 25:387-399.

Tiu RV, Gondek LP, O'Keefe CL, Elson P, Huh J, Mohamedali A, Kulasekararaj A, Advani AS, Paquette R, List AF, Sekeres MA, McDevitt MA, Mufti GJ, Maciejewski, JP. 2011. Prognostic impact of SNP array karyotyping in myelodysplastic syndromes and related myeloid malignancies. Blood 117:4552-4560.

Vardiman JW, Thiele J, Arber DA, Brunning RD, Borowitz MJ, Porwit A, Harris NL, Le Beau MM, Hellström-Lindberg E, Tefferi A, Bloomfield CD. 2009. The 2008 revision of the World Health Organization (WHO) classification of myeloid neoplasms and acute leukemia: rationale and important changes. Blood 114:937-951.

Zemanova Z, Michalova K, Buryova H, Brezinova J, Kostylkova K, Bystricka D, Novakova M, Sarova I, Izakova S, Lizcova L, Ransdorfova S, Krejcik Z, Merkerova MD, Dohnalova A, Siskova M, Jonasova A, Neuwirtova R, Cermak J. 2014. Involvement of deleted chromosome 5 in complex chromosomal aberrations in newly diagnosed myelodysplastic syndromes (MDS) is correlated with extremely adverse prognosis. Leuk Res 38:537-544.

Common and progressive gene expression changes in the progression of myelodysplastic syndromes to acute myeloid leukemia

2

María Abáigar*[1], Sara Aibar*[2], Rocío Benito[1], María Díez-Campelo[3], Fernando Ramos[4,5], Eva Lumbreras[1], Franscisco J Campos-Laborie[2], Marta Megido[6], Isabel Recio[7], Lourdes Hermosín[8], Javier Sánchez-del-Real[5], Carmen Olivier[9], Rebeca Cuello[10], Lurdes Zamora[11], Ken Mills[12], M Consuelo del Cañizo[3,13], Javier De Las Rivas[2], Jesús M Hernández-Rivas[1,3,13]

[1]*Unidad de Diagnóstico Molecular y Celular del Cáncer, Centro de Investigación del Cáncer-IBMCC (USAL-CSIC), Salamanca, Spain*

[2]*Bioinformática y Genómica Funcional, IBMCC-Centro de Investigación del Cáncer (USAL/CSIC/IBSAL), Salamanca, Spain*

[3]*Servicio de Hematología, Hospital Universitario de Salamanca, Spain*

[4]*IBIOMED, Instituto de Biomedicina, Universidad de León, Spain*

[5]*Servicio de Hematología, Hospital Universitario de León, Spain*

[6]*Servicio de Hematología, Hospital del Bierzo, Ponferrada, Spain*

[7]*Servicio de Hematología, Hospital Nuestra Señora de Sonsoles, Ávila, Spain*

[8]*Servicio de Hematología, Hospital Jerez de la Frontera, Spain*

[9]*Servicio de Hematología, Hospital General de Segovia, Spain*

[10]*Servicio de Hematología, Hospital Clínico Universitario de Valladolid, Spain*

[11]*Servicio de Hematología, Hospital Universitari Germans Trias i Pujol, Badalona, Spain*

[12]*Centre for Cancer Research and Cell Biology, Queen's University Belfast, Belfast, UK*

[13]*IBSAL, Instituto de Investigación Biomédica de Salamanca, Spain*

*These authors contributed equally to this work

Haematologica. Submitted

Common and progressive gene expression changes in the progression of myelodysplastic syndromes to acute myeloid leukemia

María Abáigar,[*1] Sara Aibar,[*2] Rocío Benito,[1] María Díez-Campelo,[3] Fernando Ramos,[4,5] Eva Lumbreras,[1] Francisco J Campos-Laborie,[2] Marta Megido,[6] Isabel Recio,[7] Lourdes Hermosín,[8] Javier Sánchez-del-Real,[5] Carmen Olivier,[9] Rebeca Cuello,[10] Lurdes Zamora,[11] Ken Mills,[12] M Consuelo del Cañizo,[3,13] Javier De Las Rivas,[2] and Jesús M Hernández-Rivas[1,3,13]

[1]Unidad de Diagnóstico Molecular y Celular del Cáncer, Centro de Investigación del Cáncer-IBMCC (USAL-CSIC), Salamanca, Spain
[2]Bioinformática y Genómica Funcional, IBMCC-Centro de Investigación del Cáncer (USAL/CSIC/IBSAL), Salamanca, Spain
[3]Servicio de Hematología, Hospital Universitario de Salamanca, Spain
[4]IBIOMED, Instituto de Biomedicina, Universidad de León, Spain
[5]Servicio de Hematología, Hospital Universitario de León, Spain
[6]Servicio de Hematología, Hospital del Bierzo, Ponferrada, Spain
[7]Servicio de Hematología, Hospital Nuestra Señora de Sonsoles, Ávila, Spain
[8]Servicio de Hematología, Hospital Jerez de la Frontera, Spain
[9]Servicio de Hematología, Hospital General de Segovia, Spain
[10]Servicio de Hematología, Hospital Clínico Universitario de Valladolid, Spain
[11]Servicio de Hematología, Hospital Universitari Germans Trias i Pujol, Badalona, Spain
[12]Centre for Cancer Research and Cell Biology, Queen's University Belfast, Belfast, UK
[13]IBSAL, Instituto de Investigación Biomédica de Salamanca, Spain
*These authors contributed equally to this work

Myelodysplastic syndromes (MDS) are a group of hematopoietic stem cell disorders at high risk of developing acute myeloid leukemia (AML). In order to gain insights into the molecular mechanism underlying the evolution of MDS to more aggressive stages, a multi-platform genome-wide expression profiling (the Human Genome U133 Plus 2.0 and the Human Exon 1.0 ST microarrays (Affymetrix)) was carried out in a series of 73 patients with normal cytogenetics and 17 controls. Expression profiling revealed that several genes and gene pathways were commonly and progressively deregulated in the transition from non-malignant bone marrow conditions through early and advanced MDS to AML. Several genes began to be up-regulated in early MDS and continued that trend towards AML (pattern1, MDS/AML-up, (*NPM1*, *MYST1*, *RPL22*, *RPS6*)), while other genes showed the biggest change in their expression level in the transition from advanced MDS to AML. (pattern3, AML-up, *HOXA9*, *MEIS1*, *FLT3*)). By contrast, other genes were progressively down-regulated during the evolution of the disease, showing the minimum levels in AML (pattern 2 and 4, MDS/AML- and AML-down, respectively, *CEACAM3*, *CRISP3*, *CAMP*, *MMP9*). Pattern 1 was related to DNA damage response and ribosome. Pattern 3 to cell proliferation and differentiation arrest, while patterns 2 and 4 with immune response. Furthermore, two transcription factors (ATF2, TAF7) were identified as controlling a large number of these dynamically deregulated genes. The present study demonstrated the presence of a progressive deregulation of several cellular functions, with common deregulated genes, in the transition from non-malignant bone marrow through early and advanced MDS to AML.

Keywords: Gene expression profiling, myelodysplastic syndromes, acute myeloid leukemia, progression

INTRODUCTION

Myelodysplastic syndromes (MDS) represents a heterogeneous group of clonal myeloid stem cell disorders characterized by abnormal differentiation and maturation of myeloid cells, bone marrow (BM) failure, and a genetic instability with an enhanced risk to transform to acute myeloid leukemia (AML).[1-3] Both hematological malignancies, MDS and AML, are characterized by high biological and clinical heterogeneity due to their complex pathophysiology, which is based on the accumulation of a wide variety of genetic and epigenetic aberrations in hematopoietic progenitor cells, resulting in altered cell growth and differentiation.[4]

Genome-wide expression profiling (GEP) based on high-density microarray platforms have been largely used in the study of transcriptional pathways involved in cancer. These studies proved to be a powerful tool to further characterize gene expression signatures associated with distinct cancer subtypes and helped to better classify these diseases, to stratify patients into different risk classes and to identify deregulated genes implicated in cancer development, including acute and chronic leukemias, and MDS.[2,5-8] It is worth to mention that, unlike for AML, 50% of MDS samples could not accurately be classified using specific markers selected from expression data.[2] However, gene expression analysis distinguished clinical MDS subgroups with prognostic relevance, in terms of time to AML transformation, and classified MDS with a more aggressive disease and those with a more indolent clinical course as different disease entities.[2,8] The analysis of the genes discriminating these prognostic MDS entities revealed several pathways with genes known to be actively involved in AML, including *HOXA9*, *FLT3*, *KIT*, and *WT1*.[2,8]

In contrast to the wealth of global expression and transcriptomic data for most leukemia and lymphoma entities, the MDS have not been yet extensively analyzed. However, several studies comparing gene expression signatures between MDS, healthy individuals, and AML have been reported providing valuable insights into the genes and biological pathways implicated in these disorders.[1-4,9-14] Most of these studies were focused in the analysis of enriched CD34+ cell populations, mainly in MDS with del(5q). Interestingly, these studies revealed that distinct gene expression signatures were associated with specific MDS subtypes and also with different cytogenetic abnormalities, since patients showing del(5q), -7/del(7q), and trisomy 8 have a characteristically different expression profile.[1,4,10,11] Thus, patients with RARS were characterized by deregulation of mitochondrial and iron-related genes.[1,14] Moreover, CD34+ cells from MDS patients with del(5q) had a decreased expression of multiple ribosomal genes including *RPS14*, contained in the deleted region.[12,13] In addition, patients with -7/del(7q) were characterrized by an activation of the JNK pathway, while patients with trisomy 8, showed a deregulation of immune response and leukocyte activation.[10,11] Overall, transcriptomic studies along recent years have revealed that early MDS (low-risk MDS) were characterized by a deregulation of the immune response and an increased apoptosis in the bone marrow, with an elevated ratio of apoptosis/proliferation, but the mechanisms for this finding are not yet established. In contrast, advanced MDS (high-risk MDS) showed a decreased apoptosis, as well as deregulation of DNA damage response and checkpoint pathways.[3,4,9,10,15] The deregulated pathways identified in MDS provide deeper insights into the molecular mechanisms underlying MDS and AML, and thereby have provided novel targets for treatment, and holds promise for further advances in predicting prognosis and response to therapy.

Thus, the use of current genome-wide expression platforms and large scale analyses have allowed to propose some gene signatures associated with patients with MDS. However, the molecular mechanisms underlying myelodysplastic pathogenesis and evolution, and its frequent progression to AML are not fully understood, and further investigation of the specific genes and pathways involved in the defective hematopoietic cell differentiation and abnormal clone expansion in those patients who undergo progression to AML is required. Therefore, we performed a multi-platform profiling of gene expression changes occurring from non-malignant marrow conditions through different stages of MDS and towards AML, frequently evolving from MDS. In this report, we identified common genes with a progressive up-/down-regulation in AML patients when compared to high-risk MDS, low-risk MDS, and non-leukemic controls.

MATERIALS AND METHODS

Samples collection

A total of 182 patients with myeloid malignancies (127 MDS and 55 AML) and 17 age-matched controls with non-malignant disorders (referred to as non-

leukemic controls, NoL) were assessed by GEP. Previously reported microarray-based GEP studies revealed that specific MDS subtypes and patients with cytogenetic abnormalities (i.e. del(5q), -7/del (7q) or trisomy 8) had distinct gene expression profiles and deregulated gene pathways.[1,10,11] For this reason MDS patients with ringed sideroblasts, as well as those patients with chromosomal rearrangements were excluded. Thus, the final cohort of the study (n=90) only comprised patients with normal cytogenetics. All MDS and AML samples were obtained from untreated patients at the time of diagnosis, while five AML samples were secondary AML following MDS (sAML). After this appropriate selection of samples, MDS were classified according to the 2008 World Health Organization (WHO) criteria,[16] and the following morphological subtypes were included: refractory cytopenia with unilineage dysplasia (RCUD, n=11), refractory cytopenia with multilineage dysplasia (RCMD, n=23), refractory anemia with excess blasts type 1 (RAEB-1, n=9), and RAEB-2 (n=10). Because of the low number of subjects within each MDS subtype and according to the WHO classification, patients with RCUD and RCMD were grouped together and considered as low-risk MDS (LR-MDS, n=34), while RAEB-1 and RAEB-2 were grouped as high-risk MDS (HR-MDS, n=19), based on the percentage of bone marrow blasts.

All cases that entered the study were clinically well characterized and bone marrow morphology was carefully reviewed before inclusion (Supplementary Table S1). Patient age ranged from 48 to 88 years (median 76) and the female/male distribution was 38/52.

All samples were collected with written informed consent for research purposes according to the guidelines of the Declaration of Helsinki and the appropriate Local Ethical Committees approvals for all patients included in the study.

Gene Expression analytic platforms

Gene expression profiling (GEP) studies were performed in three series of patients on two different analytic platforms (see Supplementary Figure S1A): the Human Genome U133 Plus 2.0 microarray (Affymetrix, Santa Clara, CA), and the Human Exon 1.0 ST microarray (Affymetrix). The last comprised 36 samples (10 low-risk MDS, 10 high-risk MDS, 5 AML, 5 sAML and 6 controls) and was considered the "main group of expression data". In the remaining two series, the Human Genome U133 Plus 2.0 microarray (Affymetrix) was used as part of the Microarray Innovations in LEukemia (MILE) Study.[7] These two last series were produced at different times and following different molecular biology protocols, and merged into a single series as explained below (see section *Bioinformatic analysis: Data preprocessing*). After merging, it comprised 54 samples (24 low-risk MDS, 9 high-risk MDS, 10 AML and 11 controls) and was considered the "secondary group of expression data".

Gene expression microarray preparation

Total RNA was isolated from bone marrow mononuclear cells according to the MILE Study protocol.[7] The Human Genome U133 Plus 2.0 microarray and the Human Exon 1.0 ST microarray (Affymetrix) were labeled and hybridized according to the MILE Study described methods[7] and following the manufacturer's protocols for the GeneChip platform by *Affymetrix,*[17] respectively. Briefly, methods included first- and second-strand cDNA synthesis, double-stranded cDNA purification, cRNA synthesis and biotin-labeling by in vitro transcription, recovery and quantification of labeled cRNA, cRNA fragmentation and subsequent hybridization onto the microarray, post-hybridization washings, staining using a streptavidin-coupled fluorescent dye, and scanning. The following Affymetrix GeneChip instruments were used: Hybridization Oven 640, Fluidics Station FS450, and Scanner GCS3000.

Bioinformatic analysis

Data preprocessing

The main series, on GeneChips Human Exon 1.0 ST, was preprocessed (background correction, normalization, and summarization) using *RMA* [18] with the *gene mapper* chip definition file (CDF) from *GATExplorer,*[19] which allows mapping the unambiguous probesets directly to Ensembl gene loci and removal of cross-hybridization noise (Ensembl Version 57, March 2010, assembly NCBI36). Three outlier samples were detected with *arrayQualityMetrics* [20] and removed.

In order to merge the other two series, on the microarray platform Human Genome U133 Plus 2.0, each dataset was normalized individually using the algorithm *Frozen RMA* (*fRMA*)[21] with the *gene mapper* CDF from *GATExplorer*.[19] To calculate the *fRMA* vectors for this CDF, we used a collection of 1,335 human samples from 163 GEO datasets, grouped into 267 batches of 5 samples each (being these batches based on study and tissue, as recommended in the paper). Once the two datasets

were preprocessed independently with *fRMA*, they were merged into a single dataset with *InSilicoMerging*,[22] using *Combat* [23] to remove the remaining batch effect.

Detection of genes with increasing/decreasing expression trends

To identify the genes with an increasing/decreasing trend along the stages of the disease we calculated the *Goodman and Kruskal's Gamma* correlation,[24] implemented in *R package Rococo*,[25] between the gene expression and the categorical value representing the stage of the disease, sorted as non-leukemic samples (NoL), low-risk MDS (LR-MDS), high-risk MDS (HR-MDS), and AML. The disease stage was defined based on MDS risk classification (4 stage contrast): NoL→ LR-MDS → HR-MDS → AML. In this way, a correlation value (Gamma value) was obtained for each gene. The genes with a significant Gamma (absolute gamma >0.50, and FDR adjusted *P* value <0.05) were selected as correlated with the progression of the disease. This correlation study provided 2 gene lists (see Supplementary Figure S1B):

- The Full List of 1163 genes correlated with the progression of the disease in the "main group of expression data" (Human Exon 1.0 microarray).
- The Core List of 266 genes, which was a subset of the genes correlated in the main series, which was also confirmed in the "secondary group of expression data" (Human Genome U133Plus2.0 microarray). This list is the intersection of genes correlated with the progression of the disease in Series 1+2, and the main Series (*P* value of the intersection <0.00001).

For subsequent analyses we only considered the Core List of 266 genes. Furthermore the Full List of 1163 genes was only used in the analysis of the transcription factors potentially associated with the evolution of MDS. Both lists are available in the Supplementary Table S2.

Detection of patterns within the genes with increasing/ decreasing trends

To identify groups of genes with similar expression patterns we used a clustering based on Self Organizing Maps (*SOM*),[26] a robust unsupervised clustering and dimensionality reduction method that allows to search for common expression patterns among groups of genes. For this purpose, the expression of each gene was normalized by subtracting its mean and dividing by the standard deviation. Then, for each gene, the expression values were sorted in ascending or descending order (ascending if the mean expression in AML samples was higher than in NoL, and descending otherwise). With this normalized and sorted gene expression data, the genes were clustered with the *SOM* implementation in *Kohonen R package* [27] using a 3x3 grid with rectangular topology to allocate up to nine possible clusters (see Supplementary Figure S1C). Within the nine groups provided by the clustering analysis, three groups were not assigned any genes, and four major patterns included a significant number of genes. To obtain the final assignment of a gene to one pattern, it was required that the gene was assigned consistently to the same pattern in both groups of the study (main and secondary groups), or to one of the main patterns in one of the groups and to the related intermediate pattern in the other group.

Functional analysis of the genes included in the patterns

A standard singular Functional Enrichment Analysis (FEA) approach, without additional filtering or clustering, was applied to the gene lists from each pattern. For each of the gene lists, and independent analysis was performed through RDAVIDWebService using DAVID's Functional Annotation Chart with the main annotations related to pathways and biological processes (GO-BP, KEGG, REACTOME, PANTHER and BIOCARTA).

Transcription factors associated to the gene patterns

We searched for enriched transcription factor binding sites (TFBS) and chip-seq tracks in the promoter regions of the genes in the Full List of 1163 genes using *iRegulon*, [28] a Cytoscape plugin that includes a method to detect enriched TFBS and their direct targets. For each pattern we splitted the genes into up-regulated and down-regulated. Then, we searched for enriched TFBS in the close promoter region (up to 500bp upstream) and also at ±10kb from the translation start site of each genes.

RESULTS

Gene expression changes correlated with MDS evolution

To find the genes whose expression may be associated to the evolution of the disease, we searched for genes whose expression levels evolved following and increasing or decreasing trend through

the different stages of MDS, with the non-leukemic samples as reference of the origin stage, and the AML samples as reference of the malignant stage they are heading to. Thus, we calculated the correlation between the expression level of each gene and the disease stage, being grouped by risk level as non-leukemic controls, LR-MDS, HR-MDS, and AML.

The results allowed us the identification of 4 major patterns with a significant number of genes: pattern 1, MDS/AML-up signature; pattern 2, MDS/AML-down signature; pattern 3, AML-up signature; and pattern 4, AML-down signature. Pattern 1 described a group of genes with progressively increasing expression levels in low-risk MDS, high-risk MDS patients, with the maximum in AML samples. Genes within this pattern showed the greatest increase in their expression level (>50% of the expression range) in low-risk MDS patients (Figure 1A, Table 1). Pattern 2 described a group of genes whose expression levels were progressively and significantly lower in low-risk MDS patients when compared to non-leukemic controls, and that these levels reduced again in high-risk MDS patients, reaching minimum in AML samples. Only in pattern 2, the genes showed the greatest decrease in their expression level in low-risk MDS patients (Figure 1B, Table 1). Pattern 3 describes a group of genes that were significantly up-regulated in AML patients when compared to MDS (patients with LR and HR-MDS) and non-leukemic controls. It should be noted that these genes showed a highly attenuated but continued up-regulation in high-risk MDS *vs* low-risk MDS, and in low-risk MDS *vs* non-leukemic controls. It was only the AML patients that showed a greater than 50% change in the expression range (Figure 1C, Table 1). Finally, pattern 4 described a group of genes with significantly lower expression in AML patients than in MDS (patients with LR and HR-MDS) and non-leukemic controls (Figure 1D, Table 1).

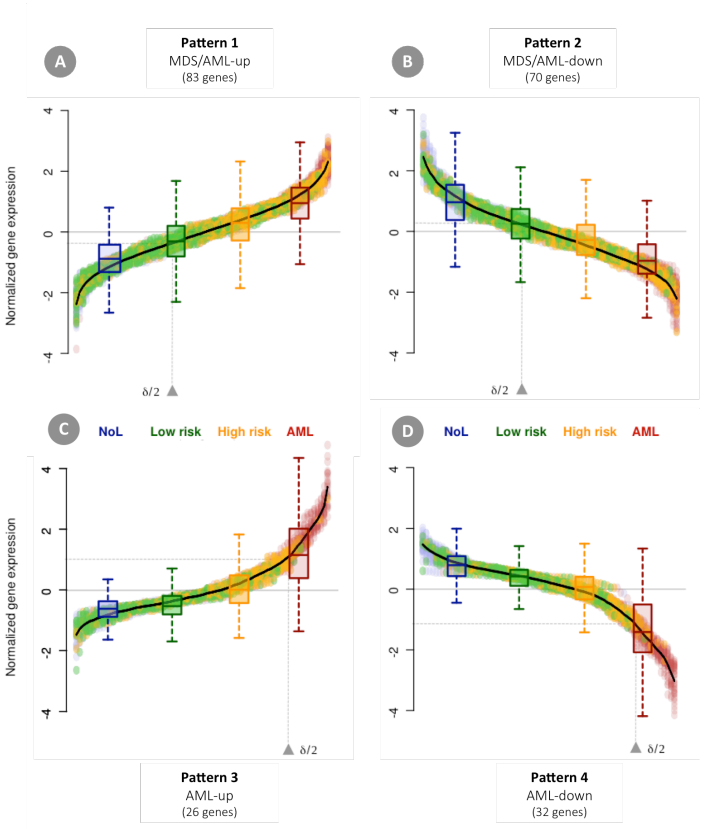


Figure 1: **Expression patterns found in the progression of MDS**. Boxplots represent the normalized expression level of the genes identified within each pattern from unified datasets grouped by disease category, including non-malignant samples, low-risk MDS, high-risk MDS and AML. In the background, gene expression values used for the *SOM* clustering (see methods): each dot represents the expression of one gene in one sample (total number of samples n=90). For each gene, samples were sorted on increasing or decreasing expression values. The color indicates the disease category (blue = Non-malignant; green = LR-MDS; orange = HR-MDS; and red = AML). The black line in the background along the expression value represents the average expression of the genes. The *delta/2* value (δ/2, horizontal grey dotted line) indicates the 50% change in the expression range (i.e. the half height between the minimum and maximum expression) and the stage of the disease in which it falls (vertical grey dotted line). Patterns 1 (**A**) and 2 (**B**) had the δ/2 value in the LR-MDS stage with a more linear behavior, while patterns 3 (**C**) and 4 (**D**) had the δ/2 in the AML stage.

Table 1. Genes and cellular functions deregulated in each pattern

	Gene Symbol				Gene Symbol				Gene Symbol			Gene Symbol
Pattern 1 (MDS/AML-up)	*AC010619.2*	*C20orf199*	*RAB37*	Pattern 2 (MDS/AML-down)	*CLEC4E*	*KCNJ15*	*SYNE2*	Pattern 3 (AML-up)	*HOXA3*	Pattern 4 (AML-down)	*CRISP3*
	RP11-209A2.1	*C22orf28*	*RAG1AP1*		*SIRPA*	*KLRK1*	*TMEM154*		*HOXA7*		*LTF*
	RPL12	*C4orf14*	*SLC25A36*		*SIRPB1*	*LILRA3*	*TMEM2*		*HOXA9*		*CAMP*
	RPL13	*C5orf26*	*SLC25A5*		*CEACAM21*	*LRRC4*	*TRPM6*		*HOXB2*		*CLEC4D*
	RPL13A	*C6orf48*	*SNHG8*		*CEACAM3*	*MANSC1*	*UCA1*		*HOXB3*		*S100A12*
	RPL18	*CAMLG*	*TFAM*		*MMP25*	*MAPK1*	*ZDHHC3*		*MEIS1*		*S100A9*
	RPL22	*CCNB1IP1*	*TFPI*		*MMP27*	*MAST3*			*FLT3*		*PGLYRP1*
	RPL28	*CFD*	*TMEM147*		*AC107883.1*	*NAV2*			*ADA*		*CEACAM6*
	RPL3	*CMTM3*	*TMEM159*		*AKTIP*	*NELL2*			*ANGPT1*		*CEACAM8*
	RPS5	*DCTD*	*TOMM20*		*AL034374.4*	*NFAM1*			*C12orf76*		*MMP8*
	RPS6	*DDAH2*	*TOP1MT*		*ARHGAP24*	*NLRC4*			*CALCOCO2*		*MMP9*
	MRPS12	*DDOST*	*TSC22D1*		*ARSG*	*NUDT19*			*CEP70*		*ANXA3*
	EEF2	*DHPS*	*TTC19*		*ATP2C2*	*ORM1*			*CLIP2*		*ARG1*
	EIF3K	*FLAD1*	*TWSG1*		*BMX*	*ORM2*			*COL24A1*		*C5orf32*
	PSMA1	*FPGS*	*UBE2E1*		*C20orf177*	*PADI2*			*EGFL7*		*CD24L4*
	PUF60	*GAMT*	*UFM1*		*CA4*	*PFKFB2*			*EIF3G*		*CDA*
	DHX30	*GLTSCR2*	*USE1*		*CABLES2*	*PLSCR1*			*FAM116B*		*CHI3L1*
	MYST1	*GNA15*	*ZBTB10*		*CHIT1*	*PLXNC1*			*GPR114*		*CKAP4*
	H3F3B	*GNB2L1*	*ZNF618*		*CLTCL1*	*PPEF1*			*ISYNA1*		*ELOVL5*
	HIST1H2BD	*GPX1*			*CRISP2*	*PRKAA1*			*KIAA0125*		*GCA*
	NPM1	*GTF2F1*			*CSGALNACT1*	*PTPRJ*			*MSI2*		*GLT1D1*
	RASSF1	*HINT1*			*DAPK2*	*RAB11FIP4*			*PPFIBP1*		*HP*
	RCC1	*HSP90AB1*			*DNAH10*	*RASGRP1*			*RBMS2*		*IQSEC1*
	CCNG1	*IMPDH2*			*E2F8*	*RETN*			*STAR*		*ITGAM*
	PDCD7	*LDHB*			*F5*	*ROGDI*			*TMTC4*		*LCN2*
	PMAIP1	*LETMD1*			*FBXW2*	*S100A11*			*TWISTNB*		*MGAM*
	APRT	*LRP5L*			*FGD4*	*SFXN1*					*PLBD1*
	ATP1B1	*LYSMD2*			*GLRX*	*SLC16A14*					*QPCT*
	ATP5G2	*NT5C*			*GPR155*	*SLC25A21*					*RAB31*
	BMP1	*PDLIM1*			*HIRA*	*SNTB2*					*STOM*
	BZW2	*PEBP1*			*HSPA1L*	*SVIL*					*TCN1*
	C12orf10	*PPTC7*			*IPCEF1*	*SYNE1*					*WIPI1*

The most representative cellular functions deregulated in each pattern			
p1 (MDS/AML-up)	• Ribosomal Proteins and translation pathways • DNA damage response and checkpoint, cell cycle • Apoptosis • Nucleosome and chromatin assembly	**p3 (AML-up)**	• Induction of cell proliferation • Repression of cell differentiation
p2 (MDS/AML-down)	• Immune response • Cell adhesion • Matrix metalloproteinases	**p4 (AML-down)**	• Growth inhibition • Immune response • Cell adhesion • Matrix metalloproteinases

Analysis of the deregulated genes included in each pattern of MDS evolution

As mentioned before, we searched for those genes whose expression levels followed a continued up- or down-regulation as the disease evolved. The, pattern 1, defined as MDS/AML-up signature, consisted of 83 genes, pattern 2, MDS/AML-down signature, contained 70 genes, whereas pattern 3 and 4 (AML-up and AML-down signatures) consisted of 26 and 32 genes, respectively (Figure1, Table 1). A Functional Enrichment Analysis was performed for the dynamically deregulated genes within each pattern in order to get the most representative cellular functions corresponding to each pattern, and to get some insights into the biological functions that were consistently deregulated as the disease progressed.

According to this, pattern 1 (MDS/AML-up signature) was characterized by the progressive up-regulation of the DNA damage response and checkpoint pathways (*NPM1*, *RASSF1*, *RCC1*, *CCNG1*), apoptosis (*PDCP7*, *PMAIP1*), as well as ribosome and translation pathways (*RPL12*, *RPL13*, *RPL13A*, *RPL18*, *RPL22*, *RPL28*, *RPL3*, *RPS5*, *RPS6*, *MRPS12*, *EEF2*, *EIF3K*, *PUF60*, *DHX30*) (Table 1). It should be noted that a large number of these progressively over-expressed genes were ribosomal protein-coding genes. Other biological functions enriched within pattern 1 were nucleosome and chromatin assembly (*MYST1*, *H3F3B*, *HIST1H2BD*), where some genes coding for histones were found up-regulated. By contrast, the functional enrichment analysis of the genes in Pattern 2 (MDS/AML-down signature) identified genes related to immune response (*CLEC4E*, *SIRPB1*, *SIRPA*), matrix metalloproteinases (*MMP25*, *MMP27*), and cell adhesion (*CEACAM3*) (Table 1).

The genes included in the Pattern 3 (AML-up signature) and Pattern 4 (AML-down signature) were mainly involved in cell differentiation, and cellular

growth and proliferation (Table 1). Pattern 3 contained several pro-proliferative genes, such as *HOX* genes (*HOXA3*, *HOXA7*, *HOXA9*, *HOXB2*, *HOXB3*), *MEIS1* and *FLT3*, while pattern 4, had some anti-proliferative genes such as *CRISP3*, all of them known to be actively involved in AML. Furthermore, *HOX* genes were also related to cell differentiation, as they act as repressors of this process. In addition, some genes related to immune response (*LTF*, *CAMP*, *CLEC4D*, *S100A12*, *PGLYRP1*), matrix metallo-proteinases (*MMP8*, *MMP9*), and cell adhesion (*CEACAM8*, *CEACAM6*) were also observed in pattern 4 (Table 1). Examples from each pattern are shown in Supplementary Figure S2.

Transcription factors potentially associated with the gene expression patterns in the evolution of MDS

Interestingly, the analysis of the genes included within each pattern revealed that some of them were transcription factors (TFs). Due to the genes studied had an increasing or decreasing trend in their expression levels, we decided to explore whether any deregulated TF could potentially be leading the expression changes we observed. For this reason we further examined the putative transcription factor binding sites (TFBS), and their probable regulators, within the promoter regions of the genes in the Full List of 1163 genes. This analysis provided a total of 660 TFs, of which 42 were included in the list of deregulated genes (Full List) (see Supplementary Table S3). We then focused on exploring which genes assigned to each pattern (from the Core List) were potentially regulated by any of these 42 TFs.

Intriguingly, only 7 TFs (TAF7, ATF2, XBP1, ETS2, NR1H3, TGIF2, CBFB) appeared to be regulating many of the genes from pattern 2, MDS/AML-up signature (Figure 2A). By contrast, the 26 genes in pattern 1 have TFBS for many TFs, some of them within the same gene list, which leads to a very connected network (Figure 2B). TAF7 and ATF2 had 34 and 35 potential targets within pattern 2, respectively, many of them being ribosomal protein-coding genes. Looking deeper into these two TFs, *TAF7* was found increasingly up-regulated during disease evolution, since early MDS stages, while *ATF2* was found continuously up-regulated mainly in the transition from advanced MDS to AML.

DISCUSSION

Myelodysplastic syndromes (MDS) are considered as a pre-leukemic state that frequently progresses to acute myeloid leukemia (AML). The presence of common molecular mechanisms affected in both diseases has been described. In fact both, MDS and AML, show mutations in common genes, such as *RUNX1*, *TET2*, *FLT3*, *DNMT3A* and others.[29-33] At transcriptional level MDS patients show a well-defined gene expression signature, mainly high-risk MDS and MDS associated with cytogenetic abnormalities in which a gene dosage effect has been reported.[1,2,4,9-14,34] Thus, the present study was restricted to MDS with normal cytogenetics and was focused on the identification of common and progressively deregulated genes, which may potentially be associated with the progression of the disease, by a multi-platform genome-wide expression profiling. This methodological approach allowed us to identify a set of genes that began to be deregulated (up-regulated/down-regulated) in early MDS stages (LR-MDS) and remained progressively altered during the progression towards advanced MDS (HR-MDS) and AML, while the gene expression level of other genes changed drastically in more advanced disease stages (HR-MDS/AML).

These common and dynamically deregulated genes implicated in the progression of the disease were related to key cellular functions of known relevance in MDS. Thus, apoptosis, DNA damage response and checkpoint pathways, ribosome and translation pathways (mainly enriched in pattern 1, MDS-up), immune response (enriched in patterns 2 and 4, MDS- and MDS/AML-down), cell differentiation, and cellular growth and proliferation (mainly enriched in pattern 3, AML-up), were progressively deregulated during the evolution of MDS. It should be noted that nucleosome and chromatin assembly were additional significant functions that were also altered.

An important group of genes that participated in the evolution of the disease was related to the DNA damage response and checkpoint pathways. DNA damage facilitates the activation of the responses to DNA damage stimuli and checkpoint pathways, in order to slow down cell cycle progression and allow cells to repair of damaged DNA.[10,35,36] The efficient repair of damaged DNA is crucial for the maintenance of genome integrity and cell survival, preventing the propagation of cells containing genomic abnormalities.[35,36] If the DNA damage is too extensive, these mechanisms can trigger apoptosis or cell senescence (via activation of p53) in order to eliminate defective cells.[10,35-37] It is well recognized that advanced MDS are characterized by high levels of genomic instability, due to defective function of the

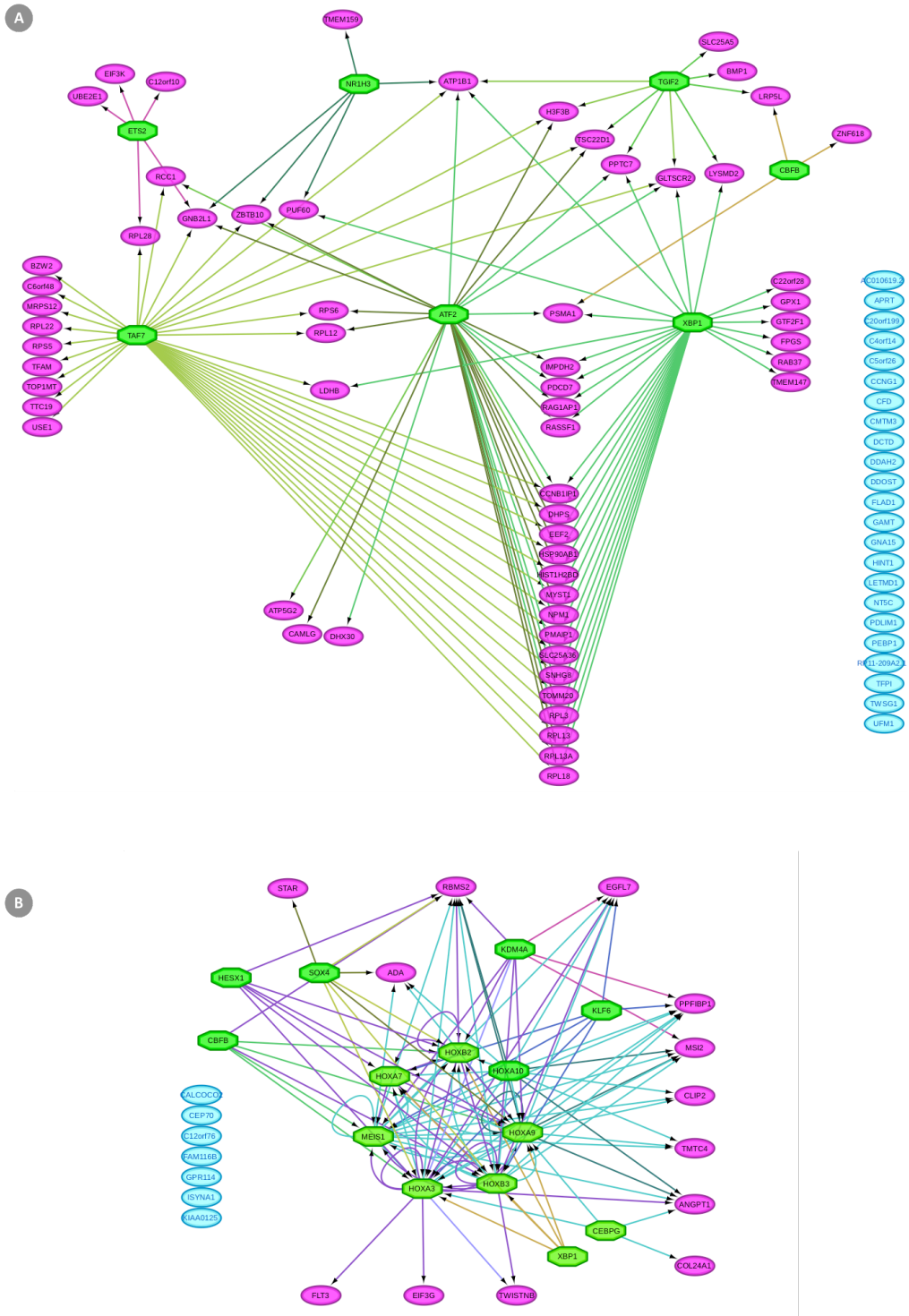


Figure 2: **More relevant transcription factors and their targets involved within pattern 1, MDS/AML-up** (**A**)**, and pattern 3, AML-up** (**B**). Pink and blue nodes represent the up-regulated genes within each pattern, and green represents the up-regulated transcription factors from the Full List. An arrow pointing from A to B indicates that A potentially regulates B.

DNA repair mechanisms.[10,36] Interestingly, in our study, those gene pathways related to the DNA damage response and checkpoints already appeared up-regulated in early MDS, suggesting that in early MDS these mechanisms are fully active, maybe to compensate for a damage generated in the DNA sequence. Moreover, and somehow expected, these genes involved in DNA damage response and checkpoints pathways were consistently up-regulated in HR-MDS with the maximum expression levels in AML patients. Some of them are implicated in chromatin assembly such as *MYST1*, *H3F3B*, *HIST1H2BD*, while others are related to DNA damage response and cell cycle arrest (*NPM1*, *RCC1*, *RASSF1*, and *CCNG1*). *MYST1* gene belongs to a family of lysine acetyltransferases with important roles in gene regulation and DNA damage response.[38] It has been reported that this protein is required for efficient repair of DNA damage through recruitment of other DNA repair mediator proteins such as Mdc1, 53bp1, and Brca1 to the sites of DNA damage.[38] Of note, less studied cellular functions were also altered in the transition from non-malignant BM through MDS to AML, involving nucleosome and chromatin assembly. Thus, *MYST1* also plays important roles in transcription regulation through the acetylation of the histone H4, by regulating higher-order chromatin structures, and also the transcription factor p53.[38,39] Overall, histones play a central role in transcription regulation, DNA repair, DNA replication and chromosomal stability. *NPM1* gene encodes a moonlighting protein that is involved in several cellular processes such as mediating ribosomal and histone (H3, H2B and H4) assembly, centrosome duplication, participating in DNA repair and regulating stability of tumor suppressors like p53. Previous observations have demonstrated that is consistently overexpressed in malignant tissues or cell lines in comparison to the benign state.[40] Mutations in this gene are associated with AML. *RCC1* (Regulator Of Chromosome Condensation 1) encodes a protein involved in the regulation of the cell cycle by detecting unreplicated DNA and inhibiting the progression through it. During the cell cycle, nuclear DNA replication (S phase) and mitosis (M phase) are linked such that replication of DNA must be complete before mitosis can begin. Thus, an active system for detecting unreplicated DNA and transducing an inhibitory signal to prevent the activation of mitotic factors could be present in the cells even at early stages of MDS. The RCC1 protein has proven to be involved in this regulatory process. RCC1 defects resulted in abnormal chromosome segregation and genomic instability, which are characteristic of many cancer cells.[41,42] *RASSF1* encodes a protein that mediates the apoptotic response to Ras activation and also plays an important role in the DNA damage response as it has been shown to inhibit the accumulation of cyclin D1, inducing cell cycle arrest,[43] playing a critical role in the DNA repair itself by interacting with the DNA repair protein XPA, essential for the nucleotide excision repair mechanism.[44] *CCNG1* encodes cyclin G1, which is an important cell cycle regulator, and is associated with cell cycle arrest in response to DNA damage.[45] The overexpression of these genes may indicate that even in the earliest stages of the disease the cells show a defensive reaction against DNA damage. However, as the disease progresses, these mechanisms are still fully active, but they are not enough stronger to efficiently repair the damaged DNA, resulting in an increased genomic instability and conferring the cell a growth advantage over its normal counterparts, leading to the progression to AML.

One of the hallmarks of MDS is the increased apoptosis, which is widely known as deregulated in MDS, especially in early MDS stages, and considered to contribute to the characteristic ineffective blood cell production in MDS, which results in peripheral blood cytopenias. In this study we found over-expression of genes involved in apoptosis-related pathways in early MDS patients when compared to non-leukemic controls, such as *PDCD7* and *PMAIP1* was observed. These findings are consistent with prior studies in which bone marrow of LR-MDS is characterized by high level of apoptosis.[3] However, in HR-MDS and AML patients the expression levels of these apoptosis-related genes showed an attenuated but continued up-regulation when compared to those of LR-MDS (Figure 1). *PDCD7* (Programmed Cell Death 7) encodes a protein that interfere the splicing of mRNA precursors, in association with the U11 small nuclear ribonucleoprotein (snRNP). This protein have been shown to promote apoptosis when overexpressed.[46] In addition, *PDCD7* overexpression has been associated with shorter overall survival (OS) and relapse-free survival (RFS) in a subset of AML patients with normal karyotype, and thereby, with more aggressive forms and poor prognosis of AML.[47] *PMAIP1* gene, also known as *NOXA*, is a pro-apoptotic member of the Bcl-2 family, and is involved in inducing p53-mediated apoptosis after DNA damage in tumor cells.[48] It has also been shown that

this protein binds specifically Mcl-1 (Myeloid Cell Leukemia 1), and displaces Mcl-1/Bak1 interaction, resulting in the release of Bak1 (BCL2-Antagonist/Killer 1) which leads to mitochondrial membrane permeabilization, efflux of apoptogenic proteins from the mitochondria, and finally the activation of caspases and apoptosis.[49]

We have observed a consistent up-regulation of genes coding for ribosomal proteins (RPs) in MDS and AML, demonstrating that a substantial number of RPs-coding genes were overexpressed in early MDS and continued to be up-regulated in advanced MDS and AML. Several studies have shown that deregulated ribosome biogenesis leads to the release of RPs, triggering the so-called nucleolar stress. These free RPs could lead to the activation of p53 (via MDM2) to induce cell cycle arrest, senescence, or apoptosis, as well as transcription/translation, mRNA processing, DNA repair, and tumorigenesis.[50-52] In our study, an up-regulation of several RPS-coding genes, as well as other translation-related genes in MDS and AML displaying normal karyotypes was observed. Of note, RPS6 has been reported as a downstream effector of the mTOR pathway, and was shown to contribute to cell growth and proliferation through the translation of specific mRNAs.[53,54] In addition, RPL22 has been shown to be associated with histone H1 and co-localize on condensed chromatin. Its overexpression results in gene transcription inhibition, while reduction of RPL22 results in transcriptional up-regulation of genes.[50] Altogether, these findings and previous reported data in several other neoplastic conditions may indicate that RPs could play a fundamental role in disease progression and aggressiveness.[9,50,51]

Two important cellular functions mainly deregulated in the transition from advanced MDS to AML are cellular differentiation and proliferation. We observed an up-regulation of genes promoting proliferation while suppressing differentiation (*HOX* genes, *FLT3*), and down-regulation of genes that inhibit cellular proliferation (*LTF*, *CRISP3*). One of the most up-regulated genes were the homeodomain containing genes, *HOX* gene family and *MEIS1*, whose role in leukemogenesis has been convincingly demonstrated.[5,55-57] In addition to their role in establishing body plan during development, *HOX* genes have been implicated in several biological processes, including cell migration, proliferation and differentiation, as well as hematopoietic stem cell self-renewal.[58-60] *HOXA9*, in particular, and its cofactor *MEIS1* have been extensively studied in AML, since these genes are frequently found up-regulated in acute leukemias.[5,55-58,60] In addition, the aberrant expression of *HOXA9* has been associated with poor prognosis in AML.[55,56]. Our results are in accordance with previous reported data in which *HOXA9* and *MEIS1*, as well as other *HOXA* and *HOXB* cluster genes, are overexpressed in MDS at high risk of developing AML or in MDS who, indeed, evolved to AML.[2,8] Interestingly, *HOXA9* overexpression has been shown to promote hematopoietic progenitor cells self-renewal and expansion. In contrast, reduced expression has been associated with hematopoietic differentiation.[60] It should be noted that the up-regulation of *HOXA9* is linked to the up-regulation of *FLT3*, which is implicated in cell proliferation.[58,60] *HOXA7*, *HOXB3*, and *HOXB2* have also been found overexpressed in HR-MDS, AML and several solid tumors, also promoting cell proliferation.[2,8,55,56,61-64] Consistent with this, the up-regulation of *HOX* genes, mainly *HOXA9* in its cofactor *MEIS1*, will promote hematopoietic progenitor cells proliferation, while suppressing myeloid differentiation, leading to the accumulation of blasts in the BM and progression to AML. In addition, in the transition form advanced MDS to AML we observed a marked down-regulation of *CRISP3* and *LTF*, both implicated in immune response and growth inhibition.[65-67] LTF has been proposed as having an anti-tumorigenic function, through the regulation of natural killer (NK) cell activity, inhibition of cell proliferation and enhancement of apoptosis. Some studies have shown that *LTF* expression is down-regulated in lung, breast, prostate and gastric cancer, glioblastoma, and leukemia.[67]

The analysis of the putative transcription factor binding sites (TFBS) in the promoter regions of our deregulated genes allowed us to infer that two main transcription factors (TFs), TAF7 and ATF2, could potentially lead the expression changes that we observed. TAF7 was found increasingly up-regulated during disease evolution, and with many ribosome-related genes as potential targets. *TAF7* (TATA binding protein associated factor 7) is a component of the TFIID protein complex, and controls the first steps of transcription, reflecting that plays an essential role in cell proliferation.[68] Moreover, *ATF2* was found up-regulated mainly in the transition from advanced MDS to AML, with many RPs-coding genes as potential targets. *ATF2* (Activating Transcription Factor 2) gene encodes a moonlighting protein based on its ability to perform multiple cellular functions, such as transcriptional control under stress

conditions, chromatin remodeling, DNA damage response and cell death, depending on its subcellular localization.[69-72] Several studies have demonstrated that overexpression, phosphorylation and altered subcellular localization of ATF2, as well as interaction with oncogenic proteins (JUN), play a crucial role in oncogenic transformation and tumorigenesis. Thus, strong nuclear ATF2 localization has been observed in more aggressive cancers and associated with metastasis.[69-72] Further functional studies are warranted to discern the consequences of the deregulation of these two TFs. However, as they seem to be deregulating a large number of genes in our study, they open up the possibility for therapeutic intervention.

In summary, the present study demonstrated the presence of a progressive deregulation of several cellular functions, with common deregulated genes, in the transition from non-malignant bone marrow through early and advanced MDS to AML. The main cellular functions directly or indirectly implicated in disease evolution were DNA damage response and checkpoint pathways, "nucleolar stress" proteins, apoptosis, cellular proliferation and suppression of myeloid differentiation, as well as deregulation of immune responses. Furthermore, a large number of the dynamically deregulated genes identified in the present study seem to be regulated by two transcription factors, ATF2 and TAF7, not previously related to myeloid malignancies, which could be potential targets for therapeutic intervention. Therefore, the evolution from normal bone marrow trough MDS to AML seems to be an orchestrated mechanism that involves common deregulated functional pathways.

FUNDING

This work was partially supported by Grants from the Spanish Fondo de Investigaciones Sanitarias FIS (PI12/00281); Proyectos de Investigación del SACYL (BIO/SA47/13; BIO/SA52/14; GRS/874/A13; GRS 994/A/14); COST Action "EuGESMA" (BM0801); Red Temática de Investigación Cooperativa en Cáncer (RTICC), Instituto de Salud Carlos III (ISCIII), Spanish Ministry of Economy and Competitiveness & European Regional Development Fund (ERDF) "*Una manera de hacer Europa*" (Innocampus, CEI-2010-1-0010) (RD12/0036/0069; RD12/0036/0029; RD12/0036/0044); and the European Union Seventh Framework Programme [FP7/2007-2013] under Grant Agreement n°306242-NGS-PTL. This study was part of the international multi-institutional Microarray Innovations in LEukemia (MILE) study, a collaborative effort centered on the European LeukemiaNet (ELN, www.leukemia-net.org), and sponsored in part by Roche Molecular Systems Inc. MA was supported by a "*Junta para Ampliación de Estudios*" fellowship [09-02402] of the Spanish National Research Council (*Consejo Superior de Investigaciones Científicas*, CSIC), co-funded by the European Social Fund. SA was supported by a grant from the "*Consejería de Educación*" of the "*Junta de Castilla y León*" (JCyL) and the European Social Fund (ESF).

ACKNOWLEDGEMENTS

The authors would like to thank Irene Rodríguez, Sara González, Teresa Prieto, Mª Ángeles Ramos, Almudena Martín, Ana Díaz, Ana Simón, María del Pozo, Vanesa Gutiérrez and Sandra Pujante from the Centro de Investigación del Cáncer, Salamanca, Spain, for their technical assistance.

CONFLICTS-OF-INTEREST

The authors declare no competing financial interests.

SUPPLEMENTARY MATERIAL

Supplementary Table S1. Summary of the clinical characteristics of the patients included in the study.

Supplementary Table S2. Information of the 1163 deregulated genes from the Full List. The first 266 rows correspond to the genes included in the Core List (Available in the CD)

Supplementary Table S3. The 42 deregulated transcription factors (TFs) from the Full List (Available in the CD)

Supplementary Figure S1. Conceptual design.

Supplementary Figure S2. Example of expression patterns associated with the progression of the disease.

REFERENCES

1. Del Rey M, Benito R, Fontanillo C, et al. Deregulation of genes related to iron and mitochondrial metabolism in refractory anemia with ring sideroblasts. PLoS One. 2015;10(5):e0126555.
2. Mills KI, Kohlmann A, Williams PM, 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.
3. Del Rey M, O'Hagan K, Dellett M, et al. Genome-wide profiling of methylation identifies novel targets with aberrant hypermethylation and reduced expression in low-risk myelodysplastic syndromes. Leukemia. 2013;27(3):610-8.
4. Theilgaard-Mönch K, Boultwood J, Ferrari S, et al. Gene expression profiling in MDS and AML: Potential and future avenues. Leukemia. 2011;25(6):909-20.

5. Gutiérrez NC, López-Pérez R, Hernández JM, et al. Gene expression profile reveals deregulation of genes with relevant functions in the different subclasses of acute myeloid leukemia. Leukemia. 2005;19(3):402-9.
6. Quackenbush J. Microarray analysis and tumor classification. N Engl J Med. 2006;354(23):2463-72.
7. Haferlach T, Kohlmann A, Wieczorek L, et al. Clinical utility of microarray-based gene expression profiling in the diagnosis and subclassification of leukemia: Report from the international microarray innovations in leukemia study group. J Clin Oncol. 2010;28(15):2529-37.
8. Bresolin S, Trentin L, Zecca M, et al. Gene expression signatures of pediatric myelodysplastic syndromes are associated with risk of evolution into acute myeloid leukemia. Leukemia. 2012;26(7):1717-9.
9. Sridhar K, Ross DT, Tibshirani R, Butte AJ, Greenberg PL. Relationship of differential gene expression profiles in CD34+ myelodysplastic syndrome marrow cells to disease subtype and progression. Blood. 2009;114(23):4847-58.
10. Pellagatti A, Cazzola M, Giagounidis A, et al. Deregulated gene expression pathways in myelodysplastic syndrome hematopoietic stem cells. Leukemia. 2010, Apr;24(4):756-64.
11. Chen G, Zeng W, Miyazato A, et al. Distinctive gene expression profiles of CD34 cells from patients with myelodysplastic syndrome characterized by specific chromosomal abnormalities. Blood. 2004;104(13):4210-8.
12. Boultwood J, Pellagatti A, Cattan H, et al. Gene expression profiling of CD34+ cells in patients with the 5q- syndrome. Br J Haematol. 2007;139(4):578-89.
13. Pellagatti A, Hellström-Lindberg E, Giagounidis A, et al. Haploinsufficiency of RPS14 in 5q- syndrome is associated with deregulation of ribosomal-and translation-related genes. Br J Haematol. 2008;142(1):57-64.
14. Boultwood J, Pellagatti A, Nikpour M, et al. The role of the iron transporter ABCB7 in refractory anemia with ring sideroblasts. PLoS One. 2008;3(4):e1970.
15. Hofmann WK, de Vos S, Komor M, Hoelzer D, Wachsman W, Koeffler HP. Characterization of gene expression of CD34+ cells from normal and myelodysplastic bone marrow. Blood. 2002;100(10):3553-60.
16. Vardiman JW, Thiele J, Arber DA, 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.
17. Rodríguez AE, Hernández JÁ, Benito R, et al. Molecular characterization of chronic lymphocytic leukemia patients with a high number of losses in 13q14. PLoS One. 2012;7(11):e48485.
18. Irizarry RA, Bolstad BM, Collin F, Cope LM, Hobbs B, Speed TP. Summaries of affymetrix genechip probe level data. Nucleic Acids Res. 2003;31(4):e15-.
19. Risueño A, Fontanillo C, Dinger ME, De Las Rivas J. GATExplorer: Genomic and transcriptomic explorer; mapping expression probes to gene loci, transcripts, exons and ncrnas. BMC Bioinformatics. 2010;11:221.
20. Kauffmann A, Gentleman R, Huber W. ArrayQualityMetrics--a bioconductor package for quality assessment of microarray data. Bioinformatics. 2009;25(3):415-6.
21. McCall MN, Bolstad BM, Irizarry RA. Frozen robust multiarray analysis (fRMA). Biostatistics. 2010;11(2):242-53.
22. Taminau J, Meganck S, Lazar C, et al. Unlocking the potential of publicly available microarray data using insilicodb and insilicomerging R/bioconductor packages. BMC Bioinformatics. 2012;13:335.
23. Johnson WE, Li C, Rabinovic A. Adjusting batch effects in microarray expression data using empirical bayes methods. Biostatistics. 2007;8(1):118-27.
24. Goodman LA, Kruskal WH. Measures of association for cross classifications*. J Am Stat Assoc. 1954;49(268):732-64.
25. Bodenhofer U, Krone M. RoCoCo: An R package implementing a robust rank correlation coefficient and a corresponding test (2011). Software Available at http://www.bioinf.jku.at/software/rococo/.
26. Kohonen T. Self-organized formation of topologically correct feature maps. Biol Cybern. 1982;43(1):59-69.
27. Wehrens R, Buydens LM. Self-and super-organizing maps in R: The kohonen package. J Stat Softw. 2007;21(5):1-19.
28. Janky R, Verfaillie A, Imrichová H, et al. IRegulon: From a gene list to a gene regulatory network using large motif and track collections. PLoS Comput Biol. 2014;10(7):e1003731.
29. Rocquain J, Carbuccia N, Trouplin V, 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.
30. Bejar R, Stevenson K, Abdel-Wahab O, et al. Clinical effect of point mutations in myelodysplastic syndromes. N Engl J Med. 2011;364(26):2496-506.
31. Cazzola M, Della Porta MG, Malcovati L. The genetic basis of myelodysplasia and its clinical relevance. Blood. 2013;122(25):4021-34.
32. Papaemmanuil E, Gerstung M, Malcovati L, et al. Clinical and biological implications of driver mutations in myelodysplastic syndromes. Blood. 2013;122(22):3616-27; quiz 3699.
33. Haferlach T, Nagata Y, Grossmann V, et al. Landscape of genetic lesions in 944 patients with myelodysplastic syndromes. Leukemia. 2014;28(2):241-7.
34. Vasikova A, Belickova M, Budinska E, Cermak J. A distinct expression of various gene subsets in CD34+ cells from patients with early and advanced myelodysplastic syndrome. Leuk Res. 2010;34(12):1566-72.
35. Zhou T, Hasty P, Walter CA, Bishop AJ, Scott LM, Rebel VI. Myelodysplastic syndrome: An inability to appropriately respond to damaged DNA? Exp Hematol. 2013;41(8):665-74.
36. Zhou T, Chen P, Gu J, et al. Potential relationship between inadequate response to DNA damage and development of myelodysplastic syndrome. Int J Mol Sci. 2015;16(1):966-89.
37. Kefala M, Papageorgiou SG, Kontos CK, et al. Increased expression of phosphorylated NBS1, a key molecule of the DNA damage response machinery, is an adverse prognostic factor in patients with de novo myelodysplastic syndromes. Leuk Res. 2013;37(11):1576-82.
38. Li X, Corsa CA, Pan PW, et al. MOF and H4 K16 acetylation play important roles in DNA damage repair by modulating recruitment of DNA damage repair protein Mdc1. Mol Cell Biol. 2010;30(22):5335-47.
39. Li X, Wu L, Corsa CA, Kunkel S, Dou Y. Two mammalian MOF complexes regulate transcription activation by distinct mechanisms. Mol Cell. 2009;36(2):290-301.
40. Wong JC, Hasan MR, Rahman M, et al. Nucleophosmin 1, upregulated in adenomas and cancers of the colon, inhibits p53-mediated cellular senescence. Int J Cancer. 2013;133(7):1567-77.
41. Dasso M. RCC1 in the cell cycle: The regulator of chromosome condensation takes on new roles. Trends Biochem Sci. 1993;18(3):96-101.
42. Moore WJ, Zhang C, Clarke PR. Targeting of RCC1 to chromosomes is required for proper mitotic spindle assembly in human cells. Curr Biol. 2002;12(16):1442-7.

43. Donninger H, Vos MD, Clark GJ. The RASSF1A tumor suppressor. J Cell Sci. 2007;120(18):3163-72.
44. Donninger H, Clark J, Rinaldo F, et al. The RASSF1A tumor suppressor regulates XPA-mediated DNA repair. Mol Cell Biol. 2015;35(1):277-87.
45. Ohtsuka T, Jensen MR, Kim HG, Kim KT, Lee SW. The negative role of cyclin G in atm-dependent p53 activation. Oncogene. 2004;23(31):5405-8.
46. Park EJ, Kim JH, Seong RH, Kim CG, Park SD, Hong SH. Characterization of a novel mouse cDNA, ES18, involved in apoptotic cell death of T-cells. Nucleic Acids Res. 1999;27(6):1524-30.
47. Tian Y, Huang Z, Wang Z, et al. Identification of novel molecular markers for prognosis estimation of acute myeloid leukemia: Over-expression of PDCD7, FIS1 and Ang2 may indicate poor prognosis in pretreatment patients with acute myeloid leukemia. PLoS One. 2014;9(1):e84150.
48. Suzuki S, Nakasato M, Shibue T, Koshima I, Taniguchi T. Therapeutic potential of proapoptotic molecule Noxa in the selective elimination of tumor cells. Cancer Sci. 2009;100(4):759-69.
49. Zaher M, Akrout I, Mirshahi M, Kolb JP, Billard C. Noxa upregulation is associated with apoptosis of chronic lymphocytic leukemia cells induced by hyperforin but not flavopiridol. Leukemia. 2009;23(3):594-6.
50. Lindström MS. Emerging functions of ribosomal proteins in gene-specific transcription and translation. Biochem Biophys Res Commun. 2009;379(2):167-70.
51. Zhang Y, Lu H. Signaling to p53: Ribosomal proteins find their way. Cancer Cell. 2009;16(5):369-77.
52. Warner JR, McIntosh KB. How common are extraribosomal functions of ribosomal proteins? Mol Cell. 2009;34(1):3-11.
53. Chaisuparat R, Rojanawatsirivej S, Yodsanga S. Ribosomal protein S6 phosphorylation is associated with epithelial dysplasia and squamous cell carcinoma of the oral cavity. Pathol Oncol Res. 2013;19(2):189-93.
54. Hutchinson JA, Shanware NP, Chang H, Tibbetts RS. Regulation of ribosomal protein S6 phosphorylation by casein kinase 1 and protein phosphatase 1. J Biol Chem. 2011;286(10):8688-96.
55. Drabkin HA, Parsy C, Ferguson K, et al. Quantitative HOX expression in chromosomally defined subsets of acute myelogenous leukemia. Leukemia. 2002;16(2):186-95.
56. Sitwala KV, Dandekar MN, Hess JL. HOX proteins and leukemia. Int J Clin Exp Pathol. 2008;1(6):461.
57. Owens BM, Hawley RG. HOX and non-HOX homeobox genes in leukemic hematopoiesis. Stem Cells. 2002;20(5):364-79.
58. Argiropoulos B, Yung E, Humphries RK. Unraveling the crucial roles of Meis1 in leukemogenesis and normal hematopoiesis. Genes Dev. 2007;21(22):2845-9.
59. Dickson GJ, Kwasniewska A, Mills KI, Lappin TR, Thompson A. Hoxa6 potentiates short-term hemopoietic cell proliferation and extended self-renewal. Exp Hematol. 2009;37(3):322-33.e3.
60. Huang Y, Sitwala K, Bronstein J, et al. Identification and characterization of Hoxa9 binding sites in hematopoietic cells. Blood. 2012;119(2):388-98.
61. Boimel PJ, Cruz C, Segall JE. A functional in vivo screen for regulators of tumor progression identifies HOXB2 as a regulator of tumor growth in breast cancer. Genomics. 2011;98(3):164-72.
62. Zhang Y, Cheng JC, Huang HF, Leung PC. Homeobox A7 stimulates breast cancer cell proliferation by up-regulating estrogen receptor-alpha. Biochem Biophys Res Commun. 2013;440(4):652-7.
63. Li Q, Zhu F, Chen P. miR-7 and miR-218 epigenetically control tumor suppressor genes RASSF1A and Claudin-6 by targeting HoxB3 in breast cancer. Biochem Biophys Res Commun. 2012;424(1):28-33.
64. Li Y, Yang XH, Fang SJ, et al. HOXA7 stimulates human hepatocellular carcinoma proliferation through cyclin E1/CDK2. Oncol Rep. 2015;33(2):990-6.
65. Pathak BR, Breed AA, Nakhawa VH, Jagtap DD, Mahale SD. Growth inhibition mediated by PSP94 or CRISP-3 is prostate cancer cell line specific. Asian J Androl. 2010;12(5):677-89.
66. Kim CW, Lee TH, Park KH, Choi SY, Kim J. Human lactoferrin suppresses TNF-α-induced intercellular adhesion molecule-1 expression via competition with NF-κB in endothelial cells. FEBS Lett. 2012;586(3):229-34.
67. Zhang H, Feng X, Liu W, et al. Underlying mechanisms for LTF inactivation and its functional analysis in nasopharyngeal carcinoma cell lines. J Cell Biochem. 2011;112(7):1832-43.
68. Gegonne A, Devaiah BN, Singer DS. TAF7: Traffic controller in transcription initiation. Transcription. 2013;4(1):29-33.
69. Bhoumik A, Ronai Z. ATF2: A transcription factor that elicits oncogenic or tumor suppressor activities. Cell Cycle. 2008;7(15):2341-5.
70. Lopez-Bergami P, Lau E, Ronai Z. Emerging roles of ATF2 and the dynamic AP1 network in cancer. Nat Rev Cancer. 2010;10(1):65-76.
71. Gozdecka M, Breitwieser W. The roles of ATF2 (activating transcription factor 2) in tumorigenesis. Biochem Soc Trans. 2012;40(1):230-4.
72. Lau E, Ronai ZA. ATF2 - at the crossroad of nuclear and cytosolic functions. J Cell Sci. 2012;125(Pt 12):2815-24.

3

Prognostic impact of the number of methylated genes in myelodysplastic syndromes and acute myeloid leukemias treated with azacytidine

María Abáigar[1,2], Fernando Ramos[3,4], Rocío Benito[1,2], María Díez-Campelo[2,5], Javier Sánchez-del-Real[3], Lourdes Hermosín[6], Juan Nicolás Rodríguez[7], Carlos Aguilar[8], Isabel Recio[9], Jose María Alonso[10], Natalia de las Heras[3], Marta Megido[11], Marta Fuertes[3], María Consuelo del Cañizo[2,5], Jesús María Hernández-Rivas[1,2,5]

[1]*Cancer Research Center-IBMCC (USAL-CSIC), Salamanca, Spain*

[2]*IBSAL, Salamanca, Spain*

[3]*Dept. of Hematology, Hospital de León, Spain*

[4]*Institute of Biomedicine (Ibiomed), University of León, Spain*

[5]*Dept. of Hematology, Hospital Universitario de Salamanca, Spain*

[6]*Dept. of Hematology, Hospital del SAS, Jerez de la Frontera, Spain*

[7]*Dept. of Hematology, Hospital Juan Ramón Jiménez, Huelva, Spain*

[8]*Dept. of Hematology, Hospital Santa Bárbara, Soria, Spain*

[9]*Dept. of Hematology, Hospital Nuestra Señora de Sonsoles, Ávila, Spain*

[10]*Dept. of Hematology, Hospital Río Carrión, Palencia, Spain*

[11]*Dept. of Hematology, Hospital del Bierzo, Ponferrada, Spain*

Ann Hematol. (2013) doi: 10.1007/s00277-013-1799-9

PubMed PMID: 23740492

Ann Hematol
DOI 10.1007/s00277-013-1799-9

ORIGINAL ARTICLE

Prognostic impact of the number of methylated genes in myelodysplastic syndromes and acute myeloid leukemias treated with azacytidine

María Abáigar · Fernando Ramos · Rocío Benito · María Díez-Campelo · Javier Sánchez-del-Real · Lourdes Hermosín · Juan Nicolás Rodríguez · Carlos Aguilar · Isabel Recio · Jose María Alonso · Natalia de las Heras · Marta Megido · Marta Fuertes · María Consuelo del Cañizo · Jesús María Hernández-Rivas

Received: 5 March 2013 / Accepted: 16 May 2013

Abstract The prognostic impact of the aberrant hypermethylation in response to azacytidine (AZA) remains to be determined. Therefore, we have analyzed the influence of the methylation status prior to AZA treatment on the overall survival and clinical response of myeloid malignancies. DNA methylation status of 24 tumor suppressor genes was analyzed by methylation-specific multiplex ligation-dependent probe amplification in 63 patients with myelodysplastic syndromes and acute myeloid leukemia treated with azacytidine. Most patients (73 %) showed methylation of at least one gene, but only 12 % of patients displayed ≥3 methylated genes. The multivariate analysis demonstrated that the presence of a high number (≥2) of methylated genes (P=0.022), a high WBC count (P=0.033), or anemia (P=0.029) were independent prognostic factors associated with shorter overall survival. The aberrant methylation status did not correlate with the response to AZA, although four of the five patients with ≥3 methylated genes did not respond. By contrast, favorable cytogenetics independently influenced the clinical response to AZA as 64.7 % of patients with good-risk cytogenetic abnormalities responded (P=0.03). Aberrant

Electronic supplementary material The online version of this article (doi:10.1007/s00277-013-1799-9) contains supplementary material, which is available to authorized users.

M. Abáigar · R. Benito · J. M. Hernández-Rivas
Cancer Research Center-IBMCC (USAL-CSIC), Salamanca, Spain

M. Abáigar · R. Benito · M. Díez-Campelo · M. C. del Cañizo · J. M. Hernández-Rivas
IBSAL, Salamanca, Spain

F. Ramos · J. Sánchez-del-Real · N. de las Heras · M. Fuertes
Department of Hematology, Hospital de León, León, Spain

F. Ramos
Institute of Biomedicine (Ibiomed), University of León, León, Spain

M. Díez-Campelo · M. C. del Cañizo · J. M. Hernández-Rivas
Department of Hematology, Hospital Universitario de Salamanca, Salamanca, Spain

L. Hermosín
Department of Hematology, Hospital del SAS, Jerez de la Frontera, Spain

J. N. Rodríguez
Department of Hematology, Hospital Juan Ramón Jiménez, Huelva, Spain

C. Aguilar
Department of Hematology, Hospital Santa Bárbara, Soria, Spain

I. Recio
Department of Hematology, Hospital Nuestra Señora de Sonsoles, Ávila, Spain

J. M. Alonso
Department of Hematology, Hospital Río Carrión, Palencia, Spain

M. Megido
Department of Hematology, Hospital del Bierzo, Ponferrada, Spain

J. M. Hernández-Rivas (✉)
Servicio de Hematología y Departamento de Medicina, Hospital Universitario de Salamanca, Paseo San Vicente 58-182, 37007 Salamanca, Spain
e-mail: jmhr@usal.es

Published online: 06 June 2013

methylation status influences the survival of patients treated with AZA, being shorter in those patients with a high number of methylated genes.

Keywords Acute myeloid leukemia · Azacytidine · Methylation · Myelodysplastic syndromes · Survival

Introduction

Myelodysplastic syndromes (MDS) are clonal hematopoietic stem-cell disorders associated with a myeloid differentiation blockade leading to bone marrow (BM) accumulation of myeloid progenitor cells and peripheral blood cytopenias. Approximately 30 % of MDS progress to acute myeloid leukemia (AML) [1, 2].

For many years, therapy for MDS was based on best supportive care consisting of red blood cell or platelet transfusions, as well as intensive chemotherapy, and allogeneic stem-cell transplantation (alloSCT) for high-risk MDS, the latter being the only curative strategy although only a few patients benefit from it. However, in recent years, new promising treatment options with proven clinical efficacy, such as hypomethylating agents (azacytidine and decitabine) and lenalidomide have emerged [3–6].

Azacytidine (AZA) is a natural analog of cytidine that irreversibly binds to DNA methyltransferases and inhibits their function, resulting in a progressive reduction of DNA methylation and gene expression reactivation [7]. Recent studies have shown that AZA induces 50–60 % responses in MDS and AML patients and significantly improves overall survival [3, 4, 7–11].

DNA methylation is an epigenetic mechanism that consists in the addition of methyl groups to cytosine residues within CpG-islands, which are located in or near gene promoter regions. Aberrant DNA methylation of CpG-islands has been associated with gene silencing, via inhibition of gene transcription [11, 12]. Interestingly, in recent years, an increasing number of hypermethylated tumor suppressor genes have been related to myeloid neoplasms. Some of these genes are related to cell cycle control (*CDKN1B*, *CDKN2B*, *HIC1*, and *FHIT*), cell adhesion (*IGSF4*, *CDH1*, and *CDH13*), apoptosis regulation (*DAPK1*), and cell growth (*ESR1*) [11, 13–17]. However, the prognostic impact of the aberrant methylation status prior to any treatment with hypomethylating agents in patients with MDS or AML has been poorly investigated. In fact, the results from the few studies that explored the relationship between aberrant DNA methylation and response to hypomethylating agents are discordant [10, 11]. In addition, no association between aberrant methylation of 10 selected genes and clinical response to decitabine was found, while there was an association with overall survival [11]. Conversely, a relationship between the hypermethylation of *P15* gene or the presence of high levels of methylation and response to AZA and survival has been reported by others [10, 13]. In addition, *TET2* mutations and *DNMT3B* gene amplification have been recently described to be associated with response to AZA treatment [18–20].

In the present study, we have analyzed the methylation status of a selected set of tumor suppressor genes in MDS and AML patients treated with azacytidine, in order to evaluate the influence of aberrant methylation on survival and clinical response to azacytidine.

Materials and methods

Patient characteristics

A total of 78 patients from 11 Spanish institutions diagnosed as MDS or AML who had been treated with AZA (VIDAZA®, Celgene Corp., Summit, NJ) were prospectively evaluated. Among these, 15 patients were excluded from the study because AZA had been used as maintenance after alloSCT (n=5), response was not assessable (n=4), or DNA quality was considered inadequate for analysis (n=6). Finally, the study cohort included a total number of 63 patients. BM samples were obtained before AZA therapy. In addition, five control BM samples were obtained from healthy individuals. This study was performed in agreement with the guidelines of *Declaration of Helsinki*, was approved by the Local Ethical Committees, and all patient samples were collected after obtaining informed consent.

Patient characteristics were collected just before AZA treatment and are summarized in Table 1. Median age was 69 years (range 49–84) and 65.1 % of the patients were males. According to the World Health Organization (WHO) 2008 classification [21], the study included 39 patients diagnosed with MDS and 24 with AML. Among the MDS patients, two had MDS associated with isolated del(5q), 14 cases had refractory cytopenia with multi-lineage dysplasia, 12 patients had refractory anemia with excess blasts (RAEB) type 1, 8 cases were diagnosed as RAEB-2, and the remaining 3 patients, as MDS unclassified (MDS-U). Median leukocyte count was 3×10^9/L (range 0.7 to 51.4×10^9/L) with 11.9 % of patients showing at least $\geq15\times10^9$/L, median hemoglobin level was 8.9 g/dL (range 5.4 to 14.0 g/dL) with 67.8 % of patients showing less than 10 g/dL, and median platelet count was 51.9×10^9/L (range 2 to 836×10^9/L) with 49.2 % having less than 50×10^9/L. Among MDS patients, one half had less than 5 % BM blasts and among patients with AML, also one half had less than 30 % BM blasts (Table 1).

Conventional cytogenetics

Conventional cytogenetic analysis was carried out with standard G-banding with trypsin-Giemsa staining in all samples included in the study, as previously described [22]. Chromosomal

Table 1 Patient characteristics before AZA treatment

	%	Median [Range]
Age (years)		69 [49–84]
< 70	52.4	
≥ 70	47.6	
Sex		
Male	65.1	
Female	34.9	
WHO diagnosis		
MDS del(5q)	3.2	
RCMD	22.2	
RAEB-1	19.0	
RAEB-2	12.7	
MDS-U	4.8	
AML with MDS related changes	17.5	
AML, other	20.6	
IPSS-R score (MDS patients)		
Very Low	2.8	
Low	16.7	
Intermediate	27.8	
High	27.8	
Very High	25.0	
WBC count ($x10^9$/L)		3 [0.7-51.4]
< 15	88.1	
≥ 15	11.9	
Hemoglobin level (g/dL)		8.9 [5.4-14.0]
< 10	67.8	
≥ 10	32.2	
Platelet count ($x10^9$/L)		51.9 [2–836]
< 50	49.2	
≥ 50	50.8	
LDH (U/L)		393 [189–4486]
< 450	60.0	
≥ 450	40.0	
BM blasts (%)		9.5 [0.6-80.0]
< 5	32.1	
5 - 20	32.1	
20 - 30	17.9	
≥ 30	17.9	
PB blasts (%)		0 [0–58]
Absent	56.1	
Present	43.9	
Cytogenetics		
Normal	55.6	
−5 / del(5q)	6.3	
−7 / del(7q)	11.1	
+ 8	7.9	
−5 / del(5q) + other abnormality	6.3	
Other abnormality	4.8	
Complex	7.9	
Nº cycles		6 [1–21]

WHO World Health Organization, *IPSS-R* Revised International Prognosis Scoring System, *WBC* white blood cells, *LDH* lactate dehydrogenase, *BM* bone marro, *PB* peripheral blood

abnormalities were classified using the International System for Human Cytogenetic Nomenclature criteria [23].

FISH analysis

Interphase fluorescence in situ hybridization (FISH) was performed on all samples using commercially available probes for the following regions: 5p15, 5q33-q34 (LSI D5S23, D5S721 probe), 7p11, 7q31 (Vysis D7S522, CEP 7 Probe), and for the centromere of chromosome 8 (CEP 8 DNA probe) all of them from Vysis/Abbott (Co, Downers Grove, IL), using the previously described methods [24]. Dual-color FISH was performed using differently labeled control and test probes. At least 400 interphase nuclei per probe were analyzed by two independent observers, and images were recorded using an E1000 microscope (Nikon, Tokyo, Japan) equipped with the Quips system (Vysis, Downers Grove, IL). Both, FISH and cytogenetic studies were carried out at the Hospital Universitario de Salamanca.

Response criteria

Clinical response in patients with MDS as well as in AML with low-blast cell count (20–30 %) was assessed according to the International Working Group (IWG) response criteria for MDS [25] in accordance to Fenaux et al. [3, 8], while patients with more than 30 % BM blasts were rated according to the International Working Group (IWG) criteria for AML following European LeukemiaNet recommendations [26].

DNA isolation

DNA from all samples was isolated with QIAamp DNA Mini Kit (Qiagen) according to manufacturer's standard protocol. Concentration and quality of extracted DNA were determined with NanoDrop spectrophotometer by measuring the ratio of absorbance at 260 and 280 nm (A260/280). DNA integrity was also evaluated by agarose gel electrophoresis.

Methylation analysis-MLPA

DNA methylation status was analyzed using the methylation specific multiplex ligation-dependent probe amplification (MS-MLPA®) technique. The MS-MLPA probe-mix used was ME001-C1 v08 0808 (MRC-Holland, Amsterdam, The Netherlands) [27] which contains 41 probes. Of these, 26 probes correspond to CpG-islands located in promoter regions of 24 different tumor suppressor genes (listed in Table 2) while the 15 remaining probes are reference probes that lack the HhaI digestion site. These 15 probes were used to quantify the methylation level.

The MS-MLPA experiments were performed following the manufacturer's recommended guidelines with small modifications [27, 28]. In brief, approximately 100 ng of

Table 2 Incidence of methylated genes in MDS and AML samples assessed by MS-MLPA

Gene symbol	Chromosome position	Gene function	Patients with methylated genes (%)		
			MDS (n=36)	AML (n=17)	Total (n=53)
TP73	1p36.32	Cell cycle control	0	4.2	1.6
CASP8	2q33.1	Apoptosis regulation	0	0	0
VHL	3p25.3	Transcription factor binding, apoptosis regulation, cell cycle control	0	0	0
RARB	3p24.2	Transcription regulation, signal transduction	7.7	12.5	9.5
MLH1	3p22.2	Cell cycle control, DNA repair	0	0	0
RASSF1	3p21.31	Cell cycle control, Ras signaling	0	0	0
FHIT	3p14.2	Cell cycle control, nucleotide metabolism	0	4.2	1.6
APC	5q22.2	Wnt signaling, cell adhesion	2.6	0	1.6
ESR1	6q25.1	Transcription regulation, signal transduction, cell growth regulation	17.9	33.3	23.8
CDKN2A	9p21.3	Cell cycle control	5.1	0	3.2
CDKN2B	9p21.3	Cell cycle control	25.6	20.8	23.8
DAPK1	9q21.33	Apoptosis regulation	0	0	0
PTEN	10q23.31	Cell adhesion, apoptosis regulation, cell cycle control	0	0	0
CD44	11p13	Cell adhesion	2.6	0	1.6
GSTP1	11q13.2	Apoptosis regulation	5.1	0	3.2
ATM	11q22.3	DNA repair, apoptosis regulation, cell cycle control	0	0	0
IGSF4 (*CADM1*)	11q23.2	Cell adhesion	17.9	41.7	27.0
CDKN1B	12p13.1	Cell cycle control	12.8	8.3	11.1
CHFR	12q24.33	Cell cycle control, ubiquitination	0	0	0
BRCA2	13q13.1	DNA repair, cell cycle control	0	0	0
CDH13	16q23.3	Cell adhesion	12.8	29.2	19.0
HIC1	17p13.3	Cell cycle control, transcription regulation	2.6	0	1.6
BRCA1	17q21.31	DNA repair, cell cycle control	0	0	0
TIMP3	22q12.3	Apoptosis regulation	0	4.2	1.6

genomic DNA from each sample were denatured and subsequently hybridized with MS-MLPA probes. After hybridization, each mixture was equally split into two tubes: in one tube, all probes were ligated to their respective target sequences; while in the other tube, the ligation reaction was combined with HhaI enzyme digestion. Subsequently, all products were amplified by PCR. Due to the fact that HhaI enzyme is a methylation-sensitive endonuclease, if the DNA is methylated, the HhaI enzyme is not able to digest the DNA-probe hybrid, resulting in the amplification by PCR of only the methylated complex. PCR products were analyzed by capillary electrophoresis and quantified with GeneMapper software (Applied Biosystems). Finally, methylation level was determined with Coffalyser MLPA software v9 (MRC-Holland), as the ratio between peak areas of the digested and the undigested samples. All reactions were carried out twice. According to the methylation level showed by the controls (Supplementary Table S1), a methylation of ≥0.30 for CDKN2B gene and ≥0.10 for the remaining genes was considered qualitatively as an aberrant methylation status.

Statistical analysis

To assess the influence of methylation status and clinical covariates in response to AZA treatment, univariate comparisons were performed using t test for continuous variables analysis and Chi-square test for categorical variables. Only those variables with a P value <0.1 in univariate analysis were introduced in a multivariate study, using the logistic regression model.

Survival analysis was performed using standard univariate and multivariate statistical tests by using Kaplan–Meier test (*log-rank*) and Cox regression method, respectively.

For both the association and survival analyses, P values ≤0.05 were considered statistically significant. The statistical package used for all analyses was SPSS version 17.0.

Results

Cytogenetics

A total of 28 patients had abnormal cytogenetics: 4 with isolated −5/del(5q), 7 with isolated −7/del(7q), 5 with trisomy 8, 4 with an additional chromosomal abnormality to chromosome 5 alteration, 1with trisomy 9, 1 with trisomy 11, 1 with trisomy 14, and 5 patients with a complex cytogenetics (≥3 abnormalities). The remaining 35 patients showed normal cytogenetics. Those cases with abnormal karyotype are summarized in the Supplementary Table S2.

All patients with complex cytogenetics or additional chromosomal abnormalities to chromosome 5 alterations were considered as having adverse cytogenetics. In contrast, trisomy of chromosomes 8, 9, 11, and 14, as well as normal cytogenetics were defined as favorable. According to previous reports, isolated −5/del(5q) or −7/del(7q) were considered as adverse abnormalities in the AML group [29, 30]. By contrast, the presence of isolated −5/del(5q) and isolated del(7q) were defined as being favorable in MDS patients, while the presence of isolated monosomy 7 or −7/del(7q) with any additional chromosomal abnormalities were considered as adverse cytogenetics [3, 31, 32].

Methylation analysis

The MS-MLPA analysis detected aberrant DNA methylation of at least one gene in 46 of all 63 patients (73 %): 23 patients displayed an aberrant methylation of only one gene, 15 samples in two genes, and in the remaining 8 patients, an aberrant methylation was detected in three or more genes. No differences in the frequency of methylation in AML vs. MDS samples were observed (79.2 vs. 69.2 %, P=0.39).

The most frequently methylated genes were *IGSF4* (also known as *CADM1*) (27 %), *CDKN2B* (23.8 %), *ESR1* (23.8 %), *CDH13* (19 %), and *CDKN1B* (11.1 %). Other genes such as*TP73*, *RARB*, *FHIT*, *APC*, *CDKN2A*, *CD44*, *GSTP1*, *HIC1*, and *TIMP3* were found methylated in less than 10 % of patients. All of these genes are listed in Table 2. No methylation was observed in the control samples.

Analysis of prognostic factors for overall survival

Among the 63 patients included in the study, 41 patients died during the study. Median survival was 11.7 months (range 0.6–40.5), with a median follow-up time of 19.2 months. In the univariate analysis, a WBC count ≥15×10^9/L (P=0.021), a hemoglobin level <10 g/dL (P=0.015) and the presence of ≥2 methylated genes (P=0.017) were associated with shorter overall survival (OS) (Table 3, Fig. 1). In contrast, the presence of an aberrant cytogenetics (P=0.085) was not a predictive factor for overall survival. It is noteworthy that no statistical differences between patients with MDS and those with AML were found (P=0.507) (Fig. 2).

Multivariate analysis revealed that the presence of a WBC count ≥15×10^9/L (P=0.033), a hemoglobin level <10 g/dL (P=0.029) and a high number of methylated genes (≥2 genes methylated) (P=0.022) were independently associated with shorter OS (Table 3). It should be noted that the eight patients with ≥3 methylated genes showed a shorter survival (Supplementary Fig. S3).

Table 3 Prognostic factors for overall survival

Patient characteristics before treatment	Univariate analysis		Multivariate analysis	
	Median OS (months)	*P* value	HR [95 % CI]	*P* value
Age (years) (<70/≥70)	12.8/16.6	0.680		
Sex (Male/Female)	12.6/20.3	0.175		
WHO diagnosis (MDS/AML)	13.5/11.8	0.507		
WBC count (×10^9/L) (<15/≥15)	13.5/9.7	0.021	2.7 [1.1–6.9]	0.033
Hemoglobin level (g/dL) (<10/≥10)	10.7/20.3	0.015	0.4 [0.2–0.9]	0.029
Platelet count (×10^9/L) (<50/≥50)	9.7/19.6	0.073		
LDH (U/L) (<450/≥450)	12.8/13.2	0.538		
BM blasts % (<5/5–20/20–30/≥30)	19.2/10.7/9.3/11.8	0.945		
PB blasts (absent/present)	13.5/11.8	0.562		
Cytogenetics (favorable/adverse)	16.6/8.9	0.085		
Normal cytogenetics (yes/no)	12.8/13.2	0.226		
Number of methylated genes (≤1/≥2)	19.2/10.7	0.017	2.2 [1.1–4.3]	0.022

OS overall survival, *WHO* World Health Organization, *WBC* white blood cells, *LDH* lactate dehydrogenase, *BM* bone marrow, *PB* peripheral blood, *HR* hazard ratio, *CI* confidence interval

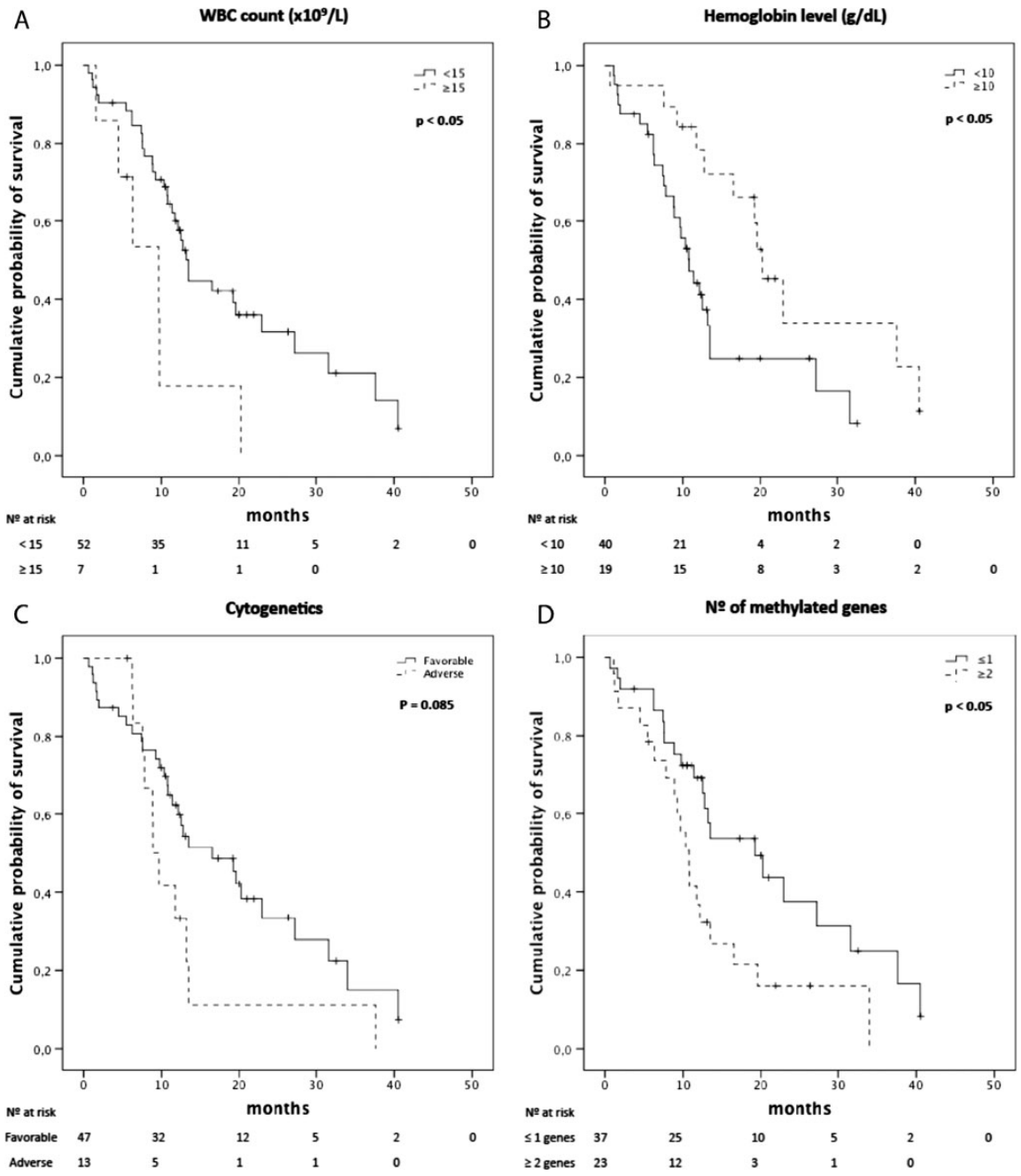


Fig. 1 Kaplan–Meier curves according to **a** WBC count, **b** hemoglobin level, **c** cytogenetic group, and **d** number of methylated genes. Patients showing a WBC count ≥15×109/L, a hemoglobin level <10 g/dL, adverse cytogenetics, and ≥2 methylated genes had a shorter OS

Analysis of predictive factors for response to azacytidine

In order to assess the factors influencing the overall response to AZA, we excluded from the study those patients who received less than 4 cycles (n=10) in which response to AZA could not be evaluated. The median number of cycles administered was 7 (range from 4 to 21). Responses are summarized in Table 4. In brief, among all patients, 49.1 %

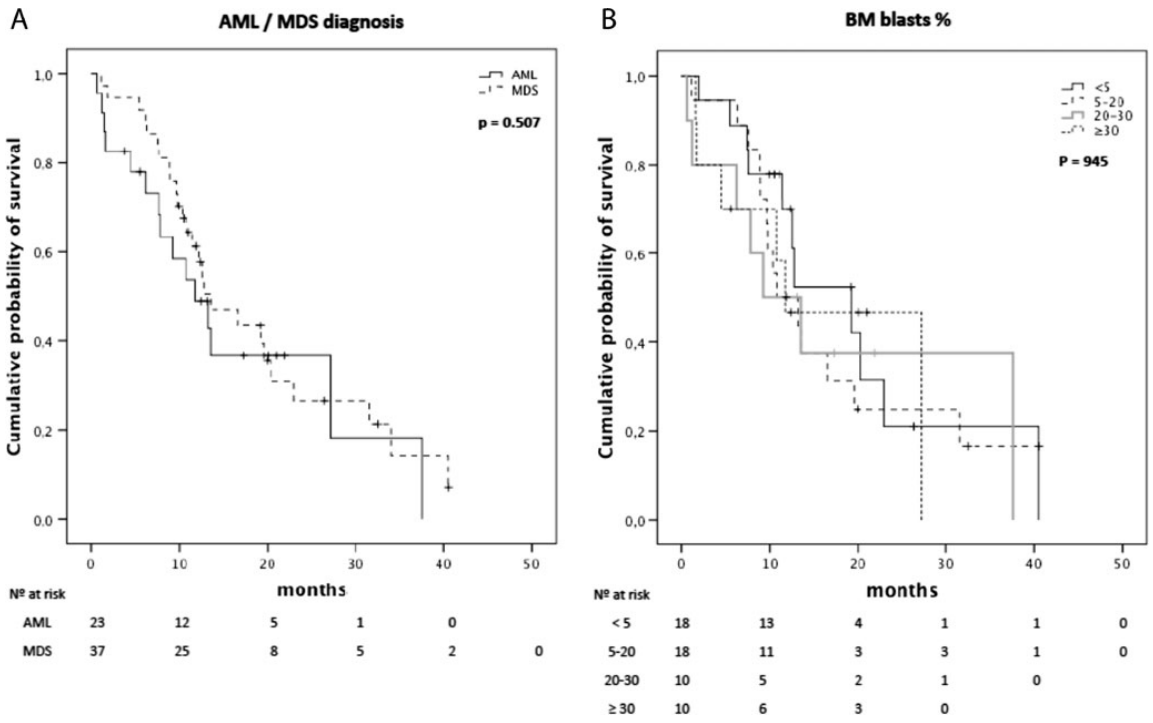


Fig. 2 Kaplan–Meier curves according to **a** MDS or AML diagnosis and **b** the BM blasts cell count (%). No statistical differences were observed between MDS and AML, as well as between patients with 5 % BM blasts compared to those showing 5–20, 20–30, and 30 % of blast cells in the bone marrow

achieved a response to AZA: 26.4 % achieved a complete remission (CR), 5.7 % of patients had a partial remission (PR), and 17 % showed a hematological improvement (HI), while 9.4 % had stable disease (SD). By contrast, 22.6 % of patients did not achieve any response and the remaining 18.9 % showed progressive disease (Table 4). WHO-defined MDS and AML patients showed similar responses as regards to overall survival, complete remission, and progression rates. We grouped those patients who achieved CR, PR, and HI as "responders", while those not achieving any response or progressing under AZA therapy were grouped as "non-responders".

Univariate analysis showed that the only factors predictive of a response to AZA treatment were the presence of less than 5.5 % of blast cells in peripheral blood (PB) (P=0.05) and favorable cytogenetics (P=0.02). However, methylation level or the methylation status of any individual gene was not significantly associated with response to AZA treatment (Table 5). WHO diagnosis of MDS vs. AML did not show any influence on clinical response to AZA.

Methylation level was not significantly associated with response. However, only one out of the five patients with ≥3 genes methylated before treatment achieved a partial remission. The other four patients, who did not respond, included one case with adverse cytogenetics, one patient showing a monosomy 7, and the remaining two cases with normal cytogenetics.

Multivariate analysis including cytogenetics and the methylation status revealed that favorable cytogenetics (P=0.026) was the only factor that independently influenced the response to AZA (Table 5). In our study cohort, 64.7 % of patients with favorable cytogenetics responded to AZA, while only 28.6 % with adverse cytogenetics did respond to the drug. Interestingly, the study comprised seven cases with isolated −7/del(7q): four AML patients, of which only one patient did show a response to AZA (SD), and three MDS patients, of which all had a response.

Discussion

Nowadays there are an increasing number of patients with hematological malignancies treated with hypomethylating agents (azacytidine and decitabine). Clinical experience has demonstrated the effectiveness of this epigenetic therapy as a high proportion of patients treated with these agents achieve a response. However, there are some unresponsive patients, and therefore, it is important to better identify the predictive factors for response to these agents. In the present

Table 4 Response to AZA treatment (based on IWG criteria)

IWG 2006	MDS		AML		TOTAL	
	n=36	Percent (%)	*n*=17	Percent (%)	*n*=53	Percent (%)
Complete remission (CR)	10	27.8	4	23.5	14	26.4
Partial remission (PR)	2	5.6	1	5.9	3	5.7
Hematological improvement (HI)	7	19.4	2	11.8	9	17.0
Stable disease (SD)	4	11.1	1	5.9	5	9.4
Failure	6	16.7	6	35.3	12	22.6
Disease progression	7	19.4	3	17.6	10	18.9

IWG criteria applied to each individual AML patient were dependent on his/her actual BM blast count (30 % cutoff) (see Materials and Methods section for additional details)

IWG International Working Group.

study, we carried out a methylation analysis of 24 tumor-related genes in a series of 63 patients with myeloid malignancies. Our data showed that AML and MDS patients had a low incidence of aberrant hypermethylation, which is even lower among MDS patients, and that the methylation status correlated with survival as an independent factor. Thus, a high number of methylated genes, in addition to a high WBC count and the presence of anemia were associated with shorter overall survival. By contrast, only cytogenetics independently influenced the response to AZA.

Furthermore, one of the goals of the present study was to evaluate the clinical application of the MS-MLPA technique for methylation analysis. Among the other techniques available today for the analysis of the aberrant methylation status of single or multiple genes, such as methylation-specific PCR, bisulfite pyrosequencing, and methylation arrays, the multiplex MS-MLPA technique allows the screening of promoter methylation status of a large number of genes in one single experiment. Both MDS and AML patients showed promoter aberrant hypermethylation of several genes [10–17, 33, 34]. Our study included the analysis of many of these previously reported genes such as *CDKN2B*, *CDKN1B*, *HIC1*, *FHIT*, *IGSF4*, *CDH13*, *TP73*, *ESR1*, and *MLH1*, and our results are in concordance with those in previous reports, as the most frequently methylated genes were *IGSF4*, *CDKN2B*, *ESR1*, *CDH13*, and *CDKN1B* [11, 15, 17, 34]. The MS-MLPA technique uses a low amount of DNA and only requires standard laboratory equipment [27, 35–38]. Thus, MS-MLPA could be a technique that could be used in the clinical setting for the analysis of DNA methylation.

Table 5 Analysis of predictive factors for response to AZA

Patient characteristics before treatment	Univariate analysis			Multivariate analysis	
	Mean or Number of patients		*P* value	OR [95 % CI]	*P* value
	Responders	Non-responders			
Age (years)	68.0	67.2	0.753		
Sex (male/female)	17/9	15/7	0.838		
WHO diagnosis (MDS/AML)	19/7	13/9	0.306		
WBC count ($\times 10^9$/L)	5.1	6.1	0.640		
Hemoglobin level (g/dL)	9.3	8.6	0.124		
Platelet count ($\times 10^9$/L)	141.7	73.6	0.147		
LDH (U/L)	449.3	541.4	0.267		
BM blasts (%)	14.4	17.3	0.600		
PB blasts (%)	1.4	10.8	0.047		
Cytogenetics (favorable/adverse)	22/4	12/10	0.022	0.36 [0.04–0.7]	0.026
Normal cytogenetics (yes/no)	15/11	11/11	0.594		
Number of methylated genes (≤1/≥2)	17/9	13/9	0.654	0.011 [−0.29–0.3]	0.942

OR overall response, *WHO* World Health Organization, *WBC* white blood cells, *LDH* lactate dehydrogenase, *BM* bone marrow, *PB* peripheral blood

The presence of gene promoter hypermethylation in MDS and AML patients has been reported in several studies to have prognostic significance [11, 13, 16]. In our study, the presence of more than two methylated genes was related to a shorter survival in patients treated with AZA. Several studies have measured the DNA methylation level of a selected set of genes, previously evaluated in MDS and AML, showing that lower levels of methylation correlated with longer survival [3, 9, 11]. Both previously reported data and our results suggest that poor-risk cytogenetics and a high number of methylated genes predict a shorter OS [11] (Herman JG, et al. Presented at. AACR 2009 [Abstract 4746]).

The relationship between aberrant methylation of single or multiple genes and response to treatment with hypomethylating agents remains a controversial issue. In accordance with previous reports, we failed to demonstrate any significant association between aberrant DNA methylation immediately *before treatment* and response to AZA [11]. By contrast, a single study suggested that hypermethylation of *CDKN2B* gene *at baseline* may predict response to AZA, as the methylation level in responders was lower than in non-responders [10].

Cytogenetics was related to response to AZA, and our data suggest that the presence of poor-risk cytogenetic abnormalities could predict a lower response. The presence of complex cytogenetics and any additional abnormality accompanying chromosome 5 alteration in either MDS or AML patients, as well as isolated −5/del(5q) and −7/del(7q) in AML patients, should be considered poor predictive features, as previously described [10]. By contrast, we have considered MDS with isolated del(7q) as a good prognostic category because in the AZA-001 trial, MDS patients with −7/del(7q) treated with AZA had a significantly higher overall survival than MDS patients with −7/del(7q) with conventional care [3]. In addition, a recent study reported that MDS patients with isolated del(7q) had a better prognosis than those with isolated monosomy 7 or those having −7/del(7q) with any other additional chromosomal abnormalities [31]. For these reasons, we considered MDS patients treated with AZA with isolated del(7q) as having favorable cytogenetics. Interestingly, among our study cohort, three MDS patients had isolated −7/del(7q), all of which responded, and by contrast, only one of four AML patients with isolated −7/del(7q) did show a response to AZA. Sample size considerations and a lack of additional biological insight preclude any speculation on the biological basis of this surprising fact.

Despite the huge number of gene mutations that have been reported in MDs and AML, the genetic factors that could predict response to hypomethylating agents as azacytidine still remain largely unknown. Nevertheless, mutations in *TET2* and *DNMT3A* genes, as well as *DNMT3A* gene amplification have been described to be associated with a higher response rate to hypomethylating drugs. However, many other epigenetic and genetic factors should be studied in larger series to better analyze their prognostic value. [18–20]

In summary, our results demonstrate that a number out of a selected panel of tumor suppressor genes are found hypermethylated in MDS and AML patients and that their aberrant methylation has a negative influence on overall survival. Moreover, our results suggest that the analysis of the methylation status before AZA treatment by a feasible methodology such as MS-MLPA could be useful for clinicians in order to identify a group of patients with poor survival with AZA therapy, in which other alternative therapeutic approaches might be considered. Further studies in a large cohort of patients are necessary to clearly elucidate the implication of the methylation status in the outcome of patients treated with azacytidine.

Acknowledgments We thank Professor Jesús F. San Miguel for the critical review of the manuscript and Irene Rodríguez, Sara González, Teresa Prieto, Ma. Ángeles Ramos, Almudena Martín, Ana Díaz, Ana Simón, María del Pozo, and Vanesa Gutiérrez of the Centro de Investigación del Cáncer, Salamanca, Spain, for their technical assistance. This work was partially supported by a grant (RD12/0036/0069) from Red Temática de Investigación Cooperativa en Cáncer (RTICC), Instituto de Salud Carlos III (ISCIII), Spanish Ministry of Economy and Competitiveness & European Regional Development Fund (ERDF) "Una manera de hacer Europa", and by grants from the Spanish Fondo de Investigaciones Sanitarias (FIS) [12/000281], Sanidad de Castilla y León (SACYL) [106/A/06, 355/A/09], Fundación Memoria de D. Samuel Solórzano Barruso, COST action-EuGESMA [BM0801], and Celgene Corporation, Spain. MA is fully supported by the Spanish Consejo Superior de Investigaciones Científicas-Junta para la Ampliación de Estudios (CSIC, JAE Predoctoral) [09–02402].

Conflict of interest Fernando Ramos received honoraria and grant support from Celgene Corporation, as well as honoraria from Novartis, Amgen, Pfizer, and Merck Sharp & Dohme. The other authors have declared no conflicts of interest.

References

1. Corey SJ, Minden MD, Barber DL et al (2007) Myelodysplastic syndromes: the complexity of stem-cell diseases. Nat Rev Cancer 7:118–129
2. Nimer SD (2008) Myelodysplastic syndromes. Blood 111:4841–4851. doi:10.1182/blood-2007-08-078139
3. Fenaux P, Mufti GJ, Hellstrom-Lindberg E et al (2009) Efficacy of azacitidine compared with that of conventional care regimens in the treatment of higher-risk myelodysplastic syndromes: a randomised, open-label, phase III study. Lancet Oncol 10:223–232. doi:10.1016/S1470-2045(09)70003-8
4. Garcia-Manero G, Fenaux P (2011) Hypomethylating agents and other novel strategies in myelodysplastic syndromes. J Clin Oncol 29:516–523. doi:10.1200/JCO.2010.31.0854
5. Kantarjian HM, O'Brien S, Shan J et al (2007) Update of the decitabine experience in higher risk myelodysplastic syndrome and analysis of prognostic factors associated with outcome. Cancer 109:265–273
6. Musolino C, Sant'antonio E, Penna G et al (2010) Epigenetic therapy in myelodysplastic syndromes. Eur J Haematol 84:463–473. doi:10.1111/j.1600-0609.2010.01433.x
7. Gotze K, Muller-Thomas C, Peschel C (2009) The role of azacitidine in the management of myelodysplastic syndromes (MDS). Cancer Manag Res 1:119–130

8. Fenaux P, Mufti GJ, Hellstrom-Lindberg E et al (2010) Azacitidine prolongs overall survival compared with conventional care regimens in elderly patients with low bone marrow blast count acute myeloid leukemia. J Clin Oncol 28:562–569. doi:10.1200/JCO.2009.23.8329
9. Itzykson R, Thepot S, Quesnel B et al (2011) Prognostic factors for response and overall survival in 282 patients with higher-risk myelodysplastic syndromes treated with azacitidine. Blood 117:403–411. doi:10.1182/blood-2010-06-289280
10. Raj K, John A, Ho A et al (2007) CDKN2B methylation status and isolated chromosome 7 abnormalities predict responses to treatment with 5-azacytidine. Leukemia 21:1937–1944
11. Shen L, Kantarjian H, Guo Y et al (2010) DNA methylation predicts survival and response to therapy in patients with myelodysplastic syndromes. J Clin Oncol 28:605–613. doi:10.1200/JCO.2009.23.4781
12. Figueroa ME, Skrabanek L, Li Y et al (2009) MDS and secondary AML display unique patterns and abundance of aberrant DNA methylation. Blood 114:3448–3458. doi:10.1182/blood-2009-01-200519
13. Aggerholm A, Holm MS, Guldberg P et al (2006) Promoter hypermethylation of p15INK4B, HIC1, CDH1, and ER is frequent in myelodysplastic syndrome and predicts poor prognosis in early-stage patients. Eur J Haematol 76:23–32
14. Aoki E, Uchida T, Ohashi H et al (2000) Methylation status of the p15INK4B gene in hematopoietic progenitors and peripheral blood cells in myelodysplastic syndromes. Leukemia 14:586–593
15. Hess CJ, Errami A, Berkhof J et al (2008) Concurrent methylation of promoters from tumor associated genes predicts outcome in acute myeloid leukemia. Leuk Lymphoma 49:1132–1141. doi:10.1080/10428190802035990
16. Lin J, Yao DM, Qian J et al (2008) Methylation status of fragile histidine triad (FHIT) gene and its clinical impact on prognosis of patients with myelodysplastic syndrome. Leuk Res 32:1541–1545. doi:10.1016/j.leukres.2008.02.008
17. Valencia A, Cervera J, Such E et al (2011) Aberrant methylation of tumor suppressor genes in patients with refractory anemia with ring sideroblasts. Leuk Res 35:479–483. doi:10.1016/j.leukres.2010.08.012
18. Itzykson R, Kosmider O, Cluzeau T et al (2011) Impact of TET2 mutations on response rate to azacitidine in myelodysplastic syndromes and low blast count acute myeloid leukemias. Leukemia. doi:10.1038/leu.2011
19. Simo-Riudalbas L, Melo SA, Esteller M (2011) DNMT3B gene amplification predicts resistance to DNA demethylating drugs. Genes Chromosomes Cancer 50:527–534. doi:10.1002/gcc.20877
20. Metzeler KH, Walker A, Geyer S et al (2012) DNMT3A mutations and response to the hypomethylating agent decitabine in acute myeloid leukemia. Leukemia 26:1106–1107. doi:10.1038/leu.2011.342
21. Vardiman JW, Thiele J, Arber DA et al (2009) The 2008 revision of the World Health Organization (WHO) classification of myeloid neoplasms and acute leukemia: rationale and important changes. Blood 114:937–951. doi:10.1182/blood-2009-03-209262
22. Hernandez JM, Gonzalez MB, Granada I et al (2000) Detection of inv(16) and t(16;16) by fluorescence in situ hybridization in acute myeloid leukemia M4Eo. Haematologica 85:481–485
23. Shaffer LG, Slovak ML, Campbell LJ (2009) An international system for human cytogenetic nomenclature (2009): recommendations of the International Standing Committee on Human Cytogenetic Nomenclature. S. Karger AG, Basel. ISBN 978-3-8055-8985-7
24. Gonzalez MB, Hernandez JM, Garcia JL et al (2004) The value of fluorescence in situ hybridization for the detection of 11q in multiple myeloma. Haematologica 89:1213–1218
25. Cheson BD, Greenberg PL, Bennett JM et al (2006) Clinical application and proposal for modification of the International Working Group (IWG) response criteria in myelodysplasia. Blood 108:419–425
26. Dohner H, Estey EH, Amadori S et al (2010) Diagnosis and management of acute myeloid leukemia in adults: recommendations from an international expert panel, on behalf of the European LeukemiaNet. Blood 115:453–474. doi:10.1182/blood-2009-07-235358
27. Nygren AO, Ameziane N, Duarte HM et al (2005) Methylation-specific MLPA (MS-MLPA): simultaneous detection of CpG methylation and copy number changes of up to 40 sequences. Nucleic Acids Res 33:e128
28. Schouten JP, McElgunn CJ, Waaijer R et al (2002) Relative quantification of 40 nucleic acid sequences by multiplex ligation-dependent probe amplification. Nucleic Acids Res 30:e57
29. Byrd JC, Mrozek K, Dodge RK et al (2002) Pretreatment cytogenetic abnormalities are predictive of induction success, cumulative incidence of relapse, and overall survival in adult patients with de novo acute myeloid leukemia: results from Cancer and Leukemia Group B (CALGB 8461). Blood 100:4325–4336
30. Grimwade D, Hills RK, Moorman AV et al (2010) Refinement of cytogenetic classification in acute myeloid leukemia: determination of prognostic significance of rare recurring chromosomal abnormalities among 5,876 younger adult patients treated in the United Kingdom Medical Research Council trials. Blood 116:354–365. doi:10.1182/blood-2009-11-254441
31. Cordoba I, Gonzalez-Porras JR, Nomdedeu B et al (2011) Better prognosis for patients with del(7q) than for patients with monosomy 7 in myelodysplastic syndrome. Cancer 118:127–133. doi:10.1002/cncr.26279
32. Greenberg P, Cox C, LeBeau MM et al (1997) International scoring system for evaluating prognosis in myelodysplastic syndromes. Blood 89:2079–2088
33. Iwai M, Kiyoi H, Ozeki K et al (2005) Expression and methylation status of the FHIT gene in acute myeloid leukemia and myelodysplastic syndrome. Leukemia 19:1367–1375
34. Jiang Y, Dunbar A, Gondek LP et al (2009) Aberrant DNA methylation is a dominant mechanism in MDS progression to AML. Blood 113:1315–1325. doi:10.1182/blood-2008-06-163246
35. Berginc G, Bracko M, Glavac D (2010) MS-MLPA reveals progressive age-dependent promoter methylation of tumor suppressor genes and possible role of IGSF4 gene in colorectal carcinogenesis of microsatellite instable tumors. Cancer Invest 28:94–102. doi:10.3109/07357900902849673
36. Castro M, Grau L, Puerta P et al (2010) Multiplexed methylation profiles of tumor suppressor genes and clinical outcome in lung cancer. J Transl Med 8:86. doi:10.1186/1479-5876-8-86
37. Dikow N, Nygren AO, Schouten JP et al (2007) Quantification of the methylation status of the PWS/AS imprinted region: comparison of two approaches based on bisulfite sequencing and methylation-sensitive MLPA. Mol Cell Probes 21:208–215
38. Marzese DM, Gago FE, Vargas-Roig LM, Roque M (2010) Simultaneous analysis of the methylation profile of 26 cancer related regions in invasive breast carcinomas by MS-MLPA and drMS-MLPA. Mol Cell Probes 24:271–280. doi:10.1016/j.mcp.2010.05.002

General Discussion

"The important thing is to never stop questioning"

Albert Einstein

Myelodysplastic syndromes (MDS) are a group of clonal stem cell disorders characterized by a high clinical and biological heterogeneity with a high risk of progression to acute myeloid leukemia (AML). No single genomic event that potentially might be responsible for the development of MDS and its frequent progression to AML has been identified. Instead, a wide variety of genomic and epigenetic abnormalities have been identified, resulting in abnormal differentiation, maturation and growth of hematopoietic myeloid cells, leading to bone marrow (BM) failure and an enhanced risk to develop AML. Considering the molecular heterogeneity of MDS, high-throughput genome-wide studies (aCGH, SNP-A, microarray-based GEP, NGS) are powerful tools to further characterize genomic abnormalities in MDS patients, rather than the study of single targets. Such analyses have provided valuable insights into our current understanding of the biology of these complex disorders, and thereby an improvement in the diagnosis and prognostic assessment of these diseases, as well as provided novel targets for treatment.

High-resolution genome-wide genotyping techniques, such as single nucleotide polymorphism array (SNP-A) and array-based comparative genomic hybridization (aCGH), are powerful tools for studying copy number abnormalities (CNAs) in the genome with higher resolution compared to conventional cytogenetics (CC).[47] The application of these genome-wide scanning techniques in the study of MDS, despite the drawback of their inability to detect balanced translocations, allowed for the identification of new, cytogenetically cryptic and recurrent CNAs that are not detected by CC. To date, several aCGH and SNP-A studies of MDS have been reported, all with similar end-point data, suggesting their potential clinical utility.[13,47,48] These techniques led to a higher detection rate of chromosomal aberrations (70-90%) compared with that picked up with standard cytogenetics alone (40-60%), including those with or without cytogenetic abnormalities, such as patients with normal or non-informative standard cytogenetic results.[13] In our study we identified that 30% of MDS patients with non-informative cytogenetics carried clonal copy number changes as revealed by aCGH, which is in accordance with previous reported series showing CNAs in 50% of this subset of patients.[48] In addition, we identified CNAs in patients with normal karyotypes (10%), mainly in patients with less than 20 good-quality metaphases (overall 13.5%; 11-19 metaphases = 10.5%; ≤10 metaphases = 21.4%). Therefore, aCGH and SNP-A studies provide additional information over CC, allowing the prognostic stratification according to the IPSS-R

that could change the clinical management of this group of patients.[48] In addition, our study confirmed most (86.7%) of the aberrations already detected by CC in patients showing an abnormal karyotype. Thus, a high concordance between aCGH data was observed, confirming the reliability of this technique for the assessment of large genetic abnormalities. Clinically, our results and previous reported data showed that microarray-based karyotyping refines the prognostic assessment of patients with MDS for known lesions (i.e., -5/5q-, -7/7q-, trisomy 8, 20q-, complex karyotype).[42,46,47,53,77] Moreover, microarray-based karyotyping enabled the identification of submicroscopic copy number changes. Thus, in the present study we identified cryptic abnormalities (27% of patients) in regions of known relevance in MDS and MDS/MPN pathogenesis,[2,34,42,50] such as deletions in 2p23.3 (*DNMT3A*), 4q24 (*TET2*), 5q33.1 (*SPARC*), 7q22.1 (*CUX1*), 21q22.12 (*RUNX1*) and Xp11.4 (*BCOR*), as well as gains in 21q22.3 (*U2AF1*). Forty five percent of patients with cryptic CNAs had normal karyotypes, while 15% were non-informative results by CC. Of note, 20% and 15% of normal karyotype patients carried *TET2* and *RUNX1* deletions, respectively. Therefore, microarrays provided new information about critical regions involved in the pathogenesis of MDS. This technique should be used in the screening of large series of MDS patients.

We would emphasize that all CNAs have in common that they either lead to a loss of function in case of the deletion of tumor suppressor genes (TSG) or to a gain of function in case of the amplification of oncogenes. Moreover, cryptic regions of CNAs (microdeletions, microamplifications) or LOH, as they often implicate individual or a few genes, are the most revealing for targeted sequencing, and indeed, these regions have been instrumental in the detection of novel gene mutations.[44,47,131] Thus, new microarray-based-detected lesions allowed for the identification of mutations in *TET2* in MDS/MPN, *CBL* in CMML, *MPL* in RARS-T, and *EZH2* in MDS, CMML and AML.[47] For this reason we decided to assess whether gene mutations occur in the non-deleted allele of the most relevant genes included within the regions affected by cryptic CNAs. Therefore, *TET2*, *DNMT3A*, *RUNX1*, and *BCOR* genes were selected for further amplicon-based deep sequencing. In addition, we also analyzed *TP53* gene in those cases showing a deletion in 17p13, due to its pathogenic significance. Our results showed that a low proportion of patients carried a simultaneous deletion and mutation in the genes investigated, except for *TP53*. In fact, only one patient carried a *DNMT3A* mutation, three patients had a *TET2* mutation, and four patients showed *TP53*

mutations. The presence of these mutations, or the concomitant deletion and mutation, may show an impact on survival. However, further studies are required.

A consideration when using whole-genome microarray-based techniques for the detection of CNAs/LOH is the incidental finding of copy number variations (CNVs) that are present in normal individuals.[132] These CNVs, which are thousands of regions in which segments of DNA are gained or lost, together with SNPs are responsible of the inherited diversity among individuals.[44,133,134] Thus, a straightforward comparison with matched normal DNA for each patient is required to reliably detect aberrations arising from acquired clonal genomic changes.[44] This is one limitation of our study, and thereby in order to reduce the number of false positives generated by the analysis of aCGH, either technical artifacts or CNVs, several measures were undertaken: 1) A commercial "pool" of female/male DNA was used as the control sample in the experiment, 2) Any detected CNA overlapping with those regions cataloged as CNV in the Database of Genomic Variants (http://dgv.tcag.ca/) was excluded from further analysis, and 3) All CNAs were further confirmed by interphase fluorescence in situ hybridization (FISH) or by an independent genome-wide analysis of CNAs with a microarray from a different platform to that used in the main study. In addition, aCGH and SNP-A studies are not able to detect balanced abnormalities (translocations, inversions) and minor pathologic clones. Thus, integrated analysis including both CC and aCGH/SNP-A will overcome this limitation.[43,46]

In the era of genomic medicine, chromosomal abnormalities will continue to have clinical significance and conventional cytogenetics will continue to have a fundamental role in the study of MDS.[2] Our results and previously reported data support the idea that CC should be complemented not only by genome-wide microarray-based analyses, but also by high-throughput sequencing of multi-gene panels, which have already been described for MDS and related myeloid neoplasms, instead of single genes, in the workflows for the study of these groups of patients. However, further studies are warranted to clarify how to integrate this increased knowledge of CNAs and gene mutations in MDS into the clinical practice.

The definition of a complex karyotype is based on the presence of at least 3 independent clonal abnormalities. However, few groups have proposed different thresholds of complexity.[16] In our study we detected three HR-MDS patients showing a much higher

complexity than expected as revealed by aCGH. These three patients displayed complex genomic rearrangements, showing multiple non-contiguous CNAs involving a single chromosome, consistent with chromothripsis.[37-40,58-60,71] Although this genetic abnormality was initially inferred from sequencing data, it could also be inferred from aCGH/SNP-A data. Thus, at least ten changes in segmental copy number between two, or occasionally three, copy number states confined to one chromosome, chromosome arm, or even a few chromosomes are required to define chromothripsis.[59] Similar patterns of copy-number alterations were associated with chromothripsis in other studies.[135] Generally, chromothripsis involves different chromosomes with a random distribution. However, we detected three MDS patients (1.2%) showing chromothripsis on chromosome 13 exclusively. To support the hypothesis that chromothripsis could affect recurrently one single chromosome or chromosomal regions, there are several works reporting chromothripsis on chromosome 16p in 3/7 multiple myeloma patients, chromosome 21 in 5/9 iAMP21 acute lymphoblastic leukemia cases, chromosome 5 in 3/8 chronic lymphocytic leukemia, and recently, chromosome 13 in three retinoblastoma cases.[60,70,136,137] Noteworthy, this last work is the first reporting chromothripsis on chromosome 13, spanning the *RB1* locus, as a recurrent abnormality. The authors suggested that a chromothripsis event, disrupting and, therefore inactivating the tumor suppressor gene *RB1*, might be responsible for the initiation of retinoblastoma by the inactivation of the *RB1* gene.[137] The three MDS cases of our study, displayed different patterns of genomic alteration on chromosome 13, but several genes were found affected in the same way. Thus, *FLT3* was commonly amplified, while *BRCA2* and *RB1* were commonly deleted. According to the suggestion by Mcevoy *et al.*,[137] whereby the inactivation of a tumor suppressor gene by chromothripsis could potentially initiate tumorigenesis, we found two TSGs disrupted in the three patients showing chromothripsis, *BRCA2* and *RB1*.[138,139] Moreover, *TP53* mutations, another TSG, have been described as involved in this complex genomic abnormality.[59,136] Of note, the three MDS cases with chromothripsis carried *TP53* mutations. All together, *FLT3* amplification, a gene promoting proliferation, *BRCA2* and *RB1* disruption, and *TP53* mutations could be implicated in the pathogenesis of some cases of MDS. Chromothripsis has been considered as a “poor risk” feature, as it has been associated with more aggressive tumors, i.e. AML.[59,60,135] In agreement to this, the three cases of the study were high-risk MDS, showing a poor outcome with less than one year survival. Several mechanisms have been suggested as responsible for

this catastrophic genome breakage.[37,38,40,59,61,63,140] One of these mechanisms is the telomere shortening, a phenomenon that has been described in MDS leading to genomic instability, and associated with disease progression.[11,77] Telomerase, TERT on 5p13.33, is essential to maintain telomeres integrity. Inhibition of telomerase activity has been shown to cause cellular senescence and cell death. By contrast, increased *TERT* expression and telomerase activity, leads to telomeres erosion and chromosome end fusion, and consequently chromosome instability.[11] Interestingly, Salaverria *et al*. observed that the 3 CLL cases with chromothripsis on chromosome 5p had the *TERT* gene locus amplified.[136] Undoubtedly, some genetic details could be missed in our study. Nevertheless, aCGH was able to detect a highly rearranged chromosome 13, which reflects many breakage events, sufficient to identify the genomic signature of chromothripsis. Therefore, this genetic abnormality could be involved in the pathogenesis of some cases of MDS.[49,141]

All cytogenetic, genetic and epigenetic abnormalities may affect gene expression by means of aberrant transcription, epigenetic regulation, or gene dosage effects.[142] Gene expression profiling has proved to be a powerful tool for the simultaneous analysis of expression levels of nearly all known genes, and has been applied in the study of MDS allowing for a better understanding of their complex biology. Such studies were mainly focused on the identification of independent gene markers for each disease subtype, by comparing gene expression signatures between MDS and healthy individuals, between different risk groups, or between different cytogenetic subclasses.[106,109-111,115-121] In addition, given that morphological dysplastic features, cytogenetic and epigenetic abnormalities, as well as several gene mutations are common lesions found in MDS and other related myeloid disorders, including AML, it has been suggested that these complex disorders might have common underlying genetic defects.[2,35,92] For these reasons we aimed to analyze common deregulated genes and gene pathways during the evolution of the disease. In this effort, we applied a multi-platform genome-wide gene expression profiling to study those genes whose expression levels evolved following and increasing or decreasing trend from non-malignant BM conditions through early (LR-MDS) and advanced MDS (HR-MDS) to AML.

We identified common genes and gene pathways that were progressively deregulated (up- or down-regulated) in the transition from non-malignant BM conditions to early MDS stages

(LR-MDS), which also remained progressively deregulated during the progression towards advanced MDS (HR-MDS) and AML. Thus, several cellular functions of known relevance in MDS were identified as progressively up-regulated, including DNA damage response and checkpoint pathways, apoptosis, ribosome and translation-related pathways, as well as nucleosome and chromatin assembly, while immune responses showed an increasing down-regulation. Additionally, we identified another group of genes whose expression levels changed drastically in the transition from HR-MDS to AML. These genes were related to promoting cellular growth and proliferation, suppressing the differentiation of myeloid progenitor cells, and down-regulating the immune response.

It is well recognized that advanced MDS are characterized by high levels of genomic instability, translated into an increased occurrence of genomic abnormalities in comparison with early MDS, due to defective function of the mechanisms responding to and repairing of DNA damage.[109,116,143] Genomic instability is a condition in which cells are prone to acquire and accumulate permanent genomic defects.[143,144] The presence of genetic abnormalities is frequent in MDS and AML. Thus, clonal chromosomal abnormalities are found in 40-60% of MDS patients at diagnosis, while cryptic lesions are detected in 70-90% patients as revealed by microarray-based karyotyping, and 90% of MDS patients have at least one mutation in one of 40 known cancer-related genes. Furthermore, the acquisition of additional genomic lesions over time is often found as the disease progresses. Therefore, genomic instability has become an important feature of MDS.[2,16,23,24,35,36,45-47,54,57,144]

Under normal conditions, our genome is constantly exposed to genotoxic stress (i.e., reactive oxygen species, γ-radiation, benzene), however, cells have several mechanisms to protect themselves against permanent DNA damage and its adverse effects, known as DNA damage response (DDR), replication and spindle checkpoints.[143,144] Interestingly, we found a progressive overexpression of several genes implicated in chromatin assembly (*MYST1*, *H3F3B*, *HIST1H2BD*), DNA damage response and cell cycle arrest (*NPM1*, *RCC1*, *RASSF1*, *CCNG1*) during the progression of MDS. Noteworthy, we detected that these DDR and checkpoint pathways already appeared up-regulated in early MDS stages, and continued that trend towards advanced MDS stages and AML. Accordingly, several authors reported an overexpression of several DDR and checkpoints-related genes (*JUN*, *TLK1*, *TP53BP1*,

MRE11A) in RAEB-2 patients.[109,116] Of note, the activation of the DDR machinery (i.e., pNBS1, pATM and γH2AX) was observed in LR-MDS patients, was higher in HR-MDS, and was further increased in AML patients when compared to MDS, suggesting that the activation of DDR is an early event during neoplastic transformation.[145] In the presence of a fully operational machinery that efficiently repairs the DNA damage, the chances of a cell to acquire and accumulate genomic alterations are low.[143] However, several genetic or epigenetic abnormalities affecting key genes of these processes, such as *TP53* and *RAS* mutations, *CDKN2B* hypermethylation, *RAD51*, *RB1*, *BRCA2*, *FANCA* and *FANCD2* down-regulation have been identified in MDS patients.[116,144] In addition, it has been suggested that the DNA repair in the hematopoietic system decreases with age,[144,146] which is of great interest since MDS is a disease that generally affects older population. In addition, several congenital BM failure syndromes (i.e., Fanconi anemia, Bloom syndrome) associated with abnormal DNA repair, as well as people who have received chemo/radiotherapy show an enhanced risk to develop MDS.[144] Based on our observations we might say that in MDS patients, cells are showing a defensive reaction against DNA damage, even in normal karyotype patients and also in the earliest stages of the disease. However, as the disease progresses from early to advanced MDS and AML, these DDR and checkpoint pathways are still overexpressed, maybe still responding to a constant damage in the DNA sequence. Unfortunately, likely due to either deficient repair mechanisms or too much DNA damage, MDS cells accumulate and/or mis-repair damaged DNA, resulting in an increased genomic instability, leading to the progression to AML.[116,144] All these observations indicate that MDS could be a disease associated with an inability to properly respond to DNA damage.

Increased apoptosis is a striking feature of early MDS.[110,115,116,147,148] In fact, previous work from our group showed that *BCL2* expression levels were lower in LR-MDS, and that the underexpression of *BCL2* was caused by aberrant hypermethylation. In addition, this work also suggested that the hypermethylation and decreased expression of the *ETS1* transcription factor, with the consequent deregulation of its apoptosis-related targets, could promote apoptosis in MDS patients.[115] In the present study we found an overexpression of some apoptosis-related genes, such as *PDCD7* and *PMAIP1*, in early MDS patients when compared to non-leukemic controls. These data reflect that initially, in early MDS stages, in response to cells carrying a cellular damage, the homeostasis is working in the right way:

eliminating defective cells.[143] However, in our study, we found increased levels of some apoptosis-related genes in LR-MDS, but also these genes continued to be progressively up-regulated in HR-MDS and AML patients.

Over the last decade, growing evidence has been accumulated regarding the importance of ribosome function and translational control in the development of several human diseases, including cancer.[149-153] Mutations and deletions of several ribosomal protein-coding genes has been found in various congenital disorders and ribosomopathies, such as *RPS19* in Diamond–Blackfan anemia (DBA), and *RPS14* in the 5q- syndrome.[116,120,149,153,154] By contrast, *RPS3* has been found up-regulated in colorectal tumors, *RPL15* and *RPL19* in gastric cancer, and *RPL7a* and *RPL37* in prostate cancer.[150,155] Furthermore, many key tumor suppressors (p53, Rb and PTEN) and oncoproteins (MYC, RAS and PI3K/AKT/mTOR signaling components) have been found to directly regulate ribosome biogenesis,[149,150,152] and therefore highlighting its association with malignant transformation and tumor progression. Various studies have shown that the disruption of ribosome biogenesis by cell stressors (i.e., DNA damage, drugs, oxidative stress, hypoxia) leads to the release of ribosome-free forms of proteins, which, beyond their role in stabilizing the ribosome, also show additional "extra-ribosomal" functions that have not yet been fully characterized, but include activation of p53 (via MDM2) to induce cell cycle arrest, senescence, or apoptosis, as well as transcription/translation, mRNA processing, DNA repair, development and tumorigenesis.[150,152,153,155-159] In addition, it has also been reported that the overexpression of various RPs can promote either tumor progression (RPS6, RPS3, RPS13, RPL13) or suppress tumorigenesis by activating tumor suppressors or inactivating oncoproteins (RPL11, RPL5, RPL23, RPS7, RPS3).[149,150,153,160] The most widespread function of RPs is the activation of p53. Several RPs have been shown to bind the MDM2 protein, and to inhibit the MDM2 E3 ligase activity towards p53. This will result in the stabilization and activation of p53, and consequently in p53-dependent cell cycle arrest and apoptosis.[149,150,152,153,155-160] By contrast, other RPs have shown to play an oncogenic role such as RPS3 in non-small cell lung cancer, when phosphorylated. Akt-mediated phosphorylation of RPS3 triggers its translocation into the nucleus where it has shown to induce the expression of prosurvival genes via association with NF-kB.[153] In the present study we observed an increasing overexpression of several genes coding for RPs during disease evolution. As these RPs-coding genes appear constantly

and progressively overexpressed, they seem promote cell proliferation and progression rather than inducing apoptosis. In fact, various RPs-coding genes were found overexpressed in MDS *vs* normal individuals, and more significantly in MDS who undergo progression to AML, suggesting that up-regulation of RPs may be associated with disease progression and aggressiveness.[110] Interestingly, since rDNA transcription can be therapeutically targeted with the small molecule CX-5461, and mTOR signaling is currently targeted by sirolimus/rapamycin,[149,150] ribosome biogenesis represents a potential target for treatment in MDS and AML patients.

Deregulation of the homeodomain-containing genes (*HOX* and *non-HOX*) has been increasingly accepted as implicated in many hematologic malignancies with a well-recognized role in leukemogenesis.[104,161] According to published data, we identified overexpression of *HOXA9*, *HOXA7*, *HOXB3*, *HOXB2* and *MEIS1* genes in the transition form advanced MDS stages to AML.[104,106,108,161-165] Intriguingly, *HOX* genes, preferentially A and B clusters, and *MEIS1* are expressed at high levels in early hematopoietic progenitor cells and are down-regulated following differentiation.[163-166] In myeloid cells, Hoxa9 and Meis1 have been shown to form a DNA-binding complex that cooperates, with other factors such as Flt3, to regulate the cellular proliferation, differentiation and self-renewal.[164,165,167] Consistent with this, the up-regulation of *HOX* genes, specially the concomitant overexpression of *HOXA9* in its cofactor *MEIS1*, promote hematopoietic progenitor cells self-renewal and proliferation, while inducing myeloid differentiation arrest. Of interest, cooperative activity of Hoxa9 and Meis1 has been shown to protect cells from pro-apoptotic signals, contributing to a general anti-apoptotic effect.[167] Altogether, these mechanisms in turn may lead to the clonal expansion of the malignant clone and explain the accumulation of blasts in the BM and progression to AML.

There is growing evidence that immunodeficiency also underlies the complex pathogenesis of MDS and could be related to disease progression. In our study we observed a down-regulation of several genes implicated in immune response, matrix metalloproteinases, and cell adhesion during the evolution of the disease. It is worth to mention that some of these genes were continuously reduced form non-malignant samples through MDS to AML, while others were drastically repressed in the transition from HR-MDS to AML. This is consistent

with the idea that a dysfunctional immune system favors an immunosuppressive BM microenvironment that allows malignant clones to survive and proliferate.[168]

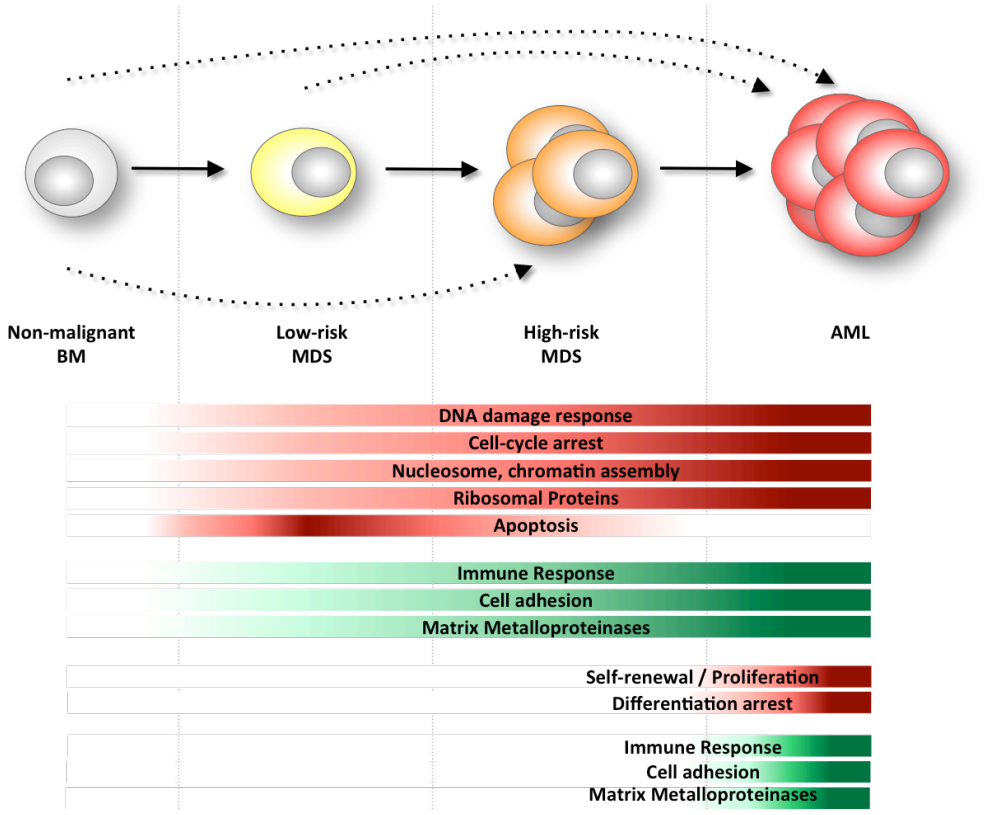


Figure 5. **Proposed model to explain genome-wide expression changes in the development of MDS and progression towards AML**. Red indicates up-regulation and green indicates down-regulation.

On the basis of our observations, we propose a model for MDS development and progression towards AML (Figure 5). Under this scenario, the current picture depicts a malignant disease, with marked genomic abnormalities even at early stages. During the initial phases the cells will show a defensive reaction against DNA damage, by activating the machinery to repair the DNA damage, leading to cell cycle arrest and apoptosis. One of the mechanisms that could activate the DNA damage response may be driven by the

deregulation of ribosomal proteins, via stabilization and activation of p53, after disruption of the nucleolar integrity. Under this conditions, cells will show increased levels of apoptosis, which means that the bone marrow cells are working in the right way, trying to eliminate the defective cells. At this point, the cells still retain their differentiation/maturation capacity. However, the ratio apoptosis/self-renewal-proliferation in the BM is high. This reflects that the derived cells from the initial clone not only continue to differentiate, but also die prematurely. At the same time, due to a down-regulation of the immune response, increased levels of pro-apoptotic cytokines are released. These dysfunctional mechanisms result in an intramedullary excessive apoptosis, which would explain, at least in part, the ineffective hematopoiesis and peripheral blood cytopenias observed in LR-MDS patients (Figure 5).

However, in the presence of accumulating damage on hematopoietic stem cells (HSCs), these cells still show an activation of the DNA damage respond and repair mechanisms, as well as an up-regulation of ribosomal proteins. In addition, due to deficient DNA repair machinery or excessive DNA damage, HSCs and progenitors acquire and accumulate additional genetic and/or epigenetic abnormalities. At this point, damaged cells will show increased proliferative and self-renewal properties, whereas they are not capable to differentiate. Furthermore, these cells show an increased resistance to apoptosis and immune evasion, demanding a high protein synthesis to survive. This could explain the up-regulation of ribosomal proteins, which in this case might contribute to progression. Altogether, these mechanisms would provide to the malignant cell a selective growth advantage, and cooperate allowing for malignant clone expansion. As a result, the proportion of blasts in the bone marrow increases over time and AML eventually develops (Figure 5). Altogether, these data suggest that MDS progression towards AML is an orchestrated process resulting as a consequence of accumulating DNA damage, likely due to that HSCs are not able to accurately respond to DNA damage.

Our study provided a biological basis for many of the alterations in cellular pathways of known significance in MDS evolving to an overt AML, and thereby allowing for a better understanding of the mechanisms underlying the pathogenesis of MDS and their propensity to progress to more aggressive stages. In order to reduce the genetic heterogeneity among

the samples included in the study, we restricted the analysis to patients with normal karyotype exclusively, and excluded RARS and RCMD-RS, which have been associated with somatic *SF3B1* mutations. A possible limitation of our study is the assessment of the mutational status of the samples analyzed. Future studies should be carried out to better identify whether the presence of specific gene mutations might drive the oncogenesis in MDS patients.

Many of MDS patients die due to complications of the disease or progress to AML if they are not treated. In recent years, two DNA methyltransferase inhibitors, decitabine and 5-azacytidine, have been approved for the treatment of patients with MDS and AML and become the standard therapy in MDS and AML patients not suitable for allogeneic stem-cell transplantation (alloSCT). These drugs are effective for the treatment of these disorders, since they have shown to induce responses in about half of patients, with 10-15% of complete responses (CRs), to improve hematopoiesis, delay disease progression and, in some cases, the elimination or reduction of the malignant clone. Therefore, hypomethylating agents (HMAs) have shown to prolong overall survival compared with best supportive care regimens or traditional chemotherapy.[80] However, these responses are temporary and there are still 40% to 50% of patients unresponsive to these agents.[8] For this reason, molecular biomarkers are needed to identify patients who are most likely to respond to treatment with HMAs.

Epigenetic changes have been frequently found in MDS and AML patients and, aberrant hypermethylation especially, has been associated with gene silencing.[80,81,169-171] In contrast to genetic abnormalities, epigenetic silencing is a reversible process.[74] In fact, several studies showed a reduction in methylation after treatment with DNMT inhibitors, and a deregulation of specific cancer-related genes, such as *CDKN2B*, *CDH1*, and *DAPK1*.[80,172] For these reason, we thought that DNA methylation could be used as a predictive factor of response to HMAs. Only few studies explored this hypothesis, but it is still unclear whether baseline methylation correlates with clinical response. Here we examined the methylation status of a panel of 24 tumor suppressor genes in MDS and AML patient who received AZA treatment by MS-MLPA. Despite one study reporting that *CDKN2B* hypermethylation may predict response to AZA,[83] another work reported that no association was found.[172] In addition, methylation reversal

after treatment was not associated with clinical response either.[172] Accordingly, we failed to demonstrate any significant correlation between the presence of an aberrant methylation prior AZA treatment of either any single gene or gene combinations and the clinical response. However, due to response to specific treatments could alter the prognostic impact of adverse disease features or genetic abnormalities,[173] we explore the relationship between methylation status and overall survival, in order to define the group of patients most likely to benefit from treatment with 5-azacytidine. First of all, we found no statistical differences in survival between patients with MDS and those with AML after AZA treatment. For this reason, we analyzed MDS and AML as a unique entity. Our data showed that baseline methylation status before treatment with AZA was associated with overall survival as an independent factor. Actually, the presence of ≥2 methylated genes was associated with shorter survival in these subset of patients treated with AZA. Noteworthy, we also analyzed the prognostic impact of the presence of 0, 1, 2 and ≥3 aberrantly methylated genes, showing that as the number of methylated genes increases, the survival decreases. This analysis revealed that those patients displaying the highest level of methylation (≥3 genes) had a very short survival. Hypermethylation of several genes such as *CDKN2B*, *HIC1*, *CDH1*, *ESR1* and *FHIT* genes were previously reported as associated with poor prognosis.[79,83,84,86-89] Therefore we also examined the relationship between the methylation status of any individual gene or gene combinations and the outcome. However, no correlation was found, likely due to sample size limitations.

Recently, *TET2* mutations have shown to predict an increased response to HMAs, especially when allele burden is above 10% and no *ASXL1* mutations are present. However, the presence of these mutations was not associated with differences in overall survival.[173,174] More importantly, the work form Bejar *et al*. also stated that no pattern of mutation was strongly associated with a lack of response to treatment to HMAs, not even those that confer a very poor prognosis, such as *RUNX1* and *TP53*. Moreover, the presence of *TP53*, *RUNX1*, *ASXL1*, *EZH2*, and *ETV6* mutations, usually associated with a poor outcome, were not found to be related to the prognosis in those MDS patients treated with HMAs.[173,174] Thus, mutational information alone should not be used as a basis for denying therapy with an HMA if treatment is indicated.[173] Since an increasing number of genetic and epigenetic abnormalities have been described in MDS and AML patients, the combination of genetic

and epigenetic scanning techniques might hold promise to select those patients who could benefit from HMAs therapy.

In summary, the biologic and molecular data presented herein, from genome-wide expression profiling, high-density microarray-based karyotyping, amplicon-based deep sequencing and methylation analysis in the study of myelodysplastic syndromes, have improved our knowledge of the complex biology underlying the pathogenesis of these disorders, trough the identification of genomic abnormalities that may have clinical relevance. The use of aCGH may be a used as a complementary tool to conventional cytogenetics in the study of MDS, mainly in those patients with poor-quality or few metaphases to be analyzed. In addition, this study shows the involvement of novel genes and gene pathways that could be considered as potential therapeutic targets in the future.

Concluding Remarks

1. The use of aCGH karyotyping in the routine evaluation of MDS patients could be used as a complementary technique to conventional cytogenetics, especially in those patients with either non-informative cytogenetic results or few metaphases available for karyotypic studies. Thus, in this subset of patients, aCGH allows the detection of genetic abnormalities enabling the prognostic stratification according to the IPSS-R, which could change the clinical management of this group of patients.

2. Genome-wide DNA copy number analysis allows the identification of cytogenetically cryptic chromosomal lesions in regions of known relevance in the pathogenesis of MDS, including deletions in *DNMT3A*, *TET2*, *SPARC*, *CUX1*, and *RUNX1* loci, as well as gains in the *U2AF1* locus. Further amplicon-based deep sequencing reveals that a low proportion of patients carry a simultaneous deletion and mutation in the investigated genes.

3. High-throughput copy number analysis allows the detection of chromothripsis involving chromosome 13 in three high-risk MDS patients. Chromothripsis patterns of genomic alteration are different among the three patients. However, some cancer-related genes such as *FLT3*, *FLT1*, and *XPO4* are commonly amplified, while *BRCA2* and *RB1* are commonly deleted. In addition, all three patients display complex karyotypes, *TP53* mutations, and show a poor outcome.

4. MDS progression from non-malignant bone marrow conditions towards AML seems to be an orchestrated process. Thus, genome-wide expression profiling during MDS evolution shows that several genes and gene pathways are commonly and progressively deregulated in the transition from non-malignant BM conditions through early and advanced MDS stages to AML.

    4.1. DNA damage response and checkpoint pathways, apoptosis, ribosome and translation-related pathways, as well as nucleosome and chromatin assembly are progressively up-regulated during the progression of the disease, even since early MDS stages. However, cell proliferation, self-renewal and differentiation arrest are up-regulated in the transition from advanced MDS to AML.

4.2. Immune response, cell adhesion and matrix metalloproteinases related genes are increasingly down-regulated during the evolution of the disease, showing a more repression in the transition from advanced MDS to AML.

5. The presence of two or more methylated genes in MDS and AML patients before treatment with 5-azacytidine is associated with a shorter overall survival. Baseline methylation does not correlate with clinical response. However, the presence of an adverse cytogenetics is associated with a worse response to treatment with 5-azacytidine.

Resumen en castellano

Tesis Doctoral

Caracterización Molecular de los Síndromes Mielodisplásicos (SMD):

Análisis de las alteraciones genómicas en el desarrollo de los SMD, la progresión a leucemia aguda mieloblástica y la respuesta al tratamiento con 5-Azacitidina

Directores:

Prof. Dr. Jesús M. Hernández Rivas
Dra. M. del Rocío Benito Sánchez

María Abáigar Alvarado

2015

"Cuando buscamos el tesoro, nos damos cuenta de que el camino es el propio tesoro"

Paulo Coelho

Introducción General

"Un viaje de miles de millas comienza con un solo paso"

Lao Tse

1. Los síndromes mielodisplásicos

1.1. Generalidades

Los síndromes mielodisplásicos (SMD) constituyen un grupo heterogéneo de enfermedades hematológicas que afectan a la célula madre hematopoyética mieloide. Se caracterizan por la presencia de alteraciones morfológicas, inmunofenotípicas, funcionales y genómicas de las células de las diferentes líneas hematopoyéticas mieloides, por la presencia de células inmaduras (blastos), por una hematopoyesis clonal e ineficaz en la médula ósea (MO) que da lugar a citopenias en sangre periférica (SP) (anemia, neutropenia y/o trombocitopenia en grado o combinaciones variables), a pesar de una médula ósea normo o hipercelular. .[1-3] La forma de presentación así como la evolución de los SMD es muy variable, desde formas indolentes con una esperanza de vida casi normal durante años a otras más agresivas que progresan rápidamente a leucemia aguda mieloblástica (LAM), aunque generalmente se asocian a mal pronóstico debido a que acortan la supervivencia.[4,5] Se manifiestan generalmente en personas mayores de 50 años y su incidencia aumenta con la edad, siendo la mediana de aparición de 70 años. Además, la enfermedad es más común en hombres que en mujeres. [3,6-11]

En el desarrollo de los SMD se han implicado varios factores de riesgo, como la edad, el sexo masculino, haber recibido un tratamiento anterior de quimioterapia o radioterapia, los agentes inmunosupresores, las infecciones víricas, el tabaco, el alcohol, la exposición prolongada a radiación ionizante, al benceno, y otras exposiciones ambientales u ocupacionales.[6,7,12,13] Aunque en algunas ocasiones se ha observado que existe una predisposición genética, como ocurre en pacientes con anemia de Fanconi, síndrome de Schwachman-Diamond, neutropenia congénita severa, disqueratosis congénita o anemia de Diamond-Blackfan.[6,7,11,13]

1.2. Clasificación diagnóstica, citogenética y pronóstica

El diagnóstico y la clasificación a día de hoy de los SMD se realiza en base a las alteraciones morfológicas y citogenéticas que presentan las células de la médula ósea.[11,14,25] La clasificación diagnóstica más reciente de la Organización Mundial de la Salud (OMS) se basa en el porcentaje de blastos en médula ósea y sangre periférica, el tipo y grado de displasia mieloide y la presencia de sideroblastos en forma de anillo.[14] En el año 2008, la OMS ha establecido las siguientes entidades para los SMD: citopenia refractaria con displasia unilínea (CRDU) que incluye los subtipos de anemia (AR), neutropenia (NR) y trombocitopenia refractaria (TR), AR con sideroblastos en anillo (AR-SA), citopenia refractaria con displasia multilínea (CRDM), anemia refractaria con exceso de blastos tipo1 y tipo2 (AREB-1 y AREB-2), síndrome 5q- (aportación citogenética) y SMD inclasificable (MDS-U). La presencia de más del 20% de blastos en médula ósea se define como LAM.[14]

Con respecto a la clasificación citogenética debemos saber que las técnicas de citogenética convencional (CC) han demostrado que un 40-60% de los pacientes con SMD presentan alteraciones cromosómicas al momento del diagnóstico.[2,3,7,13,16,18] Estas alteraciones cromosómicas pueden afectar a uno o a varios cromosomas.[19] Las alteraciones más frecuentes son la pérdida total o parcial del cromosoma 5 (-5/del(5q)), así como del cromosoma 7 (-7/del(7q)), la trisomía del cromosoma 8, las pérdidas del brazo largo del cromosoma 20 (20q-), y la monosomía del cromosoma Y.[2,7,13,16,20] Se ha visto que la frecuencia de alteraciones citogenéticas aumenta con la severidad de la enfermedad. Además, la adquisición de estas anomalías a lo largo del tiempo se ha asociado con un mayor de riesgo de transformación leucémica y una supervivencia más corta.[19,23,24] La presencia de alteraciones citogenéticas y la proporción de blastos en la médula ósea son, hasta la fecha, los indicadores pronósticos más importantes para la supervivencia, el riesgo de transformación leucémica y, en algunos casos, para la selección del tipo de tratamiento.[2,19,22]

Debido a la gran heterogeneidad pronóstica dentro de un mismo subtipo morfológico y el reconocimiento de que el cariotipo de las células de la médula ósea puede predecir la evolución de la enfermedad, es necesario disponer de sistemas que ayuden a establecer el pronóstico de cada enfermo para valorar la necesidad y el tipo de tratamiento que deben

recibir.[4,9] Por esta razón se han desarrollado diversas clasificaciones pronósticas, como el Índice Pronóstico Internacional (IPSS) y el Índice Pronóstico basado en la clasificación OMS (WPSS), que han sido los sistemas de clasificación pronóstica más empleados hasta la fecha, y actualmente también el IPSS revisado (IPSS-R), una actualización del IPSS.[5,22,25,30,31] En todos ellos, la proporción de blastos en médula ósea, los hallazgos citogenéticos, el número y grado de las citopenias en sangre periférica y la dependencia transfusional son los indicadores pronósticos más importantes para la supervivencia y el riesgo de transformación a LAM. Combinando estos factores, todos estos sistemas de estratificación han establecido diferentes grupos de riesgo, y diferencian los llamados SMD de bajo riesgo de los SMD de alto riesgo.[5,22,30] Estos grupos de pacientes se diferencian en la esperanza de vida, que es inferior en los SMD de alto riesgo, entre 4 meses y 1 año para este grupo de pacientes, y en la probabilidad de trasformación a LAM, siendo superior en los SMD de alto riesgo, en los que a 1,5 años el 25% de pacientes evolucionan a LAM.[5,22]

2. Alteraciones biológicas y moleculares de los SMD

Poco a poco se van perfilando las bases moleculares que determinan la heterogeneidad de los SMD, pero aún queda mucho por avanzar. Entre los mecanismos que pueden contribuir al desarrollo y progresión de los SMD se encuentran: las alteraciones citogenéticas, las disomías uniparentales (UPD), la haploinsuficiencia, las mutaciones, las alteraciones epigenéticas, una apoptosis anormal, una desregulación del sistema inmune y de las vías de transducción de señales y las alteraciones en el microambiente medular.[11,33] Además, en los SMD, al igual que en otros tipos de cáncer, un único evento genético no es suficiente para que una célula pueda derivar en el desarrollo de un cáncer, sino que la adquisición gradual de varias alteraciones son necesarias para el desarrollo y progresión de esta enfermedad.[37]

2.1. Alteraciones en el número de copias y anomalías crípticas

El estudio de las alteraciones estructurales del genoma ha sido de gran interés ya que a través de la adquisición de reordenamiento genómicos como deleciones, amplificaciones, translocaciones e inversiones, que afectan a determinadas regiones del genoma, pueden activar o inactivar determinados genes que pueden promover el desarrollo y la progresión tumoral.[38-41]

Los avances en las técnicas genómicas de alta resolución, como los arrays de CGH (aCGH) o de SNPs (SNP-A), y su aplicación en el estudio de las neoplasias hematológicas, como en los SMD, ha permitido la identificación de variaciones en el número de copias (ganancias y pérdidas), con una mayor resolución que la CC, a excepción de las translocaciones e inversiones o la presencia de clones minoritarios (<25-30% de células alteradas). También han permitido la detección de disomías uniparentales (UPD). Estos datos sugieren que el los estudios de aCGH y SNP-A podrían tener una gran utilidad a nivel clínico.[18,41-57] Por un lado, han permitido la detección de alteraciones en un 50% de los casos sin mitosis.[46-48] Pero además, han hecho posible la identificación de alteraciones crípticas (<5 Mb), no detectables mediante las técnicas de CC, en un 70-90% de los SMD y en >50% de los SMD con cariotipo normal.[18,41-57] Estas alteraciones crípticas afectaban a regiones importantes en la patogénesis de los SMD, siendo las más frecuentes las deleciones de los genes *TET2*, *RUNX1*, *ETV6*, *TP53*, *NF1* y *DNMT3A*.[42,44,49,50] Por tanto, los estudios de aCGH y SNP-A han permitido una mejor caracterización de las alteraciones cromosómicas que presentan los SMD y han proporcionado un nuevo punto de partida para el estudio de genes concretos mediante secuenciación, para la identificación de mutaciones genéticas. De este modo han supuesto un gran avance en el conocimiento de algunas de las rutas alteradas en los SMD que podrían llevar al desarrollo de terapias dirigidas más eficaces.[44,47,48]

2.2. Inestabilidad genómica: Cromotripsis

Tradicionalmente el cáncer se produce por la adquisición progresiva de mutaciones y alteraciones que dan lugar a la progresión tumoral. [37-40] Recientemente se ha descrito un

fenómeno denominado "cromotripsis" (del griego: "chromo" de cromosoma y "thripsis" de romperse en pedazos), en el cuál se producen decenas e incluso cientos de reordenamientos cromosómicos en una catástrofe celular única, que generalmente afecta a un cromosoma o región cromosómica, y en algunos casos, a varios cromosomas.[37-40,58-61] La firma de cromotripsis comprende múltiples y complejos reordenamientos que pueden observarse a modo de variaciones en el número de copias que oscilan entre una, dos y, ocasionalmente, 3 estados.[37-39,58,60,61] Por ejemplo, una región normal con dos copias puede estar seguida por una región con una copia, seguida por otra región normal, y ésta seguida de otra región con 3 copias.[61] La explicación más sencilla para este fenómeno es que, en algún momento durante el desarrollo del cáncer, uno o más cromosomas son completamente fragmentados y reensamblados de forma aleatoria y errónea por la maquinaria de reparación del ADN.[38-40,61] Es decir, algunos de estos fragmentos son reensamblados sin ningún orden aparente y en cualquier orientación generando un cromosoma derivativo. Otros fragmentos no son incorporados a este nuevo cromosoma y se pierden, dando lugar a pérdidas de material genético. Mientras que otros son ensamblados formando unas estructuras circulares extra-cromosómicas, que citogenéticamente se observan como "dobles minutos", que en los siguientes ciclos de división de la célula pueden ser amplificados, dando lugar a ganancias en el número de copias.[39,40,59]

El fenómeno de la cromotripsis se ha observado en un 2-3% de los cánceres en general, incluyendo las neoplasias hematológicas y los tumores sólidos. Sin embargo, se ha descrito una incidencia muy superior, del 25%, en los cánceres óseos.[37-39,58-61,64-71] Este fenómeno se ha asociado con las formas más agresivas de un tumor, un pronóstico adverso y mutaciones de *TP53*.[37,38,40,59-61,63]

2.3. Alteraciones en la metilación del ADN

Los SMD y LAM también se caracterizan por presentar alteraciones en la metilación del ADN. La metilación del DNA es un mecanismo epigenético que consiste en la adición de grupos metilo a las citosinas de las islas CpG de las regiones promotoras de los genes, constituyendo un proceso clave en la estabilización del genoma, la remodelación de la cromatina y la

regulación de la transcripción.[72-77] En particular, la hipermetilación de las islas CpG se ha asociado con el silenciamiento génico.[72,75,77,79-82] En los SMD y LAM se han descrito un gran número de genes que están hipermetilados. Algunos están relacionados con el control del ciclo celular (*CDKN1B*, *CDKN2B*, *HIC1* y *FHIT*), adhesión celular (*IGSF4*, *CDH1* y *CDH13*), regulación de la apoptosis (*DAPK1*), crecimiento celular (*ESR1*) y reparación del ADN (*MLH1*).[77,79,80,83-89] Además, la hipermetilación de algunos de estos genes como *CDKN2B*, *HIC1*, *CDH1*, *ESR1* y *FHIT* se ha asociado con un pronóstico adverso y un elevado riesgo de transformación a LAM.[78-80,83,84,86-90] Por otro lado, se han encontrado mutaciones en genes reguladores de la metilación (*DNMT3A*, *TET2*, *IDH1*, *IDH2*).[2,34,35,36,92,93,94,101] Al contrario que las alteraciones citogenéticas, el silenciamiento génico causando por una metilación aberrante del ADN es un proceso reversible.[74,77,91] Por esta razón, la metilación aberrante constituye una atractiva diana terapéutica en los SMD y LAM. De hecho, actualmente se están utilizando agentes hipometilantes, como la 5-azacitidina, que se ha observado que producen una progresiva reducción de la metilación del ADN y, consecuentemente, de la reactivación de la expresión génica, para el tratamiento de los pacientes con SMD y LAM.[72,74,77,91]

2.4. Mutaciones génicas

Se han producido grandes avances en el conocimiento de los mecanismos moleculares implicados en la patogénesis de los SMD en los últimos diez años gracias a la aplicación de las técnicas de secuenciación masiva, como la secuenciación completa del genoma y exoma, o de grandes paneles de genes.[2,31,36] De este modo se han identificado mutaciones en más de 40 genes relacionados con mecanismos implicados en la patogénesis de los SMD y otras neoplasias mieloides.[26,34-36,92] Estos genes están involucrados en las vías de señalización (*JAK2*, *KRAS*, *NRAS*, *CBL*), de regulación transcripcional (*RUNX1*, *BCOR*, *BCORL1*, *ETV6*, *EVI1*, *GATA2*, *TP53*), regulación epigenética (*DNMT3A*, *TET2*, *IDH1*/*2*, *EZH2*, *ASXL1*, *ATRX*, *KDM6A*), en la maquinaria de "splicing" del ARN (*SF3B1*, *U2AF1*, *SRSF2*, *ZRSR2*, *PRPF8*), replicación del ADN (*SETBP1*) y en el complejo de las cohesinas (*STAG2*, *RAD21*, *SMC1A*, *SMC3*). [2,34-36,93-101]
En el momento actual se estima que un 80-90% de los pacientes con SMD presentan al menos una mutación en algún gen, incluso aquellos que tienen un cariotipo normal.[35,36]

Algunas de estas alteraciones han demostrado tener valor diagnóstico y pronóstico. Se han descrito, por ejemplo, mutaciones en *TET2* en el 20-30% de los SMD, en *SF3B1* en el 15-30% de los SMD y en más del 75% de los SMD con sideroblastos en anillo. Pero además, se ha visto que mutaciones en los genes *ASXL1*, *ETV6*, *EZH2*, *RUNX1* y *TP53*, presentes en menor proporción (<10% cada una), predicen una supervivencia más corta y un alto riesgo a desarrollar leucemia en los pacientes con SMD.[2,26,34-36,95,97,99,100,102,103]

2.5. Alteraciones a nivel de la expresión génica

El análisis de los perfiles de expresión génica (PEG) mediante microarrays constituye la técnica más comúnmente utilizada para estudiar la expresión de miles de genes y una herramienta para el estudio de los distintos tipos de cáncer, ya que permiten clasificar las diversas formas de cáncer e incluso determinar la presencia de nuevos subtipos.[104-108] Además, permiten identificar nuevos marcadores para un diagnóstico más preciso, marcadores de la progresión de la enfermedad y/o de posibles dianas terapéuticas.[106,107,109,111-113] En los SMD se han realizado una gran variedad estudios de expresión génica, principalmente en poblaciones celulares separadas como CD34+, y predominantemente en SMD con del(5q). Estos estudios han permitido establecer diferencias entre los SMD y los individuos sanos, entre subtipos específicos de SMD, y entre subclases citogenéticas, permitiendo profundizar en la biología de estas enfermedades.[106,109-111,114-121] Por ejemplo, los pacientes con del(5q), -7/del(7q) o trisomía 8 presentan un PEG distinto, poniendo de manifiesto el efecto de dosis génica.[116,119-121] Así, los SMD con del(5q) presentan una desregulación de los genes ribosomales, los SMD con pérdidas del cromosoma 7 se caracterizan por la activación de la vía JNK, y los SMD con trisomía 8 por una desregulación de la respuesta inmune.[116,119-121] Del mismo modo, las AR-SA se caracterizan por una desregulación de genes relacionados con el metabolismo del hierro y mitocondria.[117,118,122] Por lo tanto, estos estudios han permitido profundizar en la patofisiología de los SMD, revelando que éstos presentan trastornos en la apoptosis, proliferación y diferenciación celular y del sistema sistema inmune.[109-111,114-121]

3. Tratamiento de los SMD: 5-Azacitidina

El tratamiento actual de los SMD es muy variable y depende del subtipo de SMD y de la severidad de la enfermedad (riesgo del paciente, estratificado según estos sistemas pronósticos), así como de la edad y el estado general del paciente.[26,123] El objetivo del tratamiento en los pacientes con SMD de bajo riesgo es mejorar las citopenias, la sintomatología y la calidad de vida, mientras que en los pacientes con SMD de alto riesgo, el objetivo es modificar el curso natural de la enfermedad, prolongando la supervivencia y reduciendo el riesgo de transformación a LAM.[7,26]

En los SMD de bajo riesgo el tratamiento consiste en transfusiones de sangre o plaquetas, agentes estimulantes de la eritropoyesis y granulopoyesis, antibióticos, análogos de la trombopoyetina y quelantes del hierro.[7,26,123] Para los SMD de alto riesgo, el único tratamiento con finalidad curativa es el trasplante de médula ósea.[7,26,123] Sin embargo, la edad avanzada de muchos pacientes, la toxicidad de dicho procedimiento y/o el no disponer de un donante compatible limitan su empleo, siendo muy pocos los pacientes (<5%) que pueden beneficiarse de él.[7,26] Por esta razón se recurre al empleo de otras medidas terapéuticas, sin capacidad curativa, como la quimioterapia. En los últimos años se han desarrollado nuevos fármacos para el tratamiento de los SMD, tales como la lenalidomida (agente de gran efectividad en los pacientes con síndrome 5q-) y los agentes hipometilantes (5-azacitidina y decitabina).[7,26,123,124] Estos fármacos han demostrado ser eficaces en corregir las citopenias, logrando remisiones citogenéticas y la reducción de la proporción de blastos en la MO.[26] De hecho, están recomendados, por la mayor parte de las guías internacionales, como primera opción de tratamiento en los pacientes con SMD de bajo riesgo y del(5q), y alto riesgo, respectivamente.[7,126,130]

La 5-azacitidina (5-AZA) es un agente que inhibe la función de las DNA-metiltransferasas, lo que resulta en una progresiva reducción de la metilación del ADN y, consecuentemente, de la reactivación de la expresión génica.[72,74,77,91] La experiencia clínica y los datos publicados han demostrado que la 5-AZA es un tratamiento eficaz para los SMD y LAM, ya que un 50-60% de estos pacientes alcanzan una respuesta, mostrando una mejoría de las cifras en SP, una reducción de los requerimientos transfusionales, así como una mejoría en la calidad de

vida. Además, ha demostrado retrasar la progresión a LAM y prolongar la supervivencia.[7,72,74,78,124,126-128,130] A pesar de su eficacia, algunos pacientes siguen siendo refractarios a estas terapias, y en el momento actual no existen factores predictivos de respuesta a estos agentes.

A pesar de los avances en el conocimiento sobre los mecanismos implicados en la patogénesis de los SMD en estos últimos años, todavía quedan cuestiones por resolver, como la utilidad clínica de los arrays genómicos (aCGH) en el estudio de los SMD con un número insuficiente de mitosis, la evaluación de la cromotripsis, así como el valor diagnóstico y/o pronóstico de las anomalías crípticas detectadas por aCGH. Además, la dinámica de los cambios en la expresión génica en los pacientes con SMD no ha sido estudiada en profundidad, por lo que el análisis del PEG podría ser una herramienta valiosa para una mejor comprensión de la biología de estas enfermedades. Por último, ya que la 5-azacitidina es uno de los fármacos más ampliamente empleados en el tratamiento de los SMD, sería de gran interés el poder definir posibles factores predictivos de respuesta a este fármaco.

Hipótesis

Los síndromes mielodisplásicos constituyen una enfermedad muy heterogénea debido a su compleja fisiopatología. Un diagnóstico preciso resulta esencial para la comprensión del comportamiento clínico y biológico de los SMD. Por lo tanto, resulta necesario complementar los datos clínicos y los estudios de morfología y citogenética con otros marcadores que definan la biología de los SMD, con la intención de predecir con más precisión la evolución de la enfermedad y diseñar nuevas estrategias terapéuticas más dirigidas y específicas.

El diagnóstico y clasificación de los SMD se establece de acuerdo a las alteraciones morfológicas y citogenéticas que presentan las células de la MO, combinadas con determinados parámetros biológicos. La citogenética convencional (CC) sigue siendo el estándar oro para la identificación de las alteraciones cromosómicas en los SMD y todavía desempeña un papel fundamental en la evaluación pronóstica de estos pacientes. Sin embargo, en ocasiones resulta difícil establecer una clasificación pronóstica correcta, especialmente en los casos que presentan pocas alteraciones morfológicas, con escasa población de blastos y sin alteraciones citogenéticas, como es el caso de los SMD con cariotipo normal o no informativo, debido a la ausencia de mitosis. De hecho, aproximadamente un 40-60% de los SMD presentan un cariotipo normal y en un 10-15% de los SMD no se obtienen suficientes metafases o son de mala calidad, lo que no excluye que presenten defectos genéticos que no hayan sido detectados por las técnicas de CC. Además, el empleo de los arrays genómicos (aCGH, SNP-A) en el estudio de los SMD ha hecho posible la detección de nuevas alteraciones y una mejor caracterización de las alteraciones cromosómicas que presentan estas enfermedades. Por esta razón, las técnicas de CC no son suficientes para el estudio en profundidad de estas enfermedades.

En la era de las técnicas genómicas de alta resolución, las mutaciones representan un nuevo parámetro a tener en cuenta en el diagnóstico y evaluación pronóstica de los SMD, ya que un 70-90% de estos pacientes presentan al menos una mutación oncogénica, incluso aquellos pacientes que tienen un cariotipo normal. Se han detectado mutaciones recurrentes en más de 40 genes en los SMD, siendo *SF3B1*, *TET2*, *SRSF2*, *ASXL1*, *DNMT3A* y *RUNX1* (>10%) los genes más frecuentemente mutados. Algunas de estas mutaciones se han asociado con subtipos concretos de SMD mientras que otras han demostrado tener un

importante valor pronóstico, ya que se han asociado con una supervivencia más corta y un mayor riesgo de transformación a LAM. Además, algunos de estos genes como *TET2*, *RUNX1*, y *TP53*, también presentan deleciones en el alelo no mutado.

Como podemos observar, los estudios mediante arrays genómicos y la secuenciación masiva han identificado un gran número de alteraciones en genes implicados en diversos procesos en los SMD y otras neoplasias hematológicas. La gran diversidad de estos defectos genéticos y sus combinaciones ponen de manifiesto la gran heterogeneidad observada en los SMD. Por esta razón, la combinación de ambas metodologías de alta resolución nos permitirá tener una visión global, y a la vez más detallada, de los defectos genéticos que se producen en los SMD y profundizar en el conocimiento de la biología de estas enfermedades.

Los estudios del perfil expresión génica (PEG) mediante microarrays han demostrado que los diferentes subtipos de SMD presentan un PEG diferente, y han permitido la identificación de algunos de los genes y mecanismos implicados en estas enfermedades. Nuestro grupo ha demostrado que los estadios más precoces de los SMD (SMD de bajo riesgo), se caracterizan por presentar trastornos de la diferenciación celular y una elevada apoptosis intramedular, lo que explicaría, al menos en parte, la hematopoyesis ineficaz y las citopenias periféricas características de los SMD. En estadios más avanzados de la enfermedad (SMD de alto riesgo) se producirían nuevas alteraciones genéticas y epigenéticas que darían lugar a un aumento de la proliferación celular y una reducción relativa de la apoptosis, lo que conduciría a la acumulación de blastos en la MO y transformación a LAM. A pesar de los avances que se han producido en el conocimiento de la biología de los SMD y LAM gracias a la aplicación de los microarrays de expresión, así como de las técnicas de alta resolución descritas anteriormente, los mecanismos moleculares que podrían determinar la progresión de SMD a LAM son en gran parte desconocidos. Por ello, el estudio del PEG en los SMD y LAM puede permitirnos profundizar en los mecanismos de su trasformación leucémica.

En los SMD y LAM se han descrito un gran número de genes que están hipermetilados y se han observado también mutaciones en genes reguladores de la metilación del ADN. Por esta razón, el empleo de agentes hipometilantes, como la 5-azacitidina (5-AZA), parece ser una buena opción terapéutica para estas hemopatías. De hecho, se ha convertido en el tratamiento de primera línea para los SMD de alto riesgo y LAM con pocos blastos,

demostrando ser un tratamiento eficaz para estas enfermedades. Sin embargo, aún no está claro si el número o el tipo de genes hipermetilados puede estar relacionado con la respuesta a estos agentes. Por lo tanto, la valoración del estado de metilación en los SMD y LAM tratados con 5-AZA podría seleccionar grupos de enfermos en los que este fármaco pudiera tener mayor o menor eficacia. Además, se ha demostrado que la existencia de alteraciones no complejas en el cromosoma 7, asociadas a un pronóstico adverso con las terapias tradicionales, define un grupo de enfermos respondedores a 5-AZA (40%) que presentan, además, una supervivencia más larga. Por ello, consideramos que el estudio de la metilación antes del tratamiento con 5-azacitidina debe complementarse con datos genómicos para definir aquellos factores que puedan estar relacionados con la respuesta a este fármaco.

Objetivos

Objetivo general:

Identificar nuevos marcadores genéticos que puedan contribuir a un mejor diagnóstico y pronóstico en la evaluación de los pacientes con SMD, y que podrían estar relacionado con los procesos biológicos implicados en la patogénesis de los SMD y en la progresión de los SMD a LAM, así como marcadores moleculares de la respuesta al tratamiento con agentes hipometilantes.

Objetivos específicos:

- Identificar nuevas alteraciones genéticas en los pacientes con SMD mediante arrays genómicos (aCGH).

- Evaluar la aplicación de las técnicas de análisis genómico masivo, tales como los arrays genómicos (aCGH) y la secuenciación masiva (NGS) en el diagnóstico de los SMD como herramienta complementaria a la citogenética convencional (CC).

- Analizar el perfil de expresión génico (PEG) en los SMD, LAM y enfermos sin hemopatías malignas con el fin de identificar aquellos genes y mecanismos que puedan estar implicados en el desarrollo de los SMD y su progresión a LAM.

- Evaluar la influencia de un estado de metilación aberrante y las alteraciones citogenéticas en la supervivencia y respuesta a 5-azacitidina

Resultados - Resúmenes

Esta sección incluye un resumen del trabajo experimental realizado en esta tesis doctoral, incluyendo Material y Métodos, Resultados y Discusión, y está dividida en tres capítulos:

Capítulo 1. M Abáigar, C Robledo, R Benito, F Ramos, M Díez-Campelo, L Hermosín, J Sánchez-del-Real, J María Alonso, R Cuello, M Megido, JN Rodríguez, G Martín-Núñez, C Aguilar, M Vargas, AA Martín, JL García, A Kohlmann, MC del Cañizo, JM Hernández-Rivas. ***Chromothripsis is a recurrent genomic abnormality in high-risk myelodysplastic syndromes***. Genes Chromosomes Cancer. En segunda revisión.

Capítulo 2. M Abáigar, S Aibar, R Benito, M Díez-Campelo, F Ramos, E Lumbreras, FJ Campos-Laborie, M Megido, I Recio, L Hermosín, J Sánchez-del-Real, C Olivier, R Cuello, L Zamora, K Mills, MC del Cañizo, J De Las Rivas, JM Hernández-Rivas. ***Common and progressive gene expression changes in the progression of myelodysplastic syndromes to acute myeloid leukemia***. Haematologica. Enviado.

Capítulo 3. M Abáigar, F Ramos, R Benito, M Díez-Campelo, J Sánchez-del-Real, L Hermosín, JN Rodríguez, C Aguilar, I Recio, JM Alonso, N de las Heras, M Megido, M Fuertes, MC del Cañizo, JM Hernández-Rivas. ***Prognostic impact of the number of methylated genes in myelodysplastic syndromes and acute myeloid leukemias treated with azacytidine***. Ann Hematol. 2013 Nov;92(11):1543-52. doi: 10.1007/s00277-013-1799-9. PubMed PMID: 23740492.

Todos ellos han sido desarrollados para lograr el objetivo general de este trabajo y dar el título de esta tesis doctoral: "Caracterización molecular de los síndromes mielodisplásicos (SMD): Análisis de las alteraciones genómicas en el desarrollo de los SMD, la progresión a leucemia aguda mieloblástica y la respuesta al tratamiento con 5-azacitidina".

Resumen del capítulo 1:

Chromothripsis is a recurrent genomic abnormality in high-risk myelodysplastic syndromes

María Abáigar, Cristina Robledo, Rocío Benito, Fernando Ramos, María Díez-Campelo, Lourdes Hermosín, Javier Sánchez-del-Real, Jose M Alonso, Rebeca Cuello, Marta Megido, Juan N Rodríguez, Guillermo Martín-Núñez, Carlos Aguilar, Manuel Vargas, Ana A Martín, Juan L García, Alexander Kohlmann, M Consuelo del Cañizo, Jesús M Hernández-Rivas

Genes Chromosomes Cancer. (En segunda revisión)

La versión original, incluyendo tablas y figuras está disponible en las páginas 40 - 60 y el material suplementario en las páginas 189 - 195.

Introducción y objetivos

El estudio de las alteraciones estructurales del genoma ha sido de gran interés ya que, a través de la adquisición de reordenamientos genómicos, se pueden activar o inactivar determinados genes que pueden promover la progresión tumoral. Un descubrimiento reciente es el fenómeno denominado "cromotripsis", que consiste en la generación de entre 10 y cientos de reordenamientos cromosómicos de manera simultánea en un cromosoma o región cromosómica. Este fenómeno ha sido descrito en un 2-3% de los cánceres en general, incluyendo diversas neoplasias hematológicas como el mieloma múltiple (1,3%) y la leucemia linfática crónica (10%), aunque en los osteosarcomas se ha observado con una incidencia más elvada (25%), y se ha asociado con mal pronóstico y mutaciones en *TP53*.

Por otro lado, la presencia de alteraciones citogenéticas junto con la proporción de blastos en la MO son los parámetros con mayor valor pronóstico para los pacientes con SMD y LMMC, asignándolos a diferentes categorías de riesgo en función de la supervivencia y probabilidad de evolución a LAM. Sin embargo, en ocasiones resulta difícil establecer una clasificación pronóstica correcta, ya que aproximadamente un 40-60% de los SMD y LMMC tienen un cariotipo normal y en un 10-15% el cariotipo es no informativo por citogenética convencional.

Por otro lado, el empleo de los arrays genómicos (aCGH, SNP-A) en el estudio de los SMD ha permitido la detección de alteraciones crípticas, como las deleciones de los genes *TET2* y *RUNX1*, en aquellos pacientes con un cariotipo normal. Recientemente, los estudios mediante secuenciación masiva han identificado una gran cantidad de genes que presentan mutaciones y que están implicados en la patogénesis de los SMD.

Por ello nos propusimos estudiar la presencia de nuevas alteraciones genéticas en los SMD y LMMC mediante un estudio integrado combinando las técnicas de aCGH y secuenciación masiva (NGS) en una serie de pacientes con SMD y LMMC.

Pacientes y métodos

Se analizaron un total de 301 pacientes al momento del diagnóstico clasificados según criterios OMS 2008 como SMD (n=240) y SMD/SMP (n=61). La mediana de edad era de 77 años (11-93) con predominio de varones (63,5%). Todos los casos fueron estudiados mediante CC y FISH de manera rutinaria para la identificación de alteraciones citogenéticas. En un 13% de los pacientes no se obtuvieron mitosis, el 15% presentaron un cariotipo patológico, mientras que el 72%, la mayor parte de los pacientes incluidos en el estudio, tenían un cariotipo normal. Este último grupo fue dividido en tres categorías en función al número de metafases de buena calidad que fueron analizadas: ≥20 metafases (n=164), entre 11-19 metafases (n=38), y ≤10 metafases (n=14).

En todos los pacientes se realizaron estudios de aCGH para identificar las variaciones en el número de copias, ganancias y pérdidas, con una mayor resolución que la CC, mediante el aCGH *Human CGH 12X135K Whole-Genome Tiling Array v3.0* (Roche-NimbleGen). Todas odas las regiones ganadas o perdidas fueron contrastadas con la información existente en la base de datos *Database of Genomic Variants* (http://dgv.tcag.ca/) para eliminar aquellas regiones descritas como variaciones en el número de copias (CNV) en la población normal. Además, los resultados fueron validados mediante FISH o un aCGH distinto (Agilent).

Además, en una serie de casos seleccionados se analizaron las mutaciones de los genes *DNMT3A*, *TET2*, *RUNX1*, *BCOR* y *TP53* mediante secuenciación masiva de amplicones (NGS) con el sistema GS-Junior (454-Roche). Concretamente se estudiaron la regiones codificantes completas de los genes *TET2*, *RUNX1* y *BCOR*, mientras que sólo los exones 7-23 de *DNMT3A* y 4-11 de *TP53*. Todas las variantes descritas como polimorfismos en la bases de datos *dbSNP* (http://www.ncbi.nlm.nih.gov/SNP/) y las variantes intrónicas fueron excluidas.

Resultados

El análisis de los perfiles de aCGH mostró la presencia de reordenamientos masivos que cumplían los criterios de cromotripsis en 3/240 (1.2%) SMD. En los tres casos, esta alteración se observó afectando a un cromosoma, no a varios, y en este estudio en particular, al cromosoma 13 exclusivamente. Estos 3 pacientes estaban diagnosticados de AREB (3/40,

7.5%) y tenían un cariotipo complejo por aCGH (3/17, 17,6%), con una mediana de 16 regiones alteradas (un rango de 11-25). Cabe destacar que los 3 pacientes presentaron pérdidas a nivel de 5q23.2-q35.3 y dos de ellos, además, tenías pérdidas en 7q22.3-q36.3 y 15q11.1-21.2. Como corresponde a la cromotripsis, no se observó un patrón común de alteración a nivel global. Sin embargo, pudimos observar que determinadas regiones aparecían ganadas o perdidas en común en los 3 pacientes, afectando a genes relacionados con cáncer como *XPO4*, *FLT1*, *FLT3*, *BRCA2* y *RB1*.

Como ya se ha indicado, la cromotripsis se ha asociado con la presencia de mutaciones en *TP53*. Por esta razón, estudiamos la presencia de mutaciones en este gen en estos tres pacientes mediante secuenciación masiva de amplicones. Los tres enfermos con SMD y cromotrisis presentaron mutaciones en *TP53*, todas ellas localizadas en el exón 5, que corresponde al dominio de unión al ADN. Cabe señalar, que la supervivencia de los tres casos fue inferior a 1 año.

Por otro lado, el análisis de la serie global reveló un total de 285 alteraciones por aCGH en 71 pacientes (23,6%). Las alteraciones más frecuentes fueron: del(5q) (35%), del(20q) (18%), del(7q) (14%), -Y (14%), trisomía 8 (10%) y del(4q24) (10%). Además, mediante en análisis de aCGH se identificaron alteraciones crípticas (≤5 Mb), por debajo del límite de detección de la CC, en 23 casos. Estas alteraciones afectaban a genes importantes en la patogénesis de los SMD, siendo las más frecuentes las deleciones en 2p23.3 (*DNMT3A*, n=2), 4q24 (*TET2*, n=7), 21q22 (*RUNX1*, n=5), y Xp11.4 (*BCOR*, n=2). Estos genes contenidos en las regiones delecionadas fueron estudiados por secuenciación masiva de amplicones. Además, también se analizaron las mutaciones de *TP53* en aquellos casos que presentaron deleciones a nivel de 17p13. El estudio de secuenciación masiva reveló que sólo 1 paciente con deleción de *DNMT3A*, 3 pacientes con pérdida de *TET2* y 4 enfermos con deleción de *TP53* presentaron mutaciones en el otro alelo. El resto de los genes estudiados no presentaron ninguna mutación en los casos analizados.

Debido a que se disponía de la información de la citogenética convencional de todos los casos, los resultados del análisis de aCGH fueron comparados con los de la CC. Así se observó que el 9,3% de los pacientes con cariotipo normal y el 30% de los pacientes sin mitosis presentaron alteraciones. El análisis de los casos con cariotipo alterado corroboró los

resultados en el 86,7% de los casos, mostrando una gran concordancia entre los datos de CC y aCGH. Además, el estudio de los resultados del análisis de aCGH de los 216 casos con citogenética normal por CC, confirmó que el 92% de los pacientes con ≥20 metafases analizadas, el 89% de aquellos casos con 11-19 metafases estudiadas, y sólo el 78% de los casos con ≤10 metafases no presentaron alteraciones por aCGH.

Conclusiones

El estudio de aCGH identificó la presencia de cromotripsis como un fenómeno novedoso y recurrente, aunque poco frecuente, en tres SMD de alto riesgo. En los tres casos afectó exclusivamente al cromosoma 13, con genes como *XPO4*, *FLT1* y *FLT3*, comúnmente amplificados, y *BRCA2* y *RB1*, como comúnmente delecionados. Los tres casos con SMD y cromotripsis presentaron además un cariotipo complejo, mutaciones en *TP53* y una supervivencia muy corta, todas ellas características que se han asociado con el fenómeno de cromotripsis.

Por otro lado, el estudio integrado de arrays genómicos y secuenciación masiva permitieron la identificación de deleciones crípticas en regiones genómicas donde se localizan genes relacionados con los SMD, como *DNMT3A* (2p23.3; n=2.8%), *TET2* (4q24; n=10%) *TP53* (17p13; n=8.5%), *RUNX1* (21q22; n=7%), and *BCOR* (Xp11.4; n=2.8%), así como mutaciones en el otro alelo en *DNMT3A* (n=1), *TET2* (n=3), y *TP53* (n=4). Estas alteraciones crípticas se observaron principalmente en pacientes con cariotipo normal (45%) o no informativo (15%) por citogenética convencional, a excepción de aquellos pacientes con una deleción y mutación simultáneas de *TP53* que tenían un cariotipo complejo.

Además, el presente estudio demuestra que los arrays genómicos pueden utilizarse como una técnica complementaria a la citogenética convencional en la evaluación de los pacientes con SMD y LMMC, principalmente en los pacientes sin crecimiento o con un cariotipo normal y ≤20 metafases analizables.

Resumen del capítulo 2:

Common and progressive gene expression changes in the progression of myelodysplastic syndromes to acute myeloid leukemia

María Abáigar, Sara Aibar, Rocío Benito, María Díez-Campelo, Fernando Ramos, Eva Lumbreras, Franscisco J Campos-Laborie, Marta Megido, Isabel Recio, Lourdes Hermosín, Javier Sánchez-del-Real, Carmen Olivier, Rebeca Cuello, Lurdes Zamora, Ken Mills, M Consuelo del Cañizo, Javier De Las Rivas, Jesús M Hernández-Rivas

Haematologica. Submitted

La versión original, incluyendo tablas y figuras está disponible en las páginas 61 - 75 y el material suplementario en las páginas 197 - 199.

Introducción y objetivos

Los síndromes mielodisplásicos (SMD) representan un grupo heterogéneo de enfermedades clonales de la células madre hematopoyéticas caracterizadas por una diferenciación y maduración anormal de las células mieloides, inestabilidad genética y un elevado riesgo de progresión a LAM. Ambas hemopatías, tanto los SMD como las LAM, presentan alteraciones genéticas, epigenéticas y mutaciones comunes que resultan en un crecimiento y diferenciación celular anómalos.

El análisis del perfil de expresión génica (PEG) mediante microarrays de alta densidad se ha utilizado en el estudio de los mecanismos transcripcionales que están alterados e implicados en cáncer, incluyendo las leucemias y los SMD. Estos estudios han proporcionado grandes avances en el conocimiento de los genes y los mecanismos implicados en la patogénesis de estas enfermedades. Su aplicación en los SMD demostró que los éstos podían ser estratificados en diferentes grupos, en términos de tiempo de transformación a LAM, en base a sus perfiles de expresión. Así, aquellos SMD con una enfermedad más agresiva eran clasificados como una entidad diferente de aquellos SMD con un curso clínico más indolente. Algunos de los genes que discriminaban estas entidades eran *HOXA9*, *FLT3*, *KIT* y *WT1*. Estos estudios, a lo largo de los últimos años, han demostrado también que los SMD, en los estadios más tempranos, se caracterizan por una desregulación de la respuesta inmune y un aumento de la apoptosis intramedular, con un ratio apoptosis/proliferación elevado. Por el contrario, en estadios más avanzados se observa una disminución de la apoptosis, así como una desregulación de la respuesta y reparación del daño en el ADN. Sin embargo, las bases moleculares que expliquen estas alteraciones y que podrían determinar la progresión de los SMD a LAM son en gran parte desconocidos.

Por ello, se analizaron los cambios de expresión génica que se producían en la transición desde una médula ósea no maligna, a través de los diferentes estadios de los SMD, hacia LAM mediante dos plataformas de microarrays de expresión de alta densidad.

Pacientes y métodos

Se estudiaron un total de 182 pacientes con hemopatías malignas, 127 eran SMD y 55 eran LAM, mediante microarrays de expresión. Además, se incluyeron 17 muestras de pacientes sin hemopatías de origen tumoral que se utilizaron como controles, denominados en este trabajo como no leucemias (NoL). Para evitar el efecto de dosis génica, excluimos del estudio todos aquellos pacientes que presentaran alteraciones citogenéticas. Asimismo, debido a que los SMD con sideroblastos en anillo presentan un PEG característico, éstos fueron también eliminados del estudio. De este modo, el grupo finalmente analizado estaba constituido por 90 pacientes con citogenética normal. Éstos fueron clasificados según criterios OMS 2008 en: CRDU (n=11), CRDM (n=23), AREB-1 (n=9), AREB-2 (n=10) y AML (n=20). Los pacientes con CRDU y CRDM fueron agrupados y considerados como SMD de bajo riesgo (SMD-BR), mientras que los pacientes con AREB-1 y AREB-2 fueron considerados como SMD de alto riesgo (SMD-AR).

Los estudios del pefil de expresión génico se realizaron en dos plataformas diferentes: el microarray *Human Genome U133 Plus 2.0* y el *Human Exon 1.0 ST* (Affymetrix).

Basándonos en la hipótesis de que los distintos subtipos de SMD son estadios consecutivos en la progresión hacia LAM, se buscaron aquellos genes cuya expresión seguía una tendencia creciente o decreciente durante la evolución de la enfermedad. Para ello se calculó la correlación (*Goodman and Kruskal's Gamma correlation*) entre la expresión génica y el estadio de la enfermedad, ordenados como NoL → SMD-BR → SMD-AR → LAM. De esta manera, para cada gen se obtuvo un valor de correlación (valor Gamma). Para continuar con el estudio de evolución, se seleccionaron aquellos genes que se correlacionaban con la progresión de la enfermedad, es decir, tenían una valor de correlación Gamma significativo (valor Gamma absoluto >0,50 y un FDR con *P* valor ajustado <0,05).

Resultados

Como ya se ha indicado, con el fin de identificar aquellos genes que podrían estar relacionados con la progresión de la enfermedad, se buscaron aquellos genes cuyos niveles de expresión seguían una tendencia creciente o decreciente a través de los diferentes

estadios de los SMD, considerando las no leucemias como el punto de origen no maligno y las LAM como el estado final de la progresión. Para ello, se empleó el método *SOM*, un método de agrupación no supervisado que identifica grupos de genes con un patrón de expresión similar. De este modo este análisis identificó una serie de genes y funciones celulares que estaban común y progresivamente desreguladas durante la transición desde una medula ósea no maligna, a través de estadios tempranos de los SMD (SMD-BR), estadios de SMD más avanzados (SMD-AR), hasta las LAM.

En concreto, se observó que algunos genes comenzaban a estar sobre-expresados en los SMD de bajo riesgo y que continuaban con esa tendencia de expresión creciente hacia las LAM (patrón 1 o patrón creciente SMD/LAM) como *NPM1*, *CCNG1*, *PMAIP1*, *MYST1*, *RPL22*, *RPS6*. Otros genes, a pesar de seguir esta misma tendencia creciente, presentaban el mayor cambio en sus nivel de expresión en la transición de los SMD de alto riego a LAM (patrón 3 o patrón creciente LAM). Algunos de estos genes eran *HOXA9*, *HOXA7*, *MEIS1* y *FLT3*. Por el contrario, se observó que otro grupo de genes se iban infra-expresando progresivamente durante la evolución de la enfermedad, alcanzando unos niveles de expresión mínimos en las LAM (patrón 2 o patrón decreciente SMD/LAM y patrón 4 o patrón decreciente LAM). Algunos de estos genes eran *MMP25*, *CEACAM3*, *LTF*, *CRISP3*, *CAMP*, *MMP9*. Los genes icluidos en el patrón 1 estaban relacionados con la respuesta al daño en el ADN y los ribosomas. Los genes del patrón 3 estaban implicados en promover la proliferación celular y suprimir la diferenciación, mientras que los genes de los patrones 2 y 4 estaban relacionados con la desregulación de la respuesta inmune.

Curiosamente, el análisis funcional de los genes incluidos en cada uno de los patrones reveló que algunos de ellos eran factores de transcripción. Debido a que los genes estudiados seguían unas tendencias crecientes o decrecientes en sus niveles de expresión, decidimos buscar si algún factor de transcripción que estuviera desregulado en nuestro estudio podría ser el responsable de los cambios de expresión observados. De manera interesante se identificaron 42 factores de transcripción, dos de los cuales, ATF2 y TAF7, parecían estar controlando un gran número de los genes del patrón 1, siendo la mayoría de ellos genes ribosomales.

Conclusiones

El presente estudio, y la metodología aplicada en este trabajo, demostró que una serie de funciones celulares están desreguladas de manera progresiva, y con genes desregulados en común, durante la transición desde una médula ósea no maligna, a través de los SMD de bajo riesgo y los SMD de alto riesgo, hasta las LAM.

Las principales funciones celulares que podrían estar directa o indirectamente implicadas en la progresión de la enfermedad fueron: una activación de la respuesta daño en el ADN y los puntos de control del ciclo celular, una sobre-expresión de los genes codificantes para proteínas ribosomales y responsables del denominado “estrés nucleolar", una elevada apoptosis, un aumento en la proliferación celular y la supresión de la diferenciación mieloide, así como una desregulación de la respuesta inmune.

Por otra parte, un gran número de los genes identificados como progresivamente desregulados en este estudio podrían estar regulados por dos factores de transcripción, ATF2 y TAF7, que no han sido previamente relacionados con las neoplasias mieloides. Estos dos factores de transcripción, o las funciones que éstos regulan, resultan dianas potencialmente atractivas para la investigación de nuevas terapias.

Por lo tanto, la evolución desde una médula ósea normal a través de los SMD hacia la leucemia parece ser un mecanismo organizado que implica genes y funciones celulares desreguladas en común.

Resumen del capítulo 3:

Prognostic impact of the number of methylated genes in myelodysplastic syndromes and acute myeloid leukemias treated with azacytidine

María Abáigar, Fernando Ramos, Rocío Benito, María Díez-Campelo, Javier Sánchez-del-Real, Lourdes Hermosín, Juan Nicolás Rodríguez, Carlos Aguilar, Isabel Recio, Jose María Alonso, Natalia de las Heras, Marta Megido, Marta Fuertes, María Consuelo del Cañizo, Jesús María Hernández-Rivas

Ann Hematol. (2013) doi: 10.1007/s00277-013-1799-9
PubMed PMID: 23740492

La versión original, incluyendo tablas y figuras está disponible en las páginas 77 - 88 y el material suplementario en las páginas 201 - 203.

Introducción y objetivos

Los agentes hipometilantes, como la 5-azacitidina, han supuesto un gran avance en el tratamiento de los síndromes mielodisplásicos (SMD) y las leucemias agudas mieloblásticas (LAM) ya que muchos de estos pacientes alcanzan una respuesta y porque aumentan significativamente la supervivencia. La 5-azacitidina (5-AZA) es un agente hipometilante que se une de forma irreversible a las DNA-metiltransferasas e inhiben su función, lo que resulta en una progresiva disminución de la metilación del ADN. La metilación aberrante de las islas CpG del ADN se ha asociado con el silenciamiento génico, por esta razón, el tratamiento con 5-azacitidina podría reactivar la expression génica. En los últimos años se ha descrito que diversos genes supresores tumorales están hipermetilados en las neoplasias mieloides. Por ello, la 5-azacitidina parece ser una buena opción para el tratamiento de estas hemopatías. Sin embargo, el significado pronóstico del estado de metilación del ADN en la respuesta terapéutica a estos agentes está poco definido.

Por esta razón analizamos el estado de metilación previo al tratamiento con 5-azacitidina de un grupo de genes supresores tumorales en pacientes con SMD y LAM, con el fin de evaluar la influencia de un estado de metilación aberrante en la supervivencia y respuesta a este fármaco.

Pacientes y métodos

Se incluyeron un total de 78 pacientes con SMD o LAM que habían sido tratados con 5-azacitidina. Sin embargo, a lo largo del estudio se excluyeron 15 pacientes porque no cumplían los criterios de inclusión, que eran: el haber recibido más de 4 ciclos de tratamiento, que la 5-azacitidina no fuese utilizada como tratamiento de mantenimiento tras un trasplante alogénico, que la respuesta fuese evaluable y el disponer de ADN de buena calidad. Así que finalmente, el estudio se centró en 63 pacientes: 39 SMD y 24 LAM, de los que se recogieron las características clínicas, de respuesta al tratamiento, supervivencia, citogenética y FISH. Además, se analizaron 5 muestras de MO procedentes de donantes sanos.

El análisis de metilación se analizó en todos los enfermos mediante la técnica MS-MLPA, una metodología basada en la PCR, para un panel de 24 genes supresores tumorales relacionados con control del ciclo celular, regulación del a apoptosis, reparación del ADN, adhesión celular y crecimiento celular. La respuesta al tratamiento se evaluó de acuerdo con los criterios del "International Working Group" (IWG) para los SMD y las LAM. Aquellos pacientes que alcanzaron una remisión completa (RC), una respuesta parcial (RP) o una respuesta hematológica (RH) fueron considerados como respondedores, mientras que aquellos pacientes en los que falló el tratamiento (FT) o progresaron (PROG) fueron considerados como no respondedores.

Resultados

En la mayoría de los pacientes (73%) se observó metilación de alguno de los genes estudiados: 23 pacientes tenían 1 sólo gen metilado, 15 tenían 2 genes metilados y los restantes 8 (12%) presentaron 3 o más genes metilados. Los genes metilados con mayor frecuencia fueron: *IGSF4* (27%), *CDKN2B* (23.8%), *ESR1* (23.8%), *CDH13* (19%) y *CDKN1B* (11.1%). Otros genes como *TP73*, *RARB*, *FHIT*, *APC*, *CDKN2A*, *CD44*, *GSTP1*, *HIC1* y *TIMP3* aparecieron metilados en menos del 10% de los pacientes. Por en contrario no se observó metilación de ninguno de los genes en los individuos control.

Nuestro objetivo era evaluar si el estado de metilación, así como otras características clínicas, podrían estar relacionadas con la supervivencia y la respuesta al tratamiento con 5-azacitina. En primer lugar, no se encontraron diferencias estadísticas en la supervivencia entre los pacientes con SMD y LAM. Por esta razón, analizamos los SMD y LAM como una única entidad. El análisis estadístico reveló que la presencia de anemia (P=0,015), leucocitosis (P=0,021), y un elevado nivel de metilación, definido como más de 2 genes metilados (P=0,017), se asociaron con una menor supervivencia en el análisis univariante. Todas estas variables mantuvieron su significación en el análisis multivariante (P=0,029, P=0,033 y P=0,022, respectivamente) como factores pronósticos independientes de la supervivencia en los pacientes tratados con 5-azacitidina. Cabe señalar que el análisis de los pacientes que no presentaron metilación de alguno de los genes estudiados, aquellos con 1 gen metilado, los que tienen 2 genes metilados y aquellos con 3 o más genes metilados,

mostró que a medida que el número de genes metilados aumenta, la supervivencia disminuye. De hecho, los pacientes del grupo con mayor número de genes metilados son los que presentan una supervivencia muy corta.

En relación con la respuesta a 5-azacitidina, se observó que el estado de metilación no se asoció de forma significativa con la respuesta a este fármaco (*P*=0.654), aunque 4 de los 5 pacientes con más de 3 genes metilados no respondieron. Por el contrario, la presencia de una citogenética favorable era el único factor que podría predecir, de forma independiente, la respuesta a 5-azacitidina (*P*=0,026). En nuestro estudio, de todos los pacientes con una citogenética favorable, un 64,7% respondieron. Cabe destacar que 3 SMD presentaron alteraciones del cromosoma 7 y todos ellos respondieron al tratamiento con 5-AZA.

Conclusiones

Los resultados del presente estudio mostraron que, además de los factores clínicos previamente descritos, la presencia de un estado de metilación aberrante influía negativamente en la supervivencia de los enfermos tratados con 5-azacitidina, siendo más corta en el grupo de pacientes con los niveles de metilación más elevados. Por el contrario, una citogenética adversa se asoció con una peor respuesta a este fármaco.

En resumen, nuestros resultados sugieren que el análisis del estado de metilación previo al tratamiento con 5-azacitidina mediante una metodología como MS-MLPA, podría ser útil para identificar el grupo de pacientes con una supervivencia más corta con los agentes hipometilantes, en los que podrían considerarse otras alternativas terapéuticas.

Discusión General

"Lo importante es no dejar de hacerse preguntas"

Albert Einstein

Los síndromes mielodisplásicos (SMD) constituyen un conjunto de enfermedades hematológicas clonales que afectan a la célula madre hematopoyética mieloide. Se caracterizan por una gran heterogeneidad clínica y biológica con un elevado riesgo de progresión a leucemia aguda mieloblástica (LAM). Se han identificado una gran variedad de alteraciones genómicas y epigenéticas que resultan en una diferenciación, maduración y crecimiento anómalo de las células hematopoyéticas mieloides, lo que da lugar a un fallo medular y a un mayor riesgo de desarrollar leucemia. Sin embargo, hasta el momento no se ha identificado una única mutación o anomalía genómica que pueda ser responsable del origen y desarrollo de los SMD así como de su frecuente progresión a LAM. Teniendo en cuenta la heterogeneidad molecular de los SMD, las técnicas genómicas de alta resolución (aCGH, SNP-A, PEG mediante microarrays, NGS) son herramientas de gran utilidad y mucho más eficaces en el análisis de las anomalías genómicas en los pacientes con SMD. Estos estudios han proporcionado grandes avances en el conocimiento de la biología de estas enfermedades tan complejas, y por lo tanto una mejora en su diagnóstico y evaluación pronóstica, así como han dianas para el desarrollo de nuevos fármacos para su tratamiento.

Hasta la fecha son muchos los estudios que se han realizado en los SMD mediante aCGH y SNP-A permitiendo la identificación de alteraciones nuevas y crípticas, no detectadas por las técnicas de citogenética convencional (CC), sugiriendo su utilidad a nivel clínico.[13,47,48] Los resultados presentados en esta tesis doctoral, así como los datos previamente publicados, han demostrado que existe una muy buena concordancia entre los datos de aCGH y CC, confirmando la fiabilidad de esta técnica para el análisis de grandes anomalías genéticas. Asimismo, los estudios de arrays genómicos han permitido redefinir el pronóstico de algunos pacientes con SMD para las alteraciones ya conocidas como -5/5q-, -7/7q-, trisomía 8, 20q-, y cariotipos complejos.[42,46,47,53,77] Cabe destacar que todas las alteraciones en el número de copias CNA conducen a una pérdida de función en el caso de la deleción de un gen supresor tumoral o a una ganancia de función en el caso de la amplificación de un oncogén. Por esta razón, las alteraciones crípticas (microdeleciones, microamplificaciones, LOH) han despertado un gran interés, ya que a menudo implican a genes individuales o muy pocos genes, que pueden ser estudiados posteriormente por secuenciación.[44,47,131] De hecho, estas regiones han sido fundamentales en el descubrimiento de nuevas mutaciones, como *TET2* en los SMD/SMP, *CBL* en las LMMC, *MPL* en las ARSA-T, y *EZH2* en los MDS, LMMC y LAM.[47] En

nuestro estudio se identificaron alteraciones crípticas en regiones importantes en la patogénesis de los SMD como deleciones en 2p23.3 (*DNMT3A*), 4q24 (*TET2*), 5q33.1 (*SPARC*), 7q22.1 (*CUX1*), 21q22.12 (*RUNX1*) y Xp11.4 (*BCOR*), y ganancias en 21q22.3 (*U2AF1*).[2,34,42,50] Así decidimos analizar si existían mutaciones en el alelo no delecionado en los genes genes más relevantes aplicando la técnica de secuenciación masiva de amplicones. Nuestros resultados mostraron que una baja proporción de pacientes presentaba una deleción y mutación simultánea en los genes analizados, a excepción de *TP53*. Por lo tanto, los arrays genómicos proporcionan información nueva sobre regiones críticas involucradas en la patogénesis de los SMD, y sugerimos que series más amplias de pacientes con SMD deberían ser estudiadas mediante aCGH.

Además, en la era de la genómica, las alteraciones cromosómicas continuarán siendo relevantes a nivel clínico y la citogenética convencional (CC) seguirá desempeñando un papel fundamental en el estudio de los SMD.[2] Nuestros resultados junto con los datos previamente publicados sugieren que la CC debe complementarse no sólo con estudios mediante arrays genómicos sino también con estudios de secuenciación masiva para la evaluación de estos enfermos con SMD, utilizando paneles de genes en lugar de genes individuales. Sin embargo, resulta necesario ampliar estos estudios para integrar toda la información generada a través de estos análisis genómicos masivos, como las alteraciones crípticas y las mutaciones, en la práctica clínica para el manejo de los pacientes con SMD.

Por otro lado, los análisis mediante aCGH nos permitieron identificar la presencia de cromotripsis en tres pacientes con SMD de alto riesgo, una subtipo de SMD que presenta un alto riesgo de desarrollar una leucemia. Generalmente, la cromotripsis afecta a diferentes cromosomas de manera aleatoria. Sin embargo, en nuestro estudio observamos múltiples y complejas alteraciones consistentes con los criterios de cromotripis en el cromosoma 13 en los tres pacientes. Otros estudios han reportado la presencia de este fenómeno en un cromosoma de manera recurrente, como en 16p en 3/7 pacientes con mieloma múltiple, en el cromosoma 21 en 5/9 iAMP21 casos de leucemia linfoblástica aguda, en el cromosoma 5 en 3/8 casos con leucemia linfocítica crónica, y más recientemente, en el cromosoma 13 en tres casos de retinoblastoma.[60,70,136,137] De acuerdo con la sugerencia de Mcevoy et al.,[137] de que la inactivación de un gen supresor tumoral por un evento de cromotripsis podría iniciar

el desarrollo de un tumor, como por ejemplo la inactivación de *RB1* en el retinoblastoma, cabe mencionar que en nuestro estudio encontramos dos genes supresores de tumores alterados en los tres pacientes con cromotripsis, *BRCA2* y *RB1*.[138,139] Además, los tres pacientes con cromotripsis presentaron mutaciones de *TP53*, otro gen supresor tumoral que ya se ha relacionado con el fenómeno de cromotripsis.[59,136] Sin duda, en nuestro podemos estar perdiendo muchos detalles genéticos, sin embargo, el estudio de aCGH identificó cromosoma 13 altamente alterado, reflejando muchos puntos de ruptura, consistentes con la firma genómica de la cromotripsis. Por lo tanto, esta compleja anomalía genética podría estar implicado en la patogénesis de algunos casos de SMD.[49,141]

Hay que tener en cuenta que todas las alteraciones citogenéticas, genéticas y epigenéticas pueden afectar la expresión génica a través de anomalías a nivel de la transcripción, de la regulación epigenética, o por el efecto de dosis génica.[142] El estudio de los perfiles de expresión génica ha demostrado ser una herramienta útil para el análisis simultáneo de los niveles de expresión de casi todos los genes conocidos, y se ha aplicado en el estudio de los SMD proporcionando una mejor comprensión de los complejos mecanismos que subyacen su patofisiología. Dichos estudios se han centrado principalmente en la identificación de marcadores genéticos específicos para cada subtipo de SMD, comparando el PEG entre los SMD e individuos sanos, entre los diferentes grupos de riesgo, o entre los diferentes subtipos de SMD.[106,109-111,115-121] Sin embargo, dado que los SMD y las LAM tienen anomalías morfológicas, citogenéticas, epigenéticas y mutaciones comunes, nos propusimos analizar los genes y funciones celulares desreguladas en común durante la evolución de los SMD.[2,35,92] Así, se identificaron una serie de genes y funciones celulares que estaban desregulados en común y de manera progresiva (creciente o decreciente) durante la transición desde una médula ósea no maligna, a través de los SMD de bajo y alto riesgo, hasta las LAM. Los genes desregulados estaban implicados en la activación de la respuesta al daño en el ADN y de los puntos de control del ciclo celular, un aumento de la apoptosis, de los ribosomas, el ensamblaje de los nucleosomas y la remodelación de la cromatina, en promover crecimiento celular y la proliferación, la supresión de la diferenciación de las células progenitoras mieloides, y una represión de respuesta inmune.

En base a nuestras observaciones, proponemos un modelo para el desarrollo de los SMD y la progresión hacia la LAM (Representado en la figura 5, página 100). En este modelo se representa una enfermedad maligna que tiene alteraciones genómicas incluso en las primeras etapas de la enfermedad. Durante las fases iniciales, las células mostrarían una reacción frente a un daño en el ADN, a través de la activación de los mecanismos de reparación, que producirían una parada del ciclo celular o apoptosis. Esta respuesta al daño en el ADN podría estar desencadenada por la desregulación de las proteínas ribosomales, a través de la estabilización y activación de p53. Ante esta situación, las células mostrarían un aumento en los niveles de apoptosis indicando que las células de médula ósea están actuando de la manera correcta, es decir, tratando de eliminar las células defectuosas. En este punto, las células todavía conservarían su capacidad de diferenciación/maduración. Sin embargo, el ratio apoptosis/auto-renovación-proliferación en la MO es elevado. Esto reflejaría que las células derivadas a partir del clon maligno inicial no sólo continuarían diferenciándose, sino que también morirían de forma prematura. Al mismo tiempo, debido a un mal funcionamiento de la respuesta inmune se produciría una liberación de citoquinas pro-apoptóticas. Estos mecanismos desregulados darían lugar a una elevada apoptosis intramedular, lo que explicaría, al menos en parte, la hematopoyesis ineficaz y citopenias periféricas observadas en los pacientes con SMD de bajo riesgo.

Sin embargo, ante la acumulación de daño en las células madre hematopoyéticas, las células todavía mostrarían una activación de los mecanismos de respuesta y reparación del daño en el ADN y responde las de reparación, así como una sobre-expresión de las proteínas ribosomales. Además, debido a una maquinaria de reparación deficiente o a un daño excesivo en el ADN, las células madre hematopoyéticas y las células progenitoras adquirirían y acumularían nuevas alteraciones genéticas y/o epigenéticas. En este punto, las células dañadas mostrarían un aumento en las propiedades proliferativas y de auto-renovación, mientras que no serían capaces de diferenciarse hacia estadios más maduros. Además, estas células mostrarían una resistencia aumentada a la apoptosis, así como una evasión al sistema inmune, exigiendo una alta síntesis de proteínas para sobrevivir. Esto podría explicar la sobre-expresión de las proteínas ribosomales, que en este caso podrían contribuir a la progresión de la enfermedad. En conjunto, estos mecanismos proporcionarían a la célula maligna una cara ventaja proliferativa, y cooperarían permitiendo la expansión del clon

maligno. Como resultado, la proporción de blastos en la médula ósea aumentaría a lo largo del tiempo, desarrollándose la leucemia. En conjunto, estos datos sugieren que la progresión de los SMD hacia LAM es un proceso organizado, resultado de la acumulación de daño en el ADN, probablemente debido a que las células madre hematopoyéticas no son capaces de responder adecuadamente a este daño.

Nuestro estudio proporciona una base biológica que explicaría muchas de las alteraciones obsevadas en funciones celulares implicadas en la progresión de los SMD a LAM. Con el objetivo de reducir la heterogeneidad genética entre las muestras incluidas en el estudio, el análisis se restringió a aquellos pacientes con cariotipo normal exclusivamente. Además, se excluyeron aquellos SMD con sideroblastos en anillo, que se han asociado con mutaciones somáticas de *SF3B1*. Una posible limitación de nuestro estudio es que desconocemos las mutaciones que puedan presentar de las muestras analizadas. Por esta razón, creemos que es necesario realizar estudios en series más amplias de pacientes para identificar si la presencia de mutaciones en genes concretos podría estar relacionado con la desregulación génica observada en los SMD.

En los últimos años, se han desarrollado nuevos fármacos para el tratamiento de los SMD y LAM, como los agentes hipometilantes (5-azacitidina y decitabina), de hecho, se han convertido en el tratamiento de primera línea para estas enfermedades.[7,126,130] Estos fármacos has demostrado ser eficaces, con una tasa de remisiones completas del 10-15%, produciendo una mejoría hematológica, e incluso retrasando la progresión a LAM y prolongando la supervivencia.[7,72,74,78,124,126-128,130] Sin embargo, un 40-50% de los pacientes siguen siendo refractarios a estas terapias. Por esta razón, en el momento actual es de gran interés el poder disponer de factores predictivos de respuesta a estos agentes, para seleccionar aquellos pacientes que más se beneficiarían con este tratamiento.[8]

En los SMD y LAM, se han descrito alteraciones en la metilación del ADN, asociadas con el silenciamiento génico.[80,81,169-171] Algunos estudios han demostrado que el tratamiento con agentes hipometilantes produce una reducción de la metilación y una re-expresión de genes como *CDKN2B*, *CDH1* y *DAPK1*.[80,172] Por esta razón, pensamos que el estado de metilación previo al tratamiento podría predecir la respuesta a estos agentes. Sin embargo, nuestro estudio, al igual que otros previamente descritos, demostraron que una metilación

aberrante para los genes estudiados no se asociaba con la respuesta a 5-AZA.[83,172] Además, el responder o no a un determinado fármaco, podría alterar el pronóstico de una enfermedad.[173] Por ello, el estudio de la supervivencia en relación con la metilación, en estos pacientes tratados con 5-AZA, es también de gran interés para definir el grupo de pacientes con mayor probabilidad de beneficiarse de este tratamiento. Nuestros datos demostraron que el estado de metilación previo a la 5-azacitidina, en concreto la presencia de ≥2 genes metilados, estaba asociada con una supervivencia más corta, siendo menor en el grupo de pacientes con los nivel de metilación más elevados (≥3 genes). Resultados similares habían sido reportados para los genes *CDKN2B*, *HIC1*, *CDH1*, *ESR1* y *FHIT*, en los que la hipermetilación se asociaba con un peor pronóstico.[79,83,84,86-89]

Estudios recientes han demostrado que los pacientes con mutaciones de *TET2* presentan mejores respuestas a los agentes hipometilantes, cuando estas mutaciones aparecen en más del 10% de las células y no existen mutaciones de *ASXL1*. Por el contrario, estas mutaciones no se asociaron con mejoras en la supervivencia.[173,174] Sin embargo, y lo que es más importante, este trabajo demostró que no existe una mutación o combinación de mutaciones que se asociara con la falta de respuesta a estos fármacos, ni siquiera aquellas mutaciones que confieren un pronóstico muy adverso como *RUNX1* y *TP53*.[173,174] Del mismo modo, ninguna de estas mutaciones se relacionó con mal pronóstico en el grupo de pacientes tratados con 5-AZA. Por ello, hasta la fecha no existen suficientes evidencias que contraindiquen el tratamiento con agentes hipometilantes en estos pacientes.[173] Debido a que el número de alteraciones genéticas y epigenéticas que se van descubriendo en los SMD y LAM aumenta cada día, estudios integrados combinando las diferentes técnicas de análisis genómico, serían de gran interés para para seleccionar aquellos en los que los agentes hipometilantes fueran más eficaces.

En resumen, los datos biológicos y moleculares presentados en esta tesis doctoral, en base a los estudios del perfil de expresión génico, de arrays genómicos, secuenciación masiva y metilación, han permitido ampliar nuestro conocimiento sobre la biología de estas hemopatías, mediante la identificación de alteraciones genómicas que podrían tener relevancia clínica. Nuestros datos resaltan la utilidad de las técnicas de análisis masivo en el estudio de los SMD, principalmente en aquellos pacientes con cariotipo normal o pocas

metafases analizables. Y además, los genes y mecanismos descritos podrían ser considerados como posibles dianas terapéuticas en el futuro.

Conclusiones

1. El empleo de la técnica de aCGH en la evaluación de los pacientes con SMD podría utilizarse como una técnica complementaria a la citogenética convencional, especialmente en aquellos pacientes con resultados no informativos, debido a la ausencia de mitosis, o con pocas metafases evaluables para los estudios citogenéticos. Por lo tanto, en este grupo de pacientes, los aCGH pueden detectar alteraciones genéticas permitiendo la estratificación pronóstica de acuerdo con el IPSS-R, lo que podría modificar la práctica clínica para el manejo de este grupo de pacientes

2. El análisis de las alteraciones en el número de copias mediante arrays genómicos permite la identificación de anomalías crípticas en regiones relevantes en la patogénesis de los SMD, como deleciones en *DNMT3A*, *TET2*, *SPARC*, *CUX1* y *RUNX1*, así como ganancias en *U2AF1*. Además, los estudios de secuenciación masiva revelan que una baja proporción de pacientes presenta una deleción y mutación simultáneas en los genes estudiados.

3. Las técnicas genómicas de alta resolución para el análisis del número de copias permiten la detección de cromotripsis en el cromosoma 13 en tres pacientes con SMD de alto riesgo. Los patrones de alteración genómica en los casos con cromotripsis son diferentes entre los tres pacientes. Sin embargo, algunos genes relacionados con cáncer como *FLT3*, *FLT1*, y *XPO4* están comúnmente amplificados, mientras que *BRCA2* y *RB1* aparecen comúnmente delecionados. Además, los tres pacientes presentan cariotipos complejos, mutaciones en *TP53*, y un pronóstico adverso.

4. La progresión de los SMD desde una médula ósea no maligna hasta las LAM parece ser un mecanismo organizado. Así, el análisis de los perfiles de expresión génica, durante la evolución de los SMD, muestra que determinados genes y funciones celulares están desregulados en común y de manera progresiva durante la transición desde una médula ósea no maligna, a través de los SMD de bajo y alto riesgo, hasta las LAM.

    4.1. Los genes relacionados con la respuesta al daño en el ADN y los puntos de control del ciclo celular, con la apoptosis y los ribosomas, así como con el ensamblaje de los nucleosomas y la remodelación de la cromatina están progresivamente sobre-expresados durante la evolución de la enfermedad, incluso desde los estadios más

tempranos de los SMD. Sin embargo, los genes implicados en la proliferación celular, la capacidad de auto-renovación y la diferenciación celular están sobre-expresados en la transición de los SMD de alto riesgo a las LAM.

4.2. Los genes relacionados con la respuesta inmune, la adhesión celular y metaloproteinasas de la matriz extracelular están progresivamente infra-expresados durante la evolución de la enfermedad, mostrando una mayor represión en la transición desde los SMD en estadios más avanzados hacia las LAM.

5. La presencia de dos o más genes metilados en pacientes con SMD y LAM antes del tratamiento con 5-azacitidina se asocia con una supervivencia global más corta. Sin embargo, el estado de metilación previo al tratamiento no se correlaciona con la respuesta clínica. Por el contrario, la presencia de una citogenética adversa se asocia con una peor respuesta al tratamiento con 5-azacitidina.

References

1. Mufti GJ, Bennett JM, Goasguen J, et al. Diagnosis and classification of myelodysplastic syndrome: International Working Group on Morphology of myelodysplastic syndrome (IWGM-MDS) consensus proposals for the definition and enumeration of myeloblasts and ring sideroblasts. Haematologica 2008;93(11):1712-7.
2. Cazzola M, Della Porta MG, Malcovati L. The genetic basis of myelodysplasia and its clinical relevance. Blood 2013;122(25):4021-34.
3. Visconte V, Selleri C, Maciejewski JP, Tiu RV. Molecular pathogenesis of myelodysplastic syndromes. Transl Med UniSa 2014;8:19-30.
4. Malcovati L, Porta MG, Pascutto C, et al. Prognostic factors and life expectancy in myelodysplastic syndromes classified according to WHO criteria: A basis for clinical decision making. J Clin Oncol 2005;23(30):7594-603.
5. Malcovati L, Della Porta MG, Strupp C, et al. Impact of the degree of anemia on the outcome of patients with myelodysplastic syndrome and its integration into the WHO classification-based Prognostic Scoring System (WPSS). Haematologica 2011;96(10):1433-40.
6. Corey SJ, Minden MD, Barber DL, Kantarjian H, Wang JC, Schimmer AD. Myelodysplastic syndromes: The complexity of stem-cell diseases. Nat Rev Cancer 2007;7(2):118-29.
7. Adès L, Itzykson R, Fenaux P. Myelodysplastic syndromes. The Lancet 2014;383(9936):2239-52.
8. Steensma DP. Myelodysplastic syndromes: Diagnosis and treatment. Mayo Clin Proc 2015;90(7):969-83.
9. Garcia-Manero G. Myelodysplastic syndromes: 2014 update on diagnosis, risk-stratification, and management. Am J Hematol 2014;89(1):97-108.
10. Cazzola M, Malcovati L. Myelodysplastic syndromes--coping with ineffective hematopoiesis. N Engl J Med 2005;352(6):536-8.
11. Deeg HJ, Bowen DT, Gore SD, Haferlach T, Le Beau MM, Niemeyer C (eds). 2006. Hematologic malignancies: Myelodysplastic syndromes. Springer.
12. Catenacci DV, Schiller GJ. Myelodysplasic syndromes: A comprehensive review. Blood Rev 2005;19(6):301-19.
13. Malcovati L, Hellström-Lindberg E, Bowen D, et al. Diagnosis and treatment of primary myelodysplastic syndromes in adults: Recommendations from the European LeukemiaNet. Blood 2013;122(17):2943-64.
14. Vardiman JW, Thiele J, Arber DA, 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.
15. Haferlach C, Rieder H, Lillington DM, et al. Proposals for standardized protocols for cytogenetic analyses of acute leukemias, chronic lymphocytic leukemia, chronic myeloid leukemia, chronic myeloproliferative disorders, and myelodysplastic syndromes. Genes Chromosomes Cancer 2007;46(5):494-9.
16. Haase D, Germing U, Schanz J, et al. New insights into the prognostic impact of the karyotype in MDS and correlation with subtypes: Evidence from a core dataset of 2124 patients. Blood 2007;110(13):4385-95.
17. Shaffer LG, McGowan-Jordan J, Schmid M (eds). 2013. ISCN 2013: An International System for Human Cytogenetic Nomenclature: Recommendation of the International Standing Committee on Human Cytogenetic Nomenclature. S Karger, Basel.
18. Gondek LP, Haddad AS, O'Keefe CL, et al. Detection of cryptic chromosomal lesions including acquired segmental uniparental disomy in advanced and low-risk myelodysplastic syndromes. Exp Hematol 2007;35(11):1728-38.
19. Schanz J, Tüchler H, Solé F, et al. New comprehensive cytogenetic scoring system for primary myelodysplastic syndromes (MDS) and oligoblastic acute myeloid leukemia after MDS derived from an international database merge. J Clin Oncol 2012;30(8):820-9.
20. Sole F, Luno E, Sanzo C, et al. Identification of novel cytogenetic markers with prognostic significance in a series of 968 patients with primary myelodysplastic syndromes. Haematologica 2005;90(9):1168-78.
21. Such E, Cervera J, Costa D, et al. Cytogenetic risk stratification in chronic myelomonocytic leukemia. Haematologica 2011;96(3):375-83.
22. Greenberg PL, Tuechler H, Schanz J, et al. Revised international prognostic scoring system for myelodysplastic syndromes. Blood 2012;120(12):2454-65.
23. Flach J, Dicker F, Schnittger S, et al. An accumulation of cytogenetic and molecular genetic events characterizes the progression from MDS to secondary AML: an analysis of 38 paired samples analyzed by cytogenetics, molecular mutation analysis and SNP

microarray profiling. Leukemia. 2011;25(4):713-8.

24. Meggendorfer M, De Albuquerque A, Nadarajah N, et al. Karyotype evolution and acquisition of FLT3 or RAS pathway alterations drive progression of myelodysplastic syndrome to acute myeloid leukemia. Haematologica 2015. [Epub ahead of print]

25. Bejar R. Prognostic models in myelodysplastic syndromes. ASH Education Program Book 2013;2013(1):504-10.

26. Estephan F, Tiu RV. Current and novel therapeutic approaches in myelodysplastic syndromes. J Community Support Oncol 2014;12(7):236-49.

27. Bennett JM, Catovsky D, Daniel MT, et al. Proposals for the classification of the myelodysplastic syndromes. Br J Haematol 1982;51(2):189-99.

28. Recommendations for a morphologic, immunologic, and cytogenetic (MIC) working classification of the primary and therapy-related myelodysplastic disorders. Report of the workshop held in Scottsdale, Arizona, USA, on February 23-25, 1987. Third MIC Cooperative Study Group. Cancer Genet Cytogenet 1988;32(1):1-10.

29. Vardiman JW, Harris NL, Brunning RD. The World Health Organization (WHO) classification of the myeloid neoplasms. Blood 2002;100(7):2292-302.

30. Greenberg P, Cox C, LeBeau MM, et al. International Scoring System for evaluating prognosis in myelodysplastic syndromes. Blood 1997;89(6):2079-88.

31. Jonas BA, Greenberg PL. MDS prognostic scoring systems - past, present, and future. Best Pract Res Clin Haematol 2015;28(1):3-13.

32. Malcovati L, Germing U, Kuendgen A, et al. Time-dependent prognostic scoring system for predicting survival and leukemic evolution in myelodysplastic syndromes. J Clin Oncol 2007;25(23):3503-10.

33. Tefferi A, Vardiman JW. Myelodysplastic syndromes. N Engl J Med 2009;361(19):1872-85.

34. Bejar R, Stevenson K, Abdel-Wahab O, et al. Clinical effect of point mutations in myelodysplastic syndromes. N Engl J Med 2011;364(26):2496-506.

35. Papaemmanuil E, Gerstung M, Malcovati L, et al. Clinical and biological implications of driver mutations in myelodysplastic syndromes. Blood 2013;122(22):3616-27; quiz 3699.

36. Haferlach T, Nagata Y, Grossmann V, et al. Landscape of genetic lesions in 944 patients with myelodysplastic syndromes. Leukemia 2014;28(2):241-7.

37. Korbel JO, Campbell PJ. Criteria for inference of chromothripsis in cancer genomes. Cell 2013;152(6):1226-36.

38. Maher CA, Wilson RK. Chromothripsis and human disease: Piecing together the shattering process. Cell 2012;148(1-2):29-32.

39. Stephens PJ, Greenman CD, Fu B, et al. Massive genomic rearrangement acquired in a single catastrophic event during cancer development. Cell 2011;144(1):27-40.

40. Forment JV, Kaidi A, Jackson SP. Chromothripsis and cancer: Causes and consequences of chromosome shattering. Nat Rev Cancer 2012;12(10):663-70.

41. Suela J, Alvarez S, Cigudosa JC. DNA profiling by arraycgh in acute myeloid leukemia and myelodysplastic syndromes. Cytogenet Genome Res 2007;118(2-4):304-9.

42. Thiel A, Beier M, Ingenhag D, et al. Comprehensive array CGH of normal karyotype myelodysplastic syndromes reveals hidden recurrent and individual genomic copy number alterations with prognostic relevance. Leukemia 2011;25(3):387-99.

43. Merkerova MD, Bystricka D, Belickova M, et al. From cryptic chromosomal lesions to pathologically relevant genes: Integration of snp-array with gene expression profiling in myelodysplastic syndrome with normal karyotype. Genes Chromosomes Cancer 2012;51(5):419-28.

44. Heinrichs S, Kulkarni RV, Bueso-Ramos CE, et al. Accurate detection of uniparental disomy and microdeletions by SNP array analysis in myelodysplastic syndromes with normal cytogenetics. Leukemia 2009;23(9):1605-13.

45. Martínez-Ramírez A, Urioste M, Melchor L, et al. Analysis of myelodysplastic syndromes with complex karyotypes by high-resolution comparative genomic hybridization and subtelomeric CGH array. Genes Chromosomes Cancer 2005;42(3):287-98.

46. O'Keefe CL, Tiu R, Gondek LP, et al. High-resolution genomic arrays facilitate detection of novel cryptic chromosomal lesions in myelodysplastic syndromes. Exp Hematol 2007;35(2):240-51.

47. Tiu RV, Gondek LP, O'Keefe CL, et al. Prognostic impact of SNP array karyotyping in myelodysplastic syndromes and related myeloid malignancies. Blood 2011;117(17):4552-60.

48. Arenillas L, Mallo M, Ramos F, et al. Single nucleotide polymorphism array karyotyping: A diagnostic and prognostic tool in myelodysplastic syndromes with unsuccessful conventional cytogenetic testing. Genes Chromosomes Cancer 2013;52(12):1167-77.

49. Slovak ML, Smith DD, Bedell V, et al. Assessing karyotype precision by microarray-based comparative genomic hybridization in the myelodysplastic/myeloproliferative syndromes. Mol Cytogenet 2010;3:23.

50. Kolquist KA, Schultz RA, Furrow A, et al. Microarray-based comparative genomic hybridization of cancer targets reveals novel, recurrent genetic aberrations in the myelodysplastic syndromes. Cancer Genet 2011;204(11):603-28.

51. Mohamedali A, Gäken J, Twine NA, et al. Prevalence and prognostic significance of allelic imbalance by single-nucleotide polymorphism analysis in low-risk myelodysplastic syndromes. Blood 2007;110(9):3365-73.

52. Wang L, Fidler C, Nadig N, et al. Genome-wide analysis of copy number changes and loss of heterozygosity in myelodysplastic syndrome with del(5q) using high-density single nucleotide polymorphism arrays. Haematologica 2008;93(7):994-1000.

53. Maciejewski JP, Mufti GJ. Whole genome scanning as a cytogenetic tool in hematologic malignancies. Blood 2008;112(4):965-74.

54. Starczynowski DT, Vercauteren S, Telenius A, et al. High-resolution whole genome tiling path array CGH analysis of CD34+ cells from patients with low-risk myelodysplastic syndromes reveals cryptic copy number alterations and predicts overall and leukemia-free survival. Blood 2008;112(8):3412-24.

55. Gondek LP, Tiu R, O'Keefe CL, Sekeres MA, Theil KS, Maciejewski JP. Chromosomal lesions and uniparental disomy detected by SNP arrays in MDS, MDS/MPD, and mds-derived AML. Blood 2008;111(3):1534-42.

56. Akagi T, Ogawa S, Dugas M, et al. Frequent genomic abnormalities in acute myeloid leukemia/myelodysplastic syndrome with normal karyotype. Haematologica 2009;94(2):213-23.

57. Makishima H, Rataul M, Gondek LP, et al. FISH and SNP-A karyotyping in myelodysplastic syndromes: Improving cytogenetic detection of del(5q), monosomy 7, del(7q), trisomy 8 and del(20q). Leuk Res 2010;34(4):447-53.

58. Kloosterman WP, Hoogstraat M, Paling O, et al. Chromothripsis is a common mechanism driving genomic rearrangements in primary and metastatic colorectal cancer. Genome Biol 2011;12(10):R103.

59. Rausch T, Jones DT, Zapatka M, et al. Genome sequencing of pediatric medulloblastoma links catastrophic DNA rearrangements with TP53 mutations. Cell 2012;148(1-2):59-71.

60. Magrangeas F, Avet-Loiseau H, Munshi NC, Minvielle S. Chromothripsis identifies a rare and aggressive entity among newly diagnosed multiple myeloma patients. Blood 2011;118(3):675-8.

61. Meyerson M, Pellman D. Cancer genomes evolve by pulverizing single chromosomes. Cell 2011;144(1):9-10.

62. Holland AJ, Cleveland DW. Chromoanagenesis and cancer: Mechanisms and consequences of localized, complex chromosomal rearrangements. Nat Med 2012, Nov;18(11):1630-8.

63. Jones MJ, Jallepalli PV. Chromothripsis: Chromosomes in crisis. Dev Cell 2012;23(5):908-17.

64. Edelmann J, Holzmann K, Miller F, et al. High-resolution genomic profiling of chronic lymphocytic leukemia reveals new recurrent genomic alterations. Blood 2012;120(24):4783-94.

65. Mackinnon RN, Campbell LJ. Chromothripsis under the microscope: A cytogenetic perspective of two cases of AML with catastrophic chromosome rearrangement. Cancer Genet 2013;206(6):238-51.

66. Nagel S, Meyer C, Quentmeier H, Kaufmann M, Drexler HG, MacLeod RA. Chromothripsis in hodgkin lymphoma. Genes Chromosomes Cancer 2013;52(8):741-7.

67. Boeva V, Jouannet S, Daveau R, et al. Breakpoint features of genomic rearrangements in neuroblastoma with unbalanced translocations and chromothripsis. PLoS One 2013;8(8):e72182.

68. Hirsch D, Kemmerling R, Davis S, et al. Chromothripsis and focal copy number alterations determine poor outcome in malignant melanoma. Cancer Res 2013;73(5):1454-60.

69. Kloosterman WP, Cuppen E. Chromothripsis in congenital disorders and cancer: Similarities

and differences. Curr Opin Cell Biol 2013;25(3):341-8.

70. Li Y, Schwab C, Ryan SL, et al. Constitutional and somatic rearrangement of chromosome 21 in acute lymphoblastic leukaemia. Nature 2014;508(7494):98-102.

71. Kloosterman WP, Koster J, Molenaar JJ. Prevalence and clinical implications of chromothripsis in cancer genomes. Curr Opin Oncol 2014;26(1):64-72.

72. Götze K, Müller-Thomas C, Peschel C. The role of azacitidine in the management of myelodysplastic syndromes (MDS). Cancer Manag Res 2009;1:119-30.

73. Musolino C, Sant'antonio E, Penna G, et al. Epigenetic therapy in myelodysplastic syndromes. Eur J Haematol 2010;84(6):463-73.

74. Navada SC, Steinmann J, Lübbert M, Silverman LR. Clinical development of demethylating agents in hematology. J Clin Invest 2014;124(1):40-6.

75. Garcia-Manero G. Modifying the epigenome as a therapeutic strategy in myelodysplasia. Hematology Am Soc Hematol Educ Program 2007:405-11.

76. Simó-Riudalbas L, Melo SA, Esteller M. DNMT3B gene amplification predicts resistance to DNA demethylating drugs. Genes Chromosomes Cancer 2011;50(7):527-34.

77. Greenberg PL. The multifaceted nature of myelodysplastic syndromes: Clinical, molecular, and biological prognostic features. J Natl Compr Canc Netw 2013;11(7):877-84; quiz 885.

78. Issa JP. The myelodysplastic syndrome as a prototypical epigenetic disease. Blood 2013;121(19):3811-7.

79. Shen L, Kantarjian H, Guo Y, et al. DNA methylation predicts survival and response to therapy in patients with myelodysplastic syndromes. J Clin Oncol 2010;28(4):605-13.

80. Figueroa ME, Skrabanek L, Li Y, et al. MDS and secondary AML display unique patterns and abundance of aberrant DNA methylation. Blood 2009;114(16):3448-58.

81. Figueroa ME, Reimers M, Thompson RF, et al. An integrative genomic and epigenomic approach for the study of transcriptional regulation. PLoS One 2008;3(3):e1882.

82. Esteller M. Aberrant DNA methylation as a cancer-inducing mechanism. Annu Rev Pharmacol Toxicol 2005;45:629-56.

83. Raj K, John A, Ho A, et al. CDKN2B methylation status and isolated chromosome 7 abnormalities predict responses to treatment with 5-azacytidine. Leukemia 2007;21(9):1937-44.

84. Aggerholm A, Holm MS, Guldberg P, Olesen LH, Hokland P. Promoter hypermethylation of p15ink4b, HIC1, CDH1, and ER is frequent in myelodysplastic syndrome and predicts poor prognosis in early-stage patients. Eur J Haematol 2006;76(1):23-32.

85. Aoki E, Uchida T, Ohashi H, et al. Methylation status of the p15ink4b gene in hematopoietic progenitors and peripheral blood cells in myelodysplastic syndromes. Leukemia 2000;14(4):586-93.

86. Hess CJ, Errami A, Berkhof J, et al. Concurrent methylation of promoters from tumor associated genes predicts outcome in acute myeloid leukemia. Leuk Lymphoma 2008;49(6):1132-41.

87. Lin J, Yao DM, Qian J, et al. Methylation status of fragile histidine triad (FHIT) gene and its clinical impact on prognosis of patients with myelodysplastic syndrome. Leuk Res 2008;32(10):1541-5.

88. Valencia A, Cervera J, Such E, et al. Aberrant methylation of tumor suppressor genes in patients with refractory anemia with ring sideroblasts. Leuk Res 2011;35(4):479-83.

89. Iwai M, Kiyoi H, Ozeki K, et al. Expression and methylation status of the FHIT gene in acute myeloid leukemia and myelodysplastic syndrome. Leukemia 2005;19(8):1367-75.

90. Jiang Y, Dunbar A, Gondek LP, et al. Aberrant DNA methylation is a dominant mechanism in MDS progression to AML. Blood 2009;113(6):1315-25.

91. Nolte F, Hofmann W-K. Myelodysplastic syndromes: Molecular pathogenesis and genomic changes. Ann Hematol 2008;87(10):777-95.

92. Rocquain J, Carbuccia N, Trouplin V, 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.

93. Roller A, Grossmann V, Bacher U, et al. Landmark analysis of DNMT3A mutations in hematological malignancies. Leukemia 2013;27(7):1573-8

94. Walter MJ, Ding L, Shen D, et al. Recurrent DNMT3A mutations in patients with

myelodysplastic syndromes. Leukemia 2011;25(7):1153-8.

95. Chen CY, Lin LI, Tang JL, et al. RUNX1 gene mutation in primary myelodysplastic syndrome--the mutation can be detected early at diagnosis or acquired during disease progression and is associated with poor outcome. Br J Haematol 2007;139(3):405-14.

96. Damm F, Chesnais V, Nagata Y, et al. BCOR and BCORL1 mutations in myelodysplastic syndromes and related disorders. Blood 2013;122(18):3169-77.

97. Kulasekararaj AG, Smith AE, Mian SA, et al. TP53 mutations in myelodysplastic syndrome are strongly correlated with aberrations of chromosome 5, and correlate with adverse prognosis. Br J Haematol 2013;160(5):660-72.

98. Thota S, Viny AD, Makishima H, et al. Genetic alterations of the cohesin complex genes in myeloid malignancies. Blood 2014;124(11):1790-8.

99. Thol F, Friesen I, Damm F, et al. Prognostic significance of ASXL1 mutations in patients with myelodysplastic syndromes. J Clin Oncol 2011;29(18):2499-506.

100. Thol F, Kade S, Schlarmann C, et al. Frequency and prognostic impact of mutations in SRSF2, U2AF1, and ZRSR2 in patients with myelodysplastic syndromes. Blood 2012;119(15):3578-84.

101. Pellagatti A, Boultwood J. The molecular pathogenesis of the myelodysplastic syndromes. Eur J Haematol 2015;95(1):3-15.

102. Papaemmanuil E, Cazzola M, Boultwood J, et al. Somatic SF3B1 mutation in myelodysplasia with ring sideroblasts. N Engl J Med 2011;365(15):1384-95.

103. Yoshida K, Sanada M, Shiraishi Y, et al. Frequent pathway mutations of splicing machinery in myelodysplasia. Nature 2011;478(7367):64-9.

104. Gutiérrez NC, López-Pérez R, Hernández JM, et al. Gene expression profile reveals deregulation of genes with relevant functions in the different subclasses of acute myeloid leukemia. Leukemia 2005;19(3):402-9.

105. Quackenbush J. Microarray analysis and tumor classification. N Engl J Med 2006;354(23):2463-72.

106. Mills KI, Kohlmann A, Williams PM, 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.

107. Haferlach T, Kohlmann A, Wieczorek L, et al. Clinical utility of microarray-based gene expression profiling in the diagnosis and subclassification of leukemia: Report from the international microarray innovations in leukemia study group. J Clin Oncol 2010;28(15):2529-37.

108. Bresolin S, Trentin L, Zecca M, et al. Gene expression signatures of pediatric myelodysplastic syndromes are associated with risk of evolution into acute myeloid leukemia. Leukemia 2012;26(7):1717-9.

109. Vasikova A, Belickova M, Budinska E, Cermak J. A distinct expression of various gene subsets in CD34+ cells from patients with early and advanced myelodysplastic syndrome. Leuk Res 2010;34(12):1566-72.

110. Sridhar K, Ross DT, Tibshirani R, Butte AJ, Greenberg PL. Relationship of differential gene expression profiles in CD34+ myelodysplastic syndrome marrow cells to disease subtype and progression. Blood 2009;114(23):4847-58.

111. Theilgaard-Mönch K, Boultwood J, Ferrari S, et al. Gene expression profiling in MDS and AML: Potential and future avenues. Leukemia 2011;25(6):909-20.

112. Wouters BJ, Löwenberg B, Delwel R. A decade of genome-wide gene expression profiling in acute myeloid leukemia: Flashback and prospects. Blood 2009;113(2):291-8.

113. Bullinger L, Döhner K, Bair E, et al. Use of gene-expression profiling to identify prognostic subclasses in adult acute myeloid leukemia. N Engl J Med 2004;350(16):1605-16.

114. Hofmann WK, de Vos S, Komor M, Hoelzer D, Wachsman W, Koeffler HP. Characterization of gene expression of CD34+ cells from normal and myelodysplastic bone marrow. Blood 2002;100(10):3553-60.

115. Del Rey M, O'Hagan K, Dellett M, et al. Genome-wide profiling of methylation identifies novel targets with aberrant hypermethylation and reduced expression in low-risk myelodysplastic syndromes. Leukemia 2013;27(3):610-8.

116. Pellagatti A, Cazzola M, Giagounidis A, et al. Deregulated gene expression pathways in myelodysplastic syndrome hematopoietic stem cells. Leukemia 2010;24(4):756-64.

117. Boultwood J, Pellagatti A, Nikpour M, et al. The role of the iron transporter ABCB7 in refractory

anemia with ring sideroblasts. PLoS One 2008;3(4):e1970.

118. Del Rey M, Benito R, Fontanillo C, et al. Deregulation of genes related to iron and mitochondrial metabolism in refractory anemia with ring sideroblasts. PLoS One 2015;10(5):e0126555.

119. Chen G, Zeng W, Miyazato A, et al. Distinctive gene expression profiles of CD34 cells from patients with myelodysplastic syndrome characterized by specific chromosomal abnormalities. Blood 2004;104(13):4210-8.

120. Boultwood J, Pellagatti A, Cattan H, et al. Gene expression profiling of CD34+ cells in patients with the 5q- syndrome. Br J Haematol 2007;139(4):578-89.

121. Pellagatti A, Hellström-Lindberg E, Giagounidis A, et al. Haploinsufficiency of RPS14 in 5q- syndrome is associated with deregulation of ribosomal-and translation-related genes. Br J Haematol 2008;142(1):57-64.

122. Nikpour M, Pellagatti A, Liu A, et al. Gene expression profiling of erythroblasts from refractory anaemia with ring sideroblasts (RARS) and effects of G-CSF. Br J Haematol 2010;149(6):844-54.

123. Visconte V, Tiu RV, Rogers HJ. Pathogenesis of myelodysplastic syndromes: An overview of molecular and non-molecular aspects of the disease. Blood Res 2014;49(4):216-27.

124. Garcia-Manero G, Fenaux P. Hypomethylating agents and other novel strategies in myelodysplastic syndromes. J Clin Oncol 2011;29(5):516-23.

125. Ridgeway JA, Fechter L, Murray C, Borràs N. Update on the science of myelodysplastic syndromes. Clin J Oncol Nurs 2012;16 Suppl:9-22.

126. Itzykson R, Thépot S, Quesnel B, et al. Prognostic factors for response and overall survival in 282 patients with higher-risk myelodysplastic syndromes treated with azacitidine. Blood 2011;117(2):403-11.

127. Silverman LR, Demakos EP, Peterson BL, et al. Randomized controlled trial of azacitidine in patients with the myelodysplastic syndrome: A study of the cancer and leukemia group B. J Clin Oncol 2002;20(10):2429-40.

128. Kornblith AB, Herndon JE, Silverman LR, et al. Impact of azacytidine on the quality of life of patients with myelodysplastic syndrome treated in a randomized phase III trial: A cancer and leukemia group B study. J Clin Oncol 2002;20(10):2441-52.

129. Fenaux P, Mufti GJ, Hellstrom-Lindberg E, et al. Efficacy of azacitidine compared with that of conventional care regimens in the treatment of higher-risk myelodysplastic syndromes: A randomised, open-label, phase III study. Lancet Oncol 2009;10(3):223-32.

130. Fenaux P, Mufti GJ, Hellström-Lindberg E, et al. Azacitidine prolongs overall survival compared with conventional care regimens in elderly patients with low bone marrow blast count acute myeloid leukemia. J Clin Oncol 2010;28(4):562-9.

131. Bacher U, Haferlach T, Schnittger S, et al. Investigation of 305 patients with myelodysplastic syndromes and 20q deletion for associated cytogenetic and molecular genetic lesions and their prognostic impact. Br J Haematol 2014;164(6):822-33.

132. Gunn SR, Mohammed MS, Gorre ME, et al. Whole-genome scanning by array comparative genomic hybridization as a clinical tool for risk assessment in chronic lymphocytic leukemia. J Mol Diagn 2008;10(5):442-51.

133. Stankiewicz P, Lupski JR. Structural variation in the human genome and its role in disease. Annu Rev Med 2010;61:437-55.

134. Haraksingh RR, Abyzov A, Gerstein M, Urban AE, Snyder M. Genome-wide mapping of copy number variation in humans: Comparative analysis of high resolution array platforms. PLoS One 2011;6(11):e27859.

135. Przybytkowski E, Lenkiewicz E, Barrett MT, et al. Chromosome-breakage genomic instability and chromothripsis in breast cancer. BMC Genomics 2014;15:579.

136. Salaverria I, Martín-Garcia D, López C, et al. Detection of chromothripsis-like patterns with a custom array platform for chronic lymphocytic leukemia. Genes Chromosomes Cancer 2015;54(11):668-80.

137. McEvoy J, Nagahawatte P, Finkelstein D, et al. RB1 gene inactivation by chromothripsis in human retinoblastoma. Oncotarget 2014;5(2):438-50.

138. Di Fiore R, D'Anneo A, Tesoriere G, Vento R. RB1 in cancer: Different mechanisms of RB1 inactivation and alterations of pRb pathway in tumorigenesis. J Cell Physiol 2013;228(8):1676-87.

139. Lee H. Cycling with BRCA2 from DNA repair to mitosis. Exp Cell Res 2014;329(1):78-84.

140. Crasta K, Ganem NJ, Dagher R, Lantermann AB, Ivanova EV, Pan Y, et al. DNA breaks and

chromosome pulverization from errors in mitosis. Nature 2012;482(7383):53-8.

141. Zemanova Z, Michalova K, Buryova H, et al. Involvement of deleted chromosome 5 in complex chromosomal aberrations in newly diagnosed myelodysplastic syndromes (MDS) is correlated with extremely adverse prognosis. Leuk Res 2014;38(5):537-44.

142. Gerstung M, Pellagatti A, Malcovati L, et al. Combining gene mutation with gene expression data improves outcome prediction in myelodysplastic syndromes. Nat Commun 2015;6:5901.

143. Zhou T, Hasty P, Walter CA, Bishop AJ, Scott LM, Rebel VI. Myelodysplastic syndrome: An inability to appropriately respond to damaged DNA? Exp Hematol 2013;41(8):665-74.

144. Zhou T, Chen P, Gu J, et al. Potential relationship between inadequate response to DNA damage and development of myelodysplastic syndrome. Int J Mol Sci 2015;16(1):966-89.

145. Kefala M, Papageorgiou SG, Kontos CK, et al. Increased expression of phosphorylated NBS1, a key molecule of the DNA damage response machinery, is an adverse prognostic factor in patients with de novo myelodysplastic syndromes. Leuk Res 2013;37(11):1576-82.

146. Rübe CE, Fricke A, Widmann TA, et al. Accumulation of DNA damage in hematopoietic stem and progenitor cells during human aging. PLoS One 2011;6(3):e17487.

147. Parker JE, Mufti GJ, Rasool F, Mijovic A, Devereux S, Pagliuca A. The role of apoptosis, proliferation, and the Bcl-2-related proteins in the myelodysplastic syndromes and acute myeloid leukemia secondary to MDS. Blood 2000;96(12):3932-8.

148. Slape CI, Saw J, Jowett JB, et al. Inhibition of apoptosis by BCL2 prevents leukemic transformation of a murine myelodysplastic syndrome. Blood 2012;120(12):2475-83.

149. Golomb L, Volarevic S, Oren M. P53 and ribosome biogenesis stress: The essentials. FEBS Lett 2014;588(16):2571-9.

150. De Las Heras-Rubio A, Perucho L, Paciucci R, Vilardell J, LLeonart ME. Ribosomal proteins as novel players in tumorigenesis. Cancer Metastasis Rev 2014;33(1):115-41.

151. Diesch J, Hannan RD, Sanij E. Perturbations at the ribosomal genes loci are at the centre of cellular dysfunction and human disease. Cell Biosci 2014;4:43.

152. Woods SJ, Hannan KM, Pearson RB, Hannan RD. The nucleolus as a fundamental regulator of the p53 response and a new target for cancer therapy. Biochim Biophys Acta 2015;1849(7):821-9.

153. Zhou X, Liao WJ, Liao JM, Liao P, Lu H. Ribosomal proteins: Functions beyond the ribosome. J Mol Cell Biol 2015;7(2):92-104.

154. Galili N, Qasim SA, Raza A. Defective ribosome biogenesis in myelodysplastic syndromes. Haematologica 2009;94(10):1336-8.

155. Lindström MS. Emerging functions of ribosomal proteins in gene-specific transcription and translation. Biochem Biophys Res Commun 2009;379(2):167-70.

156. Zhang Y, Lu H. Signaling to p53: Ribosomal proteins find their way. Cancer Cell 2009;16(5):369-77.

157. Warner JR, McIntosh KB. How common are extraribosomal functions of ribosomal proteins? Mol Cell 2009;34(1):3-11.

158. James A, Wang Y, Raje H, Rosby R, DiMario P. Nucleolar stress with and without p53. Nucleus 2014;5(5):402-26.

159. Grummt I. The nucleolus—guardian of cellular homeostasis and genome integrity. Chromosoma 2013;122(6):487-97.

160. Bursac S, Brdovcak MC, Donati G, Volarevic S. Activation of the tumor suppressor p53 upon impairment of ribosome biogenesis. Biochim Biophys Acta 2014;1842(6):817-30.

161. Owens BM, Hawley RG. HOX and non-HOX homeobox genes in leukemic hematopoiesis. Stem Cells 2002;20(5):364-79.

162. Drabkin HA, Parsy C, Ferguson K, et al. Quantitative HOX expression in chromosomally defined subsets of acute myelogenous leukemia. Leukemia 2002;16(2):186-95.

163. Sitwala KV, Dandekar MN, Hess JL. HOX proteins and leukemia. International Journal of Clinical and Experimental Pathology 2008;1(6):461.

164. Argiropoulos B, Yung E, Humphries RK. Unraveling the crucial roles of Meis1 in leukemogenesis and normal hematopoiesis. Genes Dev 2007;21(22):2845-9.

165. Huang Y, Sitwala K, Bronstein J, et al. Identification and characterization of Hoxa9 binding sites in hematopoietic cells. Blood 2012;119(2):388-98.

166. Dickson GJ, Kwasniewska A, Mills KI, Lappin TR, Thompson A. Hoxa6 potentiates short-term

hemopoietic cell proliferation and extended self-renewal. Exp Hematol 2009;37(3):322-33.e3.

167. Nakamura T. Meis and Hox: a mighty pair defeats apoptosis. Blood 2005;105(3):909-10.

168. Aggarwal S, van de Loosdrecht AA, Alhan C, Ossenkoppele GJ, Westers TM, Bontkes HJ. Role of immune responses in the pathogenesis of low-risk MDS and high-risk MDS: Implications for immunotherapy. Br J Haematol 2011;153(5):568-81.

169. Maeda T, Wakasawa T, Shima Y, Tsuboi I, Aizawa S, Tamai I. Role of polyamines derived from arginine in differentiation and proliferation of human blood cells. Biol Pharm Bull 2006;29(2):234-9.

170. Galm O, Herman JG, Baylin SB. The fundamental role of epigenetics in hematopoietic malignancies. Blood Rev 2006;20(1):1-13.

171. Martin-Subero JI, Ammerpohl O, Bibikova M, et al. A comprehensive microarray-based DNA methylation study of 367 hematological neoplasms. PLoS One 2009;4(9):e6986.

172. Fandy TE, Herman JG, Kerns P, et al. Early epigenetic changes and DNA damage do not predict clinical response in an overlapping schedule of 5-azacytidine and entinostat in patients with myeloid malignancies. Blood 2009;114(13):2764-73.

173. Bejar R, Lord A, Stevenson K, et al. TET2 mutations predict response to hypomethylating agents in myelodysplastic syndrome patients. Blood 2014;124(17):2705-12.

174. Itzykson R, Kosmider O, Cluzeau T, et al. Impact of TET2 mutations on response rate to azacitidine in myelodysplastic syndromes and low blast count acute myeloid leukemias. Leukemia 2011;25(7):1147-52.

List of Genes

APC	Adenomatosis polyposis coli
ASXL1	Additional sex combs like transcriptional regulator 1
ATF2	Activating transcription factor 2
ATRX	Alpha thalassemia/mental retardation syndrome X-linked
BCL2	B-cell CLL/lymphoma 2
BCOR	BCL6 corepressor
BCORL1	BCL6 corepressor-like 1
BRAF	B-Raf proto-oncogene, serine/threonine kinase
BRCA2	Breast cancer 2
CAMP	Cathelicidin antimicrobial peptide
CBL	Cbl proto-oncogene
CCNG1	Cyclin G1
CD44	CD44 molecule (Indian blood group)
CDH1	Cadherin 1
CDH13	Cadherin 13
CDKN1B	Cyclin-dependent kinase inhibitor 1B
CDKN2A	Cyclin-dependent kinase inhibitor 2A
CDKN2B (P15ink4b)	Cyclin-dependent kinase inhibitor 2B
CEACAM3	Carcinoembryonic antigen-related cell adhesion molecule 3
CEACAM6	Carcinoembryonic antigen-related cell adhesion molecule 6
CEACAM8	Carcinoembryonic antigen-related cell adhesion molecule 8
CEBPA	CCAAT/enhancer binding protein (C/EBP), alpha
CLEC4D	C-type lectin domain family 4, member d
CLEC4E	C-type lectin domain family 4, member e
CREBBP	CREB binding protein
CRISP3	Cysteine-rich secretory protein 3
CUX1	Cut-like homeobox 1
DAPK1	Death-associated protein kinase 1
DHX30	DEAH (Asp-Glu-Ala-His) box helicase 30
DNMT3A	DNA (cytosine-5-)-methyltransferase 3 alpha
DNMT3B	DNA (cytosine-5-)-methyltransferase 3 beta
EEF2	Eukaryotic translation elongation factor 2
EIF3K	Eukaryotic translation initiation factor 3, subunit K
EP300	E1A binding protein p300
ESR1	Estrogen receptor 1
ETS1	v-ets avian erythroblastosis virus E26 oncogene homolog 1
ETV6	Ets variant 6
EVI1	MDS1 And EVI1 Complex Locus
EZH2	enhancer of zeste 2 polycomb repressive complex 2 subunit
FANCA	Fanconi anemia, complementation group A
FANCD2	Fanconi anemia, complementation group D2
FHIT	fragile histidine triad
FLT1	Fms-related tyrosine kinase 1
FLT3	fms-related tyrosine kinase 3
GATA2	GATA binding protein 2
GNAS	GNAS complex locus
GSTP1	Glutathione S-transferase Pi-1
H3F3B	H3 histone, family 3B (H3.3B)
HIC1	Hypermethylated in cancer 1
HIST1H2BD	Histone cluster 1, H2bd
HOXA3	Homeobox A3
HOXA7	Homeobox A7
HOXA9	Homeobox A9
HOXB2	Homeobox B2
HOXB3	Homeobox B3
IDH1	Isocitrate dehydrogenase 1
IDH2	Isocitrate dehydrogenase 2
IGSF4 (CADM1)	Cell adhesion molecule 1
IRF1	Interferon regulatory factor 1

JAK2	Janus kinase 2
JUN	Jun proto-oncogene
KDM6A	Lysine (K)-specific demethylase 6A
KIT	v-kit Hardy-Zuckerman 4 feline sarcoma viral oncogene
KRAS	Kirsten rat sarcoma viral oncogene homolog
LTF	Lactotransferrin
MEIS1	Meis homeobox 1
MLH1	MutL homolog 1
MMP25	Matrix metallopeptidase 25
MMP27	Matrix metallopeptidase 27
MMP8	Matrix metallopeptidase 8
MMP9	Matrix metallopeptidase 9
MPL	MPL proto-oncogene, thrombopoietin receptor
MRE11A	MRE11 homolog A, double strand break repair nuclease
MRPS12	Mitochondrial ribosomal protein S12
MYST1	K(Lysine) Acetyltransferase 8
NF1	Neurofibromin 1
NPM1	Nucleophosmin
NRAS	Neuroblastoma RAS viral (v-ras) oncogene homolog
PDCD7	Programmed cell death 7
PGLYRP1	Peptidoglycan recognition protein 1
PMAIP1 (NOXA)	Phorbol-12-myristate-13-acetate-induced protein 1
PRPF40	Pre-mRNA processing factor 40 homolog
PRPF8	Pre-mRNA processing factor 8
PTEN	Phosphatase and tensin homolog
PTPN11	Protein tyrosine phosphatase, non-receptor type 11
PUF60	Poly-U binding splicing factor 60KDa
RAD21	RAD21 cohesin complex component
RAD51	RAD51 Recombinase
RARB	Retinoic acid receptor, beta
RASSF1	Ras association (RalGDS/AF-6) domain family member 1
RB1	Retinoblastoma 1
RCC1	Regulator Of Chromosome Condensation 1
RPL12	Ribosomal Protein L12
RPL13	Ribosomal Protein L13
RPL13A	Ribosomal Protein L13a
RPL15	Ribosomal Protein L15
RPL18	Ribosomal Protein L18
RPL19	Ribosomal Protein L19
RPL22	Ribosomal Protein L22
RPL28	Ribosomal Protein L28
RPL3	Ribosomal Protein L3
RPL37	Ribosomal Protein L37
RPL7a	Ribosomal Protein L7a
RPS14	Ribosomal Protein S14
RPS19	Ribosomal Protein S19
RPS3	Ribosomal Protein S3
RPS5	Ribosomal Protein S5
RPS6	Ribosomal Protein S6
RUNX1	Runt-related transcription factor 1
S100A12	S100 calcium binding protein A12
SETBP1	SET binding protein 1
SF1	Splicing factor 1
SF3A1	Splicing factor 3a, subunit 1
SF3B1	Splicing factor 3b, subunit 1
SIRPA	Signal-regulatory protein alpha
SIRPB1	Signal-regulatory protein beta 1
SMC1A	Structural maintenance of chromosomes 1A
SMC3	Structural maintenance of chromosomes 3
SPARC	Secreted protein, acidic, cysteine-rich (osteonectin)
SRSF2	Serine/arginine-rich splicing factor 2
STAG2	Stromal antigen 2
TAF7	TATA binding protein associated factor 7
TERT	Telomerase
TET2	Tet methylcytosine dioxygenase 2
TIMP3	TIMP metallopeptidase inhibitor 3
TLK1	Tousled-like kinase 1
TP53	Tumor protein p53
TP53BP1	Tumor protein p53 binding protein 1

TP73	Tumor protein p73	***WT1***	Wilms tumor 1
U2AF1	U2 small nuclear RNA auxiliary factor 1	***XPO4***	Exportin 4
U2AF65	U2 small nuclear RNA auxiliary factor large subunit	***ZRSR2***	Zinc finger (CCCH type), RNA-binding motif and serine/arginine rich 2

List of Tables and Figures

INTRODUCTION

RESULTS - CHAPTER 1

RESULTS - CHAPTER 2

RESULTS - CHAPTER 3

DISCUSSION

Supplementary Appendix

Chromothripsis is a recurrent genomic abnormality in high-risk myelodysplastic syndromes

María Abáigar,[1] Cristina Robledo,[1] Rocío Benito,[1] Fernando Ramos,[2,3] María Díez-Campelo,[4] Lourdes Hermosín,[5] Javier Sánchez-del-Real,[3] Jose M Alonso,[6] Rebeca Cuello,[7] Marta Megido,[8] Juan N Rodríguez,[9] Guillermo Martín-Núñez,[10] Carlos Aguilar,[11] Manuel Vargas,[12] Ana A Martín,[4] Juan L García,[1] Alexander Kohlmann,[13] M Consuelo del Cañizo,[4,14] and Jesús M Hernández-Rivas[1,4,14]

SUPPLEMENTARY METHODS

Patients

All patients were classified according to the 2008 World Health Organization criteria (Vardiman et al., 2009). The following morphological subtypes were included: MDS associated with isolated del(5q) (MDS del(5q), n=7), refractory cytopenia with unilineage dysplasia (RCUD, n=20), refractory cytopenia with multilineage dysplasia (RCMD, n=147), refractory anemia with ringed sideroblasts (RARS, n=11), refractory anemia with excess of blasts type 1 (RAEB-1, n=23), RAEB-2 (n=23), MDS unclassified (MDS-U, n=9), CMML (n=58), and RARS with thrombocytosis (RARS-T, n=3) (Table 1).

Array-based comparative genomic hybridization studies

Genomic DNA (gDNA) was isolated using the QIAamp DNA Mini Kit (Qiagen, Hilden, Germany) according to the manufacturer's standard protocol. gDNA concentration and quality were assessed with a NanoDrop spectrophotometer (ND-1000; NanoDrop Technologies, Wilmington, DE) by measuring the ratio of absorbance at 260 and 280 nm (A260/280), and also at 230 nm (A260/230). gDNA integrity (degradation) was evaluated by running a 1% agarose gel with ethidium bromide visualization.

DNA copy number abnormalities (CNAs) were studied in all samples with the Human CGH 12x135K Whole-Genome Tiling v3.0 Array (Roche NimbleGen). This platform contains around 135,000 60-mer oligonucleotide probes, with a median probe spacing of approximately 12 kb (12,524 bp) (NCBI Build 36.1; UCSC hg18, March 2006) across the human genome. Pooled male or female human commercial DNA samples were used as reference (Promega, Mannheim, Germany). Sample preparation and hybridization were performed following the NimbleGen CGH array standard protocol. Briefly, 500 ng of gDNA from each patient and sex-matched reference sample were labeled in parallel with Cy3 and Cy5, respectively. Subsequently, 20 µg of each labeled test and the corresponding sex-matched reference DNA were mixed and co-hybridized to the microarray overnight at 42°C. After hybridization, slides were washed and scanned at 2-µm resolution using the NimbleGen MS 200 Microarray Scanner (Roche NimbleGen) (Robledo et al., 2011). Raw data were extracted from the scanned images and analyzed using the segMNT algorithm from NimbleScan software (version2.6; Roche NimbleGen), which divides aCGH data into segments consisting of regions of homogenous copy number. This software produces a variable called mad1.dr, which represents the median absolute deviation of the $\log_2$ ratio difference between consecutive probes along the chromosome. This parameter was used as a measure of the noise in the hybridization and to create sample-specific cut-offs for the detection of CNAs. An objective cut-off value above (gains) and below (losses) the mad1.dr parameter and a simultaneous visual examination of whole-genome view ratio plots using SignalMap (version1.9; Roche NimbleGen) were used to distinguish regions of gain or loss. Only CNAs larger than 500 kb were considered. Genome location and size of all CNAs detected by aCGH were converted to hg38 assembly (GRCh Build 38; UCSC hg38, December 2013) (http://genome.ucsc.edu/cgi-bin/hgLiftOver).

Next-generation sequencing studies

We applied amplicon-based next-generation sequencing using 454 Titanium amplicon chemistry (Roche Applied Science) in selected cases to investigate mutations occurring in *DNMT3A*, *TET2*, *RUNX1*, *TP53* and *BCOR* genes. The complete coding region of *TET2*, *RUNX1* and *BCOR*, exons 7-23 of *DNMT3A* and exons 4-11 of *TP53* were covered by 27, 7, 29, 16 and 8 amplicons, respectively (Supplementary Table S3). Amplicon libraries were prepared following the manufacturer's recommendations and previously described methods (Kohlmann et al., 2011). In brief, individual amplicons were amplified in five preconfigured 96-well primer plates using the FastStart High Fidelity PCR System kit and GC-RICH PCR System kit (Roche Applied Science). After amplification, all amplicons of each plate were individually purified using Agencourt AMPure XP beads (Beckman Coulter, Krefeld, Germany), in order to remove short fragments, quantified with the Quant-iT PicoGreen dsDNA kit (Invitrogen, Carlsbad, CA, USA), adjusted to a final concentration of $1x10^9$ molecules/µL and combined in an equimolar ratio to generate the corresponding amplicon pools. Subsequently, each pool was adjusted to a final concentration of $2x10^6$ molecules/µL. Further, 454 Life Sciences NGS steps, such as the emulsion PCR, breaking of the emulsions, enrichment of beads carrying amplified DNA, loading DNA beads on a PicoTiterPlate and sequencing were performed following the manufacturer's recommendations (Roche Applied Science). Multiple molecular barcodes were used to identify each patient. Gene and transcript IDs, information on primer sequences and the respective amplification mixes and cycler protocols are given below (Supplementary Tables S2, S3, S4 and S5).

All sequencing data were generated during several runs on the 454 Sequencing Systems, GS FLX and GS Junior, and processed using GS Run Browser software (version 2.9; Roche Applied Science). All amplicon reads were aligned against the reference transcripts (Supplementary Table S2) of the corresponding genes (*DNMT3A*, *TET2*, *RUNX1*, *TP53* and *BCOR*) for variant detection using the Sequence Pilot software (version 3.5.2; JSI medical systems) and GS Amplicon Variant Analyzer Software (version 2.9; Roche Applied Science). For the detection of variants, filters were set to display sequence variants occurring in >2% of bidirectional reads per amplicon in at least one patient (Kohlmann et al., 2011).

SUPPLEMENTARY TABLES

Supplementary Table S1. All genomic copy number changes detected by aCGH and all conventional cytogenetic information for the whole series. (Available in the CD)

Supplementary Table S2. Ensembl gene and transcript IDs for the five genes selected for NGS.

Gene Symbol	Gene ID	Transcript ID
DNMT3A	ENSG00000119772	ENST00000264709
RUNX1	ENSG00000159216	ENST00000344691
TET2	ENSG00000168769	ENST00000380013
TP53	ENSG00000141510	ENST00000269305
BCOR	ENSG00000183337	ENST00000378444

Supplementary Table S3. PCR primer-pair sequences for all amplicons representing *DNMT3A* (16), *RUNX1* (7), *TET2* (27), *TP53* (8) and *BCOR* (29) genes.

Gene	Exon/Amplicon	Forward Sequence 5' -> 3'	Reverse Sequence 5' -> 3'	Length	PCR Mix	PCR Protocol
DNMT3A	7	TTCCTGGAGAGGTCAAGGTG	TGGAGAGAGGAGAGCAGGAC	345	1	1
DNMT3A	8	GCCTCGTGACCACTGTGTAA	ACCCACCACAGGCAGAGTAG	341	1	1
DNMT3A	9	CTCCTCTTTGCATCGGGTAA	ACCTGCACTCCAACTTCCAG	335	1	1
DNMT3A	10	TGTGCCACCCTCACTACTCA	TCCCTAAGCATGGCTTTCC	334	1	1
DNMT3A	11-12	GACCTTGGCACCTGCTTTC	CCACACTAGGAGTGCCAGAGTT	358	1	1
DNMT3A	13	GGTCACAGTGCCTCCCTTT	ACCCTGTACATGCCCAGAAG	332	1	1
DNMT3A	14	CACAGGCAGATGAGGTTTCC	CCCAGCTAAGGAGACCACTG	332	1	1
DNMT3A	15	CCCTAGCCATGCTCCAGAC	CCCACAACCAAGGCTCAG	345	1	1
DNMT3A	16	CAGGGTGTGTGGGTCTAGGA	TGCATACGTTTCCACTTCACA	343	1	1
DNMT3A	17	AAAGATAGGACTTGGGCCTACA	CTGCCTCCAGGTGCTGAG	332	1	1
DNMT3A	18	TGGTCCCGTTCTTGTTTAGG	CAAGGAGGAAGCCTATGTGC	339	1	1
DNMT3A	19	GACAGCTATTCCCGATGACC	GCTCCACAATGCAGATGAGA	346	1	1
DNMT3A	20	TGTGTGGCTCCTGAGAGAGA	CATGGCAGAGCAGCTAGTCA	335	1	1
DNMT3A	21	TGGTGGATTTGTGTCTTTGC	CATCCTGCCCTTCCTTCTC	337	1	1
DNMT3A	22	CTGCGAACTCTGCTCACTCA	AGCAAGCACAGCAATCAGAA	333	1	1
DNMT3A	23	CACTCACCCTGCCCTCTCT	AAAGCCCTCCGGTATTTCC	345	1	1
RUNX1	3	GCTGTTTGCAGGGTCCTAAC	GGCCTCCGCCTGTCCTC	348	2	2
RUNX1	4	CATTGCTATTCCTCTGCAACC	GTTTGTTGCCATGAAACGTG	342	2	2
RUNX1	5	AAATTCCGGGAGTGTTGTCA	GAAAGGTTGAACCCAAGGAA	341	2	2
RUNX1	6	TGATCTCTTCCCTCCCTCCT	CAGTTGGTCTGGGAAGGTGT	348	2	2
RUNX1	7	ATTTGAACAAGGGCCACTCA	AATGTTCTGCCAACTCCTTCA	342	3	2
RUNX1	8.01	CTCCGCAACCTCCTACTCAC	CCCACCATGGAGAACTGGTA	342	3	2
RUNX1	8.02	CCCGTTCCAAGCCAGCTC	GCTTGTCGCGAACAGGAG	342	3	2
TET2	3.01	ATTCAACTAGAGGGCAGCCTTG	ACTGTGCGTTTTATTCCTCCAT	338	1	1
TET2	3.02	GAATACCCTGTATGAAGGGAAGC	CCCACTGCAGTTATGTGTTGAA	335	1	1
TET2	3.03	TGTAGCCCAAGAAAATGCAG	TGGGTGAGTGATCTCACAGG	343	1	1
TET2	3.04	CATCTCACATAAATGCCATTAACA	AGCTTGCAAATTGCTGCTG	350	1	1
TET2	3.05	GAAAATAACATCCAGGGAACCA	CCCTCTATTTTCACTTCCCTTAAA	350	1	1
TET2	3.06	GGAGTTTTAGAAGAACACCACCA	TCGACCCTTCAGAATCTCTTG	348	1	1
TET2	3.07	CCAATTTTTGGTAGCAGTGGA	CCAGCTGTGTTGTTTTCTGG	334	1	1
TET2	3.08	TGACCTCCAAACAATACACTGG	TGAGTTTGAAAATGGCTCAGTC	350	1	1
TET2	3.09	CCCAGTGTTGAAACAGCA	ACTTCCTCCAGTCCCATTTG	339	1	1
TET2	3.10	TGGTGAAAATCAGTATTCAAAATCA	CCCTGTAGAACTGAAGCTTGTTG	336	1	1
TET2	3.11	CTTCTTCACAGGTGCTTTCAAG	ATACAGGCATGTGGCTTGC	348	1	1
TET2	3.12	TTGCCATAGTCAGATGCACAG	CTGAAGAAGTTGTTTGCTGCTCT	350	1	1
TET2	3.13	TTGACTAGACAAACCACTGCTG	TTTATGAGCCTTTACAAATTGCTG	343	1	1
TET2	4	TGGCACATTTTCTAATAGATCAGTC	CTTTGTGTGTGAAGGCTGGA	344	1	1
TET2	5	AAACCGTTCATTTCTCAGGATG	GTAATGTTCTTTTTAACTGGCATGA	335	1	1

Supplementary Table S3 *(Continued)*

Gene	Exon/Amplicon	Forward Sequence 5' -> 3'	Reverse Sequence 5' -> 3'	Length	PCR Mix	PCR Protocol
TET2	6	TGACCCTTGTTTTGTTTTGG	CGCTGAACTCTCTTCCTTTCA	355	1	1
TET2	7	ATAGACACCTATAATATCAGCTGCAC	CAGTTTGGGAAAAACTTTGATTA	349	1	1
TET2	8	CCATATATTGTGTTTGGGATTCAA	GCAGTGGTTTCAACAATTAAGAG	337	1	1
TET2	9	TGCTCTATTTTGTGTCATTCCATT	CAGTGTGAGAACAGACTCAACAG	341	1	1
TET2	10.01	GGGACCTGTAGTTGAGGCTGT	GGGGCTGACTTTTCCTTTTC	347	1	1
TET2	10.02	GAGTTTGGGAGTGTGGAAGC	GGGGGCAAAACCAAAATAAT	335	1	1
TET2	11.01	GCCTTCATAAAATAATCATCAACA	CTGCAGCTTGAGATGAGGTG	344	1	1
TET2	11.02	CCAATCCAGTTAGTCCTTATCCA	AAAACTCTGGCTATTTCCAAACC	350	1	1
TET2	11.03	CAAGCCAAGACCCTCTGTCT	GCATGAAGAGAGCTGTTGAA	352	1	1
TET2	11.04	GGTGAACATCATTCACCTTCTC	GAATTGACCCATGAGTTGGAG	355	1	1
TET2	11.05	AGACAGCGAGCAGAGCTTTC	AAGTTTCATGTGGCTCAGCA	335	1	1
TET2	11.06	AGCCCGTGAGAAAGAGGAAG	ACTGTGACCTTTCCCCACTG	336	1	1
TP53	4	ACCTGGTCCTCTGACTGCTC	CAGGCATTGAAGTCTCATGG	361	4	1
TP53	5	CACTTGTGCCCTGACTTTCA	CACTCGGATAAGATGCTGAGG	343	4	1
TP53	6	CAGATAGCGATGGTGAGCAG	TTGCACATCTCATGGGGTTA	335	4	1
TP53	7	GCACTGGCCTCATCTTGG	AAGAGGTCCCAAAGCCAGAG	334	4	1
TP53	8	GGACAGGTAGGACCTGATTTC	TCTCCATCCAGTGGTTTCTTC	346	4	1
TP53	9	AAAGGGGAGCCTCACCAC	TGTCTTTGAGGCATCACTGC	342	4	1
TP53	10	GCTGTATAGGTACTTGAAGTGCAG	CTGCCTTTGACCATGAAGG	349	4	1
TP53	11	AGGGAAAAGGGGCACAG	CCCCACAACAAAACACCAGT	339	4	1
BCOR	2	CTGTAGACAGCAGGCACAGC	CACGGTGGCTGTGAGAAGT	341	1	1
BCOR	3	GGGAGATCTGTGAAAGATGGA	CAGGAAGGCAAGACTGGAAG	333	1	1
BCOR	4.01	TAAGGCTGATGGCATGTTGA	GCAGAAGCCTCCACTGTCTC	338	1	1
BCOR	4.02	GTCCCGGGAAACATCGTC	CACCCTCCATGTAAGGATTGA	346	1	1
BCOR	4.03	CAAACAGAGCCCTCTCAACA	CTGACGCCCATCTTCCAC	360	1	1
BCOR	4.04	GAGGCTCTCGACACCTTCG	TATGGCTTTGACAGGGCAAC	333	1	1
BCOR	4.05	ACTCGGAGTTCCACAAGCAC	CACTTCCAGCCCTGCTGT	360	1	1
BCOR	4.06	AGATGCTTCCAAAGCTGACC	CACTCGACACTGACCCTGAA	341	1	1
BCOR	4.07	CACAGCAGCGGAGTTCATC	ACTCCTGGGGTAGGGAATTG	359	1	1
BCOR	4.08	GGCCTTCCACCAAGCTCTAT	GACCGGGATCTCCTCTCTG	353	1	1
BCOR	4.09	TACCACACACGCCCATAGAG	GCTCAACTGAGGGCTTGG	342	1	1
BCOR	4.1	CCAGCTTTGCAGCAGAGAG	GGGGTCACATCCACACTTG	340	1	1
BCOR	4.11	AACCCAAGAGGATCCCAAAC	GTGCCCTGCCCTACCATAC	342	1	1
BCOR	5	AGGTTGCTTAAAGGGATAGAGTATGT	CACAAACTTCCCTTTGTATATTTGG	350	1	1
BCOR	6	TTTTCCATCTCCGTTCTCTTTC	TGCCACCATTCATAAGGACA	333	1	1
BCOR	7.01	CAGATCCTGCCTTCTCATGG	GCAGACAGGGCTCACCTTTA	344	1	1
BCOR	7.02	CCAGGCGAGTATAGTGTTGGA	TCCACATCTCCTGTCCATCC	347	1	1
BCOR	8.01	GGGGGCTTTTTGATTAATTTG	ACCAGCTTCTGTTGCCTTTG	346	1	1
BCOR	8.02	CAACCTGAAGGTGTGCATTG	GCCCTTTCCTACAGAGAGTGG	346	1	1
BCOR	9.01	GTAACGCCCGCTTCTTCTC	TGGACTCCTGAGGGATCAAG	337	1	1
BCOR	9.02	AAAAGCAGGCTCAGCCAAG	AAAGCTTTGTCCCCTCAAGC	348	1	1
BCOR	10.01	GAAGCAGCGGCTGTTAGAAG	CCCACAGGGACACACCTC	331	1	1
BCOR	10.02	TGTCCGGAGATTCAGAAAGC	GAGCTCTGCTCTCCATGTCC	344	1	1
BCOR	11	GCAGTGATCTGATGGCATTC	CCCAGTGGTCAGTGTGCTTA	342	1	1
BCOR	12	AAGTGTGGCTTGCATGAGC	AGTTCAAACACTCGGCTGCT	331	1	1
BCOR	13	CCTTCCAGCCTGTCATGAAT	TCCACAGTTGTCAACAACAACA	348	1	1
BCOR	14	GCAGGACTCACTTGGGAAAG	CTGCTCAAAGCGCATTTCTA	350	1	1
BCOR	15.01	TGGGAGCAGAATGAAAGGAC	GAGCCCAGCAGAGTCTGAAT	358	1	1
BCOR	15.02	GGCAGGTTTCTGCAAGTCTC	ATGTTTCTAAGTCCTTCTGACTGGA	331	1	1

Length, including sequence-specific primer.
PCR Mix: PCR amplification mixes
PCR Protocol: PCR amplification protocol
Specific PCR amplification mixes and protocols are referred to by a number corresponding to those listed in Supplementary Tables S4 and S5, respectively.

Supplementary Table S4. PCR amplification mixes for NGS.

All PCR reactions were prepared from a starting material of 30-60 ng of genomic DNA, according to the manufacturers' recommendations. Amplification mixes #1, #2 and #4 were used with the FastStart High Fidelity PCR System, dNTPack (Roche Applied Science) and amplification mix #3 was used with the GC-RICH PCR System, dNTPack (Roche Applied Science).

Amplification Mix #1	
	For 1 well (µl)
Molecular Biology Grade Water	19.75
FastStart 10x Reaction Buffer with $MgCl_2$ (vial #2)	2.5
PCR Grade Nucleotide Mix (vial #6)	0.5
FastStart High Fidelity Enzyme Blend (vial #1)	0.25
Patient DNA (20 ng/µl)	2.0
Total volume	**25**

Amplification Mix #3	
	For 1 well (µl)
Molecular Biology Grade Water	14.5
PCR Grade Nucleotide Mix (vial #6)	0.5
GC-RICH Resolution Solution (vial #3)	2.5
GC-RICH PCR Reaction Buffer (vial #2)	5.0
GC-RICH Enzyme Mix (vial #1)	1.0
Patient DNA (20 ng/µl)	1.5
Total volume	**25**

Amplification Mix #2	
	For 1 well (µl)
Molecular Biology Grade Water	18.75
FastStart 10x Reaction Buffer with $MgCl_2$ (vial #2)	2.5
PCR Nucleotide Mix (vial #6)	0.5
FastStart High Fidelity Enzyme Blend (vial #1)	0.25
Patient DNA (20 ng/µl)	3.0
Total volume	**25**

Amplification Mix #4	
	For 1 well (µl)
Molecular Biology Grade Water	15.75
FastStart High Fidelity ReactionBuffer without $MgCl_2$ (vial #3)	2.5
$MgCl_2$ Stock Solution (vial #4)	2.5
DMSO (vial #5)	1.25
PCR Nucleotide Mix (vial #6)	0.5
FastStart High Fidelity Enzyme Blend (vial #1)	0.5
Patient DNA (20 ng/µl)	2.0
Total volume	**25**

Supplementary Table S5. PCR protocols for NGS.

All PCR reactions were performed in the 96-Well GeneAmp® PCR System 9700 (Applied Biosystems, Foster City, CA).

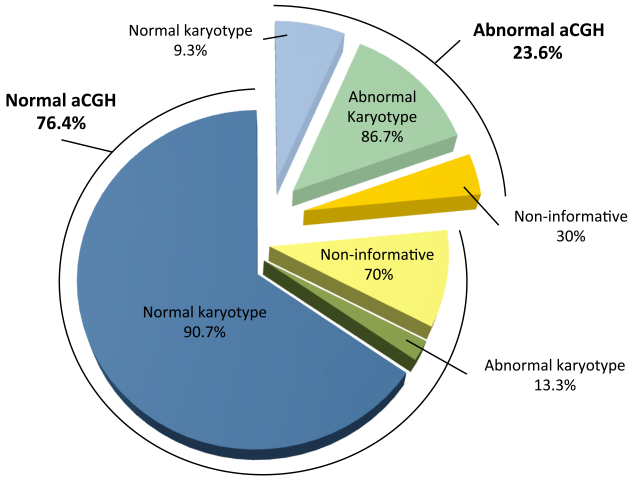
PCR Protocol #1		
Cycles	**Temperature (°C)**	**Time**
1x	95	10 min
10x	95	30 s
	*63 → touchdown, 58	30 s
	72	30 s
25x	95	30 s
	58	30 s
	72	30 s
1x	72	10 min
1x	12	on hold

*Touchdown: decreasing by 0.5°C per cycle to 58°C

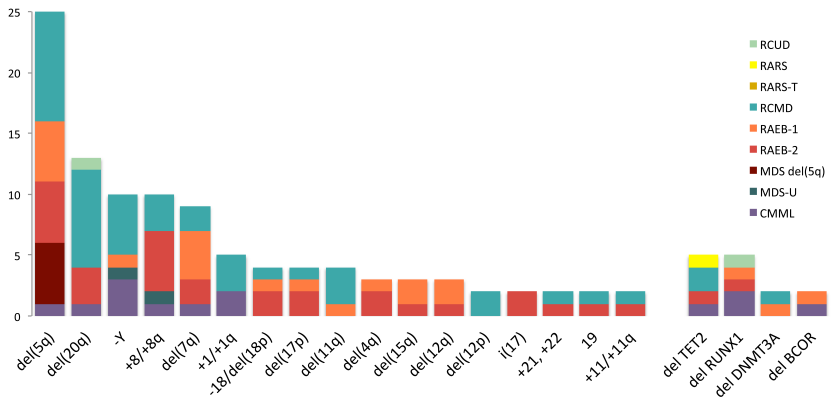
PCR Protocol #2		
Cycles	**Temperature (°C)**	**Time (min)**
1x	95	3
44x	95	1
	60	1
	72	1
1x	72	10
1x	12	on hold

Supplementary Table S6. Details of chromosome 13 rearrangements in the three high-risk MDS patients affected by chromothripsis. (Available in the CD)

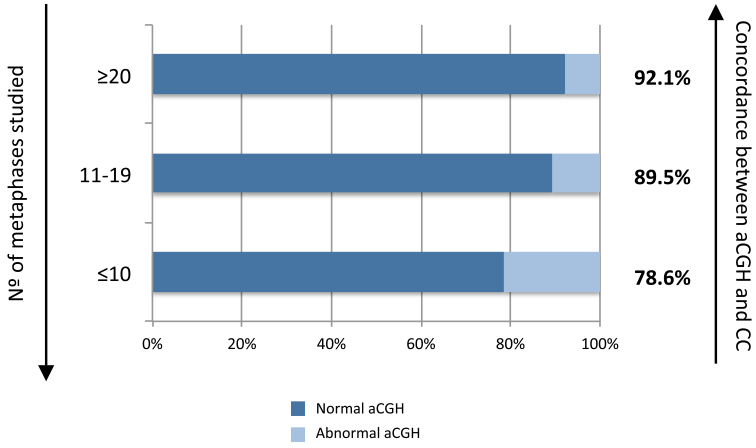
SUPPLEMENTARY FIGURES



Supplementary Figure S1. **Proportion of the whole series of patients with normal and abnormal aCGH profiles**. Each aCGH category is then divided by the cytogenetic subgroups detected by CC studies: normal, abnormal and non-informative karyotype. Percentages represent the proportion of patients from the total number of patients within each cytogenetic subgroup.



Supplementary Figure S2. **Frequency of large recurrent genomic abnormalities and frequency of cryptic recurrent CNAs involving genes of known significance in MDS and MDS/MPN patients only seen by aCGH**. All abnormalities are classified by MDS and MDS/MPN subtypes and color-coded as indicated on the right panel of the figure.



Supplementary Figure S3. **Relationship between aCGH and CC studies in the normal karyotype group**. Normal karyotype patients are divided into three categories on the basis of the number of good-quality metaphases evaluated: ≤10, 11-19 and ≥20. Patients with normal and abnormal aCGH results within each category are represented by different shades of blue.

REFERENCES

Kohlmann A, Klein HU, Weissmann S, Bresolin S, Chaplin T, Cuppens H, Haschke-Becher E, Garicochea B, Grossmann V, Hanczaruk B, Hebestreit K, Gabriel C, Iacobucci I, Jansen JH, te Kronnie G, van de Locht L, Martinelli G, McGowan K, Schweiger MR, Timmermann B, Vandenberghe P, Young BD, Dugas M, Haferlach T. 2011. The Interlaboratory RObustness of Next-generation sequencing (IRON) study: a deep sequencing investigation of TET2, CBL and KRAS mutations by an international consortium involving 10 laboratories. Leukemia 25:1840-1848.

Robledo C, García JL, Benito R, Flores T, Mollejo M, Martínez-Climent JÁ, García E, Gutiérrez NC, Piris MA, Hernández JM. 2011. Molecular characterization of the region 7q221 in splenic marginal zone lymphomas. PLoS One 6:e24939.

Vardiman JW, Thiele J, Arber DA, Brunning RD, Borowitz MJ, Porwit A, Harris NL, Le Beau MM, Hellström-Lindberg E, Tefferi A, Bloomfield CD. 2009. The 2008 revision of the World Health Organization (WHO) classification of myeloid neoplasms and acute leukemia: rationale and important changes. Blood 114:937-951

Common and progressive gene expression changes in the progression of myelodysplastic syndromes to acute myeloid leukemia

María Abáigar,[*1] Sara Aibar,[*2] Rocío Benito,[1] María Díez-Campelo,[3] Fernando Ramos,[4,5] Eva Lumbreras,[1] Francisco J Campos-Laborie,[2] Marta Megido,[6] Isabel Recio,[7] Lourdes Hermosín,[8] Javier Sánchez-del-Real,[5] Carmen Olivier,[9] Rebeca Cuello,[10] Lurdes Zamora,[11] Ken Mills,[12] M Consuelo del Cañizo,[3,13] Javier De Las Rivas,[2] and Jesús M Hernández-Rivas[1,3,13]

SUPPLEMENTARY TABLES

Supplementary Table S1. Summary of the clinical characteristics of the patients included in the study.

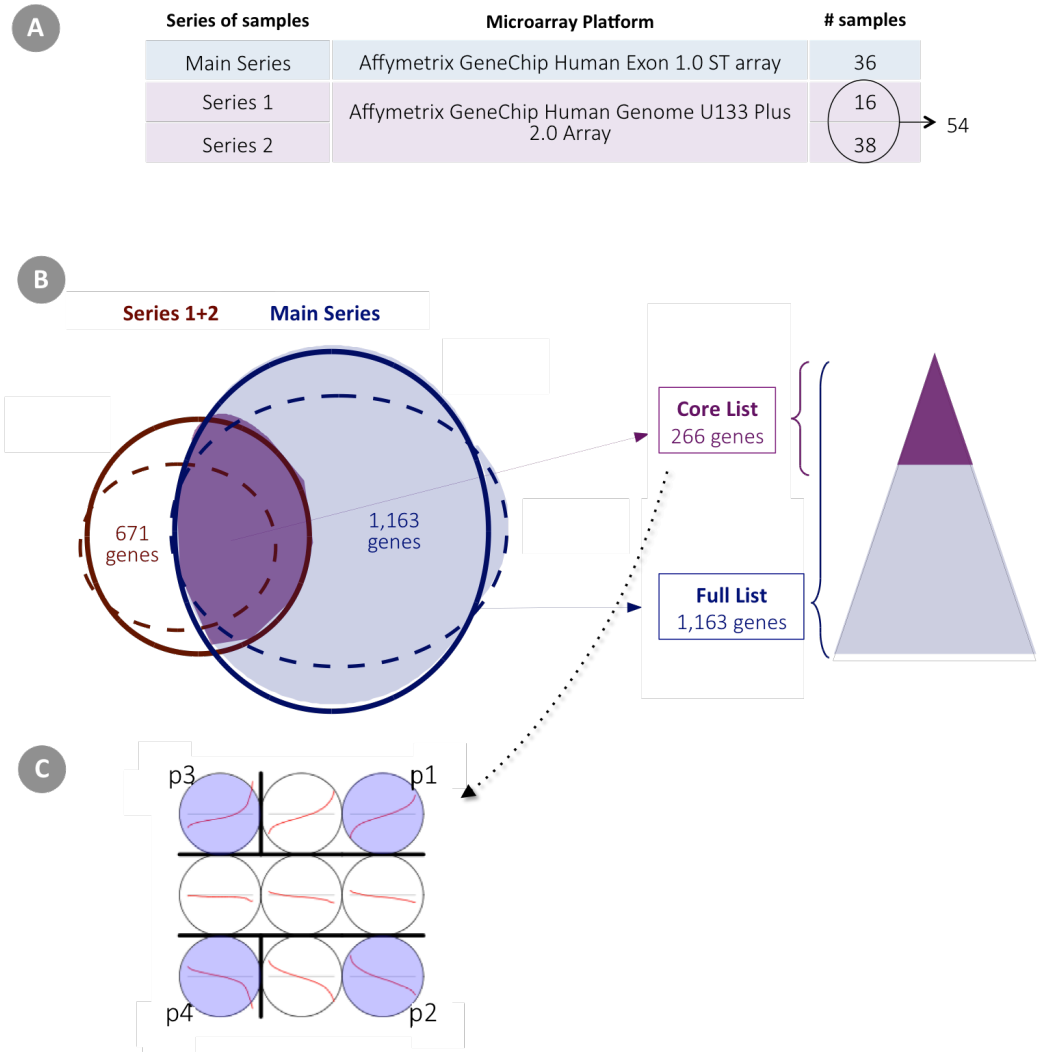
Variables	Median [Range]	% of samples
Age (years)	76 [48 - 88]	
Gender (Male/Female)		57.8 / 42.2
Peripheral blood values		
Hemoglobin level (g/dl)	7 [10.2 - 12.8]	
Neutrophil count ($x10^9$/L)	2 [0.3 - 7.5]	
Platelet count ($x10^9$/L)	176 [27 - 499]	
Bone marrow blasts (%) *	0 [6.3 - 88.8]	
< 10%		46.6
10 -20%		26
≥ 20%		27.4
Disease		
Non-malignant disorders		18.9
MDS (WHO 2008 classification)		
RCUD		12.2
RARS		0
RCMD		25.6
RAEB-1		11.1
RAEB-2		22.2
MDS-U		0
MDS del(5q)		0
AML		22.2
Conventional Cytogenetics		
Normal		100
Abnormal		0

* Calculated for MDS and AML samples

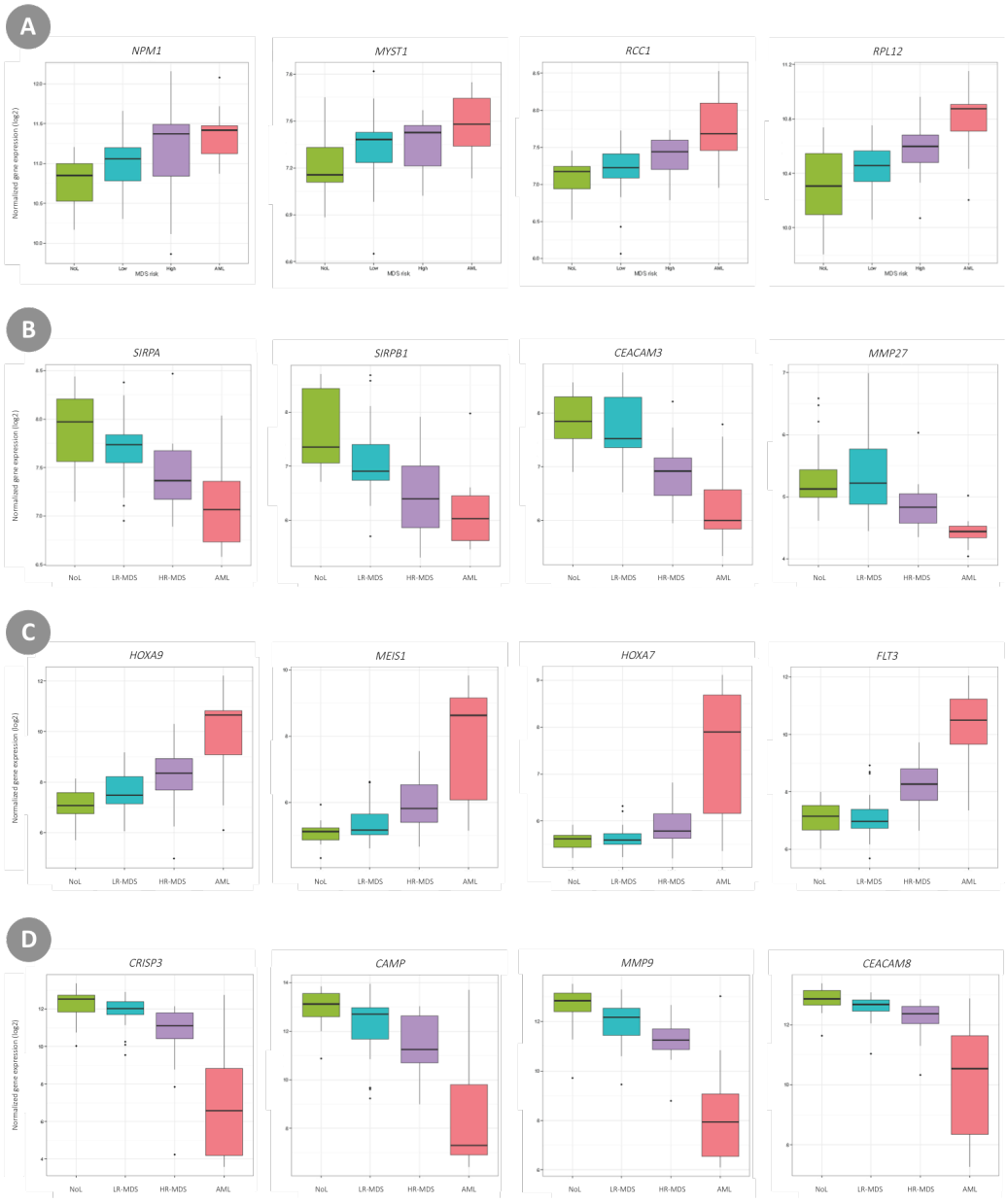
Supplementary Table S2. Information of the 1163 deregulated genes from the Full List. The first 266 rows correspond to the genes included in the Core List. Red indicates up-regulation and green down-regulation. The following columns represents the mean difference in expression level between samples and controls (mean.Diff), fold change (FC.R), gamma and adjusted p values from correlation analysis, the pattern to which each gene was assigned by the *SOM* clustering, gene type and gene description (Available in the CD)

Supplementary Table S3. The 42 deregulated transcription factors (TFs) from the Full List. FC.R, Fold change in the Full List / Main Series; Pattern, the pattern to which each TF was assigned by the *SOM* clustering (the asterisk indicates that it was also assigned to the same pattern in the Core List / Series 1+2). The top half of the table shows the TFs that are up-regulated (Red), and the bottom half, the ones down-regulated (Green). (Available in the CD)

SUPPLEMENTARY FIGURES



Supplementary Figure S1. **Conceptual design**. (S1A) Scheme illustrating the series of the study. (S1B) The genes that correlate with the progression of the disease, based on the risk level in the Mains Series and Series 1+2, and the composition of each of the gene lists (Full List and Core List). (S1C) The nine expression patterns provided by the *SOM* clustering analysis (S1C). The four patterns selected (p1, p2, p3 and p4) are highlighted in blue.



Supplementary Figure S2. **Example of expression patterns associated with the progression of the disease**. The boxplots represent the expression signal distributions (y axis, normalized intensity signal in log2) in the samples grouped by disease stage (x axis, blue = non-malignant; green = LR-MDS; orange = HR-MDS; and red = AML) for four selected genes per pattern. Pattern 1 (A), pattern 2 (B), pattern 3 (C), and pattern 4 (D).

Prognostic impact of the number of methylated genes in myelodysplastic syndromes and acute myeloid leukemias treated with azacytidine

María Abáigar · Fernando Ramos · Rocío Benito · María Díez-Campelo · Javier Sánchez-del-Real · Lourdes Hermosín · Juan Nicolás Rodríguez · Carlos Aguilar · Isabel Recio · Jose María Alonso · Natalia de las Heras · Marta Megido · Marta Fuertes · María Consuelo del Cañizo · Jesús María Hernández-Rivas

Supplementary Table S1 Methylation in control samples assessed by MS-MLPA

Gene	Control-1	Control-2	Control-3	Control-4	Control-5
TP73	0.02	0.02	0.01	0.01	0.03
CASP8	0.02	0.02	0.02	0.02	0.03
VHL	0.02	0	0.01	0.01	0
RARB	0.09	0.03	0.03	0.04	0.04
MLH1	0.02	0.02	0.01	0.02	0.03
MLH1	0.04	0	0	0	0
RASSF1	0.04	0	0.01	0	0
RASSF1	0.02	0	0	0.01	0.01
FHIT	0.02	0.01	0.01	0.01	0.01
APC	0.04	0.04	0.01	0.06	0.03
ESR1	0.03	0.04	0.05	0.06	0.01
CDKN2A	0.05	0.01	0.01	0.01	0.06
CDKN2B	**0.18**	**0.21**	**0.21**	**0.18**	**0.2**
DAPK1	0.03	0.02	0.02	0.02	0.01
PTEN	0.02	0.02	0.01	0.02	0.01
CD44	0.01	0.01	0	0.02	0.02
GSTP1	0.03	0.03	0.03	0.03	0.04
ATM	0.01	0.01	0	0.01	0.02
IGSF4	0	0	0.02	0	0.01
CDKN1B	0.14	0.03	0.04	0.01	0.02
CHFR	0.01	0	0	0	0
BRCA2	0.03	0	0	0.01	0.01
CDH13	0.05	0.03	0.02	0.03	0.08
HIC1	0.01	0.01	0	0.01	0.01
BRCA1	0.01	0.01	0.01	0.02	0.01
TIMP3	0.04	0.01	0.01	0.01	0.01

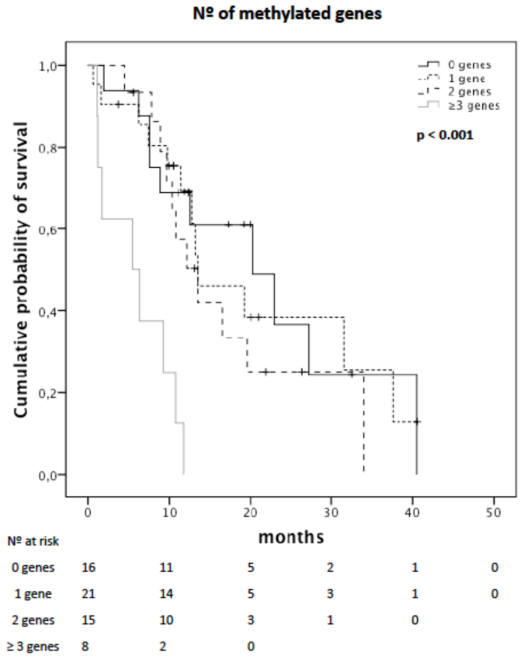
* Based on the methylation levels of *CDKN2B* gene found in our control samples (0.18 – 0.21) we considered a methylation cutoff of >0.30 for *CDKN2B* gene.

Supplementary Table S2 Summary of cytogenetic results

Code No.	MDS / AML	Karyotype
51	AML	47,XY,del(5)(q13q31),+8,ins(13)(q13q22),add(19)(p13)[cp20]
85	AML	44,XY,add(1)(p36),del(5)(q13q31),-15,-17[3]/46,XY[7]
87	AML	45,XY,-7[5]/46,XY[15]
115	AML	47,XY,+14[14]/46,XY[6]
91	AML	45,XX,-7[5]/46,XX[10]
37	AML	48,XX,add(3)(p21),del(5)(q13q31),+11,-12,add(17)(p13),+21[13]/46,XX[2]
34	AML	47,XX,+11[7]/46,XX[13]
80	AML	47,XY,+9[15]/46,XY[5]
131	MDS	47,XY,+del(1)(p13),del(7)(q21)[11]/46,XY[9]
20	MDS	46,XX,del(5)(q13q31)[22]
98	MDS	46,XY,del(7)(q21),+mar[15]/46,XY[5]
26	MDS	7,XY,+8[4]/46,XY[16]
113	MDS	47,XX,+8[4]/46,XY[14]
67	MDS	46,XX, add(14)(q32)[10]/46,XX[6]
12	MDS	46,XX,del(5)(q13q31)[20]/46,XX[5]
112	MDS	47,XY,del(5)(q21q31),+21[20]
15	MDS	47,XX,+8[4]/46,XX[12]
69	MDS	46,XX,del(5)(q21q31),del(7)(q21)[7]/46,XX[13]
94	MDS	45,XX,add(5)(q31),del(6)(q21),-7,-14,+r[17]/46,XX[3]
68	MDS	46,XX,del(5)(q13q31),del(7)(q21)[8]/46,XX[12]
9	MDS	53,XX,+3,+14,+15,+20,marx3[18]/46,XX[2]
76	MDS	46,XY,del(5)(q21q31),del(7)(q21)[5]/46,XY[15]

* In 6 additional cases abnormal FISH was found

Supplementary Figure S3



Supplementary Figure S3 Kaplan-Meier curves according to the number of methylated genes with the cutoffs of 0, 1, 2 and ≥3 methylated genes. This analysis showed that with an increasing number of methylated genes, the survival decreases. Thus patients with at least 3 methylated genes had a shorter OS than patients with 2 methylated genes. Moreover, patients with 2 methylated genes had a shorter OS than those with 1 methylated gene, and this last group in relation to those patients with no methylation (median survival of 5.5, 13.5, 13.5 and 20.3 months respectively, $p<0.001$).