

Improving molecular diagnosis in myeloid neoplasms. The role of cell-free DNA

Nieves García Gisbert

TESI DOCTORAL UPF / 2021

Thesis supervisors:

Dr. Beatriz Bellosillo Paricio

Dr. Carles Besses Raebel

Dr. Xavier Calvo González

DEPARTMENT OF EXPERIMENTAL AND HEALTH SCIENCES



**Universitat
Pompeu Fabra**
Barcelona

Acknowledgements/Agradecimientos

En primer lugar, quería agradecer a mis directores de tesis, por ayudarme siempre con su experiencia y dedicar su tiempo y su esfuerzo a que esta tesis saliera adelante. A Bea, por darme esta oportunidad, por ayudarme a crecer y por escucharme y apoyarme en los momentos complicados. Al Dr. Besses por su rigor, experiencia y compromiso con los diferentes proyectos de esta tesis. A Xavi por transmitirme siempre su ilusión por todos los proyectos que hemos hecho, por apoyarme y por enseñarme tanto. Ha sido un placer teneros como directores.

A mis compañeras, Lierni, Laura, Conchi, Marta, gracias por todos los momentos que hemos compartido estos años y por ser siempre un apoyo. Me habéis enseñado muchísimo. A Mireia, por la alegría que siempre nos transmites. Gracias por tantas y tantas comidas juntas.

También a todos los compañeros/as del Grup de Recerca con los que he coincidido durante estos años. A Joan, gracias por ayudarme tanto y por tu paciencia con los cuadraditos de cada oncoplot. A Antonio, por su apoyo en estos años.

A todos los compañeros/as del servicio de Biología Molecular por ayudarme tanto y tratarme siempre con tanto cariño (Raquel, Gabri, Mari, Mercè, Lidia, Dani, Rebeca, Erica, Ester, Gemma, Sergi... gracias!!). También a las compañeras de Citogenética (Marta, María, me ha encantado trabajar con vosotras; Blanca, gracias por el apoyo en esta etapa), y Citología Hematológica (Leo, Ana, Bea, Rosa, gracias por prestaros a ayudarme siempre).

A todos los compañeros/as del servicio de Hematología del Hospital de Mar (especialmente a Sara, Brayan, Marcio, Eugenia y Anna), sin vosotros este trabajo no habría sido posible.

Al Servei de Genòmica de la UPF por ponernos un NextSeq tras otro y facilitarnos siempre todo el trabajo.

Y por supuesto, a todos los pacientes que han colaborado en los estudios incluidos en esta tesis.

A mis padres, por educarnos siempre para ser curiosos y querer aprender y mejorar como personas. Mamá, gracias por estar ahí siempre y darme los mejores consejos de mundo. Papá, gracias por apoyarme tanto y quererme como lo haces.

A mi hermano, Víctor, porque no habría llegado hasta aquí si no hubiera crecido a tu lado. Te admiro muchísimo y es un orgullo tenerte como hermano.

A toda mi familia, en especial a mi abuela y a mi yaya por todo su cariño. También a mi abuelo, mi yayo y mi tío Javier que ya no están. Cuanto más tiempo paso lejos más cuenta me doy de la familia tan maravillosa que tengo. A todos mis primos que se van a comer el mundo, a Diego, por estar ahí siempre.

A mi otra pequeña familia aquí en Barcelona, Lucia, Mar, Samu y Rafa, tengo mucha suerte de teneros y esta tesis también es vuestra. Lucía, eres mi luz, gracias por apoyarme siempre. Mar, me hace muy feliz tenerte en mi vida. Gracias a mis compañeras de basket y a toda la gente que ha formado parte de esta etapa.

A Mercedes, Bea C, Bea S y Alberto, por quererme desde la distancia, entenderme y apoyarme en todo lo que me propongo. A Mude por todos estos años de amistad, te adoro.

Y a Maru, por ser el mejor compañero que haya podido elegir. Gracias por transmitirme siempre tu optimismo, tu alegría y tus ganas de vivir al máximo cada momento.

Gracias a todos por enseñarme tanto.

Content

ABSTRACT	I
RESUMEN	III
PREFACE	V
LIST OF FIGURES	VII
LIST OF TABLES	IX
ABBREVIATIONS	XI
INTRODUCTION	XV
1. MYELOID NEOPLASMS	1
<i>1.1. Introduction and classification of myeloid neoplasms</i>	<i>1</i>
<i>1.2 Genetic alterations in myeloid neoplasms</i>	<i>3</i>
1.2.1 Cell signaling.....	3
1.2.1.1 JAK-STAT pathway: <i>JAK2, CALR, MPL, SH2B3</i>	4
1.2.1.2 RAS pathway: <i>KRAS, NRAS, CBL, NF1, PTPN11</i>	8
1.2.2 Epigenetic regulation	10
1.2.2.1 DNA methylation: <i>TET2, DNMT3A, IDH1/IDH2</i>	11
1.2.2.2 Histone modification: <i>ASXL1, EZH2</i>	13
1.2.3 Splicing: <i>SF3B1, SRSF2, U2AF1, ZRSR2, PRPF8, DDX41</i>	14
1.2.4 Transcription factors: <i>RUNX1, SETBP1, ETV6, BCOR/BCORL1, PHF6, CUX1, GATA2</i>	18
1.2.5 DNA repair: <i>TP53, PPM1D, CHEK2, ATM</i>	21
1.2.6 Cohesins: <i>STAG2, RAD21</i>	22
1.2.7 Genes, mutational frequency and clinical implications	23
<i>1.3 Role of molecular diagnosis in myeloid neoplasms. Main limitations</i>	<i>26</i>
<i>1.4 Clonal hematopoiesis of indeterminate potential</i>	<i>27</i>
2. MYELOPROLIFERATIVE NEOPLASMS (MPN)	31
<i>2.1. Introduction</i>	<i>31</i>
<i>2.2. Classification</i>	<i>31</i>

2.2.1 PV	31
2.2.2 ET.....	33
2.2.3 PMF	35
2.3. <i>Molecular characterization in MPNs</i>	38
2.2.1. Driver Mutations	38
2.2.4. Non-driver mutations.....	42
2.2.5. Risk stratification based on molecular alterations.....	42
3. MYELODYSPLASTIC SYNDROMES (MDS).....	45
3.1 <i>Introduction</i>	45
3.2 <i>Classification</i>	47
3.3 <i>Cytogenetics in MDS</i>	49
3.4 <i>Molecular characterization in MDS</i>	51
4. CHRONIC MYELOMONOCYTIC LEUKEMIA.....	55
4.1 <i>Introduction</i>	55
4.2 <i>Classification</i>	58
4.2.1 Oligomonocytic chronic myelomonocytic leukemia (OM-CMML)	58
4.3 <i>Molecular characterization in CMML</i>	59
4.4 <i>Flow cytometry in CMML</i>	60
5. LIQUID BIOPSY AND CELL-FREE DNA.....	63
5.1 <i>Cell-free DNA biology and origin</i>	65
5.2 <i>Applications of cell-free DNA analysis</i>	67
5.2.1 cfDNA in myeloid neoplasms	67
5.3 <i>Limitations of cell-free DNA analysis</i>	69
HYPOTHESIS AND OBJECTIVES	71
RESULTS	75
ARTICLE 1: CIRCULATING CELL-FREE DNA IMPROVES THE MOLECULAR CHARACTERISATION OF PH-NEGATIVE MYELOPROLIFERATIVE NEOPLASMS	77
ARTICLE 2: MOLECULAR AND CYTOGENETIC CHARACTERIZATION OF MYELODYSPLASTIC SYNDROMES IN CELL-FREE DNA.....	89

ARTICLE 3: OLIGOMONOCYTIC AND OVERT CHRONIC MYELOMONOCYTIC LEUKEMIA SHOW SIMILAR CLINICAL, GENOMIC, AND IMMUNOPHENOTYPIC FEATURES	113
ARTICLE 4: ANALYSIS OF SALIVA SAMPLES AND CLUSTER OF DIFFERENTIATION 3 (CD3)+ LYMPHOCYTES AS A SOURCE OF GERMLINE DNA IN MYELOPROLIFERATIVE NEOPLASMS.....	127
DISCUSSION	131
CONCLUSIONS	143
BIBLIOGRAPHY.....	147
FUNDING	179
ANNEX I: PRESENTATIONS AT INTERNATIONAL CONFERENCES	181
ANNEX II: SUPPLEMENTARY INFORMATION.....	183

ABSTRACT

Myeloid malignancies are clonal diseases originated in myeloid hematopoietic stem cells that are frequently initiated by somatic mutations. The detection of genetic alterations has considerably improved the diagnostic accuracy in myeloid neoplasms, however multiple aspects should be further improved in the diagnostic and classification tools that are used in clinical practice. The main goal of the research projects presented in this thesis is to improve the accuracy of the diagnosis and classification of myeloid malignancies, focused on myeloproliferative neoplasms (MPN), myelodysplastic syndromes (MDS) and chronic myelomonocytic leukemia (CMML). We explored the role of cell free DNA (cfDNA) analysis as a new non-invasive diagnostic tool in MPN and MDS patients and detected an equivalent mutational profile in paired samples of cfDNA and tumoral cells. On the other hand, we compared the clinical, genomic, and immunophenotypic features of a series of oligomonocytic CMML (OM-CMML) and overt CMML and observed similar characteristics supporting the consideration of OM-CMML as a distinctive subtype of CMML. Finally, we assessed if saliva samples and CD3+ lymphocytes were a suitable source of germline DNA in MPN patients, and found that the use CD3+ lymphocytes was a better option for germline DNA obtention than saliva samples, which were frequently contaminated with tumoral cells.

RESUMEN

Las neoplasias mieloides son enfermedades clonales que se originan en las células madre hematopoyéticas mieloides y son iniciadas generalmente por mutaciones somáticas. La detección de estas alteraciones genéticas ha mejorado considerablemente la precisión diagnóstica en las neoplasias mieloides, sin embargo, aún se deben mejorar múltiples aspectos en las herramientas de diagnóstico y clasificación que se utilizan en la práctica clínica. El objetivo principal de los estudios presentados en esta tesis es mejorar la precisión del diagnóstico y clasificación de las neoplasias mieloides, en concreto en neoplasias mieloproliferativas (NMP), síndromes mielodisplásicos (SMD) y leucemia mielomonocítica crónica (LMMC). Hemos explorado el papel del análisis de ADN libre circulante (cfDNA) como una nueva herramienta diagnóstica no invasiva en pacientes con NMP y SMD, y hemos detectado un perfil mutacional equivalente en muestras pareadas de cfDNA y células tumorales. Por otro lado, comparamos las características clínicas, genómicas e inmunofenotípicas de una serie de LMMC oligomonocítica (OM-CMML) y LMMC clásica y observamos características similares entre ambos grupos, lo que apoya la consideración de OM-CMML como un subtipo de LMMC. Finalmente, evaluamos si las muestras de saliva y los linfocitos CD3 + eran una fuente adecuada de ADN de línea germinal en pacientes con NMP, y observamos que el uso de linfocitos CD3+ era una mejor opción para la obtención de ADN germinal que las muestras de saliva, que estaban contaminadas en su mayoría con células tumorales.

PREFACE

In the present thesis, the research has been carried out in the Group of Applied Clinical Research in Hematology, Hospital del Mar Medical Research Institute (IMIM). In addition, these studies have been conducted with the collaboration of the Pathology Department and the Hematology Department of Hospital del Mar, and with the Group of Translational Research on Hematological Neoplasms, IMIM.

The studies included in the present thesis have been published in international journals:

Garcia-Gisbert N, Fernández-Ibarrondo L, Fernández-Rodríguez C, Gibert J, Andrade-Campos M, Arenillas L, Camacho, L, Angona A, Longarón R, Salar A, Calvo X, Besses C*, Bellosillo, B*. **Circulating cell-free DNA improves the molecular characterisation of Ph-negative myeloproliferative neoplasms.** Br J Haematol. 2021;192(2):300–9. *equally contributed

Garcia-Gisbert N, Garcia-Ávila S, Merchán B, Salido M, Fernández-Rodríguez C, Gibert J, Fernández-Ibarrondo L, Camacho L, Lafuente M, Longarón R, Espinet B, Vélez P, Pujol RM, Andrade-Campos M, Arenillas L, Salar A, Calvo X, Besses C, Bellosillo B. **Molecular and cytogenetic characterization of myelodysplastic syndromes in cell-free DNA.** Manuscript submitted to Blood Advances

Calvo X*, Garcia-Gisbert N*, Parraga I, Gibert J, Florensa L, Andrade-Campos M, Merchan B, Garcia-Avila S, Montesdeoca S, Fernández-Rodríguez C, Salido M, Puiggros A, Espinet B, Colomo L, Roman-Bravo D, Bellosillo B, Ferrer A, Arenillas L. **Oligomonocytic and overt chronic myelomonocytic leukemia show similar clinical, genomic, and immunophenotypic features.** Blood Adv. 2020 Oct 27;4(20):5285–96. *equally contributed

Garcia-Gisbert N*, Camacho L*, Fernández-Ibarrondo L, Fernández-Rodríguez C, Longarón R, Gibert J, Angona A, Andrade-Campos M, Salar A, Besses C, Bellosillo B. **Analysis of saliva samples and cluster of differentiation 3 (CD3)+ lymphocytes as a source of germline DNA in myeloproliferative neoplasms.** Br J Haematol. 2020 Jun 1;189(5):e204–7. *equally contributed

LIST OF FIGURES

Figure 1. JAK-STAT Pathway.....	4
Figure 2. <i>JAK2</i> , <i>CALR</i> and <i>MPL</i> mutations in the JAK-STAT Pathway.	7
Figure 3. RAS pathway signaling.	8
Figure 4. Methylation and demethylation process.....	11
Figure 5. Histone modification by the PRC2 complex.....	13
Figure 6. Splicing machinery and most frequent mutations in <i>U2AF1</i> , <i>SRSF2</i> and <i>SF3B1</i>	15
Figure 7. CHIP frequency with age and most frequently mutated genes.....	29
Figure 8. Phylogenetic reconstruction of genetic alterations in MPNs.	40
Figure 9. Molecular classification of MPNs.....	43
Figure 10. Origin and precursor states of MDS.....	47
Figure 11. Genetics of MDS. Most frequently mutated genes and cytogenetic alterations in MDS.	52
Figure 12. Prognostic implications of gene mutations in MDS.....	53
Figure 13. Molecular alterations acquisition in CMML.....	56
Figure 14. Distribution of PB monocyte populations by flow cytometry in healthy adults, CMML and reactive monocytosis.....	61
Figure 15. Liquid biopsy applications in clinical context.....	63
Figure 16. Cell-free DNA obtention and analysis.....	65
Figure 17. Nucleosome footprint in cfDNA.....	66
Figure 18. Mutational dynamics in cfDNA of a patient receiving azacitidine.	68
Figure 19. Cumulative incidence of relapse based on the presence of mutated circulating tumor DNA (ctDNA) after allogenic hematopoietic stem cell transplantation in MDS/AML.	68

LIST OF TABLES

Table 1. WHO 2017 myeloid neoplasm classification.....	2
Table 2. Most frequently mutated genes in MPN, MDS and CMML	24
Table 3. Acronyms describing clonal hematopoiesis and related conditions.	28
Table 4. WHO PV diagnosis criteria	33
Table 5. WHO ET diagnosis criteria.....	34
Table 6. WHO prePMF diagnosis criteria.....	36
Table 7. WHO overt PMF diagnosis criteria.....	37
Table 8. WHO classification for MDS	48
Table 9. Cytogenetic alterations associated to MDS	49
Table 10. IPSS-R prognostic subgroups based on cytogenetics.....	50
Table 11. IPSS-R scores.. ..	50
Table 12. WHO CMML diagnostic criteria.....	57

ABBREVIATIONS

5-hmC	5-hydroxymethylcytosines
5-mC	5-methylcytosines
AML	Acute myeloid leukemia
ARCH	Aging Related Clonal Hematopoiesis
AUC	Area under the curve
BM	Bone marrow
bp	Base pairs
CALR	Calreticulin
CCUS	Clonal cytopenia of undetermined significance
CD2	Cluster of differentiation 2
CD3	Cluster of differentiation 3
CD56	Cluster of differentiation 56
cfDNA	Cell-free DNA
cfmiRNA	Cell-free micro RNA
cfRNA	Cell-free RNA
CHIP	Clonal hematopoiesis of indeterminate potential
CHIP-seq	Chromatin immunoprecipitation sequencing
circRNA	Circular RNA
CMA	Chromosomal microarrays
CMML	Chronic myelomonocytic leukemia
CPPS	CMML-specific prognostic scoring system
CPPS-Mol	CMML-specific prognostic scoring system with molecular data
CTC	Circulating tumor cell
ctDNA	Circulating tumor DNA
d-CMML	Dysplastic chronic myelomonocytic leukemia
del	Deletion
DNA	Deoxyribonucleic Acid
dPCR	Digital PCR
EDTA	Ethylene Diamine Tetra Acetic acid
EGFR	Epidermal growth factor receptor
EPO	Erythropoietin
ER	Endoplasmic reticulum
ET	Essential thrombocythemia
FC	Flow cytometry
FISH	Fluorescence in situ hybridization
G-CSF	Granulocyte colony-stimulating factor
GDP	Guanosine diphosphate

GM-CSF	Granulocyte-macrophage colony-stimulating factor
GTP	Guanosine triphosphate
HGVS	Human Genome Variation Society
HMA	Hypomethylating agents
HSC	Hematopoietic stem cell
HU	Hydroxyurea/hydroxycarbamide
HUMARA	Human Androgen Receptor Gene Assay
ICUS	Idiopathic Cytopenia of Undetermined Significance
idic()	Isodicentric
IFN	Interferon
IGV	Integrative Genomics Viewer
ins	Insertion
inv()	Inversion
IPSS	International Prognostic Scoring System
IPSS-R	Revised International Prognostic Scoring System
JAK	Janus kinase
JH1	Janus homology-1
JH2	Janus homology-2
LDH	Lactate dehydrogenase
LoH	Loss of heterozygosity
MAF	Mutant allele frequency
MAPK	Mitogen-activated protein kinase
MDS	Myelodysplastic syndrome
MDS-EB	Myelodysplastic syndrome with excess blasts
MDS-MLD	Myelodysplastic syndrome with multilineage dysplasia
MDS-RS	Myelodysplastic syndrome with ring sideroblasts
MDS-SLD	Myelodysplastic syndrome with single lineage dysplasia
MDS-U	Unclassifiable myelodysplastic syndrome
MF	Myelofibrosis
MO1	Classical monocytes
MO2	Intermediate monocytes
MO3	Non-classical monocytes
MPN	Myeloproliferative neoplasm
mRNA	Messenger ribonucleic acid
NGS	Next generation sequencing
OM-CMML	Oligomonocytic chronic myelomonocytic leukemia
OS	Overall survival
PB	Peripheral blood
PBMC	Peripheral blood mononuclear cell
p-CMML	Proliferative chronic myelomonocytic leukemia

PCR	Polymerase chain reaction
pDC	Plasmacytoid dendritic cells
PH	Pleckstrin homology
Ph-	Philadelphia negative
PI3K	Phosphatidylinositol 3-kinase
PMF	Primary myelofibrosis
PRC2	Polycomb repressive complex 2
PV	Polycythemia vera
qPCR	Quantitative polymerase chain reaction
RAS	Rat sarcoma virus
RNA	Ribonucleic acid
ROC	Receiver operating characteristic curve
rRNA	Ribosomal ribonucleic acid
RS	Ring sideroblasts
RTK	Receptor tyrosine kinases
SH2	Src Homology 2
snRNA	Small nuclear RNAs ribonucleic acid
STAT	Signal transducers and activators of transcription
t()	Translocation
t-CMML	Therapy related chronic myelomonocytic leukemia
TEP	Tumor educated platelets
TF	Transcription factor
TN	Triple negative
TPO	Thrombopoietin
uMPN	Unclassifiable myeloproliferative neoplasm
VAF	Variant allele frequency
vcf	Variant calling files
WBC	White blood cell
WHO	World Health Organization
WT	Wild Type

INTRODUCTION

1. MYELOID NEOPLASMS

1.1. Introduction and classification of myeloid neoplasms

Myeloid malignancies include a heterogeneous group of clonal diseases that are originated in the myeloid hematopoietic stem cells. Patients with myeloid malignancies present clonal stem cells with detectable genetic alterations that affect proliferation, differentiation and homeostasis. The presence of aberrant hematopoiesis is the main common feature of this disease group, however, the clinical phenotype is conditioned by the characteristics of the clonal hematological cells. When these aberrant clonal stem cells have the capacity to mature and expand, this proliferative disease phenotype is associated to myeloproliferative neoplasms (MPN). On the other hand, myelodysplastic syndromes (MDS) are characterized by hematological progenitors that present ineffective hematopoiesis, impeding correct cell differentiation. Those cases in which both proliferative and dysplastic features are observed are included in the myeloproliferative/myelodysplastic group (MPN/MDS).

The World Health Organization (WHO) 2017 revision of the myeloid neoplasms includes 9 categories (Table 1)(1). The research projects presented in this thesis are focused in the group of chronic myeloid neoplasms including MPN, MDS and MDS/MPN. Regarding the MDS/MPN group, this thesis is focused in chronic myelomonocytic leukemia (CMML).

Table 1. WHO 2017 myeloid neoplasm classification. Adapted from Arber et al., *Blood*, 2016.

Myeloproliferative neoplasms (MPN)
Chronic myeloid leukemia (CML), <i>BCR-ABL1</i> ⁺
Chronic neutrophilic leukemia (CNL)
Polycythemia vera (PV)
Primary myelofibrosis (PMF)
PMF, prefibrotic/early stage
PMF, overt fibrotic stage
Essential thrombocythemia (ET)
Chronic eosinophilic leukemia, not otherwise specified (NOS)
MPN, unclassifiable
Mastocytosis
Myeloid/lymphoid neoplasms with eosinophilia and rearrangement of <i>PDGFRA</i> , <i>PDGFRB</i> , or <i>FGFR1</i> , or with <i>PCM1-JAK2</i>
Myelodysplastic/myeloproliferative neoplasms (MDS/MPN)
Chronic myelomonocytic leukemia (CMML)
Atypical chronic myeloid leukemia (aCML), <i>BCR-ABL1</i> ⁻
Juvenile myelomonocytic leukemia (JMML)
MDS/MPN with ring sideroblasts and thrombocytosis (MDS/MPN-RS-T)
MDS/MPN, unclassifiable
Myelodysplastic syndromes (MDS)
MDS with single lineage dysplasia
MDS with ring sideroblasts (MDS-RS)
MDS-RS and single lineage dysplasia
MDS-RS and multilineage dysplasia
MDS with multilineage dysplasia
MDS with excess blasts
MDS with isolated del(5q)
MDS, unclassifiable
Myeloid neoplasms with germ line predisposition
Acute myeloid leukemia (AML) and related neoplasms
Blastic plasmacytoid dendritic cell neoplasm
Acute leukemias of ambiguous lineage
B-lymphoblastic leukemia/lymphoma
T-lymphoblastic leukemia/lymphoma

Adapted from Arber et al., *Blood*, 2016

1.2 Genetic alterations in myeloid neoplasms

The identification of somatic mutations has considerably improved the diagnostic accuracy of myeloid neoplasms. During the last decades, it has been possible to identify genetic alterations in the majority of patients with MPNs, MDS and CMML. The detection of molecular alterations has become part of the diagnostic criteria in several disease subgroups of myeloid neoplasms.

The most frequently mutated genes in chronic myeloid neoplasms can be classified according to the function of the protein in 6 groups:

- Cell signaling (*JAK2, CALR, MPL, SH2B3, NRAS, KRAS, CBL, NF1, PTPN11*)
- Epigenetic regulation (*TET2, DNMT3A, IDH1/IDH2, ASXL1, EZH2*)
- Splicing (*SF3B1, SRSF2, U2AF1, ZRSR2, PRPF8, DDX41*)
- Transcription factors (*RUNX1, SETBP1, ETV6, BCOR/BCORL1, PHF6, CUX1, GATA2*)
- DNA repair (*TP53, PPM1D, CHEK2, ATM*)
- Cohesins (*STAG2, RAD21*)

The function of each gene group is described in detail in the following sections (1.2.1 to 1.2.6). The frequency and clinical implications of these genes are listed in 1.2.7.

1.2.1 Cell signaling

The process of signal transduction in the cell is essential for generating a cellular response to an external stimulus that activates a cell surface receptor. When the ligand binds to the receptor, a series of biochemical reactions are performed from the cell surface to the nucleus to regulate gene expression. Mutations in proteins involved in signaling transduction generally induce the constitutive activation of the pathway in absence of ligand, leading to deregulated gene expression. In myeloid malignancies, the Janus Kinase-

Signal transducer and activator of transcription (JAK-STAT) and Rat sarcoma (RAS) are the most frequently altered pathways.

1.2.1.1 JAK-STAT pathway: *JAK2*, *CALR*, *MPL*, *SH2B3*

Three main proteins constitute the JAK-STAT signaling pathway: cell receptors, JAK kinases and STAT proteins (Figure 1).

-**Cell receptors** are located in the cell surface and are activated by type I cytokines such as erythropoietin (EPO), thrombopoietin (TPO), granulocyte-macrophage colony-stimulating factor (GM-CSF), and granulocyte colony-stimulating factor (G-CSF)(2). The binding of these ligands activates the receptors and induces their homodimerization.

-**JAK kinases** are activated following the receptor homodimerization, inducing the binding of the nearest JAK protein to the cytoplasmic domain of each receptor. Consequently, two JAK kinases are recruited into receptor complex and phosphorylate each other to activate their kinase activity(3,4).

-**STAT proteins** are phosphorylated by JAK kinases, triggering the formation of STAT homodimers, which are able to enter the cell nucleus. Once in the nucleus, STAT proteins are activated as transcription factors that bind DNA and regulate gene expression(5).

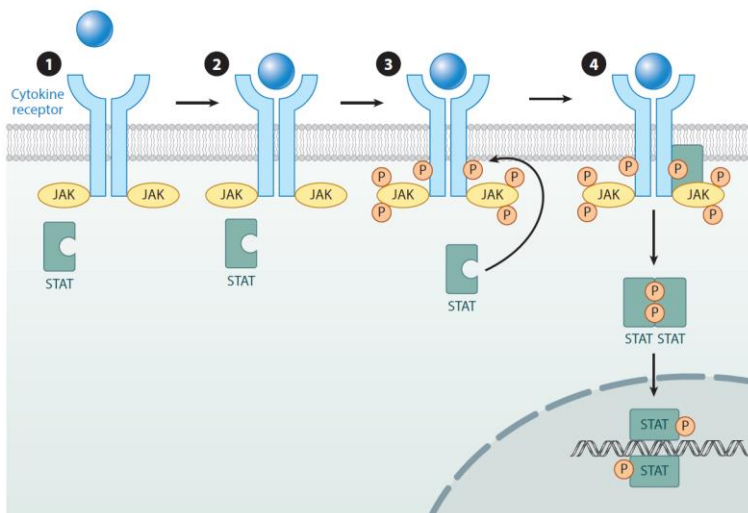


Figure 1. JAK-STAT Pathway. (O’Shea *et al.*, *Annu Rev Med*, 2015)(6)

Mutations in the JAK-STAT pathway induce the abnormal activation of STAT proteins and are causative of the MPN phenotype. These mutations are found in 90% of MPNS, however, they are not exclusive for MPN disease group. Approximately, 2-10% of CMML show mutations in this pathway, which is coherent with the proliferative features observed in this disease subgroup. In MDS, alterations in the JAK-STAT pathway are infrequent (<5%).

JAK2

JAK2 (Janus Kinase 2) gene encodes for the JAK2 protein, a tyrosine kinase that participates in the JAK-STAT pathway signaling. JAK2 activity is critical for signaling transduction from receptors to STAT proteins(3).

In 2005, the somatic *JAK2* p.V617F (*JAK2V617F*) mutation was identified for the first time in MPN patients(7–11). This mutation is a missense mutation consisting on a change from valine (V) to phenylalanine (F) in the aminoacid 617 of the *JAK2* gene. The affected gene region, located in exon 14, is a JH2 (Janus homology-2) pseudokinase domain that regulates the activity of the JH1 (Janus homology-1) tyrosine kinase domain. The *JAK2V617F* mutation in JH2 blocks the regulation of JH1, inducing a constitutive activation of the kinase domain. Therefore, *JAK2V617F* is a hyperactivating mutation as it induces a constitutive activation of JAK2 protein in absence of a ligand (Figure 2)(7,10). It is considered a driver alteration as the constant activation of the signaling pathway is causative for MPNs.

Posterior studies in 2007 found activating mutations in exon 12 of *JAK2* when analyzing MPN patients with erythrocytosis lacking the *JAK2V617F* mutation. These mutations affect the genomic region between the SH2 (Src Homology 2) and JH2 domain of *JAK2*, and induce a similar effect in protein interaction to *JAK2V617F*(12).

CALR

Calreticulin (*CALR*) gene encodes for the calreticulin protein, involved in cellular processes in the cytoplasm and in the endoplasmic reticulum (ER). In the cytoplasm, calreticulin is involved in the regulation of proliferation, apoptosis, phagocytosis and immune response(13). In the ER, it acts as a

calcium binding chaperone protein that regulates the correct folding of other proteins. Calreticulin protein has three key domains to perform its cellular activity: the N-terminal domain, the proline-rich domain and the C-terminal domain. The N terminal domain and the proline-rich domain are mainly implicated in the chaperone functions of CALR. In normal conditions, the C-terminal domain presents a KDEL (K-lysine, D-aspartic acid, E-glutamic acid, L-leucine) sequence enriched in negatively charged amino acids that is essential for protein retention in the ER(14).

In 2013, somatic mutations in *CALR* were identified for the first time by NGS in patients with MPNs without *JAK2V617F* mutation (15,16). The two most frequent mutations in *CALR* are frameshift mutations: type 1 mutation is a 52 base pairs (bp) deletion (p.L367fs*46) and type 2 mutation is a 5 bp insertion (p.K385fs*47). Other mutations are classified as type 1-like or 2-like based on the affected region of the protein. Mutated *CALR* produces the synthesis of an altered protein with a C-terminal domain that is positively charged, inhibiting its capacity to bind the ER. In this context, mutated calreticulin induces the aberrant activation the JAK-STAT pathway by interacting with the receptor of thrombopoietin (MPL), which leads to the MPN phenotype(17,18).

MPL

MPL (c-mpl) gene codifies for the TPO receptor protein, which is essential for blood cell proliferation, especially for megakaryocytes(19). When the TPO ligand binds to the receptor, this interaction induces the homodimerization of the MPL receptor and triggers JAK2 activation and the positive signal is transduced to activate the JAK-STAT pathway(20). MPL receptor presents two extracellular receptor domains, a transmembrane domain and an intracellular domain. The transmembrane domain, codified by exon 10 of *MPL* gene is essential to keep the receptor inactive as it prevents receptor homodimerization in absence of the ligand(21).

Mutations in *MPL* gene, identified for the first time in 2006(19), are frequently located in exon 10 affecting the transmembrane domain of the MPL receptor. MPL mutated receptor is aberrantly activated in absence of ligand, inducing

the receptor homodimerization and activation of the JAK-STAT pathway, inducing uncontrolled megakaryopoiesis and MPN phenotype(22).

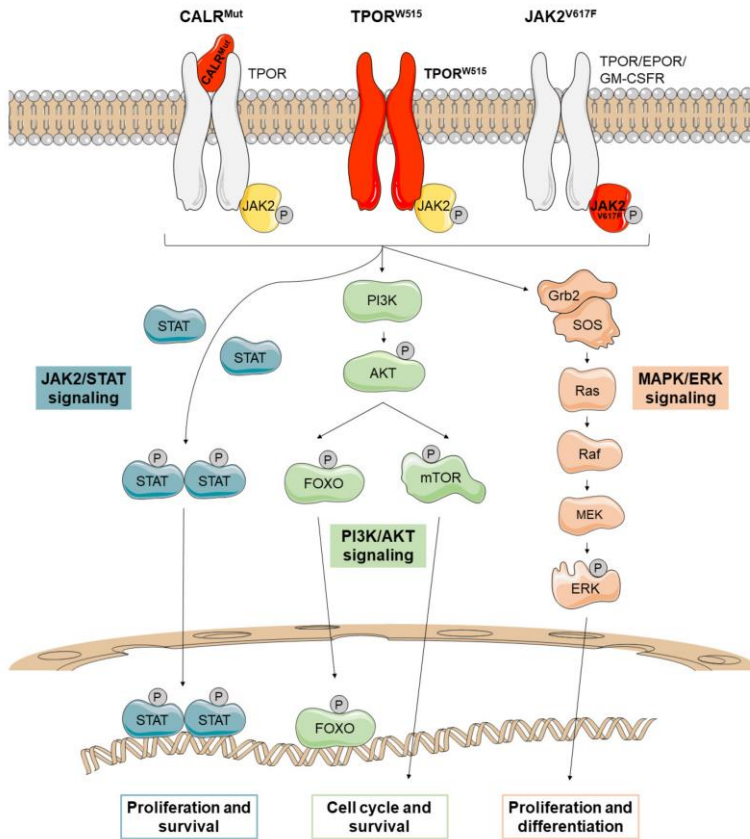


Figure 2. JAK2, CALR and MPL mutations in the JAK-STAT Pathway.
(Guijarro-Hernández *et al.*, *Cancers* 2021)(23)

SH2B3

SH2B Adaptor Protein 3 (*SH2B3*) gene, also named lymphocyte adapter protein (*LNK*), codifies for the SH3B3 protein which is a key regulator of normal hematopoiesis. SH2B3 acts as a negative regulator of the JAK-STAT pathway and therefore is involved in cell proliferation control(24).

SH2B3 mutations are missense mutations that can occur in all exons, but are more frequently found in exon 2. This exon codifies for the pleckstrin homology (PH) domain, essential for SH2B3 functions. Mutations are mainly

changes that reduce the level of activity of SH2B3, increase the activity of the JAK-STAT pathway and are associated to uncontrolled proliferation(24,25).

1.2.1.2 RAS pathway: KRAS, NRAS, CBL, NF1, PTPN11

RAS pathway (also known as Ras-Raf-MEK-ERK pathway) is implicated in signaling transduction from cellular receptors including receptor tyrosine kinases (RTKs) such as epidermal growth factor receptor (EGFR) to the cell nucleus. When the receptor is activated by the ligand, it triggers the kinase activity of the cytoplasmic domain of the receptor. This activates a cascade of protein-protein interaction that leads to SOS activation, which removes a guanosine diphosphate (GDP) molecule from RAS proteins (such as KRAS and NRAS). RAS can then bind guanosine triphosphate (GTP) and become active. Activated RAS triggers a kinase cascade leading to the phosphorylation of mitogen-activated protein kinase (MAPK). MAPK is then activated and coordinates the activity of transcription factor implicated in cell cycle regulation and proliferation (Figure 3). Mutations affecting the RAS pathway generally lead to uncontrolled proliferation of the mutated clone(26).

RAS pathway mutations are frequent in CMML (30%) and particularly in proliferative variants of CMML(27). In MDS, these mutations are found in 5-10 % of cases. In MPNs, RAS pathway mutations are identified mainly in patients with myelofibrosis (5-10%) (28–32).

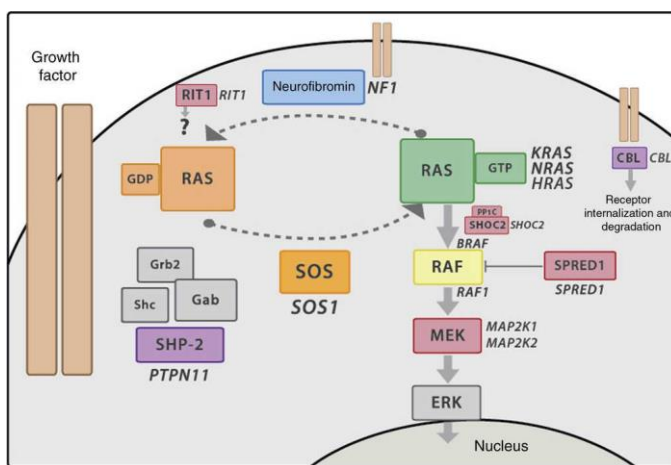


Figure 3. RAS pathway signaling. (Carcavilla et al., An Pediatr, 2020) (33)

KRAS

Kirsten rat sarcoma virus (*KRAS*) gene encodes for a GTPase, a protein that is activated and inactivated by the binding to GTP and GDP, respectively. *KRAS* gene is known to be an oncogene and is mutated in several malignancies such as lung or colorectal cancer(34).

KRAS mutations in myeloid malignancies are generally missense mutations which occur in the protein positions 12, 13, 61 and 146. These changes are activating mutations that induce aberrant activation of the RAS pathway in absence of receptor activation(35).

NRAS

Neuroblastoma ras viral oncogene homolog (*NRAS*) codifies for the NRAS protein, which is very similar to *KRAS*. NRAS is also a GTPase known to be an oncogene.

Similar to *KRAS* gene, *NRAS* activating mutations mainly occur in protein positions 12,13 and 61. However *NRAS* mutations in myeloid malignancies are slightly more frequent than *KRAS* mutations(35,36).

CBL

Casitas B-lineage Lymphoma (*CBL*) gene codifies for the CBL protein, a RING finger E3 ubiquitin ligase. CBL is involved in several signal transduction pathways that are essential for hematopoietic stem cell maintenance. In normal conditions, this protein tags the substrates by ubiquitination to trigger their degradation by the proteasome.

CBL mutations in myeloid malignancies are mainly detected in protein positions 366-420. This region codifies for the RING finger domain and hotspot mutations in *CBL* are known to disrupt its ubiquitin ligase activity(37,38). In consequence, mutated *CBL* reduces the ubiquitin mediated degradation of tyrosine kinases(39) which leads to aberrant activation of the RAS pathway(40). The JAK-STAT and PI3K pathways are also deregulated by mutated *CBL*(41,42). Overall, *CBL* mutations alter the regulation of cell proliferation and differentiation.

NF1

The *neurofibromin 1* gene (*NF1*), encodes for the protein neurofibromin. This protein is also a RAS GTPase with a key role in cell proliferation regulation. Neurofibromin is predominantly implicated in the activation and regulation of the RAS pathway(43,44).

Germline mutations in *NF1* are causative of neurofibromatosis type 1 syndrome(45). On the other hand, somatic mutations in *NF1* have also been described, and are frequently truncating or splicing mutations distributed throughout the gene. They are loss of function mutations that impede its regulatory function, thus inducing aberrant activation of the RAS pathway(46). *NF1* mutations frequently co-occur with a NF1 mutation in the other allele(47).

PTPN11

The Protein Tyrosine Phosphatase Non-Receptor Type 11 gene (*PTPN11*) encodes for the SHP2 protein, a tyrosine phosphatase that regulates signaling of the RAS pathway. The SH2 domain of SHP2 is the active site of the protein and is essential for its catalytic function.

Germline mutations in *PTPN11* are causative of Noonan syndrome(48). In myeloid malignancies, somatic mutations in *PTPN11* gene have been described, and occur along the whole codifying region, however, the regions coding for the SH2 domains are more frequently affected. These are frequently missense mutations affecting this active site and leading to gain of function changes that activate several pathways, including RAS(49).

1.2.2 Epigenetic regulation

Epigenetic cell processes are those involved in DNA modifications without the alteration of the DNA sequence itself and comprise, among others, DNA methylation (*TET2*, *DNMT3A*, *IDH1*, *IDH2*) and histone modification (*ASXL1*, *EZH2*, *KMT2A*).

1.2.2.1 DNA methylation: *TET2*, *DNMT3A*, *IDH1/IDH2*

DNA methylation is an epigenetic mechanism in which methyl groups are added to the DNA, generally to CpG islands. This process is essential to regulate gene expression. In MDS and CMML, mutations in genes implicated in DNA methylation (i.e. *TET2*, *DNMT3A* and *IDH1/2*) have been described as driver alterations in the disease pathogenesis. In this context, DNA methylation alterations in myeloid malignancies produce hypermethylation of genes implicated in proliferation and hematopoietic regulation(50).

The DNA methylation and demethylation process is a cyclic reaction to activate and inactivate gene expression. The addition of a methyl group to cytosines generates 5-methylcytosines (5-mC) which inhibits gene expression(51). To reverse this reaction, a previous step of hydroxymethylation is required: a hydroxyl group is added to 5-mC to generate 5-hydroxymethylcytosines (5-hmC). This leads to final DNA demethylation (Figure 4).

Mutations in DNA methylation genes are frequent in CMML (45-70%)(52–54). In MDS, these mutations are also common, found in 20-25% of cases (28–30). In MPN, DNA methylation genes are mutated in 10-20% patients (31,32).

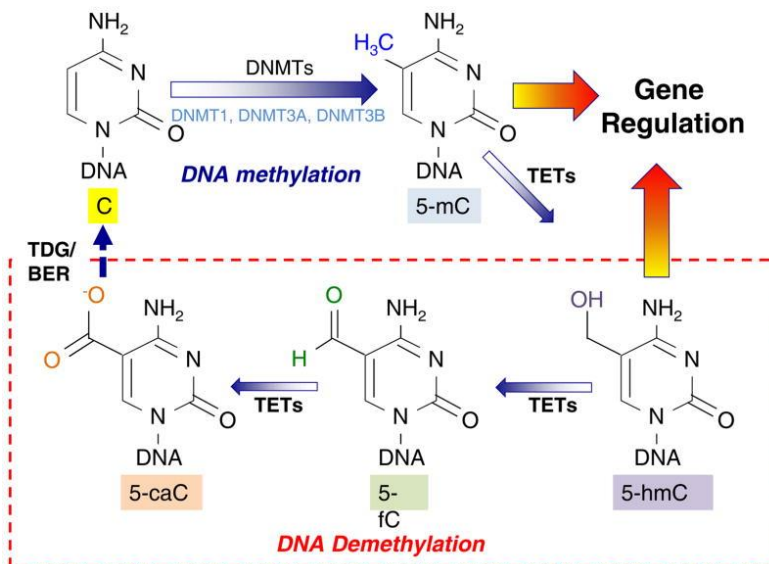


Figure 4. Methylation and demethylation process.

(Abe et al., *ADV EXP MED BIOL*, 2017)(55)

TET2

The Ten-Eleven Translocation-2 (*TET2*) gene encodes for the TET2 protein, a dioxygenase that catalyzes the switch of 5-methylcytosine to 5-hydroxymethylcytosine. TET2-mediated demethylation is required to coordinate transcription(56,57).

Mutations in *TET2* gene can be missense or truncating mutations that are distributed all over the gene. *TET2* is a large and polymorphic gene which further complicates accurate variant interpretation. *TET2* mutations are mainly loss of function mutations that impede its catalytic function in hydroxymethylation. Therefore, *TET2* loss of function leads to an increase in global methylation and in consequence those genes with methylated promoters are less expressed(58). Mutations in *TET2* are the most characteristic of CMML and frequently co-occur with *SRSF2* mutations(52–54).

DNMT3A

The DNA methyltransferase 3 alpha (*DNMT3A*) gene codifies for DNMT3A protein a DNA methyltransferase that is involved in *de novo* methylation(59). DNMT3A-mediated methylation is essential for epigenetic silencing of genes involved in hematopoietic stem cell regulation and differentiation(60).

Mutations in *DNMT3A* are generally missense or truncating mutations close to the methyltransferase domain of the protein, although mutations have been identified throughout the whole sequence. The most frequent mutation is a missense mutation in position 882 (p.Arg882His). These mutations are loss on function mutations that impede *de novo* methylation. *DNMT3A* mutations produce global hypomethylation in DNA, inducing aberrant expression of genes implicated in hematopoiesis regulation. Consequently, *DNMT3A* mutated cells are incapable of efficiently differentiate(60). Of note, *DNMT3A* mutations are the most frequently found alterations when analyzing the molecular profile of healthy cohorts, which is further explained in the next subsection *1.4 Clonal hematopoiesis of indeterminate potential*.

IDH1/IDH2

The isocitrate dehydrogenase 1 and 2 (*IDH1* and *IDH2*) encode two similar proteins (*IDH1/2*) involved in the metabolism of 2-oxoglutarate, the principal cofactor of TET2 protein. *IDH1/2* proteins catalyze the conversion of isocitrate to 2-oxoglutarate which is an essential metabolite for hydroxymethylation.

Mutations in *IDH1* and *IDH2* are generally detected in exon 4, in particular in protein positions R132 for *IDH1* and R140/R172 for *IDH2*. Mutated *IDH1/2* disrupts isocitrate conversion to 2-oxoglutarate altering cell metabolism and inhibiting correct hydroxymethylation (61). Altered IDH activity also produces the accumulation of 2-hydroxyglutarate which blocks cell differentiation and leads to an hypermethylated state of the hematopoietic cells. In line with this, mutations in *IDH* genes and *TET2* are almost always mutually exclusive(62).

1.2.2.2 Histone modification: *ASXL1*, *EZH2*

Histones are DNA binding proteins that wrap DNA to conform chromatin structure and are crucial for gene expression regulation. The process of histone modification is a post-transcriptional mechanism that produces covalent changes to histone proteins, such as methylation or acetylation (Figure 5). These changes are essential for chromatin structure regulation and therefore to coordinate gene expression. In myeloid malignancies, the most frequently altered genes related to histone modification are implicated in the mechanism of histone methylation (*ASXL1* and *EZH2*).

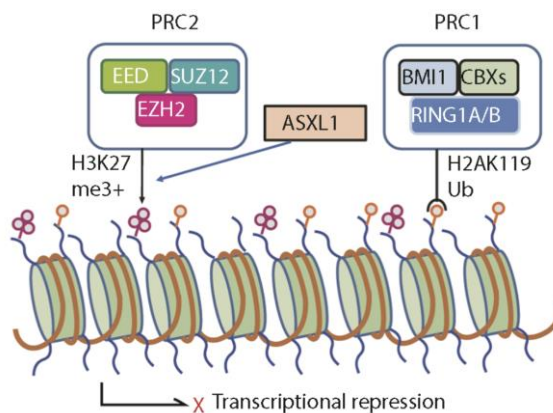


Figure 5. Histone modification by the PRC2 complex.
(Adapted from Ogawa, *Blood*, 2019) (30)

Mutations in genes implicated in histone modifications have been described in 15% of MDS and in 40-50% of CMML(28–30,52–54). In MPN, these mutations are present in 15-25% of cases with myelofibrosis, while they are infrequent in the other MPN subgroups(31,32).

ASXL1

The ASXL transcriptional regulator 1 (*ASXL1*) gene codifies for a polycomb-related protein. ASXL1 protein interacts with the polycomb repressive complex 2 (PRC2) to trimethylate histones. ASXL1-PRC2 interaction is required for trimethylation which is a key mechanism to regulate repression of genes implicated in hematopoiesis.

Mutations in *ASXL1* are mainly truncating or nonsense alterations in exon 12 that produce an aberrant C- terminal domain. These alterations result in a dysfunctional ASXL1 protein and therefore gene repression via PRC2 is defective leading to a dysregulation of hematopoiesis(63,64).

EZH2

The enhancer of zeste 2 polycomb repressive complex 2 (*EZH2*) gene encodes for a subunit of the PRC2 complex. EZH2 protein is a catalytic component of PRC2 with a key role in histone methylation as a mechanism of gene repression. Interestingly, EZH2 has also been associated with activation of DNA methyltransferases, therefore linking the mechanisms of DNA methylation and histone modifications(65).

Mutations in *EZH2* are loss of function missense or truncating mutations distributed all over the coding sequence. In consequence, *EZH2* mutations impede PRC2 activity and deregulate repression of genes implicated in hematopoiesis(66,67).

1.2.3 Splicing: *SF3B1, SRSF2, U2AF1, ZRSR2, PRPF8, DDX41*

The process of RNA modification to generate mature messenger ribonucleic acids (mRNA) is known as splicing. In this stage, immature RNA (pre-RNA) is processed to eliminate introns and join exonic regions, which is required for correct mRNA translation into proteins. This process is performed in the cell

nucleus by the spliceosome, a large protein complex formed by numerous proteins and small nuclear RNAs (snRNA) (Figure 6). Alterations in the splicing machinery can generate aberrant mRNAs, induce exon skipping and alter protein translation.

Splicing genes are altered across several subgroups of myeloid malignancies. In MDS, splicing genes are altered in the majority of patients, being the most frequently mutated gene group (45-83%) (68–71). For CMML patients, splicing machinery is also frequently altered (30-50%). However, in MPNs splicing mutations are overall infrequent, except for those cases with myelofibrosis (5-15%) in which these mutations are generally associated to poor outcomes(31,32). Splicing mutations occur early in the disease onset of MDS and CMML, while in MPNs they are associated to a later stage of the disease and is related to clonal evolution.

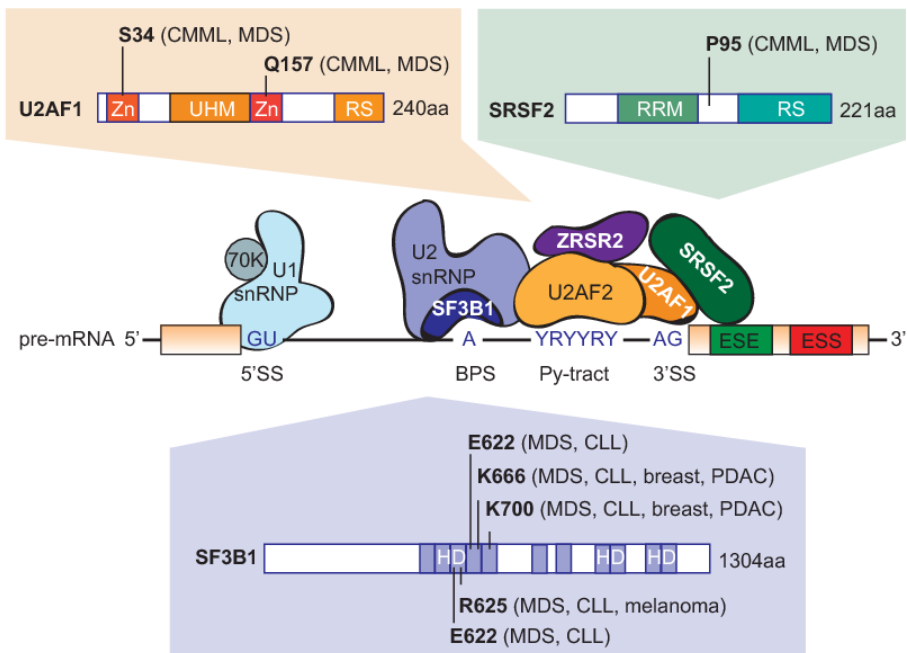


Figure 6. Splicing machinery and most frequent mutations in *U2AF1*, *SRSF2* and *SF3B1*.
(Anczukow *et al.*, *RNA*, 2016) (72)

SF3B1

SF3B1 gene codifies for SF3B1 protein, the largest protein of the spliceosome factor 3b (SF3B) complex. SF3B1 is a key component in the spliceosome, essential for intron point sequence identification and for spliceosome assembly(73).

SF3B1 mutations are predominantly located in exons 14–15, the hotspot of the gene, which codifies for the HEAT domain of the protein. *SF3B1* p.K700E is the most recurrent mutation in *SF3B1* where it is detected in 50% of *SF3B1* mutated cases. Other *SF3B1* mutations are also frequently detected in the HEAT domain, mainly in codons from 622 to 700. The HEAT domain is critical for RNA recognition and binding of the spliceosome to RNA. *SF3B1* mutations produce alterations in the splicing process of genes implicated in several pathways. Altered SF3B1 protein induces aberrant splicing in several genes implicated in iron transportation in mitochondria, which is causative of abnormal iron deposition in the cell around mitochondria. In consequence, patients with *SF3B1* mutations are strongly associated with a specific phenotype in bone marrow cells, the presence of ring sideroblasts. *SF3B1* mutations are detected in 30% of MDS patients, while the frequency rises up to 80% in the subgroup of MDS with ring sideroblasts(74,75).

SRSF2

The serine and arginine rich splicing factor 2 (*SRSF2*) gene encodes for the SRSF2 protein which is part of the spliceosome. SRSF2 recognizes the intron-exon limits, promotes spliceosome binding to splice sites and regulates alternative splicing(76,77).

Mutations in *SRSF2* are mainly missense mutations affecting the hotspot P95 in exon 1. These mutation affects SRSF2 protein in a proximal position to the RNA binding domain (RRM), which alters the spliceosome union to RNA(78). Mutant SRSF2 alters the splicing of thousands of genes, which contributes to deregulated hematopoiesis (79). *SRSF2* mutations frequently co-occur with mutations in *TET2* or *ASXL1* genes in CMML (30-50%). These mutations are

also found in 10-15% of MDS(69,70). In MPNs, they are present in 9-17% of cases with myelofibrosis but are infrequent in the other disease subgroups.

U2AF1

The U2 Small Nuclear RNA Auxiliary Factor 1(*U2AF1*) gene encodes for the U2 auxiliary factor protein which is a component of the spliceosome. *U2AF1* protein is essential for the spliceosome binding to the pre-mRNA branch site. It also mediates the interactions between the spliceosome and the subunits that bind to the mRNAs.

Mutations affecting the *U2AF1* gene are missense mutations in protein positions 34 and 157 located in the region encoding for the two zinc finger domains. These alterations induce aberrant splicing patterns that are associated to altered hematopoiesis(80,81).

ZRSR2

The Zinc Finger CCCH-Type, RNA Binding Motif And Serine/Arginine Rich 2 (*ZRSR2*) codifies for the *ZRSR2* protein, a pre-mRNA-binding protein. It recognizes the 3'-splice sites of mRNA and is essential for the spliceosome assembly.

Mutations in *ZRSR2* are detected all along the coding sequence and are frequently truncating mutations. These are loss of function mutations that impede the correct recognition of the splice sites(82).

PRPF8

The pre-mRNA Processing Factor 8 (*PRPF8*) gene encodes for the *PRPF8* subunit of the spliceosome. *PRPF8* is the most evolutionary conserved protein of the spliceosome indicating the relevance of its functionality(83). It is essential for interactions with the 5' and 3' splice sites of pre-mRNA in splicing and alternative splicing(84,85).

Mutations in *PRPF8* are predominantly missense mutations that scatter all over the gene. Mutations in *PRPF8* gene lead to altered pre-mRNA splicing, producing aberrant splicing in genes involved in iron metabolism and

hematopoiesis regulation (86). Accordingly, *PRPF8* mutations have been associated with the presence of ring sideroblasts(86).

DDX41

The DEAD-Box Helicase 41 (*DDX41*) gene codifies for the RNA helicase *DDX41*. This protein is involved in splicing, translation and RNA structure modifications.

Missense and truncating mutations in *DDX41* gene have been identified throughout the coding sequence. Germline and somatic mutations in *DDX41* have been detected in MDS. Germline mutations predispose to MDS or AML and are located close to the C-terminal domain. Somatic *DDX41* mutations frequently co-occur with germline *DDX41* mutations, and are closer to the N-terminal domain(87). Of note, cases with germline *DDX41* mutations develop MDS disease phenotype at approximately 65 years, similarly to de novo MDS or AML(88,89).

1.2.4 Transcription factors: *RUNX1*, *SETBP1*, *ETV6*, *BCOR/BCORL1*, *PHF6*, *CUX1*, *GATA2*

Transcription factors are proteins directly involved in initiation of transcription. These proteins regulate RNA synthesis by binding to DNA at the promoter sequences and recruiting the transcription complex. The most frequently altered genes encoding for transcription factors in myeloid malignancies are *RUNX1*, *SETBP1*, *ETV6*, *PHF6* and *CUX1*.

Alteration in these genes are detected in 10-20% of MDS and CMML cases (28–30,52–54), while in MPN these mutations are infrequent(<5%)(31,32).

RUNX1

The Runt-related transcription factor 1 (*RUNX1*) encodes for a protein of the core-binding factor. *RUNX* proteins present a runt-homology domain to bind DNA and promote transcription. *RUNX1* participates in regulation of the expression of several genes implicated in hematopoiesis.

RUNX1 mutations have been identified throughout the whole sequence, and can be missense or truncating. These are loss of function mutations that disrupt the transcription factor function of the protein(90). Consequently, *RUNX1* mutated cases show overexpression of several genes involved in hematopoiesis, and are associated to abnormal proliferation(91).

SETBP1

The Set-Binding Protein 1 (*SETBP1*) codifies for a nuclear protein that interacts with the SET protein, a nuclear oncogene implicated multiple cellular processes such as DNA replication and transcription regulation(92).

Mutations in *SETBP1* are distributed mainly in the hotspot encoding for the SKI-homologous domain (protein positions 858-871)(93). These mutations directly affect gene transcription and induce aberrant cell proliferation (94).

ETV6

The ETS Variant Transcription Factor 6 (*ETV6*) gene encodes for a transcription factor with DNA binding functions. *ETV6* plays a key role in hematopoietic regulation by acting as a transcription repressor(95).

ETV6 alterations in myeloid malignancies include translocations, deletions and point mutations. Point mutations are frequently truncating alterations that disrupt *ETV6* interactions either with other proteins or with DNA(96). *ETV6* mutations have been associated to uncontrolled cell proliferation of the mutated clone(97).

BCOR/BCORL1

The BCL6 Corepressor (*BCOR*) and BCL6 Corepressor Like 1 (*BCORL1*) genes encodes for two transcriptional corepressors. *BCOR* and *BCORL1* have similar functions and are co-repressors of the BCL6 oncogene(98,99).

BCOR and *BCORL1* mutations are truncating or missense mutations distributed throughout the whole sequence, with *BCOR* gene showing a hotspot position in codon N1425(100). These alterations are loss of function mutations that are associated to late expansion of myeloid

malignancies(100). *BCOR* and *BCORL1* mutations induce a deregulation of the hematopoiesis control(101,102).

PHF6

The PHD Finger Protein 6 (*PHF6*) encodes for a tumor suppressor protein involved in the regulation of rRNA synthesis. It is also involved in cell proliferation, cell cycle and DNA damage(103).

PHF6 mutations are truncating or missense mutations that have been describe throughout the whole sequence(104). These mutations predominantly affect male gender since *PHF6* is located at X chromosome. This suggests that, in females, *PHF6* mutated in heterozygosis is not sufficient to produce clonal expansion(105).

CUX1

The CUT-like homeobox 1 (*CUX1*) gene codifies for a transcription factor with a tumor suppressor function(106). *CUX1* is implicated in maintenance of hematopoietic progenitors and also in the base excision DNA repair mechanism(107).

CUX1 deletions and point mutations have been identified in myeloid malignancies. Loss of function mutations in *CUX1* lead to deficient DNA repair and therefore contribute the accumulation of other mutations(108).

GATA2

The GATA Binding Protein 2 (*GATA2*) gene encodes for a transcription factor involved in proliferation and maintenance of hematopoietic stem cells(109).

Mutation in *GATA2* are frequently germline mutations that lead to myelodysplasia at early age(110). These alterations are recurrent in the hotspot encoding for the zinc finger domains (exons 4-6). *GATA2* deficiency alters the self-renewal capacity of hematopoietic stem cells(109,111).

1.2.5 DNA repair: *TP53*, *PPM1D*, *CHEK2*, *ATM*

DNA is susceptible to biochemical alterations produced by external agents or DNA polymerase errors. In this context, DNA repair machinery, cell cycle regulators or cell death pathways (if damage is unsustainable) are activated. Dysregulations of genes implicated in these functions have been identified throughout the majority of cancer types.

In myeloid malignancies, the most frequently altered genes are *TP53*, *PPM1D*, *CHEK2* and *ATM*.

TP53

The tumor protein p53 (*TP53*) encodes for the TP53 tumor suppressor protein, which is essential for DNA damage response and regulation of cell cycle and apoptosis. *TP53* is the most frequently mutated gene in cancer since the disruption of *TP53* function affects cell proliferation and survival.

TP53 mutations are distributed all over the coding sequence and are known to alter the tumor suppressor function of the protein(112). Mutated *TP53* in myeloid malignancies is known to induce genetic instability and is generally associated to worse outcomes.

PPM1D

The Protein Phosphatase, Mg²⁺/Mn²⁺ Dependent 1D (*PPM1D*) codifies for a protein phosphatase implicated in cell cycle regulation. PPM1D acts as a negative regulator of p53 protein to stop G2 checkpoint in response to stress(113).

Mutations in *PPM1D* are frequently nonsense and frameshift mutations occurring in exons 5 and 6. These alterations are loss of function mutations that disrupt cell cycle regulation. *PPM1D* mutated cells become tolerant to DNA damage and resistant to apoptosis(113).

CHEK2

The Checkpoint Kinase 2 (*CHEK2*) gene codifies for a cell cycle checkpoint regulator protein. *CHEK2* is a tumor suppressor protein which is activated in response to DNA damage(114).

The majority of *CHEK2* mutations reported are germline alterations that are associated to an increased risk for cancer(115). However, *CHEK2* mutations can also be somatic and are distributed throughout the whole sequence. Similar to the other DNA repair genes, these mutations produce a dysregulation of DNA damage response machinery, which gives selective advantage to this cells(114).

ATM

The Ataxia Telangiectasia Mutated (*ATM*) gene encodes for a protein kinase that is implicated in cell cycle checkpoint through interaction with p53 and *CHEK2* proteins(116).

Germline mutations in *ATM* gene in homozygosis are causative of ataxia telangiectasia(117). Somatic mutations in *ATM* are found all over the sequence, but frequently affect the kinase domain. *ATM* mutated cells have selective advantage by escaping the traditional cell-cycle checkpoints(116).

1.2.6 Cohesins: *STAG2*, *RAD21*

Cohesins are multi-protein subunit complexes that maintain the sister chromatids together in mitosis and meiosis. They are also implicated in damaged DNA repair and in gene expression regulation in mitosis and proliferating cells(118).

STAG2

The Stromal Antigen 2 Gene (*STAG2*) encodes for a subunit of the cohesin complex. *STAG2* gene is the most frequently altered cohesin gene and has been found to be mutated across several cancer types, producing chromatid cohesion defects.

STAG2 mutations in myeloid malignancies are loss-of-function alterations(119,120). These mutations were thought to induce altered chromosomal separation in cell division, however, this is not observed in myeloid malignancies. It has been demonstrated that *STAG2* mutations impede hematopoietic stem cell differentiation by regulating chromatin accessibility and transcription factor activity(121).

RAD21

The *RAD21* Cohesin Complex Component (*RAD21*) gene codifies for a subunit of the cohesin complex. *RAD21* protein is involved in the repair of DNA double-strand breaks.

RAD21 mutations are generally truncating or missense mutations that impede its functionality(119,120). Mutated *RAD21* is implicated in genomic instability and induces aberrant hematopoietic self-renewal by epigenetic repression(122).

1.2.7 Genes, mutational frequency and clinical implications

The mutational frequency and clinical implications of the most frequently altered genes in MPN, MDS and CMML are listed in Table 2(123,124).

Table 2. Most frequently mutated genes in MPN, MDS and CMML.

Gene function	Gene (location)	Biological function	Frequent mutations	Mutational frequency			Clinical relevance			
				MPN	MDS	CMML	MPN	MDS	CMML	
CELL SIGNALING	JAK-STAT	JAK2 9p24.1	Tyrosine kinase, hyperactivating mutations produce abnormal myeloproliferation	JAK2V617F exon 12 mutations	PV 95-97% ET 50-60% PMF 50-60%	<5%	2-10%	High VAF associated to thrombosis and fibrosis	Associated to del(5q) and <i>SF3B1</i> (MDS/MPN with RS-T)	Associated with p-CMML
		CALR 19p13.2	Multi-functional chaperone, mutations activate JAK-STAT via MPL	Ins/del in exon 9	ET 15-30% PMF 23-35%	<1%	<1%	Associated to good prognosis	Uncertain	Uncertain
		MPL 1p34.2	TPO receptor, mutations induce uncontrolled megakaryopoiesis	Exon 10: W515K/L/A, S505N	ET 3-5% PMF 5-10%	<1%	<1%	Higher risk of thrombosis and fibrosis	Uncertain	Uncertain
		SH2B3 12q24.12	Negative regulator of JAK-STAT	Exon 2	PV 1% ET/PMF 3-6%	<1%	5-7%	Associated to AML transformation	Uncertain	Uncertain
	RAS	KRAS 12p12.1	GTPase in RAS/ MAPK pathway	Codons: 12,13,61,146	PMF 6%	5-10%	10-20%	Higher risk of fibrosis and lower overall survival	Associated to worse outcomes	Associated with p-CMML
		NRAS 1p13.2	GTPase in RAS/ MAPK pathway	Codons: 12,13,61		5-10%	10-20%		Associated to worse outcomes	Associated with p-CMML and to worse outcomes
		CBL 11q23.3	Ubiquitin ligase, mutations activate RAS pathway	Codons: 336-420	PMF 5-10%	<5%	8-18%	Associated to AML transformation	Associated to worse outcomes	Associated to lower overall survival
		NF1 17q11.2	RAS GTPase, regulates cell proliferation	Truncating, all exons	5%	1%	4%	Uncertain	Uncertain	Uncertain
		PTPN11 12q24.13	Tyrosine phosphatase, RAS signaling	Sh2 domain, all exons	PMF 5-10%	<1%	4%	Fibrotic AML transformation and lower overall survival	higher bone marrow blasts and worse outcomes	Associated to lower overall survival
		EPIGENETIC REGULATION	DNA METHYLATION	TET2 4q24	Methylcytosine, key role in DNA demethylation	Loss of function mutations, all exons	PV 10-20% ET 10-15% PMF 10-20%	20-30%	45-70%	Inconclusive
DNMT3A 2p23.3	Methyltransferase involved in <i>de novo</i> methylation			R882H, all exons	PV/ET 1-5% PMF 5-15%	12-18%	2-10%	Associated to AML transformation	Associated to worse outcomes	AML transformation and worse outcomes
IDH1/IDH2 2q34/15q26.1	Dehydrogenases, metabolism of 2-oxoglutarate, cofactor of TET2			R132 for IDH1 R140/R172 for IDH2	PMF 3-6%	<5%	5-10% for IDH2, <1% for IDH1	Lower overall survival	Associated to worse outcomes	Associated to worse outcomes
HISTONE MODIFICATION	ASXL1 20q11.21		Regulates transcription via histone methylation	truncating or nonsense alterations in exon 12	PV/ET 2-5% PMF 13-25%	5-25%	40-50%	Higher risk of thrombosis and fibrosis, lower overall survival	Associated to worse outcomes	AML transformation and worse outcomes
	EZH2 7q36.1		Key role in histone methylation for gene repression	missense or truncating mutations, all exons	PV 1-3%	5-10%	5-12%	Fibrosis, AML transformation and lower overall survival	Associated to worse outcomes	Associated to worse outcomes and CMML-2

Gene function	Gene (location)	Biological function	Frequent mutations	Mutational frequency			Clinical relevance		
				MPN	MDS	CMML	MPN	MDS	CMML
SPLICING	SF3B1 2q33.1	Key component in the spliceosome, intron sequence identification	K700E exons 14-16	ET 3% PMF 4-7%	ET 1% PMF 5-10%	5-10%	Fibrotic transformation. Evolution to MDS with RS	Associated to RS, good prognosis	Uncertain
	SRSF2 17q25.1	Recognizes intron-exon limits, key in spliceosome binding	P95, exon 1	PMF 9-17%	10-15%	30-50%	AML transformation and lower overall survival	Associated to worse outcomes	Associated to worse outcomes
	U2AF1 21q22.3	Spliceosome binding to the pre-mRNA branch site	Codons 34 and 157, exons 2-6	PMF 15%	8-12%	5-10%	Disease progression	Associated to worse outcomes	Associated to worse outcomes
	ZRSR2 Xp22.2	Recognizes the 3'-splice sites, key for spliceosome assembly	Truncating mutations, all exons	ET 3% PMF 5-10%	5-10%	5-10%	Fibrosis/AML transformation. Lower overall survival	Unknown	Unknown
	PRPF8 17p13.3	Interactions with 5' and 3' pre-mRNA splice sites	Missense, all exons	<1%	1-4%	<1%	Uncertain	Associated to RS	Uncertain
	DDX41 5q35.3	Key in splicing, translation and RNA modifications	Missense/truncating, all exons	1%	3%	2%	Uncertain	Associated to high risk MDS	Uncertain
TRANSCRIPTION FACTORS	RUNX1 21q22.12	Subunit of core-binding factor, promotes transcription	Missense/truncating, all exons	PV/ET 3-5% PMF 5%	10-15%	10-30%	AML transformation and lower overall survival	Associated to worse outcomes	AML transformation and worse outcomes
	SETBP1 18q12.3	Interacts with SET, implicated in DNA replication and transcription	Codons: 858-871	1-4%	<5%	5-10%	Uncertain	Associated to worse outcomes	AML transformation and worse outcomes
	ETV6 12p13.2	Transcription factor, binds DNA to repress transcription	Translocations / deletions/ point mutations	<1%	<5%	<1%	Uncertain	Associated to worse prognosis and del(7)	Uncertain
	BCOR/BCORL1 Xp11.4/ Xq26.1	Co-repressors of the BCL6 oncogene	Truncating/missense All exons BCOR N1425	<1%	<5%	<5%	AML transformation and lower overall survival	Inconclusive	Uncertain
	PHF6 Xq26.2	Tumor suppressor protein, regulates rRNA synthesis	Missense/truncating, all exons	<1%	~3%	~5%	AML transformation and lower overall survival	Associated to higher molecular complexity	Uncertain
	CUX1 7q22.1	Transcription factor, implicated in hematopoiesis and DNA repair	Missense/truncating, all exons	<1%	2%	3%	AML transformation and lower overall survival	Associated to lower overall survival	Associated to lower overall survival
	GATA2 3q21.3	Transcription factor: proliferation and hematopoiesis	Missense/truncating, exons 4-6	<1%	<5%	<1%	Uncertain	AML transformation and lower overall survival	Uncertain
	TP53 17p13.1	DNA damage response, cell cycle and apoptosis regulation	Loss of function mutations, all exons	PV/ET 1% PMF <5%	8-12%	<5%	AML transformation and lower overall survival	Associated to complex karyotype and worse outcomes	Associated to worse outcomes
DNA REPAIR	PPM1D 17q23.2	Phosphatase, p53 regulator	nonsense and frameshift, exons 5,6	2%	<1%	<1%	Associated to chemotherapy resistance	Associated to therapy related MDS	Uncertain
	CHEK2 22q12.1	Cell cycle checkpoint regulator	Missense/truncating, all exons	<1%	<1%	<1%	Uncertain	Associated to worse outcomes	Uncertain
	ATM 11q22.3	Kinase, cell cycle regulator	Missense/truncating, all exons	<1%	<1%	<1%	Uncertain	Uncertain	Uncertain
	STAG2 Xq25	cohesin subunit, affects chromatin conformation	Missense/truncating, all exons	<1%	5-10%	5-10%	Uncertain	Lower overall survival	Associated to worse outcomes
COHESINS	RAD21 8q24.11	cohesin subunit / hematopoietic regulation	Missense/truncating, all exons	<1%	1-3%	<1%	Uncertain	Lower overall survival	Uncertain

PV: polycythemia vera, ET: essential thrombocythemia, PMF: primary myelofibrosis, d-CMML: dysplastic CMML, p-CMML: proliferative CMML, VAF: variant allele frequency, ins: insertion, del: deletion, HMA: hypomethylating agents, RS: ring sideroblasts, RS-T: ring sideroblasts and thrombocytosis.

1.3 Role of molecular diagnosis in myeloid neoplasms. Main limitations.

Genomic characterization holds the potential to further improve clinical decision making since certain molecular alterations are associated with distinct clinical manifestations. Moreover, the finding of these molecular markers has become crucial to understand the pathogenesis of these diseases. The molecular complexity of myeloid malignancies contributes to the clinical heterogeneity observed in these patients.

Over the last decade, the widespread use of Next Generation Sequencing (NGS) has resulted in a revolution in favor of personalized medicine. The application of genomics in myeloid neoplasms enhances the capacity for personalized interventions by offering accurate diagnosis, risk stratification and personalized treatments. A greater number of patients can be studied using molecular biology techniques, which represents an important advance in clinical management.

However, molecular diagnosis still presents some clear limitations. The sensitive of NGS and other classical molecular biology techniques is limited. In consequence, subclonal or newly acquired mutations poorly represented in the sample may be missed. Higher sensitivity techniques for point mutations such as digital PCR (dPCR) are being incorporated in the routine assessment, however, myeloid malignancies show a complex molecular landscape that requires the analysis of a broad range of mutations. In this context, optimization of the molecular biology protocols to improve sensitivity is necessary until new high sensitivity sequencing approaches are available for routine analysis.

Another major limitation that may affect sensitivity of molecular studies in myeloid malignancies is the obtention of optimal samples to extract DNA and perform these analyses. In MPNs, molecular characterization is conventionally performed in isolated granulocytes, whole blood samples or bone marrow specimens. Peripheral blood (PB) granulocytes isolation is time consuming and implies that mainly mature circulating cells are analyzed.

Whole blood samples are unpurified and therefore may lead to sensitivity limitations as non-mutated lymphoid cells can be present. Finally, bone marrow (BM) samples are suitable for analyzing early hematopoietic tumor cells, but the obtention of these samples require invasive procedures for the patients. In MDS and CMML, both PB and BM samples are routinely used for molecular characterization. PB samples are overall suitable for mutation analysis in MDS and CMML, however, subclonal mutations present in BM and poorly represented in PB may be undetected(125,126). Again, similarly to MPNs, BM sample obtention is invasive specially for frail elderly patients and could be insufficient in patients with fibrotic or hypocellular BM.

Besides, further studies are required since genetic alterations of unknown clinical significance are found in some cases. Currently, there is not enough scientific evidence yet to provide a reliable clinical and biological interpretation of these variants. Moreover, some of these variants may have a germline origin and therefore additional studies are required to determine whether their origin is somatic or germline.

1.4 Clonal hematopoiesis of indeterminate potential

Some of the most frequent molecular alterations of myeloid neoplasms, described in 1.2, have been recently identified in healthy individuals. The presence of clonal genetic alterations with a variant allele frequency (VAF) \geq 2%, in the absence of hematological neoplasms, is known as clonal hematopoiesis of indeterminate potential (CHIP) (Table 3) (127). CHIP is characterized by the clonal expansion of blood cells harboring somatic mutations predominantly in three genes: *DNMT3A*, *TET2* and *ASXL1* and to a lesser extent in other relevant genes in myeloid malignancies(129,130).

The frequency of CHIP is associated with age since these mutations were observed in only 1% of individuals younger than 50 years but in 10% of those older than 65. The percentage rises up to 20% in individuals older than 90 years (Figure 7). Of note, cases with CHIP have a higher risk of developing a subsequent hematologic cancer or cardiovascular disease.

Table 3. Acronyms describing clonal hematopoiesis and related conditions.

Adapted from Bejar, Leukemia, 2015(128)

Acronym	Condition	Description
ARCH	Aging Related Clonal Hematopoiesis	Presence of detectable, benign clonal hematopoiesis (defined by the presence of somatic mutations in the blood or bone marrow) whose incidence increases with age. No formal definition involving clonal abundance or types of mutations. No clinical significance is implied
CHIP	Clonal Hematopoiesis of Indeterminate Potential	Somatic mutations of myeloid malignancy associated genes in the blood or bone marrow present at $\geq 2\%$ variant allele frequency in individuals without a diagnosed hematologic disorder
ICUS	Idiopathic Cytopenia of Undetermined Significance	Patients with one or more unexplained cytopenias who do not meet diagnostic criteria for MDS or another hematologic disorder. Can occur with or without clonal hematopoiesis although often used to refer to cytopenias without evidence of clonal hematopoiesis.
CCUS	Clonal Cytopenia of Undetermined Significance	Patients with one or more unexplained cytopenias who do not meet diagnostic criteria for hematologic disorder, but who have somatic mutations of myeloid malignancy associated genes in the blood or bone marrow present at $\geq 2\%$ variant allele frequency. Can be considered as the intersection between CHIP and ICUS

At the molecular level, mutations identified in CHIP and myeloid malignancies occur in similar genes. In this regard, two key differences have been observed. First, the presence of more than one mutation in the same individual was extremely infrequent in CHIP, while it is a frequent event in myeloid malignancies. Secondly, VAF of mutations found in healthy cohorts was much lower when compared with myeloid malignancies. Of note, the mutated gene is also a decisive feature to differentiate CHIP as mutations in other genes than *DNMT3A*, *TET2* and *ASXL1* (and to a lesser extent *JAK2*, *PPM1D*, *SF3B1* and *TP53*) are extremely infrequent(129,130).

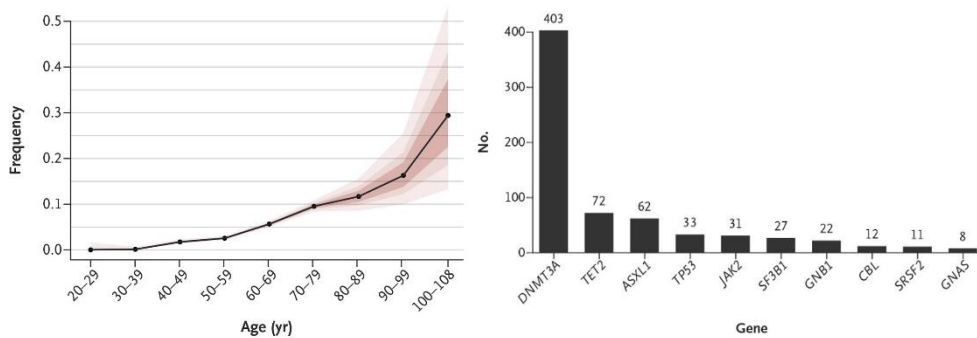


Figure 7. CHIP frequency with age and most frequently mutated genes.
(Adapted from Jaiswal *et al.*, *NEJM*, 2014) (129)

2. MYELOPROLIFERATIVE NEOPLASMS (MPN)

2.1. Introduction

Myeloproliferative neoplasms (MPN) are chronic hematological diseases characterized by the clonal expansion of mature cells from one or more myeloid lines. They are clinically characterized by a hypercellular bone marrow, splenomegaly and higher risk of thrombosis. Patients with MPN may transform to myelofibrosis and acute leukemia(131).

2.2. Classification

According to the 2017 WHO classification, MPN include the following entities(1):

- Chronic myeloid leukemia (CML), BCR-ABL1+
- Chronic neutrophilic leukemia (CNL)
- Polycythemia vera (PV)
- Primary myelofibrosis (PMF)
 - PMF, prefibrotic/early stage
 - PMF, overt fibrotic stage
- Essential thrombocythemia (ET)
- Chronic eosinophilic leukemia, not otherwise specified (NOS)
- MPN, unclassifiable

For the research projects presented in this thesis, only patients with classical (Philadelphia negative, Ph-) MPN were included: Polycythemia vera (PV), Essential thrombocythemia (ET) and Primary myelofibrosis (PMF).

2.2.1 PV

CLINICAL FEATURES

Polycythemia vera (PV) is a clonal stem cell disease characterized by proliferation of red blood cells, leukocytes and platelets with a predominance

INTRODUCTION

of red blood cell production. The disease shows a trend to present thrombotic and hemorrhagic events. Clinical management is focused to minimize the risk of thrombosis, being this complication the leading cause of morbidity and mortality in PV patients(132).

ETIOLOGY/PATHOGENESIS

Erythrocyte overproduction in PV is particular because it develops in the absence of a detectable physiologic stimulus(133). Levels of EPO, a glycoprotein responsible of stimulating red blood cell production, are lower than normal in most PV cases(134).

The discovery of a somatic gain-of-function mutation in the *JAK2* gene in 2005 led to the identification of the cause of the disease. PV is a homogeneous disease group in respect of molecular alterations, since *JAK2* mutations are identified in almost all PV patients. Approximately 96% of PV patients harbor the activating *JAK2*V617F mutation and 3% show mutations in exon 12 of *JAK2* (135,136).

DIAGNOSTIC CRITERIA

PV diagnostic criteria were revised in the 2017 World Health Organization (WHO) classification (Table 4)(1). Of note, BM biopsy is not essential for diagnosis in those cases with sustained absolute erythrocytosis, *JAK2* mutation and subnormal serum EPO level. Nevertheless, BM biopsy at diagnosis might have a predictive value regarding further evolution to myelofibrosis.

Table 4. WHO PV diagnosis criteria. Adapted from Arber *et al.*, *Blood*, 2016.**Major criteria**

1. Hemoglobin >16.5 g/dL in men
Hemoglobin >16.0 g/dL in women
or,
Hematocrit >49% in men
Hematocrit >48% in women
or,
increased red cell mass (RCM)
2. BM biopsy showing hypercellularity for age with trilineage growth (panmyelosis) including prominent erythroid, granulocytic, and megakaryocytic proliferation with pleomorphic, mature megakaryocytes (differences in size)
3. Presence of *JAK2V617F* or *JAK2* exon 12 mutation

Minor criterion

1. Subnormal serum erythropoietin level

Diagnosis of PV requires meeting either all 3 major criteria, or the first 2 major criteria and the minor criterion[†]

[†]Criterion number 2 (BM biopsy) may not be required in cases with sustained absolute erythrocytosis: hemoglobin levels >18.5 g/dL in men (hematocrit, 55.5%) or >16.5 g/dL in women (hematocrit, 49.5%) if major criterion 3 and the minor criterion are present. However, initial myelofibrosis (present in up to 20% of patients) can only be detected by performing a BM biopsy; this finding may predict a more rapid progression to overt myelofibrosis (post-PV MF).

2.2.2 ET**CLINICAL FEATURES**

Essential Thrombocythemia (ET) is a chronic myeloproliferative disorder characterized by sustained thrombocytosis, megakaryocytic hyperplasia in bone marrow and a trend to present thrombotic or hemorrhagic complications. ET patients may evolve to myelofibrosis and very rarely to acute leukemia. The disease has a female sex predominance and life expectancy is near normal when compared to a sex and age-matched population. The main goal of treatment is to prevent the appearance of thrombotic complications(137,138).

ETIOLOGY/PATHOGENESIS

ET patients are mainly characterized by isolated thrombocytosis discovered incidentally in a routine blood test. From the molecular point of view, the majority of ET patients present genetic mutations resulting in dysregulated signaling pathways that disrupt proliferation and survival of hematological cells. Patients with ET, in contrast to PV, can present driver mutations in three genes: *JAK2*, *CALR*, or *MPL*. Somatic alterations in these genes are found in about 90% of ET patients, remaining 10% without any of these driver mutations what constitutes a group named triple-negative ET. The *JAK2*V617F mutation occurs in approximately 50-60% of ET patients. *CALR* mutations occur in 15-30% of patients, and *MPL* mutations are the less frequent, found in 3-5% of all ET patients(31).

DIAGNOSTIC CRITERIA

ET diagnostic criteria according to the 2017 revision of the WHO classification are shown in Table 5(1). The main change in the new diagnostic criteria is the incorporation of *CALR* mutation into the group of driver mutations. The analysis of *JAK2*, *CALR* and *MPL* mutations is required for ET diagnosis. In patients without any of these driver mutations, known as “triple negative”, the identification of a clonal marker is considered a minor criterion.

Table 5. WHO ET diagnosis criteria. Adapted from Arber *et al.*, *Blood*, 2016.

Major criteria

1. Platelet count $\geq 450 \times 10^9/L$
2. BM biopsy showing proliferation mainly of the megakaryocyte lineage with increased numbers of enlarged, mature megakaryocytes with hyperlobulated nuclei. No significant increase or left shift in neutrophil granulopoiesis or erythropoiesis and very rarely minor (grade 1) increase in reticulin fibers
3. Not meeting WHO criteria for *BCR-ABL1*⁺ CML, PV, PMF, myelodysplastic syndromes, or other myeloid neoplasms
4. Presence of *JAK2*, *CALR*, or *MPL* mutation

Minor criterion

1. Presence of a clonal marker or absence of evidence for reactive thrombocytosis

Diagnosis of ET requires meeting all 4 major criteria or the first 3 major criteria and the minor criterion

2.2.3 PMF

CLINICAL FEATURES

Primary myelofibrosis (PMF) is a myeloproliferative neoplasm characterized by bone marrow fibrosis, proliferation of atypical and dysplastic megakaryocytes, extramedullary hematopoiesis and splenomegaly. The disease has the worst prognosis among the Ph-negative myeloproliferative neoplasms with median survival of 6-8 years. Evolution to acute leukemia appears in approximately 20% of patients and besides disease progression, infection, thromboembolic complications, and portal hypertension are the most frequent causes of death(139). The disease is clinically very heterogeneous, with patients showing a short survival and others a more prolonged clinical course. Roughly 5-25% of ET and PV patients may transform to myelofibrosis, a condition known as post-ET/post-PV myelofibrosis with a better outcome than de novo PMF but with a similar management(140,141). Treatment of PMF and secondary forms of myelofibrosis is focused on the relief of symptoms and improving quality of life.

ETIOLOGY/PATHOGENESIS

PMF is a clonal myeloid proliferation characterized by an increase of fibroblasts, mesenchymal cells, and collagen deposition secondary to cytokine stimulation and increased growth factors apparently shed from clonally expanded megakaryocytes. The abnormal excess of fibers and collagen stem from fibroblasts. Several cytokines have been implicated in the pathogenesis of bone marrow fibrosis, among them, transforming growth factor-beta, platelet derived growth factor, epidermal growth factor and vascular endothelial growth factor are the main cytokines involved.

PMF is classified in two subgroups: “prefibrotic” and “overtly fibrotic” PMF. The currently used methods for MF classification are based in the density and type of BM fibrosis. Prefibrotic PMF and ET have a similar clinical presentation and molecular profile, hence morphologic examination is required to distinguish between the two entities, as prefibrotic PMF has a worse prognosis than ET.

PMF patients present somatic driver mutations in about 90% of cases, being *JAK2* mutations the most frequent alterations (50-60% of PMF), followed by *CALR* (25-35%), and *MPL* (5-10%)(32). Approximately 10% of PMF patients are triple negative, in which no driver mutation is identified, a feature that is associated to worse prognosis (median survival of 3.2 years)(142). Large studies of PMF including molecularly annotated patients found that PMF patients with *CALR* mutations have lower risk of developing thrombocytopenia, anemia or severe leukocytosis(142,143); and a better overall survival rate than *CALR* WT (median survival was 17.7 years in *CALR* mutated, 9.2 years in *JAK2* mutated and 9.1 years in *MPL* mutated)(142).

DIAGNOSTIC CRITERIA

The diagnostic criteria of PMF are based in the 2017 revision of the WHO classification and establish two different subgroups: prefibrotic (Table 6) and overt myelofibrosis (Table 7)(1). The analysis of *JAK2*, *CALR* and *MPL* mutations is required in both subtypes of PMF as the presence of a driver mutation is a major criterion. In triple negative patients, genetic studies using NGS are recommended in order to search for the most frequent mutations found in PMF, who are helpful in determining the clonal nature of the disease.

Table 6. WHO prePMF diagnosis criteria. Adapted from Arber *et al.*, *Blood*, 2016.

Major criteria	
1.	Megakaryocytic proliferation and atypia, without reticulin fibrosis >grade 1, accompanied by increased age-adjusted BM cellularity, granulocytic proliferation, and often decreased erythropoiesis
2.	Not meeting the WHO criteria for <i>BCR-ABL1</i> ⁺ CML, PV, ET, myelodysplastic syndromes, or other myeloid neoplasms
3.	Presence of <i>JAK2</i> , <i>CALR</i> , or <i>MPL</i> mutation or in the absence of these mutations, presence of another clonal marker, [†] or absence of minor reactive BM reticulin fibrosis [‡]
Minor criteria	
1.	Presence of at least 1 of the following, confirmed in 2 consecutive determinations: <ol style="list-style-type: none"> a. Anemia not attributed to a comorbid condition b. Leukocytosis $\geq 11 \times 10^9/L$

- c. Palpable splenomegaly
- d. LDH increased to above upper normal limit of institutional reference range

Diagnosis of prePMF requires meeting all 3 major criteria, and at least 1 minor criterion

† In the absence of any of the 3 major clonal mutations, the search for the most frequent accompanying mutations (eg, *ASXL1*, *EZH2*, *TET2*, *IDH1/IDH2*, *SRSF2*, *SF3B1*) are of help in determining the clonal nature of the disease.

‡ Minor (grade 1) reticulín fibrosis secondary to infection, autoimmune disorder or other chronic inflammatory conditions, hairy cell leukemia or other lymphoid neoplasm, metastatic malignancy, or toxic (chronic) myelopathies.

Table 7. WHO overt PMF diagnosis criteria. Adapted from Arber *et al.*, *Blood*, 2016.

Major criteria

1. Presence of megakaryocytic proliferation and atypia, accompanied by either reticulín and/or collagen fibrosis grades 2 or 3
2. Not meeting WHO criteria for ET, PV, *BCR-ABL1*⁺ CML, myelodysplastic syndromes, or other myeloid neoplasms
3. Presence of *JAK2*, *CALR*, or *MPL* mutation or in the absence of these mutations, presence of another clonal marker,[†] or absence of reactive myelofibrosis[‡]

Minor criteria

1. Presence of at least 1 of the following, confirmed in 2 consecutive determinations:
 - a. Anemia not attributed to a comorbid condition
 - b. Leukocytosis $\geq 11 \times 10^9/L$
 - c. Palpable splenomegaly
 - d. LDH increased to above upper normal limit of institutional reference range
 - e. Leukoerythroblastosis

Diagnosis of overt PMF requires meeting all 3 major criteria, and at least 1 minor criterion

† In the absence of any of the 3 major clonal mutations, the search for the most frequent accompanying mutations (eg, *ASXL1*, *EZH2*, *TET2*, *IDH1/IDH2*, *SRSF2*, *SF3B1*) are of help in determining the clonal nature of the disease.

‡ BM fibrosis secondary to infection, autoimmune disorder, or other chronic inflammatory conditions, hairy cell leukemia or other lymphoid neoplasm, metastatic malignancy, or toxic (chronic) myelopathies.

2.3. Molecular characterization in MPNs

Molecular profiling is essential in MPNs to ensure the correct diagnosis and management of these patients. Genetic alterations are involved in MPN disease pathogenesis, hence gene mutations are classified as driver and non-driver depending on the affected gene. Driver alterations are known to be sufficient to generate a MPN phenotype, and they occur in three genes: *JAK2*, *CALR* and *MPL*, affecting the JAK-STAT pathway(19,144,145). They were thought to be mutually exclusive, however, there is increasing evidence that these mutations can co-occur in the same patient, especially in those with low *JAK2* allele burden(146). It has been recently demonstrated that in these cases, driver mutations are present in different cell clones(147).

Non-driver mutations occur in a high variety of genes, implicated in DNA methylation, histone modification, DNA repair, splicing, transcription factors and cell signaling, among others. Non-driver mutations can influence the phenotype of the disease, induce clonal evolution and cause disease progression.

2.2.1. Driver Mutations

JAK2

The *JAK2V617F* mutation is the most frequently identified mutation in MPNs, and was incorporated as a major diagnosis criterion for the first time in 2008 WHO classification(148). It is observed in approximately 96% of PV patients, and in 50-60% of ET and PMF patients.

Mouse models harboring the *JAK2V617F* mutation develop similar features than PV patients and secondary myelofibrosis(149–153). Of note, other studies proved that if the variant allele frequency (VAF) of *JAK2V617F* mutation was altered, a phenotype similar to ET or PMF is observed(154,155).

- *JAK2* exon 12 mutations

Patients with exon 12 mutations in *JAK2* (mainly PV) present a more erythroid dominant expansion, inducing higher hemoglobin and hematocrit values than *JAK2V617F* patients. No significant differences were found in the incidence of

thrombosis, transformation to MF/AML and overall survival between both groups(156).

- *JAK2* lost of heterozygosity (LoH)

The majority of MPN patients carry *JAK2*V617F mutation in heterozygosity, that is, clonal cells present one allele of the *JAK2* gene affected by the V617F mutation, while the other *JAK2* allele remains unmutated. In approximately 30% of MPN patients, mainly in PV and PMF, loss of heterozygosity (LOH) on the short arm of chromosome 9 affecting *JAK2* (9p24) has been observed. LoH occurs when, by a mechanism of mitotic recombination in a heterozygous cell, the WT copy of the *JAK2* gene is lost and two *JAK2*V617F mutated copies are present in homozygosis in the same cell. *JAK2*V617F mutations in homozygosis usually present a VAF higher than 50% in the granulocytic compartment, due to the loss of *JAK2* WT allele in the pathological clones.

PV patients with *JAK2*V617F in homozygosis have been correlated with increased erythropoiesis and myelopoiesis, lower platelet count, higher rate of splenomegaly and higher proportion of patients requiring cytoreductive therapy(157,158).

Mouse models with low VAF of *JAK2*V617F were associated with thrombocytosis while higher *JAK2*V617F VAFs were associated with erythrocytosis(149,154). This is in line with clinical observations in PV and ET, suggesting that the presence *JAK2*V617F mutation in homozygosis or heterozygosis plays a role in determining whether the patient phenotype is mostly erythroid (PV) or megakaryocytic (ET).

- *JAK2* variant allele frequency monitoring

JAK2 VAF may be modulated during the disease course as a consequence of treatment, clonal evolution or disease transformation. Some patients present a stable *JAK2* VAF during the disease course, while other cases present an expansion or a reduction of the *JAK2* mutated clone. These changes in *JAK2*V617F VAF have clinical implications in PV and ET. Patients with a *JAK2*V617F VAF persistently higher than 50% or with progressively increase in

INTRODUCTION

JAK2V617F VAF during the follow up have a higher risk of transformation to MF(159).

- *JAK2* early acquisition

MPNs are clonal diseases occurring mainly in elderly patients, therefore, since the discovery of *JAK2* mutation, driver mutations in MPNs were thought to occur early before the development of the MPN phenotype. The timing of acquisition of somatic *JAK2* mutations was unclear due to sensitivity limitations in the available technology. A recently published study reconstructed the phylogeny of hematopoiesis in patients with MPN (Figure 8) (160). Analyzing the whole spectrum of somatic mutations, it was possible to estimate that the acquisition of JAK2V617F mutation occurred in utero or during childhood. Moreover, the mean time period between *JAK2*V617F acquisition and MPN clinical presentation was 31 years. These findings are of outmost importance to understand the pathogenesis of MPN and indicate that the clinical diagnosis occurs after decades of clonal expansion of the driver clone. These results provide crucial information for early detection and disease prevention in the future.

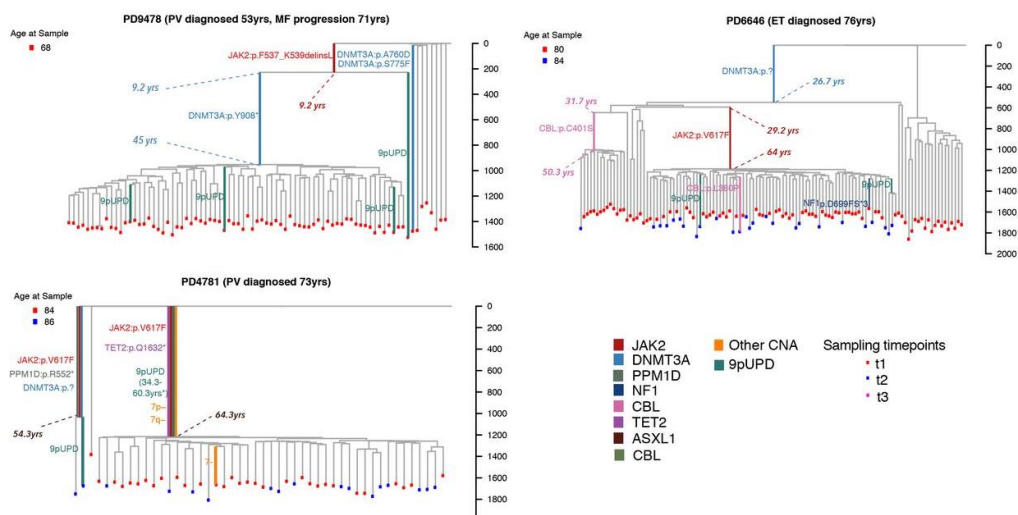


Figure 8. Phylogenetic reconstruction of genetic alterations in MPNs.
(Adapted from Williams *et al*, *BioRxiv*, 2020) (160)

CALR

Since their discovery in 2013, *CALR* mutations have been detected in approximately 15-30% of ET patients and 25-35% in PMF. ET patients with *CALR* mutations showed a lower incidence of thrombosis, lower leukocyte count, lower hemoglobin level and higher platelet count than *JAK2* mutated ET patients. (161,162). Moreover, patients with type 1 *CALR* mutations were associated to a higher risk of transformation to MF when compared to type 2 *CALR*(163).

- *CALR* variant allele frequency monitoring

In contrast to *JAK2*V617F mutations, *CALR* mutations are present in heterozygosis in most cases. The VAF of *CALR* mutations in both ET and PMF is usually of 40-50%, suggesting that the *CALR* mutated clone expands into the whole granulocytic line but in this case LOH is infrequent. Monitoring of *CALR* mutated patients showed that VAF of *CALR* mutations remains stable in time, both in untreated cases and in patients receiving cytoreductive therapy(164,165). Clinical studies with pegylated interferon and ruxolitinib showed that *CALR* VAF decreased in a minority of the cases (13/31 and 3/18, respectively) (164,166).

MPL

MPL mutations are the less frequently detected driver mutations in MPNs, observed in 3-5% of ET patients and 5-10% of PMF patients. These mutations occur mainly in exon 10 of *MPL* gene, being the most frequently identified mutations *MPL* p.W515L, p.W515K, p.W515A and p.S505N. Mutations in exons 4 and 5 of *MPL* gene have been also identified in triple negative ET. In MPNs, *MPL* mutations are acquired somatically, however, *MPL* germline mutations have been identified in two hereditary disorders: hereditary thrombocytosis and amegakaryocytic congenital thrombocytopenia.

In ET patients, *MPL* mutations have been associated to higher risk of myelofibrotic transformation (167,168).

2.2.4. Non-driver mutations

Mutations identified in MPNs in other genes than *JAK2*, *CALR* and *MPL* are known as non-driver mutations. Mutations in more than 50 non-driver genes have been identified in approximately 50% of ET and PV patients and in 80% of PMF patients. These mutations are important in MPN pathogenesis and disease progression. The most frequently mutated genes are further described in 1.2.

Non-driver mutations may co-occur in the same cell clone than driver mutations, but also in a different clone. Of note, it has been observed that in some MPNs cases that evolve to AML the transformed clone does not harbor the MPN driver mutation, suggesting that a different cell clone is expanding and producing the AML phenotype.

Furthermore, the order of mutation acquisition is relevant for the pathogenesis of MPNs. Founder *TET2* mutations induce a limited clone expansion until a *JAK2* mutation occur as a second hit. In these circumstances, *TET2* prevent *JAK2* from inducing erythroid proliferation. In contrast, when *JAK2* mutations are the first hit there is a trend to develop PV and increased risk of thrombosis(169). Mutations in *CALR* and *MPL* frequently occur early, while other mutations such as *TP53* or *NRAS* are later acquired(170).

2.2.5. Risk stratification based on molecular alterations

Larger cohorts of patients have been analyzed in order to to associate this complex molecular landscape with the disease phenotype and prognosis. A recently published study has described a classification of MPNs into 8 groups based on the molecular profile (Figure 9) (170).

Patients with mutations in *TP53* or *del(17p)/del(5q)/-5* are classified as the group with worst prognosis. The second group is defined by the presence of non-driver mutations in 18 genetic aberrations involved in epigenetics, splicing, cell signaling and transcription (*EZH2*, *IDH1*, *IDH2*, *ASXL1*, *PHF6*, *CUX1*, *ZRSR2*, *SRSF2*, *U2AF1*, *KRAS*, *NRAS*, *GNAS*, *CBL*, Chr7/7qLOH, Chr4qLOH, *RUNX1*, *STAG2*, and *BCOR*). These two groups (1 and 2) are associated to worse outcomes regardless the status of the driver genes.

The next four groups (3 to 6) of the classification are defined by the driver mutations. *CALR* mutated cases present the best outcomes followed by *JAK2* mutated in homozygosis, *MPL* mutated and *JAK2* in heterozygosis.

Patients without any of the previously described mutations were classified in two groups (7 and 8): unmutated cases (best prognosis) and cases with other mutations (such as *TET2*, *DNMT3A*).

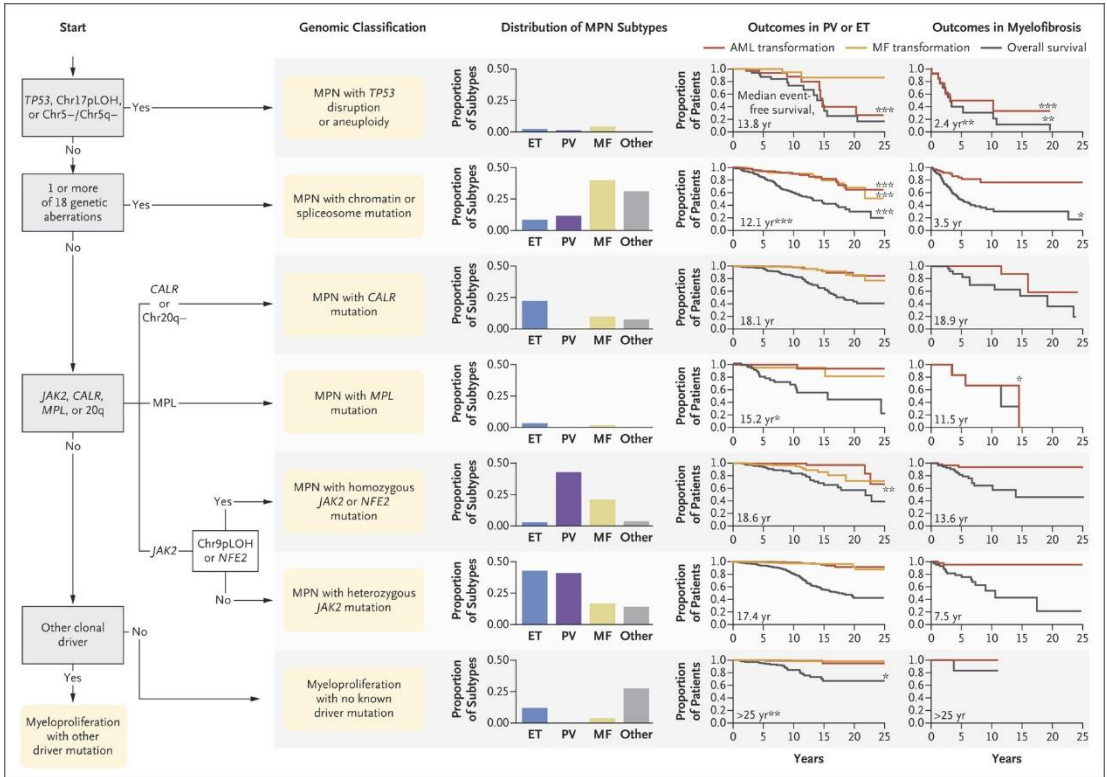


Figure 9. Molecular classification of MPNs (Grinfeld *et al.*, *NEJM*, 2018)(170).

3. MYELODYSPLASTIC SYNDROMES (MDS)

3.1 Introduction

Myelodysplastic syndromes (MDS) are clonal diseases characterized by persistent cytopenia, dysplasia, ineffective hematopoiesis and clonal proliferation of aberrant hematopoietic stem cells, together with higher risk of evolution to AML.

CLINICAL FEATURES

Clinical presentations in MDS are heterogeneous but share cytopenia in one or more cell lineages and morphologic dysplasia in one or more bone marrow myeloid lineages. The main reason for medical consulting in these patients is anemia produced by the ineffective hematopoiesis. Not all MDS patients require treatment at the time of diagnosis, and usually when treatment is required the main objective is to control anemia and improve the quality of life.

ETIOLOGY/PATHOGENESIS

The disease origin of MDS occurs in a hematopoietic stem cell that acquires genomic alterations. These genetic lesions provide selective advantage to the cell clone, which expands in the bone marrow compartment. Several driver genes have been identified in MDS, implicated in different cell activities, which is in line with the clinical and biological heterogeneity observed in these disorders(90).

The majority of *de novo* MDS cases are found in patients older than 70. Although these mutations are thought to occur with the aging process, other factors are also responsible of MDS development. A proportion of MDS patients have genetic predisposition and present germline mutations that may accelerate the acquisition of additional mutations and induce clonal expansion(171). On the other hand, cases of MDS with a prior history of cytotoxic therapy or radiotherapy are classified as therapy related MDS, in

the WHO category of therapy related myeloid neoplasms(1). Exposition to aggressive therapies like chemotherapy (i.e.: alkylating agents, topoisomerase II inhibitors) plays a double role, on the one hand it would allow the selection of pre-existing clones and on the other hand it would allow the appearance of new ones as a consequence of the direct damage of DNA (172).

DIAGNOSTIC CRITERIA

For MDS diagnosis, persistent cytopenia (>6 months) of at least one cell lineage is required (hemoglobin, <10 g/dL; platelet count, <100 × 10⁹/L; or absolute neutrophil count, <1.8 × 10⁹/L). Other common causes of cytopenia should be discarded. Bone marrow aspiration is performed to detect morphologic dysplasia, blast percentage quantification and perform conventional cytogenetics. In addition to cytopenia, at least one of the following criteria should be present for MDS diagnosis: either bone marrow dysplasia (≥ 10%) in one or more myeloid lineages, 5-19% bone marrow blasts and/or presence of MDS-defining cytogenetic alterations. The subtype of MDS is determined depending on the number of dysplastic lineages, BM and PB blast percentage, presence of ring sideroblasts and cytogenetics. MDS subtypes are further explained in 4.2.

Precursor conditions of MDS, observed in cases not meeting the MDS criteria, have been also recently described. As previously mentioned in the first chapter, the presence of a somatic mutation without detectable cytopenia is known as CHIP. Moreover, those cases with CHIP and an unexplained cytopenia but insufficient WHO criteria for MDS diagnosis are known as clonal cytopenia of undetermined significance (CCUS). These two stages, CHIP and CCUS, are thought to be the previous stages of the disease that frequently evolve to MDS (Figure 10).

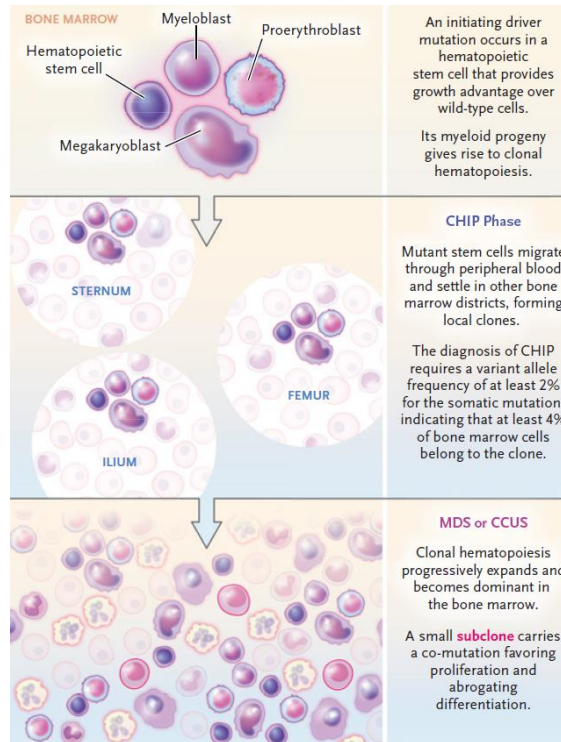


Figure 10. Origin and precursor states of MDS (173)
(Adapted from Cazzola, *NEJM*, 2020)

3.2 Classification

MDS classification and diagnostic criteria were revised in the 2017 WHO classification (Table 8). MDS disease group includes eight different MDS groups according to the number of dysplastic lineages, the presence of ring sideroblasts, the percentage of BM and PB blasts, and cytogenetics. From the molecular point of view, *SF3B1* mutations are relevant in MDS classification. In *SF3B1* wild type patients, a minimum of 15% of ring sideroblasts is required to classify these cases as MDS with ring sideroblasts (MDS-RS), whereas in *SF3B1* mutated patients $\geq 5\%$ of ring sideroblasts is sufficient to establish MDS-RS diagnosis. Cytogenetics are essential to ensure accurate MDS classification, especially the presence of del(5q) which defines a unique MDS subgroup: MDS with isolated del(5q).

INTRODUCTION

Table 8. WHO classification for MDS. Adapted from Arber *et al.*, *Blood*, 2016.

Name	Dysplastic lineages	Cytopenias*	RS (%)	BM/PB blasts	Cytogenetics (karyotype)
MDS with single lineage dysplasia (MDS-SLD)	1	1 or 2	<15%/<5%†	BM <5%, PB <1%, no Auer rods	Any, unless fulfills all criteria for MDS with isolated del(5q)
MDS with multilineage dysplasia (MDS-MLD)	2 or 3	1-3	<15%/<5%†	BM <5%, PB <1%, no Auer rods	Any, unless fulfills all criteria for MDS with isolated del(5q)
MDS with ring sideroblasts (MDS-RS)					
MDS-RS with single lineage dysplasia (MDS-RS-SLD)	1	1 or 2	≥15%/≥5%†	BM <5%, PB <1%, no Auer rods	Any, unless fulfills all criteria for MDS with isolated del(5q)
MDS-RS with multilineage dysplasia (MDS-RS-MLD)	2 or 3	1-3	≥15%/≥5%†	BM <5%, PB <1%, no Auer rods	Any, unless fulfills all criteria for MDS with isolated del(5q)
MDS with isolated del(5q)	1-3	1-2	None or any	BM <5%, PB <1%, no Auer rods	del(5q) alone or with 1 additional abnormality except -7 or del(7q)
MDS with excess blasts (MDS-EB)					
MDS-EB-1	0-3	1-3	None or any	BM 5%-9% or PB 2%-4%, no Auer rods	Any
MDS-EB-2	0-3	1-3	None or any	BM 10%-19% or PB 5%-19% or Auer rods	Any
MDS, unclassifiable (MDS-U)					
with 1% blood blasts	1-3	1-3	None or any	BM <5%, PB = 1%, ‡ no Auer rods	Any
with single lineage dysplasia and pancytopenia	1	3	None or any	BM <5%, PB <1%, no Auer rods	Any
based on defining cytogenetic abnormality	0	1-3	<15%§	BM <5%, PB <1%, no Auer rods	MDS-defining abnormality
Refractory cytopenia of childhood	1-3	1-3	None	BM <5%, PB <2%	Any

*Cytopenias defined as: hemoglobin, <10 g/dL; platelet count, <100 × 10⁹/L; and absolute neutrophil count, <1.8 × 10⁹/L. Rarely, MDS may present with mild anemia or thrombocytopenia above these levels. PB monocytes must be <1 × 10⁹/L

†If *SF3B1* mutation is present.

‡One percent PB blasts must be recorded on at least 2 separate occasions.

§Cases with ≥15% ring sideroblasts by definition have significant erythroid dysplasia, and are classified as MDS-RS-SLD.

3.3 Cytogenetics in MDS

Cytogenetic abnormalities are found in approximately 30-50% of MDS patients. They are essential at diagnosis for accurate classification, prognosis and treatment (174–178). At the time of diagnosis, the presence of a MDS defining cytogenetic alteration (Table 9, except for -Y, del(20q) and +8 which have been also associated to normal aging) together with persistent cytopenia is sufficient to establish MDS diagnosis. In addition, the presence of 5q deletions (del5q) at diagnosis defines a distinct MDS subgroup, MDS with isolated del(5q), when del(5q) is the only cytogenetic alteration or with an additional abnormality except -7 or del(7q)(1). Approximately 15% of MDS presents complex karyotype at the time of diagnosis, defined by the presence of 3 or more cytogenetic alterations (175,176,178).

Table 9. Cytogenetic alterations associated to MDS.

Adapted from Vardiman et al., Blood, 2009(179).

Unbalanced abnormalities	Balanced abnormalities
-7 or del(7q)	t(11;16)(q23;p13.3)
-5 or del(5q)	t(3;21)(q26.2;q22.1)
i(17q) or t(17p)	t(1;3)(p36.3;q21.1)
-13 or del(13q)	t(2;11)(p21;q23)
del(11q)	inv(3)(q21q26.2)
del(12p) or t(12p)	t(6;9)(p23;q34)
del(9q)	
idic(X)(q13)	
+8*	
Del(20q)*	
-Y*	

*these alterations are not sufficient to establish MDS diagnosis, since they are also present in normal aging.
del: deletion, t: translocation, idic: isodicentric, inv: inversion

Cytogenetic alterations have an independent prognostic impact on MDS. In the Revised International Prognostic Scoring System (IPSS-R), the best and widely accepted index score for prognostic stratification of MDS, cytogenetics is the variable with the highest prognostic value(178) (Table 10) (Table 11). Cases with very complex karyotype, defined by the presence of > 3

INTRODUCTION

cytogenetic abnormalities, have been identified as the group with worse prognosis.

Table 10. IPSS-R prognostic subgroups based on cytogenetics.

Adapted from Arber *et al.*, *Blood*, 2016.

Prognostic subgroups, % of patients	Cytogenetic abnormalities	Median survival, * y	Median AML evolution, 25%, * y	Hazard ratios OS/AML*	Hazard ratios OS/AML†
Very good (4%*/3%†)	-Y, del(11q)	5.4	NR	0.7/0.4	0.5/0.5
Good (72%*/66%†)	Normal, del(5q), del(12p), del(20q), double including del(5q)	4.8	9.4	1/1	1/1
Intermediate (13%*/19%†)	del(7q), +8, +19, i(17q), any other single or double independent clones	2.7	2.5	1.5/1.8	1.6/2.2
Poor (4%*/5%†)	-7, inv(3)/t(3q)/del(3q), double including -7/del(7q), complex: 3 abnormalities	1.5	1.7	2.3/2.3	2.6/3.4
Very poor (7%*/7%†)	Complex: > 3 abnormalities	0.7	0.7	3.8/3.6	4.2/4.9

OS indicates overall survival; and NR, not reached.

*Data from patients in this IWG-PM database, multivariate analysis (n = 7012).

†Data from Schanz *et al.* (n = 2754).

Table 11. IPSS-R scores. Adapted from Arber *et al.*, *Blood*, 2016.

Prognostic variable	0	0.5	1	1.5	2	3	4
Cytogenetics	Very good	—	Good	—	Intermediate	Poor	Very poor
BM blast, %	≤ 2	—	> 2% - < 5%	—	5%-10%	> 10%	—
Hemoglobin	≥ 10	—	8 - < 10	< 8	—	—	—
Platelets	≥ 100	50-< 100	< 50	—	—	—	—
ANC	≥ 0.8	< 0.8	—	—	—	—	—

— indicates not applicable.

The most frequent cytogenetic abnormalities in MDS are imbalanced alterations, therefore genomic material is loss or gained, while translocations are infrequent. In this setting, recent studies demonstrated that is possible to detect MDS cytogenetic alterations using NGS. Those genomic regions

affected by a gain of genetic material (for example trisomies) are more represented in the DNA sample than in unaffected regions, while losses of genetic material are less represented in the DNA sample. Therefore, when a gain/loss of genetic information is present in the sample, the read depth by NGS is altered in the amplicons covering that region, which allows its detection (180).

3.4 Molecular characterization in MDS

The application of molecular studies in MDS has been recently incorporated as a diagnostic tool in clinical practice. Indeed, approximately 90 % of MDS patients show at least one mutation in the most frequently affected genes in myeloid malignancies (further described in 1.2)(28,29,90). At diagnosis, mutational profile can be extremely helpful when integrated in the patient clinical context. The presence of ≥ 2 mutations or mutations with $\geq 10\%$ VAF in myeloid associated genes is useful to establish an accurate diagnosis and discriminate MDS from CCUS and CHIP(181).

In contrast to MPNs, there is not a unique driver mutation or mutations for MDS but a more complex genetic landscape has been identified at the time of diagnosis. The average number of mutations per patient at diagnosis is 2.6 (182) and a higher risk WHO subtypes are associated to the presence of a higher number of mutations(29). The most frequently altered genes in MDS are implicated in splicing (*SF3B1*, *SRSF2*, *U2AF1*, *ZRSR2*) and epigenetics (*TET2*, *DNMT3A*, *IDH1/IDH2*, *ASXL1*, *EZH2*) (Figure 11).

Mutations in particular genes are associated to specific disease phenotypes. *SF3B1* mutations are associated with the presence of ring sideroblasts in bone marrow (74,75), in fact, the detection of a *SF3B1* mutation is sufficient to classify as MDS-RS those cases with $\geq 5\%$ of ring sideroblasts(1). Mutations in splicing machinery genes are almost always mutually exclusive in the same patient. In this line, *SF3B1* and *SRSF2* mutations are generally mutually exclusive, suggesting that cells with this co-mutation are unsustainable. On the other hand, the presence of *TP53* mutations is associated to genetic instability and therefore to complex karyotype.

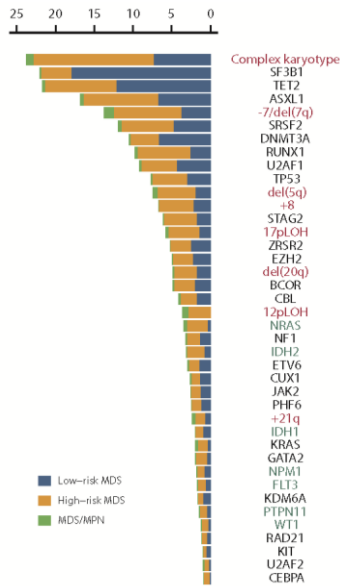


Figure 11. Genetics of MDS. Most frequently mutated genes and cytogenetic alterations in MDS. Adapted from Ogawa *et al.*, *Blood*, 2019(30)

The role of molecular alterations in MDS prognosis is not clearly defined, therefore, mutational profile has not been incorporated to international prognostic scores yet. To date, molecular information has not shown a substantial prognostic prediction improvement when added to traditional prognostic indices, based on clinical, cytological and cytogenetic variables. This is probably because molecular alterations, as founder events, are already reflected in the variables included in the prognostic index (cytopenias, cytogenetic alterations and BM blasts)(183). Recent efforts from the International Working Group for prognosis in MDS have permitted to show preliminary data about the molecular IPSS (M-IPSS)(184) that is expected to show a substantial improvement in the prognostic accuracy of MDS. Studies including large cohorts of MDS found a clear association between the number of mutations and overall survival (Figure 12a)(28,90,185). Mutations in *SF3B1* gene mutations are the only molecular alterations associated with better outcomes (Figure 12b)(74,75) whereas the accumulation of mutations in other genes is generally associated to worse outcomes. *TP53*

mutations/deletions, especially with affectation of both alleles, presented lower overall survival and higher risk of leukemic transformation(112).

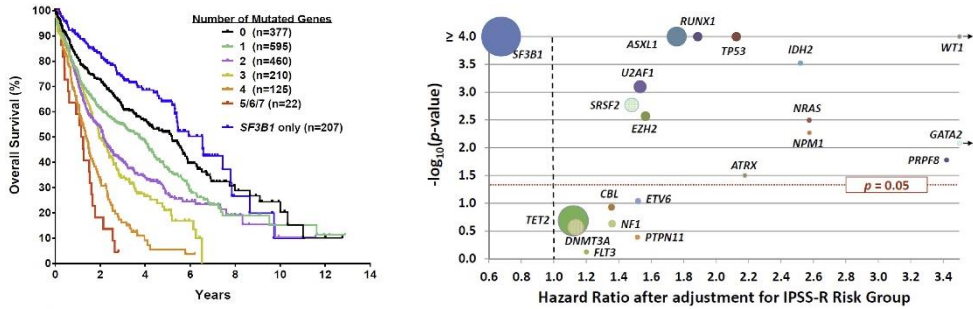


Figure 12. Prognostic implications of gene mutations in MDS. A) Survival analysis of MDS grouped by number of mutations and SF3B1 mutation as a single alteration B) Hazard ratio of death for MDS patients with ≤5% BM blasts, adjusted for IPSS-R risk group. Circle size indicated the frequency of the mutated gene. Genes above the red line show a significant association with prognosis (independent of the IPSS-R). Mutations in significant genes with a hazard ration of greater than 1 are adverse, while SF3B1 is the only prognostically favorable mutated gene. Adapted from Haider *et al.*, *ASCO Educational Book*, 2017(185).

4. CHRONIC MYELOMONOCYTIC LEUKEMIA

4.1 Introduction

Chronic myelomonocytic leukemia (CMML) is a clonal disease affecting myeloid hematopoietic stem cells, characterized by monocytosis in peripheral blood together with bone marrow dysplasia, also showing predisposition to evolve to AML. CMML is a disease subgroup included in the MDS/MPN group of the 2017 WHO myeloid neoplasms classification with shared characteristics of myeloproliferative neoplasms and myelodysplastic syndromes(1).

CLINICAL FEATURES

CMML is mainly characterized by the presence of persistent (>3 months) peripheral blood monocytosis ($\geq 1 \times 10^9/L$) with monocytes constituting $\geq 10\%$ of the white blood cell count. CMML bone marrow presents dysplastic features in most cases; however, it may not be present in a minority of cases. They frequently present cytopenias, mainly anemia and thrombocytopenia, in some cases accompanied by splenomegaly. Approximately 30% of CMML patients present autoimmune diseases such as rheumatoid arthritis or psoriasis at the time of diagnosis or previous to developing the CMML phenotype(186,187).

ETIOLOGY/PATHOGENESIS

Patients with CMML show clonal hematopoiesis with aberrant myeloid differentiation. CMML pathological clone has its origin in HSC, which acquires genomic alterations that provide selective advantage over the others. *TET2* mutations are considered one of these alterations that imply enough proliferation advantage to trigger clonal expansion in CMML(188). Mouse models demonstrated that *TET2* deficient cells are more prone to stimulate the granulocyte and monocyte expansion(189). Moreover, *SRSF2* mutations frequently co-occur with *TET2* mutations, a co-mutation that is identified as the molecular signature of CMML(188,190,191). *TET2* and *SRSF2* mutations in the same clone are thought to cooperate stimulating the expansion of

granulocyte/monocyte progenitors, inducing higher effect than *TET2* alone. Besides, mutations in *ASXL1* gene are also thought to be an initiating event in CMML, both by themselves and co-occurring with *TET2*(54,188) (Figure 13).

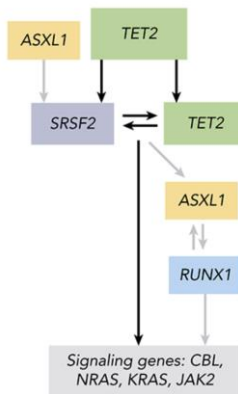


Figure 13. Molecular alterations acquisition in CMML
(Adapted from Palomo *et al.*, *Blood*, 2020) (54)

In a subgroup of CMML, known as therapy related CMML (t-CMML), mutation acquisition has been associated to prior exposure to cytotoxic therapy or ionizing radiation received due to previous neoplastic disorders in those patients. t-CMML represents approximately 10% of CMML and is associated with worse clinical outcomes when compared to de novo CMML(192–194).

DIAGNOSTIC CRITERIA

The CMML diagnostic criteria currently used were published in the 2017 WHO classification (Table 12)(1), although only minimal changes were made when compared to previous classifications. The presence of a molecular abnormality is not essential for CMML diagnosis. Nevertheless, for cases without dysplasia, CMML diagnosis can be established if a clonal molecular or cytogenetic abnormality is present or if monocytosis is persistent (> 3 months) and other causes of monocytosis are excluded. In this context, the presence of mutations in myeloid associated genes can be used to support diagnosis in certain cases, although integrated with all the other diagnostic tools as these mutations can be frequently found in the elderly as CHIP.

Table 12. WHO CMML diagnostic criteria. Adapted from Arber *et al.*, *Blood*, 2016.**CMML diagnostic criteria**

1. **Persistent PB monocytosis** $\geq 1 \times 10^9/L$, with monocytes accounting for $\geq 10\%$ of the WBC count
2. Not meeting WHO criteria for ***BCR-ABL1*⁺** CML, PMF, PV, or ET*
3. No evidence of ***PDGFRA***, ***PDGFRB***, or ***FGFR1*** rearrangement or ***PCM1-JAK2*** (should be specifically excluded in cases with eosinophilia)
4. **<20% blasts** in the blood and BM†
5. **Dysplasia in 1 or more myeloid lineages.**
If myelodysplasia is absent or minimal, the diagnosis of CMML may still be made if the other requirements are met
And
 - a) An acquired clonal cytogenetic or molecular genetic abnormality is present in hemopoietic cells‡
 Or
 - b) The monocytosis (as previously defined) has persisted for at least 3 months and all other causes of monocytosis have been excluded

* Cases of MPN can be associated with monocytosis or they can develop it during the course of the disease. These cases may simulate CMML. In these rare instances, a previous documented history of MPN excludes CMML, whereas the presence of MPN features in the BM and/or of MPN-associated mutations (***JAK2***, ***CALR***, or ***MPL***) tend to support MPN with monocytosis rather than CMML.

† Blasts and blast equivalents include myeloblasts, monoblasts, and promonocytes. Promonocytes are monocytic precursors with abundant light gray or slightly basophilic cytoplasm with a few scattered, fine lilac-colored granules, finely distributed, stippled nuclear chromatin, variably prominent nucleoli, and delicate nuclear folding or creasing. Abnormal monocytes, which can be present both in the PB and BM, are excluded from the blast count.

‡ The presence of mutations in genes often associated with CMML (eg, ***TET2***, ***SRSF2***, ***ASXL1***, ***SETBP1***) in the proper clinical context can be used to support a diagnosis. It should be noted however, that many of these mutations can be age-related or be present in subclones. Therefore, caution would have to be used in the interpretation of these genetic results.

For CMML diagnosis, exclusion of all cases of reactive monocytosis is challenging. Several inflammatory conditions, as well as infections or neoplasms, may lead to a higher monocyte count in PB. Autoimmunity is also a frequent feature in CMML that may produce monocytosis (187,195,196). In this regard, molecular studies are the best strategy to demonstrate clonality and therefore, are recommended when possible. Flow cytometry (FC) analysis of PB monocyte subsets has become also a remarkably useful diagnostic tool(197).

4.2 Classification

CMML has been historically classified as a subtype of MDS until WHO classification in 2001, when CMML was included in a new category for myelodysplastic/myeloproliferative neoplasms (MDS/MPN) due to the presence of both proliferative and dysplastic features(198). The presentation of CMML is heterogeneous, thus CMML is sub-classified in proliferative CMML (p-CMML) and dysplastic CMML (d-CMML), based on a white blood cell count of $\geq 13 \times 10^9/L$ for p-CMML and $< 13 \times 10^9/L$ for d-CMML(1,27). P-CMML frequently present leukocytosis, myeloproliferation and hepatomegaly/splenomegaly; and d-CMML is associated to PB cytopenias, transfusion dependence and recurrent infections(50,199).

On the other hand, CMML has been also classified into three categories according to PB and BM blast percentage: CMML-0 (<2% PB and <5% BM blasts), CMML-1 (2-4% PB and 5-9% BM blasts) and CMML-2 (>5% PB and 10%-19% BM blasts and/or when Auer rods are observed)(200). This classification is based on the prognostic differences observed between these three groups. CMML median survival is reduced when higher percentage of medullary blasts is observed and when presenting as p-CMML(201).

4.2.1 Oligomonocytic chronic myelomonocytic leukemia (OM-CMML)

As previously described, the CMML diagnosis requires the presence of persistent PB absolute ($\geq 1 \times 10^9/L$) and relative monocytosis ($\geq 10\%$ of the leukocytes). These diagnostic thresholds, although are necessary to distinguish between entities, may be arbitrary and lack biological significance given that other diagnostic tools are now available to define and understand CMML.

In 2017, the concept of oligomonocytic chronic myelomonocytic leukemia (OM-CMML) was defined for the first time as those MDS or MDS/MPNs cases with relative monocytosis ($\geq 10\%$ monocytes) and a total monocyte count of 0.5 to $< 1 \times 10^9/L$ (202). These patients with relative monocytosis and mild absolute monocytosis do not reach the mandatory threshold to diagnose

classical CMML ($\geq 1.0 \times 10^9/L$), thus they are currently classified as MDS. It has been recently demonstrated that OM-CMML and overt CMML present similar clinical and genomic characteristics. They frequently present *TET2*, *SRSF2* and *ASXL1* mutations, reinforcing the idea of OM-CMML and overt CMML sharing a similar genetic and biological origin. To further support this idea, the majority of OM-CMML evolve to overt CMML during the follow-up(203).

Several studies support that the OM-CMML classification should be included into clinical practice. The correct distinction between CMML and MDS is clinically relevant and previous investigations defend that OM-CMML would be better classified as a sub-group of CMML(202,203). Nevertheless, further studies are required to confirm this hypothesis as previous research molecular studies were performed in only 24 OM-CMML patients. In addition, no previous information of the immunophenotypic features of OM-CMML was provided(202). Additional studies to better characterize OM-CMML are required to support the consideration of OM-CMML as a distinctive subtype of CMML.

4.3 Molecular characterization in CMML

Molecular studies have emerged as a useful tool to identify clonality in CMML, and may be required for diagnosis in those cases without dysplastic features. Moreover, 80-90% of patients with CMML show mutations in *TET2*, *SRSF2* and/or *ASXL1*(52,191,204,205). Mutations in genes involved in the RAS pathway (i.e. *NRAS*, *KRAS*, *CBL* and less frequently *NF1* and *PTPN11*) are detected in 30% of CMML patients, especially in the p-CMML subtype(52,191). In this context, the analysis of these genes by NGS allows the detection of a clonal marker in the vast majority of CMML patients.

The molecular profile of d-CMML and p-CMML presents some interesting differences that are associated to the dysplastic or proliferative phenotype. Mutations in RAS pathway, *JAK2*, *ASXL1* and *SETBP1* genes are more frequent in p-CMML while mutations in *TET2* and *SF3B1* are more frequent in d-CMML(52).

Molecular alterations have been incorporated in prognostic scores of CMML. In 2016, a CMML-specific prognostic scoring system that incorporates molecular genetic data (CPPS-Mol) was described, where mutations in *RUNX1*, *NRAS*, *SETBP1* and *ASXL1* were independently associated with worse outcomes(52). CMML patients were classified into 4 categories according to genetic score (including cytogenetics and mutations in *RUNX1*, *NRAS*, *SETBP1* and *ASXL1*), red blood cell transfusion dependency, white blood cell count, and BM blasts. These 4 groups presented significant differences in overall survival and cumulative incidence of leukemic evolution. This score represented an evolution of the CMML-specific prognostic scoring system (CPSS), the most widespread and best accepted prognostic stratification model to date(206).

4.4 Flow cytometry in CMML

Peripheral blood flow cytometry has emerged as a useful diagnostic tool to help diagnose CMML (197). Patients with CMML show an increase in the proportion of classical monocytes (CD14+/CD16-) that can be measured by flow cytometry (Figure 14). Monocytes can be divided in three subsets: CD14+/CD16- (classical), CD14+/CD16+ (intermediate) and CD14low/CD16+ (non-classical). In normal conditions and reactive monocytosis, classical monocytes constitute approximately 85% of total monocytes(207). The presence of an increased proportion of classical monocytes, establishing a cut off value of 94%, showed a 95.1% specificity and 91.9% sensitivity for CMML diagnosis (197). Of note, this distribution was independent to the molecular profile of CMML. These observations provide the basis to incorporate this test into regular clinical practice. The analysis of monocyte distribution could be a helpful test for screening and for differentiating between reactive monocytosis and CMML(208).

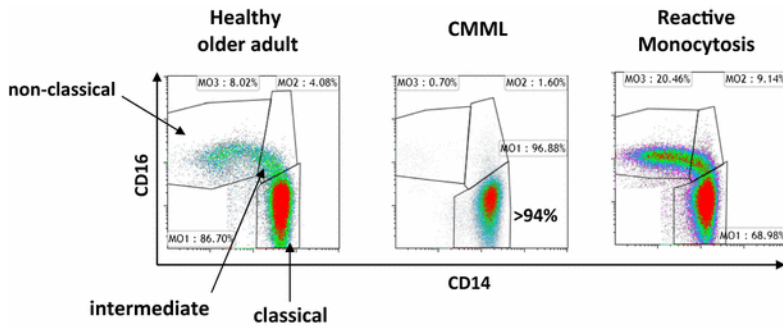


Figure 14. Distribution of PB monocyte populations by flow cytometry in healthy adults, CMML and reactive monocytosis. The presence of >94% of classical monocytes (CD14+/CD16-) is highly specific (95.1%) and sensitive (91.9%) for CMML diagnosis(197,209).

Adapted from Itzykson *et al.*, *IJH*, 2017.

5. LIQUID BIOPSY AND CELL-FREE DNA

Liquid biopsy is known as the analysis of extracellular molecules that harbor genetic information, including cell free DNA, RNA, circulating tumor cells, extracellular vesicles (including exosomes) and tumor educated platelets. It has emerged as an innovative minimally invasive approach that has been implemented in solid and hematologic tumors (Figure 15).

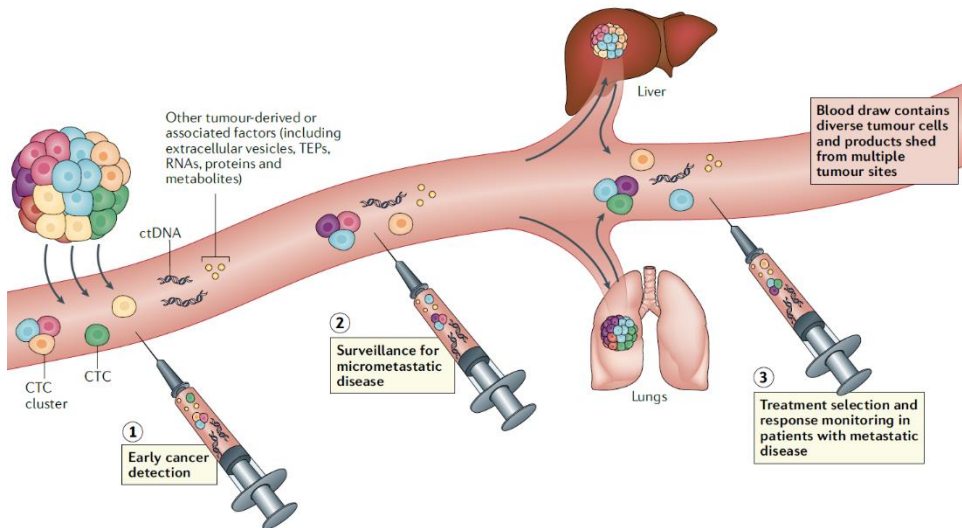


Figure 15. Liquid biopsy applications in clinical context.

Adapted from Ignatiadis et al. *Nat. Rev. Clin. Oncol.*, 2021(210).

Cell free DNA (cfDNA) molecules are double-stranded DNA fragments circulating in peripheral blood plasma, that were described for the first time in 1948(211). These molecules are mainly found in peripheral blood plasma, although they can be present in other body fluids such as saliva, urine or cerebrospinal fluid(212–214). cfDNA is approximately 166 base pairs in length which corresponds to the section of DNA wrapped around one nucleosome. To date, cfDNA is the most extended target for liquid biopsy applications in clinical practice.

In addition to DNA, cell free RNA (cfRNA) molecules are also present in peripheral blood plasma. However, RNA is less stable in blood and therefore

it has been less studied than cfDNA. Circulating cell-free miRNA (cfmiRNAs) are short molecules, more stable than other cfRNAs as they are resistant to RNase activity. Therefore, cfmiRNAs have the potential to be a new biomarker for cancer diagnosis and prognosis(215). Other recently described molecules are circular RNAs (circRNA), which are noncoding RNA molecules with no with no 5' end caps or 3' poly(A) tails and a circular covalently-bonded structure. CircRNA molecules are involved in transcription regulation, splicing and protein-protein and protein-RNA interactions(216). These molecules are promising biomarkers, as they are abundant and highly stable in peripheral blood. In addition, certain chromosomal alterations produce highly specific fusion circRNAs(217). In lung cancer, fusion circRNA can be used as non-invasive biomarkers to detect tumor at early stages and evaluate treatment response(218).

Another target for liquid biopsy studies are circulating tumor cells (CTCs), which are tumoral cells present in blood plasma that extravasate from the tumor into the circulation. These cells are of special interest as they originate the metastatic processes(219). In fact, the number of CTCs in plasma correlates with worst outcomes and metastasis in solid tumors(220,221). However, CTC analysis requires the application of high sensitivity techniques or CTC cell enrichment, since CTCs are present in a very small proportion in comparison with the high background of normal blood cells.

An additional analyte of liquid biopsies are extracellular vesicles, and particularly exosomes. Exosomes are microvesicles containing proteins, DNA, RNA, lipids and other cellular metabolites. Exosomes are implicated in intercellular communication, and in oncology, the analysis of exosome composition has been associated to metastatic processes. These vesicles participate in the preparation of the pre-metastatic niche(222), and can educate cells toward a pro-metastatic phenotype in melanoma(223).

Finally, the presence of tumor educated platelets (TEPs) in plasma has been described. Tumor cells incorporate mutant RNA molecules into platelets to generate these TEPs(224). In cancer patients, the analysis of RNA isolated from TEPs allowed the detection of the tumor genetic alterations. Moreover,

by the location of the primary tumor can be identified by the analysis of the RNA profile in TEPs(225).

The research projects presented in this thesis are focused on the analysis of cell-free DNA (cfDNA) (Figure 16).

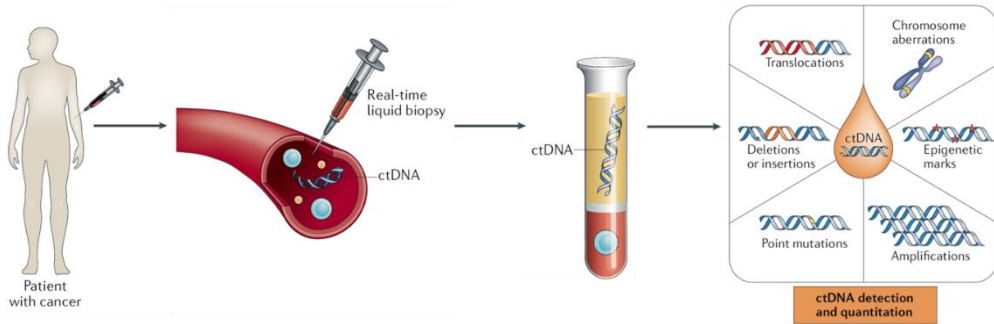


Figure 16. Cell-free DNA obtention and analysis(282).
Adapted from Pantel *et al*, *Nat Rev Clin Oncol*, 2019.

5.1 Cell-free DNA biology and origin

CfDNA is released into the bloodstream mainly through apoptosis or necrosis by both healthy and tumor cells. In cancer patients, the proportion of cfDNA that is released by tumor cells is highly variable, as non-tumoral cells are also contributing to the total cfDNA present in plasma. Large studies described that mutational VAF was highly variable (0.03%–97.6%) and disease dependent(226). The term circulating tumor DNA (ctDNA) is also frequently used and refers specifically to the cell-free DNA released exclusively by tumor cells.

In this context, several studies tried to identify the origin of non-tumoral cfDNA. In 2002, cfDNA of 22 patients was analyzed after receiving a sex-mismatched BM transplantation(227). Male patients receiving BM from female donors presented a low percentage of Y-chromosome DNA in cfDNA (6.9%); while female patients receiving BM from male donors showed a higher percentage (59.5%). These results indicated for the first time that the origin of cfDNA was predominantly hematological.

Posterior studies confirmed that cfDNA has its origin mainly in hematological cells. Since each tissue presents a distinct nucleosome disposition, Snyder *et al.* analyzed the nucleosome footprint in cfDNA and found that hematopoietic lineages were the main contributors to cfDNA in healthy individuals (Figure 17)(228). A similar conclusion was reached by Ulz *et al.*, who found that the cfDNA read depth patterns by NGS reflected the expression signature of hematopoietic cells in healthy donors(229). These two studies also demonstrated that cfDNA fragmentation is not arbitrary and therefore the genome is not equally represented in cfDNA. Similar results were later obtained when analyzing the methylation patterns of cfDNA(230). To further explain these findings, Sadeh *et al.* recently published their results about chromatin immunoprecipitation sequencing (CHIP-seq) analysis of cfDNA. They found that bone marrow megakaryocytes were the main contributors of cfDNA in healthy individuals(231).

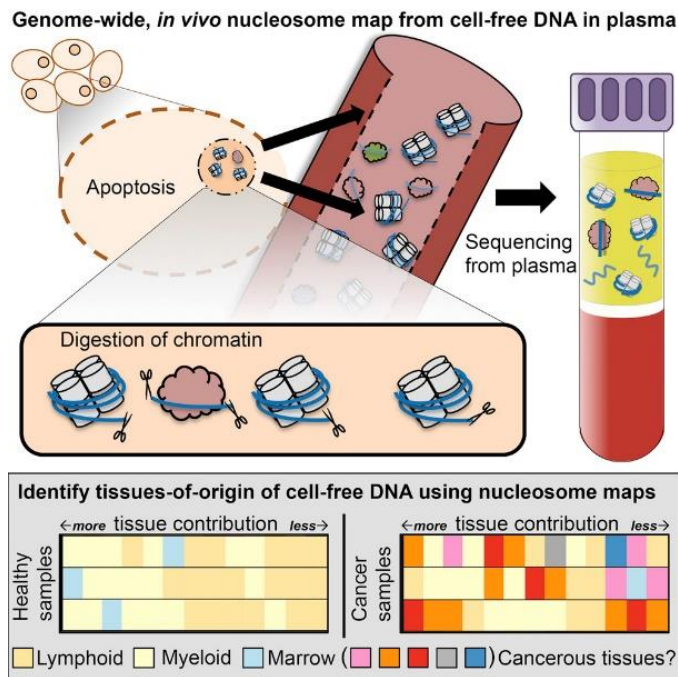


Figure 17. Nucleosome footprint in cfDNA. CfDNA was obtained from plasma and sequenced. Nucleosome distribution in cfDNA was matched to the nucleosome footprint database to determine the tissue-of origin in healthy and cancer samples(228).

Adapted from Snyder *et al.*, *Cell*, 2016

5.2 Applications of cell-free DNA analysis

In the last decade, liquid biopsy has emerged as an innovative approach for molecular characterization that has been implemented in solid and hematologic tumors. The analysis of peripheral blood plasma is a less invasive procedure for molecular profiling than tumor biopsies in solid tumors or lymphomas. In leukemias, cfDNA isolation is also a more accessible alternative than BM biopsy or aspirate to perform genetic analysis. Moreover, liquid biopsy applications are useful at diagnosis, to monitor genetic lesions and to evaluate treatment response.

5.2.1 cfDNA in myeloid neoplasms

In myeloid neoplasms, most of the liquid biopsy research has been focused on myelodysplastic syndromes and acute myeloid leukemia(232–236). Focusing on the most recent research, in 2016 cfDNA from a short series of 16 MDS patients was studied(237). CfDNA analysis showed higher detection rate and higher VAF than PB cells analysis. Later in 2017, 12 patients with MDS were monitored in different time points using cfDNA, and presented similar mutations and clonal dynamics than BM samples(238) (Figure 18). Additionally, in one case karyotype alterations were also detectable in cfDNA. In line with this results, a posterior research demonstrated that cfDNA alterations had a prognostic impact in 14 MDS and 37 AML cases receiving allogeneic hematopoietic stem cell transplantation(239) (Figure 19). Another posterior study in AML confirmed that cfDNA allowed reliable molecular characterization, however, discordant mutations with VAFs lower than 10% were found in both BM and cfDNA, suggesting that subclonal populations may be missed(240).

In contrast, there is scarce literature of cfDNA analysis in MPNs. Several groups demonstrated that *JAK2* mutations were detectable in plasma or serum DNA and presented similar VAF to PB cells, but other driver and non-driver mutations were not analyzed (241–244).

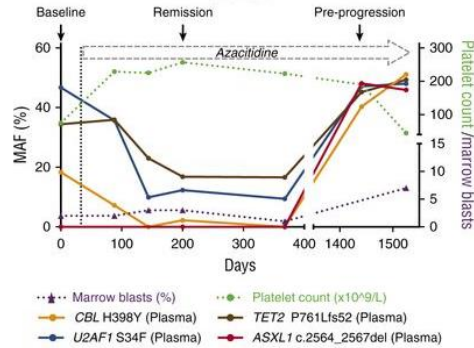


Figure 18. Mutational dynamics in cfDNA of a patient receiving azacitidine. Mutant allele frequency (MAF), platelet count, percentage of BM blasts and remission/progression time points are also shown(238). Adapted from Yeh *et al.*, *Blood*, 2017

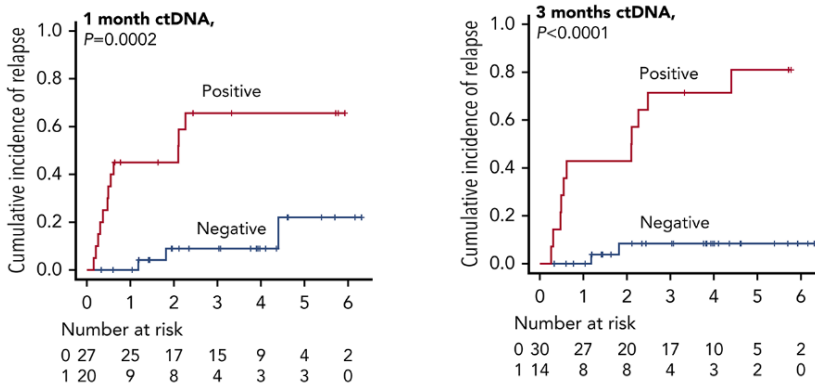


Figure 19. Cumulative incidence of relapse based on the presence of mutated circulating tumor DNA (ctDNA) after allogeneic hematopoietic stem cell transplantation in MDS/AML. Similar results were obtained at 1 month and 3 months after transplantation(239). Adapted from Nakamura *et al.*, *Blood*, 2019.

In respect to precursor states of myeloid malignancies, it has been recently demonstrated that clonal hematopoiesis in healthy individuals can be detected by the analysis of cfDNA(245). Interestingly, they found that the concordance between cfDNA and PB cells was overall very high, however, for variants with low VAF (<10%) there was poor concordance. These findings in cfDNA in CHIP are in consonance with the presence of clonal hematopoiesis that has been reported in patients with solid tumors(246,247).

5.3 Limitations of cell-free DNA analysis

In solid neoplasms and lymphomas, one of the main limitations of liquid biopsy is that tumoral cfDNA represents a small fraction of total cfDNA(226). Therefore, high sensitivity techniques are required to detect these alterations. This makes the application of liquid biopsy analysis more complex, and may lead to false negative results, especially for the identification of poorly represented mutations. In addition, the presence of mutations derived from clonal hematopoiesis (further explained in 1.4) should be considered, since hematopoietic cells are the main contributors of cfDNA(247). Other key technical limitation of liquid biopsy studies is that the obtention of pure cfDNA is dependent of pre-analytical and analytical steps. When the PB sample is obtained, plasma should be isolated within the first four hours if conventional EDTAK3 tubes are used for collection, since the apoptosis of PB cells dilutes the cfDNA with genomic DNA. Otherwise stabilizing tubes should be used. Nevertheless, the pre-analytical handling of samples should be standardized to ensure the high quality of cfDNA samples(248). For gene expression analysis, RNA from tumoral cells is currently a better option than cfRNA due its low concentration in plasma samples.

In comparison with traditional tissue biopsies, liquid biopsies are a less invasive method, allow more frequent samplings and contain a homogeneous profile of the tumoral genetic alterations. However, liquid biopsy studies are limited to genomic alterations, therefore tumor biopsies are still required to evaluate tumor morphology. Most cancer types require an initial histologic sample for diagnosis. Moreover, some mechanisms of resistance to treatment show morphology progression(249) that is not reflected in cfDNA. To date, liquid biopsies in solid tumors may be more helpful as a complementary tool to confirm malignancy, to perform molecular analysis if tissue is insufficient and for the follow up of previously identified molecular alterations.

HYPOTHESIS AND OBJECTIVES

The identification of somatic mutations provided a better understanding of the pathogenesis in myeloid neoplasms, but subsequently revealed that multiple aspects should be further improved in the diagnostic and classification tools that are used in clinical practice.

The first hypothesis of this doctoral thesis is that the analysis of cell-free DNA, mainly originated in the hematopoietic cells of the bone marrow, could be a useful strategy to evaluate the mutational profile in myeloid neoplasms. In myeloproliferative neoplasms, the analysis of cell-free DNA could reflect better the mutational profile of hematopoietic tumor cells than the study of mature populations. In myelodysplastic syndromes, cell-free DNA analysis could overcome the limitations inherent to the bone marrow analysis and accompanying cytopenias.

The second hypothesis of this doctoral thesis is that patients with OM-CMML, currently classified as MDS according to the 2017 WHO classification, could be better defined as a subtype of CMML. Moreover, the integration of clinical, genomic and immunophenotypic information could support the consideration of OM-CMML as a distinct subtype of CMML.

GENERAL OBJECTIVE

The main goal of the research projects here presented is to improve the accuracy of the diagnosis and classification of myeloid neoplasms.

SPECIFIC OBJECTIVES

1. To explore the accuracy and reliability of cfDNA analysis as a new non-invasive diagnostic tool in a cohort of MPN patients.
2. To explore the role of cfDNA analysis as a non-invasive diagnostic tool for the detection and monitoring of molecular and cytogenetic alterations in a cohort of MDS patients.
3. To compare the molecular, immunophenotypic and clinical profile of a well-annotated series of OM-CMML and CMML.
4. To ascertain if DNA obtained from saliva samples and CD3+ lymphocytes from peripheral blood is a suitable source of germline DNA for molecular studies in MPN patients.

RESULTS

Article 1:

Circulating cell-free DNA improves the molecular characterisation of Ph-negative myeloproliferative neoplasms


Garcia-Gisbert N, Fernández-Ibarrondo L, Fernández-Rodríguez C, Gibert J,
Andrade-Campos M, Arenillas L, Camacho, L, Angona A, Longarón R, Salar A,
Calvo X, Besses C*, Bellosillo, B*

*equally contributed

British Journal of Haematology, 2021; 192(2):300–9

DOI: 10.1111/bjh.17087

Circulating cell-free DNA improves the molecular characterisation of Ph-negative myeloproliferative neoplasms

Nieves Garcia-Gisbert,^{1,2}  Lierni Fernández-Ibarrondo,^{1,2}  Concepcion Fernández-Rodríguez,^{1,3} Joan Gibert,¹  Marcio Andrade-Campos,^{1,4}  Leonor Arenillas,³  Laura Camacho,^{1,3} Anna Angona,^{1,4} Raquel Longarón,^{1,3} Antonio Salar,^{1,4}  Xavier Calvo,³  Carlos Beses^{1,4,*}  and Beatriz Bellosillo^{1,2,3,*} 

¹Group of Applied Clinical Research in Haematology, Cancer Research Program-IMIM (Hospital del Mar Medical Research Institute, ²Pompeu Fabra University, ³Department of Pathology, Hospital del Mar-IMIM, and ⁴Department of Haematology, Hospital del Mar-IMIM, Barcelona, Spain

Received 18 June 2020; accepted for publication 13 August 2020

Correspondence: Beatriz Bellosillo, Group of Applied Clinical Research in Haematology, Cancer Research Program-IMIM (Hospital del Mar Medical Research Institute) Barcelona, Spain.

E-mail: bbelosillo@parcdesalutmar.cat

*These authors contributed equally.

Background

Philadelphia-negative myeloproliferative neoplasms (MPNs) are a heterogeneous group of clonal haematopoietic disorders characterised by clonal proliferation of mature myeloid cells. In MPN patients, genomic alterations are implicated in the pathogenesis of the disease and molecular profiling is mandatory to establish the correct diagnosis and to optimise their management. Currently, molecular analysis of MPN patients is conventionally performed in DNA from isolated granulocytes, whole blood samples and, less frequently, in bone marrow specimens.

In recent years, it has been demonstrated that it is possible to characterise the molecular profile of some solid tumours and haematological neoplasms by analysing the circulating

Summary

Genetic studies in patients with Philadelphia-negative myeloproliferative neoplasms (MPNs) are essential to establish the correct diagnosis and to optimise their management. Recently, it has been demonstrated that it is possible to detect molecular alterations analysing cell-free DNA (cfDNA) in plasma samples, which is known as liquid biopsy. We have assessed the molecular profile of a cohort of 107 MPN patients [33 polycythaemia vera (PV), 56 essential thrombocythaemia (ET), 14 primary myelofibrosis (PMF) and 4 unclassifiable MPN] by next-generation sequencing (NGS) using cfDNA and paired granulocyte DNA. A high concentration of cfDNA in plasma was observed in patients with high molecular complexity, in *MPL*-mutated cases, and in PMF patients. Targeted sequencing of cfDNA showed a comparable mutational profile (100% accuracy) to the one obtained in granulocytic DNA and a strong correlation was observed between the variant allele frequency (VAF) of gene mutations in both DNA sources. The median VAF detected in cfDNA (29.0%; range: 0.95–91.73%) was significantly higher than the VAF detected in granulocytes (median 25.2%; range: 0.10–95.5%), especially for *MPL* mutations (44.3% vs. 22.5%). In conclusion, our data support the use of cfDNA as a fast, sensitive and accurate strategy for performing molecular characterisation of MPN patients.

cell-free (cf)-DNA present in plasma samples, which is also known as 'liquid biopsy'. This cfDNA consists of short DNA fragments derived from both cancer cells [which is known as circulating tumoural DNA (ctDNA)] and non-tumoural cells.^{1,2} The field of ctDNA is rapidly evolving, with the expectation of liquid biopsy either complementing or even replacing invasive tissue biopsies in the near future.

It has been recently reported that the majority of the cfDNA has its origin in immature haematopoietic and bone marrow cells.^{3–6} Hence, some studies have demonstrated that it is possible to characterise the molecular profile of primary bone marrow diseases such as myelodysplastic syndromes by analysing cfDNA.^{7–9} In this context, the analysis of cfDNA in MPNs could reveal the molecular profile of early haematopoietic tumour cells. This approach has not been tested so far, as

N. Garcia-Gisbert *et al.*

mutational analysis in MPNs is routinely performed in the mature cell population circulating in peripheral blood.

The aim of the study was to assess the accuracy and reliability of cfDNA analysis in a cohort of MPN patients, comparing this technique with the genotype of peripheral blood granulocytes.

Patients and methods

Patients

Peripheral blood samples from 107 patients with MPNs were prospectively collected for this study at the time of diagnosis: 33 polycythaemia vera (PV), 56 essential thrombocythaemia (ET), 14 primary myelofibrosis (PMF) and 4 unclassifiable MPNs (uMPNs). Clinico-biological characteristics of the patients are shown in Table I. We additionally obtained peripheral blood samples from 33 healthy controls. The study was approved by the local ethics committee.

Peripheral blood processing and DNA isolation

Peripheral blood samples were collected in K3EDTA tubes and processed in the first 4 h to isolate plasma and granulocytes (Figure S1). Granulocyte DNA was extracted with Genovision M48 (Qiagen, Hilden, Germany). CfDNA was isolated with MagMAX Cell-Free DNA Isolation Kit (Thermo Fisher Scientific, Foster City, CA, USA) and quantified (Qubit, Thermo Fisher Scientific). Purity of the cfDNA was assessed by electrophoresis (4200 TapeStation system, Agilent, Santa Clara, CA, USA) to exclude the presence of genomic DNA. All the cfDNA samples analysed in this project were free of genomic DNA contamination.

Next-generation sequencing

Libraries were prepared using a custom panel covering the whole coding region of 25 myeloid-associated genes (QIAseq Custom DNA Panels, Qiagen, Hilden, Germany): *ASXL1*, *CALR*, *CBL*, *CSF3R*, *DNMT3A*, *ETV6*, *EZH2*, *IDH1*, *IDH2*, *JAK2*, *KIT*, *KRAS*, *MPL*, *NRAS*, *PRPF8*, *RUNX1*, *SETBP1*, *SF3B1*, *SH2B3*, *SRSF2*, *STAG2*, *TET2*, *TP53*, *U2AF1*, *ZRSR2*, that incorporates unique molecular identifiers to tag individual DNA molecules. Libraries were sequenced with a 3000× minimum read depth in MiSeq/NextSeq (Illumina, San Diego, CA, USA).

The GeneGlobe Data Analysis Center (Qiagen) was used for FASTQ trimming, alignment and variant calling (smCounter2, Qiagen). Variants were annotated and classified using Illumina VariantStudio 3.0 software and confirmed with Integrative Genomics Viewer (IGV) v2.4 software. Only pathogenic variants with a variant allele frequency (VAF) >2% were considered.

Digital PCR

Digital PCR analysis for point mutations was performed in duplicate using the QuantStudio 3D Digital PCR System (Applied Biosystems, Foster City, CA, USA).

Table I. Clinico-biological features of patients included in the study.

	PV n = 33	ET n = 56	PMF n = 14	uMPN n = 4	P-value		
					PV vs. TE	PV vs. MF	TE vs. MF
Age, median (range), years	71 (37–91)	63 (25–92)	62 (36–95)	78 (71–82)	0.088	0.217	0.866
Gender, male, n (%)	21 (63.3)	19 (33.9)	10 (71.4)	3 (75)	0.012	0.858	0.024
Haemoglobin, median (range), g/l	171 (126–226)	146 (102–163)	124 (84–158)	155 (145–157)	<0.001	<0.001	0.001
Leukocyte count, median (range), ×10 ⁹ /l	10.5 (5.9–21.9)	8.8 (5.02–17.3)	11.2 (3.8–22.7)	10.1 (6.4–15.8)	0.004	0.701	0.037
Platelet count, median (range), ×10 ⁹ /l	580 (214–1198)	678 (466–1800)	434 (71–1209)	632 (435–755)	0.004	0.255	0.003
Driver gene							
<i>JAK2</i> , n (%)	33 (100)	39 (69.6)	3 (21.4)	4 (100)	0.001	<0.001	0.017
<i>CALR</i> , n (%)	0	9 (16.1)	6 (42.9)	0	–	–	0.062
<i>MPL</i> , n (%)	0	2 (3.6)	3 (21.4)	0	–	–	0.051
TN, n (%)	0	6 (10.7)	2 (14.3)	0	–	–	0.656
Number of mutations, median (range)	2.0 (1–6)	1.0 (1–4)	2.5 (0–7)	2.0 (1–5)	0.008	0.404	0.008
Amount cfDNA, median (range), ng/plasma ml	17.4 (4.6–70.0)	14.3 (3.4–292)	73.1 (3.9–594)	25.8 (11.7–39.5)	0.740	<0.001	<0.001

PV, polycythaemia vera; ET, essential thrombocythaemia; PMF, primary myelofibrosis; uMPN, unclassifiable myeloproliferative neoplasm; TN, triple negative; cfDNA, cell-free DNA. Significant P-values < 0.05 are marked in bold.

Statistical analysis

IBM SPSS Statistics software was used for statistical analysis. For categorical data, comparisons of proportions were evaluated by the chi-square test or Fisher's exact test as appropriate. For continuous variables, comparisons were assessed by the non-parametric Mann–Whitney or Wilcoxon test when appropriate. We assessed Spearman's rank correlation coefficient to evaluate the strength of association between two variables. P -values < 0.05 were considered statistically significant. Coverage metrics were obtained using the DeCovA library.¹⁰ Variant analysis was performed in R version 3.6.2 (R Foundation for Statistical Computing, Vienna, Austria; <https://cran.r-project.org>) using the maftools package.¹¹ The code used in R 3.6.2 to create the figures is displayed in the Supplemental methods.

Results

The amount of cell-free DNA in plasma varies among MPN disease phenotypes

Circulating cfDNA was isolated from plasma samples of 107 MPN patients and 33 healthy controls. The amount of cfDNA obtained per ml of plasma was significantly higher in patients with PMF (median 73.0 ng/ml, range: 3.95–594) than in PV (median 17.4 ng/ml, range: 4.62–69.96) or ET patients (median 14.3 ng/ml, range: 3.44–292.60) ($P < 0.001$). In addition, a significantly lower concentration was obtained in healthy controls (median 5.16 ng/ml, range: 2.0–11.76), when compared to MPN patients ($P < 0.001$ in all comparisons) (Table 1, Fig 1).

We analysed the correlation of the amount of cfDNA with clinical and biological characteristics. No association was observed with haemoglobin, haematocrit or platelets. However, a significant positive correlation was observed between the amount of cfDNA and the leukocyte count ($P = 0.023$, $r = 0.220$). To assess whether the high cfDNA concentration observed in PMF patients was directly related to the peripheral blood leukocyte count, the cfDNA concentration values were normalised by the leukocyte count. The adjusted values (ratio: cfDNA amount per ml of plasma/leukocyte count) were still significantly higher in patients with PMF than in PV ($P < 0.001$) and ET patients ($P < 0.001$). A positive correlation was also observed between the amount of cfDNA and the serum lactate dehydrogenase (LDH) levels ($P < 0.001$, $r = 0.532$). This correlation was still observed when excluding PMF patients from the analysis ($P < 0.001$, $r = 0.442$) and among the group of PMF patients ($P = 0.008$, $r = 0.679$). Finally, a higher concentration of cfDNA was observed in patients who suffered a thrombotic event at the time of diagnosis or during follow-up ($n = 10$; median 37.0 ng/ml of plasma, range: 3.44–292.6) than in the group of patients without thrombotic events (17.0 ng/ml, range: 3.61–594) ($P = 0.038$) (median follow-up: 15 months,

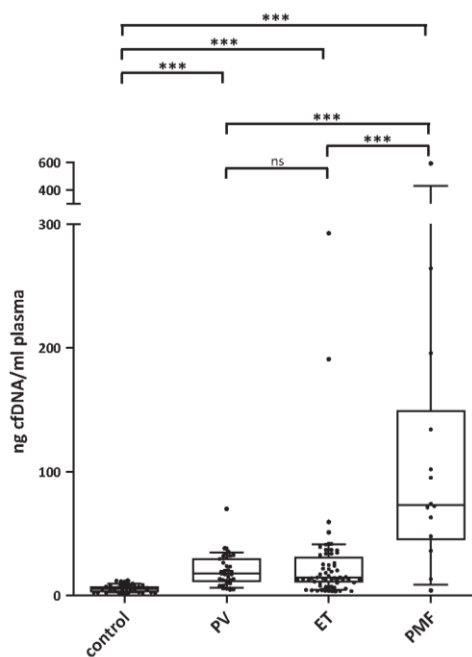


Fig 1. Amount of cfDNA obtained according to disease phenotype. Total cfDNA quantity (ng/ml plasma) obtained for each patient ($n = 107$) grouped by disease phenotype. For healthy controls ($n = 33$) the cfDNA quantity per ml was significantly lower than the PV, ET and PMF group. The PMF group presented a higher quantity of total cfDNA than PV and ET ($P < 0.001$, Mann–Whitney). *** $P < 0.001$, ns: non-significant.

range: 1–60). However, this difference did not reach statistical significance when analysing the disease phenotypes independently.

cfDNA and granulocyte DNA show an equivalent mutational profile

We next performed mutational profiling of granulocyte DNA in the 107 patients by next-generation sequencing (NGS). We observed a median of 2.0 mutations/patient in PV (range: 1–6 mutations), 1.0 mutation/patient in ET (range: 0–4), 2.5 mutations/patient in PMF (range: 0–7) and 2.0 mutations/patient in unclassifiable MPNs (range: 1–5) (Table 1). Eight cases (six ET and two PMF) were classified as triple negative (TN) as no mutations were identified in any of the driver genes; however, mutations in non-driver genes were observed in one ET and one PMF patient (Fig 2A). The most frequently mutated non-driver genes in the whole cohort were *TET2* (20/107, 18.7%), *ASXL1* (16/107, 15.0%), *DNMT3A* (11/107, 10.3%), *SRSF2* (7/107,

N. Garcia-Gisbert *et al.*

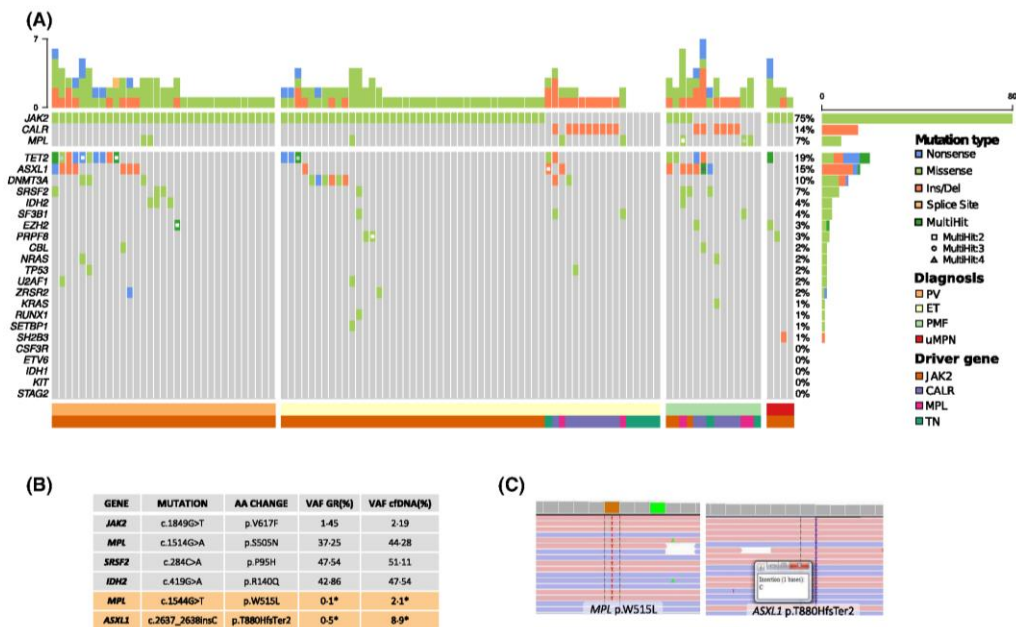


Fig 2. (A) Distribution of mutations identified in the 107 MPN patients. Results of the sequencing of the 25 genes are shown in the plot where each column represents a patient and each row represents a gene. The number of mutations identified per patient is represented as columns in the top row. Genes are ordered from the most to the least frequently mutated, and frequencies for each gene are displayed at the right, as well as the mutation type (nonsense, missense, insertion/deletion, splice site or multihit). Patients with more than one mutation in the same gene are represented as shown in the legend (two, three or four mutations in the same gene). The phenotype (PV, ET, PMF, uMPNs) and the driver gene (JAK2, CALR, MPL, TN) for each patient are depicted in the bottom rows. (B) Mutations detected in a PMF patient by NGS in granulocyte DNA (GR) and cfDNA. *Mutations confirmed by digital PCR. (C) Images show the mutations detected in MPL and ASXL1 genes by NGS in cfDNA visualised using the IGV software. Ins: insertion, Del: deletion, PV: polycythemia vera, ET: essential thrombocythemia, PMF: primary myelofibrosis, uMPN: unclassifiable myeloproliferative neoplasm, TN: triple negative. [Colour figure can be viewed at wileyonlinelibrary.com]

6.5%), IDH2 (4/107, 3.7%) and SF3B1 (4/107, 3.7%). A detailed list of the variants detected is provided in Table S1.

Targeted sequencing of cfDNA showed, overall, a comparable mutational profile to the one obtained in granulocytic DNA. All mutations detected in the granulocytic fraction, considered as the gold standard, were also detected in the paired cfDNA sample, thus resulting in 100% accuracy. Interestingly, in one PMF patient, two previously undetected mutations in MPL and ASXL1 were detected in cfDNA. We confirmed these mutations in cfDNA with digital PCR as an orthogonal method, and its high sensitivity allowed us to detect these mutations also in the granulocyte fraction with a VAF below the NGS limit of detection (Fig 2B).

The median VAF for the gene mutations detected in cfDNA was 29.0% (range: 0.95–91.73%), which was significantly higher than the granulocyte VAFs (median 25.2%; range: 0.10–95.5%) ($P < 0.001$). A strong correlation was observed between the VAFs of granulocytic DNA and cfDNA ($r = 0.897$, $P < 0.001$) (Fig 3).

JAK2 and MPL mutations present a higher VAF in cfDNA than in granulocytes

Most JAK2V617F cases (58/80) presented a higher VAF in cfDNA (median 37.4%; range: 2.19–91.7) than in granulocytes (median 32.2%; range: 1.45–95.5) ($P < 0.001$). This higher allele frequency in cfDNA was mainly observed in ET patients (Fig S2). Interestingly, all MPL mutations presented higher VAF in cfDNA (median 44.3%; range: 2.10–69.4) than in granulocytes (median 22.5%; range: 0.10–38.0; $P = 0.003$) (Fig 4A, Fig S2). These VAFs were confirmed by digital PCR in both granulocytic DNA and cfDNA extracted from an extra plasma tube from the same patient and time point. No significant differences were observed for CALR mutations. Regarding non-driver genes, a slightly higher VAF of SRSF2 mutations was observed in cfDNA (median 47.63; range: 5.58–51.1) than in granulocytic DNA (median 44.14; range: 7.39–47.91, $P = 0.043$). No significant differences were observed for any other non-driver gene. Overall, a higher VAF was observed in cfDNA compared to

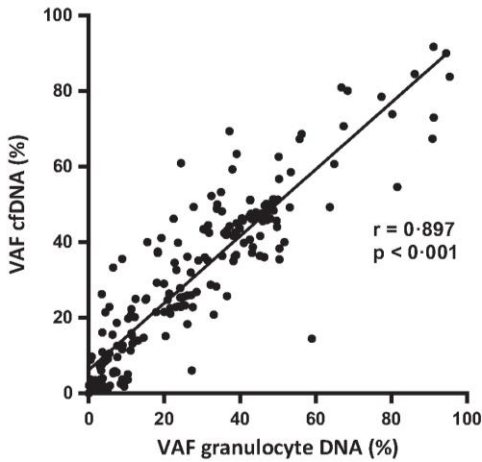


Fig 3. Scatter plot of the 202 variants detected showing the correlation between the variant allele frequency (VAF) of the mutations in cfDNA and DNA from granulocytes.

granulocyte DNA for *MPL* (median 66% increase), *JAK2* (median 20% increase) and *SRSF2* (median 6% increase) mutations.

We next assessed the representation of the myeloid genes in cfDNA and in granulocyte DNA, as it has been reported that nucleosome position affects DNA fragmentation and some genomic regions are overrepresented, hence the cell genome might not be equally represented in cfDNA.⁵ We compared the depth of coverage obtained for the independent gene regions, and we observed a higher representation of *MPL* fragments in cfDNA libraries than in granulocyte DNA libraries (Fig 4B, Fig S3), which suggests that *MPL* regions are better represented in the cfDNA from MPNs patients.

Total cfDNA correlates with genomic complexity

Regarding the amount of cfDNA obtained per ml of plasma, we observed a positive correlation between the number of mutations identified per patient and the amount of plasma cfDNA ($r = 0.242$, $P = 0.012$). In our cohort, patients with more mutations (considering both driver and non-driver genes) presented a higher concentration of cfDNA/ml (Fig 5). When analysing only PMF patients, who have a higher incidence of mutations and a higher concentration of cfDNA, this correlation was maintained ($r = 0.572$, $P = 0.033$).

In line with the previous observations for gene mutation VAFs, *MPL*-mutated patients presented a higher quantity of cfDNA (median 36.75 ng/plasma ml; range: 12.9–264) than the *MPL*-wild-type MPNs (17.0 ng/plasma ml; range: 3.44–594; $P = 0.034$), supporting a higher shedding of cfDNA in *MPL*-mutated cases.

For non-driver genes, patients with *ASXL1* mutations presented a larger amount of cfDNA (median 35.0 ng/ml; range: 4.62–594) than non-mutated patients (median 16.9 ng/ml; range: 3.44–292; $P = 0.019$). No significant differences were observed when the remaining non-driver genes were studied.

Finally, we studied whether there was any correlation between the amount of plasma cfDNA and VAF of the driver mutation or the higher mutation detected. Of note, no correlation was found between the total cfDNA circulating in plasma and the cfDNA mutation VAF or the cfDNA VAF/granulocyte VAF.

cfDNA may be useful to monitor response to treatment

Finally, we analysed whether cfDNA allele frequencies are modulated in a similar way to those in granulocyte DNA during cytoreductive treatment. In two *JAK2V617F*-mutated cases (one PV and one ET), the molecular response to treatment was assessed during a follow-up of 35 and 32 months respectively (Fig 6). For the PV case, who received hydroxycarbamide, the *JAK2V617F* VAF remained stable in both the granulocytes and cfDNA during the follow-up. For the ET case, who was treated with interferon, a proportional decrease in the *JAK2V617F* VAF was observed in granulocytes and cfDNA.

Discussion

In the present study, we have assessed the genomic characterisation of MPNs by targeted NGS of plasma cfDNA. This is, to the best of our knowledge, the first report analysing the mutational profile of MPN patients in cfDNA. Of note, all samples were taken at diagnosis before any kind of cytoreductive therapy, thus excluding any potentially modifying effect on the results. Our data demonstrate an equivalent repertoire of variants in both driver and non-driver genes supporting the feasibility of performing the molecular characterisation of MPN patients by cfDNA, since all mutations initially identified in the granulocyte fraction were also detected in plasma. It is cheaper to use cfDNA than granulocyte DNA: fewer reagents and hands-on time are needed for sample processing (in our hands 25 min for cfDNA processing including two rounds of centrifugation *versus* a minimum of 1 h for granulocyte isolation). Therefore, the analysis of cfDNA represents a novel strategy that would be useful for routine testing as cfDNA is obtained fast and easily from blood plasma, when compared with granulocyte purification.

Standardisation of cfDNA manipulation is still ongoing; however, it is generally accepted that samples should be processed within the first 6 h after blood collection to avoid leukocyte lysis and contamination of cfDNA with genomic DNA. Nevertheless, recent studies have reported that cfDNA levels are stable in K3EDTA tubes at room temperature for

N. Garcia-Gisbert *et al.*

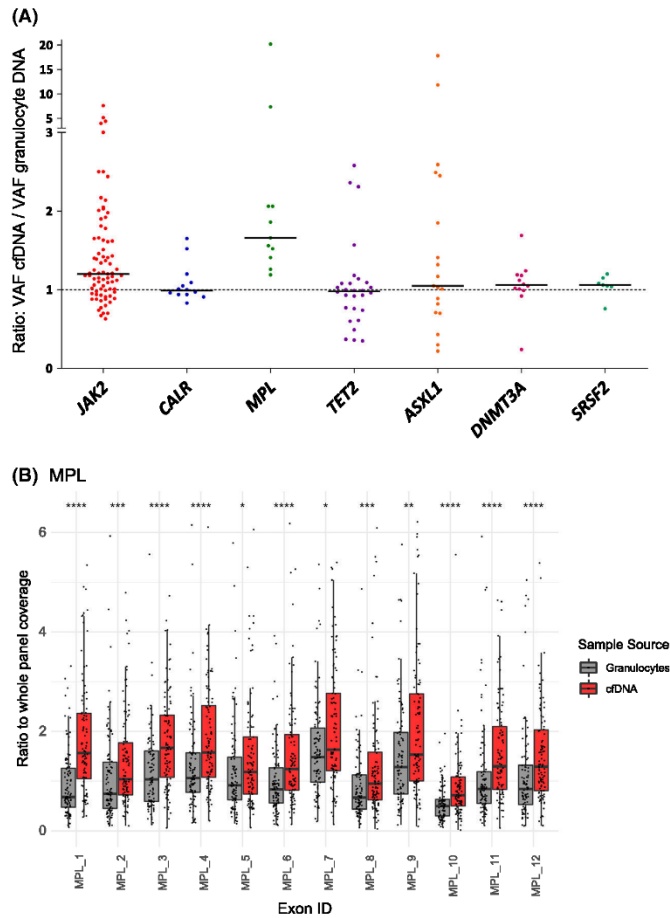


Fig 4. (A) Ratio (cDNA VAF / granulocyte VAF) of the detected variants of the most frequently mutated genes in our cohort. Median VAF for each gene is shown as a black line. Variants situated in the plot above the line have a higher VAF in cDNA than in granulocyte DNA and variants below the line have a higher VAF in granulocyte DNA. *MPL* mutations were detected in all the patients with a higher VAF in cDNA. (B) *MPL* read depth in granulocytes and cDNA. The ratio (read depth for the *MPL* exon/whole panel read depth for that sample) for the 12 exons of the *MPL* gene is shown for the granulocyte DNA samples and the cDNA samples. All *MPL* exons were overrepresented in cDNA libraries when compared to granulocytes DNA. * $P \leq 0.05$, ** $P \leq 0.01$, *** $P \leq 0.001$, **** $P \leq 0.0001$. [Colour figure can be viewed at wileyonlinelibrary.com]

up to 24 h, which would facilitate the incorporation of cfDNA in routine practice.^{12,13}

One of the main caveats of liquid biopsy in the study of solid tumours is that very highly sensitivity methods are required to assess gene alterations since tumoural cfDNA represents a small fraction that is frequently challenging to identify.¹⁴ To overcome this difficulty, NGS analysis requires expensive molecular tagging and high-depth sequencing. In contrast to solid tumours, the VAFs of gene mutations in cfDNA from MPN patients were similar or superior to those

found in granulocytic DNA. This finding is of utmost importance as it supports the potential use of cfDNA for molecular profiling by NGS at an affordable cost and with limited analysis complexity. However, either in solid tumours or in haematological neoplasms, if cfDNA is used as a screening procedure, the confounding possibility of clonal haematopoiesis of indeterminate potential (CHIP) should be taken into account and the detection of mutations in myeloid-associated genes should be integrated in the patient clinical context.¹⁵

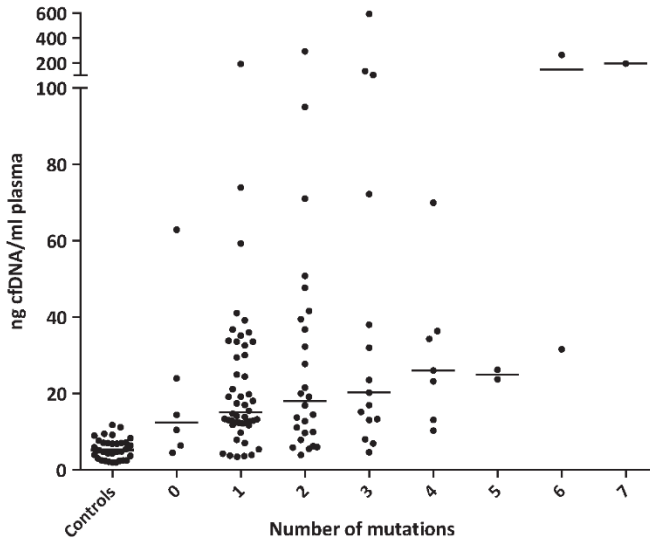


Fig 5. cfDNA quantity (ng/ml plasma) represented and grouped by the number of mutations (0 – 7) identified in each patient by NGS. Each dot represents a patient. Lines represent medians for each group.

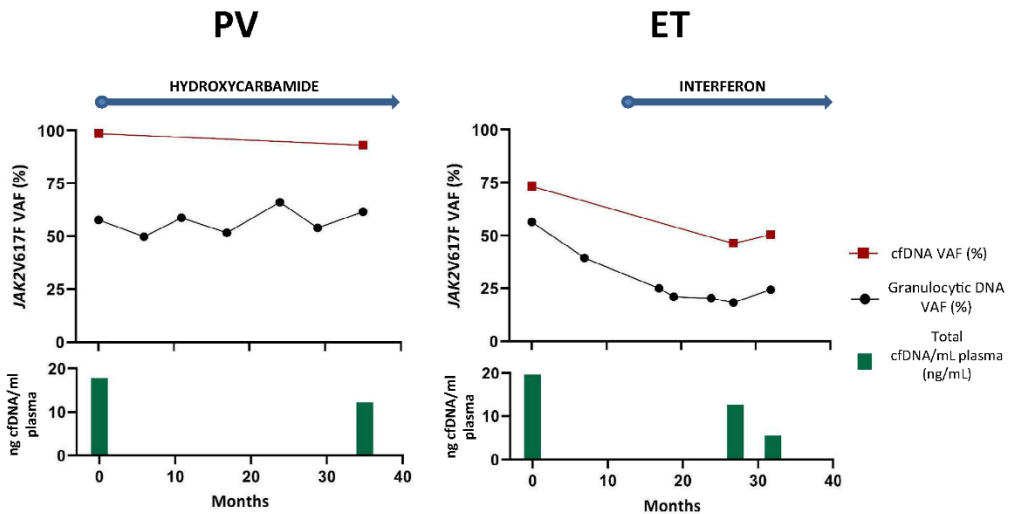


Fig 6. Monitoring of molecular response in cfDNA in two *JAK2V617F* mutated cases. Molecular response to treatment was assessed during a follow up of 35 months in the PV case and 32 months in the ET case. The PV case started hydroxycarbamide treatment one month after the diagnosis. The ET patient started interferon treatment 14 months after the diagnosis. Squares represent the VAF in cfDNA, black dots represent the VAF in granulocyte DNA and columns in the last row show the concentration of cfDNA in plasma expressed in ng/ml of plasma. [Colour figure can be viewed at wileyonlinelibrary.com]

In our study, the VAFs of mutations detected were higher in cfDNA than in granulocytes for *MPL* and *JAK2* mutations (66% and 20% increase respectively), and to a lower extent

for *SRSF2* (6% increase). These results indicate that the analysis of cfDNA could improve mutation detection in MPNs. Our findings are in line with previous studies showing that

N. Garcia-Gisbert *et al.*

plasma or serum samples are a better option than peripheral blood cells to detect *JAK2V617F* mutation in MPN patients.^{16,17} Moreover, in one PMF patient included in our cohort, two additional mutations in *MPL* and *ASXL1* were identified in cfDNA that had not been initially seen in granulocytic DNA, suggesting that cfDNA analysis provides a higher sensitivity to detect both driver and non-driver mutations in MPN patients, especially in *MPL*-mutated cases. Our group has already described that the analysis of granulocytes is not always the best strategy to perform mutational analysis, especially in ET patients, in whom platelets represented the mutated population better at least for *JAK2V617F*-mutated patients.^{18,19} However, the use of platelets is cumbersome, whereas the use of cfDNA is an easy approach to implement in routine practice that could potentially become the method of choice.

Regarding the total amount of cfDNA, we observed that this value differs depending on the clinical or disease phenotype and genotype of patients. cfDNA is a variable parameter, that even in healthy individuals presents significant variations due to inflammation, injury or exercise.^{20–22} However, irrespective of disease phenotype our MPN samples showed a significantly higher amount of cfDNA than those of healthy controls, indicating a higher release of cfDNA in plasma from MPN clonal cells. Of note, PMF patients' plasma contained a higher quantity of cfDNA than that of PV and ET patients, even when adjusting the cfDNA concentration values with the leukocyte count. This finding supports that the cfDNA in PMF is not released by peripheral blood leukocytes, but by haematopoietic cells. This higher shedding of cfDNA is in line with the known increased release of CD34⁺ cells into circulation and the altered bone marrow stem cell niche in PMF. Moreover, these differences in the amount of cfDNA are concordant with the fact that PMF patients show higher levels of circulating nucleosomes in blood than those with PV, ET or healthy controls.²³

MPN patients with greater molecular complexity showed a higher concentration of cfDNA in plasma than those patients with fewer mutations. This higher shedding of cfDNA could be explained in PMF because patients acquire more mutations and release more cfDNA than in PV and ET; however, inside the PMF group, the correlation was still observed. In this regard, it has been reported that cfDNA in plasma has the capacity to enter cell nuclei and insert itself into the host cell genome. Previous studies *in vivo* have reported that it is possible to generate tumours by injecting cfDNA molecules in mice,²⁴ and have found evidence that cfDNA sequences are capable of inserting into the genome of healthy cells,^{25,26} which would contribute to genetic instability.²⁷ A possible hypothesis to explain the correlation between the amount of cfDNA and the number of mutations is that this aberrant production of higher quantities of cfDNA is inducing genetic instability in haematopoietic cells of MPN patients by the insertion of DNA molecules and hence, inducing a higher rate of mutations. Nevertheless, we cannot rule out that

genomic instability induced by other mechanisms (e.g. alterations of the DNA repair machinery) could prompt a higher shedding of cfDNA to peripheral blood.

Interestingly, all *MPL* mutations were found with a higher VAF in cfDNA than in granulocytes both in ET and PMF patients. In addition, we observed that *MPL*-mutated patients had higher amounts of cfDNA than *MPL*-wild-type patients, suggesting that this higher shedding of cfDNA improves the detection rate in these patients. Moreover, unlike for other genes, we observed a higher number of reads covering the *MPL* gene in cfDNA samples than in granulocyte DNA samples. CfDNA fragments result from nucleosome fragmentation that does not occur randomly. In fact, it has been reported that specific fragmentation profiles are different between cancer patients and healthy individuals.²⁸ No profiles so far have been described for haematological malignancies, but it seems possible that this fragmentation pattern may somehow favour the detection of *MPL* mutations.

The improvement in the detection of *MPL* mutations with the use of cfDNA may have direct clinical impact, as mutated ET patients have a higher risk of myelofibrotic transformation. cfDNA genotyping especially in *JAK2*- and *CALR*-wild-type MPN patients will detect with more reliability low-burden *MPL* mutations and improve the recognition of relevant prognostic factors.^{29,30}

Finally, as molecular monitoring may be of interest in some MPN patients, we assessed the molecular profile in sequential samples from two patients who received cytoreductive therapy. In both cases, the VAF in cfDNA was higher than in granulocytes in all samples, and cfDNA profiling mirrored the modulation of the granulocytic *JAK2V617F* burden. Additional studies are required to further explore this application.

In conclusion, our data show that the analysis of cfDNA allows the characterisation of the molecular abnormalities of patients with MPNs. The sensitivity and accuracy for mutation detection in driver and non-driver genes were equal or even superior to that obtained when studying the isolated granulocytic population, especially regarding the detection of *MPL* mutations.

Acknowledgements

This study was supported in part by ISCIII, PI16/0153, PI19/0005, 2017SGR205, PT17/0015/0011, Grant Gilead 2016 and Xarxa de Banc de Tumors de Catalunya.

Author contributions

NGG performed the research and the statistical analysis, analysed and interpreted the results and wrote the manuscript. CB and BB designed the research study, analysed and interpreted the results and wrote the manuscript. LFI, CFR and JG performed the research and analysed and interpreted the results. MA, LC, AA, RL, LA, AS, and XC

collected and analysed the data. All authors reviewed and approved the final version of the manuscript.

Conflicts of interest

AS: Roche (Research Funding, Speakers Bureau), Janssen Pharmaceuticals (Consultancy, Speakers Bureau), Gilead (Consultancy, Speakers Bureau), Celgene (Consultancy). CB: Gilead (Research Funding). BB: Thermo Fisher Scientific (Consultancy, Speakers Bureau), Qiagen (Consultancy, Speakers Bureau), Roche (Research Funding, Speakers Bureau).

Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Fig S1. Sample workflow for DNA extraction and mutational analysis.

Fig S2. Ratio [cell-free (cf)-DNA variant allele frequency (VAF)/granulocyte VAF] of the detected variants of the most frequently mutated genes in our cohort, grouped by disease phenotype. Median VAF for each gene is shown as a black line. Variants situated in the plot above the line have a higher VAF in cfDNA than in granulocytes and variants below the line have a higher VAF in granulocytes.

Fig S3. Read depth for each gene in granulocytes and cell-free (cf)-DNA. The ratio (read depth for the exon/whole panel read depth for that sample) for each exon of the genes included in the next-generation sequencing (NGS) panel is shown for the granulocyte samples and the cfDNA samples. *, $P \leq 0.05$; **, $P \leq 0.01$; ***, $P \leq 0.001$; ****, $P \leq 0.001$.

Table S1. Genetic variants detected by next-generation sequencing [NGS; Human Genome Variation Society (HGVS) nomenclature].

References

- Siravegna G, Marsoni S, Siena S, Bardelli A. Integrating liquid biopsies into the management of cancer. *Nat Rev Clin Oncol*. 2017;14(9):531–48.
- Wan JCM, Massie C, Garcia-Corbacho J, Mouliere F, Brenton JD, Caldas C, et al. Liquid biopsies come of age: towards implementation of circulating tumour DNA. *Nat Rev Cancer*. 2017;17(4):223–38.
- Lui YYN, Chik K-W, Chiu RWK, Ho C-Y, Lam CWK, Lo YMD. Predominant hematopoietic origin of cell-free DNA in plasma and serum after sex-mismatched bone marrow transplantation. *Clin Chem*. 2002;48(3):421–7.
- Snyder MW, Kircher M, Hill AJ, Daza RM, Shendure J. Cell-free DNA comprises an In Vivo Nucleosome Footprint that Informs Its Tissues-Of-Origin. *Cell*. 2016;164(1–2):57–68.
- Ulz P, Thallinger GG, Auer M, Graf R, Kashofer K, Jahn SW, et al. Inferring expressed genes by whole-genome sequencing of plasma DNA. *Nat Genet*. 2016;48:1273–8.
- Moss J, Magenheimer J, Neiman D, Zemmour H, Loyfer N, Korach A, et al. Comprehensive human cell-type methylation atlas reveals origins of circulating cell-free DNA in health and disease. *Nat Commun*. 2018;9(1):5068.
- Iriyama C, Tomita A, Hoshino H, Adachi-Shirahata M, Furukawa-Hibi Y, Yamada K, et al. Using peripheral blood circulating DNAs to detect CpG

- global methylation status and genetic mutations in patients with myelodysplastic syndrome. *Biochem Biophys Res Commun*. 2012;419(4):662–9.
- Nakamura S, Yokoyama K, Shimizu E, Yusa N, Kondoh K, Ogawa M, et al. Prognostic impact of circulating tumor DNA status post-allogeneic hematopoietic stem cell transplantation in AML and MDS. *Blood*. 2019;133(25):2682–95.
- Yeh P, Dickinson M, Ftouni S, Hunter T, Sinha D, Wong SQ, et al. Molecular disease monitoring using circulating tumor DNA in myelodysplastic syndromes. *Blood*. 2017;129(12):1685–90.
- Dimassi S, Simonet T, Labalme A, Boutry-Kryza N, Campan-Fournier A, Lamy R, et al. Comparison of two next-generation sequencing kits for diagnosis of epileptic disorders with a user-friendly tool for displaying gene coverage. *DeCovA. Appl Transl Genomics*. 2015;1(7):19–25.
- Mayakonda A, Lin D-C, Assenov Y, Plass C, Koeffler HP. Maftools: efficient and comprehensive analysis of somatic variants in cancer. *Genome Res*. 2018;28(11):1747–56.
- Parpart-Li S, Bartlett B, Popoli M, Adleff V, Tucker L, Steinberg R, et al. The effect of preservative and temperature on the analysis of circulating tumor DNA. *Clin Cancer Res*. 2017;23(10):2471–7.
- Risberg B, Tsui DWY, Biggs H, Martin R-V, de Almagro A, Dawson SF, et al. Effects of collection and processing procedures on plasma circulating cell-free DNA from cancer patients. *J Mol Diagnostics*. 2018;20:883–92.
- Zill OA, Banks KC, Fairclough SR, Mortimer SA, Vowles JV, Mokhtari R, et al. The landscape of actionable genomic alterations in cell-free circulating Tumor DNA from 21,807 advanced cancer patients. *Clin Cancer Res*. 2018;24(15):3528–38.
- Steenma DP. Clinical consequences of clonal hematopoiesis of indeterminate potential. *Blood Adv*. 2018;2(22):3404–10.
- Ma W, Kantarjian H, Zhang X, Sun W, Buller AM, Jilani I, et al. Higher detection rate of JAK2 mutation using plasma. *Blood*. 2008;111(7):3906–7.
- Barra GB, Santa Rita TH, Almeida ALSC, Jácomo RH, Nery LFA. Serum Has Higher Proportion of Janus Kinase 2 V617F Mutation Compared to Paired EDTA-Whole Blood Sample: A Model for Somatic Mutation Quantification Using qPCR and the 2-AAcCq Method. *Diagnostics*. 2020;10(3):153.
- Bellosillo B, Martínez-Avilés L, Gimeno E, Florensa L, Longarón R, Navarro G, et al. A higher JAK2 V617F-mutated clone is observed in platelets than in granulocytes from essential thrombocythemia patients, but not in patients with polycythemia vera and primary myelofibrosis. *Leukemia*. 2007;21(6):1331–2.
- Angona A, Fernández-Rodríguez C, Alvarez-Larrán A, Camacho I, Longarón R, Torres E, et al. Molecular characterisation of triple negative essential thrombocythemia patients by platelet analysis and targeted sequencing. *Blood Cancer J*. 2016;6(8):e463.
- Atamaniuk J, Vidotto C, Tschan H, Bachl N, Stuhlmeier KM, Müller MM. Increased concentrations of cell-free plasma DNA after exhaustive exercise. *Clin Chem*. 2004;50:1668–70.
- Paunel-Görgülü A, Wacker M, El Aita M, Hassan S, Schlachtenberger G, Deppe A, et al. cfDNA correlates with endothelial damage after cardiac surgery with prolonged cardiopulmonary bypass and amplifies NETosis in an intracellular TLR9-independent manner. *Sci Rep*. 2017;7(1):17421.
- Jackson Chornenki NL, Coke R, Kwong AC, Dwivedi DJ, Xu MK, McDonald E, et al. Comparison of the source and prognostic utility of cfDNA in trauma and sepsis. *Intensive Care Med Exp*. 2019;7(1):29.
- Marin Oyarzún CP, Carestia A, Lev PR, Glembofsky AC, Castro Ríos MA, Moiraghi B, et al. Neutrophil extracellular trap formation and circulating nucleosomes in patients with chronic myeloproliferative neoplasms. *Sci Rep*. 2016;6:38738.
- Trejo-Becerril C, Pérez-Cárdenas E, Taja-Chayeb L, Anker P, Herrera-Goeppfert R, Medina-Velázquez LA, et al. Cancer progression mediated by horizontal gene transfer in an in vivo model. *PLoS One*. 2012;7(12):e52754.
- Bergsmedh A, Szeles A, Henriksson M, Bratt A, Folkman MJ, Spetz A-L, et al. Horizontal transfer of oncogenes by uptake of apoptotic bodies. *Proc Natl Acad Sci*. 2001;98(11):6407–11.

N. Garcia-Gisbert *et al.*

26. Mittra I, Khare NK, Raghuram GV, Chaubal R, Khambatti F, Gupta D, et al. Circulating nucleic acids damage DNA of healthy cells by integrating into their genomes. *J Biosci.* 2015;**40**(1):91–111.
27. Basak R, Nair NK, Mittra I. Evidence for cell-free nucleic acids as continuously arising endogenous DNA mutagens. *Mutat Res Mol Mech Mutagen.* 2016;**1**(793–794):15–21.
28. Cristiano S, Leal A, Phallen J, Fiksel J, Adleff V, Bruhm DC, et al. Genome-wide cell-free DNA fragmentation in patients with cancer. *Nature.* 2019;**570**(7761):385–9.
29. Haider M, Elala YC, Gangat N, Hanson CA, Tefferi A. MPL mutations and palpable splenomegaly are independent risk factors for fibrotic progression in essential thrombocythemia. *Blood Cancer J.* 2016;**6**(10):e487.
30. Asp J, Andréasson B, Hansson U, Wasslavik C, Abellsen J, Johansson P, et al. Mutation status of essential thrombocythemia and primary myelofibrosis defines clinical outcome. *Haematologica.* 2016;**101**(4):e129–e132.

Article 2:

Molecular and Cytogenetic Characterization of Myelodysplastic Syndromes in Cell-free DNA

Garcia-Gisbert N, Garcia-Ávila S, Merchán B, Salido M, Fernández-Rodríguez C, Gibert J, Fernández-Ibarrondo L, Camacho L, Lafuente M, Longarón R, Espinet B, Vélez P, Pujol RM, Andrade-Campos M, Arenillas L, Salar A, Calvo X, Besses C, Bellosillo B.

Manuscript submitted to *Blood Advances*

MOLECULAR AND CYTOGENETIC CHARACTERIZATION OF MYELODYSPLASTIC SYNDROMES IN CELL-FREE DNA

Garcia-Gisbert N^{1,2}, Garcia-Ávila S^{1,3}, Merchán B¹, Salido M^{4,5}, Fernández-Rodríguez C^{1,5}, Gibert J¹, Fernández-Ibarrondo L^{1,2}, Camacho L^{1,5}, Lafuente M^{1,2}, Longarón R^{1,5}, Espinet B^{4,5}, Vélez P^{1,3}, Pujol RM⁶, Andrade-Campos M¹, Arenillas L^{4,5}, Salar A^{1,3}, Calvo X^{4,5}, Besses C¹, Bellosillo B^{1,5}

¹Group of Applied Clinical Research in Hematology, IMIM-Hospital del Mar, Barcelona, Spain ²Pompeu Fabra University, Barcelona, Spain ³Department of Hematology, Hospital del Mar, Barcelona, Spain ⁴Group of Translational Research on Hematological Neoplasms, IMIM-Hospital del Mar, Barcelona, Spain ⁵Department of Pathology, Hospital del Mar, Barcelona, Spain ⁶Department of Dermatology, Hospital del Mar, Barcelona, Spain

ABSTRACT

Molecular and cytogenetic studies are essential in patients with myelodysplastic syndromes (MDS) for diagnosis and prognosis. Cell-free DNA (cfDNA) analysis has been reported as a reliable non-invasive approach for detecting molecular abnormalities in MDS, however, there is limited information about cytogenetic alterations and monitoring in cfDNA. We have assessed the molecular and cytogenetic profile of a cohort of 70 patients with MDS by next-generation sequencing (NGS) using cfDNA and compared the results to paired bone marrow (BM) DNA. Sequencing of BM DNA and cfDNA showed a comparable mutational profile (92.1% concordance) and variant allele frequencies (VAF) strongly correlated between both sample types. Of note, *SF3B1* mutations were detected with significantly higher VAF in cfDNA than in BM DNA. NGS and microarrays were highly concordant to detect chromosomal alterations although with lower sensitivity than karyotype/FISH. Nevertheless, all cytogenetic aberrations detected by NGS in BM DNA were also detected in cfDNA. Additionally, molecular and cytogenetic alterations were monitored and we observed an excellent correlation between the VAF of mutations in BM DNA and cfDNA across multiple matched time points. A decrease in the cfDNA VAF was detected in patients responding to therapy, but not in non-responding patients. Of note, cfDNA analysis also showed cytogenetic evolution in 2 cases not responding to treatment. In conclusion, our results support the analysis of cfDNA as an

accurate strategy for performing molecular characterization, detection of chromosomal aberrations and monitoring of MDS patients.

INTRODUCTION

Myelodysplastic syndromes (MDS) are hematopoietic stem cell disorders characterized by dysplasia and ineffective hematopoiesis that are driven by somatically acquired genomic alterations. Molecular studies and conventional cytogenetics are essential in MDS to establish a correct diagnosis and to set up accurate risk stratification(1). Routinely, these analyses are performed in bone marrow (BM) samples, in particular cytogenetic analysis as it is difficult to obtain metaphases from peripheral blood (PB) samples(2).

In recent years, it has been demonstrated that molecular profiling can be performed robustly using cell-free DNA (cfDNA) analysis in solid tumors and lymphomas. CfDNA molecules are short DNA fragments present in plasma samples that are mainly released by immature hematopoietic and bone marrow cells(3–6). As MDS are characterized by an excessive apoptosis in bone marrow (7,8), so an increased release of cfDNA to plasma is expected in these patients. Indeed, several groups have reported that it is possible to identify the genetic alterations in MDS by analyzing cfDNA(9–12).

However, there is limited information regarding the detection of cytogenetic alterations in MDS patients by cfDNA analysis. To this aim, we have designed a targeted gene panel to detect in a single test both molecular and cytogenetic alterations by NGS and investigated its potential use with cfDNA in comparison to bone marrow samples in a cohort of MDS patients.

PATIENTS AND METHODS

Patients. BM aspirates and PB samples were prospectively collected from 70 newly diagnosed MDS patients or cases who only received erythropoietin with the following diagnoses: MDS with single lineage dysplasia (SLD; n=1), MDS with multilineage dysplasia (MLD; n=35), MDS with ring sideroblasts (RS)-SLD (n=5), MDS-RS-MLD (n=17), MDS with isolated del(5q) (n=2), MDS with excess blasts (EB)-1(n=6), MDS-EB-2(n=2), MDS-unclassifiable (MDS-U;

n=2) (Table 1). The IPSS-R score was calculated for each patient(13). We additionally analyzed PB samples from 21 healthy controls and 18 acute myeloid leukemia (AML) patients (**Supplementary Table 1**). The study was approved by the local ethics committee (2016/6768/l).

Peripheral blood and bone marrow processing and DNA isolation. BM aspirates were collected and BM DNA was extracted with MagAttract DNA Blood Mini M48 Kit (Qiagen, Hilden, Germany). Peripheral blood samples were collected in K3EDTA tubes and processed in the first 4 hours to isolate plasma (Supplementary Fig 1). CfDNA was isolated automatically using QIAAsymphony SP (QIAAsymphony DSP Virus/Pathogen Kit, Qiagen, Hilden, Germany) and quantified with Qubit 3.0 (Thermo Fisher Scientific, Eugene, USA). Purity of cfDNA was assessed by electrophoresis (4200 TapeStation system, Agilent, Santa Clara, USA) to discard the presence of genomic DNA. All the cfDNA samples analyzed in this project were free of genomic DNA contamination.

Next Generation Sequencing (NGS). Genomic characterization was performed in paired samples of BM DNA and cfDNA by NGS in all patients. Libraries were prepared using a custom panel including 48 myeloid-associated genes (*ASXL1, ATM, BCOR, BCORL1, CALR, CBL, CEBPA, CHEK2, CSF3R, CSNK1A1, CUX1, DDX41, DLEU7, DNMT3A, EGR1, ETV6, EZH2, FLT3, GATA2, IDH1, IDH2, JAK2, KIT, KMT2A, KRAS, MPL, NF1, NPM1, NRAS, PHF6, PPM1D, PRPF8, PTPN11, RAD21, RUNX1, SETBP1, SF3B1, SH2B3, SRSF2, STAG2, TET2, TNFSF11, TP53, TP53RK, TP53TG5, U2AF1, WT1, ZRSR2*) and genomic regions localized at the most frequently altered chromosomes in MDS (QIAseq Custom DNA Panels, Qiagen) Genomic regions included in the NGS Panel are included in **Supplementary Table 2**. Unique molecular identifiers were incorporated before targeted amplification to tag individual DNA molecules. Libraries were sequenced with a 3000x minimum read depth in MiSeq/NextSeq (Illumina, San Diego, CA, USA).

The GeneGlobe Data Analysis Center (Qiagen) was used for FASTQ trimming, alignment and variant calling (smCounter2)(14). Variants were annotated and classified using Illumina VariantStudio 3.0 software and visualized with Integrative Genomics Viewer (IGV) v2.11 software. Only pathogenic and likely

pathogenic variants with a variant allele frequency (VAF)>2% were considered.

Copy number variant (CNV) analysis was performed by NGS to detect cytogenetic alterations in both cfDNA and BM DNA. Samples from healthy individuals (2 to 4) were included in all sequencing runs and used as coverage controls. Gene coverage was compared to each sample using GeneGlobe Data Analysis Center to identify regions affected by CNVs, where the normalized coverage is significantly different from the controls(15).

Chromosomal Microarrays (CMA). Cytoscan 750K cytogenetic Solution (Thermo Fisher Scientific) was used to obtain a genetic gain, loss and regions of homozygosity (ROH) profile following the manufacturer's recommendations. This chip consists of more than 750,000 markers for copy number analysis with 550,000 unique non-polymorphic probes and approximately 200,000 SNPs that fully genotype with greater than 99 percent accuracy. Chromosome Analysis Suite v.4.1 (ChAS) software (Thermo Fisher Scientific) and the hg38 genome version (NA36 annotations) was used to analyze the results. Gains with a minimum of 25 altered markers in a 150 Kb region, losses with at least 35 altered markers in a 75 Kb region, and regions with telomeric copy neutral loss of heterozygosity (CN-LOH) greater than 10 Mb or affecting relevant genes have been collected.

Fluorescence in situ hybridization (FISH) analyses. FISH studies were performed according to the standard methods used in our laboratory(16). FISH studies were performed on BM cells from cytogenetic cultures using the following probes: Vysis CEP8, Vysis EGFR1 FISH probe kit (Abbott Molecular, IL, US) and XL 20q12/20qter (Metasystems, Altlußheim, Germany).

Statistical Analysis. The IBM SPSS Statistics software was used for statistical analysis. For categorical data, comparisons of proportions were evaluated by Chi-square test or Fisher's exact test as appropriate. For continuous variables, comparisons were assessed by nonparametric Mann–Whitney or Wilcoxon test when appropriate. We assessed the Spearman's rank correlation coefficient to evaluate the strength of association between two variables. P-values <0.05 were considered statistically significant. Coverage metrics were obtained using DeCovA library (17). Variant analysis was performed in R 3.6.2

version using maftools package (18). The code used in R 3.6.2 to create the figures is displayed in **Supplemental Methods** and the required files to generate the figures and full list of variants identified is shown in **Supplemental Data 1-3**.

RESULTS

The amount of cell-free DNA in plasma is higher in MDS than in controls.

A total of 70 plasma samples from MDS patients at diagnosis or in the absence of any therapy were analyzed. The clinical and biological features of patients are shown in Table 1. The amount of total cfDNA obtained in MDS patients (median: 58.4 ng/ml) was significantly higher than that obtained from healthy controls (median: 32.4 ng/ml) ($P = 0.023$, Mann-Whitney) (Figure 1). No significant differences were observed in cfDNA concentration among MDS patients when comparing by disease subtype or by risk category. Nevertheless, even low risk MDS patients had a significantly higher cfDNA concentration than the healthy control group ($P = 0.023$). On the contrary, a significantly lower cfDNA concentration was observed in the MDS group when compared to the cohort of AML patients ($P = 0.017$).

We analyzed the correlation of the concentration of cfDNA with clinical and biological characteristics. A positive correlation was observed between the amount of cfDNA and the serum lactate dehydrogenase (LDH) levels ($P=0.027$, $r_s=0.273$). No statistically significant association was observed with hematological parameters (hemoglobin, leukocytes, monocytes, platelets or blast percentage).

cfDNA and BM DNA show an equivalent mutational profile.

Mutational profiling of BM DNA and cfDNA showed comparable results: 187 mutations were detected in BM DNA and cfDNA, with a 92.1% concordance (Figure 2). The most frequently mutated genes were *TET2* (45.7%), *SF3B1* (37.1%), *ASXL1* (21.4%), *DNMT3A* (20.0%), *SRSF2* (15.7%), *ZRSR2* (11.4%) and *U2AF1* (11.4%). A strong correlation was observed between the VAF of BM and cfDNA ($r_s = 0.797$, $P < 0.001$, Spearman) (Figure 3). There were 16 discordant mutations: 8 were only detected in cfDNA and 8 were only

detected in BM (Figure 4A). These discordant mutations presented a lower variant allele frequency (VAF) (median 5.60%, range 2.5–25.53%) when compared to the VAF observed in the whole cohort (median 28.27%, range 0.74–98.28%) ($P < 0.001$) (Figure 4B). These cases showed that the correlation between BM DNA and cfDNA mutations may decrease when studying low-represented subclones.

***SF3B1* mutations present a higher VAF in cfDNA than BM DNA.**

We compared the VAF of the detected mutations in cfDNA and BM DNA grouped by gene and observed that VAFs of *SF3B1* mutations were significantly higher in cfDNA than in BM DNA ($P=0.016$, Wilcoxon) (Figure 5A). No significant differences were observed in the concentration of total cfDNA between the *SF3B1* mutated and the *SF3B1* wild type patients. Mutations in exon 15 of *SF3B1* (NP_036565.2: p.(Lys700Glu) in all cases) presented a tendency towards a higher VAF cfDNA/BM ratio than mutations in other *SF3B1* exons (median ratio of 1.82 vs. 1.09, $P = 0.08$) (Figure 5B). In this context, we assessed the representation of *SF3B1* exons in cfDNA, as it has been reported that nucleosome distribution affects DNA fragmentation, and as a consequence some genomic regions are overrepresented in cfDNA (5). We compared the depth of coverage obtained for the exons, and we observed a higher representation of *SF3B1* exon 15 in cfDNA libraries than in BM libraries (Figure 5C). This finding was not observed in other *SF3B1* exons, which suggests that exon 15 is better represented in cfDNA from MDS patients, thus producing a higher VAF in cfDNA of the *SF3B1* p.(Lys700Glu) mutation. In line with previous studies, we observed that the percentage of ring sideroblasts in bone marrow correlated with the VAF of *SF3B1* mutations, in both BM DNA and cfDNA ($r_s = 0.684$, $P < 0.001$ in BM DNA and $r_s = 0.602$, $P = 0.002$ in cfDNA) (Supplementary Figure 2).

Moreover, we identified a *SF3B1* p.(Lys700Glu) mutation only detectable in cfDNA in one patient. In this case, the quantification of ring sideroblasts was not assessable due to lack of cellularity in the BM aspirate, being the analysis of cfDNA a useful non-invasive alternative to identify the presence of this pathological clone. This *SF3B1* mutation was later confirmed by NGS in a subsequent PB sample.

CMA and NGS were highly concordant to detect cytogenetic aberrations

In addition to gene mutations, we assessed the detection of cytogenetic alterations by NGS. Cytogenetic/FISH alterations were detected at the time of diagnosis in 20/70 (28.6%) MDS patients (Figure 6A). From those, 2/20 were infrequent alterations in MDS and were not covered by the design of the NGS panel and in 6/20 cases chromosome Y loss was the only alteration detected, neither covered by the NGS panel. So, the cohort included cytogenetic alterations potentially detectable by our gene panel in 12/70 patients.

NGS analysis detected abnormalities in 10/70 MDS patients, in both BM DNA and cfDNA. Interestingly, in a patient without analyzable metaphases in bone marrow karyotype, del(20q) was found by NGS and further confirmed by CMA. Overall, CMA and NGS were highly concordant to detect chromosomal aberrations although they did not reach the sensitivity achieved by conventional cytogenetic analysis (karyotype/FISH) (Figure 6B) (Supplementary Figure 3). Nevertheless, as previously stated, all cytogenetic aberrations detected by NGS in BM DNA were also detected in cfDNA.

cfDNA is useful to predict transformation and monitor response to treatment.

Molecular and cytogenetic alterations were monitored in sequential samples from 7 cases (median follow up: 13 months, range 10-30). We observed an excellent correlation between the VAF of mutations in BM and cfDNA across multiple-matched time points. Both sample types showed similar clonal dynamics irrespectively of the treatment and allowed the monitoring of both mutations and chromosomal aberrations (Figure 7).

In those cases treated with hypomethylating agents (i.e. azacitidine), a VAF decrease was detected in patients responding to therapy, but not in non-responding patients. Of note, cfDNA analysis also showed cytogenetic evolution in 2 cases not responding to azacitidine (del(12p) and +21), who had to stop treatment due to lack of response. In the patient treated with FLAG-IDA followed by hematopoietic cell transplantation (HCT), the 5 mutations identified at diagnosis were undetectable in cfDNA in a sample collected 7

months after the HCT. One patient treated with hypoxia-inducible factor (HIF) inhibitor showed a VAF decrease of *DNMT3A* and *SF3B1* mutations and a concomitant increase in the *RUNX1* and *SETBP1* VAF during the follow-up, who later transformed to chronic myelomonocytic leukemia. In addition, the emergence of a mutation in *ASXL1*, undetectable at diagnosis, was identified in the latest sample available in both cfDNA and BM DNA.

Two patients who were not receiving treatment were also monitored. One patient, who evolved to AML, showed a clonal expansion of the *NF1* mutant clone at the time of AML transformation. The second patient acquired a subclonal del(7q) not detected by NGS and observed only by karyotype in 2/20 metaphases. Although our cohort of AML cases (15/18 were *de novo* AML) presented a higher cfDNA concentration than MDS at diagnosis, we did not observe an increase in the concentration of cfDNA in the 2 MDS patients that progressed to AML.

DISCUSSION

In the present study we have assessed the genomic characterization of MDS by targeted NGS of plasma cfDNA compared to bone marrow DNA. This is, to the best of our knowledge, the largest series of cfDNA analysis in MDS patients. Of note, all samples were taken at diagnosis or before treatment, thus excluding any potentially modifying effect on the results. We designed an NGS gene panel to detect with a single test both molecular and cytogenetic alterations and investigated its potential use in cfDNA, which would be particularly useful in several cases such as: non-fit or fragile elderly patients, in patients with fibrotic or hypocellular BM and in patients with contraindication or difficult access to bone marrow samples. Our data demonstrate that the analysis of cfDNA represents a novel strategy that would be useful for routine testing as cfDNA is obtained fast and easily from blood plasma, when compared with bone marrow aspirates or purified CD34+ cells(19).

In patients with solid tumors, cfDNA is being incorporated as a non-invasive strategy to assess molecular alterations in routine clinical practice. It has been reported that the majority of the cfDNA is released by hematopoietic cells in

health and disease (3–6). However, MDS patients showed a significantly higher amount of cfDNA than healthy controls, indicating a higher release of cfDNA to PB plasma from MDS clonal cells. Of note, even low risk MDS contained a higher quantity of cfDNA than controls. The ineffective hematopoiesis in bone marrow stem cell niche and the increased apoptosis of these cells in MDS (8,9) is in line with this higher shedding of cfDNA to PB. A significant correlation was observed between the cfDNA concentration and LDH values, in accordance with previous studies (12,20). However, contrary to previous findings (12), we did not find a higher concentration of cfDNA in International Prognostic Scoring System (IPSS) high-risk groups than in low-risk groups. This discordant observation could be explained because most MDS patients included in this study were low-risk IPSS-R patients. The successful analysis of cfDNA in low-risk patients is especially relevant as the incorporation of this technology in clinical practice would allow frequent monitoring of these patients without requiring sequential bone marrow aspirates.

Previous studies comparing the reliability of cfDNA and total PB cellular DNA analysis to detect molecular abnormalities by NGS (21)(22) showed that cfDNA analysis was a better option, as additional mutations were detected in cfDNA and the VAF in cfDNA was significantly higher than in PB DNA(21).

In our study, we observed a similar mutational profile in cfDNA and BM DNA (93% concordance) and the VAFs of the mutations identified in both sample types were highly correlated. However, some discordant mutations were also identified in a small proportion of patients, in some cases mutations that may have prognostic relevance, such as *SF3B1* mutations or mutations in damage DNA repair (DDR) genes. Cases with mutations detected in cfDNA and not in BM DNA were all cases in which plasma and BM samples had been collected at different time points; in addition, some of these discordant cases showed low VAFs that could reflect small clones emerging or slowly expanding, or *de novo* acquired mutations. Overall, cfDNA and BM DNA showed a high concordance, although they may have a worse correlation in subclonal alterations, as it has been previously reported for cfDNA in AML (23). Also, it cannot be ruled out that some of these alterations could derive from clonal

hematopoiesis of indeterminate potential (CHIP) which reinforces the need to always integrate the molecular information with the patients' morphological studies and clinical context (24,25).

Interestingly, the VAFs of *SF3B1* mutations were significantly higher in cfDNA than in BM DNA, especially for exon 15 *SF3B1* mutations (i.e. p.(Lys700Glu)). This observation is clinically relevant, as the analysis of cfDNA could be the best alternative to detect these mutations when low quality BM aspirates are obtained or to detect small mutant clones. In these cases, previous studies agree that cfDNA is a better option than PB cells to assess the molecular profile in MDS (21). We identified a *SF3B1* mutation in cfDNA and not in BM DNA in a patient in which the presence of ring sideroblasts was not assessable due to lack of cellularity in the BM aspirate. We hypothesize that the position of the nucleosomes in exon 15 of *SF3B1* could be facilitating the detection of the *SF3B1* p.(Lys700Glu) mutation, as we observed that exon 15 of *SF3B1* was more represented in cfDNA than in BM DNA, whereas exons 14 and 16 were not. This finding may indicate that chromatin is more condensed in exon 14 of *SF3B1*, and therefore these fragments are preserved when released into the bloodstream.

Regarding the detection of cytogenetic alterations in cfDNA of MDS patients, our results confirm that gains or losses of genetic material are reflected in the cfDNA and thus we can identify most of them by NGS. This is the first study assessing the detection of cytogenetic alterations in cfDNA by NGS in a cohort of MDS patients. Only, one previously published study identified the loss of chromosome 9 in cfDNA in a patient with MDS using low-coverage whole-genome sequencing (LC-WGS) (11). Other studies had used NGS to detect these chromosomal alterations using PB or BM samples (26). We believe that the identification of cytogenetic alterations in cfDNA using NGS is a useful technique for those patients without analyzable metaphases in the karyotype, with a hypocellular BM aspirate or as a non-invasive tool to monitor these alterations in the follow-up of MDS patients, as it is known that clonal dynamic acquisition during the follow up of has a significant prognostic impact on MDS(27). On the one side, we detected a del(20q) in a patient without analyzable metaphases at diagnosis, which was confirmed by CMA.

Moreover, in two patients receiving treatment with hypomethylating agents, cfDNA mirrored the cytogenetic alterations acquired during therapy. Further validation of these results including high-risk cytogenetic subgroups would support the value of cfDNA analysis as a useful tool to be implemented in routine clinical practice that could improve the identification of alterations required for accurate risk classification.

However, we have also identified some limitations in our CNA analysis. It should be noted that the design of the NGS panel only covers a selected part of the genome, so those alterations occurring outside of the covered region will not be detected. The chromosome Y deletion was not included in the NGS panel designed since it has been associated with normal aging(28) and is considered a very good prognosis cytogenetic alteration in IPSS-R score(13). As NGS comprehensive molecular profiling with broad NGS targeted panels or even exome or whole genome analyses are implemented in clinical practice, these limitations will be overcome. We also observed that, in our experience, patients with subclonal cytogenetic alterations, poorly represented in the sample, could only be detected by karyotype or FISH, due to limitations of sensitivity of NGS or CMA. We must recall that, at present, cytogenetic techniques are still besides morphology the backbone elements of MDS diagnosis.

One of the plausible applications of cfDNA is disease monitoring. To evaluate this, we analyzed sequential samples from 7 patients with MDS using NGS. Previous studies already proved that the clonal dynamics of BM VAFs were mimicked by cfDNA VAFs (11). This fact, as we observed, indicates whether the patient is responding to treatment or not in a non-invasive assessment. The most remarkable result when monitoring these MDS patients was the finding that in two patients not responding to hypomethylating agents, NGS of cfDNA allowed us to detect the occurrence of cytogenetic evolution. Monitoring patients using this less invasive technique could allow early detection of clinically relevant genomic changes, using only a PB sample. One of the limitations of this technique is that, as it has occurred in one of the untreated patients, we do not have sufficient sensitivity to detect alterations that are present in a low proportion of cells, not being detected by cfDNA

until the clone has expanded. However, this approach would facilitate frequent assessment of disease evolution or response to treatment, especially in fragile elderly patients or cases with hypocellular or fibrotic BM.

In conclusion, we have shown that cfDNA mirrors the molecular profile of BM in MDS. In our cohort, enriched with low risk IPSS-R patients, cytogenetic alterations were detectable in most cases by NGS in both BM DNA and cfDNA. These data support that the analysis of cfDNA is a reliable and feasible method to characterize and monitor the molecular abnormalities present in patients with MDS.

Acknowledgements. Study supported by ISCIII-FEDER, PI16/0153, PI19/0005, 2017SGR205, PT20/00023 and Xarxa de Banc de Tumors de Catalunya.

Author contributions: NGG performed the research and the statistical analysis, analyzed and interpreted the results and wrote the manuscript. CB and BB designed the research study, analyzed and interpreted the results and wrote the manuscript. SGA, BM, MS, CFR and JG performed the research, collected the data and interpreted the results. LFI, LC, ML, RL, BE, PV, RMP and MAC, collected and analyzed the data. LA, AS and XC collected the data, analyzed and interpreted the results. All authors reviewed and approved the final version of the manuscript.

BIBLIOGRAPHY

1. Arber DA, Orazi A, Hasserjian R, Thiele J, Borowitz MJ, Le Beau MM, et al. The 2016 revision to the World Health Organization classification of myeloid neoplasms and acute leukemia. *Blood*. 2016 May 19;127(20):2391–405.
2. Cherry AM, Slovak ML, Campbell LJ, Chun K, Eclache V, Haase D, et al. Will a peripheral blood (PB) sample yield the same diagnostic and prognostic cytogenetic data as the concomitant bone marrow (BM) in myelodysplasia? *Leuk Res*. 2012 Jul 1;36(7):832–40.
3. Lui YYN, Chik K-W, Chiu RWK, Ho C-Y, Lam CWK, Lo YMD. Predominant hematopoietic origin of cell-free DNA in plasma and serum after sex-mismatched bone marrow transplantation. *Clin Chem*. 2002 Mar;48(3):421–7.
4. Snyder MW, Kircher M, Hill AJ, Daza RM, Shendure J. Cell-free DNA Comprises an In Vivo Nucleosome Footprint that Informs Its Tissues-Of-Origin. *Cell*. 2016 Jan 14;164(1–2):57–68.
5. Ulz P, Thallinger GG, Auer M, Graf R, Kashofer K, Jahn SW, et al.

- Inferring expressed genes by whole-genome sequencing of plasma DNA. *Nat Genet.* 2016 Oct 29;48(10):1273–8.
6. Moss J, Magenheim J, Neiman D, Zemmour H, Loyfer N, Korach A, et al. Comprehensive human cell-type methylation atlas reveals origins of circulating cell-free DNA in health and disease. *Nat Commun.* 2018 Dec 29;9(1):5068.
 7. Clark DM, Lampert IA. Apoptosis is a common histopathological finding in myelodysplasia: The correlate of ineffective haematopoiesis. *Leuk Lymphoma.* 1990;2(6):415–8.
 8. Shetty V, Hussaini S, Alvi S, Joshi L, Shaher A, Dangerfield B, et al. Excessive apoptosis, increased phagocytosis, nuclear inclusion bodies and cylindrical confronting cisternae in bone marrow biopsies of myelodysplastic syndrome patients. *Br J Haematol.* 2002 Mar 1;116(4):817–25.
 9. Iriyama C, Tomita A, Hoshino H, Adachi-Shirahata M, Furukawa-Hibi Y, Yamada K, et al. Using peripheral blood circulating DNAs to detect CpG global methylation status and genetic mutations in patients with myelodysplastic syndrome. *Biochem Biophys Res Commun.* 2012 Mar 23;419(4):662–9.
 10. Nakamura S, Yokoyama K, Shimizu E, Yusa N, Kondoh K, Ogawa M, et al. Prognostic impact of circulating tumor DNA status post–allogeneic hematopoietic stem cell transplantation in AML and MDS. *Blood.* 2019 Jun 20;133(25):2682–95.
 11. Yeh P, Dickinson M, Ftouni S, Hunter T, Sinha D, Wong SQ, et al. Molecular disease monitoring using circulating tumor DNA in myelodysplastic syndromes. *Blood.* 2017 Mar 23;129(12):1685–90.
 12. Suzuki Y, Tomita A, Nakamura F, Iriyama C, Shirahata-Adachi M, Shimada K, et al. Peripheral blood cell-free DNA is an alternative tumor DNA source reflecting disease status in myelodysplastic syndromes. *Cancer Sci.* 2016 Sep 25;107(9):1329–37.
 13. Greenberg PL, Tuechler H, Schanz J, Sanz G, Garcia-Manero G, Solé F, et al. Revised International Prognostic Scoring System for Myelodysplastic Syndromes. *Blood.* 2012 Sep 20;120(12):2454.
 14. Xu C, Gu X, Padmanabhan R, Wu Z, Peng Q, DiCarlo J, et al. smCounter2: an accurate low-frequency variant caller for targeted sequencing data with unique molecular identifiers. *Biol I, editor. Bioinformatics.* 2019 Apr 15;35(8):1299–309.
 15. Reinecke F, Satya RV, DiCarlo J. Quantitative analysis of differences in copy numbers using read depth obtained from PCR-enriched samples and controls. *BMC Bioinformatics.* 2015 Jan 28;16(1):1–9.
 16. Solé F, Salido M, Espinet B, Garcia JL, Martinez Climent JA, Granada I, et al. Splenic marginal zone B-cell lymphomas: two cytogenetic subtypes, one with gain of 3q and the other with loss of 7q. *Haematologica.* 2001 Jan;86(1).
 17. Dimassi S, Simonet T, Labalme A, Boutry-Kryza N, Campan-Fournier A,

- Lamy R, et al. Comparison of two next-generation sequencing kits for diagnosis of epileptic disorders with a user-friendly tool for displaying gene coverage, DeCovA. *Appl Transl Genomics*. 2015 Dec 1;7:19–25.
18. Mayakonda A, Lin D-C, Assenov Y, Plass C, Koeffler HP. Maftools: efficient and comprehensive analysis of somatic variants in cancer. *Genome Res*. 2018 Oct 19;28(11):1747–56.
 19. Martin R, Acha P, Ganster C, Palomo L, Dierks S, Fuster-Tormo F, et al. Targeted deep sequencing of CD34+ cells from peripheral blood can reproduce bone marrow molecular profile in myelodysplastic syndromes. *Am J Hematol*. 2018 Jun 1;93(6):E152–4.
 20. Garcia-Gisbert N, Fernández-Ibarrondo L, Fernández-Rodríguez C, Gibert J, Andrade-Campos M, Arenillas L, et al. Circulating cell-free DNA improves the molecular characterisation of Ph-negative myeloproliferative neoplasms. *Br J Haematol*. 2021;192(2):300–9.
 21. Albitar F, Ma W, Diep K, De Dios I, Agersborg S, Thangavelu M, et al. Deep Sequencing of Cell-Free Peripheral Blood DNA as a Reliable Method for Confirming the Diagnosis of Myelodysplastic Syndrome. *Genet Test Mol Biomarkers*. 2016 Jul 1;20(7):341–5.
 22. Zhao P, Qin J, Liu W, Zhu Q, Fan T, Xiao H, et al. Using circulating tumor DNA to monitor myelodysplastic syndromes status. Vol. 37, *Hematological Oncology*. John Wiley and Sons Ltd; 2019. p. 531–3.
 23. Short NJ, Patel KP, Albitar M, Franquiz M, Luthra R, Kanagal-Shamanna R, et al. Targeted next-generation sequencing of circulating cell-free DNA vs bone marrow in patients with acute myeloid leukemia. *Blood Adv*. 2020 Apr 28;4(8):1670–7.
 24. Steensma DP. Clinical consequences of clonal hematopoiesis of indeterminate potential. *Blood Adv*. 2018 Nov 27;2(22):3404–10.
 25. Gutierrez-Rodrigues F, Beerman I, Groarke EM, Patel BA, Spitofsky N, Dillon LW, et al. Utility of plasma cell-free DNA for de novo detection and quantification of clonal hematopoiesis. *Haematologica*. 2021 Sep 30;
 26. Liquori A, Lesende I, Palomo L, Avetisyan G, Ibáñez M, González-Romero E, et al. A Single-Run Next-Generation Sequencing (NGS) Assay for the Simultaneous Detection of Both Gene Mutations and Large Chromosomal Abnormalities in Patients with Myelodysplastic Syndromes (MDS) and Related Myeloid Neoplasms. *Cancers (Basel)*. 2021 Apr 18;13(8).
 27. Jabbour E, Takahashi K, Wang X, Cornelison AM, Abruzzo L, Kadia T, et al. Acquisition of cytogenetic abnormalities in patients with IPSS defined lower-risk myelodysplastic syndrome is associated with poor prognosis and transformation to acute myelogenous leukemia. *Am J Hematol*. 2013 Oct 1;88(10):831–7.
 28. Stone JF, Sandberg AA. Sex chromosome aneuploidy and aging. *Mutat Res*. 1995 Oct;338(1–6):107–13.

TABLES AND FIGURES

Table 1. Clinical and biological features of MDS patients included in the study.

Characteristic	MDS N=70
Age, median (range), years	81 (54-94)
Male gender, n (%)	51 (72.9)
Hemoglobin, median (range), g/dL	11.75 (7.6-17.8)
WBC count, median (range), x10 ⁹ /L	4.58 (1.44-12.28)
Neutrophil count, median (range), x10 ⁹ /L	2.41 (0.31-7.68)
Platelet count, median (range), x10 ⁹ /L	154 (28-584)
BM blasts %, median (range)	2 (0-19)
Presence PB blasts, n (%)	2 (2.86)
LDH, median (range)	292 (111-487)
Altered karyotype, n (%)	20 (28.6)
-Y as single alteration, n (%)	6 (8.6)
Altered karyotype (other than -Y), n (%)	14 (20)
IPSS-R risk group	
Very low, n (%)	28 (40)
Low, n (%)	31 (44.3)
Intermediate, n (%)	7 (10)
High, n (%)	2 (2.9)
Very high, n (%)	2 (2.9)
MDS subtype (WHO 2017)	
MDS-SLD, n (%)	1 (1.4)
MDS-MLD, n (%)	35 (50)
MDS-RS-SLD, n (%)	5 (7.1)
MDS-RS-MLD, n (%)	17 (24.3)
MDS-del(5q), n (%)	2 (2.9)
MDS-EB-1, n (%)	6 (8.6)
MDS-EB-2, n (%)	2 (2.9)
MDS-U, n (%)	2 (2.9)
Number of patients with mutations, n (%)	66 (94.3)

Abbreviations: WBC: white blood cells; BM: bone marrow; PB: peripheral blood; LDH: Lactate dehydrogenase; IPSS-R: Revised International Prognostic Scoring System; MDS: myelodysplastic syndrome; WHO: World Health Organization; SLD: single lineage dysplasia; MLD: multilineage dysplasia; RS: ring sideroblasts; EB: excess blasts; U: unclassifiable.

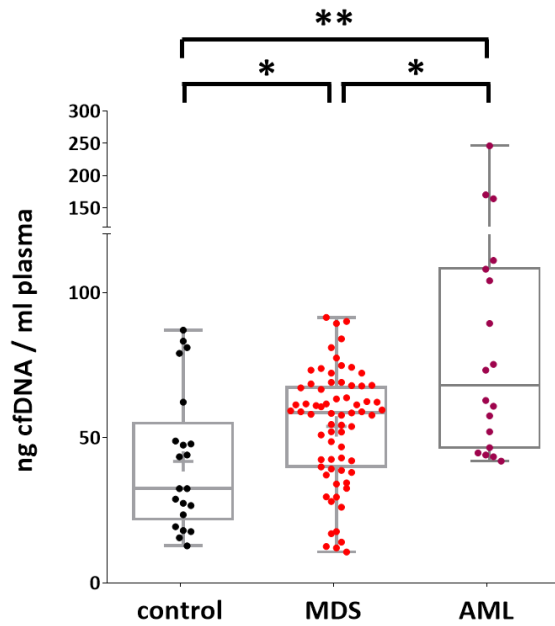


Figure 1. cfDNA concentration (ng cfDNA/ml) in plasma samples from healthy controls, MDS and AML. * $P \leq 0.05$, ** $P \leq 0.01$

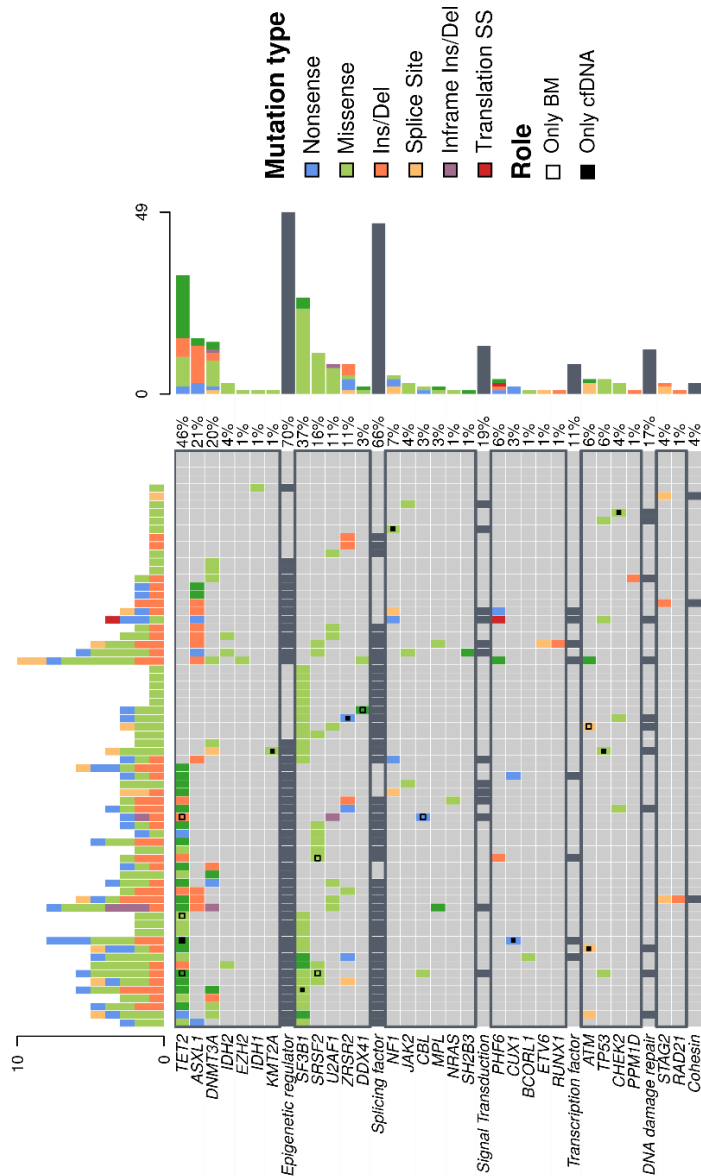


Figure 2. Distribution of mutations identified in both BM DNA and cfDNA in the 70 MDS patients. Results of the sequencing are shown in the plot where each column represents a patient and each row represents a gene. The number of mutations identified per patient is represented as columns in the top row. Genes are grouped by function and ordered from the most to the least frequently mutated. Frequencies for each gene are displayed at the right, as well as the mutation type (nonsense, missense, insertion/deletion, splice site or translation start site). Discordant mutations are represented with a square as shown in the legend: filled squares show mutations only identified in cfDNA and empty squares show mutations only identified in BM DNA. Ins: insertion, Del: deletion, SS: start site, BM: bone marrow, cfDNA: cell-free DNA

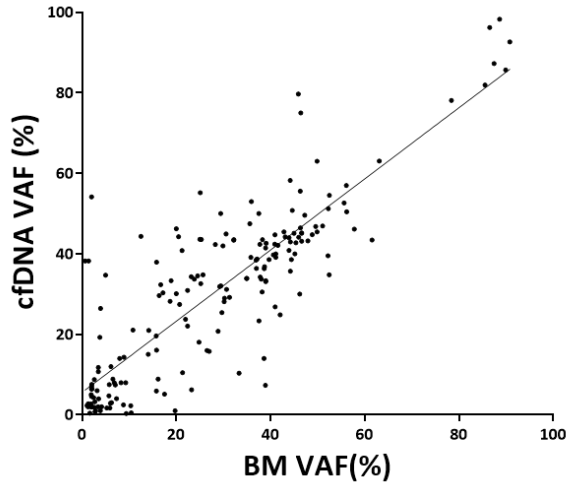


Figure 3. Scatter plot of the 187 variants detected in cfDNA and BM DNA showing the correlation between the variant allele frequency (VAF) ($r_s = 0.797$, $P < 0.001$, Spearman).

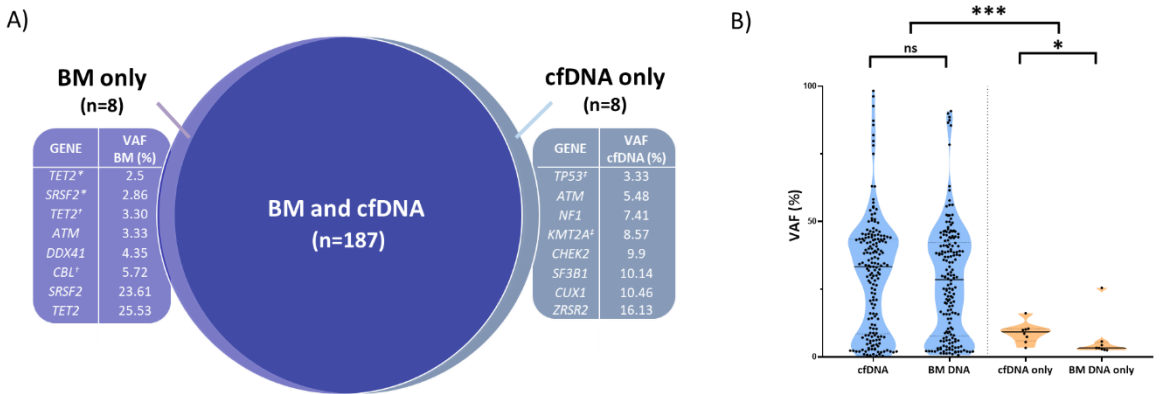


Figure 4. A) Discordant mutations identified in BM DNA and cfDNA. One patient presented two mutations only detected in cfDNA (marked as ‡ in the figure) and two patients showed two mutations only detected in BM (marked as * and † in the figure). The 10 remaining discordant alterations were identified in 10 different patients. **B)** Variant allele frequencies identified in concordant (blue) and discordant (orange) mutations in BM DNA and cfDNA. BM: bone marrow, cfDNA: cell-free DNA, VAF: variant allele frequency, ns: not significant, * $P \leq 0.05$, ** $P \leq 0.01$, *** $P \leq 0.001$

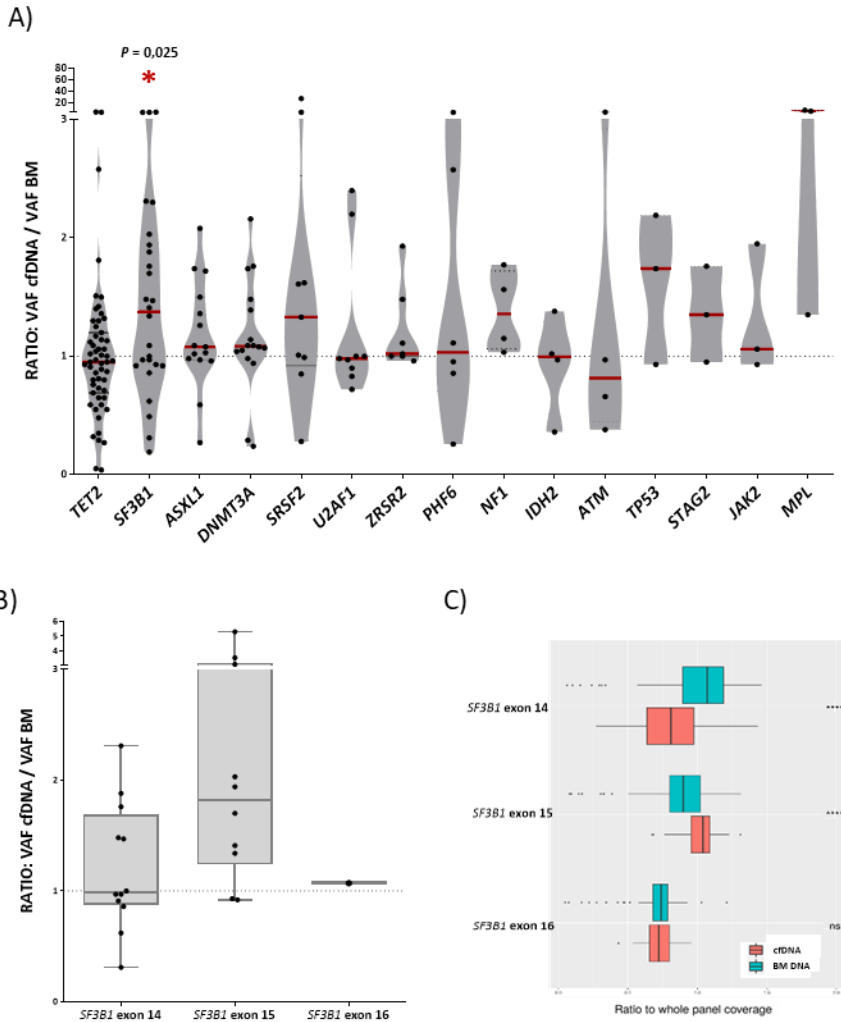


Figure 5. A) Ratio (cfDNA VAF / BM VAF) of the detected variants of the most frequently mutated genes in our cohort. Median VAF ratio for each gene is shown as a red line. Variants situated in the plot above the line have a higher VAF in cfDNA than in BM DNA and variants below the line have a higher VAF in BM DNA. *SF3B1* mutations were detected with a significantly higher VAF in cfDNA. **B)** Ratio (cfDNA VAF / BM VAF) of the *SF3B1* mutations in exons 14, 15 and 16. Mutations in exon 15 of *SF3B1* (p.(Lys700Glu) in all cases) presented a higher VAF cfDNA/BM ratio than mutations in exons 14 and 16 *SF3B1* exons. **c)** *SF3B1* read depth in BM and cfDNA. The ratio (read depth for the *SF3B1* exon/whole panel read depth for that sample) for exons 14,15 and 16 of the *SF3B1* gene is shown for BM DNA samples and the cfDNA samples. *SF3B1* exon 15 is overrepresented in cfDNA libraries when compared to BM libraries. BM: bone marrow, cfDNA: cell-free DNA, VAF: variant allele frequency, * $P \leq 0,05$, ** $P \leq 0,01$, *** $P \leq 0,001$, **** $P \leq 0,0001$, ns: not significant

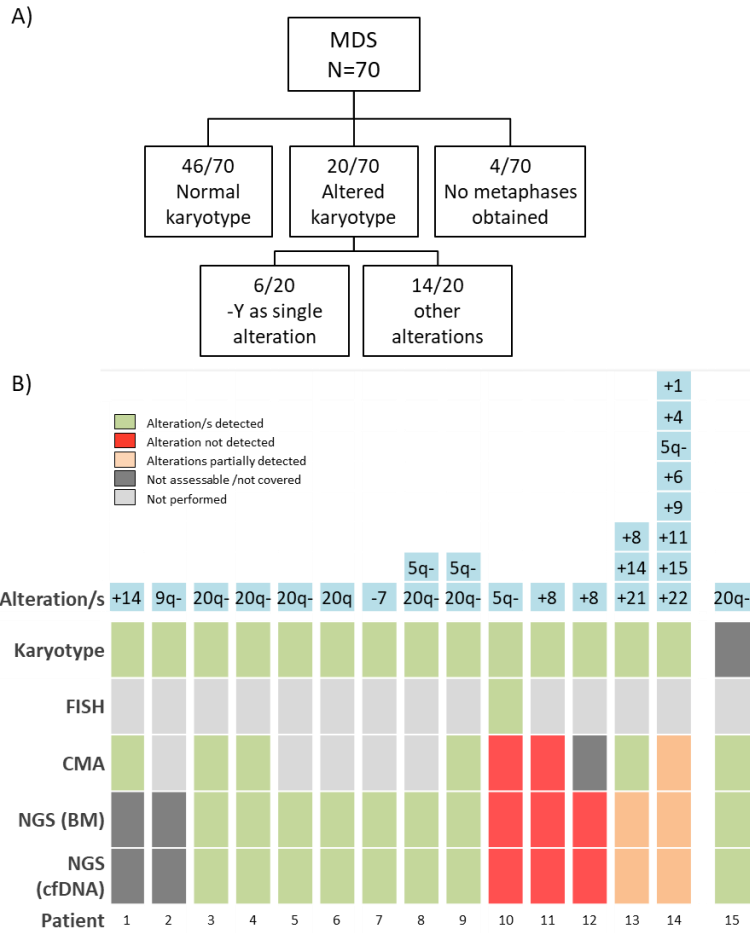


Figure 6. A) Cytogenetic results obtained by karyotype at diagnosis **B)** Detection of cytogenetic alterations by conventional karyotype, FISH, CMA and NGS. Two cases presented alterations not covered by the NGS panel (patients 1 and 2). 9/12 (75%) remaining cases with altered karyotype/FISH (patients 3-14 in the figure) were detected by NGS. Patient 10 presented a 5q- detected in few metaphases and confirmed by FISH, and patients 11 and 12 presented a +8 detected by karyotype in few metaphases. In patient 13, +8 and +21 alterations were detected by NGS, while chromosome 14 is not covered by the design of the NGS panel. In patient 14, 5q- was the only alteration detected by both CMA and NGS due to sensitivity limitations. In a patient without analyzable metaphases (patient 15), 20q- was found by NGS and confirmed by CMA. FISH: Fluorescence In Situ Hybridization, CMA: chromosomal microarrays; NGS: next generation sequencing, BM: bone marrow, cfDNA: cell-free DNA

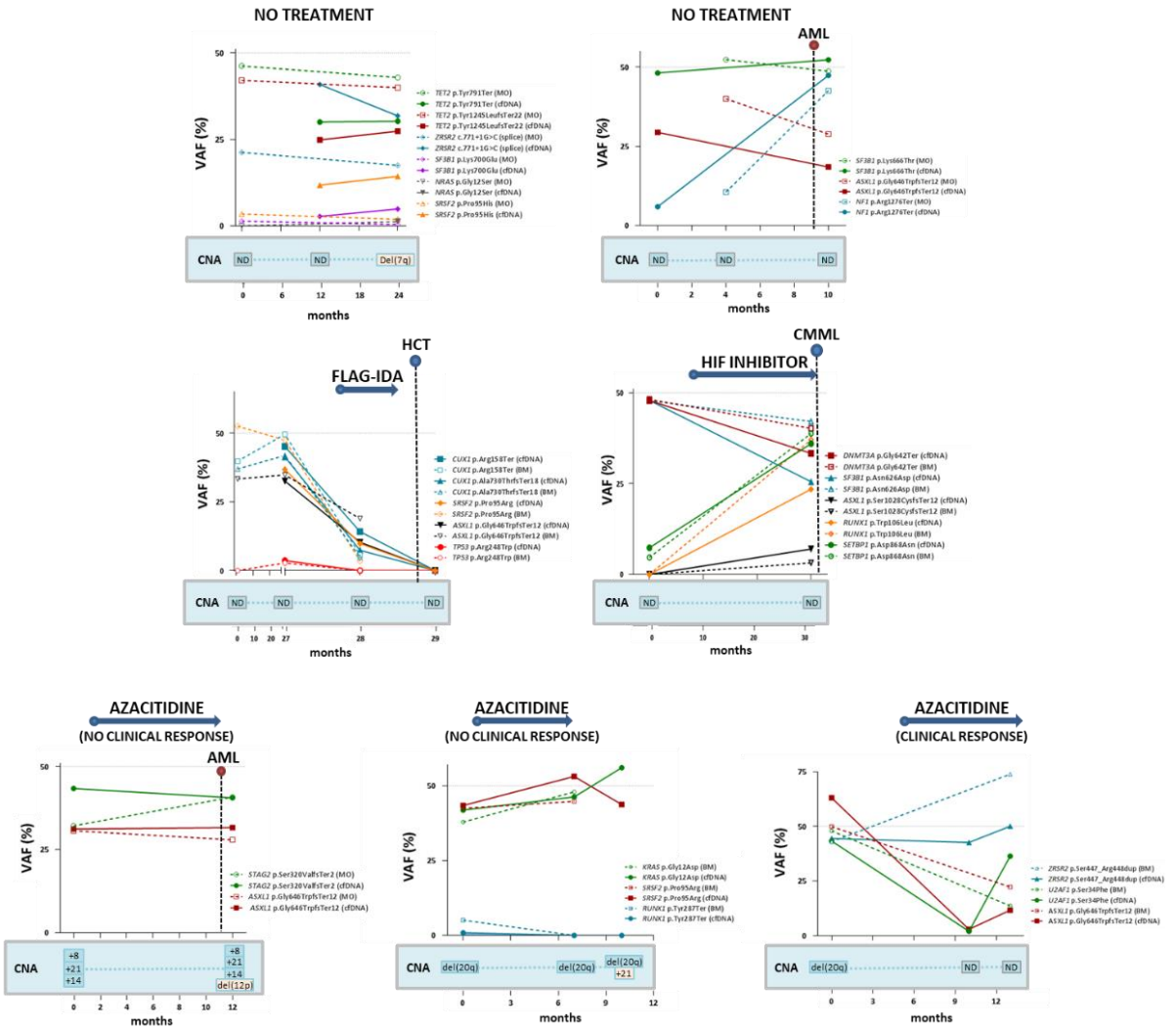


Figure 7. Monitoring of molecular and cytogenetic alterations in 7 patients with MDS. 5 patients receiving treatment (3 azacitidine, 1 FLAG-IDA + HCT, 1 HIF (hypoxia-inducible factor) inhibitor) and 2 untreated cases were included. BM VAF dynamics are shown with a dotted line and cfDNA dynamics are shown with a solid line. VAF: variant allele frequency, AML: acute myeloid leukemia, CMML: chronic myelomonocytic leukemia, CNA: copy number alteration; BM: bone marrow, cfDNA: cell-free DNA, HCT: hematopoietic cell transplantation; ND: not detected

Article 3:

Oligomonocytic and overt chronic myelomonocytic leukemia show similar clinical, genomic, and immunophenotypic features

Calvo X*, Garcia-Gisbert N*, Parraga I, Gibert J, Florensa L, Andrade-Campos M, Merchan B, Garcia-Avila S, Montesdeoca S, Fernández-Rodríguez C, Salido M, Puiggros A, Espinet B, Colomo L, Roman-Bravo D, Bellosillo B, Ferrer A, Arenillas L

* equally contributed

Blood Advances, 2020; 4 (20): 5285–5296

DOI: 10.1182/bloodadvances.2020002206

Oligomonocytic and overt chronic myelomonocytic leukemia show similar clinical, genomic, and immunophenotypic features

Xavier Calvo,^{1,*} Nieves Garcia-Gisbert,^{2,3,*} Ivonne Parraga,¹ Joan Gibert,² Lourdes Florensa,¹ Marcio Andrade-Campos,⁴ Brayan Merchan,⁴ Sara Garcia-Avila,⁴ Sara Montesdeoca,¹ Concepción Fernández-Rodríguez,² Marta Salido,⁵ Anna Puiggros,⁵ Blanca Espinet,⁵ Luis Colomo,^{3,6} David Roman-Bravo,¹ Beatriz Bellosillo,^{2,3} Ana Ferrer,^{1,3} and Leonor Arenillas¹

¹Laboratori de Citologia Hematològica, Servei de Patologia, Grup de Recerca Translacional en Neoplàsies Hematològiques (GRETNHE), and ²Laboratori de Biologia Molecular, Servei de Patologia, Grup de Recerca Clínica Aplicada en Neoplàsies Hematològiques, IMIM Hospital del Mar Research Institute, Barcelona, Spain; ³Departament de Ciències Experimentals i de la Salut, Universitat Pompeu Fabra, Barcelona, Spain; and ⁴Servei d'Hematologia Clínica, Grup de Recerca Clínica Aplicada en Neoplàsies Hematològiques, ⁵Laboratori de Genètica Molecular, and ⁶Secció d'Hematopatologia, Servei de Patologia, GRETNHE, IMIM Hospital del Mar Research Institute, Barcelona, Spain

Key Points

- OM-CMML and overt CMML show a similar clinical, morphological, cytogenetic, molecular, and immunophenotypic profile.
- The results support the consideration of OM-CMML as a distinctive subtype of CMML.

Oligomonocytic chronic myelomonocytic leukemia (OM-CMML) is defined as those myelodysplastic syndromes (MDSs) or myelodysplastic/myeloproliferative neoplasms, unclassifiable with relative monocytosis ($\geq 10\%$ monocytes) and a monocyte count of 0.5 to $<1 \times 10^9/L$. These patients show clinical and genomic features similar to those of overt chronic myelomonocytic leukemia (CMML), although most of them are currently categorized as MDS, according to the World Health Organization 2017 classification. We analyzed the clinicopathologic features of 40 patients with OM-CMML with well-annotated immunophenotypic and molecular data and compared them to those of 56 patients with overt CMML. We found similar clinical, morphological, and cytogenetic features. In addition, OM-CMML mirrored the well-known complex molecular profile of CMML, except for the presence of a lower percentage of RAS pathway mutations. In this regard, of the different genes assessed, only *CBL* was found to be mutated at a significantly lower frequency. Likewise, the OM-CMML immunophenotypic profile, assessed by the presence of $>94\%$ classical monocytes (MO1s) and CD56 and/or CD2 positivity in peripheral blood monocytes, was similar to overt CMML. The MO1 percentage $>94\%$ method showed high accuracy for predicting CMML diagnosis (sensitivity, 90.7%; specificity, 92.2%), even when considering OM-CMML as a subtype of CMML (sensitivity, 84.9%; specificity, 92.1%) in our series of 233 patients (39 OM-CMML, 54 CMML, 23 MDS, and 15 myeloproliferative neoplasms with monocytosis and 102 reactive monocytosis). These results support the consideration of OM-CMML as a distinctive subtype of CMML.

Introduction

Based on the World Health Organization (WHO) 2017 classification, chronic myelomonocytic leukemia (CMML) diagnosis requires the presence of persistent peripheral blood monocytosis $\geq 1 \times 10^9/L$, with monocytes accounting for $\geq 10\%$ of the leukocytes.^{1,2} Although most CMML cases display dysmyelopoiesis, it may not be present. In the absence of dysplasia, a diagnosis of CMML can still be made by the demonstration of clonality by an acquired clonal cytogenetic or molecular abnormality. If no

Submitted 1 May 2020; accepted 20 September 2020; published online 27 October 2020. DOI 10.1182/bloodadvances.2020002206.

*X.C. and N.G.-G. contributed equally to this study.

Presented at the 62nd annual meeting of the American Society of Hematology, Orlando, FL, 7-10 December 2019 and at the 24th Congress of the European Hematology Association, Amsterdam, The Netherlands, 13-16 June 2019.

Original data are available by e-mail request to the corresponding author, Xavier Calvo (e-mail: xcalvo@parcdesalutmar.cat).

The full-text version of this article contains a data supplement.

© 2020 by The American Society of Hematology

clonal marker can be found and dysplasia is not present, the diagnosis of CMML may also be established if the monocytosis persists for at least 3 months and all causes of reactive monocytosis have been excluded.^{1,3} In this context, a wide spectrum of neoplastic, infectious, or inflammatory conditions should be ruled out before the diagnosis of CMML is established. Nevertheless, an autoimmune or neoplastic disease may appear concomitantly, and its presence does not exclude a diagnosis of CMML. Next-generation sequencing (NGS) has emerged as the best tool for establishing diagnostic certainty, because it allows for the demonstration of clonality in most cases of CMML, but this technology is not accessible worldwide. Approximately 90% of patients with CMML display mutations of the *TET2*, *SRSF2*, and/or *ASXL1* gene.⁴⁻⁷ By contrast, an accessible method such as flow cytometry (FC) analysis of peripheral blood (PB) monocyte subsets has attracted interest as a means of diagnosing CMML, because an increase in the classical monocyte (MO1) fraction to >94% shows high sensitivity and specificity for predicting CMML diagnosis.⁸

Recently, Geyer et al defined oligomonocytic chronic myelomonocytic leukemia (OM-CMML) as cases of myelodysplastic syndrome (MDSs) or MDS/myeloproliferative neoplasm (MPN), unclassifiable, with relative monocytosis ($\geq 10\%$ monocytes) and a total monocyte count of 0.5 to $< 1 \times 10^9/L$.⁹ According to the WHO 2017 classification, most of these patients are currently classified within the different categories of MDS. The researchers demonstrated that these cases share clinical and genomic features with overt CMML. To the best of our knowledge, there are no FC data about the distribution of the PB monocyte subset in OM-CMML. Selimoglu-Buet et al indicated that the accumulation of MO1 >94% defined a subgroup of patients with MDS that frequently evolved into CMML.¹⁰ Although some of those patients met OM-CMML criteria, there are no series that explore this aspect in a homogeneous group of patients with OM-CMML.

The purpose of this study was to provide a comprehensive comparison between a large series of well-annotated patients with OM-CMML or CMML, with particularly novel data concerning the immunophenotypic and molecular characteristics of OM-CMML. In addition, we assessed the accuracy of the MO1 percentage >94% method of predicting CMML and OM-CMML diagnosis in a large series ($n = 233$) of patients.

Methods

Patients

We prospectively studied 236 patients and assessed the PB distribution of monocyte subsets by FC in 233 of them. This assessment has been part of the diagnostic routine in our laboratory since 2016. The patients were either initially diagnosed or followed up during this period. All diagnostics were established according to WHO 2017 criteria.

Of the 236 patients, 56 were diagnosed with CMML (16 with "proliferative type" CMML [p-CMML] and 40 with "dysplastic type" [d-CMML]), and 40 met OM-CMML diagnostic criteria. According to the WHO 2017 MDS classification, patients with OM-CMML were classified into the following categories: 1 MDS with single-lineage dysplasia (MDS-SLD), 16 MDS with multilineage dysplasia (MDS-MLD), 4 MDS with ring sideroblasts and single-lineage dysplasia (MDS-RS-SLD), 9 MDS with ring sideroblasts and multilineage dysplasia (MDS-RS-MLD), 9 MDS with excess blasts-1

(MDS-EB-1), and 1 MDS with excess blasts-2 (MDS-EB-2). In addition, we identified 23 patients with MDS who did not meet OM-CMML diagnostic criteria (1 MDS-SLD, 11 MDS-MLD, 3 MDS-RS-SLD, 6 MDS-RS-MLD, 1 MDS-EB-1, and 1 MDS with isolated deletion of 5q), 15 had Ph-negative MPNs with $\geq 1 \times 10^9/L$ monocytes (7 essential thrombocytosis, 6 polycythemia vera, and 2 primary myelofibrosis), and 102 patients had absolute monocytosis of reactive origin. The study was conducted according to the biomedical research guidelines of the Declaration of Helsinki.

NGS

Molecular characterization by targeted NGS was performed on DNA extracted from total PB or bone marrow (BM). Targeted amplicon libraries (QIAseq Custom DNA Panels; Qiagen, Hilden, Germany) were prepared from a custom panel covering the full exonic regions of 25 genes associated with myeloid malignancies (*ASXL1*, *CALR*, *CBL*, *CSF3R*, *DNMT3A*, *ETV6*, *EZH2*, *IDH1*, *IDH2*, *JAK2*, *KIT*, *KRAS*, *MPL*, *NRAS*, *PRPF8*, *RUNX1*, *SETBP1*, *SF3B1*, *SH2B3*, *SRSF2*, *STAG2*, *TET2*, *TP53*, *U2AF1*, and *ZRSR2*).¹¹ Libraries were sequenced with MiSeq or NextSeq (Illumina, San Diego, CA) with a 2000 \times minimum coverage. The variant allele frequencies (VAF; proportion of mutated reads out of the total NGS reads) for each mutation were recorded. The NGS methodology is described in further detail in the supplemental Material.

Flow cytometry analysis of monocyte subsets in peripheral blood

Multiparametric FC analysis of monocyte subsets was performed on whole PB collected on EDTA. Based on Euroflow Consortium recommendations we followed the stain-lyse-wash procedure with FACS Lysing Solution (BD Biosciences, San Jose, CA).¹² Cell surface staining of 2×10^6 cells was performed, and at least 500 000 total events were acquired per tube (FACS Canto II; BD Biosciences). Analysis was performed with Infinicyt, version 1.7 (Cytognos SL, Salamanca, Spain). The strategy of analysis and the 5-tube experimental panel are described in the supplemental Data and supplemental Figure 1.

Statistical analysis

Categorical variables are described by frequencies and percentages and continuous variables as means, medians, and ranges. For categorical data, comparisons of proportions were evaluated by χ^2 test, χ^2 test with Yates continuity correction, or Fisher's exact test, as appropriate. For continuous variables, comparisons were assessed by nonparametric Mann-Whitney *U* test. No adjustments were made to *P*-values for multiple tests. We assessed the Spearman rank correlation or the Φ coefficient to evaluate the strength of association between 2 variables. The area under the receiver operating characteristic (ROC) curve (AUC) of the percentage of MO1s and MO3s was calculated to assess its accuracy for predicting CMML diagnosis. We used the Youden index ($J = \text{sensitivity} + \text{specificity} - 1$) for evaluating the balance between sensitivity and specificity. Survival curves were constructed by the Kaplan-Meier method, using the interval from the date of diagnosis to the date of last contact or death and compared by log-rank test. Differences were considered statistically significant when $P < .05$ in a 2-tailed test. The code used in R v3.6.2 to create the figures is displayed in supplemental Data 2.

Results

OM-CMML and overt CMML show similar clinical, morphological, and cytogenetic features

The clinical findings for the 40 patients with OM-CMML and the 56 patients with CMML are compared in Table 1. As shown, we observed no significant differences in age, sex, platelet count, BM dysgranulopoiesis, BM dysthrombopoiesis, percentage of BM blasts, percentage of abnormal karyotypes, distribution of the Spanish cytogenetic risk groups,¹³ and proportion of patients showing blasts in PB. Patients with OM-CMML showed lower absolute leukocyte and monocyte count, a predictable finding, given the definition of OM-CMML. Moreover, they showed a lower percentage of PB and BM monocytes and BM promonocytes.

Patients with OM-CMML were also more anemic and showed more evident dyserythropoiesis. In this sense, we observed a higher proportion of OM-CMML showing *SF3B1* mutation and $\geq 5\%$ ring sideroblasts (28% vs 12%; $P = .056$). Patients with OM-CMML or CMML who displayed this feature showed a significantly lower hemoglobin level and a higher median percentage of dyserythropoiesis (Hb, median: 11 vs 12 g/dL; $P = .010$; dyserythropoiesis: median, 60% vs 22%; $P < .001$, *SF3B1* mutated vs unmutated).

OM-CMML and overt CMML show a similar mutational profile

Molecular characterization by NGS was performed in all patients with OM-CMML and in 53 of 56 patients with CMML. As depicted in Table 1, there were no significant differences in the proportion of patients showing at least 1 mutation (40 of 40 vs 52 of 53; $P = .99$; OM-CMML vs CMML) in the median number of mutated genes per patient (2 vs 3; $P = .407$, OM-CMML vs CMML) or in the median number of mutations per patient (3 vs 3; $P = .134$, OM-CMML vs CMML).

The mutation patterns of OM-CMML and CMML are depicted in Figure 1. The genes mutated at a frequency $>10\%$ in patients with OM-CMML were *TET2* (72%), *SRSF2* (30%), *SF3B1* (27.5%), *ZRSR2* (20%), *ASXL1* (17.5%), *DNMT3A* (15%), and *RUNX1* (12.5%). In patients with CMML, the genes mutated at a frequency $>10\%$ were *TET2* (75%), *ASXL1* (28.3%), *SRSF2* (26.4%), *CBL* (20.8%), *SF3B1* (17%), *NRAS* (11.3%), and *KRAS* (11.3%). The VAFs were similar between both groups in all genes, except for *DNMT3A* (supplemental Table 1). In line with the literature, the 3 most frequently mutated genes in CMML group were *TET2*, *ASXL1*, and *SRSF2*.^{4-6,14} Remarkably, no significant difference was observed in the proportion of patients showing concurrent *TET2* and *SRSF2* mutations, the gene signature of CMML¹⁵ (27.5% vs 22.6%, $P = .591$, OM-CMML vs CMML). Only 1 gene mutated at a significantly different frequency: *CBL* (2.5% vs 20.8%; $P = .011$, OM-CMML vs CMML; Figure 2A). Notably, we found no gene mutated at a significantly different proportion when comparing OM-CMML and d-CMML (Table 2). As expected, CMML showed a higher percentage of mutations in RAS pathway genes (mutations in *CBL*, *NRAS*, and/or *KRAS*) than did OM-CMML, given that these genes have been associated with proliferative features^{16,17} (37.7% vs 5%; $P < .001$). Although d-CMML showed a significantly higher percentage of RAS-pathway mutations than OM-CMML (27% vs 5%; $P = .011$), this difference was especially evident in p-CMML (62.5% vs 5%; $P < .001$), in which genes associated with

proliferation were present at higher frequencies¹⁶⁻¹⁸: *CBL* (2.5% vs 31.3%; $P = .006$), *NRAS* and/or *KRAS* (2.5% vs 31.3%; $P = .006$), and *ASXL1*¹⁴ (17.5% vs 62.5%; $P = .003$) (Table 2). These mutations were also more frequent in p-CMML than in d-CMML (*ASXL1*: 62.5% vs 13.5, $P = .001$; RAS-pathway: 62.5% vs 27%, $P = .014$) (Table 2). It is also worth noting that the proliferative condition associated with the presence of *ASXL1* mutations in our series could be explained in part by the presence of concomitant RAS-pathway mutations. In this sense, we found a positive correlation between mutations in the RAS pathway and *ASXL1* (Φ coefficient, 0.23; $P = .029$). It would be interesting to explore this association in larger series of patients with CMML, because *ASXL1* mutation is a well-established independent adverse prognostic factor in CMML,^{5,6,14} but it also seems to be partially interrelated with RAS mutations and p-CMML type, 2 other well-accepted independent adverse prognostic factors in this disease.^{6,19}

As previously reported, mutations in the hydroxymethylation pathway (mutations in *IDH1*, *IDH2*, and/or *TET2*) are almost mutually exclusive in acute myeloid leukemia²⁰ and CMML.²¹ In our CMML series, we found no concomitant mutations in *TET2* and *IDH1* or *IDH2*, but surprisingly, 3 patients with OM-CMML showed simultaneous mutations in these genes, 2 with *TET2* and *IDH2* mutations and 1 with *TET2* and *IDH1* mutations (Figures 1A and 3). The impairment of the hydroxymethylation pathway was present in the majority of these patients (78%, 83%; OM-CMML, CMML) and, remarkably, in all p-CMML cases in our series. Moreover, a high proportion of patients with OM-CMML or CMML showed more than 1 *TET2* mutation (37.5% vs 52.8%, $P = .142$, OM-CMML vs CMML) and the distribution of the different *TET2* subtype of mutations was almost identical in a comparison of both groups (Figure 2B). These findings suggest that the impairment of this pathway could be the pathophysiological hallmark of these entities.

Finally, in line with published data, mutations in the assessed splicing factors in our series (*SF3B1*, *SRSF2*, *ZRSR2*, *U2AF1*, and *PRPF8*) were almost mutually exclusive.²²⁻²⁶ In the OM-CMML group, we observed 1 patient with concomitant *SRSF2* and *SF3B1* mutations and another with simultaneous *SRSF2* and *ZRSR2* mutations (Figures 1A, and 3). Only 1 patient with CMML showed simultaneous *SF3B1* and *ZRSR2* mutations (Figures 1B and 3).

Graphic representations of the mutations are depicted in supplemental Figure 2, and the full list of variants identified is shown in supplemental Data 3.

The increase in the fraction of MO1 $>94\%$ is shown as the approach with the highest accuracy for predicting CMML diagnosis

The repartition of monocyte subsets in PB was assessed in 233 patients (39 OM-CMML, 54 CMML, 23 MDS that did not meet OM-CMML diagnostic criteria, 15 MPN with $\geq 1 \times 10^9/L$ monocytes, and 102 with reactive monocytosis). The percentage of MO1s in these groups of patients is shown in Figure 4. As Selimoglu-Buet et al and other later studies have shown, the increase in MO1 fraction $>94\%$ is a very sensitive and specific predictor of CMML diagnosis.^{8,27,28} We explored the sensitivity and specificity of this method in our series. Because the minimum diagnostic criterion for considering the diagnosis of CMML is the presence of at least $1 \times 10^9/L$ monocytes in PB, we first analyzed the 171 patients in our

Table 1. Comparison of clinical findings between patients with OM-CMML and CMML

Characteristic	OM-CMML (n = 40)	CMML (n = 56)	d-CMML (n = 40)	P-CMML (n = 16)	P (OM-CMML vs CMML)	P (OM-CMML vs d-CMML)	P (OM-CMML vs P-CMML)	P (d-CMML vs P-CMML)
Age, median (range), y	77.5 (44-91)	76.5 (60-94)	76.5 (60-88)	77 (61-94)	.997	.95	.913	.835
Male sex, n (%)	27 (68)	29 (52)	21 (52.5)	8 (50)	.124	.171	.222	.866
Hemoglobin, median (range), g/L	11.35 (7.3-14.8)	12.3 (5.4-15.7)	12.55 (6.6-15.7)	11.85 (5.4-14.6)	.019	.005	.650	.198
WBC count, median (range), ×10 ⁹ /L	4.58 (2.49-7.02)	9.03 (2.81-34.82)	7.95 (2.81-12.62)	17.11 (13.18-34.82)	<.001	<.001	<.001	<.001
Neutrophil count, median (range), ×10 ⁹ /L	2.15 (0.38-4.42)	4.86 (0.47-19.62)	3.35 (0.47-7.02)	8.77 (4.67-19.62)	<.001	<.001	<.001	<.001
Platelet count, median (range), ×10 ⁹ /L	138.5 (96-318)	138.5 (15-559)	137 (33-559)	140 (15-294)	.565	.462	.986	.618
Monocyte count, median (range), ×10 ⁹ /L	0.71 (0.53-0.96)	1.96 (1-13.33)	1.67 (1-3.01)	5.02 (2.69-13.33)	<.001	<.001	<.001	<.001
PB monocyte %, median (range)	17 (10-26)	24.85 (13-60)	23.85 (13-42)	26.7 (20-60)	<.001	<.001	<.001	.002
BM monocyte %, median (range)	5 (1-13)	10 (2-27)	9 (2-16)	12 (2-27)	<.001	<.001	.001	.175
BM promonocyte %, median (range)	1 (0-6)	3 (0-14)	2 (0-7)	4 (1-14)	<.001	<.001	<.001	.012
BM blast %, median (range)	3 (0-8)	3 (1-15)	3 (1-9)	4 (1-15)	.838	.798	.338	.252
Dyserythropoiesis, median (range), %	31 (0-80)	23 (0-90)	22 (0-90)	24 (3-77)	.04	.026	.411	.252
≥10%, %	74	42.7	71.4	80	.027	.03	.329	.728
Dysgranulopoiesis, median (range), %	46 (0-100)	44 (0-92)	40 (0-92)	55 (9-78)	.917	.636	.530	.233
≥10%, %	87.2	88.5	86.5	93.3	1	.929	1	.659
Dys thrombopoiesis, median (range), %	13 (0-64)	8 (0-69)	7 (0-69)	16 (0-54)	.15	.063	.794	.142
≥10%, %	65.7	46.7	42.9	60	.089	.055	.726	.476
Karyotype, abnormal/total cases (%)	10/38 (26)	9/50 (18)	7/36 (19.4)	2/14 (14.3)	.498	.482	.475	.99
Spanish cytogenetic risk group, n (%)					.384	.559	.546	.809
Low	33 (87)	45 (90)	32 (88.9)	13 (92.9)				
Intermediate	3 (8)	1 (2)	1 (2.8)	0 (0)				
High	2 (5)	4 (8)	3 (8.3)	1 (7.1)				
Presence PB blasts, n (%)	2/40 (5)	8/50 (16)	3/36 (8.3)	5/15 (33.3)	.176	.663	.013	.039
2017 WHO classification, n (%)					.013	.121	.002	.023
CMML-0	22 (55)	13 (26)	12 (32.4)	1 (6.7)	.003	.078	.002	.078
CMML-1	14 (35)	29 (56)	21 (56.8)	8 (53.3)	.048	.055	.354	.822
CMML-2	4 (10)	10 (19)	4 (10.8)	6 (40)	.256	.99	.018	.024
Patients with an associated autoimmune disease, n (%) ^a	4 (10)	5 (9)	4 (10)	1 (6.3)	.99	.99	.99	.99
Number of patients with mutations (%)	40 (100)	52/53 (98)	36/37 (97.3)	16 (100)	.99	.481	.99	.99
Number of mutated genes, median (range)	2 (1-8)	3 (0-5)	2 (0-5)	4 (2-5)	.407	.46	.001	<.001
Number of mutations, median (range)	3 (1-9)	3 (0-9)	3 (0-9)	4.5 (2-7)	.134	.859	.002	.003

Bold P values are statistically significant.
WBC, white blood cell; Spanish cytogenetic risk group: low, normal, and isolated Y chromosome loss; intermediate, other abnormalities except those mentioned in low- and high-risk categories; and high, +8, abnormalities of chromosome 7 and complex karyotype (≥3 abnormalities).
^aFour patients with OM-CMML presented with a concomitant autoimmune disease: 1, systemic sclerosis; and 1, immune thrombocytopenic purpura. Five patients with CMML presented with a concomitant autoimmune disease: 1, systemic sclerosis; 1, ankylosing spondylitis HLA-B27⁺; 1, antiphospholipid syndrome; 1, immune thrombocytopenic purpura.

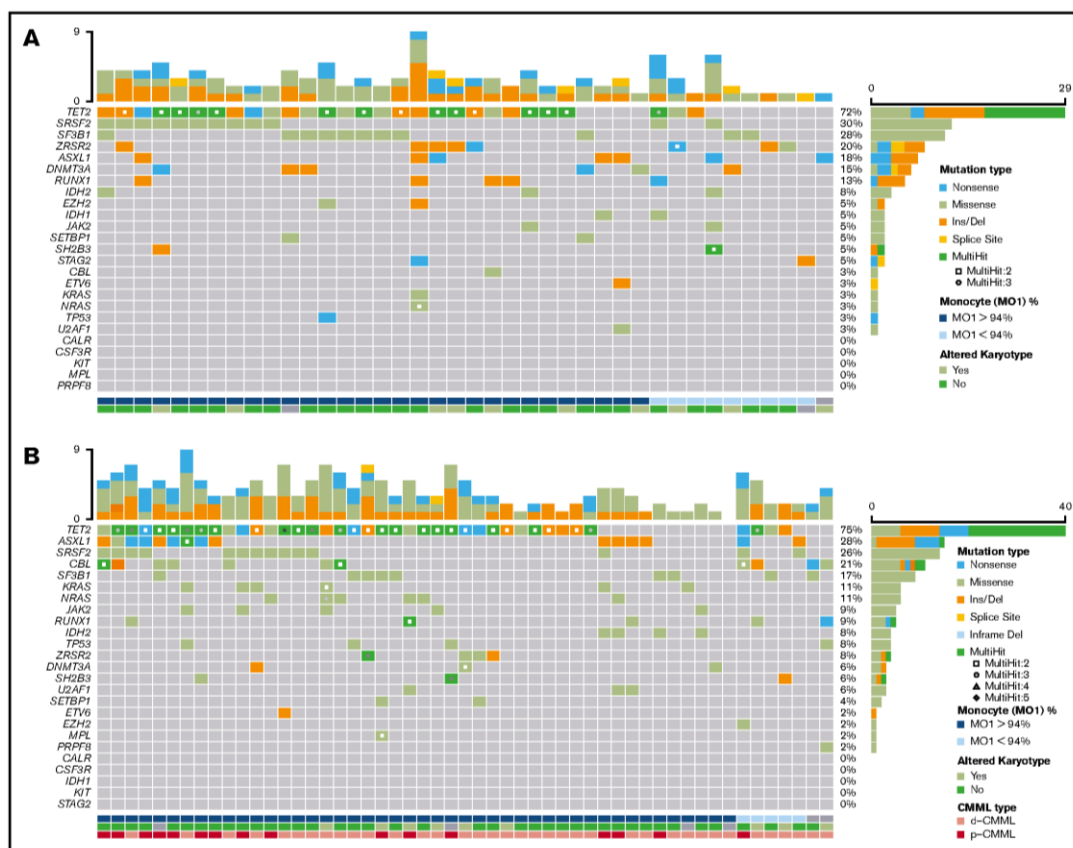


Figure 1. Mutational profile in patients with OM-CMML and CMML. Mutations were identified by NGS in 40 patients with OM-CMML (A) and in 53 patients with CMML (B). Results of the sequencing of the 25 genes are shown in the plot, where each column represents a patient and each row represents a gene. The number of mutations identified per patient is represented as columns in the top row. Genes are ordered from the most to the least frequently mutated, and frequencies for each gene are displayed (right), as well as the mutation type (nonsense, missense, insertion/deletion, splice site, or multihit). Patients with more than 1 mutation in the same gene are represented as shown in the key (2, 3, 4, or 5 mutations in the same gene). The immunophenotypic profile, assessed by the presence of MO1s upper 94%, is shown (bottom; MO1 >94%, blue, MO1 ≤94%, light blue, nonanalyzed, gray). Cytogenetic results are also displayed (bottom row; altered karyotype, lime green; normal karyotype, light green; nonanalyzed, gray). CMML types are also shown (bottom row: d-CMML, light red; p-CMML, red).

cohort with $\geq 1 \times 10^9/L$ monocytes (54 CMML and 15 MPN with monocytosis and 102 reactive monocytosis). The presence of MO1 percentage >94% predicted the diagnosis of CMML with a high sensitivity (90.7%) and specificity (92.2%). Because another group proposed MO1 percentage >95% as the best cutoff for predicting CMML diagnosis,²⁹ we assessed according to that criterion in our series. MO1 percentage >95% showed a lower sensitivity (83.3%) and a slightly better specificity (95.7%), and the balance between sensitivity and specificity calculated by the Youden index ($J = 79$) was worse than the 94% cutoff ($J = 82.9$). The AUC of the percentage of MO1 in our series was 0.941 (Figure 5), in line with the previous literature.^{8,27,28} Other authors have proposed the reduction of the percentage of MO3 as the best predictor for CMML diagnosis.³⁰ In our series, the MO1 population showed a better AUC than did the MO3 population (0.933). Moreover, we

found the cutoff in the percentage of MO3s under 3.18% to have the best predictive capacity in our series, but it performed worse than did the MO1 >94% cutoff (sensitivity, 92.6%; specificity, 83.8%; $J = 76.4$).

OM-CMML and overt CMML show similar immunophenotypic features

The comparison between OM-CMML and CMML showed that the MO1 percentage was significantly lower in OM-CMML, but it is noteworthy that the median and mean MO1 percentages in OM-CMML were above the 94% cutoff (median, 96.11 vs 97.96; mean, 94.76 vs 96.93; $P = .001$, OM-CMML vs CMML). Moreover, the proportion of patients with MO1 percentage >94% was not significantly different when OM-CMML was compared with CMML

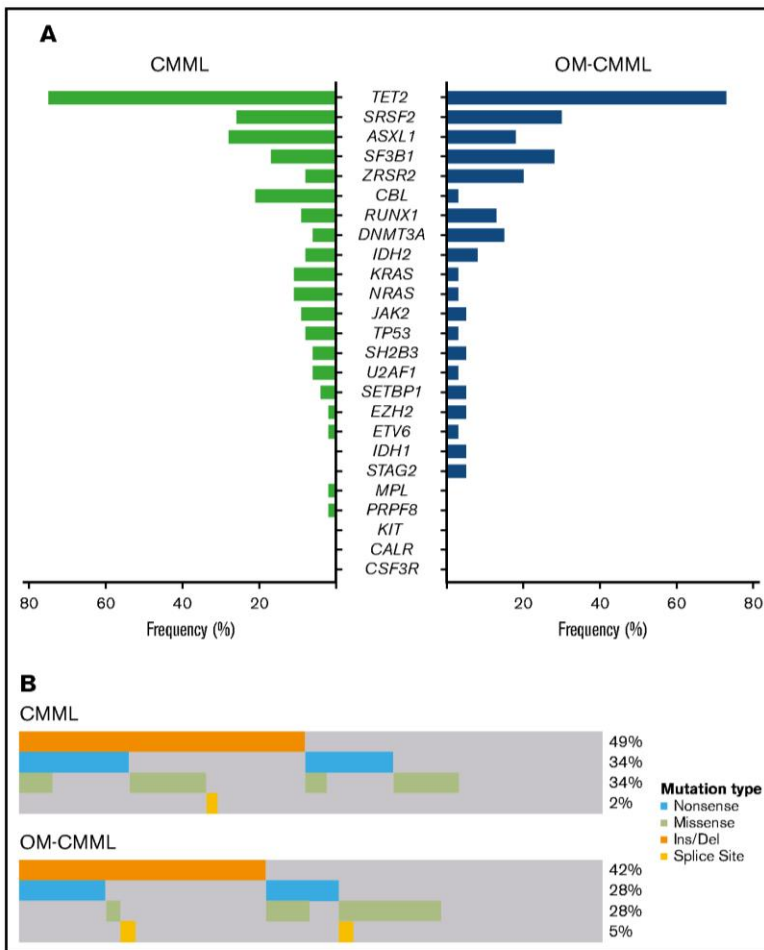


Figure 2. Distribution of mutated genes in CMML and OM-CMML. (A) Frequencies of the 25 genes analyzed by NGS in the CMML and OM-CMML groups. Genes are ordered from the most to the least frequently mutated, combining the CMML and OM-CMML cases. *CBL* was the only gene mutated at a significantly different frequency in the groups (2.5% vs 20.8%; OM-CMML vs CMML; $P = .011$). (B) The plot represents all the mutations identified in the *TET2* gene classified by the type of alteration, with insertions or deletions of nucleotides (orange) being the most frequent mutations identified. Nonsense mutations, producing a stop codon in the sequence (light blue), and missense mutations, producing a change in 1 amino acid (lime green), were the second most commonly identified. The least common alterations in our cohort were splice site mutations (yellow). No significant differences were observed in the distribution of mutations in *TET2* when both disease groups were compared.

(76.9% vs 90.7%; $P = .122$; Figure 4). Although probably achieved in the context of a type II error, this result is impressive because, as previously mentioned, the specificity of the MO1 percentage >94% test is ~90% to 95% and, therefore, only a 5% to 10% false-positive rate should be expected. However, in the group of patients with OM-CMML, a 76.9% false-positive rate was observed, because these patients had a current diagnosis of MDS according to the 2017 WHO recommendation. Likewise, no differences were observed in the percentage of patients showing CD56 positivity in monocytes (61.5% vs 63%; $P = .889$, OM-CMML vs CMML) or in the percentage of them showing CD2 (28.2% vs 35.2%; $P = .477$, OM-CMML vs CMML; supplemental Table 2). On the contrary, we found significant differences between patients with OM-CMML and those with MDS who did not meet OM-CMML diagnostic criteria in MO1 percentage (median: 96.11 vs 89.95; mean: 94.76 vs 89.01; $P < .001$, OM-CMML vs MDS), the proportion of patients with MO1 percentage >94% (76.9% vs 8.7%; $P < .001$, OM-CMML vs MDS), and the percentage of patients showing CD56 (61.5% vs

8.7%; $P < .001$, OM-CMML vs MDS) or CD2 (28.2% vs 0; $P = .005$, OM-CMML vs MDS; supplemental Table 2). No patient with OM-CMML, CMML, or MDS showed CD7 positivity. It is remarkable that we found no significant difference in the distribution of patients with OM-CMML and the comparator group of patients with MDS ($P = .433$), among the MDS categories stipulated by the WHO 2017 classification. Therefore, the differences detected between the OM-CMML and MDS groups are not attributable to their primary WHO 2017 classification.

Interestingly, a significantly higher proportion of patients with OM-CMML with *TET2* mutations had a MO1 percentage >94% (89.7% vs 40%, $P = .004$). Notably, patients with OM-CMML with *TET2* mutations demonstrated this feature in a percentage similar to overt CMML (89.7% vs 90.7%; $P = .99$). This mutation was the only one of the assessed mutations that enabled division of the OM-CMML series into 2 groups, which showed a significant difference in the proportion of patients with MO1 percentage >94% (supplemental Table 3).

RESULTS: ARTICLE 3 (OM-CMML IS A DISTINCTIVE SUBTYPE OF CMML)

Table 2. Distribution of somatic mutations in patients with OM-CMML, d-CMML, or p-CMML

	OM-CMML, n = 40, %	d-CMML, n = 37, %	p-CMML, n = 16, %	P (OM-CMML vs d-CMML)	P (d-CMML vs p-CMML)	P (OM-CMML vs p-CMML)
ASXL1	17.5	13.5	62.5	.757	.001	.003
CALR	—	—	—	—	—	—
CBL	2.5	16.2	31.3	.051	.275	.006
CSF3R	—	—	—	—	—	—
DNMT3A	15	5.4	6.3	.266	.99	.660
ETV6	2.5	2.7	—	.99	.99	.99
EZH2	5	—	6.3	.494	.302	.99
IDH1	5	—	—	.494	—	.99
IDH2	7.5	2.7	18.8	.616	.077	.338
JAK2	5	10.8	6.3	.419	.99	.99
KIT	—	—	—	—	—	—
KRAS	2.5	10.8	12.5	.189	.99	.193
MPL	—	—	6.3	—	.302	.286
NRAS	2.5	8.1	18.8	.346	.351	.066
RUNX1	12.5	10.8	6.3	.99	.99	.662
PRPF8	—	2.7	—	.481	.99	—
SETBP1	5	2.7	6.3	.99	.517	.99
SF3B1	27.5	16.2	18.8	.279	.99	.734
SH2B3	5	2.7	12.5	.99	.213	.570
SRSF2	30	18.9	43.8	.299	.123	.362
STAG2	5	—	—	.494	—	.99
TET2	72	73	81.3	.99	.731	.734
TP53	2.5	8.1	6.3	.346	.99	.494
U2AF1	2.5	2.7	12.5	.99	.213	.193
ZRSR2	20	10.8	—	.352	.303	.089
NRAS and/or KRAS	2.5*	16.2	31.3	.051	.275	.006
RAS pathway	5	27	62.5	.011	.014	<.001

Bold *P* values are statistically significant.

NRAS and/or KRAS, mutations in both genes or one of them; RAS pathway, mutations in CBL, NRAS, and/or KRAS genes.

*One patient showed concurrent NRAS and KRAS mutations (Figure 1A).

As published by Cargo et al,³¹ CD56 positivity in monocytes correlated positively with *TET2* mutation in our series, both as a binary value (Φ coefficient, 0.45; $P < .001$) or as a continuous variable (ρ Spearman, 0.4; $P < .001$). Likewise, the median expression of CD56 in monocytes was significantly higher in the patients with *TET2* mutations (34% vs 3%; $P < .001$), and the proportion of patients showing CD56 positivity was also higher in the *TET2*-mutated group (75% vs 24%; $P < .001$).

Given the similarities observed between patients with OM-CMML or CMML, we placed them together in a single category (93 recoded CMML: 39 OM-CMML and 54 overt CMML) and assessed the strength of the MO1 >94% method in all 233 patients of our series (93 recoded CMML, 23 MDS, 15 MPN with monocytosis, and 102 with reactive monocytosis). The presence of MO1 percentage >94% predicted the diagnosis of these patients with high sensitivity (84.9%) and specificity (92.1%; $J = 77$). The AUC of the percentage of MO1 was 0.908, and the best MO1 cutoff was >94% (Figure 5). Because a similar proportion of patients with OM-CMML or CMML showed CD56 and CD2 positivity and these

findings were rarely seen in the other groups of patients analyzed (supplemental Table 2), we tried to improve the performance of the method by using a combined approach: the presence of MO1 percentage >94% and CD56 positivity and/or CD2 positivity. The presence of at least 1 of these features presented a better sensitivity (94.6%) with a slightly lower specificity (87.8%), and the balance between sensitivity and specificity was clearly better ($J = 82.4$). The sensitivity of this approach when evaluating patients with OM-CMML was 89.7%, whereas the sensitivity in patients with CMML increased to 98.1%. Thus, given its sensitivity, this combined assay may be of high utility as a screening test in this context.

As previously reported by Tari et al,³² we observed a significantly higher false-negative rate of the MO1 >94% test in those patients with a concomitant autoimmune disease (44.4% vs 11.9%, patients with OM-CMML or CMML, analyzed together, with and without an associated autoimmune disease; $P = .027$).

Finally, we compared the percentage of plasmacytoid dendritic cells (pDCs) in PB from total leukocytes among the OM-CMML,

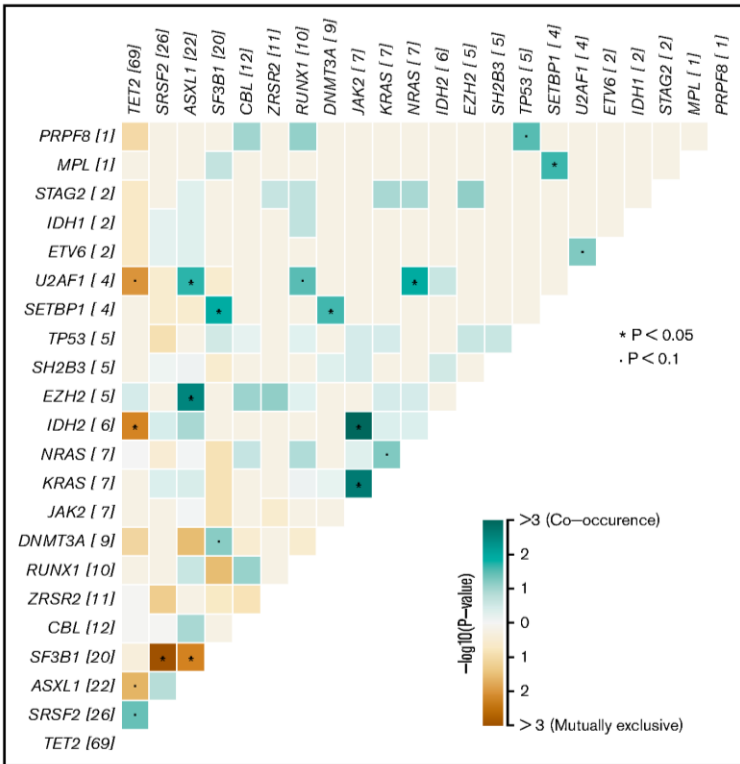


Figure 3. Co-occurrence or mutual exclusivity of genes in both OM-CMML and CMML. The plot shows, for the whole group of OM-CMML and CMML, all genes found to be altered in our cohort ordered by the number of mutations identified. In those genes that were frequently found to be mutated in the same patient, the interactions are depicted in lime green. In genes that were observed to be mutually exclusive and thus not frequently altered in the same patient, the interactions are depicted in brown. * $P < .05$; * $P < .1$.

d-CMML, and p-CMML groups (median, 0.05%, 0.04%, and 0.015%, respectively). We observed that p-CMML had a significantly lower percentage of pDCs than OM-CMML ($P = .022$). Likewise, we observed a trend when comparing OM-CMML with CMML as a whole group (median, 0.05% vs 0.02%; $P = .067$). In this regard, progression from low- to high-risk categories or even leukemic transformation in MDS patients has been associated with a progressive decrease in pDCs.^{33,34} These data enable us to infer that the transition of one stage to another may be partially favored by the progressive decline of pDCs, which would lead to a decrease in immune surveillance.

Patients with OM-CMML that evolved to CMML showed inferior survival

At a median follow-up of 31.1 months, 18% of patients with OM-CMML that evolved to CMML showed a median time to evolution of 34.3 months. The overall survival (OS) and cumulative incidence of evolution to CMML at 3 years of the 40 patients with OM-CMML were 85.9% and 15.7%, respectively. Seven patients with OM-CMML that evolved to CMML had a significantly shorter OS than did those in whom it did not evolve (median OS: not reached vs 64.62 months; $P = .026$). Patients in whom the disease evolved showed no significant differences regarding immunophenotypic or molecular patterns. In this regard, we did not find any variable showing a significant influence in predicting time to CMML (number of mutations, number of mutated genes, RAS-pathway mutations,

number of *TET2* mutations, truncating vs nontruncating type *TET2* mutations, and molecular CMML-specific prognostic scoring system). Notably, 4 of 7 patients with OM-CMML that evolved to CMML died, showing a very short median OS from the moment of progression (3.42 months; 95% CI, 0.6-6.2). One patient progressed to acute myeloid leukemia and the other 3 patients died of severe infections. Although this finding deserves to be taken into consideration, larger series of patients are needed before generating warnings in this area.

Discussion

We analyzed the clinicopathologic features of the largest series of patients with OM-CMML reported to date, with extensively studied clinical, morphological, cytogenetic, molecular, and immunophenotypic data. In addition, we compared the features of these patients to those of a large series of patients with CMML, with data concerning immunophenotypic characteristics of OM-CMML being especially novel. In this sense, we compared the utility of the MO1 >94% test between these 2 groups of patients and collected a large series of patients with MPN with absolute monocytosis and reactive monocytosis and a subset of patients with MDS who did not fulfill OM-CMML diagnostic criteria. Notably, we report one of the largest published series to assess the MO1 >94% criterion, in either the total number of patients assessed ($n = 233$) or in the number of patients with CMML analyzed ($n = 54$). The increase of MO1 >94% provided high sensitivity (90.7%) and specificity

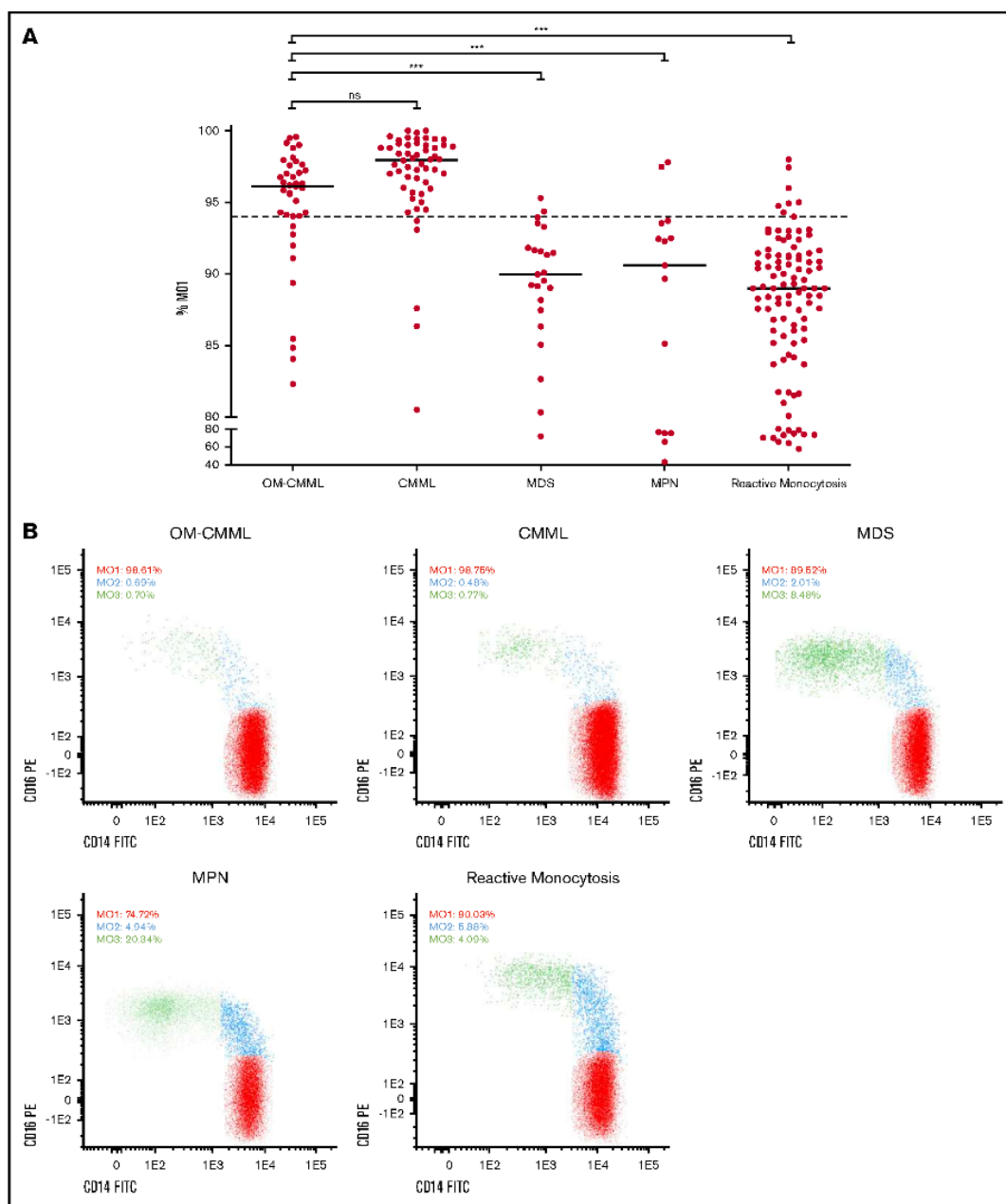


Figure 4. Percentage of MO1s in the 233 cases grouped by disease. (A) Flow cytometry results are shown as the percentage of MO1 identified for the 5 groups analyzed (39 OM-CMML, 54 CMML, and 23 MDS that did not meet OM-CMML diagnostic criteria; 15 MPN with $\geq 1 \times 10^9/L$ monocytes, and 102 reactive monocytosis). Each dot represents a patient result for the MO1 test, and lines represent the median percentage for each disease group. The dotted line indicates the 94% cutoff of the test. *** $P < .001$. (B) The distribution of monocyte subsets is shown for 5 examples, each one from a different disease group. The percentage of the 3 monocyte subsets (MO1, MO2, and MO3) out of the total monocytes is displayed for each example. ns, not significant.

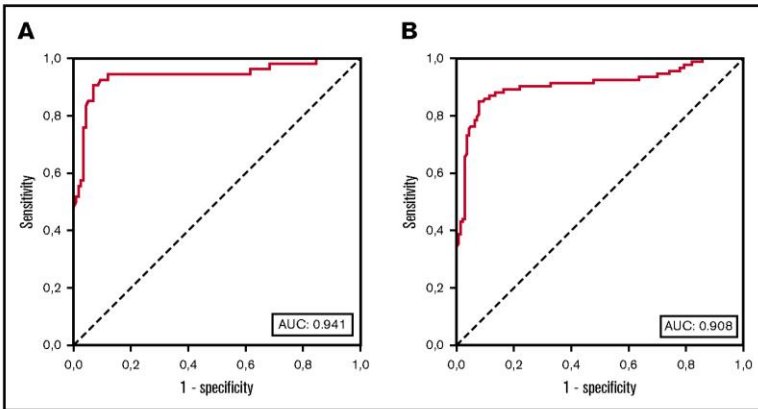


Figure 5. ROC AUC curves of the percentage of MO1s in our series. (A) ROC curve analysis of diagnostic sensitivity and specificity of the MO1 percentage in 171 patients with $\geq 1 \times 10^9/L$ PB monocytes (54 CMML, 15 MPN with monocytosis, and 102 with reactive monocytosis). (B) ROC curve analysis of diagnostic sensitivity and specificity of the MO1 percentage in PB monocytes of 233 patients (93 recoded CMML, including 39 OM-CMML and 54 overt CMML; 23 MDS not meeting OM-CMML diagnostic criteria; 15 MPN with monocytosis; and 102 with reactive monocytosis).

(92.2%) for CMML diagnosis in our series. Although the 94% threshold was initially validated by 2 studies,^{27,28} some controversial results have recently appeared in the literature. Picot et al²⁹ detected the 95% cutoff as the one with the best sensitivity and specificity, and later, Hudson et al³⁰ found that the MO3 percentage $< 1.13\%$ was the best predictor of a diagnosis of CMML. Although valuable, these studies were based on a small number of patients with CMML (15 in Picot et al and 16 in Hudson et al). In addition, the different series in the literature assessing the performance of the MO1 $> 94\%$ criterion are not well studied from a molecular point of view.^{8,10,27-30} In contrast, molecular characterization by targeted NGS was performed in all patients with OM-CMML and in 53 of 56 patients with CMML in our series. The lack of molecular data could diminish the accuracy of the results of the MO1 $> 94\%$ test because, as previously stated, some uncertainty may exist when establishing a CMML diagnosis in some cases (eg, absence of dysmyelopoiesis, absence of clonality assessed by cytogenetics, and coexistence of autoimmune or neoplastic diseases). In our series, the best cutoff in MO1 percentage was $> 94\%$ and the MO1 population showed the best predictive capacity for the diagnosis of CMML, validating the results of the French group.^{8,27}

Focusing on the comparison between OM-CMML and CMML, we found no significant differences in the proportion of patients with MO1 percentage $> 94\%$ or in those who showed CD56 or CD2 positivity in monocytes. Based on this, we tried to improve the performance of the MO1 $> 94\%$ method by using a combined approach: the presence of a percentage of MO1s $> 94\%$ and CD56 and/or CD2 positivity in monocytes. This method afforded better sensitivity (94.6%) with slightly lower specificity (87.8%) than the MO1 $> 94\%$ cutoff, and the balance between sensitivity and specificity was clearly superior. Thus, given its high sensitivity, this combined assay emerged as an excellent screening test in this context. Interestingly, as previously reported by Tarfi et al,³² we observed a significantly higher false-negative rate of the MO1 $> 94\%$ test in those patients with a concomitant autoimmune disease. They showed that a decrease in the 6-sulfo lac-nac (slan) + MO3 monocytes below 1.7% is characteristic of CMML and persists in those exhibiting an associated inflammatory state.³² Therefore, in future studies, it would be interesting to dispose of

the anti-slan antibody, to further improve the precision of the method.

OM-CMML and overt CMML showed similar clinical, morphological, and cytogenetic features, with the exception that patients with OM-CMML showed lower hemoglobin levels and more evident dyserythropoiesis. This finding could be partially explained by a higher proportion of patients with OM-CMML showing *SF3B1* mutation and $\geq 5\%$ ring sideroblasts in BM. In our series, patients with OM-CMML or CMML displaying this feature showed a significantly lower hemoglobin level and a higher median percentage of dyserythropoiesis.

OM-CMML and overt CMML show a similar mutational profile. We found no significant difference in the proportion of patients showing concurrent *TET2* and *SRSF2* mutations, the well-accepted gene signature of CMML.¹⁵ As previously shown, the impairment of the hydroxymethylation pathway (mutations in *IDH1*, *IDH2*, and/or *TET2* genes) is present in most of these patients. Moreover, in line with previous data,³⁵ a high proportion of patients with OM-CMML or CMML showed multiple *TET2* mutations. Interestingly, patients with OM-CMML who had *TET2* mutations had MO1 percentage $> 94\%$ in a rate similar to those with overt CMML. Moreover, as previously reported by Cargo et al,³¹ CD56 positivity in monocytes was significantly associated with *TET2* mutation in our series. These findings suggest that the impairment of this pathway could be the pathophysiological hallmark of these entities. Hydroxymethylation has been recognized as a physiological passive DNA demethylation process.³⁶⁻³⁸ Therefore, it is expected that patients with OM-CMML or CMML will present aberrant DNA hypermethylation states mediated by an ineffective production of 5-hydroxymethylcytosines.^{20,39,40} In future studies, it would be interesting to explore the implication of DNA methylation, and especially 5-hydroxymethylcytosine levels, in prognosis, disease progression, and prediction of response to hypomethylating agents,²¹ in this group of patients.

The only gene mutated at a significantly lower frequency when comparing OM-CMML with CMML was *CBL*. This finding was expected, because mutations in genes of the RAS pathway (*CBL*, *NRAS*, *KRAS*, *NF1*, and *PTPN11*) are well-known secondary events in CMML and have been associated with proliferative

features.^{16,17} This finding is in line with that of Geyer et al,⁹ who found a significantly lower proportion of patients with OM-CMML with *CBL* mutations (0% vs 28%; OM-CMML vs CMML). It is also remarkable that a significantly higher proportion of patients with p-CMML carried *ASXL1* mutations when compared to those with OM-CMML and d-CMML. This result agrees with the published data showing that patients with CMML harboring *ASXL1* mutations have more prominent leukocytosis than the group not displaying this mutation.¹⁴

Finally, in our series, 18% of patients had OM-CMML that evolved to CMML. This observation supports considering OM-CMML as an early stage of d-CMML.⁹ If true, it would allow for the establishment of a continuum of OM-CMML, d-CMML, and p-CMML. In this sense, as previously mentioned, we inferred that second genetic hits, such as the acquisition of RAS-pathway mutations, could promote the transition from one stage to another.¹⁶ In addition, immune dysregulation, together with a progressive decrease in immune surveillance, could play a pivotal role in the progression of the disease. In this regard, as previously reported by other researchers,^{33,34} we found a decline in pDCs when comparing OM-CMML with CMML, and this was especially evident when OM-CMML was compared to p-CMML. Interestingly, in our series, the patients with OM-CMML that evolved to CMML showed a significantly shorter overall survival than did those in whom it did not evolve.

In summary, OM-CMML and overt CMML show similar clinical, morphological, cytogenetic, molecular, and immunophenotypic features. In addition, the MO1 percentage >94% method showed a high accuracy for predicting CMML and OM-CMML diagnosis in

our series. The results reinforce the consideration of OM-CMML as a distinctive subtype of CMML.

Acknowledgments

This study was supported in part by grants from Instituto de Salud Carlos III (ISCIII) and Spanish Ministry of Health grants FIS PI16/0153, FIS PI19/0005, and FEDER 2017SGR205, 2017SGR437, and PT17/0015/0011; Beca Gilead 2016; and Xarxa de Banc de Tumors de Catalunya.

Authorship

Contribution: X.C. designed the study; X.C., I.P., N.G.-G., L.A., M.A.-C., B.M., S.G.-A., D.R.-B., B.B., L.C., and L.F. collected and assembled data from the study patients; X.C., N.G.-G., L.A., C.F.-R., J.G., M.S., A.P., B.E., S.M., and A.F. analyzed the data; X.C., N.G.-G., L.A., and J.G. interpreted the data; X.C. wrote the final version of the manuscript; and all authors reviewed and approved the final version of the manuscript.

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

ORCID profiles: X.C., 0000-0001-7934-9130; N.G.-G., 0000-0002-8185-9786; J.G., 0000-0002-0742-0759; L.F., 0000-0001-5906-3311; M.A.-C., 0000-0003-1637-7112; B.M., 0000-0002-6733-6006; M.S., 0000-0001-9988-5977; A.P., 0000-0001-9627-4978; B.E., 0000-0002-4294-8145; L.C., 0000-0001-5236-5085; B.B., 0000-0002-5335-2726; A.F., 0000-0002-3381-9472; L.A., 0000-0002-8176-7179.

Correspondence: Xavier Calvo, Paseo Marítimo, 25, 08003 Barcelona, Spain; e-mail: xcalvo@parcdesalutmar.cat.

References

1. Swerdlow SH, Campo E, Harris NL, et al. WHO classification of tumours of haematopoietic and lymphoid tissues. 4th ed, revised. Lyon, France: International Agency for Research on Cancer, WHO; 2017.
2. Arber DA, Orazi A, Hasserjian R, et al. The 2016 revision to the World Health Organization classification of myeloid neoplasms and acute leukemia. *Blood*. 2016;127(20):2391-2405.
3. Valent P, Orazi A, Savona MR, et al. Proposed diagnostic criteria for classical chronic myelomonocytic leukemia (CMML), CMML variants and pre-CMML conditions. *Haematologica*. 2019;104(10):1935-1949.
4. Meggendorfer M, Haferlach T, Alpermann T, et al. Specific molecular mutation patterns delineate chronic neutrophilic leukemia, atypical chronic myeloid leukemia, and chronic myelomonocytic leukemia. *Haematologica*. 2014;99(12):e244-e246.
5. Itzykson R, Kosmider O, Renneville A, et al. Prognostic score including gene mutations in chronic myelomonocytic leukemia. *J Clin Oncol*. 2013;31(19):2428-2436.
6. Elena C, Galli A, Such E, et al. Integrating clinical features and genetic lesions in the risk assessment of patients with chronic myelomonocytic leukemia. *Blood*. 2016;128(10):1408-1417.
7. Patnaik MM, Tefferi A. Cytogenetic and molecular abnormalities in chronic myelomonocytic leukemia. *Blood Cancer J*. 2016;6(2):e393.
8. Selimoglu-Buet D, Wagner-Ballon O, Saada V, et al; Francophone Myelodysplasia Group. Characteristic repartition of monocyte subsets as a diagnostic signature of chronic myelomonocytic leukemia. *Blood*. 2015;125(23):3618-3626.
9. Geyer JT, Tam W, Liu Y-C, et al. Oligomonocytic chronic myelomonocytic leukemia (chronic myelomonocytic leukemia without absolute monocytosis) displays a similar clinicopathologic and mutational profile to classical chronic myelomonocytic leukemia. *Mod Pathol*. 2017;30(9):1213-1222.
10. Selimoglu-Buet D, Badaoui B, Benayoun E, et al; Groupe Francophone des Myélodysplasies. Accumulation of classical monocytes defines a subgroup of MDS that frequently evolves into CMML. *Blood*. 2017;130(6):832-835.
11. Palomo L, Ibáñez M, Abáigar M, et al; Spanish Group of MDS (GESMD). Spanish Guidelines for the use of targeted deep sequencing in myelodysplastic syndromes and chronic myelomonocytic leukaemia. *Br J Haematol*. 2020;188(5):605-622.
12. Kalina T, Flores-Montero J, van der Velden VHJ, et al; EuroFlow Consortium (EU-FP6, LSHB-CT-2006-018708). EuroFlow standardization of flow cytometer instrument settings and immunophenotyping protocols. *Leukemia*. 2012;26(9):1986-2010.

13. Such E, Cervera J, Costa D, et al. Cytogenetic risk stratification in chronic myelomonocytic leukemia. *Haematologica*. 2011;96(3):375-383.
14. Patnaik MM, Itzykson R, Lasho TL, et al. ASXL1 and SETBP1 mutations and their prognostic contribution in chronic myelomonocytic leukemia: a two-center study of 466 patients. *Leukemia*. 2014;28(11):2206-2212.
15. Cazzola M, Della Porta MG, Malcovati L. The genetic basis of myelodysplasia and its clinical relevance. *Blood*. 2013;122(25):4021-4034.
16. Itzykson R, Solary E. An evolutionary perspective on chronic myelomonocytic leukemia. *Leukemia*. 2013;27(7):1441-1450.
17. Itzykson R, Kosmider O, Renneville A, et al. Clonal architecture of chronic myelomonocytic leukemias. *Blood*. 2013;121(12):2186-2198.
18. Patnaik MM, Padron E, LaBorde RR, et al. Mayo prognostic model for WHO-defined chronic myelomonocytic leukemia: ASXL1 and spliceosome component mutations and outcomes. *Leukemia*. 2013;27(7):1504-1510.
19. Such E, Germing U, Malcovati L, et al. Development and validation of a prognostic scoring system for patients with chronic myelomonocytic leukemia. *Blood*. 2013;121(15):3005-3015.
20. Figueroa ME, Abdel-Wahab O, Lu C, et al. Leukemic IDH1 and IDH2 mutations result in a hypermethylation phenotype, disrupt TET2 function, and impair hematopoietic differentiation. *Cancer Cell*. 2010;18(6):553-567.
21. Meldi K, Qin T, Buchi F, et al. Specific molecular signatures predict decitabine response in chronic myelomonocytic leukemia. *J Clin Invest*. 2015;125(5):1857-1872.
22. Bejar R. Myelodysplastic syndromes diagnosis: what is the role of molecular testing? *Curr Hematol Malig Rep*. 2015;10(3):282-291.
23. Bejar R, Steensma DP. Recent developments in myelodysplastic syndromes. *Blood*. 2014;124(18):2793-2804.
24. Papaemmanuil E, Gerstung M, Malcovati L, et al; Chronic Myeloid Disorders Working Group of the International Cancer Genome Consortium. Clinical and biological implications of driver mutations in myelodysplastic syndromes. *Blood*. 2013;122(22):3616-3627, quiz 3699.
25. Haferlach T, Nagata Y, Grossmann V, et al. Landscape of genetic lesions in 944 patients with myelodysplastic syndromes. *Leukemia*. 2014;28(2):241-247.
26. Kurtovic-Kozaric A, Przychodzen B, Singh J, et al. PRPF8 defects cause missplicing in myeloid malignancies. *Leukemia*. 2015;29(1):126-136.
27. Tari S, Harrivel V, Dumezy F, et al; Groupe Francophone des Myélodysplasies (GFM). Multicenter validation of the flow measurement of classical monocyte fraction for chronic myelomonocytic leukemia diagnosis. *Blood Cancer J*. 2018;8(11):114.
28. Talati C, Zhang L, Shaheen G, et al. Monocyte subset analysis accurately distinguishes CMML from MDS and is associated with a favorable MDS prognosis. *Blood*. 2017;129(13):1881-1883.
29. Picot T, Aanei CM, Flandrin Gresta P, et al. Evaluation by flow cytometry of mature monocyte subpopulations for the diagnosis and follow-up of chronic myelomonocytic leukemia. *Front Oncol*. 2018;8:109.
30. Hudson CA, Burack WR, Leary PC, Bennett JM. Clinical utility of classical and nonclassical monocyte percentage in the diagnosis of chronic myelomonocytic leukemia. *Am J Clin Pathol*. 2018;150(4):293-302.
31. Cargo C, Cullen M, Taylor J, et al. The use of targeted sequencing and flow cytometry to identify patients with a clinically significant monocytosis. *Blood*. 2019;133(12):1325-1334.
32. Tari S, Badaoui B, Freynet N, et al; Groupe Francophone des Myélodysplasies (GFM). Disappearance of slan-positive non-classical monocytes for diagnosis of chronic myelomonocytic leukemia with associated inflammatory state. *Haematologica*. 2019;105(4):e147-e152.
33. Saft L, Björklund E, Berg E, Hellström-Lindberg E, Porwit A. Bone marrow dendritic cells are reduced in patients with high-risk myelodysplastic syndromes. *Leuk Res*. 2013;37(3):266-273.
34. Kerkhoff N, Bontkes HJ, Westers TM, de Grijijl TD, Kordasti S, van de Loosdrecht AA. Dendritic cells in myelodysplastic syndromes: from pathogenesis to immunotherapy. *Immunotherapy*. 2013;5(6):621-637.
35. Patnaik MM, Zahid MF, Lasho TL, et al. Number and type of TET2 mutations in chronic myelomonocytic leukemia and their clinical relevance. *Blood Cancer J*. 2016;6(9):e472-e472.
36. Tahiliani M, Koh KP, Shen Y, et al. Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science*. 2009;324(5929):930-935.
37. Ito S, D'Alessio AC, Taranova OV, Hong K, Sowers LC, Zhang Y. Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. *Nature*. 2010;466(7310):1129-1133.
38. Guo JU, Su Y, Zhong C, Ming GL, Song H. Emerging roles of TET proteins and 5-hydroxymethylcytosines in active DNA demethylation and beyond. *Cell Cycle*. 2011;10(16):2662-2668.
39. Liu X, Zhang G, Yi Y, et al. Decreased 5-hydroxymethylcytosine levels are associated with TET2 mutation and unfavorable overall survival in myelodysplastic syndromes. *Leuk Lymphoma*. 2013;54(11):2466-2473.
40. 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):3448-3458.

Article 4

Analysis of saliva samples and cluster of differentiation 3 (CD3)+ lymphocytes as a source of germline DNA in myeloproliferative neoplasms

Garcia-Gisbert N*, Camacho L*, Fernández-Ibarrondo L, Fernández-Rodríguez C, Longarón R, Gibert J, Angona A, Andrade-Campos M, Salar A, Besses C, Bellosillo B

* equally contributed

British Journal of Haematology, 2020; 189(5):e204–7.

DOI: 10.1111/bjh.16624

Analysis of saliva samples and cluster of differentiation 3 (CD3)+ lymphocytes as a source of germline DNA in myeloproliferative neoplasms

The widespread use of genetic studies in patients with myeloproliferative neoplasms (MPN) by next-generation sequencing (NGS) has led to the identification of new genetic variants. To establish the potential pathogenic role of these

variants with scarce or inexistent literature, it is important to assess whether their origin is somatic or germline,¹ as germline variants in myeloid-related genes have been associated with hereditary disorders^{2,3} or germline predisposition, while

e204

© 2020 British Society for Haematology and John Wiley & Sons Ltd
British Journal of Haematology, 2020, **189**, e194–e221

somatic variants are associated with clonal expansion of haematopoietic stem cells not only in MPNs, but also in most types of myeloid malignancies.

In this context, germline DNA is required to perform these analyses, which is frequently obtained from several sources, e.g. saliva, buccal mucosa, cluster of differentiation 3 (CD3)⁺ lymphocytes, fibroblasts or hair. The aim of the present study was to ascertain if DNA obtained from saliva samples and CD3⁺ lymphocytes from peripheral blood is a suitable source of germline DNA for molecular studies of patients with Philadelphia chromosome (Ph)-negative MPNs.

Paired samples of saliva and granulocytes from 191 patients with MPNs [76 polycythaemia vera (PV), 109 thrombocythaemia (ET), five primary myelofibrosis (PMF), one unclassifiable MPN] harbouring mutations in Janus kinase 2 (*JAK2*; $n = 159$), calreticulin (*CALR*; $n = 26$) and *MPL* proto-oncogene, thrombopoietin receptor (*MPL*; $n = 6$) were analysed. Saliva DNA was obtained using the Oragene-DNA kit (DNA Genotek). *JAK2V617F* was assessed by allele-specific real-time polymerase chain reaction (PCR), *CALR* mutations were determined by fragment analysis electrophoresis and *MPL* mutations by digital PCR. Methods and materials are detailed in Data S1.

Analysis of DNA from saliva showed that in 69/191 (36.1%) cases, neither the wild type nor the mutated form of the *JAK2*, *CALR* or *MPL* genes were detected, although a large amount of DNA was present in the sample. The absence of amplifiable human DNA suggested that it could come from the microorganisms present in the oral mucosa, and was confirmed by the amplification of the 16s ribosomal gene. Thus, these 69 samples were considered as uninformative.

In the remaining 122 (63.9%) saliva DNA samples, driver gene mutations were detected in 89/122 cases (73%) and a strong correlation between the variant allele frequency (VAF) of the mutation in granulocytes and saliva was observed ($r = 0.706$, $P < 0.001$, Spearman) (Fig 1).

Of note, in 50 of these 122 saliva samples the VAF detected was >30%, meaning that in these cases (41% of the cohort) the use of saliva as a control for germline mutations, would result in considering acquired variants as germline.

To get further insight into the origin of this mutated DNA present in saliva, we hypothesised that it might be cell-free tumoral DNA (ctDNA), as it has been described that saliva is a possible source of ctDNA. We assessed the size of the DNA fragments by capillary electrophoresis and found that most DNA fragments corresponded to high molecular weight DNA with no evidence of fragments around 166 base pairs, indicating that the isolated DNA was not ctDNA. Next, we analysed the relation with the leucocyte count. Patients with positive saliva had higher leucocyte counts in peripheral blood [median (range) 8.48 (2.71–34.8) $\times 10^9$ cells/l in positive saliva vs 6.25 (2.59–21.75) $\times 10^9$ cells/l in negative saliva; $P = 0.001$, Mann–Whitney], supporting the hypothesis that the positivity for the driver mutation in saliva samples may be due to leucocyte presence in the oral mucosa.

Previous studies have described that saliva DNA from patients with MPN can be positive for *JAK2V617F* mutation.^{4,5} Moreover, a strong correlation has also been described between the VAF of *JAK2V617F* in granulocytes and saliva in a limited number of patients.⁶ Our present study confirms these observations in a larger cohort and includes for the first time cases with driver mutations in *CALR* and *MPL*, indicating that this effect would be affecting all patients with MPN. This is of utmost importance as broad gene panels are being performed in triple-negative patients, in which the detection of new variants should be carefully evaluated in germline DNA.

Searching for an alternative sample type for germline studies, we isolated CD3⁺ lymphocytes by immunomagnetic selection in 64 *JAK2V617F*-positive patients. In all cases *JAK2* was amplifiable, and only one sample showed a high *JAK2V617F* VAF of 27.3% [median (range) VAF 2.39 (0.01–27.3)%]. Previous studies described that CD3⁺ lymphocytes are negative or present a very low VAF for *JAK2V617F*,⁵ but in some exceptional cases *JAK2V617F* is detected with a high VAF, suggesting homozygosity in the lymphoid compartment. It is possible that in these positive cases, the *JAK2* mutation appears at an earlier stage of the haematopoiesis, affecting several blood lineages.⁷

Additionally, CD3⁺ cells from 12 *CALR*- and five *MPL*-mutated cases were isolated. For the *CALR*-mutated cases, the VAF in CD3⁺ was higher than that observed in the *JAK2V617F*-positive cases [median (range) VAF 11.2 (3.66–28.4)%], whereas for the *MPL*-mutated cases the VAF in CD3⁺ was lower [median (range) VAF 1.45 (0.43–3.54)%]

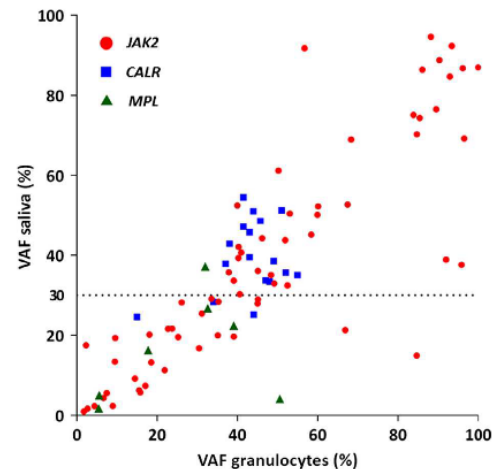


Fig 1. Scatter plot of the 89 cases with positive driver detection in saliva showing the correlation between the variant allele frequency (VAF) of the mutation in saliva and granulocytes. In cases above the line at 30% VAF ($n = 50$) the use of saliva as a germline control could result in considering acquired variants as germline.

Correspondence

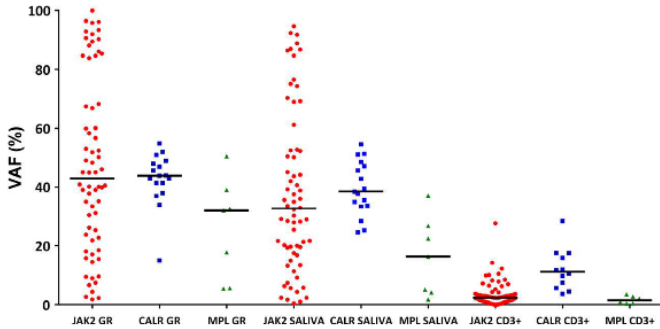


Fig 2. VAF in granulocytes (GR), saliva and CD3+ lymphocytes for *JAK2*, *CALR* and *MPL* mutated cases.

(Fig 2). The higher VAF in the CD3+ fraction of *CALR*-mutated cases than in *JAK2*- or *MPL*-mutated cases could be indicative of an earlier origin of these mutations in the haematopoietic lineage. *CALR* mutations have been described to originate in the haematopoietic stem cell⁸ and it has been previously postulated that *CALR* mutations could be an early event in the development of the disease. Mutations in *CALR* are generally present in the whole granulocytic cell compartment with a VAF of around 50%, suggesting a greater proliferative advantage of the mutated progenitor *CALR* clone compared to the *JAK2*-mutated patients.⁹ In favour of this hypothesis, *CALR* mutations are acquired before additional mutations in other non-driver genes, while *JAK2* mutations can be preceded by first hits in other genes such as ten-eleven translocation 2 (*TET2*), DNA methyltransferase 3 alpha (*DNMT3A*) or additional sex combs like-1 (*ASXL1*).^{8,10}

As a whole, although we detected some positive cases for the somatic driver mutation in the CD3+ fraction, the presence of the mutated clone was significantly lower [median (range) VAF 2.83 (0.00–28.37)%] than in the saliva samples [median (range) VAF 21.7 (0.00–94.64)%; $P < 0.001$, Wilcoxon].

In conclusion, in patients with MPN with somatic mutations, saliva samples are not a reliable source of germline control DNA. Moreover, in an important set of cases in our present cohort (36%), no human DNA was detected in the saliva samples. Therefore, the use of CD3+ lymphocytes is a better option than saliva for the study of germline DNA variants in Ph-negative MPNs.

Acknowledgments

This study was supported in part by grants from El Instituto de Salud Carlos III (ISCIII) and Spanish Ministry of Health, PI16/0153, 2017SGR205, PI17/0015/0011. Beca Gilcad 2016 and Xarxa de Banc de Tumors de Catalunya. We thank Roger Anglada and Xavier Calvo for technical support.


Author contributions

Nieves Garcia-Gisbert, Laura Camacho, Lierni Fernández-Ibarrondo and Raquel Longarón performed the research.

Concepcion Fernández-Rodríguez, Carlos Besses and Beatriz Bellosillo designed the research study. Joan Gibert, Anna Angona and Marcio Andrade-Campos contributed essential reagents or tools. Nieves Garcia-Gisbert, Laura Camacho and Joan Gibert analysed the data. Nieves Garcia-Gisbert, Concepcion Fernández-Rodríguez, Antonio Salar, Carlos Besses and Beatriz Bellosillo wrote the paper.

Conflict of interest

Antonio Salar received honoraria for speaker, consultancy or advisory role from Roche, Janssen Pharmaceuticals, Gilead and Celgene. Carlos Besses received honoraria for speaker, consultancy or advisory role from Gilead. Beatriz Bellosillo received honoraria for speaker, consultancy or advisory role from Astra-Zeneca, Biocartis, Merck-Serono, Novartis, Qiagen, Hoffman-La Roche, ThermoFisher, Pfizer and BMS.

Nieves Garcia-Gisbert^{1,2,*} 

Laura Camacho^{1,3,*}

Lierni Fernández-Ibarrondo¹

Concepcion Fernández-Rodríguez^{1,3}

Raquel Longarón^{1,3}

Joan Gibert¹

Anna Angona^{1,4}

Marcio Andrade-Campos^{1,4}

Antonio Salar^{1,4}

Carlos Besses^{1,4}

Beatriz Bellosillo^{1,2,3}

¹Grup de Recerca Clínica, Aplicada en Neoplàsies Hematològiques-Hospital del Mar IMIM, ²Pompeu Fabra University, ³Department of Pathology, Hospital del Mar-IMIM and ⁴Department of Hematology, Hospital del Mar-IMIM, Barcelona, Spain.

E-mail: bbelosillo@parcdesahutmar.cat

*equally contributed

First published online 31 March 2020

doi: 10.1111/bjh.16624

e206

© 2020 British Society for Haematology and John Wiley & Sons Ltd
British Journal of Haematology, 2020, **189**, e194–e221

Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Data S1. Supplementary Material 1 – Methods.

References

- Li MM, Datto M, Duncavage EJ, Kulkarni S, Lindeman NI, Roy S, et al. Standards and guidelines for the interpretation and reporting of sequence variants in cancer: a joint consensus recommendation of the association for molecular pathology, American society of clinical oncology, and college of American Pathologists. *J Mol Diagn.* 2017;**19**:4–23.
- Bellanné-Chantelot C, Mosca M, Marty C, Favier R, Vainchenker W, Plo I. Identification of MPL R102P mutation in hereditary thrombocytosis. *Front Endocrinol.* 2017;**8**:1–4.
- Marty C, Saint-Martin C, Pecquet C, Grosjean S, Saliba J, Mouton C, et al. Germ-line JAK2 mutations in the kinase domain are responsible for hereditary thrombocytosis and are resistant to JAK2 and HSP90 inhibitors. *Blood.* 2014;**123**:1372–83.
- Barnholt KE, Hinds DA, Kiefer AK, Do CB, Eriksson N, Mountain JL, et al. Estimation of JAK2 V617F prevalence by detection of the mutation in saliva samples from online MPN and general population cohorts. *Blood.* 2012;**120**:1737.
- Tenedini E, Bernardis I, Artusi V, Artuso L, Roncaglia E, Guglielmelli P, et al. Targeted cancer exome sequencing reveals recurrent mutations in myeloproliferative neoplasms. *Leukemia.* 2013;**28**:1052–9.
- Hinds DA, Barnholt KE, Mesa RA, Kiefer AK, Do CB, Eriksson N, et al. Germ line variants predispose to both JAK2 V617F clonal hematopoiesis and myeloproliferative neoplasms. *Blood.* 2016;**128**:1121–8.
- Larsen TS, Christensen JH, Hasselbalch HC, Pallisgaard N. The JAK2 V617F mutation involves B- and T-lymphocyte lineages in a subgroup of patients with Philadelphia-chromosome negative chronic myeloproliferative disorders. *Br J Haematol.* 2007;**136**:745–51.
- Nangalia J, Massie CE, Baxter EJ, Nice FL, Gundem G, Wedge DC, et al. Somatic CALR mutations in myeloproliferative neoplasms with nonmutated JAK2. *N Engl J Med.* 2013;**369**:2391–405.
- Rumi E, Pietra D, Ferretti V, Klampfl T, Harutyunyan AS, Milosevic JD, et al. JAK2 or CALR mutation status defines subtypes of essential thrombocythemia with substantially different clinical course and outcomes. *Blood.* 2014;**123**:1544–51.
- Klampfl T, Gisslinger H, Harutyunyan AS, Nivarthi H, Rumi E, Milosevic JD, et al. Somatic mutations of calreticulin in myeloproliferative neoplasms. *N Engl J Med.* 2013;**369**:2379–90.

DISCUSSION

Myeloid neoplasms are a heterogeneous group of diseases that present some common features such as aberrant hematopoiesis and presence of genetic alterations. In the last decades, molecular alterations have been implemented in MPN, CMML and MDS as part of their diagnostic criteria, and are also useful for disease classification and prognosis. Nevertheless, molecular diagnosis still presents some limitations that need to be overcome to ensure accurate genetic characterization. In this context, the analysis of cfDNA is a fast and easy strategy that may potentially improve molecular assessment in routine clinical practice. This will lead to a better understanding of the pathogenesis of these diseases and facilitate the analysis of each patient in detail; which is essential to optimize diagnostic groups and define new entities such as OM-CMML.

In the first article, we analyzed the molecular profile of MPN patients in plasma cfDNA and compared the results to peripheral blood granulocytes. In the second article assesses the molecular and cytogenetic profile of MDS patients in cfDNA and compared the result to bone marrow DNA. In the third article, we analyzed the molecular and immunophenotypic features of a well-annotated series of OM-CMML and compared the result to CMML. In the fourth article, we evaluated the presence of MPN driver mutations in saliva samples and CD3+ lymphocytes from peripheral blood.

The first part of this thesis is focused on the role of liquid biopsy in myeloproliferative and myelodysplastic diseases. Overall, we found that cfDNA mirrors the genetic landscape present in tumoral cells of both disease types. In MPN patients, we compared the molecular alterations in cfDNA and in paired granulocytes, which are considered the gold standard, and we found an equivalent mutational profile in both driver and non-driver genes. In MDS patients, we compared the genomic profile of cfDNA with BM DNA, including the analysis of cytogenetic alterations in addition to molecular profiling. We found a highly concordant genomic profile, although this concordance was lower for subclonal alterations.

In MPNs, this was the first published report describing the mutational profile of a well-annotated series of MPN patients using liquid biopsy. Previous

analysis of plasma or serum demonstrated that *JAK2* mutations were detectable in plasma or serum samples (241–244). In 2006, *JAK2* mutations were detected in plasma of MPN patients for the first time(241). This study was later expanded and found that plasma DNA presented a higher proportion of mutated DNA in comparison to paired peripheral blood cells(242). These results were later discussed by Salama *et al.* who argued that *JAK2* mutated DNA in plasma was an artifact of cell lysis(250). They concluded that the presence of *JAK2* mutation in plasma was due to the progressive lysis of granulocytes during PB storage and did not recommend plasma for *JAK2* analysis. However, they found that this artifact was produced after 2-11 days of PB storage at room temperature(250). In our study, we have processed PB samples in the first 4 hours to isolate plasma and granulocytes; and found that *JAK2* mutations were detected with a slightly higher VAF in plasma cfDNA than in granulocytes (medians of 29.0% vs 25.2%, respectively). Moreover, we performed a quality control by electrophoresis to ensure that pure cfDNA was obtained (approximately 166 base pairs). In our plasma samples, we did not find cellular DNA fragments produced by cell lysis, which present higher molecular weight.

For *MPL* mutations in MPN patients, we found that cfDNA presented a significantly higher VAF than granulocyte DNA (medians of 44.3% vs 22.5%, respectively) in all *MPL* mutated patients. This difference was higher than that observed for *JAK2* mutations. We confirmed *MPL* VAFs by high sensitive digital PCR in all *MPL* mutated cases. To try to help understand what was occurring with *MPL* gene, we analyzed the representation of the genomic regions covered by the NGS panel in cfDNA and granulocyte DNA. It has been described that cfDNA fragments result from nucleosome fragmentation that does not occur randomly. In fact, it has been reported that specific fragmentation profiles are different between cancer patients and healthy individuals(251). Fragmentation patterns of healthy patients indicated that cfDNA is mostly of hematological origin. No fragmentation profiles have been described yet for hematological malignancies, but we hypothesized that this fragmentation pattern may facilitate the detection of *MPL* mutations, when compared to cellular DNA. In this regard, we observed a higher number of

reads covering the *MPL* gene in cfDNA than in granulocyte DNA, unlike in other genes. Previous results of our group described that platelet RNA presented a higher proportion of *JAK2* mutation than granulocytes(252,253), however, cfDNA analysis is faster and easier to implement in routine practice than the use of platelets. On the other hand, in 2021, Sadeh *et al.* performed CHIP-seq of plasma cell-free DNA nucleosomes and found that nucleosomes footprint indicated a major megakaryocytic origin of cfDNA(231). Megakaryocyte maturation process is mediated by numerous DNA replication rounds without cell division, resulting in polyploidy cells (2N to 128N)(254). Mutated megakaryocytes in MPNs may be the main contributors to cfDNA in plasma, releasing a higher proportion of mutated DNA due to their polyploidy. This is in line with our results and could explain the high presence of *JAK2/MPL* mutations in cfDNA and platelets. Of note, in one patient with PMF we detected two mutations, in *ASXL1* and *MPL* genes, that were not detected by NGS in DNA from PB granulocytes. To confirm these findings, we performed digital PCR. This finding indicates that cfDNA analysis may provide higher sensitivity detection of driver and non-driver mutations, especially in *MPL*-mutated cases.

In this context, we found interesting similarities with MPNs when we analyzed the genomic profile of cfDNA in MDS patients. Although *MPL* mutations are infrequent in MDS, we found three *MPL* mutations in our MDS cohort, and all cases presented higher VAF in cfDNA than in BM DNA (median VAF of 26.9% in cfDNA and 4.5% in BM DNA) which is concordant with our results in MPN patients. In MDS, one of the main findings was that the VAF of *SF3B1* mutations were significantly higher in cfDNA than in BM DNA. This difference has not been previously reported, which may be due to the lower number of *SF3B1* mutated patients analyzed by other groups(237–239). In the MPN cohort, we found *SF3B1* mutations in 4 patients, and 3/4 presented a higher VAF in cfDNA than in granulocytes (median VAF of 22% in cfDNA and 13% in granulocyte DNA), which is concordant with MDS results. Additionally, for the MDS patients, we confirmed that it is possible to detect cytogenetic alterations by NGS in cfDNA. Yeh *et al.* found that, in one MDS patient, the loss of chromosome 9 was detectable by low-coverage whole-genome

sequencing of cfDNA, however, this was the only reported case in the literature. We observed that cfDNA revealed the gain or loss of genetic material in the clonal MDS cells. Genomic regions affected by a deletion were less represented in cfDNA while genomic regions in chromosomes affected by a trisomy were more represented in cfDNA. This finding supports that using a single NGS test it is possible to detect both molecular and cytogenetic alterations in cfDNA, although we observed sensitivity limitations in comparison to karyotype or FISH. On the contrary, in MPN patients, we used a targeted NGS panel that was not designed to detect cytogenetic alterations. Despite this, the detection of cytogenetic and molecular alterations using a single non-invasive test, like we did in cfDNA from MDS, would be of special interest in MPN. Obtaining sufficient analyzable metaphases by conventional cytogenetics may be difficult in MPN, especially in PMF cases, in which cytogenetic alterations are prognostically relevant. In particular, Grinfeld *et al.* classification of MPN included the following cytogenetic alterations: aberrations at chromosomes 17p and 5q in the highest risk group, LOH at chromosome 4q and 7q/7 in the second risk group, and deletion at chromosome 20q in the same group as *CALR* mutated cases. In this context, the inclusion of these genomic regions in a NGS panel and its application using cfDNA in MPN patients may be a useful non-invasive approach for accurate risk classification.

Moreover, we found that in MPN and MDS patients, the VAF of mutations in cfDNA were equal or even higher than VAF in cellular DNA, indicating that a high proportion of tumoral circulating DNA was present in plasma samples. Contrary to solid tumors, in which tumoral circulating DNA is a minority in total cfDNA, in myeloid neoplasms mutations were identified with higher VAF. These results are concordant with the recently published results about the hematopoietic origin of cfDNA, especially released by immature myeloid cells. In consequence, one of the main limitations of liquid biopsy analysis in solid tumors and lymphomas is that tumoral cfDNA may represent a small fraction of the total cfDNA(255). Therefore, accurate liquid biopsy studies require the application of high sensitivity techniques to detect mutations. In our

experience, liquid biopsy analysis in myeloid neoplasms is easier due to the high percentage of mutated cfDNA present in plasma.

We observed that the concentration of cfDNA in plasma was higher in MPN and MDS patients than in plasma samples from healthy patients (control group). Moreover, in MPN, PMF patients presented higher cfDNA concentration than PV and ET cases. These findings are in accordance with the results reported by Marin-Oyarzún *et al.*, who evaluated the presence of circulating nucleosomes in MPNs by enzyme linked immunosorbent assay (ELISA)(256). They found that MPNs presented a higher presence of circulating nucleosomes than healthy controls, and PMF presented higher nucleosome concentration than PV and ET. Interestingly, we found that cfDNA concentration was higher in cases with greater molecular complexity. This correlation was still observed in the PMF group. In contrast, for MDS patients we did not find any correlation between cfDNA concentration and IPSS-R risk groups, neither with molecular complexity. Suzuki *et al.* found that higher risk MDS patients presented a higher cfDNA concentration in plasma than lower risk cases(257), an observation that was not replicated in our cohort probably due to the low proportion of higher risk cases. Of note, we found a significant correlation between lactate dehydrogenase (LDH) levels and total cfDNA concentration in both the MDS and the MPN group. Accordingly, high levels of LDH and high cfDNA concentrations in serum or plasma are both indicators of cell damage(162,257–259).

CfDNA analysis in myeloid neoplasms emerges as a non-invasive strategy to characterize the molecular profile of early hematopoietic cells. Overall, the results obtained from cfDNA analysis in MPN and MDS support the application of liquid biopsy analysis as an accurate, fast and easy method for routine testing. Moreover, our data, together with previous studies(238,260), demonstrated that it is possible to monitor molecular dynamics in patients receiving treatment using cfDNA analysis in MPN and MDS patients. We hope that liquid biopsy analysis is implemented in clinical routine practice of MDS and MPN as a helpful tool to characterize and monitor genetic alterations, in the same way as in other hematological malignancies such as lymphomas. From a translational perspective, cfDNA analysis is used in lymphomas as a

real-time minimally invasive approach to evaluate clonal evolution and track treatment response(261–267).

On the other hand, we analyzed the largest cohort of OM-CMML reported to date with well annotated clinical, morphological, cytogenetic, molecular and immunophenotypic data. OM-CMML and overt CMML shared overall similar characteristics, which supports that those patients fulfilling the criteria for OM-CMML are mostly being misclassified as MDS.

From the molecular point of view, OM-CMML and CMML showed a similar mutational profile. The *TET2* and *SRSF2* co-mutation was found equally in both groups (27.5% in OM-CMML vs 22.6% in CMML). This co-mutation has been described as the genetic signature for the development of CMML disease(188,190,191). Therefore, those *TET2* and *SRSF2* mutated cases that do fulfill the diagnostic criteria for OM-CMML showed strong biological evidence to support their consideration as a distinctive subtype of CMML.

We found that *CBL* was the only gene mutated with a significantly different frequency in CMML and OM-CMML. This is in line with previous observations made by Geyer *et al.*, who did not find any patient with *CBL* mutations in their OM-CMML cohort(202). This could be explained by the fact that mutations in the RAS pathway have been associated to the proliferative subtype of CMML and are later acquired in disease evolution(52,191). We believe that OM-CMML is a preliminary step of d-CMML and that is the reason why RAS mutations are infrequent in OM-CMML. Other groups have reported their results of cases with OM-CMML. Montalban-Bravo *et al.* described the clinical and biological characteristics of 30 patients who met the criteria for OM-CMML(268). RAS pathway mutations (*NRAS*, *KRAS*, *CBL*, *NF1*, *SETBP1*, *PTPN11*) were found in a significantly higher proportion in CMML and were more likely to appear as minor clones in patients with OM-CMML compared to CMML.

Genes involved in DNA hydroxymethylation pathways (*TET2*, *IDH1*, *IDH2*) were affected in 83% of CMML and 75% of OM-CMML. We observed that both OM-CMML and CMML frequently presented multiple *TET2* mutations in the same patient. We are currently further exploring these observations as we

believe that multiple *TET2* mutations in the same patient could be other genetic signature of these entities, similar to the co-mutation of *TET2* and *SRSF2*. Clonal cells with both alleles affected by *TET2* mutations would produce the homozygotic loss of TET2 function severely affecting hydroxymethylation.

We analyzed the utility of the monocyte distribution immunophenotypic test in a cohort including OM-CMML, CMML, MDS, MPN with relative monocytosis and cases with reactive monocytosis. We found that the 94% of MO1 threshold was better predictor for CMML diagnosis, consistent with previous reports(269,270). In our series, the presence of >94% of MO1 provided high sensitivity and specificity (90.7% and 92.2%, respectively) for CMML diagnosis. Additionally, we reported for the first time the immunophenotypic profile of cases with OM-CMML. Most OM-CMML patients (76.9%) presented more than 94% of MO1 monocytes, with a median of 96.11% of MO1 monocytes. In contrast, only 8.7% of MDS patients (not fulfilling OM-CMML criteria) presented more than 94% of MO1 monocytes. This is of outmost importance as the immunophenotypic results, as well as the molecular profile, reinforce the idea of OM-CMML being more similar to CMML than to MDS.

Overall, these results are in accordance with findings reported by other studies, however, we described for the first time a well-annotated series from the immunophenotypic and molecular point of view, which allowed us to find some interesting associations between genotype and immunophenotype. We found that CD56 expression was associated with the presence of *TET2* mutations. In addition, OM-CMML cases with *TET2* mutations presented a similar percentage of patients with >94% of MO1 monocytes to those with overt CMML, which suggest that *TET2* mutations may drive to monocyte expansion and differentiation in CMML and OM-CMML. Focusing on the comparison between OM-CMML and CMML, we found no significant differences in the proportion of patients with MO1 percentage >94% or in CD56 expression. In accordance with these results, CD2 positivity in monocytes was identified in a similar proportion in OM-CMML and CMML.

The clinical outcomes of the OM-CMML cohort are currently being analyzed by our group. We believe that OM-CMML is the first step in the proliferative continuum of CMML (OM-CMML, d-CMML and p-CMML), which is supported by the molecular alterations (RAS pathway mutations) acquired in the transformation from OM-CMML to d-CMML, and from d-CMML to p-CMML. In line with these results, we observed a reduction of plasmacytoid dendritic cells (pDC) in CMML when compared to OM-CMML. Previous studies have reported that a progressive decrease in pDCs in MDS is associated to progression to higher-risk categories or AML(271,272). Likewise, immune dysregulation and decrease of pDC in OM-CMML could be implicated in transition to d-CMML, and in a later stage to p-CMML.

Overall, we hope that, in the future, the presence or absence of genetic alterations and the immunophenotypic profile play a stronger role in diagnosis, classification and prognosis of these patients. Disease classifications should become less dependent of restrictive thresholds that, although are required for categorization, frequently leave some of these patients in diagnostic groups that do not correspond to their biological reality, such as the case of OM-CMML.

Finally, in the last publication included in this thesis we analyzed the presence of driver mutations in saliva samples and CD3+ lymphocytes from peripheral blood in a large cohort of MPN patients. These sample types are generally considered as non-tumoral. The use of saliva samples DNA as a source of non-tumoral cells is extended, as they are broadly thought to mainly contain epithelial cells from the mouth mucosa. CD3+ cells are also thought to be non-tumoral in MPN patients, due to the clear myeloid differentiated profile of MPN tumoral cells. We analyzed saliva samples from 191 MPN patients to determine if the DNA obtained from this samples was free of contamination with tumoral DNA. In a high proportion of saliva samples (36%) human housekeeping genes were undetectable while bacterial 16S ribosomal DNA was amplifiable, therefore, these samples were considered as non-informative. In the remaining samples containing human DNA, *JAK2*, *CARL* or *MPL* mutations were frequently detected and showed a VAF higher than 30%,

indicating that clonal cells are abundant in saliva samples. In line with these results, several studies described that saliva DNA from patients with MPN can be positive for *JAK2V617F* mutation (273–275). In fact, in 2012 a large group of MPN patients were tested for the *JAK2V617F* mutation using saliva DNA as a non-invasive source of DNA, and *JAK2* mutations were detected with high specificity and sensitivity(273). Later studies found that saliva samples were contaminated by myeloid cells (274–276). In this context, these results suggest that saliva samples in MDS and CMML are probably contaminated with clonal cells similarly to MPN patients, although no specific studies have been done. Therefore, the use of saliva DNA as germline control should be avoided in myeloid malignancies as it could result in considering acquired variants as germline; and saliva DNA could be more useful as non-invasive approach for molecular characterization although further studies are required. Of note, saliva analysis in healthy individuals also showed the presence of lymphoid cells (276,277), hence it is probable that saliva samples are also contaminated with tumoral cells in lymphoid malignancies, an issue that should be explored in future research.

We isolated PB CD3+ lymphocytes by immunomagnetic selection in MPN cases to find an alternative sample type for germline studies. Overall, low VAF of driver mutations was detected in CD3+ cells that could help to accurately discriminate between acquired and germline mutation. The isolation of CD3+ cells is a better option than saliva DNA, however, in a minority of cases we found high VAF of the driver mutation. These observations can be explained because some mutations occur in early progenitors that will lead to both myeloid and lymphoid mutated cells. We found that mutations in *CALR* occurred earlier and therefore were represented with higher VAF in CD3+ cells. It has also been reported that in some cases, such as patients with *TET2* mutations, genetic alterations can affect both myeloid and lymphoid lineages (278). This is an important limitation as the identification of mutations in CD3+ cells could lead to considering somatic variants as germline in these patients with mutations in common progenitors. Historically, CD3+ cells had been a target of analysis to identify clonality in MPN female patients based on X-chromosome inactivation. Analysis of the human androgen receptor

gene (HUMARA) compares the methylation status of granulocytes (or peripheral blood mononuclear cells [PBMCs]) and control cells, such as CD3+ cells(279). This technique allows clonality detection in females with MPN, however since the identification of other markers of clonality, mainly somatic mutations, HUMARA assay has gradually lost its clinical utility.

Other alternatives have been explored to obtain germline DNA, such as fibroblasts(278,280,281). Cultured skin fibroblasts are a reliable option to isolate DNA free from hematologic contamination, however, obtention of sufficient cells in the culture biopsy is time-consuming and requires cell culture equipment and experience(280).

The identification of a reliable germline sample type is important, as NGS is evolving to analyze wider regions of the genome and therefore a higher number of alterations with unknown significance are identified. To clarify the role of these variants in the disease, a proper germline control is essential and may have clinical implications for the patient and his family. In this context, further studies are required to determine and standardize the better option for germline DNA testing.

CONCLUSIONS

1. The analysis of cfDNA allows the characterization of the molecular abnormalities of patients with MPN. The sensitivity and accuracy for mutation detection in driver and non-driver genes were equal or even superior to that obtained when studying the isolated granulocytic population, especially regarding the detection of MPL mutations.
2. The analysis of cfDNA allows the detection and monitoring of molecular and cytogenetic abnormalities of patients with MDS. In our cohort, enriched with low risk IPSS-R MDS patients, cytogenetic alterations were detectable in most cases by NGS in both BM DNA and cfDNA.
3. OM-CMML and CMML presented similar molecular, immunophenotypic and clinical features, which supports the consideration of OM-CMML as a distinct subtype of CMML.
4. Saliva samples are not a reliable source of germline DNA in MPNs. The use of CD3+ lymphocytes is a better option than saliva for the study of germline DNA variants in MPN patients.

BIBLIOGRAPHY

1. Arber DA, Orazi A, Hasserjian R, Thiele J, Borowitz MJ, Le Beau MM, et al. The 2016 revision to the World Health Organization classification of myeloid neoplasms and acute leukemia. *Blood*. 2016 May 19;127(20):2391–405.
2. Parganas E, Wang D, Stravopodis D, Topham DJ, Marine JC, Teglund S, et al. Jak2 is essential for signaling through a variety of cytokine receptors. *Cell*. 1998 May 1;93(3):385–95.
3. Witthuhn BA, Quelle FW, Silvennoinen O, Yi T, Tang B, Miura O, et al. JAK2 associates with the erythropoietin receptor and is tyrosine phosphorylated and activated following stimulation with erythropoietin. *Cell*. 1993 Jul 30;74(2):227–36.
4. Remy I, Wilson IA, Michnick SW. Erythropoietin receptor activation by a ligand-induced conformation change. *Science*. 1999 Feb 12;283(5404):990–3.
5. Shuai K, Ziemiecki A, Wilks AF, Harpur AG, Sadowski HB, Gilman MZ, et al. Polypeptide signalling to the nucleus through tyrosine phosphorylation of Jak and Stat proteins. *Nat* 1993 3666455. 1993;366(6455):580–3.
6. O’Shea JJ, Schwartz DM, Villarino A V., Gadina M, McInnes IB, Laurence A. The JAK-STAT pathway: impact on human disease and therapeutic intervention. *Annu Rev Med*. 2015 Jan 14;66:311–28.
7. James C, Ugo V, Le Couédic J-P, Staerk J, Delhommeau F, Lacout C, et al. A unique clonal JAK2 mutation leading to constitutive signalling causes polycythaemia vera. *Nature*. 2005 Apr 28;434(7037):1144–8.
8. Baxter EJ, Scott LM, Campbell PJ, East C, Fourouclas N, Swanton S, et al. Acquired mutation of the tyrosine kinase JAK2 in human myeloproliferative disorders. *Lancet*. 2005 Mar;365(9464):1054–61.
9. Kralovics R, Passamonti F, Buser AS, Teo S-S, Tiedt R, Passweg JR, et al. A gain-of-function mutation of JAK2 in myeloproliferative disorders. *N Engl J Med*. 2005 Apr 28;352(17):1779–90.
10. Levine RL, Wadleigh M, Cools J, Ebert BL, Wernig G, Huntly BJP, et al. Activating mutation in the tyrosine kinase JAK2 in polycythemia vera, essential thrombocythemia, and myeloid metaplasia with myelofibrosis. *Cancer Cell*. 2005 Apr;7(4):387–97.

BIBLIOGRAPHY

11. Zhao R, Xing S, Li Z, Fu X, Li Q, Krantz SB, et al. Identification of an Acquired JAK2 Mutation in Polycythemia Vera. *J Biol Chem*. 2005 Jun 17;280(24):22788–92.
12. Scott LM, Tong W, Levine RL, Scott MA, Beer PA, Stratton MR, et al. JAK2 Exon 12 Mutations in Polycythemia Vera and Idiopathic Erythrocytosis. *N Engl J Med*. 2007 Feb;356(5):459–68.
13. Krause K-H, Michalak M. Calreticulin. *Cell*. 1997 Feb 21;88(4):439–43.
14. Wang WA, Groenendyk J, Michalak M. Calreticulin signaling in health and disease. *Int J Biochem Cell Biol*. 2012 Jun 1;44(6):842–6.
15. Nangalia J, Massie CEE, Baxter EJJ, Nice FLL, Gundem G, Wedge DCC, et al. Somatic CALR Mutations in Myeloproliferative Neoplasms with Nonmutated JAK2. *N Engl J Med*. 2013 Dec 19;369(25):2391–405.
16. Klampfl T, Gisslinger H, Harutyunyan AS, Nivarthi H, Rumi E, Milosevic JD, et al. Somatic Mutations of Calreticulin in Myeloproliferative Neoplasms. *N Engl J Med*. 2013 Dec 19;369(25):2379–90.
17. Lau WWY, Hannah R, Green AR, Göttgens B. The JAK-STAT signaling pathway is differentially activated in CALR-positive compared with JAK2V617F-positive ET patients. *Blood*. 2015;125(10):1679.
18. Lim K-H, Chang Y-C, Chiang Y-H, Lin H-C, Chang C-Y, Lin C-S, et al. Expression of CALR mutants causes mpl-dependent thrombocytosis in zebrafish. *Blood Cancer J* 2016 6:10. 2016 Oct 7;6(10):e481–e481.
19. Pikman Y, Lee BH, Mercher T, McDowell E, Ebert BL, Gozo M, et al. MPLW515L is a novel somatic activating mutation in myelofibrosis with myeloid metaplasia. *PLoS Med*. 2006 Jul;3(7):e270.
20. Gurney AL, Wong SC, Henzel WJ, Sauvage FJ de. Distinct regions of c-Mpl cytoplasmic domain are coupled to the JAK-STAT signal transduction pathway and Shc phosphorylation. *Proc Natl Acad Sci*. 1995 Jun 6;92(12):5292–6.
21. Ding J, Komatsu H, Iida S, Yano H, Kusumoto S, Inagaki A, et al. The Asn505 mutation of the c-MPL gene, which causes familial essential thrombocythemia, induces autonomous homodimerization of the c-Mpl protein due to strong amino acid polarity. *Blood*. 2009 Oct 8;114(15):3325–8.
22. Chaligné R, James C, Tonetti C, Besancenot R, Le Couédic JP, Fava F, et

- al. Evidence for MPL W515L/K mutations in hematopoietic stem cells in primitive myelofibrosis. *Blood*. 2007 Nov 15;110(10):3735–43.
23. Guijarro-hernández A, Vizmanos JL. A Broad Overview of Signaling in Ph-Negative Classic Myeloproliferative Neoplasms. *Cancers* 2021, Vol 13, Page 984. 2021 Feb 26;13(5):984.
24. Oh ST, Simonds EF, Jones C, Hale MB, Goltsev Y, Gibbs KD, et al. Novel mutations in the inhibitory adaptor protein LNK drive JAK-STAT signaling in patients with myeloproliferative neoplasms. *Blood*. 2010 Aug 12;116(6):988–92.
25. Maslah N, Cassinat B, Verger E, Kiladjian J-J, Velazquez L. The role of LNK/SH2B3 genetic alterations in myeloproliferative neoplasms and other hematological disorders. *Leuk* 2017 318. 2017 May 9;31(8):1661–70.
26. Tyner JW, Erickson H, Deininger MWN, Willis SG, Eide CA, Levine RL, et al. High-throughput sequencing screen reveals novel, transforming RAS mutations in myeloid leukemia patients. *Blood*. 2009 Feb 19;113(8):1749–55.
27. Ricci C, Fermo E, Corti S, Molteni M, Faricciotti A, Cortelezzi A, et al. RAS Mutations Contribute to Evolution of Chronic Myelomonocytic Leukemia to the Proliferative Variant. *Clin Cancer Res*. 2010 Apr 15;16(8):2246–56.
28. Papaemmanuil E, Gerstung M, Malcovati L, Tauro S, Gundem G, Van Loo P, et al. Clinical and biological implications of driver mutations in myelodysplastic syndromes. *Blood*. 2013 Nov 21;122(22):3616–27; quiz 3699.
29. Haferlach T, Nagata Y, Grossmann V, Okuno Y, Bacher U, Nagae G, et al. Landscape of genetic lesions in 944 patients with myelodysplastic syndromes. *Leukemia*. 2014 Feb 13;28(2):241–7.
30. Ogawa S. Genetics of MDS. *Blood*. 2019 Mar 7;133(10):1049–59.
31. Tefferi A, Lasho TL, Guglielmelli P, Finke CM, Rotunno G, Elala Y, et al. Targeted deep sequencing in polycythemia vera and essential thrombocythemia. *Blood Adv*. 2016 Nov 29;1(1):21.
32. Tefferi A, Lasho TL, Finke CM, Elala Y, Hanson CA, Ketterling RP, et al. Targeted deep sequencing in primary myelofibrosis. *Blood Adv*. 2016

- Dec 13;1(2):105–11.
33. Carcavilla A, Suárez-Ortega L, Sánchez AR, Gonzalez-Casado I, Ramón-Krauel M, Labarta JI, et al. Noonan syndrome: genetic and clinical update and treatment options. *An Pediatría (English Ed.* 2020 Jul 1;93(1):61.e1-61.e14.
 34. Ihle NT, Byers LA, Kim ES, Saintigny P, Lee JJ, Blumenschein GR, et al. Effect of KRAS Oncogene Substitutions on Protein Behavior: Implications for Signaling and Clinical Outcome. *JNCI J Natl Cancer Inst.* 2012 Feb 8;104(3):228–39.
 35. Buradkar A, Bezerra E, Coltro G, Lasho TL, Finke CM, Gangat N, et al. Landscape of RAS pathway mutations in patients with myelodysplastic syndrome/myeloproliferative neoplasm overlap syndromes: a study of 461 molecularly annotated patients. *Leuk* 2020 352. 2020 Jun 8;35(2):644–9.
 36. Parikh C, Subrahmanyam R, Ren R. Oncogenic NRAS rapidly and efficiently induces CMML- and AML-like diseases in mice. *Blood.* 2006 Oct 1;108(7):2349–57.
 37. Caligiuri MA, Briesewitz R, Yu J, Wang L, Wei M, Arnoczky KJ, et al. Novel c-CBL and CBL-b ubiquitin ligase mutations in human acute myeloid leukemia. *Blood.* 2007 Aug 1;110(3):1022–4.
 38. Makishima H, Cazzolli H, Szpurka H, Dunbar A, Tiu R, Huh J, et al. Mutations of E3 Ubiquitin Ligase Cbl Family Members Constitute a Novel Common Pathogenic Lesion in Myeloid Malignancies. *J Clin Oncol.* 2009 Dec 20;27(36):6109–16.
 39. Sanada M, Suzuki T, Shih L-Y, Otsu M, Kato M, Yamazaki S, et al. Gain-of-function of mutated C-CBL tumour suppressor in myeloid neoplasms. *Nature.* 2009 Aug 20;460(7257):904–8.
 40. Levkowitz G, Waterman H, Zamir E, Kam Z, Oved S, Langdon WY, et al. c-Cbl/Sli-1 regulates endocytic sorting and ubiquitination of the epidermal growth factor receptor. *Genes Dev.* 1998 Dec 1;12(23):3663–74.
 41. Wang L, Rudert WA, Loutaev I, Roginskaya V, Corey SJ. Repression of c-Cbl leads to enhanced G-CSF Jak-STAT signaling without increased cell proliferation. *Oncogene.* 2002 Aug 2;21(34):5346–55.

42. Ueno H, Sasaki K, Honda H, Nakamoto T, Yamagata T, Miyagawa K, et al. c-Cbl Is Tyrosine-Phosphorylated by Interleukin-4 and Enhances Mitogenic and Survival Signals of Interleukin-4 Receptor by Linking With the Phosphatidylinositol 3'-Kinase Pathway. *Blood*. 1998 Jan 1;91(1):46–53.
43. Ballester R, Marchuk D, Boguski M, Saulino A, Letcher R, Wigler M, et al. The NF1 locus encodes a protein functionally related to mammalian GAP and yeast IRA proteins. *Cell*. 1990 Nov 16;63(4):851–9.
44. Buchberg AM, Cleveland LS, Jenkins NA, Copeland NG. Sequence homology shared by neurofibromatosis type-1 gene and IRA-1 and IRA-2 negative regulators of the RAS cyclic AMP pathway. *Nature*. 1990 Sep;347(6290):291–4.
45. Rasmussen SA, Friedman JM. NF1 Gene and Neurofibromatosis 1. *Am J Epidemiol*. 2000 Jan 1;151(1):33–40.
46. Largaespada DA, Brannan CI, Jenkins NA, Copeland NG. Nf1 deficiency causes Ras-Dediated granulocyte/macrophage colony stimulating factor hypersensitivity and chronic myeloid leukaemia. *Nat Genet*. 1996 Feb;12(2):137–43.
47. Haferlach C, Grossmann V, Kohlmann A, Schindela S, Kern W, Schnittger S, et al. Deletion of the tumor-suppressor gene NF1 occurs in 5% of myeloid malignancies and is accompanied by a mutation in the remaining allele in half of the cases. *Leuk* 2012 264. 2011 Oct 21;26(4):834–9.
48. Tartaglia M, Mehler EL, Goldberg R, Zampino G, Brunner HG, Kremer H, et al. Mutations in PTPN11, encoding the protein tyrosine phosphatase SHP-2, cause Noonan syndrome. *Nat Genet* 2001 294. 2001 Nov 12;29(4):465–8.
49. Loh ML, Vattikuti S, Schubert S, Reynolds MG, Carlson E, Lieuw KH, et al. Mutations in PTPN11 implicate the SHP-2 phosphatase in leukemogenesis. *Blood*. 2004 Mar 15;103(6):2325–31.
50. Shih AH, Abdel-Wahab O, Patel JP, Levine RL. The role of mutations in epigenetic regulators in myeloid malignancies. *Nat Rev Cancer*. 2012 Sep;12(9):599–612.
51. Herman JG, Baylin SB. Gene Silencing in Cancer in Association with Promoter Hypermethylation. *N Engl J Med*. 2003 Nov

BIBLIOGRAPHY

- 20;349(21):2042–54.
52. Elena C, Galli A, Such E, Meggendorfer M, Germing U, Rizzo E, et al. Integrating clinical features and genetic lesions in the risk assessment of patients with chronic myelomonocytic leukemia. *Blood*. 2016;128(10):1408–17.
 53. Mason C, Khorashad J, Tantravahi S, Kelley T, Zabriskie M, Yan D, et al. Age-related mutations and chronic myelomonocytic leukemia. 2016;
 54. Palomo L, Meggendorfer M, Hutter S, Twardziok S, Ademà V, Fuhrmann I, et al. Molecular landscape and clonal architecture of adult myelodysplastic/myeloproliferative neoplasms. *Blood*. 2020 Oct 15;136(16):1851–62.
 55. Abe JI, Sandhu UG, Hoang NM, Thangam M, Quintana-Quezada RA, Fujiwara K, et al. Coordination of Cellular Localization-Dependent Effects of Sumoylation in Regulating Cardiovascular and Neurological Diseases. 2017 Feb 1;963:337–58.
 56. Tahiliani M, Koh KP, Shen Y, Pastor WA, Bandukwala H, Brudno Y, et al. Conversion of 5-Methylcytosine to 5-Hydroxymethylcytosine in Mammalian DNA by MLL Partner TET1. *Science* (80-). 2009 May 15;324(5929):930–5.
 57. Zhang H, Zhang X, Clark E, Mulcahey M, Huang S, Shi YG. TET1 is a DNA-binding protein that modulates DNA methylation and gene transcription via hydroxylation of 5-methylcytosine. *Cell Res* 2010 2012. 2010 Nov 16;20(12):1390–3.
 58. Solary E, Bernard OA, Tefferi A, Fuks F, Vainchenker W. The Ten-Eleven Translocation-2 (TET2) gene in hematopoiesis and hematopoietic diseases. *Leuk* 2014 283. 2013 Nov 13;28(3):485–96.
 59. Okano M, Bell DW, Haber DA, Li E. DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell*. 1999 Oct 29;99(3):247–57.
 60. Challen GA, Sun D, Jeong M, Luo M, Jelinek J, Berg JS, et al. Dnmt3a is essential for hematopoietic stem cell differentiation. *Nat Genet* 2011 441. 2011 Dec 4;44(1):23–31.
 61. Reitman ZJ, Jin G, Karoly ED, Spasojevic I, Yang J, Kinzler KW, et al. Profiling the effects of isocitrate dehydrogenase 1 and 2 mutations on

- the cellular metabolome. *Proc Natl Acad Sci U S A*. 2011 Feb 22;108(8):3270–5.
62. Figueroa ME, Abdel-Wahab O, Lu C, Ward PS, Patel J, Shih A, et al. Leukemic IDH1 and IDH2 mutations result in a hypermethylation phenotype, disrupt TET2 function, and impair hematopoietic differentiation. *Cancer Cell*. 2010 Dec 14;18(6):553–67.
 63. Abdel-Wahab O, Adli M, LaFave LM, Gao J, Hricik T, Shih AH, et al. ASXL1 mutations promote myeloid transformation through loss of PRC2-mediated gene repression. *Cancer Cell*. 2012 Aug 14;22(2):180–93.
 64. Abdel-Wahab O, Gao J, Adli M, Dey A, Trimarchi T, Chung YR, et al. Deletion of *Asxl1* results in myelodysplasia and severe developmental defects in vivo. 2013 Nov 18;210(12):2641–59.
 65. Viré E, Brenner C, Deplus R, Blanchon L, Fraga M, Didelot C, et al. The Polycomb group protein EZH2 directly controls DNA methylation. *Nat* 2005 4397078. 2005 Dec 14;439(7078):871–4.
 66. Ernst T, Chase AJ, Score J, Hidalgo-Curtis CE, Bryant C, Jones A V, et al. Inactivating mutations of the histone methyltransferase gene *EZH2* in myeloid disorders. *Nat Genet* 2010 428. 2010 Jul 4;42(8):722–6.
 67. Nikoloski G, Langemeijer SMC, Kuiper RP, Knops R, Massop M, Tönnissen ERLTM, et al. Somatic mutations of the histone methyltransferase gene *EZH2* in myelodysplastic syndromes. *Nat Genet* 2010 428. 2010 Jul 4;42(8):665–7.
 68. Yoshida K, Sanada M, Shiraishi Y, Nowak D, Nagata Y, Yamamoto R, et al. Frequent pathway mutations of splicing machinery in myelodysplasia. *Nat* 2011 4787367. 2011 Sep 11;478(7367):64–9.
 69. Thol F, Kade S, Schlarman C, Löffeld P, Morgan M, Krauter J, et al. Frequency and prognostic impact of mutations in *SRSF2*, *U2AF1*, and *ZRSR2* in patients with myelodysplastic syndromes. *Blood*. 2012 Apr 12;119(15):3578–84.
 70. Meggendorfer M, Roller A, Haferlach T, Eder C, Dicker F, Grossmann V, et al. *SRSF2* mutations in 275 cases with chronic myelomonocytic leukemia (CMML). *Blood*. 2012 Oct 11;120(15):3080–8.
 71. Graubert TA, Shen D, Ding L, Okeyo-Owuor T, Lunn CL, Shao J, et al.

BIBLIOGRAPHY

- Recurrent mutations in the U2AF1 splicing factor in myelodysplastic syndromes. *Nat Genet.* 2012 Jan 11;44(1):53–7.
72. Anczukow O, Krainer AR. Splicing-factor alterations in cancers. *RNA.* 2016 Sep 1;22(9):1285–301.
73. CL W, R L. Spliceosome structure and function. *Cold Spring Harb Perspect Biol.* 2011 Jul;3(7):1–2.
74. Papaemmanuil E, Cazzola M, Boultonwood J, Malcovati L, Vyas P, Bowen D, et al. Somatic SF3B1 Mutation in Myelodysplasia with Ring Sideroblasts. *N Engl J Med.* 2011 Oct 13;365(15):1384–95.
75. Malcovati L, Karimi M, Papaemmanuil E, Ambaglio I, Jädersten M, Jansson M, et al. SF3B1 mutation identifies a distinct subset of myelodysplastic syndrome with ring sideroblasts. *Blood.* 2015 Jul 9;126(2):233–41.
76. Wahl MC, Will CL, Lührmann R. The Spliceosome: Design Principles of a Dynamic RNP Machine. *Cell.* 2009 Feb 20;136(4):701–18.
77. Kim E, Ilgan JO, Liang Y, Daubner GM, Lee SC-W, Ramakrishnan A, et al. SRSF2 Mutations Contribute to Myelodysplasia by Mutant-Specific Effects on Exon Recognition. *Cancer Cell.* 2015 May 11;27(5):617–30.
78. Komeno Y, Huang Y-J, Qiu J, Lin L, Xu Y, Zhou Y, et al. SRSF2 Is Essential for Hematopoiesis, and Its Myelodysplastic Syndrome-Related Mutations Dysregulate Alternative Pre-mRNA Splicing. *Mol Cell Biol.* 2015 Sep 1;35(17):3071–82.
79. Masaki S, Ikeda S, Hata A, Shiozawa Y, Kon A, Ogawa S, et al. Myelodysplastic Syndrome-Associated SRSF2 Mutations Cause Splicing Changes by Altering Binding Motif Sequences. *Front Genet.* 2019;10(MAR).
80. Graubert TA, Shen D, Ding L, Okeyo-Owuor T, Lunn CL, Shao J, et al. Recurrent mutations in the U2AF1 splicing factor in myelodysplastic syndromes. *Nat Genet* 2011 441. 2011 Dec 11;44(1):53–7.
81. Przychodzen B, Jerez A, Guinta K, Sekeres MA, Padgett R, Maciejewski JP, et al. Patterns of missplicing due to somatic U2AF1 mutations in myeloid neoplasms. *Blood.* 2013 Aug 8;122(6):999.
82. Madan V, Kanojia D, Li J, Okamoto R, Sato-Otsubo A, Kohlmann A, et al. Aberrant splicing of U12-type introns is the hallmark of ZRSR2

- mutant myelodysplastic syndrome. *Nat Commun* 2015 61. 2015 Jan 14;6(1):1–14.
83. RJ G, JD B. Prp8 protein: at the heart of the spliceosome. *RNA*. 2005 May;11(5):533–57.
84. Galej WP, Oubridge C, Newman AJ, Nagai K. Crystal structure of Prp8 reveals active site cavity of the spliceosome. *Nature*. 2013 Jan 23;493(7434):638–43.
85. Schellenberg MJ, Wu T, Ritchie DB, Fica S, Staley JP, Atta KA, et al. A conformational switch in PRP8 mediates metal ion coordination that promotes pre-mRNA exon ligation. *Nat Struct Mol Biol*. 2013 Jun 19;20(6):728–34.
86. Kurtovic-Kozaric A, Przychodzen B, Singh J, Konarska MM, Clemente MJ, Otrrock ZK, et al. PRPF8 defects cause missplicing in myeloid malignancies. *Leuk* 2015 291. 2014 Apr 30;29(1):126–36.
87. Quesada AE, Routbort MJ, DiNardo CD, Bueso-Ramos CE, Kanagal-Shamanna R, Khoury JD, et al. DDX41 mutations in myeloid neoplasms are associated with male gender, TP53 mutations and high-risk disease. *Am J Hematol*. 2019 Jul 1;94(7):757–66.
88. Polprasert C, Schulze I, Sekeres MA, Makishima H, Przychodzen B, Hosono N, et al. Inherited and Somatic Defects in DDX41 in Myeloid Neoplasms. *Cancer Cell*. 2015 May 11;27(5):658–70.
89. Bannon SA, Routbort MJ, Montalban-Bravo G, Mehta RS, Jelloul FZ, Takahashi K, et al. Next-Generation Sequencing of DDX41 in Myeloid Neoplasms Leads to Increased Detection of Germline Alterations. *Front Oncol*. 2021 Jan 28;0:2689.
90. Bejar R, Stevenson K, Abdel-Wahab O, Galili N, Nilsson B, Garcia-Manero G, et al. Clinical Effect of Point Mutations in Myelodysplastic Syndromes. *N Engl J Med*. 2011 Jun 30;364(26):2496–506.
91. Growney JD, Shigematsu H, Li Z, Lee BH, Adelsperger J, Rowan R, et al. Loss of Runx1 perturbs adult hematopoiesis and is associated with a myeloproliferative phenotype. *Blood*. 2005 Jul 15;106(2):494–504.
92. Piazza R, Magistrini V, Redaelli S, Mauri M, Massimino L, Sessa A, et al. SETBP1 induces transcription of a network of development genes by acting as an epigenetic hub. *Nat Commun* 2018 91. 2018 Jun

BIBLIOGRAPHY

- 6;9(1):1–13.
93. Thol F, Suchanek KJ, Koenecke C, Stadler M, Platzbecker U, Thiede C, et al. SETBP1 mutation analysis in 944 patients with MDS and AML. *Leuk* 2013 2710. 2013 May 7;27(10):2072–5.
 94. Hou H-A, Kuo Y-Y, Tang J-L, Chou W-C, Yao M, Lai Y-J, et al. Clinical implications of the SETBP1 mutation in patients with primary myelodysplastic syndrome and its stability during disease progression. *Am J Hematol*. 2014 Feb 1;89(2):181–6.
 95. Lopez RG, Carron C, Oury C, Gardellin P, Bernard O, Ghysdael J. TEL Is a Sequence-specific Transcriptional Repressor. *J Biol Chem*. 1999 Oct 15;274(42):30132–8.
 96. Wang Q, Dong S, Yao H, Wen L, Qiu H, Qin L, et al. ETV6 mutation in a cohort of 970 patients with hematologic malignancies. *Haematologica*. 2014 Oct 1;99(10):e176.
 97. Padron E, Yoder S, Kunigal S, Mesa T, Teer JK, Ali N Al, et al. ETV6 and signaling gene mutations are associated with secondary transformation of myelodysplastic syndromes to chronic myelomonocytic leukemia. *Blood*. 2014;123(23):3675.
 98. Huynh KD, Fischle W, Verdin E, Bardwell VJ. BCoR, a novel corepressor involved in BCL-6 repression. *Genes Dev*. 2000 Jul 15;14(14):1810.
 99. Li M, Collins R, Jiao Y, Ouilllette P, Bixby D, Erba H, et al. Somatic mutations in the transcriptional corepressor gene BCORL1 in adult acute myelogenous leukemia. *Blood*. 2011 Nov 24;118(22):5914.
 100. Damm F, Chesnais V, Nagata Y, Yoshida K, Scourzic L, Okuno Y, et al. BCOR and BCORL1 mutations in myelodysplastic syndromes and related disorders. *Blood*. 2013 Oct 31;122(18):3169–77.
 101. Wamstad JA, Corcoran CM, Keating AM, Bardwell VJ. Role of the transcriptional corepressor Bcor in embryonic stem cell differentiation and early embryonic development. *PLoS One*. 2008 Jul 30;3(7):e2814.
 102. Tiacci E, Grossmann V, Martelli MP, Kohlmann A, Haferlach T, Falini B. The corepressors BCOR and BCORL1: two novel players in acute myeloid leukemia. *Haematologica*. 2012 Jan 1;97(1):3–5.
 103. Fliedner A, Gregor A, Ferrazzi F, Ekici AB, Sticht H, Zweier C. Loss of PHF6 leads to aberrant development of human neuron-like cells. *Sci*

- Reports 2020 101. 2020 Nov 4;10(1):1–15.
104. Przychodzen B, Gu X, You D, Hirsch CM, Clemente MJ, Viny AD, et al. PHF6 - Somatic Mutations and Their Role in Pathophysiology of MDS and AML. *Blood*. 2015 Dec 3;126(23):1259.
 105. Chien KS, Kanagal-Shamanna R, Naqvi K, Sasaki K, Alvarado Y, Takahashi K, et al. The Impact of PHF6 Mutations in Myelodysplastic Syndromes, Chronic Myelomonocytic Leukemia, and Acute Myeloid Leukemia. *Blood*. 2019 Nov 13;134(Supplement_1):1436–1436.
 106. McNerney ME, Brown CD, Wang X, Bartom ET, Karmakar S, Bandlamudi C, et al. CUX1 is a haploinsufficient tumor suppressor gene on chromosome 7 frequently inactivated in acute myeloid leukemia. *Blood*. 2013 Feb 7;121(6):975–83.
 107. Ramdzan ZM, Vadnais C, Pal R, Vandal G, Cadieux C, Leduy L, et al. RAS Transformation Requires CUX1-Dependent Repair of Oxidative DNA Damage. *PLOS Biol*. 2014;12(3):e1001807.
 108. Aly M, Ramdzan ZM, Nagata Y, Balasubramanian SK, Hosono N, Makishima H, et al. Distinct clinical and biological implications of CUX1 in myeloid neoplasms. *Blood Adv*. 2019 Jul 23;3(14):2164–78.
 109. Pater E de, Kaimakis P, Vink CS, Yokomizo T, Yamada-Inagawa T, Linden R van der, et al. Gata2 is required for HSC generation and survival. *J Exp Med*. 2013 Dec;210(13):2843.
 110. Dickinson RE, Griffin H, Bigley V, Reynard LN, Hussain R, Haniffa M, et al. Exome sequencing identifies GATA-2 mutation as the cause of dendritic cell, monocyte, B and NK lymphoid deficiency. *Blood*. 2011 Sep 8;118(10):2656–8.
 111. Ezoë S. GATA-2/estrogen receptor chimera regulates cytokine-dependent growth of hematopoietic cells through accumulation of p21WAF1 and p27Kip1 proteins. *Blood*. 2002 Nov 15;100(10):3512–20.
 112. Bernard E, Nannya Y, Hasserjian RP, Devlin SM, Tuechler H, Medina-Martinez JS, et al. Implications of TP53 allelic state for genome stability, clinical presentation and outcomes in myelodysplastic syndromes. *Nat Med* 2020 2610. 2020 Aug 3;26(10):1549–56.
 113. Nannenga B, Lu X, Dumble M, Van Maanen M, Nguyen TA, Sutton R,

BIBLIOGRAPHY

- et al. Augmented cancer resistance and DNA damage response phenotypes in PPM1D null mice. *Mol Carcinog*. 2006 Aug 1;45(8):594–604.
114. Bartek J, Lukas J. Chk1 and Chk2 kinases in checkpoint control and cancer. *Cancer Cell*. 2003 May 1;3(5):421–9.
115. Walsh T, Casadei S, Coats KH, Swisher E, Stray SM, Higgins J, et al. Spectrum of Mutations in BRCA1, BRCA2, CHEK2, and TP53 in Families at High Risk of Breast Cancer. *JAMA*. 2006 Mar 22;295(12):1379–88.
116. Banin S, Moyal L, Shieh SY, Taya Y, Anderson CW, Chessa L, et al. Enhanced Phosphorylation of p53 by ATM in Response to DNA Damage. *Science (80-)*. 1998 Sep 11;281(5383):1674–7.
117. Sandoval N, Platzer M, Rosenthal A, Dörk T, Bendix R, Skawran B, et al. Characterization of ATM Gene Mutations in 66 Ataxia Telangiectasia Families. *Hum Mol Genet*. 1999 Jan 1;8(1):69–79.
118. Peters J-M, Tedeschi A, Schmitz J. The cohesin complex and its roles in chromosome biology. *Genes Dev*. 2008 Nov 15;22(22):3089–114.
119. Kon A, Shih L-Y, Minamino M, Sanada M, Shiraishi Y, Nagata Y, et al. Recurrent mutations in multiple components of the cohesin complex in myeloid neoplasms. *Nat Genet* 2013 4510. 2013 Aug 18;45(10):1232–7.
120. Thota S, Viny AD, Makishima H, Spitzer B, Radivoyevitch T, Przychodzen B, et al. Genetic alterations of the cohesin complex genes in myeloid malignancies. *Blood*. 2014;124(11):1790.
121. Mazumdar C, Shen Y, Xavy S, Zhao F, Reinisch A, Li R, et al. Leukemia-Associated Cohesin Mutants Dominantly Enforce Stem Cell Programs and Impair Human Hematopoietic Progenitor Differentiation. *Cell Stem Cell*. 2015 Dec 3;17(6):675–88.
122. Fisher JB, Peterson J, Reimer M, Stelloh C, Pulakanti K, Gerbec ZJ, et al. The cohesin subunit Rad21 is a negative regulator of hematopoietic self-renewal through epigenetic repression of Hoxa7 and Hoxa9. *Leukemia*. 2017 Mar 24;31(3):712–9.
123. Grupo Español de Síndromes Mielodisplásicos (GESMD). Guía de aplicación clínica de la secuenciación masiva en síndromes mielodisplásicos y leucemia mielomonocítica crónica. 2017.

124. Grupo Español de Enfermedades Mieloproliferativas Crónicas Filadelfia Negativas (GEMFIN). Manual de Recomendaciones en Neoplasias Mieloproliferativas Crónicas Filadelfia Negativas. 2020.
125. Martin R, Acha P, Ganster C, Palomo L, Dierks S, Fuster-Tormo F, et al. Targeted deep sequencing of CD34+ cells from peripheral blood can reproduce bone marrow molecular profile in myelodysplastic syndromes. *Am J Hematol*. 2018 Jun 1;93(6):E152–4.
126. Mohamedali AM, Gäken J, Ahmed M, Malik F, Smith AE, Best S, et al. High concordance of genomic and cytogenetic aberrations between peripheral blood and bone marrow in myelodysplastic syndrome (MDS). *Leuk* 2015 299. 2015 May 6;29(9):1928–38.
127. Steensma DP, Bejar R, Jaiswal S, Lindsley RC, Sekeres MA, Hasserjian RP, et al. Clonal hematopoiesis of indeterminate potential and its distinction from myelodysplastic syndromes. *Blood*. 2015 Jul 2;126(1):9.
128. Bejar R. CHIP, ICUS, CCUS and other four-letter words. *Leuk* 2017 319. 2017 Jun 8;31(9):1869–71.
129. Jaiswal S, Fontanillas P, Flannick J, Manning A, Grauman P V., Mar BG, et al. Age-Related Clonal Hematopoiesis Associated with Adverse Outcomes. *N Engl J Med*. 2014 Dec 25;371(26):2488–98.
130. Genovese G, Kähler AK, Handsaker RE, Lindberg J, Rose SA, Bakhoun SF, et al. Clonal Hematopoiesis and Blood-Cancer Risk Inferred from Blood DNA Sequence. *N Engl J Med*. 2014 Dec 25;371(26):2477–87.
131. Spivak JL. The chronic myeloproliferative disorders: Clonality and clinical heterogeneity. *Semin Hematol*. 2004 Apr 1;41(2 SUPPL. 3):1–5.
132. Barbui T, Finazzi G, De Gaetano G, Marchioli R, Tognoni G, Patrono C, et al. Polycythemia vera: The natural history of 1213 patients followed for 20 years. *Ann Intern Med*. 1995 Nov 1;123(9):656–64.
133. Prchal JF, Axelrad AA. Letter: Bone-marrow responses in polycythemia vera. *N Engl J Med*. 1974 Jun 13;290(24):1382.
134. Birgegård G, Wide L. Serum erythropoietin in the diagnosis of polycythaemia and after phlebotomy treatment. *Br J Haematol*. 1992 Aug;81(4):603–6.

BIBLIOGRAPHY

135. Pardanani A, Lasho TL, Finke C, Hanson CA, Tefferi A. Prevalence and clinicopathologic correlates of JAK2 exon 12 mutations in JAK2V617F-negative polycythemia vera. *Leukemia*. 2007 Sep 28;21(9):1960–3.
136. Vannucchi AM, Antonioli E, Guglielmelli P, Pardanani A, Tefferi A. Clinical correlates of JAK2V617F presence or allele burden in myeloproliferative neoplasms: a critical reappraisal. *Leuk* 2008 227. 2008 May 22;22(7):1299–307.
137. Cortelazzo S, Viero P, Finazzi G, D’Emilio A, Rodeghiero F, Barbui T. Incidence and risk factors for thrombotic complications in a historical cohort of 100 patients with essential thrombocythemia. *J Clin Oncol*. 1990 Mar;8(3):556–62.
138. Besses C, Cervantes F, Pereira A, Florensa L, Solé F, Hernández-Boluda JC, et al. Major vascular complications in essential thrombocythemia: a study of the predictive factors in a series of 148 patients. *Leukemia*. 1999 Feb;13(2):150–4.
139. Tefferi A, Mudireddy M, Mannelli F, Begna KH, Patnaik MM, Hanson CA, et al. Blast phase myeloproliferative neoplasm: Mayo-AGIMM study of 410 patients from two separate cohorts. *Leuk* 2018 325. 2018 Feb 2;32(5):1200–10.
140. Barosi G, Mesa RA, Thiele J, Cervantes F, Campbell PJ, Verstovsek S, et al. Proposed criteria for the diagnosis of post-polycythemia vera and post-essential thrombocythemia myelofibrosis: a consensus statement from the international working group for myelofibrosis research and treatment. *Leukemia*. 2008 Feb;22(2):437–8.
141. Tefferi A, Saeed L, Hanson CA, Ketterling RP, Pardanani A, Gangat N. Application of current prognostic models for primary myelofibrosis in the setting of post-polycythemia vera or post-essential thrombocythemia myelofibrosis. *Leukemia*. 2017;31(12):2851.
142. Rumi E, Pietra D, Pascutto C, Guglielmelli P, Martínez-Trillos A, Casetti I, et al. Clinical effect of driver mutations of JAK2, CALR, or MPL in primary myelofibrosis. *Blood*. 2014 Aug 14;124(7):1062.
143. Tefferi A, Lasho TL, Finke CM, Knudson RA, Ketterling R, Hanson CH, et al. CALR vs JAK2 vs MPL-mutated or triple-negative myelofibrosis: clinical, cytogenetic and molecular comparisons. *Leukemia*. 2014 Jul 9;28(7):1472–7.

144. Lundberg P, Takizawa H, Kubovcakova L, Guo G, Hao-Shen H, Dirnhofer S, et al. Myeloproliferative neoplasms can be initiated from a single hematopoietic stem cell expressing JAK2-V617F. *J Exp Med*. 2014 Oct 20;211(11):2213–30.
145. Li J, Prins D, Park HJ, Grinfeld J, Gonzalez-Arias C, Loughran S, et al. Mutant calreticulin knockin mice develop thrombocytosis and myelofibrosis without a stem cell self-renewal advantage. *Blood*. 2018 Feb 8;131(6):649–61.
146. Nussenzveig RH, Pham HT, Perkins SL, Prchal JT, Agarwal AM, Salama ME. Increased frequency of co-existing JAK2 exon-12 or MPL exon-10 mutations in patients with low JAK2 V617F allelic burden. *Leuk Lymphoma*. 2016 Jun 2;57(6):1429–35.
147. Thompson ER, Nguyen T, Kankanige Y, Yeh P, Ingbritsen M, McBean M, et al. Clonal independence of JAK2 and CALR or MPL mutations in comutated myeloproliferative neoplasms demonstrated by single cell DNA sequencing. *Haematologica*. 2021 Jan 1;106(1):313.
148. Swerdlow S, Campo E, Harris N, Jaffe E, Pileri S, Stein H. WHO classification of tumours of haematopoietic and lymphoid tissues. Vol. 2, IARC. 2008. 439 p.
149. Lacout C, Pisani DF, Tulliez M, Gachelin FM, Vainchenker W, Villeval J-L. JAK2V617F expression in murine hematopoietic cells leads to MPD mimicking human PV with secondary myelofibrosis. *Blood*. 2006 Sep 1;108(5):1652–60.
150. Wernig G, Mercher T, Okabe R, Levine RL, Lee BH, Gilliland DG. Expression of Jak2V617F causes a polycythemia vera–like disease with associated myelofibrosis in a murine bone marrow transplant model. *Blood*. 2006 Jun 1;107(11):4274.
151. Xing S, Wanting TH, Zhao W, Ma J, Wang S, Xu X, et al. Transgenic expression of JAK2V617F causes myeloproliferative disorders in mice. *Blood*. 2008 May 15;111(10):5109.
152. Akada H, Yan D, Zou H, Fiering S, Hutchison RE, Mohi MG. Conditional expression of heterozygous or homozygous Jak2V617F from its endogenous promoter induces a polycythemia vera–like disease. *Blood*. 2010 Apr 29;115(17):3589.
153. Marty C, Lacout C, Martin A, Hasan S, Jacquot S, Birling M-C, et al.

BIBLIOGRAPHY

- Myeloproliferative neoplasm induced by constitutive expression of JAK2V617F in knock-in mice. *Blood*. 2010 Aug 5;116(5):783–7.
154. Tiedt R, Hao-Shen H, Sobas MA, Looser R, Dirnhofer S, Schwaller J, et al. Ratio of mutant JAK2-V617F to wild-type Jak2 determines the MPD phenotypes in transgenic mice. *Blood*. 2008 Apr 15;111(8):3931–40.
155. Shide K, Shimoda HK, Kumano T, Karube K, Kameda T, Takenaka K, et al. Development of ET, primary myelofibrosis and PV in mice expressing JAK2 V617F. *Leukemia*. 2008 Jan 22;22(1):87–95.
156. Passamonti F, Elena C, Schnittger S, Skoda RC, Green AR, Girodon F, et al. Molecular and clinical features of the myeloproliferative neoplasm associated with JAK2 exon 12 mutations. *Blood*. 2011 Mar 10;117(10):2813–6.
157. Vannucchi AM, Antonioli E, Guglielmelli P, Rambaldi A, Barosi G, Marchioli R, et al. Clinical profile of homozygous JAK2 617V>F mutation in patients with polycythemia vera or essential thrombocythemia. *Blood*. 2007 Aug 1;110(3):840–6.
158. Tefferi A, Lasho TL, Schwager SM, Strand JS, Elliott M, Mesa R, et al. The clinical phenotype of wild-type, heterozygous, and homozygous JAK2V617F in polycythemia vera. *Cancer*. 2006 Feb 1;106(3):631–5.
159. Alvarez-Larrán A, Bellosillo B, Pereira A, Kerguelen A, Carlos Hernández-Boluda J, Martínez-Avilés L, et al. JAK2 V617F monitoring in polycythemia vera and essential thrombocythemia: Clinical usefulness for predicting myelofibrotic transformation and thrombotic events. *Am J Hematol*. 2014 May;89(5):517–23.
160. Williams N, Lee J, Moore L, Baxter EJ, Hewinson J, Dawson KJ, et al. Phylogenetic reconstruction of myeloproliferative neoplasm reveals very early origins and lifelong evolution. *bioRxiv*. 2020 Nov 9;2020.11.09.374710.
161. Tefferi A, Wassie EA, Guglielmelli P, Gangat N, Belachew AA, Lasho TL, et al. Type 1 versus Type 2 calreticulin mutations in essential thrombocythemia: A collaborative study of 1027 patients. *Am J Hematol*. 2014 Aug 1;89(8):E121–4.
162. Boyd EM, Bench AJ, Goday-Fernández A, Anand S, Vaghela KJ, Beer P, et al. Clinical utility of routine MPL exon 10 analysis in the diagnosis of essential thrombocythaemia and primary myelofibrosis. *Br J*

- Haematol. 2010 Apr 1;149(2):250–7.
163. Pietra D, Rumi E, Ferretti V V, Buduo CA Di, Milanese C, Cavalloni C, et al. Differential clinical effects of different mutation subtypes in CALR-mutant myeloproliferative neoplasms. *Leukemia*. 2016 Feb;30(2):431–8.
 164. Verger E, Cassinat B, Chauveau A, Dosquet C, Giraudier S, Schlageter M-H, et al. Clinical and molecular response to interferon- α therapy in essential thrombocythemia patients with CALR mutations. *Blood*. 2015 Dec 10;126(24):2585–91.
 165. DYNAMICS OF CALR MUTANT ALLELE BURDEN IN ESSENTIAL THROMBOCYTHEMIA.... EHA Library. Bellosillo B. Jun 13 2014; 53859 [Internet]. [cited 2021 Sep 2]. Available from: https://library.ehaweb.org/eha/2014/19th/53859/beatriz.bellosillo.dynamics.of.calr.mutant.allele.burden.in.essential.html?f=listing%3D1%2Abrowseby%3D8%2Asortby%3D2%2Amedia%3D2%2Ace_id%3D717%2Aot_id%3D8332
 166. Guglielmelli P, Rotunno G, Bogani C, Mannarelli C, Giunti L, Provenzano A, et al. Ruxolitinib is an effective treatment for CALR-positive patients with myelofibrosis. *Br J Haematol*. 2016 Jun 1;173(6):938–40.
 167. Haider M, Elala YC, Gangat N, Hanson CA, Tefferi A. MPL mutations and palpable splenomegaly are independent risk factors for fibrotic progression in essential thrombocythemia. *Blood Cancer J*. 2016;6(10):e487.
 168. Asp J, Andréasson B, Hansson U, Wasslavik C, Abellsson J, Johansson P, et al. Mutation status of essential thrombocythemia and primary myelofibrosis defines clinical outcome. *Haematologica*. 2016 Apr;101(4):e129–32.
 169. Ortmann CA, Kent DG, Nangalia J, Silber Y, Wedge DC, Grinfeld J, et al. Effect of Mutation Order on Myeloproliferative Neoplasms. *N Engl J Med*. 2015 Feb 12;372(7):601.
 170. Grinfeld J, Nangalia J, Baxter EJ, Wedge DC, Angelopoulos N, Cantrill R, et al. Classification and Personalized Prognosis in Myeloproliferative Neoplasms. *N Engl J Med*. 2018 Oct 11;379(15):1416–30.
 171. Kennedy AL, Shimamura A. Genetic predisposition to MDS: clinical

- features and clonal evolution. *Blood*. 2019 Mar 7;133(10):1071–85.
172. Kuendgen A, Nomdedeu M, Tuechler H, Garcia-Manero G, Komrokji RS, Sekeres MA, et al. Therapy-related myelodysplastic syndromes deserve specific diagnostic sub-classification and risk-stratification—an approach to classification of patients with t-MDS. *Leuk* 2020 353. 2020 Jun 29;35(3):835–49.
173. Cazzola M. Myelodysplastic Syndromes. Longo DL, editor. *N Engl J Med*. 2020 Oct 1;383(14):1358–74.
174. Vallespí T, Imbert M, Mecucci C, Preudhomme C, Fenaux P. Diagnosis, classification, and cytogenetics of myelodysplastic syndromes. *Haematologica*. 1998 Mar;83(3):258–75.
175. Solé F, Luño E, Sanzo C, Espinet B, Sanz GF, Cervera J, et al. Identification of novel cytogenetic markers with prognostic significance in a series of 968 patients with primary myelodysplastic syndromes. *Haematologica*. 2005 Sep;90(9):1168–78.
176. Haase D, Germing U, Schanz J, Pfeilstöcker M, Nösslinger T, Hildebrandt B, 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 Dec 15;110(13):4385–95.
177. Schanz J, Tüchler H, Solé F, Mallo M, Luño E, Cervera J, 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 Mar 10;30(8):820–9.
178. Greenberg PL, Tuechler H, Schanz J, Sanz G, Garcia-Manero G, Solé F, et al. Revised International Prognostic Scoring System for Myelodysplastic Syndromes. *Blood*. 2012 Sep 20;120(12):2454.
179. Vardiman JW, Thiele J, Arber DA, Brunning RD, Borowitz MJ, Porwit A, et al. The 2008 revision of the World Health Organization (WHO) classification of myeloid neoplasms and acute leukemia: rationale and important changes. *Blood*. 2009 Jul 30;114(5):937–51.
180. Liquori A, Lesende I, Palomo L, Avetisyan G, Ibáñez M, González-Romero E, et al. A Single-Run Next-Generation Sequencing (NGS) Assay for the Simultaneous Detection of Both Gene Mutations and Large Chromosomal Abnormalities in Patients with Myelodysplastic

- Syndromes (MDS) and Related Myeloid Neoplasms. *Cancers* (Basel). 2021 Apr 18;13(8).
181. Malcovati L, Gallì A, Travaglino E, Ambaglio I, Rizzo E, Molteni E, et al. Clinical significance of somatic mutation in unexplained blood cytopenia. *Blood*. 2017 Jun 22;129(25):3371–8.
 182. Bejar R. Implications of molecular genetic diversity in myelodysplastic syndromes. *Curr Opin Hematol*. 2017 Mar 1;24(2):73–8.
 183. Kennedy JA, Ebert BL. Clinical Implications of Genetic Mutations in Myelodysplastic Syndrome. *J Clin Oncol*. 2017 Mar 20;35(9):968–74.
 184. Papaemmanuil E. Lessons learned from the IWG-PM Development of IPSS-Molecular. 16th Int Congr Myelodysplastic Syndr. 2021;
 185. Haider M, Duncavage EJ, Afaneh KF, Bejar R, List AF. New Insight Into the Biology, Risk Stratification, and Targeted Treatment of Myelodysplastic Syndromes. *Am Soc Clin Oncol Educ B*. 2017 May 29;(37):480–94.
 186. Zahid MF, Barraco D, Lasho TL, Finke C, Ketterling RP, Gangat N, et al. Spectrum of autoimmune diseases and systemic inflammatory syndromes in patients with chronic myelomonocytic leukemia. *Leuk Lymphoma*. 2017 Jun 3;58(6):1488–93.
 187. Peker D, Padron E, Bennett JM, Zhang X, Horna P, Epling-Burnette PK, et al. A Close Association of Autoimmune-Mediated Processes and Autoimmune Disorders with Chronic Myelomonocytic Leukemia: Observation from a Single Institution. *Acta Haematol*. 2015 Feb 17;133(2):249–56.
 188. Itzykson R, Kosmider O, Renneville A, Morabito M, Preudhomme C, Berthon C, et al. Clonal architecture of chronic myelomonocytic leukemias. *Blood*. 2013 Mar 21;121(12):2186–98.
 189. Li Z, Cai X, Cai C-L, Wang J, Zhang W, Petersen BE, et al. Deletion of Tet2 in mice leads to dysregulated hematopoietic stem cells and subsequent development of myeloid malignancies. *Blood*. 2011 Oct 27;118(17):4509–18.
 190. Cazzola M, Della Porta MG, Malcovati L. The genetic basis of myelodysplasia and its clinical relevance. *Blood*. 2013 Dec 12;122(25):4021–34.

BIBLIOGRAPHY

191. Itzykson R, Kosmider O, Renneville A, Gelsi-Boyer V, Meggendorfer M, Morabito M, et al. Prognostic score including gene mutations in chronic myelomonocytic leukemia. *J Clin Oncol*. 2013 Jul 1;31(19):2428–36.
192. Takahashi K, Pemmaraju N, Strati P, Noguerras-Gonzalez G, Ning J, Bueso-Ramos C, et al. Clinical characteristics and outcomes of therapy-related chronic myelomonocytic leukemia. *Blood*. 2013 Oct 17;122(16):2807–11; quiz 2920.
193. Subari S, Patnaik M, Alfakara D, Gangat N, Elliott M, Hogan W, et al. Patients With Therapy-Related CMML Have Shorter Median Overall Survival Than Those With De Novo CMML: Mayo Clinic Long-Term Follow-Up Experience. *Clin Lymphoma Myeloma Leuk*. 2015 Sep 1;15(9):546–9.
194. Patnaik MM, Vallapureddy R, Yalniz FF, Hanson CA, Ketterling RP, Lasho TL, et al. Therapy related-chronic myelomonocytic leukemia (CMML): Molecular, cytogenetic, and clinical distinctions from de novo CMML. *Am J Hematol*. 2018 Jan 1;93(1):65–73.
195. Elbæk MV, Sørensen AL, Hasselbalch HC, MV E, AL S, HC H. Chronic inflammation and autoimmunity as risk factors for the development of chronic myelomonocytic leukemia? *Leuk Lymphoma*. 2016 Feb 17;57(8):1793–9.
196. Montoro J, Gallur L, Merchán B, Molero A, Roldán E, Martínez-Valle F, et al. Autoimmune disorders are common in myelodysplastic syndrome patients and confer an adverse impact on outcomes. *Ann Hematol*. 2018 Aug 1;97(8):1349–56.
197. Selimoglu-Buet D, Wagner-Ballon O, Saada V, Bardet V, Itzykson R, Bencheikh L, et al. Characteristic repartition of monocyte subsets as a diagnostic signature of chronic myelomonocytic leukemia. *Blood*. 2015 Jun 4;125(23):3618–26.
198. Jaffe ES, Harris NL, Stein H VJ. Pathology and genetics of tumours of haematopoietic and lymphoid tissues. *WHO Classification of Tumours*. Vol. 3, IARC. 2001.
199. Bennett JM, Catovsky D, Daniel MT, Flandrin G, Galton DA, Gralnick HR, et al. Proposals for the classification of the acute leukaemias. French-American-British (FAB) co-operative group. *Br J Haematol*. 1976 Aug;33(4):451–8.

-
200. Swerdlow SH, Campo E, Harris NL, Jaffe ES, Pileri SA, Stein H TJ. WHO Classification of Tumours of Haematopoietic and Lymphoid Tissues. 4th ed. Vol. 2, IARC. 2017.
201. Schuler E, Schroeder M, Neukirchen J, Strupp C, Xicoy B, Kündgen A, et al. Refined medullary blast and white blood cell count based classification of chronic myelomonocytic leukemias. *Leuk Res.* 2014 Dec 1;38(12):1413–9.
202. Geyer JT, Tam W, Liu Y-C, Chen Z, Wang SA, Bueso-Ramos C, et al. Oligomonocytic chronic myelomonocytic leukemia (chronic myelomonocytic leukemia without absolute monocytosis) displays a similar clinicopathologic and mutational profile to classical chronic myelomonocytic leukemia. *Mod Pathol.* 2017 Sep 1;30(9):1213–22.
203. Valent P, Orazi A, Savona MR, Patnaik MM, Onida F, Loosdrecht AA van de, et al. Proposed diagnostic criteria for classical chronic myelomonocytic leukemia (CMML), CMML variants and pre-CMML conditions. *Haematologica.* 2019;104(10):1935.
204. Meggendorfer M, Haferlach T, Alpermann T, Jeromin S, Haferlach C, Kern W, et al. Specific molecular mutation patterns delineate chronic neutrophilic leukemia, atypical chronic myeloid leukemia, and chronic myelomonocytic leukemia. *Haematologica.* 2014 Dec 1;99(12):e244–6.
205. Patnaik MM, Tefferi A. Cytogenetic and molecular abnormalities in chronic myelomonocytic leukemia. *Blood Cancer J.* 2016;6(January):1–8.
206. Such E, Germing U, Malcovati L, Cervera J, Kuendgen A, Della Porta MG, et al. Development and validation of a prognostic scoring system for patients with chronic myelomonocytic leukemia. *Blood.* 2013 Apr 11;121(15):3005–15.
207. Ziegler-Heitbrock L, Ancuta P, Crowe S, Dalod M, Grau V, Hart DN, et al. Nomenclature of monocytes and dendritic cells in blood. *Blood.* 2010 Oct 21;116(16):e74–80.
208. Solary E, Itzykson R. How I treat chronic myelomonocytic leukemia. *Blood.* 2017 Jul 13;130(2):126–36.
209. Itzykson R, Duchmann M, Lucas N, Solary E. CMML: Clinical and molecular aspects. *Int J Hematol* 2017 1056. 2017 Apr 28;105(6):711–

BIBLIOGRAPHY

- 9.
210. Ignatiadis M, Sledge GW, Jeffrey SS. Liquid biopsy enters the clinic — implementation issues and future challenges. *Nat Rev Clin Oncol* 2021 185. 2021 Jan 20;18(5):297–312.
211. Mandel P, Metais P. Nuclear Acids In Human Blood Plasma. *C R Seances Soc Biol Fil.* 1948 Feb 1;142(3–4):241–3.
212. Wang Y, Springer S, Mulvey CL, Silliman N, Schaefer J, Sausen M, et al. Detection of somatic mutations and HPV in the saliva and plasma of patients with head and neck squamous cell carcinomas. *Sci Transl Med.* 2015 Jun 24;7(293):293ra104.
213. Ding S, Song X, Geng X, Liu L, Ma H, Wang X, et al. Saliva-derived cfDNA is applicable for EGFR mutation detection but not for quantitation analysis in non-small cell lung cancer. *Thorac Cancer.* 2019 Oct 1;10(10):1973–83.
214. Ponti G, Manfredini M, Tomasi A. Non-blood sources of cell-free DNA for cancer molecular profiling in clinical pathology and oncology. *Crit Rev Oncol Hematol.* 2019 Sep 1;141:36–42.
215. Schwarzenbach H, Nishida N, Calin GA, Pantel K. Clinical relevance of circulating cell-free microRNAs in cancer. *Nat Rev Clin Oncol.* 2014 Mar;11(3):145–56.
216. Huang A, Zheng H, Wu Z, Chen M, Huang Y. Circular RNA-protein interactions: functions, mechanisms, and identification. *Theranostics.* 2020;10(8):3503.
217. Guarnerio J, Bezzi M, Jeong JC, Paffenholz S V., Berry K, Naldini MM, et al. Oncogenic Role of Fusion-circRNAs Derived from Cancer-Associated Chromosomal Translocations. *Cell.* 2016 Apr 7;165(2):289–302.
218. Tan S, Gou Q, Pu W, Guo C, Yang Y, Wu K, et al. Circular RNA F-circEA produced from EML4-ALK fusion gene as a novel liquid biopsy biomarker for non-small cell lung cancer. *Cell Res* 2018 286. 2018 Apr 8;28(6):693–5.
219. Nguyen DX, Bos PD, Massagué J. Metastasis: from dissemination to organ-specific colonization. *Nat Rev Cancer.* 2009 Apr;9(4):274–84.
220. De Bono JS, Scher HI, Montgomery RB, Parker C, Miller MC, Tissing H,

- et al. Circulating Tumor Cells Predict Survival Benefit from Treatment in Metastatic Castration-Resistant Prostate Cancer. *Clin Cancer Res.* 2008 Oct 1;14(19):6302–9.
221. Cristofanilli M, Budd GT, Ellis MJ, Stopeck A, Matera J, Miller MC, et al. Circulating tumor cells, disease progression, and survival in metastatic breast cancer. *N Engl J Med.* 2004 Aug 19;351(8):781–91.
222. Hoshino A, Costa-Silva B, Shen TL, Rodrigues G, Hashimoto A, Tesic Mark M, et al. Tumour exosome integrins determine organotropic metastasis. *Nat* 2015 5277578. 2015 Oct 28;527(7578):329–35.
223. Peinado H, Alečković M, Lavotshkin S, Matei I, Costa-Silva B, Moreno-Bueno G, et al. Melanoma exosomes educate bone marrow progenitor cells toward a pro-metastatic phenotype through MET. *Nat Med* 2012 186. 2012 May 27;18(6):883–91.
224. Nilsson RJA, Balaj L, Hulleman E, Van Rijn S, Pegtel DM, Walraven M, et al. Blood platelets contain tumor-derived RNA biomarkers. *Blood.* 2011 Sep 29;118(13):3680–3.
225. Best MG, Sol N, Kooi I, Tannous J, Westerman BA, Rustenburg F, et al. RNA-Seq of Tumor-Educated Platelets Enables Blood-Based Pan-Cancer, Multiclass, and Molecular Pathway Cancer Diagnostics. *Cancer Cell.* 2015 Nov 9;28(5):666–76.
226. Zill OA, Banks KC, Fairclough SR, Mortimer SA, Vowles J V., Mokhtari R, et al. The Landscape of Actionable Genomic Alterations in Cell-Free Circulating Tumor DNA from 21,807 Advanced Cancer Patients. *Clin Cancer Res.* 2018 Aug 1;24(15):3528–38.
227. Lui YYN, Chik K-W, Chiu RWK, Ho C-Y, Lam CWK, Lo YMD. Predominant hematopoietic origin of cell-free DNA in plasma and serum after sex-mismatched bone marrow transplantation. *Clin Chem.* 2002 Mar;48(3):421–7.
228. Snyder MW, Kircher M, Hill AJ, Daza RM, Shendure J. Cell-free DNA Comprises an In Vivo Nucleosome Footprint that Informs Its Tissues-Of-Origin. *Cell.* 2016 Jan 14;164(1–2):57–68.
229. Ulz P, Thallinger GG, Auer M, Graf R, Kashofer K, Jahn SW, et al. Inferring expressed genes by whole-genome sequencing of plasma DNA. *Nat Genet.* 2016 Oct 29;48(10):1273–8.

BIBLIOGRAPHY

230. Moss J, Magenheim J, Neiman D, Zemmour H, Loyfer N, Korach A, et al. Comprehensive human cell-type methylation atlas reveals origins of circulating cell-free DNA in health and disease. *Nat Commun.* 2018 Dec 29;9(1):5068.
231. Sadeh R, Sharkia I, Fialkoff G, Rahat A, Gutin J, Chappleboim A, et al. ChIP-seq of plasma cell-free nucleosomes identifies gene expression programs of the cells of origin. *Nat Biotechnol.* 2021;39(5):586–98.
232. Vasioukhin V, Anker P, Maurice P, Lyautey J, Lederrey C, Stroun M. Point mutations of the N-ras gene in the blood plasma DNA of patients with myelodysplastic syndrome or acute myelogenous leukaemia. *Br J Haematol.* 1994 Apr 1;86(4):774–9.
233. Jilani I, Estey E, Manshuri T, Caligiuri M, Keating M, Giles F, et al. Better detection of FLT3 internal tandem duplication using peripheral blood plasma DNA. *Leukemia.* 2003 Jan 3;17(1):114–9.
234. Rogers A, Joe Y, Manshoury T, Dey A, Jilani I, Giles F, et al. Relative increase in leukemia-specific DNA in peripheral blood plasma from patients with acute myeloid leukemia and myelodysplasia. *Blood.* 2004 Apr 1;103(7):2799–801.
235. Iriyama C, Tomita A, Hoshino H, Adachi-Shirahata M, Furukawa-Hibi Y, Yamada K, et al. Using peripheral blood circulating DNAs to detect CpG global methylation status and genetic mutations in patients with myelodysplastic syndrome. *Biochem Biophys Res Commun.* 2012 Mar 23;419(4):662–9.
236. Mohamedali AM, Alkhatibi H, Kulasekararaj A, Shinde S, Mian S, Malik F, et al. Utility of peripheral blood for cytogenetic and mutation analysis in myelodysplastic syndrome. *Blood.* 2013 Jul 25;122(4):567–70.
237. Albitar F, Ma W, Diep K, De Dios I, Agersborg S, Thangavelu M, et al. Deep Sequencing of Cell-Free Peripheral Blood DNA as a Reliable Method for Confirming the Diagnosis of Myelodysplastic Syndrome. *Genet Test Mol Biomarkers.* 2016 Jul 1;20(7):341–5.
238. Yeh P, Dickinson M, Ftouni S, Hunter T, Sinha D, Wong SQ, et al. Molecular disease monitoring using circulating tumor DNA in myelodysplastic syndromes. *Blood.* 2017 Mar 23;129(12):1685–90.
239. Nakamura S, Yokoyama K, Shimizu E, Yusa N, Kondoh K, Ogawa M, et

- al. Prognostic impact of circulating tumor DNA status post–allogeneic hematopoietic stem cell transplantation in AML and MDS. *Blood*. 2019 Jun 20;133(25):2682–95.
240. Short NJ, Patel KP, Albitar M, Franquiz M, Luthra R, Kanagal-Shamanna R, et al. Targeted next-generation sequencing of circulating cell-free DNA vs bone marrow in patients with acute myeloid leukemia. *Blood Adv*. 2020 Apr 28;4(8):1670–7.
241. Ma W, Kantarjian H, Verstovsek S, Jilani I, Gorre M, Giles F, et al. Hemizygous/homozygous and heterozygous JAK2 mutation detected in plasma of patients with myeloproliferative diseases: correlation with clinical behaviour. *Br J Haematol*. 2006 Aug 1;134(3):341–3.
242. Ma W, Kantarjian H, Zhang X, Sun W, Buller AM, Jilani I, et al. Higher detection rate of JAK2 mutation using plasma. *Blood*. 2008 Apr 1;111(7):3906–7.
243. Choi MY, Kato S, Wang H-Y, Lin JH, Lanman RB, Kurzrock R. JAK2 V617F mutation in plasma cell-free DNA preceding clinically overt myelofibrosis: Implications for early diagnosis. *Cancer Biol Ther*. 2018;19(8):664–8.
244. Barra GB, Santa Rita TH, Almeida ALSC, Jácomo RH, Nery LFA. Serum Has Higher Proportion of Janus Kinase 2 V617F Mutation Compared to Paired EDTA-Whole Blood Sample: A Model for Somatic Mutation Quantification Using qPCR and the 2- $\Delta\Delta$ Cq Method. *Diagnostics*. 2020 Mar 12;10(3):153.
245. Gutierrez-Rodriguez F, Beerman I, Groarke EM, Patel BA, Spitofsky N, Dillon LW, et al. Utility of plasma cell-free DNA for de novo detection and quantification of clonal hematopoiesis. *Haematologica*. 2021 Sep 30;
246. Razavi P, Li BT, Brown DN, Jung B, Hubbell E, Shen R, et al. High-intensity sequencing reveals the sources of plasma circulating cell-free DNA variants. *Nat Med* 2019 2512. 2019 Nov 25;25(12):1928–37.
247. Bellosillo B, Montagut C. High-accuracy liquid biopsies. *Nat Med* 2019 2512. 2019 Nov 25;25(12):1820–1.
248. Risberg B, Tsui DWY, Biggs H, Ruiz-Valdepenas Martin de Almagro A, Dawson SJ, Hodgkin C, et al. Effects of Collection and Processing Procedures on Plasma Circulating Cell-Free DNA from Cancer Patients.

- J Mol Diagnostics. 2018 Nov 1;20(6):883–92.
249. Junker K, Langner K, Klinke F, Bosse U, Thomas M. Grading of Tumor Regression in Non-small Cell Lung Cancer: Morphology and Prognosis. *Chest*. 2001 Nov 1;120(5):1584–91.
250. Salama ME, Swierczek SI, Hickman K, Wilson A, Prchal JT. Plasma quantitation of JAK2 mutation is not suitable as a clinical test: an artifact of storage. *Blood*. 2009 Jul 2;114(1):223–4.
251. Cristiano S, Leal A, Phallen J, Fiksel J, Adleff V, Bruhm DC, et al. Genome-wide cell-free DNA fragmentation in patients with cancer. *Nature*. 2019 Jun 29;570(7761):385–9.
252. Bellosillo B, Martínez-Avilés L, Gimeno E, Florensa L, Longarón R, Navarro G, et al. A higher JAK2 V617F-mutated clone is observed in platelets than in granulocytes from essential thrombocythemia patients, but not in patients with polycythemia vera and primary myelofibrosis. *Leukemia*. 2007 Jun 15;21(6):1331–2.
253. Angona A, Fernández-Rodríguez C, Alvarez-Larrán A, Camacho L, Longarón R, Torres E, et al. Molecular characterisation of triple negative essential thrombocythaemia patients by platelet analysis and targeted sequencing. *Blood Cancer J*. 2016 Aug 26;6(8):e463–e463.
254. Machlus KR, Italiano JE. The incredible journey: From megakaryocyte development to platelet formation. *J Cell Biol*. 2013 Jun 10;201(6):785–96.
255. Wan JCM, Massie C, Garcia-Corbacho J, Mouliere F, Brenton JD, Caldas C, et al. Liquid biopsies come of age: towards implementation of circulating tumour DNA. 2017;17(4):223–38.
256. Marin Oyarzún CP, Carestia A, Lev PR, Glembotsky AC, Castro Ríos MA, Moiraghi B, et al. Neutrophil extracellular trap formation and circulating nucleosomes in patients with chronic myeloproliferative neoplasms. *Sci Rep*. 2016;6:38738.
257. Suzuki Y, Tomita A, Nakamura F, Iriyama C, Shirahata-Adachi M, Shimada K, et al. Peripheral blood cell-free DNA is an alternative tumor DNA source reflecting disease status in myelodysplastic syndromes. *Cancer Sci*. 2016 Sep 25;107(9):1329–37.
258. Wimazal F, Sperr WR, Kundi M, Vales A, Fonatsch C, Thalhammer-

- Scherrer R, et al. Prognostic significance of serial determinations of lactate dehydrogenase (LDH) in the follow-up of patients with myelodysplastic syndromes. *Ann Oncol.* 2008 May 1;19(5):970–6.
259. Shah S, Mudireddy M, Hanson CA, Ketterling RP, Gangat N, Pardanani A, et al. Marked elevation of serum lactate dehydrogenase in primary myelofibrosis: clinical and prognostic correlates. *Blood Cancer J.* 2017 Dec 18;7(12):657.
260. Zhao P, Qin J, Liu W, Zhu Q, Fan T, Xiao H, et al. Using circulating tumor DNA to monitor myelodysplastic syndromes status. Vol. 37, *Hematological Oncology.* John Wiley and Sons Ltd; 2019. p. 531–3.
261. Rossi D, Diop F, Spaccarotella E, Monti S, Zanni M, Rasi S, et al. Diffuse large B-cell lymphoma genotyping on the liquid biopsy. *Blood.* 2017 Apr 6;129(14):1947–57.
262. Suehara Y, Sakata-Yanagimoto M, Hattori K, Nanmoku T, Itoh T, Kaji D, et al. Liquid biopsy for the identification of intravascular large B-cell lymphoma. *Haematologica.* 2018 Jun 3;103(6):e241.
263. Delfau-Larue M-H, van der Gucht A, Dupuis J, Jais J-P, Nel I, Beldi-Ferchiou A, et al. Total metabolic tumor volume, circulating tumor cells, cell-free DNA: distinct prognostic value in follicular lymphoma. *Blood Adv.* 2018 Apr 10;2(7):807–16.
264. Bohers E, Viailly P-J, Becker S, Marchand V, Ruminy P, Maingonnat C, et al. Non-invasive monitoring of diffuse large B-cell lymphoma by cell-free DNA high-throughput targeted sequencing: analysis of a prospective cohort. *Blood Cancer J* 2018 88. 2018 Aug 1;8(8):1–13.
265. Hickmann A-K, Frick M, Hadaschik D, Battke F, Bittl M, Ganslandt O, et al. Molecular tumor analysis and liquid biopsy: a feasibility investigation analyzing circulating tumor DNA in patients with central nervous system lymphomas. *BMC Cancer* 2019 191. 2019 Mar 1;19(1):1–12.
266. Nagy Á, Bátaí B, Balogh A, Illés S, Mikala G, Nagy N, et al. Quantitative Analysis and Monitoring of EZH2 Mutations Using Liquid Biopsy in Follicular Lymphoma. *Genes* 2020, Vol 11, Page 785. 2020 Jul 13;11(7):785.
267. Qi F, Cao Z, Chen B, Chai Y, Lin J, Ye J, et al. Liquid biopsy in extranodal NK/T-cell lymphoma: a prospective analysis of cell-free DNA

BIBLIOGRAPHY

- genotyping and monitoring. *Blood Adv.* 2021 Jun 8;5(11):2505–14.
268. Montalban-Bravo G, Kanagal-Shamanna R, Guerra V, Ramos-Perez J, Hammond D, Shilpa P, et al. Clinical outcomes and influence of mutation clonal dominance in oligomonocytic and classical chronic myelomonocytic leukemia. *Am J Hematol.* 2021 Feb 1;96(2):E50–3.
269. Talati C, Zhang L, Shaheen G, Kuykendall A, Ball M, Zhang Q, et al. Monocyte subset analysis accurately distinguishes CMML from MDS and is associated with a favorable MDS prognosis. *Blood.* 2017 Mar 30;129(13):1881–3.
270. Tarfi S, Harrivel V, Dumezy F, Guy J, Roussel M, Mimoun A, et al. Multicenter validation of the flow measurement of classical monocyte fraction for chronic myelomonocytic leukemia diagnosis. *Blood Cancer J* 2018 811. 2018 Nov 14;8(11):1–10.
271. Saft L, Björklund E, Berg E, Hellström-Lindberg E, Porwit A. Bone marrow dendritic cells are reduced in patients with high-risk myelodysplastic syndromes. *Leuk Res.* 2013 Mar;37(3):266–73.
272. Kerkhoff N, Bontkes HJ, Westers TM, de Gruijl TD, Kordasti S, van de Loosdrecht AA. Dendritic cells in myelodysplastic syndromes: from pathogenesis to immunotherapy. *Immunotherapy.* 2013 Jun;5(6):621–37.
273. Barnholt KE, Hinds DA, Kiefer AK, Do CB, Eriksson N, Mountain JL, et al. Estimation of JAK2 V617F Prevalence by Detection of the Mutation in Saliva Samples From Online MPN and General Population Cohorts. *Blood.* 2012 Nov 16;120(21):1737–1737.
274. Tenedini E, Bernardis I, Artusi V, Artuso L, Roncaglia E, Guglielmelli P, et al. Targeted cancer exome sequencing reveals recurrent mutations in myeloproliferative neoplasms. *Leukemia.* 2013;28(5):1052–9.
275. Hinds DA, Barnholt KE, Mesa RA, Kiefer AK, Do CB, Eriksson N, et al. Germ line variants predispose to both JAK2 V617F clonal hematopoiesis and myeloproliferative neoplasms. *Blood.* 2016 Aug 25;128(8):1121–8.
276. Cianga CM, Antohe I, Zlei M, Constantinescu D, Cianga P. Saliva leukocytes rather than saliva epithelial cells represent the main source of DNA. *Rev Rom Med Lab.* 2016;24(1):31–44.

277. Theda C, Hwang SH, Czajko A, Loke YJ, Leong P, Craig JM. Quantitation of the cellular content of saliva and buccal swab samples. *Sci Rep*. 2018 Dec 1;8(1).
278. Duployez N, Goursaud L, Fenwarth L, Bories C, Marceau-Renaut A, Boyer T, et al. Familial myeloid malignancies with germline TET2 mutation. *Leuk* 2019 345. 2019 Dec 11;34(5):1450–3.
279. Zamora L, Espinet B, Florensa L, Besses C, Bellosillo B, Solé F. Clonality analysis by HUMARA assay in Spanish females with essential thrombocythemia and polycythemia vera. *Haematologica*. 2005 Feb;90(2):259–61.
280. DeRoin L, Cavalcante De Andrade Silva M, Petras K, Arndt K, Phillips N, Wanjari P, et al. Assessing the Feasibility and Limitations of Cultured Skin Fibroblasts for Germline Genetic Testing in Hematologic Disorders. *Blood*. 2020 Nov 5;136(Supplement 1):35–6.
281. Kraft IL, Godley LA. Identifying potential germline variants from sequencing hematopoietic malignancies. *Hematology*. 2020 Dec 4;2020(1):219–27.
282. Pantel K, Alix-Panabières C. Liquid biopsy and minimal residual disease — latest advances and implications for cure. *Nat Rev Clin Oncol*. 2019 Jul 22;16(7):409–24.

FUNDING

The research included in the present thesis has been supported with funding by Instituto de Salud Carlos III (ISCIII-FEDER): PI16/0153, PI19/0005, FI17/00017, Grant Gilead 2016 and Xarxa de Bancs de Tumors de Catalunya (XBTC).

ANNEX I: PRESENTATIONS AT INTERNATIONAL CONFERENCES DERIVED FROM THE THESIS

MOLECULAR CHARACTERIZATION OF PH-NEGATIVE MYELOPROLIFERATIVE NEOPLASMS IN TUMOR CIRCULATING NUCLEIC ACIDS

García-Gisbert N, Fernández-Ibarrondo L, Fernández-Rodríguez C, Camacho L, Angona A, Gibert J, Calvo X, Arenillas L, Longarón R, Andrade M, Salar A, Besses C, Bellosillo B.

Poster presentation

61st American Society of Hematology (ASH) Annual Meeting. December 7-10, 2019. Orlando, EEUU. Blood 134 (suppl 1) Abstract 2953, 2019

OLIGOMONOCYTIC CHRONIC MYELOMONOCYTIC LEUKEMIA (O-CMML) AND CHRONIC MYELOMONOCYTIC LEUKEMIA (CMML) SHOW SIMILAR CLINICAL, MORPHOLOGICAL, IMMUNOPHENOTYPIC AND MOLECULAR FEATURES

Calvo X, **García-Gisbert N**, Parraga I, Florensa L, Montesdeoca S, Fernández Rodríguez C, Salido M, Puiggrós A, Espinet B, Bellosillo B, Colomo L, Roman D, Andrade M, Merchan B, Ferrer del Álamo A, Arenillas L.

Poster presentation

61st American Society of Hematology (ASH) Annual Meeting. December 7-10, 2019. Orlando, EEUU. Blood 134 (suppl 1) Abstract 4266, 2019

HYDROXYUREA REDUCES NEUTROPHIL EXTRACELLULAR TRAP FORMATION IN MYELOPROLIFERATIVE NEOPLASMS.

Cuenca EJ, **García-Gisbert N**, Los Reyes-García A, Aroca C, Martínez-Montesinos L, Andrade M, Aguila S, Guijarro-Carrillo PJ, Cifuentes-Riquelme R, González-Conejero R, Martínez C, Vicente V, Alvarez-Larrán A, Bellosillo B, Teruel-Montoya R, Ferrer-Marin F.

Poster presentation

62nd American Society of Hematology (ASH) Annual Meeting. December 6-9, 2020. Virtual, Blood 134 (suppl 1) Abstract 1267, 2020

MOLECULAR AND CYTOGENETIC CHARACTERIZATION OF MYELOYDYSPLASTIC SYNDROMES IN CELL-FREE DNA.

Garcia-Gisbert N, Merchán B, Garcia-Ávila S, Salido M, Fernández-Rodríguez C, Fernández-Ibarrondo L, Camacho L, Gibert J, Lafuente M, Longarón R, Espinet B, Andrade-Campos M, Arenillas L, Calvo X, Besses C, Salar A, Bellosillo B.

Poster presentation

European Hematology Association (EHA) Congress 2021, Virtual: June 9-17, EHA library 325648; EP890

NON-INVASIVE GENETIC PROFILING AND MONITORING IN MYELOYDYSPLASTIC SYNDROMES

Garcia-Gisbert N, Merchán B, Garcia-Ávila S, Salido M, Fernández-Rodríguez C, Gibert J, Fernández-Ibarrondo L, Camacho L, Lafuente M, Longarón R, Espinet B, Vélez P, Pujol RM, Andrade-Campos M, Arenillas L, Calvo X, Besses C, Salar A, Bellosillo B

Poster presentation

63rd American Society of Hematology (ASH) Annual Meeting. December 11-14, 2021, Atlanta

MULTIPLE *TET2* MUTATIONS AS A NEW BIOLOGICAL CLUE FOR DIFFERENTIATING OLIGOMONOCYTIC CHRONIC MYELOMONOCYTIC LEUKEMIA FROM MYELOYDYSPLASTIC SYNDROMES

Garcia-Gisbert N, Arenillas L, Roman D, Rodríguez-Sevilla JJ, Merchán B, Garcia-Ávila S, Fernández-Rodríguez C, Gibert J, Salar A, Bellosillo B, Ferrer A, Calvo X

Poster presentation

63rd American Society of Hematology (ASH) Annual Meeting. December 11-14, 2021, Atlanta

ANNEX II: SUPPLEMENTARY INFORMATION

Supplementary information.

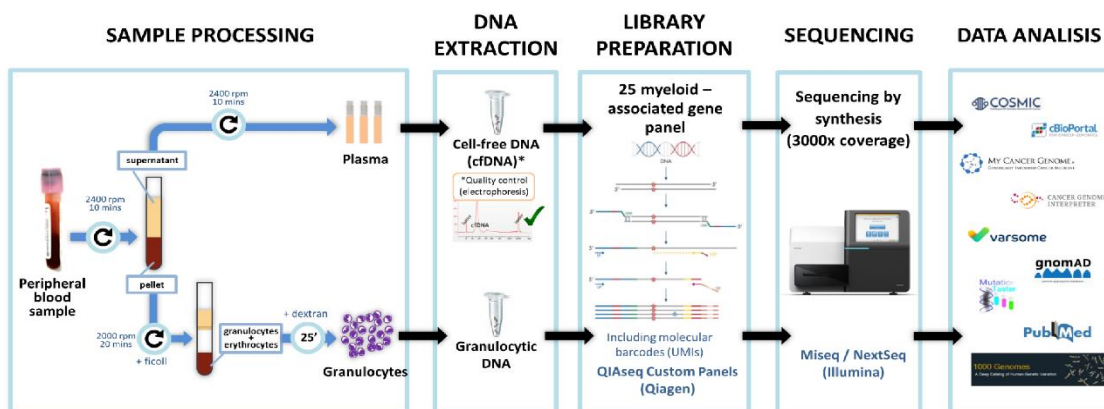
Circulating cell-free DNA improves the molecular characterisation of Ph-negative myeloproliferative neoplasms.

García-Gisbert N^{1,2}, Fernández-Ibarrondo L^{1,2}, Fernández-Rodríguez C^{1,3}, Gibert J¹, Andrade-Campos M^{1,4}, Arenillas L³, Camacho L^{1,3}, Angona A^{1,4}, Longarón R^{1,3}, Salar A^{1,4}, Calvo X³, Besses C^{1,4*}, Bellosillo B^{1,2,3*}

¹Group of Applied Clinical Research in Haematology, Cancer Research Program-IMIM (Hospital del Mar Medical Research Institute) Barcelona, Spain ² Pompeu Fabra University, Barcelona, Spain ³ Department of Pathology, Hospital del Mar-IMIM, Barcelona, Spain ⁴ Department of Haematology, Hospital del Mar-IMIM, Barcelona, Spain

*equally contributed

Supplementary Figure 1. Sample workflow for DNA extraction and mutational analysis.



ANNEX II: SUPPLEMENTARY INFORMATION (CELL-FREE DNA IN MPN)

Supplementary Table 1. Genetic variants detected by NGS (HGVS nomenclature).

PATIENT NUMBER	PHENOTYPE	GENE	MUTATION	PROTEIN CHANGE	TYPE	CONSEQUENCE
1	PV	JAK2	NM_004972.3:c.1849G>T	NP_004963.1:p.(Val617Phe)	snv	missense_variant
2	PV	JAK2	NM_004972.3:c.1849G>T	NP_004963.1:p.(Val617Phe)	snv	missense_variant
3	PV	JAK2	NM_004972.3:c.1849G>T	NP_004963.1:p.(Val617Phe)	snv	missense_variant
4	PV	MPL	NM_005373.2:c.1544G>T	NP_005364.1:p.(Trp515Leu)	snv	missense_variant
4	PV	JAK2	NM_004972.3:c.1849G>T	NP_004963.1:p.(Val617Phe)	snv	missense_variant
4	PV	DNMT3A	NM_175629.2:c.2317C>A	NP_783328.1:p.(Leu773Ile)	snv	missense_variant
5	PV	JAK2	NM_004972.3:c.1849G>T	NP_004963.1:p.(Val617Phe)	snv	missense_variant
6	PV	JAK2	NM_004972.3:c.1849G>T	NP_004963.1:p.(Val617Phe)	snv	missense_variant
7	PV	MPL	NM_005373.2:c.1775G>A	NP_005364.1:p.(Arg592Gln)	snv	missense_variant
7	PV	IDH2	NM_002168.2:c.419G>A	NP_002159.2:p.(Arg140Gln)	snv	missense_variant
7	PV	JAK2	NM_004972.3:c.1849G>T	NP_004963.1:p.(Val617Phe)	snv	missense_variant
8	PV	JAK2	NM_004972.3:c.1849G>T	NP_004963.1:p.(Val617Phe)	snv	missense_variant
9	PV	IDH2	NM_002168.2:c.419G>A	NP_002159.2:p.(Arg140Gln)	snv	missense_variant
9	PV	SRSF2	NM_003016.4:c.284C>G	NP_003007.2:p.(Pro95Arg)	snv	missense_variant
9	PV	JAK2	NM_004972.3:c.1849G>T	NP_004963.1:p.(Val617Phe)	snv	missense_variant
10	PV	ASXL1	NM_015338.5:c.1934dupG	NP_056153.2:p.(Gly646TrpfsTer12)	insertion	frameshift_variant
10	PV	JAK2	NM_004972.3:c.1849G>T	NP_004963.1:p.(Val617Phe)	snv	missense_variant
10	PV	CBL	NM_005188.3:c.1286T>A	NP_005179.2:p.(Ile429Asn)	snv	missense_variant
11	PV	IDH2	NM_002168.2:c.419G>A	NP_002159.2:p.(Arg140Gln)	snv	missense_variant
11	PV	JAK2	NM_004972.3:c.1849G>T	NP_004963.1:p.(Val617Phe)	snv	missense_variant
12	PV	ASXL1	NM_015338.5:c.2535dupC	NP_056153.2:p.(Ser846GlnfsTer5)	insertion	frameshift_variant
12	PV	TEF2	NM_001127208.2:c.4139dupA	NP_001120680.1:p.(His1380GlnfsTer21)	insertion	frameshift_variant
12	PV	JAK2	NM_004972.3:c.1849G>T	NP_004963.1:p.(Val617Phe)	snv	missense_variant
13	PV	ASXL1	NM_015338.5:c.1900_1922delAGAGAGGGCCACCCTGCCAT	NP_056153.2:p.(Glu635ArgfsTer15)	deletion	frameshift_variant
13	PV	JAK2	NM_004972.3:c.1849G>T	NP_004963.1:p.(Val617Phe)	snv	missense_variant
13	PV	ZRSR2	NM_005088.3:c.1119_1120delCT	NP_005080.1:p.(Tyr373Ter)	deletion	stop_gained,frameshift_variant
14	PV	TEF2	NM_001127208.2:c.2881G>T	NP_001120680.1:p.(Glu961Ter)	snv	stop_gained
14	PV	JAK2	NM_004972.3:c.1849G>T	NP_004963.1:p.(Val617Phe)	snv	missense_variant
15	PV	JAK2	NM_004972.3:c.1849G>T	NP_004963.1:p.(Val617Phe)	snv	missense_variant
15	PV	EZH2	NM_004456.4:c.2105C>T	NP_004447.2:p.(Ala702Val)	snv	missense_variant
15	PV	EZH2	NM_004456.4:c.654dupT	NP_004447.2:p.(Pro219SerfsTer3)	insertion	frameshift_variant
16	PV	TEF2	NM_001127208.2:c.5347C>T	NP_001120680.1:p.(Gln1783Ter)	snv	stop_gained
16	PV	JAK2	NM_004972.3:c.1849G>T	NP_004963.1:p.(Val617Phe)	snv	missense_variant
17	PV	ASXL1	NM_015338.5:c.4127dupG	NP_056153.2:p.(Pro1377SerfsTer3)	insertion	frameshift_variant
17	PV	JAK2	NM_004972.3:c.1849G>T	NP_004963.1:p.(Val617Phe)	snv	missense_variant
18	PV	UZF1	NM_006758.2:c.470A>C	NP_006749.1:p.(Gln157Pro)	snv	missense_variant
18	PV	TEF2	NM_001127208.2:c.4138C>T	NP_001120680.1:p.(His1380Tyr)	snv	missense_variant
18	PV	JAK2	NM_004972.3:c.1849G>T	NP_004963.1:p.(Val617Phe)	snv	missense_variant
18	PV	ASXL1	NM_015338.5:c.2955delC	NP_056153.2:p.(Asn985ThrfsTer7)	deletion	frameshift_variant
19	PV	SRSF2	NM_003016.4:c.284C>T	NP_003007.2:p.(Pro95Leu)	snv	missense_variant
19	PV	ASXL1	NM_015338.5:c.1249C>T	NP_056153.2:p.(Arg417Ter)	snv	stop_gained
19	PV	TEF2	NM_001127208.2:c.3232_3233delA/C	NP_001120680.1:p.(Thr1078ProfsTer25)	deletion	frameshift_variant
19	PV	TEF2	NM_001127208.2:c.3733_3737delTACTC	NP_001120680.1:p.(Tyr1245GlyfsTer21)	deletion	frameshift_variant
19	PV	JAK2	NM_004972.3:c.1849G>T	NP_004963.1:p.(Val617Phe)	snv	missense_variant
19	PV	TEF2	NM_001127208.2:c.3662G>C	NP_001120680.1:p.(Cys1221Ser)	snv	missense_variant
20	PV	JAK2	NM_004972.3:c.1849G>T	NP_004963.1:p.(Val617Phe)	snv	missense_variant
21	PV	JAK2	NM_004972.3:c.1849G>T	NP_004963.1:p.(Val617Phe)	snv	missense_variant
22	PV	ASXL1	NM_015338.5:c.2959_2960insT	NP_056153.2:p.(Gly987ValfsTer4)	insertion	frameshift_variant
22	PV	TEF2	NM_001127208.2:c.5079C>G	NP_001120680.1:p.(Tyr1693Ter)	snv	stop_gained
22	PV	JAK2	NM_004972.3:c.1849G>T	NP_004963.1:p.(Val617Phe)	snv	missense_variant
23	PV	JAK2	NM_004972.3:c.1849G>T	NP_004963.1:p.(Val617Phe)	snv	missense_variant
23	PV	TEF2	NM_001127208.2:c.541delA	NP_001120680.1:p.(Ile181PhefsTer2)	deletion	frameshift_variant
24	PV	JAK2	NM_004972.3:c.1849G>T	NP_004963.1:p.(Val617Phe)	snv	missense_variant
24	PV	TEF2	NM_001127208.2:c.5693C>G	NP_001120680.1:p.(Ser1898Cys)	snv	missense_variant
25	PV	JAK2	NM_004972.3:c.1849G>T	NP_004963.1:p.(Val617Phe)	snv	missense_variant
25	PV	TEF2	NM_001127208.2:c.3500+1G>A		snv	splice_donor_variant
26	PV	JAK2	NM_004972.3:c.1849G>T	NP_004963.1:p.(Val617Phe)	snv	missense_variant
27	PV	TP53	NM_000546.5:c.731G>A	NP_000537.3:p.(Gly244Asp)	snv	missense_variant
27	PV	TEF2	NM_001127208.2:c.4138C>T	NP_001120680.1:p.(His1380Tyr)	snv	missense_variant
27	PV	JAK2	NM_004972.3:c.1849G>T	NP_004963.1:p.(Val617Phe)	snv	missense_variant
27	PV	DNMT3A	NM_175629.2:c.878G>T	NP_783328.1:p.(Gly293Val)	snv	missense_variant
28	PV	JAK2	NM_004972.3:c.1849G>T	NP_004963.1:p.(Val617Phe)	snv	missense_variant
29	PV	JAK2	NM_004972.3:c.1849G>T	NP_004963.1:p.(Val617Phe)	snv	missense_variant
30	PV	JAK2	NM_004972.3:c.1849G>T	NP_004963.1:p.(Val617Phe)	snv	missense_variant
31	PV	JAK2	NM_004972.3:c.1849G>T	NP_004963.1:p.(Val617Phe)	snv	missense_variant
32	PV	JAK2	NM_004972.3:c.1849G>T	NP_004963.1:p.(Val617Phe)	snv	missense_variant
32	PV	SRSF2	NM_003016.4:c.130T>C	NP_003007.2:p.(Tyr44His)	snv	missense_variant
33	PV	DNMT3A	NM_175629.2:c.2644C>T	NP_783328.1:p.(Arg882Cys)	snv	missense_variant
33	PV	TEF2	NM_001127208.2:c.1669C>T	NP_001120680.1:p.(Gln557Ter)	snv	stop_gained
33	PV	JAK2	NM_004972.3:c.1849G>T	NP_004963.1:p.(Val617Phe)	snv	missense_variant

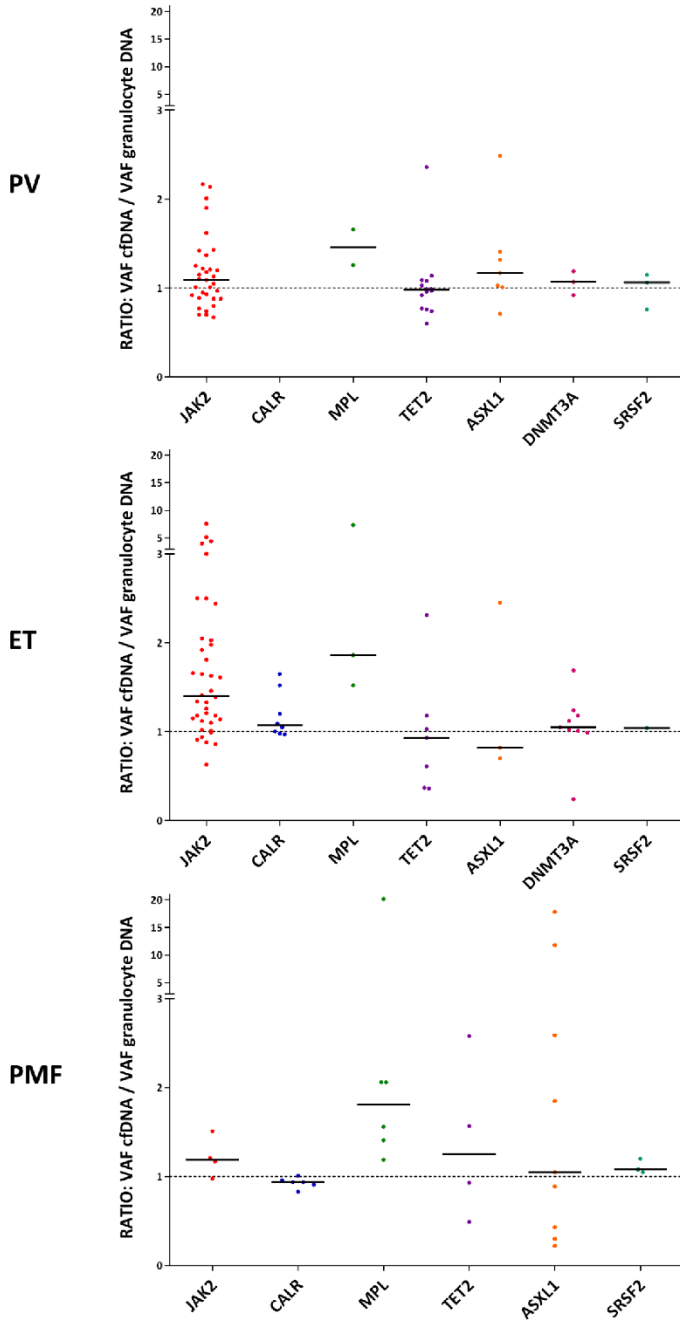
ANNEX II: SUPPLEMENTARY INFORMATION (CELL-FREE DNA IN MPN)

33	PV	TE72	NM_001127208.2:c.1207C>T	NP_001120680.1:p.(Cln403Ter)	snv	stop_gained
33	PV	NRAS	NM_002524.4:c.34G>A	NP_002515.1:p.(Gly12Ser)	snv	missense_variant
1	ET	JAK2	NM_004972.3:c.1849G>T	NP_004963.1:p.(Val617Phe)	snv	missense_variant
2	ET	JAK2	NM_004972.3:c.1849G>T	NP_004963.1:p.(Val617Phe)	snv	missense_variant
3	ET	JAK2	NM_004972.3:c.1849G>T	NP_004963.1:p.(Val617Phe)	snv	missense_variant
4	ET	JAK2	NM_004972.3:c.1849G>T	NP_004963.1:p.(Val617Phe)	snv	missense_variant
5	ET	MPL	NM_005364.1:c.1544G>T	NP_005364.1:p.(Trp515Leu)	snv	missense_variant
5	ET	SF3B1	NM_012433.2:c.1876A>G	NP_036565.2:p.(Asn626Asp)	snv	missense_variant
6	ET	SRSF2	NM_003016.4:c.284C>A	NP_003007.2:p.(Pro95His)	snv	missense_variant
6	ET	SF3B1	NM_012433.2:c.1998G>T	NP_036565.2:p.(Lys666Asn)	snv	missense_variant
6	ET	JAK2	NM_004972.3:c.1849G>T	NP_004963.1:p.(Val617Phe)	snv	missense_variant
6	ET	RUNX1	NM_001754.4:c.493G>T	NP_001745.2:p.(Gly165Cys)	snv	missense_variant
7	ET	JAK2	NM_004972.3:c.1849G>T	NP_004963.1:p.(Val617Phe)	snv	missense_variant
8	ET	TP53	NM_000546.5:c.818G>A	NP_000537.3:p.(Arg273His)	snv	missense_variant
8	ET	CALR	NM_004343.3:c.1154_1155insTTGTC	NP_004334.1:p.(Lys385AsnfsTer47)	insertion	frameshift_variant
9	ET	ASXL1	NM_015338.5:c.1900_1922delAGAGCGGCCACCACCGCCAT	NP_056153.2:p.(Glu635ArgfsTer15)	deletion	frameshift_variant
9	ET	ASXL1	NM_015338.5:c.1934dupG	NP_056153.2:p.(Gly646TrpfsTer12)	insertion	frameshift_variant
9	ET	TE72	NM_001127208.2:c.5602C>T	NP_001120680.1:p.(His1868Tyr)	snv	missense_variant
10	ET	JAK2	NM_004972.3:c.1849G>T	NP_004963.1:p.(Val617Phe)	snv	missense_variant
10	ET	TE72	NM_001127208.2:c.1945C>T	NP_001120680.1:p.(Cln469Ter)	snv	stop_gained
11	ET	JAK2	NM_004972.3:c.1849G>T	NP_004963.1:p.(Val617Phe)	snv	missense_variant
12	ET	CALR	NM_004343.3:c.1092_1143del	NP_004334.1:p.(Leu367ThrfsTer46)	deletion	frameshift_variant
13	ET	JAK2	NM_004972.3:c.1849G>T	NP_004963.1:p.(Val617Phe)	snv	missense_variant
14	ET	DNMT3A	NM_175629.2:c.2390A>G	NP_783328.1:p.(Asn797Ser)	snv	missense_variant
14	ET	CALR	NM_004343.3:c.1092_1143del	NP_004334.1:p.(Leu367ThrfsTer46)	deletion	frameshift_variant
15	ET	CALR	NM_004343.3:c.1154_1155insTTGTC	NP_004334.1:p.(Lys385AsnfsTer47)	insertion	frameshift_variant
16	ET	JAK2	NM_004972.3:c.1849G>T	NP_004963.1:p.(Val617Phe)	snv	missense_variant
16	ET	DNMT3A	NM_175629.2:c.2602T>C	NP_783328.1:p.(Phe868Leu)	snv	missense_variant
17	ET	DNMT3A	NM_175629.2:c.709C>T	NP_783328.1:p.(Cln237Ter)	snv	stop_gained
17	ET	JAK2	NM_004972.3:c.1849G>T	NP_004963.1:p.(Val617Phe)	snv	missense_variant
18	ET	CALR	NM_004343.3:c.1092_1143del	NP_004334.1:p.(Leu367ThrfsTer46)	deletion	frameshift_variant
19	ET	JAK2	NM_004972.3:c.1849G>T	NP_004963.1:p.(Val617Phe)	snv	missense_variant
20	ET	JAK2	NM_004972.3:c.1849G>T	NP_004963.1:p.(Val617Phe)	snv	missense_variant
21	ET	MPL	NM_005364.1:c.1544G>T	NP_005364.1:p.(Trp515Leu)	snv	missense_variant
21	ET	SETBP1	NM_015559.2:c.2602G>A	NP_056374.2:p.(Asp868Asn)	snv	missense_variant
21	ET	U2AF1	NM_006758.2:c.470A>G	NP_006749.1:p.(Cln157Arg)	snv	missense_variant
21	ET	JAK2	NM_004972.3:c.1849G>T	NP_004963.1:p.(Val617Phe)	snv	missense_variant
22	ET	CALR	NM_004343.3:c.1154_1155insTTGTC	NP_004334.1:p.(Lys385AsnfsTer47)	insertion	frameshift_variant
23	ET	JAK2	NM_004972.3:c.1849G>T	NP_004963.1:p.(Val617Phe)	snv	missense_variant
24	ET	JAK2	NM_004972.3:c.1849G>T	NP_004963.1:p.(Val617Phe)	snv	missense_variant
25	ET	ASXL1	NM_015338.5:c.2083delC	NP_056153.2:p.(Cln695AsnfsTer8)	deletion	frameshift_variant
25	ET	JAK2	NM_004972.3:c.1849G>T	NP_004963.1:p.(Val617Phe)	snv	missense_variant
26	ET	JAK2	NM_004972.3:c.1849G>T	NP_004963.1:p.(Val617Phe)	snv	missense_variant
27	ET	JAK2	NM_004972.3:c.1849G>T	NP_004963.1:p.(Val617Phe)	snv	missense_variant
27	ET	PRPF8	NM_006445.3:c.2146G>T	NP_006436.3:p.(Ala716Ser)	snv	missense_variant
28	ET	JAK2	NM_004972.3:c.1849G>T	NP_004963.1:p.(Val617Phe)	snv	missense_variant
29	ET	JAK2	NM_004972.3:c.1849G>T	NP_004963.1:p.(Val617Phe)	snv	missense_variant
30	ET	PRPF8	NM_006445.3:c.4780T>C	NP_006436.3:p.(Cys1594Arg)	snv	missense_variant
30	ET	JAK2	NM_004972.3:c.1849G>T	NP_004963.1:p.(Val617Phe)	snv	missense_variant
30	ET	PRPF8	NM_006445.3:c.2071C>T	NP_006436.3:p.(His691Tyr)	snv	missense_variant
31	ET	JAK2	NM_004972.3:c.1849G>T	NP_004963.1:p.(Val617Phe)	snv	missense_variant
32	ET	DNMT3A	NM_175629.2:c.2645G>A	NP_783328.1:p.(Arg882His)	snv	missense_variant
32	ET	JAK2	NM_004972.3:c.1849G>T	NP_004963.1:p.(Val617Phe)	snv	missense_variant
33	ET	JAK2	NM_004972.3:c.1849G>T	NP_004963.1:p.(Val617Phe)	snv	missense_variant
34	ET	JAK2	NM_004972.3:c.1849G>T	NP_004963.1:p.(Val617Phe)	snv	missense_variant
35	ET	TE72	NM_001127208.2:c.694C>T	NP_001120680.1:p.(Cln232Ter)	snv	stop_gained
35	ET	JAK2	NM_004972.3:c.1849G>T	NP_004963.1:p.(Val617Phe)	snv	missense_variant
36	ET	JAK2	NM_004972.3:c.1849G>T	NP_004963.1:p.(Val617Phe)	snv	missense_variant
37	ET	JAK2	NM_004972.3:c.1849G>T	NP_004963.1:p.(Val617Phe)	snv	missense_variant
38	ET	JAK2	NM_004972.3:c.1849G>T	NP_004963.1:p.(Val617Phe)	snv	missense_variant
39	ET	CALR	NM_004343.3:c.1154delA	NP_004334.1:p.(Lys385ArgfsTer45)	deletion	frameshift_variant
40	ET	DNMT3A	NM_175629.2:c.1958delT	NP_783328.1:p.(Leu653TrpfsTer52)	deletion	frameshift_variant
40	ET	JAK2	NM_004972.3:c.1849G>T	NP_004963.1:p.(Val617Phe)	snv	missense_variant
41	ET	JAK2	NM_004972.3:c.1849G>T	NP_004963.1:p.(Val617Phe)	snv	missense_variant
41	ET	ZRSR2	NM_005089.3:c.840C>G	NP_005080.1:p.(Cys280Trp)	snv	missense_variant
42	ET	DNMT3A	NM_175629.2:c.2645G>A	NP_783328.1:p.(Arg882His)	snv	missense_variant
42	ET	JAK2	NM_004972.3:c.1849G>T	NP_004963.1:p.(Val617Phe)	snv	missense_variant
43	ET	JAK2	NM_004972.3:c.1849G>T	NP_004963.1:p.(Val617Phe)	snv	missense_variant
44	ET	SF3B1	NM_012433.2:c.1997A>G	NP_036565.2:p.(Lys666Arg)	snv	missense_variant
44	ET	TE72	NM_001127208.2:c.4787dupA	NP_001120680.1:p.(Asn1596LysfsTer18)	insertion	frameshift_variant
44	ET	DNMT3A	NM_175629.2:c.2600dupT	NP_783328.1:p.(Phe868IlefsTer9)	insertion	frameshift_variant
44	ET	CALR	NM_004343.3:c.1092_1143del	NP_004334.1:p.(Leu367ThrfsTer46)	deletion	frameshift_variant
45	ET	DNMT3A	NM_175629.2:c.1903delC	NP_783328.1:p.(Arg635GlyfsTer16)	deletion	frameshift_variant
45	ET	JAK2	NM_004972.3:c.1849G>T	NP_004963.1:p.(Val617Phe)	snv	missense_variant

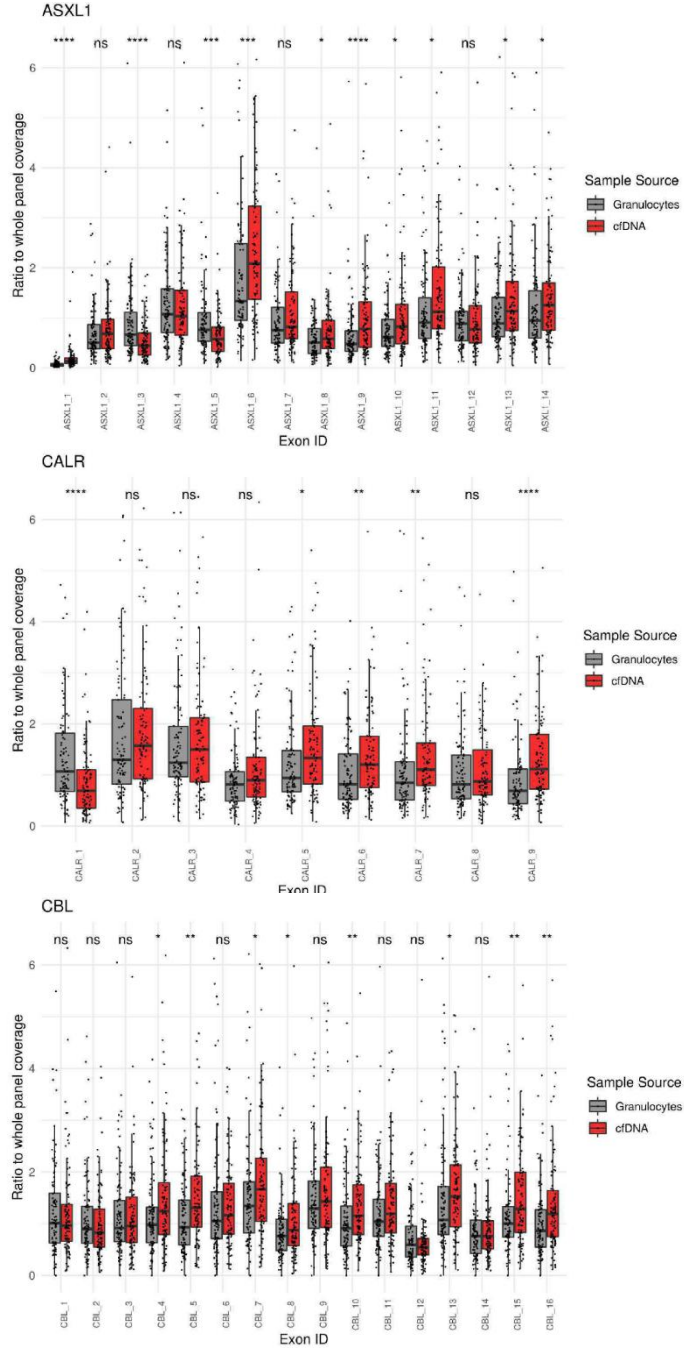
ANNEX II: SUPPLEMENTARY INFORMATION (CELL-FREE DNA IN MPN)

46	ET	TET2	NM_001127208.2:c.2340dupA	NP_001120680.1:p.(Val781SerfsTer8)	insertion	frameshift_variant
46	ET	TET2	NM_001127208.2:c.4636C>T	NP_001120680.1:p.(Gln1546Ter)	snv	stop_gained
46	ET	TET2	NM_001127208.2:c.5011delA	NP_001120680.1:p.(Ser1671ValfsTer24)	deletion	frameshift_variant
46	ET	JAK2	NM_004972.3:c.1849G>T	NP_004963.1:p.(Val617Phe)	snv	missense_variant
47	ET	MPL	NM_005373.2:c.1544G>T	NP_005364.1:p.(Trp515Leu)	snv	missense_variant
47	ET	ASXL1	NM_015338.5:c.1900_1922delAGAGAGGCCGCCACCTGCCAT	NP_056153.2:p.(Glu635ArgfsTer15)	deletion	frameshift_variant
48	ET	CALR	NM_004343.3:c.1092_1143del	NP_004334.1:p.(Leu367ThrfsTer46)	deletion	frameshift_variant
49	ET	JAK2	NM_004972.3:c.1849G>T	NP_004963.1:p.(Val617Phe)	snv	missense_variant
50	ET	JAK2	NM_004972.3:c.1849G>T	NP_004963.1:p.(Val617Phe)	snv	missense_variant
51	ET	JAK2	NM_004972.3:c.1849G>T	NP_004963.1:p.(Val617Phe)	snv	missense_variant
52	ET		no mutations found			
53	ET		no mutations found			
54	ET		no mutations found			
55	ET		no mutations found			
56	ET		no mutations found			
1	PMF	CALR	NM_004343.3:c.1092_1143del	NP_004334.1:p.(Leu367ThrfsTer46)	deletion	frameshift_variant
1	PMF	TET2	NM_001127208.2:c.2746C>T	NP_001120680.1:p.(Gln916Ter)	snv	stop_gained
1	PMF	ASXL1	NM_015338.5:c.1752_1755dupGGTT	NP_056153.2:p.(Lys586GlyfsTer2)	insertion	frameshift_variant
1	PMF	EZH2	NM_004456.4:c.1682G>A	NP_004447.2:p.(Arg561His)	snv	missense_variant
2	PMF	MPL	NM_005373.2:c.1544G>T	NP_005364.1:p.(Trp515Leu)	snv	missense_variant
2	PMF	MPL	NM_005373.2:c.775G>A	NP_005364.1:p.(Glu259Lys)	snv	missense_variant
2	PMF	MPL	NM_005373.2:c.1459A>G	NP_005364.1:p.(Thr487Ala)	snv	missense_variant
3	PMF	CALR	NM_004343.3:c.1092_1143del	NP_004334.1:p.(Leu367ThrfsTer46)	deletion	frameshift_variant
4	PMF	ASXL1	NM_015338.5:c.1934dupG	NP_056153.2:p.(Gly646TrpfsTer12)	insertion	frameshift_variant
4	PMF	TET2	NM_001127208.2:c.3640C>T	NP_001120680.1:p.(Arg1214Trp)	snv	missense_variant
4	PMF	JAK2	NM_004972.3:c.1849G>T	NP_004963.1:p.(Val617Phe)	snv	missense_variant
5	PMF	CALR	NM_004343.3:c.1092_1143del	NP_004334.1:p.(Leu367ThrfsTer46)	deletion	frameshift_variant
6	PMF	ASXL1	NM_015338.5:c.1900_1922delAGAGAGGCCGCCACCTGCCAT	NP_056153.2:p.(Glu635ArgfsTer15)	deletion	frameshift_variant
6	PMF	ASXL1	NM_015338.5:c.1926dupA	NP_056153.2:p.(Gly643ArgfsTer15)	insertion	frameshift_variant
6	PMF	ASXL1	NM_015338.5:c.1969G>T	NP_056153.2:p.(Glu657Ter)	snv	stop_gained
6	PMF	ASXL1	NM_015338.5:c.3187C>T	NP_056153.2:p.(Gln1063Ter)	snv	stop_gained
6	PMF	TET2	NM_001127208.2:c.3947delC	NP_001120680.1:p.(Pro1316GlnfsTer47)	deletion	frameshift_variant
6	PMF	CBL	NM_005188.3:c.1259G>A	NP_005179.2:p.(Arg420Gln)	snv	missense_variant
6	PMF	CALR	NM_004343.3:c.1092_1143del	NP_004334.1:p.(Leu367ThrfsTer46)	deletion	frameshift_variant
7	PMF	NRAS	NM_002524.4:c.35G>A	NP_002515.1:p.(Gly12Asp)	snv	missense_variant
7	PMF	KRAS	NM_00360.2:c.34G>A	NP_203524.1:p.(Gly12Ser)	snv	missense_variant
7	PMF	CALR	NM_004343.3:c.1092_1143del	NP_004334.1:p.(Leu367ThrfsTer46)	deletion	frameshift_variant
8	PMF	CALR	NM_004343.3:c.1092_1143del	NP_004334.1:p.(Leu367ThrfsTer46)	deletion	frameshift_variant
9	PMF	SRSF2	NM_003016.4:c.284C>T	NP_003007.2:p.(Pro95Leu)	snv	missense_variant
9	PMF	ASXL1	NM_015338.5:c.1762C>T	NP_056153.2:p.(Gln588Ter)	snv	stop_gained
10	PMF	MPL	NM_005373.2:c.1514G>A	NP_005364.1:p.(Ser505Asn)	snv	missense_variant
10	PMF	MPL	NM_005373.2:c.1544G>T	NP_005364.1:p.(Trp515Leu)	snv	missense_variant
10	PMF	IDH2	NM_002168.2:c.419G>A	NP_002159.2:p.(Arg140Gln)	snv	missense_variant
10	PMF	SRSF2	NM_003016.4:c.284C>T	NP_003007.2:p.(Pro95His)	snv	missense_variant
10	PMF	JAK2	NM_004972.3:c.1849G>T	NP_004963.1:p.(Val617Phe)	snv	missense_variant
10	PMF	ASXL1	NM_015338.5:c.2637_2638insC	NP_056153.2:p.(Thr880HisfsTer2)	insertion	frameshift_variant
11	PMF	JAK2	NM_004972.3:c.1849G>T	NP_004963.1:p.(Val617Phe)	snv	missense_variant
11	PMF	TET2	NM_001127208.2:c.3880T>C	NP_001120680.1:p.(Tyr1294His)	snv	missense_variant
12	PMF	SRSF2	NM_003016.4:c.284C>A	NP_003007.2:p.(Pro95His)	snv	missense_variant
12	PMF	ASXL1	NM_015338.5:c.1934dupG	NP_056153.2:p.(Gly646TrpfsTer12)	insertion	frameshift_variant
12	PMF	JAK2	NM_004972.3:c.1849G>T	NP_004963.1:p.(Val617Phe)	snv	missense_variant
13	PMF	MPL	NM_005373.2:c.1544G>T	NP_005364.1:p.(Trp515Leu)	snv	missense_variant
13	PMF	SF3B1	NM_012433.2:c.1997A>C	NP_036565.2:p.(Lys666Thr)	snv	missense_variant
14	PMF		no mutations found			
1	uMPN	JAK2	NM_004972.3:c.1849G>T	NP_004963.1:p.(Val617Phe)	snv	missense_variant
2	uMPN	JAK2	NM_004972.3:c.1849G>T	NP_004963.1:p.(Val617Phe)	snv	missense_variant
2	uMPN	PRPF8	NM_006445.3:c.4792G>A	NP_006436.3:p.(Asp1598Asn)	snv	missense_variant
3	uMPN	TET2	NM_001127208.2:c.2890C>T	NP_001120680.1:p.(Gln964Ter)	snv	stop_gained
3	uMPN	TET2	NM_001127208.2:c.5353A>T	NP_001120680.1:p.(Lys1785Ter)	snv	stop_gained
3	uMPN	JAK2	NM_004972.3:c.1849G>T	NP_004963.1:p.(Val617Phe)	snv	missense_variant
3	uMPN	TET2	NM_001127208.2:c.5618T>C	NP_001120680.1:p.(Ile1873Thr)	snv	missense_variant
3	uMPN	EZH2	NM_004456.4:c.2069G>A	NP_004447.2:p.(Arg690His)	snv	missense_variant
4	uMPN	SH2B3	NM_005475.2:c.1566dupC	NP_005466.1:p.(Glu523ArgfsTer23)	insertion	frameshift_variant
4	uMPN	JAK2	NM_004972.3:c.1849G>T	NP_004963.1:p.(Val617Phe)	snv	missense_variant

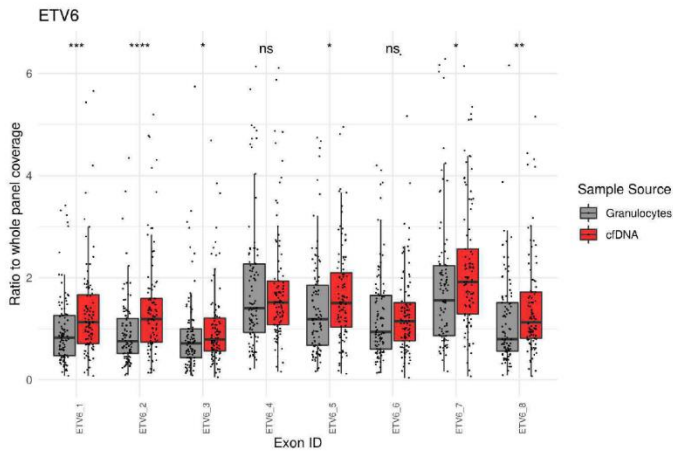
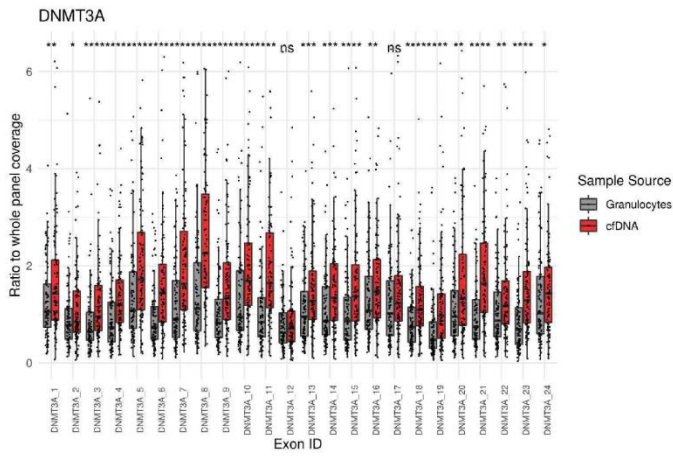
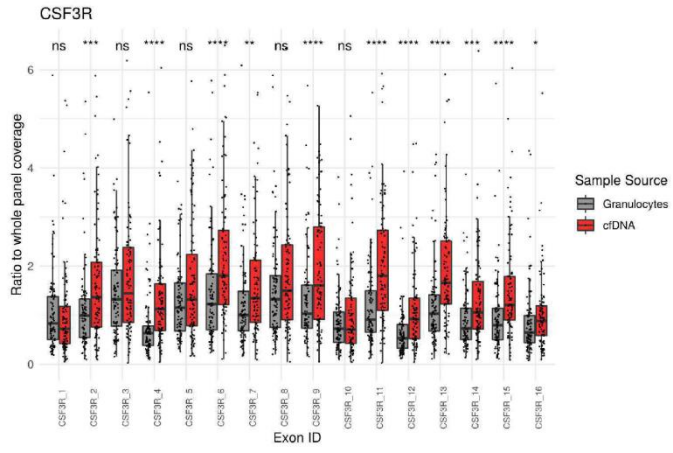
Supplementary Figure 2. Ratio (cfDNA VAF / granulocyte VAF) of the detected variants of the most frequently mutated genes in our cohort, grouped by disease phenotype. Median VAF for each gene is shown as a black line. Variants situated in the plot above the line have a higher VAF in cfDNA than in granulocytes and variants below the line have a higher VAF in granulocytes.

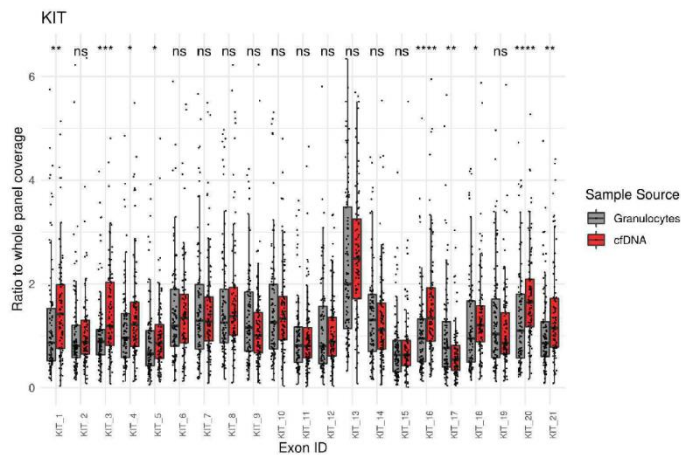
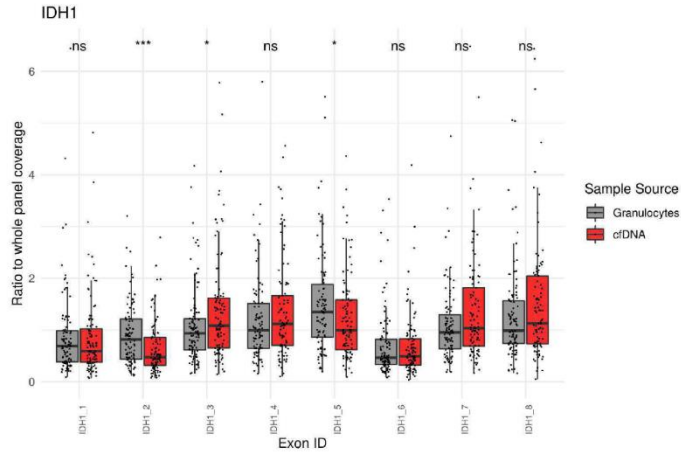
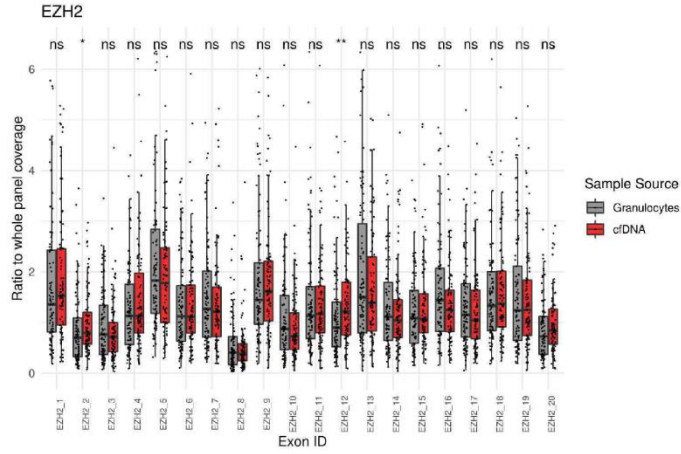


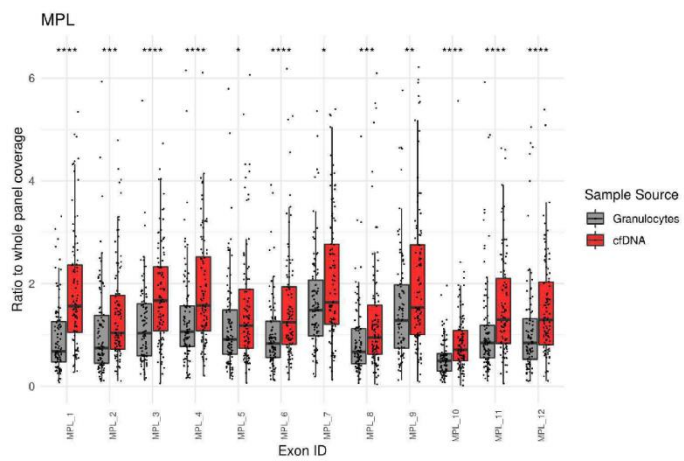
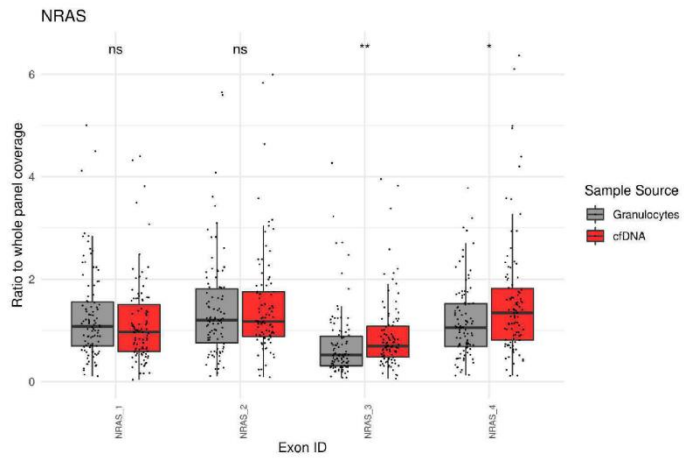
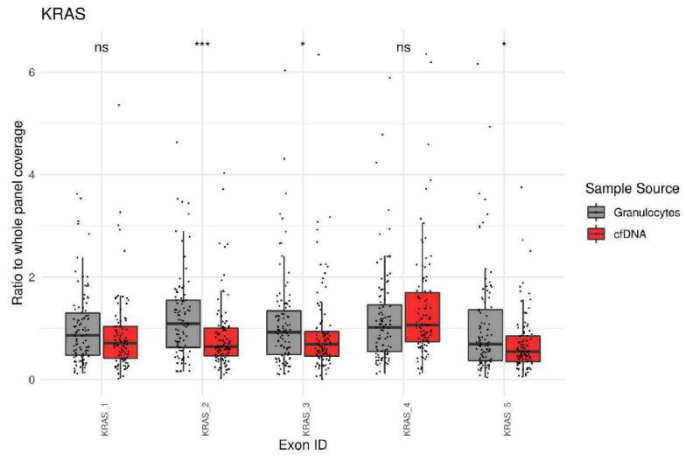
Supplementary Figure 3. Read depth for each gene in granulocytes and cfDNA. The ratio (read depth for the exon/whole panel read depth for that sample) for each exon of the genes included in the NGS panel is shown for the granulocyte samples and the cfDNA samples. * $P \leq 0.05$, ** $P \leq 0.01$, *** $P \leq 0.001$, **** $P \leq 0.0001$.



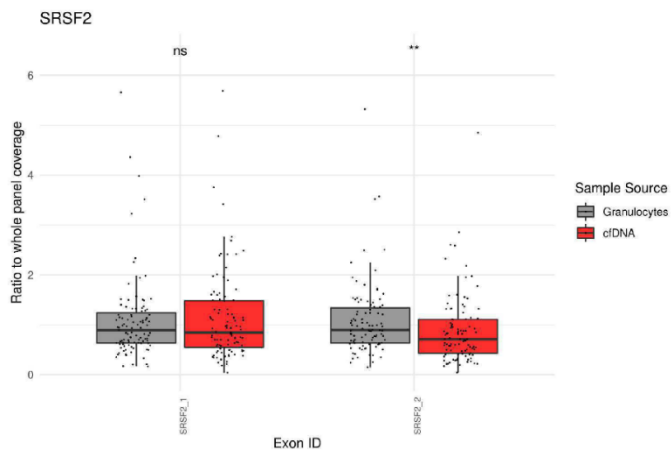
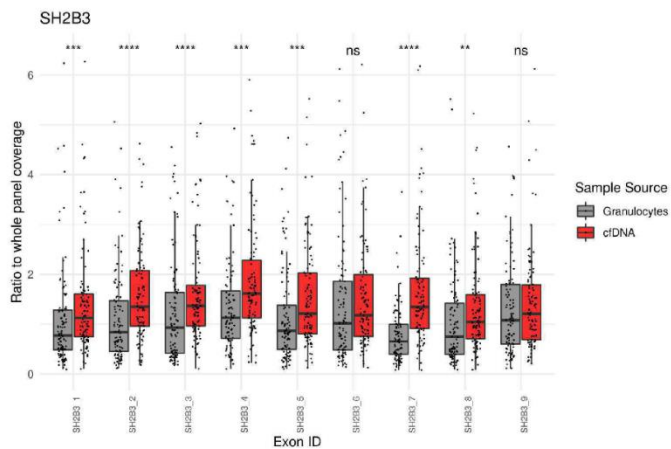
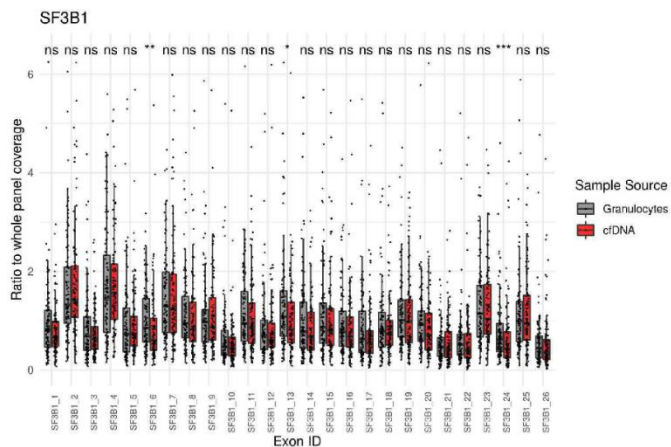
ANNEX II: SUPPLEMENTARY INFORMATION (CELL-FREE DNA IN MPN)



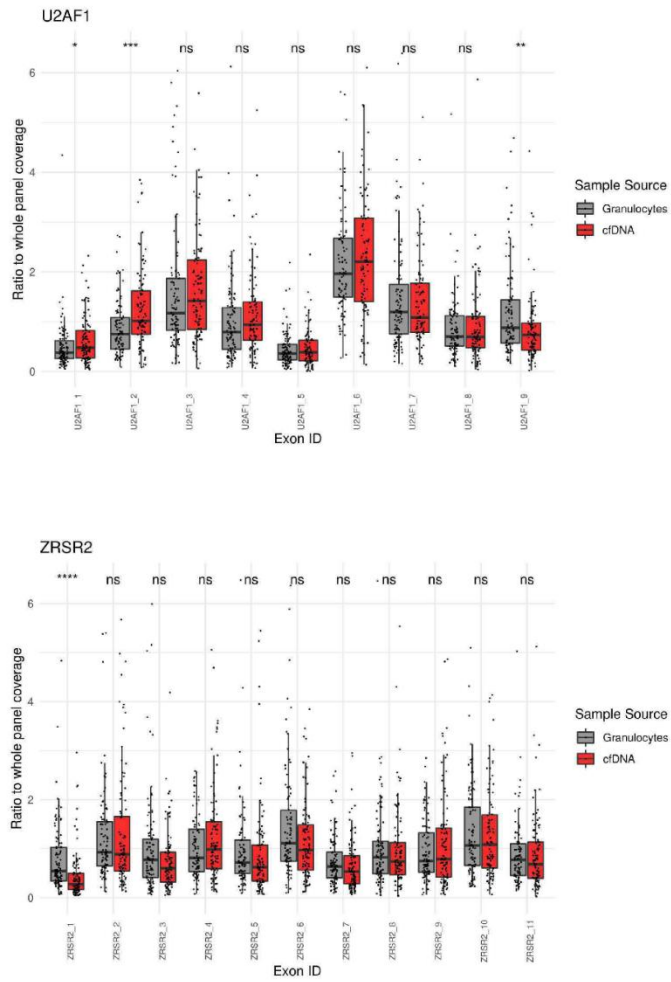




ANNEX II: SUPLEMETARY INFORMATION (CELL-FREE DNA IN MPN)



ANNEX II: SUPPLEMENTARY INFORMATION (CELL-FREE DNA IN MPN)



Supplemental methods. Code used in R 3.6.2 to create the figures.

```
#####Load packages needed for analysis#####
library(readr)
library(maftools)
library(haven)
library(readxl)
library(tidyr)
library(ggplot2)
library(dplyr)

#####Load datasets #####
Mutaciones_cfDNA_MPNS <- read_excel("/Volumes/UUI/Nieves_mielo/Mutaciones_cfDNA_MPNS.xlsx")

MPNS_INCLUIDOSFIS16_TODOS <- read_sav("/Volumes/UUI/Nieves_mielo/MPNS_INCLUIDOSFIS16_TODOS.sav")
MPNS_INCLUIDOSFIS16_TODOS <- MPNS_INCLUIDOSFIS16_TODOS[-1]
MPNS_coverage <- gather(MPNS_INCLUIDOSFIS16_TODOS, condition, Tumor_Sample_Barcode,
DNA_GR:DNA_cfDNA, factor_key=TRUE)
MPNS_INCLUIDOSFIS16_TODOS$DX <- factor(MPNS_INCLUIDOSFIS16_TODOS$DX, levels = c(0,1,2,4, 100), labels =
c("No class", "PV", "ET", "MFP", "Control"))

MPNS_INCLUIDOSFIS16_TODOS$Driver_GEN <- factor(MPNS_INCLUIDOSFIS16_TODOS$Driver_GEN, levels =
c(0,1,2,3,4), labels = c("TN", "JAK", "CALR", "MPL", "MultiDriver"))

colnames(MPNS_INCLUIDOSFIS16_TODOS)[6] <- "Tumor_Sample_Barcode"
#####Recode variables to fit in maftools object#####
Mutaciones_cfDNA_MPNS$Sample <- gsub("QIaseq-DNA-smCounter2.", "", Mutaciones_cfDNA_MPNS$Sample)
Mutaciones_cfDNA_MPNS$Sample <- gsub(".smCounter.anno.", "", Mutaciones_cfDNA_MPNS$Sample)
Mutaciones_cfDNA_MPNS$Sample <- gsub("^.*\\.", "", Mutaciones_cfDNA_MPNS$Sample)
Mutaciones_cfDNA_MPNS$Sample <- gsub("_.*$", "", Mutaciones_cfDNA_MPNS$Sample)

colnames(Mutaciones_cfDNA_MPNS)[2] <- "Hugo_Symbol"
colnames(Mutaciones_cfDNA_MPNS)[3] <- "Chromosome"
colnames(Mutaciones_cfDNA_MPNS)[4] <- "Start_Position"
Mutaciones_cfDNA_MPNS$End_Position <- Mutaciones_cfDNA_MPNS$Start_Position
Mutaciones_cfDNA_MPNS$Reference_Allele <- gsub(">.*$", "", Mutaciones_cfDNA_MPNS$Variant)
Mutaciones_cfDNA_MPNS$Tumor_Seq_Allele2 <- gsub("^.*>.*\\.", "", Mutaciones_cfDNA_MPNS$Variant)
colnames(Mutaciones_cfDNA_MPNS)[19] <- "Variant_Classification"
colnames(Mutaciones_cfDNA_MPNS)[9] <- "Variant_Type"
colnames(Mutaciones_cfDNA_MPNS)[1] <- "Tumor_Sample_Barcode"
unique(Mutaciones_cfDNA_MPNS$Variant_Classification)
Mutaciones_cfDNA_MPNS$Variant_Classification[grepl("missense_variant",
Mutaciones_cfDNA_MPNS$Variant_Classification)] <- "Missense_Mutation"
Mutaciones_cfDNA_MPNS$Variant_Classification[grepl("stop_gained",
Mutaciones_cfDNA_MPNS$Variant_Classification)] <- "Nonsense_Mutation"
Mutaciones_cfDNA_MPNS$Variant_Classification[grepl("frameshift_variant",
Mutaciones_cfDNA_MPNS$Variant_Classification)] <- "Frame_Shift_Ins"
Mutaciones_cfDNA_MPNS$Variant_Classification[grepl("splice_acceptor_variant",
Mutaciones_cfDNA_MPNS$Variant_Classification)] <- "Splice_Site"
Mutaciones_cfDNA_MPNS$Variant_Classification[grepl("splice_donor_variant",
Mutaciones_cfDNA_MPNS$Variant_Classification)] <- "Splice_Site"

Mutaciones_cfDNA_MPNS$Variant_Type[grepl("snv", Mutaciones_cfDNA_MPNS$Variant_Type)] <- "SNP"
Mutaciones_cfDNA_MPNS$Variant_Type[grepl("deletion", Mutaciones_cfDNA_MPNS$Variant_Type)] <- "DEL"
Mutaciones_cfDNA_MPNS$Variant_Type[grepl("insertion", Mutaciones_cfDNA_MPNS$Variant_Type)] <- "INS"
Mutaciones_cfDNA_MPNS$Variant_Type[grepl("mnp", Mutaciones_cfDNA_MPNS$Variant_Type)] <- "SNP"
unique(Mutaciones_cfDNA_MPNS$Classification)
#####Add multihit data for plotting#####
library(plyr); library(dplyr)
multihit_data <- ddpby(Mutaciones_cfDNA_MPNS,.(Tumor_Sample_Barcode,Hugo_Symbol),nrow)
colnames(multihit_data)[3] <- "MultiHit"

Mutaciones_cfDNA_MPNS <- left_join(Mutaciones_cfDNA_MPNS, multihit_data)
#####Load MAF object#####
MPN_MAF <- read.maf(Mutaciones_cfDNA_MPNS, clinicalData = MPNS_INCLUIDOSFIS16_TODOS)

genes <- c('JAK2', 'CALR', 'MPL', 'TET2', 'ASXL1', 'DNMT3A', 'SRSF2', 'IDH2',
'SF3B1', 'EZH2', 'PRPF8', 'CBL', 'NRAS', 'TP53', 'U2AF1',
'ZRSR2', 'KRAS', 'RUNX1', 'SETBP1', 'SH2B3', 'CSF3R', 'ETV6', 'IDH1',
'KIT', 'STAG2')

library(RColorBrewer)
vc_cols = RColorBrewer::brewer.pal(n = 8, name = 'Paired')
```

ANNEX II: SUPLEMETARY INFORMATION (CELL-FREE DNA IN MPN)

```
vc_cols = c("#A6CEE3", "darkolivegreen3", "cornflowerblue", "#33A02C", "coral", "#E31A1C", "#FDBF6F", "#FF7F00")
names(vc_cols) = c(
  'Frame_Shift_Del',
  'Missense_Mutation',
  'Nonsense_Mutation',
  'Multi_Hit',
  'Frame_Shift_Ins',
  'In_Frame_Ins',
  'Splice_Site',
  'In_Frame_Del'
)
CL_colors = c("#D53E4F", "#99D594")
CL_colors = RColorBrewer::brewer.pal(n = 5, name = 'Spectral')
DG_colors = RColorBrewer::brewer.pal(n = 5, name = 'Dark2')

names(CL_colors) = c("No class", "PV", "ET", "MFP", "Control")
names(DG_colors) = c("TN", "JAK", "CALR", "MPL", "MultiDriver")

annot_colors = list(DX = CL_colors, Driver_GEN = DG_colors)

#####Generate Oncoplot#####
pdf("/Volumes/UUI/Nieves_mielo/Oncoplot_ordered.pdf", width = 14)
oncoplot(maf = MPN_MAF, top = 25, colors = vc_cols,
  annotationColor = annot_colors, keepGeneOrder = TRUE,
  sortByAnnotation = TRUE, groupAnnotationBySize = FALSE,
  genes = genes, annotationOrder = c("PV", "ET", "MFP", "No class"),
  removeNonMutated = FALSE,
  SampleNamefontSize = 0.7,
  clinicalFeatures = c("DX", "Driver_GEN"),
  additionalFeature = list(c("MultiHit", "2"), c("MultiHit", "3"),
    c("MultiHit", "4")),
  additionalFeatureCol = c("white", "gray80", "gray40"),
  additionalFeaturePch = c(15, 16, 17))
dev.off()

####Coverage analysis####
colnames(MPNs_coverage)
MPNs_coverage$PANEL <- gsub("CDHS-13593Z-", "", MPNs_coverage$PANEL)
MPNs_coverage$PANEL <- gsub(" miel", "", MPNs_coverage$PANEL)

QIAseq_DNA_panel_CDHS_13593Z_900_roi <- read_table2("/Volumes/UUI/Nieves_mielo/Panels/QIAseq_DNA-
panel_CDHS-13593Z-900_mieloide/QIAseq_DNA_panel_CDHS-13593Z-900.roi.bed",
  col_names = FALSE, skip = 1)
QIAseq_DNA_panel_CDHS_21326Z_924_roi <-
read_table2("/Volumes/UUI/Nieves_mielo/Panels/QIAseqTargetedDnaCustomPanel.CDHS-21326Z-924_mieloide
rutina/QIAseq_DNA_panel_CDHS-21326Z-924.roi.bed",
  col_names = FALSE, skip = 1)

library(dplyr)
QIAseq_joined <- inner_join(QIAseq_DNA_panel_CDHS_13593Z_900_roi,
QIAseq_DNA_panel_CDHS_21326Z_924_roi)
#The only difference between panels is PPM1D so coverage analysis is performed using panel 900 bed file
anti_join(QIAseq_DNA_panel_CDHS_21326Z_924_roi, QIAseq_DNA_panel_CDHS_13593Z_900_roi)
QIAseq_joined$Lenght <- QIAseq_joined$X3 - QIAseq_joined$X2
ggplot(QIAseq_joined, aes(x = X4, y = Lenght))+
  geom_boxplot()+ theme(axis.text.x = element_text(size=7, angle = 90))+
  ylim(0,500)

QIAseq_DNA_panel_CDHS_13593Z_900_roi_cov <-
read_delim("/Volumes/UUI/Nieves_mielo/QIAseq_DNA_panel_CDHS-13593Z-900.roi.cov.txt",
  "\t", escape_double = FALSE, trim_ws = TRUE)

Coverage <- QIAseq_DNA_panel_CDHS_13593Z_900_roi_cov %>% group_by(X4) %>% mutate(id = row_number())
Coverage$Amplicon_ID <- paste0(Coverage$X4, "_", Coverage$id)
Coverage <- Coverage[-c(1,2,3,4)]
columns_names <- colnames(Coverage)
column_names_clean <- columns_names[!columns_names %in% grep("X", columns_names, value = T)]

Coverage_clean <- Coverage[, -grep("min|%", Coverage[1,])]
colnames(Coverage_clean) <- column_names_clean
Coverage_clean <- Coverage_clean[-c(1,2),]
Coverage_clean$Gene <- gsub("_*", "", Coverage_clean$Amplicon_ID)
```

```

Coverage_clean[1:199] <- mutate_all(Coverage_clean[1:199], function(x) as.numeric(as.character(x)))
Panel_mean <- t(as.data.frame(colMeans(Coverage_clean[1:199])))

Coverage_normalized <- Coverage_clean[1:199]/Panel_mean
Coverage_normalized <- cbind(Coverage_clean$Amplicon_ID, Coverage_normalized)

Genes_to_analyze <- t(Coverage_normalized)
colnames(Genes_to_analyze) <- Genes_to_analyze[1,]
Genes_to_analyze <- Genes_to_analyze[-1,]
Genes_to_analyze <- as.data.frame(Genes_to_analyze)
Genes_to_analyze$Sample_names <- rownames(Genes_to_analyze)

Genes_to_analyze$Sample_names <- gsub("_S[0-9]+", "", Genes_to_analyze$Sample_names)

Conditions <- MPNs_coverage[c(203,204)]
Conditions$Tumor_Sample_Barcode <- as.character(Conditions$Tumor_Sample_Barcode)

Genes_to_analyze <- left_join(Genes_to_analyze, Conditions, by=c("Sample_names" = "Tumor_Sample_Barcode"))

Genes_to_analyze$condition <- factor(Genes_to_analyze$condition)
str(Genes_to_analyze$condition)
table(Genes_to_analyze$condition)

data_long <- gather(Genes_to_analyze, Amplicon_ID, measurement, CSF3R_1:STAG2_33, factor_key=TRUE)
data_long$Gene <- gsub("_[0-9]+", "", data_long$Amplicon_ID)
data_long$measurement <- as.numeric(data_long$measurement)
unique(data_long$Gene)
str(data_long)
library(ggpubr)
uniq_species = unique(data_long$Gene)
####Plot coverage for each genen of the panel####
for (i in uniq_species) {

temp_plot <- ggplot(data= subset(data_long, Gene == i), aes(x=Amplicon_ID, y=measurement, fill=condition))+
  geom_boxplot(outlier.shape = NA)+
  geom_jitter(size=0.01, position = position_jitterdodge(jitter.width = 0.3, dodge.width = 0.8))+
  # coord_flip()+
  ylim(0,6.5)+
  stat_compare_means(paired = TRUE, size = 4,
    aes(label = .p.signif.), label.x = 0.5, label.y = 6.4)+
  ylab("Ratio to whole panel coverage")+
  xlab("Exon ID")+
  ggtitle(i) + theme_minimal() +
  theme(axis.text.x =element_text(size=7, angle = 90))+
  scale_fill_manual(values=c("#999999", "firebrick1"), name = "Sample Source", labels = c("Granulocytes", "cfDNA"))

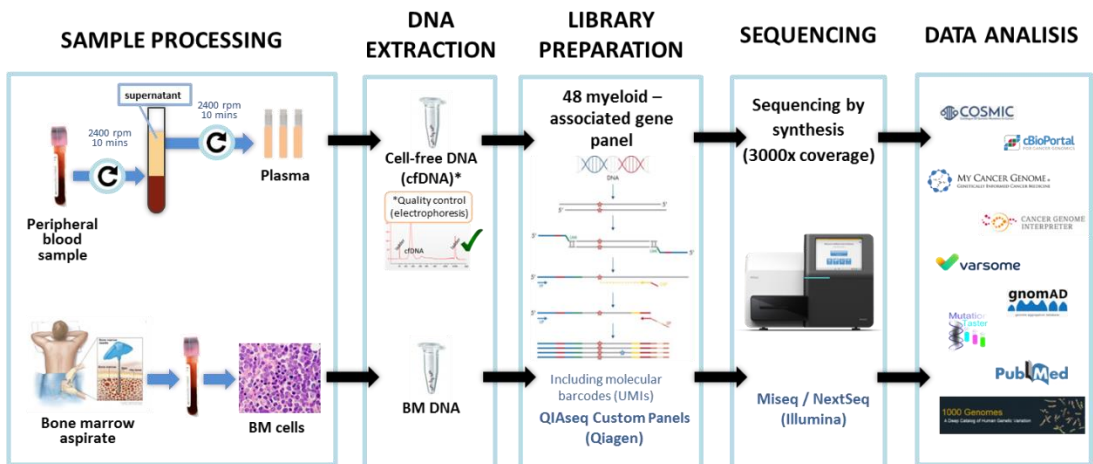
ggsave(temp_plot, file=paste0("/Volumes/UUI/Nieves_mielo/Coverage_", i, ".pdf"))
}

```

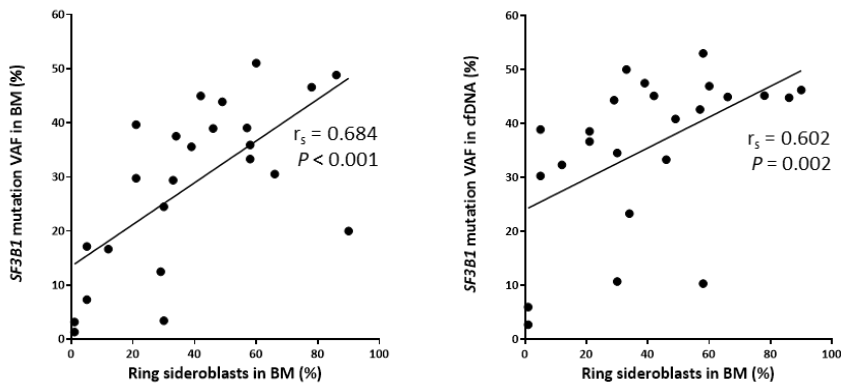

SUPPLEMENTARY INFORMATION

MOLECULAR AND CYTOGENETIC CHARACTERIZATION OF MYELODYSPLASTIC SYNDROMES IN CELL-FREE DNA

Garcia-Gisbert N^{1,2}, Garcia-Ávila S^{1,3}, Merchán B¹, Salido M^{4,5}, Fernández-Rodríguez C^{1,5}, Gibert J¹, Fernández-Ibarrondo L^{1,2}, Camacho L^{1,5}, Lafuente M^{1,2}, Longarón R^{1,5}, Espinet B^{4,5}, Vélez P^{1,3}, Pujol RM⁶, Andrade-Campos M¹, Arenillas L^{4,5}, Salar A^{1,3}, Calvo X^{4,5}, Besses C¹, Bellosillo B^{1,5}



Supplementary Figure 1. Sample workflow for DNA extraction and mutational analysis.



Supplementary Figure 2. Correlation between percentage of ring sideroblasts in bone marrow and VAFs of *SF3B1* mutations in BM and cfDNA.



Supplementary Figure 3. CNV results by NGS in a patient with 20q- and 5q- alterations. Results of the coverage analysis of *EGR1* (chr5) and *TP53TG5* (chr20) genes are shown, which were included in the design of the NGS gene panel to cover chr5 and chr20 chromosomal aberrations. Each dot in the plot represents a genomic region covered by the gene panel. The green line shows the normal values (two copies of the genomic region). Dots above 2 indicate a potential gain of genetic material and dots below 2 indicate a potential loss of genetic material. A) CNV analysis of *EGR1* and *TP53TG5* in BM DNA. B) CNV analysis of *EGR1* and *TP53TG5* in cfDNA C) CNV analysis of a patient with normal karyotype. D)CMA results confirming the 5q- and 20q- in the patient.

Supplementary Table 1. WHO 2017 classification of AML patients (Arber et al, Blood, 2016).

PATIENT	WHO 2017 CLASSIFICATION
1	AML with minimal differentiation
2	AML with myelodysplasia-related changes
3	AML with myelodysplasia-related changes
4	AML with recurrent genetic abnormalities (<i>biallelic mutations of CEBPA</i>)
5	AML with recurrent genetic abnormalities (<i>NPM1</i> mutated)
6	AML with recurrent genetic abnormalities (<i>NPM1</i> mutated)
7	AML with recurrent genetic abnormalities (<i>NPM1</i> mutated)
8	AML with recurrent genetic abnormalities (<i>NPM1</i> mutated)
9	AML with recurrent genetic abnormalities (<i>NPM1</i> mutated)
10	AML with recurrent genetic abnormalities, t(8;21)(q22;q22.1);RUNX1-RUNX1T1
11	AML, NOS
12	AML, NOS, Pure erythroid leukemia
13	AML, NOS, with maturation
14	AML, NOS, without maturation
15	AML, NOS, without maturation
16	Therapy-related myeloid neoplasm
17	Therapy-related myeloid neoplasm
18	Therapy-related myeloid neoplasm

Supplementary Table 2. Genes and genomic regions included in the NGS panel design.

GENES	CHROMOSOME LOCATION	COVERED REGION
<i>ASXL1</i>	20q11.21	Full exonic region
<i>ATM</i>	11q22.3	Full exonic region
<i>BCOR</i>	Xp11.4	Full exonic region
<i>BCORL1</i>	Xq26.1	Full exonic region
<i>CALR</i>	19p13.13	Exon 9
<i>CBL</i>	11q23.3	Full exonic region
<i>CEBPA</i>	19q13.11	Full exonic region
<i>CHEK2</i>	22q12.1	Full exonic region
<i>CSF3R</i>	1p34.3	Full exonic region
<i>CSNK1A1</i>	5q32	Full exonic region
<i>CUX1</i>	7q22.1	Full exonic region
<i>DDX41</i>	5q35.3	Full exonic region
<i>DLEU7</i>	13q14.3	Full exonic region
<i>DNMT3A</i>	2p23.3	Full exonic region
<i>EGR1</i>	5q31.2	Full exonic region
<i>ETV6</i>	12p13.2	Full exonic region
<i>EZH2</i>	7q36.1	Full exonic region
<i>FLT3</i>	13q12.2	Full exonic region
<i>GATA2</i>	3q21.3	Full exonic region
<i>IDH1</i>	2q34	Exon 4
<i>IDH2</i>	15q26.1	Exon 4
<i>JAK2</i>	9p24.1	Full exonic region
<i>KIT</i>	4q12	Exon 17
<i>KMT2A</i>	11q23.3	Full exonic region
<i>KRAS</i>	12p12.1	Full exonic region
<i>MPL</i>	1p34.2	Full exonic region
<i>NF1</i>	17q11.2	Full exonic region
<i>NPM1</i>	5q35.1	Full exonic region
<i>NRAS</i>	1p13.2	Full exonic region
<i>PHF6</i>	Xq26.2	Full exonic region
<i>PPM1D</i>	17q23.2	Full exonic region
<i>PRPF8</i>	17p13.3	Full exonic region
<i>PTPN11</i>	12q24.13	Full exonic region
<i>RAD21</i>	8q24.11	Full exonic region
<i>RUNX1</i>	21q22.12	Full exonic region
<i>SETBP1</i>	18q12.3	Full exonic region
<i>SF3B1</i>	2q33.1	Exons 14,15,16
<i>SH2B3</i>	12q24.12	Full exonic region
<i>SRSF2</i>	17q25.1	Full exonic region
<i>STAG2</i>	Xq25	Full exonic region
<i>TET2</i>	4q24	Full exonic region
<i>TNFSF11</i>	13q14.11	Full exonic region
<i>TP53</i>	17p13.1	Full exonic region
<i>TP53RK</i>	20q13.12	Full exonic region
<i>TP53TG5</i>	20q13.12	Full exonic region
<i>U2AF1</i>	21q22.3	Full exonic region
<i>WT1</i>	11p13	Full exonic region
<i>ZRSR2</i>	Xp22.2	Full exonic region
Polymorphic region close to <i>EGR1</i>	5q31.2	chr5:137805574-137805662
Polymorphic region in locus <i>D7S486 (1)</i>	7q3.1	chr7:115814732-115815082
Polymorphic region in locus <i>D7S486 (2)</i>	7q3.1	chr7:115825276-115825301
Polymorphic region in locus <i>D7S486 (3)</i>	7q3.1	chr7:115900252-115900830
Polymorphic region in locus <i>D7S486 (4)</i>	7q3.1	chr7:115948439-115949024
Polymorphic region in locus <i>D7S486 (5)</i>	7q3.1	chr7:115953508-115953582

OLIGOMONOCYTIC AND OVERT CHRONIC MYELOMONOCYTIC LEUKEMIA SHOW SIMILAR CLINICAL, GENOMIC AND IMMUNOPHENOTYPIC FEATURES

Xavier Calvo^{1*}, Nieves Garcia-Gisbert^{2,3*}, Ivonne Parraga¹, Joan Gibert², Lourdes Florensa¹, Marcio Andrade-Campos⁴, Brayan Merchan⁴, Sara Garcia-Avila⁴, Sara Montesdeoca¹, Concepción Fernández-Rodríguez², Marta Salido⁵, Anna Puiggros⁵, Blanca Espinet⁵, Luís Colomo^{6,3}, David Roman-Bravo¹, Beatriz Bellosillo^{2,3}, Ana Ferrer^{1,3}, Leonor Arenillas¹

SUPPLEMENTAL DATA

Morphological Studies

At least, two bone marrow and one peripheral blood May–Grünwald–Giemsa-stained smears were used for conducting the morphologic analysis. In addition, a Prussian blue-stained bone marrow smear was used for assessing the percentage of ring sideroblasts. The WHO 2017 proposals for evaluating the morphological diagnosis of myelodysplastic syndrome and chronic myelomonocytic leukemia were followed strictly. As recommended, peripheral blood and bone marrow differential counts were performed on at least 200 and 500 cells, respectively. Following the 2017 WHO recommendations, the threshold used for considering a myeloid cell line as dysplastic was the presence of $\geq 10\%$ abnormal cells in the corresponding myeloid lineage. For the evaluation of dysplasia, at least 200 neutrophils, 200 erythroblasts, and 30 megakaryocytes were assessed in bone marrow. Multilineage dysplasia was defined by dysplasia involving two or more lineages. Morphological evaluation of monocytes and their precursors was made following the current consensus document¹, allowing the differentiation of 4 subtypes: monoblast, promonocyte, immature monocyte, and mature monocyte. As currently recommended by the Spanish Guidelines for the diagnosis and treatment of myelodysplastic syndromes and chronic myelomonocytic leukemia, bone marrow biopsy was conducted only in those cases where fibrosis, hypoplastic myelodysplastic syndromes, or idiopathic cytopenias of undetermined significance were suspected.

1. Goasguen JE, Bennett JM, Bain BJ, et al. Morphological evaluation of monocytes and their precursors. *Haematologica*. 2009;94(7):994-997.

Conventional Cytogenetic Studies

Cytogenetic analyses were performed on G-banded chromosomes obtained from 24 h unstimulated bone marrow cultures. When possible, at least 20 metaphases per sample were studied. Karyotypes were described according to the International System for Human Cytogenetic Nomenclature.

Next-Generation Sequencing

The DNA obtained from total PB or BM was quantified by Qubit fluorometer (Thermo Fisher Scientific, Carlsbad, USA). 40 ng of DNA were required for library preparation. Targeted amplicon libraries (QIAseq Custom DNA Panels, Qiagen, Hilden, Germany) were prepared using a custom panel covering the full exonic regions of 25 genes associated with myeloid

malignancies (*ASXL1*, *CALR*, *CBL*, *CSF3R*, *DNMT3A*, *ETV6*, *EZH2*, *IDH1*, *IDH2*, *JAK2*, *KIT*, *KRAS*, *MPL*, *NRAS*, *PRPF8*, *RUNX1*, *SETBP1*, *SF3B1*, *SH2B3*, *SRSF2*, *STAG2*, *TET2*, *TP53*, *U2AF1*, *ZRSR2*). Library preparation incorporated molecular barcoding technology to tag individual DNA molecules, which enables variant detection with high confidence by avoiding false positives, PCR artifacts and library bias. Libraries were sequenced with 2x150-bp paired-end reads using either MiSeq or NextSeq (Illumina, San Diego, CA, USA) with a 2000x minimum coverage.

Sequencing files were processed using the GeneGlobe Data Analysis Center (Qiagen) for FASTQ trimming, alignment to the reference genome and generation of variant calling files (.vcf) (smCounter2, Qiagen). The obtained variants were then annotated and classified using Illumina VariantStudio 3.0 software according to genomic databases (GenomAD, Varsome, cBioPortal, dbSNP, COSMIC, My Cancer Genome, Cancer Genome Interpreter) and evidence of pathogenicity in the literature. Variants were classified into five groups: benign, likely benign, unknown significance, likely pathogenic and pathogenic^{1,2}. Only variants classified as pathogenic or likely pathogenic were included in this study. The limit of detection established for variant detection was 2% variant allele frequency (VAF). In cases with low VAF, variants were confirmed visually using the Integrative Genomics Viewer (IGV) v2.4 software.

1. Palomo L, Ibáñez M, Abáigar M, et al. Spanish Guidelines for the use of targeted deep sequencing in myelodysplastic syndromes and chronic myelomonocytic leukaemia. *Br. J. Haematol.* 2020;188(5):605–622.
2. Li MM, Datto M, Duncavage EJ, et al. Standards and Guidelines for the Interpretation and Reporting of Sequence Variants in Cancer: A Joint Consensus Recommendation of the Association for Molecular Pathology, American Society of Clinical Oncology, and College of American Pathologists. *J Mol Diagn.* 2017;19(1):4-23.

Flow cytometry analysis of monocyte subsets in peripheral blood

Multiparametric flow cytometry analysis of monocyte subsets was performed on whole peripheral blood collected on EDTA. Based on Euroflow Consortium recommendations we follow the stain-lyse-wash procedure with FACS Lysing Solution (BD Biosciences, CA, USA). Cell surface staining of 2×10^6 cells was performed and at least 500,000 total events were acquired per tube (FACS Canto II, BD Biosciences). A 4-color experimental panel with five tubes was run for all samples. Tube 1: CD14 (FITC; clone MΦP9), CD16 (PE; clone B73.1), CD45 (PerCP-Cy5.5; clone 2D1), CD33 (APC; clone P67.6); tube 2: CD64 (FITC; clone 10.1), CD56 (PE; clone NCAM16.2), HLA-DR (PerCP-Cy5.5; clone G46-6), CD33 (APC; clone P67.6); tube 3: CD2 (FITC; clone S5.2), CD7 (PE; clone M-T701), CD45 (PerCP-Cy5.5; clone 2D1), CD33 (APC; clone P67.6); tube 4: CD56 (FITC; clone NCAM16.2), CD123 (PE; clone 9F5), HLA-DR (PerCP-Cy5.5; clone G46-6), CD45 (APC; clone 2D1); tube 5: CD2 (FITC; clone S5.2), CD14 (PE; clone MΦP9), HLA-DR (PerCP-Cy5.5; clone G46-6), CD33 (APC; clone P67.6) (all antibodies from BD Biosciences, San Jose, CA). Analysis was performed with Infinicyt version 1.7 software (Cytognos SL).

Briefly, we excluded doublets (FSC-A/FSC-H dot plot), *debris* (FSC/SSC and CD45/SSC dot plots) and NK-cells (CD16+, CD33-, and FSC/SSC lymphocyte-gate). A proper strategy for excluding NK-cells is crucial, since some of these may be difficult to differentiate from the nonclassical monocyte subset (CD14- or dim/CD16+) (Supplemental Figure 1a). For this purpose, we selected the CD16-B73.1 antibody since this binds to CD16-positive neutrophils with lower intensity when compared with some other CD16-specific antibodies (e.g.: 3G8, VEP13, NKP15, GO22)¹. By using this, the CD16 positive populations of monocytes and NK-cells are better distinguished from neutrophils. The monocyte gate was determined based on a CD45 vs SSC plot and CD33 vs SSC plot. Since nonclassical monocytes show dimmer CD33 and brighter CD45 expression than the rest of monocytes, it is important to draw a wide gate on CD33 vs SSC plot to avoid the loss of this population (Supplemental Figure 1b). Next, as in Selimoglu-Buet et al, CD14- or dim/CD16- cells were excluded (mainly myeloid dendritic cells are removed in this step) and finally, the resulting monocyte population was assessed for CD14 and CD16 expression. By this method, we were able to establish the percentage of classical monocytes (MO1: CD14+/CD16-), intermediate monocytes (MO2: CD14+/CD16+), and nonclassical monocytes (MO3: CD14- or dim/CD16+) from the total monocyte population (Supplemental Figure 1c). In addition, we assessed the expression of CD56 (Supplemental Figure 1c), CD7 and CD2 in monocytes (cutoff positivity $\geq 20\%$). The fifth tube of our protocol was especially designed to quantify myeloid dendritic cells (CD33+ bright, HLA-DR+ bright, CD14-, CD2+). To properly assess CD2 expression in monocytes, myeloid dendritic cells must be removed, since these are CD33+ bright, like monocytes, and express CD2 (Supplemental Figure 1d). The fourth tube was designed to quantify plasmacytoid dendritic cells and to evaluate CD56 expression on these (Supplemental Figure 1e).

1. Perussia B, Trinchieri G, Jackson A, et al. The Fc receptor for IgG on human natural killer cells: phenotypic, functional, and comparative studies with monoclonal antibodies. *J Immunol.* 1984 Jul;133(1):180-9.

Supplemental Figure 1. Gating strategy used to identify monocyte subsets in peripheral blood and to assess CD56 and CD2 expression in monocytes (a-d). Gating strategy used to identify myeloid dendritic cells (mDCs) in peripheral blood (d). Gating strategy used to identify plasmacytoid dendritic cells (pDCs) in peripheral blood (e).

Figure 1a

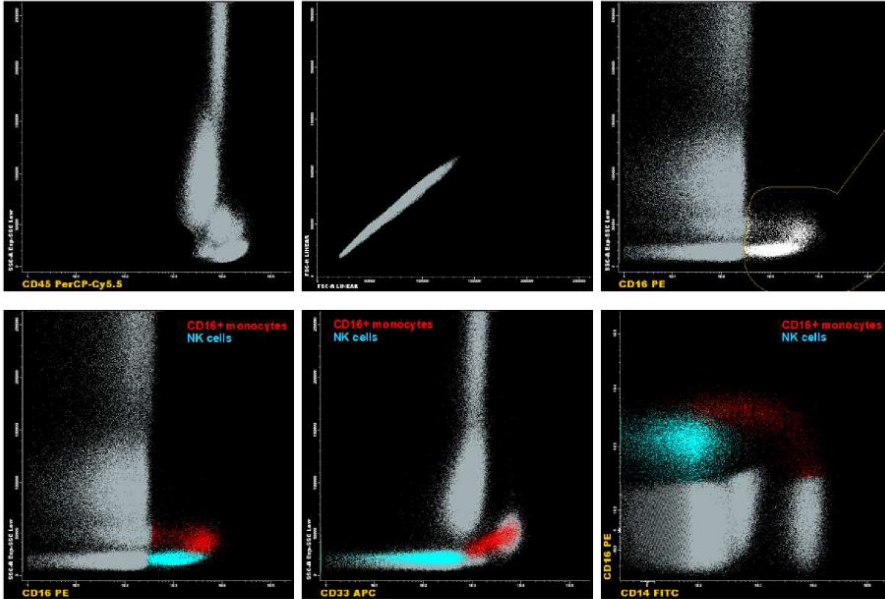


Figure 1b

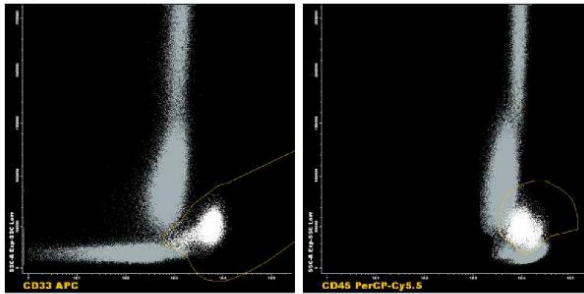


Figure 1c

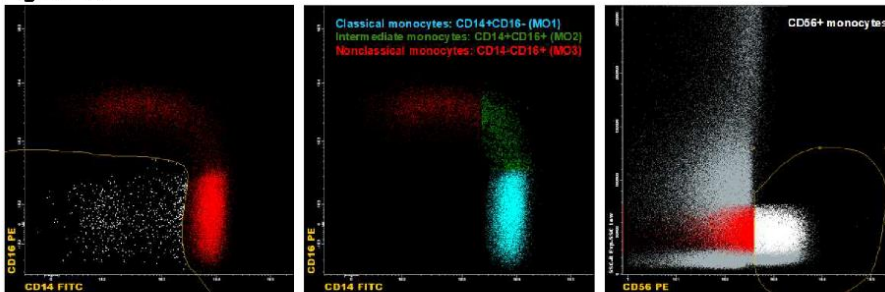


Figure 1d

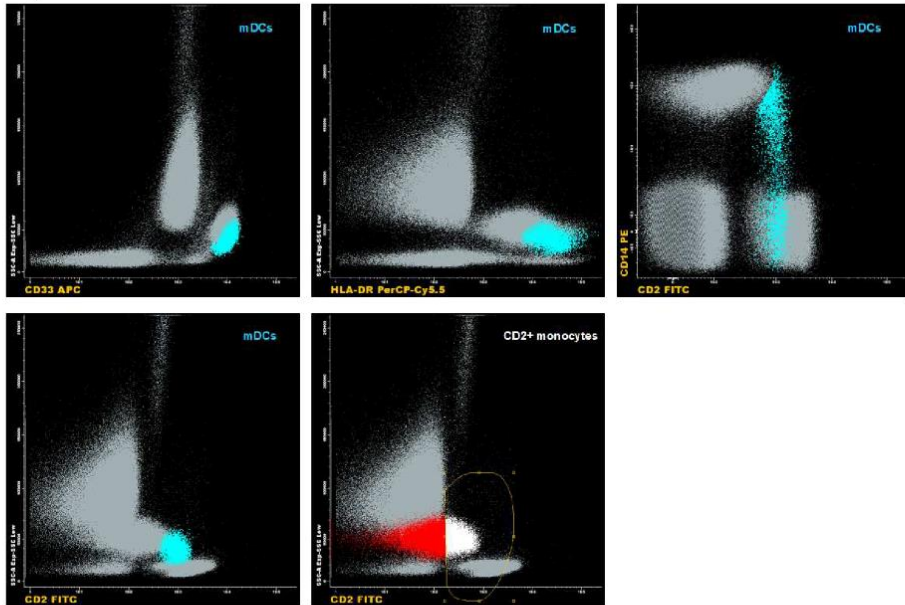
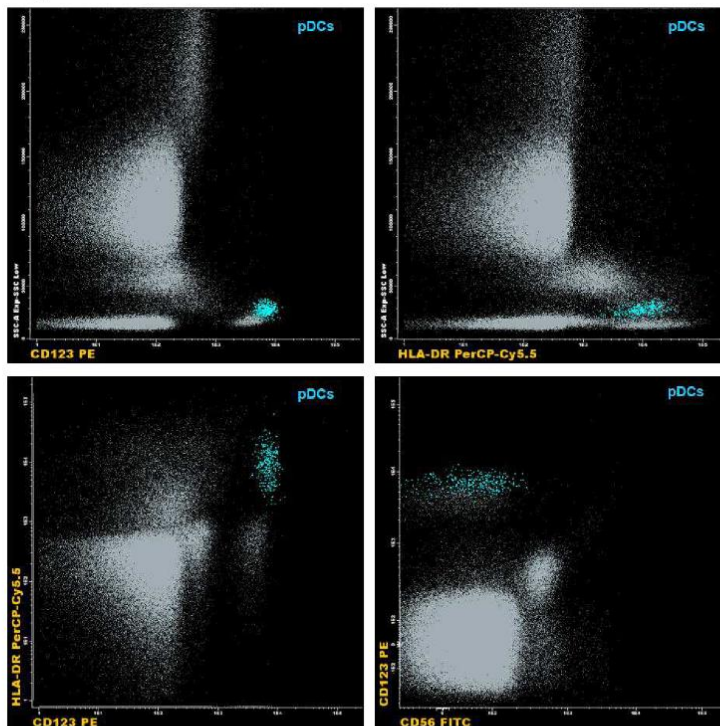
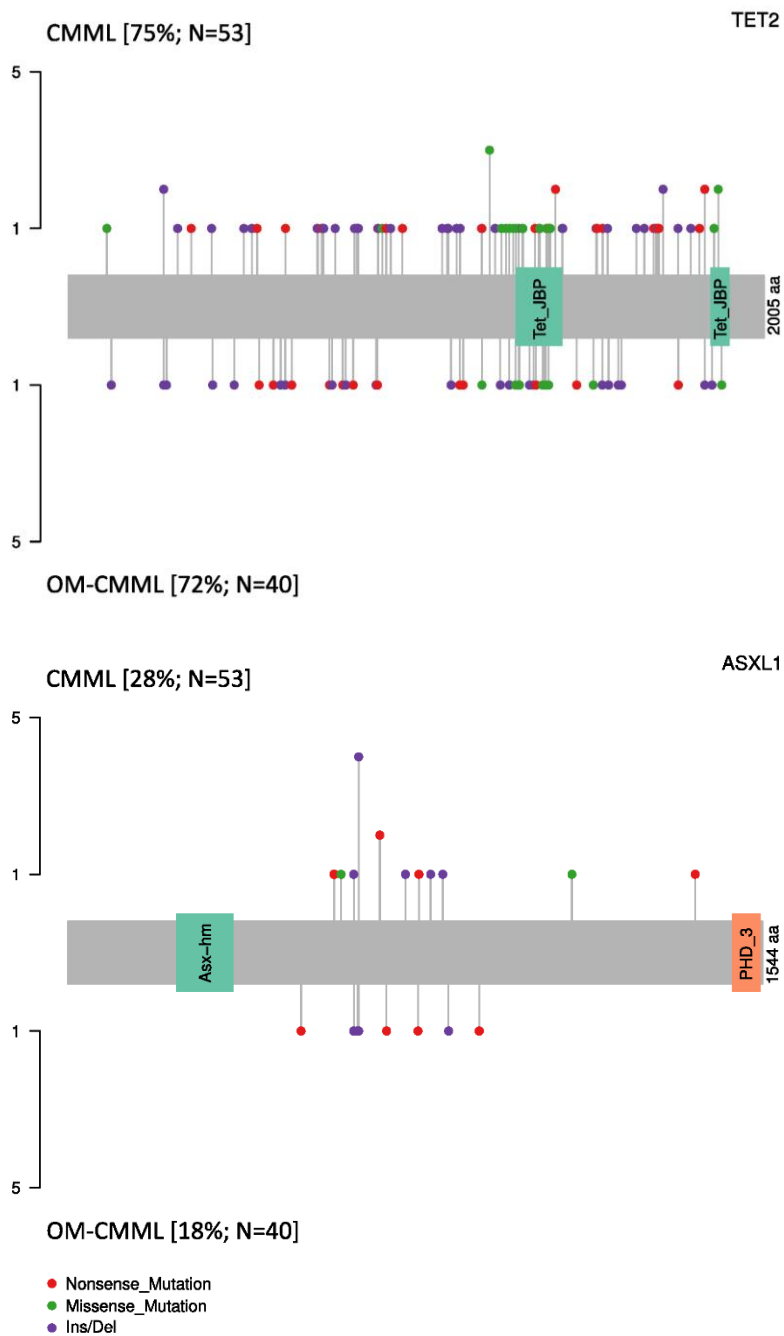
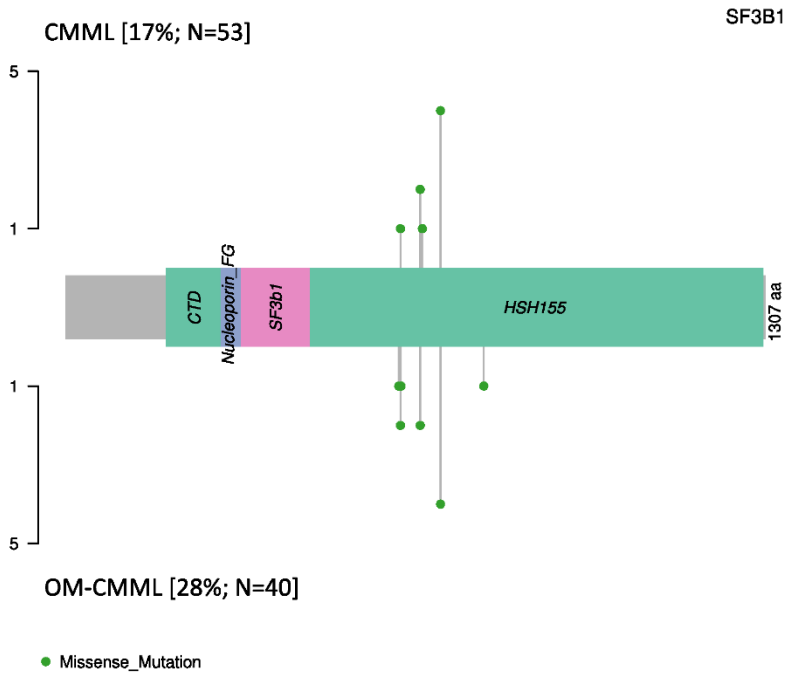
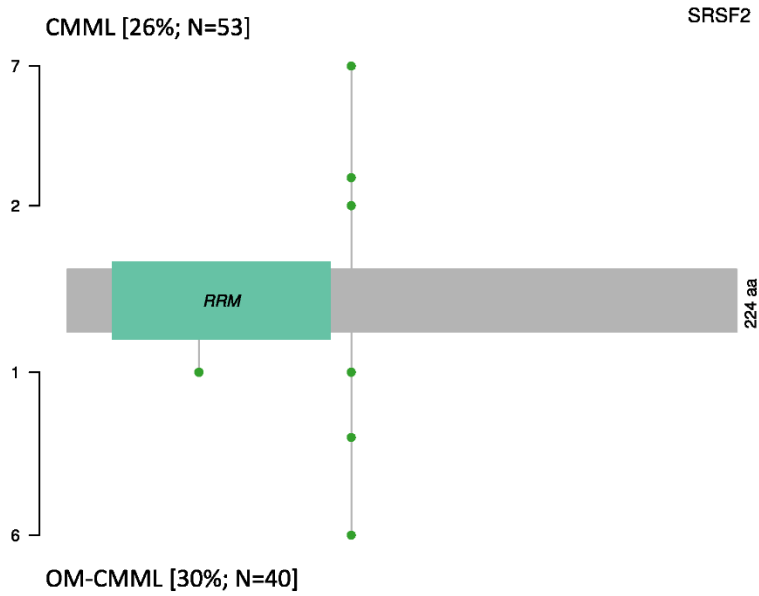


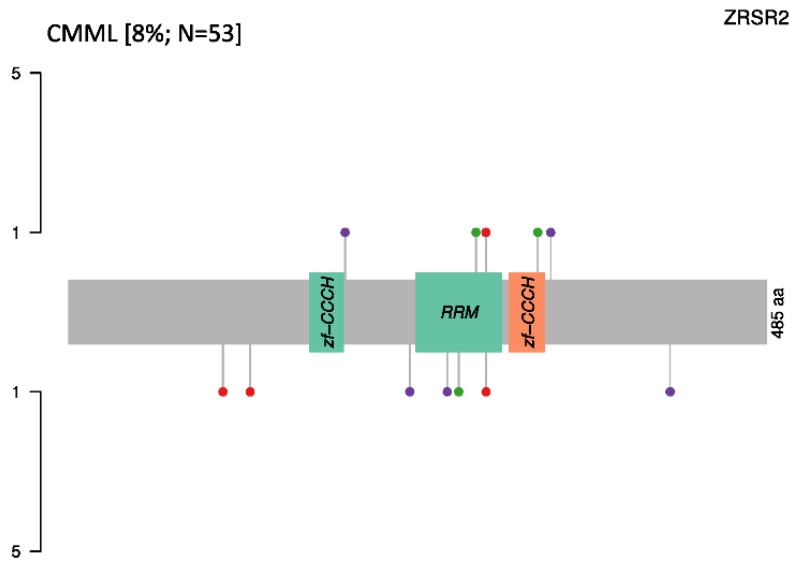
Figure 1e



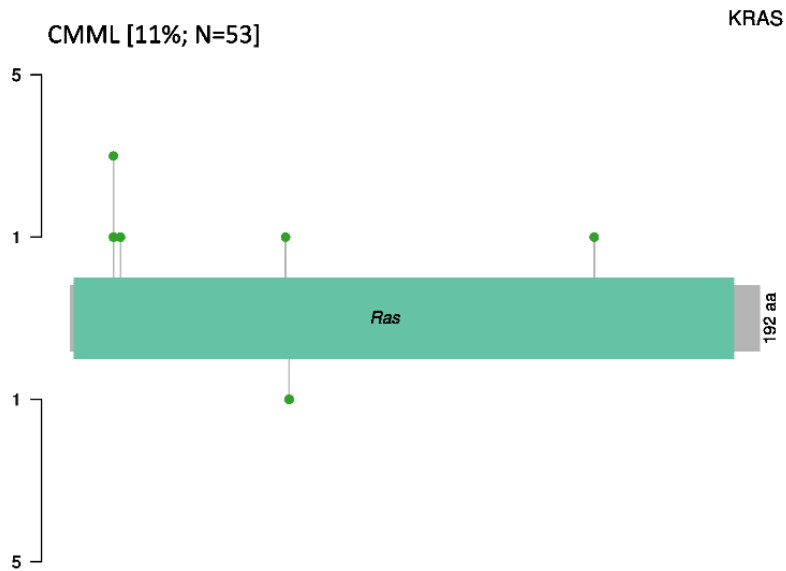
Supplemental Figure 2. Lollipop figures of mutations detected in *TET2*, *SRSF2*, *ASXL1*, *SF3B1*, *ZRSR2*, *NRAS*, *KRAS*, and *DNMT3A* genes in 40 OM-CMML and 53 CMML patients.





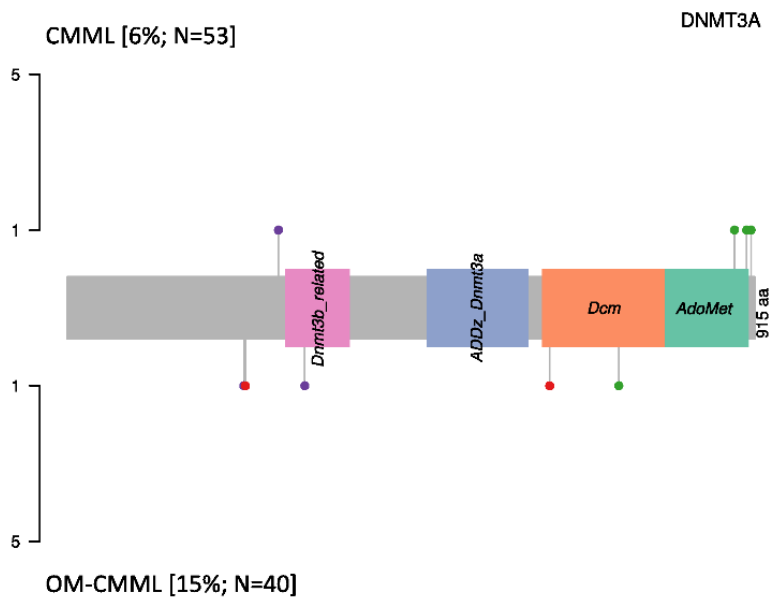
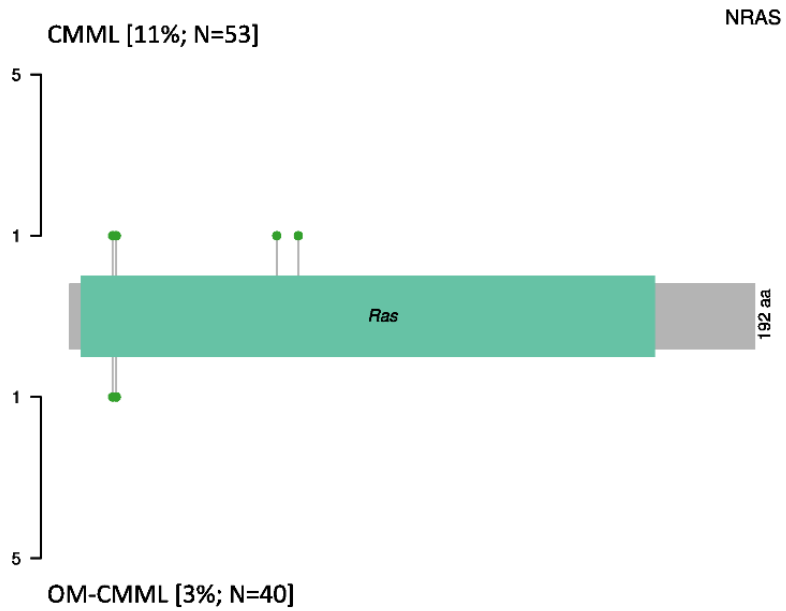


OM-CMML [20%; N=40]



OM-CMML [3%; N=40]

- Nonsense_Mutation
- Missense_Mutation
- Ins/Del



- Nonsense_Mutation
- Missense_Mutation
- Ins/Del

Supplemental Table 1. Comparison of the variant allele frequencies (VAFs) of the different somatic mutations assessed between 40 OM-CMML and 53 CMML patients. The table shows the median VAFs and the ranges of the different gene mutations analyzed. The only gene that showed a significantly different VAF when comparing both groups of patients was *DNMT3A* ($P=0.024$).

	OM-CMML	CMML	<i>P</i> value
<i>ASXL1</i>	32.7 (12-44.6)	36.9 (11.3-48.5)	0.581
<i>CALR</i>	-	-	-
<i>CBL</i>	2.8	37.6 (1.5-87.2)	0.545
<i>CSF3R</i>	-	-	-
<i>DNMT3A</i>	40.5 (6.7-48.2)	3 (2-4.7)	0.024
<i>ETV6</i>	43.9	46.6	1
<i>EZH2</i>	68.2 (48.5-87.9)	11.9	0.667
<i>IDH1</i>	24 (20.6-27.5)	-	-
<i>IDH2</i>	12.2 (10.2-39.7)	43.9 (1.9-48.5)	0.400
<i>JAK2</i>	9.4 (3.6-15.3)	7.9 (2.7-36.6)	0.857
<i>KIT</i>	-	-	-
<i>KRAS</i>	3.4	29 (12.9-43.3)	0.286
<i>MPL</i>	-	8.2	-
<i>NRAS</i>	22.8	11.6 (2.6-19.1)	0.286
<i>RUNX1</i>	29.7 (2.2-37.1)	34.1 (6.3-37.7)	0.841
<i>PRPF8</i>	-	44	-
<i>SETBP1</i>	10.3 (4.7-16)	23.4 (2.7-44.2)	1
<i>SF3B1</i>	41.3 (5.4-49)	39.8 (7.3-49)	0.882
<i>SH2B3</i>	39.6 (30.7-48.5)	20.1 (14.8-35.2)	0.400
<i>SRSF2</i>	40.3 (2.1-46.1)	40.9 (9.6-52.5)	0.392
<i>STAG2</i>	56.8 (17.8-95.7)	-	-
<i>TET2</i>	40.8 (16.6-95.4)	40.4 (2.4-82.9)	0.722
<i>TP53</i>	27.9	14.3 (2.9-71.5)	0.800
<i>U2AF1</i>	47.7	47.4 (43.8-50.6)	1
<i>ZRSR2</i>	61 (9.9-91.7)	38.2 (8.9-73.9)	0.368

Supplemental Table 2. Proportion of patients showing CD56 and/or CD2 positivity in monocytes ($\geq 20\%$ antigenic expression) among the different groups of patients analyzed (OM-CMML, CMML, MDS, MPN with monocytosis, and reactive monocytosis).

	OM-CMML	CMML	MDS	MPN	Reactive monocytosis	<i>P</i> value
CD56+	61.5% (24/39)	63% (34/54)	8.7% (2/23)	20% (3/15)	3.9% (4/102)	< 0.001
CD2+	28.2% (11/39)	35.2% (19/54)	0	6.7% (1/15)	0	< 0.001

ANNEX II: SUPPLEMENTARY INFORMATION (OM-CMML IS A DISTINCTIVE SUBTYPE OF CMML)

Supplemental Table 3. Influence of the different mutations assessed in the proportion of OM-CMML patients showing MO1 percentage >94%. A significantly higher proportion of OM-CMML patients with *TET2* mutation presented MO1 percentage >94% (P=0.004). This was the only one of the assessed mutations that permitted the splitting of the OM-CMML series into two groups, which showed a significant difference in the proportion of patients with MO1 percentage >94%.

	<i>ASXL1</i> mut	<i>ASXL1</i> wt	<i>CBL</i> mut	<i>CBL</i> wt	<i>DNMT3A</i> mut	<i>DNMT3A</i> wt	<i>ETV6</i> mut	<i>ETV6</i> wt
MO1 >94%	83.3% (5/6)	75.8% (25/33)	100% (1/1)	76.3% (29/38)	83.3% (5/6)	75.8% (25/33)	100% (1/1)	76.3% (29/38)
MO1 ≤94%	16.7% (1/6)	24.2% (8/33)	0	23.7% (9/38)	16.7% (1/6)	24.2% (8/33)	0	23.7% (9/38)

	<i>EZH2</i> mut	<i>EZH2</i> wt	<i>IDH1</i> mut	<i>IDH1</i> wt	<i>IDH2</i> mut	<i>IDH2</i> wt	<i>JAK2</i> mut	<i>JAK2</i> wt
MO1 >94%	100% (2/2)	75.7% (28/37)	50% (1/2)	78.4% (29/37)	66.7% (2/3)	77.8% (28/36)	50% (1/2)	78.4% (29/37)
MO1 ≤94%	0	24.3% (9/37)	50% (1/2)	21.6% (8/37)	33.3% (1/3)	22.2% (8/36)	50% (1/2)	21.6% (8/37)

	<i>KRAS</i> mut	<i>KRAS</i> wt	<i>NRAS</i> mut	<i>NRAS</i> wt	<i>RUNX1</i> mut	<i>RUNX1</i> wt	<i>SETBP1</i> mut	<i>SETBP1</i> wt
MO1 >94%	100% (1/1)	76.3% (29/38)	100% (1/1)	76.3% (29/38)	80% (4/5)	76.5% (26/33)	100% (2/2)	75.7% (28/37)
MO1 ≤94%	0	23.7% (9/38)	0	23.7% (9/38)	20% (1/5)	23.5% (8/34)	0	24.3% (9/37)

	<i>SF3B1</i> mut	<i>SF3B1</i> wt	<i>SH2B3</i> mut	<i>SH2B3</i> wt	<i>SRSF2</i> mut	<i>SRSF2</i> wt	<i>STAG2</i> mut	<i>STAG2</i> wt
MO1 >94%	81.8% (9/11)	75% (21/28)	50% (1/2)	78.4% (29/37)	83.3% (10/12)	74.1% (20/27)	50% (1/2)	78.4% (29/37)
MO1 ≤94%	18.2% (2/11)	25% (7/28)	50% (1/2)	21.6% (8/37)	16.7% (2/12)	25.9% (7/27)	50% (1/2)	21.6% (8/37)

	<i>TET2</i> mut	<i>TET2</i> wt	<i>TP53</i> mut	<i>TP53</i> wt	<i>U2AF1</i> mut	<i>U2AF1</i> wt	<i>ZRSR2</i> mut	<i>ZRSR2</i> wt
MO1 >94%	89.7% (26/29)	40% (4/10)	100% (1/1)	76.3% (29/38)	100% (1/1)	76.3% (29/38)	62.5% (5/8)	80.6% (25/31)
MO1 ≤94%	10.3% (3/29)	60% (6/10)	0	23.7% (9/38)	0	23.7% (9/38)	37.5% (3/8)	19.4% (6/31)

ANNEX II: SUPPLEMENTARY INFORMATION (OM-CMML IS A DISTINCTIVE SUBTYPE OF CMML)

Supplemental Data 3. Genetic variants detected by NGS (HGVS nomenclature) in OM-CMML and CMML patients.

PACIENT ID	PHENOTYPE	GENE	MUTATION	PROTEIN CONSEQUENCE	VARIANT ALLELE FREQUENCY (VAF)	TYPE	CONSEQUENCE
1	OM-CMML	<i>SF3B1</i>	NM_012433.2:c.1986C>G	NP_036565.2:p.His662Gln	45.23	snv	missense_variant
1	OM-CMML	<i>DNMT3A</i>	NM_175629.2:c.1015-2A>G		42.23	snv	splice_acceptor_variant
2	OM-CMML	<i>ZRSR2</i>	NM_005089.3:c.320T>G	NP_005080.1:p.Leu107Ter	54.47	snv	stop_gained
2	OM-CMML	<i>TET2</i>	NM_001127208.2:c.4781delC	NP_001120680.1:p.Pro1594LeufsTer2	37.67	deletion	frameshift_variant
2	OM-CMML	<i>TET2</i>	NM_001127208.2:c.1835delC	NP_001120680.1:p.Pro612LeufsTer27	37.41	deletion	frameshift_variant
3	OM-CMML	<i>ZRSR2</i>	NM_005089.3:c.827+1G>A		89.28	snv	splice_donor_variant
3	OM-CMML	<i>TET2</i>	NM_001127208.2:c.4393C>T	NP_001120680.1:p.Arg1465Ter	45.45	snv	stop_gained
3	OM-CMML	<i>ASXL1</i>	NM_015338.5:c.2740G>T	NP_056153.2:p.Glu914Ter	44.65	snv	stop_gained
3	OM-CMML	<i>TET2</i>	NM_001127208.2:c.1433_1449delAAAGGCTCAGAATAA	NP_001120680.1:p.Glu478ValfsTer6	38.42	deletion	frameshift_variant
4	OM-CMML	<i>TET2</i>	NM_001127208.2:c.4126G>T	NP_001120680.1:p.Asp1376Tyr	41.43	snv	missense_variant
4	OM-CMML	<i>TET2</i>	NM_001127208.2:c.1648C>T	NP_001120680.1:p.Arg550Ter	40.81	snv	stop_gained
5	OM-CMML	<i>STAG2</i>	NM_001042749.1:c.2542C>T	NP_001036214.1:p.Gln848Ter	95.71	snv	stop_gained
5	OM-CMML	<i>ZRSR2</i>	NM_005089.3:c.789delA	NP_005080.1:p.Glu263AspfsTer4	91.72	deletion	frameshift_variant
5	OM-CMML	<i>EZH2</i>	NM_004456.4:c.479delA	NP_004447.2:p.Asp160ValfsTer7	87.92	deletion	frameshift_variant
5	OM-CMML	<i>ASXL1</i>	NM_015338.5:c.1934dupG	NP_056153.2:p.Gly646TrpfsTer12	26.53	insertion	frameshift_variant
5	OM-CMML	<i>NRAS</i>	NM_002524.4:c.35G>C	NP_002515.1:p.Gly12Ala	22.79	snv	missense_variant
5	OM-CMML	<i>TET2</i>	NM_001127208.2:c.1246_1247insGAACC	NP_001120680.1:p.Pro416ArgfsTer13	20.48	insertion	frameshift_variant
5	OM-CMML	<i>NRAS</i>	NM_002524.4:c.38G>A	NP_002515.1:p.Gly13Asp	9.93	snv	missense_variant
5	OM-CMML	<i>RUNX1</i>	NM_001754.4:c.292delC	NP_001745.2:p.Leu98SerfsTer24	6.86	deletion	frameshift_variant
5	OM-CMML	<i>KRAS</i>	NM_033602.2:c.182A>G	NP_203524.1:p.Gln61Arg	3.4	snv	missense_variant
6	OM-CMML	<i>STAG2</i>	NM_001042749.1:c.2265+1G>A		17.89	snv	splice_donor_variant
7	OM-CMML	<i>SF3B1</i>	NM_012433.2:c.2098A>G	NP_036565.2:p.Lys700Glu	49.02	snv	missense_variant
7	OM-CMML	<i>EZH2</i>	NM_004456.4:c.2233G>A	NP_004447.2:p.Glu745Lys	48.5	snv	missense_variant
7	OM-CMML	<i>TET2</i>	NM_001127208.2:c.2461C>T	NP_001120680.1:p.Gln821Ter	46.6	snv	stop_gained
7	OM-CMML	<i>TET2</i>	NM_001127208.2:c.4115C>T	NP_001120680.1:p.Thr1372Ile	46	snv	missense_variant
7	OM-CMML	<i>TP53</i>	NM_000546.5:c.1024C>T	NP_000537.3:p.Arg342Ter	27.86	snv	stop_gained
8	OM-CMML	<i>TET2</i>	NM_001127208.2:c.4100C>T	NP_001120680.1:p.Pro1367Leu	94.56	snv	missense_variant
8	OM-CMML	<i>SRSF2</i>	NM_003016.4:c.284C>T	NP_003007.2:p.Pro95Leu	47.39	snv	missense_variant
9	OM-CMML	<i>TET2</i>	NM_001127208.2:c.3575G>T	NP_001120680.1:p.Gly1192Val	95.43	snv	missense_variant
9	OM-CMML	<i>RUNX1</i>	NM_001754.4:c.127_128insGCCGC	NP_001745.2:p.Pro43ArgfsTer7	37.07	insertion	frameshift_variant
9	OM-CMML	<i>CBL</i>	NM_005188.3:c.1211G>A	NP_005179.2:p.Cys404Tyr	2.83	snv	missense_variant
11	OM-CMML	<i>UZAF1</i>	NM_006758.2:c.470A>C	NP_006749.1:p.Gln157Pro	47.67	snv	missense_variant
11	OM-CMML	<i>ETV6</i>	NM_001987.4:c.33+1G>C		43.93	snv	splice_donor_variant
11	OM-CMML	<i>ASXL1</i>	NM_015338.5:c.2122C>T	NP_056153.2:p.Gln708Ter	42.01	snv	stop_gained
11	OM-CMML	<i>ASXL1</i>	NM_015338.5:c.2535dupC	NP_056153.2:p.Ser846GlnfsTer5	41.42	insertion	frameshift_variant
12	OM-CMML	<i>ZRSR2</i>	NM_005089.3:c.812A>G	NP_005080.1:p.Tyr271Cys	67.65	snv	missense_variant
13	OM-CMML	<i>TET2</i>	NM_001127208.2:c.3412C>T	NP_001120680.1:p.Gln1138Ter	33.33	snv	stop_gained
13	OM-CMML	<i>JAK2</i>	NM_004972.3:c.1849G>T	NP_004963.1:p.Val617Phe	15.28	snv	missense_variant
13	OM-CMML	<i>TET2</i>	NM_001127208.2:c.4753_4754delAAC	NP_001120680.1:p.Thr1585PhefsTer28	14.77	deletion	frameshift_variant
13	OM-CMML	<i>IDH2</i>	NM_002168.2:c.419G>A	NP_002159.2:p.Arg140Gln	12.21	snv	missense_variant
14	OM-CMML	<i>TET2</i>	NM_001127208.2:c.3893G>A	NP_001120680.1:p.Cys1298Tyr	16.57	snv	missense_variant
14	OM-CMML	<i>SF3B1</i>	NM_012433.2:c.1873C>T	NP_036565.2:p.Arg625Cys	16.36	snv	missense_variant
15	OM-CMML	<i>TET2</i>	NM_001127208.2:c.3410-2A>G		42.09	snv	splice_acceptor_variant
15	OM-CMML	<i>TET2</i>	NM_001127208.2:c.3812dupG	NP_001120680.1:p.Cys1271TrpfsTer29	40.81	insertion	frameshift_variant
16	OM-CMML	<i>RUNX1</i>	NM_001754.4:c.676_677delAAG	NP_001745.2:p.Ser226Ter	29.73	deletion	frameshift_variant
16	OM-CMML	<i>TET2</i>	NM_001127208.2:c.2656_2660dupCAGGA	NP_001120680.1:p.Glu887AspfsTer36	18.97	insertion	frameshift_variant
17	OM-CMML	<i>TET2</i>	NM_001127208.2:c.4615dupC	NP_001120680.1:p.Gln1539ProfsTer39	36.12	insertion	frameshift_variant
17	OM-CMML	<i>SRSF2</i>	NM_003016.4:c.283C>A	NP_003007.2:p.Pro95Thr	2.14	snv	missense_variant
18	OM-CMML	<i>SF3B1</i>	NM_012433.2:c.2342A>G	NP_036565.2:p.Asp781Gly	41.26	snv	missense_variant
18	OM-CMML	<i>TET2</i>	NM_001127208.2:c.3866G>A	NP_001120680.1:p.Cys1289Tyr	40.9	snv	missense_variant
18	OM-CMML	<i>DNMT3A</i>	NM_175629.2:c.939_945dupGTGGATG	NP_783328.1:p.Thr316ValfsTer10	34.15	insertion	frameshift_variant
19	OM-CMML	<i>SH2B3</i>	NM_005475.2:c.685_691delGGCCCCG	NP_005466.1:p.Gly229MetfsTer47	30.72	deletion	frameshift_variant
19	OM-CMML	<i>SRSF2</i>	NM_003016.4:c.284C>T	NP_003007.2:p.Pro95Leu	24.38	snv	missense_variant
19	OM-CMML	<i>TET2</i>	NM_001127208.2:c.4042C>T	NP_001120680.1:p.Gln1348Ter	23.99	snv	stop_gained
19	OM-CMML	<i>TET2</i>	NM_001127208.2:c.848_861delIAGCTGCCAACAAG	NP_001120680.1:p.Glu283AlafsTer7	19.68	deletion	frameshift_variant
19	OM-CMML	<i>DNMT3A</i>	NM_175629.2:c.709C>T	NP_783328.1:p.Gln237Ter	6.65	snv	stop_gained
20	OM-CMML	<i>DNMT3A</i>	NM_175629.2:c.1924G>T	NP_783328.1:p.Gly642Ter	48.2	snv	stop_gained
20	OM-CMML	<i>SF3B1</i>	NM_012433.2:c.1876A>G	NP_036565.2:p.Asn626Asp	47.83	snv	missense_variant
20	OM-CMML	<i>SETBP1</i>	NM_015559.2:c.2602G>A	NP_056374.2:p.Asp868Asn	4.69	snv	missense_variant
21	OM-CMML	<i>IDH1</i>	NM_005896.2:c.394C>T	NP_005887.2:p.Arg132Cys	27.5	snv	missense_variant
21	OM-CMML	<i>ASXL1</i>	NM_015338.5:c.1900_1922del23	NP_056153.2:p.Glu635ArgfsTer15	12.01	deletion	frameshift_variant
22	OM-CMML	<i>ZRSR2</i>	NM_005089.3:c.868C>T	NP_005080.1:p.Arg290Ter	52.92	snv	stop_gained
22	OM-CMML	<i>TET2</i>	NM_001127208.2:c.4150G>C	NP_001120680.1:p.Asp1384His	32.32	snv	missense_variant
22	OM-CMML	<i>ZRSR2</i>	NM_005089.3:c.376C>T	NP_005080.1:p.Arg126Ter	4.58	snv	stop_gained
23	OM-CMML	<i>DNMT3A</i>	NM_175629.2:c.2201T>C	NP_783328.1:p.Phe734Ser	38.7	snv	missense_variant
24	OM-CMML	<i>SRSF2</i>	NM_003016.4:c.284C>T	NP_003007.2:p.Pro95Leu	37.57	snv	missense_variant
24	OM-CMML	<i>TET2</i>	NM_001127208.2:c.1771C>T	NP_001120680.1:p.Gln591Ter	35.1	snv	stop_gained
24	OM-CMML	<i>RUNX1</i>	NM_001754.4:c.1097delIT	NP_001745.2:p.Ile366ThrfsTer228	34.45	deletion	frameshift_variant
24	OM-CMML	<i>ASXL1</i>	NM_015338.5:c.1926dupA	NP_056153.2:p.Gly643ArgfsTer15	32.74	insertion	frameshift_variant
25	OM-CMML	<i>SF3B1</i>	NM_012433.2:c.1866G>T	NP_036565.2:p.Glu622Asp	27.27	snv	missense_variant
25	OM-CMML	<i>TET2</i>	NM_001127208.2:c.369dupPT	NP_001120680.1:p.Asn124Ter	22.28	insertion	frameshift_variant
25	OM-CMML	<i>IDH2</i>	NM_002168.2:c.419G>A	NP_002159.2:p.Arg140Gln	10.19	snv	missense_variant
25	OM-CMML	<i>SRSF2</i>	NM_003016.4:c.284C>A	NP_003007.2:p.Pro95His	9.02	snv	missense_variant
26	OM-CMML	<i>TET2</i>	NM_001127208.2:c.2257A>T	NP_001120680.1:p.Lys753Ter	56.72	snv	stop_gained

ANNEX II: SUPPLEMENTARY INFORMATION (OM-CMML IS A DISTINCTIVE SUBTYPE OF CMML)

26	OM-CMML	TET2	NM_001127208.2:c.3899T>G	NP_001120680.1:p.Phe1300Cys	21.18	snv	missense_variant
26	OM-CMML	SF3B1	NM_012433.2:c.2098A>G	NP_036565.2:p.Lys700Glu	5.43	snv	missense_variant
27	OM-CMML	DNMT3A	NM_175629.2:c.703delG	NP_783328.1:p.Glu235SerfsTer81	46.05	deletion	frameshift_variant
27	OM-CMML	TET2	NM_001127208.2:c.4668_4671delTGTC	NP_001120680.1:p.Val1557ThrfsTer13	44.56	deletion	frameshift_variant
27	OM-CMML	SF3B1	NM_012433.2:c.2098A>G	NP_036565.2:p.Lys700Glu	44.06	snv	missense_variant
27	OM-CMML	SETBP1	NM_015559.2:c.2608G>A	NP_056374.2:p.Gly870Ser	15.96	snv	missense_variant
28	OM-CMML	TET2	NM_001127208.2:c.3985delC	NP_001120680.1:p.Leu1329CysfsTer34	49.15	deletion	frameshift_variant
29	OM-CMML	SH2B3	NM_005475.2:c.622G>C	NP_005466.1:p.Glu208Gln	48.49	snv	missense_variant
29	OM-CMML	SRSF2	NM_003016.4:c.284C>A	NP_003007.2:p.Pro95His	44.38	snv	missense_variant
29	OM-CMML	IDH2	NM_002168.2:c.419G>A	NP_002159.2:p.Arg140Gln	39.69	snv	missense_variant
29	OM-CMML	SH2B3	NM_005475.2:c.1283_1284delIAC	NP_005466.1:p.His428ProfsTer27	22.63	deletion	frameshift_variant
29	OM-CMML	ASXL1	NM_015338.5:c.1552G>T	NP_056153.2:p.Glu518Ter	22.05	snv	stop_gained
29	OM-CMML	JAK2	NM_004972.3:c.1849G>T	NP_004963.1:p.Val617Phe	3.6	snv	missense_variant
30	OM-CMML	TET2	NM_001127208.2:c.822delC	NP_001120680.1:p.Asn275IlefsTer18	52.27	deletion	frameshift_variant
30	OM-CMML	SRSF2	NM_003016.4:c.284C>T	NP_003007.2:p.Pro95Leu	42.29	snv	missense_variant
30	OM-CMML	TET2	NM_001127208.2:c.3384T>A	NP_001120680.1:p.Tyr1128Ter	8.64	snv	stop_gained
30	OM-CMML	TET2	NM_001127208.2:c.2400_2401delTA	NP_001120680.1:p.His800GlnfsTer15	2.86	deletion	frameshift_variant
31	OM-CMML	SF3B1	NM_012433.2:c.1986C>G	NP_036565.2:p.His662Gln	36.74	snv	missense_variant
32	OM-CMML	SF3B1	NM_012433.2:c.2098A>G	NP_036565.2:p.Lys700Glu	42.29	snv	missense_variant
32	OM-CMML	TET2	NM_001127208.2:c.1873delA	NP_001120680.1:p.Thr625HisfsTer14	41.95	deletion	frameshift_variant
32	OM-CMML	TET2	NM_001127208.2:c.5562dupT	NP_001120680.1:p.Leu1855SerfsTer4	12.08	insertion	frameshift_variant
33	OM-CMML	SRSF2	NM_003016.4:c.284C>T	NP_003007.2:p.Pro95Leu	40.3	snv	missense_variant
33	OM-CMML	TET2	NM_001127208.2:c.4537G>C	NP_001120680.1:p.Glu1513Gln	39.29	snv	missense_variant
33	OM-CMML	TET2	NM_001127208.2:c.2279_2280delITT	NP_001120680.1:p.Phe760SerfsTer8	37.62	deletion	frameshift_variant
34	OM-CMML	TET2	NM_001127208.2:c.2373T>A	NP_001120680.1:p.Tyr791Ter	41.84	snv	stop_gained
34	OM-CMML	TET2	NM_001127208.2:c.3732_3733delICT	NP_001120680.1:p.Tyr1245LeufsTer22	41.07	deletion	frameshift_variant
34	OM-CMML	ZRSR2	NM_005089.3:c.771+1G>C		30.92	snv	splice_donor_variant
35	OM-CMML	SRSF2	NM_003016.4:c.130T>C	NP_003007.2:p.Tyr44His	31.07	snv	missense_variant
35	OM-CMML	TET2	NM_001127208.2:c.4020dupT	NP_001120680.1:p.Ala1341CysfsTer3	31.01	insertion	frameshift_variant
35	OM-CMML	TET2	NM_001127208.2:c.3308_3309delIAT	NP_001120680.1:p.Asn1103IlefsTer26	29.2	deletion	frameshift_variant
35	OM-CMML	ZRSR2	NM_005089.3:c.709delC	NP_005080.1:p.Leu237Ter	9.89	deletion	frameshift_variant
36	OM-CMML	TET2	NM_001127208.2:c.2671C>T	NP_001120680.1:p.Gln891Ter	45.34	snv	stop_gained
36	OM-CMML	SRSF2	NM_003016.4:c.284C>A	NP_003007.2:p.Pro95His	44	snv	missense_variant
36	OM-CMML	IDH1	NM_005896.2:c.394C>T	NP_005887.2:p.Arg132Cys	20.59	snv	missense_variant
36	OM-CMML	TET2	NM_001127208.2:c.1930C>T	NP_001120680.1:p.Gln644Ter	18.48	snv	stop_gained
36	OM-CMML	TET2	NM_001127208.2:c.5502_5518delGGGTG TGGCTTCGGTG	NP_001120680.1:p.Gln1834HisfsTer6	4.38	deletion	frameshift_variant
36	OM-CMML	RUNX1	NM_001754.4:c.1174C>T	NP_001745.2:p.Gln392Ter	2.15	snv	stop_gained
37	OM-CMML	TET2	NM_001127208.2:c.4112T>A	NP_001120680.1:p.Val1371Asp	47.53	snv	missense_variant
37	OM-CMML	TET2	NM_001127208.2:c.3594+2T>C		46.38	snv	splice_donor_variant
37	OM-CMML	SRSF2	NM_003016.4:c.284C>T	NP_003007.2:p.Pro95Leu	42.89	snv	missense_variant
38	OM-CMML	ZRSR2	NM_005089.3:c.1252delC	NP_005080.1:p.His418ThrfsTer?	88.89	deletion	frameshift_variant
39	OM-CMML	TET2	NM_001127208.2:c.5273C>G	NP_001120680.1:p.Ser1758Ter	56.95	snv	stop_gained
39	OM-CMML	SRSF2	NM_003016.4:c.284C>G	NP_003007.2:p.Pro95Arg	46.11	snv	missense_variant
40	OM-CMML	TET2	NM_001127208.2:c.5647A>C	NP_001120680.1:p.Thr1883Pro	39.77	snv	missense_variant
40	OM-CMML	SF3B1	NM_012433.2:c.1873C>T	NP_036565.2:p.Arg625Cys	37.85	snv	missense_variant
41	CMML	TET2	NM_001127208.2:c.3570delT	NP_001120680.1:p.Gln1191ArgfsTer35	43.6	deletion	frameshift_variant
41	CMML	SF3B1	NM_012433.2:c.2098A>G	NP_056374.2:p.Lys700Glu	42.9	snv	missense_variant
41	CMML	CBL	NM_005188.3:c.1211G>A	NP_005179.2:p.Cys404Tyr	40.5	snv	missense_variant
41	CMML	TET2	NM_001127208.2:c.4570C>T	NP_001120680.1:p.Gln1524Ter	40	snv	stop_gained
41	CMML	ASXL1	NM_015338.5:c.1900_1922delI23	NP_056153.2:p.Glu635ArgfsTer15	27.45	deletion	frameshift_variant
42	CMML	CBL	NM_005188.3:c.1255T>A	NP_005179.2:p.Cys419Ser	46.34	snv	missense_variant
42	CMML	SRSF2	NM_003016.4:c.284C>A	NP_003007.2:p.Pro95His	45.67	snv	missense_variant
42	CMML	TET2	NM_001127208.2:c.3743T>G	NP_001120680.1:p.Leu1248Arg	42.64	snv	missense_variant
42	CMML	ASXL1	NM_015338.5:c.2498_2499delIGT	NP_056153.2:p.Ser833ThrfsTer17	40.83	deletion	frameshift_variant
42	CMML	CBL	NM_005188.3:c.1618C>T	NP_005179.2:p.Arg540Ter	34.86	snv	stop_gained
43	CMML	DNMT3A	NM_175629.2:c.2663T>A	NP_783328.1:p.Leu888Gln	4.65	snv	missense_variant
44	CMML	SF3B1	NM_012433.2:c.1873C>T	NP_036565.2:p.Arg625Cys	40.84	snv	missense_variant
44	CMML	CBL	NM_005188.3:c.826G>T	NP_005179.2:p.Glu276Ter	2.87	snv	stop_gained
45	CMML	TET2	NM_001127208.2:c.3954+1G>A		29.82	snv	splice_donor_variant
45	CMML	JAK2	NM_004972.3:c.1849G>T	NP_004963.1:p.Val617Phe	7.88	snv	missense_variant
45	CMML	TET2	NM_001127208.2:c.4977_4978dupTA	NP_001120680.1:p.Arg1660IlefsTer36	4.18	insertion	frameshift_variant
46	CMML	UZAF1	NM_006758.2:c.470A>G	NP_006749.1:p.Gln157Arg	43.79	snv	missense_variant
46	CMML	RUNX1	NM_001754.4:c.602G>A	NP_001745.2:p.Arg201Gln	37.71	snv	missense_variant
46	CMML	ASXL1	NM_015338.5:c.1934dupG	NP_056153.2:p.Gly646TrpfsTer12	28.35	insertion	frameshift_variant
47	CMML	CBL	NM_005188.3:c.1290_1292delGGT	NP_005179.2:p.Val431del	87.16	deletion	inframe_deletion
47	CMML	ASXL1	NM_015338.5:c.3359A>T	NP_056153.2:p.Lys1120Met	48.48	snv	missense_variant
47	CMML	TET2	NM_001127208.2:c.3385delIG	NP_001120680.1:p.Asp1129IlefsTer8	45.57	deletion	frameshift_variant
47	CMML	TET2	NM_001127208.2:c.3781C>T	NP_001120680.1:p.Arg1261Cys	44.63	snv	missense_variant
47	CMML	SRSF2	NM_003016.4:c.284C>T	NP_003007.2:p.Pro95Leu	40.86	snv	missense_variant
47	CMML	TET2	NM_001127208.2:c.5079C>G	NP_001120680.1:p.Tyr1693Ter	1.99	snv	stop_gained
48	CMML	SRSF2	NM_003016.4:c.284C>G	NP_003007.2:p.Pro95Arg	52.54	snv	missense_variant
48	CMML	ASXL1	NM_015338.5:c.1934dupG	NP_056153.2:p.Gly646TrpfsTer12	33.33	insertion	frameshift_variant
49	CMML	SF3B1	NM_012433.2:c.1986C>G	NP_036565.2:p.His662Gln	47.18	snv	missense_variant
50	CMML	ZRSR2	NM_005089.3:c.91C>T	NP_005080.1:p.Arg31Trp	73.94	snv	missense_variant
50	CMML	TET2	NM_001127208.2:c.331A>T	NP_001120680.1:p.Lys111Ter	29.51	snv	stop_gained
50	CMML	SETBP1	NM_015559.2:c.2959C>T	NP_056374.2:p.Arg987Trp	2.95	snv	missense_variant
51	CMML	SRSF2	NM_003016.4:c.284C>T	NP_003007.2:p.Pro95Leu	24.32	snv	missense_variant
51	CMML	TET2	NM_001127208.2:c.2494delIG	NP_001120680.1:p.Val832PhefsTer9	20.28	deletion	frameshift_variant
51	CMML	TET2	NM_001127208.2:c.333A>T	NP_001120680.1:p.Lys111Asn	17.32	snv	missense_variant
51	CMML	TET2	NM_001127208.2:c.333delA	NP_001120680.1:p.Lys111AsnfsTer2	17	deletion	frameshift_variant
51	CMML	TET2	NM_001127208.2:c.3690_3703delCCTGG TGIGGGAAG	NP_001120680.1:p.Ile1230MetfsTer8	13.04	deletion	frameshift_variant
52	CMML	TET2	NM_001127208.2:c.1585dupT	NP_001120680.1:p.Cys529LeufsTer38	24.09	insertion	frameshift_variant
52	CMML	TET2	NM_001127208.2:c.1236delIT	NP_001120680.1:p.Pro413HisfsTer14	19.73	deletion	frameshift_variant

ANNEX II: SUPPLEMENTARY INFORMATION (OM-CMML IS A DISTINCTIVE SUBTYPE OF CMML)

53	CMML	IDH2	NM_002168.2:c.419G>A	NP_002159.2:p.Arg140Gln	48.46	snv	missense_variant
53	CMML	U2AF1	NM_006758.2:c.470A>G	NP_006749.1:p.Gln157Arg	47.41	snv	missense_variant
53	CMML	ASXL1	NM_015338.5:c.1934dupG	NP_056153.2:p.Gly646TrpfsTer12	32.98	insertion	frameshift_variant
53	CMML	NRAS	NM_002524.4:c.38G>A	NP_002515.1:p.Gly13Asp	2.63	snv	missense_variant
54	CMML	TET2	NM_001127208.2:c.4618C>T	NP_001120680.1:p.Gln1540Ter	45.1	snv	stop_gained
54	CMML	CBL	NM_005188.3:c.1211G>A	NP_005179.2:p.Cys404Tyr	36.6	snv	missense_variant
54	CMML	ASXL1	NM_015338.5:c.2338C>T	NP_056153.2:p.Gln780Ter	15.5	snv	stop_gained
54	CMML	EZH2	NM_004456.4:c.2216T>C	NP_004447.2:p.Leu739Pro	11.9	snv	missense_variant
54	CMML	SRSF2	NM_003016.4:c.284C>A	NP_003007.2:p.Pro95His	9.61	snv	missense_variant
54	CMML	CBL	NM_005188.3:c.1246T>C	NP_005179.2:p.Cys416Arg	8.8	snv	missense_variant
55	CMML	TET2	NM_001127208.2:c.2713_2716del	NP_001120680.1:p.Asp905CysfsTer15	44.4	snv	missense_variant
56	CMML	SRSF2	NM_003016.4:c.284C>G	NP_003007.2:p.Pro95Arg	26.11	snv	missense_variant
56	CMML	TET2	NM_001127208.2:c.5582G>T	NP_001120680.1:p.Gly1861Val	25.97	snv	missense_variant
56	CMML	TET2	NM_001127208.2:c.3812dupG	NP_001120680.1:p.Cys12711TrpfsTer29	15.64	insertion	frameshift_variant
57	CMML	TET2	NM_001127208.2:c.4909delC	NP_001120680.1:p.Leu1637TrpfsTer58	31.2	deletion	frameshift_variant
57	CMML	SF3B1	NM_012433.2:c.2098A>G	NP_056374.2:p.Lys700Glu	30.2	snv	missense_variant
57	CMML	TET2	NM_001127208.2:c.2887C>T	NP_001120680.1:p.Gln963Ter	12.9	snv	stop_gained
58	CMML	TET2	NM_001127208.2:c.3640C>T	NP_001120680.1:p.Arg1214Trp	48.23	snv	missense_variant
58	CMML	TET2	NM_001127208.2:c.4164_4165delGCinsA T	NP_001120680.1:p.MetGln1388IleTer	40.02	mnp	stop_gained
58	CMML	TET2	NM_001127208.2:c.1876C>T	NP_001120680.1:p.Gln626Ter	7.38	snv	stop_gained
59	CMML	SF3B1	NM_012433.2:c.1986C>G	NP_036565.2:p.His662Gln	7.31	snv	missense_variant
59	CMML	IDH2	NM_002168.2:c.352C>T	NP_002159.2:p.Pro118Ser	1.89	snv	missense_variant
60	CMML	TET2	NM_001127208.2:c.3230dupA	NP_001120680.1:p.His1077GlnfsTer27	42.45	insertion	frameshift_variant
60	CMML	SH2B3	NM_005475.2:c.393delG	NP_005466.1:p.Cys133AlafsTer64	35.23	deletion	frameshift_variant
61	CMML	IDH2	NM_002168.2:c.419G>A	NP_002159.2:p.Arg140Gln	46.52	snv	missense_variant
61	CMML	JAK2	NM_004972.3:c.1849G>T	NP_004963.1:p.Val617Phe	35.11	snv	missense_variant
62	CMML	TET2	NM_001127208.2:c.5093delA	NP_001120680.1:p.Asn1698TrpfsTer21	55.42	deletion	frameshift_variant
62	CMML	ASXL1	NM_015338.5:c.4180G>T	NP_056153.2:p.Glu1394Ter	41.51	snv	stop_gained
62	CMML	TET2	NM_001127208.2:c.3640C>T	NP_001120680.1:p.Arg1214Trp	25.62	snv	missense_variant
62	CMML	SH2B3	NM_005475.2:c.1174C>T	NP_005466.1:p.Arg392Trp	14.81	snv	missense_variant
62	CMML	TET2	NM_001127208.2:c.3780delT	NP_001120680.1:p.Arg1261AlafsTer5	12.14	deletion	frameshift_variant
63	CMML	TET2	NM_001127208.2:c.5456T>G	NP_001120680.1:p.Leu1819Ter	31.93	snv	stop_gained
63	CMML	TET2	NM_001127208.2:c.943delT	NP_001120680.1:p.Ser315ProfsTer32	30.95	deletion	frameshift_variant
63	CMML	CBL	NM_005188.3:c.2599C>T	NP_005179.2:p.Gln867Ter	16.28	snv	stop_gained
63	CMML	TET2	NM_001127208.2:c.5618T>C	NP_001120680.1:p.Ile1873Thr	9.95	snv	missense_variant
63	CMML	NRAS	NM_002524.4:c.35G>A	NP_002515.1:p.Gly12Asp	8.1	snv	missense_variant
64	CMML	ASXL1	NM_015338.5:c.1934dupG	NP_056153.2:p.Gly646TrpfsTer12	36.29	insertion	frameshift_variant
65	CMML	TET2	NM_001127208.2:c.2507_2510del	NP_001120680.1:p.Asn836IlefsTer4	78.9	deletion	frameshift_variant
65	CMML	CBL	NM_005188.3:c.1101_1103del	p.Gln367_Tyr368delinsHis	42.6	deletion	frameshift_variant
65	CMML	RUNX1	NM_001754.4:c.167T>C	NP_001745.2:p.Leu56Ser	34.1	snv	missense_variant
65	CMML	TET2	NM_001127208.2:c.4138C>T	NP_001120680.1:p.His1380Tyr	13.5	snv	missense_variant
65	CMML	TET2	NM_001127208.2:c.3928A>C	NP_001120680.1:p.Lys1310Gln	5	snv	missense_variant
66	CMML	TET2	NM_001127208.2:c.4035T>G	NP_001120680.1:p.Tyr1345Ter	29.64	snv	stop_gained
66	CMML	TET2	NM_001127208.2:c.822delIC	NP_001120680.1:p.Asn275IlefsTer18	25.69	deletion	frameshift_variant
66	CMML	NRAS	NM_002524.4:c.173C>T	NP_002515.1:p.Thr58Ile	4.85	snv	missense_variant
67	CMML	SRSF2	NM_003016.4:c.284C>T	NP_003007.2:p.Pro95Leu	41.48	snv	missense_variant
67	CMML	KRAS	NM_033360.2:c.437C>T	NP_203524.1:p.Ala146Val	41.36	snv	missense_variant
67	CMML	TET2	NM_001127208.2:c.3869G>C	NP_001120680.1:p.Ser1290Ter	40.35	snv	stop_gained
67	CMML	JAK2	NM_004972.3:c.1849G>T	NP_004963.1:p.Val617Phe	14.41	snv	missense_variant
68	CMML	ZRSR2	NM_005089.3:c.575delA	NP_005080.1:p.Asn192IlefsTer46	54.55	deletion	frameshift_variant
68	CMML	TET2	NM_001127208.2:c.3892T>G	NP_001120680.1:p.Cys1298Gly	38.3	snv	missense_variant
68	CMML	TET2	NM_001127208.2:c.4210C>T	NP_001120680.1:p.Arg1404Ter	38.27	snv	stop_gained
69	CMML	ASXL1	NM_015338.5:c.1774C>T	NP_056153.2:p.Gln592Ter	39.1	snv	stop_gained
69	CMML	CBL	NM_005188.3:c.1111T>A	NP_005179.2:p.Tyr371Asn	38.62	snv	missense_variant
69	CMML	TET2	NM_001127208.2:c.5059C>T	NP_001120680.1:p.Gln1687Ter	34.86	snv	stop_gained
69	CMML	TET2	NM_001127208.2:c.3273dupA	NP_001120680.1:p.Pro1092TrpfsTer12	33.61	insertion	frameshift_variant
70	CMML	TET2	NM_001127208.2:c.2191C>T	NP_001120680.1:p.Gln731Ter	82.86	snv	stop_gained
70	CMML	TET2	NM_001127208.2:c.2190_2191insGACA	NP_001120680.1:p.Gln731AspfsTer24	77.06	insertion	frameshift_variant
70	CMML	ASXL1	NM_015338.5:c.2249_2250delICC	NP_056153.2:p.Pro750ArgfsTer23	43.52	deletion	frameshift_variant
71	CMML	TET2	NM_001127208.2:c.5140_5141delIAA	NP_001120680.1:p.Asn1714CysfsTer14	42.62	deletion	frameshift_variant
72	CMML	SF3B1	NM_012433.2:c.1996A>G	NP_036565.2:p.Lys666Glu	45.46	snv	missense_variant
72	CMML	TET2	NM_001127208.2:c.2674C>T	NP_001120680.1:p.Gln892Ter	32.59	snv	stop_gained
72	CMML	TP53	NM_000546.5:c.817C>T	NP_000537.3:p.Arg273Cys	17.43	snv	missense_variant
73	CMML	SRSF2	NM_003016.4:c.284C>T	NP_003007.2:p.Pro95Leu	34.99	snv	missense_variant
73	CMML	TET2	NM_001127208.2:c.4561delG	NP_001120680.1:p.Val1521SerfsTer50	34.2	deletion	frameshift_variant
73	CMML	TET2	NM_001127208.2:c.4272delIT	NP_001120680.1:p.Asp1425TrpfsTer23	22.67	deletion	frameshift_variant
73	CMML	KRAS	NM_033360.2:c.35G>A	NP_203524.1:p.Gly12Asp	14.67	snv	missense_variant
73	CMML	DNMT3A	NM_175629.2:c.839_840dupAC	NP_783328.1:p.Glu281TrpfsTer36	2.71	insertion	frameshift_variant
74	CMML	ASXL1	NM_015338.5:c.1773C>G	NP_056153.2:p.Tyr591Ter	47.36	snv	stop_gained
74	CMML	TET2	NM_001127208.2:c.5500C>T	NP_001120680.1:p.Gln1834Ter	46.06	snv	stop_gained
74	CMML	ASXL1	NM_015338.5:c.1819G>T	NP_056153.2:p.Gly607Cys	29.72	snv	missense_variant
74	CMML	TET2	NM_001127208.2:c.4210C>T	NP_001120680.1:p.Arg1404Ter	24.53	snv	stop_gained
74	CMML	KRAS	NM_033360.2:c.40G>A	NP_203524.1:p.Val14Ile	19.25	snv	missense_variant
74	CMML	TET2	NM_001127208.2:c.3921delIG	NP_001120680.1:p.Lys1308SerfsTer55	15.62	deletion	frameshift_variant
74	CMML	TP53	NM_000546.5:c.818G>A	NP_000537.3:p.Arg273His	10.99	snv	missense_variant
74	CMML	JAK2	NM_004972.3:c.1849G>T	NP_004963.1:p.Val617Phe	3.63	snv	missense_variant
74	CMML	TET2	NM_001127208.2:c.3845G>A	NP_001120680.1:p.Gly1282Asp	2.88	snv	missense_variant
75	CMML	ZRSR2	NM_005089.3:c.847G>C	NP_005080.1:p.Ala283Pro	21.75	snv	missense_variant
75	CMML	TET2	NM_001127208.2:c.5500C>T	NP_001120680.1:p.Gln1834Ter	14.36	snv	stop_gained
75	CMML	TET2	NM_001127208.2:c.1061C>A	NP_001120680.1:p.Ser354Ter	14.06	snv	stop_gained
75	CMML	DNMT3A	NM_175629.2:c.2729C>T	NP_783328.1:p.Ala910Val	3.03	snv	missense_variant
75	CMML	DNMT3A	NM_175629.2:c.2711C>T	NP_783328.1:p.Pro904Leu	1.6	snv	missense_variant
76	CMML	TET2	NM_001127208.2:c.4165C>T	NP_001120680.1:p.Gln1389Ter	49.5	snv	missense_variant
76	CMML	SF3B1	NM_012433.2:c.2098A>G	NP_056374.2:p.Lys700Glu	47.2	snv	missense_variant

ANNEX II: SUPLEMETARY INFORMATION (OM-CMML IS A DISTINCTIVE SUBTYPE OF CMML)

76	CMML	TET2	NM_001127208.2:c.5381delC	NP_001120680.1:p.Ala1794ValfsTer26	45.9	deletion	frameshift_variant
76	CMML	SETBP1	NM_015559.2:c.2900T>C	NP_056374.2:p.Phe9675Ser	44.2	snv	missense_variant
76	CMML	MPL	NM_005373.2:c.1544G>T	NP_005364.1:p.Trp515Leu	8.2	snv	missense_variant
76	CMML	MPL	NM_005373.2:c.1514G>A	NP_005364.1:p.Ser505Asn	6.1	snv	missense_variant
77	CMML	ASXL1	NM_015338.5:c.2077C>T	NP_056153.2:p.Arg693Ter	41.4	snv	stop_gained
77	CMML	SRSF2	NM_003016.4:c.284C>A	NP_003007.2:p.Pro95His	40.25	snv	missense_variant
77	CMML	TET2	NM_001127208.2:c.5104C>T	NP_001120680.1:p.Gln1702Ter	40	snv	stop_gained
77	CMML	TET2	NM_001127208.2:c.3574G>T	NP_001120680.1:p.Gly1192Ter	39.9	snv	stop_gained
78	CMML	KRAS	NM_033360.2:c.34G>C	NP_203524.1:p.Gly12Arg	38.82	snv	missense_variant
79	CMML	SRSF2	NM_003016.4:c.284C>A	NP_003007.2:p.Pro95His	42.57	snv	missense_variant
79	CMML	TET2	NM_001127208.2:c.4075C>T	NP_001120680.1:p.Arg1359Cys	39.28	snv	missense_variant
79	CMML	CBL	NM_005188.3:c.1210T>C	NP_005179.2:p.Cys404Arg	4.23	snv	missense_variant
80	CMML	TET2	NM_001127208.2:c.822delC	NP_001120680.1:p.Asn275IlefsTer18	21.87	deletion	frameshift_variant
80	CMML	TET2	NM_001127208.2:c.4063_4064delGCinsT A	NP_001120680.1:p.Ala1355Ter	21.39	mnp	stop_gained
81	CMML	NRAS	NM_002524.4:c.35G>A	NP_002515.1:p.Gly12Asp	15.07	snv	missense_variant
81	CMML	NRAS	NM_002524.4:c.35G>T	NP_002515.1:p.Gly12Val	15.07	snv	missense_variant
81	CMML	KRAS	NM_033360.2:c.34G>C	NP_203524.1:p.Gly12Arg	12.93	snv	missense_variant
81	CMML	NRAS	NM_002524.4:c.190T>G	NP_002515.1:p.Tyr64Asp	4.9	snv	missense_variant
81	CMML	TET2	NM_001127208.2:c.5268_5274delITCATT A	NP_001120680.1:p.His1757LeufsTer4	3.93	deletion	frameshift_variant
81	CMML	JAK2	NM_004972.3:c.1849G>T	NP_004963.1:p.Val617Phe	2.7	snv	missense_variant
81	CMML	CBL	NM_005188.3:c.1211G>A	NP_005179.2:p.Cys404Tyr	2.47	snv	missense_variant
81	CMML	KRAS	NM_033360.2:c.179G>A	NP_203524.1:p.Gly60Asp	2.32	snv	missense_variant
82	CMML	SRSF2	NM_003016.4:c.284C>A	NP_003007.2:p.Pro95His	48.15	snv	missense_variant
82	CMML	ETV6	NM_001987.4:c.305_306insG	NP_001978.1:p.Phe102LeufsTer10	46.62	insertion	frameshift_variant
82	CMML	TET2	NM_001127208.2:c.1516delA	NP_001120680.1:p.Arg506AspfsTer27	42.99	deletion	frameshift_variant
82	CMML	TET2	NM_001127208.2:c.3869C>T	NP_001120680.1:p.Ser1290Leu	37.17	snv	missense_variant
82	CMML	TET2	NM_001127208.2:c.2474delC	NP_001120680.1:p.Ser825Ter	2.54	deletion	frameshift_variant
82	CMML	TET2	NM_001127208.2:c.3813C>G	NP_001120680.1:p.Cys1271Trp	2.16	snv	missense_variant
82	CMML	TET2	NM_001127208.2:c.3640C>T	NP_001120680.1:p.Arg1214Trp	1.66	snv	missense_variant
83	CMML	TET2	NM_001127208.2:c.2784delT	NP_001120680.1:p.Pro929LeufsTer24	30.43	deletion	frameshift_variant
83	CMML	TET2	NM_001127208.2:c.3356dupT	NP_001120680.1:p.Leu1119PhefsTer11	10.53	insertion	frameshift_variant
84	CMML	UZAF1	NM_006758.2:c.467G>A	NP_006749.1:p.Arg156His	50.61	snv	missense_variant
84	CMML	TET2	NM_001127208.2:c.4132T>C	NP_001120680.1:p.Cys1378Arg	47.6	snv	missense_variant
84	CMML	NRAS	NM_002524.4:c.37G>C	NP_002515.1:p.Gly13Arg	15.64	snv	missense_variant
84	CMML	RUNX1	NM_001754.4:c.592G>A	NP_001745.2:p.Asp198Asn	14.07	snv	missense_variant
84	CMML	NRAS	NM_002524.4:c.35G>C	NP_002515.1:p.Gly12Ala	8.18	snv	missense_variant
84	CMML	RUNX1	NM_001754.4:c.1022_1023insTTGGC	NP_001745.2:p.Ile342TrpfsTer254	5.98	insertion	frameshift_variant
85	CMML	TET2	NM_001127208.2:c.2305delC	NP_001120680.1:p.Gln769SerfsTer44	42.01	deletion	frameshift_variant
85	CMML	TET2	NM_001127208.2:c.5140_5141delIAA	NP_001120680.1:p.Asn1714CysfsTer14	9.09	deletion	frameshift_variant
86	CMML	SF3B1	NM_012433.2:c.1997A>G	NP_036565.2:p.Lys666Arg	34.73	snv	missense_variant
86	CMML	TET2	NM_001127208.2:c.4268T>A	NP_001120680.1:p.Val1423Asp	2.37	snv	missense_variant
87	CMML	TP53	NM_000546.5:c.818G>A	NP_000537.3:p.Arg273His	71.46	snv	missense_variant
87	CMML	PRPF8	NM_006445.3:c.4792G>A	NP_006436.3:p.Asp1598Asn	6.99	snv	missense_variant
87	CMML	RUNX1	NM_001754.4:c.422C>A	NP_001745.2:p.Ser141Ter	34.33	snv	stop_gained
87	CMML	CBL	NM_005188.3:c.1259G>A	NP_005179.2:p.Arg420Gln	1.54	snv	missense_variant
88	CMML	SRSF2	NM_003016.4:c.284C>A	NP_003007.2:p.Pro95His	49.76	snv	missense_variant
88	CMML	TET2	NM_001127208.2:c.2207delC	NP_001120680.1:p.Ser736TyrfsTer15	48.55	deletion	frameshift_variant
88	CMML	TET2	NM_001127208.2:c.4661_4664delCAGA	NP_001120680.1:p.Thr1554SerfsTer16	27.45	deletion	frameshift_variant
88	CMML	TET2	NM_001127208.2:c.1630C>T	NP_001120680.1:p.Arg544Ter	12.07	snv	stop_gained
88	CMML	ASXL1	NM_015338.5:c.2077C>T	NP_056153.2:p.Arg693Ter	11.29	snv	stop_gained
88	CMML	RUNX1	NM_001754.4:c.425C>T	NP_001745.2:p.Ala142Val	6.21	snv	missense_variant
88	CMML	TET2	NM_001127208.2:c.2147dupC	NP_001120680.1:p.His717ThrfsTer6	5.05	insertion	frameshift_variant
89	CMML	TET2	NM_001127208.2:c.4156C>G	NP_001120680.1:p.His1386Asp	47.43	snv	missense_variant
89	CMML	SRSF2	NM_003016.4:c.284C>A	NP_003007.2:p.Pro95His	39.48	snv	missense_variant
89	CMML	NRAS	NM_002524.4:c.34G>C	NP_002515.1:p.Gly12Arg	19.14	snv	missense_variant
90	CMML	TET2	NM_001127208.2:c.3283delA	NP_001120680.1:p.Arg1095GlufsTer11	35.4	deletion	frameshift_variant
90	CMML	TET2	NM_001127208.2:c.2155_2156delITT	NP_001120680.1:p.Leu719AlafsTer3	28.68	deletion	frameshift_variant
90	CMML	SF3B1	NM_012433.2:c.2098A>G	NP_036565.2:p.Lys700Glu	22.44	snv	missense_variant
90	CMML	ZRSR2	NM_005089.3:c.976T>C	NP_005080.1:p.Cys326Arg	8.95	snv	missense_variant
90	CMML	ZRSR2	NM_005089.3:c.772-1G>A		7.66	snv	splice_acceptor_variant
90	CMML	ZRSR2	NM_005089.3:c.868C>T	NP_005080.1:p.Arg290Ter	5.14	snv	stop_gained
90	CMML	ZRSR2	NM_005089.3:c.1003_1013delCCCAACAA TGA	NP_005080.1:p.Pro335IlefsTer5	2.5	deletion	frameshift_variant
91	CMML	SRSF2	NM_003016.4:c.284C>A	NP_003007.2:p.Pro95His	46.79	snv	missense_variant
91	CMML	IDH2	NM_002168.2:c.419G>T	NP_002159.2:p.Arg140Leu	45.18	snv	missense_variant
91	CMML	KRAS	NM_033360.2:c.34G>C	NP_203524.1:p.Gly12Arg	43.31	snv	missense_variant
91	CMML	ASXL1	NM_015338.5:c.2415dupC	NP_056153.2:p.Thr806HisfsTer16	36.9	insertion	frameshift_variant
92	CMML	TET2	NM_001127208.2:c.5618T>C	NP_001120680.1:p.Ile1873Thr	43.25	snv	missense_variant
92	CMML	TET2	NM_001127208.2:c.2677delG	NP_001120680.1:p.Ala893LeufsTer28	42.41	deletion	frameshift_variant
92	CMML	SH2B3	NM_005475.2:c.933_940delGAGCACAG	NP_005466.1:p.Thr313ArgfsTer11	20.09	deletion	frameshift_variant
92	CMML	SH2B3	NM_005475.2:c.1204G>A	NP_005466.1:p.Val402Met	12.64	snv	missense_variant
92	CMML	SH2B3	NM_005475.2:c.1038dupG	NP_005466.1:p.Leu347AlafsTer38	6.68	insertion	frameshift_variant
92	CMML	SH2B3	NM_005475.2:c.20_21insAA	NP_005466.1:p.Pro85SerfsTer29	5.43	insertion	frameshift_variant
92	CMML	TP53	NM_000546.5:c.476C>G	NP_000537.3:p.Ala159Gly	2.06	snv	missense_variant

Supplementary Material 1 – Methods

Peripheral blood samples were centrifuged on a density gradient (Biocoll Separating Solution, 1.077g/ml; BiochromAG, Berlin, Germany) for PBMC separation. Then, granulocytes were obtained from remaining material by sedimentation on a 2% dextran solution. Genomic DNA was extracted with an automated procedure using the BioRobot M48 (Qiagen, Hilden, Germany).

Saliva DNA was obtained using the Oragene-DNA kit (DNA Genotek, Ottawa, Canada) and quantified using the Nanodrop spectrophotometer (Thermo Fisher Scientific, Wilmington, USA). The presence of bacterial DNA was studied by 16s ribosomal gene amplification in 32 saliva samples using the Platinum PCR SuperMix High Fidelity (Thermo Fisher Scientific) and PCR products were visualized by agarose gel electrophoresis (E-Gel Agarose Gels 2% SYBR Safe, Thermo Fisher Scientific, Carlsbad, USA).

CD3+ T lymphocytes were isolated from PBMC by positive selection using MACS immunomagnetic columns (Miltenyi Biotec, Bergisch Gladbach, Germany) from cases in which the saliva sample was positive for the driver mutation (64 *JAK2*, 12 *CALR*, 5 *MPL*). Purity of the isolated CD3+ fraction was assessed by flow cytometry in available samples and was >98% (FACSCanto II System 8-color, BD Biosciences, San José, USA).

The mutational status of *JAK2* gene was determined by allele-specific real time PCR (7500 Fast real time PCR System, Applied Biosystems, Foster City, USA), *CALR* gene by fragment analysis electrophoresis (3500 Genetic Analyzer, Applied Biosystems, Foster City, USA) and *MPL* gene by digital PCR (QuantStudio 3D, Applied Biosystems, Foster City, USA) for the W515L, S204P, S505N and R592Q mutations.

Statistical analyses were performed using the statistical package for the social sciences software version 22.0 (SPSS, Chicago, USA). In order to assess the association between variant allele frequency (VAF) of the mutation in granulocytes and saliva we used Spearman's correlation. To analyze leukocyte counts in patients with positive and negative saliva, the Mann-Whitney test was used. After testing for normality of the data, a Wilcoxon signed rank test was used for non-parametric paired data. Statistical significance was set as $P < 0.05$.