UNIVERSITY OF MODENA AND REGGIO EMILIA

Doctor of Philosophy in Molecular and Regenerative Medicine

Single-cell genomics unmasks clonal hierarchy and heterogeneity in Myeloproliferative Neoplasms

Candidate:
SELENE MALLIA

Supervisor:
Prof. ROSSELLA MANFREDINI
Ph.D. School Director:
Prof. MICHELE DE LUCA

XXXIV DOCTORATE CYCLE

ABSTRACT

Somatic mutations in Hematopoietic Stem Cells (HSCs) cause Myeloproliferative Neoplasms (MPNs), including Polycythemia Vera, Essential Thrombocythemia and Primary Myelofibrosis (PMF).

PMF is a heterogeneous disorder consisting of bone marrow fibrosis, megakaryocyte hyperplasia and extramedullary hematopoiesis and is characterized by the worst prognosis among MPNs. About 15-20% of patients are unresponsive to conventional therapies and develop Acute Myeloid Leukemia (AML).

In HSCs the main mutations, identified as "driver mutations" during MPNs pathogenesis, involve *JAK2*, *CALR* and *MPL* genes; in addition, many other genetic alterations contribute to the prognosis worsening and the development of AML.

Disease progression and leukemic evolution in PMF results from an increase of the genomic complexity and clonal heterogeneity.

Many studies confirmed that the mutational acquisition order affects the clinical outcome. However, the clonal architecture determining disease evolution and the clones guiding leukemic transformation are poorly understood.

Recent studies demonstrate that single-cell (sc-) genomics is a sensitive technique suitable to study clonal heterogeneity and to detect the evolution of the malignant cells in hematological neoplasms.

For this reason, we used the sc-genomics approach to clarify the clonal complexity in PMF.

Firstly, we developed a workflow for $CD34^+$ Hematopoietic Stem Progenitor Cells (HSPCs) isolation from cord blood, fixation and immunostaining for CD34, in order to singularly separate the cells by DEPArray system (Menarini Silicon Biosystem) and to obtain a cell population suitable for sc-analysis.

Then, we compared different whole genome amplification (WGA) protocols for single cells in order to obtain a uniform DNA amplification for Sanger sequencing and minimize allele drop out effect.

Based on this method, we analyzed the $CD34^+$ HSPCs of a PMF patient carrying JAK2V617F and other MPN frequent mutations. This patient was treated with JAK2-inhibitor Ruxolitinib but he was unresponsive to therapy and evolved to AML.

In order to reconstruct the clonal hierarchy and architecture, we analyzed $CD34^+$ cells during chronic phase (T1), the accelerated phase (T2) and the AML phase (T3).

By means to sc-analysis, we established that *TET2* was the first mutated gene, preceding *JAK2* mutation, and this probably conferred a lower sensitivity to treatment.
Moreover, we identified an increase of the allele burden of the *TP53* mutation during disease progression, suggesting that *TP53*-mutated clones supported the accelerated phase (T2).
Interestingly, we already detected in T1 phase a small cell fraction, undetectable by bulk NGS and carrying the leukemogenic *FLT3* mutation, probably driving the T3 phase.
Finally, we characterized *SRSF2* homozygous mutation that has not been described yet.
Altogether our data demonstrate that sc-genomics is a promising method to uncover clonal heterogeneity in MPNs, highlighting the early occurrence of pro-leukemic mutations and to describe the real scenario of mutational events in hematological diseases.

Keywords:

Stem Cells, Single-cell, Genomics, Myelofibrosis, Leukemic Evolution

SINTESI

Le Neoplasie Mieloproliferative (MPN) Croniche Philadelphia-negative sono disordini ematologici caratterizzati dalla presenza di mutazioni somatiche che colpiscono le cellule staminali ematopoietiche e dalla proliferazione incontrollata di una o più linee mieloidi; esse comprendono la Policitemia Vera, la Trombocitemia Essenziale e la Mielofibrosi Primaria (PMF).

La PMF è una neoplasia eterogenea, contraddistinta dalla presenza di fibrosi midollare, iperplasia megacariocitaria e ematopoiesi extramidollare, e mostra la peggiore prognosi tra tutte le MPN. I pazienti spesso non rispondono ai trattamenti e nel 15-20% dei casi sviluppano una Leucemia Mieloide Acuta (LMA).

Le mutazioni ricorrenti, conosciute come "mutazioni *drivers*", interessano i geni *JAK2*, *CALR* e *MPL*, ma a complicare il profilo mutazionale intervengono altre alterazioni che sono spesso responsabili del peggioramento del quadro clinico e della trasformazione leucemica.

La progressione della malattia e l'evoluzione leucemica nella PMF è accompagnata da un aumento della complessità genomica e dell'eterogeneità clonale.

Molti studi hanno confermato come l'ordine di acquisizione delle mutazioni influenzi il decorso clinico. Tuttavia sono ancora poco conosciute le caratteristiche dei cloni che determinano la malattia e che guidano la trasformazione leucemica.

Studi recenti hanno dimostrato come la genomica su singola cellula sia una tecnica sensibile per studiare l'eterogeneità clonale e l'evoluzione delle leucemie.

Per questa ragione, abbiamo adottato un approccio di genomica su singola cellula per studiare la complessità clonale della PMF.

Dapprima abbiamo sviluppato un metodo di isolamento delle cellule staminali e progenitori ematopoietici $CD34^+$ dal sangue cordonale, di fissazione e marcatura delle cellule $CD34^+$, al fine di ottenere una popolazione cellulare adatta alla separazione in singole cellule, sfruttando il sistema del DEPArray (Menarini Silicon Biosystem).

In seguito, abbiamo confrontato diversi protocolli di amplificazione dell'intero genoma su singola cellula al fine di ottenere un'amplificazione omogenea, minimizzando l'effetto di *allele drop out*, per proseguire con il sequenziamento Sanger.

Usando questa procedura, abbiamo analizzato le cellule $CD34^+$ di un paziente affetto da PMF, positivo per la mutazione JAK2V617F e per altre alterazioni genetiche, caratteristiche delle MPN.

Il paziente, nonostante il trattamento con il JAK2-inibitore Ruxolitinib, ha sviluppato una LMA.

Al fine di ricostruire la gerarchia e l'architettura clonale, abbiamo analizzato le cellule $CD34^+$ alla diagnosi (T1), durante la fase accelerata (T2) e nella fase di LMA (T3).

Grazie alla analisi su singola cellula, abbiamo stabilito che il primo evento mutazionale investe *TET2*, precedendo la mutazione di *JAK2*, e probabilmente influenzando negativamente la risposta alla terapia.

Abbiamo osservato, inoltre, un aumento dei cloni mutati per *TP53* durante la progressione della malattia, suggerendo che siano stati questi cloni a supportare la fase T2.

Inaspettatamente, già nella fase T1, abbiamo riscontrato una piccola popolazione cellulare recante una mutazione pro-leucemica su *FLT3*, alterazione che non era stata evidenziata dall'analisi in NGS ma che, verosimilmente, ha guidato lo sviluppo della fase T3.

Infine, abbiamo evidenziato una mutazione omozigote su *SRSF2* non ancora descritta.

Tutti i nostri dati, confermano quindi come la genomica su singola cellula sia una tecnologia promettente di analisi della eterogeneità clonale delle MPN e che permetta sia di evidenziare precocemente caratteristiche leucemiche sia di ottenere un quadro chiaro degli eventi mutazionali che interessano i disordini ematologici.

Parole chiave:

Cellule staminali , Singola cellula, Genomica, Mielofibrosi, Evoluzione leucemica

Index

1. INTRODUCTION

1.1 Myeloproliferative Neoplasms

1.1.1 *Definition and classification*

Philadelphia (Ph)-negative Chronic Myeloproliferative neoplasms (MPNs) are a group of clonal hematological diseases caused by somatic mutations that affect hematopoietic stem and progenitor cells (HSPCs) in bone marrow (BM). Consistent with the chronic nature of the disorders, MPNs are characterized by the slow accumulation of abnormal myeloid cells in peripheral blood (PB) [1].

For the first time in 1951 William Dameshek described these hematological diseases characterized by hyperproliferation of precursors in BM and overproduction of mature blood cells and he provided a classification based on their clinical features [2]. In 1960 Nowell and Hungerford discovered a new chromosomal abnormality, which is known as the Philadelphia chromosome. In particular, it is a small chromosome derived from a reciprocal chromosomal translocation, involving chromosomes 9 and 22, t(9;22)(q34.1;q11.2) with the creation of the fusion gene *BCR-ABL1*. This acquired somatic alteration became the marker to diagnose chronic myeloid leukemia (CML) [3,4].

Until 2008 the MPNs were classically called "Chronic Myeloproliferative Diseases" and subsequently were renamed "Myeloproliferative Neoplasm" by the World Health Organization (WHO) in relation to their neoplastic behavior. According to the WHO guidelines defined in 2008 and revised in 2016, MPNs include several subtypes: CML, chronic neutrophilic leukemia (CNL), chronic eosinophilic leukemia (CEL), polycythemia vera (PV), essential thrombocythemia (ET) and primary myelofibrosis (PMF). In particular, three main MPNs are classified as Philadelphia negative disorders given by the absence of the *BCR-ABL* cytogenetic marker: PV, ET and PMF [5,6] .

Table 1 synthesizes the WHO classification of MPNs and leukemia revised in 2016. The definition of PMF is particularly worthy of attention; in 2016 the WHO indicated two possible stages of PMF, the prefibrotic/early (pre-PMF) stage and the overt fibrotic stage. Pre-PMF is thus characterized as a PMF phase and is distinguished from ET; the two disorders may be confused but have different outcomes [6].

In this thesis work, Philadelphia negative disease will be generally called MPNs.

WHO Myeloid Neoplasm and Leukemia classification (revised in 2016)
Myeloproliferative neoplasms (MPN)
• Chronic myeloid leukemia (CML), *BCR-ABL1*-positive
• Chronic neutrophilic leukemia (CNL)
• Polycythemia vera (PV)
• Primary Myelofibrosis (PMF) *PMF, prefibrotic/early stage* *PMF, overt fibrotic stage*
• Essential thrombocythemia (ET)
• Chronic eosinophilic leukemia (CEL)
• Mastocytosis
Myeloid/lymphoid neoplasms with eosinophilia and rearrangement of PDGFRA, PDGFRB, or FGFR1, or with PCM1-JAK2
Myelodysplastic/myeloproliferative neoplasms (MDS/MPN)
Myelodysplastic syndromes (MDS)
Acute myeloid leukemia (AML) and related neoplasms

Table 1: **World Health Organization (WHO) classification of myeloid malignancies revised in 2016.**

MPNs are rare hematological diseases with an incidence rate lower than 6 per 100.000 cases diagnosed each year; PV from 0.4 to 2.8, ET from 0.38 to 1.7, and PMF from 0.1 to 1.0 in detail [7].

Several characteristics are common among MPNs, such as extramedullary hematopoiesis (EMH) or spontaneous transformation in Acute Myeloid Leukemia (AML) [8]. Notable is the possibility of each MPNs to transform into another; PV can evolve to ET and *vice versa*, PMF can progress to PV, PV and ET patients can develop a myelofibrosis (MF) condition post-PV and post-ET [9]. This ability to transform to each other generated the hypothesis that MPNs are a "biological *continuum*" [10]. Nevertheless, PV, ET and PMF are heterogeneous diseases with different clinical features and prognosis [11], as it will be examined in detail.

MPNs arise from the transformation of one or small group of hematopoietic stem cells (HSCs). This event compromises the normal hematopoiesis, leading to the overproduction of mature blood cells, in particular expanding erythrocytes in PV, increasing platelets count in ET and generating bone marrow fibrosis in PMF. The clinical consequences of this hematological

disturbance concern marrow hypercellularity, increase of thrombosis and hemorrhage risk and possibility of leukemic transformation in the long term [12].

In normal conditions, several cytokines contribute to the control of hematopoietic proliferation, differentiation and survival. Despite the stimulus and cellular response must be accurate, as in all biological signalling, hematopoietic progenitors in MPNs have hypersensitivity and/or independence to endogenous growth factors in lineage-specific manner and this is closely related to clinico-pathological features in MPNs [13,14].

It is necessary to understand, however, the biological basis of MPNs, to determine a cause-and-effect relationship between the abnormal behaviour of myeloid cells and the pathogenesis. For this purpose, it may be worth asking the question: what do MPNs have in common? The answer could be into the genesis of the diseases: an alteration in HSCs normal biology.

HSCs generate different cell lineages (lymphoid and myeloid) and all cells of the hematopoietic system [15]. It is easy to think that the key to understanding the MPNs pathobiology is in hematopoiesis. Therefore, whenever self-renewal or differentiation of HSCs is alterated, does it always generate MPNs?

The discovery of clonal hematopoiesis (CH) made the scenario more complex; CH is a condition found in some individuals where a small "abnormal" cellular population, presenting one or more somatic mutations, origines a large cellular clone that dominates the unmutated cellular counterparts. More specifically, CH is defined as a fitness advantage of a cellular clone on the other. Individuals affected by CH had a wide range of phenotypic conditions, riding from no effect on health to a permissive state of myeloid disease and/or pre-leukemic state [16]. Despite CH is a clonal hematological disorder and it shares common characteristics, MPNs and CH are different entities, in fact, MPNs always show neoplastic behaviour, as it has already explained [17, 18]. Such behaviour has long been recognized as a consequence of DNA mutation. In particular, I am talking about three genetic alterations, known as "driver" mutations, affecting Janus kinase 2 (*JAK2*), calreticulin (*CALR*) and Myeloproliferative leukemia virus oncogene (*MPL*) genes [19-21].

A chapter of this thesis work will be focused on the mutation profile of MPNs.

1.1.2 Polycythemia Vera (PV)

PV is the most common disease among MPNs, the incidence rate is about 0.68-2.6 per 100.000 individuals per year in the European population. In general, MPNs distribution in population is influenced by gender and in male population, PV and PMF had a higher prevalence than ET [7,22-24].

MPNs are considered disorders of adulthood with an onset about 65-70 years and the median age diagnosis in PV-patients is 65 years [7,22].

In recent years, MPNs diagnosis is more frequent in pediatric and young patients than in the past. Interestingly, a recent study, including a cohort of 396 (84%) ET-patients and 75 (16%) PV-patients, demonstrated that all of them are very young with a median age at diagnosis of 9.3 years for ET patients and 12 years for PV. All of these patients were identified as negative for common mutations in MPNs [25].

The increase of red-cells in PB, known as erythrocytosis (Figure 1, a), is the hallmark of PV differently than ET and PMF. This effect in PB is determined by BM hypercellularity and mature myeloid lineage overproduction, resulting in red cells, platelets and granulocytes hyperproliferation (panmyelosis) (Figure 1, b) [26,27].

(a) *(b)*

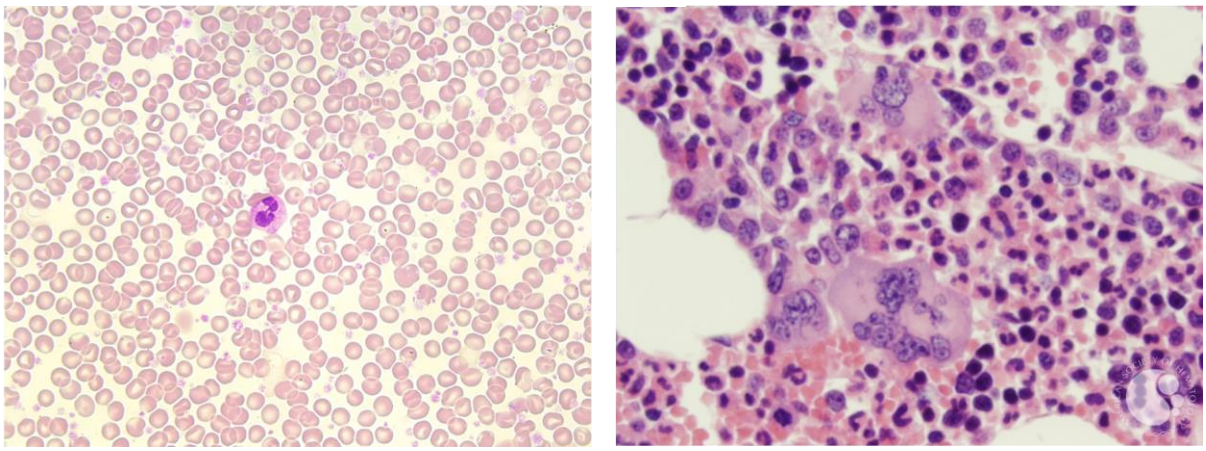


Figure 1: Polycythemia vera. a. Peripheral blood smear from PV patient characterized by moderate leukocytosis with neutrophilia, elevated RBC count, and normal platelet count. Image is taken from MLS Collection, University of Alberta. **b.** The trephine core shows a hypercellular marrow for age with panmyelosis (proliferation of the erythroid, granulocytic, and megakaryocytic lineages). Megakaryocytes are increased and include frequent hyperlobated forms. Image taken from American Society of Hematology Image Bank.

In PV, erythrocytosis is frequently associated with thrombocytosis and/or leukocytosis, increasing the thrombotic risk [7]. The reduction of erythropoietin (EPO) endogenous levels was observed in PV patients, resulting in the hypersensitivity of erythroid precursors to EPO and in the increase of red-cells count. In normal conditions, the progenitor cells must respond to EPO signaling by differentiating in erythroid cells. In contrast, in PV condition erythroid progenitors showed the ability to proliferate even in absence of EPO endogenous stimulation. [28, 29]. Moreover, in PV, erythroid cells had more sensitivity to several other cytokines and growth factors, such as stem cell factor (SCF), interleukin (IL)-3 and granulocyte-macrophage colony-stimulating factor (GM-CSF) [30-32]. Erythropoiesis is considered a major criteria to diagnose PV and subnormal serum EPO level is a minor criteria (Table 2) [6, 7].

2016 WHO PV diagnostic criteria
Major criteria
1. Hemoglobin >16.5 g/dL in men or >16.0 g/dL in women, OR hematocrit >49% in men or >48% in women, OR increased red cell mass (more than 25% above mean normal predicted value)
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 JAK2 (V617F) or JAK2 exon 12 mutation
Minor criteria
Subnormal serum erythropoietin level
Diagnosis of PV requires meeting all 3 major criteria or the first 2 major criteria and the minor criterion.

Table 2: 2016 WHO diagnostic criteria for PV. Notes: 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 myelofibrosis)

The phenomenon of EPO-independent erythroid proliferation has been well known for a long time [28, 29]. However, the cause does not result from a genetic alteration of the EPO receptor (EPO-R). EPO-R mutations have not been observed in PV patients [33]. In 2005, the discovery of causative mutation in *JAK2* gene was of great help in understanding of EPO hypersensitivity

and/or independence in PV erythroid progenitors. Besides, IL-3 dependent cell lines, carrying JAK2V617F mutations, showed the cytokine hypersensitivity or independence mimicking the EPO hypersensitivity or independence observed in PV [34]. The positivity to the *JAK2* mutation, in particular JAK2V617F (considered as "driver" mutation) and *JAK2* exon 12 mutations, is essential for PV diagnosis. In over 95% of PV cases, at least one *JAK2* mutation is found and this observation is one of the major diagnostic criteria for PV (Table 2) [6,7]. The biological correlation between *JAK2* mutation and EPO hypersensitivity and/or independence in PV depends on the role of JAK2 in erythroid progenitors proliferation. As it will be discussed in detail, *JAK2* mutations cause a gain-of function of JAK2 and, consequently, the JAK/STAT pathway constitutive activation [35].

The expansion of red-cells fraction and the alteration in blood parameters are deeply related to the symptoms of PV patients. Although the onset of the disease may be asymptomatic, PV patients frequently develop clinical manifestations including headache, loss of weight, blurred vision, mucous membrane bleeding, erythromelalgia and pruritus. During the PV progression, HSCs migrate from canonical hematopoietic sites to new sites involving liver and spleen, causing hepato-splenomegaly, in a process called EMH [9, 27, 36,37]. Moreover, the excessive increase of blood hyperviscosity enhances the risk of arterial or venous thrombosis and/or bleeding in PV patients [38].

Thrombotic events contribute to shortening survival in PV patients in combination with advanced age (greater than 60 years). The PV median survival is of about 14 years, but many other events occur influencing this data. For instance, allele burden of JAK2V617F mutation >50% are associated with evolution to post-PV MF. Old age, alterations in blood parameters such as leukocytosis and abnormal karyotype are involved in AML development [37, 39, 40].

The main therapeutic strategies in PV are based on prevention of clinical complications and symptoms affecting the circulatory system. In PV patients, phlebotomy and aspirin are the primary treatments, considering that the risk of death is related to thrombotic events. Moreover, the administration of low-dose aspirin reduces microvascular effects, headache, thrombotic events and the risk of death for cardiovascular causes [37,41].

In high-risk PV patients (>60 years), a first-choice agent is hydroxyurea (HU). This cytoreductive therapeutic strategy is well tolerated in long-time treatment. On the other hand,

interferon-α (IFN-α) and busulfan are second-line drugs [9, 37, 42, 43]. Dosing of JAK2-inhibitors, such as Ruxolitinib, are recommended only for the splenomegaly treatment [9, 37].

1.1.3 Essential Thrombocythemia (ET)

ET has the slowest progression among the MPNs, therefore survival is higher in ET patients than PV or PMF patients and it is about 20 years[37]. However, if on one hand sex and age do not influence survival rate [44], in the other one ET has an higher prevalence in female patients compared to PV or PMF. [23, 24]. In particular, the gender disparity is more evident in young and midlife ET patients, suggesting a potential hormonal influence in the pathology [45].

Generally, the ET diagnosis occurs at the median age of 68 years, although cases have been described in childhood and youth, and the incidence rate in the populations is between 0.38 and 1.7 per 100.000 people per year [7, 25, 46].

The abnormal proliferation of megakaryocyte lineage characterizes ET, causing the increase of platelets count [47,48]. The presence of malignant megakaryocytic clones induces a persistent non-reactive thrombocytosis. For this reason thrombocytosis is one of major diagnostic criteria for ET (Table 3) [6, 48, 49].

2016 WHO ET diagnostic criteria
Major criteria
1. Platelet count ≥450x10^9/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 criteria
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.

Table 3: 2016 WHO diagnostic criteria for ET

Given the thrombocytosis in ET, the platelets are increased and clustered, observing the PB smear (Figure 2, a) [12]. In BM, cellularity is normal, the cavity is enriched with mature megakaryocytes with hyperlobulated nuclei (Figure 2, b) [50, 51].

(a) *(b)*

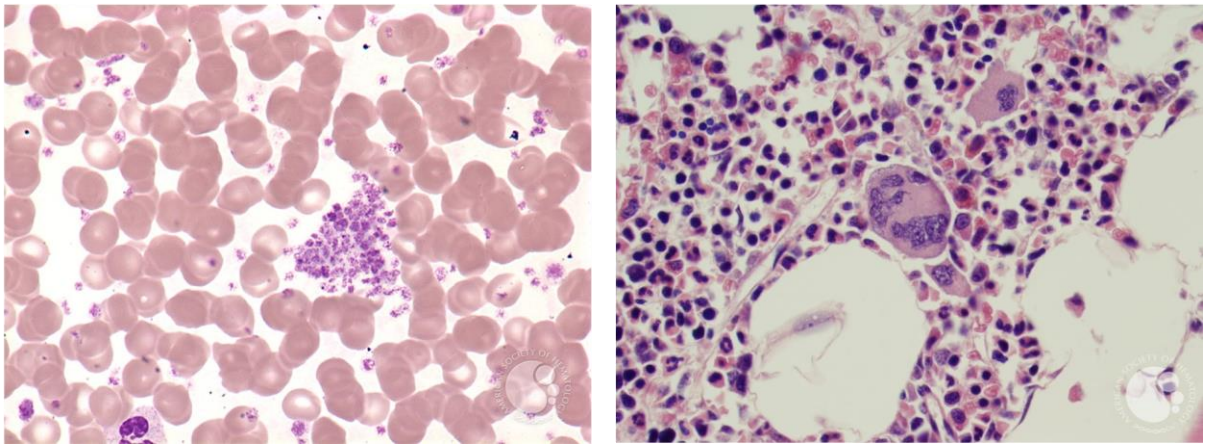


Figure 2: Essential thrombocythemia: a. Peripheral blood smear from ET patient. Platelets are markedly increased in the peripheral smear. Some platelet forms cluster together. Images are taken from the American Society of Hematology Image Bank. **b.** Trephine core biopsy representing marrow hypercellularity with increased megakaryocytes in ET patient. The megakaryocytes are dispersed throughout the marrow and include frequent large forms with abundant cytoplasm and deeply lobated nuclei.

As regards the ET diagnosis, it is important to remember the classification of the MPNs proposed by the WHO in 2016, in which ET is clearly distinguishable from pre-PMF (Table 4). ET and pre-PMF show different clinical outcomes; pre-PMF patients have a better prognosis compared with PMF patients but worse respect to ET patients. Furthermore, pre-PMF is considered an earlier stage of PMF [6, 7, 44]. If on one side both ET and pre-PMF diagnosis is accompanied by thrombocytosis [52], on the other side, BM biopsy helps to clarify disease natures, in fact, BM morphology is considered as the gold standard of diagnosis [52]. In contrast to "true" ET, BM biopsy in pre-PMF shows hypercellularity, high number of neutrophils and atypical megakaryocyte with increased nucleo–cytoplasmic ratio, bulbous, cloud-like, hyperchromatic nuclei (Figure 3) [7, 53, 54]. Moreover, ET and pre-PMF patients show not only different phenotype at diagnosis but also different outcomes; pre-PMF patients had shorter overall survival and higher incidence of AML transformation than ET patients [55].

2016 WHO prePMF diagnostic criteria
Major criteria
1. Megakaryocytic proliferation and atypia, without reticulin fibrosis >grade 1, accompanied by increased age-adjusted bone marrow cellularity, granulocytic proliferation, and often decreased erythropoiesis
2. Not meeting WHO criteria for BCR-ABL1⁺ CML, PV, ET, 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 minor reactive BM reticulin fibrosis†
Minor criteria
Presence of at least 1 of the following, confirmed in 2 consecutive determinations: • Anemia not attributed to a comorbid condition • Leukocytosis (WBC count ≥11x10^9/L) • Palpable splenomegaly • LDH level increased to above upper normal limit of institutional reference range
Diagnosis of prefibrotic myelofibrosis requires meeting all 3 major criteria, and at least 1 minor criterion.

Table 4: 2016 WHO diagnostic criteria for pre-PMF. Notes: * 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) reticulin fibrosis secondary to infection, autoimmune disorder, or other chronic inflammatory conditions, hairy cell leukemia or other lymphoid neoplasm, metastatic malignancy, or toxic myelopathies

(a) *(b)*

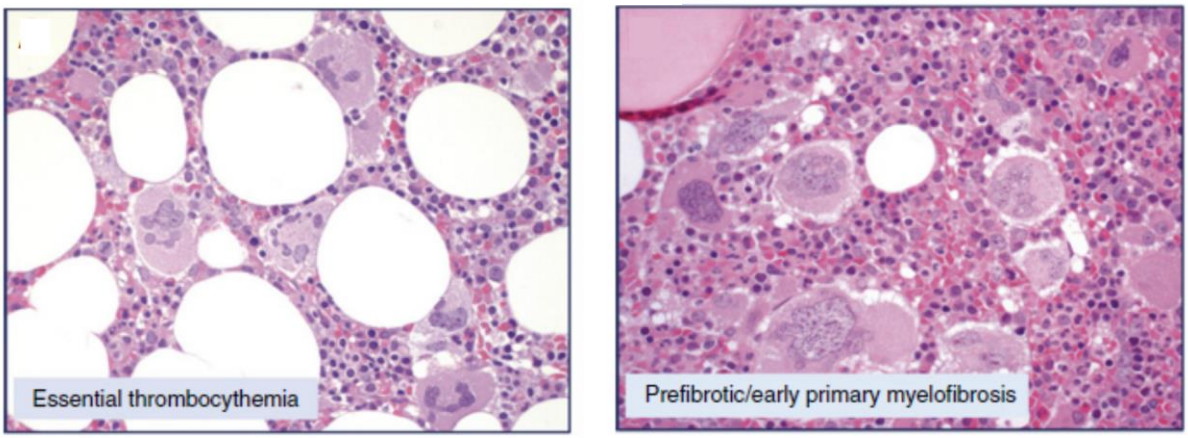


Figure 3: Comparison between ET and pre-PMF BM biopsies. a. ET biopsy showing normocellular marrow characterized by the proliferation of giant megakaryocytes with hyperlobulated nuclei, scattered or in loose clusters. **b.** Pre-PMF biopsy displaying hypercellular marrow with granulocytic proliferation and large megakaryocytes with atypical bulbous nuclei. Adapted from: Rumi E. et al, Blood 2017

In ET both the platelets morphology and number are altered. Moreover, this phenomenon is associated with elevated levels of platelet-specific proteins and thromboxane generation, reflecting an inappropriate state of platelet activation and resulting in the increased risk of thrombosis and bleeding in patients [56, 57]. It is interesting to highlight that in male ET patients, the risk of venous thrombosis event increases, showing a correlation beetween risk and vascular anatomy and inflamatory response sex-dependent [58]. Moreover, thrombotic risk in ET is associated with specific common mutations in MPNs, in particular *CALR*-mutation is a protective factor compared to JAK2V617F in ET patients [59].

Actually, molecular studies demonstrated that ET is frequently associated with the presence of at least one of MPNs common "driver" mutations. The percentage distribution of *JAK2*, *CALR* or *MPL* gene mutations is respectively 55%, 25% and 3% in the ET-affected population, although 17% of ET patients do not carry any of these and they are considered triple-negative (TN) [37].

The ET occurrence is often asymptomatic or is not associated with a specific symptomatology such as increased fatigue, peripheral paresthesia, headache, light-headedness and blurred vision. Other clinical evidence is related to thrombocytosis and abnormal platelets activation, involving the vascular system and causing erythromelalgia and major thrombosis [60-62]. Interesting to note is the risk of patients developing hemorrhagic events, despite the large number of platelets. This is probably due to a depletion and sequestration of the Von Willebrand Factor (VWF), leading to the acquired von Willebrand syndrome (AVWF). Moreover, AVWS in ET patients is mainly associated with JAK2V617F mutation [63, 64].

Splenomegaly is a common feature in MPNs and is one of the main criteria in MPNs diagnosis [6, 7, 62]. In particular, 10-20% of ET patients show spleen enlargement, predominantly, in younger individuals, in male, with higher platelets number and JAK2V617F allele burden. Splenomegaly impacts on overall survival and thrombotic risk, providing a prognostic negative factor in ET patients [65-66]. Nevertheless, as in PV, many other risk factors affect survival and are related to advanced age (>60 years), leukocytosis, history of thrombosis [37, 67].

ET can progress to PV or MF [7,9,37]. Although ET shows increased survival among MPNs, MF transformation may expose patients to the risk of fatal outcome. ET transformation to MF occurs in 0.8-4.9% of patients at 10 years and 4-11% at 15 years [68].

As other MPNs, the chronic diseases can transform into an acute disorder developing AML in patients. The cases of leukemic transformation from ET correspond to 0.7-3% of cases at 10 years and 2.1-5.3% at 15 years [68]. Recent studies demonstrate that additional mutations are necessary for disease evolution such as mutations of *TP53, RUNX1,* and *SRSF2* genes, suggesting a time-dependent molecular mechanism [69].

The main therapeutic indications in patients with ET concern thrombo-hemorrhagic complications prevention, limiting circulatory symptoms. To this purpose, low-dose aspirin is used in low-risk ET patients, considering that ET is an indolent disease. In addition, cytoreductive treatments, such as HU combined with once-daily aspirin therapy, are administered to high-risk and older patients. On the contrary, drugs considered myelosuppressive, such as anagrelide and IFN-α, should be used as "second-line" therapies in patients refractory to HU and in younger patients [37, 70-72].

1.1.4 Primary Myelofibrosis (PMF)

1.1.4.1 Epidemiology

PMF is the most severe disease among MPNs. Therefore, PMF patients have median survival about 6 years [11,39,73]. The PMF incidence rate is 0.3-2.0 per 100.000 per year, classifying as rarer compared to PV and ET. In contrast to ET and like PV, PMF had a higher distribution in male individuals than female. Moreover, male have a higher rate of myelofibrosis transformation post-PV/ET [7, 22-24, 73].

The median age at diagnosis is 70 years. [7, 22]. PMF in pediatric age is extremely rare with few cases reported in literature data and unknown incidence. The median age in PMF children patients is between 0 and 17 years. Most of these children showed different histologic, clinical and genetic features from PMF adult patients and had a good prognosis and spontaneous resolution of the disease. Moreover, more PMF cases in pediatric age are associated with familiar PMF, suggesting a kind of hereditary myelofibrosis [74, 75]

1.1.4.2 Diagnosis

PMF is characterized by stem cell-derived clonal myeloproliferation, involving the megakaryocyte lineage with an high pathogenetic relevance. In PMF, hyperplastic and

dysplastic megakaryocytes are accompanied by BM fibrosis, which is the main feature of the disease. BM fibrosis is correlated to other important characteristics of PMF, such as osteosclerosis, angiogenesis, abnormal cytokine production, inflammation, hepatosplenomegaly and EMH [76-77].

Given the wide impact of the disease on the structure and functions of the BM, the BM morphology analysis is the gold standard for PMF diagnosis and more in general for all chronic MPNs. According to the WHO system revised in 2016, BM histology is the most important criterion to distinguish the pre-PMF from "true" ET (respectively Table 3 and 4) and to diagnose over-PMF (Table 5) [6,7].

2016 WHO overtPMF diagnostic criteria
Major criteria
1. Presence of megakaryocytic proliferation and atypia, accompanied by either reticulin and/or collagen fibrosis grades 2 or 3
2. Not meeting WHO criteria for BCR-ABL1⁺ CML, PV, ET, 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 minor reactive BM reticulin fibrosis†
Minor criteria
Presence of at least 1 of the following, confirmed in 2 consecutive determinations: • Anemia not attributed to a comorbid condition • Leukocytosis (WBC count ≥11x10^9/L) • Palpable splenomegaly • LDH level increased to above upper normal limit of institutional reference range • Leukoerythroblastosis
Diagnosis of overt PMF requires meeting all 3 major criteria, and at least 1 minor criterion.

Table 5: 2016 WHO diagnostic criteria for overt-PMF. Notes: * 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 myelopathies

In pre-PMF, BM biopsy shows megakaryocytic proliferation and atypia; megakaryocytes have abnormal maturation with hyperchromatic and irregularly folded nuclei. This phenomenon is combined with the increase of BM cellularity and the granulocytic proliferation (Figure 4). Megakaryocytic proliferation and atypia are equally present in overt-PMF (Figura 5 a) and are typically combined with either reticulin and/or collagen fibrosis (Figura 5 b).

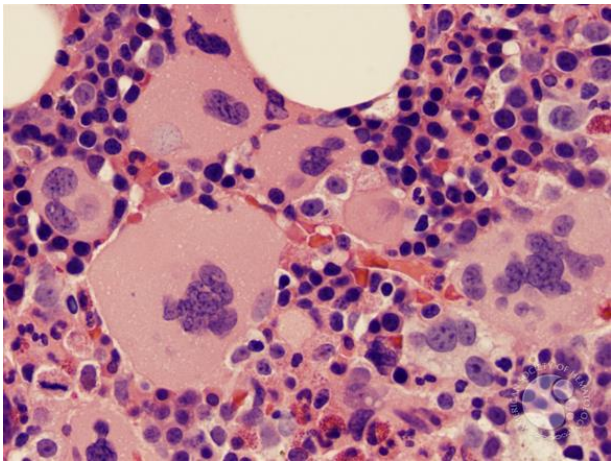


Figure 4: Myelofibrosis BM biopsy. Bone marrow biopsy, hematoxylin and eosin stain, high power show a megakaryocyte clustering. The image represents hyperplastic and dysplastic megakaryocytes in a patient with a diagnosis of PMF. The megakaryocytes are of variable size, show dysplastic nuclear changes and are clustered together. Image taken from American Society of Hematology Image Bank.

(a) *(b)*

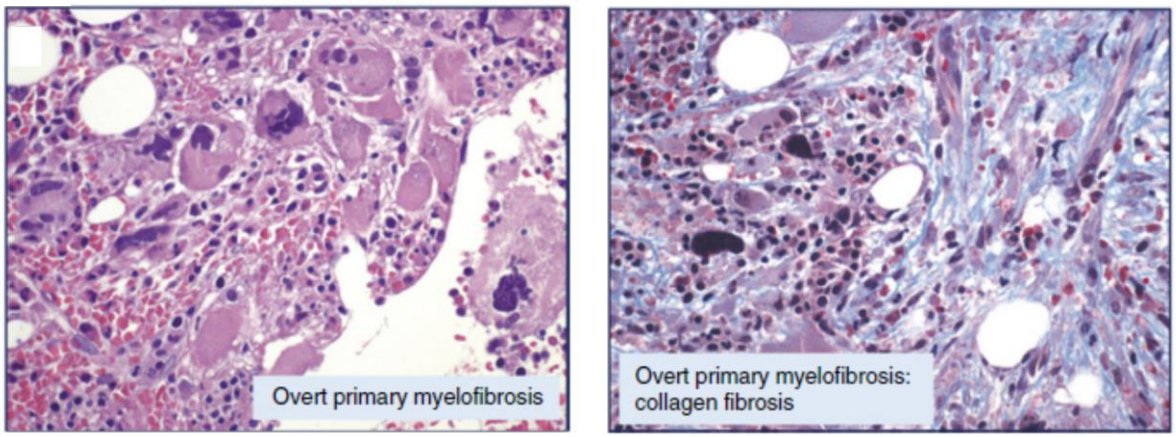


Figure 5: overt-PMF BM biopsies. a. overt-PMF: Hypercellular marrow, proliferation of atypical megakaryocytes forming dense clusters, and dilated vessels with intraluminal hematopoiesis **b.** overt-PMF (collagen fibrosis): Bands of collagen fibrosis within hematopoietic lacunae. Adapted from: Rumi E. et al, Blood 2017

Moreover, it is important to point out the classification of different fibrosis grades, ranking from no crossovers in reticulin, corresponding to normal BM condition, to diffuse and dense increase to reticulin intersections (Table 6) [7,76-78].

Myelofibrosis grading
MF-0 Scattered linear reticulin with no intersections (crossovers) corresponding to normal bone marrow
MF-1 Loose networks of reticulin with many intersections especially in perivascular areas
MF-2 Diffuse and dense increase in reticulin with extensive intersections, occasionally with focal bundles of thick fibers mostly consistent with collagen, and/or focal osteosclerosis*.
MF-3 Diffuse and dense increase in reticulin with extensive intersections and coarse bundles of thick fibers consistent with collagen, usually associated with osteosclerosis*.

Table 6: Grading of myelofibrosis. Semiquantitative grading of BM fibrosis (MF) with minor modifications concerning collagen and osteosclerosis. Fiber density should be assessed only in hematopoietic areas. Note: *In grade MF-2 or MF3 an additional trichrome stain is recommended.

Considering the impact on the functions and morphology of BM in PMF they will be described in more detail in a subparagraph dedicated to the "pathobiology".

The PB smear in PMF is different from PV or ET; it may show immature erythroblasts and leukocytes, teardrop-shaped erythrocytes, circulating megakaryocytes and altered platelets. (Figure 6 a). Moreover, observing a PB smear it is possible to identify the disease progression. To this purpose, the presence of 10-19% of immature circulating blasts indicates an accelerated PMF phase (PMF-AP) and the blasts percentage >20% denotes a leukemic evolution (Figure 6 b) [79, 80].

(a) *(b)*

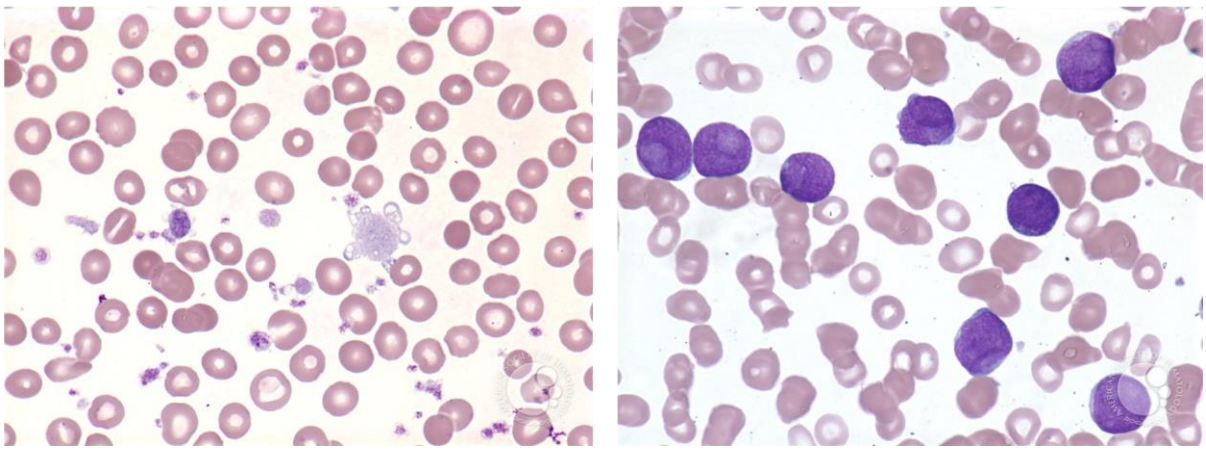


Figure 6: PMF smear and circulating myeloid blast. a. PB smear show large, irregular platelet forms are also seen. **b.** Determining the cell lineage of blasts may be difficult particularly when they lack distinguishing morphologic features such as granules or Auer rods. In this picture, the blasts have characteristic indentations or imprints in the nucleus.

This feature is known as "thumbprinting" and is associated with myeloid lineage. Images are taken from American Society of Hematology Image Bank.

Another sign of disease progression detectable through the PB smear is the increase of immature cells and blasts. The early stage of PMF is characterized by the increase of erythrocytes, platelets and white blood cells [81]. During PMF-AP, leukopenia and thrombocytopenia can be observed or a particular condition, can be found, known as monocytosis. Development of monocytosis in PMF patients is associated with anemia, white blood cells increase, platelets count decrease and low-levels of hemoglobin, and it is considered a high-risk factor related to decrease in overall survival [81-83].

In association with BM biopsy and PB smear observation, the analysis of the mutational profile are crucial for the diagnosis. Three main mutations in MPNs are considered "drivers" mutations, involving *JAK2, CALR* and *MPL* genes. In PMF, *JAK2* (V617F) alteration is detected in 60-65%, *CALR* (exon 9 indels) in 25-30% and *MPL* (exon 10 mutations) in 4-5% of patients. Nevertheless, about 5-10% of PMF patients do not carry any of the "driver" mutations and these cases are TN. TN-PMF patients show a poor clinical outcome [7]. According to WHO criteria, in the absence of at least one of "driver" mutations, the occurrence of PMF in TN-patients could be related to the precence of other common genetic alterations in MPN patients, affecting the Additional Sex Combs Like 1 (*ASXL1*), Enhancer Of Zeste 2 (*EZH2*), Tet Methylcytosine Dioxygenase 2 (*TET2*), Isocitrate Dehydrogenase 1/2 (*IDH1/2*), Splicing Factor 3b Subunit 1 (*SF3B1*) and Serine and Arginine Rich Splicing Factor 2 (*SRSF2*) genes [6]. The mutational profile of MPN patients will be better detailed in a specific chapter of this work thesis.

Beyond the absence of "driver mutations", TN patients have diagnostic features similar to the myelodysplastic syndrome, including BM fibrosis and hypercellularity, multilineage dysplasia, severe cytopenia, causing high transfusion requirement and poor survival [7].

1.1.4.3 Pathogenesis

As previously mentioned, a stem cell-derived clonal myeloproliferation, involving the megakaryocyte lineage, characterizes the PMF pathogenesis. The hyperplasticity and dysplasticity of megakaryocytes is one of the most common features in association with BM fibrosis. Megakaryocytes are the main characters of PMF with inflammation, that involves

mainly monocytes activation and release of inflammatory cytokines, leading to development of the stromal changes in BM, typical of this disease. Chronic inflammation impacts the initiation of the disease, also neutrophilic granulocytes play a major role in the secretion of inflammatory molecules, interacting with thrombocytes and the endothelium, contributing significantly to the pathogenesis. All together these events contribute to the disruption of BM stroma and to the mobilization of HSPCs, leaving the medullary site to colonize secondary hematopoiesis sites, such as spleen and liver, resulting in EMH. EMH is associated with hepatosplenomegaly in PMF. Moreover, BM fibrosis, inflammation and EMH are combined with osteosclerosis and angiogenesis, making the patho-biology of the disease more complex (Figure 7) [76, 77, 84-86].

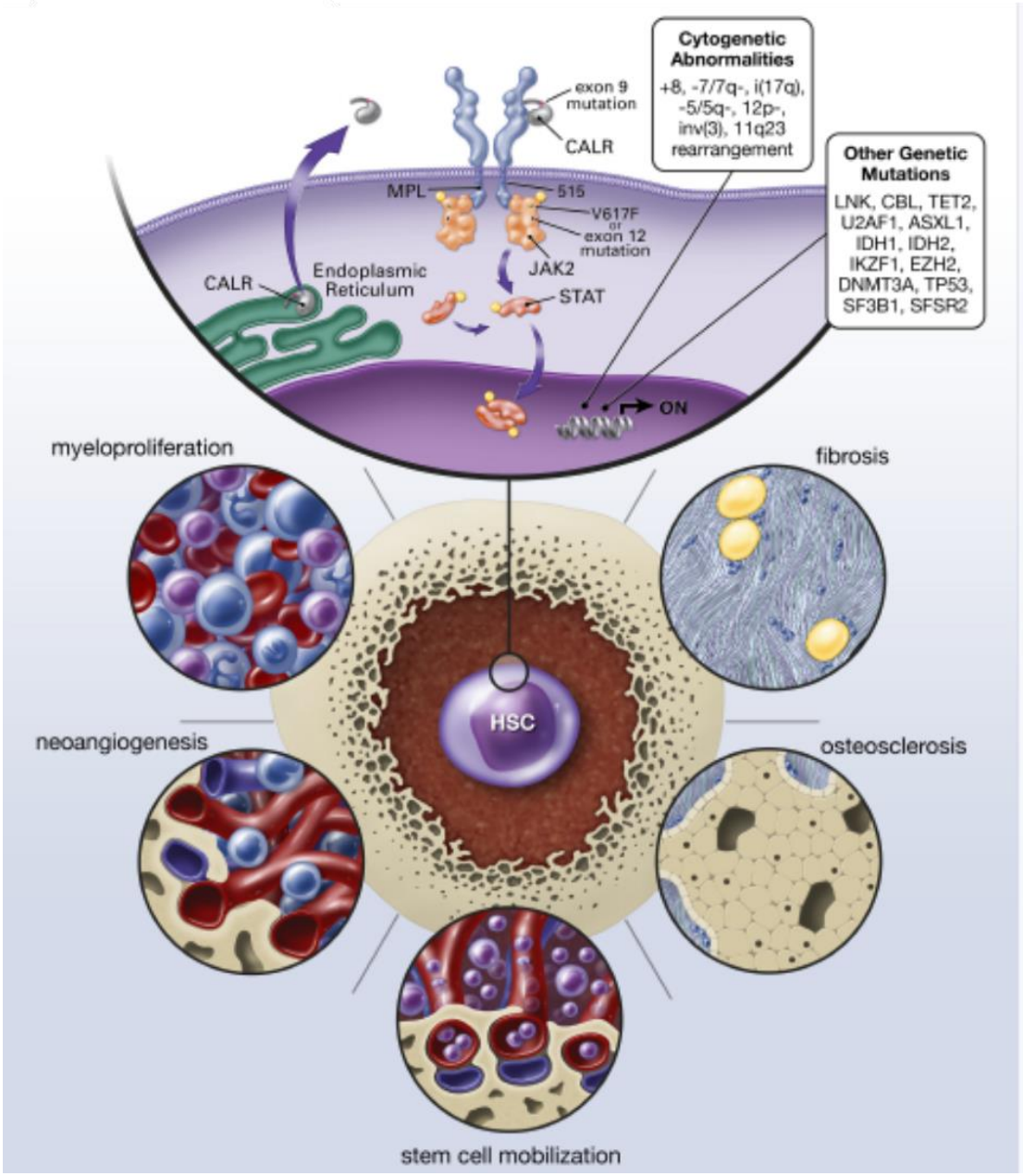


Figure 7: Pathophysiology of Primary Myelofibrosis. The figure highlights the important mutations, cytogenetic abnormalities, and their relationship with the crucial JAK/STAT signaling pathway in MF. Moreover, the biological consequences of HSC mutation are represented. Image is adapted by Benjamin Garmezy et al., Blood Reviews 2021

Considering the crucial role of megakaryocytes in PMF pathogenesis, many researches have focused on the alterations of the megakaryopoiesis in the disease. *In vitro* studies demonstrated the capability of $CD34^+$ HSPCs, isolated from PB of PMF patients, to differentiate in megakaryocytes, generating greater numbers of megakaryocytes than normal $CD34^+$ cells. Moreover, these megakaryocytes developed a delayed response to the apoptosis. The large

megakaryocyte differentiation and the reduction of apoptosis rate could explain the megakaryocytic hyperplasia in PMF [87].

As mentioned, megakaryocytes in PMF show alteration of the morphology, they are smaller than normal, their nuclei are hypolobulated with a bulbous aspect, show clustering features and produce abnormal proplatelets with structural alterations [77, 88]. In addition, it should be noted that megakaryocyte dysplasia is one of the features used to distinguish pre-PMF and ET. In particular, in pre-PMF immature megakaryocytes are located in the trabecular zone of BM, producing atypical megakaryocytes, with altered nuclear-cytoplasmic ratio, bulbous and hyperchromatic “cloudlike” nuclei [89].

Regarding the PMF pathogenesis, the prevailing hypothesis is that BM fibrosis is a secondary process. BM fibrosis, consisting of an excessive production and deposition of matrix fibers in the BM is the hallmark of overt-PMF. The process starts from the production and release of reticulin fibers in the BM stroma and continues with the fibers substitution by collagen I, III, IV, V, and VI. The advanced stage of the disease is characterized by the deposition of lamin and adhesive glycoproteins, such as vitronectin, fibronectin and tenascin. In addition, the transforming growth factor-beta 1 (TGF-β1) increases matrix biosynthesis and reduces the matrix degradation. TGF-β1 also affects the hematopoietic cells in the microenvironment, inducing their differentiation [90].

On the one hand TGF-β1 plays a pivotal role in the BM fibrotic process, on the other it also has a role in megakaryocyte differentiation. Indeed, the “normal” megakaryocytes release TGF-β1 and express its receptors, activating in megakaryocytes extending proplatelets, a type of autocrine stimulation of megakaryocyte development. PMF megakaryocytes express high levels of TGF-β1 but this effect does not seem to enhance the activation of downstream pathways, supporting a role of TGF-β1 in the aberrant megakaryopoiesis but not in proplatelets formation [91].

During the fibrotic process, several cytokines, such as platelet-derived growth factor (PDGF), basic fibroblast growth factor (bFGF) and interleukin-1β (IL-1β), are released by neoplastic megakaryocytes, platelets and monocytes. All of these signalling factors contribute to the disruption of the normal biology of BM. On the basis of this process there are the common PMF mutations in HPSCs. It is demonstrated that JAK2V617F-mutated HPSCs secrete interleukin-1β, activating the apoptosis in the mesenchymal cells and destroying the

interactions between the mesenchymal cells and normal HSPCs. In this way, mutated-cell favor their clonal expansion at the expense of normal cells. After the creation of a favourable tumoural microenvironment, these cytokines contribute to the maintenance of the neoplastic clone, constituting the niches with stromal cells in which malignant cells are deregulated [92, 93].

As mentioned, according to histological BM observation, four different stages of MF are defined as MF-0, MF1, MF-2 and MF-3, with increasing implications for the BM (Table 6). If on the one hand MF-0/1 represents early/prefibrotic myelofibrosis, on the other MF-2/3 displays a larger impairment of BM, increasing collagen fibers and osteosclerosis [7,76-78]. Recent studies highlighted the prognostic significance of BM fibrosis grade in PMF patients. Higher risk of death is associated with fibrosis grade greater than 1. These grades correlate with clinical characteristics that indicate the advanced phase of the disease. Moreover, additional somatic mutations, such as in *ASXL1* and *EZH2* genes, are found in patients with higher grade of fibrosis [94]. Excessive extracellular matrix deposition is closely related to disease progression leading to fibrosis and also osteosclerosis and neoangiogenesis [93]

In PMF, the abnormal bone growth in BM results in osteosclerosis. Osteosclerosis is characterized by increased bone density and abnormal hardening (Figure 8) [95].

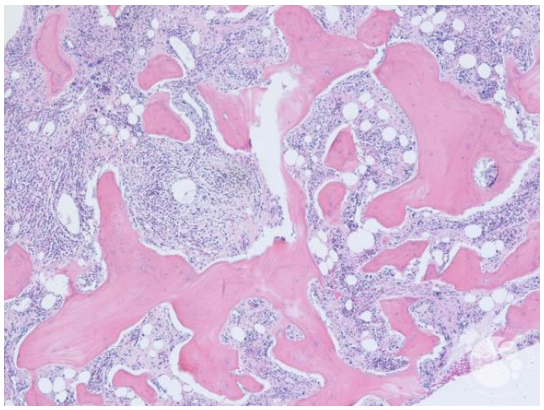


Figure 8: Osteosclerosis. Photomicrograph of a bone marrow biopsy with osteosclerosis. Osteosclerosis is identified as thickened anastomosing bony trabeculae, often seen in conjunction with a myeloproliferative neoplasm or with metastatic carcinoma in the bone marrow. In this case, the marrow is extensively replaced by a myeloproliferative neoplasm with cords of cells growing in a swirling pattern. Image taken from American Society of Hematology Image Bank.

In PMF patients, it was demonstrated that the strong alteration of BM induces dramatic change in bone remodeling, in particular showing a positive balance in bone remodelling, increasing bone volume and decreasing bone erosion [96]. This effect is due to two different but correlated mechanisms, in one hand the proliferation of osteoblasts induced by TGF-β and, in the other one the inhibition of osteoclastogenesis mediated by osteoprotegerin (OPG) [95]. The interaction between the cytokine receptor activator of nuclear factor-kB ligand (RANKL) and its receptor activator of nuclear factor-kB (RANK) is required to activate the osteoclasts formation. In this regard, OPG is considered the most important negative regulator of osteoclastogenesis, given that it blocks RANKL/RANK activation, limiting osteoclastogenesis [97]. The increase of OPG expression by stromal and endothelial cells is demonstrated in PMF patients, in particular higher production of OPG was correlated to high grade of fibrosis [98]. High level of OPG is, however, detected in the serum of PMF patients and OPG expression was observed in megakaryocytes derived from patients, supposing the OPG involvement in PMF also through overexpression by megakaryocyte [99].

The relationship between PMF and the altered bone remodeling not only depends on stromal cells, megakariocytes and therefore the production of OPG. The monocytes seem to play an important role in PMF osteosclerosis. It is known that the monocyte/macrophage lineage generates osteoclasts by the activation of the RANKL/RANK pathway. After the binding of RANKL, monocytes can fuse and give rise to multinucleated osteoclasts, able to cause bone resorption [100]. Interestingly, a recent study found that osteoclasts are generated by neoplastic monocytes in PMF patients. In PMF patients, the osteoclasts showed alteration in their structures, have fewer nuclei and decreased fusion events, compatible with a reduction of the capability to resorption of mineral matrix in bone. These peculiarities in PMF osteoclasts suggest their impairment in the altered bone remodeling and their contribution to the induction of osteosclerosis in PMF [101].

Several other niche-associated pathways are involved in the PMF pathogenesis, including the increase of hypoxia, angiogenesis and inflammation. In particular, hypoxia and angiogenesis are two associated features in the disease; malignant haematopoietic cells highly consume oxygen and secrete pro-angiogenic factors, such as vascular endothelial growth factor A (VEGFA). In this way, neovascularization is induced. It can provide nutrients, oxygen and factors, promoting the survival and proliferation of malignant cells [102].

Worthy of mention, VEGF has been reported as increased in the serum of PMF patients compared to PV and ET, supporting a stronger impact of neoangiogenesis in PMF pathogenesis than in PV and ET [103] .

In recent years, it has emerged that inflammation plays a crucial role in the PMF pathogenesis. The tumor microenvironment is progressively transformed during disease progression. In the early stage, immune cells infiltrate the tumor microenvironment likely to kill cancer cells. During disease progression the immune system shows the inability to inhibit the abnormal growth of cancer cells and induces a cascade of inflammatory signals that play a role in the maintenance of the disease [104]. In view of the close link between cancer and inflammation, talking about PMF, an interesting question is: does PMF trigger inflammation or does inflammation drive PMF development? If in one hand malignant cells can trigger inflammatory responses, in the other chronic inflammation itself may lead to an MPN [105-106].

In PMF, several studies have demonstrated the dysregulation in inflammatory cytokines production, showing the increase in the plasma of molecules such as IL-8, C-C Motif Chemokine Ligand 2 (CCL2), CCL3, CCL4, IL-1RA and tumor necrosis factor (TNF) [107]. The C-reactive protein (CRP) is another inflammation marker and it has been demonstrated that blood levels of CRP are significantly higher in PMF patients, in particular, higher CRP levels correlate with JAK2V617F homozygosity [108]. Treatment with JAK2 inhibitors, such as Ruxolinitib, confirmed the reduction in the plasma levels of inflammatory cytokines [109].

It is more evident that chronic inflammation characterizes PMF and in general MPNs, leading to the increase of circulating cytokines and chemokines levels. In addition, the accumulation of reactive oxygen species increases genetic instability. To this purpose, MPNs are considered a unique model to describe the relationship between hematologic malignancy and chronic inflammation. The occurrence of one of "driver" mutations leads to a progressive inflammation and this hypothesis is supported by the observation before allogeneic hematopoietic stem cell transplantation (alloHSCT). After alloHSCT is performed, chronic inflammation decreases in response to the eradication of the neoplastic clone and the replacement with normal cells. BM fibrosis shows a progressive healing [110].

In the end, EMH is certainly one of the most important hallmarks of the PMF patogenesis. PMF is characterized by the presence of circulating hematopoietic $CD34^+$ cells [81]. A proteolytic microenvironment is formed in BM of PMF patients, resulting in the mobilization of $CD34^+$

cells that move from BM to PB. In this mechanism several matrix metalloproteinase (MMPs), such as MMP-9 play a pivotal role. Increased level of MMP-9 and soluble vascular cellular adhesion molecule (VCAM-1), supporting the activity of neutrophil elastase, are demonstrated in the plasma of PMF patients [111-112].

The stromal cell-Derived Factor-1 (SDF-1, also known as CXC chemokine ligand 12, CXCL12) is the most important chemo-attractant factor for CD34+ cells which express its receptor Chemokine Receptor 4 (CXCR4). In PMF patients the alteration of the CXCL12/CXCR4 axis is reported as a promoter of $CD34^+$ migration. In particular, a decrease in the CXCR4 transcription, related to the hypermethylation of the promoter, is described in PMF $CD34^+$ cells and its effect may result in the mobilization of $CD34^+$ cells [113]. The increase in circulating $CD34^+$ cells is involved in EMH, determining the most common clinical manifestation in PMF: the splenomegaly [62].

1.1.4.4 Clinical manifestations

The most common presenting symptoms of PMF are mainly related to EMH, in particular to its role in causing splenomegaly [62, 114]. Palpable splenomegaly is one of the minor criteria to diagnose PMF established by WHO in 2016 (Table 4 and 5) [6]. Splenomegaly is manifested by abdominal fullness, early satiety and left upper quadrant pain, related often to the development of spontaneous splenic infarction. EMH can affect other organs and body areas including the liver, causing hepatomegaly and portal hypertension; the lymph nodes, resulting in lymphadenopathy; the serosal surfaces, generating ascites or pleural effusions; the epidural spaces, making nerve or cord compression.

PMF patients show common constitutional symptoms such as fatigue, night sweats, fever and weight loss, that, when manifested with severity, require treatments.

Another common clinical manifestation in PMF is the anemia, caused by ineffective erythropoiesis and affecting abot of 50% of PMF patients. The platelets and white blood cells count is, on the contrary, a variable parameter in PMF patients, presenting with the cells count increase or decrease.

The patients have leukopenia (white blood cells < 4×10^9/dl) in 7-8% of cases or important leukocytosis in 9-11% (white blood cells > 30×10^9/dl). Similarly, about 13% of PMF patients show thrombocytosis (platelet count > 500.000/mm^3) and, in contrast, about 37% thrombocytopenia (platelet count < 150.000/mm^3) [114]. In particular, thrombocytopenia is an adverse prognostic factor in PMF patients, and is correlated to several clinical manifestations, such as high rates of anemia and transfusion dependency, high peripheral blast counts and greater risk of leukemic transformation [115].

Causes of death in PMF include leukemic progression, that occurs in 20% of patients, and comorbid conditions, including cardiovascular events, infection or bleeding related to cytopenias [76, 79].

1.1.4.5 Prognosis and risk stratification

PMF patients have the worst prognosis compared to PV and ET, with a median survival of about 6 years [11,39,76].

The main cause of death in PMF is the leukemic transformation (LT). LT can occur in all MPNs with different incidence ranges; PMF shows significantly higher incidence up to 20 % at 10 years, followed by PV with 2.3-14.4% of incidence and 0.7-3% for ET. Different factors are associated with LT, including the advanced age, leukocytosis, exposure to myelosuppressive therapy, cytogenetic abnormalities, as well as increased number of mutations in genes associated with myeloid neoplasms. TN patients, harbouring *ASXL1*, *SRSF2*, *IDH1* or *IDH2* mutations, show increased risk of LT. The prognosis after LT is adverse, long-term survival is rare with a medium overall survival ranging from 2.6–7.0 months [116-117].

To predict LT, the hemoglobin level (<10 g/dL) and the number of circulating blasts (≥1%) were evaluated as universal detectable diagnostic parameters [118].

Risk stratification in PMF is based on prognostic models used in clinical practice to predict the prognosis and to make therapeutic decisions [7].

In 2009 International Working Group for Myeloproliferative Neoplasm Research and Treatment (IWG-MRT) developed the first prognostic scoring system called International Prognostic Scoring System (IPSS). IPSS provide to evaluate the risk stratification in PMF

patients at the time of initial diagnosis basing on five independent predictors of inferior survival: (1) the age >65 years, (2) the hemoglobin level <10 g/dL, (3) the leukocyte count >25 x 109/L, (4) the percentage of circulating blasts ≥1% and (5) the presence of constitutional symptoms. According to the presence of 0, 1, 2 or ≥3 adverse factors, IPSS defines low, intermediate-1, intermediate-2 and high-risk disease categories, that respectively show median survivals of 11.3, 7.9, 4 and 2.3 years [119].

Next, IWG-MRT modified IPSS and introduced a time-dependent prognostic model known as Dynamic International Prognostic Scoring System (DIPSS). This new system maintains the same prognostic variables of IPSS to score the progression of disease during the follow-up [120].

More recently, DIPSS was modified into DIPSS-plus by adding other three DIPSS-independent risk factors: (6) the platelet count <100 x 10^9/L, (7) the need of red cell transfusion and (8) the unfavourable karyotype, meaning a complex karyotype with one or two abnormalities including +8, -7/7q-, i(17q), inv(3), 25/5q-, 12p- or 11q23 chromosomal rearrangement. DIPSS-plus identifies four risk categories, low (no risk factors), intermediate-1 (one risk factor), intermediate-2 (two or 3 risk factors) and high (four or more risk factors), that respectively show median survivals of 15.4, 6.5, 2.9 and 1.3 years [121].

The impact of "driver" mutations on the clinical course, leukemic transformation and survival of patients has, increasingly, emerged in the past few years. It was found that patients with *CALR* mutation had a lower risk of developing anemia, leukocytosis, thrombocytopenia and lower risk of thrombosis. Moreover, *CALR*-mutant patients show better overall survival compared with *JAK2*-mutant or TN patients. TN patients had a higher incidence of leukemic transformation than *CALR*- or *JAK2*-mutant patients [122].

Therefore, *CALR* mutation (in particular a gene alteration known as *CALR* type 1) was considered a favourable prognostic factor in PMF patients [123, 124].

In addition, recent studies demonstrated that *ASXL1, EZH2, SRSF2* and *IDH1/2* mutations represent unfavourable prognostic factors for PMF patients.However, the presence of at least one of these mutations is associated with the worst prognosis with higher risk for premature death and LT. To this purpose, these genetic alterations identify the "high molecular risk" (HMR) category of PMF patients [125, 126].

Considering the important role played by mutations in PMF prognosis and risk stratification, in recent years, scoring systems have been developed assessing the mutational profile of PMF patients. These systems for PMF prognostication and risk stratification evaluation are developed incorporating clinical, cytogenetic and mutation data. In particular, Mutation-Enhanced International Prognostic Score System for transplant-age patients (MIPSS70) is based on molecular information and individuates three risk categories: low-risk in which overall survival is 95% in 5 years, intermediate-risk with 70% and high-risk categories with 29%, corresponding to median survival of 27.7 years, 7.1 years and 2.3 years, respectively. The MIPSS70-plus model, that also considers cytogenetic information, defines four risk categories: low risk with overall survival 91% in 5 years, intermediate-risk with 66%, high-risk with 42% and very high-risk categories with 7% [127, 128]. Ultimately, the genetically inspired prognostic scoring system (GIPSS) is developed exclusively based on genetic markers. This model shows a comparable predictive accuracy compared to MIPSS70-plus and offers a low-complexity referring exclusively on karyotype and on a limited number of mutations, including *ASXL1*, *SRSF2*, *U2AF1*, and *CALR* [129].

1.1.4.6 Therapeutic strategies

The clinical application of allogeneic stem cell transplantation (alloHSCT) originated in the 1950s and early 1960s and was used for patients with advanced acute leukemia in the late 1960s and early 1970s with clinical success. AlloHSCT is indicated as standard treatment for several hematologic malignancies including chronic myeloid leukemia, acute myeloid leukemia and MPNs [130].

To date, alloHSCT is the only treatment approach capable of prolonging survival and care potentially PMF, despite many alloHSCT result in substantial comorbidity and high transplant-related mortality [76].

The alloHSCT is generally considered in patients with DIPSS-plus intermediate-2 and high-risk aged <70 years. Recent studies demonstrated the importance to optimize patient-related factors that can impact on outcome. To this purpose, patients with high risk disease or those at high risk of LT should be considered for early transplantation. In contrast, intermediate-2 risk patients can address alloHSCT at disease progression. [131].

Despite transplantation being the best therapeutic option, initially intensive chemotherapy is required to reduce the disease burden in patients [117]. Recently, it has emerged the possibility to use ruxolitinib as pre-transplant treatment, optimizing myeloablative regimens to decrease toxicity. Patients pre-treated with Ruxolinitib with positive response to therapy show favourable outcomes after alloHSCT [132].

However, in most patients alloHSCT fail due to advanced age, significant comorbidities and poor performance status, consequently, fewer than 10% of patients undergo a transplant. Patients ineligible for transplantation are treated with supportive therapy and non-intensive chemotherapy. The clinical decision to treat with intensive therapy is now limited only to patients eligible for alloHCST, considering the benefits [117].

The most common symptom of PMF is the splenomegaly caused by EMH. Splenomegaly results in several disorders and symptoms that have a negative impact on health and life quality of PMF patients [62, 114]. Considering that symptomatic splenomegaly is associated with constitutional symptoms and disorders such as anemia and cytopenias, the treatment is difficult and can afford to worsen clinical conditions of patients. Cytoreductive treatment, usually HU, is the first-line therapy, being effective in around 40% of the patients [133]. HU is an effective and generally well-tolerated therapy. Nevertheless, HU treatment accentuates anemia and for this reason, concomitant treatment is needed. Patients with anemia, besides transfusion therapy, take drugs as androgens, erythropoietin or erythropoietin stimulating agents of immunomodulatory drugs such as thalidomide, lenalidomide or pomalidomide. However, a substantial proportion of patients do not respond to the therapy [134].

Patients refractory to treatment can be treated with the splenectomy. The procedure involves substantial morbidity as well as a certain mortality risk. Patients not eligible for splenectomy can be treated with local radiotherapy which leads to a transient relief of the symptoms but may induce severe and long-lasting cytopenias [133].

In 2005, the discovery of *JAK2* mutation in MPNs [35] provided a great contribution to the understanding of pathogenesis and the clinical manifestations in PMF. Considering the activanting role of JAK2V617F on JAK/STAT pathway, inhibitor of kinase activity of JAK1/2 were developed and in 2011 Ruxolitinib was approved by Food and Drug Administration (FDA) as the first drug for the treatment of splenomegaly and/or systemic symptoms of patients with intermediate or high-risk PMF [135]. In fact, ruxolitinib treatment often results in spleen

shrinkage and symptom improvement. However, JAK inhibitors do not improve cytopenias, have relatively modest effects on bone marrow fibrosis and driver mutation allele burden and do not modify the risk of LT. For this reason, newer JAK inhibitors are developed to improve anemia (eg, momelotinib) or safety and efficacy in severely thrombocytopenic patients (eg, pacritinib) [136].

Of note, hypomethylating agents, 5-azacytidine (azacytidine) and 5-aza-20deoxycitidine (decitabine), are approved to treat patients who are ineligible for alloHSCT, acting as adjuvant therapy in contrast with the common adverse effects of myelosuppression. Several therapeutic protocols require the combination of JAK2 inhibitor and hypomethylating agents [117].

1.2 Mutational profile in MPNs

The knowledge of the genetic origin of MPNs is very recent.

In 2005, the discovery of gain-of-function mutation affecting the tyrosine kinase JAK2, known as JAK2V617F, [19, 35] was the turning point in the MPNs research. Over the years, subsequent studies have highlighted other important pathogenic mutations affecting *MPL*, encoding the thrombopoietin (TPO) receptor and *CALR* [20, 21]. *JAK2*, *MPL* and *CALR* occur in mutually exclusive manner and are considered "driver" mutations, considering their role in driving the myeloproliferative phenotype [17]. All "driver" mutations impact on the JAK/STAT signalling pathway, as will be described in more detail [137].

Recent studies demonstrate the involvement of other genetic alterations in MPNs pathogenesis involving the regulation of different cellular processes, such as epigenetic regulation, in particular DNA methylation (*TET2*, *DNMT3A*, *IDH1/2*) and histone modifications (*ASXL1*, *EZH2*), or RNA splicing (*SF3B1*, *SRSF2*). These mutations co-occur with "driver" mutations and probably cooperate to disease progression. To this purpose, it is interesting to highlight that in the majority of PV and ET cases, only one mutation is detected. In contrast, in PMF it is much more frequent to find many somatic mutations (three or more) and this may well explain the major complexity and severity of the disease [138].

In the next part of this chapter, "drivers" and "non-drivers" will be distinguished and the main mutations that characterise MPNs will be described in detail.

1.2.1 Driver mutations

The somatic mutations described in MPNs are functionally categorised as "driver" and "other/non-driver" mutations, based on the responsibility to give clonal stem-cell expansion. Three are currently recognized as "driver" mutations in PMF, involving *JAK2* (located on chromosome 9p24), *MPL* (located on chromosome 1p34) and *CALR* (located on chromosome 19p13.2) genes [139].

The frequency of these mutations is typical of each MPNs and specific to each gene. The main common genetic alteration in MPNs regards *JAK2* and it is known as JAKV617F. This genetic alteration is harboured by 95% of PV patients and 50-60% of patients with ET and PMF. Other

mutations involving *JAK2* in exon 12 mutations affect the remaining 3% of PV patients. *CALR* mutations are detected in 20-25% of ET and 25-30% of PMF patients. Finally, 2-3% of patients with ET and 3-5% of patients with PMF carry one activating mutation in *MPL* gene [6, 138, 140]. Except for very few sporadic cases, the "driver" mutations are mutually exclusive [138]. Worthy of note, about 10% of MPN patients do not carry any of *JAK2*, *CALR* or *MPL* genetic lesions and are considered TN [7]. TN patients, however, may harbour atypical *JAK2* or *MPL* mutations [141].

1.2.1.1 JAK2

The year 2005 marks a crucial time for the "history" of MPNs. A *JAK2* somatic mutation, causing a gain-of function and, consequently, the JAK/STAT pathway constitutive activation, was discovered in MPNs. The importance of this mutation was immediately clear, considering that it explained the hyperproliferative nature of the diseases [34, 35, 142]. Further studies have fully confirmed the involvement of this mutation in the pathogenesis of MPNs, to the point that *JAK2* mutations (JAK2V617F considered the "driver" mutation or the mutation of exon 12) are considered a major criteria to diagnose the diseases [6, 7].

The family of Janus kinase (JAK) is a non-receptor tyrosine kinase proteins including JAK1, JAK2, JAK3 and TYK2, that mediate signalling of several cytokines and hormones. The name of these kinases is curious, because they are termed Janus kinases, considering the presence of two apparent kinase domains, which is reminiscent of the two-faced depiction of the Roman God Janus [143].

The *JAK2* gene is located on chromosome 9p24 and encodes a protein composed of 1132 amino acids. Similar to the other proteins of the family, JAK2 structure protein is composed of seven conserved domains, known as JAK homology domains (JH1-JH7). In particular, JH1 and JH2 domains are located in the carboxyl terminus of the protein and have a pivotal role in the function of JAK2. JH1 is the tyrosine kinase domain activated by the binding between cytokine and its receptor. JH2 is a catalytically inactive pseudokinase domain and acts as the autoinhibitory domain. The band 4.1, ezrin, radixin, moesin (FERM) (JH5-JH7) in the N-terminus of the protein is responsible for the interaction with the cytosolic domain of the cytokine receptors. JAK2 contains also an SH2-like domain (JH3-JH4) which does not bind phosphotyrosine [144, 145].

The cytokine binding induces the receptor rearrangement, allowing the dimerization and the activation of the two receptor-associated JAKs through trans-phosphorylation of the tyrosine kinase domain. The receptor activation results in the stimulation of kinase activity and in the phosphorylation of other substrates including signal transducers and activators of transcription (STAT) proteins. In this way the signalling is finally triggered; the STATs phosphorylation induces their homo- or heterodimerization, enabling the migration to the nucleus where they can regulate transcription of their target genes, mainly involved in control of proliferation and survival [138,144].

The JAK2 signalling pathway primarily involves cytokine receptors such as EPO-R, TPO receptor (TPO-R), and granulocyte-colony stimulating factor receptor (GCSF-R). TPO-R can also activate other JAK family members, such as TYK2, however, with low signalling output. In particular, EPO-R activates mainly STAT5 and GCSF-R strongly activates STAT3. In addition, JAK/STAT activation regulates other signalling pathways such as mitogen-activated protein kinase (MAPK), phosphatidylinositol-3'-kinase (PI3K) and AKT/mammalian target of rapamycin (mTOR) (Figure 9) [146, 147].

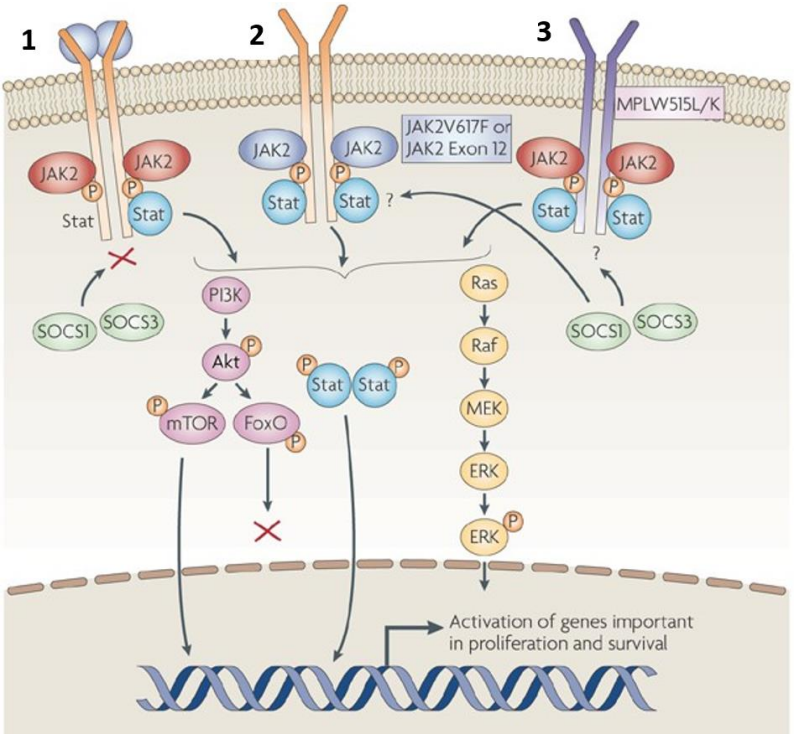


Figure 9: Mechanism of activation JAK2 signalling pathway. 1. Cytokine ligands normally bind cytokine receptors, which results in Janus kinase 2 (JAK2) phosphorylation, recruitment of signal transducer and activator of transcription (Stat) signalling proteins and phosphorylation and activation of downstream signalling pathways including Stat transcription factors, mitogen activated protein kinase (MAPK) signalling proteins, and

the phosphotidylinositol 3-kinase (PI3K)–Akt pathway. **2.** The JAK2V617F and JAK2 exon 12 mutant kinases bind cytokine receptors and are phosphorylated in the absence of ligand, and lead to ligand-independent activation of downstream signalling pathways. **3.** By contrast, MPLW515L/K mutant thrombopoietin receptors are able to phosphorylate wild-type JAK2 in the absence of thrombopoietin, and result in the activation of signalling pathways downstream of JAK2. Negative regulation of JAK2 signalling is normally mediated by suppressor of cytokine signalling (Socs) proteins, most notably SOCS1 and SOCS3; recent data indicate that the *JAK2V617F* allele might escape negative feedback by SOCS3. Adapted from: Levine, R. *et al. Nat Rev Cancer* 2007.

The main genetic alteration affecting *JAK2* in MPNs is called JAK2V617F. It is a somatic mutation occurring at nucleotide 1849 of the exon 14 and it consists in the substitution of guanine with a thymine in codon 617. The G>T substitution results in an amino acid switch from valine (V) to phenylalanine (F). The JAK2V617F mutation affects the JH2 autoinhibitory domain of the protein. In particular, the aminoacid change induces a conformational change which interferes with its negative regulatory action and determines the loss of JH1 inhibition. In this regard, the mutation plays a gain of function leading to the constitutive activation of the kinase and ultimately the JAK/STAT pathway (Figure 10) [19, 34, 35, 142, 148].

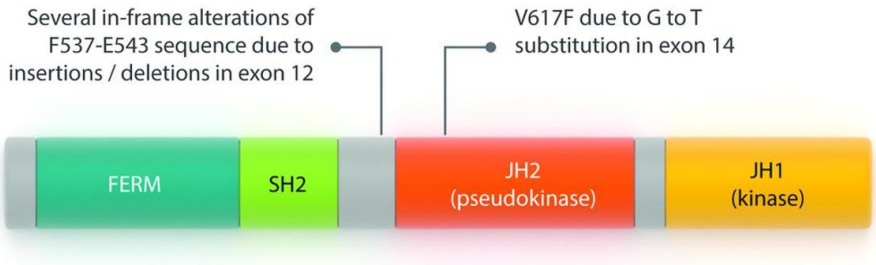


Figure 10: JAK2 structure and mutation localization. In figure it is shown a schematic representation of JAK2 structure highlighting the most important protein domains. Figure also displays the position of the most frequent mutations affecting JAK2 in MPNs. Adapted from: Grinfeld J. et al., Haematologica 2017

Particularly interesting is the epigenetic impact of JAK2V617F mutation, given that JAK2 functions also as an epigenetic modifier. JAK2 may locates in the nucleus of HSPCs and phosphorylates tyrosine 41 on histone protein H3 (H3Y41), inhibiting the recruitment of the repressor Heterochromatin Protein 1α (HP1-α) and leading to the expression of several oncogenes [149-151].

The knowledge of the pathogenetic role of JAK2V617F mutation comes from the *in vivo* experiments with mouse models.

In 2006, the first mouse models were developed by three different groups, using lethally irradiated mice transplanted with BM cells infected *ex vivo* with retroviral vectors expressing JAK2V617F. These studies demonstrated that mice developed a myeloproliferative disorder similar to human PV and which can finally progress to MF [152-154].

In vivo studies also provide information on the allele burden impact on MPNs development. The generation of heterozygous, hemizygous and homozygous JAK2V617F mice, by using conditional *JAK2* knock-out and JAK2V617F knock-in mice, demonstrated that heterozygous JAK2V617F expression results in a PV-like, hemizygous or homozygous JAK2V617F results in accelerated myelofibrosis in mice. The loss of *JAK2* wild-type allele is related to the severity of the disorders [155]. The correlation between JAK2V617F mutation allele burden and clinical phenotype is confirmed by the observation that the JAK2V617F allele burden is higher in PV and PMF, and in patients developing secondary MF [156].

The JAK2V617F mutation is not the only genetic alteration that may affect *JAK2* in MPNs pathogenesis. A gain-of-function mutation involving *JAK2* exon 12 was identified in V617F-negative patients [157]. In particular, several genetic alterations may occur in *JAK2* exon 12 and affect a region adjacent to the pseudokinase domain, spanning from residues 536 to 547. The most frequent mutations are N542-E543 and E543-D544 deletions which affect the linker between SH2 and JH2 domains. These genetic alterations result in the constitutive activation of the kinase [158].

1.2.1.2 MPL

An important new step in MPNs studies was the discovery of the participation of *MPL* gene, located on chromosome 1p34 [139], to disease pathogenesis [20]. The encoded protein is the thrombopoietin receptor, a member of the type 1 homodimeric cytokine receptor family including EPOR and GCSF-R. TPO-R is unique compared to the other members of the family because it has two distal extracellular cytokine receptor homology domain (CRHD) to bond TPO [159].

Since its discovery in the late 1950s, TPO was described as a humoral factor induced in response to thrombocytopenia and capable of increasing circulating platelet count. TPO/TPO-R binding actually plays a crucial role not only in the regulation of megakaryocyte

differentiation and platelet homeostasis, but also in HSC self-renewal and proliferation [160]. *MPL* therefore acts in two processes indispensable during the hematopoiesis, firstly, take action on HSCs and, secondly on the megakaryocyte/platelet lineage [161]. It is therefore simple to hypothesise that a mutation of *MPL* can cause an MPNs.

The binding domain CRHD is a hotspot for mutation impaired in the MPNs [159]. Among MPNs, *MPL* mutations are restricted to ET patients (2-3%) and PMF patients (3-5%) and are quite rare [6, 138, 140].

Most of the genetic alterations detected in *MPL* affect the exon 10. In particular, the most frequent alteration results in the substitution of tryptophan 515 (W515) in mature protein. The W515 switch depends on the substitution of the nucleotide 1544 in codon 515, frequently a G to T substitution, leading to the replacement of tryptophan with another amino acido, that is the leucine (MPLW515L). MPLW515L is the driver mutation and it is responsible for the constitutive activation of the receptor. In mouse models it was demonstrated that it promotes the ET-like myeloproliferative phenotype characterised by thrombocytosis which quickly evolved to MF. In addition, other nucleotide substitutions lead to the amino acid change, generally, in lysine (MPLW515K), asparagine (W515N), arginine (W515R), alanine (W515A) or glycine (W515G) [20, 162, 163]. The W515 is located between the cytosolic and the transmembrane domains of the receptor in a 5 amino acids sequence which prevents the spontaneous activation of the receptor (Figure 11) [20]. The W515 substitution generates the spontaneous activation of receptor in the absence of TPO [138].

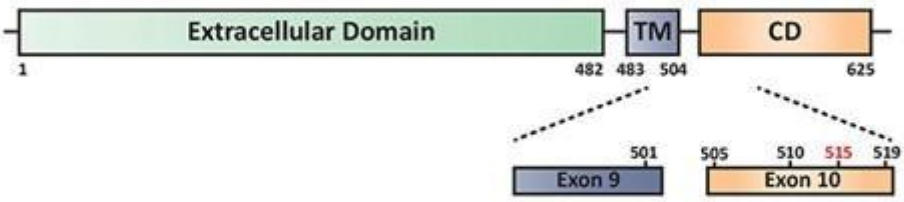


Figure 11: MPL structure and mutations localization. Structural domains and mutations of the MPL. Extracellular domain indicates NH2 amino terminal region, TM transmembrane domain and CD cytoplasmic domain. Dashed lines highlight exons reported mutations of MPL. W515L or W515K are indicated in red while the others hot spots variant mutations in black. Numbers represent amino acid positions. Adapted from: Palumbo G. A. et al., Front. Oncol. 2019

Another more rare substitution, namely S505N, located in the transmembrane domain, induces the constitutive activation of the receptor by stabilising its dimeric orientation. This mutation was first discovered in hereditary thrombocytosis and it is also detected in <1% of ET cases [163, 164].

MPL, *CALR* or *JAK2* mutations can impact on MPL cell-surface expression, increasing MPL turnover and post-translational processing. In particular, MPL is produced as an incompletely-glycosylated protein, which is fully glycosylated in the Golgi with JAK2 as its obligate chaperone. In this context, if JAK2 is responsible for enhancing MPL stability and recycling, the JAK2V617F mutation increases MPL ubiquitination and proteasomal degradation [159]. Treatments with drugs that inhibit JAK2 demonstrate this close relationship between MPL and mutated JAK2. JAK2 inhibition restores cell-surface levels of MPL. In addition, JAK2 mutation also affects the activation of MPL and for this reason the MPL/JAK2 axis plays a crucial role in the development of MPNs [161].

TPO drives MPL activation via receptor dimerization or by inducing conformational changes in a pre-formed receptor dimer. In physiological conditions, in absence of TPO, MPL and JAK2 are monomeric and inactive [160].

JAK2V617F exhibits the ability to require, bind and activate MPL. MPL is pre-associated with JAK2. When the MPL/JAK2 complex becomes mature in the Golgi, they are expressed at the cell surface as a partially pre-dimerized receptor. JAK2V617F activates MPL TPO-independent signalling. It is interesting to underline that also *CALR* mutants activate the MPL/JAK2 axis. *CALR* mutants acquire the ability to stably bind, dimerize, stabilise and transport out of the ER dimers of MPL. The binding of *CALR* mutants to MPL leads to the close apposition of MPL and JAK2, inducing JAK2 activation [161].

1.2.1.3 CALR

Since 2013, CALR has become a keystone in MPNs study, when its involvement in the pathogenesis was discovered [21].

Calreticulin is a Ca^{2+}-binding protein of 417 amino acids and is localised in the endoplasmic reticulum (ER) lumen. As other ER proteins, CALR acts as a chaperone and is mainly implicated in the folding of proteins destined to secretion and insertion in the plasma membrane. CALR consists of three domains: an N-terminal domain lectin-like globular domain, a central a central region proline-rich domain with high-affinity and low-capacity for Ca^{2+} binding and C-terminal domain with low-affinity and high-capacity for Ca^{2+} binding.In particular, the C-terminal region includes the KDEL retention sequence that ensures CALR retrieval from the Golgi apparatus to the ER [165]. The C-terminal domain also modulates the Ca^{2+} storage and homeostasis. The majority of intracellular Ca2+ is stored in the lumen of the ER. CALR participates in several biological processes, including the antigen processing and presentation for the adaptive immune response, cell adhesion/migration, cell proliferation, immunogenic cell death. CALR plays a role in the nucleus, inhibiting the interactions between retinoic acid receptors and its DNA response elements, thereby modulating gene expression (Figure 12) [166].

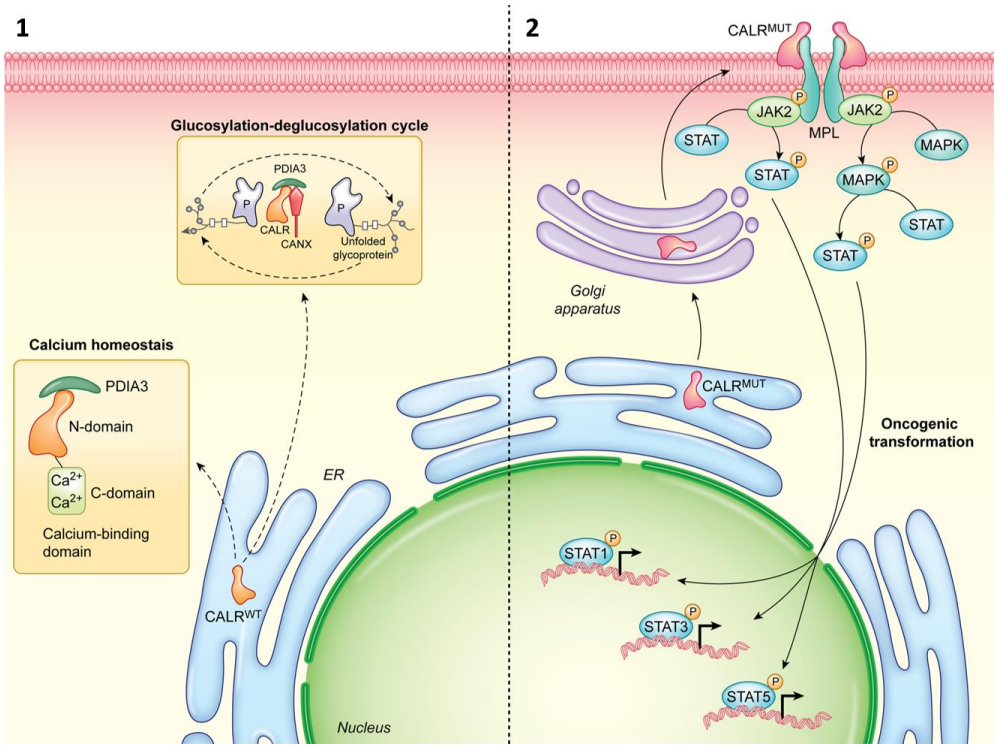


Figure 12: Cellular functions of wild-type and mutant calreticulin. 1. Wild-type calreticulin (CALR-WT) mediates key functions not only as it cooperates with CANX and PDIA3 in the control of protein folding within the ER, but also as it contributes to reticular Ca2+ buffering. **2**. Cancer-associated CALR mutations compromise the capacity of mutant CALR (CALR-MUT) to be retained in the ER, resulting in anterograde CALR transport from the ER to the plasma membrane via the Golgi apparatus, constitutive association with

MPL, and TPO-independent oncogenic signaling (at least in some cells) via JAK2 and MAPKs. P, phosphate. Adapted from: Fucikova, J. et al., Cell Res 2021

CALR is located on chromosome 19p13.2 and its mutations affect the majority of MF (25-30%) and ET (20-25%) patients that are negative for *JAK2* and *MPL* mutations [20, 139, 167]. More than 50 mutations of *CALR* have been described but all of them affect the exon 9 of the gene and induce a +1 (-1 +2) frameshift which generates a new C-terminus of the protein that lacks the KDEL motif (Figure 13)[168].

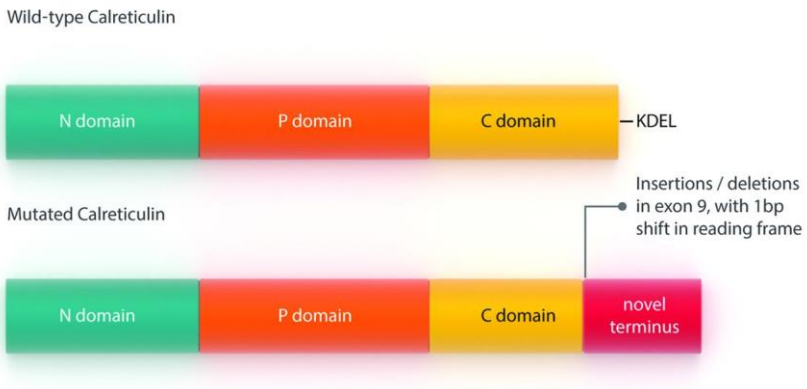


Figure 13: CALR structure and mutations localization. In figure it is shown a schematic representation of wild-type and mutated CALR structure highlighting the most important protein domains. Figure highlight the position of the most frequent mutations affecting CALR in MPNs and their effects on protein structure. Adapted from: Grinfeld J. et al., Haematologica 2017

The most frequent mutations are a deletion of 52-base pair (bp) (del52), named type 1, and an insertion of 5-bp (ins5), called type 2. The type 1 eliminates almost all the negatively charged amino acids in the calcium binding domain, whereas the type 2 eliminates around 50% of these charges. In ET patients type 1 and type 2 mutations have similar frequencies (55% and 35% respectively), in contrast to the PMF cases where type 1 is much more frequent then type 2 (75% vs 15%) [138].

Considering the physiological role of CALR, it is not immediately clear how a *CALR* mutation may induce the MPNs. CALR mutated can form a stable complex with MPL, before its processing in a N-glycosylated mature protein. The CALR/MPL complex is present in Golgi apparatus and can migrate to the cell surface, where MPL becomes able to dimerize in

thrombopoietin-independent condition, thanks to the bond with CALR mutated. The loss of C-terminal KDEL sequence creates new protein interaction and abnormal CALR localization, including into the nucleus [169]. In addition, in Ba/F3 cells it was demonstrated that mutated CALR can induce ligand-independent activation of JAK/STAT pathway via MPL. The pathological interaction between CALR mutated and MPL leads to the dimerization of JAK2 kinase domain and to its activation [170].

CALR mutations can drive the oncogenesis through a mechanism that involves its ability to Ca^{2+} binding. The mutations lead to the substitution of negatively charged amino acids with a stretch of positively charged residues, perturbing the ability of CALR to retain Ca^{2+}ions within the ER lumen. In this way, CALR is responsible for enhanced Ca^{2+}signalling and thus, for cell proliferation [171].

The clinical impact of *CALR* mutations in MPN patients is interesting. *CALR*-mutated ET patients were preferentially male and showed higher platelet count and lower hemoglobin and leukocyte count and have a lower risk of thrombosis, compared with *JAK2*- and *MPL*-mutated patients [172]. In PMF patients, *CARL* mutations are correlated to a lower frequency of spliceosome mutations, to a lower rate of red blood cell transfusion dependency and to an improved overall survival, compared to *JAK2*-mutated patients [173].

However, mutations type1 and type 2 do not show exactly the same behaviour. Type 1 mutation is more common in PMF than type 2 [169]. Patients carrying the type 2 mutation show higher platelet counts, low risk of thrombosis, a reduced overall survival and decreased frequency of fibrosis than type 1. Furthermore, *CALR* deletion is primarily associated with a myelofibrosis phenotype and a significantly higher risk of myelofibrotic transformation in ET [123, 174, 175].

Studies with mouse models demonstrate that in heterozygous and homozygous conditional inducible knock-in (KI) mice, type 1 mutation induces thrombocytosis, leukocytosis, splenomegaly, BM hypocellularity, megakaryocytic lineage amplification, expansion and competitive advantage of the hematopoietic stem cell compartment, compared to type 2 mutation. *CARL* mutated KI mice recapitulate the phenotypic differences detectable in patients [176].

1.2.2 Non driver mutations

The pathological role of the "driver" mutations have demonstrated that JAK2V617F, MPLW515L/K and *CALR* indels are able to activate the JAK/STAT signalling, inducing the myeloproliferative phenotype, that is the hallmark of MPNs. Nevertheless, "driver" mutations are not the only genetic alterations that can affect MPN patients. When considering the mutational profile related to MPNs pathogenesis, it should be recalled that in PMF it is much more frequent to find many somatic mutations (three or more). In contrast, the co-occurence of different somatic mutations is rarely described in ET and PV [7].

Recent studies based on Next Generation Sequencing (NGS) have highlighted other frequently mutated genes involved in different pathways. Some of these are epigenetic regulators that act as DNA (TET2, DNMT3A, IDH1/2) or histones (ASXL1, EZH2) modificators. Also factors involved in the spliceosome complex function and the splicing regulation (SF3B1, SRSF2) can be affected by genetic alterations. Considering the co-occurrence with "driver" mutations, it was studied their impact on disease onset and progression. [138].

Worthy of mention, some of these mutations are considered HMR (*ASXL1*, *EZH2*, *SRSF2*, or *IDH1/2*), given that they represent unfavourable prognostic factors for PMF patients, increasing the risk for premature death and LT [125, 126].

1.2.2.1 TET2

In 2009, mutations in *TET2* were identified in several patients affected by different myeloid malignancies (myelodysplastic syndromes, MPNs and AMLs), candidating, for the first time, *TET2* as a gene involved in these hematological diseases.

The *TET2* gene is located on chromosome 4q24 and encodes for an enzyme that acts as a DNA modulator. In particular, as other TET family members, TET2 is a methylcytosine dioxygenase, able to catalyze the DNA hydroxymethylation by converting 5-methylcytosine (5 mC) to 5-hydroxymethylcytosine (5 hmC). TET dioxygenases require oxygen, α-ketoglutarate and Fe (II). TET2 promotes an alternative pathway of DNA demethylation through oxidation, influencing the gene expression [178].

TET2 is also able to impact the histone regulation via the recruitment of O-linked-N-acetylglucosamine (O-GlcNAc) transferase (OGT) to chromatin. The interaction of TET2 with

OGT enhances the O-GlcNAcylation of Histone 2B at Ser112 and this event is associated with high levels of transcription. Moreover, the complex TET2-OGT colocalizes on chromatin and promotes the tri-methylation of lysine 4 on histone H3 (H3K4me3), an "activating" histone modification resulting (forse) in gene transcription [179].

TET2 loss of heterozygosity (LOH) and somatic mutations are described in 30–50% of myelodysplastic syndrome and MPN patients, whereas 32% of secondary AML. In general, deletions, nonsense and missense mutations were found across multiple exons [180].

In particular, the loss of function mutations affect *TET2* in MPNs with a frequency of 4% in ET, 10-16% in PV and 8-15% in PMF [177].

Several studies demonstrate that *TET2* mutations do not seem to affect survival, the risk of thrombosis and LT and do not impact on the cytogenetic profile of patients. Therefore, their contribution to pathogenesis is still not completely defined [177, 181, 182]. The hypothesis is that the pathogenetic effect of *TET2* mutations might include abnormal DNA hypermethylation [183].

1.2.2.2 DNMT3A

TET2 is not the only factor involved in DNA methylation that can be mutated in MPNs. Genetic lesions in *DNMT3A* are found in patients; they are rare in PV and ET (5%) and are more common in patients with MF (15%) or AML (20%) [184].

Worthy of note is the early appearance of *DNMT3A* mutations in MPNs, also prior to acquisition of JAK2V617F. This observation is in accordance with the low allele burdens of JAK2V617F in ET patients and to the reports of *DNMT3A* mutations in clonal hematopoiesis in normal individuals [185].

DNMT3A is a *de novo* DNA methyltransferase that catalyses the addition of methyl groups into active chromatin in CpG-rich regions, leading to gene inactivation. Mutations of *DNMT3A* consist of loss of DNA binding and reduced catalytic activity [184]. *DNMT3A* mutations were associated with poor outcome and, thus, are of clinical relevance but the exact mechanisms of action of *DNMT3A* mutations are still unclear [186].

Recently it was demonstrated that *DNMT3A* loss results in transformation of PV into lethal MF. It was hypothesised that *DNMT3A* mutations cooperate with JAK2V617F to induce myelofibrosis, probably through the increasing chromatin accessibility at active enhancers,

leading to increased stemness, inflammatory signalling, and consequent myelofibrotic transformation mediated by DNMT3A [184].

1.2.2.3 IDH1/2

IDH1 and *IDH2* mutations are extremely rare in MPNs and are described in about 2%, 1%, 4% respectively of PV, ET and PMF cases. On the contrary, they are more frequent in the blastic phase (BP) of MPNs, about 20%. IDH1/2 are enzymes that convert, for oxidative decarboxylation, the isocitrate in α-ketoglutarate. The mutations induce loss-of-activity in the conversion of isocitrate and gain-of-function in the conversion of α-ketoglutarate to 2-hydroxyglutarate [183].

IDH1/2 mutations lead to elevated levels of 2-hydroxyglutarate that inhibit the function of TET2. *TET2* and *IDH1/2* mutations are mutually exclusive in AML [187].

IDH mutations are correlated to an adverse effect on survival, the presence of an *IDH* mutation can predict worse prognosis. This suggests a pathogenetic contribution to leukemic but not fibrotic disease transformation [188].

1.2.2.4 ASXL1

ASXL proteins are mammalian homologues of additional sex combs (ASX), a regulator of Trithorax and Polycomb complexes in *Drosophila* that is crucial to maintain homeotic gene activation and silencing, regulates *Hox* gene expression and drives random X inactivation. ASXL1 is involved in other biological pathways, including genomic and autosomal locus imprinting, RNA interference, cell cycle regulation, oncogenesis, and stem cell maintenance [189].

Somatic *ASXL1* mutations occur in patients with hematological diseases. The mutational frequencies were 13% in PMF, 23% in post-PV/ET MF, 18% in blast-phase of MPNs. The mutations generally affect exon 12, in particular it was reported the high mutation prevalence c.1934dupG (p.Gly646TrpfsX12) in PMF patients [183].

ASXL1 mutations occur as somatic nonsense mutations and insertion/deletion mutations in a clustered region adjacent to the highly conserved pleckstrin-homology domain, result in loss of ASXL1 protein expression or in expression of a truncated isoform. The ASXL1 loss in hematopoietic cells causes the reduction of polycomb repressive complex 2 (PRC2)-mediated histone H3 lysine 27 (H3K27) tri-methylation. The H3K27me3 through inhibition of PRC2 recruitment leads to the repression of ASXL1 targets, including the posterior HOXA cluster, known as the gene involved in myeloid transformation [190].

ASXL1 mutations have been associated with a poor prognosis since they are associated with an increased risk of AML transformation in MPN patients [138]. A study of 570 patients with PMF confirms the prognostic advantage of harbouring a *CALR* mutation in absence of an *ASXL1* and it shows the poor prognosis in presence of an *ASXL1* mutation. The presence of *CALR* mutations appears to attenuate, but not fully overcome, the unfavourable prognosis in *ASXL1*-mutated patients [191].

It is interesting to note that *ASXL1* and/or *EZH2* double mutations in HSC act as a synergistic interplay. The double *ASXL1/EZH2* mutated neoplastic HSPC led to the highest engraftment levels in mouse model and the engrafted neoplastic clones resulted in the reproduction of the chronic prefibrotic disease parameters, such as anemia, atypical megakaryopoiesis, splenomegaly and even fibrosis development. In this way, *ASXL1/EZH2* seem to have a synergistic effect on the impairing erythropoiesis in PMF [192].

1.2.2.5 EZH2

Gene silencing is one of the cornerstones in epigenetic regulation. The establishment and maintenance of this key process is controlled by two different molecular mechanisms; the heritable repression of gene activity that is the objective of the Polycomb group (PcG) proteins and the DNA methylation systems. The PcG protein EZH2 is a connecting point between these biological systems because it is associated with PRC2/3 and interacts with DNMTs [193].

EZH2 mutations are frequent in 3% of PV and 7% of PMF cases [183].

EZH2 mutations may occur with JAK2V617F and contribute synergistically to the development of MPNs, accentuating the phenotype. Mice double mutant had high platelet and

neutrophil counts, more advanced myelofibrosis and reduced survival. Moreover, these displayed expansion of the HSPCs which shift the differentiation toward megakaryopoiesis at the expense of erythropoiesis [194].

EZH2-mutated PMF patients have higher leukocyte and blast-cell counts, showed larger spleens at diagnosis and most of them are in the high-risk group according to IPSS. Leukemia-free survival and overall survival are significantly reduced in PMF patients that carry an *EZH2* mutation [195].

1.2.2.6 SRSF2

SRSF2 is a member of the serine/arginine-rich (SR) protein family and is involved in splicing process and spliceosome assembly. SRSF2 is able to bind to exonic splicing enhancer (ESE) sequences within pre-mRNA through its RNA recognition motif domain (RRM). The most common mutation is known as SRSF2P95H and its biological significance is currently unclear [196].

It is important to highlight that splicing factors are the most frequently mutated genes in myeloid neoplasms. In particular, *SRSF2* mutations have been found in about 20% of chronic myelomonocytic leukemia (CMML) and in about 40-50% of myelodysplastic syndrome (MDS) patients. *SRSF2* mutations have been associated with a poor prognosis and shortened overall survival [197, 198].

About 17% of PMF patients harbour *SRSF2* monoallelic mutations P95H. PMF *SRSF2* mutation is associated with advanced age and high-risk category per DIPSS-plus, with shortened overall and leukemia-free survival [199].

1.3 Single-Cell Technologies

The scientists who deal with biology know it is a complex matter. The understanding of the pathological processes and in particular cancer is very complex. In this case, knowing the identity of the affected cell population is essential to examine what changes occur during the neoplastic transformation process.

It has been said since the beginning of this work that MPNs are caused by somatic mutations that affect HSCs and HSPCs. It has been stressed the complexity of the mutational profile and consequently, the pathogenesis and the pathobiology, the symptoms and the clinical features, the risk stratification and the disease evolution of MPN patients.

It is highlighted that, although MPNs have many similarities, they are complex and heterogeneous and they have specific characteristics. In addition, the clinicians well know that each patient, affected by one of the MPNs, is unique and each new diagnosis writes a new story. This makes it difficult to choose the best therapeutic strategy and establish an accurate prognosis. In order to simplify, many classification and prognostic scores were born.

Considering all these reflection points, a question arises: how can we understand this complex matter? How can we discern this wide heterogeneity? To deeply understand the MPNs, is it enough to look at what happens to the affected cell population?

As in a study of population, although all of its individuals have many common characteristics and live in the same conditions, it is likely to generalize, losing the information about the single individual. In the same way, investigating the MNPs, is it risky to generalize by studying the HSCs or HSCPs populations? The sc-study allows us to examine the distinguishing marks of the molecular change that cause MPNs pathogenesis, to determine the identity card and to track the history of the cell, which gave rise to the disease.

In recent years, sc-technologies are becoming essential tools for the study of cancer biology. In contrast to bulk analyses that measure the biological processes based on cell populations, sc-studies provide a finer picture, deciphering the heterogeneity in a cell population and identifying the cellular subclones. This approach is sensitive and it is ideal for studying rare clones.

Considering the large possibility to apply sc-analysis, sc-“omics” technologies are developed: sc-genomics, sc-transcriptomics, sc-epigenomics and sc-proteomics. Cancer is characterised by genetic modifications that impact on epigenetics and gene expression. Sc-omics approaches

contribute to understanding the characteristics of cancer, to unmask tumour heterogeneity with higher resolution and accuracy and to identify molecular mechanisms of tumorigenesis, evolution, metastasis and immune responses. In this way, sc-studies can provide important information to improve the accuracy of diagnosis, treatment, and prognosis [200].

Nowadays, the central requirement for precision medicine is the characterization of the biological and cellular events, resulting in intratumoral clonal heterogeneity. Sc-genomics, in particular, unravelled the cellular hierarchy, discovering rare cells that could drive metastatic disease, drug resistance or disease progression [201].

In general, the term sc-genomics can include transcriptomic or genetic studies that are based on RNA-sequencing (-seq) or DNA-seq. Hereafter, the analysis of genetic variability based on sc-DNA-seq will be called sc-genomics.

1.3.1 Single-Cell Genomics: an introduction

Sc-genomic approaches provide a crucial resolution for understanding the interaction between different genetic events that determine the disease initiation, drive the evolution, lead the relapse and metastasis. Although the massive sequencing of the genome has brought up the importance of driver mutations in tumorigenesis, the knowledge about genetic diversification and clonal selection in cancer has been simplified leaking information useful to diagnosis, prognostic risk stratification, targeted therapy and minimal residual disease (MRD) monitoring [201].

The field of single cell genomics has progressed rapidly over the last 10 years. The initial studies were limited to sequencing only a few cells at a time, but today, high-throughput systems are developed, making it possible to sequence thousands of cells in parallel, reducing costs [202].

“Bulk” analysis has limitations in the understanding of clonal heterogeneity. Some of these are related based on a cell population for the reconstruction of clonality, because tumour populations are often complex and contain several cell populations. “Bulk” approach does not consider small subclones that might be important to disease progression. In cancer studies, it may be necessary also to detect 1% of clones and not only 99% . In this way, “bulk” approach has provided mixed signals, amalgamating clones with different genotypes [201-202].

Noteworthy, sc-technologies are not only useful for determining the contributions of genetic heterogeneity in cancer development, but they are employed in several fields. For example, it is possible to use sc-genomics to identify and assemble the genomes of unculturable microorganisms. Sc-genomics studies support cytogenetics analyses in the characterization of mosaic genetic variations [203].

Interesting applications in neurobiology include the study of genomic variability between individual neurons of the brain, especially in germline mutations or chromosomal aberrations that cause neurological diseases [204]. Sc-genomics is also an instrument to highlight the alteration in germ cell biology, highlighting the genome instability and aneuploidy. This technology can be employed in fertility evaluation, in prenatal screening and genetic testing [205].

The success of the sc-method is based on, certainly, an efficient physical isolation of the cells and an high-performance DNA amplification, especially given that it is crucial to acquire sufficient genomic material from a single cell, useful for downstream analysis. The impact on the final data of these two aspects, as will be discussed, must not be underestimated. The introduction of errors and biases in the steps of a sc-protocol method may strongly compromise the results.

In the next paragraphs these topics will be described in detail.

1.3.2 Single-Cell Isolation

The sc-isolation from a sample is the basis, as well as the first step, of a sc-protocol.

The physical properties of the cells, including density, size, electric charges and deformability are some of the principles exploited by sc-isolation techniques. Considering these characteristics, techniques based on density gradient centrifugation, membrane filtration, microchip-based capture platforms that do not require labelling of the cells were developed (Figure 14) .

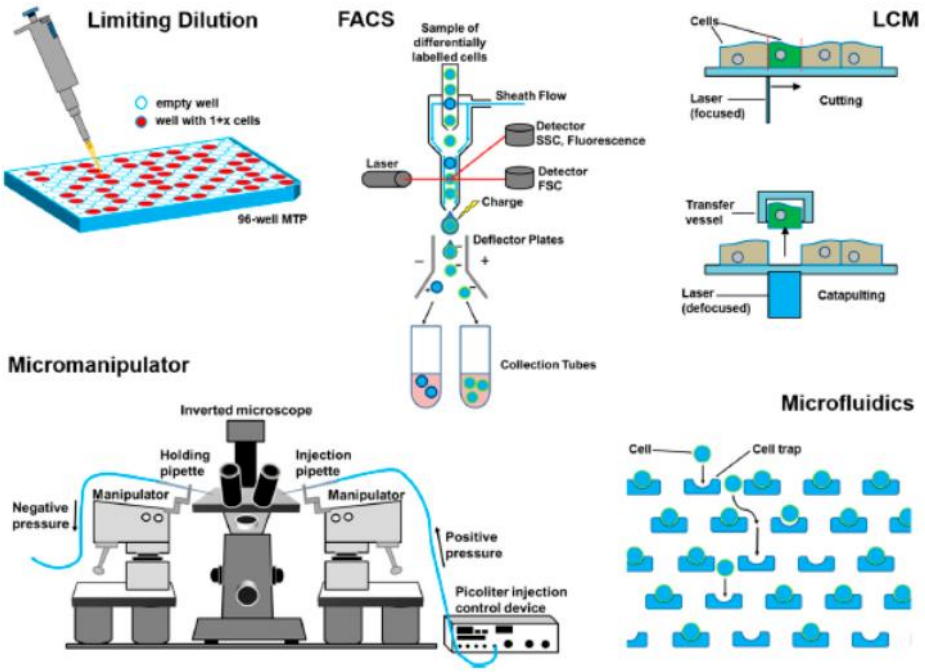


Figure 14: Schematic overview of single-cell separation technologies discussed in the following. Image taken from Gross A.. et al., Int. J. Mol. Sci. 2015

Another subtype of sc-isolation uses the biological characteristics of the cell for the sorting, employing the affinity-based methods, including fluorescence-activated cell sorting (FACS) and magnetic activated cell sorting (MACS) [204].

Beyond the guiding principles for choosing the appropriate method, adopting the specific technique must consider the sample origin. Working with solid tissues requires mechanical or enzymatic dissociation. The use of laser-capture microdissection provides a low-throughput isolating DNA from single cells. Microfluidic and bead-based methods, developed for circulating tumour cells, allow to enrich the sample. Working with samples in suspension, several approaches can be used to isolate single cells, including manual manipulation with serial dilution, micropipetting, microwell dilution and optical tweezers. Some protocols require the use of FACS, ensuring the isolation of intact cells or nuclei. In addition, automated micromanipulation methods use droplets or micromechanical valves in microfluidic devices, accomplishing the obtaining of microscopy data from each chamber or well containing a single cell [203].

Laser-capture microdissection (LCM) is a method to use tissue samples from a microscope slide, working under direct microscopic visualization. The cells can be directly harvested or

isolated by cutting away unwanted cells, giving a pure and enriched histological preparation. LCM does not destroy adjacent tissues after initial microdissection, maintaining the morphology of both the captured cells and the residual tissue. The major limitation of LCM is the identification of cells that must be accurate and require the knowledge of morphological characteristics [204, 206].

The dilution of a cell suspension can provide the isolation of a single cell by using hand-pipettes or pipetting robots. The limiting dilution is well known for decades and was employed mainly for the production of monoclonal cell cultures. This procedure is simple and based on a statistical distribution process. [207].

The development of microfluidics allows the control of fluid range manipulating micro- to pico-litres in networks of channels with dimensions from tens to hundreds of micrometres. This method offers some advantages, especially, the higher throughput, smaller sample volume, automatic sample processing and lower contamination risk. The principles employed are hydrodynamic or electrical manipulation. Cell sorting by using a microfluidic chip is based on cell-affinity chromatography, immunomagnetic beads, on physical characteristics of a cell, on differences between dielectric properties of various cell types in cell suspension. Cell-affinity chromatography–based microfluidic is the method most frequently used. Currently, the microfluidics methods can be combined with different separation principles [204].

Flow cytometry is an ancient technique used since the 1970s. The creation of FACS systems provided the ability to isolate single cells. FACS systems employ laser excitation. The cell stream is rapidly passed by a laser beam to provide optical excitation and then optical detectors are used downstream to capture cell specific signals and the bypassing cells can also be sorted. Popular systems like the FACS-Aria™ III (Becton, Dickinson and Company, Franklin Lakes, NJ, USA) provide up to six different colored excitation lasers and simultaneous fluorescent read-out in up to 18 colour channels and is able to generate up to 100,000 droplets per second and analyse up to 70,000 events per second [207].

1.3.2.1 DEPArray

The DEPArray system from Silicon Biosystems is an automatic sort that can enrich a cell suspension isolating labelled cells by immunofluorescence. The term is derived by the working principle of technology which is dielectrophoresis (DEP) [209].

DEPArray technology is a microchip-based digital sorter, it combines the microfluidic with the microelectronic in order to isolate single cells. The core of this technology is the microsystem cartridge, composed of a microelectronic silicon chip, a microfluidic chamber and valves. The DEP employs an electro-kinetic principle, meaning the ability of a nonuniform electric field to exert forces on neutral or polarizable particles suspended in a liquid (Figure 15) [210].

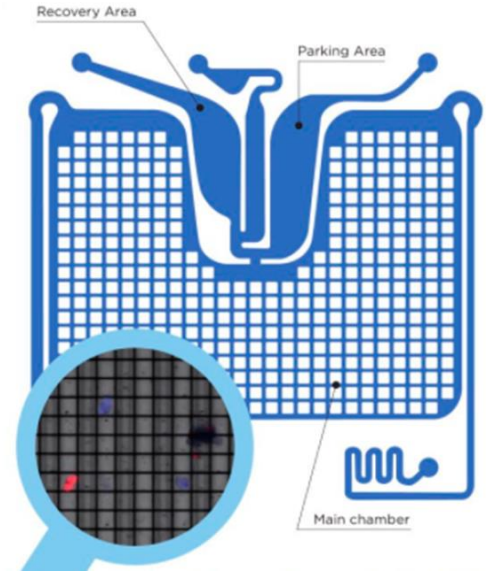


Figure 15: Schematic representation of DEPArray flow cell. The sample is loaded in the main chamber, the selected cells are moved first to the parking area and from there, single or in pools, into the recovery area for delivery into a collection tube Image taken from Menarini silicon-biosystems

The DEP is applied to areas or so called “cages” that contain, hold and move the cell [209].

The array can generate up to 30,000 “DEP cages”. Assuming a random Poisson distribution, when the sample is unlimited, the optimal loading to obtain the maximum number of cells in a cage would be equal to 30,000 cells. In practice, considering a pseudo-random distribution and the use of a non-ideal sample, such as clusters, the best results are obtained with cell numbers as 20,000 cells [210]. The efficiency of the sorting process is very high with a specificity of 100 % [209].

In addition, one of the characteristics of DEPArray is the capability to provide an image-based isolation of a single cell. The DEPArray has a fluorescent microscope integrated, able to acquire high-resolution images for each individual cell in the sample. The detection is possible through the expression or the absence of specific markers correlated to the cell population object of the study. In this way, the fluorescence signals provide an accurate selection based

on the combination of physical and biological parameters, including cell size, shape, circularity and fluorescence intensity [210].

1.3.3 Single-Cell Whole Genome Amplification

In a sc-genomics analysis, the cells must be recovered from a heterogeneous sample, sorted and isolated in individual units, or better to say, single cells. At this point, a genomics experiment requires the genetic material to proceed. The membrane must be lysed and DNA extracted. This step is crucial because an incomplete recovery of sc-DNA, due to DNA damage or chromosomal breaks, can strongly compromise the following steps and the quality of the analysis [204].

The amount of DNA from a single human cell is a few picograms [211]. For this reason, in order to obtain genetic information from single cells, genome amplification is mandatory. During the last 10 years, different technologies of sc- whole genome amplification (WGA) were developed, aiming to minimize the introduction of artefacts, such as amplification bias, genome loss, mutations and chimaeras.

The WGA methods can be categorized in two groups. The first amplifies the entire genomes using the PCR amplification and is based on isothermal methods, using random primers and high processive polymerase. This method owes its success to the PCR efficiency, and, in this regard, it is necessary to highlight that this is closely related to the molecular characteristics of the loci. For this reason, these approaches risk providing a non uniform genome amplification. The second approaches utilize the higher fidelity of polymerase to produce greater genome coverage and lower error rates [203].

1.3.3.1 Linker adapter technique PCR (LA-PCR)

The first PCR-based genome amplification approach was the interspersed repetitive sequence (IRS) PCR. IRS-PCR uses primers, designed to anneal to repetitive sequences as Alu repeats. The uniform distribution of Alu repeats in the genome made this unfavourable, considering that the repeats position was often unsuitable to generate PCR products. For this reason the linker adapter technique PCR (LA-PCR) was developed. LA-PCR was considered an alternative approach that minimizes the limits of the IRS-PCR approach. The method involves

a restricted digestion of the DNA and ligation of adaptors to serve as priming sites for subsequent PCR. This approach was used for the first time for the cloning of microdissected chromosomes [212-213].

1.3.3.2 Degenerate oligonucleotide-primed PCR (DOP-PCR)

The degenerate oligonucleotide-primed PCR (DOP-PCR) was employed for the first time in 1992 for genome mapping studies. The novelty consisted in the use of oligonucleotides of partially degenerate sequence [214].

The degenerate primers contain a random six-base sequence at the 3' end and a fixed sequence at the 5' end. For the initial amplification, a low annealing temperature is used. During the second amplification, the previous products are amplified with a primer targeting the 5' fixed sequence at a higher annealing temperature. DOP-PCR has been used to amplify picogram quantities of genomic DNA, but provides low genome coverage (Figure 16) [211].

Recently, the DOP-PCR became more modern, termed as improved DOP-PCR (iDOP-PCR).A new thermostable DNA polymerase (SD polymerase) was introduced in the protocol, using its strong strand-displacement activity. iDOP-PCR can provide a better quality of the amplified DNA libraries [215].

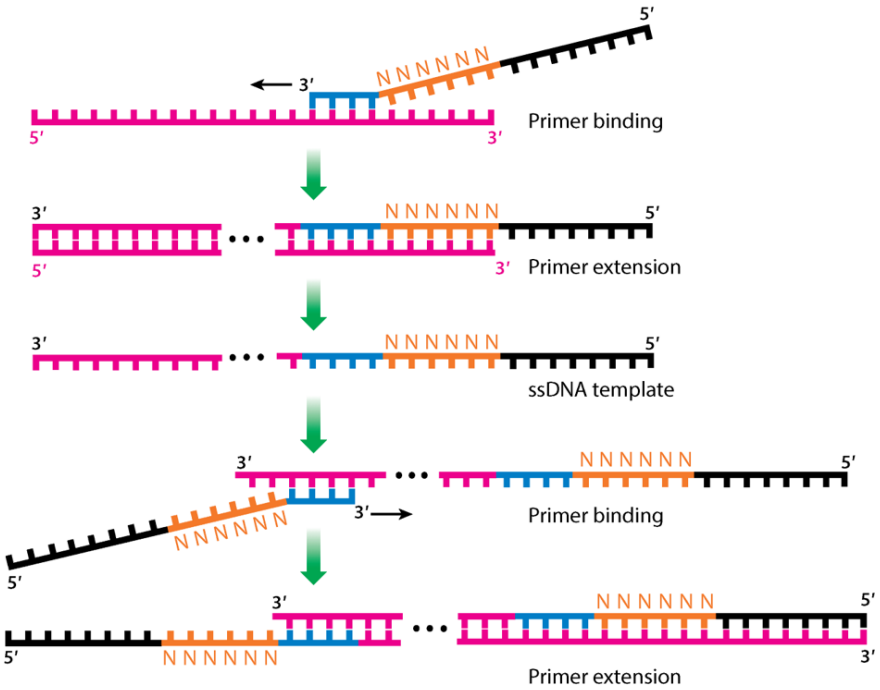


Figure 16: Schematic representation of the degenerate oligonucleotide–primed polymerase chain reaction (DOP-PCR). DOP-PCR uses degenerate oligonucleotide primers for whole-genome amplification. Image taken from Huang L, et al., Annu Rev Genomics Hum Genet. 2015

1.3.3.3 Primer extension preamplification (PEP)

Primer extension preamplification was firstly used to amplify DNA from single sperm nuclei. The original protocol involves a 50 cycle PCR program, using Taq polymerase and a 15 base random oligonucleotide primer. The annealing was performed in a low stringency temperature (37°C), which is then gradually ramped (10 s/°C) up to 55°C prior to a 4 minute elongation step (55°C). Using this approach, approximately 78% of the genome from a single haploid cell can be copied at least 30 times [216].

The improved PEP (I-PEP) PCR approach employed a DNA polymerase cocktail to emphasize the proofreading DNA polymerase activity, resulting in more efficient WGA, with increased fidelity [217].

1.3.3.4 Multiple Displacement Amplification (MDA)

The Multiple Displacement Amplification (MDA) is a WGA method developed in 2001 that introduces a random hexamer as a primer and uses the φ29 DNA polymerase. φ29 is highly processive DNA polymerase, with high fidelity and with strong strand displacement activity [211].

MDA is considered an optimal approach of sc-WGA because it is suitable for efficient amplification of small amounts of DNA. The protocol was developed for preimplantation genetic diagnosis and it can also be useful for the amplification of other minute quantities of DNA, such as from forensic material or microdissected tissue. The product of MDA can be used for multiple downstream applications, such as sequencing, short tandem repeat analysis or array comparative genomic hybridization (Figure 17) [218].

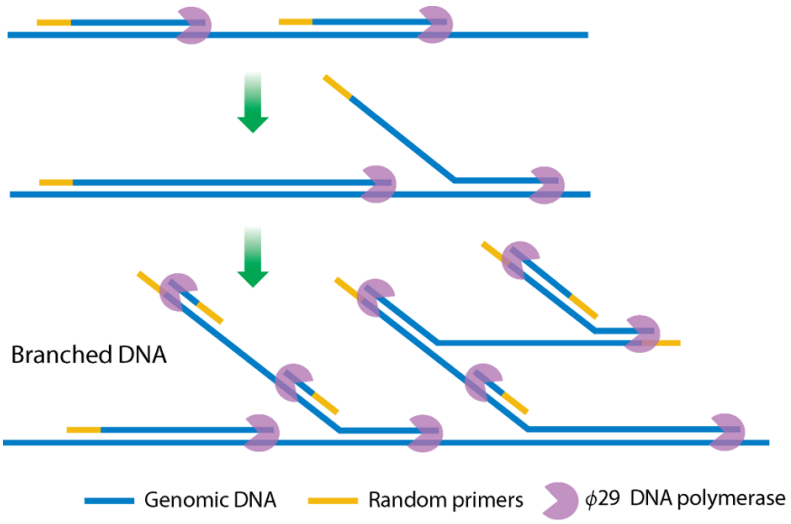


Figure 17: Schematic representation of the multiple displacement amplification (MDA). Random primers are used for isothermal amplification with φ29 DNA polymerase, which has strong strand displacement activity. Image taken from Huang L, et al., Annu Rev Genomics Hum Genet. 2015

MDA works in isothermal conditions, like DOP-PCR, it is an exponential amplification process of random primers exenton [211].

After the annealing of random hexamers to denatured DNA, a strand-displacement synthesis at a constant temperature starts. For the synthesis of DNA, a gradually increasing number of priming events occur, resulting in DNA products of high molecular weight [218].

The main limit of the technique is the overamplification and the possibility to introduce sequence-dependent bias in certain genomic regions [211].

1.3.3.5 Multiple Annealing and Looping Based Amplification Cycles (MALBAC)

In 2012, the multiple annealing looping based amplification cycles (MALBAC) was described as a new sc-WGA protocol with unique characteristics. The novelty was the introduction of a quasi-linear amplification, able to reduce sequence-dependent bias exacerbated by exponential amplification [211].

The MALBAC method introduces multiple cycles of low-temperature annealing to promote efficient priming and quasi-linear preamplification with the looping strategy. During the pre-

amplification, more than 10 copies of amplicons are produced. In this way, the errors dependent by the initial PCR based on a single copy of DNA fragments are significantly reduced [219].

The primers used by MALBAC strategy have a common 27-nucleotide sequence at the 5 end and 8 random nucleotides at the 3 end, which can evenly hybridize to the template when the temperature is lowered (to 15–20 °C). Elevating the temperature (to 70–75 °C), semiamplicons with variable lengths are made. These are melted off from the templates (at 95 °C). At this point, full amplicons with complementary ends are synthesized, but by reducing the temperature (to 58 °C), hairpins are formed, preventing their further amplification. This cycle is repeated 8–12 times [211].

In general, five cycles of preamplification are followed by exponential amplification of the full amplicons by PCR to generate micrograms of DNA (Figure 18) [218].

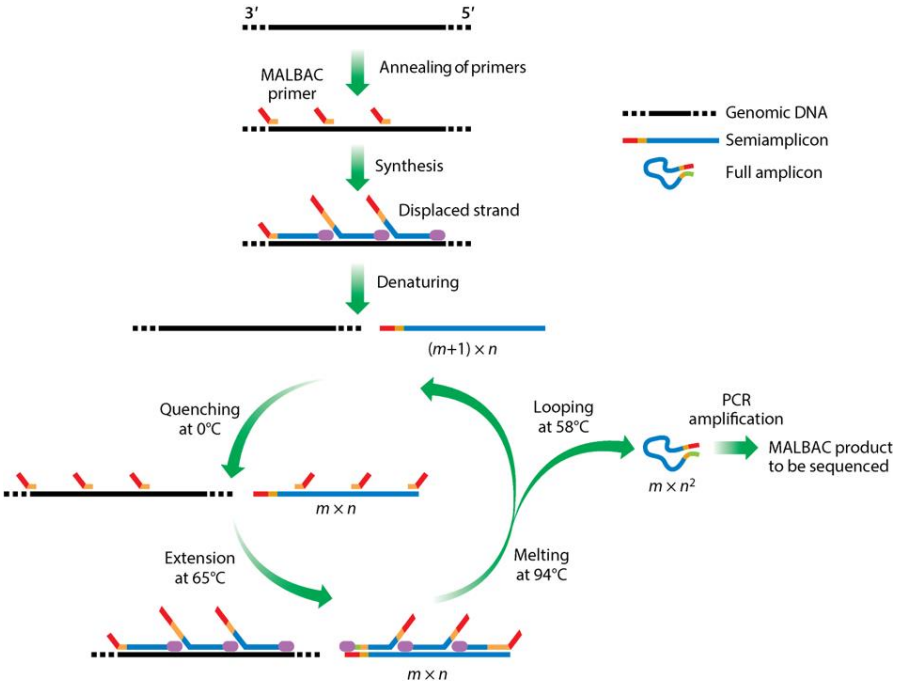


Figure 18: Schematic representation of the multiple annealing and looping–based amplification cycles (MALBAC). Random primers with a fixed sequence are used in a temperature cycle in which only the original genomic DNA and semiamplicons are linearly amplified, and full amplicons are protected from further amplification by the formation of DNA loops owing to the complementarity of the fixed sequences at the 3 and 5 ends. The DNA loops are polymerase chain reaction (PCR) amplified at the final stage.. Here, m is the number of temperature cycles (m = 0 ∼ 10), and n is the number of primers bound; (m + 1) × n is the number of semiamplicons present at the mth cycle, and m × n2 is the number of full amplicons generated in the mth cycle. Image taken from Huang L, et al., Annu Rev Genomics Hum Genet. 2015

In contrast to PCR-based methods and MDA based on exponential amplification, MALBAC uses quasi-linear amplification, providing a higher accuracy for copy number variations (CNVs) detection and a low false negative rate for SNV detection [211].

In a recent study MDA and MALBAC methods were compared and it was demonstrated that MALBAC showed higher uniformity, specificity and reproducibility than MDA [220].

1.3.4 Single-Cell data and errors determination

After the obtaining of sc-DNA, how can we interrogate the genetic material?

The targeting of specific locations of the genomes can be a solution. Target-specific amplification using microfluidic devices can significantly increase coverage. However, as the size of the genome region interrogated increases, also the probability of false variants increases. If the query is the entire genomes of single cells, whole genome sequencing (WGS) can facilitate the answer.

In general, it is necessary to compare the variant alleles detected in the sc-analysis with "bulk" data to exclude bias selection. Also the employment of molecular barcoding can reduce error rate [203].

In addition, although as it has been shown, there are various methods to obtain DNA from a single cell, it is necessary to highlight that the WGA method is not faultless. The genome has regions difficult to solve, like GC-rich regions, that can produce mainly allelic drop outs (ADO), preferential allelic amplifications (PA), chimeric DNA-molecules and nucleotide copy errors [221].

Other errors can be introduced during WGA, such as loss of coverage, decreased coverage uniformity. Some of the quality control to use include visual confirmation of an individually isolated cell, WGA product qualification and/or quantification, genome coverage and ADO. [203].

In addition, the application of automated cell capture, amplification and library preparation systems may increase the affordability of sc-analyses [221].

1.3.5 Single-Cell studies in hematological diseases

HSCs are the founder cells of the blood and of the immune systems. Despite, classically, HSCs and HSPCs have been considered homogeneous populations, incredibly, these cells hide a significant heterogeneity. Beyond the ability of HSCs to differentiate into multipotent progenitors (MPPs), they conserve the potential to generate new HSCs in a process called self-renewal. population. However, recent studies have brought to light molecular and functional heterogeneity in the HSC pool. Scientific works about the transplantation of HSC subpopulations revealed that individual HSCs exhibit a wide range of potentials upon transplantation, differing in lineage contributions and reconstitution kinetics of individual lineages, with two HSCs, rarely showing identical patterns.

The cause of HSCs heterogeneity have different bases; firstly the different localization in BM niches that results in a different exposition to biochemical and biophysical environmental stimuli. The presence of mutations may cause genetic mosaics of HSCs and differential epigenetic modifications provide a strong heterogeneity. In addition, the capability to segregate asymmetrically can contribute to various cell component distribution, leading to different cellular behaviours [222].

Recent technologies have employed sc-methods to resolve cellular heterogeneity, reconstructing the molecular profiles, the lineage differentiation and the regulatory relationships. Sc-analysis helped to rebuild hematopoietic lineage hierarchy and to identify specific markers of hematopoietic subpopulations. Moreover, new scenarios were opened about the different ageing of HSCs, distinguishing “old” and “young” HSCs [223-224].

All these new data have contributed to the deep understanding of hematopoiesis. Therefore, can sc-approaches be used to study the alterations in hematopoiesis and hematological disease?

Considering that a strong heterogeneity was found in normal hematopoiesis, studies of malignant hematopoiesis require high-precision instrument to unravel the intratumoral heterogeneity, the cancer cell hierarchical organization and the subclones distribution in neoplastic population. Hematological diseases are characterized by the acquisition of genetic lesions. More recently, sc-WGS were used to study different hematological diseases [225]. A study at sc-level have revealed the genetic architecture and the clone responsible for cancer maintenance and propagation in childhood acute lymphoblastic leukaemia (ALL), identifying

the *ETV6–RUNX1* gene fusion as the early "driver" of the disease [226]. An integrative approach between sc-DNA and sc-RNA analysis has resolved the genotype-phenotype relationship in chronic lymphocytic leukemia (CLL), uncovering the mutated *LCP1* and *WNK1* as novel CLL drivers [227]. Sc-genotyping has demonstrated that mutations of *FLT3* and *NPM1* occur in at least 9 distinct clonal populations in AML [228]. Recently, through sc-DNA-seq the clonal architecture and mutational profiles of 123 AML patients were described. Data obtained by sc-analysis reveals the co-occurrence of mutations and provides reconstruction of clonal architecture characterized by linear and branching patterns. Integrating the analysis with cell surface protein analysis, the genetic and phenotypic evolution in AML was described [229].

Sc-approaches are also employed to study different characteristics of MPNs. Sc-genetic analysis of MPNs has provided important insights into the genetic evolution and clonal selection of mutant cells during disease development and progression, demonstrating the importance of the order of mutation acquisition during disease pathogenesis. On the other hand, sc-transcriptomic has revealed aberrant megakaryocyte differentiation trajectories in MPN patients [230].

MPNs can be considered an excellent disease model to describe cancer biology, considering that they are heterogeneous disorders, characterized by complex mutational profile and are able to change themselves in other MPNs or evolve to AML, as has been described since the beginning of this work. In this regard, MPNs are the ideal candidate to fully exploit the potential of sc-technologies. In particular, sc-genomics can respond to key questions about the role of "driver" and "no driver" mutation in clonal evolution. The knowledge of the mutation or specific combination of genetic lesions that offer a fitness advantage to mutated clones can be amplified, identifying also the genetic lesions responsible for disease evolution and leukemic transformation. Sc-analysis can provide important information about the clonal distribution and the zygosities of mutant alleles.

For this reason, sc-genomics of MPNs may be an interesting journey that has to be made to investigate MPNs pathogenesis.

2. EXPERIMENTAL DESIGN

MPNs are rare, chronic and clonal disorders caused by somatic mutations that affect HSCs and HSPCs. MPNs share common features, including EMH or the spontaneous transformation to AML and the possibility of each MPN to transform into another [8-10]. However, they are heterogenous diseases with different clinical features and prognosis [11]. PMF is the most severe among MPNs and patients show an incidence of LT between 15 and 20% [7].

Classically, *JAK2*, *MPL* and *CALR* mutations are defined as "driver" mutations [17]. In addition, several other genetic lesions co-occur with "driver" mutations and cooperate to disease progression, making the phenotype complex and worsening patients' prognosis [7]. Many studies confirmed that the mutation acquisition order impacts pathogenesis, influencing disease progression and the response to therapy [231]. In recent years, sc-genomics provide important information for the study of clonal evolution and help to reconstruct the cellular hierarchy and architecture in hematological diseases. Thanks to its accuracy and sensitivity, sc-technologies are the elective methods to unmask tumor heterogeneity [201].

We used a sc-genomic approach to unmask the clonal heterogeneity in PMF and to shed light on the mutations that can drive disease onset and leukemic transformation.

Firstly, we have set up a protocol of HSCs isolation, immunostaining for CD34 and CD38 and fixing, in order to sort $CD34^{+}CD38^{-}$ stem cells by DEPArray (Menarini Silicon Biosystems). We tested paraformaldehyde (PFA) and methanol (MetOH) as fixative agents, assessing their ability to preserve cells immunostaining.

To obtain an uniforme genome amplification with high coverage and low rate of errors and bias, we compared six commercial WGA kits for sc-analysis, based on different WGA strategies. We evaluated the compatibility of the different WGA methods and cell fixation in terms of sequence quality, yield and ADO rate. In this way, we set up an optimized workflow for sc-genomics of the HSCs and HSPCs (Figure 19) [232].

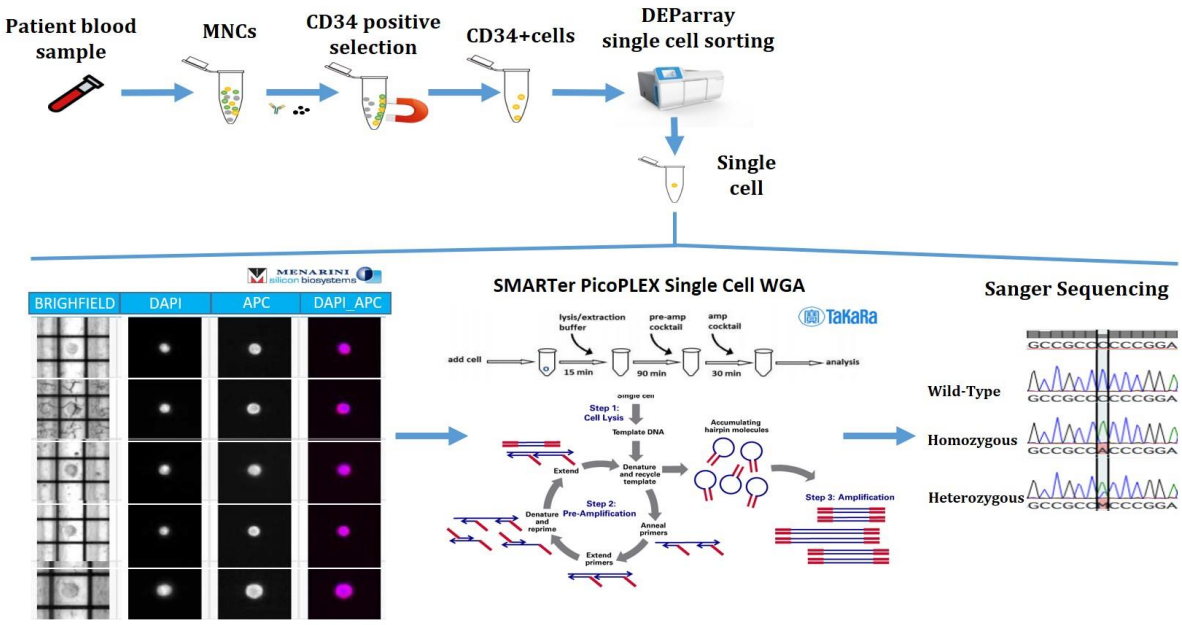


Figure 19: Schematic representation our workflow for cell immunostaining, fixation and WGA

Then, in order to understand the molecular events involved in MPNs progression, we employed sc-genomic approach to characterize the clonal hierarchy and heterogeneity of HSPCs in a PMF patient, analyzed at three different stages of the disease: at diagnosis (T1), during the accelerated phase (T2), and in the AML phase (T3).

We employed single cell genomic analysis to dermine the clonal evolution during disease progression, since the single cell analysis is more sensitive in identifying rare clones that could be undetectable by bulk analysis. (Figure 20). Firstly, we sorted at the single cell level the $CD34^{+}$ cells and performed WGA. We designed sequencing strategies based on the mutations identified by bulk NGS performed at the different timepoints. According to the genomic data, we reconstructed the clonal phylogenetic tree and the mutation acquisition order through the R and R Studio software.

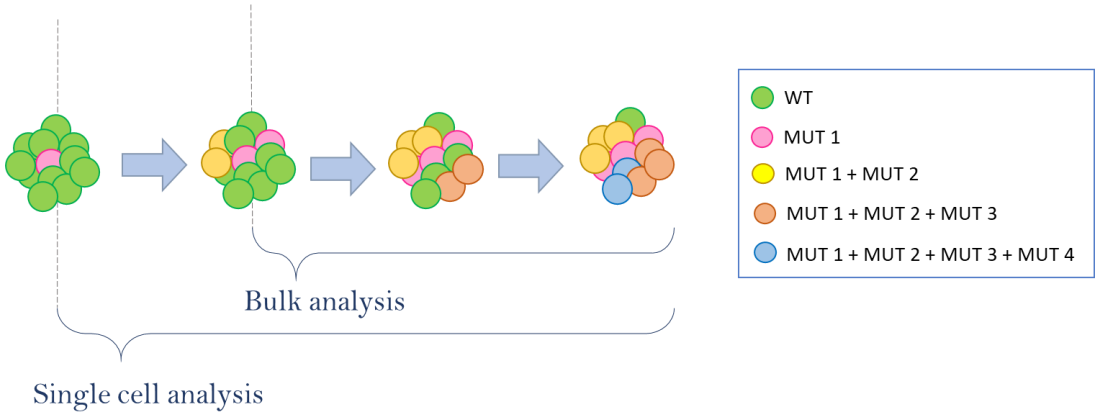


Figure 20: Schematic representation of the comparison between bulk and single cell approach sensitivity. Single cell-technology allows the detection of mutations with low frequency. A rare mutated clone could be undetected by bulk analysis and identified only by sc-analysis in the early stage of the disease of small mutated cell fraction.

Moreover, in order to investigate the presence of potential CNVs in the *TP53* gene, related to disease progression, we performed multiplex ligation-dependent probe amplification (MLPA) in bulk CD34^{+} cells. Since MLPA in bulk did not highlight any CNVs, we sequenced all of *TP53* gene exons at the single cell level, in order to detect genetic rearrangements even in a small cell subpopulation (Figure 21).

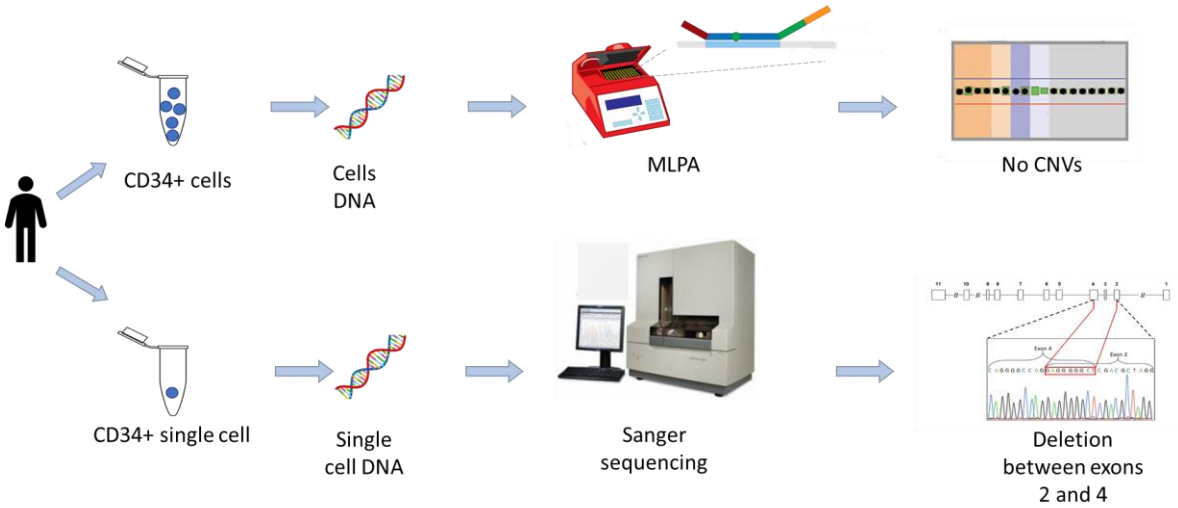


Figure 21: Schematic representation of the employment of single cell technology to identify a small cell fraction harbored CNV. Bulk MLPA on CD34^{+} do not detect any CNVs. In contrast, single cell genomic analysis on CD34^{+} by means of Sanger sequencing is able to identify a small deletion of the *TP53* gene.

Finally, we performed the study of the zygosity of *SRSF2* mutations with a single cell approach. To exclude possible bias introduced by WGA, we performed mutation analysis by means of Sanger Sequencing of DNA from $CD34^+$ cell-derived colonies. In addition, we performed a SNP genotyping assay on both single-cell WGAs and DNA extracted from colonies (Figure 22).

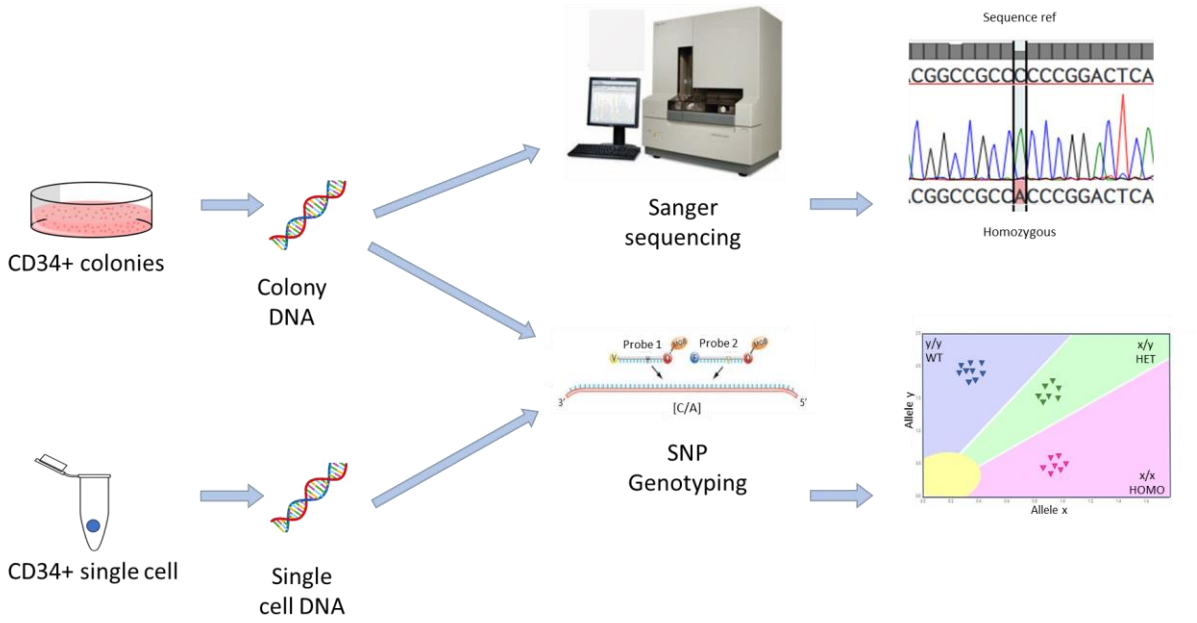


Figure 22: Schematic representation of the characterization of *SRSF2* zygosity mutation. Mutation analysis performed by means of Sanger Sequencing of DNA from $CD34^+$ cells-derived colonies. SNP genotyping assay on both single-cell WGAs and DNA extracted from colonies is also shown.

In this way, we investigated the power of single cell technology in order to identify rare clone early during disease progression, to detect genetic alteration harbored only in small cell population and to characterize the zygosity state of genetic variants uncommon in hematological neoplasms [233].

3. MATERIALS AND METHODS

3.1 Ethics Statement

Human $CD34^+CD38^-$ HSCs were purified upon donor's informed written consent from umbilical Cord Blood samples, collected after normal deliveries, according to the institutional guidelines for discarded material (Clearance of Ethical Committee for Human experimentation of Florence: Comitato Etico Area Vasta dell'Azienda Ospedaliero-Universitaria Careggi, approval date: 22 April 2011, approval file number # 2011/ 0014777).
A PMF patient who evolved to AML was studied at three timepoints: during the chronic phase (Time1, T1), after 8 months of Ruxolitinib treatment (Time2, T2), and after 11 months of Ruxolitinib treatment at AML diagnosis (Time3, T3). The diagnosis was performed in agreement with the World Health Organization (WHO) criteria updated in 2016. The study was conducted in accordance with the Declaration of Helsinki and with ethical approval obtained from the local ethics committee of the Humanitas Research Hospital – Milan, Italy. (Approval date: 28 Jan 2019; approval file # 2175). The patient provided written informed consent to take part in the study.

3.2 $CD34^+CD38^-$ Cell Purification, Immunostaining and Fixation

Cord blood $CD34^+CD38^-$ were purified by means of $CD34^+CD38^-$ Cell Isolation Kit, human (Miltenyi Biotec; Bergisch Gladbach, Germany). Cell purity was assessed by means of flow cytometry after immunostaining with anti-human CD34-APC (1:50) and Labeling Check Reagent-PE (1:33), Labeling Check Reagent-FITC (1:33) or antibodies (Miltenyi Biotec). The labeling check reagent-PE identifies CD38 expression by binding the beads bond to the CD38 antigen. Immunostained cells were then fixed with either Paraformaldehyde (PFA) 2% or 0.5% in PBS1× at 4 ◦C for 15 min, 45 min, 2 h or overnight or MetOH 100% 2 h at 4 ◦C and washed with PBS1× before being resuspended in autoMACS Running Buffer (Miltenyi Biotec).

3.3 $CD34^+CD38^-$ Cell Sorting

$CD34^+CD38^-$ cells were then subjected to single-cell sorting by means of DEPArray™ System (Menarini Silicon Biosystems, Castel Maggiore, BO, Italy) [210]. The sorted population was made up of $CD34^+CD38^-$ cells with a purity of 100%. All cells have been subjected to DAPI staining in order to assess viability. Each single cell was resuspended in 1 µL of 1X PBS.

3.4 Cell Lines Culture and Sorting

In order to assess the ADO rate of each WGA kit, the evaluation of known Single Nucleotide Variants was performed by means of Sanger sequencing on HEL and K562 cell lines, since they carry known genomic variants either in homozygous or heterozygous state. HEL and K562 cell lines were obtained from the ATCC and cultured in RPMI medium (Euroclone S.p.A, Pero, MI, Italy), supplemented with 10% heat-inactivated FBS (Sigma-Aldrich, St. Louis, MO, USA) and 1 mmol/L l-Glutamine (Euroclone). HEL cells (JAK2V617F homozygous mutant) were sorted as single cells in 1 µL of 1×PBS by means of limiting dilution. K562 cells (*ASXL1* c.1773C>G heterozygous mutant) were fixed with PFA 0.5% in PBS 1× at 4 ◦C for 15'.

3.5 Single-Cell Whole Genome Amplification

Whole Genome Amplification was performed on 20 single $CD34^{+}CD38^{-}$, HEL or K562 cells by means of SMARTer PicoPLEX Single Cell WGA kit (Takara Bio Inc, Kusatsu, Japan), REPLI-g Single Cell Kit (Qiagen, Hilden, Germany), Ampli1™ WGA Kit (Menarini Silicon Biosystems), MALBAC Single Cell WGA Kit (Yikon Genomics, Beijing, China) or GenomePlex® Single Cell Whole Genome Amplification Kit (Sigma-Aldrich) following manufacturer's protocol. For each WGA experiment, a genomic DNA sample and 1uL of Low Tris-EDTA (TE) buffer (Thermo Fisher Scientific, Waltham, MA, USA) were included as positive and negative controls, respectively. The WGA product was then purified by means of AMPure XP (Beckman Coulter, Brea, CA, USA) immunomagnetic beads and Int. J. Mol. Sci. 2020, 21, 7366 10 of 11 eluted in 20 µL of Low TE buffer. Quality control on WGA yield and amplicons' size was performed through Bioanalyzer High Sensitivity DNA Analysis (Agilent Technologies, Santa Clara, CA, USA).

3.6 Sanger Sequencing

PCR and sequencing analysis were performed by using BigDye® Direct Cycle Sequencing Kit (Applied Biosystem®, Foster City, CA, USA) following the manufacturer's protocol. Primers used for PCR reactions (Thermo Fisher Scientific and Integrated DNA Technologies, Coralville, IA) were synthesized with the addition of the M13 tail (M13 FW: 50 TGTAAAACGACGGCCAGT 30 ; M13 RV: 50 CAGGAAACAGCTATGACC 30).

Sequencing products were purified by ethanol/EDTA precipitation. Sequencing was performed by capillary electrophoresis on 3130xl Genetic Analyzer (Applied Biosystems®).

3.7 Bulk next generation sequencing (NGS) analysis

Targeted DNA-sequencing on genomic DNA extracted from whole peripheral blood of the patient was performed through a Capture-based target enrichment kit—CE-IVD Myeloid Solution™ by Sophia Genetics. Sequencing was performed on the Illumina MiSeq instrument. The minimum required coverage was set to 1000×. After demultiplexing the FASTQ files were further processed using the Sequence Pilot software version 4.1.1 Build 510 (JSI Medical Systems, Ettenheim, Germany) for alignment and variant calling. Analysis parameters were set according to the manufacturer's default recommendation. The validity of the somatic mutations was checked against the publicly accessible COSMIC v69 database (http://cancer.sanger.ac.uk/cancergenome/projects/cosmic) and functional interpretation was performed using SIFT 1.03 (http://sift.jcvi.org), PolyPhen 2.0 (http://genetics.bwh.harvard.edu/pph2) and MutationTaster 1.0 algorithms (http://www.mutationtaster.org). Additionally, *TP53* variants were verified using the IARC repository. Single nucleotide polymorphisms (SNP) were annotated according to the NCBI dbSNP (http://www.ncbi.nlm. nih.gov/snp; Build 137) and gnomAD (http://gnomad.broadinstitute.org; gnomAD r2.0.1) databases.

3.8 Purification of $CD34^+$ cells

Frozen peripheral blood mononuclear cells (PBMCs) were thawed following 10× Genomics® "Sample Preparation Demonstrated Protocol" (10× Genomics, Pleasanton, CA, USA). Immunomagnetic selection of $CD34^+$ cell population was then performed by means of "CD34 Microbead kit, human" (Miltenyi Biotec, Bergisch Gladbach, Germany) following the protocol provided by the manufacturer. The sample was then split between genomic and transcriptomic analyses. Moreover, a fraction of the $CD34^+$ cell population was seeded in MethoCult™ GF H4434 (StemCell Technologies Inc.; Vancouver). Colonies were picked and genotyped for *SRSF2* p.P95H mutation.

3. 9 $CD34^+$ cell immunostaining, fixation, and single-cell sorting

Cells were stained with anti-human CD34 (AC136, APC, Miltenyi Biotec, Cat No. 130-113-738, 1:50) and anti-human CD38 (IB6, PE, Miltenyi Biotec, Cat No. 130-113-989, 1:50) antibodies. Immunostained cells were then fixed with Paraformaldehyde (PFA) 0.5% in PBS 1× at 4 °C for 15′ and washed with PBS 1× before being resuspended in autoMACS Running Buffer (Miltenyi Biotec). Immunostained $CD34^+$ cells were then subjected to single-cell sorting by means of DEPArray™ Technology (Menarini Silicon Biosystems). Every single cell was resuspended in 1 μl PBS 1×.

3.10 Whole genome amplification

Whole genome amplification (WGA) was performed on 300 single cells for each Time by means of SMARTer PicoPLEX Single Cell WGA kit (Takara)following the manufacturer's protocol as previously described. For each WGA experiment, a genomic DNA sample and 1 μl Low TE (Tris-EDTA) buffer (ThermoFisher Scientific) were included as a positive and negative control, respectively. WGA product was then purified by means of AMPure XP (Beckman Coulter) immunomagnetic beads and eluted in 20 μl Low TE buffer. Quality control on WGA yield and amplicons' size was performed through Bioanalyzer High Sensitivity DNA Analysis (Agilent) on 1 μl of 1:20 diluted WGAs.

3.11 Mutation detection

PCR and sequencing analysis were performed by using BigDye® Direct Cycle Sequencing Kit (Applied Biosystem®) following the manufacturer's protocol. Primers used for PCR reactions were synthesized (Table 7) ThermoFisher Scientific and Integrated DNA Technologies). Sequencing products were purified by ethanol/EDTA precipitation.

Gene	ID	Mutation (genomic coordinate hg38)	F	R
JAK2	Hs00532916_CE	g.5073770	TGAAGCAGCAAGTATGATGAGCAA	CTGACACCTAGCTGTGATCCTG
ASXL1	custom	g.32434646	CGGATCATCCCCACCACGGAGT	GGCGGCAGTAGTTGTGTTCGC
TET2	Hs00732331_CE	g.105243718	ATTGTGATTCTCATCCTGGTGTGG	AATGTAAAAGTGCACGCTGAACTC
TET2	Hs00629106_CE	g.105237351_105237358	ATTGGATACACCTGTCAAGACTCAATATG	CAAGACACAAGCATCGGTAACTTG
SRSF2	Hs00170120_CE	g.76736877	CGGCTGTGGTGTGAGTCCG	GCTGCCCGCCCAGTTGTTA
TP53	custom	g.7674250	AAGCCACAGGTTAAGAGGTCCC	AAAAGGCCTCCCCTGCTTGCC
FLT3	Hs00785316_CE	g.28018505	CCACAGTGAGTGCAGTTGTTTAC	AGAAGAAAGATTGCACTCCAGGATAATAC

Table 7: Primers used for PCR reactions in single cell sequencing

Sequencing was performed by capillary electrophoresis on 3130xl Genetic Analyzer (Applied Biosystems®). *SRSF2* p.P95H variant was also analyzed by means of SNP genotyping through the Custom TaqMan™ SNP Genotyping Assay (rs751713049_g_t,ANAAJWD, SNP, ThermoFisher Scientific) in picked colonies and in single-cell WGAs.

3.12 Clonal hierarchy reconstruction

The clonal tree reconstruction of the mutations' acquisition order was performed through the CellScape R package (script retrieved from https://github.com/shahcompbio/cellscape/blob/511a8b6a2d7c6eb28fddb83f1b29c2c1092d5bae/R/cellscape.R). For the trees representing the presence/absence of the variants, variant allele frequency (VAF) = 0 was assigned to a wild-type status and VAF =1 was assigned to a mutant status in the mutational matrix. Clonal evolution analysis was performed through the TimeScape R package (script retrieved from https://github.com/shahcom bio/cellscape/blob/511a8b6a2d7c6eb28fddb83f1b29c2c1092d5bae/R/cellscape.R). Each clone was assigned to a letter and Timepoints were indicated as T1, T2, and T3. For the trees representing the zygosity of the mutations, VAF = 0 was assigned to a wild-type status, VAF = 0.5 was assigned to heterozygosity and VAF = 1 was assigned to homozygosity. All analyses were performed using R (version 3.6.2) and R Studio (version 1.3.959) software.

3.13 Genomic data analysis

The fishplot representing mutational frequencies was performed through the Fishplot R package (script retrieved from https://github.com/chrisamiller/fishplot). The mutational frequency for each variant was computed as the percentage of mutated cells (either heterozygous or homozygous) within the totality of the cells of each timepoint. Joy Plots of the frequencies of *TP53*, *FLT3*, and *SRSF2* mutations through time were performed by means of ggridges R package (retrieved from https://CRAN.R-project.org/package=ggridges; script retrieved from https://cran.r-project.org/web/packages/ggridges/vignettes/introduction.html). The bubble-plot of the frequencies of mutations through time was built through ggplot2 R package (script retrieved from https://www.r-graphgallery. com/320-the-basis-of-bubble-plot.html). All analyses were performed using R (version 3.6.2) and R Studio (version 1.3.959) software.

3.14 *TP53* multiplex ligation-dependent probe amplification (MLPA)

DNA coming from $CD34^+$ cells of the patient and Mononuclear cells of Healthy Donors was extracted by means of DNeasy Blood & Tissue Kit (Qiagen, Hilden, Germany). 50 ng of DNA were analyzed through P056 MLPA kits (MRC-Holland, Amsterdam, Netherlands), following the manufacturer's instructions. At least five controls were included in each run. Fragment analysis was performed by capillary electrophoresis on the 3130xl Genetic Analyzer device (Applied Biosystems). The software Coffalyser.net (MRC-Holland), was used to analyze all MLPA data.

4. RESULTS

4.1 Cell fixation is necessary for DEPArray single-cell sorting

Cell fixation can represent a limiting step in single-cell sorting. Indeed, it is useful to preserve single cell preparation without altering cellular protein and nucleic acid fractions for downstream experiments.

We employed the DEPArray system that uses an immunophenotype analysis in order to sort the cells at single cell level. For this reason, firstly, we have set up the cell immunostaining protocol suitable to identify $CD34^{+}CD38^{-}$ cells during cell sorting. We isolated $CD34^{+}CD38^{-}$ HSCs from cord blood, separating them from the $CD34^{+}CD38^{+}$ by means of immunomagnetic system, as reported in "materials and methods". We performed the immunostaining to identify the expression of CD34 and CD38 surface markers. We controlled the immunostaining efficiency by using unfixed $CD34^{+}CD38^{+}$ cells (Figure 23).

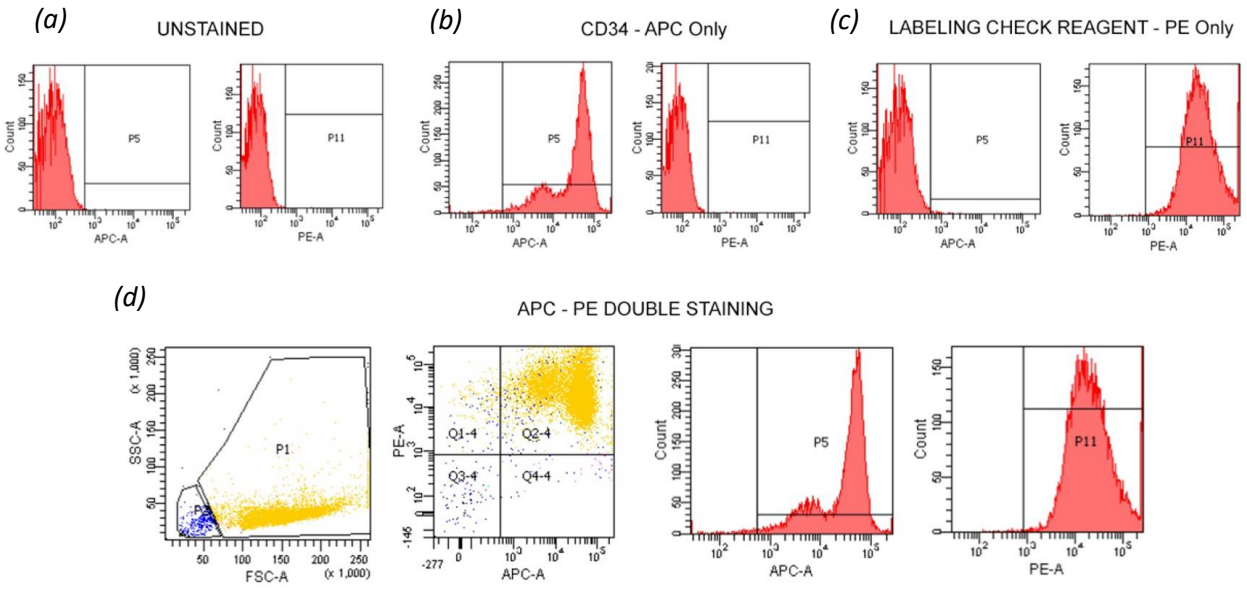


Figure 23: Cytofluorimetric controls referring to CD34-APC/Labeling check reagent-PE immunostaining. a. Unstained control of unfixed $CD34^{+}CD38^{+}$ cells. **b.** CD34-APC single staining performed on unfixed $CD34^{+}CD38^{+}$ cells. **c.** Labeling check reagent-PE single staining performed on unfixed $CD34^{+}CD38^{+}$ cells. The Labeling check reagent labels the beads bond to CD38 during immunomagnetic separation. **d.** CD34-APC/Labeling check reagent-PE double staining performed on unfixed $CD34^{+}CD38^{+}$ cells.

Then, we have set up the fixation conditions and evaluated the impact of cell fixation on immunostaining. We used 2% PFA overnight to fix the immunostained cells. This fixation

strategy proved to be compatible with APC and PE fluorochromes as shown in the flow cytometry analysis of the $CD34^{+}CD38^{+}$ cell fraction (Figure 24 a).

Fixed $CD34^{+}CD38^{-}$ cell population was separated using the DEPArray technique. The immunostaining for CD38 antigen was useful to obtain pure $CD34^{+}CD38^{-}$ cells during the DEPArray sorting, allowing the cell selection for CD34 positivity and CD38 negativity. In addition, during DEPArray separation, nuclei staining was performed using DAPI in order to verify the presence of just one cell inside the DEPArray cage (Figure 24 b).

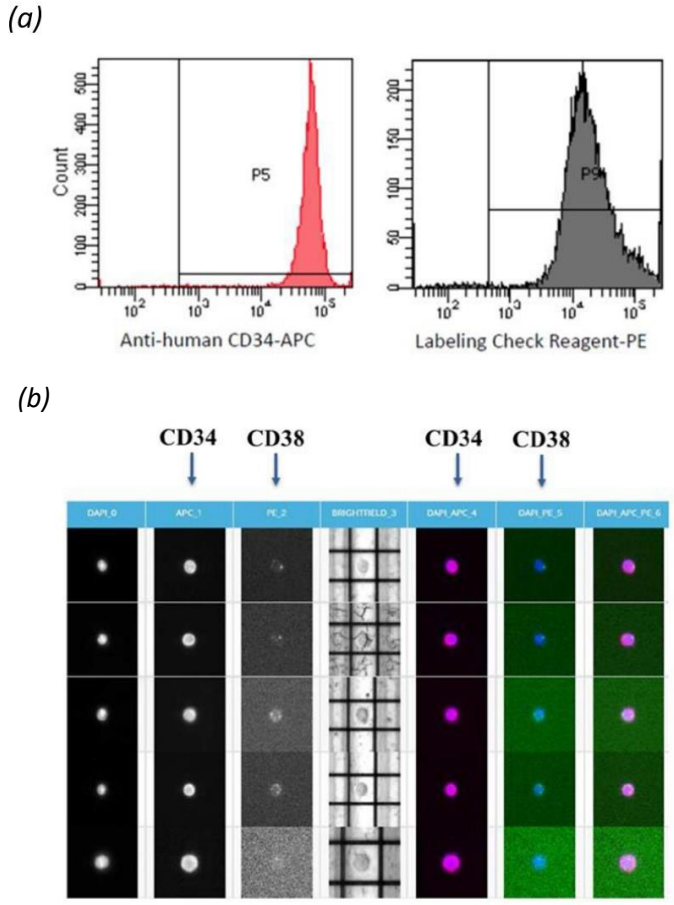


Figure 24: Single-cell sorting using DEPArray. **a.** Flow cytometry analysis of $CD34^{+}CD38^{+}$ cells fixed with 2% PFA overnight. As shown, APC and PE fluorochromes are compatible with PFA. **b**. Single $CD34^{+}CD38^{-}$ cells sorted by DEPArray, labeled with anti-CD34-APC and Labeling check reagent-PE. The labeling check reagent-PE identifies CD38 expression by binding the beads bond to the CD38 antigen during immunomagnetic separation. All cells have been subjected to DAPI staining in order to verify the presence of just one cell inside the DEPArray cage

4.2 Cell fixation method can affect WGA efficiency

After the finalization of sc-sorting, in order to obtain DNA from a single cell and perform an uniforme amplification of sc-genome for downstream sequencing, we tested three commercially available sc-WGA kits. We employed the Ampli1™ WGA Kit, SMARTer PicoPLEX Single Cell WGA kit and REPLI-g Single Cell Kit on single $CD34^{+}CD38^{-}$cells.
The WGA kits that we have selected are based on different principles. Ampli1 WGA kit protocol results on DNA fragmentation through Mse1 restriction enzyme followed by genome amplification, PicoPLEX performs a quasi-random priming pre-amplification step that precedes a uniform isothermal genome amplification and RepliG protocol is based on MDA by φ29 DNA polymerase.
The amplified DNA was analysed for size distribution and for its suitability to be genotyped by Sanger sequencing of *JAK2* and *ASXL1*.
Results showed that Ampli1 kit was able to amplify single-cell DNA (Figure 25 a), but the Mse1-mediated DNA fragmentation determined important limitations due to its high cutting frequency. This type of fragmentation affects the design of sequencing strategies in regions with known mutations in MPNs as in the case of *JAK2* and *ASXL1* (Figure 25 b). Moreover, Sanger sequencing of *ASXL1* exon 13 coming from PicoPLEX- and RepliG-amplified DNA was inefficient in these experimental conditions, despite the absence of the enzyme digestion step (Figure 25 c).

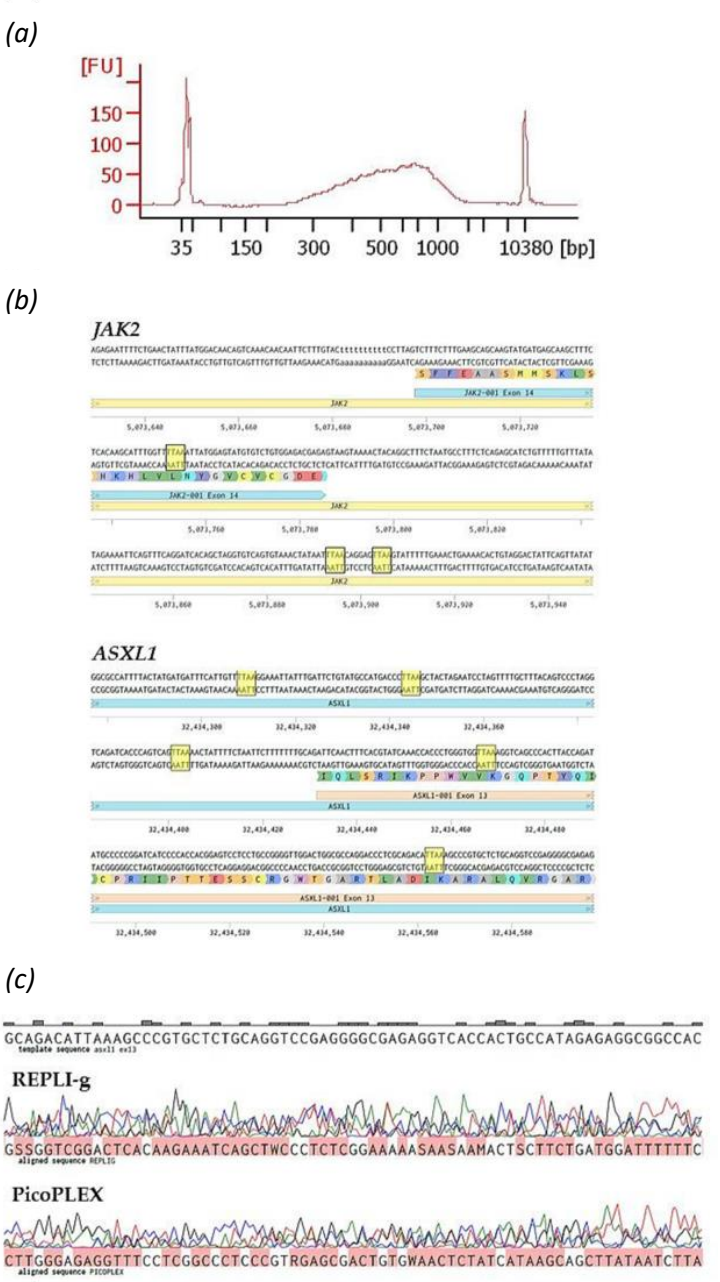


Figure 25: WGA kits amplification efficiency on cells fixed overnight with 2% PFA. **a.** Bioanalyzer profile of amplified DNA, obtained through Ampli1 WGA kit, coming from a single $CD34^{+}CD38^{-}$ cell. **b.** Cutting frequency of MseI enzyme, used for DNA fragmentation in Ampli1 single-cell WGA kit, in two representative regions of *JAK2* and *ASXL1* genes. 5'-TTAA-3' restriction sites are highlighted in yellow. **c.** Two representative sequences of the *ASXL1* gene deriving from $CD34^{+}CD38^{-}$ cells (n = 20) fixed with PFA 2% overnight. A single-cell genome was amplified with either RepliG or PicoPLEX WGA kit. Both sequences were unreadable highlighting the incompatibility between the kits and the fixation protocol.

4.3 MetOH or low-concentration PFA could be alternative fixation methods

In order to understand if the cell fixation method was responsible for the sequencing failure, we sorted at the single-cell level unfixed HEL cells through limiting dilution and subjected them to WGA with PicoPLEX and RepliG kits. Subsequently, we performed Sanger sequencing on *JAK2* exon 14 of amplified DNA. As expected, WGAs obtained from DNA of unfixed cells resulted in the efficient amplification and sequencing of the target region with both PicoPlex and RepliG kits (Figure 26 a).

Considering these results, in order to find a compatible fixation strategy with both immunostaining and WGA, we tried another fixative agent. For this reason, we performed 100% MetOH fixation for 2 h at 4 °C on immunostained $CD34^{+}CD38^{+}$ cells. In addition, at the same time, we decided to employ also MALBAC and DOP-PCR protocols, respectively MALBAC and Genomeplex kits.

The cytometric analysis immediately shows that, unlike PFA, MetOH fixation is not able to preserve PE staining, therefore proving to be unsuitable for downstream DEPArray or FACS separation (Figure 24 b).

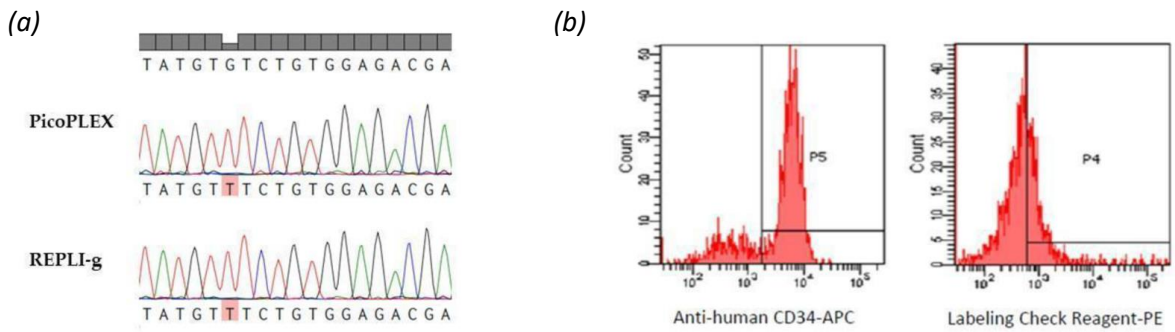
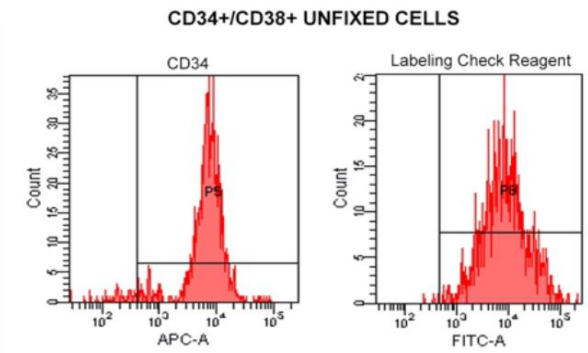


Figure 26: Assays of alternative cells fixation methods. a. Two representative sequences of *JAK2* gene coming from unfixed HEL cells whose genome was amplified through either RepliG or PicoPLEX WGA kit (n = 20 for each kit). In all cells, the homozygous *JAK2* c.1849G>T carried by HEL cells was detected. **b.** $CD34^{+}CD38^{+}$ cells fixed with MetOH 100% for 2 h. Flow cytometry analysis shows that MetOH fixation preserves APC fluorochrome but decreases PE signal. The labeling check reagent-PE identifies CD38 expression.

When we tested an alternative immunostaining strategy, employing Labeling check reagent-FITC, we observed that also FITC fluorescence signal was incompatible with MetOH fixation (Figure 27).

(a)



(b)

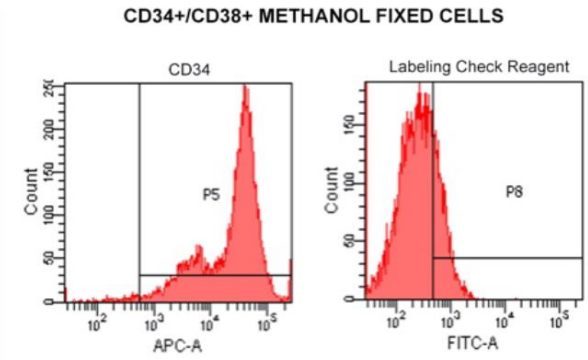


Figure 27: Cytofluorimetric analysis of CD34-APC/Labeling check reagent-FITC immunostaining. **a.** Double staining performed on unfixed $CD34^{+}/CD38^{+}$ cells. **b.** Double staining performed on 100% MetOH fixed $CD34^{+}/CD38^{+}$ cells.

Nevertheless, it is evident that WGAs performed on MetOH fixed cells through PicoPLEX, RepliG, Genomeplex or MALBAC kits resulted in efficient amplification and sequencing of target regions (Figure 28).

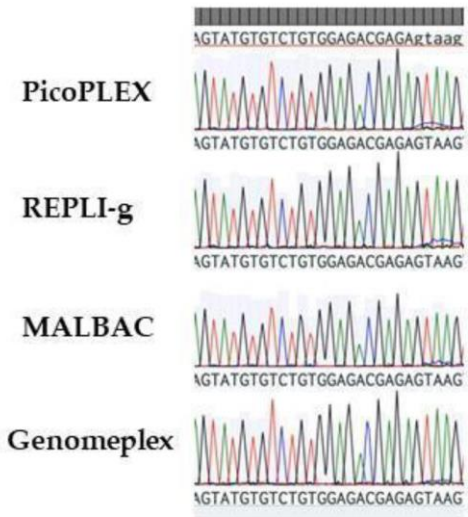


Figure 28: CD34^{+}CD38^{+} cells fixed with 100% MetOH for 2 h at 4 °C. Single-cell WGAs were obtained with PicoPlex, RepliG, GenomePlex and MALBAC kits (*n* = 20 for each kit). Electropherograms represent the analysis of a sequence of *JAK2* gene.

Therefore, although MetOh fixation preserves the quality of the amplified genome, it can disturb cell immunostaining. MetOH fixation may be compatible with WGA applications that do not require immunostained cell populations for sc-sorting.

At this point, we modified the PFA fixation protocol changing the concentration and the time of incubation. In particular, we used PFA at lower concentrations and for shorter incubation times in order to limit its cross-linking activity upon DNA and histone residues that could hamper optimal genomic amplification. We tested 0.5% PFA incubation for 15 min, 45 min and 2 h. In all three conditions, CD34^{+}CD38^{+} cells immunostaining with APC and PE fluorochromes was preserved (data not shown). Results obtained from longer incubation showed that WGA performed in 45 min and 2 h on fixed cells was inefficient with either RepliG or Genomeplex kit and low quality WGAs were obtained through PicoPLEX and MALBAC kits. On the contrary, when using PicoPLEX and MALBAC kits reducing the time of incubation, WGAs showed good efficiency for downstream Sanger sequencing after 15 min of 0.5% PFA (Figure 29). Considering these data, we selected 0.5% PFA incubation for 15 min as the optimal fixation protocol for our experiments.

Figure 29: CD34^{+}CD38^{+} cells fixed with 0.5% PFA for 15 min. Single-cell WGAs were obtained with PicoPlex and MALBAC kits (*n* = 20 for each kit). Electropherograms represent the analysis of a sequence of *JAK2* gene.

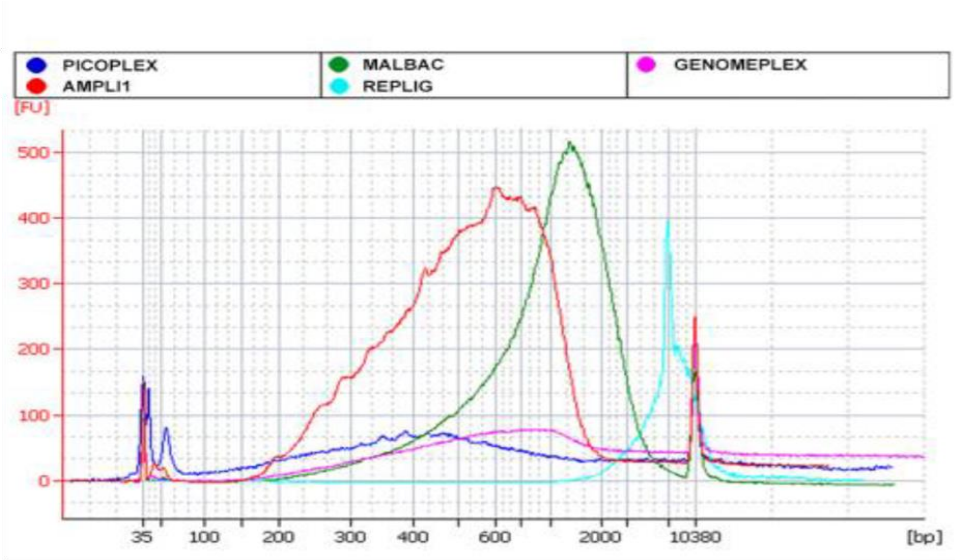
4.4 Comparison of ADO effect in single-cell Whole Genome Amplification kits

Once we optimized the fixation conditions, we started to analyze the performance of the different WGA kits in terms of yield, efficiency for downstream Sanger sequencing and ADO

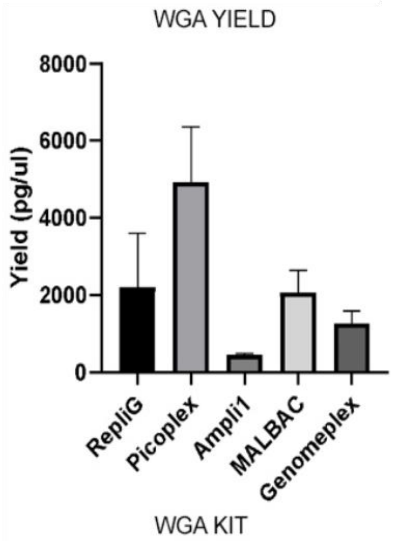
rate. To this aim, K562 cells were fixed in PFA 0.5% for 15 min and WGA were performed with the five different kits (Ampli1, PicoPlex, RepliG, GenomePlex and MALBAC).

We evaluated "quality" and "quantity" of WGA products, obtained by different kits, using the bioanalyzer. Results showed that PicoPLEX and MALBAC amplification protocols were characterized by the highest yield and the larger size distribution. On the contrary, Ampli1 and Genomeplex had a markedly lower yield. In addition, Ampli1 fragments were shorter than the ones produced by the other kits, highlighting once more the limitations of this WGA protocol. Moreover, RepliG fragments were longer than the ones generated by other kits, due to the high processivity of φ29 DNA polymerase (Figure 30 a). Statistical analysis, comparing WGA yield and average fragment length of the different WGA kits, shows that PicoPlex results in greater WGA yield and that RepliG results in the highest average fragment length (Figure 30 b and c).

(a)



(b)



(c)

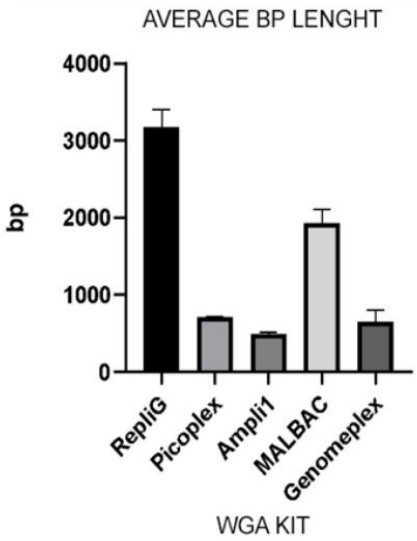


Figure 30: Comparison of single-cell WGA quality, yield and average fragment length of the different WGA kits. a. Bioanalyzer profiles of amplified DNA coming from single K562 cells fixed with PFA 0.5% for 15 min. Color code: PicoPlex in red, Ampli1 in blue, Malbac in green, RepliG in light blue and GenomePlex in purple. **b**. Analysis of the WGA yield of PFA 0.5% 15 min fixed K562 cells subjected to WGA with five different WGA kits (n=5 for each kit). **c.** Analysis of the average length of the genomic fragments coming from PFA 0.5% 15' fixed K562 cells subjected to WGA with five different WGA kits (n=5 for each kit). Data are shown as mean ± SEM.

Then, we assessed the suitability of WGA products for Sanger sequencing analysis. We sequenced a portion of *ASXL1* exon 13 in WGA products obtained from each kit in K562 cells. The region of interest was not amplified due to Mse1 digestion using the Ampli1 kit (Figure 25 b). Despite the good yield, RepliG derived WGA showed to be quite inefficient for Sanger sequencing of this region (37.50% of sequencing efficiency). On the contrary, PicoPLEX and MALBAC showed both a good yield and a high Sanger sequencing efficiency (95%) (Figure 31).

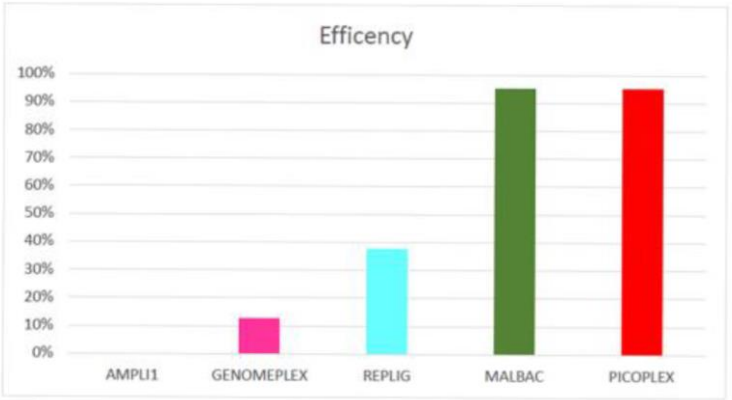


Figure 31: Comparison of sequencing efficiency of the different WGA kits. Percentage of sequencing efficiency, performed on a region of the *ASXL1* gene, obtained with single-cell WGA kits (*n* = 20 cells for each kit).

Finally, we evaluated the ADO rate resulting from the kits (RepliG, PicoPLEX and MALBAC) as an index of good sequencing efficiency. For this reason, we performed Sanger sequencing of a heterozygous variant of the *ASXL1* gene (c.1773C>G) carried by K562 cells. Both RepliG and PicoPLEX show a low ADO rate, on the contrary, MALBAC kit produced a higher ADO rate detecting either wild-type or homozygous cells (Figure 32).

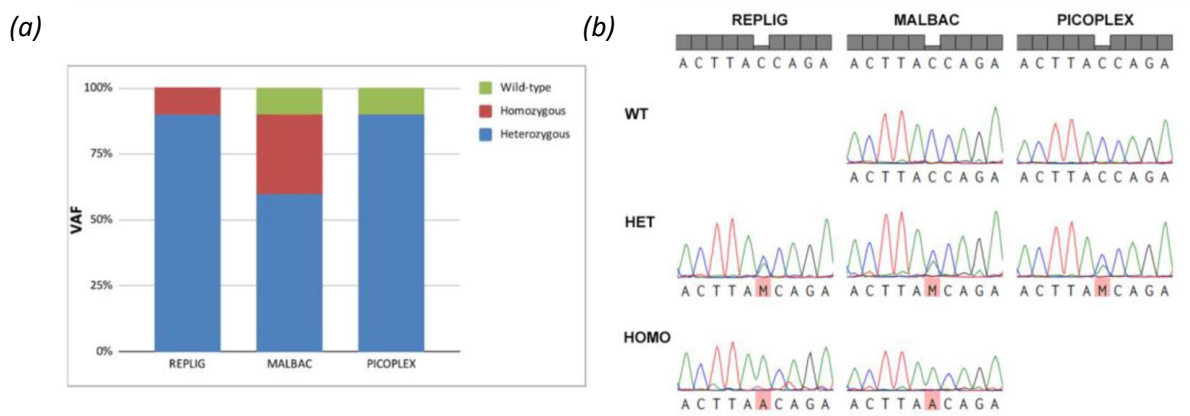


Figure 32: Comparison of ADO effect of the different WGA kits. a. Histogram representing the rate of ADO effect obtained with the analysis of the heterozygous variant of the *ASXL1* gene (c.1773C>G) carried by K562 cell line (*n* = 20 cells for each kit). **b.** Representative peaks of homozygous, heterozygous and wild-type alleles identified using Sanger sequencing. Sequences were obtained by different single-cell WGA kits (*n* = 20 cells for each kit) and were related to the heterozygous variant of *ASXL1* gene (c.1773C>G) carried by K562 cell line.

Altogether our results showed that PicoPLEX is the only kit that combines the highest amplification yield, the highest suitability for downstream sequencing and the lowest ADO rate, compared to the other WGA kits (Table 8).

WGA Method	Compatible with PFA Fixation (High Concentration and Long Time)	Compatible with PFA Fixation (Low Concentration and Short Time)	Compatible with MetOH Fixation	Efficiency (%)	ADO Effect	Sequence Quality
Ampli1	Yes	Yes	Yes	0	-	-
Genomeplex	No	No	Yes	12.5	-	-
REPLI-g	No	No	Yes	37.5	Low	Low
PicoPLEX	No	Yes	Yes	95	Low	High
MALBAC	No	Yes	Yes	95	High	Medium

Table 8: The characteristics and the compatibility of WGA kits

4.5 Single-cell analysis allowed to reconstuct the clonal hierarchy and evolution in a patient with PMF evolved to secondary AML

The set up of a workflow suitable for sc-genomic analysis of primary HSCs has enabled us to employ a sc-approach to study the clonal hierarchy and heterogeneity in MPNs.

Our goal was to reconstruct the clonal architecture of the stem cell compartment during MPNs evolution of a PMF patient. In this regard, we analyzed $CD34^+$ HSPCs isolated from the peripheral blood of a PMF patient at three different stages of the disease: at diagnosis (T1), during the accelerated phase (T2), and in the AML phase (T3). At diagnosis the patient was 78 years old and showed severe clinical features, such as an high rate of peripheral $CD34^+$ cells, hepatosplenomegaly, grade 3 BM fibrosis and he was classified as high risk according to the DIPSS prognostic system. During the accelerated phase he was treated with Ruxolitinb but he was unresponsive to therapy, except for a minimal reduction of splenomegaly. The clinical status subsequently worsened and the patient progressed to secondary AML at T3 as demonstrated by blast frequency (Table 9).

Clinical-hematological data	T1	T2	T3
White blood cells (×109/l)	31.03	105.2	151.18
Neutrophils absolute count (×109/l)	27.8	95.6	118.5
Peripheral blasts (%)	2	2	20
Peripheral CD34+ (%)	1	1	15
Leukoerythroblastosis	Yes	Yes	Yes
Hemoglobin (g/l)	9.9	9.4	8.4
MCV (fl)	87.7	92.6	91.3
Platelets (×109/l)	79	84	41
LDH	1153	2630	1848
Hepatosplenomegaly	Yes	Yes	Yes
Transfusion dependency	No	Yes	Yes
Bone marrow cellularity (%)	55	50	NA
Fibrosis (EUMNET consensus grade)	3	3	Fibrotic substitution
Karyotype (ISCN)	46,XY[25]	Failed	47,XY,+21[3]/46, XY,i(21)(q10)[2]/ 46,XY[25]
DIPSS risk group	High	High	/

Table 9: The clinical and hematological features during time

According to the protocol previously described, we sorted the $CD34^+$ cells at single cell level, performed WGA and sequenced the amplified DNA by means of Sanger sequencing. We designed sequencing strategies based on mutations identified by bulk NGS analysis performed at different timepoints (Table 10). The patient carried the canonical JAK2V617F "driver" mutation.

Mutational state

Gene	Transcript	Variant	Protein mutation	T1 VAF (%)	T2 VAF (%)	T3 VAF (%)
TET2 (a)	NM_001127208	c.3743T>C	p.Leu1248Pro	48.1	42.4	46.1
JAK2	NM_004972	c.1849G>T	p.Val617Phe	75.8	71.5	51.6
TET2 (b)	NM_017628	c.3409_3416delGGTAATGT	p.Gly1137Profs*5	37.7	40.3	47
ASXL1	NM_015338	c.1934dupG	p.Gly646Trpfs*12	43.6	40.6	40.6
SRSF2	NM_003016	c.284C>A	p.Pro95His	46.1	45.6	45.3
TP53	NM_000546	c.713G>C	p.Cys238Ser	5.7	25.6	13.7
FLT3	NM_004119	c.2503G>C	p.Asp835Tyr	/	6.3	35.3

Table 10: The mutational profile of the patient and the changes of VAF mutation during time. The data were obtained through bulk NGS analysis

We analyzed the mutation profile at sc-level of 900 cells (300/timepoint) and reconstructed a single-cell phylogeny according to the presence/absence of the mutations (Figure 33 a). The reconstruction of the phylogenetic tree revealed that just a small number of parental clones seem to generate clonal complexity. In detail, we identified fifty-two clones during evolution (Figure 33 b).

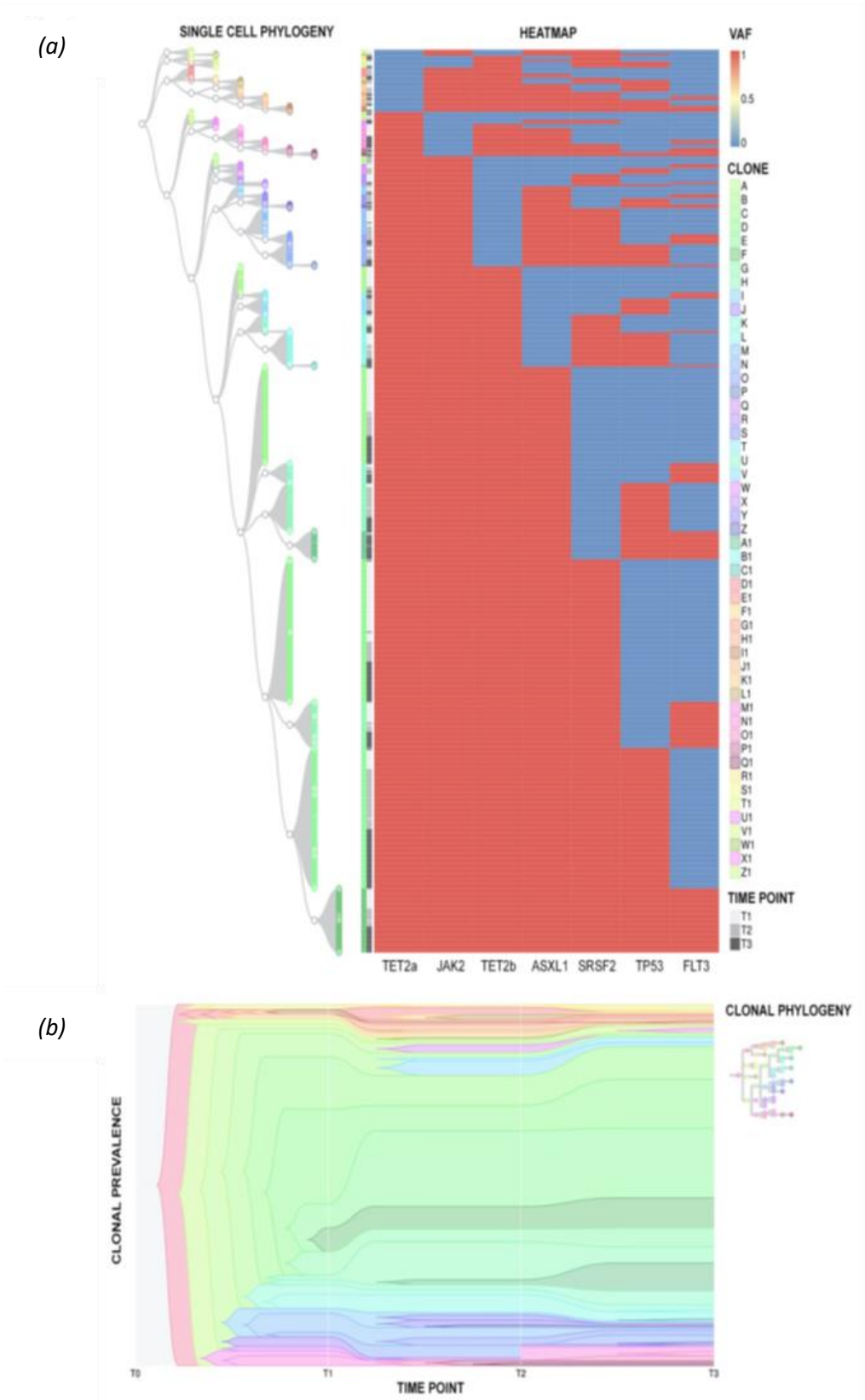


Figure 33: Single cell phylogeny and temporal clone evolution. a. The phylogenetic tree of 900 single cells analyzed in T1, T2 and T3 is shown. Left part: tree starting from a group of founder clones. Each clone (A-Z1) is represented by a color; nodes are represented as white circles. Right part: heatmap describing a mutational event (blue: wild-

type, red: mutated) of the genes indicated in the bottom part of the graph. Each cell in the phylogenetic tree corresponds to a row in the heatmap, identifying its mutational profile. **b**. the Fishplot shows the abundance of the 52 clones (A-Z1), identified in (**a**), and their temporal evolution. Each clone is represented by the same color used in (**a**).

In order to better understand the hierarchical distribution of the clones, we focused on 9 clones that constitute the main branch of clonal architecture and showed significant changes during disease progression. We observed that all clones derived from a progenitor cell carrying the *TET2* mutation p.Leu1248Pro (hereafter called *TET2a*) (Figure 34 a). As shown in Figure 34 a the *TET2*-mutated clone acquired additional somatic mutations resulting in the outgrowth of other clones. We detected the clonal prevalence during the phylogenesis (Figure 34 b). From this analysis emerged that the clone harboring the *TP53* mutation clearly expanded during disease progression (Figure 34 a, b). In addition, we were able to identify the presence of *FLT3*-mutated cells already in the chronic phase (Figure 34 a, b). Of note, *FLT3* mutation was not detected by bulk NGS at diagnosis (Table 10). *FLT3*-mutated clones were detected with low frequency in T1 (9%), then they expanded in T2 (15%) and in T3 (23%). In addition, *FLT3* mutation occurred in a clone displaying all mutations detected in this patient (Figure 34 a, b) Moreover, analyzing all cells, we traced the mutation acquisition order (Figure 34 c) and we noticed that *TET2a* mutation is the first to be acquired, followed by the "driver" JAK2V617F and by another *TET2* mutation (p.Gly1137Profs*5, hereafter called *TET2b*) (Table 10). Then *ASXL1*, *SRSF2* and *TP53* mutations occur in chronological order. *FLT3* mutation represents the last evolutionary genetic lesion acquired. Notably, the percentage of *TP53*-mutated cells increases from 20.8% in T1 to 63% in T2, indicating an involvement of *TP53* mutation in the accelerated phase of the disease. The frequency of *FLT3*-mutated cells grew linearly from 16.6% in T1 to 25% in T2 up to 35.6% in T3, suggesting a role in AML onset.

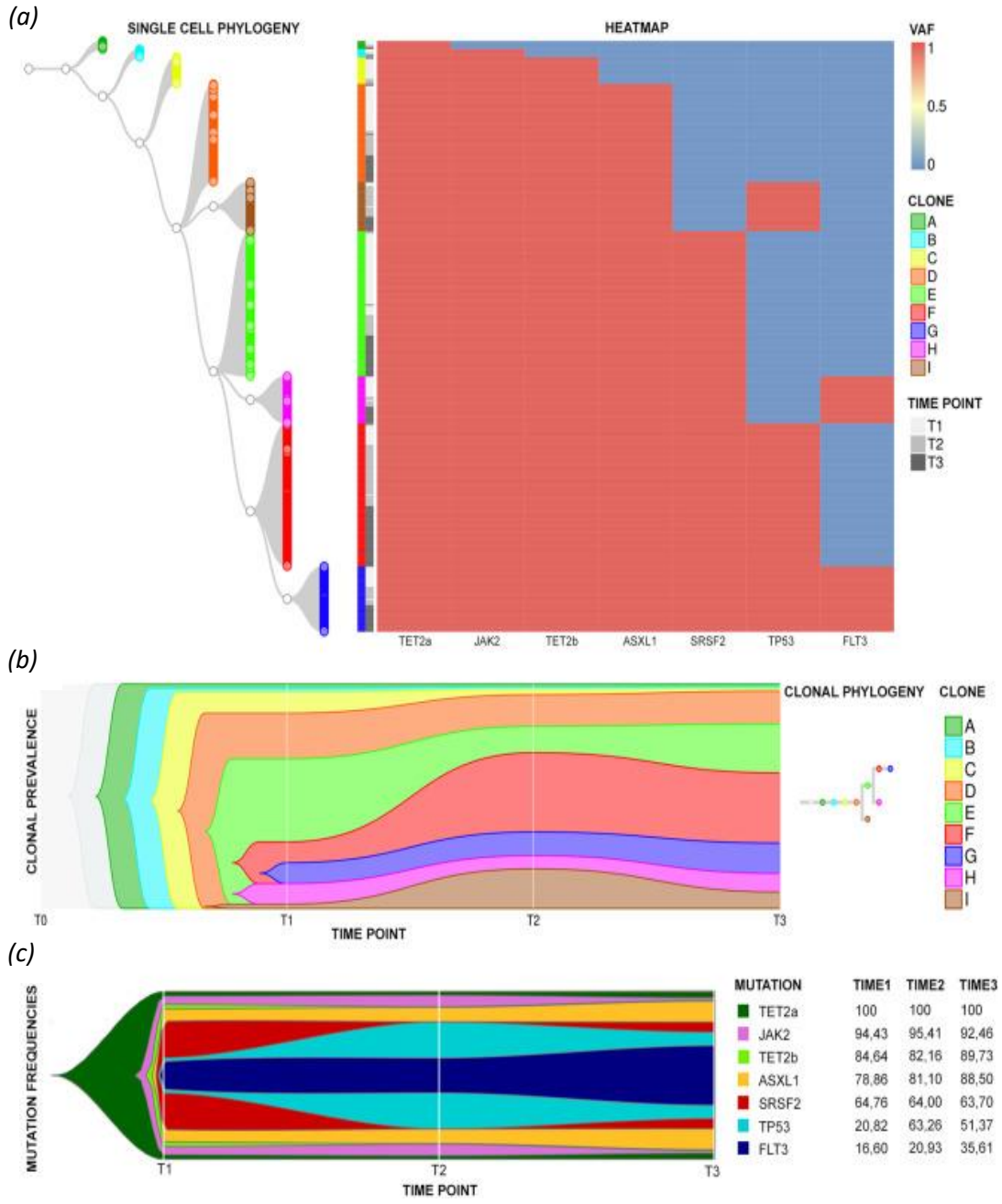


Figure 34: Single-cell phylogeny and mutation acquisition order. **a.** The phylogenetic tree of the most represented clones identified in T1, T2, and T3. On the left is shown the tree starting from the founder cell carrying a single *TET2a* mutation (clone A). From the parental *TET2a*-mutated cell, 8 clones originated (B–I) and each clone is represented by a specific color; nodes are represented as white circles. On the right is the heatmap showing the mutational event (blue: wild-type, red: mutated) of the genes indicated at the bottom. Each cell in the phylogenetic tree corresponds to a row in the heatmap, identifying its mutational profile. In **b**, the fishplot indicates the abundance of 9 clones (A–I), identified in a, and their temporal evolution. Each clone is represented by the same color used in a. In particular, F clone in red and I clone in brown represent *TP53*-mutated clones, while G clone in blue and H clone in purple represent *FLT3*-mutated ones. **c.** The fishplot indicated the variation of mutation frequencies through time and the mutational acquisition order. Each variant is represented by a color. On the right, the table summarizing the mutation frequencies in single cells during time is shown.

4.6 Clonal heterogeneity increases during disease progression from chronic phase to AML phase

Sc-approach can provide information about the zygosity of mutations. We reconstructed the phylogenetic trees for each timepoint, considering the zygosity of the different mutations (Figure 35, 36, 37). Comparing the results, we observed an increase of clonal heterogeneity during disease progression, as regards the combination of different genetic lesions and the zygosity of mutations.

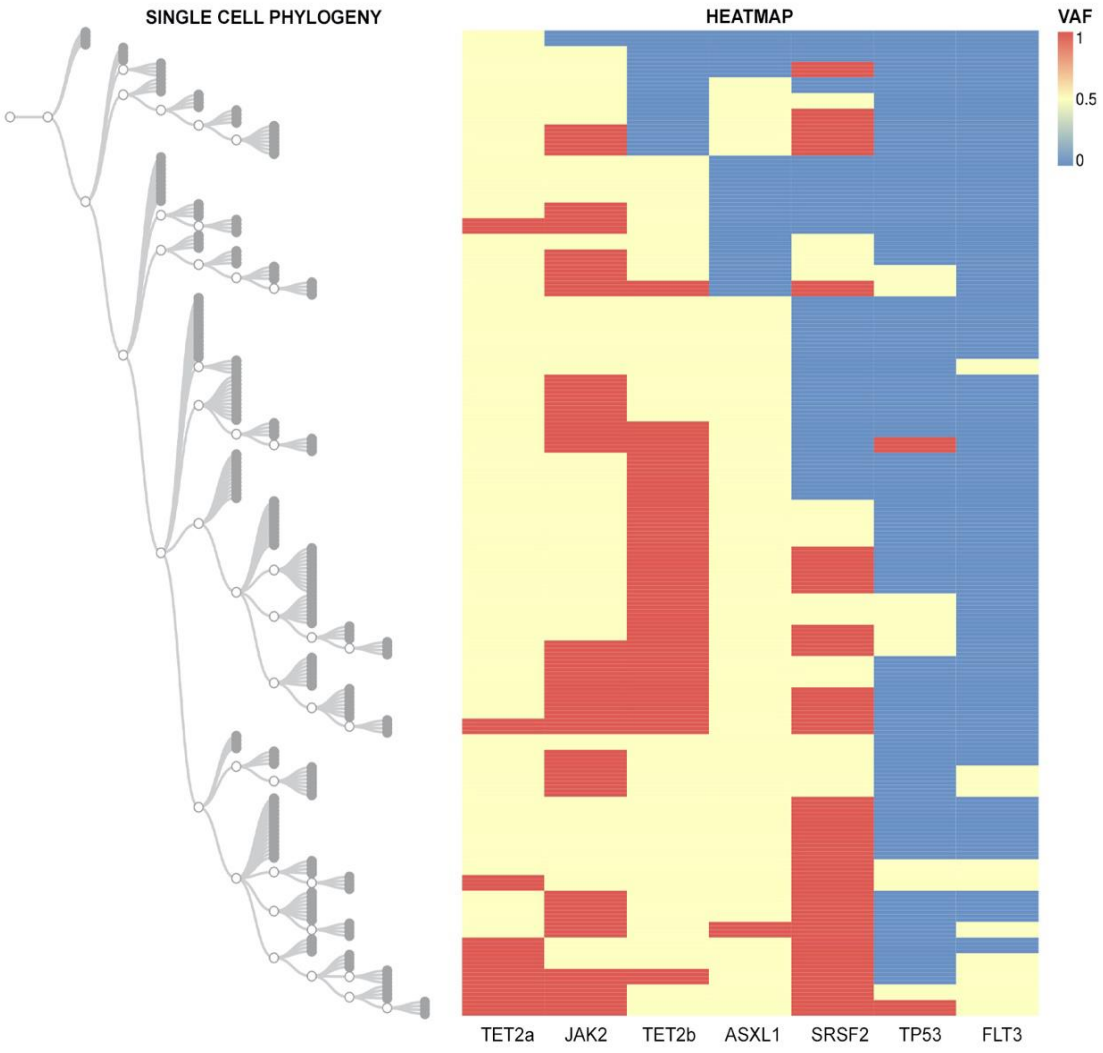


Figure 35: Single cell phylogeny of Time 1. Phylogenetic tree reconstruction built based on the zygosity of the different mutations in Time 1 and performed using CellScape R package. Left part: phylogenetic tree, black circles represents every single cell, the white circles represents the nodes of different clones. Right part: heatmap describing mutational event (blue: wild-type, yellow: heterozygous, red: homozygous) of the genes indicated in the bottom part of the graph. Each cell in the phylogenetic tree corresponds to a row in the heatmap, identifying its mutational profile.

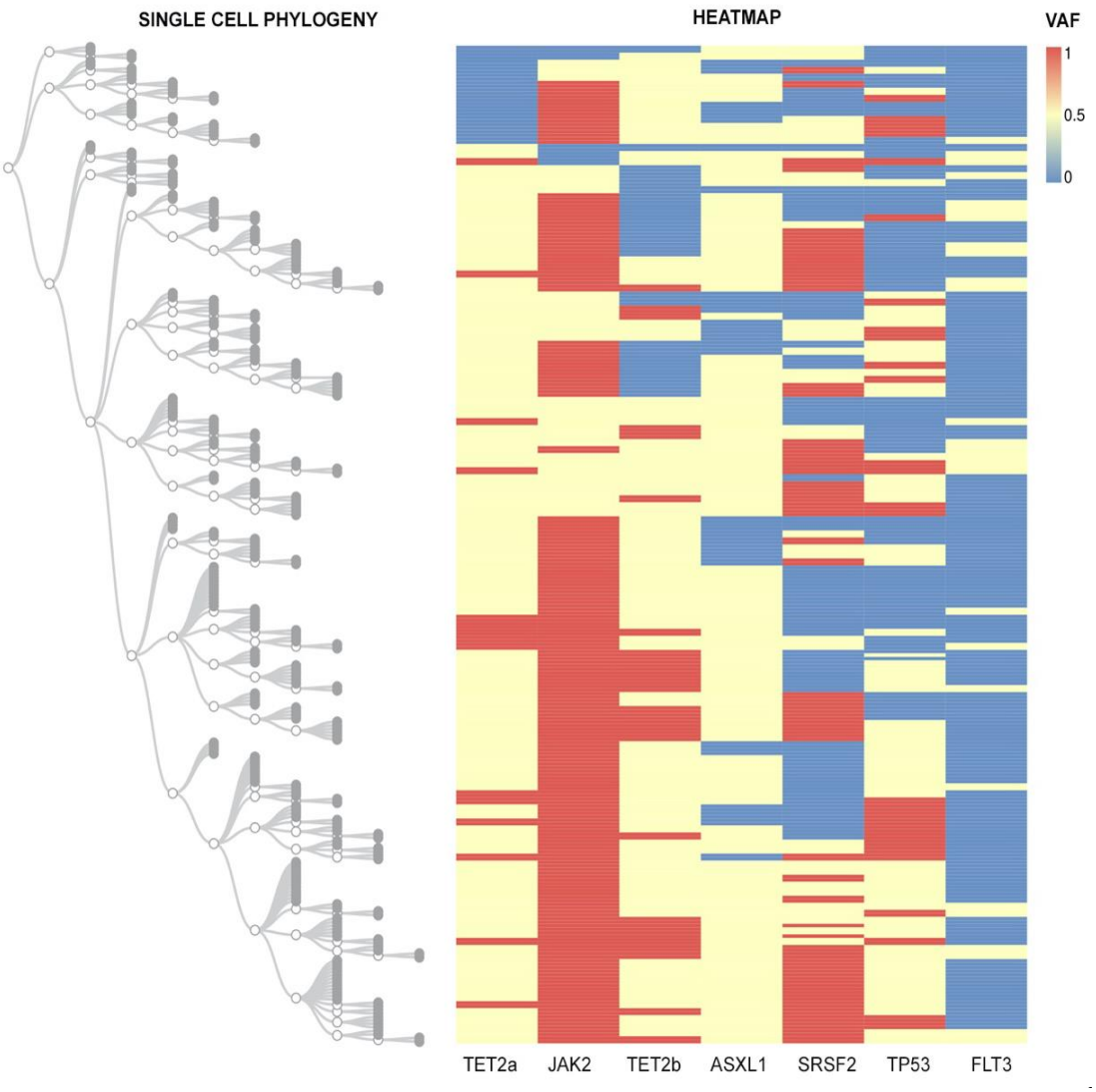


Figure 36: Single cell phylogeny of Time 2. Phylogenetic tree reconstruction built based on the zygosity of the different mutations in Time 2 and performed using CellScape R package. Left part: phylogenetic tree, black circles represents every single cell, the white circles represents the nodes of different clones. Right part: heatmap describing mutational event (blue: wild-type, yellow: heterozygous, red: homozygous) of the genes indicated in the bottom part of the graph. Each cell in the phylogenetic tree corresponds to a row in the heatmap, identifying its mutational profile.

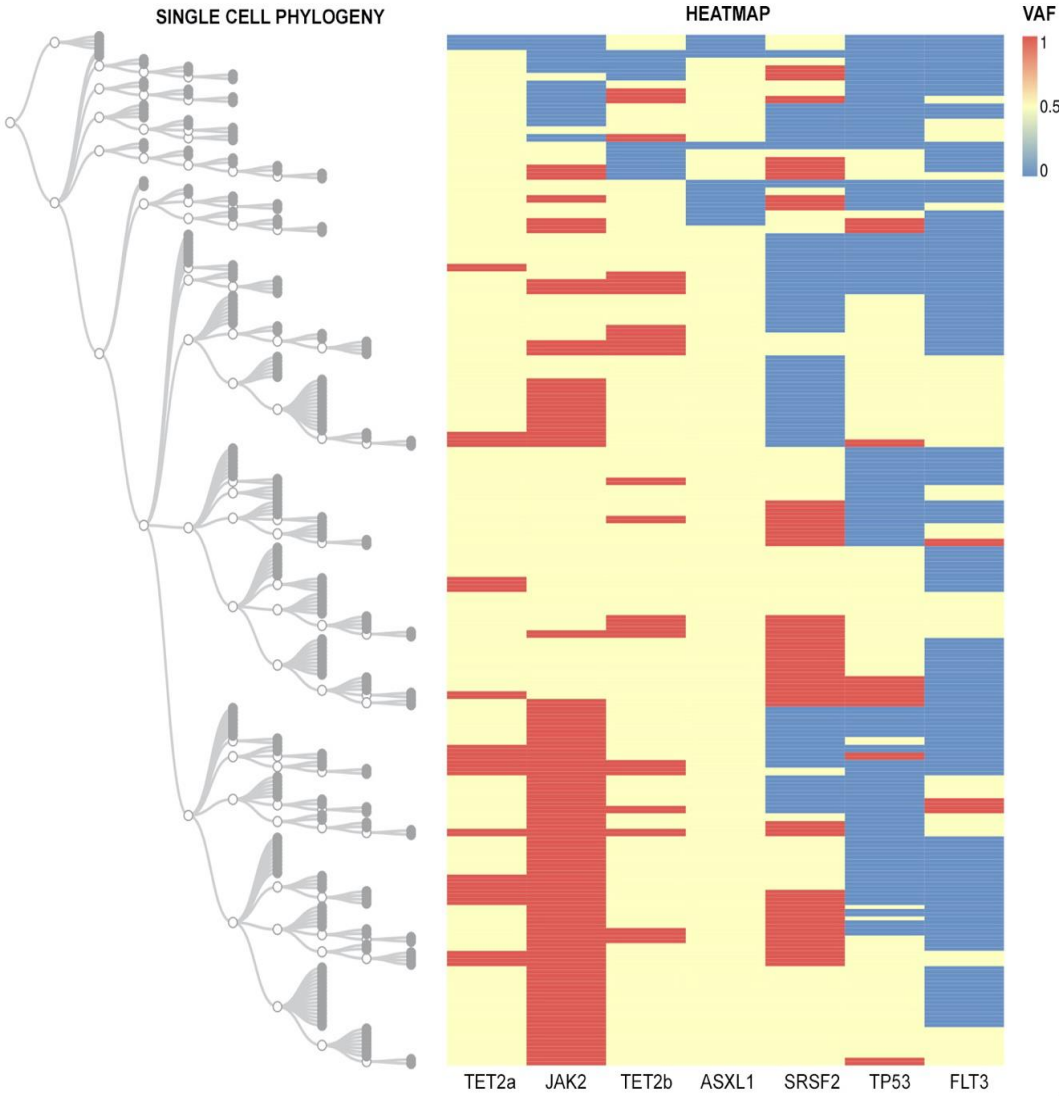


Figure 37: Single cell phylogeny of Time 3. Phylogenetic tree reconstruction built based on the zygosity of the different mutations in Time 3 and performed using CellScape R package. Left part: phylogenetic tree, black circles represents every single cell, the white circles represents the nodes of different clones. Right part: heatmap describing mutational event (blue: wild-type, yellow: heterozygous, red: homozygous) of the genes indicated in the bottom part of the graph. Each cell in the phylogenetic tree corresponds to a row in the heatmap, identifying its mutational profile.

Particularly interesting is the zygosity modulation of the *TP53* and *FLT3* variants. The number of *TP53* and *FLT3*-mutated cells in heterozygous configuration increases during disease progression (Fig. 38 a, b and Table 11). Another very intriguing result concerns the presence of *FLT3* homozygous mutation only in the leukemic phase.

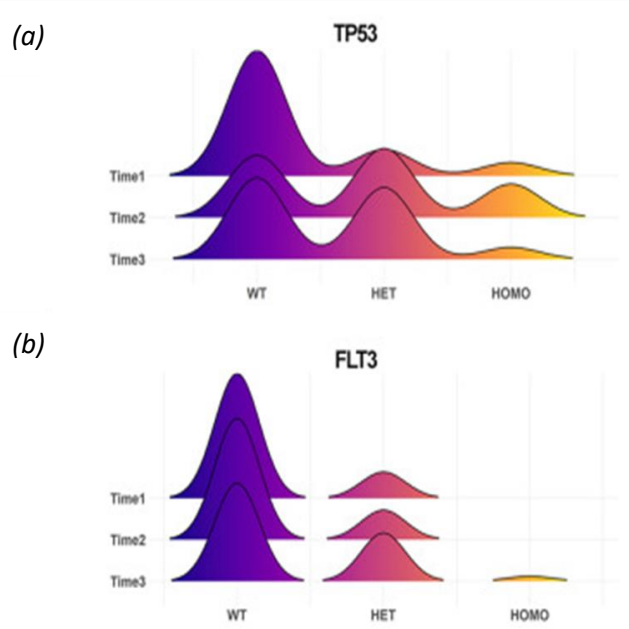


Figure 38: *TP53* and *FLT3* mutations evolution during time. The joy plot shows the distribution of zygosity of *TP53* (**a**) and *FLT3* (**b**) in T1, T2, and T3, respectively

TP53	WT	HET	HOMO
Time1	76	16	8
Time2	38.02	42.25	19.71
Time3	49.34	44.07	6.57

FLT3	WT	HET	HOMO
Time1	82.6	17.3	0
Time2	80.4	19.5	0
Time3	64.9	31.8	3.2

Table 11: Percentage of TP53 and FLT3 mutated cells during disease progression

The other genetic alterations do not show relevant change in zygosity during time (Figure 39).

Figure 39: Zygosity distribution of all mutations in each time point. Bubble plot of the frequencies of mutations through time performed by ggplot2 R package. On the y axis are represented all variants in Time 1, 2 and 3; on the x axis is shown the mutational status (WT: wild-type; HET: heterozygous; HOMO: homozygous). Each colour indicates a different variant. Circles' dimension indicates the relative abundance of the event.

4.7 Single-cell sequencing allows the characterization of a novel *TP53* variant

Given the importance of *TP53* loss-of-function in the activation of mechanisms underlying tumor evolution and the expansion of *TP53*-mutated cells during disease progression, we evaluated the presence of other genetic alterations, such as CNVs. In order to assess the occurrence of CNVs during disease progression, bulk NGS data were analyzed using the CNV

algorithm implemented in the SOPHIA DDM platform, which did not reveal detectable alterations.

In addition, considering that clones carrying *TP53* p.Cys238Ser mutation increased during disease progression, we analyzed this region by MLPA in bulk $CD34^+$ cells of T2 and T3, in order to identify potential CNVs in the *TP53* gene. We chose to study cells of T2 and T3 according to the significant expansion of the *TP53*-mutated clones in these timepoints. Also this analysis did not highlight any gene rearrangement (Figure 40 a-d).

Then, we employed sc-genomics in order to detect any smaller CNVs affecting *TP53*. In this regard, we sequenced all of *TP53* gene exons at single cell level for each timepoint. Interestingly, Sanger sequencing identified a 454 bp deletion between *TP53* exons 2 and 4 (Figure 40 e). This new likely pathogenic variant was found in a small cell subpopulation, already harboring the p.Cys238Ser homozygous mutation. Moreover, its frequency increased in T2 following the trend of p.Cys238Ser mutated cells (Figure 40 f).

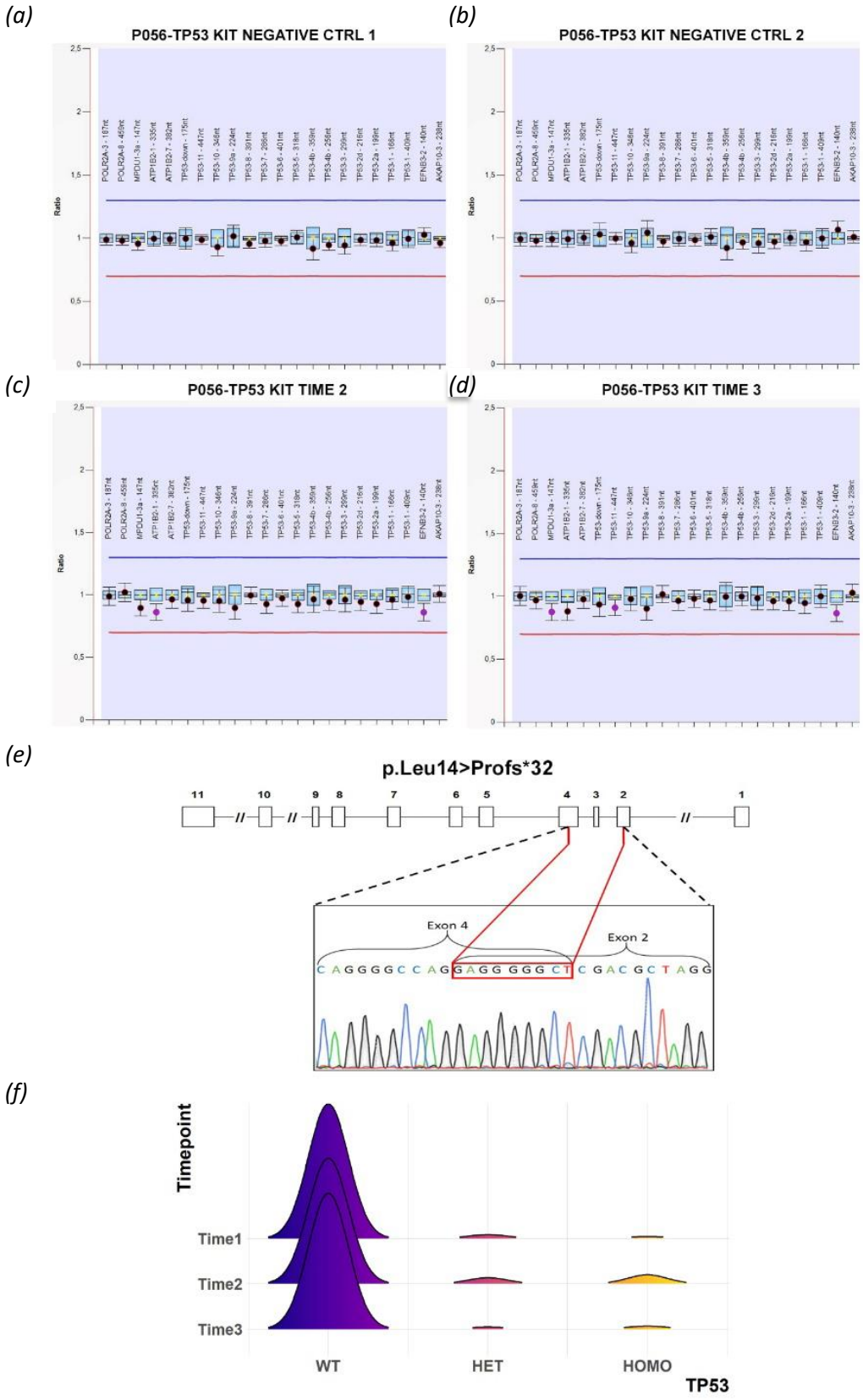


Figure 40: Characterization of *TP53* deletion. Panels **a-d** shows the results of Multiplex Ligation-dependent Probe Amplification (MLPA) analysis of *TP53* gene in bulk CD34^{+} cells. The data obtained from two healthy controls (Ctrl) (**a**, **b**), T2 (**c**) and T3 (**d**) do not show any CNVs. Panel **e** shows a deletion of 454bp in *TP53* gene (c.41_152del,

p.Leu14>Profs*32) affecting a region between exons 2 and 4, detected by Sanger sequencing. In Panel **f** the Joy Plot shows the distribution of 454bp *TP53* deletion referred to all $CD34^+$ cells analyzed in T1 (wt 96%, het 2,6%, homo 1,4%), T2 (wt 89,5%, het 4,2%, homo 6,3%) and T3 (wt 97,4%, het 0,65%, homo 1,95%). Wt: wild-type; het: heterozygous; homo: homozygous.

4.8 Single-cell sequencing allows the characterization of *SRSF2* homozygous mutation

The VAF of SRSF2 P95H in the bulk NGS analysis was about 50% in this patient (Table 10). Surprisingly, sc-study identified a consistent fraction of homozygous cells. In particular, sc-analysis detected P95H variant as homozygous in 39.4%, heterozygous in 26.7%, and wild-type in 33.8% of T1 cells (Figure 41). The zygosity of this variant does not change significantly during disease evolution (Table 12). Nevertheless, it is interesting that P95H homozygosity is detectable only at the single-cell level, given that VAF about 50% in NGS might be wrongly associated with heterozygous configuration.

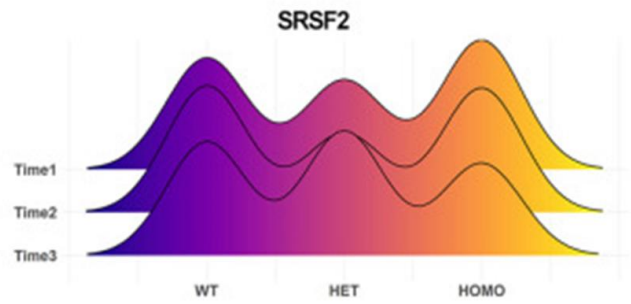


Figure 41:· SRSF2 mutations evolution during time. The joy plot shows the distribution of zygosity of *SRSF2* in T1, T2, and T3, respectively

SRSF2	WT	HET	HOMO
Time1	33.8	26.7	39.4
Time2	37.8	25	37.1
Time3	34.4	37.6	27.9

Table 12: Percentage of SRSF2 mutated cells during disease progression

To exclude a possible ADO effect and verify our results, we performed the same mutational analysis on $CD34^+$ cells-derived colonies, therefore not subjected to WGA. In this way, we could exclude the possibility of errors related to WGA. As a result, Sanger sequencing detected only heterozygous and homozygous colonies (Figure 42).

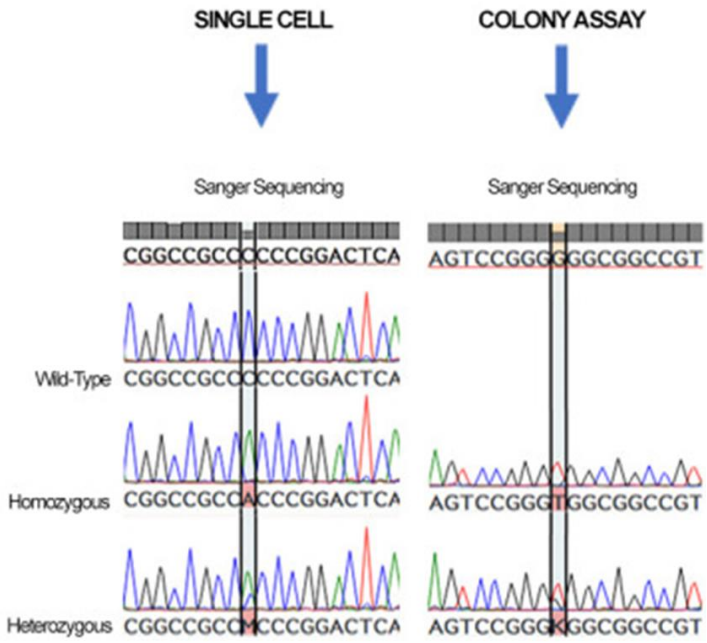


Figure 42: Characterization of *SRSF2* homozygous mutation. The results of Sanger sequencing on single cells (SC) and colonies obtained from PB CD34^{+} cells. On the left are shown the electropherograms representing wild-type, heterozygous, and homozygous *SRSF2* gene variants in SC, while on the right there are electropherograms from colonies indicating the heterozygous and homozygous mutational state

To further validate these findings, we performed a SNP genotyping assay on both single-cell WGAs and DNA extracted from colonies. The results confirmed the presence of homozygous cells, in detail 36% of wild-type, 28% of heterozygous, and 36% of homozygous cells in T2 (Figure 43 a and Figure 39). In addition, we identified 16% of heterozygous and 84% of homozygous colonies in T2 (Figure 43 b). Surprisingly, no wild-type colonies are detected. This difference between single-cells and colony analysis could be explained by the possibility that the semisolid culture could introduce some bias due to the prolonged growth of selected clones.

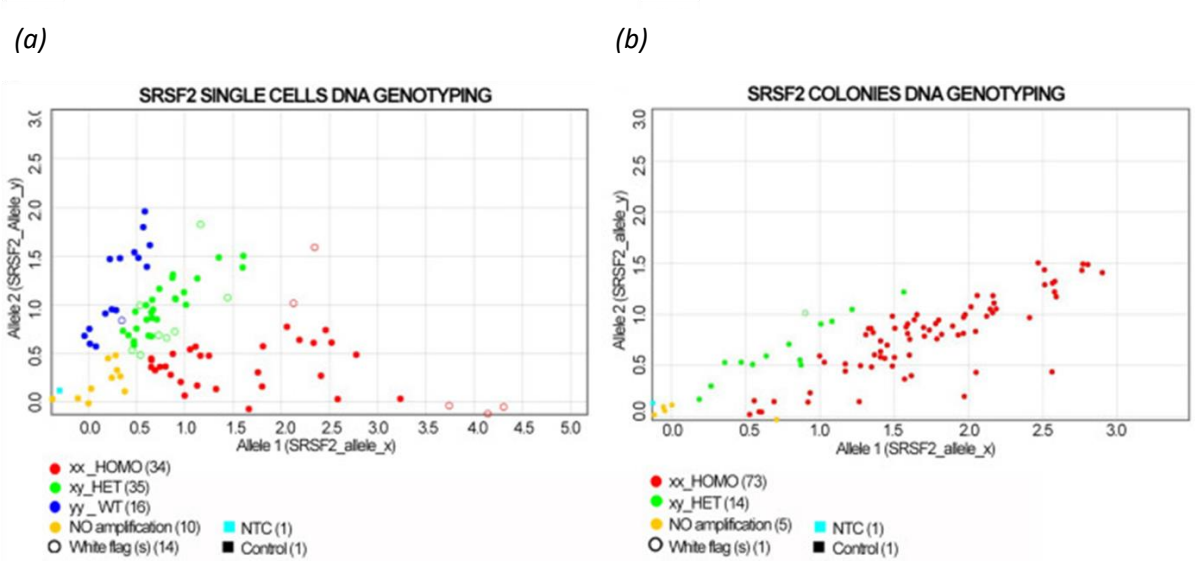


Figure 43: *SRSF2* genotyping confirmed an *SRSF2* homozygous mutation. Representative dot plots show genotyping results obtained from SC (**a**) and colonies (**b**). Allele 1 (x) indicates *SRSF2* mutated gene and allele 2 (y) refers to *SRSF2* wild-type gene. The graph represents the distribution of amplified DNA according to a wild-type (y/y, blue dot), heterozygous (x/y, green dot), and homozygous (x/x, red dot) configuration. In a, the dot plots highlight the presence of wild-type (blue), heterozygous (green), and homozygous (red) single cells. In b, the dot plot shows heterozygous (green) and homozygous (red) colonies.

5. DISCUSSION

The application of single-cell technologies has recently given new and essential information for understanding the molecular mechanisms involved in oncogenesis and cancer progression. In particular, the high resolution that is provided by the sc-genomics approach allows to better characterize the genetic events involved in disease initiation and evolution, in relapse, in metastasis and in MRD. In addition, sc-studies help to unravel the clonal diversity, giving useful information for diagnosis, for risk stratification and personalized therapy [201]. For these reasons, many molecular studies in oncology employed sc-approach to study the cellular subpopulations involved in many solid tumours, such as breast [234], kidney [235], colon [236] and prostate [237] cancers. In the onco-hematology field, sc-technologies contributed to the understanding of molecular events that guide disease onset and progression, and to the identification of clonal characteristics that determine the selection and expansion of neoplastic populations [225-230].

PMF is a complex clonal disorder and represents a good candidate for the study of clonal heterogeneity in hematologic neoplasms. PMF is characterized by a molecular and cellular complexity derived from the acquisition of several somatic mutations that impact at multiple levels on the biology of cancer cell populations. From the clinical point of view, this results in different clinical histories of PMF patients, regarding both disease progression and response to therapy [76]. In this view, we decided to use sc-genomic analysis to characterize the genomic landscape and to determine the clonal architecture of PMF during disease progression.

As previously described, when developing a sc-study, two aspects of the protocols are crucial for the success of the experiment: the sc-sorting and the WGA.

For this reason, first of all, we aimed at optimizing a workflow for cell immunostaining and fixation in order to perform sc-sorting of hematopoietic $CD34^{+}CD38^{-}$ stem cells, in which originate hematological malignancies such as PMF. In our experimental design, we chose to employ the DEPArray system to sort $CD34^{+}CD38^{-}$ cells at single cell level.

The DEPArray from Silicon Biosystems is an automated sorting system that employs immunofluorescence to identify the cells to be isolated. In order to provide an accurate cell selection, this device includes a fluorescent microscope, which analyses the cell immunophenotype and assess the cell viability [210].

In order to preserve cell immunostaining, therefore, it was necessary to set up the cell fixation conditions, ensuring an efficient sc-sorting. Many protocolos of cell fixation are available. Each method has advantages and disadvantages and, for this reason, it is important to choose the most compatible fixation agents with their own experimental design. In general, cross linking fixatives, such as formaldehyde and glutaraldehyde, or low molecular weight alcohols (methanol, ethanol) are preferred. Fixation with formaldehyde is advantageous to "fix" the cell in a life-like state, given that it preserves proteins in native conformation both with or without post-transcriptional modifications [238]. In addition, cross linking fixatives, such as formaldehyde, create covalent chemical bonds between proteins in tissues and arrest all the cellular activity [239]. Another cross-linking fixative, PFA, is able to anchor soluble proteins to the cytoskeleton conferring additional rigidity to the tissue [240]. PFA fixation is particularly recommended for immunostaining and for further analysis through flow cytometry [241] and for cell sorting [242].

Our results demonstrated that PFA fixation was compatible with APC and PE fluorochromes (Figure 24 a). Moreover, the subsequent DEPArray separation confirmed the immunophenotype and provided a good single cell separation (Figure 24 b).

The other limiting step in sc-genomic protocol is the single-cell WGA, given that a single cell contains a low amount of DNA (<10 pg) [242], that could be insufficient for a mutational analysis. In recent years, however, several WGA methods, such as DOP-PCR, MDA, MALBAC, have been developed based on different molecular principles to optimise DNA amplification reducing bias and errors [211]. Despite its widespread use, PFA can alter nucleic acid structure and integration, impacting on genome uniform amplification, given that PFA generates a cross-linking effect between DNA and histone residues [244].

In order to test the compatibility of PFA fixation and WGA, we assayed several commercial kits. Our results demonstrated that PFA fixation was compatible with Ampli1 WGA kit. However, this strategy was not appropriate for our experiment given that Ampli1 employs Mse1 restriction enzyme, which cuts on some hotspot regions of genes frequently mutated in PMF, such as *JAK2* and *ASXL1* (Figure 25). In particular, many Mse1 restriction sites are located in close proximity to the exon 14 codon affected by JAK2V617F "driver" mutation, making the variant discrimination very difficult. For this reason a WGA strategy based on enzymatic fragmentation is strongly limited for the purpose of our work.

In contrast, PFA fixation was not compatible with WGA kits that bypass the enzymatic fragmentation, such as PicoPlex and RepliG (Figure 25). We supposed that the fixation conditions were too strong, probably due to the crosslinking effect introduced by PFA between DNA and the complexed proteins. We corroborated this hypothesis by performing WGA of unfixed HEL single cells. In this case WGA was efficient and Sanger sequencing of selected regions was successful, confirming the deleterious effect of strong PFA cell fixation on downstream analyses (Figure 26 a). Thanks to these results, we considered the use of other fixative agents.

For many years alcohol-based reagents have been employed for the fixation in biological preparations, such as methanol (MetOH) and, in fact, it can be considered a good cytological fixative agent [245]. Alcohol fixatives work by dehydration and cause protein denaturation and precipitation. By this activity they damage the cellular structure, determine changes in protein conformation and, for this reason, they are not ideal to preserve sample imaging but they are useful for experiments that require nucleic acid preservation [246]. In addition, MetOH fixation is considered a gentle method of fixation that can be compatible with scRNA-seq application, given that it preserves the nucleic acids integrity, preventing RNA degradation [247]. MetOH is suitable for droplet-based transcriptional analysis without compromising scRNA sequencing data [248].

Our results demonstrated that the use of MetOH as an alternative cell fixative resulted in efficient WGA and Sanger sequencing with most of the commercially available WGA kits (PicoPlex, RepliG, Genomeplex and MALBAC) (Figure 28). However, MetOH negatively affected the preservation of cell immunostaining, in particular it strongly reduced PE fluorescence, proving that MetOH is unsuitable for single-cell sorting with the DEPArray system (Figure 26). Our data confirmed the use of MetOH as good fixative for sc-genomics analysis, as already described [247, 248]. However, we found that MetOH fixation compromise sc-sorting of cell populations based on immunostaining.

Therefore, we performed $CD34^{+}CD38^{-}$ cell fixation with lower PFA concentrations and for a shorter time in order to mitigate its cross-linking effect. The adoption of this protocol resulted in efficient WGA using PicoPlex and MALBAC kits, both based on the MALBAC amplification strategy (Figure 29). On the contrary, the use of RepliG and GenomePlex kits, based on MDA and DOP-PCR approaches, were unsuccessful.

Finally, we found that the optimal fixation protocol for $CD34^{+}CD38^{-}$ immunostained cells is 0.5% PFA for 15 min.

In order to assess DNA yield, ADO rate and the efficiency of downstream Sanger sequencing, we tested in these experimental conditions the five WGA commercial kits on single K562 cells (Figure 30, 31 and 32). As expected, the WGA yields and fragment size distribution reflected the molecular approaches employed by the different kits. Ampli1 WGA kit produced markedly lower yield and shorter fragment size (Figure 30), confirming the limitations associated with Mse1 digestion. PicoPlex and MALBAC showed the higher yield and the larger fragment size distribution, according to their quasi-random priming strategy and uniform amplification of the hairpin molecules generated during the pre-amplification step. In addition, MALBAC-based approaches showed a high sequencing efficiency of the region of interest, suggesting that the looping-based amplification allows a uniform coverage of genomic DNA. On the contrary, RepliG protocol, based on MDA, generated the longest WGA fragments, but failed to provide DNA suitable for Sanger sequencing. Our results are in agreement with a previous study that demonstrated a higher level of uniformity, specificity and reproducibility of MALBAC than MDA technologies [220].

Finally, we performed ADO rate assessment. ADO is a common effect of the genome amplification techniques that reduces the efficiency of PCR-based targeted sequencing. ADO was firstly described as a "partial amplification failure," involving in a misdiagnosis for both dominant and recessive diseases. This phenomenon consists in a selective allele amplification during PCR [252]. The ADO rate is a crucial parameter to evaluate a correct single-cell genotyping in terms of variant zygosity. Sanger sequencing performed on a heterozygous variant of the *ASXL1* gene carried by K562 cells demonstrated that RepliG and PicoPlex kits are able to detect heterozygous mutations with a higher accuracy than MALBAC kits. Using MALBAC kits we detected a higher rate of false positives (homozygous mutants) and negatives (wild-type) (Figure 32).

Comparing the efficiency of genome amplification, the ADO effect and the sequence quality of all of the tested kits we confirmed previous studies that used different WGA approaches to detect CNVs. The high reproducibility, fidelity and great uniformity of genome amplification are strongly required to highlight CNVs. The strategy adopted by PicoPlex seems to be the most advantageous, given that it balances efficiency, quality of genome amplification and mutation analysis [249-251]. As a whole, our study demonstrated that a mild PFA fixation (0.5% PFA for 15 min) is compatible with $CD34^{+}CD38^{-}$ cell immunostaining, DEPArray sorting and single-cell WGA. The PicoPlex WGA kit was found to be the better performing kit in terms of WGA yield, fragment size, sequencing eciency and ADO rate.

In this work we propose a robust and reproducible protocol to conduct single-cell experiments on HSCs and HSPCs [232]. In addition, we demonstrate the power of sc-technologies in unmasking clonal heterogeneity that characterizes hematological neoplasms. We applied the sc-genomic technology to study the genetic landscape of the HSPCs in MPNs and to understand the molecular events that guide MPNs progression. In particular, we studied a case of PMF that developed secondary AML. We followed the disease evolution during the time, by analysing the $CD34^+$ cell population at three different stages of the disease: diagnosis (T1), accelerated phase (T2) and AML phase (T3).

Through NGS bulk analysis we characterized the mutations carried by the patient in the different timepoints (Table 10). The patient harboured the canonical JAK2V617F mutation, two different *TET2* mutations and *ASXL1*, *SRSF2* and *TP53* genetic lesions at the diagnosis.

The mutation profile at sc-level allowed us to reconstruct the clonal phylogenetic tree and to detail the mutation acquisition order. Firstly, we observed that every clone carried the *TET2* mutation, demostrating that *TET2* genetic lesion represents the first mutational event, followed by the "driver" JAK2V617F (Figure 34).

Even if JAK2V617F is considered a "driver"mutation, it has been reported that other somatic genetic alterations, such as those affecting *TET2*, can precede JAK2V17F in HSPCs of MPN patients [177, 253]. *TET2* is commonly mutated in MPNs and confers an increased risk for reduced overall survival and acute leukemia transformation [254]. Colony genotypization showed that, in some patients, *TET2* mutations occurred before JAK2V617F, while in others *TET2* and JAK2V617F mutations can arise in different clones. According to the authors, *TET2* genetic alterations could represent a predisposing event for the acquisition of JAK2 mutations. [255]. In addition, many studies demonstrated that mutations in epigenetic modifier genes, such as *DNMT3A*, *EZH2* and *ASXL1*, can occur before the "driver" mutation [185, 192].

In a work published in 2015, Ortmann and collegues attributed a clinical significance to the mutation order in patient derived HSPCs carrying *TET2* and JAK2V617F mutations,. The authors found that when a *TET2* is the first genetic lesion, HSPC population undergoes a significant expansion. In this work it is shown that the mutation order influences therapy response and clinical outcome; for example, "*TET2*-first patients" are less sensitive to Ruxolitinib treatment [231]. Our results are in accordance with this research, in fact the patient we analyzed was treated with Ruxolitinib, but the therapy did not significantly change his clinical history and he progressed to leukemia. Our work confirmed the importance of

understanding the order in which genetic lesions occur because it can strongly affect how patients respond to therapies.
Understanding the mutation order might improve the risk stratification and delivery of targeted therapy in order to intervene early on the driver events of the disease evolution. The prediction of disease progression based on the presence of heterogeneous and detrimental mutations may help the drug discovery and the development of predictive biomarkers [256].

In this regard, through sc-analysis we identified the presence of *FLT3*-mutant cells already in the chronic phase, although the *FLT3* mutation has not been detected by the bulk NGS (Figure 34 and Table 10). *FLT3* mutations were found in approximately 30% of AML cases and are a negative prognostic factor [257]. *FLT3* mutations are predominant in *de novo* AML, but represent rare event in MPNs; in fact, it is estimated that the mutation affects 3% of MPN-AP/BP cases. In both MPN-AP and MPN-BP *FLT3* mutations are more frequently associated with other genetic alterations such as *ASXL1*, *IDH1/2*, *TET2*, *SRSF2* and *TP53* [258].

The patient we studied carried a FLT3-tyrosine kinase domain (TKD) mutation as the last mutational event, which promotes cell proliferation by constitutively activating the receptor, as well as the canonical FLT3-internal tandem duplication (ITD) [259]. We showed the progressive expansion of *FLT3*-mutated clones during the progression evolution from chronic phase to BP (Figure 34), suggesting a probable support in AML onset.

As a whole, sc-technology allowed us to detect in advance some leukemic markers in the chronic phase, that were undetectable by the current bulk analysis due to their low allelic frequency. The ability to identify earlier leukemic markers empathizes the importance of our approach given that it could provide better prognostic evaluation and risk stratification.

Several genes were commonly involved in leukemic transformation, such as *TET2* and *ASXL1*. PMF patients showed different mutational profiles and the mutational order is different from one patient to another. This genetic heterogeneity therefore complicates the identification of the specific mutation role in leukemic progression [260].
Our approach could help to understand the events contributing to leukemic transformation that are still not well characterized. The identification of patients at higher risk of leukemic transformation could suggest to change the therapeutic approaches before the clinical manifestation of BP. The treatment options in MPN-BP are limited to conventional approaches, like standard chemotherapy, with limited efficacy. In fact, median survival for patients with secondary AML is about 2.4 months [261, 262]. The MPN patients who evolved to AML

showed a resistance to anti-leukemic treatment adopted in *de novo* AML, probably because they have distinct clinical and genetic features. The alloHSCT is the best therapeutic option and patients ineligible for alloHSCT are treated with hypomethylating agents or combined treatment strategies. Therefore, in chronic phase it is crucial to detect mutations that could predispose and drive sAML development. The study at sc-level not only provided information about the AML transformation but also allowed us to globally monitor the disease progression. By reconstructing the clonal phylogeny we demonstrated that *TP53*-mutated clones expanded during disease progression, in particular from the chronic to AP. In addition, they showed a slight decrease from AP to BP (Figure 34).
MPN-AP is considered the transition phase between chronic phase and MPN-BP. [258]. AP is defined as a condition showing 10%- 19% myeloid blasts in PB or BM [263]. MPN-AP has a poor prognosis and a high risk for progression to MPN-BP [264].

Risk factors for leukemic transformation include rising blast counts, acquisition of HMR mutations (*ASXL1*, *IDH1/2*, *SRSF2*, *EZH2*, *U2AF1Q157*), karyotype alterations (e.g 17p deletion and gain of 1q), type 1 *CALR*-unmutated status, severe anemia and high-grade bone marrow fibrosis [258]. The genetic alteration of *TP53* is more frequently associated with MPN-BP (27%) than with chronic phase (3%). The genetic lesion might involve both *TP53* alleles. In particulat, the acquired uniparental disomy of chromosome 17p have been described as being correlated to homozygous mutations [265]. The *TP53* mutation is correlated with shorter overall survival, increased circulating blasts and cytogenetic abnormalities in MPNs [266]. The changes of the *TP53* VAF have been described as associated with AML evolution [267]. Comparing the clonal expansion and the mutation frequency of *TP53* and *FLT3* variants, we suggested that *TP53* mutations could more strictly promote the AP while its mutation was not sufficient to support leukemic transformation. On one hand *FLT3*-mutated clones showed a linear expansion during disease progression, on the other the greater growth of *TP53*-mutated clones was observed from chronic to AP. In this way, our study contributes to distinguish the molecular events that guide AP from genetic alterations related to BP.

Kubesova *et al.* demonstrated that *TP53* mutation is associated with old age, it may be undetectable at diagnosis and slowly increase during disease course. In this way, the mutations may persist at low levels for years without an immediate risk of progression [267]. Our results could be in accordance with this work, given that in our patient *TP53* does not seem directly responsible for leukemic transformation.

In a recent study it was demonstrated that *TP53* impairment determines CNVs, which is correlated with poor prognosis in MDS patients [268]. Therefore, we deeply analyzed *TP53* mutational state at sc-level, identifying a novel *TP53* deletion (Figure 40). Moreover, this mutation was found in co-occurrence with the *TP53* p.Cys238Ser homozygous mutation. Although a small cell subpopulation carried this variant, we suggested that it could support the genomic instability that characterizes the AP. This result represents a further contribution of our understanding of the impact of *TP53* mutation on prognosis.

In addition, our sc-approach is helpful to obtain important information about clonal heterogeneity and the changes in the clonal architecture during disease progression. According to the sc-study of Mylonas *et al.*, we observed a significant increase in the clonal heterogeneity during disease evolution by considering the zygosity of the different mutations in each time point (Figure 35-37). Interestingly, the number of mutated cells carrying *TP53* and *FLT3* genetic alterations in heterozygous configuration increases. Moreover, we detected the *FLT3* homozygous mutation only in the BP (Figure 38 and Table 11 and 12). Moreover, we demonstrated that VAFs of *TP53* and *FLT3* are the most modulated during disease progression.

Finally, the zygosity information, provided by sc-sequencing, allowed for the first time the characterization of *SRSF2* homozygous mutation (Figure 41 and 42). *SRSF2* is frequently mutated in AML derived from MPNs, although the presence of *SRSF2* homozygous mutation has not been still well characterized in MPN patients [270]. Homozygous germline *SRSF2* knockout (KO) are embryonically lethal in mice and the downregulation in mouse embryonic fibroblasts results in G2/M cell cycle arrest and genomic instability. In addition, *SRSF2*-deficient fetal liver cells showed significantly enhanced apoptosis and decreased levels of hematopoietic stem/progenitor cells [271, 272]. Spliceosome components are recurrently mutated in MDS, CMML, CLL and AML. These mutations are described as heterozygous point mutations and mutually exclusive, suggesting a limit on cellular tolerance of disrupted spliceosome function [273]. Despite in our patient the VAF of SRSF2 P95H in the diagnostic NGS was 50%, suggesting an heterozygous mutation, sc-approach unraveled the presence of a homozygous mutation. Only sc-analysis could provide detailed zygosity information, individuating for each cell the correct allelic asset and distinguishing precisely the wild-type, heterozygous or homozygous state.

In order to validate our data, we performed a clonogenic assay and genotyping of colonies to study the zygosity of *SRSF2* mutation The data confirmed the presence of homozygous-

mutated colonies (Figure 42 e 43), but, at the same time, highlighted the biases introduced by the colony approach. Therefore, to prevent the clone selection effect, we propose sc-genomic analysis as an accurate method to determine the zygosity of mutations. Our data also demonstrated that sc-technology could contribute to the discovery of homozygous mutations in genes described as mutated only in heterozygosity.

6. CONCLUSION AND FUTURE PROSPECTIVES

In order to perform single-cell experiments, it is important to pay attention to robustness and reproducibility of the protocol, since the biological material used in these procedures is often precious or not recoverable.

Our study has developed a protocol for sc-genomics analysis able to individuate mutations undetectable by bulk NGS analysis and to provide crucial information about the clonal complexity and evolution of hematological malignancies.

Sc-approach could bypass the colony-assays, overcoming their bias and errors, restoring an image of clonal architecture more representative of patient cellular heterogeneity.

Sc-approach unmasks clonal heterogeneity and can reveal early molecular markers related to disease evolution and poor prognosis, thus potentially contributing to patient risk stratification.

In addition, the sc-genomics describes the allele distribution in a cell population, understanding the cell zygosity and highlighting the presence of rare or novel genetic variants. Therefore, our work suggests that sc-genomics is a powerful instrument with a possible predictive value of disease evolution.

Recently, droplet-based sc-genomics strategies were developed using targeted NGS at single cell level, such as the Tapestri platform (Mission Bio). By employing an automated strategy, the experimental procedure is certainly faster. In addition, despite the actual clinical applicability of single cell analysis is expensive, an automated single cell technology could be suitable in the future in the clinical practice.

Single cell genomics could allow a patient personalized approach, permitting a better prognostic evaluation, addressing target therapies and finally, monitoring the response to treatments and the tumor relapse.

7. BIBLIOGRAPHY

1. Mead AJ, Mullally A. Myeloproliferative neoplasm stem cells. Blood. 2017 Mar 23;129(12):1607-1616. doi: 10.1182/blood-2016-10-696005. Epub 2017 Feb 3. PMID: 28159736; PMCID: PMC5413298.
2. William Dameshek, Some Speculations on the Myeloproliferative Syndromes, Blood, Volume 6, Issue 4, 1951, Pages 372-375, ISSN 0006-4971,
3. Nowell PC, Hungerford DA. Chromosome studies on normal and leukemic human leukocytes. J Natl Cancer Inst. 1960;25:85-109.
4. Tefferi A, Gilliland G. Classification of chronic myeloid disorders: from Dameshek towards a semi-molecular system. Best Pract Res Clin Haematol. 2006;19(3):365-385.
5. Vardiman JW, Thiele J, Arber DA, et al. The 2008 revision of the World Health Organization (WHO) classification of myeloid neoplasms and acute leukemia: rationale and important changes. Blood. 2009;114(5):937-951.
6. 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.
7. Rumi E, Cazzola M. Diagnosis, risk stratification, and response evaluation in classical myeloproliferative neoplasms. Blood. 2017;129(6):680-692
8. Delhommeau F, Pisani DF, James C, Casadevall N, Constantinescu S, Vainchenker W. Oncogenic mechanisms in myeloproliferative disorders. Cell Mol Life Sci. 2006;63(24):2939-2953.
9. Spivak JL. Myeloproliferative Neoplasms. N Engl J Med. 2017;376(22):2168-2181.
10. Campbell PJ, Scott LM, Buck G, et al. Definition of subtypes of essential thrombocythaemia and relation to polycythaemia vera based on JAK2 V617F mutation status: a prospective study. Lancet. 2005;366(9501):1945-1953.
11. Vannucchi AM, Guglielmelli P, Tefferi A. Advances in understanding and management of myeloproliferative neoplasms. CA Cancer J Clin. 2009;59(3):171-191.
12. Campbell PJ, Green AR. The myeloproliferative disorders. N Engl J Med. 2006 Dec 7;355(23):2452-66. doi: 10.1056/NEJMra063728. PMID: 17151367.
13. Axelrad AA, Eskinazi D, Correa PN, Amato D. Hypersensitivity of circulating progenitor cells to megakaryocyte growth and development factor (PEG-rHu MGDF) in essential thrombocythemia. Blood. 2000;96(10):3310-3321.

14. Casadevall N, Vainchenker W, Lacombe C, et al. Erythroid progenitors in polycythemia vera: demonstration of their hypersensitivity to erythropoietin using serum free cultures. Blood. 1982;59(2):447-451.
15. Lee-Six H, Kent DG. Tracking hematopoietic stem cells and their progeny using whole-genome sequencing. Exp Hematol. 2020;83:12-24. doi:10.1016/j.exphem.2020.01.004
16. Bowman RL, Busque L, Levine RL. Clonal Hematopoiesis and Evolution to Hematopoietic Malignancies. Cell Stem Cell. 2018 Feb 1;22(2):157-170. doi: 10.1016/j.stem.2018.01.011. PMID: 29395053; PMCID: PMC5804896
17. Nangalia J, Green AR. Myeloproliferative neoplasms: from origins to outcomes. Hematology Am Soc Hematol Educ Program. 2017;2017(1):470-479. doi:10.1182/asheducation-2017.1.470
18. Silver AJ, Bick AG, Savona MR. Germline risk of clonal haematopoiesis. Nat Rev Genet. 2021 Sep;22(9):603-617. doi: 10.1038/s41576-021-00356-6. Epub 2021 May 13. Erratum in: Nat Rev Genet. 2021 Jun 17;: PMID: 33986496; PMCID: PMC8117131.
19. Levine RL, Wadleigh M, Cools J, et al. Activating mutation in the tyrosine kinase JAK2 in polycythemia vera, essential thrombocythemia, and myeloid metaplasia with myelofibrosis. Cancer Cell. 2005;7(4):387-397.
20. Pikman Y, Lee BH, Mercher T, et al. MPLW515L is a novel somatic activating mutation in myelofibrosis with myeloid metaplasia. PLoS Med. 2006;3(7):e270.
21. Nangalia J, Massie CE, Baxter EJ, et al. Somatic CALR mutations in myeloproliferative neoplasms with nonmutated JAK2. N Engl J Med. 2013;369(25):2391-2405.
22. Moulard O, Mehta J, Fryzek J, Olivares R, Iqbal U, Mesa RA. Epidemiology of myelofibrosis, essential thrombocythemia, and polycythemia vera in the European Union. Eur J Haematol. 2014;92(4):289-297.
23. Geyer HL, Kosiorek H, Dueck AC, et al. Associations between gender, disease features and symptom burden in patients with myeloproliferative neoplasms: an analysis by the MPN QOL International Working Group. Haematologica. 2017;102(1):85-93.
24. Rory M. Shallis, Amer M. Zeidan, Rong Wang, Nikolai A. Podoltsev, Epidemiology of the Philadelphia Chromosome-Negative Classical Myeloproliferative Neoplasms, Hematology/Oncology Clinics of North America, Volume 35, Issue 2, 2021, Pages 177-189
25. Ianotto JC, Curto-Garcia N, Lauermanova M, Radia D, Kiladjian JJ, Harrison CN. Characteristics and outcomes of patients with essential thrombocythemia or

polycythemia vera diagnosed before 20 years of age: a systematic review. Haematologica. 2019;104(8):1580-1588. doi:10.3324/haematol.2018.200832

26. Bilgrami S, Greenberg BR. Polycythemia rubra vera. Semin Oncol. 1995;22(4):307-326.
27. McMullin MF, Wilkins BS, Harrison CN. Management of polycythaemia vera: a critical review of current data. Br J Haematol. 2016 Feb;172(3):337-49. doi: 10.1111/bjh.13812. Epub 2015 Oct 22. PMID: 26492433.
28. Prchal JF, Axelrad AA. Letter: Bone-marrow responses in polycythemia vera. N Engl J Med. 1974;290(24):1382.
29. Green AR. Pathogenesis of polycythaemia vera. Lancet. 1996;347(9005):844-845.
30. De Wolf JT, Beentjes JA, Esselink MT, et al. In polycythemia vera human interleukin 3 and granulocyte-macrophage colony-stimulating factor enhance erythroid colony growth in the absence of erythropoietin. Exp Hematol. 1989;17(9):981-983.
31. Dai CH, Krantz SB, Dessypris EN, Means RT, Jr., Horn ST, Gilbert HS. Polycythemia vera. II. Hypersensitivity of bone marrow erythroid, granulocyte-macrophage, and megakaryocyte progenitor cells to interleukin-3 and granulocyte-macrophage colony-stimulating factor. Blood. 1992;80(4):891-899.
32. Ugo V, Marzac C, Teyssandier I, et al. Multiple signaling pathways are involved in erythropoietin-independent differentiation of erythroid progenitors in polycythemia vera. Exp Hematol. 2004;32(2):179-187.
33. Hess G, Rose P, Gamm H, Papadileris S, Huber C, Seliger B. Molecular analysis of the erythropoietin receptor system in patients with polycythaemia vera. Br J Haematol. 1994;88(4):794-802.
34. James C, Ugo V, Le Couedic JP, et al. A unique clonal JAK2 mutation leading to constitutive signalling causes polycythaemia vera. Nature. 2005;434(7037):1144-1148.
35. Baxter EJ, Scott LM, Campbell PJ, East C, Fourouclas N, Swanton S, Vassiliou GS, Bench AJ, Boyd EM, Curtin N, Scott MA, Erber WN, Green AR; Cancer Genome Project. Acquired mutation of the tyrosine kinase JAK2 in human myeloproliferative disorders. Lancet. 2005 Mar 19-25;365(9464):1054-61. doi: 10.1016/S0140-6736(05)71142-9. Erratum in: Lancet. 2005 Jul 9-15;366(9480):122. PMID: 15781101.
36. Tefferi A. Polycythemia vera and essential thrombocythemia: 2012 update on diagnosis, risk stratification, and management. Am J Hematol. 2012 Mar;87(3):285-93. doi: 10.1002/ajh.23135. PMID: 22331582

37. Tefferi A, Barbui T. Polycythemia vera and essential thrombocythemia: 2017 update on diagnosis, risk-stratification, and management. Am J Hematol. 2017;92(1):94-108.
38. Marchioli R, Finazzi G, Specchia G, et al. Cardiovascular events and intensity of treatment in polycythemia vera. N Engl J Med. 2013;368(1):22-33.
39. Tefferi A, Guglielmelli P, Larson DR, et al. Long-term survival and blast transformation in molecularly annotated essential thrombocythemia, polycythemia vera, and myelofibrosis. Blood. 2014;124(16):2507-2513
40. Vannucchi AM, Antonioli E, Guglielmelli P, et al. Clinical profile of homozygous JAK2 617V>F mutation in patients with polycythemia vera or essential thrombocythemia. Blood. 2007;110(3):840-846.
41. Landolfi R, Marchioli R, Kutti J, Gisslinger H, Tognoni G, Patrono C, Barbui T; European Collaboration on Low-Dose Aspirin in Polycythemia Vera Investigators. Efficacy and safety of low-dose aspirin in polycythemia vera. N Engl J Med. 2004 Jan 8;350(2):114-24. doi: 10.1056/NEJMoa035572. PMID: 14711910.
42. Cortelazzo S, Finazzi G, Ruggeri M, Vestri O, Galli M, Rodeghiero F, Barbui T. Hydroxyurea for patients with essential thrombocythemia and a high risk of thrombosis. N Engl J Med. 1995 Apr 27;332(17):1132-6. doi: 10.1056/NEJM199504273321704. PMID: 7700286.
43. Antonioli E, Guglielmelli P, Pieri L, Finazzi M, Rumi E, Martinelli V, Vianelli N, Luigia Randi M, Bertozzi I, De Stefano V, Za T, Rossi E, Ruggeri M, Elli E, Cacciola R, Cacciola E, Pogliani E, Rodeghiero F, Baccarani M, Passamonti F, Finazzi G, Rambaldi A, Bosi A, Cazzola M, Barbui T, Vannucchi AM; AGIMM Investigators. Hydroxyurea-related toxicity in 3,411 patients with Ph'-negative MPN. Am J Hematol. 2012 May;87(5):552-4. doi: 10.1002/ajh.23160. Epub 2012 Apr 4. PMID: 22473827.
44. Barbui T, Thiele J, Passamonti F, et al. Survival and disease progression in essential thrombocythemia are significantly influenced by accurate morphologic diagnosis: an international study. *J Clin Oncol*. 2011;29(23):3179-3184.
45. Srour SA, Devesa SS, Morton LM, Check DP, Curtis RE, Linet MS, Dores GM. Incidence and patient survival of myeloproliferative neoplasms and myelodysplastic/myeloproliferative neoplasms in the United States, 2001-12. Br J Haematol. 2016 Aug;174(3):382-96
46. Moulard O, Mehta J, Fryzek J, Olivares R, Iqbal U, Mesa RA. Epidemiology of myelofibrosis, essential thrombocythemia, and polycythemia vera in the European Union. *Eur J Haematol*. 2014;92(4):289-297.

47. Elisa Rumi, Mario Cazzola; How I treat essential thrombocythemia. *Blood* 2016; 128 (20): 2403–2414. doi: https://doi.org/10.1182/blood-2016-05-643346
48. Fialkow PJ, Faguet GB, Jacobson RJ, Vaidya K, Murphy S. Evidence that essential thrombocythemia is a clonal disorder with origin in a multipotent stem cell. *Blood*. 1981;58(5):916-919
49. Schafer AI. Thrombocytosis and thrombocythemia. *Blood Rev*. 2001;15(4):159-166.
50. Thiele J, Orazi A, Kvasnicka HM, et al. European Bone Marrow Working Group trial on reproducibility of World Health Organization criteria to discriminate essential thrombocythemia from prefibrotic primary myelofibrosis. Haematologica 2012;97(3):360-5--comment. *Haematologica*. 2012;97(3):e5-6; discussion e7-8.
51. Thiele J, Kvasnicka HM, Vardiman JW, et al. Bone marrow fibrosis and diagnosis of essential thrombocythemia. *J Clin Oncol*. 2009;27(34):e220-221; author reply e222-223.
52. Schalling, M., Gleiss, A., Gisslinger, B. *et al.* Essential thrombocythemia vs. pre-fibrotic/early primary myelofibrosis: discrimination by laboratory and clinical data. *Blood Cancer Journal* 7, 643 (2017). https://doi.org/10.1038/s41408-017-0006-y
53. Barosi G, Rosti V, Bonetti E, et al. Evidence that prefibrotic myelofibrosis is aligned along a clinical and biological continuum featuring primary myelofibrosis. *PLoS One*. 2012;7(4):e35631.
54. Brousseau M, Parot-Schinkel E, Moles M-P, Boyer F, Hunault M & Rousselet M-C, Practical application and clinical impact of the WHO histopathological criteria on bone marrow biopsy for the diagnosis of essential thrombocythemia versus prefibrotic primary myelofibrosis. Histopathology. 2010;56, 758–767
55. Rumi E, Boveri E, Bellini M, et al. Clinical course and outcome of essential thrombocythemia and prefibrotic myelofibrosis according to the revised WHO 2016 diagnostic criteria. *Oncotarget*. 2017;8(60):101735-101744. Published 2017 Oct 6. doi:10.18632/oncotarget.21594
56. Balduini CL, Bertolino G, Noris P, Piletta GC. Platelet aggregation in platelet-rich plasma and whole blood in 120 patients with myeloproliferative disorders. Am J Clin Pathol. 1991 Jan;95(1):82-6. doi: 10.1093/ajcp/95.1.82. PMID: 1987755.
57. Wehmeier A, Südhoff T, Meierkord F. Relation of platelet abnormalities to thrombosis and hemorrhage in chronic myeloproliferative disorders. Semin Thromb Hemost. 1997;23(4):391-402. doi: 10.1055/s-2007-996114. PMID: 9263357.

58. Alessandra Carobbio, Juergen Thiele, Francesco Passamonti, Elisa Rumi, Marco Ruggeri, Francesco Rodeghiero, Maria Luigia Randi, Irene Bertozzi, Alessandro M. Vannucchi, Elisabetta Antonioli, Heinz Gisslinger, Veronika Buxhofer-Ausch, Guido Finazzi, Naseema Gangat, Ayalew Tefferi, Tiziano Barbui; Risk factors for arterial and venous thrombosis in WHO-defined essential thrombocythemia: an international study of 891 patients. *Blood* 2011; 117 (22): 5857–5859. doi: https://doi.org/10.1182/blood-2011-02-339002
59. Pérez Encinas MM, Sobas M, Gómez-Casares MT, Abuin Blanco A, Noya Pereira MS, Raya JM, Andrade-Campos MM, Álvarez Larrán A, Lewandowski K, Łukasz S, Hernández Boluda JC, Ferrer-Marín F, Fox ML, Gołos A, Gasior Kabat M, Magro Mazo E, Czyż A, Martín Martín A, Bellosillo Paricio B, Quinteiro García C, González Martín JM, Stuckey R. The risk of thrombosis in essential thrombocythemia is associated with the type of CALR mutation: A multicentre collaborative study. Eur J Haematol. 2021 Mar;106(3):371-379. doi: 10.1111/ejh.13561. Epub 2020 Dec 30. PMID: 33275803.
60. Campbell PJ, Green AR. Management of polycythemia vera and essential thrombocythemia. Hematology Am Soc Hematol Educ Program. 2005:201-8. doi: 10.1182/asheducation-2005.1.201. PMID: 16304381.
61. Michiels JJ, Berneman Z, Van Bockstaele D, van der Planken M, De Raeve H, Schroyens W. Clinical and laboratory features, pathobiology of platelet-mediated thrombosis and bleeding complications, and the molecular etiology of essential thrombocythemia and polycythemia vera: therapeutic implications. Semin Thromb Hemost. 2006 Apr;32(3):174-207. doi: 10.1055/s-2006-939431. PMID: 16673274.
62. Scherber R, Dueck AC, Johansson P, Barbui T, Barosi G, Vannucchi AM, Passamonti F, Andreasson B, Ferrari ML, Rambaldi A, Samuelsson J, Birgegard G, Tefferi A, Harrison CN, Radia D, Mesa RA. The Myeloproliferative Neoplasm Symptom Assessment Form (MPN-SAF): international prospective validation and reliability trial in 402 patients. Blood. 2011 Jul 14;118(2):401-8. doi: 10.1182/blood-2011-01-328955. Epub 2011 May 2. PMID: 21536863.
63. Rottenstreich A, Kleinstern G, Krichevsky S, Varon D, Lavie D, Kalish Y. Factors related to the development of acquired von Willebrand syndrome in patients with essential thrombocythemia and polycythemia vera. Eur J Intern Med. 2017 Jun;41:49-54. doi: 10.1016/j.ejim.2016.11.011. Epub 2016 Dec 2. PMID: 27919526.

64. Awada H, Voso MT, Guglielmelli P, Gurnari C. Essential Thrombocythemia and Acquired von Willebrand Syndrome: The Shadowlands between Thrombosis and Bleeding. *Cancers (Basel)*. 2020;12(7):1746. Published 2020 Jun 30. doi:10.3390/cancers12071746
65. Andriani A, Latagliata R, Anaclerico B, Spadea A, Rago A, Di Veroli A, Spirito F, Porrini R, De Muro M, Crescenzi Leonetti S, Villivà N, De Gregoris C, Montefusco E, Polverelli N, Santoro C, Breccia M, Cimino G, Majolino I, Mazzucconi MG, Vianelli N, Alimena G, Montanaro M, Palandri F. Spleen enlargement is a risk factor for thrombosis in essential thrombocythemia: Evaluation on 1,297 patients. Am J Hematol. 2016 Mar;91(3):318-21. doi: 10.1002/ajh.24269. PMID: 26748894.
66. Accurso V, Santoro M, Raso S, Contrino AD, Casimiro P, Piazza FD, Perez A, Russo A, Siragusa S. Splenomegaly impacts prognosis in essential thrombocythemia and polycythemia vera: A single center study. Hematol Rep. 2019 Dec 4;11(4):8281. doi: 10.4081/hr.2019.8281. PMID: 31871612; PMCID: PMC6902304.
67. Barbui T, Finazzi G, Carobbio A, et al. Development and validation of an International Prognostic Score of thrombosis in World Health Organization-essential thrombocythemia (IPSET-thrombosis). *Blood*. 2012;120(26):5128-5133;
68. Cerquozzi, S., Tefferi, A. Blast transformation and fibrotic progression in polycythemia vera and essential thrombocythemia: a literature review of incidence and risk factors. *Blood Cancer Journal* 5, e366 (2015). https://doi.org/10.1038/bcj.2015.95
69. Luque Paz D, Jouanneau-Courville R, Riou J, Ianotto JC, Boyer F, Chauveau A, Renard M, Chomel JC, Cayssials E, Gallego-Hernanz MP, Pastoret C, Murati A, Courtier F, Rousselet MC, Quintin-Roué I, Cottin L, Orvain C, Thépot S, Chrétien JM, Delneste Y, Ifrah N, Blanchet O, Hunault-Berger M, Lippert E, Ugo V. Leukemic evolution of polycythemia vera and essential thrombocythemia: genomic profiles predict time to transformation. Blood Adv. 2020 Oct 13;4(19):4887-4897. doi: 10.1182/bloodadvances.2020002271. Erratum in: Blood Adv. 2020 Nov 24;4(22):5651. PMID: 33035330; PMCID: PMC7556129.
70. Finazzi G, Harrison C. Essential thrombocythemia. *Semin Hematol*. 2005;42(4):230-238
71. Rumi E, Cazzola M. How I treat essential thrombocythemia. Blood. 2016 Nov 17;128(20):2403-2414. doi: 10.1182/blood-2016-05-643346. Epub 2016 Aug 25. PMID: 27561316.

72. Tefferi, A., Vannucchi, A.M. & Barbui, T. Essential thrombocythemia treatment algorithm 2018. *Blood Cancer Journal* 8, 2 (2018). https://doi.org/10.1038/s41408-017-0041-8
73. Shallis RM, Zeidan AM, Wang R, Podoltsev NA. Epidemiology of the Philadelphia Chromosome-Negative Classical Myeloproliferative Neoplasms. Hematol Oncol Clin North Am. 2021 Apr;35(2):177-189. doi: 10.1016/j.hoc.2020.11.005. Epub 2021 Jan 4. PMID: 33641862.
74. DeLario MR, Sheehan AM, Ataya R, Bertuch AA, Vega C 2nd, Webb CR, Lopez-Terrada D, Venkateswaran L. Clinical, histopathologic, and genetic features of pediatric primary myelofibrosis--an entity different from adults. Am J Hematol. 2012 May;87(5):461-4. doi: 10.1002/ajh.23140. Epub 2012 Mar 3. PMID: 22389089.
75. Hofmann I. Myeloproliferative Neoplasms in Children. *J Hematop*. 2015;8(3):143-157. doi:10.1007/s12308-015-0256-1
76. Tefferi, A. Primary myelofibrosis: 2021 update on diagnosis, risk-stratification and management. *Am J Hematol*. 2021; 96: 145– 162. https://doi.org/10.1002/ajh.26050
77. Melo-Cardenas J, Migliaccio AR, Crispino JD. The Role of Megakaryocytes in Myelofibrosis. Hematol Oncol Clin North Am. 2021 Apr;35(2):191-203. doi: 10.1016/j.hoc.2020.11.004. Epub 2021 Jan 11. PMID: 33641863; PMCID: PMC7921544.
78. Zahr AA, Salama ME, Carreau N, et al. Bone marrow fibrosis in myelofibrosis: pathogenesis, prognosis and targeted strategies. *Haematologica*. 2016;101(6):660-671. doi:10.3324/haematol.2015.141283
79. Tefferi A. Primary myelofibrosis: 2012 update on diagnosis, risk stratification, and management. *Am J Hematol*. 2011;86(12):1017-1026.
80. Mudireddy M, Gangat N, Hanson CA, Ketterling RP, Pardanani A, Tefferi A. Validation of the WHO-defined 20% circulating blasts threshold for diagnosis of leukemic transformation in primary myelofibrosis. *Blood Cancer J*. 2018;8(6):57. Published 2018 Jun 11. doi:10.1038/s41408-018-0095-2
81. Tefferi A. Primary myelofibrosis: 2017 update on diagnosis, risk-stratification, and management. *Am J Hematol*. 2016;91(12):1262-1271.
82. Boiocchi, L., Espinal-Witter, R., Geyer, J. *et al.* Development of monocytosis in patients with primary myelofibrosis indicates an accelerated phase of the disease. *Mod Pathol* 26, 204–212 (2013). https://doi.org/10.1038/modpathol.2012.165

83. Julia T. Geyer, Elizabeth Margolskee, Spencer A. Krichevsky, Daniele Cattaneo, Leonardo Boiocchi, Paola Ronchi, Francesca Lunghi, Joseph M. Scandura, Maurilio Ponzoni, Robert P. Hasserjian, Umberto Gianelli, Alessandra Iurlo, Attilio Orazi. Disease progression in myeloproliferative neoplasms: comparing patients in accelerated phase with those in chronic phase with increased blasts (<10%) or with other types of disease progression. Haematologica 2020;105(5):e221-e224; https://doi.org/10.3324/haematol.2019.230193
84. Koschmieder, S., Mughal, T., Hasselbalch, H. *et al.* Myeloproliferative neoplasms and inflammation: whether to target the malignant clone or the inflammatory process or both. *Leukemia* 30, 1018–1024 (2016). https://doi.org/10.1038/leu.2016.12
85. Lussana F, Rambaldi A. Inflammation and myeloproliferative neoplasms. J Autoimmun. 2017 Dec;85:58-63. doi: 10.1016/j.jaut.2017.06.010. Epub 2017 Jun 30. PMID: 28669446.
86. Kiem D, Wagner S, Magnes T, Egle A, Greil R, Melchardt T. The Role of Neutrophilic Granulocytes in Philadelphia Chromosome Negative Myeloproliferative Neoplasms. *Int J Mol Sci*. 2021;22(17):9555. Published 2021 Sep 3. doi:10.3390/ijms22179555
87. Ciurea SO, Merchant D, Mahmud N, et al. Pivotal contributions of megakaryocytes to the biology of idiopathic myelofibrosis. *Blood*. 2007;110(3):986-993. doi:10.1182/blood-2006-12-064626
88. Balduini A, Badalucco S, Pugliano MT, et al. In vitro megakaryocyte differentiation and proplatelet formation in Ph-negative classical myeloproliferative neoplasms: distinct patterns in the different clinical phenotypes. *PLoS One*. 2011;6(6):e21015.
89. Thiele J, Kvasnicka HM, Vardiman J. Bone marrow histopathology in the diagnosis of chronic myeloproliferative disorders: a forgotten pearl. *Best Pract Res Clin Haematol*. 2006;19(3):413-437.
90. Agarwal A, Morrone K, Bartenstein M, Zhao ZJ, Verma A, Goel S. Bone marrow fibrosis in primary myelofibrosis: Pathogenic mechanisms and the role of TGF-β. Stem Cell Investigation. 2016 Feb 1;2016(FEB). 5. https://doi.org/10.3978/j.issn.2306-9759.2016.02.03
91. Badalucco S, Di Buduo CA, Campanelli R, Pallotta I, Catarsi P, Rosti V, Kaplan DL, Barosi G, Massa M, Balduini A. Involvement of TGFβ1 in autocrine regulation of proplatelet formation in healthy subjects and patients with primary myelofibrosis. Haematologica. 2013 Apr;98(4):514-7. doi: 10.3324/haematol.2012.076752. Epub 2013 Feb 12. PMID: 23403314; PMCID: PMC3659980.

92. Le Bousse-Kerdilès MC. Primary myelofibrosis and the "bad seeds in bad soil" concept. Fibrogenesis Tissue Repair. 2012 Jun 6;5(Suppl 1):S20. doi: 10.1186/1755-1536-5-S1-S20. PMID: 23259918; PMCID: PMC3368798.
93. Leiva, O., Ng, S., Chitalia, S. *et al.* The role of the extracellular matrix in primary myelofibrosis. *Blood Cancer Journal* 7, e525 (2017). https://doi.org/10.1038/bcj.2017.6
94. Guglielmelli P, Rotunno G, Pacilli A, et al. Prognostic impact of bone marrow fibrosis in primary myelofibrosis. A study of the AGIMM group on 490 patients. *Am J Hematol*. 2016;91(9):918-922.
95. Spampinato, M.; Giallongo, C.; Romano, A.; Longhitano, L.; La Spina, E.; Avola, R.; Scandura, G.; Dulcamare, I.; Bramanti, V.; Di Rosa, M.; Vicario, N.; Parenti, R.; Li Volti, G.; Tibullo, D.; Palumbo, G.A. Focus on Osteosclerotic Progression in Primary Myelofibrosis. *Biomolecules* 2021, *11*, 122. https://doi.org/10.3390/biom11010122
96. Mellibovsky L, Mariñoso ML, Cervantes F, Besses C, Nacher M, Nogués X, Florensa L, Munné A, Diez-Perez A, Serrano S. Relationship among densitometry, bone histomorphometry, and histologic stage in idiopathic myelofibrosis. Bone. 2004 Feb;34(2):330-5. doi: 10.1016/j.bone.2003.10.010. PMID: 14962811.
97. Boyce B. Advances in the Regulation of Osteoclasts and Osteoclast Functions. *J Dent Res* (2013) 92:860–7. doi: 10.1177/0022034513500306
98. Bock O, Loch G, Schade U, Büsche G, Wasielewski R, Wiese B, Kreipe H. Osteosclerosis in advanced chronic idiopathic myelofibrosis is associated with endothelial overexpression of osteoprotegerin. Br J Haematol. 2005 Jul;130(1):76-82. doi: 10.1111/j.1365-2141.2005.05573.x. PMID: 15982347.
99. Piotr Centkowski, Joanna Sawczuk-Chabin, Monika Prochorec-Sobieszek, Joanna Dobrzanska, Ilona Seferynska, Iwona Solarska, Miroslaw Majewski, Krzysztof Warzocha; Bone Marrow Fibrosis in Chronic Idiopathic Myelofibrosis Is Associated with Megakaryocyte Expression of Osteoprotegerin.. *Blood* 2005; 106 (11): 4955. doi: https://doi.org/10.1182/blood.V106.11.4955.4955
100. Lampiasi N, Russo R, Zito F. The Alternative Faces of Macrophage Generate Osteoclasts. *Biomed Res Int.* 2016;2016:9089610. doi:10.1155/2016/9089610
101. Veletic I, Manshouri T, Multani AS, Yin CC, Chen L, Verstovsek S, Estrov Z. Myelofibrosis osteoclasts are clonal and functionally impaired. Blood. 2019 May 23;133(21):2320-2324. doi: 10.1182/blood-2018-10-878926. Epub 2019 Feb 11. PMID: 30745304; PMCID: PMC6533604.

102. Méndez-Ferrer S, Bonnet D, Steensma DP, Hasserjian RP, Ghobrial IM, Gribben JG, Andreeff M, Krause DS. Bone marrow niches in haematological malignancies. Nat Rev Cancer. 2020 May;20(5):285-298. doi: 10.1038/s41568-020-0245-2. Epub 2020 Feb 28. PMID: 32112045.

103. Di Raimondo F, Azzaro MP, Palumbo GA, et al. Elevated vascular endothelial growth factor (VEGF) serum levels in idiopathic myelofibrosis. *Leukemia*. 2001;15(6):976-980.

104. Longhitano L, Li Volti G, Giallongo C, et al. The Role of Inflammation and Inflammasome in Myeloproliferative Disease. *J Clin Med*. 2020;9(8):2334. Published 2020 Jul 22. doi:10.3390/jcm9082334

105. Koschmieder S, Mughal TI, Hasselbalch HC, Barosi G, Valent P, Kiladjian JJ, Jeryczynski G, Gisslinger H, Jutzi JS, Pahl HL, Hehlmann R, Maria Vannucchi A, Cervantes F, Silver RT, Barbui T. Myeloproliferative neoplasms and inflammation: whether to target the malignant clone or the inflammatory process or both. Leukemia. 2016 May;30(5):1018-24. doi: 10.1038/leu.2016.12. Epub 2016 Feb 8. PMID: 26854026.

106. Hasselbalch HC. Perspectives on chronic inflammation in essential thrombocythemia, polycythemia vera, and myelofibrosis: is chronic inflammation a trigger and driver of clonal evolution and development of accelerated atherosclerosis and second cancer? *Blood*. 2012;119(14):3219-3225.

107. Tefferi A, Vaidya R, Caramazza D, Finke C, Lasho T, Pardanani A. Circulating interleukin (IL)-8, IL-2R, IL-12, and IL-15 levels are independently prognostic in primary myelofibrosis: a comprehensive cytokine profiling study. J Clin Oncol. 2011;29(10):1356-1363.

108. Barbui T, Carobbio A, Finazzi G, et al. Elevated C-reactive protein is associated with shortened leukemia-free survival in patients with myelofibrosis. *Leukemia*. 2013;27(10):2084-2086.

109. Verstovsek S, Kantarjian H, Mesa RA, et al. Safety and efficacy of INCB018424, a JAK1 and JAK2 inhibitor, in myelofibrosis. *N Engl J Med*. 2010;363(12):1117-1127.

110. Lussana F, Rambaldi A. Inflammation and myeloproliferative neoplasms. J Autoimmun. 2017 Dec;85:58-63. doi: 10.1016/j.jaut.2017.06.010. Epub 2017 Jun 30. PMID: 28669446.

111. Xu M, Bruno E, Chao J, et al. Constitutive mobilization of $CD34^+$ cells into the peripheral blood in idiopathic myelofibrosis may be due to the action of a number of proteases. *Blood*. 2005;105(11):4508-4515.

112. Janowska-Wieczorek A, Marquez LA, Nabholtz JM, et al. Growth factors and cytokines upregulate gelatinase expression in bone marrow $CD34(^+)$ cells and their transmigration through reconstituted basement membrane. *Blood*. 1999;93(10):3379-3390.

113. Bogani C, Ponziani V, Guglielmelli P, et al. Hypermethylation of CXCR4 promoter in $CD34^+$ cells from patients with primary myelofibrosis. *Stem Cells*. 2008;26(8):1920-1930.

114. Abdel-Wahab OI, Levine RL.. Primary myelofibrosis: update on definition, pathogenesis, and treatment. Annu Rev Med, 2008, 60: 233-245. Annual review of medicine. 60. 233-45. 10.1146/annurev.med.60.041707.160528.

115. Masarova L, Mesa RA, Hernández-Boluda JC, Taylor JA. Severe thrombocytopenia in myelofibrosis is more prevalent than previously reported. Leuk Res. 2020 Apr;91:106338. doi: 10.1016/j.leukres.2020.106338. Epub 2020 Feb 26. PMID: 32200188.

116. Vallapureddy, R.R., Mudireddy, M., Penna, D. *et al.* Leukemic transformation among 1306 patients with primary myelofibrosis: risk factors and development of a predictive model. *Blood Cancer Journal* 9, 12 (2019). https://doi.org/10.1038/s41408-019-0175-y

117. Iurlo A, Cattaneo D, Gianelli U. Blast Transformation in Myeloproliferative Neoplasms: Risk Factors, Biological Findings, and Targeted Therapeutic Options. *Int J Mol Sci*. 2019;20(8):1839. Published 2019 Apr 13. doi:10.3390/ijms20081839

118. Rago A, Latagliata R, Montanaro M, Montefusco E, Andriani A, Crescenzi SL, Mecarocci S, Spirito F, Spadea A, Recine U, Cicconi L, Avvisati G, Cedrone M, Breccia M, Porrini R, Villivà N, De Gregoris C, Alimena G, D'Arcangelo E, Guglielmelli P, Lo-Coco F, Vannucchi A, Cimino G. Hemoglobin levels and circulating blasts are two easily evaluable diagnostic parameters highly predictive of leukemic transformation in primary myelofibrosis. Leuk Res. 2015 Mar;39(3):314-7. doi: 10.1016/j.leukres.2015.01.001. Epub 2015 Jan 12. PMID: 25636356.

119. Cervantes F, Dupriez B, Pereira A, et al. New prognostic scoring system for primary myelofibrosis based on a study of the International Working Group for Myelofibrosis Research and Treatment. *Blood*. 2009;113(13):2895-2901

120. Passamonti F, Cervantes F, Vannucchi AM, et al. A dynamic prognostic model to predict survival in primary myelofibrosis: a study by the IWG-MRT (International Working Group for Myeloproliferative Neoplasms Research and Treatment). Blood. 2010;115(9):1703-1708.

121. Gangat N, Caramazza D, Vaidya R, et al. DIPSS plus: a refined Dynamic International Prognostic Scoring System for primary myelofibrosis that incorporates prognostic information from karyotype, platelet count, and transfusion status. *J Clin Oncol*. 2011;29(4):392-397.

122. Rumi E, Pietra D, Pascutto C, et al. Clinical effect of driver mutations of JAK2, CALR, or MPL in primary myelofibrosis. *Blood*. 2014;124(7):1062-1069. doi:10.1182/blood-2014-05-578435

123. Tefferi A, Lasho TL, Tischer A, et al. The prognostic advantage of calreticulin mutations in myelofibrosis might be confined to type 1 or type 1-like CALR variants. *Blood*. 2014;124(15):2465-2466.

124. Guglielmelli P, Rotunno G, Fanelli T, et al. Validation of the differential prognostic impact of type 1/type 1-like versus type 2/type 2-like CALR mutations in myelofibrosis. *Blood Cancer J*. 2015;5:e360.

125. Vannucchi AM, Lasho TL, Guglielmelli P, et al. Mutations and prognosis in primary myelofibrosis. *Leukemia*. 2013;27(9):1861-1869.

126. Guglielmelli P, Lasho TL, Rotunno G, et al. The number of prognostically detrimental mutations and prognosis in primary myelofibrosis: an international study of 797 patients. *Leukemia*. 2014;28(9):1804-1810.

127. Alessandro M. Vannucchi, Paola Guglielmelli, Terra L. Lasho, Giada Rotunno, Carmela Mannarelli, Annalisa Pacilli, Animesh Pardanani, Elisa Rumi, Vittorio Rosti, Curtis A. Hanson, Francesco Mannelli, Rhett P. Ketterling, Naseema Gangat, Alessandro Rambaldi, Francesco Passamonti, Giovanni Barosi, Tiziano Barbui, Ayalew Tefferi; MIPSS70: Mutation-Enhanced Prognostic System for Transplant Age Patients with Primary Myelofibrosis. *Blood* 2017; 130 (Supplement 1): 200. doi: https://doi.org/10.1182/blood.V130.Suppl_1.200.200

128. Guglielmelli P, Lasho TL, Rotunno G, Mudireddy M, Mannarelli C, Nicolosi M, Pacilli A, Pardanani A, Rumi E, Rosti V, Hanson CA, Mannelli F, Ketterling RP, Gangat N, Rambaldi A, Passamonti F, Barosi G, Barbui T, Cazzola M, Vannucchi AM, Tefferi A. MIPSS70: Mutation-Enhanced International Prognostic Score System for Transplantation-Age Patients With Primary Myelofibrosis. J Clin Oncol. 2018 Feb

1;36(4):310-318. doi: 10.1200/JCO.2017.76.4886. Epub 2017 Dec 9. PMID: 29226763.

129. Tefferi A, Guglielmelli P, Nicolosi M, Mannelli F, Mudireddy M, Bartalucci N, Finke CM, Lasho TL, Hanson CA, Ketterling RP, Begna KH, Naseema Gangat, Pardanani A, Vannucchi AM. GIPSS: genetically inspired prognostic scoring system for primary myelofibrosis. Leukemia. 2018 Jul;32(7):1631-1642. doi: 10.1038/s41375-018-0107-z. Epub 2018 Mar 23. PMID: 29654267; PMCID: PMC6035151.

130. Giralt S, Bishop MR. Principles and overview of allogeneic hematopoietic stem cell transplantation. *Cancer Treat Res*. 2009;144:1-21. doi:10.1007/978-0-387-78580-6_1

131. Babushok D, Hexner E. Allogeneic transplantation for myelofibrosis: for whom, when, and what are the true benefits?. *Curr Opin Hematol*. 2014;21(2):114-122. doi:10.1097/MOH.0000000000000015

132. Tiribelli, M., Palandri, F., Sant'Antonio, E. *et al.* The role of allogeneic stem-cell transplant in myelofibrosis in the era of JAK inhibitors: a case-based review. *Bone Marrow Transplant* 55, 708–716 (2020). https://doi.org/10.1038/s41409-019-0683-1

133. Cervantes F. How I treat splenomegaly in myelofibrosis. *Blood Cancer J*. 2011;1(10):e37. doi:10.1038/bcj.2011.36

134. Martínez-Trillos, A., Gaya, A., Maffioli, M. *et al.* Efficacy and tolerability of hydroxyurea in the treatment of the hyperproliferative manifestations of myelofibrosis: results in 40 patients. *Ann Hematol* 89, 1233–1237 (2010). https://doi.org/10.1007/s00277-010-1019-9

135. Deisseroth A, Kaminskas E, Grillo J, Chen W, Saber H, Lu HL, Rothmann MD, Brar S, Wang J, Garnett C, Bullock J, Burke LB, Rahman A, Sridhara R, Farrell A, Pazdur R. U.S. Food and Drug Administration approval: ruxolitinib for the treatment of patients with intermediate and high-risk myelofibrosis. Clin Cancer Res. 2012 Jun 15;18(12):3212-7. doi: 10.1158/1078-0432.CCR-12-0653. Epub 2012 Apr 27. PMID: 22544377.

136. Bose P, Verstovsek S. JAK Inhibition for the Treatment of Myelofibrosis: Limitations and Future Perspectives. *Hemasphere*. 2020;4(4):e424. Published 2020 Jul 21. doi:10.1097/HS9.0000000000000424

137. O'Sullivan JM, Harrison CN. JAK-STAT signaling in the therapeutic landscape of myeloproliferative neoplasms. Mol Cell Endocrinol. 2017 Aug 15;451:71-79. doi: 10.1016/j.mce.2017.01.050. Epub 2017 Feb 3. PMID: 28167129.

138. Vainchenker W, Kralovics R. Genetic basis and molecular pathophysiology of classical myeloproliferative neoplasms. Blood. 2017 Feb 9;129(6):667-679. doi: 10.1182/blood-2016-10-695940. Epub 2016 Dec 27. PMID: 28028029.

139. Szuber, Natasha; Tefferi, Ayalew Driver mutations in primary myelofibrosis and their implications, Current Opinion in Hematology: March 2018 - Volume 25 - Issue 2 - p 129-135 doi: 10.1097/MOH.0000000000000406

140. Jacob Grinfeld, Jyoti Nangalia, Anthony R. Green. Molecular determinants of pathogenesis and clinical phenotype in myeloproliferative neoplasms. Haematologica 2017;102(1):7-17; https://doi.org/10.3324/haematol.2014.113845.

141. Milosevic Feenstra JD, Nivarthi H, Gisslinger H, et al. Whole-exome sequencing identifies novel MPL and JAK2 mutations in triple-negative myeloproliferative neoplasms. Blood. 2016;127(3):325-332

142. Kralovics R, Passamonti F, Buser AS, Teo SS, Tiedt R, Passweg JR, Tichelli A, Cazzola M, Skoda RC. A gain-of-function mutation of JAK2 in myeloproliferative disorders. N Engl J Med. 2005 Apr 28;352(17):1779-90. doi: 10.1056/NEJMoa051113. PMID: 15858187.

143. Bousoik E, Montazeri Aliabadi H. "Do We Know Jack" About JAK? A Closer Look at JAK/STAT Signaling Pathway. Front Oncol. 2018 Jul 31;8:287. doi: 10.3389/fonc.2018.00287. PMID: 30109213; PMCID: PMC6079274.

144. Silvennoinen O, Hubbard SR. Molecular insights into regulation of JAK2 in myeloproliferative neoplasms. Blood. 2015 May 28;125(22):3388-92. doi: 10.1182/blood-2015-01-621110. Epub 2015 Mar 30. PMID: 25824690; PMCID: PMC4447858.

145. Toms AV, Deshpande A, McNally R, Jeong Y, Rogers JM, Kim CU, Gruner SM, Ficarro SB, Marto JA, Sattler M, Griffin JD, Eck MJ. Structure of a pseudokinase-domain switch that controls oncogenic activation of Jak kinases. Nat Struct Mol Biol. 2013 Oct;20(10):1221-3. doi: 10.1038/nsmb.2673. Epub 2013 Sep 8. PMID: 24013208; PMCID: PMC3863620.

146. Silvennoinen O, Ungureanu D, Niranjan Y, Hammaren H, Bandaranayake R, Hubbard SR. New insights into the structure and function of the pseudokinase domain in JAK2. *Biochem Soc Trans*. 2013;41(4):1002-1007.

147. Vainchenker, W., Constantinescu, S. JAK/STAT signaling in hematological malignancies. *Oncogene* 32, 2601–2613 (2013). https://doi.org/10.1038/onc.2012.347

148. Jones AV, Kreil S, Zoi K, et al. Widespread occurrence of the JAK2 V617F mutation in chronic myeloproliferative disorders. *Blood*. 2005;106(6):2162-2168.

149. Shi S, Calhoun HC, Xia F, Li J, Le L, Li WX. JAK signaling globally counteracts heterochromatic gene silencing. *Nat Genet*. 2006;38(9):1071-1076.

150. Shi S, Larson K, Guo D, et al. Drosophila STAT is required for directly maintaining HP1 localization and heterochromatin stability. *Nat Cell Biol*. 2008;10(4):489-496.

151. Dawson MA, Bannister AJ, Gottgens B, et al. JAK2 phosphorylates histone H3Y41 and excludes HP1alpha from chromatin. *Nature*. 2009;461(7265):819-822.

152. Lacout C, Pisani DF, Tulliez M, Gachelin FM, Vainchenker W, Villeval JL. JAK2V617F expression in murine hematopoietic cells leads to MPD mimicking human PV with secondary myelofibrosis. *Blood*. 2006;108(5):1652-1660.

153. 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;107(11):4274-4281.

154. Zaleskas VM, Krause DS, Lazarides K, et al. Molecular pathogenesis and therapy of polycythemia induced in mice by JAK2 V617F. *PLoS One*. 2006;1:e18

155. Akada H, Akada S, Hutchison RE, Mohi G. Loss of wild-type Jak2 allele enhances myeloid cell expansion and accelerates myelofibrosis in Jak2V617F knock-in mice. Leukemia. 2014 Aug;28(8):1627-35. doi: 10.1038/leu.2014.52. Epub 2014 Jan 31. PMID: 24480985; PMCID: PMC4117831.

156. Vannucchi, A., Antonioli, E., Guglielmelli, P. *et al.* Clinical correlates of *JAK2*V617F presence or allele burden in myeloproliferative neoplasms: a critical reappraisal. *Leukemia* 22, 1299–1307 (2008). https://doi.org/10.1038/leu.2008.113

157. Scott LM, Tong W, Levine RL, Scott MA, Beer PA, Stratton MR, Futreal PA, Erber WN, McMullin MF, Harrison CN, Warren AJ, Gilliland DG, Lodish HF, Green AR. JAK2 exon 12 mutations in polycythemia vera and idiopathic erythrocytosis. N Engl J Med. 2007 Feb 1;356(5):459-68. doi: 10.1056/NEJMoa065202. PMID: 17267906; PMCID: PMC2873834.

158. Scott LM. The JAK2 exon 12 mutations: a comprehensive review. Am J Hematol. 2011 Aug;86(8):668-76. doi: 10.1002/ajh.22063. Epub 2011 Jun 14. PMID: 21674578.

159. Spivak JL, Moliterno AR. The Thrombopoietin Receptor, MPL, Is a Therapeutic Target of Opportunity in the MPN. *Front Oncol*. 2021;11:641613. Published 2021 Mar 10. doi:10.3389/fonc.2021.641613

160. Hitchcock IS, Hafer M, Sangkhae V, Tucker JA. The thrombopoietin receptor: revisiting the master regulator of platelet production. Platelets. 2021 Aug 18;32(6):770-778. doi: 10.1080/09537104.2021.1925102. Epub 2021 Jun 7. PMID: 34097561; PMCID: PMC8292222.

161. Vainchenker W, Plo I, Marty C, Varghese LN, Constantinescu SN. The role of the thrombopoietin receptor MPL in myeloproliferative neoplasms: recent findings and potential therapeutic applications. Expert Rev Hematol. 2019 Jun;12(6):437-448. doi: 10.1080/17474086.2019.1617129. Epub 2019 May 22. PMID: 31092065.

162. Pardanani AD, Levine RL, Lasho T, et al. MPL515 mutations in myeloproliferative and other myeloid disorders: a study of 1182 patients. Blood. 2006;108(10):3472-3476.

163. Beer PA, Campbell PJ, Scott LM, et al. MPL mutations in myeloproliferative disorders: analysis of the PT-1 cohort. *Blood*. 2008;112(1):141-149.

164. Ding J, Komatsu H, Wakita A, et al. Familial essential thrombocythemia associated with a dominant-positive activating mutation of the c-MPL gene, which encodes for the receptor for thrombopoietin. *Blood*. 2004;103(11):4198-4200

165. Fucikova, J., Spisek, R., Kroemer, G. *et al.* Calreticulin and cancer. *Cell Res* 31, 5–16 (2021). https://doi.org/10.1038/s41422-020-0383-9

166. Merlinsky TR, Levine RL, Pronier E. Unfolding the Role of Calreticulin in Myeloproliferative Neoplasm Pathogenesis. Clin Cancer Res. 2019 May 15;25(10):2956-2962. doi: 10.1158/1078-0432.CCR-18-3777. Epub 2019 Jan 17. PMID: 30655313; PMCID: PMC6522317.

167. Klampfl T, Gisslinger H, Harutyunyan AS, et al. Somatic mutations of calreticulin in myeloproliferative neoplasms. *N Engl J Med*. 2013;369(25):2379-2390.

168. Cazzola M, Kralovics R. From Janus kinase 2 to calreticulin: the clinically relevant genomic landscape of myeloproliferative neoplasms. *Blood*. 2014;123(24):3714-3719.

169. How J, Hobbs GS, Mullally A. Mutant calreticulin in myeloproliferative neoplasms. Blood. 2019 Dec 19;134(25):2242-2248. doi: 10.1182/blood.2019000622. PMID: 31562135; PMCID: PMC6923668.

170. Chachoua I, Pecquet C, El-Khoury M, Nivarthi H, Albu RI, Marty C, Gryshkova V, Defour JP, Vertenoeil G, Ngo A, Koay A, Raslova H, Courtoy PJ, Choong ML, Plo I, Vainchenker W, Kralovics R, Constantinescu SN. Thrombopoietin receptor activation by myeloproliferative neoplasm associated calreticulin mutants. Blood. 2016 Mar 10;127(10):1325-35. doi: 10.1182/blood-2015-11-681932. Epub 2015 Dec 14. PMID: 26668133.

171. Prins D, González Arias C, Klampfl T, Grinfeld J, Green AR. Mutant Calreticulin in the Myeloproliferative Neoplasms. Hemasphere. 2020 Jan 15;4(1):e333. doi: 10.1097/HS9.0000000000000333. PMID: 32382708; PMCID: PMC7000472.

172. Giada Rotunno, Carmela Mannarelli, Paola Guglielmelli, Annalisa Pacilli, Alessandro Pancrazzi, Lisa Pieri, Tiziana Fanelli, Alberto Bosi, Alessandro M. Vannucchi; on behalf of the Associazione Italiana per la Ricerca sul Cancro Gruppo Italiano Malattie Mieloproliferative Investigators, Impact of calreticulin mutations on clinical and hematological phenotype and outcome in essential thrombocythemia. Blood 2014; 123 (10): 1552–1555. doi: https://doi.org/10.1182/blood-2013-11-538983

173. Tefferi A, Lasho TL, Finke CM, Knudson RA, Ketterling R, Hanson CH, Maffioli M, Caramazza D, Passamonti F, Pardanani A. CALR vs JAK2 vs MPL-mutated or triple-negative myelofibrosis: clinical, cytogenetic and molecular comparisons. Leukemia. 2014 Jul;28(7):1472-7. doi: 10.1038/leu.2014.3. Epub 2014 Jan 9. PMID: 24402162.

174. Tefferi, A., Lasho, T., Finke, C. *et al.* Type 1 vs type 2 calreticulin mutations in primary myelofibrosis: differences in phenotype and prognostic impact. *Leukemia* 28, 1568–1570 (2014). https://doi.org/10.1038/leu.2014.83

175. Cabagnols, X., Defour, J., Ugo, V. *et al.* Differential association of calreticulin type 1 and type 2 mutations with myelofibrosis and essential thrombocytemia: relevance for disease evolution. *Leukemia* 29, 249–252 (2015). https://doi.org/10.1038/leu.2014.270

176. Benlabiod, C., Cacemiro, M.d.C., Nédélec, A. *et al.* Calreticulin del52 and ins5 knock-in mice recapitulate different myeloproliferative phenotypes observed in patients with MPN. *Nat Commun* 11, 4886 (2020). https://doi.org/10.1038/s41467-020-18691-3

177. Delhommeau F, Dupont S, Della Valle V, et al. Mutation in TET2 in myeloid cancers. *N Engl J Med*. 2009;360(22):2289-2301.

178. Feng Y, Li X, Cassady K, Zou Z, Zhang X. TET2 Function in Hematopoietic Malignancies, Immune Regulation, and DNA Repair. Front Oncol. 2019 Apr 2;9:210. doi: 10.3389/fonc.2019.00210. PMID: 31001476; PMCID: PMC6454012.

179. Solary E, Bernard OA, Tefferi A, Fuks F, Vainchenker W. The Ten-Eleven Translocation-2 (TET2) gene in hematopoiesis and hematopoietic diseases. Leukemia. 2014 Mar;28(3):485-96. doi: 10.1038/leu.2013.337. Epub 2013 Nov 13. PMID: 24220273.

180. Bowman RL, Levine RL. TET2 in Normal and Malignant Hematopoiesis. Cold Spring Harb Perspect Med. 2017 Aug 1;7(8):a026518. doi: 10.1101/cshperspect.a026518. PMID: 28242787; PMCID: PMC5538403.

181. Tefferi A, Pardanani A, Lim KH, Abdel-Wahab O, Lasho TL, Patel J, Gangat N, Finke CM, Schwager S, Mullally A, Li CY, Hanson CA, Mesa R, Bernard O, Delhommeau F, Vainchenker W, Gilliland DG, Levine RL. TET2 mutations and their clinical correlates in polycythemia vera, essential thrombocythemia and myelofibrosis. Leukemia. 2009 May;23(5):905-11. doi: 10.1038/leu.2009.47. Epub 2009 Mar 5. PMID: 19262601; PMCID: PMC4654629.

182. Hussein K, Abdel-Wahab O, Lasho TL, Van Dyke DL, Levine RL, Hanson CA *et al*. Cytogenetic correlates of TET2 mutations in 199 patients with myeloproliferative neoplasms. *Am J Hematol* 2010; 85: 81–83.

183. Tefferi, A., Abdel-Wahab, O., Cervantes, F. *et al.* Mutations with epigenetic effects in myeloproliferative neoplasms and recent progress in treatment: Proceedings from the 5th International Post-ASH Symposium. *Blood Cancer Journal* 1, e7 (2011). https://doi.org/10.1038/bcj.2011.4

184. Jacquelin S, Straube J, Cooper L, Vu T, Song A, Bywater M, Baxter E, Heidecker M, Wackrow B, Porter A, Ling V, Green J, Austin R, Kazakoff S, Waddell N, Hesson LB, Pimanda JE, Stegelmann F, Bullinger L, Döhner K, Rampal RK, Heckl D, Hill GR, Lane SW. Jak2V617F and Dnmt3a loss cooperate to induce myelofibrosis through activated enhancer-driven inflammation. Blood. 2018 Dec 27;132(26):2707-2721. doi: 10.1182/blood-2018-04-846220. Epub 2018 Oct 26. PMID: 30366920.

185. Nangalia J, Nice FL, Wedge DC, Godfrey AL, Grinfeld J, Thakker C, Massie CE, Baxter J, Sewell D, Silber Y, Campbell PJ, Green AR. DNMT3A mutations occur early or late in patients with myeloproliferative neoplasms and mutation order influences phenotype. Haematologica. 2015 Nov;100(11):e438-42. doi:

10.3324/haematol.2015.129510. Epub 2015 Aug 6. PMID: 26250577; PMCID: PMC4825297.

186. Stegelmann F, Bullinger L, Schlenk RF, Paschka P, Griesshammer M, Blersch C, Kuhn S, Schauer S, Döhner H, Döhner K. DNMT3A mutations in myeloproliferative neoplasms. Leukemia. 2011 Jul;25(7):1217-9. doi: 10.1038/leu.2011.77. Epub 2011 May 3. PMID: 21537334.

187. Figueroa ME, Abdel-Wahab O, Lu C, Ward PS, Patel J, Shih A, Li Y, Bhagwat N, Vasanthakumar A, Fernandez HF, Tallman MS, Sun Z, Wolniak K, Peeters JK, Liu W, Choe SE, Fantin VR, Paietta E, Löwenberg B, Licht JD, Godley LA, Delwel R, Valk PJ, Thompson CB, Levine RL, Melnick A. 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. doi: 10.1016/j.ccr.2010.11.015. Epub 2010 Dec 9. PMID: 21130701; PMCID: PMC4105845.

188. Tefferi A, Lasho TL, Abdel-Wahab O, Guglielmelli P, Patel J, Caramazza D, Pieri L, Finke CM, Kilpivaara O, Wadleigh M, Mai M, McClure RF, Gilliland DG, Levine RL, Pardanani A, Vannucchi AM. IDH1 and IDH2 mutation studies in 1473 patients with chronic-, fibrotic- or blast-phase essential thrombocythemia, polycythemia vera or myelofibrosis. Leukemia. 2010 Jul;24(7):1302-9. doi: 10.1038/leu.2010.113. Epub 2010 May 27. PMID: 20508616; PMCID: PMC3035975.

189. Fisher CL, Randazzo F, Humphries RK, Brock HW. Characterization of Asxl1, a murine homolog of Additional sex combs, and analysis of the Asx-like gene family. Gene. 2006 Mar 15;369:109-18. doi: 10.1016/j.gene.2005.10.033. Epub 2006 Jan 18. PMID: 16412590.

190. Abdel-Wahab O, Adli M, LaFave LM, et al. ASXL1 mutations promote myeloid transformation through loss of PRC2-mediated gene repression. *Cancer Cell*. 2012;22(2):180-193

191. Tefferi A, Guglielmelli P, Lasho TL, Rotunno G, Finke C, Mannarelli C, Belachew AA, Pancrazzi A, Wassie EA, Ketterling RP, Hanson CA, Pardanani A, Vannucchi AM. CALR and ASXL1 mutations-based molecular prognostication in primary myelofibrosis: an international study of 570 patients. Leukemia. 2014 Jul;28(7):1494-500. doi: 10.1038/leu.2014.57. Epub 2014 Feb 5. PMID: 24496303.

192. Triviai, I., Zeschke, S., Rentel, J. *et al.* ASXL1/EZH2 mutations promote clonal expansion of neoplastic HSC and impair erythropoiesis in PMF. *Leukemia* 33, 99–109 (2019). https://doi.org/10.1038/s41375-018-0159-0

193. Viré, E., Brenner, C., Deplus, R. *et al.* The Polycomb group protein EZH2 directly controls DNA methylation. *Nature* 439, 871–874 (2006). https://doi.org/10.1038/nature04431

194. Shimizu T, Kubovcakova L, Nienhold R, Zmajkovic J, Meyer SC, Hao-Shen H, Geier F, Dirnhofer S, Guglielmelli P, Vannucchi AM, Feenstra JD, Kralovics R, Orkin SH, Skoda RC. Loss of Ezh2 synergizes with JAK2-V617F in initiating myeloproliferative neoplasms and promoting myelofibrosis. J Exp Med. 2016 Jul 25;213(8):1479-96. doi: 10.1084/jem.20151136. Epub 2016 Jul 11. PMID: 27401344; PMCID: PMC4986524.

195. Paola Guglielmelli, Flavia Biamonte, Joannah Score, Claire Hidalgo-Curtis, Francisco Cervantes, Margherita Maffioli, Tiziana Fanelli, Thomas Ernst, Nils Winkelman, Amy V. Jones, Katerina Zoi, Andreas Reiter, Andrew Duncombe, Laura Villani, Alberto Bosi, Giovanni Barosi, Nicholas C. P. Cross, Alessandro M. Vannucchi; *EZH2* mutational status predicts poor survival in myelofibrosis. *Blood* 2011; 118 (19): 5227–5234. doi: https://doi.org/10.1182/blood-2011-06-363424

196. Kim E, Ilagan JO, Liang Y, Daubner GM, Lee SC, Ramakrishnan A, Li Y, Chung YR, Micol JB, Murphy ME, Cho H, Kim MK, Zebari AS, Aumann S, Park CY, Buonamici S, Smith PG, Deeg HJ, Lobry C, Aifantis I, Modis Y, Allain FH, Halene S, Bradley RK, Abdel-Wahab O. SRSF2 Mutations Contribute to Myelodysplasia by Mutant-Specific Effects on Exon Recognition. Cancer Cell. 2015 May 11;27(5):617-30. doi: 10.1016/j.ccell.2015.04.006. PMID: 25965569; PMCID: PMC4429920.

197. Gabriele Todisco, Anna Gallì, Erica Travaglino, Ettore Rizzo, Elisabetta Molteni, Chiara Elena, Ilaria Ambaglio, Silvia Catricalà, Paolo Bernasconi, Virginia Valeria Ferretti, Luca Malcovati, Mario Cazzola; *SRSF2* Mutations Identify a Distinct Subtype of Myeloid Neoplasm Across Myelodysplastic Syndromes and Acute Myeloid Leukemia. *Blood* 2016; 128 (22): 4337. doi: https://doi.org/10.1182/blood.V128.22.4337.4337

198. Chiara Elena, Luca Malcovati, Elli Papaemmanuil, Anna Gallì, Matteo G Della Porta, Ilaria Ambaglio, Marta Ubezio, Matteo C Da Vià, Elisa Bono, Anna Cattaneo, Erica Travaglino, Rosangela Invernizzi, Peter J. Campbell, Mario Cazzola; Clinical and

Molecular Features Of Myeloid Neoplasms Carrying *SRSF2* Mutation. *Blood* 2013; 122 (21): 1532. doi: https://doi.org/10.1182/blood.V122.21.1532.1532

199. Terra L. Lasho, Thitina Jimma, Christy M. Finke, Mrinal Patnaik, Curtis A. Hanson, Rhett P. Ketterling, Animesh Pardanani, Ayalew Tefferi; *SRSF2* mutations in primary myelofibrosis: significant clustering with *IDH* mutations and independent association with inferior overall and leukemia-free survival. *Blood* 2012; 120 (20): 4168–4171. doi: https://doi.org/10.1182/blood-2012-05-429696

200. Peng A, Mao X, Zhong J, Fan S, Hu Y. Single-Cell Multi-Omics and Its Prospective Application in Cancer Biology. Proteomics. 2020 Jul;20(13):e1900271. doi: 10.1002/pmic.201900271. Epub 2020 Apr 30. PMID: 32223079.

201. Wills QF, Mead AJ. Application of single-cell genomics in cancer: promise and challenges. *Hum Mol Genet*. 2015;24(R1):R74-R84. doi:10.1093/hmg/ddv235

202. Lim B, Lin Y, Navin N. Advancing Cancer Research and Medicine with Single-Cell Genomics. Cancer Cell. 2020 Apr 13;37(4):456-470. doi: 10.1016/j.ccell.2020.03.008. PMID: 32289270; PMCID: PMC7899145.

203. Gawad, C., Koh, W. & Quake, S. Single-cell genome sequencing: current state of the science. *Nat Rev Genet* 17, 175–188 (2016). https://doi.org/10.1038/nrg.2015.16

204. Kaur, Raman Preet (2019). *Single-Cell Omics || Single-Cell Genomics: Technology and Applications, 179–197.* doi:10.1016/B978-0-12-814919-5.00009-9

205. Evrony GD, Hinch AG, Luo C. Applications of Single-Cell DNA Sequencing. *Annu Rev Genomics Hum Genet*. 2021;22:171-197. doi:10.1146/annurev-genom-111320-090436

206. Espina, V., Wulfkuhle, J., Calvert, V. *et al.* Laser-capture microdissection. *Nat Protoc* 1, 586–603 (2006). https://doi.org/10.1038/nprot.2006.854

207. Gross A, Schoendube J, Zimmermann S, Steeb M, Zengerle R, Koltay P. Technologies for Single-Cell Isolation. *Int J Mol Sci*. 2015;16(8):16897-16919. Published 2015 Jul 24. doi:10.3390/ijms160816897

208. Luo T, Fan L, Zhu R, Sun D. Microfluidic Single-Cell Manipulation and Analysis: Methods and Applications. *Micromachines (Basel)*. 2019;10(2):104. Published 2019 Feb 1. doi:10.3390/mi10020104

209. Swennenhuis JF, Terstappen L. Sample Preparation Methods Following CellSearch Approach Compatible of Single-Cell Whole-Genome Amplification: An Overview. Methods Mol Biol. 2015;1347:57-67. doi: 10.1007/978-1-4939-2990-0_4. PMID: 26374309.

210. Di Trapani M, Manaresi N, Medoro G. DEPArray™ system: An automatic image-based sorter for isolation of pure circulating tumor cells. *Cytometry A*. 2018;93(12):1260-1266. doi:10.1002/cyto.a.23687

211. Huang L, Ma F, Chapman A, Lu S, Xie XS. Single-Cell Whole-Genome Amplification and Sequencing: Methodology and Applications. Annu Rev Genomics Hum Genet. 2015;16:79-102. doi: 10.1146/annurev-genom-090413-025352. Epub 2015 Jun 11. PMID: 26077818.

212. Hughes S, Arneson N, Done S, Squire J. The use of whole genome amplification in the study of human disease. Prog Biophys Mol Biol. 2005 May;88(1):173-89. doi: 10.1016/j.pbiomolbio.2004.01.007. PMID: 15561304.

213. Lüdecke, HJ., Senger, G., Claussen, U. et al. Cloning defined regions of the human genome by microdissection of banded chromosomes and enzymatic amplification. Nature 338, 348–350 (1989). https://doi.org/10.1038/338348a0

214. Telenius H, Carter NP, Bebb CE, Nordenskjöld M, Ponder BA, Tunnacliffe A. Degenerate oligonucleotide-primed PCR: general amplification of target DNA by a single degenerate primer. Genomics. 1992 Jul;13(3):718-25. doi: 10.1016/0888-7543(92)90147-k. PMID: 1639399.

215. Blagodatskikh KA, Kramarov VM, Barsova EV, et al. Improved DOP-PCR (iDOP-PCR): A robust and simple WGA method for efficient amplification of low copy number genomic DNA. *PLoS One*. 2017;12(9):e0184507. Published 2017 Sep 11. doi:10.1371/journal.pone.0184507

216. Zhang, L., Cui, X., Schmitt, K., Hubert, R., Navidi, W., Arnheim, N., 1992. Whole genome amplification from a single cell: implications for genetic analysis. Proc. Natl. Acad. Sci. U.S.A. 89, 5847–5851.

217. Arneson N, Hughes S, Houlston R, Done S. Whole-Genome Amplification by Improved Primer Extension Preamplification PCR (I-PEP-PCR). CSH Protoc. 2008 Jan 1;2008:pdb.prot4921. doi: 10.1101/pdb.prot4921. PMID: 21356675.

218. Spits, C., Le Caignec, C., De Rycke, M. *et al*. Whole-genome multiple displacement amplification from single cells. *Nat Protoc* 1, 1965–1970 (2006). https://doi.org/10.1038/nprot.2006.326

219. Zong C. (2017) Multiple Annealing and Looping-Based Amplification Cycles (MALBAC) for the Analysis of DNA Copy Number Variation. In: Frade J., Gage F. (eds) Genomic Mosaicism in Neurons and Other Cell Types. Neuromethods, vol 131. Humana Press, New York, NY. https://doi.org/10.1007/978-1-4939-7280-7_7

220. Chen M, Song P, Zou D, Hu X, Zhao S, Gao S, Ling F. Comparison of multiple displacement amplification (MDA) and multiple annealing and looping-based amplification cycles (MALBAC) in single-cell sequencing. PLoS One. 2014 Dec 8;9(12):e114520. doi: 10.1371/journal.pone.0114520. Erratum in: PLoS One. 2015;10(4):e0124990. PMID: 25485707; PMCID: PMC4259343.

221. Macaulay IC, Voet T. Single cell genomics: advances and future perspectives. PLoS Genet. 2014 Jan 30;10(1):e1004126. doi: 10.1371/journal.pgen.1004126. PMID: 24497842; PMCID: PMC3907301.

222. Haas S, Trumpp A, Milsom MD. Causes and Consequences of Hematopoietic Stem Cell Heterogeneity. Cell Stem Cell. 2018 May 3;22(5):627-638. doi: 10.1016/j.stem.2018.04.003. PMID: 29727678.

223. Hamey FK, Nestorowa S, Wilson NK, Göttgens B. Advancing haematopoietic stem and progenitor cell biology through single-cell profiling. FEBS Lett. 2016 Nov;590(22):4052-4067. doi: 10.1002/1873-3468.12231. Epub 2016 Jun 21. PMID: 27259698.

224. Ye F, Huang W, Guo G. Studying hematopoiesis using single-cell technologies. *J Hematol Oncol*. 2017;10(1):27. Published 2017 Jan 21. doi:10.1186/s13045-017-0401-7

225. Povinelli BJ, Rodriguez-Meira A, Mead AJ. Single cell analysis of normal and leukemic hematopoiesis. Mol Aspects Med. 2018 Feb;59:85-94. doi: 10.1016/j.mam.2017.08.006. Epub 2017 Sep 8. PMID: 28863981; PMCID: PMC5771467.

226. Anderson K, Lutz C, van Delft FW, Bateman CM, Guo Y, Colman SM, Kempski H, Moorman AV, Titley I, Swansbury J, Kearney L, Enver T, Greaves M. Genetic variegation of clonal architecture and propagating cells in leukaemia. Nature. 2011 Jan 20;469(7330):356-61. doi: 10.1038/nature09650. Epub 2010 Dec 15. PMID: 21160474.

227. Wang L, Fan J, Francis JM, Georghiou G, Hergert S, Li S, Gambe R, Zhou CW, Yang C, Xiao S, Cin PD, Bowden M, Kotliar D, Shukla SA, Brown JR, Neuberg D, Alessi DR, Zhang CZ, Kharchenko PV, Livak KJ, Wu CJ. Integrated single-cell genetic and transcriptional analysis suggests novel drivers of chronic lymphocytic leukemia. Genome Res. 2017 Aug;27(8):1300-1311. doi: 10.1101/gr.217331.116. Epub 2017 Jul 5. PMID: 28679620; PMCID: PMC5538547.

228. Paguirigan AL, Smith J, Meshinchi S, Carroll M, Maley C, Radich JP. Single-cell genotyping demonstrates complex clonal diversity in acute myeloid leukemia. *Sci Transl Med*. 2015;7(281):281re2. doi:10.1126/scitranslmed.aaa0763

229. Morita, K., Wang, F., Jahn, K. *et al.* Clonal evolution of acute myeloid leukemia revealed by high-throughput single-cell genomics. *Nat Commun* 11, 5327 (2020). https://doi.org/10.1038/s41467-020-19119-8

230. Royston D, Mead AJ, Psaila B. Application of Single-Cell Approaches to Study Myeloproliferative Neoplasm Biology. *Hematol Oncol Clin North Am*. 2021;35(2):279-293. doi:10.1016/j.hoc.2021.01.002

231. Ortmann CA, Kent DG, Nangalia J, Silber Y, Wedge DC, Grinfeld J, Baxter EJ, Massie CE, Papaemmanuil E, Menon S, Godfrey AL, Dimitropoulou D, Guglielmelli P, Bellosillo B, Besses C, Döhner K, Harrison CN, Vassiliou GS, Vannucchi A, Campbell PJ, Green AR. Effect of mutation order on myeloproliferative neoplasms. N Engl J Med. 2015 Feb 12;372(7):601-612. doi: 10.1056/NEJMoa1412098. PMID: 25671252; PMCID: PMC4660033.

232. Carretta C, Mallia S, Genovese E, Parenti S, Rontauroli S, Bianchi E, Fantini S, Sartini S, Tavernari L, Tagliafico E, Manfredini R. Genomic Analysis of Hematopoietic Stem Cell at the Single-Cell Level: Optimization of Cell Fixation and Whole Genome Amplification (WGA) Protocol. Int J Mol Sci. 2020 Oct 6;21(19):7366. doi: 10.3390/ijms21197366. PMID: 33036143; PMCID: PMC7582552.

233. Parenti S, Rontauroli S, Carretta C, Mallia S, Genovese E, Chiereghin C, Peano C, Tavernari L, Bianchi E, Fantini S, Sartini S, Romano O, Bicciato S, Tagliafico E, Della Porta M, Manfredini R. Mutated clones driving leukemic transformation are already detectable at the single-cell level in CD34-positive cells in the chronic phase of primary myelofibrosis. NPJ Precis Oncol. 2021 Feb 4;5(1):4. doi: 10.1038/s41698-021-00144-9. PMID: 33542466; PMCID: PMC7862275.

234. Navin N., Kendall J., Troge J., Andrews P., Rodgers L., McIndoo J., Cook K., Stepansky A., Levy D., Esposito D., et al. Tumour evolution inferred by single-cell sequencing. *Nature.* 2011;472:90-U119. doi: 10.1038/nature09807

235. Xu X., Hou Y., Yin X.Y., Bao L., Tang A.F., Song L.T., Li F.Q., Tsang S., Wu K., Wu H.J., et al. Single-Cell Exome Sequencing Reveals Single-Nucleotide Mutation Characteristics of a Kidney Tumor. *Cell.* 2012;148:886–895. doi: 10.1016/j.cell.2012.02.025.

236. Yu C., Yu J., Yao X.T., Wu W.K.K., Lu Y.Y., Tang S.W., Li X.C., Bao L., Li X.X., Hou Y., et al. Discovery of biclonal origin and a novel oncogene SLC12A5 in colon cancer by single-cell sequencing. *Cell Res.* 2014;24:701–712. doi: 10.1038/cr.2014.43.

237. Alexander J., Kendall J., McIndoo J., Rodgers L., Aboukhalil R., Levy D., Stepansky A., Sun G.L., Chobardjiev L., Riggs M., et al. Utility of Single-Cell Genomics in Diagnostic Evaluation of Prostate Cancer. *Cancer Res.* 2018;78:348–358. doi: 10.1158/0008-5472.CAN-17-1138.

238. Cossarizza A, et al. Guidelines for the use of flow cytometry and cell sorting in immunological studies. Eur J Immunol. 2017 Oct;47(10):1584-1797. doi: 10.1002/eji.201646632. PMID: 29023707.

239. Thavarajah R, Mudimbaimannar VK, Elizabeth J, Rao UK, Ranganathan K. Chemical and physical basics of routine formaldehyde fixation. J Oral Maxillofac Pathol. 2012 Sep;16(3):400-5. doi: 10.4103/0973-029X.102496. PMID: 23248474; PMCID: PMC3519217.

240. Melan MA. Overview of cell fixatives and cell membrane permeants. Methods Mol Biol. 1999;115:45-55. doi: 10.1385/1-59259-213-9:45. PMID: 10098165.

241. Lanier LL, Warner NL. Paraformaldehyde fixation of hematopoietic cells for quantitative flow cytometry (FACS) analysis. J Immunol Methods. 1981;47(1):25-30. doi: 10.1016/0022-1759(81)90253-2. PMID: 7310138.

242. Thomsen ER, et al. Fixed single-cell transcriptomic characterization of human radial glial diversity. Nat Methods. 2016;13(1):87–93.

243. Hu, P.; Zhang, W.; Xin, H.; Deng, G. Single Cell Isolation and Analysis. *Front. Cell Dev. Biol.* **2016**, *4*, 116.

244. Russell JN, Clements JE, Gama L. Quantitation of gene expression in formaldehyde-fixed and fluorescence-activated sorted cells. PLoS One. 2013 Sep 2;8(9):e73849. doi: 10.1371/journal.pone.0073849. PMID: 24023909; PMCID: PMC3759445.

245. Kumarasinghe MP, Constantine SR, Hemamali RL. Methanol as an alternative fixative for cytological smears. Malays J Pathol. 1997 Dec;19(2):137-40. PMID: 10879255.

246. Andrey N, Kuzmin A, Pliss, Paras N, Prasad. Change in biomolecular profile in a single nucleolus during cell fixation. Analytical chemistry. 2014. https://doi.org/10.1021/ac503172b.

247. Wang, X., Yu, L. & Wu, A.R. The effect of methanol fixation on single-cell RNA sequencing data. *BMC Genomics* 22, 420 (2021). https://doi.org/10.1186/s12864-021-07744-6

248. Alles, J., Karaiskos, N., Praktiknjo, S.D. *et al.* Cell fixation and preservation for droplet-based single-cell transcriptomics. *BMC Biol* 15, 44 (2017). https://doi.org/10.1186/s12915-017-0383-5

249. Zhang X, Liang B, Xu X, et al. The comparison of the performance of four whole genome amplification kits on ion proton platform in copy number variation detection. *Biosci Rep*. 2017;37(4):BSR20170252. Published 2017 Jul 7. doi:10.1042/BSR20170252

250. Deleye L, Tilleman L, Vander Plaetsen AS, Cornelis S, Deforce D, Van Nieuwerburgh F. Performance of four modern whole genome amplification methods for copy number variant detection in single cells. Sci Rep. 2017 Jun 13;7(1):3422. doi: 10.1038/s41598-017-03711-y. PMID: 28611458; PMCID: PMC5469777.

251. Chen D, Zhen H, Qiu Y, Liu P, Zeng P, Xia J, Shi Q, Xie L, Zhu Z, Gao Y, Huang G, Wang J, Yang H, Chen F. Comparison of single cell sequencing data between two whole genome amplification methods on two sequencing platforms. Sci Rep. 2018 Mar 21;8(1):4963. doi: 10.1038/s41598-018-23325-2. PMID: 29563514; PMCID: PMC5862989.

252. Shestak AG, Bukaeva AA, Saber S, Zaklyazminskaya EV. Allelic Dropout Is a Common Phenomenon That Reduces the Diagnostic Yield of PCR-Based Sequencing of Targeted Gene Panels. Front Genet. 2021 Feb 1;12:620337. doi: 10.3389/fgene.2021.620337. PMID: 33633783; PMCID: PMC7901947.

253. Campbell PJ, Baxter EJ, Beer PA, et al. Mutation of JAK2 in the myeloproliferative disorders: timing, clonality studies, cytogenetic associations, and role in leukemic transformation. *Blood*. 2006;108(10):3548-3555.

254. Lundberg P, Karow A, Nienhold R, Looser R, Hao-Shen H, Nissen I, et al. Clonal evolution and clinical correlates of somatic mutations in myeloproliferative neoplasms. Blood. 2014;123(14):2220–8.

255. Schaub FX, Looser R, Li S, Hao-Shen H, Lehmann T, Tichelli A, Skoda RC. Clonal analysis of TET2 and JAK2 mutations suggests that TET2 can be a late event in the progression of myeloproliferative neoplasms. Blood. 2010 Mar 11;115(10):2003-7. doi: 10.1182/blood-2009-09-245381. Epub 2010 Jan 8. PMID: 20061559.

256. Yap TA, Gerlinger M, Futreal PA, Pusztai L, Swanton C. Intratumor heterogeneity: seeing the wood for the trees. Sci Transl Med. 2012 Mar 28;4(127):127ps10. doi: 10.1126/scitranslmed.3003854. PMID: 22461637

257. Daver, N., Venugopal, S. & Ravandi, F. FLT3 mutated acute myeloid leukemia: 2021 treatment algorithm. *Blood Cancer J.* 11, 104 (2021). https://doi.org/10.1038/s41408-021-00495-3

258. Shahin OA, Chifotides HT, Bose P, Masarova L, Verstovsek S. Accelerated Phase of Myeloproliferative Neoplasms. Acta Haematol. 2021;144(5):484-499. doi: 10.1159/000512929. Epub 2021 Apr 21. PMID: 33882481; PMCID: PMC8458218.

259. Moore, A. S. et al. Selective FLT3 inhibition of FLT3-ITD+ acute myeloid leukaemia resulting in secondary D835Y mutation: a model for emerging clinical resistance patterns. Leukemia 26, 1462–1470 (2012).

260. Abdel-Wahab O, Manshouri T, Patel J, Harris K, Yao J, Hedvat C, Heguy A, Bueso-Ramos C, Kantarjian H, Levine RL, Verstovsek S. Genetic analysis of transforming events that convert chronic myeloproliferative neoplasms to leukemias. Cancer Res. 2010 Jan 15;70(2):447-52. doi: 10.1158/0008-5472.CAN-09-3783. Epub 2010 Jan 12. PMID: 20068184; PMCID: PMC2947340.

261. Khan M, Siddiqi R, Gangat N. Therapeutic options for leukemic transformation in patients with myeloproliferative neoplasms. Leuk Res. 2017 Dec;63:78-84. doi: 10.1016/j.leukres.2017.10.009. Epub 2017 Oct 27. PMID: 29121538.

262. Rampal R, Mascarenhas J. Pathogenesis and management of acute myeloid leukemia that has evolved from a myeloproliferative neoplasm. Curr Opin Hematol. 2014 Mar;21(2):65-71. doi: 10.1097/MOH.0000000000000017. PMID: 24366192.

263. Jain, T., & Rampal, R. K. (2021). *Accelerated and Blast Phase Myeloproliferative Neoplasms. Hematology/Oncology Clinics of North America, 35(2), 325–335.* doi:10.1016/j.hoc.2020.12.008

264. Marcellino B, Mascarenhas J. Management of advanced phase myeloproliferative neoplasms. Clin Adv Hematol Oncol. 2019 Jul;17(7):405-411. PMID: 31449507.

265. Harutyunyan A, Klampfl T, Cazzola M, Kralovics R. p53 lesions in leukemic transformation. N Engl J Med. 2011 Feb 3;364(5):488-90. doi: 10.1056/NEJMc1012718. PMID: 21288114.

266. McNamara CJ, Panzarella T, Kennedy JA, et al. The mutational landscape of accelerated- and blast-phase myeloproliferative neoplasms impacts patient outcomes. *Blood Adv*. 2018;2(20):2658-2671. doi:10.1182/bloodadvances.2018021469

267. Kubesova B, Pavlova S, Malcikova J, Kabathova J, Radova L, Tom N, Tichy B, Plevova K, Kantorova B, Fiedorova K, Slavikova M, Bystry V, Kissova J, Gisslinger B, Gisslinger H, Penka M, Mayer J, Kralovics R, Pospisilova S, Doubek M. Low-burden TP53 mutations in chronic phase of myeloproliferative neoplasms: association with age, hydroxyurea administration, disease type and JAK2 mutational status. Leukemia. 2018 Feb;32(2):450-461. doi: 10.1038/leu.2017.230. Epub 2017 Jul 24. PMID: 28744014; PMCID: PMC5808067.

268. Bernard, E. et al. Implications of TP53 allelic state for genome stability, clinical presentation and outcomes in myelodysplastic syndromes. Nat. Med. 26, 1549–1556 (2020).

269. Mylonas E, Yoshida K, Frick M, Hoyer K, Christen F, Kaeda J, Obenaus M, Noerenberg D, Hennch C, Chan W, Ochi Y, Shiraishi Y, Shiozawa Y, Zenz T, Oakes CC, Sawitzki B, Schwarz M, Bullinger L, le Coutre P, Rose-Zerilli MJJ, Ogawa S, Damm F. Single-cell analysis based dissection of clonality in myelofibrosis. Nat Commun. 2020 Jan 7;11(1):73. doi: 10.1038/s41467-019-13892-x. PMID: 31911629; PMCID: PMC6946829.

270. Zhang SJ, Rampal R, Manshouri T, et al. Genetic analysis of patients with leukemic transformation of myeloproliferative neoplasms shows recurrent SRSF2 mutations that are associated with adverse outcome. *Blood*. 2012;119(19):4480-4485. doi:10.1182/blood-2011-11-390252

271. Xiao R, Sun Y, Ding JH, Lin S, Rose DW, Rosenfeld MG, Fu XD, Li X. 2007. Splicing regulator SC35 is essential for genomic stability and cell proliferation during mammalian organogenesis. *Mol Cell Biol* 27:5393–5402.

272. Komeno Y, Huang YJ, Qiu J, Lin L, Xu Y, Zhou Y, Chen L, Monterroza DD, Li H, DeKelver RC, Yan M, Fu XD, Zhang DE. 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. doi: 10.1128/MCB.00202-15. Epub 2015 Jun 29. PMID: 26124281; PMCID: PMC4525309.

273. Lee SC, North K, Kim E, Jang E, Obeng E, Lu SX, Liu B, Inoue D, Yoshimi A, Ki M, Yeo M, Zhang XJ, Kim MK, Cho H, Chung YR, Taylor J, Durham BH, Kim YJ,

Pastore A, Monette S, Palacino J, Seiler M, Buonamici S, Smith PG, Ebert BL, Bradley RK, Abdel-Wahab O. Synthetic Lethal and Convergent Biological Effects of Cancer-Associated Spliceosomal Gene Mutations. Cancer Cell. 2018 Aug 13;34(2):225-241.e8. doi: 10.1016/j.ccell.2018.07.003. PMID: 30107174; PMCID: PMC6373472.