UNIVERSITÀ DEGLI STUDI DI MILANO

DEPARTMENT OF MEDICAL BIOTECHNOLOGY AND TRANSLATIONAL MEDICINE (BIOMETRA)

PhD Course in

EXPERIMENTAL MEDICINE AND MEDICAL BIOTECHNOLOGIES

XXX cycle

Molecular heterogeneity of Multiple Myeloma: The biological and clinical relevance of novel gene mutations by Next Generation Sequencing

[Med/15 - Malattie del Sangue]

Tutor: Professor Luca BALDINI

Coordinator: Professor Massimo LOCATI

PhD student:

Martina MANZONI

Matriculation n. R10983

Academic Year: 2016/2017

Table of contents

ABSTRACT……3

SINTESI……7

INTRODUCTION……11

AIM of STUDY……29

MATERIALS and METHODS……32

RESULTS……43

DISCUSSION and CONCLUSIONS……81

BIBLIOGRAPHY……92

Abstract

Abstract

Multiple myeloma (MM) is a fatal malignant proliferation of antibody-secreting bone marrow (BM) plasma cells (PCs) characterized by a wide clinical spectrum and a profound genomic instability. Concerning in particular the mutational landscape, recent next generation sequencing (NGS) studies in MM patients indicated the lack of a universal driver mutation but several recurrently mutated genes belonging to key pathways involved in the disease, such as the MAP-kinase pathway.

MM characterization currently depends on BM aspirates for mutational analysis. However, this approach might not capture the putative spatial and genetic heterogeneity of the disease, and imposes technical hurdles that are limiting its transfer in the routine and clinical grade diagnostic laboratory, restricting also the possibility of longitudinal monitoring of disease molecular markers. Nevertheless, recent data produced in solid cancers indicate that circulating tumor DNA into the peripheral blood (PB) can be used as source of tumor DNA to provide information about tumor mass, residual disease and tumor genotype, with obvious advantages in terms of accessibility.

Based on the previous observations, the aims of the project were: (i) the evaluation of the incidence of mutations in MM driver genes (*KRAS*, *NRAS*, *TP53*, *BRAF*, *FAM46C* and *DIS3*) in a large and representative cohort of patients at different stages of PC dyscrasia (132 MM and 24 primary PC leukemia (pPCL) cases, both at onset, and 11 secondary PCL (sPCL) patients); (ii) the comparison of the mutational profiling of circulating cell free DNA (cfDNA) and BM derived DNA in a small series of patients representative of different clinical stages of PC tumors.

Specifically, we evaluated, by means of deep NGS, the incidence of mutations in *BRAF* (exons 11 and 15), *NRAS* (exons 2 and 3) and *KRAS* (exons 2-4), *DIS3* in the PIN (exons 1-4) and *RNB* (exons 10-18) functional domains, *TP53* (exons 4-9) and *FAM46C* (exon 2).

Overall, the MAPK pathway resulted affected in 60.1% of the patients (63.6% of those with sPCL, 59.8% of those with MM, and 41.7% of those with pPCL). In particular, 12% of patients were found mutated in *BRAF*, 23.9% in *NRAS*, and 29.3% in *KRAS*. *DIS3* mutations were found, respectively, in 18.5% of MM patients at diagnosis, 25% of pPCLs and 30% of sPCLs, occurring more frequently in *IGH*-translocated/nonhyperdiploid patients. Twenty-four tumour-specific mutations were identified in *FAM46C* affecting 18/162 (11%) patients. In particular, the frequency of *FAM46C* mutations was 11.7% in newly diagnosed MM, 4.2% and 20% in primary and secondary PCL, respectively. *TP53* was mutated in 4/129 (3%) MM, 6/24 (25%) pPCL, and 2/10 (20%) sPCL cases. A similar increase in prevalence associated with disease aggressiveness (5%, 29.2% and 44%, respectively) was observed for *TP53* deletion.

In all analyzed genes, co-existing mutations tended to occur at different variant allele frequencies (VAFs), thus supporting the occurrence of tumor subclones.

Longitudinal analysis at diagnosis and relapse in a subset of 19 cases including both MM and PCL cases, revealed different mutation patterns in *BRAF*/*NRAS*/*KRAS*. We observed the presence of clonal variants at both time-points; the acquisition/clonal expansion of variants in the later sample; and even the disappearance of variants at relatively high VAF values, but always concurrently with the emergence/clonal expansion of an additional mutation in another gene of the MAPK pathway. These longitudinally analysis highlighted some instances of increasing *DIS3* mutation burden during disease progression. On the contrary, in *FAM46C* we observed a reduction or disappearance of three primary mutations with a quite constant VAF in cases at onset; instead, we noticed the acquisition of *TP53* mutations in three of the nineteen cases analyzed at relapse.

The majority of *BRAF*/*DIS3*/*FAM46C* variants in mutated cases were comparably detectable at transcriptional level. Overall, the finding that *FAM46C* mutated alleles had detectable biological expression, and in some cases seemed to be even preferentially transcribed, suggests that the mutations identified herein may have functional implications.

The study on cfDNA as an accessible source of genomic material in MM patients was based on a series of 28 patients with PC disorders, of whom we collecteded: (1) cfDNA isolated from plasma; (2) tumor genomic (g)DNA from CD138+ purified BM PCs for comparative purposes, and (3) germline gDNA extracted from PB granulocytes, to filter out polymorphisms. CAPP-seq ultra-deep targeted NGS approach was performed to genotype a gene panel specifically designed to maximize the mutation recovery in PC tumors. Overall, within the interrogated genes, 18/28 (64%) patients had at least one non-synonymous somatic mutation detectable in cfDNA. Consistent with the spectrum of mutated genes in PC disorders, plasma cfDNA genotyping revealed somatic variants of *NRAS* (25%); *KRAS* (14%); *TP53*, *TRAF3* and *FAM46C* (11%, respectively); *CYLD* and *DIS3* (7%, respectively); and *BRAF* and *IRF4* (4%, respectively). cfDNA genotyping correctly identified 72% of mutations (n=28/39) discovered in tumor PCs and, overall, the VAFs in plasma samples correlated with those in tumor biopsies. Notably, the remaining mutations not discovered in cfDNA had a low representation in the purified BM PCs (median VAF=2.5%). ROC analysis showed that cfDNA genotyping had the highest sensitivity (100%) if mutations were represented with a VAF >5% in purified BM PCs.

The results of our study confirm and extend previous published evidence that MAPK pathway activation is recurrent in MM, and the finding that it is mediated by *BRAF* mutations in a significant fraction of patients has potentially immediate clinical implications. Furthermore, gene expression profiling analysis in *DIS3*-mutated patients identified a transcriptional signature suggestive for impaired RNA exosome function. Our data further support the pathological relevance of *DIS3* mutations in PC dyscrasia and suggest that *DIS3* may represent a potential tumor suppressor gene in such disorders. It is also strengthened the growing evidence that *FAM46C* is involved in the

pathogenesis of MM as a potential tumour suppressor, although its role remains to be clarified. Our outcomes also confirm that *TP53* mutations are rare in MM at presentation and rather represent a marker of progression, similarly to del(17p); however, their occurrence even in absence of deletions supports the importance of their assessment in patients with PC dyscrasia, in terms of both risk stratification and therapeutic implications.

Our results demonstrate as well that cfDNA genotyping is a feasible, non-invasive, real-time approach able to detect clonal and subclonal somatic mutations represented in at least 5% of alleles in tumor PCs, thus supporting the advisable introduction of this method in clinical trials monitoring PC dyscrasia.

Sintesi

Il mieloma multiplo (MM) è una proliferazione maligna fatale delle plasmacellule (PC) secernenti anticorpi del midollo osseo (BM), caratterizzata da un ampio spettro clinico e da una profonda instabilità genomica. Per quanto riguarda in particolare il panorama mutazionale, recenti studi di sequenziamento di nuova generazione (NGS) nei pazienti con MM hanno indicato la mancanza di una mutazione pilota universale, ma diversi geni mutati ricorrenti appartenenti a percorsi chiave coinvolti nella malattia, come la via MAP-chinasi. La caratterizzazione del MM attualmente si basa sull'analisi mutazionale effettuata sugli aspirati di BM. Tuttavia, questo approccio potrebbe non catturare la presunta eterogeneità spaziale e genetica della malattia e imporre ostacoli tecnici che ne limitano il trasferimento nel laboratorio diagnostico di routine e clinico, riducendo anche la possibilità di un monitoraggio longitudinale dei marcatori molecolari della malattia. Ciononostante, recenti dati relativi a cancri solidi indicano che il DNA circolante nel sangue periferico (PB) può essere usato come fonte di DNA tumorale per fornire informazioni sulla massa tumorale, sulla malattia residua e sul genotipo tumorale, con ovvi vantaggi in termini di accessibilità. Sulla base delle osservazioni precedenti, gli obiettivi del progetto erano: (i) la valutazione dell'incidenza delle mutazioni nei geni "driver" del MM (*KRAS*, *NRAS*, *TP53*, *BRAF*, *FAM46C* e *DIS3*) in una coorte ampia e rappresentativa di pazienti a diversi stadi di discrasia plamacellulare (132 MM e 24 casi di leucemia plasmacellulare primaria (pPCL), entrambi all'esordio, e in 11 pazienti con PCL secondaria (sPCL)); (ii) il confronto del profilo mutazionale del DNA libero da cellule circolante (cfDNA) e del DNA derivato dal BM in una piccola serie di pazienti rappresentativa di diversi stadi clinici di tumori plasmacellulari.

Nello specifico, abbiamo valutato, mediante NGS ad alta profondità, l'incidenza di mutazioni in *BRAF* (esoni 11 e 15), *NRAS* (esoni 2 e 3) e *KRAS* (esoni 2-4), *DIS3* nei domini funzionali PIN (esoni 1-4) e RNB (esoni 10-18), *TP53* (esoni 4-9) e *FAM46C* (esone 2).

Complessivamente, la via MAPK è risultata affetta da mutazioni nel 60,1% dei pazienti (63,6% di quelli con sPCL, 59,8% di quelli con MM e 41,7% di quelli con pPCL). In particolare, il 12% dei pazienti è stato trovato mutato in *BRAF*, il 23,9% in *NRAS* e il 29,3% in *KRAS*. Mutazioni di *DIS3* sono state riscontrate, rispettivamente, nel 18,5% dei pazienti con MM alla diagnosi, nel 25% delle pPCL e nel 30% delle sPCLs, con una maggior frequenza nei pazienti *IGH*-traslocati/non iperdiploidi. In *FAM46C* sono state identificate 24 mutazioni tumore-specifiche interessanti 18/162 (11%) pazienti. In particolare, la frequenza delle mutazioni di *FAM46C* era dell'11,7% nei MM di nuova diagnosi, e del 4,2% e del 20%, rispettivamente, nelle PCL primarie e secondarie. *TP53* era mutato in 4/129 (3%) MM, 6/24 (25%) casi di pPCL e 2/10 (20%) di sPCL. Una frequenza simile associata all'aggressività della malattia (5%, 29,2% e 44% rispettivamente) è stata osservata per la delezione di *TP53*.

In tutti i geni analizzati, le mutazioni coesistenti tendevano a verificarsi a diverse frequenze di varianti alleliche (VAFs), sostenendo in tal modo la presenza di subcloni tumorali.

L'analisi longitudinale alla diagnosi e alla recidiva in un sottogruppo di 19 pazienti, che includeva casi di MM e PCL, ha rivelato diversi modelli di mutazione in *BRAF*/*NRAS*/*KRAS*. Abbiamo osservato la presenza di varianti clonali ad entrambe le tempistiche; l'acquisizione/espansione clonale delle varianti nel campione successivo; e persino la scomparsa di varianti a valori di VAF relativamente alti, ma sempre in concomitanza con l'emergenza/l'espansione clonale di un'ulteriore mutazione in un altro gene della via MAPK. Queste analisi longitudinali hanno evidenziato alcuni casi di aumento del carico di mutazione di *DIS3* durante la progressione della malattia. Al contrario, in *FAM46C* abbiamo osservato una riduzione o scomparsa di tre mutazioni primarie con una VAF abbastanza costante nei casi all'esordio; invece, abbiamo notato l'acquisizione di mutazioni di *TP53* in tre dei diciannove casi analizzati in recidiva.

La maggior parte delle varianti di *BRAF*/*DIS3*/*FAM46C* nei casi mutati erano comparabilmente rilevabili a livello trascrizionale. Complessivamente, la scoperta che gli alleli mutati di *FAM46C* avevano un'espressione biologica rilevabile e, in alcuni casi, sembrava addirittura trascritta preferenzialmente, suggerisce che le mutazioni qui identificate possono avere implicazioni funzionali.

Lo studio sul cfDNA come fonte accessibile di materiale genomico nei pazienti con MM è stato condotto su una serie di 28 pazienti con disordini plasmacellulari, dei quali abbiamo raccolto: (1) cfDNA isolato dal plasma; (2) DNA genomico (gDNA) tumorale da PC di BM CD138+ purificate, per scopi comparativi e (3) gDNA germinale estratto da granulociti da PB, per filtrare i polimorfismi. L'approccio "NGS CAPP-seq ultra-deep targeted" è stato eseguito per genotipizzare un pannello genico specificamente progettato per massimizzare l'identificazione delle mutazioni nei tumori plasmacellulari. Complessivamente, all'interno dei geni interrogati, 18/28 (64%) pazienti avevano almeno una mutazione somatica non sinonima rilevabile nel cfDNA.

Coerentemente con lo spettro dei geni mutati nei disordini delle PC, la genotipizzazione del cfDNA nel plasma ha rivelato varianti somatiche di *NRAS* (25%); *KRAS* (14%); *TP53*, *TRAF3* e *FAM46C* (11%, rispettivamente); *CYLD* e *DIS3* (7%, rispettivamente); e *BRAF* e *IRF4* (4%, rispettivamente). La genotipizzazione del cfDNA ha identificato correttamente il 72% delle mutazioni (n=28/39) rilevate nelle PC tumorali e, nel complesso, le VAFs nei campioni di plasma correlavano con quelle delle biopsie tumorali. Va notato che le rimanenti mutazioni non rilevate nel cfDNA avevano una bassa rappresentazione nelle PC midollari purificate (VAF mediana=2,5%). L'analisi ROC ha mostrato che la genotipizzazione di cfDNA aveva la massima sensibilità (100%) se le mutazioni erano rappresentate con una VAF> 5% nelle PC di BM purificate.

I risultati del nostro studio confermano ed estendono le evidenze pubblicate in precedenza che l'attivazione del pathway delle MAPK è ricorrente nel MM, e la scoperta che essa è mediata da mutazioni in *BRAF* in una frazione significativa di pazienti ha implicazioni cliniche potenzialmente

immediate. Inoltre, l'analisi del profilo di espressione genica in pazienti *DIS3*-mutati ha identificato una “signature” trascrizionale suggestiva per la riduzione della funzione dell’esosoma dell'RNA. I nostri dati supportano ulteriormente la rilevanza patologica delle mutazioni di *DIS3* nelle discrasie plasmacellulari e suggeriscono che *DIS3* possa rappresentare un potenziale gene soppressore del tumore in tali disordini. E 'anche rafforzata la crescente evidenza che *FAM46C* sia coinvolto nella patogenesi del MM come potenziale gene oncosoppressore, anche se il suo ruolo rimane da chiarire. I nostri risultati confermano anche che le mutazioni di *TP53* sono rare nel MM all’esordio e rappresentano piuttosto un marker di progressione, analogamente alla del(17p); tuttavia, il loro verificarsi anche in assenza di delezioni supporta l'importanza della loro valutazione in pazienti con discrasia plasmacellulare, in termini sia di stratificazione del rischio sia di implicazioni terapeutiche. I nostri risultati dimostrano anche che la genotipizzazione di cfDNA è un approccio fattibile, non invasivo, in tempo reale in grado di rilevare mutazioni somatiche clonali e subclonali rappresentate in almeno il 5% degli alleli nelle PC tumorali, supportando così la consigliabile introduzione di questo metodo negli studi clinici che monitorano le discrasie plasmacellulari.

Introduction

Multiple myeloma (MM) represents a heterogeneous hematologic malignancy that occurs generally in the elderly population (median age at diagnosis ~70 years). Similarly to many other neoplastic diseases, also MM incidence is progressively increasing worldwide, mainly due to the major improvements in people life expectancy, leading to a marked increase in longevity. Symptomatic MM is generally preceded by an indolent expansion of clonal plasma cells, known as monoclonal gammopathy of undetermined significance (MGUS) and/or smoldering MM (sMM)[1-4]. These entities are completely asymptomatic and usually their diagnosis is made occasionally. Generally, the prevalence of MGUS and sMM progressively increases in the elderly with more than 5% prevalence among general population older than 70 years[1, 3]. The majority of MGUS and a part of sMM remains commonly stable and asymptomatic over the time. Nevertheless, in some patients, those indolent entities progress in symptomatic MM. The MGUS/sMM transformation into MM is clinically characterized by lytic bone disease, anaemia, renal failure, hypercalcemia, and susceptibility to bacterial infections. Which biological features and why some MGUS/sMM remain totally asymptomatic for decades whereas others progress in symptomatic MM is at present unknown, but the current hypothesis is the accumulation by the time of "malignant" genetic events, similarly to what happens in other cancers such as melanoma[1, 3].

Cell of Origin: The Plasma Cell

The MM normal counterpart is the plasma cell (PC), which represents the final differentiation stage of B cells. The first steps of differentiation occur within the bone marrow (BM) where different complex processes lead to the rearrangements and editing of the heavy chain immunoglobulin (Ig) gene (*IGH*)[5]. This gene (2 Mb) encodes a very large protein, presenting four major domains: the variability domain (VH), the diversity domain (DH), the joining domain (JH), and the constant domain. The first rearrangements are DNA stochastic and antigen independent deletions combining one of the DH segments to one of the six JH segments. If molecularly productive, the pro-B cell continues its differentiation processes by combining this DH-JH segment with a VH segment. This process is regulated by the recombination activating genes (*RAG1* and *RAG2*), encoding for specific enzymes that recognize precise DNA motifs within the DH, JH, and VH segments. If these rearrangements are in frame, or "productive", the pre-B cell will then rearrange the light chain genes, Ig kappa locus (*IGK*) and Ig lambda locus (*IGL*). The *IGK* is rearranged for first and, if the result is productive, the mature B cell will then be able to produce IgMk, generally expressed at the B-cell surface. If this process fails, the B-cell will then rearrange the *IGL* gene, leading to the production of an IgMλ. At the end of V(D)J recombination, B cells are able to express a functional B cell receptor (BCR) and subsequently they migrate to the secondary lymphoid organs (such as lymph nodes,

tonsils, and spleen). Here, interacting with an antigen, naive B cells are activated by interaction with CD4+ T cells in the T cell-rich area of the lymphoid tissues and aggregate into primary follicles to form germinal centres (GCs)[6, 7]. In the GC, B-cells are involved in two main immunoglobulin gene remodelling processes — namely, somatic hypermutation (SHM) and class-switch recombination (CSR) — in order to generate cells with the highest antigen-affinity antibodies of different isotype classes[8]. The GC emerges as a crucible in genesis of MGUS and MM via a distinctive pathway in which the cell of origin is targeted by on-going SHM[9].

SHM and CSR are driven by a specific enzyme called activation induced deaminase (AID) which represents the pivotal player of all GC reactions. In the first process, AID generates a variable and stochastic number of mutations within the VDJ segment. Only B-cells with mutations improving the specificity of the antibody for the antigen will survive, the others are removed via apoptosis. In the CSR process, AID recombines specific DNA segments known as switch regions with deletion of the inter switch region DNA, guarantying the expression of different Ig, either IgG, IgA, or IgE. After these processes, the B-cell maturation ends with the generation of both PCs and memory B cells (**Figure 1**). The dysregulation of AID hyper mutation activity and ability to break and recombine different DNA segments may generate a wide range of abnormal B cells that, if not promptly removed, may develop different post GC lymphoproliferative disorders such as different B-cell lymphomas, chronic lymphocytic leukemia and MM[5].

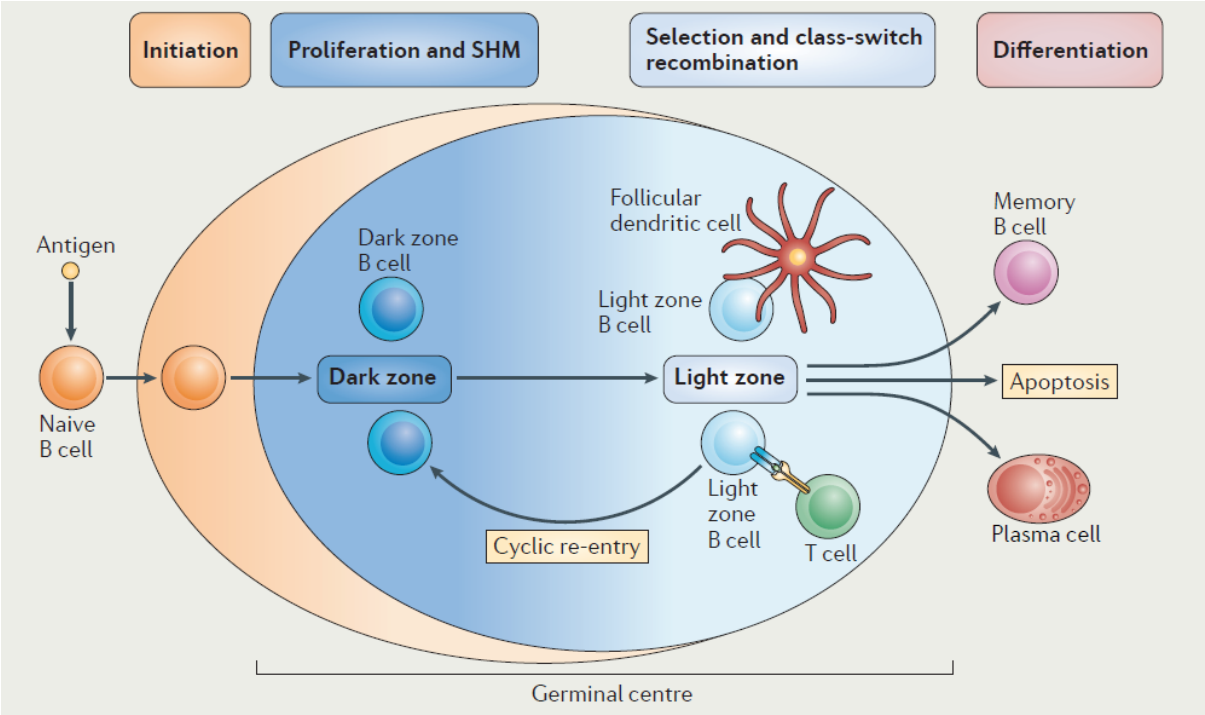

Figure 1. B-cell maturation model through the germinal centre[5].

PCs are characterized by strong BM dependence for survival and growth. The immunophenotype of MM cells resembles that of normal, terminally differentiated, long-lived BMPC (CD19–, CD20–, CD45–, CD138+). In contrast to normal long-lived PCs, MM cells maintain some potential for an extremely low rate of proliferation, usually with only just a few per cent of cycling cells, until advanced stages of the disease[10].

Multi-step clinical course and pathogenesis

MM affects antibody-secreting BMPCs and is characterized by a wide clinical spectrum ranging from the presumed pre-malignant condition of MGUS to sMM, frank and symptomatic MM, and extra-medullary myeloma/plasma cell leukemia (PCL). MGUS, like MM, produces a typical M-spike (generally non-IgM) by serum protein electrophoresis or free light chain in the urine. The MGUS condition is age dependent, affects about 4% of individuals over the age of 50, and can progress to MM at average rates of 1% per year[11, 12]. MGUS can be discriminated from MM for having a M-spike of <30 g/L, less than 10% of tumor cells among BM mononuclear cells, and no end organ damage or other symptoms. Progression of MGUS to sMM and symptomatic MM is combined with an expanding BM tumor bulk and increasingly severe organ impairment or symptoms[2]. Notwithstanding the marked dependence of MM cells on the BM tumor microenvironment, in the advanced stages of the disease the more aggressive tumors may sometimes spread to extramedullary locations, such as spleen, liver, and extra-cellular spaces. Extramedullary MM (EMM) can also manifest a leukemic phase, which is classified as secondary or primary PCL, based on the presence or absence of a preceding intramedullary MM. Most of the available human myeloma cell lines (HMCLs) have been generated from EMM or PCL tumors and represent a renewable source of the oncogenic events involved in initiation and progression of the most aggressive end-stage MM tumors[13, 14] (**Figure 2**).

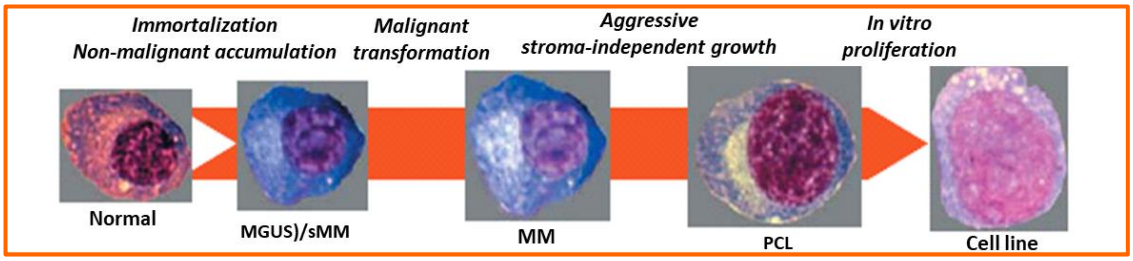

Figure 2. Multi-step molecular pathogenesis of MM. Progression through the different stages of the desease.

To date, MM diagnosis follows the criteria established by the International Myeloma Working Group in 2003[15] subsequently updated in 2009[16]. Osteolytic bone lesions and/or compression fractures detected on routine radiographs, magnetic resonance imaging, or computed tomographic scans are the most characteristic markers of end-organ damage in myeloma. Anemia (occurring in 70% of patients), renal failure (50%), and hypercalcemia (25%) are the other established markers of end-organ damage. The presence of one or more of these four markers that is felt to be related to the underlying PC disorder is required for the diagnosis of the disease (**Table 1**).

Table 1. Mayo Clinic diagnostic criteria for selected clonal plasma cell disorders.

DISORDER	DISEASE DEFINITION
Monoclonal Gammopathy of Undetermined Significance (MGUS)	▪ Serum monoclonal protein < 3 g/dL ▪ Bone-marrow plasma cells <10% ▪ Absence of end-organ damage such as lytic bone lesions, anemia, hypercalcemia or renal failure
Smoldering Multiple Myeloma (sMM; asymptomatic MM)	▪ Serum monoclonal protein (IgG or IgA) ≥3 g/dL and/or ▪ Bone-marrow plasma cell ≥10% ▪ Absence of end-organ damage such as lytic bone lesions, anemia, hypercalcemia or renal failure
Multiple Myeloma (MM)	▪ Bone marrow plasma cells ≥10% ▪ Presence of serum and/or urinary monoclonal protein (except in patients with true non-secretory multiple myeloma) ▪ Evidence of lytic bone lesions, anemia, hypercalcemia or renal failure

Plasma Cell Leukemia

PCL is the most aggressive presentation of the PC tumors, characterized by circulating PCs >2×10^9/L in peripheral blood or by a relative plasmacytosis >20% of blood leukocytes[17, 18]. While primary disease (pPCL) shows as de novo leukemia, secondary leukemic transformation (sPCL) arises from pre-existing MM, probably as a consequence of clonal transformation. The occurrence of PCL is a rare event: the incidence of sPCL in MM patients is only of 1.8–4%, similar to the occurrence of pPCL. Around 60% of patients present a primary form and the remaining 40% occur as complication during the course of MM[19]. The median age of patients with pPCL ranges between 52 and 65 years, about 10 years younger than the median age of 65 to 70 years observed in the general myeloma population and in sPCL.

Genetic architecture of Multiple Myeloma

Several pieces of evidence support the hypothesis that the oncogenic transformation in MM occurs within secondary lymphoid organs. MM is characterized by an abnormal proliferation of aberrant

PCs that generally feature a high rate of somatic mutations on *IGH*/*IGK*/*IGL*, suggesting that these cells were exposed to the AID activity and SHM GC process. In addition, the nature of monoclonal Ig is essentially IgG and IgA, rarely IgD or IgM, being another indirect evidence of MM plasma cell exposure to CSR. A further proof that straightened the MM post GC origin, is represented by the recurrence of genomic events, in particular translocations, involving the *IGH* gene[3] (see **Table 2**). Interestingly, from the analysis of *IGH* translocation breakpoints, it was shown that the t(4;14) largely involves the switch regions, suggesting errors during the CSR process[6, 20] whereas the t(11;14) may result from errors during the SMH[20]. Basing on the breakpoints analysis, a recent paper from the UK group[3] suggested that ~20% of t(11;14) and t(14;16) translocations may take place within the BM during the maturation *IGH* rearrangements (**Figure 3**).

Table 2. Genetic events underlying the initiation and progression of myeloma to plasma cell leukaemia[3].

PRIMARY GENETIC EVENTS (% OF TUMORS)	SECONDARY GENETIC EVENTS (% OF TUMORS)
IGH@ translocations and genes affected	**Molecular hallmarks**
• t(4;14): FGFR3 and MMSET (11%)	• Immortalization
• t(6;14): CCND3 (< 1%)	• G1/S abnormality (CDKN2C, RB1 (3%), CCND1 (3%) and CDKN2A)
• t(11;14): CCND1 (14%)	• Proliferation (NRAS (21%), KRAS (28%), BRAF (5%) and MYC (1%))
• t(14;16): MAF (3%)	• Resistance to apoptosis (PI3K and AKT)
• t(14;20): MAFB (1.5%)	• NF-κB pathway (TRAF3 (3%), CYLD (3%) and I-kB)
Hyperdiploidy	• Abnormal localization and bone disease (DKK1, FRZB and DNAH5 (8%))
• Trisomies of chr 3, 5, 7, 9, 11, 15, 19 and 21	• Abnormal plasma cell differentiation (XBP1 (3%), BLIMP1/PRDM1 (6%) and IRF4 (5%))
	• Abnormal DNA repair (TP53 (6%), MRE11A (1%) and PARP1)
	• RNA editing (DIS3 (13%), FAM46C (10%) and LRRK2 (5%))
SECONDARY GENETIC EVENTS (% OF TUMORS)	• Epigenetic abnormalities (UTX (10%), MLL (1%), MMSET (8%), HOXA9 and KDM6B)
	• Abnormal immune surveillance
Gains	• Abnormal energy metabolism and ADME events (absorption, distribution, metabolism and excretion)
• 1q: CKS1B and ANP32E (40%)	**Deletions**
• 12p: LTBR	• 1p: CDKN2C, FAF1 and FAM46C (30%)
• 17q: NIK	• 6q (33%)
Secondary translocations	• 8p (25%)
• t(8;14): MYC	• 11q: BIRC2 and BIRC3 (7%)
• Other non-IGH@ translocations	• 13: RB1 and DIS3 (45%)
Epigenetic events	• 14q: TRAF3 (38%)
• Global hypomethylation (MGUS to MM)	• 16q: CYLD and WWOX (35%)
• Gene-specific hypermethylation (MM to PCL)	• 17p: TP53 (8%)

It is still not clear which and how specific pre GC events may lead to a post GC disease. However, it was described that >4% of MM patients have two unrelated clones with different VDJ rearrangements, suggesting the potential existence of an aberrant genomic process during the early phase of B-cell maturation. Differently from other cancers, MM is not characterized by common and distinct "driver" events (i.e substitution). The only recurrent and potentially driver MM events are hyperdiploidy, *IGH* translocations and specific deletions. The hyperdiploid (HRD)

cytogenetic profile is observed in up to 55% of the patients[21-25], with 48–75 (mostly 49–56) chromosomes, usually with extra copies of three or more specific odd chromosomes (3, 5, 7, 9, 11, 15, 19, and 21).

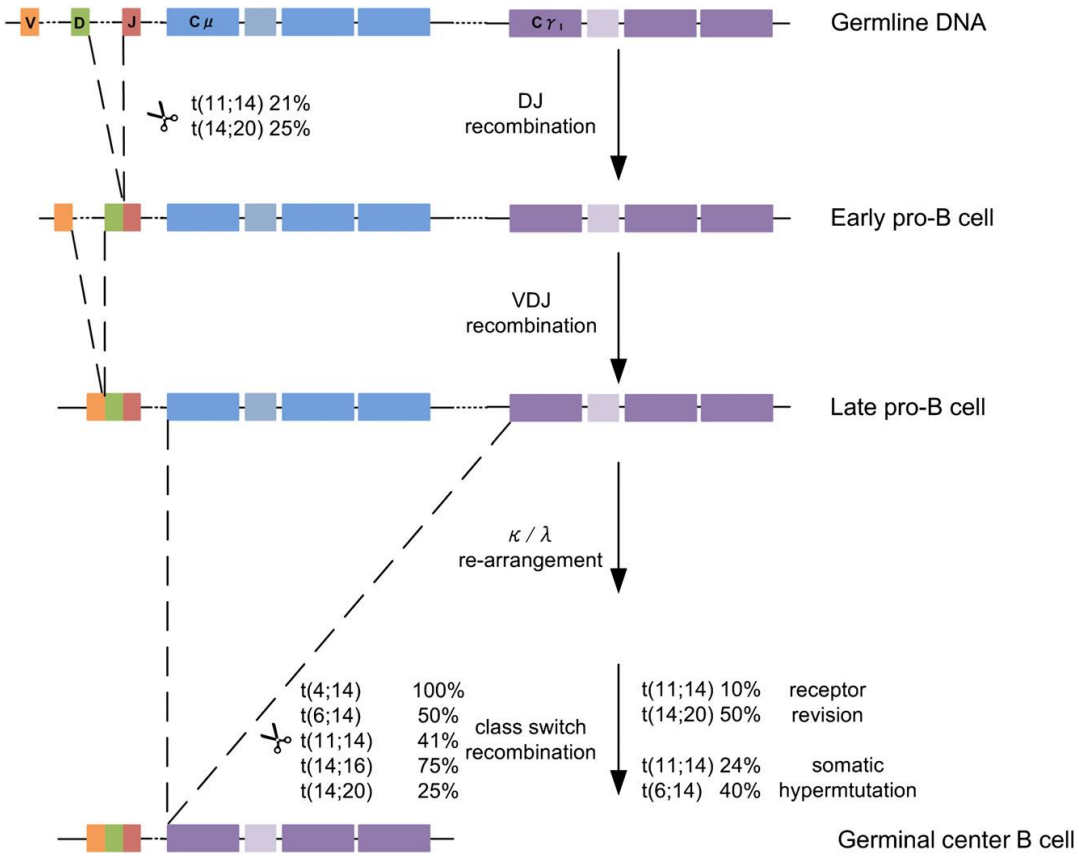

Figure 3. Summary of translocation events in myeloma. Translocations occur between the *IGH* locus and all main partner chromosomes through CSR in the germinal center. t(11;14) and t(14;20) subgroups also have other mechanisms of developing translocations through receptor revision in the germinal center or through DH-JH recombination at the early pro–B-cell stage[20].

Hyperdiploidy is probably due to an aberrant missegregation of chromosomes during mitosis, but the mechanisms responsible for these events are still unknown. In any case, HRD patients seem to have a better prognosis than non-HRD patients[26].

The second pathway is based on *IGH* translocations, observed in 40% to 50% of the patients and almost all the HMCLs[27, 28]. Generally, *IGH* translocations are mutually exclusive with HRD cytogenetic profile. Some of these 14q32 translocations are recurrent (**Table 2**), others being apparently random[29]. These rearrangements generally determine an overexpression of the partner genes under dependence of the strong *IGH* enhancers, leading to overexpression of the targeted proteins. The most frequent one is the t(11;14)(q13;q32), observed in 15%-20% of the patients[30]. The t(11;14) dysregulates the *CCND1* gene, leading to its overexpression. In frequency, the second

recurrent *IGH* translocation is the t(4;14)(p16;q32), observed in 10%-15% of the patients. This translocation is very peculiar in B-cell malignancies, by disrupting and upregulating two genes, *FGFR3* and *MMSET*/*WHSC1*, creating a fusion transcript with this latter gene[31, 32]. *FGFR3* is an oncogene activated by mutations in several solid tumor types, and it is upregulated only in 70% of MM patients harbouring the t(4;14), owing to an unbalanced translocation with loss of the telomeric part of chromosome 4, bearing *FGFR3*[33-35]. This finding suggests that the main molecular target of the translocation is *MMSET* whose upregulation leads to methylation of several proteins in the genome[36]. Other recurrent translocations are rarely observed in more than 3% of the patients, such as the t(14;16)(q32;q23), which dysregulates the *MAF* oncogene[31, 37]; the t(14;20)(q32;q11), dysregulating the *MAFB* oncogene[38]; or the t(6;14)(p21;q32), which upregulates the *CCND3* gene[21]. Interestingly, some recent next generation sequencing (NGS) data have suggested that other translocations *IGH*-independent may be involved in MM pathogenesis. The most important example is represented by translocations involving *MYC* locus, that are generally considered as secondary events[39]. Translocation between 8q24 (*MYC*) and 14q32 (*IGH*) is detectable in 5-10% of patients, but approximately 50% of all MM are characterized by a rearrangement involving *MYC*. Apart *IGH*, the most frequent *MYC*-translocation partners are *IGK*, *IGL*, *XBP1*, *FAM46C*, *CCND1* and *KRAS*[39]. Overall, these translocations lead to an aberrant *MYC* overexpression. This is one of the main proofs of the presence of other *IGH*-independent translocations that significantly contribute to the final proliferation and behaviour of MM. All these non-*IGH* translocations are generally supposed to be secondary events in the MM oncogenesis[3]. The main reason supporting this hypothesis is that most of them are often observed only at subclonal levels, in contrast to *IGH* translocations, which are usually present in virtually 100% of patient's PCs. The third recurrent aberration event in MM is represented by distinct deletions/gains. The most important and frequent are: the monosomy 13 (45% of the patients)[40], duplication of the long arm of chromosome 1 (1q gains, 30% to 35% of the patients)[41], and deletion of chromosome 1p. These two recurrent copy number changes on chromosome 1 are particularly relevant being associated with an inferior survival. Different other deletions involving the 6q, 8p, 12p, 14q, 16q, 17p, or 20p regions have been described in a significant fraction of patients[42]. All these abnormalities have been known for a long time because they are visible on the conventional karyotype and on fluorescence in situ hybridization (FISH). Thanks to comparative genomic hybridization (CGH) and single nucleotide polymorphism (SNP)–array technologies other important chromosomal changes have been revealed, especially homozygotic deletions (**Table 2**). Taken together, all these genomic data have been the scientific proof of evidence of the current model of MM pathogenesis (**Figure 4**).

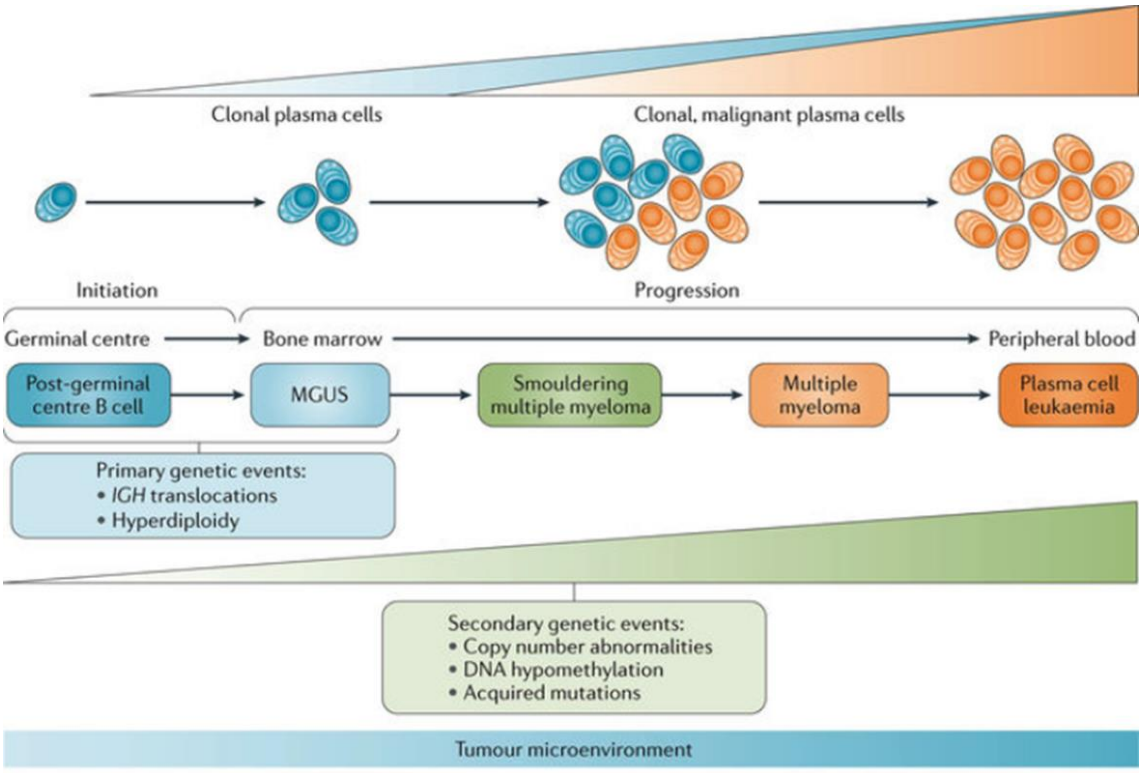

Figure 4. Multiple Myeloma oncogenetic model. Quickly after the oncogenetic event(s), differential mutations occur, creating subclones. Secondary events occur later during evolution.

Clonal heterogeneity and clonal evolution

MM has been considered as a clonal malignancy characterized by the secretion of a monoclonal Igs or monoclonal free light chains. Although we are able to detect and to describe different clonal genomic aberrations (*IGH*/*IGK*/*IGL* hypermutation, *IGH* translocations, clonal copy number changes etc), a more complex situation has been recently revealed. In 2012, three publications reported a certain degree of heterogeneity within the dominant clone, by comparing DNA abnormalities present at diagnosis and at relapse[43-45]. Based on SNP array or high-throughput sequencing, all three reports demonstrated that the clone observed at relapse could be significantly different from the one observed at diagnosis. In some cases, the relapsed clone was characterized by an increased genomic complexity, representing a classical clonal model. During this darwinian selection, the therapy resistant subclones survive and proliferate, becoming the dominant clone responsible of disease relapse[3]. More strikingly, in some patients, the relapsed clone presented pieces of DNA, which were absent at the time of diagnosis. Because *ex nihilo* creation of DNA is impossible, the conclusion of these studies was that minor subclones, undetectable at diagnosis by classical methods, were responsible for relapse. The DNA abnormalities defining these subclones were mainly supposed to be secondary changes. However, a study focused on the t(4;14), supposed to

be a primary driver genetic event, showed that this translocation could be present only in a small minor subclone at diagnosis, but be the main clone at relapse, and *vice versa*[46]. Thus, the growth of subclones is a very early event in oncogenesis, occurring probably shortly after the cell transformation (**Figure 5**).

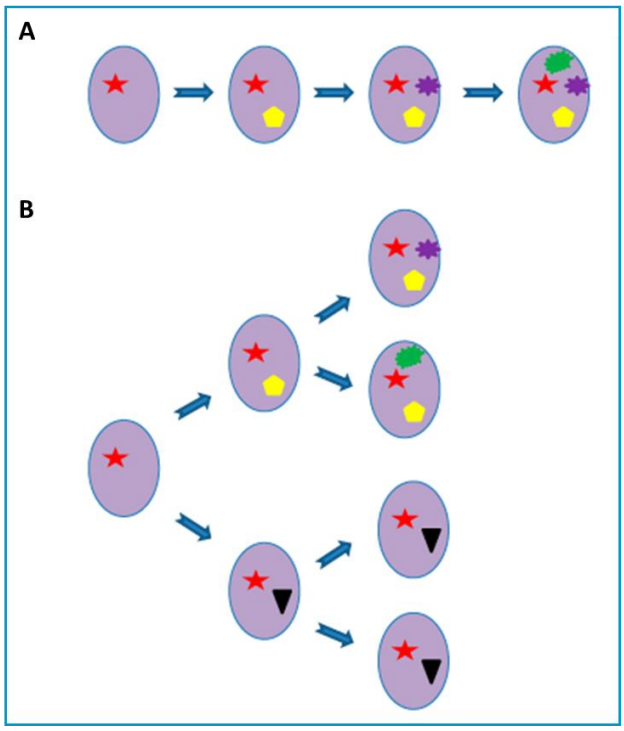

Figure 5. Subclonal evolution. Tumor progression may be characterized by two different evolution models: a "linear" evolution with accumulation of genetic events **(A)** and a "branching" evolution with early separation of subclones with different mutations, which undrego differential selection during evolution **(B)**.

Nonetheless, based on combined whole exome sequencing (WES) and single-cell genetic analysis it was suggested that up to six different major malignant clones, related by either linear or branching evolution, may be present in MM at onset (**Figure 5**)[47]. Interesting to note, the study of Melchor [48] indicated that the activation of the mitogen-activated protein kinase (MAPK) pathway could be the result of a convergent phenotype and be explained as a form of parallel evolution that leads independent clones, originating from a common ancestor, to acquire activating mutations in the same pathway or in genes included in the same pathway[48].

Despite all these progress in MM knowledge, there are still many unanswered questions. First, it is still unknown whether several subclones coexisting at the same time in a single patient have differential locations. This situation is well described in solid tumors, in fact a branching clonal evolution is probably a common feature of various solid[49, 50] and hematological tumors (such as acute myeloid leukemia)[51]. However, no data are currently available in MM. Another unresolved question is if the subclones selection can solely occur under therapeutic pressure, or does it may

occur following a spontaneous pattern during the natural history of the disease[1]. In the "Darwinian model" of intraclonal heterogeneity, multiple clones originate from the same progenitor, each of which has the same main feature but also harbors a set of distinctively acquired mutations: the prevalence of specific clones ("clonal tiding") over the others depends on the selective pressure determined by the microenvironment (the BM niche) or the treatment. This hypothesis further supports the idea that the impact of MM therapy is also related to the underlying biology of the disease, and may also explain why high-risk patients have significantly shorter progression-free and overall survival than low-risk patients despite their identical response rates[3, 52].

Multiple myeloma: a single or multiple disease?

As widely described above, MM represent a very heterogeneous disease and this was confirmed by all possible levels of investigation (clinical, cytogenetical, or molecular). Missing a specific common driver genomic hit and considering this genomic complexity, different groups investigated the existence of MM sub entities. Similarly to other haematological disease, gene expression profiling (GEP) has represented the most used approach for these investigations. In particular, global gene and miRNA expression and genome-wide DNA profiling have been used alone or as integrated tools to investigate the genomic instability underlying the biological and clinical heterogeneity of MM[21, 53, 54]. Generally, all GEP studies described the existence of different MM subgroups, and the majority of them were associated with one of the recurrent cytogenetic features[55-58] (**Figure 6**).

For instance, several global miRNA-profiling studies with impact on gene expression, biological relevance, and survival have been published, and imply a possible association with myeloma pathogenesis and molecular sub-entities in terms of specific chromosomal aberrations or gene expression-based high-risk groups[59].

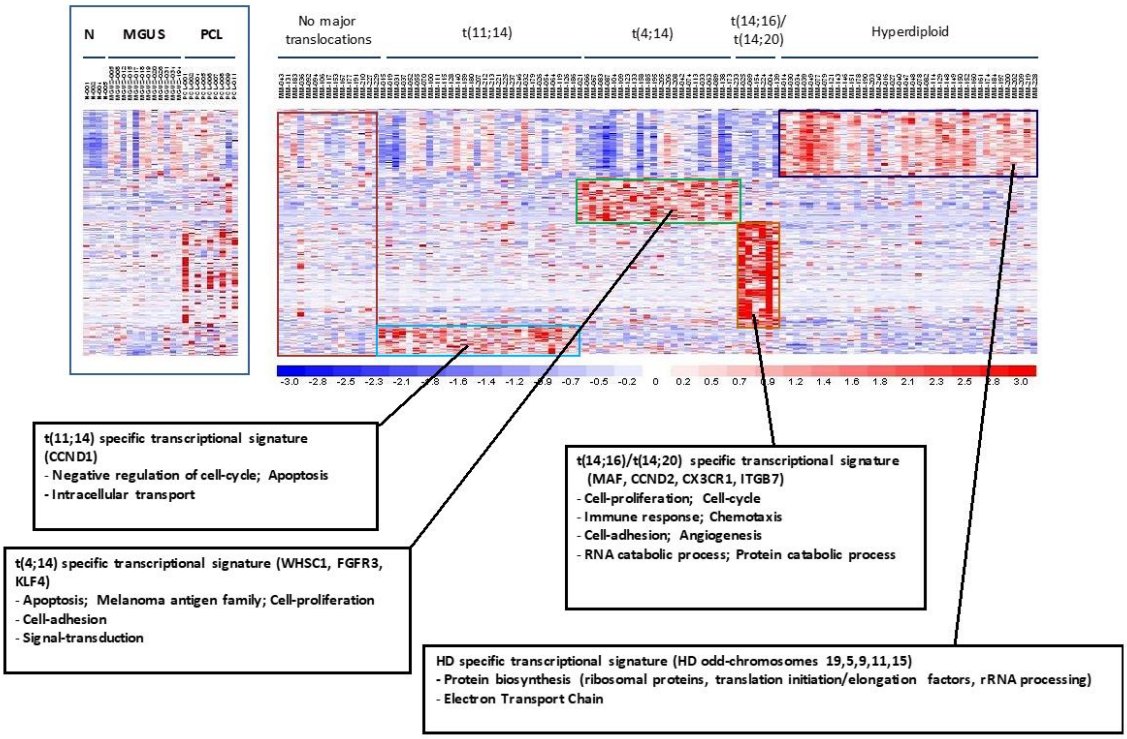

Figure 6. Distinct molecular types of MM defined by FISH are associated with specific gene expression profiles[60, 61].

The biological heterogeneity of MM was further confirmed by CGH/SNP-array analyses. Apart from the HRD group (which appears also heterogeneous based on the trisomy combinations), some other distinct subgroups were identified. Similarly to GEP, all these groups were generally related with the most common copy number changes such as monosomies 13 and 14, 1p deletions and 1q gains. Whether these subgroups represent different myeloma entities remains to be explored. However, it is quite clear that these approaches confirmed the MM heterogeneity and generally reproduced the different effect of known recurrent aberrations.

Sequencing analyses

Recently, thanks to novel parallel sequencing technology, we significantly improved our understanding of different types of cancers, including MM. In the last years, several studies have shown that whole-genome sequencing (WGS) and WES are cost effective means of identifying protein-coding mutations. While WGS allows the unguided entire DNA sequencing of an organism's genome, WES is the most efficient way to identify genetic variants in protein-coding genes[62]. Three major papers investigated by NGS techniques the mutational landscape of MM[63-65]. The first study of Chapman *et al.*[63] analysed whole genome and exome sequence data of 38 patients at different disease stages (diagnosis, relapse) and indicated that the distribution of mutations was widely

variable between patients, confirming the large heterogeneity observed at the clinical or biological level. Larger exomes studies confirmed this mutational landscape[64, 65] (**Figure 7**): the high frequency of nonsynonymous variants in MM and the progressive increase of non-silent mutations from asymptomatic to extramedullary disease (from <10 to almost 60 median mutations across MGUS, high-risk sMM, MM, and PCL)[66]. A further increased mutational load has been observed in myeloma cell lines[67]. When compared with other tumors, MM is in the middle, between low-mutated tumors, such as leukemia, and highly mutated carcinogen-induced tumors[68].

Differently from other haematological cancers where NGS approaches allowed to detect a universal common driver mutation event (*BRAF* for Hairy Cell Leukaemia and *MYD88* for Waldenstrom macroglobulinemia)[69, 70], in MM all the three studies failed to detected a specific driver event shared by all patients. Few mutations were described as recurrent without exceeding more than 30% of all patients (i.e *NRAS* and *KRAS*)[63-65, 71]. Several others mutations were observed in a few percent of the patients, but 30 genes were found recurrently mutated in at least 5% of the patients. Of note, many of these mutations are observed only in subclones, including genes supposed to act as drivers (*NRAS*, *KRAS*, *BRAF*, for instance)[64, 65] (**Figure 7**).

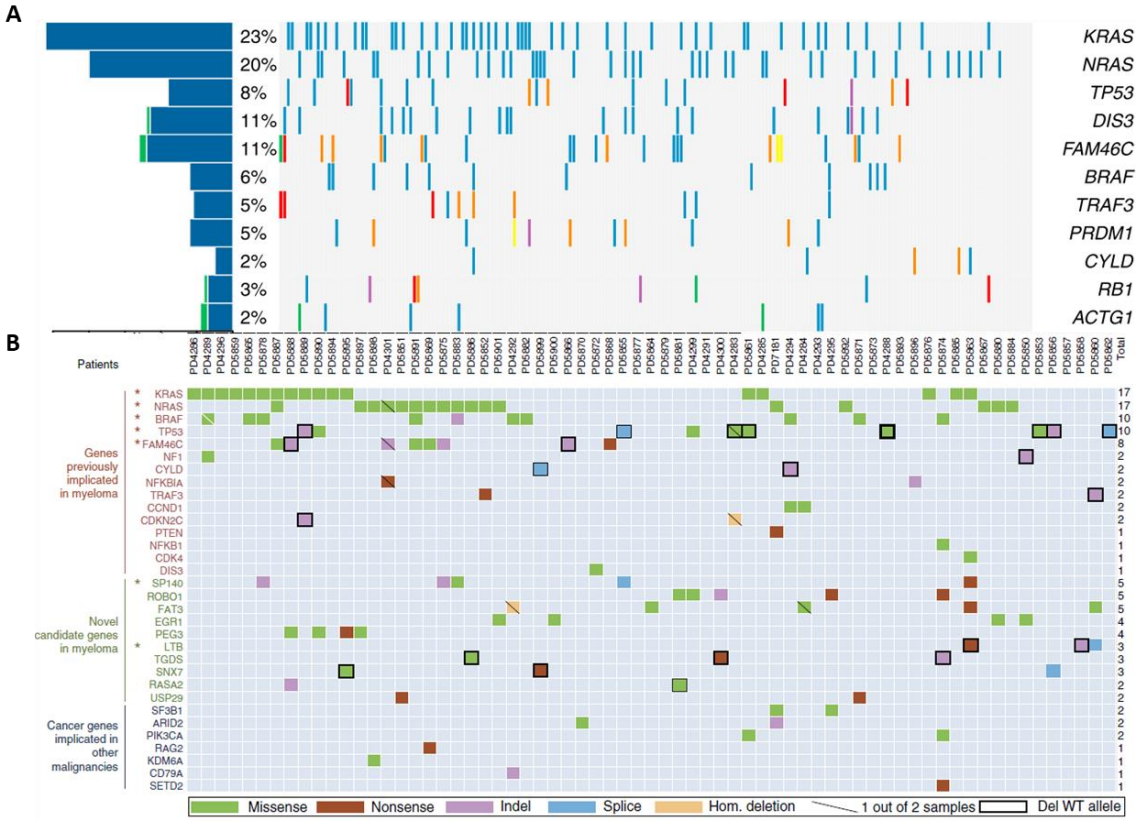

Figure 7. Landscape of genetic alterations and recurrently mutated genes in multiple myeloma in 2 datasets: **(A)** Lohr *et al.*[65] and **(B)** Bolli *et al.*[64].

Among the recurrently mutated genes with putative oncongenic role, the most frequent are *KRAS, NRAS, TP53, BRAF, TRAF3, FAM46C* and *DIS3*, confirming the main involvement of both tirosine kinase receptor and NF-kB pathways. Mutations in *NRAS, KRAS* and *TP53* genes are recognized markers of advanced disease in MM, and they had been already described as frequent by conventional sequencing[3]; overall, they occur in up to 50% of MM patients if not more.

BRAF mutations in MM

Genome-wide NGS studies have provided a rationale for molecularly aimed treatment approaches by identifying specifically targetable mutations such as those in the RAS/MAPK pathway, which are the most prevalent mutations in MM[63-65, 72]. The *BRAF* gene maps to chromosome 7 (7q34) and encodes a protein belonging to the RAF family of serine/threonine protein kinases, involved in the regulation of the MAPK/ERK signaling pathway, that affects cell division, differentiation, and secretion. Along with already known *NRAS* and *KRAS* mutations, *BRAF* mutations, recently reported to occur in 4-15% of patients[63-65], may be of potentially immediate clinical relevance because of the availability of effective BRAF inhibitors that are also being investigated in MM treatment[73-75]. The substitution of valine 600 to glutamic acid (V600E) has been reported pathognomonic in B-cell hairy cell leukemia[69], accounting for about 30%[64] to 55%[65] of observed *BRAF* variants.

DIS3 mutations in MM

DIS3 encodes a highly conserved ribonuclease, which serves as the catalytic component of the exosome complex, prominently participating to RNA processing and turnover, and it maps to chromosome 13 (13q22.1), that is found deleted in 50% of MM patients. The protein is composed of two functional domains: the N-terminal PIN (spanning from exon 1 to exon 4), with endonuclease activity, and the RNB (spanning from exon 10 to exon 18) at the C-terminal, with exonuclease activity, this the most important for cell physiology. The exosome plays a fundamental role in most RNA metabolic pathways. DIS3 is indispensable for survival in vertebrates and its function cannot be substituted by other homolog exosome-associated nucleases[76]. Somatic mutations in *DIS3* and altered expression have been reported in several cancer types (COSMIC catalogue, Release v71, http://cancer.sanger.ac.uk/cancergenome/projects/cosmic/)[77, 78] and it is recurrently mutated on average in 11% of MM patients[63, 65, 79, 80]. *DIS3* mutations in MM prevalently affect the RNB domain. Preliminary functional analysis suggested that MM-associated mutations in *DIS3* RNB domain (inhibiting its exonuclease activity) are indeed synthetically lethal with inactivation of DIS3 endonucleolytic activity, consequent to the disruption of endoribonucleolytic domain functionality. Furthermore, very recently our group demonstrated that DIS3 facilitates the maturation of the tumor suppressor *let-7* miRNAs by reducing in the cytoplasm the RNA stability of the pluripotency

factor LIN28B, an inhibitor of *let-7* processing. Through the reduction of mature *let-7*, DIS3 enhances the translation of let-7 targets such as *MYC* and *RAS* leading to enhanced tumorigenesis[81].

FAM46C mutations in MM

FAM46C recently emerged as a commonly mutated gene in MM (~5-13%)[63-65, 71, 79]. Although its function is still unknown, correlations have been found between the expression of *FAM46C* and that of ribosomal proteins and eukaryotic initiation and elongation factors involved in protein translation, prompting for a functional role in the regulation of translation. *FAM46C* maps to 1p12, frequently deleted in MM (more than 20% patients), in most of the cases as part of a larger deleted region extending telomerically to the 1p32.3 region, associated with poorer prognosis[82]. *FAM46C* alterations (either including loss-of-heterozygosity or mutations) were described to be associated with impaired overall and progression-free survival[83].

FAM46C appears to be a type I interferon-stimulated gene (ISG) and it encodes a protein that seems to silence the replication of viruses and to have a role in both protein translation and as a stability factor for mRNA.

Undoubtedly, recurrent mutations in *DIS3* and *FAM46C* prompted for further investigation of altered RNA-processing machinery and protein homeostasis mechanisms in MM.

TP53 mutations in MM

The tumor suppressor protein p53, encoded by the *TP53* gene at chromosome 17p13, mediates the response to various stress signals (including DNA damage, oxidative stress, ribonucleotide depletion, and deregulated oncogene expression), many of which are encountered during tumor development and malignant progression[84]. Loss of p53 function, due to *TP53* deletions and/or mutations or by defects in the signaling pathways upstream or downstream of p53, is associated with oncogenesis, cancer progression and drug resistance. *TP53* is mutated in about half of human cancers, and the prevalence of gene mutations greatly varies between different tumor types.

Recently, WES analyses in MM[64, 65, 71, 79], albeit reporting slightly higher mutational frequencies (probably for the extension of the analysis to the entire coding sequence and the greater sensitivity), confirmed the findings of the early studies[85-88]: *TP53* mutations are relatively rare at presentation (mutation prevalence ranging from 0% to 9.7% in representative MM patients' cohorts). The frequency of mutations increases with disease stage, reaching 25-30% in PCL[89, 90] and 80% in HMCLs[91].

A strong association has been described between *TP53* mutation and del(17p)[92]. Deletions, predominantly monoallelic, of chromosome 17p13 region containing the *TP53* gene locus occur in about 10% of untreated MM cases[40, 93, 94]; the incidence rate reported in PCL ranges from 35% to

75%[90, 95]. 17p13 deletion confers a very negative effect on survival[96], displaying the most powerful cutoff for predicting survival if the deletion is carried by more than 50% of malignant PCs[97]. Finally, a recent study identified *TP53* as the critical gene of 17p13 deletion in MM[98].

The liquid biopsy

Cancer treatment relies on the availability of accurate molecular profiles of individual tumors, which are currently determined by genetic analysis of DNA samples extracted from the neoplastic tissue. Tumor genotyping of mature B-cell tumors lacking a leukemic phase has so far relied on the analysis of the diagnostic tissue biopsy; in fact, MM characterization currently depends on BM aspirates for mutational analysis. However, multiregional sequencing of cancers showed that this approach might not capture the putative spatial and genetic heterogeneity of this multi-focal disease because the diagnostic tissue biopsy might be subjected to a selection bias resulting from spatial heterogeneity and, therefore, might not be representative of the entire tumor genetics. Indeed, different areas of the same tumor may show different genetic profiles (i.e. intratumoral heterogeneity). A biopsy from one part of a tumor may miss mutations occurring in subclones, including clinically relevant genetic biomarkers for treatment tailoring or anticipation of resistance[99].

In particullary, MM genotyping has so far relied on the analysis of purified PCs from the BM aspirate, which may fail in capturing the postulated spatial heterogeneity of the disease and imposes technical hurdles that are limiting its transfer in the routine and clinical grade diagnostic laboratory. In addition, longitudinal monitoring of disease molecular markers may be limited by patient discomfort caused by repeated BM samplings during disease course. Data produced in solid cancers indicate that circulating tumor DNA into the peripheral blood (PB) can be used as source of tumor DNA to provide information about tumor mass, residual disease and tumor genotype, with obvious advantages in terms of accessibility.

Cell-free DNA (cfDNA) fragments are shed into the bloodstream by cells undergoing apoptosis (**Figure 8**). Since these processes occur ubiquitously in the body, the blood plasma of healthy individuals by default contains circulating cfDNA at steady-state levels and primarily derived from hematopoietic cells[100]. In cancer patients, cfDNA also contains, besides cfDNA derived from nonmalignant cells, a fraction of circulating tumor DNA (ctDNA) as tumor cells also shed DNA into the circulation[101].

Information acquired from a single biopsy provides a spatially and temporally limited snapshot of a tumour and might fail to reflect its heterogeneity. Tumour cells release ctDNA into the blood, but

the majority of circulating DNA is often not of cancerous origin, and detection of cancer-associated alleles in the blood has long been impossible to achieve.

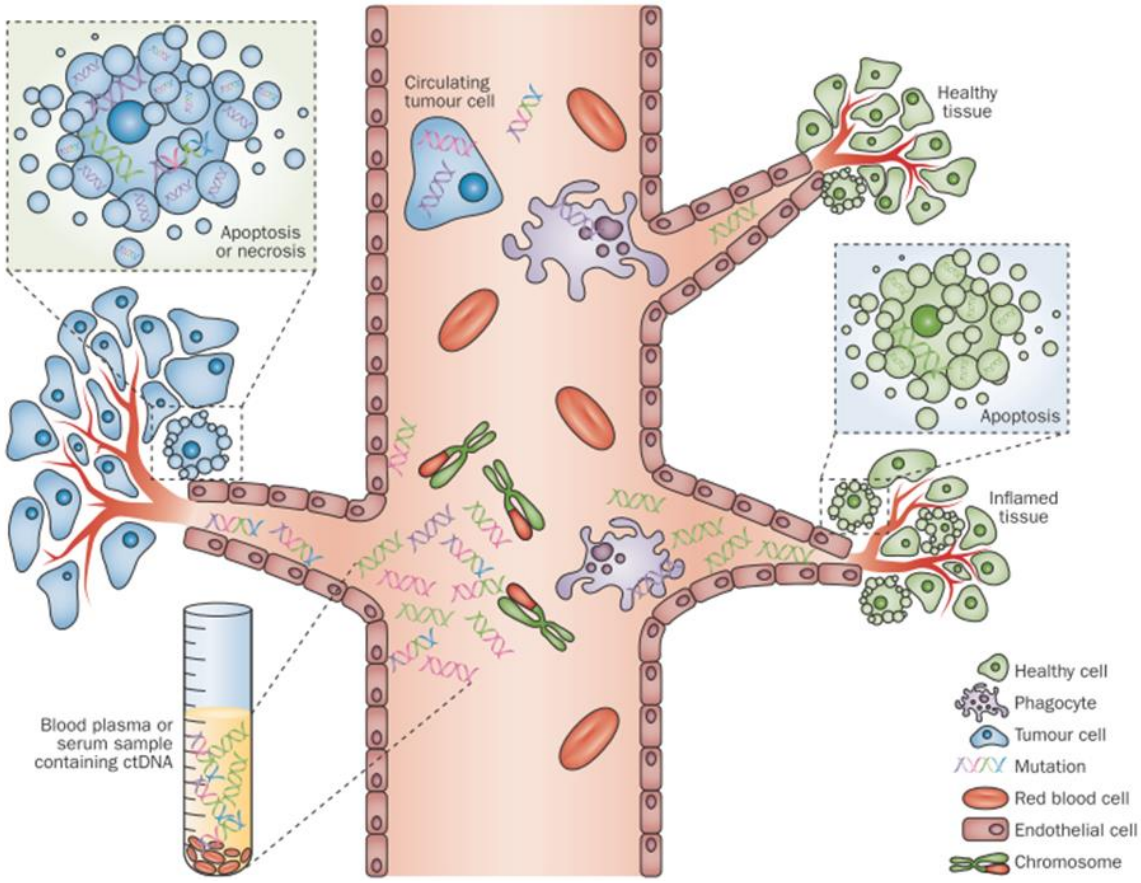

Figure 8. Release and extraction of cfDNA from the blood. cfDNA is released from healthy, inflamed or diseased (cancerous) tissue from cells undergoing apoptosis or necrosis[102].

Recent advances in the sensitivity and accuracy of DNA analysis have allowed for genotyping of cfDNA for the identification of somatic cancer gene mutations[103]. The capability to detect and quantify tumor mutations has being proven effective in tracking tumor dynamics in real time as well as serving as a "liquid biopsy" that can be used for a variety of clinical and investigational applications not previously possible. Circulating DNA is representative of the entire tumor heterogeneity, thus allowing to bypass the bias imposed by BM biopsies in the reconstruction of the entire cancer clonal architecture, and to identify resistant clones that are dormant in non-accessible tumor sites. Accessing the blood stream has also clear sampling advantages in the serial monitoring of treatment emergent resistant mutations in real time[103].

In recent months, the first promising studies on the circulating ctDNA in MM patients have appeared, tracking the clonotypic V(D)J rearrangement as disease fingerprint[104], or genotyping a

highly restricted set of cancer genes that were not specifically addressed to resolve typical MM mutational landscape[105-107].

Conversely, these studies have been the proof that cfDNA can be also used as a "liquid biopsy" to gain a representative snapshot of the underlying tumor genetics for diagnostic or prognostic purposes, to inform treatment decision, to track clonal evolution-driven resistance upon treatment, and in the end, to inform patient management in real time, still lacking in many mature B-cell tumors and, in particular, in MM.

Aim of study

In the last decade, important advances in molecular cytogenetics and global genomic studies brought a significant progress in the comprehension of MM pathogenesis. By combining global GEP and FISH analyses, our laboratory and other groups have contributed to define the transcriptional profiles associated with MGUS, MM and PCL, and those related to the major molecular subclasses of MM[53, 61, 108, 109]. Furthermore, this integrated genomics approach was developed to detect allelic imbalances in MM[110], thus revealing a wide gene dosage effect.

Recently, our group performed genome-wide studies, along with global gene and miRNA expression profiling analyses on a prospective series of pPCL at diagnosis, enrolled in an Italian multicenter clinical trial, thus revealing genomic and miRNA/gene expression patterns distinguishing pPCL from MM, and specific molecular signatures associated with clinical outcome[95, 111, 112]. More recently, a whole exome sequencing study has been performed on cohort fraction of the same pPCL cases, thus evidencing a heterogeneous mutational pattern in pPCL, involving several pathways with a central role in the pathogenesis of PC dyscrasia[89].

The first aim of this project was the evaluation of the incidence of mutations in MM driver genes (*KRAS*, *NRAS*, *TP53*, *BRAF*, *FAM46C* and *DIS3*), as suggested by several recently previous massive DNA sequencing studies[63-65], in a large and representative cohort of patients at different stages of PC dyscrasia. DNA sequencing of gene mutation hot-spots was performed by means of deep NGS on Roche Junior 454 platform, able to detect also mutations present at subclonal level and undetectable by conventional sequencing. Obtained data have been subsequently validated by means of Sanger sequencing or independent 454 re-sequencing. The research activity also aimed to identify possible correlations between the presence of driver mutations and the main MM genetic lesions, namely *IGH* translocations, hyperdiploidy, deletions of chromosomes 1p, 13q and 17p, and 1q gain; as well as the occurrence of mutations in major and minor subclones. Mutational patterns were also related with gene and miRNA expression profiles, available for the great majority of patients, to understand better their role in the biology of the disease and clinical outcome. The integration of mutational, genomic and transcriptomic data should hopefully allow to better stratifying MM patients at diagnosis.

Additionally, to investigate the longitudinal state of driver gene mutations, we analyzed a small set of MM patients for whom BM specimens were collected at two different time points, to gain further insight on the evolution of MM clonal structure. We also performed deep sequencing of cDNA of the mutated cases to verify if the mutations detected on genomic DNA are also expressed at transcriptomic level.

The second aim of the project was based on the observation that circulating DNA is a powerful non-invasive biomarker currently used to identify somatic cancer gene mutations, also due to its

representativeness of the entire tumor heterogeneity. In MM, this peculiarity would allow to bypass the bias imposed by BM biopsies in the reconstruction of the entire cancer clonal architecture, and to identify resistant clones that are dormant in not accessible tumor sites.

Our working hypothesis was that cfDNA represents an accessible source of tumor DNA for the sensitive identification of genetic biomarkers that refine the diagnostic workup, stratify prognosis and inform treatment decision in MM. The project aimed at comprehensively validating the "liquid biopsy" as a diagnostic tool to inform in real time MM diagnosis, prognostication and treatment monitoring. The task was designed as a retrospective observational study by the analysis of a consecutive series of MM patients, whose plasma cfDNA, tumor and germline genomic DNA (gDNA) were analysed, by means of targeted ultra-deep-NGS. An updated targeted resequencing gene panel including the coding exons and splice sites of genes that are recurrently mutated in MM had been specifically designed for this project in collaboration with Prof. Davide Rossi at the Institute of Oncology Research (IOR) in Bellinzona (CH). Circulating tumor DNA analysis could represent a less expensive and sustainable approach to determine MM patients' subgroups that could obtain the greatest benefit with specific chemoimmunotherapy regimens.

Materials and Methods

Patients

Pathological BM specimens from 132 patients with newly diagnosed MM, 24 cases of pPCL at onset and 11 cases of sPCL, admitted from July 2001 to April 2014, were obtained during standard diagnostic procedures after the patients had given their informed consent according to institutional guidelines (clearance from Ethic Committee, Fondazione Ospedale Maggiore di Milano). Seventy-nine patients were males; median age was 66 years (range: 42–85). MM and PCL were diagnosed based on the previously described criteria[15, 17, 113]. Many of the patients included in this study have been described in our previous reports[53, 108, 112], and 16 of the 24 pPCL patients were participants in a multicentre clinical trial (RV-PCL-PI-350, EudraCT No. 2008-003246-28)[114], aimed at evaluating safety and anti-tumor activity of combined lenalidomide and dexamethasone as first-line treatment[115].

One hundred patients had an IgG protein monoclonal component; 34 IgA; one IgG/ IgA; and one IgM protein; 101 cases had the light chain κ; 63 λ; two λ+κ. Twenty-five MM cases were in stage IA, 58 IIA/B and 49 IIIA/B. The diagnoses were made by means of conventional morphology and immunophenotype analyses and the patients were clinically staged according to Durie and Salmon criteria [Durie BG, Salmon SE. A clinical staging system for multiple myeloma. Correlation of measured myeloma cell mass with presenting clinical features, response to treatment, and survival. Cancer. 1975; 36:842-854]. No conventional cytogenetic (G-banding) analyses were available.

For the study on cfDNA a total of 28 patients were recruited from September 2016 to May 2017 in Foundation IRCCS Ca' Granda Ospedale Maggiore Policlinico (Milan, Italy) and in Oncology Institute of Southern Switzerland (Bellinzona, Switzerland) if they fulfilled the inclusion criteria: (1) male or female adults; (2) diagnosis of MM or MGUS after pathological revision; (3) evidence of signed informed consent. Patients provided informed consent in accordance with local institutional review board requirements and the Declaration of Helsinki. The clinical prognostic system for the patients were, according to International Staging System (ISS), based on b2-microglobulin and albumin, with stage I representing limited disease, stage II intermediate, and stage III the most unfavorable disease[116, 117].

The following biological material was collected: (1) cfDNA isolated from plasma, (2) tumor gDNA from the CD138+ purified PCs from BM aspiration and (3) normal germline gDNA extracted from PB granulocytes.

Sample preparation and FISH analyses

BM specimens were collected from patients at the time of diagnosis; 19 cases (14 with MM and two with pPCL at onset and relapse; two with MM at onset and at the time of leukemic transformation; and one at early and relapsed leukemic phase) were re-sampled at relapse/MM leukemic transformation, with a median time interval between the two samplings of 30 months.

Tumor gDNA was isolated from PCs purified from the BM samples using CD138 immunomagnetic microbeads as previously described[61, 118]. The purity of the selected PC populations was assessed by means of morphology and flow cytometry, and was >90% in all cases.

PB granulocytes were separated by Ficoll gradient density centrifugation as a source of normal germline gDNA.

All samples were characterized by FISH for the main genomic aberrations: *IGH* translocations (t(4;14)(p16.3;q32); t(11;14)(q13;q32); t(14;16)(q32;q23); t(6;14)(p21;q32) and t(14;20)(q32;q11)), del(13)(q14), del(17)(p13), hyperdiploidy, 1p33 (*CDKN2C*) loss and 1q21.3 (*CKS1B*) gain. FISH analysis for the detection of 13q14 and 17p13 deletions, hyperdiploidy, t(4;14), t(11;14), t(14;16), t(14;20) and aberrations of chromosome 1 were performed using the protocol provided by the manufacturer of the multicolor probes LSI D13S19/LSI 13q34, LSI *TP53*/CEP17, LSI D5S23/D5S721/CEP9/CEP15, LSI *IGH*/*FGFR3*, LSI *IGH*/*CCND1* (XT), LSI *IGH/MAF* (Vysis Inc., DownersGrove,IL), *IGH/MAFB Plus* and *CKS1B/CDKN2C* (P18) (Cytocell Ltd, Cambridge, UK), respectively. To investigated the t(6;14), we performed dual-color FISH using the 973N23 clone (specific for *CCND*3) in combination with PAC clone 998D24 (specific for IGH *locus*). The clones were selected by browsing the University of California, Santa Cruz (Santa Cruz, CA) Genome Database (http://genome.ucsc.edu/)[110].

The cut-off levels considered were 10% for fusion or break-apart probes and 20% for numerical abnormalities, as recommended by the European Myeloma Network (EMN).

DNA extraction and quantification

Genomic DNA was extracted using Wizard genomic purification DNA kit (Promega Corporation, Madison, WI, USA) according to the instructions of the manufacturer and eluted in TE. The concentration of DNA extracted was assessed by spectrophotometric quantification with Nanodrop (Thermo Fisher Scientific Inc, Waltham, MA, USA) and by Quant-iT™ PicoGreen dsDNA Assay kit (ThermoFisher Scientific). All genomic DNA was stored at −20°C until use.

Isolation and analysis of plasma cfDNA

PB (20 ml) was collected in Cell-Free DNA BCT tubes that allow obtaining stable cfDNA samples while preventing gDNA contamination that may occur due to nucleated cell disruption during sample storage, thus avoiding pre-analytical issues affecting cfDNA genotyping. PB was centrifuged at 820 g for 10 min to separate plasma from cells. Plasma was then further centrifuged at 20000 g for 10 min to pellet and remove any remaining cells and stored at -80°C until DNA extraction. CfDNA was extracted from 1-3 ml aliquots of plasma (to allow the recovery of enough genomic equivalents of DNA to reach a genotyping sensitivity of 10^{-3}) using the QIAamp circulating nucleic acid kit (Qiagen) and quantized using Quant-iT™ PicoGreen dsDNA Assay kit (ThermoFisher Scientific). Contamination of plasma cfDNA from gDNA released by blood nucleated cell disruption was ruled out by checking, through the Bioanalyzer (Agilent Technologies) instrument, the size of the DNA extracted from plasma.

Mutation analyses by the 454 GS Junior System

After DNA extraction and quantification, all samples were subjected to polymerase chain reaction (PCR) amplification and deep sequencing by the 454 GS Junior System (454 Life Sciences, Branford, CT, and Roche Applied Sciences, Indianapolis, IN). The 454 GS Junior technology is derived from the technological convergence of emulsion PCR and pyrosequencing. Genomic DNA was amplified using special fusions primers containing genome-specific sequences (reported in **Table 3** for each gene) (Primer-BLAST at http://www.ncbi.nlm.nih.gov/tools/primer-blast/index.cgi?LINK_LOC=BlastHome), and a universal tails at the 5' end to allow the addition of specific MIDs (multiplex identifier sequences used to differentiate samples being run together on the same plate) in a second amplification step (**Figure 9**).

PCR reactions were run in 25µl reaction volumes, containing 10 mmols dNTPs, 10 µM of each primer, 2.5 µl PCR buffer, 12.5 ng DNA, and 0.25 µl of FastStart High Fidelity Polymerase (Roche). Amplicon library A and B sequencing adapters and multiplex identifier (MID) tags were then added to both tails of amplicons by a second amplification step. PCR conditions were as follows: in the first amplification step, denaturating step at 94°C for 5 min followed by 25 cycles at 94°C (30 sec. per cycle), annealing step at 53°C (for DIS3 exon 1), 57°C (for *BRAF*, *NRAS*, *KRAS* and *DIS3* exons 10, 11 and 17), at 58°C (for the remaining exons of *DIS3*, *FAM46C* and *TP53*) (30 sec. per cycle), and extension at 72°C (45 sec. per cycle), followed by a final 7 min extension at 72°C. In the second amplification step, denaturating step

at 94°C for 5 min followed by 25 cycles at 94°C (20 sec. per cycle), annealing step at 55°C (20 sec. per cycle), and extension at 72°C (45 sec. per cycle), followed by a final 10 min. extension at 72°C.

PCR products were visualized on agarose gel, purified using Ampure-XPDNA-binding paramagnetic beads (Agencourt Bioscience Corp., Beckman Coulter S.p.A, Milan, Italy), and quantified in 96-well format using picogreen dye (Life Technologies, Carlsbad, California) and the Victor X2 (Perkin Elmer, Waltham, Massachusetts) fluorometer.

Samples were then diluted to an approximate concentration of 1×10^9 molecules/µL and pooled at equimolar concentrations to create a highly multiplexed amplicon library. After pooling, the library was further diluted to 10^6 molecules/µl and subjected to emulsion PCR (emPCR) at a ratio of 0.8 molecules per bead using the 454 GS Junior Titanium Series Lib-A emPCR Kit (Roche Diagnostics), according to the manufacturer's protocols. Following emPCR, the captured beads with bound DNA were enriched with a second DNA capture mechanism to separate out beads with and without bound emPCR products.

A mean of 500.000 enriched beads was used for massively parallel pyrosequencing in a Titanium PicoTiterPlate® (PTP) with Titanium reagents (Roche Diagnostics), on the GS Junior instrument, according to the 454 GS Junior Titanium Series Amplicon Library Preparation Method Manual (available online: www.454.com). The obtained sequencing reads were mapped to the reference sequence (indicated in table Primer) (RefSeq GRCh37.p13) and analyzed by the Amplicon Variant Analyzer (AVA) software version 3.0 (454 Life Sciences) to establish the mutant allele frequency.

The presence of each obtained non-synonymous variant was verified in an independent PCR product, analyzed by conventional sequencing whenever the sensitivity of the Sanger's method was consistent with the variant allele frequency (VAF). NGS analysis was repeated (run median depth of coverage=200x) in cases of mutations in less than 10% of the DNA molecules to differentiate real mutations from low-level errors introduced during PCR amplification and sequencing. To exclude germline variants we sequenced the matched normal DNA, when available, or consulted the Human dbSNP Database at NCBI (Build 142) (http://www.ncbi.nlm.nih.gov/snp) and removed from the analysis changes present in germline samples or with a corresponding rs entry (unless the same variant was also included in NCBI ClinVar database). For TP53 mutatios the International Agency for Research on Cancer (IARC) Database (http://www-p53.iarc.fr) was consulted to obtain informations about the functional relevance of the detected mutations. Specifically, a TP53 mutant was classified as "non-functional" if the median of its transcriptional activities measured in reporter assays on eight different promoters containing TP53-binding sequences was lower than 20% compared to the wild type protein. We acquired additional information about the functional relevance of the detected single-nucleotide

variants (SNVs) using five tools which predicts possible impact of amino acid substitutions on the structure and function of human proteins using straightforward physical and evolutionary comparative considerations. The functional predictors used were Polyphen2 (http://genetics.bwh.harvard.edu/pph2/), Mutation Taster (http://www.mutationtaster.org/), Mutation Assessor (http://mutationassessor.org/), SIFT (http://sift.jcvi.org/) and LRT (http://www.genetics.wustl.edu/jflab/lrt_query.html).

Table 3. Genome-specific primers for amplicon library preparation. Pubblished primers for *BRAF* [119], *NRAS* [113] and *KRAS* [120].

GENE	REFSEQ GENE	REFSEQ mRNA	EXON	PRIMER FORWARD	PRIMER REVERSE
BRAF	NG_007873.2	NM_004333.4	ex_11	5'-TCCCTCTCAGGCATAAGGTAA-3'	5'-CGAACAGTGAATATTTCCTTTGAT-3'
BRAF			ex_15	5'-TCATAATGCTTGCTCTGATAGGA-3'	5'-GGCCAAAAATTTAATCAGTGGA-3'
NRAS	NG_007572.1	NM_002524	ex_2	5'-ATGGAAGGTCACACTAGGGTT-3'	5'-TCCTTTAATACAGAATATGGG-3'
NRAS			ex_3	5'-ACCCCCAGGATTCTTACAGA-3'	5'-AGAGGTTAATATCCGCAAATG-3'
KRAS	NG_007524.1	NM_033360.3	ex_2	5'-TTAACCTTATGTGTGACATGTTCTAA-3'	5'-GAAAGTAAAGTTCCCATATTAATGGT-3'
KRAS			ex_3	5'-GCACTGTAATAATCCAGACT-3'	5'-CATGGCATTAGCAAAGACTC-3'
KRAS			ex_4	5'-AGTTGTGGACAGGTTTTGAAAGA-3	5'-ATTTCAGTGTTACTTACCTGTCTTG-3'
DIS3	NC_000013.10	NM_014953.4	ex_1	5'-GAAAGGGAAGAACCTCCGGG-3'	5'- TCCCTGTCATACCCCCTTGT-3'
DIS3			ex_2	5'- GCTTCTTGGCTTAACTTATTCAGTG-3'	5'- ACAGGAACCCTCTCCCGAA-3'
DIS3			ex_3	5'- TGCTAAGAGTTTTCACATATCCTTG-3'	5'- TCACATGAAGTTATATAGGACTACGA-3'
DIS3			ex_4	5'- AGGCTGTAGTGATGTGTAATTGC-3'	5'- GCTTACCCACCGACATTCC-3'
DIS3			ex_10	5'- TATGTTGTAGTTGTGCTTTGGAAAT-3'	5'- CAATATGCTTGACTGGGTAAATGTA-3'
DIS3			ex_11	5'- TTTGCTTGTTAATTACTCTTGTGAAG -3'	5'- TTCATGGCCGTTTAAGAATC-3'
DIS3			ex_12-13	5'- TTATGGCTAAGTAATCTGTGGTTCTA -3'	5'-AAATTAGAGATTAATAGCCATGAAACG-3'
DIS3			ex_14	5'- GTAGTGAAAGTAGGAGGACATATTG-3'	5'- GAAGCTAGAAGAAACAGGAGTCT-3'
DIS3			ex_15	5'- ACACTTGCTGTAGTCATTGTCTT-3'	5'- GCAAGCCAAATAAAGTAGAAATCAT-3'
DIS3			ex_16	5'- GCGGAGTAACTGAGAGATGAAAG-3'	5'- CAGGTAGATCAAAACACAAATAGATG-3'
DIS3			ex_17	5'- GCCGAATCTCCTACTTTTCCA-3'	5'- CCAAAAGCCGATGAACAATGA-3'
DIS3			ex_18	5'- CTTAGCGTCTCCAATATACACACA-3'	5'- CTAGCAGTATCGACAAAAGGCA-3'
FAM46C	NC_000001.11	NM_017709.3	ex_2_A	5'-GCTTCTCACCCCTCACTTTCA-3'	5'- CAGAAGGGAACACAGAACCAC-3'
FAM46C			ex_2_B	5'- CAGAGGCAGAATTTCAGCTGG-3'	5'- GAAGTGCTCAGAGATGGGATTA-3'
FAM46C			ex_2_C	5'- CCTGATCTCCCTCTCCAACAA-3'	5'- CACATGTAGCGCTCTAGAGTTT-3'
FAM46C			ex_2_D	5'- CCCACAGACCAGGAAGAAATC-3'	5'- CACGCAAGGCCAGGAGGGA-3'
FAM46C			ex_2_E	5'- CAGGCAGACTCTGAACCTCAT-3'	5'- CCTAGCCCTATTGGGGTTCC-3'
TP53	NC_000017.11	NM_000546.5	ex_4	5'-CCTGGTCCTCTGACTGCTCT-3'	5'-TTCTGGGAAGGGACAGAAGA-3'
TP53			ex_4	5'-GTCCAGATGAAGCTCCCAGA-3'	5'-GCCAGGCATTGAAGTCTCAT-3'
TP53			ex_5	5'-TCTGTCTCCTTCCTCTTCCTACA-3'	5'-AACCAGCCCTGTCGTCTCT-3'
TP53			ex_6	5'-CAGGCCTCTGATTCCTCACT-3'	5'-GCCACTGACAACCACCCTTA-3'
TP53			ex_7	5'-CCTGCTTGCCACAGGTCT-3'	5'-GTGATGAGAGGTGGATGGGT-3
TP53			ex_8	5'-GGGAGTAGATGGAGCCTGGT-3'	5'-GCTTCTTGTCCTGCTTGCTT-3'
TP53			ex_9	5'-AAAGGGGAGCCTCACCAC-3'	5'-TGTCTTTGAGGCATCACTGC-3'

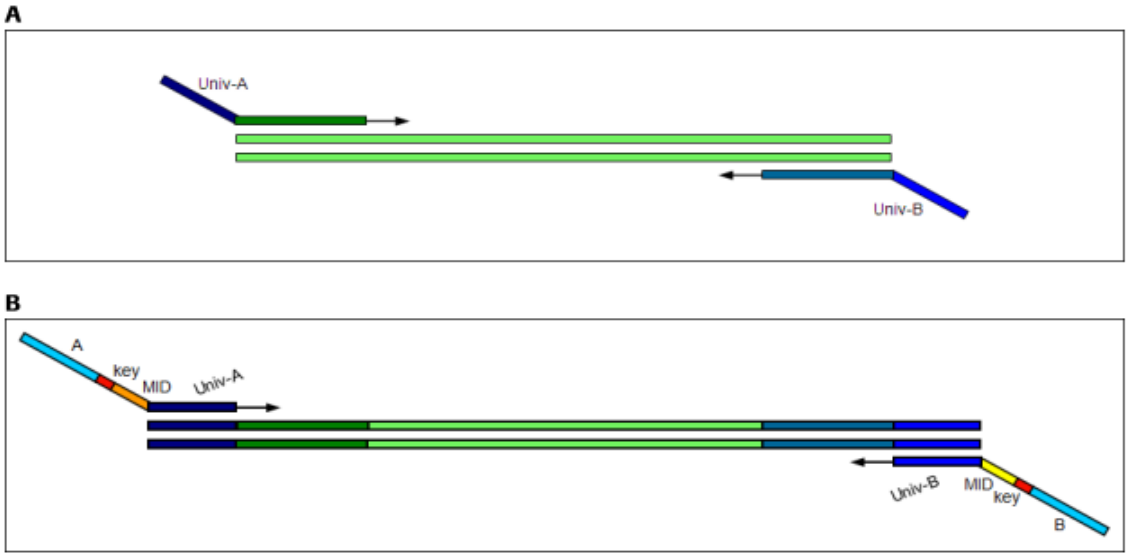

Figure 9. Schematic representation of the PCR reaction components for the preparation of a two step Universal Tailed Amplicon Library. **A)** First step, targeting the template-specific sequences and adding the universal tail. **B)** Second step, targeting the universal tail and adding the 454 Sequencing System primers and MIDs.

Validation of variants at transcriptomic level

To verify the occurrence of the obtained variants at transcriptomic level, total RNA of mutated samples was converted to cDNA using M-MLV reverse transcriptase (Invitrogen) and random hexamers, and subjected to deep sequencing of the exon harboring the variant/s detected on genomic DNA. Sequence-specific exon PCR primers were designed in the Primer 3 program (http://frodo.wi.mit.edu/primer3/) as reported in **Table 4**.

Table 4. Sequence-specific exonic PCR primers for validation of variants.

GENE	AMPLICON	PRIMER FORWARD	PRIMER REVERSE
BRAF	ex_11	5'-CTCAGCGAGAAAGGAAGTCA-3'	5'-GGAATAGCCCATGAAGAGTAGGA-3'
BRAF	ex_15	5'-AGATATTGCACGACAGACTGC-3'	5'-TTCTGATGACTTCTGGTGCC-3'
DIS3	ampl_A	5'-CCGGGGTTAGGCGTATTCTA -3'	5'- AAGTGAGAAATCGCAGTGCC -3'
DIS3	ampl_B	5'- TACTTGCTGCCCGACACTAA -3'	5'- GGAATACCAGCTTTCACTTGTG -3'
DIS3	ampl_C	5'- ATGTTCCCCATCAGCCTTTT -3'	5'- AAAATGTGACGTGGACAGGC -3'
DIS3	ampl_D	5'- GGGAAATGAATCACAATGCTGA -3'	5'- TCTGAACATGCTCTGCTTCG -3'
DIS3	ampl_E	5'- AGTCTTTGGATCAGGCCGAA -3'	5'- GGCTATTGGGGCTGACTGTA -3'
DIS3	ampl_F	5'- ACTATGGCTTAGCGTCTCCA -3'	5'- AACGTGCATCAGTGGCTTTT-3'

Library design and CAPP-seq approach for ultra-deep NGS by Illumina MiSeq

A targeted resequencing gene panel, including coding exons and splice sites of 14 genes that are recurrently mutated in MM patients was specifically designed for this project (target region of 30989bp: *BRAF*, *CCND1*, *CYLD*, *DIS3*, *EGR1*, *FAM46C*, *IRF4*, *KRAS*, *NRAS*, *PRDM1*, *SP140*, *TP53*, *TRAF3*, *ZNF462*) in collaboration with Prof. Davide Rossi at IOR in Bellinzona (CH). Inclusion *criteria* of gene panel design were based on publicly available sequencing data from three distinct datasets[64, 65, 121] and were as follows: (i) genes that were recurrently mutated in ≥ 3% of MM tumors; (ii) genes that were cross-validated in at least two of the considered genomic datasets. An *in silico* validation of the designed gene panel in the three aforementioned patients cohorts resulted in the recovery of at least one clonal mutation in 68% (95% confidence interval [CI]: 58% to 76%) of MM cases (**Figure 10**).

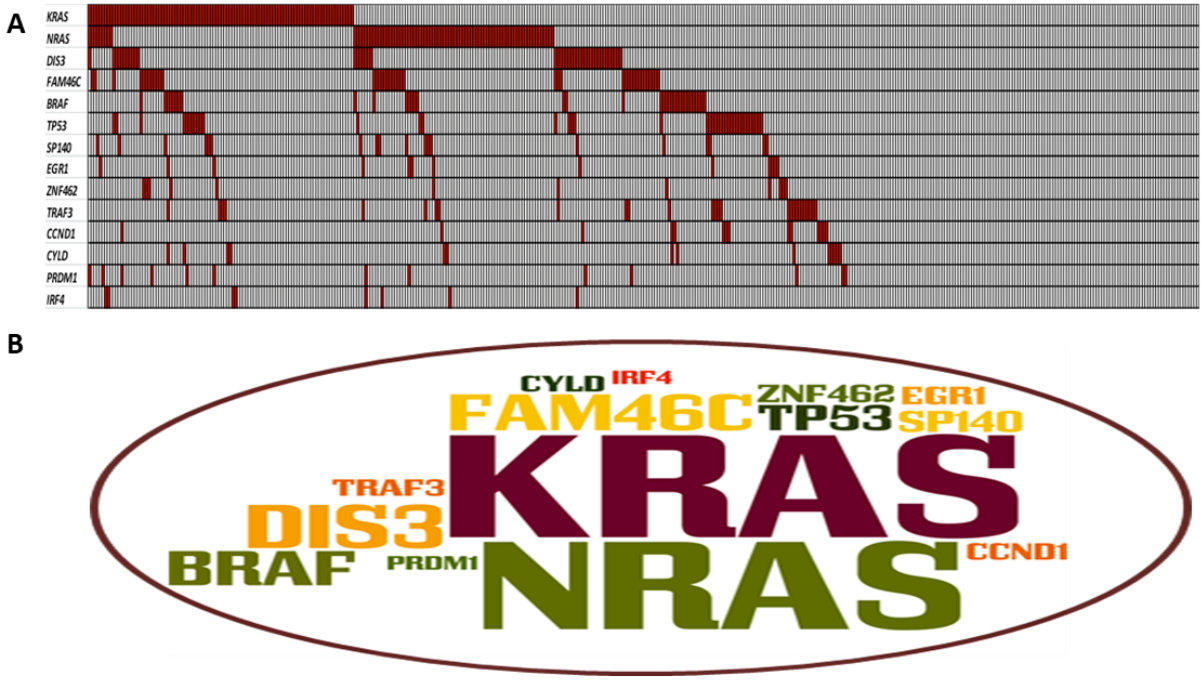

Figure 10. (A) Heatmap showing MM patients (in total 412 cases) of the three works[64, 65, 121] on which the study design was based; in red mutated cases. **(B)** Word clouds (created on http://www.wordle.net) representing the most mutated genes chosen for the panel. The percentage of the gene's prevalence generally in the distinct datases[64, 65, 121] is shown by gene name size.

The gene panel was analyzed in plasma cfDNA and in normal gDNA from the paired granulocytes, as source of germline material, for comparative purposes to filter out polymorphisms. The gDNA from the paired CD138+ purified PCs from BM aspiration was also investigated in the same cases to assess the accuracy of plasma ctDNA genotyping. Tumor and germline gDNA (median 400 ng) were sheared

through sonication before library construction to obtain 150-bp fragments. Plasma cfDNA, which is naturally fragmented, was used (median 59 ng) for library construction without additional fragmentation. Targeted ultra-deep-NGS was performed on plasma cfDNA, tumor and germline gDNA by using CAncer Personalized Profiling by deep Sequencing (CAPP-Seq) approach (**Figure 11**), which has been validated for plasma ctDNA genotyping[122]. The NGS libraries were constructed using the KAPA Library Preparation Kit (Kapa Biosystems) and hybrid selection was performed with the custom SeqCap EZ Choice Library (Roche NimbleGen). The manufacturer's protocols were modified as previously reported [122]. Multiplexed libraries were sequenced using 300-bp paired-end runs on an Illumina MiSeq sequencer. Each run included 24 multiplexed samples in order to allow >2000x coverage in >80% of the target region.

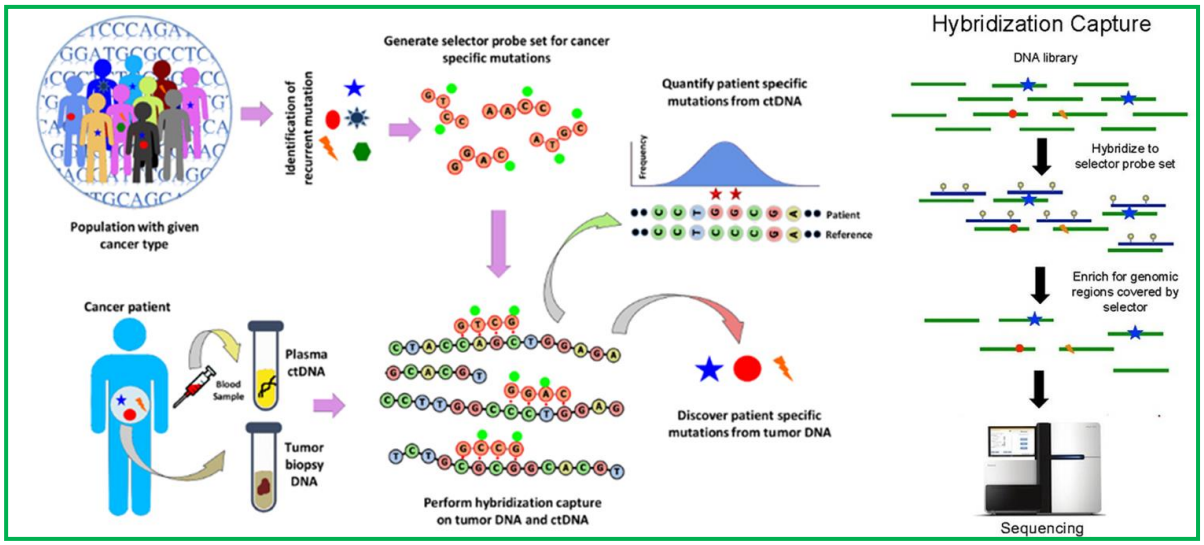

Figure 11. An overview of the workflow of CAPP-Seq and hybridization capture protocol.

Bioinformatic pipeline for variant calling and statistical analysis

Mutation calling in plasma cfDNA was performed separately and in blind from mutation calling in tumor gDNA from purified PCs. We deduped FASTQ sequencing reads by utilizing FastUniq v1.1. The deduped FASTQ sequencing reads were locally aligned to the hg19 version of the human genome using BWA v.0.6.2, and sorted, indexed and assembled into a mpileup file using SAMtools v.1. The aligned reads were processed with mpileup. Single nucleotide variations and indels were called in plasma cfDNA *vs* normal gDNA, and tumor gDNA *vs* normal gDNA, respectively, by using the somatic function of VarScan2. The variant called by VarScan 2 were annotated by using SeattleSeq Annotation 138. Variants

annotated as SNPs, intronic variants mapping >2 bp before the start or after the end of coding exons, and synonymous variants were filtered out. To filter out variants below the base-pair resolution background frequencies, the Fisher's exact test was used to test whether the variant frequency called by VarScan 2 in cfDNA or tumor gDNA, respectively, was significantly higher from that called in the corresponding paired germline gDNA, after adjusting for multiple comparison by Bonferroni test (Bonferroni-adjusted *P*=4.03252e-7). To further filter out systemic sequencing errors, a database containing all background allele frequencies in all the specimens analyzed was assembled. Based on the assumption that all background allele fractions follow a normal distribution, a Z-test was employed to test whether a given variant differs significantly in its frequency from typical DNA background at the same position in all the other DNA samples, after adjusting for multiple comparison by Bonferroni. Variants that did not pass this filter were not further considered. Variant allele frequencies for the resulting candidate mutations and the background error rate were visualized using IGV.

The sensitivity and specificity of plasma cfDNA genotyping were calculated in comparison with tumor gDNA genotyping as the gold standard. The analyses were performed with the Statistical Package for the Social Sciences (SPSS) software (Chicago, IL) and with R statistical package (http://www.r-project.org).

For analyses of correlation were used Pearson correlation's method and Mann-Whitney test.

Gene expression profiling and statistical analysis

In order to identify the transcriptional profiles related to the mutations in primary tumors, we used microarray technology to investigate GEP of a large number of the samples analyzed by NGS. Total RNA was purified using the RNeasy® Total RNA Isolation Kit (Qiagen, Valencia, CA). Preparation of DNA single-stranded sense target, hybridization to GeneChip® Gene 1.0 ST arrays (Affymetrix, Santa Clara, CA) and scanning of the arrays (7G Scanner, Affymetrix) were performed according to the manufacturer's protocols[112]. The raw intensity expression values were processed by Robust Multi-array Average (RMA) procedure[123], with the re-annotated Chip Definition Files from BrainArray libraries version 18.0.0[124], available at http://brainarray.mbni.med.umich.edu. Supervised analyses were performed using Significant Analysis of Microarrays software (SAM version 4.00; Excel front-end publicly available at http://www-stat.stanford.edu/tibs/SAM/index.html)[108, 125]. The cutoff point for statistical significance was determined by tuning the Δ parameter on the false discovery rate (FDR) and controlling the *q*-value of the selected probes (q-value=0). Differentially expressed genes were also

evaluated at the highest stringency level (median FDR 0%, 90th percentile FDR = 0). The functional annotation analysis on the selected lists was performed by means of DAVID 6.7 (http://david.abcc.ncifcrf.gov/) and the ToppFun option in ToppGene Suite (https://toppgene.cchmc.org/) tools, using the default parameters[126].

The principal component analysis (PCA) of the samples was performed by singular value decomposition of the considered data expression matrix using the *prcomp* function in the *stats* package, and the results were visualized using the *plot3d* function in the *rgl* package for R software.

The GEP data have been deposited in the NCBI Gene Expression Omnibus database (GEO; http//www.ncbi.nlm.nih.gov/geo; accession No. GSE66293).

All contingency analyses were performed by Fisher's exact test. A P value <0.05 was considered significant for all statistical calculations. A *P*-value cut-off of 0.05 and FDR Benjamini and Hochberg correction method were applied to all the annotation terms in the default parameter set, and those with a *q*-value <0.05 were defined as significantly enriched.

Results

Multiple myeloma and plasma cell leukemia samples

To estimate the frequency of *KRAS*, *NRAS*, *BRAF*, *DIS3*, *FAM46C* and *TP53* mutations in PC dyscrasia, we investigated by NGS a retrospective cohort of 132 cases with MM at onset and 24 patients with pPCL, 11 with sPCL (see Materials and Methods for patient populations). In order to gain further insights into the longitudinal status of mutations, we analyzed specimens taken from 19 patients at two different times: 14 with MM and two with pPCL at onset and relapse; two with MM at onset and at the time of leukemic transformation; and one at early and relapsed leukemic phase. Pathological samples from all patients were characterized by FISH for the presence of the main molecular alterations occurring in MM (**Table 5**). The cohort analyzed was the same for all sequenced genes, except for three cases (2 MM and 1 pPCL) not longer available for *DIS3*, 4 cases (another MM in addition) for *TP53* and 5 cases for *FAM46C* (another MM too).

Specifically, we performed deep sequencing of the following mutational hot-spots: exons 11 and 15 in *BRAF*, exons 2-3 in *NRAS* and exons 2-4 in *KRAS* (Cosmic Release v70, at http://cancer.sanger.ac.uk/cancergenome/projects/cosmic/); the two functional domains of *DIS3*, PIN (exons 1-4) and RNB (exons 10-18) domains; exon 2 in *FAM46C* and exons 4-9 in *TP53*. These mutational hotspots contained almost 98% of mutations, identified in the main previously published whole genome and exome sequencing studies in MM[65 64 79]. The median depth of coverage was ~200x (range 100-1000x).

Table 5. Clinical details of the 167 patients analyzed by the 454 GS Junior System. [Δ]The Durie clinical staging system was adopted; [‡]ND: newly diagnosed, RR: relapsed/refractory; [†]Paraprotein; *del(13), del(17), 1p loss and 1q gain were determined by FISH. °HD= presence of the hyperdiploid status based on FISH evaluation criteria; na: not applicable.

Sample	Sex	Age	Disease	Stage[Δ]	Phase[‡]	PP[†]	t(4;14)	t(11;14)	t(14;16)	t(14;20)	del(13)*	del(17p)*	1q gain*	1p loss*	HD°
MM-004	F	58	MM	IA	ND	Gκ	-	-	+	-	-	na	na	na	na
MM-015	M	71	MM	IIA	ND	Gκ	-	+	-	-	-	-	-	-	-
MM-016	M	66	MM	IIIB	ND	Gκ	-	-	-	-	-	-	+	-	+
MM-026	F	72	MM	IIIB	ND	κ	-	+	-	-	-	-	-	na	na
MM-027	M	60	MM	IA	ND	Gκ	-	-	-	-	+	-	na	na	na
MM-030	M	69	MM	IIIA	ND	Gλ	-	-	-	-	-	-	-	+	+
MM-031	M	58	MM	IIIA	ND	Aκ	-	+	-	-	-	-	+	+	-
MM-034	M	71	MM	IA	ND	Gκ	-	-	-	-	-	-	-	-	+
MM-036	M	65	MM	IIA	ND	Gκ	-	-	-	-	+	-	+	na	-
MM-037	F	50	MM	IIA	ND	Gκ+Aκ	-	+	-	-	+	-	-	-	-
MM-038	F	67	MM	IIA	ND	κ	-	-	-	-	+	-	-	-	+
MM-039	M	50	MM	IIA	ND	Gλ	-	-	-	-	-	-	-	na	+
MM-042	M	54	MM	IIIA	ND	Aλ	+	-	-	-	+	-	+	-	+
MM-043	F	73	MM	IA	ND	Gκ+Gλ	-	-	-	-	-	-	+	-	-
MM-049	M	62	MM	IIIB	ND	κ	-	-	-	-	-	-	-	-	+
MM-055	F	69	MM	IIIA	ND	Gκ	-	+	-	-	-	-	-	na	-
MM-066	F	77	MM	IIIA	ND	Aκ	+	-	-	-	-	-	-	na	-
MM-069	M	63	MM	IIA	ND	Gκ	-	-	-	+	+	-	+	na	-
MM-078	F	59	MM	IIIA	ND	κ	-	-	-	-	+	-	-	na	+
MM-079	F	74	MM	IIA	ND	Gκ+Gλ	-	-	-	-	-	-	-	-	+
MM-087	F	84	MM	IIIA	ND	Gλ	+	-	-	-	+	-	+	na	-
MM-115	F	53	MM	IIIA	ND	Gλ	-	+	-	-	+	-	na	na	-
MM-123	M	55	MM	IIIA	ND	Gκ	+	-	-	-	+	-	+	-	-
MM-131	M	73	MM	IA	ND	Gκ	-	-	-	-	-	-	-	-	-
MM-140	M	62	MM	IIIB	ND	Gκ	-	+	-	-	-	-	-	na	-
MM-143	M	61	MM	IIA	ND	Gκ	-	-	-	-	-	-	+	-	+
MM-146	M	68	MM	IIA	ND	Gλ	-	-	-	-	-	-	+	na	+
MM-148	F	55	MM	IA	ND	Aκ	-	-	-	-	-	+	+	na	+
MM-149	F	52	MM	IA	ND	Gλ	-	-	-	-	-	-	-	-	+
MM-150	F	68	MM	IIA	ND	Gλ	-	-	-	-	-	-	+	-	+
MM-151	F	71	MM	IA	ND	Gλ	-	-	-	-	-	-	-	-	+
MM-152	M	66	MM	IA	ND	Gκ	-	-	-	-	-	-	-	na	+
MM-154	F	71	MM	IIA	ND	Gκ	-	-	+	-	+	-	+	-	-
MM-159	M	56	MM	IIA	ND	κ	-	+	-	-	+	-	-	-	-
MM-174	M	85	MM	IIA	ND	Aκ	-	-	-	-	-	-	-	-	+
MM-177	M	73	MM	IIIA	ND	Gκ	-	-	-	-	+	-	+	-	-
MM-179	M	50	MM	IIIA	ND	Gλ	-	+	-	-	-	+	-	-	+
MM-195	M	62	MM	IIA	ND	Gκ	+	-	-	-	+	-	-	-	-
MM-200	F	63	MM	IA	ND	Aλ	-	-	-	-	-	-	-	-	+
MM-202	M	64	MM	IIIA	ND	Gκ	-	-	-	-	-	-	+	na	+
MM-206	F	73	MM	IIA	ND	Gκ	+	-	-	-	+	-	-	+	-
MM-207	F	68	MM	IA	ND	Aκ	-	-	-	-	+	-	+	-	-
MM-208	M	74	MM	IIA	ND	Aλ	+	-	-	-	-	-	-	-	-
MM-209	F	65	MM	IIA	ND	Gλ	-	-	-	-	-	-	-	-	+
MM-210	M	65	MM	IIB	ND	Aλ	-	-	-	-	-	-	-	-	-
MM-212	F	55	MM	IIIA	ND	Gκ	-	+	-	-	-	-	-	-	-
MM-213	M	66	MM	IIIA	ND	λ	-	+	-	-	+	-	-	-	-
MM-219	M	73	MM	IIIA	ND	Aλ	-	-	-	-	+	-	-	-	+
MM-224	F	52	MM	IIIA	ND	Gκ	-	-	+	-	-	-	+	-	-
MM-229	M	75	MM	IIA	ND	Gκ	-	-	-	-	-	-	-	-	-
MM-238	M	58	MM	IIB	ND	Gκ	-	-	-	-	+	-	+	-	-
MM-239	F	72	MM	IIA	ND	Aκ	-	-	-	-	+	-	-	-	+
MM-240	M	70	MM	IA	ND	Gλ	-	-	-	-	-	-	-	-	+
MM-241	M	54	MM	IIA	ND	Gκ	-	-	-	-	-	-	-	-	+
MM-242	M	69	MM	IIIA	ND	Aκ	-	-	-	-	+	-	-	-	+

Sample	Sex	Age	Disease	Stage[Δ]	Phase[‡]	PP[†]	t(4;14)	t(11;14)	t(14;16)	t(14;20)	del(13)*	del(17p)*	1q gain*	1p loss*	HD°
MM-243	F	68	MM	IIA	ND	Gκ	-	-	-	-	-	-	+	-	+
MM-246	M	71	MM	IIIA	ND	Aλ	-	+	-	-	-	-	-	-	-
MM-252	F	77	MM	IIA	ND	Aλ	-	+	-	-	-	-	+	-	-
MM-253	F	45	MM	IIA	ND	λ	-	-	-	-	+	-	+	-	+
MM-256	M	58	MM	IIIA	ND	Aκ	-	-	+	-	+	-	-	-	-
MM-261	F	66	MM	IIB	ND	Gκ	-	-	-	-	+	-	+	-	+
MM-262	M	74	MM	IA	ND	λ	-	-	-	-	-	-	+	-	-
MM-263	F	65	MM	IIB	ND	κ	+	-	-	-	+	-	+	-	-
MM-267	M	74	MM	IIIA	ND	Gλ	-	-	-	-	+	-	-	-	+
MM-268	M	77	MM	IIIB	ND	Gκ	-	-	-	-	-	-	-	-	-
MM-269	F	67	MM	IIA	ND	Gκ	-	-	-	-	+	-	-	-	+
MM-271	M	76	MM	IA	ND	Gκ	-	-	-	-	+	-	-	-	-
MM-274	M	59	MM	IIA	ND	Gκ	+	-	-	-	-	-	-	-	-
MM-276	F	70	MM	IIB	ND	Gκ	+	-	-	-	+	-	-	-	-
MM-278	M	73	MM	IIA	ND	Gλ	-	-	-	-	-	-	-	+	+
MM-279	F	71	MM	IIIA	ND	Gλ	-	-	-	na	+	-	+	-	na
MM-280	M	62	MM	IIIA	ND	Gλ	-	+	-	-	-	-	-	-	-
MM-281	F	77	MM	IIA	ND	Gκ	-	-	-	-	-	-	-	-	-
MM-282	M	66	MM	IIIA	ND	Gκ	-	-	-	-	-	-	+	-	+
MM-284	M	49	MM	IIIA	ND	κ	-	+	-	-	na	-	-	-	-
MM-286	F	71	MM	IIA	ND	Gκ	-	-	-	-	+	-	-	-	-
MM-295	F	74	MM	IIIA	ND	λ	-	-	-	-	+	-	na	na	na
MM-300	M	65	MM	IIB	ND	Gκ	-	-	-	-	-	-	na	na	na
MM-301	F	72	MM	IIA	ND	Aλ	-	-	-	-	-	-	+	-	+
MM-302	M	65	MM	IA	ND	Aκ	-	-	-	-	-	-	+	-	+
MM-308	M	58	MM	IA	ND	Aκ	-	-	-	-	-	-	-	-	-
MM-313	M	66	MM	IIIA	ND	Gλ	-	+	-	-	-	-	-	-	-
MM-310	M	67	MM	IIIA	ND	Gλ	-	+	-	-	-	-	-	-	-
MM-314	F	70	MM	IIIA	ND	Aλ	-	+	-	-	-	-	+	-	-
MM-317	M	56	MM	IIA	ND	Gλ	-	+	-	-	-	-	+	-	-
MM-321	M	63	MM	IIIA	ND	Gκ	-	-	-	-	+	-	-	-	+
MM-327	F	62	MM	IIA	ND	absent	-	-	-	-	+	-	+	-	+
MM-330	F	61	MM	IIIA	ND	Aκ	+	-	-	-	+	-	+	-	-
MM-334	F	45	MM	IA	ND	Gλ	-	-	-	-	+	+	+	+	+
MM-335	F	68	MM	IIA	ND	Aκ	-	-	+	-	+	-	-	-	-
MM-340	M	46	MM	IIIA	ND	Gλ	-	+	-	-	+	-	-	-	-
MM-341	M	65	MM	IIIA	ND	Gκ	-	-	-	-	-	-	+	-	-
MM-343	M	74	MM	IIIA	ND	Aλ	-	+	-	-	-	-	-	-	-
MM-351	M	na	MM	IIA	ND	Aκ	-	-	-	-	-	-	-	-	-
MM-362	F	80	MM	IIA	ND	Gλ	-	-	-	-	+	+	+	-	-
MM-372	M	54	MM	IA	ND	Gκ	-	+	-	-	-	-	+	-	-
MM-375	F	78	MM	IA	ND	Gκ	+	-	-	-	-	+	-	-	-
MM-381	F	69	MM	IA	ND	Gκ	-	-	-	na	-	-	-	-	na
MM-382	M	85	MM	IIA	ND	Gκ	-	-	-	-	+	+	na	na	na
MM-385	F	76	MM	IIIA	ND	Gλ	-	-	-	-	-	-	-	-	+
MM-386	F	76	MM	IIA	ND	Aλ	-	na	na	na	-	-	na	na	na
MM-387	F	44	MM	IIIA	ND	Gκ	-	-	-	-	-	-	+	-	+
MM-392	F	53	MM	IA	ND	Gλ	-	-	-	-	+	-	+	-	+
MM-398	F	65	MM	IIA	ND	Gκ	-	+	-	-	+	-	+	-	-
MM-402	M	67	MM	IIA	ND	Gκ	-	+	-	-	+	-	-	-	-
MM-405	M	69	MM	IA	ND	Gκ	-	+	-	-	-	-	-	-	-
MM-406	M	62	MM	IIB	ND	Aκ	-	+	-	-	-	-	+	-	na
MM-407	F	79	MM	IIIA	ND	λ	-	-	-	-	+	-	-	-	-
MM-410	F	79	MM	IIIA	ND	λ	-	-	-	-	-	-	+	-	-
MM-411	M	67	MM	IIIA	ND	Gλ	-	-	-	-	-	-	-	-	+
MM-413	F	51	MM	IIA	ND	Gκ	-	-	-	-	+	-	+	-	+
MM-414	F	67	MM	IIIA	ND	Gλ	-	+	-	-	-	-	-	-	-

Sample	Sex	Age	Disease	Stage[Δ]	Phase[‡]	PP[†]	t(4;14)	t(11;14)	t(14;16)	t(14;20)	del(13)*	del(17p)*	1q gain*	1p loss*	HD°
MM-422	F	74	MM	IIB	ND	κ	-	-	-	-	+	-	-	-	-
MM-423	F	53	MM	IIA	ND	λ	+	-	-	-	+	-	-	-	-
MM-424	F	84	MM	IA	ND	Gκ	-	-	-	-	+	-	-	+	+
MM-425	F	65	MM	IIA	ND	Aλ	-	-	-	-	+	-	+	-	-
MM-428	F	65	MM	IIA	ND	Aλ	-	-	-	-	-	-	+	+	-
MM-429	F	63	MM	IIIA	ND	Gκ	-	-	-	-	-	-	+	-	+
MM-430	na	62	MM	IIIA	ND	Gλ	-	-	-	-	-	-	-	-	-
MM-431	F	70	MM	IA	ND	Gλ	+	-	-	-	+	-	+	-	-
MM-433	M	72	MM	IIB	ND	κ	-	+	-	-	-	-	-	-	-
MM-434	F	53	MM	IIA	ND	κ	-	-	-	-	+	-	-	+	-
MM-435	na	42	MM	IIIA	ND	Gλ	-	-	-	-	-	-	-	-	+
MM-437	F	72	MM	IIB	ND	λ	-	-	-	-	+	-	+	-	-
MM-440	na	61	MM	IIA	ND	Aκ	-	-	-	-	-	-	-	+	+
MM-441	na	na	MM	IA	ND	Gλ	-	-	-	-	-	-	-	-	-
MM-442	F	65	sPCL	/	ND	Gκ	-	-	-	-	+	-	+	+	-
MM-445	na	65	MM	IIA	ND	κ	-	+	-	-	+	-	+	-	-
MM-446	na	52	MM	IIIA	ND	Gλ	-	-	-	-	-	-	+	-	+
MM-447	na	57	MM	IIA	ND	Gκ	-	+	-	-	-	-	+	-	-
MM-448	na	59	MM	IIA	ND	Aκ	-	-	-	-	+	-	+	-	-
MM-449	na	54	MM	IIA	ND	Gκ	-	-	-	-	+	-	-	+	-
MM-464	M	75	MM	IIIA	ND	Gλ	-	-	-	-	+	-	+	-	-
PCL-001	F	51	pPCL	/	ND	κ	-	-	-	+	+	-	na	na	na
PCL-002	F	69	sPCL	/	RR	λ	-	+	-	-	+	na	-	na	na
PCL-004	M	72	pPCL	/	ND	Gκ	-	-	+	-	-	-	-	-	-
PCL-005	M	76	sPCL	/	RR	Aκ	-	+	-	-	-	-	-	-	-
PCL-007	M	65	sPCL	/	RR	κ	-	-	-	-	+	+	+	na	na
PCL-008	M	57	pPCL	/	ND	Gλ	-	-	+	-	+	-	-	na	na
PCL-009	F	77	sPCL	/	RR	Gκ	+	-	-	-	+	-	+	-	-
PCL-011	M	76	sPCL	/	RR	Gκ	-	-	+	-	+	+	+	-	-
PCL-012	F	62	sPCL	/	RR	Aκ	-	-	+	-	-	+	-	-	-
PCL-014	M	72	pPCL	/	ND	λ	-	+	-	-	+	+	-	+	-
PCL-015	M	78	pPCL	/	ND	κ	-	-	+	-	+	-	+	-	-
PCL-016	F	57	pPCL	/	ND	Gκ	-	-	+	-	+	-	+	+	-
PCL-017	F	68	pPCL	/	ND	Gκ	-	-	+	-	+	+	+	+	-
PCL-018	F	59	pPCL	/	ND	κ	-	+	-	-	+	+	-	-	-
PCL-019	F	67	pPCL	/	ND	Mκ	-	-	+	-	+	+	+	+	-
PCL-020	F	79	pPCL	/	ND	Gλ	-	-	-	-	-	-	+	-	-
PCL-021	M	48	pPCL	/	ND	Gλ	+	-	-	-	+	-	+	-	-
PCL-023	M	60	pPCL	/	ND	Gκ	-	-	+	-	+	+	+	+	-
PCL-026	F	59	pPCL	/	ND	Gκ	-	-	+	-	+	-	+	-	-
PCL-027	M	65	pPCL	/	ND	λ	-	+	-	-	-	+	-	-	-
PCL-028	F	57	pPCL	/	ND	κ	-	+	-	-	+	-	-	-	-
PCL-029	M	51	pPCL	/	ND	Aλ	-	-	+	-	-	-	-	-	-
PCL-030	F	52	pPCL	/	ND	κ	-	-	-	+	+	+	-	-	-
PCL-031	F	59	sPCL	/	ND	Gλ	-	-	+	-	-	+	+	-	-
PCL-032	M	51	pPCL	/	ND	Gκ	+	-	-	-	+	-	+	-	-
PCL-035	F	76	pPCL	/	ND	κ	-	+	-	-	-	-	-	-	-
PCL-036	M	72	pPCL	/	ND	Gκ	-	+	-	-	+	-	-	-	-
PCL-037	M	72	pPCL	/	ND	Aλ	-	+	-	-	+	-	-	-	-
PCL-038	M	57	pPCL	/	ND	Gκ	-	-	-	-	+	-	+	+	+
PCL-039	M	54	sPCL	/	ND	Aκ	-	-	-	-	-	+	-	+	-
PCL-041	M	na	sPCL	/	ND	Gκ	-	na	na	na	-	-	na	na	na
PCL-042	F	69	sPCL	/	ND	Gλ	-	+	-	-	-	-	-	-	-
PCL-043	F	68	pPCL	/	ND	Gλ	-	+	-	-	-	-	-	-	-
PCL-046	F	50	pPCL	/	ND	κ	-	-	-	-	+	-	+	-	+

Assessment of mutations in PC dyscrasias

After exclusion of intronic, synonymous and germline variants, deep sequencing revealed different coding somatic mutations as follows: 8 in *BRAF*, 13 in *NRAS*, 23 in *KRAS*, 30 in *DIS3*, 24 in *FAM46C*, 14 in *TP53* (**Table 6** and **Figure 12**). The presence of each non-synonymous SNV was verified in an independent PCR product, analyzed by conventional sequencing whenever the sensitivity of the Sanger's method (i.e. about 10% in our experimental conditions) was consistent with the variant allele frequency VAF, or by an additional and ultra-deep pyrosequencing run (median depth of coverage=1110x) in case of variants at low allele frequency.

The high sensitivity of the method allowed mutation detection even at a very low variant allele frequency. In fact, a considerable proportion of patients presented with a small mutation load.

BRAF and *RAS* (*NRAS-KRAS*)

Eight distinct single nucleotide variations in *BRAF* were detected in 20 out of 167 (12%) analyzed patients. All of the identified mutations were missense substitutions (**Table 6**) and the VAFs ranged from 0.86% to 70.7% of total reads per sample. Mutations were detected in 10.6% of the patients with MM at onset (14/132), 20.8% of pPCLs (5/24) and 9.1% of sPCLs (1/11) and as the whole in the 12% (20/167) of the cases. Five of these patients (four with MM and one with pPCL) carried a mutation whose VAF was below the Sanger detection limit. The variants mainly targeted exon 15 (17/20, 85%) (**Figure 12A**). The most frequent mutation V600E was found in seven cases; the mutantion G469Awas found in three cases, E586K and G596R substitutions are detected in only one sample The most frequently mutated residue in our patients was at D594: with D594N in four cases, D594G in three, and D594E in one.

In order to obtain a more complete picture of MEK/ERK signalling activation in MM, we also investigated the entire dataset for mutations in *NRAS* and *KRAS*, encoding two GTPases that are known for long time to be involved in myelomagenesis[64, 65, 127, 128]. Sequencing of the mutation hot-spots of *NRAS* (exons 2 and 3) and *KRAS* (exons 2, 3 and 4) revealed, respectively, the occurrence of mutations in 23.9% (40/167) and 29.3% (49/167) of the cases. As a whole, *NRAS* and *KRAS* mutations were detected in 55.3% of MM cases, 20.8% of pPCL, and 54.5% of sPCL. In particular, *NRAS* mutations were found in 26.5% of the MMs at onset, 4.2% of the pPCLs, and 36.4% of the sPCLs (P = 0.008) (**Table 6**). In line with previous studies (particularly in MM), virtually all of the identified mutations affected hot-spot codons 61 (36/50, 72%), 13 (5/50, 10%) and 12 (3/50, 6.5%). *KRAS* mutations were detected in 32.6% of the MMs at onset, 16.7% of the pPCLs, and 18.2% of the sPCLs; the most frequently targeted residues were Gly12 and Gln61, followed by Gly13 and other codons in exons 2, 3, and 4, each mutated in one or two cases (**Table 6**).

Table 6. Summary of non-synonymous variants identified by NGS in *BRAF-NRAS-KRAS-DIS3-FAM46C-TP53* genes. *Reference number in bibliography.

Gene	Chr	Variant (GRCh38)	Mutation Type	AA change	dbSNP ID (v142)	COSMIC ID (v71)/IARC	ID MM literature*	Mutated patients (percentage of mutated sequencing reads)
BRAF	7	140753336A>T	Missense	V600E	rs113488022	COSM476	63,64,65,71,79	MM-313 (44.59%) PCL-043 (44.59%) PCL-015 (41.18%) MM-441 (35.74%) MM-268 (29.57%) MM-177 (27.18%) MM-446 (26.48%)
BRAF	7	140753355C>T	Missense	D594N	rs397516896	COSM27639	71	MM-295 (50.00%) PCL-026 (43.05%) MM-036 (5.61%) MM-140 (1.68%)
BRAF	7	140781602C>G	Missense	G469A	rs121913355	COSM460	63	PCL-023 (70.70%) PCL-042 (51.26%) MM-335 (50.52%)
BRAF	7	140753354T>C	Missense	D594G	rs121913338	COSM467	64,65	MM-435 (22.22%) MM-219 (7.78%) MM-411 (0.86%)
BRAF	7	140753379C>T	Missense	E586K	rs121913340	COSM463	64	PCL-026 (42.15%)
BRAF	7	140753349C>G	Missense	G596R	rs121913361	COSM469	/	MM-224 (35.44%)
BRAF	7	140753333T>G	Missense	K601T	rs397507484	COSM3878760	/	MM-039 (28.53%)
BRAF	7	140753353A>C	Missense	D594E	/	COSM253330	71	PCL-028 (4.65%)
NRAS	1	114713908T>C	Missense	Q61R	rs11554290	COSM584	63,65,79	MM-424 (95.95%) PCL-041 (90.50%) MM-174 (80.77%) PCL-039 (63.37%) MM-212 (46.70%) MM-034 (46.37%) MM-049 (34.44%) MM-435 (29.87%) MM-241 (28.98%) MM-381 (21.99%) MM-261 (21.63%) MM-229 (11.68%) MM-387 (4.56%)
NRAS	1	114713909G>T	Missense Missense	Q61K	rs121913254	COSM580	63,64,65,79	MM-434 (88.27%) MM-449 (84.36%) PCL-005 (45.05%) MM-246 (40.98%) MM-330 (33.91%) MM-414 (31.87%) MM-087 (31.74%) MM-448 (26.60%) MM-037 (12.22%) PCL-039 (8.26%) MM-327 (4.28%) MM-284 (5.86%)
NRAS	1	114713907T>G	Missense	Q61H	/	COSM586	63,64,65	MM-398 (47.45%) MM-405 (43.42%) MM-341 (31.48%) MM-387 (18.46%) PCL-027 (65.62%) MM-302 (15.73%)
NRAS	1	114716124C>G	Missense	G13R	rs121434595	COSM569	63,64,65	MM-372 (52.60%) MM-066 (30.19%) MM-302 (17.44%) MM-284 (4.61%)
NRAS	1	114716126C>T	Missense	G12D	rs121913237	COSM564	65	MM-334 (96.49%) MM-422 (11.95%) MM-295 (8.89%)
NRAS	1	114713907delT	Fram_Shift_Del	E62Kfs*6	/	/	/	MM-405 (2.67%) MM-212 (1.28%) MM-229 (1.27%)
NRAS	1	114713907T>A	Missense	Q61H	rs121913255	COSM585	63,64,65	MM-317 (48.82%) MM-146 (6.11%)
NRAS	1	114713908T>A	Missense	Q61L	rs11554290	COSM583	65	MM-276 (45.41%) MM-448 (10.37%)
NRAS	1	114716123C>T	Missense	G13D	rs121434596	COSM573	65	PCL-011 (93.06%)
NRAS	1	114713908_114713909TG>GT	Missense	Q61T	/	/	/	MM-314 (82%)
NRAS	1	114713878T>C	Missense	Y71C	/	/	/	MM-276 (45.95%)
NRAS	1	114713831T>A	Missense	S87C	/	/	/	MM-411 (45.45%)
NRAS	1	114713831delT	Fram_Shift_Del	S87Afs*17	/	/	/	MM-411 (2.24%)
KRAS	12	25227341T>G	Missense	Q61H	rs17851045	COSM554	63,64,65,79	MM-055 (52.53%) MM-442 (45.93%) MM-430 (39.77%) MM-440 (38%) MM-327 (37.34%) MM-202 (35.54%) MM-151 (35.29%) MM-229 (12.07%) MM-381 (9.51%) MM-441 (8.45%) PCL-046 (7.73%) MM-300 (7.61%) MM-402 (34.62%) MM-200 (23.74%)
KRAS	12	25245350C>T	Missense	G12D	rs121913529	COSM521	65	MM-154 (45.28%) MM-030 (45%) MM-152 (31.69%) MM-433 (31.78%) MM-026 (30.59%) MM-429 (21.47%) PCL-046 (9.01%)
KRAS	12	25245351C>G	Missense	G12R	rs121913530	COSM518	64,65,79	MM-039 (46.94%) PCL-038 (42.42%) MM-445 (28.85%) MM-159 (27.78%)
KRAS	12	25245347C>T	Missense	G13D	rs112445441	COSM532	63,65,79	MM-385 (40.68%) MM-308 (39.42%) MM-151 (9.35%) MM-219 (5.88%)
KRAS	12	25227342T>C	Missense	Q61R	rs121913240	COSM552	64,65,79	MM-038 (40.82%) MM-351 (28.66%) MM-159 (3.75%)
KRAS	12	25245321G>T	Missense	Q22K	rs121913236	COSM543	64	MM-280 (50%) MM-026 (32.94%)
KRAS	12	25225628C>T	Missense	A146T	rs121913527	COSM19404	63	MM-413 (51.24%) MM-131 (40.52%)
KRAS	12	25245348C>A	Missense	G13C	rs121913535	COSM527	/	PCL-001 (48.64%) MM-149 (29.13%)
KRAS	12	25225713T>G	Missense	K117N	/	COSM19940	63,65	MM-310 (47.09%) MM-269 (3.07%)
KRAS	12	25245328C>G	Missense	L19F	/	COSM12703	63	MM-386 (42.73%) MM-335 (18.6%)
KRAS	12	25245350C>A	Missense	G12V	rs121913529	COSM520	64,65	MM-146 (24.58%) MM-150 (8.93%)
KRAS	12	25227334A>T	Missense	Y64N	/	/	65	MM-115 (54.2%)
KRAS	12	25227266A>T	Missense	N86K	/	/	/	MM-115 (44.27%)
KRAS	12	25245332G>T	Missense	A18D	/	COSM542	65	PCL-021 (43.8%)
KRAS	12	25227334A>C	Missense	Y64D	/	/	63	MM-410 (41.96%)
KRAS	12	25245351C>A	Missense	G12C	rs121913530	COSM516	64	MM-243 (40.5%)
KRAS	12	25227348G>C	Missense	A59G	/	COSM28518	64	MM-079 (23.13%)
KRAS	12	25227330C>A	Missense	S65I	/	/	/	PCL-009 (15.84%)
KRAS	12	25245350C>G	Missense	G12A	rs121913529	COSM522	64,65	MM-140 (14.81%)
KRAS	12	25227342T>A	Missense	Q61L	rs121913240	COSM553	64,65	MM-284 (9.54%)
KRAS	12	25245351C>T	Missense	G12S	rs121913530	COSM517	/	MM-004 (9.18%)
KRAS	12	25225657C>T	Missense	S136N	/	/	/	MM-202 (4.86%)
KRAS	12	25225627G>A	Missense	A146V	/	COSM19900	64	MM-445 (2.53%)
DIS3	13	73346338C>T	Missense	D488N	/	COSM158635	65,71,80	MM-343 (46.2%) PCL-041 (45.7%) MM-398 (3.8%) MM-042 (4.6%)
DIS3	13	73336064C>T	Missense	R780K	/	COSM329311	63,65,71,79,80	PCL-019 (47.7%) PCL-042 (45.9%) MM-281 (36.8%)
DIS3	13	73355093T>C	Missense	T93A	/	/	/	PCL-021 (100%) MM-213 (56.6%)
DIS3	13	73335837G>A	Missense	R820W	rs372878316	COSM3469577	71	MM-445 (98.4%) MM-207 (2.1%)
DIS3	13	73335930G>A	Missense	R789W	/	/	71	MM-335 (94.2%) PCL-011 (91.1%)
DIS3	13	73336078A>C	Missense	F775L	/	/	65,71	MM-036 (92.6%) MM-414 (4.7%)
DIS3	13	73336064C>G	Missense	R780T	/	/	65,71	PCL-035 (45.%) MM-381 (16.4%)

Gene	Chr	Variant (GRCh38)	Mutation Type	AA change	dbSNP ID (v142)	COSMIC ID (v71)/IARC ID	MM literature*	Mutated patients (percentage of mutated sequencing reads)
DIS3	13	73336112T>C	Missense	H764R	/	/	/	PCL-036 (97.6%)
DIS3	13	73337707G>T	Missense	A670D	/	/	/	PCL-001 (97%)
DIS3	13	73355010G>C	Missense	F120L	/	/	/	MM-123 (95.3%)
DIS3	13	73355048G>A	Missense	R108C	/	/	71	MM-424 (94.3%)
DIS3	13	73355018T>C	Missense	K118E	/	/	/	PCL-015 (90.1%)
DIS3	13	73355804_73355830del	In_Frame_Del	L48_A56del	/	/	/	MM-340 (54.9%)
DIS3	13	73355968C>T	In_Frame_Del	M1_V16del	/	/	/	MM-464 (51%)
DIS3	13	73336151G>T	Missense	A751D	/	/	/	MM-340 (45.7%)
DIS3	13	73346340T>A	Missense	D487V	/	/	71	MM-317 (40.3%)
DIS3	13	73346341C>G	Missense	D487H	/	/	/	MM-372 (35.5%)
DIS3	13	73345240G>A	Missense	S550F	/	/	/	MM-386 (35.1%)
DIS3	13	73342930C>T	Missense	E626K	/	/	/	PCL-042 (33.2%)
DIS3	13	73346400C>T	Missense	R467Q	rs201674523	COSM1944518	65	MM-263 (31.1%)
DIS3	13	73336113G>A	Missense	H764Y	/	/	71	MM-207 (24.6%)
DIS3	13	73346400del	Fram_Shift_Del	R467Qfs*4	/	/	/	MM-263 (5.9%)
DIS3	13	73355110T>C	Missense	N87S	/	/	/	MM-279 (2.9%)
DIS3	13	73337650C>T	Missense	R689Q	/	COSM1367629	65	MM-340 (2.2%)
DIS3	13	73336113G>C	Missense	H764D	/	/	/	MM-207 (1.9%)
DIS3	13	73336077T>G	Missense	T776P	/	/	/	MM-207 (0.8%)
DIS3	13	73343049_73343050insA	Fram_Shift_Ins	A586Vfs*7	/	/	/	MM-055 (0.8%)
DIS3	13	73355891T>C	Missense	D27G	/	/	/	MM-143 (0.6%)
DIS3	13	73345094T>C	Missense	H568R	/	/	/	MM-310 (0.5%)
DIS3	13	73345097T>C	Missense	N567S	/	/	/	MM-150 (0.5%)
FAM46C	1	117623505_117623506insG	Fram_Shift_Ins	E213Gfs*8	/	/	/	MM-334 (100%)
FAM46C	1	117623702C>A	Missense	F278L	/	/	/	MM-314 (76.3%)
FAM46C	1	117623359A>C	Missense	N164T	/	/	/	PCL-011 (69.17%)
FAM46C	1	117623757G>A	Missense	A297T	rs149654076	COSM423576	/	MM-066 (57.97%)
FAM46C	1	117623335C>T	Missense	S156F	/	/	/	PCL-027 (47.8%)
FAM46C	1	117623145C>G	Missense	L93V	rs139168147	/	/	MM-150 (43.52%)
FAM46C	1	117623773_117623779del(7)‡	Fram_Shift_Del		/	/	/	MM-049 (41.01%)
FAM46C	1	117623783_117623801del(19)‡	Fram_Shift_Del	S302_R312delinsTK‡	/	/	/	MM-049 (41.01%)
FAM46C	1	117623805delG‡	Fram_Shift_Del		/	/	/	MM-049 (41.01%)
FAM46C	1	117623844C>T	Missense	R326C	/	/	/	MM-242 (40.13%)
FAM46C	1	117623252_117623253delAA	Fram_Shift_Del	K128Nfs*22	/	/	/	MM-448 (35.45%)
FAM46C	1	117623136G>C	Missense	D90H	/	/	65	MM-149 (33.73%)
FAM46C	1	117623740A>G	Missense	Y291C	/	/	71,83	MM-242 (33.55%)
FAM46C	1	117623741C>A	Nonsense	Y291*	/	/	/	MM-200 (25.13%)
FAM46C	1	117623587_117623588delGA	Fram_Shift_Del	G241Rfs*42	/	/	/	MM-448 (24.81%)
FAM46C	1	117623859_117623860delAA	Fram_Shift_Del	N331Pfs*33	/	/	/	MM-049 (24.72%)
FAM46C	1	117623030delC	Fram_Shift_Del	V55Sfs*17	/	/	/	MM-414 (22.5%)
FAM46C	1	117623579_117623584delAGAAAT	In_Frame_Indel	E237_I239delinsD	/	/	/	PCL-002 (21.07%)
FAM46C	1	117623674A>G	Missense	Y269C	/	/	/	MM-301 (14%)
FAM46C	1	117623268C>G	Missense	L134V	/	/	/	MM-004 (12.55%)
FAM46C	1	117623362T>G	Missense	V165G	/	/	/	MM-414 (8.99%)
FAM46C	1	117623366_117623380del(15)	In_Frame_Del	E166_V170del	/	/	/	MM-016 (5.11%)
FAM46C	1	117623923delA	Fram_Shift_Del	N352Tfs*45	/	/	/	MM-212 (6.77%)
FAM46C	1	117623290A>C	Missense	K141T	/	/	71	MM-246 (3.8%)
TP53	17	7673803G>A	Missense	R273C	rs121913343	IARC3730	/	PCL-030 (100%)
TP53	17	7674947A>G	Missense	I195T	rs760043106	IARC2446	64,65	PCL-027 (90.9%)
TP53	17	7673787G>A	Missense	P278L	rs876659802	IARC3817	/	PCL-018 (96.1%)
TP53	17	7670714_7670723del	Altered_Splicing	G245_splice	/	/	/	PCL-017 (70%)
TP53	17	7674245T>C	Missense	S240G	/	IARC3146	/	MM-262 (100%)
TP53	17	7673781C>T	Missense	R280K	rs121912660	IARC3849	/	PCL-031 (98.52%)
TP53	17	7675074C>T	Missense	E180K	rs879253911	IARC2229	/	PCL-012 (95.82%)
TP53	17	7675232G>T	Missense	S127Y	rs730881999	IARC1339	/	MM-375 (93.33%)
TP53	17	7673767C>T	Missense	E285K	rs112431538	IARC3932	71	PCL-004 (72.71%)
TP53	17	7675144_7675149delGCGGGT	In_Frame_Del	T155_R156del	/	IARC1785	/	PCL-037 (49.25%)
TP53	17	7676096C>T	Nonsense	W91*	rs876660548	IARC922	/	PCL-037 (47.34%)
TP53	17	7674149_7674230del	Altered_Splicing	I332_splice	/	/	/	MM-213 (8.22%)
TP53	17	7674238C>T	Missense	C242Y	rs121912655	IARC3185	/	MM-343 (10.96%)
TP53	17	7673794_7673795delCA	Frame_Shift_Del	A276Lfs*29	/	/	/	MM-213 (2.65%)

DIS3

Deep sequencing of *DIS3* revealed 30 different coding somatic mutations globally targeting 27 genic positions (**Table 6**). At DNA level, 27 out of 30 variants (90%) were SNVs, all resulting in missense mutations but one (73355968C>T, M1_V16del), which replaces one base within the first coding ATG, thus relocating the translation start site 16 codons downstream without altering the reading frame. The remaining three variants (10%) were indels: two involved a single different nucleotide causing a frameshift, while the other was a 27bp-deletion resulting in an in-frame deletion of nine amino acids (L48_A56del). Nonsense mutations were not observed.

DIS3 mutations globally affected 33 patients (33/164, 20.1%): 24 MMs at diagnosis (24/130, 18.5%), six primary PCLs (6/24, 25%), and three secondary PCLs (3/10, 30%).Twenty out of 30 mutations (70%) affected the RNB domain, while only eight variants (8/30, 26.7%) were located in the PIN domain. The remaining SNV mapped in a region just downstream of the RNB domain.

Overall, the most affected residue was the arginine at position 780 (R780), replaced by a lysine or by a threonine in three and two cases, respectively (**Figure 12B**). The arginine at position 780 was found involved in RNA binding function[129]. The second most affected residue in our dataset was the aspartic acid at position 488 (D488), involved in magnesium ion binding at the active site[129]. This residue was mutated in four cases, all carrying the D488N substitution, which also represents the globally most frequent variant in our series. Notably, two missense variants involving D487 residue were detected in our dataset: D487V and D487H, each in one patient. Other recurrently mutated positions in our series were T93, H764, F775, R789 and R820, each targeted in two samples by the same substitution, except for H764 (reported to be involved in RNA binding[129]) showing different amino acid changes. The 17 remaining genic positions were found privately mutated in our cohort. Notably, four patients carried multiple *DIS3* mutations; only in two of these cases (MM-340 and PCL-042) two variants were both at high allele frequency, while in the remaining samples only one of the mutations was above the detection limit of Sanger method, accompanied by other variants at low allele frequency. In particular, PCL-042 harbored the variants R780K and E626K, both in the RNB domain, in 46% and 33% of reads, respectively, and MM-340 carried a nine amino acid in-frame deletion in the PIN domain (L48_A56del) and a missense substitution in the RNB domain (A751D) in 55% and 46% of reads, respectively.

Beyond the aforementioned R780 and D487 residues, whose importance for the catalytic activity of the protein has been functionally proven, a potential detrimental effect on DIS3 nucleolytic activity can be suggested also for many of the other mutations we identified here, on the basis of their recurrence (as described above), and/or the high conservation of affected amino acid residues, and/or the functional predictions returned by different algorithms. In fact, from the alignment of human DIS3 protein sequence to mouse and yeast paralogs and to *E.coli* RNases II and

R sequences, we found that 13 of the 24 genic positions targeted by mutations resulting in amino acid substitutions or in-frame deletions in our primary patients encoded amino acid residues conserved in all the analyzed species.

FAM46C

We identified 24 tumor-specific mutations (**Table 6**) affecting 18/162 (11%) MM-pPCL-sPCL patients. Of these 24 mutations, 13 (54%) were single nucleotide variations (SNVs), all but one (Y291*) of which resulting in missense mutations, while the remaining 11 mutations (46%) were indels, six of which causing a frameshift. Globally, the predominantly inactivating nature of these identified alterations is consistent with the putative tumor suppressor role of *FAM46C* in MM.

It would appear that the *FAM46C* coding sequence is almost entirely altered by mutations, with the exception of small portions, especially in the N-terminal region. In fact, based on the recurrence of affected codons, it is difficult to identify mutational hotspots: only 20 out of the 84 residues are targeted by mutations in more than one patient (at most four in codon Y291), while the remainder are exclusively mutated (**Figure 12C**).

Concomitant *FAM46C* mutations occurred in four MM patients of our series, three of whom harboured two or more variants at high variant allele frequency. Notably, MM-049 showed four nucleotide deletions, three of which [117623773_11762377 9del(7), 117623783_117623801del(19) and 117623805delG] were carried by the same allele at an identical VAF (41.01%), and should generally result in the alteration S302_R312delinsTK.

The prevalence of *FAM46C* mutations (11.7%) identified in our newly diagnosed MM cohort was similar to that described in previous studies[63-65], but higher than the one more recently reported by Walker et al (2015). *FAM46C* mutations also occurred in extramedullary myeloma; however, based on the mutation frequency (4.2% and 20% in primary and secondary PC leukaemia, respectively), it appears difficult to outline their role in high-risk myeloma.

TP53

We identified 14 non-synonymous somatic variants in 7.3% (12/163) of cases (**Table 6**, **Figure 12D**): ten mutations were single nucleotide variations, all of which but one (introducing a premature stop codon in PCL-037) were missense mutations. The remaining four mutations were nucleotide deletions involving 2, 6, 10 and 82 base pairs respectively, only one of which (6-bp deletion in PCL-037) caused an amino acid deletion without alteration of the reading frame.

Mutation frequencies were 3.1% (4/129), 25% (6/24), and 20% (2/10) in MM, pPCL and sPCL patients, respectively (Freeman-Halton test, P = 0.0005), in line with the notion that *TP53* mutations

are associated with aggressive forms of PC dyscrasias and have a role in tumor progression than initiation[85, 86].

Exon 8 was the most frequently affected by mutations (5 variants), followed by exons 5 and 7 (3 variants each), and exons 4, 6 and 10 (one variant each), whereas no variants were found in exon 9. With the exception of the nonsense mutation W91* in case PCL-037 (localized in the SH3-like/Pro-rich structural motif) and the 10-nt deletion spanning intron 9-exon 10 junction in case PCL-017, all other variants targeted the DNA binding domain. In regards to the four nucleotide deletions identified, only one (T155_R156del in PCL-037) is reported in the IARC p53 website (http://www-p53.iarc.fr/index.html)[130], and none has been described in other MM series. In contrast, all ten single nucleotide variations found in our study are listed in the dataset of somatic mutations in sporadic cancers of the IARC database. Each of the variants is carried by a single patient in our cohort.

We evidenced that mutations cluster in the DNA binding domain; I195 and R273 are the most frequently mutated residues (I195T detected in three cases and I195M in one; R273H in two patients, and R273L and R273C each in one case), followed by P278 (P278S recurring twice, and P278L once), R248 (R248Q and R248W in one patient each) and E285 (E285K in two cases). Furthermore, virtually all the missense mutations identified in our study are classified as non-functional[130] in the IARC database (see Materials and Methods).

Noteworthy, we detected the occurrence of two concomitant *TP53* mutations in two patients of our dataset. In particular, PCL-037 carried the nonsense mutation W91* at a VAF of 47.3% and a 2-aa deletion (T155_R156del) at a VAF of 49.2%, suggesting the occurrence of both mutations virtually in the entire tumor population (due to the experimental design, it is not possible to determine whether the mutations are carried on two different chromosomes or on the same allele). MM-213 harbored a 82 nt-deletion (at VAF of 8.2%) comprising the last 50 nucleotides of exon 7 and the first 32 nucleotides of intron 7 and putatively originating an in frame amino acid deletion and altered splicing (due to the loss of a splice donor site), and a two-nucleotide deletion (at VAF of 2.6%) in exon 8 causing frameshift (A276Lfs*29).

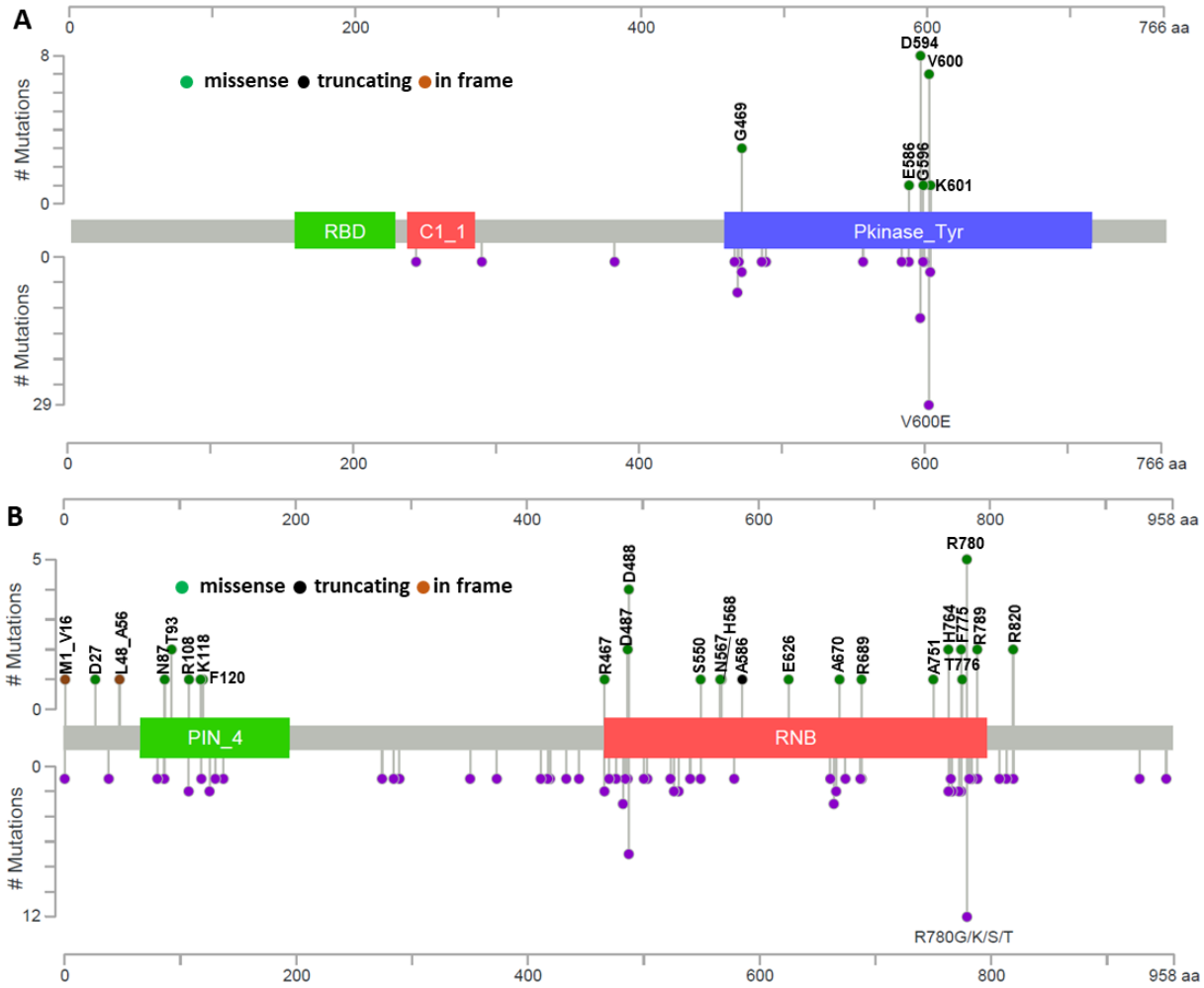

Figure 12. Representations of the mutations identified in MM/PCL patients (in the upper part of the figure) including overview of the literature (in the lower part of the figures). Amino acid changes caused by mutations found in our series are indicated. **(A)** *BRAF* mutations. RBD: Raf-like Ras-binding domain (156 – 226aa); C1_1: Phorbol esters/diacylglycerol binding domain (C1 domain) (235 – 282aa); Pkinase_Tyr: Protein tyrosine kinase (457 – 714aa). **(B)** *DIS3* mutations. PIN_4: PIN domain (66–195aa); RNB: RNB domain(467 – 797aa). To be continued.

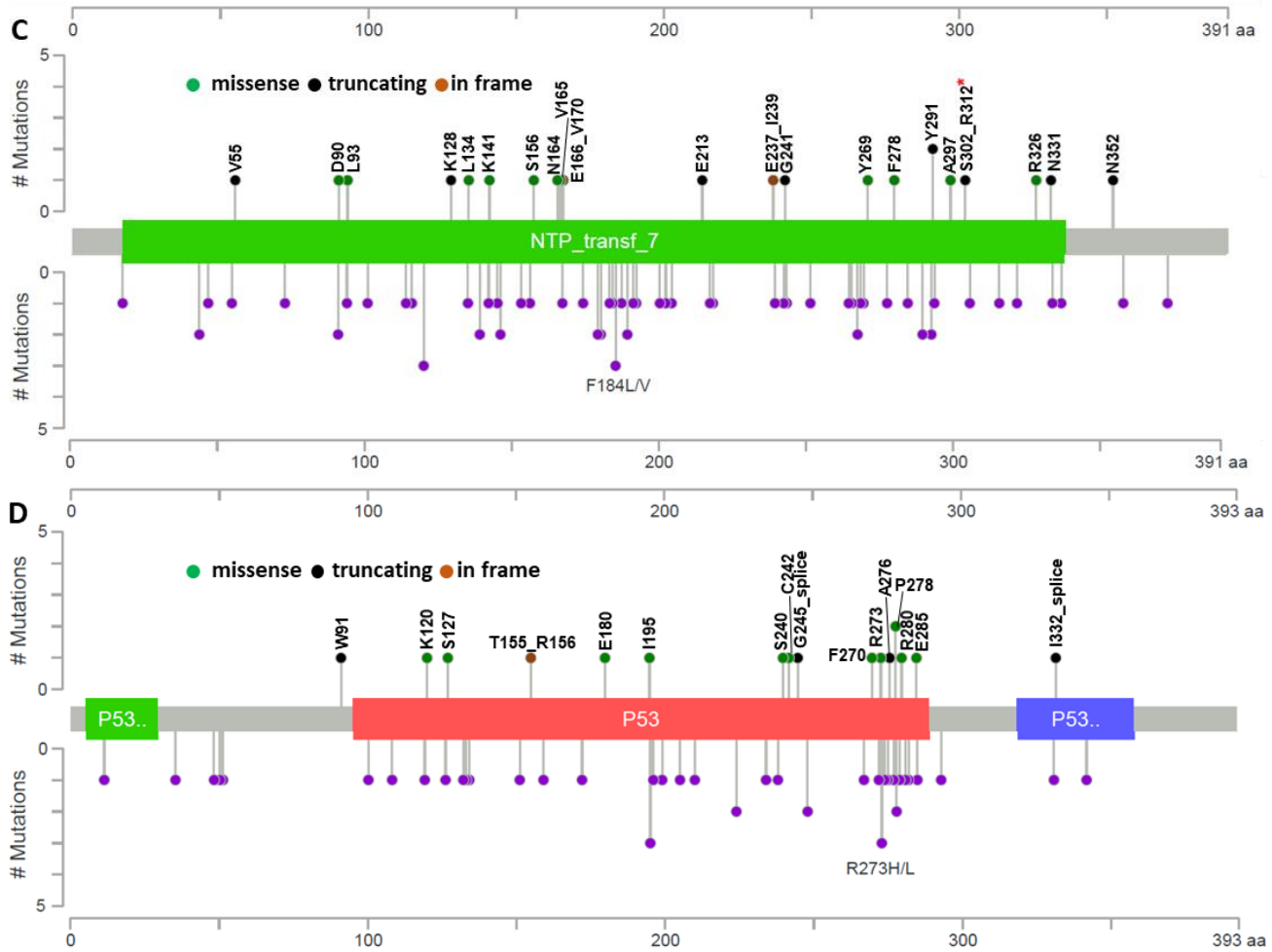

Figure 12_continued. (C) *FAM46C* mutations. NTP_transf_7: Nucleotidyltransferase (17 – 336aa). *The alteration S302_R312delinsTK is the result of three different deletions carried on the same allele in MM-049 patient (see text). **(D)** *TP53* mutations. P53_TAD: P53 transactivation motif (5 – 29aa); P53: P53 DNA-binding domain (95 – 289aa) P53_tetramer: P53 tetramerisation motif (319 – 358aa). Mutation types and corresponding color codes are as follows: Missense Mutations green, Truncating Mutations black; Inframe Mutations brown. Mutations depicted in the main dataset[83 63 65 64 71, 79, 80] purple.

Correlation with major chromosomal aberrations

Correlations of the analysed mutations with clinical and molecular charachteristics of the patients are reported in Table 7.

There was no significant association for *BRAF* mutations with any molecular lesion.

The *NRAS* mutations showed only a weak negative association with *MAF*/*MAFB* translocations (P=0.0775), and chromosome 17p deletions were less frequent in *KRAS*-mutated samples (P=0.0019). The latter were also characterized by more frequent hyperdiploidy (P=0.0124), although this association disappeared when the analysis was restricted to the MM samples alone (data not shown), probably because of the low prevalence of hyperdiploidy in the PCL samples.

Next, based on VAF, sample purity, and, if available, DNA copy number at *BRAF*, *NRAS* or *KRAS* genomic loci, we estimated the expected fraction of MM cells harboring each identified mutation. We then compared it with the percentage of CD138+ cells carrying main cytogenetic alterations (such as *IGH* translocations, hyperdiploidy, deletion of chromosome 13 or 1p, or 1q gain) as assessed by FISH (data not shown). In such a way, we found that one of the following scenarios (observed at quite comparable frequencies in our series) occurred in most of the cases: (i) the same fraction of cells (corresponding to the whole tumor clone or to a part thereof) was affected both by mutations and other chromosomal lesions, or (ii) mutations occurred in a subclone of MM cells harboring other alterations.

A positive association between *DIS3* mutations and *IGH* translocations was observed (P value = 0.0046), particularly with the t(11;14) (P value = 0.0198), and a lower occurrence of hyperdiploidy among *DIS3*-mutated cases (P value = 0.025). Conversely, we were unable to find any association with 17p deletion, 1q gain or 1p loss. Although not significantly associated with 13q14 deletion, *DIS3* mutations occurred in several cases carrying this chromosomal alteration. Specifically, 19 of the 33 (57.6%) *DIS3*-mutated patients harbored del(13q14). For eight of them, genome-wide DNA profiling data were also available[95, 110], indicating that in all but one case (PCL-019) the chromosomal deletion included the *DIS3* locus. Interestingly, VAFs around 100% were exclusively observed in a fraction of patients (10/19) with 13q deletion (and in all likelihood with heterozygous deletion of the *DIS3* locus). In the other cases, i.e. the remaining del(13q14) samples (9/19) and all the 14 patients disomic for chromosome 13, the VAFs we found were indicative of mutations occurring in hemizygosis in tumor subclones or carried in heterozygosis either by all or by a fraction of myelomatous cells.

We were unable to find any association of *FAM46C* mutations with major cytogenetic lesions, not even hyperdiploidy. In our series, 1p loss was investigated in 15/18 mutated cases by FISH using a probe mapping at 1p32.3. Only one case (MM-334) tested positive. Genome-wide DNA profiling analysis[110], performed in 11/15 patients, showed a complete concordance with cytogenetic data;

furthermore, it indicated that loss of *FAM46C* locus occurred in MM-334 and in an additional case (PCL-011) found to be disomic by FISH, but showing a del(1p11.2-1p31.3) by single nucleotide polymorphism array[110]. In line with these data, MM-334 and PCL-011 had VAFs of 100% and 70%, respectively.

As regards the correlations with the occurrence of *TP53* mutations, *IGH* translocations (particularly *MAF*-translocations) were found to be more frequent in *TP53*-mutated cases compared with the whole series (92% versus 50%, P=0.0024). *TP53* mutations were negatively associated with hyperdiploid status: no hyperdiploid cases among the *TP53*-mutated ones versus 33% in the entire cohort, P=0.009. However, the low *TP53* mutation prevalence in MM did not allow limiting contingency analyses to this class. As regards 17p deletion, a strong association of gene mutation with the deletion of the remaining allele was observed: in fact, deletion of *TP53* locus globally affected 10.6% of cases [4.7% (6/128) of MM, 29.2% (7/24) of pPCL and 44.4% (4/9) of sPCL patients respectively], while its prevalence reached 58% in *TP53*-mutated subset (P<0.0001). Overall, *TP53* inactivation as a result of gene deletion and/or mutation occurred in 7% of MM, 37.5% of pPCL and 44.4% of sPCL patients; in particular, the two lesions were simultaneously present respectively in 0.8%, 16.7% and 22.2% of cases, causing biallelic gene inactivation.

Table 7. Clinical and molecular characteristics of the total MM/PCL patients analyzed for *BRAF-NRAS-KRAS-DIS3-FAM46C-TP53* mutations.*Significance was assessed by Freeman-Halton extension of Fisher's exact test for disease type, and by Fisher's exact test for all other variables.

Charac-teristic	All patients (n=167)		BRAF wild type (n=147)		BRAF mutated (n=20)		P value*	NRAS wild type (n=127)		NRAS mutated (n=40)		P value*	KRAS wild type (n=118)		KRAS mutated (n=49)		P value*
	N	%	N	%	N	%		N	%	N	%		N	%	N	%	
MM	132	79	118	80.3	14	70		97	76.4	35	87.5		89	75.4	43	87.7	
pPCL	24	14.4	19	12.9	5	25	n.s.	23	18.1	1	2.5	0.008	20	17	4	8.2	n.s.
sPCL	11	6.6	10	6.8	1	5		7	5.5	4	10		9	7.6	2	4.1	
del(13q)	79	47.6	70	47.9	9	45		64	50.4	15	38.5		62	52.5	17	35.4	
13q disomic patients	87	52.4	76	52.1	11	55	n.s.	63	49.6	24	61.5	n.s.	56	47.5	31	64.6	n.s.
del(17p)	18	10.9	17	11.7	1	5		14	11.2	4	10		18	15.4	0	0	
17p disomic patients	147	89.1	128	88.3	19	95	n.s.	111	88.8	36	90	n.s.	99	84.6	48	100	0.0019
1q gain	68	43	61	43.9	7	36.8		53	44.2	15	39.5		53	46.5	15	34.1	
1q disomic patients	90	57	78	56.1	12	63.2	n.s.	67	55.8	23	60.5	n.s.	61	53.5	29	55.9	n.s.
1p loss	18	12.7	17	13.5	1	6.25		13	12.1	5	14.3		14	13.3	4	10.8	
1p disomic patients	124	87.3	109	86.5	15	93.8	n.s.	94	87.9	30	85.7	n.s.	91	86.7	33	89.2	n.s.
IGH trx	82	49.7	72	49.6	10	50		64	49.8	18	46.1		64	54.7	18	37.5	
no IGH trx	83	50.3	73	50.4	10	50	n.s.	62	50.2	21	53.8	n.s.	53	45.3	30	62.5	n.s.
HRD	50	32.9	45	33.8	5	26.3		37	32.2	13	35.1		29	26.6	21	48.8	
non-HRD	102	67.1	88	66.2	14	73.7	n.s.	78	67.8	24	64.9	n.s.	80	73.4	22	51.2	0.0124

Charac-teristic	All patients (n=164)		DIS3 wild type (n=131)		DIS3 mutated (n=33)		P value*	All patients (n=162)		FAM46C wild type (n=144)		FAM46C mutated (n=18)		P value*	All patients (n=163)		TP53 wild type (n=151)		TP53 mutated (n=12)		P value*
	N	%	N	%	N	%		N	%	N	%	N	%		N	%	N	%	N	%	
MM	130	79	106	81	24	73		128	79	113	78	15	83		129	79	125	83	4	33	
pPCL	24	15	18	14	6	18	n.s.	24	15	22	15	2	11	n.s.	24	15	18	12	6	50	0.0005
sPCL	10	6	7	5	3	9		10	6	9	6	1	6		10	6	8	5	2	17	
del(13q)	77	47	58	45	19	58		76	47	71	49	5	28		76	47	71	47	5	42	
13q disomic patients	86	53	72	55	14	42	n.s.	85	52	72	50	13	72	n.s.	86	53	79	53	7	58	n.s.
del(17p)	17	10.5	15	12	2	6		17	10	14	10	3	17		17	11	10	7	7	58	
17p disomic patients	145	89.5	114	88	31	94	n.s.	143	88	130	90	13	72	n.s.	144	89	139	93	5	42	<0.0001
1q gain	66	43	49	39	17	57		64	40	57	40	7	39		65	42	62	44	3	25	
1q disomic patients	89	57	76	61	13	43	n.s.	89	55	79	55	10	56	n.s.	89	58	80	56	9	75	n.s.
1p loss	18	13	16	14	2	7		18	11	17	12	1	6		18	13	17	13	1	8	
1p disomic patients	123	87	97	86	26	93	n.s.	122	75	108	75	14	78	n.s.	123	87	112	87	11	92	n.s.
IGH trx	81	50	58	44	23	74		80	49	71	49	9	50		80	50	69	46	11	92	
no IGH trx	81	50	73	56	8	26	0.0046	80	49	71	49	9	50	n.s.	81	50	80	54	1	8	0.0024
HRD	49	33	45	37	4	14		48	30	40	28	8	44		49	33	49	36	0	0	
non-HRD	101	67	77	63	24	86	0.025	100	62	92	64	8	44	n.s.	100	67	88	64	12	100	0.009

Longitudinal analysis of mutations

In order to gain further insights into the longitudinal status of mutations, we analyzed 19 patients for whom BM specimens were collected at two different timepoints: in particular, fourteen MM and two pPCL cases collected both at onset and at relapse; two MM patients at onset and at leukemic transformation, and one patient at early and relapsed leukemic phase of MM (**Table 8** and **Figure 13**). The second specimen in each case was collected after at least one line of treatment with various regimens.

In six cases (MM-239, M-263, MM-271, MM-282, MM-286, and MM-281), *BRAF/NRAS/KRAS* were wild-type at both timepoints. Four of the remaining patients had one mutated gene (*BRAF* in PCL-026, *NRAS* in MM-334, and *KRAS* in PCL-038 and MM-442) at a quite constant load throughout disease course. Others showed the appearance or disappearance of *NRAS* and *KRAS* variants with low VAFs (the occurrence of *KRAS* G12A and G12R in MM-340; the occurrence of *KRAS* Q61L in MM-200, which also retained *KRAS* Q61H at a quite constant VAF; the loss of *KRAS* G12S in MM-004; the loss of *KRAS* G13D in MM-151, which also stably carried *KRAS* Q61H; the loss of *NRAS* Q61K in MM-327, which also harbored the *KRAS* Q61H mutation in about 40% of the sequencing reads at both timepoints) (**Figure 13A-B-C**). A reduction in the mutational load of a fully clonal variant (*KRAS* Q22K, with an allele frequency of 50% at MM onset and 31% at relapse) was found in one case (MM-280) (**Figure 13B**). Interestingly, the disappearance of a high frequency mutation was associated in all cases with the occurrence/clonal expansion of a further mutation in another gene of the pathway (the loss of *KRAS* G12D in MM-429, which acquired the *NRAS* Q61R mutation; and the loss of *KRAS* G12V in MM-146, which showed a concurrent increase in the allele frequency of *NRAS* Q61H from 8% to 41%) (**Figure 13A-B**). Notably, the leukemic transformation in MM-295 was accompanied by a dramatic increase in the VAF of *NRAS* G12D (9% to 100%), whereas the mutation burden of the co-existing *BRAF* D594N variant remained stable at about 50% (**Figure 13A**).

Fifteen out of the 19 cases displayed a wild type *DIS3* status at both timepoints, while one patient (MM-281) carrying the R780K variant showed a quite constant VAF during disease course (36.8% mutated reads at diagnosis, and 47.56% at leukemic transformation) (**Figure 13D**). MM-004, wild type at onset, acquired a missense mutation affecting a fully conserved amino acid residue downstream the RNB domain and suggested to be of functional relevance (A827P); this SNV was detected in 97.1% of the sequencing reads at relapse. This patient showed two copies of chromosome 13 both at onset and at relapse, and the other cytogenetic lesions remained stable during disease progression. A considerable increase in mutation burden was observed in MM-263, with R467Q mutation occurring in 31.1% of reads at onset and 81.9% at relapse (this case also carried a frameshift mutation at the same position with VAF=5.93% at diagnosis, which disappeared at relapse). Finally, MM-340 carried three variants at onset (L48_A56del, A751D and R689Q at VAF

of 54.86%, 45.71% and 2.20%, respectively), two of which (L48_A56del and A751D) remained stable while the third (R689Q) was no longer detectable at relapse (**Figure 13D**). The disappearance of mutations at high allele frequency at onset was never observed in our cohort.

For *FAM46C* mutations, in three cases, all mutated at onset, we observed the presence of primary mutation at a quite constant VAF, or its clonal reduction/disappearance (**Figure 13E**). Unlike the findings reported by Kortum et al (2015) and Bolli et al (2014), the acquisition of *FAM46C* mutations over time was never observed in our cohort.

The analysis of these 19 cases revealed the acquisition of non-functional missense mutations (undetectable in the corresponding samples analyzed by NGS at diagnosis) in the DNA binding domain in three cases (**Figure 13F**). Specifically, the longitudinal analysis detected the substitutions P278T (VAF = 26.3%) in the myeloma patient MM-280 at relapse, F270L (VAF=97%) at the time of leukemic transformation in the myeloma patient MM-281, and K120E (VAF=22.1%) in the pPCL patient PCL-038 at relapse, respectively (**Figure 13F**). Notably, genome-wide DNA copy number data were also available for MM-281, indicating two interstitial deletions of 17p arm not involving *TP53* locus present at sub-clonal level in the sample at onset, and then detecting a larger deletion spanning from 17p12 to 17pter in virtually all cells of the later sample.

Table 8. Mutation status of analyzed genes 19 sequentially analyzed patients.

Sample name	Disease stage	*BRAF* status	*KRAS* status	*NRAS* status	*DIS3* status	*FAM46C* status	*TP53* status
MM-004	MM onset	wild type	G12S, 9.18%	wild type	wild type	L134, 12.55%	wild type
	MM relapse	wild type	wild type	wild type	A827P, 97.1%	wild type	wild type
MM-146	MM onset	wild type	G12V, 24.58%	Q61H, 6.11%	wild type	wild type	wild type
	MM relapse	wild type	wild type	Q61H, 40.78%	wild type	wild type	wild type
MM-151	MM onset	wild type	G13D, 9.35%; Q61H, 35.29%	wild type	wild type	wild type	wild type
	MM relapse	wild type	Q61H, 42.79%	wild type	wild type	wild type	wild type
MM-200	MM onset	wild type	Q61H, 23.74%	wild type	wild type	Y291*, 25.13%	wild type
	MM relapse	wild type	Q61H, 28.49%; Q61L, 5.07%	wild type	wild type	Y291*, 32.54%	wild type
MM-239	MM onset	wild type	wild type	wild type	wild type	wild type	wild type
	MM relapse	wild type	wild type	wild type	wild type	wild type	wild type
MM-263	MM onset	wild type	wild type	wild type	R467Q, 31.1%; R467Qfs*4, 5.9%	wild type	wild type
	MM relapse	wild type	wild type	wild type	R467Q, 81.9%	wild type	wild type
MM-271	MM onset	wild type	wild type	wild type	wild type	wild type	wild type
	MM relapse	wild type	wild type	wild type	wild type	wild type	wild type
MM-280	MM onset	wild type	Q22K, 50%	wild type	wild type	wild type	wild type
	MM relapse	wild type	Q22K, 31.41%	wild type	wild type	wild type	P278T, 26.3%
MM-282	MM onset	wild type	wild type	wild type	wild type	wild type	wild type
	MM relapse	wild type	wild type	wild type	wild type	wild type	wild type
MM-286	MM onset	wild type	wild type	wild type	wild type	wild type	wild type
	MM relapse	wild type	wild type	wild type	wild type	wild type	wild type
MM-327	MM onset	wild type	Q61H, 37.34%	Q61K, 4.28%	wild type	wild type	wild type
	MM relapse	wild type	Q61H, 38.15%	wild type	wild type	wild type	wild type
MM-334	MM onset	wild type	wild type	G12D, 96.49%	wild type	E213Gfs*8, 100%	wild type
	MM relapse	wild type	wild type	G12D, 92.94%	wild type	E213Gfs*8, 77.87%	wild type
MM-340	MM onset	wild type	wild type	wild type	L48_A56del, 54.9%; A751D, 45.7%; R689Q, 2.2%	wild type	wild type
	MM relapse	wild type	G12A, 4.23%; G12R, 10.05%	wild type	L48_A56del, 61%; A751D, 42.3%	wild type	wild type
MM-429	MM onset	wild type	G12D, 21.47%	wild type	wild type	wild type	wild type
	MM relapse	wild type	wild type	Q61R, 46%; E62Kfs*6, 4.09%	wild type	wild type	wild type
MM-295	MM onset	D594N, 50%	wild type	G12D, 8.89%	wild type	wild type	wild type
	MM leukemic transformation	D594N, 51.88%	wild type	G12D, 100%	wild type	wild type	wild type
MM-281	MM onset	wild type	wild type	wild type	R780K, 36.8%	wild type	wild type
	MM leukemic transformation	wild type	wild type	wild type	R780K, 47.6%	wild type	F270L, 97%
PCL-026	pPCL onset	D594N, 43.05%; E586K, 42.15%	wild type	wild type	wild type	wild type	wild type
	pPCL relapse	D594N, 40%; E586K, 42.8%	wild type	wild type	wild type	wild type	wild type
PCL-038	pPCL onset	wild type	G12R, 42.42%	wild type	wild type	wild type	wild type
	pPCL relapse	wild type	G12R, 43.17%	wild type	wild type	wild type	K120E, 22.1%
MM-442	MM leukemic transformation	wild type	Q61H, 45.93%	wild type	wild type	wild type	wild type
	sPCL relapse	wild type	Q61H, 52.75%	wild type	wild type	wild type	wild type

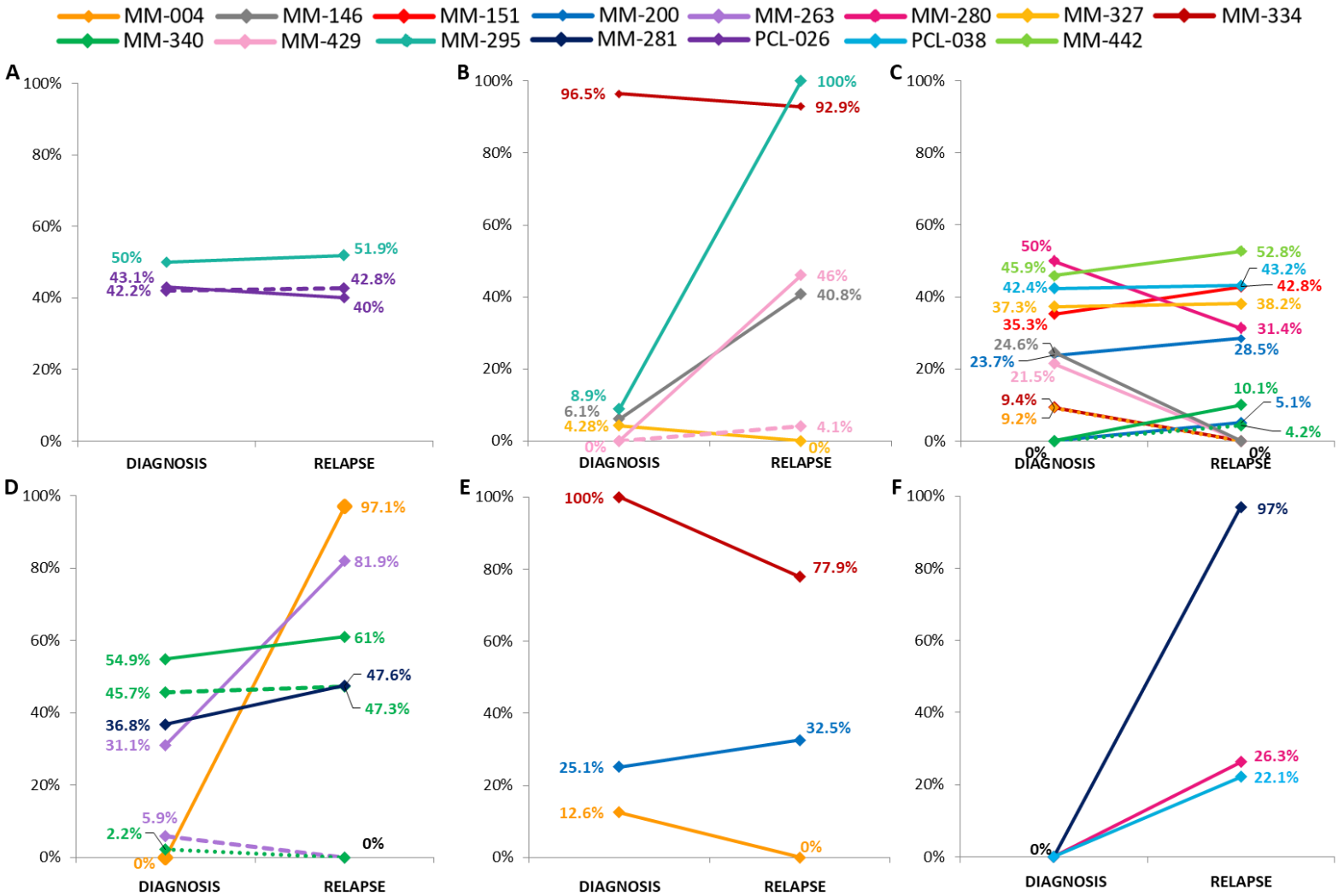

Figure 13. Changes of *BRAF* **(A)**, *NRAS* **(B)**, *KRAS* **(C)**, *DIS3* **(D)**, *FAM46C* **(E)** and *TP53* **(F)** mutational burden during disease progression. Allele frequencies of each variant are plotted at both time points for patients found mutated at diagnosis and/or relapse. The VAFs of mutations co-occurring in an individual patient are identified by the same color.

Expression of variants on cDNA

In order to assess whether the identified mutations were actually expressed, we sequenced those of our mutated cases for whom biological material was available. In particular, we were able to evaluate at transcriptional level 7 *BRAF*, 20 *DIS3* and 17 *FAM46C* variants, respectively (**Figure 14**). Specifically for *BRAF* gene, all variants but one were detected in RNA transcript, the only variant not identified had a VAF of 4.7% in genomic DNA in PCL-028 (D594E) (**Figure 14A**). In relation to *DIS3* gene, 16/20 (80%) evaluated variants were confirmed at transcriptional level (**Figure 14B**). Notably, all the variants detected at VAFs higher than 6% on genomic DNA were also identified at RNA level, while none of the five tested low-frequency genomic mutations (VAFs ranging from 1.95% to 5.96%) was detected on cDNA (**Figure 14A-B**). Overall, as regards expressed variants, mutant allele frequencies detected at gene and transcript level showed a good correlation (Pearson correlation coefficient=0.96 for both *BRAF* and *DIS3* gene).

Notably for *DIS3*, the aforementioned increase in mutation burden from diagnosis to relapse observed in MM-263 for the variant R467Q was accompanied by a doubling of the amount of expressed mutant allele (**Figure 14B**).

Specifically for *FAM46C* gene, all evaluated variants were detected and NGS analysis showed higher percentages of variant alleles in RNA transcripts compared to genomic DNA in some cases, and revealed that also mutations at low VAFs were expressed (**Figure 14C**). The finding that *FAM46C* mutated alleles had detectable biological expression, and in some cases seemed to be even preferentially transcribed, suggests that the mutations identified herein may have functional implications.

Overall, the NGS results indicated that the VAFs detected in genomic DNA and retro-transcribed RNA were significantly linearly correlated with a Pearson correlation coefficient of 0.85 (**Figure 14D**).

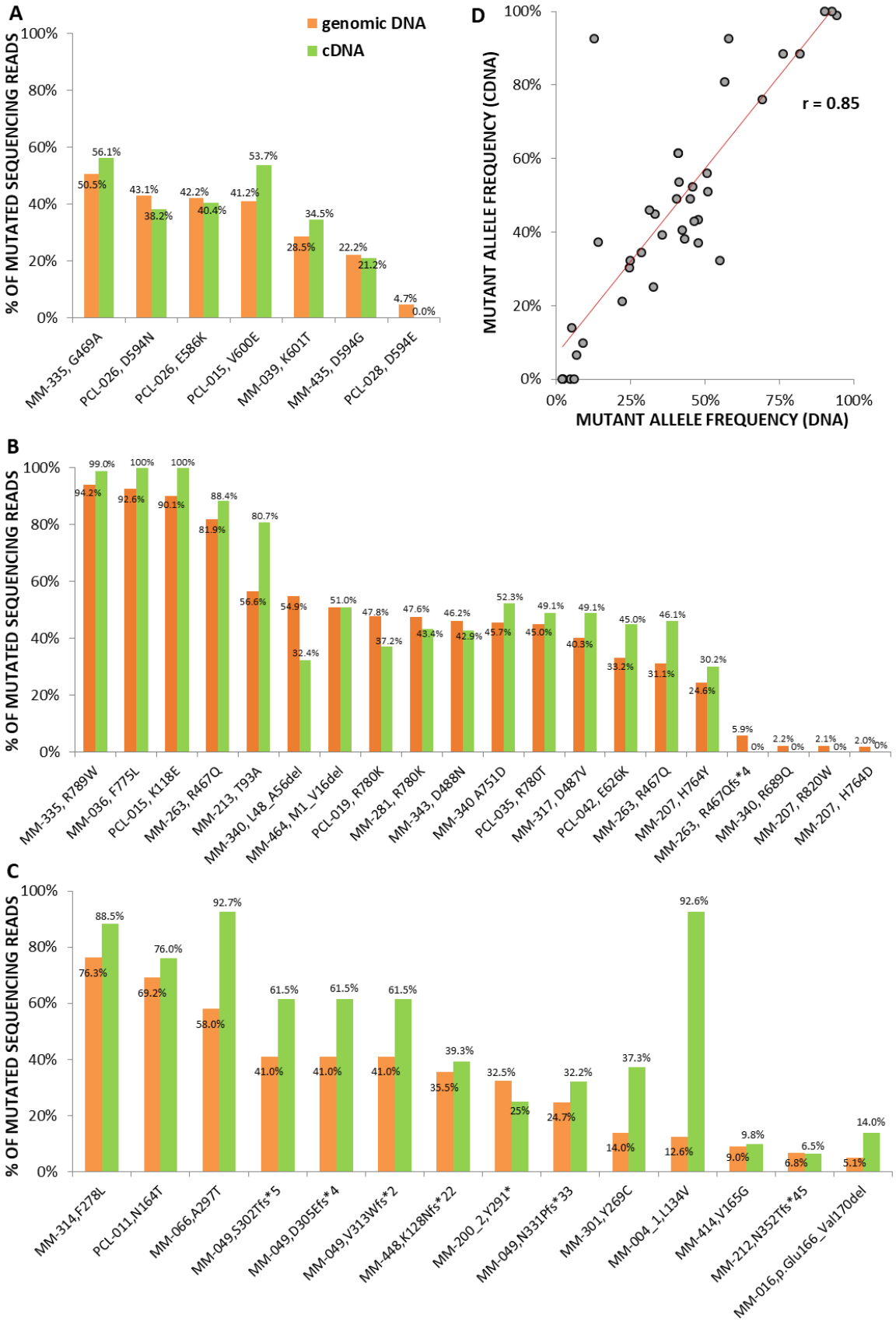

Figure 14. Mutations in *BRAF* **(A)**, *DIS3* **(B)** and *FAM46C* **(C)** detected on genomic DNA and cDNA. Percentages of variant sequencing reads identified by NGS analyses of genomic DNA and retrotranscribed total RNA. **(D)** Correlation between VAFs detected on genomic DNA and cDNA in all genes.

Distribution of *BRAF*, *NRAS*, *KRAS*, *DIS3*, *FAM46C* and *TP53* gene mutations among MM/PCL patients

Considering on the whole the three analyzed genes, *BRAF*, *KRAS* and *NRAS*, the MEK/ERK signalling pathway was affected by mutational events in more than half of the cases (96/167, 57.5%), being more frequent in sPCL (7/11, 63.6%) and MM (79/132, 59.8%), and relatively less frequent in pPCL (10/24, 41.7%). We identified 13 samples with concomitant mutations in two genes: three cases in *BRAF* and *NRAS*, five in *BRAF* and *KRAS*, and five in *NRAS* and *KRAS*. In all patients, the co-existing mutations had different VAFs, thus supporting the occurrence of tumor subclones. Notably, five of the eight *BRAF* variants in D594 that are predicted to cause *BRAF* inactivation[131] co-existed with a *NRAS* or *KRAS* mutation.

Similar to *FAM46C*, *DIS3*, another gene commonly mutated in MM, is an RNA-binding protein. The molecular mechanism(s) in which *FAM46C* and *DIS3* are involved and the putative targets are yet to be determined, although they have been described to couple RNA translation with degradation[132]. In our cohort, three out of 18 *FAM46C* mutated cases harbored simultaneous *DIS3* mutations (only in one case both alterations had high VAF)[133]. Overall, we observed mutations in these two RNA-binding proteins in 30% of patients, further supporting the hypothesis of the involvement of translational control in MM pathogenesis[63].

A fraction of *TP53* mutations co-occurred with variants affecting other MM driver genes (**Figure 15**). Indeed, two *TP53*-mutated cases carried also mutations in *DIS3*, and another one was simultaneously mutated in *TP53*, *NRAS* and *FAM46C* genes. However, *TP53* mutations were not positively associated with those of any other gene; on the contrary, the absence of *TP53* mutations in *KRAS*-mutated cases was deemed statistically significant (P=0.019).

Concerning variant allele frequencies, in *TP53*-mutated patients positive for del(17p) VAFs were almost always indicative of the exclusive presence of the mutated allele. On the contrary, in mutated cases at onset and disomic for chromosome 17, VAFs ranged from 2.65% to 100%, suggesting the occurrence of minor mutated subclones in MM-213 and MM-343. Furthermore, we observedthe presence of clonal heterozygous mutations in PCL-037; the predominance of mutated allele in PCL-004; and apparent copy number-neutral loss of heterozigosity in MM-262. As stated above, in the three patients longitudinally analyzed and showing in the late sample the emergence of *TP53* mutations not visible in the earlier sample, VAFs were of 22.1%, 26.3% and 97%, respectively.

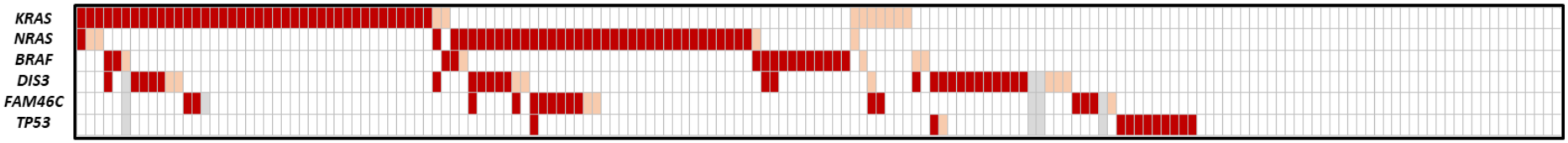

Figure 15. Heat map distribution of *TP53*, *DIS3*, *FAM46C*, *BRAF*, *KRAS* and *NRAS* gene mutations among MM/PCL patients. The rows correspond to the indicated genes, and columns represent individual MM or PCL samples, which are color-coded on the basis of gene status (white: wild-type; light red: Sanger-undetectable mutations; dark red: Sanger-detectable mutations).

Transcriptomic profile of patients with *BRAF/NRAS/KRAS/DIS3*-mutations

In order to identify the transcriptional profiles related to *BRAF*, *NRAS* and *KRAS* mutations in primary tumors, we used microarray technology to investigate a large number (n = 142) of the samples analyzed by NGS. Assuming that alterations in a limited number of myeloma cells do not appreciably affect gene expression, we arbitrarily chose a lower VAF cut-off value of 20%, thus resulting in the comparison of 60 wild-type versus 68 mutated patients. Eighty-six genes resulted differentially expressed (**Table 9A**): 27 up- and 59 down-regulated in the mutated cases. Particularly, 18 genes emerged at the highest stringency level. Interestingly, functional enrichment analysis revealed the involvement of a statistically significant fraction of modulated transcripts in MAPK signalling (*PRKD2, SPRED2, MAPKAPK2, CD300A, ARL6IP5, DUSP6, PPM1L, GRB2, LAMTOR3, SPRED1, LYN, EDN1, RASGRP1* and *ACVR1B*) at both biological process (GO:0000165, q-value=3.29E-03) and pathway level (198779 WikiPathway, q-value=1.33E-02) (**Table 10**).

Principal component analysis based on the expression of the 18 most statistically significant genes allowed a good separation between mutated and wild-type patients, without any apparent gene-specific discrimination (**Figure 16**).

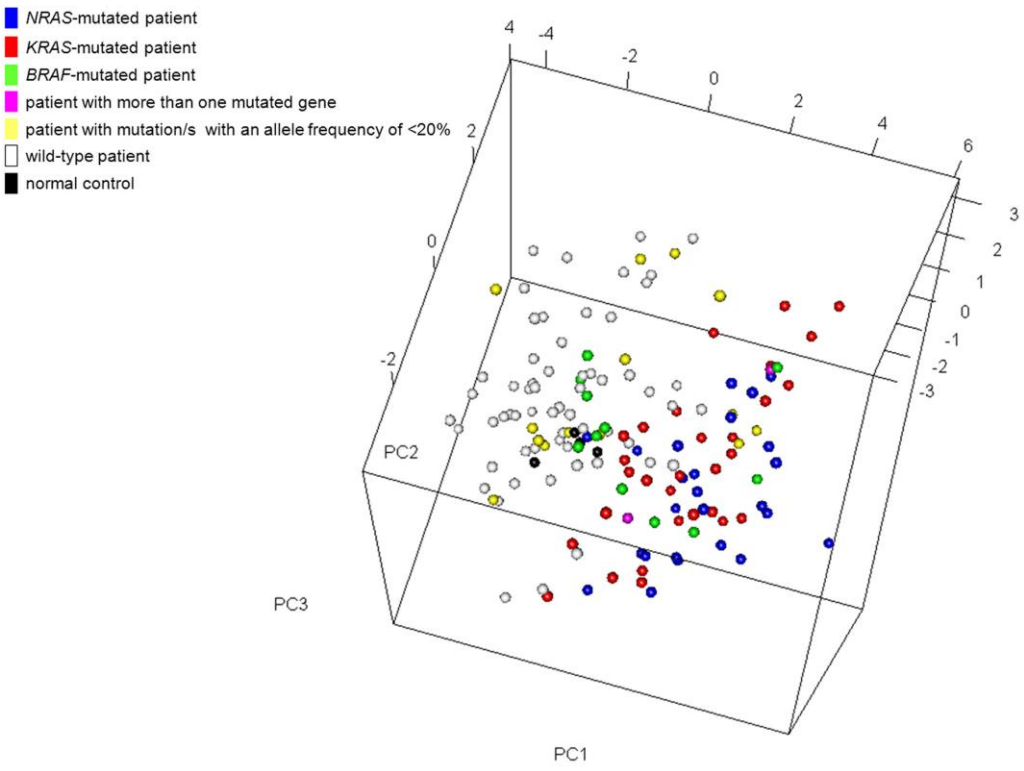

Figure 16. Multidimensional scaling plot using the most significant genes (n = 18) differentially expressed between *BRAF/NRAS/KRAS*-mutated and wild-type patients. Each point represents a single sample and is coloured on its type (patient/normal control) and mutation status as measured by sequence analysis. In the case of co-existing mutations of which only one had an allele frequency of >20%, the color corresponds to the gene with the highest mutational load.

Table 9. Modulated genes between **(A)** *BRAF/NRAS/KRAS* mutated and wild type patients and **(B)** *DIS3*-wild type and mutated patients. *Differentially expressed gene at the highest stringency level. ‡Negative scores indicate downregulated genes in mutated cases.

A

Gene	SAM score‡	Gene	SAM score
*PLEKHA1**	-5.163765	*ETV5**	6.24113
*ARL6IP5**	-4.342202	*CD300A**	5.77917
*ZHX1**	-4.311072	*SPRED2**	5.37656
*SMAD3**	-4.287826	*DUSP6**	4.36551
*AHNAK**	-4.268764	*NRXN1**	4.23105
*TBC1D1**	-4.237025	*MORC1**	4.20022
*MTMR4**	-4.167526	*F12**	4.15845
*COPZ2**	-4.103426	*RRBP1**	4.07138
*RASGRP1**	-4.07288	*TDO2**	4.03133
FLNB	-3.961337	*LGALS3BP*	3.97907
BMP2K	-3.918772	*PRKD2*	3.94335
RIOK1	-3.907891	*ACVR1B*	3.94093
TSHZ1	-3.904952	*ASPHD2*	3.92401
NOTCH2	-3.883492	*GPR56*	3.89699
ZNF681	-3.859508	*ACOXL*	3.85393
DHCR24	-3.82483	*SPRED1*	3.8092
LIFR	-3.819499	*LY9*	3.76526
ATL3	-3.818015	*ABI3BP*	3.74618
TGIF2	-3.794277	*DPEP1*	3.70693
GLO1	-3.776065	*TMEM184B*	3.70255
C6orf89	-3.76636	*PDE4D*	3.69971
ZNF570	-3.738105	*P2RY6*	3.68381
CYB5R1	-3.727057	*CASP4*	3.66862
PHTF2	-3.721193	*MIR221*	3.65834
DENND5A	-3.687141	*SEMA4D*	3.65632
CD52	-3.67123	*MAPKAPK2*	3.64609
ZADH2	-3.659897	*SIX4*	3.64407
RRP8	-3.635111		
BCL2	-3.627357		
JAZF1	-3.621626		
CD86	-3.612113		
LAMTOR3	-3.611917		
SAMSN1	-3.587501		
NTAN1	-3.586428		
PLCL1	-3.549426		
LRIG3	-3.547287		
PPM1L	-3.545335		
GRB2	-3.537068		
NECAP2	-3.522351		
KDM1B	-3.513606		
KAT6A	-3.51165		
BCL11A	-3.499956		
NEU1	-3.491002		
KIAA1147	-3.487048		
NUDT7	-3.455732		
SLC23A2	-3.446583		
LSS	-3.442877		
LYPD6B	-3.440096		
ARHGEF3	-3.436137		
CLN3	-3.423259		
INSIG1	-3.421565		
LYN	-3.382145		
APBB1IP	-3.378885		
MYL2	-3.372364		
TOP3A	-3.359096		
AKAP5	-3.358922		
RPSAP58	-3.357769		
EDN1	-3.355883		
SQLE	-3.353543		

B

Gene	SAM score‡	Gene	SAM score
LOC100132356	-7.27998	*GVINP1*	-3.16364
DKFZP434I0714	-7.24885	*KHDRBS2*	-3.14396
ILF3-AS1	-5.8891	*MTG1*	-3.1248
FLJ30403	-5.73335	*FAM154B*	-3.12346
ARHGAP5-AS1	-5.63017	*ST3GAL1*	-3.12014
HEXA-AS1	-5.52727	*LOC100128288*	-3.10885
C5orf54	-5.41888	*C7orf13*	-3.10235
C11orf71	-5.0731	*TTC30B*	-3.10072
ZNF594	-5.01186	*CXXC4*	-3.09534
FLJ38717	-4.90146	*IFIT1*	-3.08992
LOC100049716	-4.82276	*MIR320A*	-3.08972
CXorf21	-4.79455	*RNU4ATAC*	-3.08157
LOC648987	-4.75759	*PAPSS2*	-3.07687
TAPT1-AS1	-4.65361	*HSD17B7*	-3.07454
C10orf111	-4.55991	*MATN1-AS1*	-3.0743
APOBEC4	-4.54461	*LRRC23*	-3.05173
PRORSD1P	-4.35225	*ZNF141*	-3.04218
LINC00167	-4.34833	*MAP2*	-3.02161
DDX60	-4.29767	*MED31*	-3.01969
CATSPER2P1	-4.21904	*SFR1*	-3.01663
SPRYD4	-4.19583	*APOBEC3F*	-3.01397
POP1	-4.13861	*GDF9*	-3.01336
PFN4	-4.08844	*MPZ*	-2.99457
LINC00173	-3.94784	*LOC645212*	-2.98619
LOC153684	-3.93605	*CDNF*	-2.98353
APOLD1	-3.92105	*C1orf74*	-2.97908
LOC100128398	-3.87278	*TP53TG1*	-2.97753
NUDT9P1	-3.83567	*RASL11B*	-2.97566
C1orf220	-3.6607	*LOC100129726*	-2.97513
RNU12	-3.64519	*GNRHR2*	-2.97302
HTATSF1P2	-3.62834	*GPR135*	-2.96435
DHX58	-3.61677	*OASL*	-2.96341
RNU11	-3.61058	*CCDC87*	-2.94704
CENPM	-3.58134	*HCG27*	-2.93779
STPG2	-3.54494	*RNU5F-1*	-2.92498
FAM218A	-3.50969	*FAM227B*	-2.91766
FREM3	-3.49656	*MIR23A*	-2.91476
FIGN	-3.47013	*MIRLET7BHG*	-2.91089
IFIT3	-3.46587	*RIBC2*	-2.8987
ZNF268	-3.46146	*ZBTB26*	-2.89665
NS3BP	-3.40966	*PAN2*	-2.89352
ANXA2R	-3.36509	*FZD7*	-2.87535
LDOC1L	-3.3559	*CCDC85C*	-2.87272
C21orf119	-3.34986	*DLGAP1-AS1*	-2.86802
C11orf65	-3.33332	*HIST4H4*	-2.85887
FLJ37201	-3.32968	*KANSL1L*	-2.85357
LOC148696	-3.32702	*IFI6*	-2.84554
DPM3	-3.30907	*EID2B*	-2.83955
SLC4A5	-3.2872	*CCDC53*	-2.83361
LOC100506571	-3.28623	*DDX60L*	-2.82877
FLJ38576	-3.26721	*TTC23*	-2.81725
PLA2G6	-3.24009	*ZNF70*	-2.81337
ZNRD1-AS1	-3.23349	*COQ2*	-2.81171
PCDHGA1	-3.22484	*ODF3B*	-2.80322
PATL2	-3.21516	*AGBL2*	-2.79022
C9orf43	-3.21034	*TMEM234*	-2.77988
RNU5D-1	-3.20768		
IL22RA1	-3.20672		
C16orf54	-3.20079		
SCIN	-3.1948		
C5orf56	-3.18962		
LINC00528	-3.17856		
RBM45	-3.1685		

Table 10. Enriched biological processes in the MAPK mutation-associated 86 gene signature (q value<0.005).

ID	Name	q value	Hit in Query List
GO:0010563	negative regulation of phosphorus metabolic process	1.29E+03	*SEMA4D, AKAP5, PDE4D, SPRED2, CD300A, DUSP6, SPRED1, SAMSN1, LYN, EDN1, SMAD3*
GO:0010564	negative regulation of phosphate metabolic process	1.29E+03	*SEMA4D, AKAP5, PDE4D, SPRED2, CD300A, DUSP6, SPRED1, SAMSN1, LYN, EDN1, SMAD3*
GO:0010565	regulation of phosphorylation	1.29E+03	*PRKD2, SEMA4D, AKAP5, PDE4D, SPRED2, MAPKAPK2, CD300A, ARL6IP5, DUSP6, GRB2, LAMTOR3, SPRED1, BCL2, PLCL1, ACVR1B, SAMSN1, LYN, EDN1, SMAD3*
GO:0010566	regulation of protein phosphorylation	1.29E+03	*PRKD2, SEMA4D, AKAP5, PDE4D, SPRED2, MAPKAPK2, CD300A, DUSP6, LAMTOR3, SPRED1, BCL2, PLCL1, ACVR1B, SAMSN1, LYN, EDN1, SMAD3*
GO:0010567	sterol biosynthetic process	1.29E+03	*DHCR24, INSIG1, SQLE, LSS, CYB5R1*
GO:0010568	regulation of intracellular signal transduction	1.29E+03	*PRKD2, SEMA4D, RASGRP1, AKAP5, SPRED2, MAPKAPK2, TBC1D1, CD300A, ARL6IP5, DUSP6, GRB2, LAMTOR3, GPR56, PLEKHA1, SPRED1, BCL2, DENND5A, ARHGEF3, LYN, EDN1, NOTCH2*
GO:0010569	positive regulation of signaling	1.39E+03	*PRKD2, SEMA4D, AKAP5, SPRED2, NRXN1, MAPKAPK2, CD300A, ARL6IP5, GRB2, LAMTOR3, GPR56, SPRED1, BCL2, ACVR1B, LYN, EDN1, NOTCH2, SMAD3*
GO:0010570	positive regulation of cell communication	1.39E+03	*PRKD2, SEMA4D, AKAP5, SPRED2, NRXN1, MAPKAPK2, CD300A, ARL6IP5, GRB2, LAMTOR3, GPR56, SPRED1, BCL2, ACVR1B, LYN, EDN1, NOTCH2, SMAD3*
GO:0010571	positive regulation of intracellular signal transduction	1.39E+03	*PRKD2, SEMA4D, AKAP5, SPRED2, MAPKAPK2, CD300A, ARL6IP5, LAMTOR3, GPR56, SPRED1, BCL2, LYN, EDN1, NOTCH2*
GO:0010572	positive regulation of response to stimulus	1.39E+03	*PRKD2, SEMA4D, AKAP5, PDE4D, SPRED2, MAPKAPK2, CD86, CD300A, ARL6IP5, DUSP6, GRB2, LAMTOR3, GPR56, PLEKHA1, SPRED1, BCL2, ACVR1B, LYN, EDN1, NOTCH2, SMAD3*
GO:0010573	negative regulation of protein phosphorylation	1.39E+03	*SEMA4D, PDE4D, SPRED2, CD300A, DUSP6, SPRED1, SAMSN1, LYN, SMAD3*
GO:0010574	regulation of phosphate metabolic process	1.63E+03	*PRKD2, SEMA4D, RASGRP1, AKAP5, PDE4D, SPRED2, MAPKAPK2, TBC1D1, CD300A, ARL6IP5, DUSP6, GRB2, LAMTOR3, SPRED1, BCL2, PLCL1, DENND5A, ARHGEF3, ACVR1B, SAMSN1, LYN, EDN1, SMAD3*
GO:0010575	immune response activating signal transduction	1.63E+03	*PRKD2, PDE4D, MAPKAPK2, CD86, CD300A, DUSP6, GRB2, PLEKHA1, BCL2, LYN*
GO:0010576	regulation of phosphorus metabolic process	1.63E+03	*PRKD2, SEMA4D, RASGRP1, AKAP5, PDE4D, SPRED2, MAPKAPK2, TBC1D1, CD300A, ARL6IP5, DUSP6, GRB2, LAMTOR3, SPRED1, BCL2, PLCL1, DENND5A, ARHGEF3, ACVR1B, SAMSN1, LYN, EDN1, SMAD3*
GO:0010577	positive regulation of signal transduction	1.63E+03	*PRKD2, SEMA4D, AKAP5, SPRED2, MAPKAPK2, CD300A, ARL6IP5, GRB2, LAMTOR3, GPR56, SPRED1, BCL2, ACVR1B, LYN, EDN1, NOTCH2, SMAD3*
GO:0010578	immune response regulating signaling pathway	2.03E+03	*PRKD2, RASGRP1, PDE4D, MAPKAPK2, CD86, CD300A, DUSP6, GRB2, PLEKHA1, BCL2, LYN*
GO:0010579	positive regulation of developmental process	2.03E+03	*PRKD2, SEMA4D, TGIF2, AKAP5, NRXN1, CD86, ETV5, BCL2, BCL11A, ACVR1B, SIX4, LYN, EDN1, NOTCH2, SMAD3*
GO:0010580	signal transduction by phosphorylation	2.03E+03	*PRKD2, SPRED2, MAPKAPK2, CD300A, ARL6IP5, DUSP6, PPM1L, GRB2, LAMTOR3, SPRED1, ACVR1B, LYN, EDN1*
GO:0010581	negative regulation of cellular metabolic process	2.74E+03	*SEMA4D, TGIF2, AKAP5, PDE4D, SPRED2, CLN3, CD300A, INSIG1, DUSP6, KAT6A, JAZF1, SPRED1, RRP8, BCL2, BCL11A, SAMSN1, LYN, ZHX1, EDN1, NOTCH2, SMAD3*
GO:0010582	negative regulation of MAP kinase activity	2.74E+03	*SPRED2, CD300A, DUSP6, SPRED1, LYN*
GO:0010583	negative regulation of phosphorylation	2.74E+03	*SEMA4D, PDE4D, SPRED2, CD300A, DUSP6, SPRED1, SAMSN1, LYN, SMAD3*
GO:0010584	negative regulation of metabolic process	2.74E+03	*SEMA4D, TGIF2, AKAP5, PDE4D, SPRED2, CLN3, CD300A, INSIG1, DUSP6, KAT6A, JAZF1, SPRED1, RRP8, BCL2, BCL11A, ACVR1B, SAMSN1, LYN, ZHX1, EDN1, NOTCH2, SMAD3*
GO:0010585	protein phosphorylation	2.74E+03	*PRKD2, SEMA4D, RIOK1, AKAP5, PDE4D, SPRED2, BMP2K, MAPKAPK2, CD300A, DUSP6, LAMTOR3, SPRED1, BCL2, PLCL1, ACVR1B, SAMSN1, LYN, EDN1, SMAD3*
GO:0010586	regulation of multicellular organismal development	2.76E+03	*PRKD2, SEMA4D, TGIF2, NEU1, AKAP5, NRXN1, BMP2K, CD86, DUSP6, ETV5, GPR56, BCL2, BCL11A, ACVR1B, SIX4, LYN, EDN1, NOTCH2, SMAD3*
GO:0010587	activation of immune response	2.77E+03	*PRKD2, PDE4D, MAPKAPK2, CD86, CD300A, DUSP6, GRB2, PLEKHA1, BCL2, LYN*
GO:0010588	B cell receptor signaling pathway	3.23E+03	*CD300A, PLEKHA1, BCL2, LYN*
GO:0010589	immune response regulating cell surface receptor signaling pathway	3.29E+03	*PRKD2, RASGRP1, PDE4D, CD86, CD300A, GRB2, PLEKHA1, BCL2, LYN*
GO:0010590	regulation of cell migration	3.29E+03	*DPEP1, PRKD2, SEMA4D, P2RY6, CD300A, GPR56, BCL2, ACVR1B, LYN, EDN1, SMAD3*
GO:0010591	regulation of cellular component movement	3.29E+03	*DPEP1, PRKD2, SEMA4D, PDE4D, P2RY6, CD300A, GPR56, BCL2, ACVR1B, LYN, EDN1, SMAD3*
GO:0010592	MAPK cascade	3.29E+03	*PRKD2, SPRED2, MAPKAPK2, CD300A, ARL6IP5, DUSP6, PPM1L, GRB2, LAMTOR3, SPRED1, LYN, EDN1*
GO:0010593	negative regulation of cellular protein metabolic process	3.62E+03	*SEMA4D, PDE4D, SPRED2, CLN3, CD300A, DUSP6, SPRED1, SAMSN1, LYN, EDN1, SMAD3*
GO:0010594	phosphorylation	3.73E+03	*PRKD2, SEMA4D, RIOK1, AKAP5, PDE4D, SPRED2, BMP2K, MAPKAPK2, CD300A, ARL6IP5, DUSP6, PPM1L, GRB2, LAMTOR3, SPRED1, BCL2, PLCL1, ACVR1B, SAMSN1, LYN, EDN1, SMAD3*
GO:0010595	cholesterol biosynthetic process	3.73E+03	*DHCR24, INSIG1, SQLE, LSS*
GO:0010596	enzyme linked receptor protein signaling pathway	4.06E+03	*LIFR, PRKD2, TGIF2, MTMR4, MAPKAPK2, CD86, DUSP6, PPM1L, GRB2, PLEKHA1, ARHGEF3, ACVR1B, LYN, NOTCH2, SMAD3*
GO:0010597	ear development	4.06E+03	*LRIG3, MAPKAPK2, INSIG1, TSHZ1, BCL2, SIX4, EDN1*
GO:0010598	antigen receptor mediated signaling pathway	4.06E+03	*PRKD2, PDE4D, CD300A, PLEKHA1, BCL2, LYN*
GO:0010599	regulation of protein modification process	4.06E+03	*PRKD2, SEMA4D, AKAP5, PDE4D, SPRED2, MAPKAPK2, CD300A, DUSP6, LAMTOR3, SPRED1, BCL2, PLCL1, ACVR1B, SAMSN1, LYN, EDN1, SMAD3*
GO:0010600	regulation of MAPK cascade	4.06E+03	*PRKD2, SPRED2, MAPKAPK2, CD300A, ARL6IP5, DUSP6, GRB2, LAMTOR3, SPRED1, LYN, EDN1*
GO:0010601	regulation of cell motility	4.08E+03	*DPEP1, PRKD2, SEMA4D, P2RY6, CD300A, GPR56, BCL2, ACVR1B, LYN, EDN1, SMAD3*
GO:0010602	regulation of immune response	4.11E+03	*PRKD2, RASGRP1, PDE4D, MAPKAPK2, CD86, CD300A, DUSP6, GRB2, PLEKHA1, BCL2, SAMSN1, LYN, SMAD3*
GO:0010603	peptidyl amino acid modification	4.90E+03	*PRKD2, SEMA4D, PDE4D, SPRED2, MAPKAPK2, CD300A, KAT6A, SPRED1, ASPHD2, BCL2, PLCL1, ACVR1B, SAMSN1, LYN*

Notably, a few tumors not affected by mutations in any of the three genes (but possibly mutated in other genes of the pathway) showed the same activated MAPK pathway transcriptional profile as the mutated cases, as it has been found in other cancers[134]. Conversely, some mutated samples had a wild-type-like expression profile, including 10 cases carrying mutations with a low VAF; MM-295 (*BRAF* D594N); and PCL-026, harboring both D594N and E586K *BRAF* mutations. Interestingly, NGS analysis of PCL-026 indicated that both variants were carried on the same allele, thus suggesting that the putative abrogation of *BRAF* negative regulation generated by the E586K substitution may not lead to increased *BRAF* signaling because of the concomitant occurrence of the D594N mutation.

To identify transcriptional gene profiles related to *DIS3* mutations, we performed a two-classes supervised analysis, choosing to set at 20% the lower VAF cut-off, thus leading the comparison of 89 wild type *versus* 13 mutated patients. This analysis revealed that 119 genes (q-value < 0.1) (**Table 9B**) were differentially expressed between the two groups; among them, 28 genes were identified at stringent level (q-value=0) and they are depicted in the heatmap (**Figure 17**).

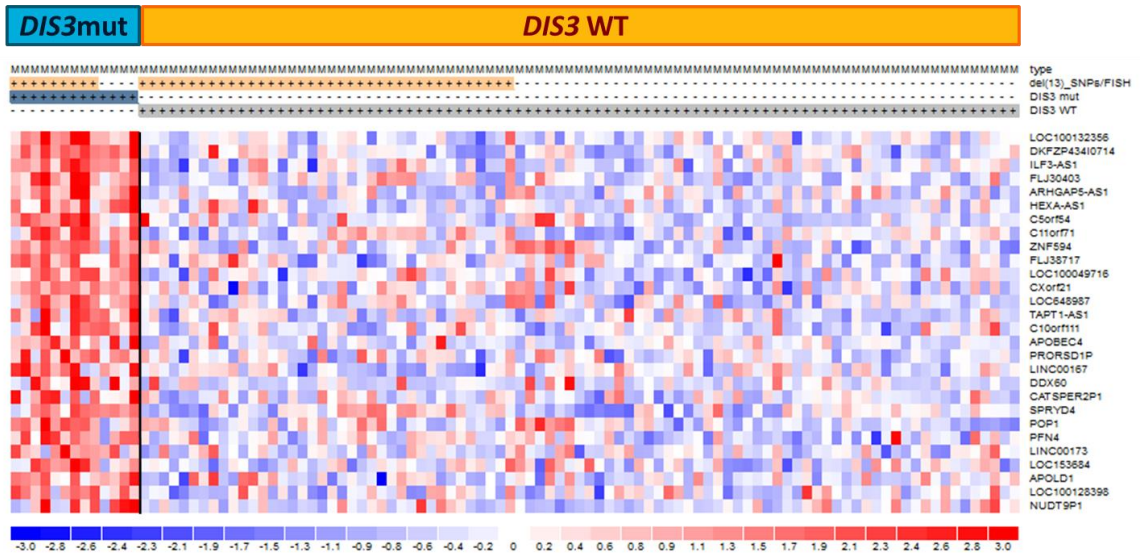

Figure 17. Heatmap of the 28 differentially expressed genes identified at q-value=0 by SAM two-class analysis of 102 MM patients stratified based on the presence of DIS3 mutations. In the heatmap, the color-scale bar represents the relative gene expression changes normalized by the standard deviation, and the color changes in each row represent gene expression level relative to the mean across the samples.

Notably, all transcripts were up-regulated in mutated cases. Interestingly, 74 were protein-coding genes, while one gene was of unknown type and 44 were non-coding RNAs/pseudogenes.

In particular, this latter class of modulated transcripts included nine antisense RNAs (*ARHGAP5-AS1, C21orf119, DLGAP1-AS1, HEXA-AS1, ILF3-AS1, LOC645212, MATN1-AS1, TAPT1-AS1*, *ZNRD1-*

AS1), seven pseudogenes (*CATSPER2P1, FLJ37201, GNRHR2, GVINP1, HTATSF1P2, NUDT9P1*, *PRORSD1P*), six long intergenic non-protein-coding RNAs (*FLJ30403, TP53TG1, LINC00167, LINC00173, LINC00528*, *LOC100129726*), five small nuclear RNAs (*RNU11, RNU12, RNU4ATAC, RNU5D-1*, *RNU5F-1*), two microRNAs (*MIR23A* and *MIR320A*), one microRNAhost gene (*MIRLET7BHG*), and a group mainly consisting of uncharacterized loci and open reading frames. We then explored whether these transcripts belonged to specific functional subgroups. Remarkably, a significant fraction of them was involved, like *DIS3* itself, in single-stranded RNA binding function (namely, *PAT2L*, *DHX58*, *DDX60* and *KHDRBS2*; P value = 1.01E-04) (**Table 11**). Beyond this specific molecular function, several other up-regulated genes in *DIS3*-mutated patients are implicated in RNA physiology: in particular, *PAN2, POP1, RBM45* and the aforementioned *PATL2* and *KHDRBS2* genes are involved in RNA metabolism; *APOBEC3F* and *APOBEC4* encode RNA editing enzymes; whereas *RNU11*, *RNU12* and *RNU4ATAC* are spliceosome subunits. Notably, *DHX58*, *DDX60* and *APOBEC3F* genes, along with *IFIT1, IFIT3, IFI6*, and *OASL*, are part of interferon-mediated anti-viral response.

Table 11. Selected significantly enriched gene ontology terms in the *DIS3* mutation-associated gene signature (q-value < 0.05, Benjamini Hochberg correction).

Category	Name	q-value	Genes
GO: Molecular Function	single-stranded RNA binding	2.01E-02	*PATL2*, *DHX58*, *DDX60*, *KHDRBS2*
GO: Biological Process	positive regulation of MDA-5 signaling pathway	3.32E-02	*DHX58*, *DDX60*
GO: Biological Process	regulation of MDA-5 signaling pathway	3.32E-02	*DHX58*, *DDX60*
GO: Biological Process	positive regulation of RIG-I signaling pathway	3.32E-02	*DHX58*, *DDX60*
GO: Biological Process	defense response to virus	3.32E-02	*DHX58*, *DDX60*, *IFIT1*, *IFIT3*, *APOBEC3F*, *OASL*
GO: Biological Process	cellular response to type I interferon	3.32E-02	*IFIT1*, *IFIT3*, *OASL*, *IFI6*
GO: Biological Process	type I interferon signaling pathway	3.32E-02	*IFIT1*, *IFIT3*, *OASL*, *IFI6*
GO: Biological Process	MDA-5 signaling pathway	3.32E-02	*DHX58*, *DDX60*
GO: Biological Process	response to type I interferon	3.32E-02	*IFIT1*, *IFIT3*, *OASL*, *IFI6*
GO: Biological Process	regulation of defense response to virus	3.76E-02	*DHX58*, *DDX60*, *IFIT1*, *APOBEC3F*
Pathway (REACTOME)	Interferon alpha/beta signaling	1.01E-04	*IFIT1*, *IFIT3*, *OASL*, *IFI6*

Circulating tumor DNA as a liquid biopsy in plasma cell dyscrasias

A CAPP-seq ultra-deep targeted NGS approach was developed to genotype a large gene panel specifically designed to maximize the mutation recovery in PC tumors, and compared the mutational profiling of cfDNA and tumor gDNA of purified PCs from BM aspirates in a consecutive series of 28 patients representative of different clinical stages of PC tumors, with clinical and molecular characteristics consistent with an unselected cohort of PC dyscrasia patients (**Table 12**). In particular, we collected the three type of biological material from two had MGUS, five sMM, and 21 symptomatic MM patients. The sampling was done at diagnosis in 25 patients, and during the course of the disease in three MM cases (**Table 12**). The targeted resequencing gene panel included coding exons and splice sites of 14 genes: *BRAF, CCND1, CYLD, DIS3, EGR1, FAM46C, IRF4, KRAS, NRAS, PRDM1, SP140, TP53, TRAF3, ZNF462* (target region: 31 kb).

Table 12. Clinical details of the 28 patients analyzed by Illumina Miseq. [Δ]The ISS was adopted; [‡]ND: newly diagnosed, RR: relapsed/refractory; [†]Paraprotein; FLC, free light-chain; *del(13), del(17), 1p loss and 1q gain were determined by FISH. °HD= presence of the hyperdiploid status on the basis of FISH evaluation criteria; na: not applicable.

Sample ID	Sex	Age	Disease	Stage[Δ]	Phase[‡]	PP[†]	% of PCs in BM biopsy	FLC ratio	t(4;14)	t(11;14)	t(14;16)	del(13)*	del(17p)*	1q gain*	1p loss*	HD°
1	F	46	MM	3	ND	λ	50	λ/κ FLC = 753	-	+	-	-	-	-	-	-
2	M	52	MM	3	ND	Gκ	90	κ/λ FLC = 708	-	-	-	+	-	+	-	+
3	M	70	MM	1	ND	Gκ	30	κ/λ FLC = 104	-	-	-	-	+	-	-	+
4	M	53	MM	2	RR	Aκ	25	κ/λ FLC = 3	na	na	na	na	na	+	-	+
5	M	56	MM	3	ND	Gκ	80	κ/λ FLC = 36	-	-	-	-	-	+	+	-
6	M	66	MM	1	ND	Aλ	28	λ/κ FLC = 182	-	-	-	-	-	+	-	+
7	F	46	MM	1	RR	λ	30	λ/κ FLC > 27000	-	-	-	+	-	+	-	+
8	F	52	MM	1	ND	λ	45	λ/κ FLC = 446	-	-	-	+	-	+	-	+
9	M	76	sMM	/	ND	Gκ	55	κ/λ FLC = 4	-	-	-	-	-	+	-	+
10	F	76	MM	1	ND	Aλ	30	λ/κ FLC = 6	-	-	-	na	-	+	-	-
11	M	54	MM	2	ND	Aκ	40	κ/λ FLC = 19	-	-	-	+	-	-	-	+
12	M	77	MM	2	ND	Gκ	30	κ/λ FLC = 17	-	-	-	-	-	-	-	+
13	M	64	MM	3	ND	Aλ	60	λ/κ FLC = 210	-	-	-	na	-	+	-	-
14	F	61	sMM	/	ND	Gκ	55	κ/λ FLC = 29	-	-	-	+	-	-	-	+
15	M	68	MM	3	ND	Aλ	70	λ/κ FLC = 65	+	-	-	+	+	+	-	-
16	F	76	sMM	/	ND	Gκ	18	κ/λ FLC = 44	-	-	-	na	-	-	-	+
17	M	59	MM	3	ND	λ	90	λ/κ FLC = 129	-	+	-	+	+	-	+	+
18	F	68	MM	1	ND	Gλ	40	λ/κ FLC = 338	-	-	-	na	-	+	-	-
19	F	64	MM	2	RR	Gκ	65	κ/λ FLC = 15	-	+	-	-	-	-	-	-
20	F	82	MM	1	ND	Gκ	11	κ/λ FLC = 21	-	+	-	na	-	-	-	-
21	F	59	sMM	/	ND	Gλ	10	λ/κ FLC = 5	-	+	-	na	-	-	-	-
22	M	78	MGUS	/	ND	Mκ	8	κ/λ FLC = 3	-	+	-	-	-	-	-	-
26	F	47	MM	1	ND	Aλ	50	λ/κ FLC = 6	-	-	-	na	-	-	-	-
27	M	71	MM	1	ND	Gκ	28	κ/λ FLC = 109	-	-	-	na	-	+	-	-
28	M	68	sMM	/	ND	Aλ	38	λ/κ FLC = 23	-	-	-	na	-	-	-	+
29	M	69	MM	2	ND	Gκ	70	κ/λ FLC = 200	-	-	-	na	-	+	-	-
30	F	50	MM	1	ND	Aκ	60	κ/λ FLC = 108	-	-	-	na	-	-	-	-
31	F	61	MGUS	/	ND	Gκ	7	κ/λ FLC = 4	-	-	-	-	-	-	-	+

cfDNA was detectable in plasma samples with an average of ~11 000 haploid genome-equivalents per mL of plasma. The amount of cfDNA correlated with clinic-pathological parameters reflecting tumor load/extension, including BM PC infiltration (Pearson correlation coefficient=0.6, *P*=0.0007; **Figure 18A**), and clinical stage. Indeed, patients presenting with ISS stage 3 had significantly higher amounts of cfDNA compared with MGUS/sMM samples and MM cases at ISS stages 1-2 (*P*=0.01; **Figure 18B**, Mann-Whitney test). The application of our targeted ultra-deep NGS approach for plasma cfDNA genotyping resulted in ≥90% of the target region covered >1 000X in all plasma samples, and ≥90% of the target region covered >2 000X in 23/28 (**Figure 19 and Table 13**).

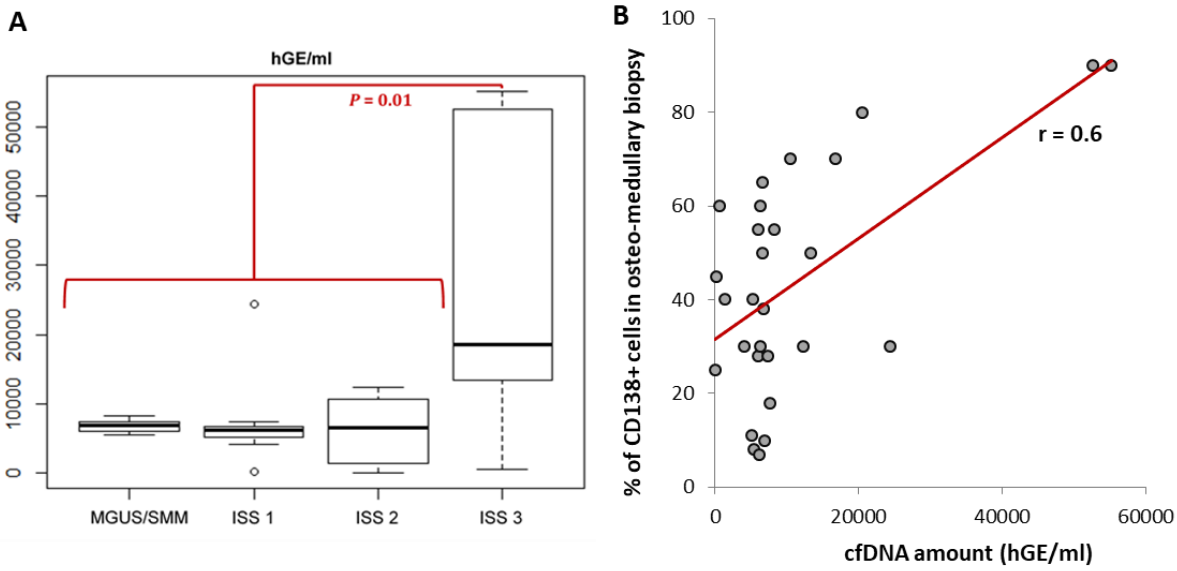

Figure 18. (A) Correlation between cfDNA amount and bone marrow plasma cell infiltration. **(B)** cfDNA amount according to diagnosis/risk stratification: the levels of cfDNA are significantly higher in MM patients at ISS stage 3 compared with MGUS/sMM samples and MM cases at ISS stages 1-2 (*P*=0.01; Mann-Whitney test).

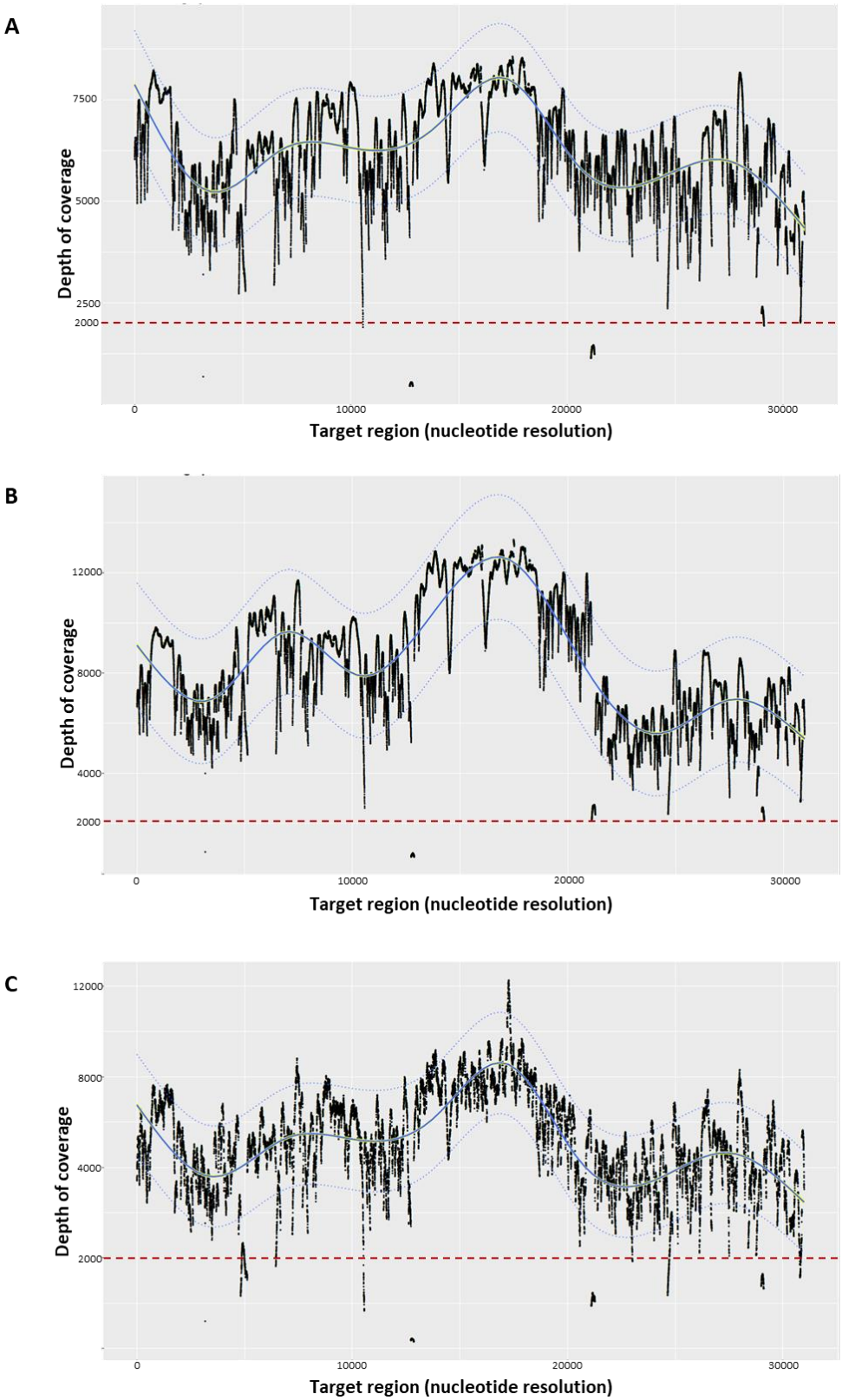

Figure 19. Coverage across the target region. Depth of coverage (y axis) across the target region (x axis) by CAPP-seq of **(A)** gDNA from the germline (granulocytes) samples, **(B)** tumor gDNA from bone marrow plasma cells, and **(C)** plasma cfDNA. Each dot represents the sequencing depth on that specific position of the target region of one single individual sample. The solid blue line shows the median depth of coverage, while the dash blue lines show the interquartile range. The dashed red line shows the 2000X coverage.

Table 13. Percentage of target region covered ≥1000X and ≥2000X in distinct patient samples. Abbreviations: GL, normal germline DNA from granulocytes; PCS, tumor genomic DNA from plasma cells; PL, cfDNA from plasma.

Sample ID	Target Region Coverage (%)		Sample ID	Target Region Coverage (%)	
	≥ 1000X	≥ 2000X		≥ 1000X	≥ 2000X
ID1 GL	100.0	99.5	ID15 GL	99.5	98.7
ID1 PCS	99.6	98.3	ID15 PCS	99.5	98.6
ID1 PL	99.4	97.9	ID15 PL	99.4	98.0
ID2 GL	99.5	99.5	ID16 GL	99.5	98.9
ID2 PCS	99.4	96.9	ID16 PCS	99.5	98.3
ID2 PL	97.8	83.5	ID16 PL	97.3	62.8
ID3 GL	98.7	85.9	ID17 GL	98.6	97.5
ID3 PCS	99.9	99.5	ID17 PCS	99.5	95.2
ID3 PL	98.9	95.8	ID17 PL	98.5	97.0
ID4 GL	98.3	43.0	ID18 GL	98.7	96.3
ID4 PCS	100.0	100.0	ID18 PCS	94.7	84.0
ID4 PL	97.6	58.5	ID18 PL	98.7	97.8
ID5 GL	97.7	63.4	ID19 GL	98.4	92.0
ID5 PCS	100.0	99.6	ID19 PCS	96.4	74.9
ID5 PL	98.2	93.3	ID19 PL	98.7	97.5
ID6 GL	100.0	99.5	ID20 GL	98.6	96.5
ID6 PCS	99.5	99.5	ID20 PCS	97.7	95.2
ID6 PL	99.4	97.4	ID20 PL	98.7	96.8
ID7 GL	99.6	99.5	ID21 GL	98.5	94.9
ID7 PCS	99.6	98.8	ID21 PCS	98.0	94.9
ID7 PL	99.5	97.9	ID21 PL	98.6	96.7
ID8 GL	99.5	98.6	ID22 GL	98.7	97.5
ID8 PCS	99.5	99.5	ID22 PCS	98.7	97.8
ID8 PL	98.6	91.9	ID22 PL	98.7	96.5
ID9 GL	99.5	98.8	ID26 GL	98.9	98.2
ID9 PCS	99.5	98.9	ID26 PCS	99.5	99.5
ID9 PL	99.4	97.8	ID26 PL	99.4	97.1
ID10 GL	99.5	99.0	ID27 GL	99.5	99.0
ID10 PCS	99.3	96.9	ID27 PCS	99.5	98.8
ID10 PL	99.5	98.5	ID27 PL	99.5	98.8
ID11 GL	99.5	98.7	ID28 GL	99.5	99.0
ID11 PCS	99.5	99.3	ID28 PCS	99.5	99.5
ID11 PL	99.3	95.5	ID28 PL	98.7	90.3
ID12 GL	99.5	99.0	ID29 GL	98.2	94.5
ID12 PCS	99.5	99.1	ID29 PCS	99.6	99.5
ID12 PL	94.5	21.9	ID29 PL	99.6	99.4
ID13 GL	99.5	98.9	ID30 GL	99.5	98.9
ID13 PCS	99.5	98.7	ID30 PCS	100.0	99.5
ID13 PL	98.7	86.0	ID30 PL	99.5	98.1
ID14 GL	99.5	98.8	ID31 GL	99.6	98.9
ID14 PCS	99.6	98.8	ID31 PCS	99.6	98.9
ID14 PL	99.4	96.5	ID31 PL	99.4	98.3

Overall, within the interrogated genes, 18/28 (64%) patients had at least one non-synonymous somatic mutation that was detectable in cfDNA (**Figure 20A** and **Table 14A**); 28 total variants were identified, with a range of 1-4 mutations per patient. Quite consistent with the typical spectrum of mutated genes in MM, plasma cfDNA genotyping revealed somatic variants of *NRAS* in 25%; *KRAS* in 14%; *TP53*, *TRAF3* and *FAM46C* in 11%, respectively, *CYLD* and *DIS3* in 7%, respectively, and *BRAF* and *IRF4* in 4% of cases, respectively. Variants in *NRAS*, *KRAS* and *BRAF* genes occurred in a mutually exclusive manner. *TP53* mutations were positively associated with the deletion of the remaining allele as revealed by FISH on purified PCs (*P*=0.02, Fisher-exact test).

To validate the tumor origin of mutations discovered in cfDNA and to derive the accuracy of our approach in resolving tumor genetics, the genotype of cfDNA was matched with that of gDNA from purified BM PCs in all the 28 patients. Sequencing of tumor gDNA identified 39 somatic mutations in 20/28 (71.4%) patients (**Figure 20B**). cfDNA genotyping correctly identified 72% of mutations (n=28/39) that were discovered in tumor PCs (**Figure 22A**), and overall the variant allele frequencies in plasma samples correlated with those in tumor biopsies (Pearson correlation coefficient=0.58, *P*=9.6e-05; **Figure 22B**). Notably, the remaining BM PCs' confirmed mutations not discovered in cfDNA (n=11) had a low representation in the tumor (median allelic frequency: 2.5%; range: 1.1-4.96%) (**Table 14B** and **Figure 20C**). Consistently, based on ROC analysis, cfDNA genotyping has the highest sensitivity in detecting tumor PCs' confirmed mutations when they are represented in at least 5% of the tumor alleles (**Figure 22C**). Above this threshold, the sensitivity of cfDNA genotyping reached 100%. Noteworthy, cfDNA genotyping was still able to detect almost half (10/21) of mutations of low-abundance in tumor PC (i.e. allelic frequency <20%), indicating a good capacity of cfDNA to mirror also the subclonal composition of the tumor. In none of the cases cfDNA genotyping identified additional somatic mutations not detected in the purified BM PCs.

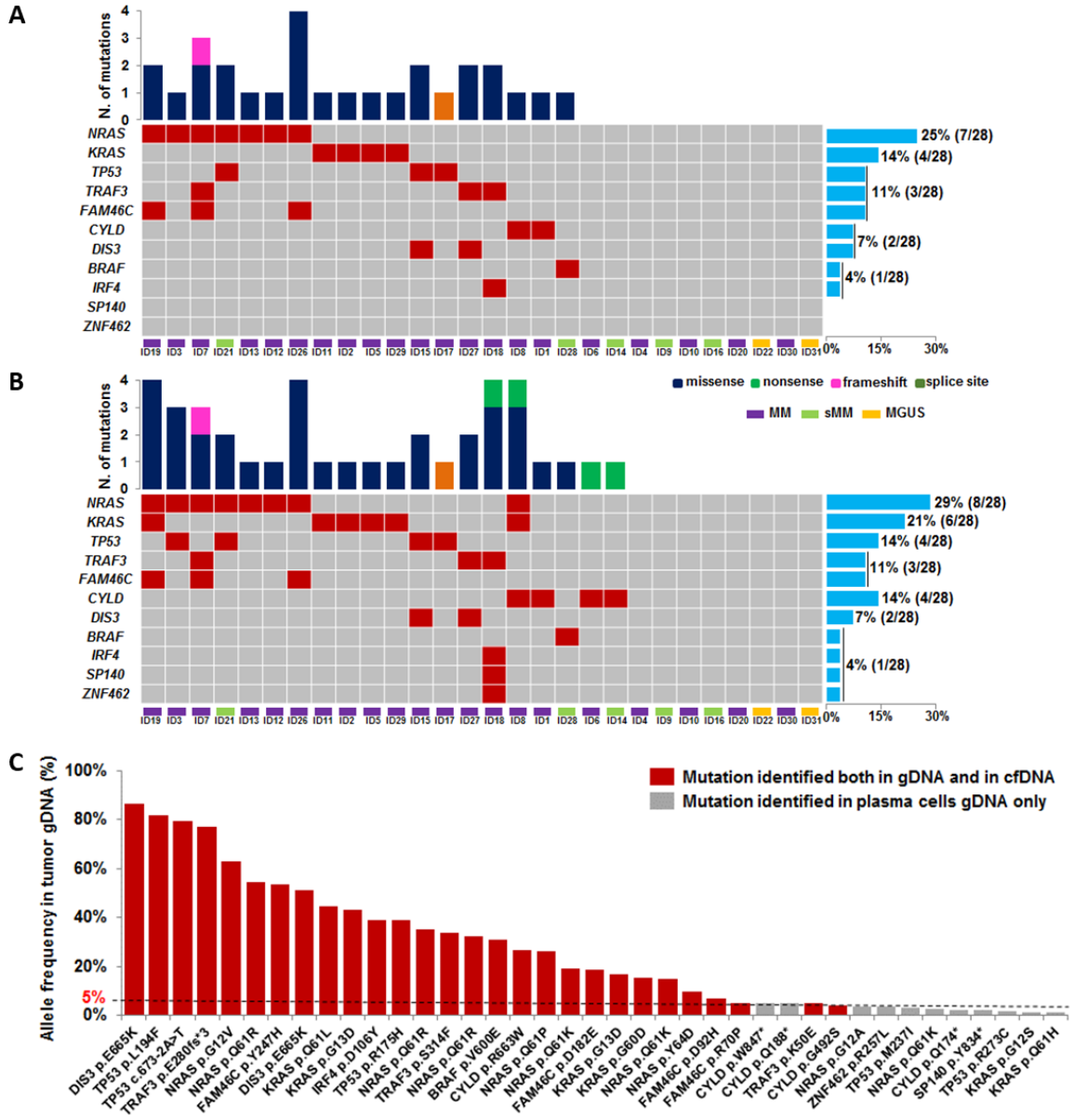

Figure 20. Overview of the mutations identified in the PC dyscrasia series. **(A)** Mutations detected in plasma cfDNA; **(B)** mutations detected in tumor gDNA. Each column represents one tumor sample and each row represents one gene. The fraction of tumors with mutations in each gene is plotted (right). The number and the type of mutations in a given tumor are plotted above the heat map. **(C)** Bar graph of the allele frequencies in tumor gDNA of the variants that were discovered in plasma cfDNA (red bars) or missed in plasma cfDNA (gray bars). The dashed line tracks the 5% allelic frequency threshold.

Table 14. Somatic non-synonymous mutations discovered by cfDNA genotyping and their validation in tumor gDNA **(A)** and those discovered in tumor gDNA genotyping and missed in plasma cfDNA **(B)**. Abbreviations: CHR, chromosome; REF, reference allele; VAR, variant allele.*Absolute chromosome coordinates of each variant based on the hg19 version of the human genome assembly.

A

ID Sample	Gene	RefSeq mRNA	CHR	Absolute position*	REF	VAR	cDNA change	Protein change	cfDNA allele fraction	gDNA allele fraction
ID1	*CYLD*	NM_015247.2	chr16	50820803	A	T	c.1987A>T	p.R663W	0.95%	26.75%
ID2	*KRAS*	NM_033360.3	chr12	25380276	T	A	c.182A>T	p.Q61L	25.01%	44.72%
ID3	*NRAS*	NM_002524.4	chr1	115258747	C	A	c.35G>T	p.G12V	3.08%	63.07%
ID5	*KRAS*	NM_033360.3	chr12	25380279	C	T	c.179G>A	p.G60D	1.05%	15.42%
ID7	*FAM46C*	NM_017709.3	chr1	118166229	T	C	c.739T>C	p.Y247H	3.82%	53.38%
ID7	*NRAS*	NM_002524.4	chr1	115256529	T	C	c.182A>G	p.Q61R	6.72%	54.57%
ID7	*TRAF3*	NM_003300.3	chr14	103363617	A	-	c.839_839delA	p.E280fs*3	9.66%	76.97%
ID8	*CYLD*	NM_015247.2	chr16	50813911	G	A	c.1474G>A	p.G492S	0.87%	3.93%
ID11	*KRAS*	NM_033360.3	chr12	25398281	C	T	c.38G>A	p.G13D	4.39%	16.82%
ID12	*NRAS*	NM_002524.4	chr1	115256529	T	C	c.182A>G	p.Q61R	3.33%	35.14%
ID13	*NRAS*	NM_002524.4	chr1	115256530	G	T	c.181C>A	p.Q61K	32.52%	19.11%
ID15	*DIS3*	NM_014953.3	chr13	73337723	C	T	c.1993G>A	p.E665K	37.86%	86.29%
ID15	*TP53*	NM_000546.5	chr17	7578269	G	A	c.580C>T	P.L194F	36.29%	81.79%
ID17	*TP53*	NM_000546.5	chr17	7577610	T	A	c.673-2A>T	p.224?	8.84%	79.53%
ID18	*IRF4*	NM_002460.3	chr6	394920	G	T	c.316G>T	p.D106Y	1.48%	39.08%
ID18	*TRAF3*	NM_003300.3	chr14	103336686	A	G	c.148A>G	p.K50E	0.29%	4.86%
ID19	*FAM46C*	NM_017709.3	chr1	118165764	G	C	c.274G>C	p.D92H	0.68%	6.98%
ID19	*NRAS*	NM_002524.4	chr1	115256521	A	C	c.190T>G	p.Y64D	0.65%	9.97%
ID21	*NRAS*	NM_002524.4	chr1	115256529	T	G	c.182A>C	p.Q61P	0.54%	26.06%
ID21	*TP53*	NM_000546.5	chr17	7578406	C	T	c.524G>A	p.R175H	0.73%	38.91%
ID26	*FAM46C*	NM_017709.3	chr1	118165699	G	C	c.209G>C	p.R70P	1.22%	5.16%
ID26	*FAM46C*	NM_017709.3	chr1	118166036	C	G	c.546C>G	p.D182E	5.35%	18.83%
ID26	*NRAS*	NM_002524.4	chr1	115256529	T	C	c.182A>G	p.Q61R	16.08%	32.59%
ID26	*NRAS*	NM_002524.4	chr1	115256530	G	T	c.181C>A	p.Q61K	11.55%	15.04%
ID27	*DIS3*	NM_014953.3	chr13	73337723	C	T	c.1993G>A	p.E665K	0.64%	51.36%
ID27	*TRAF3*	NM_003300.3	chr14	103363719	C	T	c.941C>T	p.S314F	0.42%	33.81%
ID28	*BRAF*	NM_004333.4	chr7	140453136	A	T	c.1799T>A	p.V600E	1.43%	32.88%
ID29	*KRAS*	NM_033360.3	chr12	25398281	C	T	c.38G>A	p.G13D	11.36%	43.40%

B

ID Sample	Gene	RefSeq mRNA	CHR	Absolute position*	REF	VAR	cDNA change	Protein change	cfDNA allele fraction	gDNA allele fraction
ID3	*TP53*	NM_000546.5	chr17	7577570	C	T	c.711G>A	p.M237I	-	3.31%
ID3	*TP53*	NM_000546.5	chr17	7577121	G	A	c.817C>T	p.R273C	-	1.83%
ID6	*CYLD*	NM_015247.2	chr16	50785530	C	T	c.520C>T	p.174Q*	-	2.44%
ID8	*CYLD*	NM_015247.2	chr16	50785572	C	T	c.562C>T	p.188Q*	-	4.88%
ID8	*KRAS*	NM_033360.3	chr12	25380275	T	A	c.183A>T	p.Q61H	-	1.14%
ID8	*NRAS*	NM_002524.4	chr1	115256530	G	T	c.181C>A	p.Q61K	-	2.55%
ID14	*CYLD*	NM_015247.2	chr16	50828193	G	A	c.2540G>A	p.W847*	-	4.96%
ID18	*SP140*	NM_007237.4	chr2	231176307	C	A	c.2502C>A	p.Y834*	-	2.43%
ID18	*ZNF462*	NM_021224.4	chr9	109686963	G	T	c.770G>T	p.R257L	-	3.50%
ID19	*KRAS*	NM_033360.3	chr12	25398285	C	T	c.34G>A	p.G12S	-	1.46%
ID19	*NRAS*	NM_002524.4	chr1	115258747	C	G	c.35G>C	p.G12A	-	3.58%

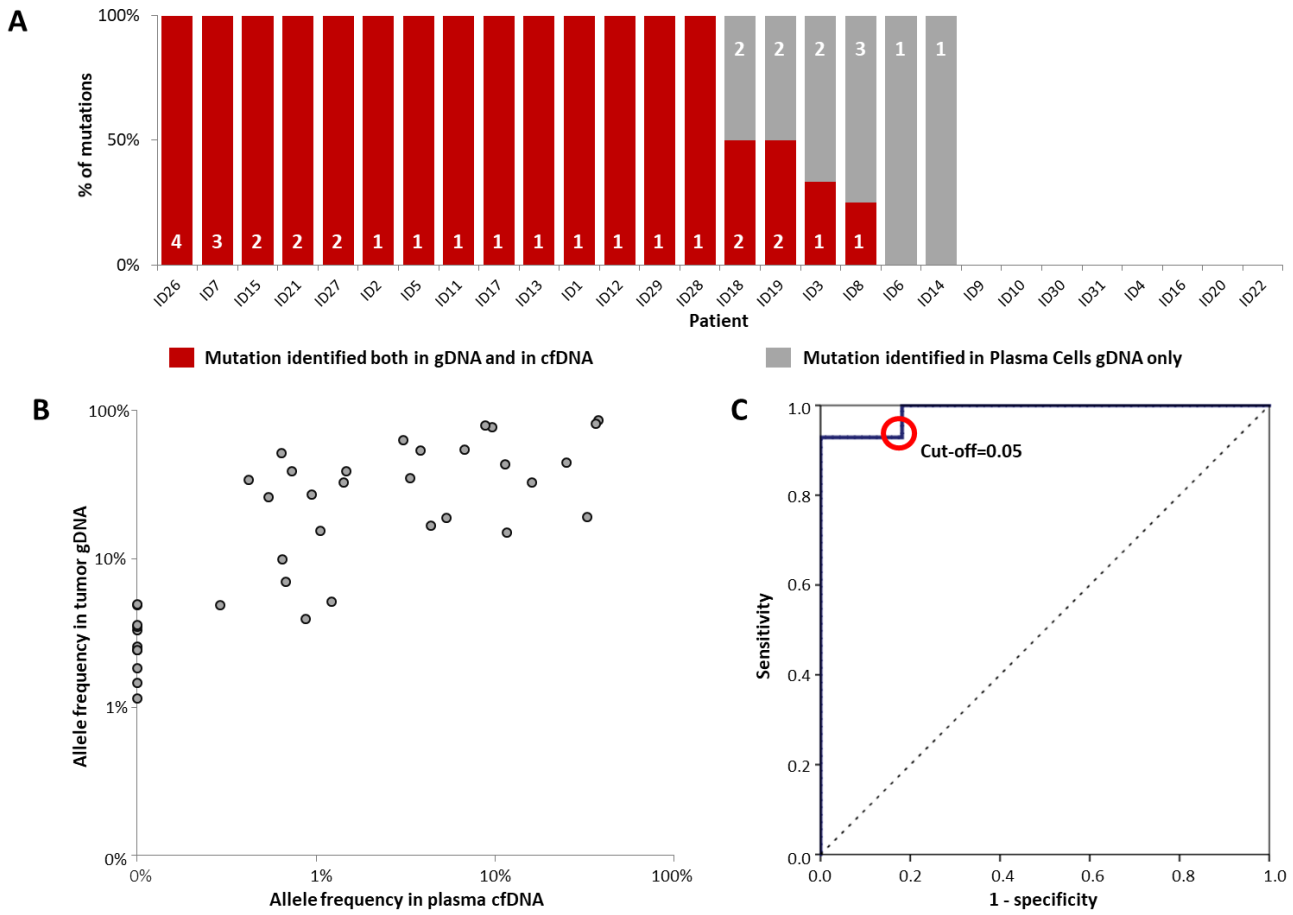

Figure 22. Concordance between plasma cfDNA and tumor gDNA genotyping. **(A)** The fraction of tumor biopsy–confirmed mutations that were detected in plasma is shown. Patients are ordered by decreasing detection rates. The red portion of the bars indicates the prevalence of tumor biopsy–confirmed mutations that were detected in plasma cfDNA. The gray portion of the bars indicates the prevalence of tumor biopsy-confirmed mutations that were not detected in plasma cfDNA. **(B)** The mutation abundance in plasma cfDNA vs the mutation abundance in tumor gDNA is comparatively represented in the scatter plot for each identified variant. **(C)** ROC analysis illustrating the performance of gDNA genotyping in discriminating the ability of cfDNA genotyping to detect biopsy-confirmed mutations according to the variant allele frequency of mutations in tumor gDNA.

Discussion and Conclusions

MM is an incurable hematological malignancy, accounting for 10% of all hematological tumors. It is characterized by a wide clinical spectrum, and by a highly heterogeneous genetic background featuring a deep instability, at both karyotypic and mutational level. As regards in particular the mutational landscape, the emerging scenario from recent NGS studies in MM patients indicated the lack of a universal driver mutation but several recurrently mutated genes belonging to key pathways involved in the disease, such as the MAP kinase pathway. What is most impacting are the clinical implications of NGS-derived data. In fact, a number of mutations have been identified that are targetable with commercially available agents, associated with poor or favorable prognosis, or responsible for drug resistance. Therefore, somatic mutations may represent novel markers to improve risk prediction and guide towards the choice of the most appropriate therapeutic strategy, taking full advantage of the available treatment options, in the perspective of applying precision medicine as ultimate goal.

MM genotyping has so far relied on the analysis of purified PCs from the BM aspirate from a single anatomical site, which may fail in capturing the entire mutational repertoire of PCs as they can localize into multiple and genetically heterogeneous foci. Furthermore, it is associated with technical hurdles and patient discomfort, which are limiting its transfer in the routine diagnostic workup of patients. For these reasons, alternative and complementary approaches that could overcome both the abovementioned limitations are required. In this context, the analysis of cfDNA, that is shed into the PB by tumor cells and contains a representation of the entire tumor genome, is one promising possibility, implying obvious advantages in terms of accessibility and ability of capturing the whole tumor heterogeneity.

The plan of this project was to test the hypothesis that cfDNA may represents an accessible source of tumor DNA for the sensitive identification of genetic biomarkers in PC dyscrasia by pursuing the following specific aims:

- evaluation of the incidence of mutations in some driver genes (*KRAS*, *NRAS*, *TP53*, *BRAF*, *FAM46C* and *DIS3*), as suggested by several recent massive DNA sequencing studies [63 65 64, 71], in a large and representative cohort of patients at different stages of PC dyscrasia;
- comparison of the mutational profiling of cfDNA and BM testing in a small series of patients representative of different clinical stages of PC tumors to assess the proof of principle that the MM mutations are detectable in plasma, so that this approachcan be used as a powerful non-invasive biomarker.

Regarding the first point, in the study population of 167 cases with different PC tumors, *BRAF* mutations were identified in 12%. The mutation frequency in our series of patients with MM at onset (10.6%) was more consistent with that found in WGS/WES studies[64, 65, 71] (on average in 8.7%

of cases) (**Figure 23**) compared with other studies[73, 135]. Our NGS analysis identified mutations with a low VAF and a number of variants other than the classic V600E (found in seven cases), which is carried by less than half of the *BRAF*-mutated patients, unlike what observed in other tumor types, i.e. melanoma, colorectal cancer, and papillary thyroid carcinoma. This substitution is the most frequent mutation, reported to destabilize the hydrophobic interaction between the glycine-rich loop and the activation segment, thus locking the protein in its active conformation and increasing BRAF kinase activity[119, 136] (**Figure 12A**). Published functional data indicate that most *BRAF* mutational events are responsible for the activation of MEK-ERK pathway, which is also caused by alterations in *NRAS* or *KRAS* (whose involvement in disease pathogenesis has already been extensively proven)[113, 128, 135] in an even more substantial percentage of patients.

We detected genetic alterations in ERK signalling in 57% of our patients, thus confirming the crucial role this pathway plays in MM development. In line with the findings of the most recent sequencing studies[64, 65, 71], our analysis revealed concomitant substitutions in more than one of the investigated genes in 13 patients, although their simultaneous occurrence at high VAF was rare. Interestingly, the concomitant occurrence of five of the eight ^{D594}*BRAF* variants with mutations in *KRAS* or *NRAS* may be explained by the experimentally described synergy of kinase-dead *BRAF* mutations involving D594 and oncogenic *RAS.* Their cooperation in inducing tumor progression, indeed, has been demonstrated in a murine model of melanoma[137], thus confirming the hypothesis that the high frequency of this inactivating mutation (also observed in our MM patient cohort) is incompatible with a random event. Furthermore, serial analyses of *BRAF, N-* and *K-RAS* mutations during disease progression revealed a slightly more complex scenario than that found by Bolli *et al.*[64], who described either clonal variants at both timepoints or the presence of acquired/increased variants in the later sample, in line with the expected positive selection of mutated subclones. Besides these patterns, we also observed the disappearance of variants with relatively high VAF values, but always occurring concurrently with the emergence/clonal expansion of an additional mutation in another gene of the pathway.

In line with recent WES studies, these data indicate that, although involving driver genes, MAPK mutations can be clonal (compatible with early events) in some patients or subclonal (compatible with late events) in others[64, 65], and may occur at variable times during tumor evolution compared to the other molecular lesions. The activation of the MEK-ERK pathway as a result of *BRAF* and *RAS* mutations identified by NGS was confirmed by the GEP data of the same patients, indicating that the MAPK cascade is one of the most enriched biological processes involving genes that are differentially expressed between mutated and wild-type patients. Together with the detection of the majority of *BRAF* genomic variants at cDNA level, this highlights the particular importance of

alterations in this pathway in terms of both occurrence and phenotypical impact in the context of the heterogeneous mutational pattern of MM.

DIS3 belongs to the human exosome complex and is endowed with both exo- and endonucleolytic activities. DIS3 regulates the processing and abundance of all RNA species[138]. *DIS3* mutations were globally identified in 20.1% of patients: specifically, mutation prevalence was 18.5% in newly diagnosed MM patients, 25% in pPCL at onset, and 30% in sPCL. Interestingly, *DIS3* mutation frequency seemed to be slightly higher in extramedullary phases (although this difference did not reach statistical significance, likely due to the relatively small size of pPCL and sPCL series). The prevalence of *DIS3* mutations identified here in representative MM cases at onset (18.5%) was higher than the prevalence reported for unselected patients' cohorts in other studies[63, 65, 71, 80] (**Figure 23**). Most likely, this could be due to the higher depth of coverage of our sequencing analysis, especially compared to that usually obtained in WGS and WES experiments. Compared to the lower frequency (11%) reported by Weißbach *et al.*[80], who used our same targeted resequencing approach, this difference could be due to the exclusion in that study of the variants with VAF lower than 8%. We obtained similar results as of Walker *et al.*[79] (18%), who exclusively analyzed *IGH*-translocated patients. Indeed, also in the present series, *DIS3* mutations were confirmed preferentially associated with *IGH* translocations and nonhyperdiploid status.

Of the two functional domains investigated in this study, the RNB domain was slightly more densely affected by mutations. The mutations of the RNB domain initially identified in MM by Chapman *et al.*[63] interfere with DIS3 ribonucleolytic activity[76]. This functional effect could be envisaged also for many of the mutations identified by us, based on the bioinformatics predictions and the pattern of differential expression between *DIS3*-mutated and wild type MM cases emerged from the gene expression profiling of our cohort. *DIS3*-mutated MM samples, in fact, displayed the up-regulation of several transcripts, including genes involved in RNA interactions and many non-coding transcripts. The finding that all the involved transcripts were over-expressed in *DIS3*-mutated patients reflects what reported by several studies analyzing the transcriptional output resulting from the mutation or depletion of exosome subunits in other species, and is consistent with the RNA degradative activities ascribed to the exosome[139-141]. In this regard, it is also worth noting that substitutions at residues R780, D487 and D488 (demonstrated or supposed to be particularly detrimental for protein function) were never observed at genomic VAFs above 50%, neither in the present nor in previously published datasets[65, 71, 79, 80], suggesting that the exclusive presence of such *DIS3* mutant forms in the absence of a *DIS3* wild type protein might have deleterious cellular effects. In general, the lack of a *DIS3* wild type allele in mutated cases was uniquely observed in some 13q14-deleted MM patients carrying other types of variants. The most affected residue R780, replaced by a lysine or by a threonine in three and two cases, respectively, was in line with what

reported in previous MM sequencing studies, where this amino acid position was already indicated as the major *DIS3* mutational hotspot[63, 65, 79, 80] and it was also confirmed by the last Walker's work[71](**Figure 12B**). The arginine at position 780 was found involved in RNA binding function[129]. Both R780K and R780T substitutions are recurrent in MM patients of our and previously published datasets[65, 79, 80], and the former was demonstrated to cause significant aberrations of DIS3 exonucleolytic activity and growth inhibition in yeast and human cellular models[76]. Of note, D488 represents the second most affected residue in *DIS3*-mutated MM patients when combining the results of the previous studies[63-65, 71, 79, 80]. Interestingly, the adjacent conserved aspartic acid at position 487 (D487H mutation found each in one patient) is equally essential for magnesium binding, and when substituted with asparagine (D487N) it was reported to completely abolish DIS3 exonuclease activity in yeast and human cellular models, causing molecular phenotypes comparable to the *DIS3*-R780K mutant[76]. Other recurrently mutated positions in our series were T93, H764, F775, R789 and R820, each targeted in two samples by the same substitution, except for H764 (reported to be involved in RNA binding[129]) showing different amino acid changes. Notably, the majority of these recurrently altered residues represent novel *DIS3* mutations in MM. In fact only the variant F775L has already been reported in another series[65]. Some of them have been already reported as altered in MM (R108, R467, S550)[65] or in other cancer types (R820, COSM3469577). Among the mutated residues never described before, A751 is involved in RNA binding [129]. Furthermore, looking at *DIS3* mutations in longitudinally analyzed patients (i.e sampled at diagnosis and relapse), we observed a similar scenario to that described by Bolli *et al.*[64] for other known driver myeloma genes, i.e. the occurrence of clonal variants at both timepoints or the acquisition/clonal expansion of variants at the later timepoint, consistently with an expected positive selection for the subclones harboring them. Of note, the disappearance at relapse of variants found at diagnosis at low allele frequency was only observed in patients carrying additional stable/clonally expanding mutations. The limited sample size and the different treatment regimens to which the patients were subjected do not allow making assumptions on the possible impact of specific therapies on the chances of survival of *DIS3*-mutated subclones.

The involvement in MM of *DIS3* and *FAM46C*, which are two of the most recurrently mutated genes in the disease (average mutation frequency of 7% and 9.1%, respectively) (**Figure 23**), was totally unexpected prior to the application of massively parallel sequencing.

Although the significance of its alteration in the disease still needs to be further elucidated, *DIS3* is considered a potential tumor suppressor in MM based on several findings: the loss of enzymatic activity caused by the MM-associated *DIS3* mutations that have been functionally characterized[76]; the loss of heterozigosity often involving *DIS3* due to simultaneous gene mutations and chromosome 13 deletion; and the reported enhanced translation of crucial oncogenes following

DIS3 inactivation[81]. Similar to DIS3, FAM46C is an RNA-binding protein whose exact functions remain to be characterized, but potentially regulating gene expression[142]. *FAM46C* too has a putative tumor suppressor role in MM, as suggested by the deletion of its genomic locus at 1p12, occurring in approximately 20% of MM patients[83], and the predominantly inactivating nature of MM-associated gene mutations. Overall, the significantly recurrent alteration of *DIS3* and *FAM46C* suggests a role of translational control processes in the pathogenesis of MM[63]. The prevalence of *FAM46C* mutations (11.7%) identified in our newly diagnosed MM cohort was similar to that described in previous studies[63-65], but higher than the one more recently reported by Walker *et al.*[71]. Interestingly, except for a few codons (D90, L93, L134,K141, E166, G241 and Y291) already reported as mutated in independent series[65, 71, 83], the majority of altered residues identified in the present study represented novel *FAM46C* mutations in MM. Combining our data with those available in the literature for MM[63-65, 71, 79, 83], it would appear that the *FAM46C* coding sequence is almost entirely altered by mutations, with the exception of small portions, especially in the N-terminal region. Overall, the finding that *FAM46C* mutated alleles had detectable biological expression, and in some cases seemed to be even preferentially transcribed, suggests that the mutations identified herein may have functional implications. In conclusion, our data strengthen the growing evidence that *FAM46C* is involved in the pathogenesis of MM as a potential tumor suppressor, although its role remains to be clarified.

In the present study, the use of deep NGS in a large representative cohort of MM at diagnosis and more aggressive disease stages substantially confirmed the accepted notion[85-88] that *TP53* mutations in MM at presentation are rarely detected (3.1% in MM), while their frequency increases with disease progression (25% in pPCL and 20% in sPCL). This would be consistent with the hypothesis that defective TP53 may cause extramedullary disease[96]. The enhanced sensitivity of our analysis, in fact, was not accompanied by a substantial increase in the mutation rate observed in MM cases, consistent with the fact that we did not find many mutations at particularly low allele frequency, which could be undetectable by the traditionally used Sanger method. Although we cannot exclude that an even greater depth of coverage would eventually bring out further smaller mutated subclones, this finding is in line with what reported by other recent NGS studies[64, 65, 143]. In extra-medullary MM, *TP53* inactivation was significantly more frequent, as a result of both gene mutations and deletions, confirming that these lesions represent late-progression events. Surely, alterations of *TP53* identify a high-risk population of patients, even in the context of molecularly comprehensive risk scoring models[71]. Although the close association between mutations and deletions makes it difficult to understand the prognostic contribution of individual lesions[64, 92, 144], the recently demonstrated haploinsufficiency of the gene[98] may be responsible of TP53 pathway deficiency whenever the cells do not have two intact copies of the gene (as a result of both deletion

and mutations). Furthermore, taking into consideration the dominant negative effect or gain of function of many of *TP53* mutations, TP53 functionality appears to be even more compromised. In general, the occurrence of mutations even in cases not carrying the deletion supports the importance of their assessment in MM patients, given that many anti-MM drugs induce apoptosis through multiple pathways that are at least in part dependent upon functional TP53 activation. Consistently, we noticed the emergence of a mutation in three of 19 MM/PCL patients analyzed at relapse; despite the NGS analysis of the earlier samples, we could not detect at onset minor mutant subclones potentially responsible for relapse. Concerning deletion of chromosome arm 17p, its frequency in intra-medullary MM patients of our series appears to be lower than that generally reported, a finding that could be related to different sample size and patient selection criteria. A strong, albeit not absolute, association of gene mutation with the deletion of the remaining allele was observed also in our series, in line with previously reported data[144]. Concerning the type of variants, unlike Chng *et al*.[144] and more similarly to what reported by Lodé and colleagues[92] and other more recent papers[64, 65, 71], most of the mutations identified here were single nucleotide missense mutations. Differently from many tumor suppressor genes, in human cancers *TP53* is actually most commonly altered by missense mutations that not only lead to a loss/diminution of the wild type TP53 activity, but, since TP53 normally acts as a tetramer, may also function as dominant negative inhibitors over any remaining wild type TP53 activity, or even give rise to a more aggressive tumor profile through gain of function activity. Gain of function roles in different cellular processes have indeed been demonstrated for several *TP53* mutants also detected in MM/PCL of the present and other published datasets[64, 65, 71, 79] (*P151S, R248Q, R248W, R273H, R273C, V274F, P278S*, and *R280K*)[145]. MM-associated *TP53* mutations are concentrated in the DNA binding domain, and the frequency at which amino acid residues are targeted only in part reflects the one globally observed in human cancers. Indeed, contrary to what is globally observed in MM, somatic mutations affected codon 195 are not among the most recurrent in human tumors. Notably, the substitution I195T, like other tumorigenic mutations in the β-sandwich affecting hydrophobic residues in the protein core, was reported to be severely destabilizing[146], determining a half-life of the unfolding of the mutant protein of less than one minute at 37°C[147]. Conversely, R273 is one of the six residues identified as hotspots for tumorigenic mutations in human cancers; it is generally targeted by two predominant substitutions, i.e. R273H and R273C, both of which identified in MM cases and preventing DNA binding by the isolated core domain[146]. With regard to the amino acid residue 278, it is one of the essential amino acids for major groove contacts in the pentamer sequence of the consensus DNA binding site[148]. (**Figure 12D**) The substitution E285K is the most common *TP53* temperature-sensitive mutation(its wild type TP53 activity is reconstituted at about 32°C) and is also carried by the RPMI-8226 MM cell line[91, 149, 150]. Amino acid substitutions R→W and

R→Q at the hot spot codon 248 both inactivate the transactivating functions of wild type *TP53*, although it has been demonstrated that they differently contribute to *in vitro* malignant behavior of tumor cells through gain-of-function properties[151].

TP53 is mutated at a significant recurrence rate in all three studies[64, 65, 71](on average 8.3%), although the mutational frequencies observed in each series are significantly different, and higher in series where advanced disease was over-represented[64, 65](**Figure 23**). In fact, as highlighted by an NGS-based analysis in a representative cohort including patients with newly diagnosed MM and more advanced stages of PC dyscrasia, *TP53* mutations are generally rare in MM at onset, and conversely constitute a marker of progression, analogously to 17p-deletion.

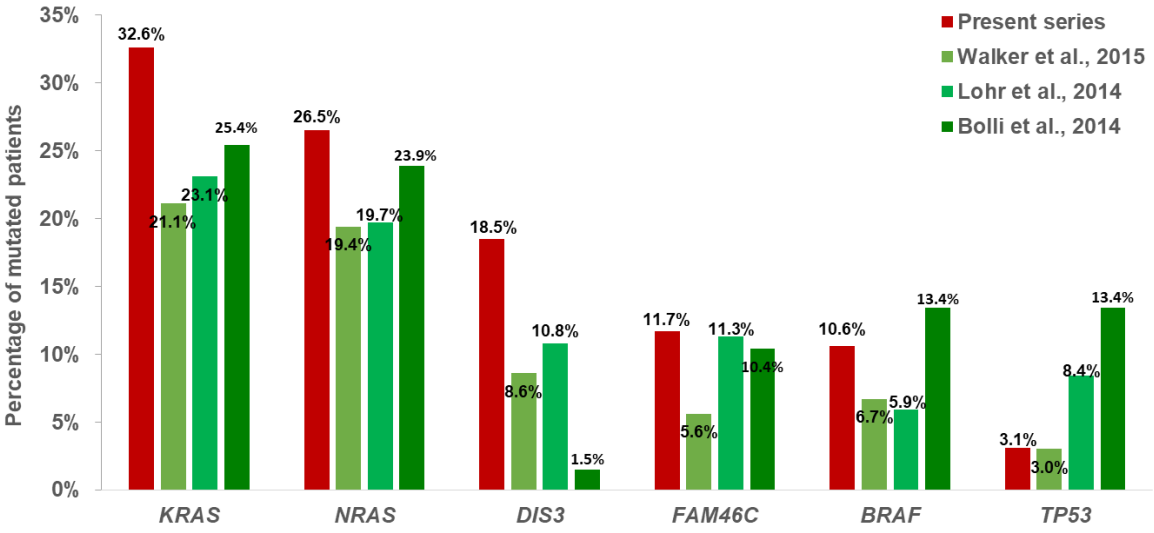

Figure 23: Most significantly mutated genes in MM. Mutation frequencies observed in our MM patients is reported in Bordeaux in comparison with frequencies observed in each of the three main WGS/WES studies in MM[64, 65, 71].

To gain further information about the pattern of evolution underlying relapse in MM, a recent longitudinal study investigated by WES, gene expression profiling and high-resolution copy number arrays 33 patients enrolled in Total Therapy protocols with the aim to define specific drivers of relapse in the precise context of a standard treatment[152]. Notably, the acquisition/expansion of mutations in known MM driver genes (*NRAS, KRAS, BRAF*) and the biallelic inactivation of tumor suppressor genes, especially *TP53*, proved to be critical events driving relapse through the escape form treatment-induced apoptosis and increased proliferation, leading to drug resistance and tumor progression. In particular, biallelic events resulting in complete inactivation of TP53 defined the group with the worst outcome after relapse.

Considering on the whole the three analyzed genes, *BRAF*, *KRAS* and *NRAS*, the MEK/ERK signalling pathway was affected by mutational events in more than half of the cases (96/167, 57.5%), being

more frequent in sPCL (63.6%) and MM (59.8%), and relatively less frequent in pPCL (41.7%). This latter finding is consistent with what newly emerged from a recent WES study in a smaller fraction of the pPCL patients of the present series[89]. This studyinvestigating the mutational profiles of a fraction of the pPCL patients of the present series (12 pPCL patients enrolled in a multicenter prospective GIMEMA Phase II trial www.clinicaltrials.gov as #RV-PCL-PI-350, www.clinicaltrialsregister.eu as #EudraCT N° 2008-003246-28) found few genes recurrently mutated in two or more samples, confirming a substantial heterogeneity of mutational patterns, as observed in intramedullary disease. Mutations in *KRAS* and *NRAS* were three-fold less frequent, while those in *DIS3* and *TP53* were till to two-fold more recurrent in pPCL than in MM, and this was perhaps at least partly related to different representativeness of cytogenetic abnormalities in pPCL and MM series. This work identified *DIS3* as one of the 14 statistically significant recurrently affected genes with potential driver role in the disease[89] and indicated that *TP53* impairment seemed to lie within a more generalized alteration of cell cycle checkpoints, which involved other genes among which the most frequently mutated were *ATM*, *ATR*, *CDKN1A* and *BRCA1*[89].

As far as our study is concerned, we observed that genomic and transcriptional VAFs of *BRAF* were highly concordant in mutated patients. A recent RNA-seq analysis of a subset of patients who had previously undergone WES, indeed, has shown that most of the identified mutations occur in genes whose expression is very low or undetectable[153]; in addition, the mutation frequencies at genomic and transcriptional levels were found not comparable for many genes[153]. This new finding further strengthens the rationale for *BRAF*-targeted therapeutic strategies.

Activating alterations of the *BRAF* serine/threonine kinase gene have also been recently described in WGS/WES studies of MM further extending the evidence of a widespread dysregulation of MAPK signalling in the disease[63-65, 79]. Over the last few years, *BRAF* has received considerable attention as a result of the success of targeted malignant melanoma therapy[154], and encouraging preliminary results have emerged from exploratory studies of the use of BRAF inhibitors in patients with *BRAF*-mutated MM[73, 75]. Moreover, it is to be hoped that more precise indications concerning its efficacy will emerge from the ongoing phase II Basket study (NCT01524978) of vemurafenib in patients with *BRAF* V600E mutation-positive cancers, which also include MM patients[155]. However, the clinical relevance of *BRAF* and *NRAS/KRAS* mutations is still unclear. Bolli *et al.*[64] found no significant survival difference between cases with and without *KRAS*, *NRAS* or *BRAF* mutations, whereas Andrulis *et al.*[73] found that overall survival was significantly shorter in their MM patients carrying *BRAF* V600E. Nevertheless, the paradoxically tumor-enhancing effects of BRAF inhibitors in the case of sub-clonal *BRAF* or co-existent *BRAF* and *RAS* mutations indicates the need for the accurate molecular characterization of patients in order to obtain the most from targeted therapeutic strategies.

The gene expression patterns observed in our study suggested that identified mutations here are actually expressed, as directly assessed by cDNA sequencing of *DIS3*-mutated patients showing comparable mutation burdens at gene and transcript levels. Overall, these data demonstrate the ultimate significance of the *DIS3* gene mutations, differently from what has been suggested for the majority of the genetic changes detected by genome-wide approaches in MM, which are found in genes that have low or no detectable biological expression[153].

The clinical implications of *DIS3* mutations in MM remain to be elucidated. Recently, Weißbach *et al.*[80] reported that the occurrence of *DIS3* mutations in minor subclones was significantly associated with a weaker chemotherapy response as compared to the *DIS3* mutations in major subclones, while globally there was only a slight trend towards shorter median overall survival for *DIS3*-mutated MM patients as compared to *DIS3* wild type ones. In our series, we had the possibility to evaluate the prognostic impact of *DIS3* mutations in the group of pPCLs included in a prospective multicenter clinical trial[114] for whom a 2.8-month follow-up was available, observing no significant association with survival or response rate (data not shown). Furthermore, although less prevalent in MM, mutations in *ATM* or *ATR* genes were lately reported to be associated with a negative impact on survival of MM patients enrolled onto the NCRI Myeloma XI Trial, in which the inability to deliver an effective apoptotic response to DNA damage represented the most significantly prognostic mutational marker[71].

As regards the validation of the liquid bopsy as a diagnostic tool, the application of targeted ultra-deep NGS approach in the prospectively collected cohort of 28 unselected PC dyscrasia patients detected circulating tumor cfDNA with a median of 11000 haploid genome-equivalents per mL and more than 90% of the target region was covered over 1000X in all plasma samples. Overall, within the interrogated genes, 64% of patients harbored somatic mutations that were detectable in plasma cfDNA. Overall, the molecular spectrum of mutations discovered in cfDNA (*NRAS* in 25%; *KRAS* in 14%; *TP53*, *TRAF3* and *FAM46C* in 11%, respectively; *CYLD* and *DIS3* in 7%, respectively; and *BRAF* and *IRF4* in 4% of cases, respectively) reflected previus observations in genomic studies based on PC genotypinf, thus supporting the tumor origin of the mutations identified in cfDNA .

Our results provide the proof of principle that cfDNA genotyping is a feasible, non-invasive, real-time approach that reliably detects clonal and subclonal somatic mutations represented in at least 5% of alleles in tumor PCs. This limit is explained by the fact that the ctDNA is diluted in cfDNA from normal cells[102, 156], and variants that are already rare in tumor gDNA are much less represented in plasma and may fall below the sensitivity threshold of the CAPP-seq.

Despite the genetic heterogeneity characterizing MM, and the inclusion in the study cohort of seven patients at pre-malignant/asymptomatic disease stages, the designed gene-panel employed

in our study proved to be very effective, in that it allowed the recovery of at least one mutation in tumor gDNA of 20/28 (71%) cases. In none of the cases cfDNA genotyping identified additional somatic mutations not detected in the purified BM PCs, thus suggesting that, as far as our limited patient cohort is concerned, the genotype of PC collected from a single tumor site is already representative of the entire tumor genetics. Alternatively, spatial genomic heterogeneity, supported by very recent findings in MM[157], may exist but involving minor subclones not sufficiently represented to be detectable in plasma.

One of the original findings of the study is that cfDNA genotyping can resolve tumor genetics also in cases at early disease stages as MGUS and sMM patients, who may benefit the most from this minimally-invasive approach. Indeed, among asymptomatic patients cfDNA genotyping could allow a non-invasive longitudinal molecular monitoring of clonal evolution of the neoplastic clone and the identification of the switch point on which the disease acquires high-risk genetic features. This has been prevented so far by the unfeasibility of serial BM sampling in the clinical routine. In my opinion, this study enhances the growing evidence supporting the use of the liquid biopsy for a variety of clinical and investigational applications not previously possible. An immediate clinical application of cfDNA genotyping in MM could be the incorporation of this minimally invasive method in clinical trials for the identification of patients carrying actionable mutations and their longitudinal genetic monitoring during targeted therapy administration or for the estimation of minimal residual disease.

In conclusion, the results of our study confirm and extend the mutational landscape of PC tumors, sustaining also the clonal heterogeneity of these diseases. These characteristics may obviously be a hindrance in the perspective of therapies directed against specific single targets, expected to act only on the subclone harbouring the particular genetic lesion. To overcome this obstacle, it would be as minimum necessary to combine different kind of treatments, each targeting different subclones. From this point of view, the application of the of NGS-based technologies in the diagnosis an follow-up of PC dyscrasias may help to clarify the mutational landscape of the diseases,.allowing to select specific treatment that may be beneficial to the patient, or, if not, to be avoided. In this scenario, the introduction of circulating tumor DNA genotyping would represent a feasible, non-invasive, real-time approach to the disease monitoring.

Bibliography

1. Corre J, Munshi N, Avet-Loiseau H. Genetics of multiple myeloma: another heterogeneity level? *Blood* 2015 Mar 19; **125**(12)**:** 1870-1876.

2. Landgren O, Kyle RA, Pfeiffer RM, Katzmann JA, Caporaso NE, Hayes RB*, et al.* Monoclonal gammopathy of undetermined significance (MGUS) consistently precedes multiple myeloma: a prospective study. *Blood* 2009 May 28; **113**(22)**:** 5412-5417.

3. Morgan GJ, Walker BA, Davies FE. The genetic architecture of multiple myeloma. *Nature reviews Cancer* 2012 May; **12**(5)**:** 335-348.

4. Weiss BM, Abadie J, Verma P, Howard RS, Kuehl WM. A monoclonal gammopathy precedes multiple myeloma in most patients. *Blood* 2009 May 28; **113**(22)**:** 5418-5422.

5. Basso K, Dalla-Favera R. Germinal centres and B cell lymphomagenesis. *Nature reviews Immunology* 2015 Mar; **15**(3)**:** 172-184.

6. Kuehl WM, Bergsagel PL. Early genetic events provide the basis for a clinical classification of multiple myeloma. *Hematology / the Education Program of the American Society of Hematology American Society of Hematology Education Program* 2005**:** 346-352.

7. Jacob J, Kelsoe G, Rajewsky K, Weiss U. Intraclonal generation of antibody mutants in germinal centres. *Nature* 1991 Dec 5; **354**(6352)**:** 389-392.

8. Berek C, Berger A, Apel M. Maturation of the immune response in germinal centers. *Cell* 1991 Dec 20; **67**(6)**:** 1121-1129.

9. Cowan G, Weston-Bell NJ, Bryant D, Seckinger A, Hose D, Zojer N*, et al.* Massive parallel IGHV gene sequencing reveals a germinal center pathway in origins of human multiple myeloma. *Oncotarget* 2015 May 30; **6**(15)**:** 13229-13240.

10. Rajkumar SV. Multiple myeloma: 2011 update on diagnosis, risk-stratification, and management. *American journal of hematology* 2011 Jan; **86**(1)**:** 57-65.

11. Kyle RA, Therneau TM, Rajkumar SV, Larson DR, Plevak MF, Offord JR*, et al.* Prevalence of monoclonal gammopathy of undetermined significance. *The New England journal of medicine* 2006 Mar 30; **354**(13)**:** 1362-1369.

12. Dispenzieri A, Katzmann JA, Kyle RA, Larson DR, Melton LJ, 3rd, Colby CL*, et al.* Prevalence and risk of progression of light-chain monoclonal gammopathy of undetermined significance: a retrospective population-based cohort study. *Lancet* 2010 May 15; **375**(9727)**:** 1721-1728.

13. Drexler HG, Matsuo Y. Malignant hematopoietic cell lines: in vitro models for the study of multiple myeloma and plasma cell leukemia. *Leukemia research* 2000 Aug; **24**(8)**:** 681-703.

14. Gabrea A, Martelli ML, Qi Y, Roschke A, Barlogie B, Shaughnessy JD, Jr.*, et al.* Secondary genomic rearrangements involving immunoglobulin or MYC loci show similar prevalences in hyperdiploid and nonhyperdiploid myeloma tumors. *Genes, chromosomes & cancer* 2008 Jul; **47**(7)**:** 573-590.

15. International Myeloma Working G. Criteria for the classification of monoclonal gammopathies, multiple myeloma and related disorders: a report of the International Myeloma Working Group. *BrJHaematol* 2003; **121**(5)**:** 749-757.

16. Kyle RA, Rajkumar SV. Criteria for diagnosis, staging, risk stratification and response assessment of multiple myeloma. *Leukemia* 2009 Jan; **23**(1)**:** 3-9.

17. Albarracin F, Fonseca R. Plasma cell leukemia. *Blood reviews* 2011 May; **25**(3)**:** 107-112.

18. Kyle RA, Maldonado JE, Bayrd ED. Plasma cell leukemia. Report on 17 cases. *Archives of internal medicine* 1974 May; **133**(5)**:** 813-818.

19. Garcia-Sanz R, Orfao A, Gonzalez M, Tabernero MD, Blade J, Moro MJ*, et al.* Primary plasma cell leukemia: clinical, immunophenotypic, DNA ploidy, and cytogenetic characteristics. *Blood* 1999 Feb 1; **93**(3)**:** 1032-1037.

20. Walker BA, Wardell CP, Johnson DC, Kaiser MF, Begum DB, Dahir NB*, et al.* Characterization of IGH locus breakpoints in multiple myeloma indicates a subset of translocations appear to occur in pregerminal center B cells. *Blood* 2013 Apr 25; **121**(17)**:** 3413-3419.

21. Bergsagel PL, Kuehl WM. Molecular pathogenesis and a consequent classification of multiple myeloma. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 2005 Sep 10; **23**(26)**:** 6333-6338.

22. Calasanz MJ, Cigudosa JC, Odero MD, Ferreira C, Ardanaz MT, Fraile A*, et al.* Cytogenetic analysis of 280 patients with multiple myeloma and related disorders: primary breakpoints and clinical correlations. *Genes, chromosomes & cancer* 1997 Feb; **18**(2)**:** 84-93.

23. Dewald GW, Kyle RA, Hicks GA, Greipp PR. The clinical significance of cytogenetic studies in 100 patients with multiple myeloma, plasma cell leukemia, or amyloidosis. *Blood* 1985 Aug; **66**(2)**:** 380-390.

24. Lai JL, Zandecki M, Mary JY, Bernardi F, Izydorczyk V, Flactif M*, et al.* Improved Cytogenetics in Multiple-Myeloma - a Study of 151 Patients Including 117 Patients at Diagnosis. *Blood* 1995 May 1; **85**(9)**:** 2490-2497.

25. Sawyer JR, Waldron JA, Jagannath S, Barlogie B. Cytogenetic findings in 200 patients with multiple myeloma. *Cancer genetics and cytogenetics* 1995 Jul 1; **82**(1)**:** 41-49.

26. Kumar S, Fonseca R, Ketterling RP, Dispenzieri A, Lacy MQ, Gertz MA*, et al.* Trisomies in multiple myeloma: impact on survival in patients with high-risk cytogenetics. *Blood* 2012 Mar 01; **119**(9)**:** 2100-2105.

27. Avet-Loiseau H, Li JY, Facon T, Brigaudeau C, Morineau N, Maloisel F*, et al.* High incidence of translocations t(11;14)(q13;q32) and t(4;14)(p16;q32) in patients with plasma cell malignancies. *Cancer research* 1998 Dec 15; **58**(24)**:** 5640-5645.

28. Bergsagel PL, Chesi M, Nardini E, Brents LA, Kirby SL, Kuehl WM. Promiscuous translocations into immunoglobulin heavy chain switch regions in multiple myeloma.

Proceedings of the National Academy of Sciences of the United States of America 1996 Nov 26; **93**(24)**:** 13931-13936.

29. Fonseca R, Debes-Marun CS, Picken EB, Dewald GW, Bryant SC, Winkler JM*, et al.* The recurrent IgH translocations are highly associated with nonhyperdiploid variant multiple myeloma. *Blood* 2003 Oct 1; **102**(7)**:** 2562-2567.

30. Chesi M, Bergsagel PL, Brents LA, Smith CM, Gerhard DS, Kuehl WM. Dysregulation of cyclin D1 by translocation into an IgH gamma switch region in two multiple myeloma cell lines. *Blood* 1996 Jul 15; **88**(2)**:** 674-681.

31. Chesi M, Bergsagel PL, Shonukan OO, Martelli ML, Brents LA, Chen T*, et al.* Frequent dysregulation of the c-maf proto-oncogene at 16q23 by translocation to an Ig locus in multiple myeloma. *Blood* 1998 Jun 15; **91**(12)**:** 4457-4463.

32. Chesi M, Nardini E, Lim RS, Smith KD, Kuehl WM, Bergsagel PL. The t(4;14) translocation in myeloma dysregulates both FGFR3 and a novel gene, MMSET, resulting in IgH/MMSET hybrid transcripts. *Blood* 1998; **92**(9)**:** 3025-3034.

33. Hebraud B, Magrangeas F, Cleynen A, Lauwers-Cances V, Chretien ML, Hulin C*, et al.* Role of additional chromosomal changes in the prognostic value of t(4;14) and del(17p) in multiple myeloma: the IFM experience. *Blood* 2015 Mar 26; **125**(13)**:** 2095-2100.

34. Keats JJ, Reiman T, Maxwell CA, Taylor BJ, Larratt LM, Mant MJ*, et al.* In multiple myeloma, t(4;14)(p16;q32) is an adverse prognostic factor irrespective of FGFR3 expression. *Blood* 2003 Feb 15; **101**(4)**:** 1520-1529.

35. Santra M, Zhan F, Tian E, Barlogie B, Shaughnessy J, Jr. A subset of multiple myeloma harboring the t(4;14)(p16;q32) translocation lacks FGFR3 expression but maintains an IGH/MMSET fusion transcript. *Blood* 2003 Mar 15; **101**(6)**:** 2374-2376.

36. Martinez-Garcia E, Popovic R, Min DJ, Sweet SM, Thomas PM, Zamdborg L*, et al.* The MMSET histone methyl transferase switches global histone methylation and alters gene expression in t(4;14) multiple myeloma cells. *Blood* 2011 Jan 6; **117**(1)**:** 211-220.

37. Hurt EM, Wiestner A, Rosenwald A, Shaffer AL, Campo E, Grogan T, *et al.* Overexpression of c-maf is a frequent oncogenic event in multiple myeloma that promotes proliferation and pathological interactions with bone marrow stroma. *Cancer Cell* 2004 Feb; **5**(2)**:** 191-199.

38. Hanamura I, Iida S, Akano Y, Hayami Y, Kato M, Miura K, *et al.* Ectopic expression of MAFB gene in human myeloma cells carrying (14;20)(q32;q11) chromosomal translocations. *Japanese journal of cancer research : Gann* 2001 Jun; **92**(6)**:** 638-644.

39. Walker BA, Wardell CP, Brioli A, Boyle E, Kaiser MF, Begum DB, *et al.* Translocations at 8q24 juxtapose MYC with genes that harbor superenhancers resulting in overexpression and poor prognosis in myeloma patients. *Blood Cancer J* 2014; **4:** e191.

40. Avet-Loiseau H, Attal M, Moreau P, Charbonnel C, Garban F, Hulin C, *et al.* Genetic abnormalities and survival in multiple myeloma: the experience of the Intergroupe Francophone du Myelome. *Blood* 2007 Apr 15; **109**(8)**:** 3489-3495.

41. Hanamura I, Stewart JP, Huang YS, Zhan FH, Santra M, Sawyer JR, *et al.* Frequent gain of chromosome band 1q21 in plasma-cell dyscrasias detected by fluorescence in situ hybridization: incidence increases from MGUS to relapsed myeloma and is related to prognosis and disease progression following tandem stem-cell transplantation. *Blood* 2006 Sep 1; **108**(5)**:** 1724-1732.

42. Avet-Loiseau H, Li C, Magrangeas F, Gouraud W, Charbonnel C, Harousseau JL, *et al.* Prognostic significance of copy-number alterations in multiple myeloma. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 2009 Sep 20; **27**(27)**:** 4585-4590.

43. Egan JB, Shi CX, Tembe W, Christoforides A, Kurdoglu A, Sinari S, *et al.* Whole-genome sequencing of multiple myeloma from diagnosis to plasma cell leukemia reveals genomic initiating events, evolution, and clonal tides. *Blood* 2012 Aug 02; **120**(5)**:** 1060-1066.

44. Keats JJ, Chesi M, Egan JB, Garbitt VM, Palmer SE, Braggio E, *et al.* Clonal competition with alternating dominance in multiple myeloma. *Blood* 2012 Aug 2; **120**(5)**:** 1067-1076.

45. Magrangeas F, Avet-Loiseau H, Gouraud W, Lode L, Decaux O, Godmer P*, et al.* Minor clone provides a reservoir for relapse in multiple myeloma. *Leukemia* 2013 Feb; **27**(2)**:** 473-481.

46. Hebraud B, Caillot D, Corre J, Marit G, Hulin C, Leleu X*, et al.* The translocation t(4;14) can be present only in minor subclones in multiple myeloma. *Clinical cancer research : an official journal of the American Association for Cancer Research* 2013 Sep 1; **19**(17)**:** 4634-4637.

47. Rollig C, Knop S, Bornhauser M. Multiple myeloma. *Lancet* 2015 May 30; **385**(9983)**:** 2197-2208.

48. Melchor L, Brioli A, Wardell CP, Murison A, Potter NE, Kaiser MF*, et al.* Single-cell genetic analysis reveals the composition of initiating clones and phylogenetic patterns of branching and parallel evolution in myeloma. *Leukemia* 2014 Aug; **28**(8)**:** 1705-1715.

49. Ding L, Ellis MJ, Li S, Larson DE, Chen K, Wallis JW*, et al.* Genome remodelling in a basal-like breast cancer metastasis and xenograft. *Nature* 2010 Apr 15; **464**(7291)**:** 999-1005.

50. Yachida S, Jones S, Bozic I, Antal T, Leary R, Fu B*, et al.* Distant metastasis occurs late during the genetic evolution of pancreatic cancer. *Nature* 2010 Oct 28; **467**(7319)**:** 1114-1117.

51. Ding L, Ley TJ, Larson DE, Miller CA, Koboldt DC, Welch JS*, et al.* Clonal evolution in relapsed acute myeloid leukaemia revealed by whole-genome sequencing. *Nature* 2012 Jan 11; **481**(7382)**:** 506-510.

52. Brioli A, Melchor L, Cavo M, Morgan GJ. The impact of intra-clonal heterogeneity on the treatment of multiple myeloma. *British journal of haematology* 2014 May; **165**(4)**:** 441-454.

53. Agnelli L, Bicciato S, Mattioli M, Fabris S, Intini D, Verdelli D*, et al.* Molecular classification of multiple myeloma: a distinct transcriptional profile characterizes

patients expressing CCND1 and negative for 14q32 translocations. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 2005 Oct 10; **23**(29)**:** 7296-7306.

54. Shaughnessy JD, Zhan FH, Burington BE, Huang YS, Colla S, Hanamura I*, et al.* A validated gene expression model of high-risk multiple myeloma is defined by deregulated expression of genes mapping to chromosome 1. *Blood* 2007 Mar 15; **109**(6)**:** 2276-2284.

55. Broyl A, Hose D, Lokhorst H, de Knegt Y, Peeters J, Jauch A*, et al.* Gene expression profiling for molecular classification of multiple myeloma in newly diagnosed patients. *Blood* 2010 Oct 7; **116**(14)**:** 2543-2553.

56. Zhan F, Hardin J, Kordsmeier B, Bumm K, Zheng M, Tian E*, et al.* Global gene expression profiling of multiple myeloma, monoclonal gammopathy of undetermined significance, and normal bone marrow plasma cells. *Blood* 2002 Mar 1; **99**(5)**:** 1745-1757.

57. Zhan F, Huang Y, Colla S, Stewart JP, Hanamura I, Gupta S*, et al.* The molecular classification of multiple myeloma. *Blood* 2006 Sep 15; **108**(6)**:** 2020-2028.

58. Hose D, Seckinger A, Jauch A, Reme T, Moreaux J, Bertsch U*, et al.* The role of fluorescence in situ hybridization and gene expression profiling in myeloma risk stratification. *Srpski arhiv za celokupno lekarstvo* 2011 Dec; **139 Suppl 2:** 84-89.

59. Seckinger A, Meissner T, Moreaux J, Benes V, Hillengass J, Castoldi M*, et al.* miRNAs in multiple myeloma--a survival relevant complex regulator of gene expression. *Oncotarget* 2015 Nov 17; **6**(36)**:** 39165-39183.

60. Agnelli L, Tassone P, Neri A. Molecular profiling of multiple myeloma: from gene expression analysis to next-generation sequencing. *Expert opinion on biological therapy* 2013 Jun; **13 Suppl 1:** S55-68.

61. Mattioli M, Agnelli L, Fabris S, Baldini L, Morabito F, Bicciato S*, et al.* Gene expression profiling of plasma cell dyscrasias reveals molecular patterns associated with distinct IGH translocations in multiple myeloma. *Oncogene* 2005 Apr 7; **24**(15)**:** 2461-2473.

62. Johnsen JM, Nickerson DA, Reiner AP. Massively parallel sequencing: the new frontier of hematologic genomics. *Blood* 2013 Nov 07; **122**(19)**:** 3268-3275.

63. Chapman MA, Lawrence MS, Keats JJ, Cibulskis K, Sougnez C, Schinzel AC*, et al.* Initial genome sequencing and analysis of multiple myeloma. *Nature* 2011 Mar 24; **471**(7339)**:** 467-472.

64. Bolli N, Avet-Loiseau H, Wedge DC, Van Loo P, Alexandrov LB, Martincorena I*, et al.* Heterogeneity of genomic evolution and mutational profiles in multiple myeloma. *Nature communications* 2014; **5:** 2997.

65. Lohr JG, Stojanov P, Carter SL, Cruz-Gordillo P, Lawrence MS, Auclair D*, et al.* Widespread genetic heterogeneity in multiple myeloma: implications for targeted therapy. *Cancer Cell* 2014 Jan 13; **25**(1)**:** 91-101.

66. Walker BA, Wardell CP, Melchor L, Brioli A, Johnson DC, Kaiser MF*, et al.* Intraclonal heterogeneity is a critical early event in the development of myeloma and precedes the development of clinical symptoms. *Leukemia* 2014 Feb; **28**(2)**:** 384-390.

67. Leich E, Weissbach S, Klein HU, Grieb T, Pischimarov J, Stuhmer T*, et al.* Multiple myeloma is affected by multiple and heterogeneous somatic mutations in adhesion- and receptor tyrosine kinase signaling molecules. *Blood Cancer J* 2013 Feb; **3:** e102.

68. Alexandrov LB, Nik-Zainal S, Wedge DC, Aparicio SA, Behjati S, Biankin AV*, et al.* Signatures of mutational processes in human cancer. *Nature* 2013 Aug 22; **500**(7463)**:** 415-421.

69. Tiacci E, Trifonov V, Schiavoni G, Holmes A, Kern W, Martelli MP*, et al.* BRAF mutations in hairy-cell leukemia. *The New England journal of medicine* 2011 Jun 16; **364**(24)**:** 2305-2315.

70. Treon SP, Xu L, Yang G, Zhou Y, Liu X, Cao Y*, et al.* MYD88 L265P somatic mutation in Waldenstrom's macroglobulinemia. *The New England journal of medicine* 2012 Aug 30; **367**(9)**:** 826-833.

71. Walker BA, Boyle EM, Wardell CP, Murison A, Begum DB, Dahir NM*, et al.* Mutational Spectrum, Copy Number Changes, and Outcome: Results of a Sequencing Study of Patients With Newly Diagnosed Myeloma. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 2015 Nov 20; **33**(33)**:** 3911-3920.

72. Agnelli L, Neri A. Next-generation sequencing in multiple myeloma: insights into the molecular heterogeneity of the disease. *International Journal of Hematologic Oncology* 2014; **3**(5)**:** 367-376.

73. Andrulis M, Lehners N, Capper D, Penzel R, Heining C, Huellein J*, et al.* Targeting the BRAF V600E mutation in multiple myeloma. *Cancer discovery* 2013 Aug; **3**(8)**:** 862-869.

74. Bohn OL, Hsu K, Hyman DM, Pignataro DS, Giralt S, Teruya-Feldstein J. BRAF V600E mutation and clonal evolution in a patient with relapsed refractory myeloma with plasmablastic differentiation. *Clinical lymphoma, myeloma & leukemia* 2014 Apr; **14**(2)**:** e65-68.

75. Sharman JP, Chmielecki J, Morosini D, Palmer GA, Ross JS, Stephens PJ*, et al.* Vemurafenib response in 2 patients with posttransplant refractory BRAF V600E-mutated multiple myeloma. *Clinical lymphoma, myeloma & leukemia* 2014 Oct; **14**(5)**:** e161-163.

76. Tomecki R, Drazkowska K, Kucinski I, Stodus K, Szczesny RJ, Gruchota J*, et al.* Multiple myeloma-associated hDIS3 mutations cause perturbations in cellular RNA metabolism and suggest hDIS3 PIN domain as a potential drug target. *Nucleic acids research* 2014 Jan; **42**(2)**:** 1270-1290.

77. de Groen FL, Krijgsman O, Tijssen M, Vriend LE, Ylstra B, Hooijberg E*, et al.* Gene-dosage dependent overexpression at the 13q amplicon identifies DIS3 as candidate oncogene in colorectal cancer progression. *Genes ChromosomesCancer* 2014; **53**(4)**:** 339-348.

78. Pils D, Tong D, Hager G, Obermayr E, Aust S, Heinze G*, et al.* A combined blood based gene expression and plasma protein abundance signature for diagnosis of epithelial ovarian cancer--a study of the OVCAD consortium. *BMC cancer* 2013; **13:** 178.

79. Walker BA, Wardell CP, Melchor L, Hulkki S, Potter NE, Johnson DC*, et al.* Intraclonal heterogeneity and distinct molecular mechanisms characterize the development of t(4;14) and t(11;14) myeloma. *Blood* 2012 Aug 2; **120**(5)**:** 1077-1086.

80. Weiβbach S, Langer C, Puppe B, Nedeva T, Bach E, Kull M*, et al.* The molecular spectrum and clinical impact of DIS3 mutations in multiple myeloma. *British journal of haematology* 2015 Apr; **169**(1)**:** 57-70.

81. Segalla S, Pivetti S, Todoerti K, Chudzik MA, Giuliani EC, Lazzaro F*, et al.* The ribonuclease DIS3 promotes let-7 miRNA maturation by degrading the pluripotency factor LIN28B mRNA. *Nucleic acids research* 2015 Apr 29.

82. Hebraud B, Leleu X, Lauwers-Cances V, Roussel M, Caillot D, Marit G*, et al.* Deletion of the 1p32 region is a major independent prognostic factor in young patients with myeloma: the IFM experience on 1195 patients. *Leukemia* 2014 Mar; **28**(3)**:** 675-679.

83. Boyd KD, Ross FM, Walker BA, Wardell CP, Tapper WJ, Chiecchio L*, et al.* Mapping of chromosome 1p deletions in myeloma identifies FAM46C at 1p12 and CDKN2C at 1p32.3 as being genes in regions associated with adverse survival. *Clinical cancer research : an official journal of the American Association for Cancer Research* 2011 Dec 15; **17**(24)**:** 7776-7784.

84. Vousden KH, Lu X. Live or let die: the cell's response to p53. *Nature reviews Cancer* 2002 Aug; **2**(8)**:** 594-604.

85. Corradini P, Inghirami G, Astolfi M, Ladetto M, Voena C, Ballerini P*, et al.* Inactivation of tumor suppressor genes, p53 and Rb1, in plasma cell dyscrasias. *Leukemia* 1994 May; **8**(5)**:** 758-767.

86. Neri A, Baldini L, Trecca D, Cro L, Polli E, Maiolo AT. p53 gene mutations in multiple myeloma are associated with advanced forms of malignancy. *Blood* 1993 Jan 1; **81**(1)**:** 128-135.

87. Portier M, Moles JP, Mazars GR, Jeanteur P, Bataille R, Klein B*, et al.* p53 and RAS gene mutations in multiple myeloma. *Oncogene* 1992 Dec; **7**(12)**:** 2539-2543.

88. Preudhomme C, Facon T, Zandecki M, Vanrumbeke M, Lai JL, Nataf E, *et al.* Rare occurrence of P53 gene mutations in multiple myeloma. *British journal of haematology* 1992 Jul; **81**(3)**:** 440-443.

89. Cifola I, Lionetti M, Pinatel E, Todoerti K, Mangano E, Pietrelli A, *et al.* Whole-exome sequencing of primary plasma cell leukemia discloses heterogeneous mutational patterns. *Oncotarget* 2015 Jul 10; **6**(19)**:** 17543-17558.

90. Tiedemann RE, Gonzalez-Paz N, Kyle RA, Santana-Davila R, Price-Troska T, Van Wier SA, *et al.* Genetic aberrations and survival in plasma cell leukemia. *Leukemia* 2008 May; **22**(5)**:** 1044-1052.

91. Mazars GR, Portier M, Zhang XG, Jourdan M, Bataille R, Theillet C, *et al.* Mutations of the p53 gene in human myeloma cell lines. *Oncogene* 1992 May; **7**(5)**:** 1015-1018.

92. Lode L, Eveillard M, Trichet V, Soussi T, Wuilleme S, Richebourg S, *et al.* Mutations in TP53 are exclusively associated with del(17p) in multiple myeloma. *Haematologica* 2010 Nov; **95**(11)**:** 1973-1976.

93. Chang H, Qi C, Yi QL, Reece D, Stewart AK. p53 gene deletion detected by fluorescence in situ hybridization is an adverse prognostic factor for patients with multiple myeloma following autologous stem cell transplantation. *Blood* 2005 Jan 1; **105**(1)**:** 358-360.

94. Schultheis B, Kramer A, Willer A, Hegenbart U, Goldschmidt H, Hehlmann R. Analysis of p73 and p53 gene deletions in multiple myeloma. *Leukemia* 1999 Dec; **13**(12)**:** 2099-2103.

95. Mosca L, Musto P, Todoerti K, Barbieri M, Agnelli L, Fabris S, *et al.* Genome-wide analysis of primary plasma cell leukemia identifies recurrent imbalances associated with changes in transcriptional profiles. *American journal of hematology* 2013 Jan; **88**(1)**:** 16-23.

96. Fonseca R, Bergsagel PL, Drach J, Shaughnessy J, Gutierrez N, Stewart AK, *et al.* International Myeloma Working Group molecular classification of multiple myeloma: spotlight review. *Leukemia* 2009 Dec; **23**(12)**:** 2210-2221.

97. An G, Li Z, Tai YT, Acharya C, Li Q, Qin X*, et al.* The impact of clone size on the prognostic value of chromosome aberrations by fluorescence in situ hybridization in multiple myeloma. *Clinical cancer research : an official journal of the American Association for Cancer Research* 2015 May 1; **21**(9)**:** 2148-2156.

98. Teoh PJ, Chung TH, Sebastian S, Choo SN, Yan J, Ng SB*, et al.* p53 haploinsufficiency and functional abnormalities in multiple myeloma. *Leukemia* 2014 Oct; **28**(10)**:** 2066-2074.

99. Gerlinger M, Rowan AJ, Horswell S, Math M, Larkin J, Endesfelder D*, et al.* Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. *The New England journal of medicine* 2012 Mar 08; **366**(10)**:** 883-892.

100. Lui YY, Chik KW, Chiu RW, Ho CY, Lam CW, Lo YM. Predominant hematopoietic origin of cell-free DNA in plasma and serum after sex-mismatched bone marrow transplantation. *Clinical chemistry* 2002 Mar; **48**(3)**:** 421-427.

101. Buedts L, Vandenberghe P. Circulating cell-free DNA in hematological malignancies. *Haematologica* 2016 Sep; **101**(9)**:** 997-999.

102. Crowley E, Di Nicolantonio F, Loupakis F, Bardelli A. Liquid biopsy: monitoring cancer-genetics in the blood. *Nature reviews Clinical oncology* 2013 Aug; **10**(8)**:** 472-484.

103. Diaz LA, Jr., Bardelli A. Liquid biopsies: genotyping circulating tumor DNA. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 2014 Feb 20; **32**(6)**:** 579-586.

104. Oberle A, Brandt A, Voigtlaender M, Thiele B, Radloff J, Schulenkorf A*, et al.* Monitoring multiple myeloma by next-generation sequencing of V(D)J rearrangements from circulating myeloma cells and cell-free myeloma DNA. *Haematologica* 2017 Jun; **102**(6)**:** 1105-1111.

105. Kis O, Kaedbey R, Chow S, Danesh A, Dowar M, Li T*, et al.* Circulating tumour DNA sequence analysis as an alternative to multiple myeloma bone marrow aspirates. *Nature communications* 2017 May 11; **8:** 15086.

106. Mithraprabhu S, Khong T, Ramachandran M, Chow A, Klarica D, Mai L*, et al.* Circulating tumour DNA analysis demonstrates spatial mutational heterogeneity that coincides with disease relapse in myeloma. *Leukemia* 2017 Jan 03.

107. Rustad EH, Coward E, Skytoen ER, Misund K, Holien T, Standal T*, et al.* Monitoring multiple myeloma by quantification of recurrent mutations in serum. *Haematologica* 2017 Jul; **102**(7)**:** 1266-1272.

108. Agnelli L, Fabris S, Bicciato S, Basso D, Baldini L, Morabito F*, et al.* Upregulation of translational machinery and distinct genetic subgroups characterise hyperdiploidy in multiple myeloma. *British journal of haematology* 2007 Feb; **136**(4)**:** 565-573.

109. Fabris S, Ronchetti D, Agnelli L, Baldini L, Morabito F, Bicciato S*, et al.* Transcriptional features of multiple myeloma patients with chromosome 1q gain. *Leukemia* 2007 May; **21**(5)**:** 1113-1116.

110. Agnelli L, Mosca L, Fabris S, Lionetti M, Andronache A, Kwee I*, et al.* A SNP microarray and FISH-based procedure to detect allelic imbalances in multiple myeloma: an integrated genomics approach reveals a wide gene dosage effect. *Genes, chromosomes & cancer* 2009 Jul; **48**(7)**:** 603-614.

111. Lionetti M, Musto P, Di Martino MT, Fabris S, Agnelli L, Todoerti K*, et al.* Biological and clinical relevance of miRNA expression signatures in primary plasma cell leukemia. *Clinical cancer research : an official journal of the American Association for Cancer Research* 2013 Jun 15; **19**(12)**:** 3130-3142.

112. Todoerti K, Agnelli L, Fabris S, Lionetti M, Tuana G, Mosca L*, et al.* Transcriptional characterization of a prospective series of primary plasma cell leukemia revealed signatures associated with tumor progression and poorer outcome. *Clinical cancer research : an official journal of the American Association for Cancer Research* 2013 Jun 15; **19**(12)**:** 3247-3258.

113. Chng WJ, Gonzalez-Paz N, Price-Troska T, Jacobus S, Rajkumar SV, Oken MM*, et al.* Clinical and biological significance of RAS mutations in multiple myeloma. *Leukemia* 2008 Dec; **22**(12)**:** 2280-2284.

114. Musto P, Simeon V, Martorelli MC, Petrucci MT, Cascavilla N, Di Raimondo F*, et al.* Lenalidomide and low-dose dexamethasone for newly diagnosed primary plasma cell leukemia. *Leukemia* 2014 Jan; **28**(1)**:** 222-225.

115. Musto P, D'Auria F, Petrucci MT, Levi A, Cascavilla N, Falcone A*, et al.* Final Results of a Phase II Study Evaluating Lenalidomide in Combination with Low Dose Dexamethasone As First Line Therapy for Primary Plasma Cell Leukemia. *Blood* 2011 Nov 18; **118**(21)**:** 1261-1261.

116. Greipp PR, San Miguel J, Durie BG, Crowley JJ, Barlogie B, Blade J*, et al.* International staging system for multiple myeloma. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 2005 May 20; **23**(15)**:** 3412-3420.

117. Kuiper R, van Duin M, van Vliet MH, Broijl A, van der Holt B, El Jarari L*, et al.* Prediction of high- and low-risk multiple myeloma based on gene expression and the International Staging System. *Blood* 2015 Oct 22; **126**(17)**:** 1996-2004.

118. Fabris S, Agnelli L, Mattioli M, Baldini L, Ronchetti D, Morabito F*, et al.* Characterization of oncogene dysregulation in multiple myeloma by combined FISH and DNA microarray analyses. *Genes, chromosomes & cancer* 2005 Feb; **42**(2)**:** 117-127.

119. Davies H, Bignell GR, Cox C, Stephens P, Edkins S, Clegg S*, et al.* Mutations of the BRAF gene in human cancer. *Nature* 2002 Jun 27; **417**(6892)**:** 949-954.

120. Chang YS, Yeh KT, Hsu NC, Lin SH, Chang TJ, Chang JG. Detection of N-, H-, and KRAS codons 12, 13, and 61 mutations with universal RAS primer multiplex PCR and N-, H-, and KRAS-specific primer extension. *Clinical biochemistry* 2010 Feb; **43**(3)**:** 296-301.

121. Kortuem KM, Braggio E, Bruins L, Barrio S, Shi CS, Zhu YX*, et al.* Panel sequencing for clinically oriented variant screening and copy number detection in 142 untreated multiple myeloma patients. *Blood Cancer J* 2016 Feb 26; **6:** e397.

122. Newman AM, Bratman SV, To J, Wynne JF, Eclov NC, Modlin LA*, et al.* An ultrasensitive method for quantitating circulating tumor DNA with broad patient coverage. *Nature medicine* 2014 May; **20**(5)**:** 548-554.

123. Irizarry RA, Hobbs B, Collin F, Beazer-Barclay YD, Antonellis KJ, Scherf U*, et al.* Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics* 2003 Apr; **4**(2)**:** 249-264.

124. Dai MH, Wang PL, Boyd AD, Kostov G, Athey B, Jones EG*, et al.* Evolving gene/transcript definitions significantly alter the interpretation of GeneChip data. *Nucleic acids research* 2005; **33**(20)**:** e175.

125. Tusher VG, Tibshirani R, Chu G. Significance analysis of microarrays applied to the ionizing radiation response. *Proceedings of the National Academy of Sciences of the United States of America* 2001 Apr 24; **98**(9)**:** 5116-5121.

126. Chen J, Bardes EE, Aronow BJ, Jegga AG. ToppGene Suite for gene list enrichment analysis and candidate gene prioritization. *Nucleic acids research* 2009 Jul; **37**(Web Server issue)**:** W305-311.

127. Bezieau S, Devilder MC, Avet-Loiseau H, Mellerin MP, Puthier D, Pennarun E*, et al.* High incidence of N and K-Ras activating mutations in multiple myeloma and primary plasma cell leukemia at diagnosis. *Hum Mutat* 2001; **18**(3)**:** 212-224.

128. Neri A, Murphy JP, Cro L, Ferrero D, Tarella C, Baldini L*, et al.* Ras oncogene mutation in multiple myeloma. *The Journal of experimental medicine* 1989 Nov 1; **170**(5)**:** 1715-1725.

129. Lorentzen E, Basquin J, Tomecki R, Dziembowski A, Conti E. Structure of the active subunit of the yeast exosome core, Rrp44: diverse modes of substrate recruitment in the RNase II nuclease family. *Molecular cell* 2008 Mar 28; **29**(6)**:** 717-728.

130. Petitjean A, Mathe E, Kato S, Ishioka C, Tavtigian SV, Hainaut P*, et al.* Impact of mutant p53 functional properties on TP53 mutation patterns and tumor phenotype: lessons from recent developments in the IARC TP53 database. *Hum Mutat* 2007 Jun; **28**(6)**:** 622-629.

131. Heidorn SJ, Milagre C, Whittaker S, Nourry A, Niculescu-Duvas I, Dhomen N*, et al.* Kinase-Dead BRAF and Oncogenic RAS Cooperate to Drive Tumor Progression through CRAF. *Cell* 2010 Jan 22; **140**(2)**:** 209-221.

132. Loreni F, Mancino M, Biffo S. Translation factors and ribosomal proteins control tumor onset and progression: how? *Oncogene* 2014 Apr 24; **33**(17)**:** 2145-2156.

133. Lionetti M, Barbieri M, Todoerti K, Agnelli L, Fabris S, Tonon G*, et al.* A compendium of DIS3 mutations and associated transcriptional signatures in plasma cell dyscrasias. *Oncotarget* 2015 Sep 22; **6**(28)**:** 26129-26141.

134. Tian S, Simon I, Moreno V, Roepman P, Tabernero J, Snel M*, et al.* A combined oncogenic pathway signature of BRAF, KRAS and PI3KCA mutation improves colorectal cancer classification and cetuximab treatment prediction. *Gut* 2013 Apr; **62**(4)**:** 540-549.

135. Mulligan G, Lichter DI, Di Bacco A, Blakemore SJ, Berger A, Koenig E*, et al.* Mutation of NRAS but not KRAS significantly reduces myeloma sensitivity to single-agent bortezomib therapy. *Blood* 2014; **123**(5)**:** 632-639.

136. Wan PT, Garnett MJ, Roe SM, Lee S, Niculescu-Duvaz D, Good VM*, et al.* Mechanism of activation of the RAF-ERK signaling pathway by oncogenic mutations of B-RAF. *Cell* 2004 Mar 19; **116**(6)**:** 855-867.

137. Heidorn SJ, Milagre C, Whittaker S, Nourry A, Niculescu-Duvas I, Dhomen N, et al. Kinase-Dead *BRAF* and Oncogenic *RAS* Cooperate to Drive Tumor Progression through CRAF. *Cell* 2010 Jan 22; **140**(2): 209-221.

138. Dziembowski A, Lorentzen E, Conti E, Seraphin B. A single subunit, Dis3, is essentially responsible for yeast exosome core activity. *Nature structural & molecular biology* 2007 Jan; **14**(1)**:** 15-22.

139. Chekanova JA, Gregory BD, Reverdatto SV, Chen H, Kumar R, Hooker T*, et al.* Genome-wide high-resolution mapping of exosome substrates reveals hidden features in the Arabidopsis transcriptome. *Cell* 2007 Dec 28; **131**(7)**:** 1340-1353.

140. Houalla R, Devaux F, Fatica A, Kufel J, Barrass D, Torchet C*, et al.* Microarray detection of novel nuclear RNA substrates for the exosome. *Yeast* 2006 Apr 30; **23**(6)**:** 439-454.

141. Kiss DL, Andrulis ED. Genome-wide analysis reveals distinct substrate specificities of Rrp6, Dis3, and core exosome subunits. *Rna* 2010 Apr; **16**(4)**:** 781-791.

142. Kuchta K, Muszewska A, Knizewski L, Steczkiewicz K, Wyrwicz LS, Pawlowski K*, et al.* FAM46 proteins are novel eukaryotic non-canonical poly(A) polymerases. *Nucleic acids research* 2016 May 05; **44**(8)**:** 3534-3548.

143. Kortum KM, Langer C, Monge J, Bruins L, Egan JB, Zhu YX*, et al.* Targeted sequencing using a 47 gene multiple myeloma mutation panel (M(3) P) in -17p high risk disease. *British journal of haematology* 2015 Feb; **168**(4)**:** 507-510.

144. Chng WJ, Price-Troska T, Gonzalez-Paz N, Van Wier S, Jacobus S, Blood E*, et al.* Clinical significance of TP53 mutation in myeloma. *Leukemia* 2007 Mar; **21**(3)**:** 582-584.

145. Muller PA, Vousden KH. Mutant p53 in cancer: new functions and therapeutic opportunities. *Cancer Cell* 2014 Mar 17; **25**(3)**:** 304-317.

146. Bullock AN, Henckel J, Fersht AR. Quantitative analysis of residual folding and DNA binding in mutant p53 core domain: definition of mutant states for rescue in cancer therapy. *Oncogene* 2000 Mar 2; **19**(10)**:** 1245-1256.

147. Friedler A, Veprintsev DB, Hansson LO, Fersht AR. Kinetic instability of p53 core domain mutants: implications for rescue by small molecules. *The Journal of biological chemistry* 2003 Jun 27; **278**(26)**:** 24108-24112.

148. Brachmann RK, Vidal M, Boeke JD. Dominant-negative p53 mutations selected in yeast hit cancer hot spots. *Proceedings of the National Academy of Sciences of the United States of America* 1996 Apr 30; **93**(9)**:** 4091-4095.

149. Christgen M, Noskowicz M, Heil C, Schipper E, Christgen H, Geffers R*, et al.* IPH-926 lobular breast cancer cells harbor a p53 mutant with temperature-sensitive functional activity and allow for profiling of p53-responsive genes. *Lab Invest* 2012 Nov; **92**(11)**:** 1635-1647.

150. Jia LQ, Osada M, Ishioka C, Gamo M, Ikawa S, Suzuki T*, et al.* Screening the p53 status of human cell lines using a yeast functional assay. *Molecular carcinogenesis* 1997 Aug; **19**(4)**:** 243-253.

151. Yoshikawa K, Hamada J, Tada M, Kameyama T, Nakagawa K, Suzuki Y*, et al.* Mutant p53 R248Q but not R248W enhances in vitro invasiveness of human lung cancer NCI-H1299 cells. *Biomedical research* 2010 Dec; **31**(6)**:** 401-411.

152. Weinhold N, Ashby C, Rasche L, Chavan SS, Stein C, Stephens OW*, et al.* Clonal selection and double-hit events involving tumor suppressor genes underlie relapse in myeloma. *Blood* 2016 Sep 29; **128**(13)**:** 1735-1744.

153. Rashid NU, Sperling AS, Bolli N, Wedge DC, Van Loo P, Tai YT*, et al.* Differential and limited expression of mutant alleles in multiple myeloma. *Blood* 2014 Nov 13; **124**(20)**:** 3110-3117.

154. Flaherty KT, Puzanov I, Kim KB, Ribas A, McArthur GA, Sosman JA*, et al.* Inhibition of mutated, activated BRAF in metastatic melanoma. *The New England journal of medicine* 2010 Aug 26; **363**(9)**:** 809-819.

155. O'Donnell E, Raje NS. Targeting BRAF in multiple myeloma. *Cancer discovery* 2013 Aug; **3**(8)**:** 840-842.

156. Abbosh C, Birkbak NJ, Wilson GA, Jamal-Hanjani M, Constantin T, Salari R*, et al.* Phylogenetic ctDNA analysis depicts early-stage lung cancer evolution. *Nature* 2017 Apr 26; **545**(7655)**:** 446-451.

157. Rasche L, Chavan SS, Stephens OW, Patel PH, Tytarenko R, Ashby C*, et al.* Spatial genomic heterogeneity in multiple myeloma revealed by multi-region sequencing. *Nature communications* 2017 Aug 16; **8**(1)**:** 268.