UNIVERSITÀ DEGLI STUDI DI MILANO

Scuola di Dottorato in
Scienze Biomediche Cliniche e Sperimentali
Corso di Dottorato in Ematologia Sperimentale

Dipartimento di
Scienze Cliniche e di Comunità

HIGH-THROUGHPUT SEQUENCING FOR THE IDENTIFICATION OF *DIS3* MUTATIONS IN MULTIPLE MYELOMA

Tesi di dottorato di:
Marzia BARBIERI
Matricola nr. R09512

Relatore: Chiar.mo Prof. Antonino NERI
Coordinatore: Chiar.mo Prof. Paolo CORRADINI

A.A. 2013/2014

Table of contents

Summary

Introduction and purposes.

Multiple myeloma (MM) is an incurable malignancy of mature plasma cells (PCs), and its pathogenesis is only partially understood. Recently, whole exome sequencing studies have identified mutations of *DIS3,* a catalytic subunit of the human exosome complex, in about 10% of MM patients. This study was aimed at investigating the spectrum of *DIS3* mutations indifferent stages of plasma cell dyscrasia.

Experimental design.

To analyze *DIS3* mutations, we investigated by next generation sequencing (NGS) a retrospective cohort of 130 cases with MM at onset and 17 at relapse (of whom 15 were also tested at onset). Moreover, we examined 24 patients with primary PC leukemia (pPCL), 12 with secondary PCL and 20 multiple myeloma cell lines. Deep sequencing of the PIN (exons 1-4) and RNB (exons 10-18) *DIS3* functional domains was performed by Roche 454 pyrosequencing on the Genome Sequencer Junior instrument. Mutations were validated by conventional Sanger sequencing or independent ultra-deep 454 pyrosequencing. All samples were characterized by fluorescence *in situ* hybridization (FISH) for the main genomic aberrations, such as *IGH* translocations, 13q and 17p deletions, hyperdiploidy, 1p33 (*CDKN2C*) loss and 1q21.3 (*CKS1B*) gain. In order to verify if the mutations detected on genomic DNA were expressed at transcriptomic level, *DIS3* cDNA of mutated cases was subjected to deep sequencing. Global gene expression of MM patients was also profiled, looking for transcriptional patterns possibly related to *DIS3* mutations. Additionally, to investigate *DIS3* mutation status longitudinally, we analyzed 20 patients whose bone marrow specimens were collected at two different time points. Finally, the association of *DIS3* mutations with clinical outcome was tested, comparing wild-type patients (n=12) with those characterized by mutations in PIN or RNB functional domains of *DIS3* (n=4).

Results.

NGS analysis revealed the presence of 41 coding non-synonymous variants, with a mutant allele frequency ranging from 0.38% to 100% of total reads (median depth of coverage 245x, range: 64-1160). Among the 41 tumor-specific mutations, 30 (73%) were single nucleotide variations, and the remaining 11 (27%) were indels. Nine of these variants have been already reported by others, eight of which also specifically in MM patients, while 32 were novel. The great majority of mutations identified in our cohort of patients affected the RNB domain (30/41, 73.2%). The mutations affected 26 MM patients at diagnosis (26/130, 20%), four at relapse (4/17, 23.5%), six pPCL (6/24, 25%), four sPCL (4/12, 33.3%) cases, and three multiple myeloma cell lines (3/20, 15%). We observed a positive

association of *DIS3* mutations with the occurrence of translocations involving IGH@ *locus*, particularly with the t(11;14), and a negative association with hyperdiploid cases, but no association with 13q and 17p deletion, gain of chromosome 1q and loss of chromosome 1p. Additionally, we showed that mutant allele frequencies detected on genomic DNA and on retrotranscribed total RNA are linearly correlated. Moreover, in serially analyzed patients, we detected mutations at constant allele frequency at both time points, or acquired/increased variants in the late sample, consistent with the expected positive selection of mutated subclones. Furthermore, global gene expression profiling analysis revealed 119 differentially expressed genes (all up-regulated in mutated cases) between *DIS3* mutated and wild-type cases. Finally, we tested the association of *DIS3* mutations with clinical outcome in 16 pPCL patients for whom the follow-up was available, showing that mutational events did not show any impact neither on PFS nor OS.

Conclusion.

Our data confirm *DIS3* as frequently mutated in MM, and importantly, seem to indicate an even greater involvement of *DIS3* alteration in more advanced stages of PC dyscrasias. Furthermore, these data, although requiring confirmation in independent patients' series, support the hypothesis that *DIS3* may play a role in development and progression of MM.

Further studies are required to elucidate the role of this gene in the pathogenesis of MM, as well as its potential use as a drug target.

Introduction

MULTIPLE MYELOMA

Multiple myeloma is a malignant disease characterized by proliferation of clonal plasma cells (PC) in the bone marrow and typically accompanied by the secretion of monoclonal immunoglobulins detectable in serum and/or urine and the presence of lytic bone lesions [1]. MM accounts for about 10% of all hematological cancers [2;3], and the median age at diagnosis is 69 years, with three-quarters of patients being diagnosed above the age of 55 years and two of three patients being men. Multiple myeloma is the second most frequent haematological malignancy with an age-adjusted incidence of six per 100 000 per year in the USA and Europe [4]. MM is incurable disease, but with the advent of more effective therapeutic strategies and improvements in supportive care, the median survival has increased from 3 years to 6 years in the past two decades [5].

Cellular origin of myeloma cells

Multiple myeloma cells are similar to long-lived, post-germinal centre plasma cells with extensive somatic hypermutation of immunoglobulin genes. PCs are characterized by strong bone marrow dependence for survival and growth. The immunophenotype of MM cells resembles that of normal, terminally differentiated, long-lived BM PC (CD19–CD20–CD45–CD138+). In contrast to normal long-lived PCs, MM tumors retain some potential for an extremely low rate of proliferation, usually with no more than a few per cent of cycling cells, until advanced stages of MM [6].

Pathogenesis

MM affects antibody-secreting bone marrow (BM) plasma cells (PCs) and shows a wide clinical presentation ranging from the presumed pre-malignant condition of monoclonal gammopathy of undetermined significance (MGUS) to smoldering MM, truly overt and symptomatic MM, and extra-medullary myeloma or plasma cell leukemia (PCL). To date, MM diagnosis is based on the criteria established by the International Myeloma Working Group in 2003 [7] subsequently updated in 2009 [8].

Osteolytic bone lesions and/or compression fractures detected on routine radiographs, magnetic resonance imaging (MRI), or computed tomographic (CT) 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 plasma cell disorder is required for the diagnosis of the disease (Table 1) [9].

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 multiple myeloma)	▪ 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

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

Role of the microenvironment in MM

Similar to long-lived PCs, MGUS and MM cells are dependent on the BM microenvironment: interactions between myeloma cells and bone marrow cells or extracellular matrix proteins that are mediated through cell-surface receptors (e.g., integrins, cadherins, selectins, and cell-adhesion molecules) increase tumor growth, survival, migration, and drug resistance. The adhesion of myeloma cells to hematopoietic and stromal cells induces the secretion of cytokines and growth factors, including interleukin-6, vascular endothelial growth factor (VEGF), insulin-like growth factor 1, members of the superfamily of tumor necrosis factor, transforming growth factor β1, and interleukin-10. These cytokines and growth factors are produced and secreted by cells in the bone marrow microenvironment, including myeloma cells, and regulated by autocrine and paracrine loops [10]. The adhesion of myeloma cells to extracellular matrix proteins (e.g., collagen, fibronectin, laminin, and vitronectin) triggers the up-regulation of cell-cycle regulatory proteins and anti-apoptotic proteins [11]. Bone lesions are caused by an imbalance in the function of osteoblasts and osteoclasts [12]. The induction of proangiogenic molecules (e.g., VEGF) enhances the microvascular density of bone marrow and accounts for the abnormal structure of myeloma tumor vessels [11]. (Figure 1).

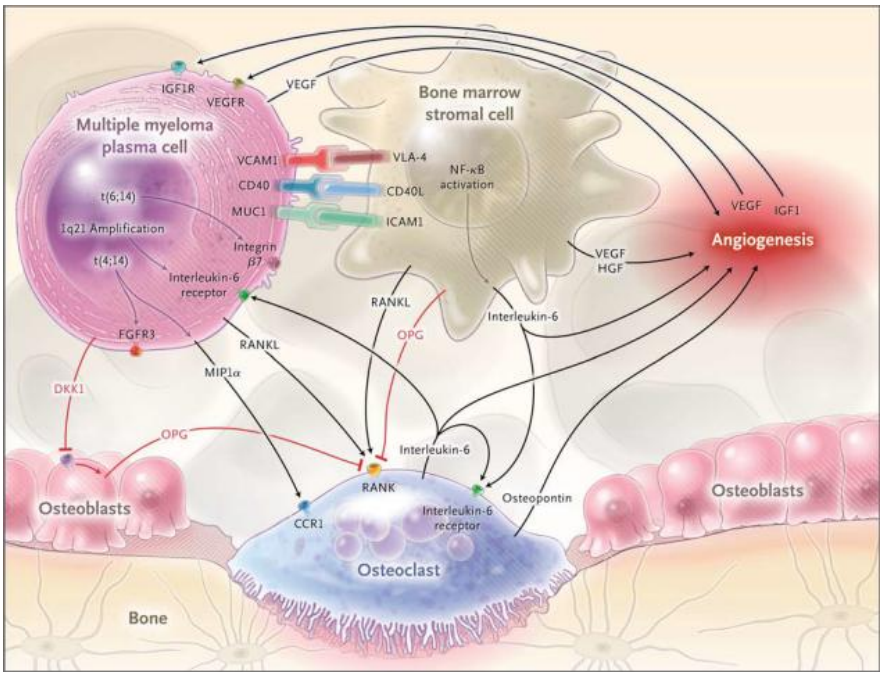

Figure 1. Interaction between plasma cells and bone marrow in multiple myeloma [13]

Multi-step clinical course of multiple myeloma

Virtually every case of MM is preceded by a pre-malignant PC tumor called monoclonal gammopathy of undetermined significance (MGUS) [14;15] that, like MM, produces a typical M-spike (almost always non-IgM) by serum protein electrophoresis (SPEP) or free light chain in the urine. PC MGUS is age dependent, is present in about 4% of individuals over the age of 50, and can progress to MM at average rates of 1% per year [16;17]. MGUS is distinguished from MM by having a M-spike of <30 g/L, with no more than 10 % of BM mononuclear cells being tumor cells, and no end organ damage or other symptoms. Progression of MGUS to smoldering MM and symptomatic MM is associated with an expanding BM tumor mass and increasingly severe organ impairment or symptoms [15]. Despite the recent advances in the understanding of the MM pathogenesis, it is still largely impossible to predict which MGUS patient will and which one will never progress to MM. Although MM cells are characterized by a strong dependence on the BM tumor microenvironment, at late stages of the disease the more aggressive tumor may sometimes extend to extramedullary locations, such as spleen, liver, and extra-cellular spaces. Extramedullary MM (EMM) can also present with a leukemic phase, that is classified as secondary or primary plasma cell leukemia (PCL), depending on whether or not a preceding intramedullary MM was recognized. Most of the available human MM cell lines (HMCLs) have been generated from EMM or PCL tumors and represent a renewable repository of the oncogenic events involved in initiation and progression of the most aggressive end-stage MM tumors [18;19] (Figure 2).

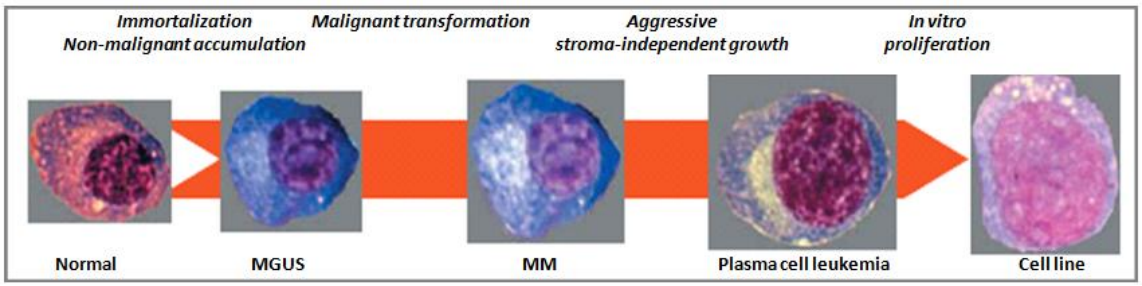

Figure 2. Multi-step molecular pathogenesis of MM. Progression through the different stages of multiple myeloma.

Clonal evolution of multiple myeloma

The recent acquisitions based on sequencing data [20-22] suggest that MM is genomically even more complex: it is now clear that the process of transformation is not due to a single genetic change but, at the time of diagnosis, there are frequently multiple independent clones whose dominance alternates during the subsequent phases of the disease in a treatment-related manner. This knowledge has also changed the view on the progression of MM from the premalignant condition of MGUS to the final stages of plasma cell leukemia.

Recent studies suggest that greater tumorigenicity and susceptibility to relapse rely on a 'reserve' of a number of clones, and demand for more detailed studies of the effects of different therapies. This is particularly important also in light of emerging evidence that a branching clonal evolution is probably a common characteristic of various solid [23;24] and hematological tumors (such as acute myeloid leukemia) [25].

Next generation sequencing analyses of paired samples that have undergone transformation from high-risk smoldering MM to overt MM have defined a number of genetic hits that contribute to subclonal growth and a survival advantage, and this (together with evidence that the complexity of the genome increases with disease stage) has suggested that disease progression is mediated by the competition of different subclones, the fittest of which emerges [26].

Nonetheless, combined whole exome sequencing (WES) and single-cell genetic analysis allowed to suggest that MM at onset may present up to six different major malignant clones, which are related by either linear or branching evolution (Figure 3). Interestingly, the study of Melchor *et al*. [27] provided evidence that the activation of the *RAS*/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 [27].

The fine characterization of subclones allowed by NGS may also provide important insights into the impact of therapy on MM. In the 'Darwinian' model of intraclonal heterogeneity, multiple clones descend from the same progenitor, each of which has the same main feature but also harbors a set of uniquely 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 [28;29].

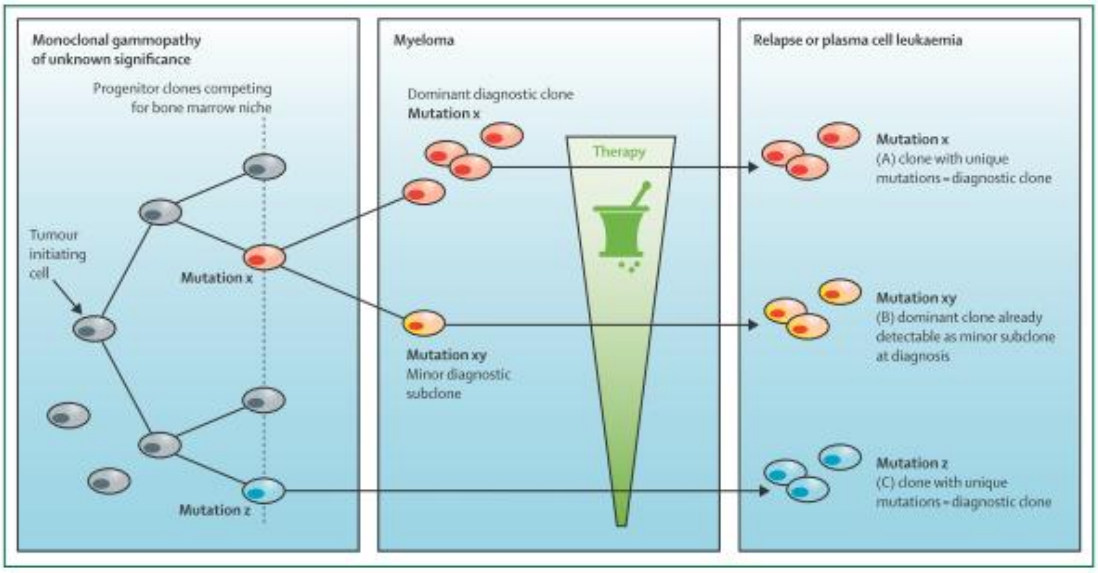

Figure 3. Clonal composition of multiple myeloma during disease progression and therapy [30].

Genetic architecture and disease progression

At the cytogenetic level, MM genome is complex and more reminiscent of epithelial cancers than of more simple leukemias [28]. Many of the genetic lesions that lead to MM have been defined, and can be categorized as inherited variation, translocations, copy number abnormalities, mutations, methylation and miRNA abnormalities (Table 2).

Chromosomal translocations

Translocations involving the immunoglobulin heavy chain (IgH) locus are present in at least half of MM cases and are thought to result from errors during the physiological process of class-switch recombination (CSR) or somatic hypermutation. It is presumed that these translocations represent primary – perhaps initiating – oncogenic events as normal B-cells pass through GCs. These translocations result in dysregulated expression of an oncogene (such as cyclin D1 (*CCND1*), *CCND3*, fibroblast growth factor receptor 3 (*FGFR3*), multiple myeloma SET domain (*MMSET*; also known as *WHSC1*), *MAF* and *MAFB*) that is juxtaposed to the strong Ig enhancers [31].

It is thought that cyclin D translocations only dysregulate expression of a cyclin D gene. By contrast *MAF* translocations dysregulate expression of a *MAF* transcription factor that causes increased expression of many genes, including cyclin D2 and adhesion molecules that are thought to enhance the

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

Table 2. Genetic events underlying the initiation and progression of MM to PCL [28].

ability of the tumor cell to interact with the BM microenvironment [31;32]. The contributions of the two genes dysregulated by t(4;14) remain controversial. *MMSET* is a chromatin-remodeling factor that is over-expressed in all tumors with a t(4;14), whereas about 20 % of tumors lack der(14) and *FGFR3* expression. The rare acquisition of *FGFR3* activating mutations during progression confirms a role for

FGFR3 in MM pathogenesis. Although an activated mutant *FGFR3* can be oncogenic, it recently was shown that wild-type *FGFR3* (as is found in most t(4; 14)) can contribute to B cell oncogenesis [33].

The patterns of spiked expression of genes deregulated by primary IgH translocations and the universal over-expression of *CCND*s genes led to the translocations and cyclin D (TC) classification that includes five groups: TC1 tumors (18%) express high levels of either cyclin D1 (11q13) or cyclin D3 (6p21) as a result of an Ig translocation; TC2 tumors (37%) ectopically express low to moderate levels of cyclin D1 despite the absence of a t(11;14) translocation; TC3 tumors (22%) are a mixture of tumors that do not fall into one of the other groups, with most expressing cyclin D2, but a few also expressing low levels of cyclin D1 or cyclin D3; TC4 tumors (16%) express high levels of cyclin D2, and also *MMSET* (and in most cases *FGFR3*) as a result of a t(4;14) translocation; and TC5 tumors (7%) express the highest levels of cyclin D2, and also high levels of either *MAF* or *MAFB*, consistent with evidence that both *MAF* transcription factors up-regulate the expression of cyclin D2 [31;34].

Copy number abnormalities

Gains and losses of DNA, resulting in copy number alterations, are common events in myeloma.

Nearly half of MGUS and MM tumors are hyperdiploid (HRD), with 48–75 (mostly 49–56) chromosomes, usually with extra copies of three or more specific chromosomes (3, 5, 7, 9, 11, 15, 19, 21). Non-hyperdiploid (NHRD) tumors have <48 and/or>75 chromosomes. Strikingly, HRD tumors rarely (~10%) have a primary IgH translocation, whereas NHRD tumors usually (~70%) have an IgH translocation [35]. Although it has been proposed that NHRD and HRD tumors represent different pathways of pathogenesis, the timing, mechanism, and molecular consequences of hyperdiploidy are unknown. In any case, HRD patients seem to have a better prognosis than NHRD patients [36].

Chromosome 1 aberrations are the most common structural aberrations in MM and mostly involve deletions in 1p and amplifications in 1q21 [37-39]. These genomic events frequently occur together in MM, and each is associated with a poor prognosis [39;40].

Gain of the chromosome 1q arm (+1q) is an event observed in 35–40% of presenting myeloma cases [37;39]; however, the prognostic value of 1q21 gain in MM remain controversial. To date, the relevant genes on 1q21 remain unclear. putative targets of this amplifications include *CKS1B* and *PMSD4* genes, mediating cell cicle progression and resistance to bortezomib, respectively [41;42].

Whole arm deletion or interstitial deletions of the 1p chromosome arm are observed in approximately 30% of myeloma patients [37;39]. Molecular genetics has revealed that two regions of 1p, 1p12 and 1p32.3, are particularly important in myeloma pathogenesis when deleted. 1p12 may be hemi or homozygously deleted and contains the candidate tumour suppressor gene *FAM46C* [43]. The function of *FAM46C* is unknown, although recent sequencing and homology studies have shown that its expression is correlated to both that of ribosomal proteins and eukaryotic initiation/elongation factors involved in protein translation [43]. 1p32.3 may also be deleted and contains the two target genes, *FAF1* and *CDKN2C. CDKN2C* is a cyclin-dependent kinase 4 inhibitor involved in negative regulation of the

cell cycle, whereas *FAF1* encodes a protein involved in initiation and/or enhancement of apoptosis through the Fas pathway. Homozygous deletion of 1p32.3 is associated with a poor prognosis in those receiving ASCT whereas in those receiving nonintensive treatment its prognostic impact is neutral [44].

Chromosome 13 deletion is observed in approximately 50% of myeloma cases and is commonly associated with nonhyperdiploid tumours [45-47]. In approximately 85% of cases, deletion of chromosome 13 constitutes a monosomy or loss of the q arm, whereas in the remaining 15% various interstitial deletions occur [45;46]. With this, the identification of key genes contributing to myeloma pathogenesis is challenging as often a level of gene function remains from the residual allele(s). Despite this, molecular studies have shown that the tumour suppressor gene *RB1* is significantly underexpressed in del(13/13q) and may therefore result in inferior negative cell cycle regulation [39]. Historically, del(13) has been associated with an unfavorable prognosis in MM, but there is now increasing evidence that its prognostic relevance may be related to its association with other genetic aberrations. Recent studies suggest that the prognostic significance of monosomy 13 may emanate from its close association with the t(4;14)(p16;q32) [48;49]. Fonseca *et al*. [49] suggest a crucial role for chromosome 13 deletions as a prerequisite for the clonal expansion of myeloma tumors.

The deletion of 17p13 in MM presumably leads to the loss of heterozygosity of *TP53*, a well-characterized tumor suppressor gene that transcriptionally regulates cell-cycle progression and apoptosis [50;51]. Deletions of chromosome 17p occur in ~10% of untreated MM tumors, and the prevalence increases with disease stage [49;52]. The deletion 17p13 has been identified clinically as an indicator of very poor prognosis [49;53;54].

Many other chromosomal deletions, focal copy number losses, and regions of LOH are seen in myeloma. Other important regions of LOH include 11q (the site of the *BIRC2* and *BIRC3* genes), 16q (the site of *CYLD*) and 14q32 (the site of *TRAF3*) [55;56]. All of these genes are involved in the nuclear factor-κB (NF-κB) pathway, indicating that up-regulation of NF-κB signalling is important in myeloma.

Mutations

Multiple myeloma (MM) is a genomically heterogeneous disease. Multiple mutations in different pathways collaborate to alter the biology of plasma cells, changing it in ways that lead to overt disease. Over the last 10 years, the use of high-throughput expression technologies has provided insights into the molecular basis of the disease individual patients. In particular, global gene and miRNA expression and genome-wide DNA profiling have been used alone or in an integrated fashion to investigate the genomic instability underlying the biological and clinical heterogeneity of MM [57-59]. However, although it is true that high-throughput profiling technologies have undoubtedly improved our understanding of the biology of MM, next-generation sequencing (NGS) has led to nothing less than a sea change in the field. Over the last year, a number of studies have shown that whole-genome (WGS) and whole-exome sequencing (WES) are cost effective means of identifying protein-coding mutations. Where the first 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 [60]. The NGS studies in MM are currently mainly based on WES and WGS. In addition to characterizing the recurrent alterations affecting MM, NGS analyses have made it possible to investigate genomic changes over time, and led to the detection of molecular alterations associated with malignant transformation and tumor progression, suggesting the alternation of multiple independent clones during the various disease phases from diagnosis to progression.

Over the last few years, NGS has extended the spectrum of simultaneously detectable alterations and provided a comprehensive picture of the disease's genetic landscape. The general scenario emerging from such studies indicates: a considerable number of nonsynonymous variants per patient; few recurrently mutated genes of probable pathogenetic significance and a heterogeneous subclonal structure at the time of diagnosis [20;22;43;61]. Two very recent studies ([20;61]) made a significant contribution to defining reference MM datasets by substantially increasing the number of cases that could be queried in order to determine the frequency of the somatic gene mutations revealed by NGS studies in larger cohorts (Figure 4).

Chapman *et al.* provided the earliest data [43] in a dataset of 38 MM tumors, 23 of which were profiled by means of WGS and 16 by means of WES (one by both). Their analysis showed the occurrence of approximately 35 WGS detected and 28 WES detected pointwise nonsynonymous mutations in each MM tumor that lead to intermediate behavior between that of genetically less-mutated leukemias (<10 nonsynonymous mutations per sample) [62] and that of more complex solid tumors (up to >500 nonsynonymous mutations) [63]. The high frequency of nonsynonymous variants in MM was confirmed by subsequent investigations, which also revealed the greater presence of nonsynonymous mutations (missense, nonsense or indel over silent ones (intronic, untranslated or synonimous) [20] and the progressive increase of non-silent mutations from asymptomatic to extramedullary disease (namely, from <10 to almost 60 median mutations across MGUS, high-risk smoldering MM, MM, and PCL) [26].

Study (year)	NGS approach	Patients (n)	Major mutations/findings, clonal portrait
Chapman *et al.* (2011)	WGS/WES	38	*KRAS* (26.5%), *NRAS* (23.5%), *TP53* (8%), *DIS3* (10.5%), *FAM46C* (13%); *BRAF* (4.5% of 161 additional cases tested); significant mutations in NF-κB pathway and *HOXA9*-regulating genes
Walker *et al.* (2012)	WES	22	*KRAS* (32%), *NRAS* (18%), *DIS3* (18%), *DNAH5* (13.5%); higher mutational load in t(4;14) than t(11;14); intratumor and intraclonal heterogeneity
Egan *et al.* (2012)	WGS	1†	Increasing number of somatic variants as the aggressiveness of the disease increases; hypothesized four divergent subclones from common ancestor
Egan *et al.* (2013)	WGS	1†	Truncating mutations in *CRBN, PSMG2* and *NR3C1* genes in drug-resistant disease
Leich *et al.* (2013)	WES	5 (and 6 HMCLs‡)	Multiple mutations in adhesion- and RTK-associated signaling molecules with interindividual pathway redundancy
Lohr *et al.* (2014)	WGS/WES	203	*KRAS* (23%), *NRAS* (20%), *DIS3* (11%), *FAM46C* (11%), *TP53* (8%), *BRAF* (6%), *TRAF3* (5%), *PRDM1* (5%); evidence that MM tumor samples may contain on average up to at least five subclones
Bolli *et al.* (2014)	WES	67	*KRAS* (25.5%), *NRAS* (25.5%), *TP53* (15%), *BRAF* (15%), *FAM46C* (12%); *TRAF3* and *DIS3* <3%; at least four patterns of clonal evolution, including no change over time, linear evolution, differential clonal response and branching evolution
Walker *et al.* (2014)	WGS/WES	28 (and 4 paired smoldering MM, and 4 HMCLs)	Intraclonal heterogeneity observed at all stages of MM from the MGUS stage; the transition from smoldering MM to MM is the result of the outgrowth of subclones present at earlier stage
Melchor *et al.* (2014)	WES	6	Linear and branching pattern of evolution shown at a single cell level. For the first time, clear evidence of parallel evolution in myeloma is shown at the single cell level

†The genome of the single patients has been sequenced at different stages of the disease.
‡Human Myeloma Cell Lines.
HMCL: Human myeloma cell line; MGUS: Monoclonal gammopathy of unknown significance; MM: Multiple myeloma; NGS: Next-generation sequencing; RTK: Tyrosine kinase receptor; WES: Whole-exome sequencing; WGS: Whole-genome sequencing.

Figure 4. Summary of the main publications involving next-generation sequencing analysis for variants identification and definition of disease clonality in multiple myeloma [64]

A further increased mutational load has been observed in myeloma cell lines [65]. 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 well-known markers of advanced disease in MM, and they had been already described as frequent by conventional sequencing [28]; overall, they occur in up to 50% MM patients if not more. *BRAF* mutations have been described in 4-14% patients [20;20;43;61]; the data on the percentage of *BRAF* mutated cases carrying V600E mutation of the gene, pathognomonic in B-cell hairy cell leukemia [66] range from about 30 [20] to 55% [61]. In addition, NGS studies further support previous evidence of recurrent (~11-13%) *FAM46C* mutations in MM [20;43;61]; although its function is still unknown, correlations have been found with the expression of ribosomal proteins, 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 that extends telomerically to the 1p32.3 region, associated with poorer prognosis [67]. *FAM46C* alterations (either incorporating loss-of-heterozigosity or mutations) were similarly correlated with impaired overall and progression-free survival [44]. Finally, NGS studies have shown that *DIS3*, a prevalently nuclear subunit of the human exosome complex whose catalytic activity is governed by the endoribonucleolytic (PIN) or 3′–5′

exoribonucleolyitic (RNB) domains, is also recurrently mutated in 11% of MM patients [20;43;61]. *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), and that *DIS3* function cannot be replaced by other homolog exosome-associated nucleases [68]. Undoubtedly, recurrent mutations in *DIS3* and *FAM46C* prompted for further investigation of altered RNA-processing machinery and protein homeostasis mechanisms in MM. Additional putative activating mutations that might be relevant in MM pathobiology have been described in *IRF4* gene, involved in sustaining survival in tumor PCs [69], in subset of genes (*HOXA9, MLL, MLL2, MLL3, UTX, WHSC1,* and *WHSC1L1*) involved in histone modifications [43], in *RB1* and *CYLD* [21;61].

Epigenetic changes

Although there has been substantial work on the genetics of myeloma, little is known about the epigenetic changes that lead to disease progression and their impact on treatment resistance [70]. DNA can be modified by methylation of cytosine residues in CpG dinucleotides; in addition, chromatin structure may be modified by histone modifications, such as methylation, acetylation, phosphorylation and ubiquitylation [71]. The most important epigenetic change that is relevant to the pathogenesis of myeloma is global DNA hypomethylation and gene-specific DNA hypermethylation during the transformation of MGUS to myeloma [72]. The most pronounced DNA methylation change is seen in the 15% of patients with the t(4;14) translocation, whose myeloma samples have increased gene-specific DNA hypermethylation compared with myeloma samples of other cytogenetic subgroups. This t(4;14) subgroup overexpresses *MMSET*, which encodes a histone methyltransferase transcriptional repressor [73-75] and its deregulation leads to global changes in histone modifications that promote cell survival, cell cycle progression and DNA repair [74]. Other chromatin modifiers are also deregulated in myeloma, including lysine-specific demethylase 6A (*KDM6A*; also known as *UTX*), *KDM6B*, mixed-lineage leukaemia (*MLL*) and homeobox A9 (*HOXA9*), but the full relevance of any resultant chromatin modifications needs further validation [43].

Micro-RNAs

MicroRNA (miRNA) genes encode a class of small RNAs (17–25 base pairs) which function to regulate the translation of other proteins by forming complementary base parings to specific mRNA transcripts. Studies have shown that miRNAs can act as both tumour suppressors and oncogenes in a range of cancers and that their transcriptional control is regulated by promoter methylation changes [76]. A substantial amount of work has been completed to investigate which miRNAs are differentially expressed in myeloma [77;78], and it has been shown that miRNA changes can deregulate genes and pathways relevant to myeloma pathogenesis including cell cycle progression, *TP53* and *MYC* [79-81].

Treatment

The treatment of MM has so far been unrelated to any particular molecular or genetic feature. The cornerstones of conventional anti-myeloma therapy have long been alkylators and corticosteroids, more specifically oral melphalan and prednisone (MP), leading to a median post-treatment overall survival (OS) of 3-4 years. Subsequently, high-dose therapy followed by autologous stem cell transplantation (ASCT) has been introduced into standard practice as it prolongs survival in comparison with conventional chemotherapy (median OS 5–7 years) [82] [83]. However, over the last few years, the treatment of MM has changed considerably as a result of drug development [82]: in particular, the immunomodulatory drugs (IMiDs) thalidomide [84] and lenalidomide [3] [85], and the proteasome inhibitor bortezomib [86;87], are effective in treating newly diagnosed or relapsed myeloma [83], and have further improved overall and event-free survival. However, there is still a lack of effective therapies targeting the deregulated biological pathways specifically associated with the disease.

PLASMA CELL LEUKEMIA

Definition and epidemiology

Plasma cell leukemia (PCL) is the most aggressive presentation of the PC neoplasms and is characterized by circulating plasma cells > 2×10^9/L in peripheral blood or by a relative plasmacytosis > 20% of blood leukocytes [88;89]. While primary disease (pPCL) presents as de novo leukemia, secondary leukemic transformation (sPCL) arises from pre-existing multiple myeloma (MM) [90;91] probably as a consequence of clonal transformation. PCL is rare. sPCL occurs in only 1.8–4% of MM patients and pPCL occurs with a comparable incidence. Consequently, few large studies of PCL patients have been reported [90-92] and the molecular basis of PCL remains poorly understood. Around 60% of patients presented a primary form and the remainder 40% occurred as complication during the course of MM [93]. The median age of patients with pPCL ranged between 52 and 65 years, about 10 years younger than the median age of 65 to 70 years observed in the general myeloma population [2] and in secondary PCL [52].

Presenting clinical features

Primary PCL has a more aggressive clinical presentation than MM including a higher tumor burden. Patients may present with symptoms due to profound anemia, hypercalcemia and thrombocytopenia. On physical examination, patients may exhibit a higher prevalence of organomegaly with involvement of the liver, spleen, lymph nodes, central nervous system and palpable extramedullary soft-tissue plasmacytomas. In contrast, the presence of lytic bone lesions is lower than that observed in MM [52]. Also, reflecting this aggressive clinical presentation, a higher proportion of patients with primary PCL have significant leukocytosis, as well as elevated lactate dehydrogenase (LDH) and β2-microglobulin serum levels.

Genetics of PCL

Cytogenetic studies show that plasma cells in primary PCL have a number of genetic abnormalities.

Ploidy status

More than 80% of patients with PCL have hypodiploid or diploid cells which is associated with poor prognosis whereas about 60% of patients with MM display hyperdiploidy, a favorable finding [94].

IGH translocations

Chromosome 14q32 translocations are found in a great number of pPCL and sPCL patients (82–87%), as would be expected in NH-MM cases. Based on the study by Tiedemann and colleagues [52], in pPCL IgH translocations almost exclusively involve 11q13 (*CCND*1), supporting a central etiological role, while in sPCL multiple partner oncogenes are involved, including 11q13, 4p16 (*FGFR3*/*MMSET*) and

16q23 (*MAF*), recapitulating MM [52]. Genomic aberrations such as t(4;14), del(13q14), del(17p), del(1p21) and 1q21 gains are adverse risk factors in MM but their significance in PCL is unclear (as they usually indicate more aggressive behavior which is ubiquitous in PCL) [54;59;95;96].

Chromosome 13

In a series of 26 patients Garcia-Sanz reported deletion of chromosome 13 in 84% of patients with PCL versus 26% for those with MM [94]. In the Mayo study loss of 13q by FISH was found very common in pPCL (85%), more so than in MM (54%) (P = 0.02); but with no difference in prevalence between pPCL and sPCL [52]. As would be expected in NH-MM Avet-Loiseau reported a high frequency of monosomy 13 (85%) [92].

Deletions of 17p13 and TP53 inactivation

Previous studies reported 17p13.1 deletion in MM as a late event found only in 10% of patients [53], and *TP53* coding mutations in only 3% [97]. In the group of 80 patients reported by Tiedemann there was a prevalence of deletion of 17p13.1, causing allelic loss of *TP53*, in 50% of pPCL and a remarkable 75% of sPCL. Moreover, functionally relevant *TP53* coding mutations were identified in 24% of PCL patients tested, contributing to a substantial prevalence of allelic *TP53* inactivation (by mutation or deletion) of 56% in pPCL and 83% in sPCL. Eleven percent of pPCL and 33% of sPCL tumors showed biallelic inactivation of *TP53* with simultaneous allelic deletion and mutation [52]. Interestingly, monoallelic or biallelic inactivation of *TP53* did not correlate significantly with survival in sPCL, unlike MM, where del(17p13.1) predicts adverse survival [98]. Lack of correlation between *TP53* status and survival may reflect ubiquitous targeting of the *TP53* pathway in sPCL.

MYC abnormalities and RAS mutation

Rearrangement of *MYC* was identified by 3'FISH break apart in 33% of pPCL and sPCL tumors and was complemented by *MYC* amplification or 5'*MYC* translocations in 8 and 17% of patients, respectively [52]. *MYC* rearrangements were associated with a trend toward inferior OS in pPCL (median OS of 8.6 versus 27.8 months without rearrangements, P = 0.006). On the other hand, mutations of *K-RAS* or *N-RAS* at codons 12, 13 or 61, previously characterized as functionally activating [99;100] [101], were found in 27% pPCL and 15% sPCL [52]. Activating mutations of *RAS* were associated with a trend toward poorer outcome in pPCL (P = 0.069). However, the prevalence of *K-* or *N-RAS* mutation in sPCL is comparable to that reported in MM (21%) [101], suggesting little, if any, selective pressure for *RAS* activation in secondary leukemic transformation from MM.

Survival

The survival of patients with pPCL is short. In seven series, historically median survival, without novel therapies, has ranged from 6.8 to 12.6 months [52;90;91;94]. Furthermore, the median survival of 231

patients from a recently published epidemiology study was only 4 months [102]. Of note, the survival rate at 5 years from diagnosis is less than 10% in all series.

The best survival data, incorporating hematopoietic stem-cell transplantation, reported a median survival longer than 3 years [103]. Unfortunately, the significant improvement in survival observed in MM in the past decade has not been seen in PCL [102].

These discouraging survival results in primary PCL are due to two facts: 1) its aggressive presentation with severe complications leading to early death within the first months from diagnosis, and 2) the lack of effective therapy to achieve sustained responses. Early mortality is still of concern and reflects the aggressiveness of the disease [92].

Treatment of PCL

Treatment of both pPCL and sPCL, as in MM, is aimed at prolonging survival and maximizing quality of life, as there are no recognized curative regimens [104]. In general, patients are treated with aggressive induction therapy followed by ASCT in those who are appropriate candidates for this approach. Chemotherapy alone is the principal option for those ineligible for ASCT. The best induction regimen for PCL is not known and there is great variability in clinical practice. Several combinations that include an alkylating agent have been used as sole treatments of PCL or as induction therapy prior to anticipated transplant. Combinations that include melphalan are generally avoided as induction therapy in potential transplant candidates to allow adequate stem cell collection. Most patients need to be treated with the most effective agents currently available [105;106] and historic data with older agents show poor rates of control.

DIS3: A KEY COMPONENT OF THE EXOSOME COMPLEX

DIS3 gene encodes a highly conserved ribonuclease which serves as the catalytic component of the exosome complex involved in regulating the processing and abundance of all RNA species. *DIS3* gene map on chromosome 13q22.1 and it is composed by 22 exons.

The multisubunit RNA exosome complex is a major ribonuclease of eukaryotic cells that participates in the processing, quality control and degradation of virtually all classes of RNA in Eukaryota [107].

The exosome structure

The eukaryotic exosome is composed of ten to eleven subunits, which can be divided into two major groups with regard to their structural and functional contributions. The first group encompasses nine proteins that form the 'exosome core'. They are all rather small polypeptides with molecular masses of 20–50 kDa. Three of these are built entirely of RNA binding domains and motifs, whereas the remaining six are single-domain proteins homologous to RNase PH, a homomultimeric enzyme found in Escherichia coli. The second group of exosome components comprises two subunits, *DIS3* and Rrp6, which associate with the core in different combinations, giving rise to isoforms of the complex. *DIS3* and Rrp6 are both large multi-domain polypeptides with molecular masses around 100 kDa. They are quite different from each other and carry all catalytic activity of a typical eukaryotic exosome. The repertoire of catalytic subunits shows considerable diversity in different groups of Eukaryota, much more than the core subunits. The schematic structure of the eukaryotic RNA exosome is reported in Figure 5 [107].

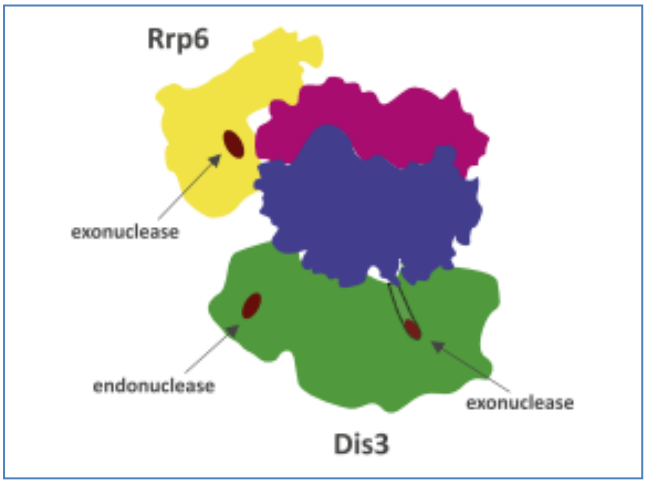

Figure 5. Side view of the eukaryotic RNA exosome with the hexamer in blue, the cap in magenta and the associated ribonucleolytic subunits in green (*DIS3*) and yellow (Rrp6). Endo- and exonuclease active sites are denoted in deep red [107]

The exosome core

The structure of the exosome core is well established, owing to the crystal structure obtained for the Homo sapiens complex [108]. Its nine subunits are arranged in a two-layered ring; the bottom layer is formed by six subunits homologous to bacterial RNase PH: Rrp41, Rrp42, Rrp43, Rrp45, Rrp46 and

Mtr3, and referred to as the 'hexamer', while the top layer consists of the three RNA binding subunits, Rrp4, Rrp40 and Csl4, and is called the 'cap' (Figure 6). Rrp4 and Rrp40 each contain one S1 and one KH RNA binding domain, and Csl4 consists of a KH domain and a zinc ribbon RNA binding motif. The cap subunits do not form contacts with one another but instead each is firmly attached to two subunits of the hexamer [108]. The central channel that traverses the ring is only wide enough to accommodate single-stranded RNA (ssRNA).

The architecture of the exosome core is ancient and perhaps even predates the exosome itself as two complexes of prokaryotic origin have the exact same domain composition and spatial arrangement. One is the exosome found in most Archaea, sometimes called the 'exosome-like complex' [109;110] and the other is polynucleotide phosphorylase (PNPase) found in Eubacteria as well as in plant and vertebrate mitochondria [111].

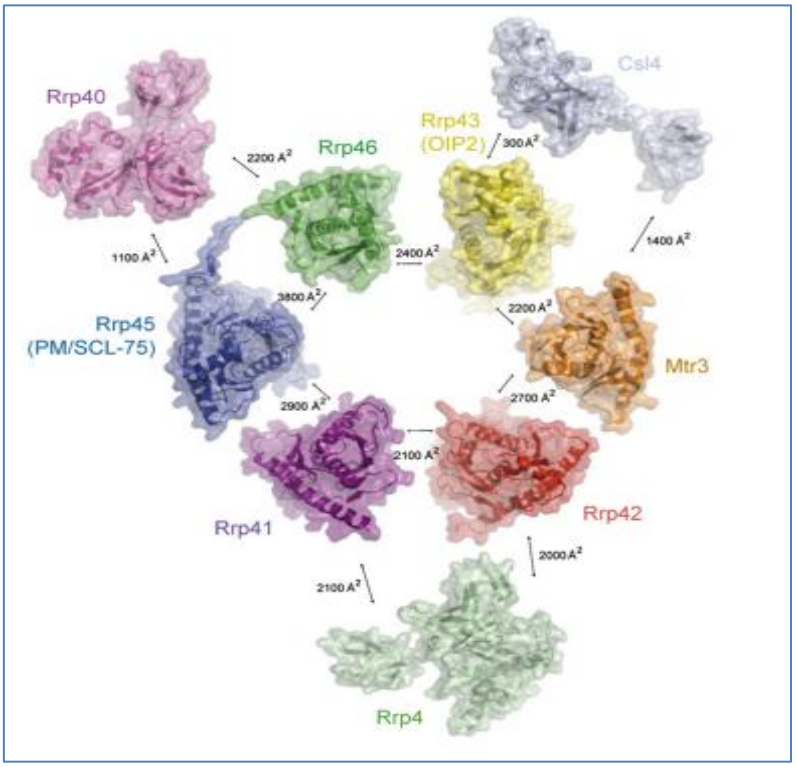

Figure 6. Architecture and subunit Interactions within the human exosome [108]

The exosome ribonucleases

DIS3

DIS3 proteins are homologous to RNase II and RNase R of E. coli, which makes them members of the RNR superfamily [112]. The RNR homology region of *DIS3* proteins has the typical domain layout: two cold-shock domains (CSDs), an RNB domain and an S1 domain. CSD and S1 are ancient folds that non-specifically bind RNA, while RNB is a catalytic domain endowed with processive exoribonuclease activity. A distinguishing feature of the *DIS3* family is a CR3 motif and a PIN domain located N-

terminally to the RNR region [113-115]. The PIN active site consists of four acidic amino acid residues but, unlike in RNB, they are dispersed throughout the otherwise weakly conserved sequence of the domain. The CR3 motif is named for its three conserved cysteine residues, but it also contains a conserved histidine residue. Although details are not clear, the motif also plays a role in the core–*DIS3* interaction and is required for some of the cellular functions of the exosome [113]. *DIS3* is an endoribonuclease and a 3'–5' exoribonuclease. The endonuclease activity resides in the PIN domain, the active site of which is exposed to the solvent and consists of four acidic amino acid residues that coordinate two divalent metal cations (preferably manganese) [114-116]. The enzyme cannot cleave dsRNA but both circular and linear ssRNA are eligible substrates. A preference for substrates with phosphorylated 5' termini has been reported [116]. The exonuclease activity of *DIS3* depends on four conserved aspartate residues that coordinate two magnesium cations to hydrolytically cleave the RNA backbone. Mutation of any one of these residues completely abolishes the activity [108;117]. The active site is located in the RNB domain, buried at the bottom of a narrow channel and can only be reached by ssRNA at least 7 nt long [118]. *DIS3* exonucleases are highly processive. They attack 3' ends of ssRNA and once bound they degrade substrate completely, releasing nucleoside 5'-monophosphates and leaving end products a few nucleotides long (this length varies slightly from one homologue to another)[119].

The human genome encodes three *DIS3* homologues, of which only h*DIS3* and h*DIS3*L were found to associate with the exosome [119;120]. Notably, both of them are processive 3'–5' hydrolytic exonucleases, whereas only h*DIS3* has also retained endonuclease activity in its PIN domain. In vivo localization studies and analyses of substrate specificities revealed that h*DIS3*L is restricted to the cytoplasmic exosome, whereas h*DIS3* is mainly a nucleoplasmic protein, with a small fraction present in the cytoplasm [119;120].

The third human *DIS3* homologue —h*DIS3*L2— does not interact with the exosome core, due to the lack of PIN domain and it has been recently demonstrated by several independent studies that it is responsible for an alternative 3'–5' RNA decay pathway in the cytoplasm [121;122]. An important role of h*DIS3*L2 in the maintenance of proper cellular metabolism is supported by the fact that mutations in its gene were found to be associated with Perlman syndrome of overgrowth and predisposition to Wilm's tumor development [123].

Rrp6 ribonucleases

Rrp6 proteins belong to the large DEDD superfamily of nucleases, which includes enzymes of various functions, e.g. oligoribonuclease and proofreading domains of DNA polymerases [112]. Rrp6 proteins are distributive 3'–5' exoribonucleases that cleave RNA hydrolytically and release nucleoside 5'-monophosphates [124]. Activity comes from the conserved DEDD domain, whose active site is solvent exposed [125;126] and consists of four conserved acidic amino acid residues coordinating two magnesium cations. The combination of a distributive mode of action with selectivity for ssRNA should preclude digestion of structured substrates. Surprisingly, Rrp6 proteins are able to break down structured

substrates provided there is a single stranded region upstream of the structure [125]. This effect is slightly stronger for the human protein than for the yeast one. The mechanistic background for this activity is unknown.

Isoforms of the exosome

The exosome core displays little structural variability across Eukaryota. All core subunits have the same subcellular localization in *S. cerevisiae* and *H. sapiens* [119;120]. They are found in the cytoplasm and the nucleus, more concentrated in the latter, and enriched in nucleoli. In contrast to core subunits, some of the associated subunits are compartment-specific and give rise to different isoforms of the exosome. In *S. cerevisiae DIS3*p has the same localization pattern as the core, whereas Rrp6p is restricted to the nucleus [124]. This implies that the *S. cerevisiae* exosome exists in two isoforms: a ten-subunit cytoplasmic one (core +*DIS3*p) and an eleven-subunit nuclear one (core + Rrp6p +*DIS3*p).

The *H. sapiens* genome contains three homologues of the yeast *DIS3* gene. However, since one of them (*DIS3*L2) does not encode a full PIN domain, the *DIS3* protein family only has two human members: *DIS3* and *DIS3*-like (*DIS3*L). *DIS3* has greater homology to yeast *DIS3*p than *DIS3*L does, but the domain composition is the same in both paralogues [119;120].

Moreover, the subcellular localization of the two human paralogues is radically different: while *DIS3*L is strictly cytoplasmic, *DIS3* is largely limited to the nucleus with only a small fraction in the cytoplasm. Notably, while the core subunits adhere to the usual subcellular distribution of exosome proteins, *DIS3* is absent from nucleoli[119].

As in *S. cerevisiae*, human RPR6 is nuclear with the typical nucleolar enrichment. In addition, small quantities are also present in the cytoplasm [119].

Taken together, three major exosome isoforms appear to exist in human cells: core+RRP6 in the nucleolus, core+RRP6+*DIS3* in the nucleoplasm and core+*DIS3*L in the cytoplasm. Other possible, less abundant isoforms are: nucleoplasmic core+*DIS3* and cytoplasmic core+*DIS3*, core+RRP6+*DIS3* and core+RRP6+*DIS3*L and core+RRP6.

Specific functions of the RNA exosome

The exosome plays a fundamental role in most RNA metabolic pathways [114;127] (Figure 7).

These pathways include turnover of normal mRNAs and AU-rich element-regulated decay of unstable mRNAs [128;129]; nuclear processing of stable RNA classes (small nuclear RNA (snRNA), small nucleolar RNA (snoRNA), rRNA, tRNA)[130;131]; degradation of unstable transcripts arising from intergenic regions [122] and complete degradation of mRNA after endonucleolytic cleavage initiated by small interfering RNA (siRNA) in RNA interference [132].

Furthermore, the exosome plays a crucial role in the quality control, as it is a primary enzyme degrading unwanted molecules in the nucleus and the cytoplasm, including incorrectly processed pre-mRNAs, rRNAs and tRNAs as well as translationally incompetent mRNAs [130;131].

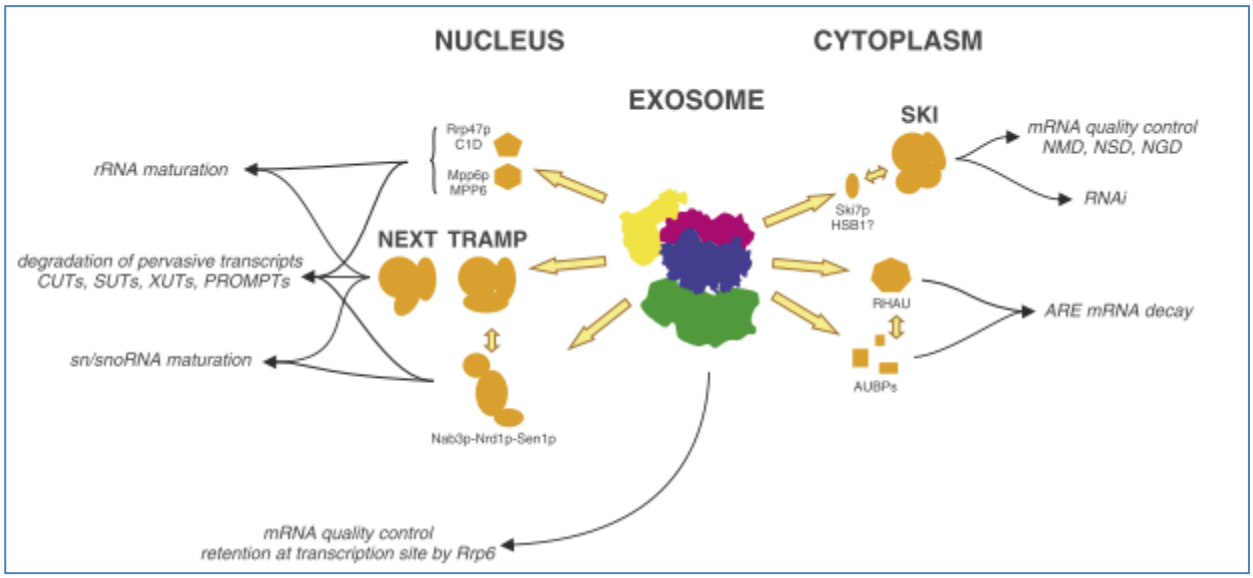

Figure 7. Map of the exosome's cellular functions [107]

Apart from its well-documented role in RNA metabolism, a recent study has implicated the exosome in the activation-induced immunoglobulin heavy-chain class switch recombination and immunoglobulin variable region somatic hypermutation in human B lymphocytes [133]. It was also suggested that the exosome is essential for activation-induced cytidine deaminase (AID) recruitment to chromatin, but the mechanistic details of this process remain elusive [133].

Impact of *DIS3* mutations in multiple myeloma

Recent next generation sequencing (NGS) studies observed an increase of single nucleotide variants (SNVs) and a change in the genetic landscape of multiple myeloma [20;21;26]. However, specific gene mutations occurring in a significant percentage of cases are rare events in MM.

One of the novel findings of the first myeloma genome publication was the identification of *DIS3* mutations in 11% of myeloma samples. Chapman *et al*. provided the earliest data [43] in a dataset of 38 MM tumors, showing an occurrence of *DIS3* mutations of 10.5%. The four mutations observed by Chapman *et al* occur at highly conserved regions and cluster within the RNB domain facing the enzyme's catalytic pocket (Figure 8).

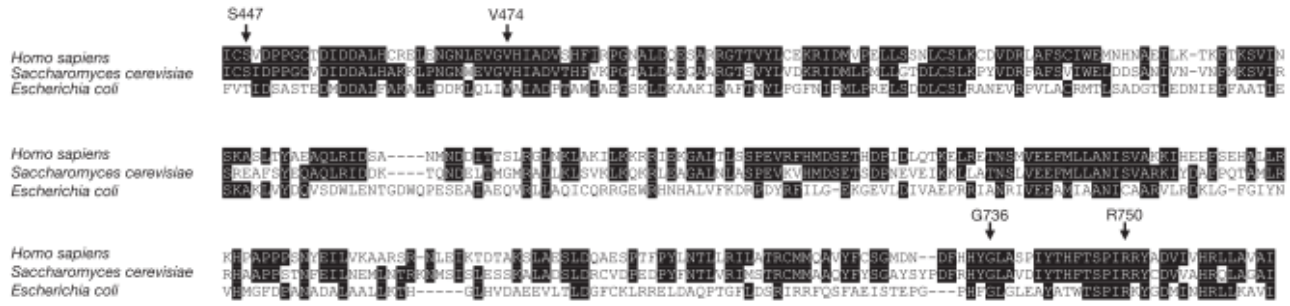

Figure 8. Alignment of human, yeast and bacterial RNB domain of DIS3. Positions of observed mutations are indicated with respect to the human sequence [43]

Two lines of evidence indicate that the *DIS3* mutations result in loss of function. First, three of the four tumours with mutations exhibited loss of heterozygosity via deletion of the remaining *DIS3* allele. Second, two of the mutations have been functionally characterized in yeast and bacteria, where they result in loss of enzymatic activity leading to the accumulation of their RNA targets [134]. Given that a key role of the exosome is the regulation of the available pool of mRNAs available for translation [135], these results indicate that *DIS3* mutations may dysregulate protein translation as an oncogenic mechanism in multiple myeloma.

The presence of mutations in *DIS3* was subsequently confirmed by Walker *et al* [26], that performed NGS analysis on 22 MM patients identifying 4 mutations in this gene. They found that the proportion of affected samples was higher (18%) than Chapman *et al* [43]. This difference was due of the samples analyzed, because *DIS3* mutations have been found exclusively in t(4;14) or t(11;14) samples, of which their dataset was composed.

A further study [61] reported that *DIS3* gene is subject to point mutations with LOH, strongly implicating *DIS3* as a tumor suppressor gene in 11% of MM patients (Figure 9). Interestingly, they reported that *DIS3* aberrations were more commonly seen among nonhyperdiploid MM cases compared to the hyperdiploid cases.

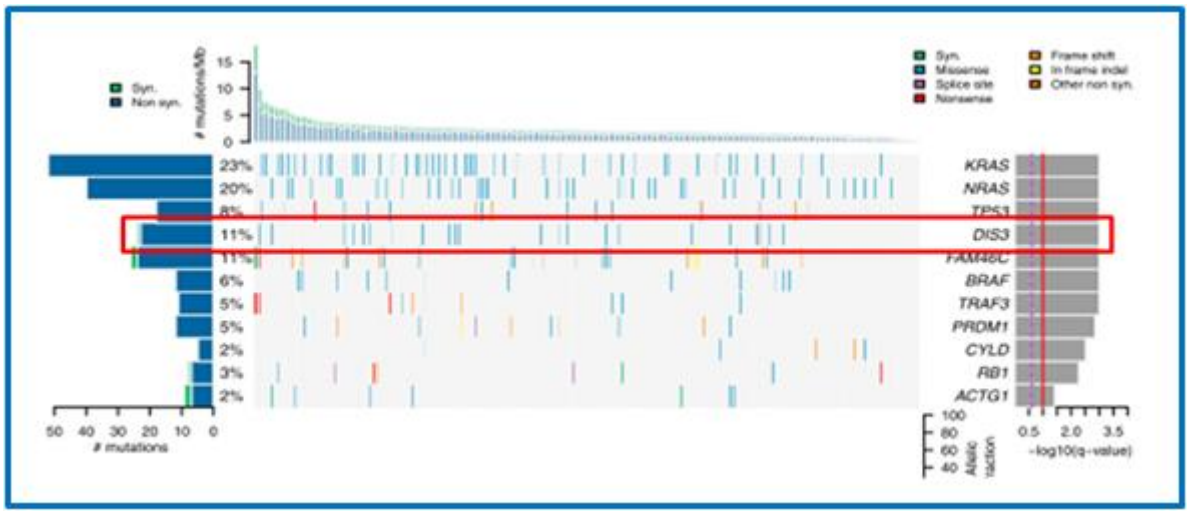

Figure 9. Significantly Mutated Genes in 203 Patients with MM [61]

Bolli *et al* [20] studied by NGS a large cohort of myeloma patients, with overrepresentation of advanced disease, hyperdiploid and del(17p) cases and underrepresentation of t(11;14) and t(4:14). This may explain the low frequency of *DIS3* mutations (1.5%) found, mostly present in IGH rearranged cases [43;136].

Finally, Weissbach *et al* [137]performed deep amplicon sequencing of entire coding region of *DIS3* in a uniformly-treated cohort of 81 MM patients. They detected a frequency of *DIS3* mutations of 11%, in agreement with previous sequencing studies that did not select for cases with IGH-translocations [43;61].

The majority of mutations found in MM affect conserved residues in the h*DIS3* RNB domain, suggesting their impact on the exonucleolytic activity of the enzyme [68] (Figure 10). In contrast, point mutations in the PIN domain appear to have only low effects on cell growth [68], although the PIN domain possesses endonucleolytic function and is necessary as a structural element for the association of the exosome in yeast [68;115;116].

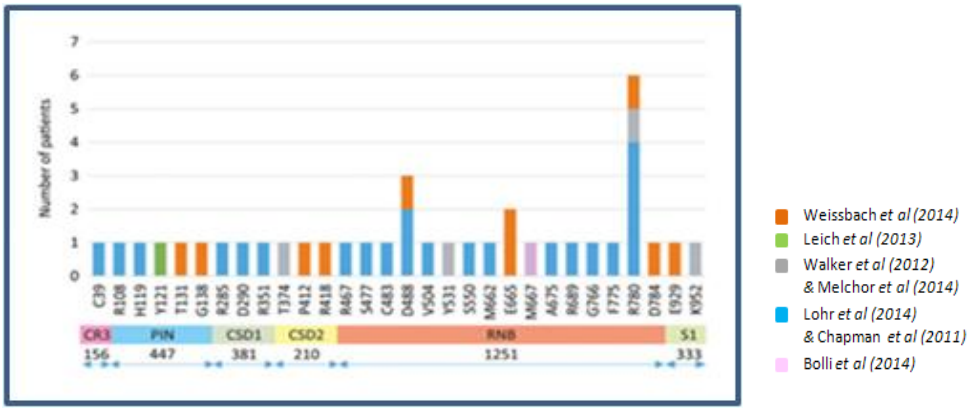

Figure 10. Frequency of SNVs in domain of *DIS3* from the literature [137].

Experimental results from the yeast model made by Tomecki *et al* [68], revealed the importance of *DIS3* and intactness of its RNB domain for RNA metabolism, protein function and cell physiology, leading to vulnerabilities that can be used for anti-cancer drug development.

Among the molecular phenotypes caused by MM mutations, there are strong effects on the ability of the enzyme to exonucleolytically digest RNA substrates, leading to increased levels of extended 5.8S precursor molecules, improperly processed at the 3'-end, and a decreased degradation rate of RNA molecules belonging to other RNA classes, such as tRNAs and their precursors, snRNAs and other small RNAs, including RNase P RNA and 7SL RNA. Moreover, experimental results from *DIS3* mutations in the yeast model corresponding to those found in MM patients, cause cell growth defects and molecular phenotypes strongly suggesting impaired exosome function [68].

Next, Tomecki *et al* [68] checking whether, as in yeast, there is a synergistic effect of MM-associated RNB domain mutations and dysfunction of h*DIS3* PIN domain endonuclease activity: the metabolic activity of the cells bearing mutations in the RNB domain was also additionally reduced on accompanying inactivation of PIN domain endonuclease, while the inactivation of the PIN domain catalytic activity alone did not lead to accumulation of 5.8S rRNA precursors.

In conclusion, although the molecular events linking the inactivation of h*DIS3* nucleolytic properties and its inability to efficiently degrade transcripts with reduced cell viability remain to be elucidated, the negative synthetic interaction between MM-associated RNB domain mutations and PIN domain dysfunction may suggest the latter as a promising drug target for cancers bearing mutations that affect h*DIS3* RNB domain exonucleolytic activity.

Involvement of *DIS3* gene in others cancers

DIS3 gene mutations were found in global screens not only in multiple myeloma patients, but also in other type of cancers, such as adenoid cystic carcinoma [138] and colorectal cancer [139]. Recently, Ding *et al* [25] reported that *DIS3* is among the most recurrently mutated genes in AML. Additionally, h*DIS3* was identified in transcriptomic analyses as one of the genes, whose expression differentiates superficial spreading melanoma from nodular melanoma [140]. *DIS3* overexpression was earlier observed in human colorectal cancer and in a mouse model of this cancer, where elevated levels of respective mRNA and protein positively correlated with the incidence of metastasis [141]. Expression profiling revealed that h*DIS3* is among several genes whose loss-of-function significantly reduces viability of colorectal carcinoma cell lines [141]. Increased levels of h*DIS3* mRNA have been also recently proposed as one of the characteristics of the epithelial ovarian cancer [142].

All examples presented above strongly suggest the existence of possible molecular link between h*DIS3* functions and development of different cancers [143]. More specifically, it appears likely that exonucleolytic activity of h*DIS3* protein —the major catalytic subunit of the exosome— might be somehow involved in this association.

Aim of the study

To gain more insights into the molecular relevance of *DIS3* mutations in MM pathogenesis and progression, we performed deep sequencing of the PIN (exons 1-4) and RNB (exons 10-18) functional domains in a retrospective cohort of 130 cases with MM at onset, 17 at relapse (of whom 15 were also tested at onset), 24 patients with primary PC leukemia (pPCL), 12 with secondary PCL and 20 multiple myeloma cell lines. Moreover, we correlated the overall *DIS3* mutation status in relation to the most recurrent chromosomal abnormalities (as assessed by fluorescence *in situ* hybridization) as well as the occurrence of *DIS3* mutations in major and minor subclones. Additionally, to investigate the state of *DIS3* mutation longitudinally, we analyzed samples of 20 patients for whom bone marrow specimens were collected at two different time points. We also performed deep sequencing of cDNA of the *DIS3*-mutated cases to verify if the mutations detected on genomic DNA are also expressed at transcriptomic level. *DIS3* mutation was also integrated with gene expression profiles in MM samples in order to identify transcriptomic profiles possibly related to *DIS3* mutations. Finally, in the 16pPCL patients for whom the follow-up was available, the association of *DIS3* mutations with clinical outcome was tested.

Materials and Methods

Patients and cell lines

Pathological bone marrow specimens from 130 cases with MM at onset (66 males, 64 females; median age 67 years, range 42–85 years), 17 at relapse (8 males, 9 females; median age 67 years, range 47–79 years) and 12 MM patients at leukemic transformation (4 males, 8 females; median age 69 years, range 54–82 years) were obtained during standard diagnostic procedures after the patients had given their informed consent according to institutional guidelines. Many of the patients included in this study have been described in our previous reports [58;144]. Additionally, we analyzed 24 primary plasma cell leukemia (pPCL) cases at onset, of whom eight (4 males, 3 females; median age 57 years, range 50–72 years) admitted to our hematology service and 16 included in a multicenter clinical trial (RV-PCL-PI-350, EudraCT N°2008-003246-28) aimed at evaluating safety and anti-tumor activity of combined lenalidomide and dexamethasone as first-line treatment [145].

Seventy-eight patients had an immunoglobulin (Ig) G protein monoclonal component, 33 IgA and one IgG/IgA protein; 48 cases expressed λ light chain and 78 κ chain. The diagnoses were made by means of conventional morphology and immunophenotype analyses and the patients were clinically staged according to previously described criteria [146]. No conventional cytogenetic (G-banding) analyses were available. Plasma cells were purified from the bone marrow samples as previously described [147;148]. The purity of the selected PC populations was assessed by means of morphology and flow cytometry, and was >90% in all cases.

Human myeloma cell lines (HMCLs) were purchased from DSMZ-German Collection of Microorganisms and cell Culture, Germany (NCI-H929, OPM2, JJN3,and KMS-12), kindly provided by Dr. T. Otsuki, Kawasaki Medical School, Okayama, Japan (KMS-28, KMS-34, KMS-18,KMS-11, KMS-26 and KMS-27), Dr S. Iida, Nagoya City University Graduate School of Medical Sciences, Nagoya, Japan(AMO1), Dr. G. Tonon, San Raffaele Scientific Institute, Milan (RPMI 8226, delta-47, SK-MM-1 and UTMC-2), Dr. F. Malavasi, Department of Genetics, University of Torino, Italy (LP-1), and Dr. P.F. Tassone, Magna Graecia University, Catanzaro, Italy (MM.1S and U266), or established in our own laboratory (CMA-03 and CMA-03/06) [149].

FISH analyses

All samples were characterized by fluorescence *in situ* hybridization (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/*CCND*1 (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/).

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). All genomic DNA was stored at −20°C until use.

Mutation analyses

After DNA extraction and quantification, all samples were subjected to 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 (spanning exons from 1to 4 for the PIN domain and exons from 10 to 18 for the RNB domain; RefSeqNM_014953.4) containing genome-specific sequences (reported in Table 3), and an 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 PCR round (Figure 10).

exon	Primer FWD	Primer REV
ex_1	5'-GAAAGGGAAGAACCTCCGGG-3'	5'- TCCCTGTCATACCCCCTTGT-3'
ex_2	5'- GCTTCTTGGCTTAACTTATTCAGTG-3'	5'- ACAGGAACCCTCTCCCGAA-3'
ex_3	5'- TGCTAAGAGTTTTCACATATCCTTG-3'	5'- TCACATGAAGTTATATAGGACTACGA-3'
ex_4	5'- AGGCTGTAGTGATGTGTAATTGC-3'	5'- GCTTACCCACCGACATTCC-3'
ex_10	5'- TATGTTGTAGTTGTGCTTTGGAAAT-3'	5'- CAATATGCTTGACTGGGTAAATGTA-3'
ex_11	5'- TTTGCTTGTTAATTACTCTTGTGAAG -3'	5'- TTCATGGCCGTTTAAGAATC-3'
ex_12-13	5'- TTATGGCTAAGTAATCTGTGGTTCTA -3'	5'-AAATTAGAGATTAATAGCCATGAAACG-3'
ex_14	5'- GTAGTGAAAGTAGGAGGACATATTG-3'	5'- GAAGCTAGAAGAAACAGGAGTCT-3'
ex_15	5'- ACACTTGCTGTAGTCATTGTCTT-3'	5'- GCAAGCCAAATAAAGTAGAAATCAT-3'
ex_16	5'- GCGGAGTAACTGAGAGATGAAAG-3'	5'- CAGGTAGATCAAAACACAAATAGATG-3'
ex_17	5'- GCCGAATCTCCTACTTTTCCA-3'	5'- CCAAAAGCCGATGAACAATGA-3'
ex_18	5'- CTTAGCGTCTCCAATATACACACA-3'	5'- CTAGCAGTATCGACAAAAGGCA-3'

Table 3. Genome-specific primers for amplicon library preparation.

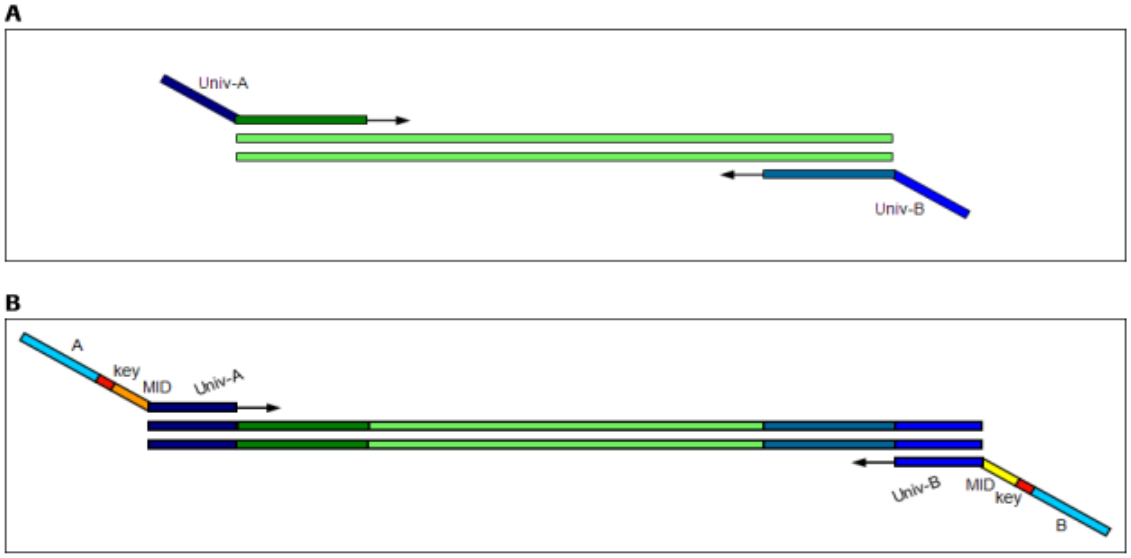

Figure 10. Schematic representation of the PCR reaction components for the preparation of a two step Universal Tailed Amplicon Library. A) First step, targeting the template-specific sequences and adding the universal tail. B) Second step, targeting the universal tail and adding the 454 Sequencing System primers and MIDs.

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 HiFi Polymerase (Roche Diagnostics). 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 exon 1), 57°C (for exons 10, 11 and 17), at 58°C (for the remaining exons) (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) and the Victor X2 (Perkin Elmer) 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 (RefSeqGRCh37.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. NGS analysis was repeated 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). Additional information about the functional relevance of the detected SNVs were obtained 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).

Validation of *DIS3* variants at transcriptomic level

To verify the occurrence of *DIS3* variants at transcriptomic level, total RNA of mutated samples was converted to cDNA using M-MLV reverse transcriptase (Invitrogen) and random examers, and subjected to deep sequencing of the exon harboring the variant/s detected on genomic DNA. Sequence-specific exonic PCR primers were designed in the Primer 3 program (http://frodo.wi.mit.edu/primer3/) and are as follows:

amplicon	Primer FWD	Primer REV
ampl_A	5'-CCGGGGTTAGGCGTATTCTA -3'	5'- AAGTGAGAAATCGCAGTGCC -3'
ampl_B	5'- TACTTGCTGCCCGACACTAA -3'	5'- GGAATACCAGCTTTCACTTGTG -3'
ampl_C	5'- ATGTTCCCCATCAGCCTTTT -3'	5'- AAAATGTGACGTGGACAGGC -3'
ampl_D	5'- GGGAAATGAATCACAATGCTGA -3'	5'- TCTGAACATGCTCTGCTTCG -3'
ampl_E	5'- AGTCTTTGGATCAGGCCGAA -3'	5'- GGCTATTGGGGCTGACTGTA -3'
ampl_F	5'- ACTATGGCTTAGCGTCTCCA -3'	5'- AACGTGCATCAGTGGCTTTT-3'

Gene expression profiling

102 MM samples underwent gene expression profiling (GEP) analysis. 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 tothe manufacturer's protocols. The raw intensity expression values were processed by Robust Multi-array Average procedure, with the re-annotated Chip Definition Files from BrainArray libraries version 18.0.0 [150], 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)[151] as previously described [152]. The cutoff point for statistical significance (q-value <10%) was determined by tuning the Δ parameter on the false discovery rate (FDR) and controlling the q-value of the selected probes. Differentially expressed genes were also evaluated at the highest stringency level (median FDR 0%, 90th perc 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.

Statistical analysis

All contingency analyses were performed by Fisher's exact test. A P value <0.05 was considered significant for all statistical calculations.

For survival analyses in pPCL, patients were stratified into two groups according to the occurrence of mutations in *DIS3*. The groups identified by this approach were tested for association with progression free survival (PFS) or overall survival (OS) using the log-rank test and the curves estimated by Kaplan–Meier method; P values were calculated according to the standard normal asymptotic distribution. Survival analysis was conducted with the km.coxph.plot function of the survcomp package in R software.

Results

Multiple myeloma and plasma cell leukemia samples

To estimate the frequency of *DIS3* mutation in plasma cell dyscrasia, we investigated by next generation sequencing (NGS) a retrospective cohort of 130 cases with MM at onset and 17 at relapse (of whom 15 were also tested at onset). Moreover, we examined 24 patients with primary PC leukemia (pPCL), 12 with secondary PCL and 20 multiple myeloma cell lines (see Materials and Methods for patient populations).

Pathological samples from all patients (n=187) were characterized by fluorescence in situ hybridization (FISH) for the presence of the main molecular alterations, namely IGH@ chromosomal translocations and hyperdidploidy status, del(13)(q14), del(17)(p13), 1p33 (CDKN2C) loss and 1q21.3 (CKS1B) gain aberrations as previously described [153;154]. 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).

Deletion of chromosome 13q was investigated in 186 cases and showed a positive result in 63 MM and 26 PCL patients while del(17) was analyzed in 185 patients and showed a positive result in 22 cases (10 MM and 12 PCL). Translocations at 14q32 involving IGH@ *locus* were detected in the 50% of cases (92/184): in particular, 15 MM and three PCL patients had t(4;14) while t(11;14), investigated in 184 cases, was present in 33 MM and in 12 PCL. T(14;16) was analyzed in 184 patients and showed a positive result in 20 cases (seven MM and 13 PCL), while t(14;20), investigated in 182 samples, was identified in one MM and two PCL. T(6;14) was analyzed in 166 samples and showed a positive result in 6 samples (all MM). Of the 172 analyzed cases, 55 MM and 3 PCL showed a hyperdiploid status according to FISH evaluation criteria [152]. Deletions of chromosome 1p were analyzed in 163 patients and were identified in 22 cases (11 MM and 11 PCL), while gain in 1q region, investigated in 178 cases, was found in 62 MM and 20 PCL patients. Twelve cases (3 MM and 9 PCL) showed the concomitant presence of both aberrations.

According to the TC classification, our MM series was stratified as follows: 28 patients in TC1, including 14 carrying t(11;14) and expressing high levels of *CCND*1 and four harboring a t(6;14) strongly deregulating *CCND*3; 25 patients in TC2, all of whom were negative for IGH translocations and expressing *CCND*1 at moderate levels than TC1, but not *CCND*2 or *CCND*3; 36 patients in TC3, all of whom were negative for the IGH translocation and *CCND*1 expression, most of whom expressed *CCND*2 at variable levels; 13 patients in TC4, all harboring t(4;14); and five patients in TC5, all carrying deregulated MAF or MAFB genes. The patients in TC4 and TC5 were characterized by a consistent expression of *CCND*2, with the TC5 patients generally showing the highest levels.

Assessment of *DIS3* mutation by NGS

To estimate the frequency of *DIS3* mutation, all samples (n. 203) were subjected to deep sequencing of the two functional domains of the gene: the PIN domain, endowed of endonucleolytic activity, spanning from exon 1 to exon 4 while the RNB domain, with a exonucleolytic activity, spanning from exon 10 to exon 18.

The obtained sequencing reads were mapped to the reference sequence (RefSeqGRCh37.p13) and analyzed by the Amplicon Variant Analyzer (AVA) software to establish the mutant allele frequency. *DIS3* exons 1-4 and 10-18 reference sequences were extracted from hg19 Human Genome Version (Transcript Variant 1, NM_014953.4) together with both neighbor intronic regions. Such sequences were used as reference sequences to align every reads and the final alignments were checked manually.

Using the AVA-Software, we combined the raw data obtained from the single sequencing runs and detected a total of 198 deletions, insertions and single nucleotide variations (SNVs), with an average and a median depth of coverage of 253x and 245x, respectively (range: 64-1160).

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 (reference SNP ID number) entry (unless the same variant was also included in NCBI ClinVar database). We found only one germline SNV (P635S) carried by eight patients, also reported in dbSNP (rs35017269) with a global Minor Allele Frequency (MAF) of 0.008/18.

After exclusion of intronic, synonymous and germline variants, deep sequencing of *DIS3* revealed 41 different coding mutations. 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 was consistent with the variant allele frequency, 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.

Among the 41 tumor-specific mutations identified, 30 (73%) were SNVs and the remaining 11 (27%) were indels (Table 1). At the amino acid level, 29 mutations (71%) were missense, 10 (24%) introduced a frameshift, and two (5%) an in-frame deletion. Nine of these variants have been already reported by others, eight of which also specifically in MM patients, while 32 were novel.

DIS3 mutations globally affected 43 samples: 26 MM at diagnosis (26/130, 20%), four at relapse (4/17, 23.5%), six primary PCL (6/24, 25%), four sPCL (4/12, 33.3%) and three multiple myeloma cell lines (3/20, 15%). The major molecular features of the 43 mutated cases are reported in Table 4.

Table 4. Summary of non-synonymous *DIS3* variants identified by NGS.

Variant (GRCh37.p13)	Exon	Type‡	Mutation	AA change	dbSNP ID	COSMIC ID (v71)	MM literature
NC_000013.10:g. 73355804_73355830del	1	deletion	IN FRAME DELETION	del AA 48-56	/	/	/
NC_000013.10:g.73355968C>T	1	SNV	IN FRAME DELETION	del AA 1-16	/	/	/
NC_000013.10:g.73355891T>C	1	SNV	MISSENSE	D27G	/	/	/
NC_000013.10:g.73355093T>C	2	SNV	MISSENSE	T93A	/	/	/
NC_000013.10:g.73355010G>C	2	SNV	MISSENSE	F120L	/	/	/
NC_000013.10:g.73355048G>A	2	SNV	MISSENSE	R108C	/	/	different variant at the same position reported by Lohr[61]
NC_000013.10:g.73355008T>G	2	SNV	MISSENSE	Y121S	/	/	YES, by Keats in the same cell line
NC_000013.10:g.73355018T>C	2	SNV	MISSENSE	K118E	/	/	/
NC_000013.10:g.73355110T>C	2	SNV	MISSENSE	N87S	/	/	/
NC_000013.10:g.73352393A>C	3	SNV	MISSENSE	V171G	/	/	YES, by Keats in the same cell line
NC_000013.10:g.73352416_73352417insT	3	insertion	FRAMESHIFT	M163Ifs*15	/	/	/
NC_000013.10:g.73346400C>T	10	SNV	MISSENSE	R467Q	rs201674523	/	YES, by Lohr[61]
NC_000013.10:g.73346338C>T	10	SNV	MISSENSE	D488N	/	COSM158635	YES, by Lohr[61],Weissbach[137]
NC_000013.10:g.73346340T>A	10	SNV	MISSENSE	D487V	/	/	position previously shown to be critical for exoribonucleotydic activity of h*DIS3*[117]
NC_000013.10:g.73346341C>G	10	SNV	MISSENSE	D487H	/	/	position previously shown to be critical for exoribonucleotydic activity of h*DIS3*[117]
NC_000013.10:g.73346338del	10	deletion	FRAMESHIFT	D488Mfs*22	/	/	different variant at the same position reported by Lohr[61]
NC_000013.10:g.73346340del	10	deletion	FRAMESHIFT	D487Afs*23	/	/	/
NC_000013.10:g.73346400del	10	deletion	FRAMESHIFT	R467Qfs*4	/	/	/
NC_000013.10:g.73345240G>A	12	SNV	MISSENSE	S550F	/	/	different variant at the same position reported by Lohr[61]
NC_000013.10:g.73345094T>C	13	SNV	MISSENSE	H568R	/	/	/
NC_000013.10:g.73345097T>C	13	SNV	MISSENSE	N567S	/	/	/
NC_000013.10:g. 73342930C>T	14	SNV	MISSENSE	E626K	/	/	/
NC_000013.10:g. 73342930del	14	deletion	FRAMESHIFT	E626Kfs*5	/	/	/
NC_000013.10:g.73343049_73343050insA	14	insertion	FRAMESHIFT	A586Vfs*7	/	/	/
NC_000013.10:g.73337707G>T	16	SNV	MISSENSE	A670D	/	/	/
NC_000013.10:g.73337683_73337684insT	16	insertion	FRAMESHIFT	I678Nfs*3	/	/	/
NC_000013.10:g.73337692_73337693insT	16	insertion	FRAMESHIFT	A675Dfs*6	/	/	different variant at the same position reported by Lohr[61]
NC_000013.10:g.73337693_73337694insT	16	insertion	FRAMESHIFT	A675Sfs*6	/	/	different variant at the same position reported by Lohr[61]
NC_000013.10:g.73336112T>C	17	SNV	MISSENSE	H764R	/	/	/
NC_000013.10:g.73336078A>C	17	SNV	MISSENSE	F775L	/	/	YES, by Lohr[61]
NC_000013.10:g.73336064C>T	17	SNV	MISSENSE	R780K	/	COSM329311	YES, by Chapman[43], Walker[136],Lohr[61] and Weissbach[137]
NC_000013.10:g.73336064C>G	17	SNV	MISSENSE	R780T	/	/	YES, by Lohr[61]
NC_000013.10:g.73336113G>A	17	SNV	MISSENSE	H764Y	/	/	/
NC_000013.10:g.73337650C>T	17	SNV	MISSENSE	R689Q	/	/	YES, by Lohr[61]
NC_000013.10:g.73336077T>G	17	SNV	MISSENSE	T776P	/	/	/
NC_000013.10:g.73336064del	17	deletion	FRAMESHIFT	R780Kfs*25	/	/	different variant at the same position reported by Chapman[43], Walker[136], Lohr[61],Weissbach[137]
NC_000013.10:g.73336113G>C	17	SNV	MISSENSE	H764D	/	/	/
NC_000013.10:g.73335837G>A	18	SNV	MISSENSE	R820W	rs372878316	/	/
NC_000013.10:g.73335816C>G	18	SNV	MISSENSE	A827P	/	/	/
NC_000013.10:g.73335930G>A	18	SNV	MISSENSE	R789W	/	/	/
NC_000013.10:g.73336151G>T	18	SNV	MISSENSE	A751D	/	/	/

Mutant allele frequency ranged from 0.38% to 100% (Figure 11): 14 cases show an allele frequency >80% while in all the remaining cases, the percentage of variant reads was suggestive of a mutation present in heterozygosis or in a tumor subclone. Of the 41 variants identified in our cohort of patients, 25 (25/41, 61%) are present at clonal level while 16 (16/41, 39%) only at low allele frequency (below the detection limit of Sanger, i.e. about 10%).

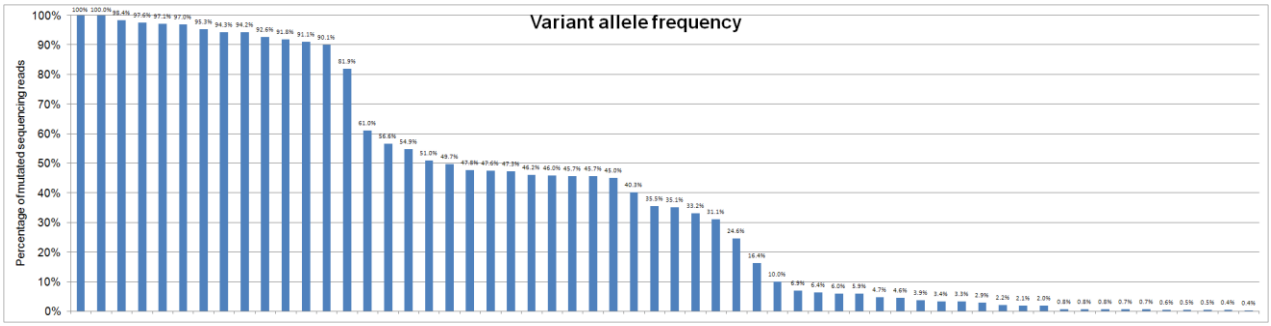

Figure 11. Variant allele frequency of *DIS3*-mutated cases.

As regards subclonal mutations, these affected only MM at onset (10/130, 7.7%) and were present with allele frequency comprised between 0.38 and 4.71%; seven of these mutations were SNVs while the remaining three were insertions causing a frameshift. Eight cases showed both clonal and subclonal mutations: seven subclonal mutations causes a frameshift and only one is a SNV.

Additionally, of the 30 SNVs, all but one (a mutation affecting the ATG start codon causing a frameshit) were analyzed by five damage prediction databases, namely Polyphen2, Mutation Taster, Mutation Assessor, LRT and SIFT. These databases predict possible impact of an amino acid substitution on the structure and function of protein. According to all the databases, 23 SNVs were designated 'damaging' (mutations that occur at conservative sites or lead to structural changes) and three 'benign' while, among the remaining three variants, there was not overlap of predictions between the databases (Table 5).

chromosomic position (37)	AA change	Polyphen2	Mutation Taster	Mutation Assessor	SIFT PREDICTION	LRT (Likelihood Ratio test)
NC_000013.10:g.73355093T>C	T93A	POSSIBLY DAMAGING MUTATION	DISEASE CAUSING	PREDICTED FUNCTIONAL (MEDIUM) FI score: 3.34	DAMAGING (SIFT score: 0)	DELETERIOUS
NC_000013.10:g.73352393A>C	V171G	PROBABLY DAMAGING MUTATION	DISEASE CAUSING	PREDICTED FUNCTIONAL (MEDIUM) FI score: 3.185	DAMAGING (SIFT score: 0)	DELETERIOUS
NC_000013.10:g.73335837G>A	R820W	PROBABLY DAMAGING MUTATION	DISEASE CAUSING	PREDICTED FUNCTIONAL (MEDIUM) FI score: 3.275	DAMAGING (SIFT score: 0)	DELETERIOUS
NC_000013.10:g.73336112T>C	H764R	PROBABLY DAMAGING MUTATION	DISEASE CAUSING	PREDICTED FUNCTIONAL (HIGH) FI score: 4.77	DAMAGING (SIFT score: 0)	DELETERIOUS
NC_000013.10:g.73335816C>G	A827P	PROBABLY DAMAGING MUTATION	DISEASE CAUSING	PREDICTED FUNCTIONAL (MEDIUM) FI score: 3.24	DAMAGING (SIFT score: 0)	DELETERIOUS
NC_000013.10:g.73337707G>T	A670D	PROBABLY DAMAGING MUTATION	DISEASE CAUSING	PREDICTED FUNCTIONAL (HIGH) FI score: 4.74	DAMAGING (SIFT score: 0)	DELETERIOUS
NC_000013.10:g.73355010G>C	F120L	PROBABLY DAMAGING MUTATION	DISEASE CAUSING	PREDICTED FUNCTIONAL (MEDIUM) FI score: 3.065	DAMAGING (SIFT score: 0.05)	DELETERIOUS
NC_000013.10:g.73355048G>A	R108C	PROBABLY DAMAGING MUTATION	DISEASE CAUSING	PREDICTED FUNCTIONAL (MEDIUM) FI score: 3.45	DAMAGING (SIFT score: 0)	NEUTRAL
NC_000013.10:g.73335930G>A	R789W	POSSIBLY DAMAGING MUTATION	DISEASE CAUSING	PREDICTED FUNCTIONAL (HIGH) FI score: 4.77	DAMAGING (SIFT score: 0)	DELETERIOUS
NC_000013.10:g.73336078A>C	F775L	PROBABLY DAMAGING MUTATION	DISEASE CAUSING	PREDICTED FUNCTIONAL (HIGH) FI score: 4.77	DAMAGING (SIFT score: 0)	DELETERIOUS
NC_000013.10:g.73355008T>G	Y121S	PROBABLY DAMAGING MUTATION	DISEASE CAUSING	PREDICTED FUNCTIONAL (MEDIUM) FI score: 3.24	DAMAGING (SIFT score: 0.04)	NEUTRAL
NC_000013.10:g.73355018T>C	K118E	POSSIBLY DAMAGING MUTATION	DISEASE CAUSING	PREDICTED FUNCTIONAL (MEDIUM) FI score: 3.24	DAMAGING (SIFT score: 0.01)	NEUTRAL
NC_000013.10:g.73346400C>T	R467Q	PROBABLY DAMAGING MUTATION	DISEASE CAUSING	PREDICTED FUNCTIONAL (HIGH) FI score: 4.675	DAMAGING (SIFT score: 0)	DELETERIOUS
NC_000013.10:g.73346338C>T	D488N	PROBABLY DAMAGING MUTATION	DISEASE CAUSING	PREDICTED FUNCTIONAL (HIGH) FI score: 4.74	DAMAGING (SIFT score: 0)	DELETERIOUS
NC_000013.10:g.73336064C>T	R780K	PROBABLY DAMAGING MUTATION	DISEASE CAUSING	PREDICTED FUNCTIONAL (HIGH) FI score: 4.77	DAMAGING (SIFT score: 0)	DELETERIOUS
NC_000013.10:g.73336151G>T	A751D	PROBABLY DAMAGING MUTATION	DISEASE CAUSING	PREDICTED FUNCTIONAL (HIGH) FI score: 4.77	DAMAGING (SIFT score: 0)	DELETERIOUS
NC_000013.10:g.73336064C>G	R780T	PROBABLY DAMAGING MUTATION	DISEASE CAUSING	PREDICTED FUNCTIONAL (HIGH) FI score: 4.77	DAMAGING (SIFT score: 0)	DELETERIOUS
NC_000013.10:g.73346340T>A	D487V	PROBABLY DAMAGING MUTATION	DISEASE CAUSING	PREDICTED FUNCTIONAL (HIGH) FI score: 4.74	DAMAGING (SIFT score: 0)	DELETERIOUS
NC_000013.10:g.73346341C>G	D487H	PROBABLY DAMAGING MUTATION	DISEASE CAUSING	PREDICTED FUNCTIONAL (HIGH) FI score: 4.74	DAMAGING (SIFT score: 0)	DELETERIOUS
NC_000013.10:g.73345240G>A	S550F	PROBABLY DAMAGING MUTATION	DISEASE CAUSING	PREDICTED FUNCTIONAL (HIGH) FI score: 4.775	DAMAGING (SIFT score: 0)	DELETERIOUS
NC_000013.10:g. 73342930C>T	E626K	BENIGN MUTATION	DISEASE CAUSING	PREDICTED NON-FUNCTIONAL (LOW) FI score: 0.835	TOLERATED (SIFT score: 0.63)	DELETERIOUS
NC_000013.10:g.73336113G>A	H764Y	PROBABLY DAMAGING MUTATION	DISEASE CAUSING	PREDICTED FUNCTIONAL (HIGH) FI score: 4.77	DAMAGING (SIFT score: 0)	DELETERIOUS
NC_000013.10:g.73355891T>C	D27G	PROBABLY DAMAGING MUTATION	DISEASE CAUSING	PREDICTED FUNCTIONAL (MEDIUM) FI score: 3.165	DAMAGING (SIFT score: 0.01)	DELETERIOUS
NC_000013.10:g.73355110T>C	N87S	PROBABLY DAMAGING MUTATION	DISEASE CAUSING	PREDICTED FUNCTIONAL (MEDIUM) FI score: 2.98	DAMAGING (SIFT score: 0.03)	DELETERIOUS
NC_000013.10:g.73337650C>T	R689Q	PROBABLY DAMAGING MUTATION	DISEASE CAUSING	PREDICTED FUNCTIONAL (HIGH) FI score: 4.77	DAMAGING (SIFT score: 0)	DELETERIOUS
NC_000013.10:g.73345094T>C	H568R	BENIGN MUTATION	DISEASE CAUSING	PREDICTED NON-FUNCTIONAL (NEUTRAL) FI score: -0.455	TOLERATED (SIFT score: 0.26)	NEUTRAL
NC_000013.10:g.73336077T>G	T776P	PROBABLY DAMAGING MUTATION	DISEASE CAUSING	PREDICTED FUNCTIONAL (HIGH) FI score: 4.77	DAMAGING (SIFT score: 0)	DELETERIOUS
NC_000013.10:g.73345097T>C	N567S	BENIGN MUTATION	DISEASE CAUSING	PREDICTED NON-FUNCTIONAL (NEUTRAL) FI score: -0.345	TOLERATED (SIFT score: 0.19)	DELETERIOUS
NC_000013.10:g.73336113G>C	H764D	PROBABLY DAMAGING MUTATION	DISEASE CAUSING	PREDICTED FUNCTIONAL (HIGH) FI score: 4.77	DAMAGING (SIFT score: 0)	DELETERIOUS

Table 5. Results of damage predictions softwares.

The great majority of mutations identified in our cohort of patients affected the RNB domain (30/41, 73.2%), while only nine variants (22%) were present in the PIN domain. The remaining two SNVs were found in a region downstream of the RNB domain. Eight of these mutations were recurrent, each affecting two or more patients: of these, three SNVs were not described previously (T93A, R789W and R820W) while five targeted conserved residues (F775, D488 and R780) already reported by others authors [43;61;72;137].

In agreement with Weissbach *et al* [137], we found two SNVs, R780K (indicated as R750K by Chapman *et al* and Walker *et al*) and D488N that are located in recurrent mutational hot-spots within fully conserved amino acid in RNB domain. This is an important finding because recurrent site-specific somatic SNVs in MM are rare [65]. Overall, of the other mutations that occurred in the RNB domain, 11 targeted highly conserved amino acids. Particularly, one of the most frequent mutation affected conserved amino acid R780, and was carried by five patients (two MM and three PCL): this mutation

affected an amino acid residue involved in substrate binding and cause a strong reduction of the exonucleolytic activity [68]. Similar effects were caused by the D487 substitution (two mutated MMs) [68]. D488 was another of the most frequently mutated residue in our patients' cohort (four cases carrying a mutation in its specific codon: three MMs and one PCL, respectively). These variations affected an aspartic acid residue involved in substrate catalysis essential for magnesium ion coordination in the catalytic centre of the protein. It could be possible that a *DIS3* mutation at position D488, which lies next to D487, might have a similar effect to those displayed by D487 variants. Otherwise, the mutations in the PIN domain did not target particularly conserved amino acid residues, predicted to coordinate an efficient degradation/processing of several natural exosome substrates.

Interestingly, inactivating variants that affected conserved amino acid residues R780, D488 and D487, were always detected (both in disomic both in 13q deleted patients) with an allele frequency lower than 50% in our dataset (and in according to Weissbach *et al* [137] and Lohr *et al* [61], suggesting that the exclusive expression of such *DIS3* proteins may have harmful effects on cell.

Furthermore, one sample (PCL-042) displayed more than one mutation in the RNB domain at high allelic frequency (R780K and E626K in 45.95% and 33.15% of mutated sequencing reads, respectively), while a MM patient harbored a 9aa in-frame deletion in the PIN domain (Ala47_Pro55del) and a missense substitution in the RNB domain (A751D). Additionally, these samples displayed, in addition to high frequency mutations, also low-level mutations.

Finally, only three (KMS26, KMS 27 and OPM2) out of the 21 HMCLs tested showed *DIS3* mutations; in agreement with Weissbach *et al* [137], also we founded the same SNV (Y121S) in the OPM2 cell line.

FISH analysis

The results of the FISH analysis of our mutated cases are reported in Table 6. Among the mutated patients (n=40), we observed deletion of chromosome 13 in 22 patients (22/40, 55%), whereas del(17)(p13) was found in three cases (7.5%). Translocations at 14q32 involving IGH@ *locus* were detected in the large majority of cases (28/40, 70%); in particular, t(4;14) was present in six patients (15%) while t(11;14) and t(14;16), each investigated in 38 cases, were present in 15 (39.5%) and in five (13.2%) patients respectively. The t(6;14) was analyzed in 35 cases, all but one found to be negative for this lesion. The t(14;20) was investigated in 36 cases and showed a positive result in one case (2.8%). Deletions of chromosome 1p were analyzed in 35 patients and were identified in 3 cases (8.6%), while gain in 1q region, investigated in 37 cases, was found in 20 (54.1%) patients. Only one case (PCL-019) showed the concomitant presence of 1q gain and 1p loss. Finally, of the 35 analyzed cases, five showed a hyperdiploid status.

Table 6. Clinical and molecular characteristics of 40 MM/PCL patients and three HMCLs carrying non-synonymous *DIS3* mutations.

Sample name	Disease stage‡	Variant (GRCh37.p10)	% of mutated sequencing reads	AA change	del(13)	del(17p)	1q gain	1p loss	t(4;14)	t(11;14)	t(14;16)	t(14;20)	t(6;14)	HD
MM-004_2	MM, R	NC_000013.10:g.73335816C>G	97.12%	A827P	-	-	+	-	-	-	+	-	-	-
MM-034	MM, E	NC_000013.10:g.73337692_73337693insT	0.44%	A675Dfs*6	-	-	-	-	-	-	-	-	-	+
MM-036	MM, E	NC_000013.10:g.73336078A>C	92.57%	F775L	+	-	+	nd	-	-	-	-	-	-
MM-042	MM, E	NC_000013.10:g.73346338C>T	4.57%	D488N	+	-	+	-	+	-	-	-	-	+
MM-055	MM, E	NC_000013.10:g.73343049_73343050insA	0.83%	A586Vfs*7	-	-	-	nd	-	+	-	-	-	-
MM-123	MM, E	NC_000013.10:g.73355010G>C	95.30%	F120L	+	-	+	-	+	-	-	-	-	-
MM-143_1	MM, E	NC_000013.10:g.73355891T>C	0.62%	D27G	-	-	+	-	-	-	-	-	-	+
MM-150	MM, E	NC_000013.10:g.73345097T>C	0.51%	N567S	-	-	+	-	-	-	-	-	-	+
MM-207	MM, E	NC_000013.10:g.73336113G>A NC_000013.10:g.73335837G>A NC_000013.10:g.73336113G>C NC_000013.10:g.73336077T>G	24.58% 2.05% 1.95% 0.79%	H764Y R820W H764D T776P	+	-	+	-	-	-	-	-	+	-
MM-208	MM, E	NC_000013.10:g.73352416_73352417insT	0.38%	M163Ifs*15	-	-	-	-	+	-	-	-	-	-
MM-213	MM, E	NC_000013.10:g.73355093T>C	56.57%	T93A	+	-	-	-	-	+	-	-	-	-
MM-263_1	MM, E	NC_000013.10:g.73346400C>T NC_000013.10:g.73346400del	31.14% 5.93%	R467Q R467Qfs*4	+	-	+	-	+	-	-	-	-	-
MM-263_2	MM, R	NC_000013.10:g.73346400C>T	81.90%	R467Q	+	-	+	-	+	-	-	-	-	-
MM-279	MM, E	NC_000013.10:g.73355110T>C	2.93%	N87S	+	-	+	-	-	-	-	nd	nd	nd
MM-281_1	MM, E	NC_000013.10:g.73336064C>T	36.80%	R780K	-	-	-	-	-	-	-	-	-	-
MM-310	MM, E	NC_000013.10:g.73345094T>C	0.50%	H568R	-	-	-	-	-	+	-	-	-	-
MM-335	MM, E	NC_000013.10:g.73335930G>A	94.20%	R789W	+	-	-	-	-	-	+	-	-	-
MM-340_1	MM, E	NC_000013.10:g.73355804_73355830del NC_000013.10:g.73336151G>T NC_000013.10:g.73337650C>T	54.86% 45.71% 2.2%	del AA 48-56 A751D R689Q	+	-	-	-	-	+	-	-	-	-
MM-340_2	MM, R	NC_000013.10:g.73355804_73355830del NC_000013.10:g.73336151G>T	61% 47.27%	del AA 48-56 A751D	+	-	-	-	-	+	-	-	-	-
MM-343	MM, E	NC_000013.10:g.73346338C>T	46.18% 5.96%	D488N D488Mfs*22	-	-	-	-	-	+	-	-	-	-
MM-372_1	MM, E	NC_000013.10:g.73346341C>G	35.48%	D487H	-	-	+	-	-	+	-	-	-	-
MM-372_3	MM, CPT	NC_000013.10:g.73346341C>G	10%	D487H	+	-	+	-	-	+	-	-	-	-
MM-381	MM, E	NC_000013.10:g.73336064C>G	16.38%	R780T	-	-	-	-	-	-	-	nd	nd	nd
MM-386	MM, E	NC_000013.10:g.73345240G>A	35.11%	S550F	-	-	nd	nd	-	nd	nd	nd	nd	nd
MM-398	MM, E	NC_000013.10:g.73346338C>T	3.85%	D488N	+	-	+	-	-	+	-	-	-	-
MM-414	MM, E	NC_000013.10:g.73336078A>C	4.71%	F775L	-	-	-	-	-	+	-	-	-	-

Sample name	Disease stage‡	Variant (GRCh37.p10)	% of mutated sequencing reads	AA change	del(13)	del(17p)	1q gain	1p loss	t(4;14)	t(11;14)	t(14;16)	t(14;20)	t(6;14)	HD
MM-424	MM, E	NC_000013.10:g.73355048G>A	94.33%	R108C	+	-	-	+	-	-	-	-	-	+
MM-445	MM, E	NC_000013.10:g.73335837G>A	98.36%	R820W	+	-	+	-	-	+	-	-	nd	-
MM-464	MM, E	NC_000013.10:g.73355968C>T	50.99%	del AA 1-16	+	-	+	-	-	-	-	-	-	-
sMM-317	MM, E	NC_000013.10:g.73346340T>A NC_000013.10:g.73346340del	40.3% 3.32%	D487V D487Afs*23	-	-	+	-	-	+	-	-	-	-
PCL-001	pPCL, E	NC_000013.10:g.73337707G>T	97%	A670D	+	-	nd	nd	-	-	-	+	-	nd
PCL-015	pPCL, E	NC_000013.10:g.73355018T>C	90.05%	K118E	+	-	+	-	-	-	+	-	-	-
PCL-019	pPCL, E	NC_000013.10:g.73336064C>T NC_000013.10:g.73336064del	47.75% 0.78%	R780K R780Kfs*25	+	+	+	+	-	-	+	-	-	-
PCL-021	pPCL, E	NC_000013.10:g.73355093T>C	100%	T93A	+	-	+	-	+	-	-	-	-	-
PCL-035	pPCL, E	NC_000013.10:g.73336064C>G	45.04%	R780T	-	-	-	-	-	+	-	-	-	-
PCL-036	pPCL, E	NC_000013.10:g.73336112T>C	97.59%	H764R	+	-	-	-	-	+	-	-	-	-
MM-281_2	sPCL	NC_000013.10:g.73336064C>T	47.56%	R780K	-	+	-	+	-	-	-	-	-	-
PCL-011	sPCL	NC_000013.10:g.73335930G>A	91.1%	R789W	+	+	+	-	-	-	+	-	-	-
PCL-041	sPCL	NC_000013.10:g.73346338C>T NC_000013.10:g.73346338del	45.68% 6.4%	D488N D488Mfs*22	-	-	nd	nd	-	nd	nd	nd	nd	nd
PCL-042	sPCL	NC_000013.10:g.73336064C>T NC_000013.10:g.73342930C>T NC_000013.10:g.73342930del	45.95% 33.15% 3.42%	R780K E626K E626Kfs*5	-	-	-	-	-	+	-	-	-	-
KMS26	HMCL	NC_000013.10:g.73352393A>C	100%	V171G	nd	nd	nd	nd	+	-	+	-	nd	nd
KMS27	HMCL	NC_000013.10:g.73346338C>T NC_000013.10:g.73346338del	49.68% 6.92%	D488N D488Mfs*22	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd
OPM2	HMCL	NC_000013.10:g.73355008T>G NC_000013.10:g.73337683_73337684insT NC_000013.10:g.73337693_73337694insT	91.79% 0.73% 0.73%	Y121S I678Nfs*3 A675Sfs*6	nd	nd	nd	nd	+	-	-	+	nd	nd

As regard the association of *DIS3* mutations and chromosomal aberrations, we observed a positive association with the occurrence of translocations involving IGH@ *locus*, particularly with the t(11;14) (Fisher's exact test *P*=0.0208) (Figure 12A), and a negative association with hyperdiploid cases (Fisher's exact test *P* =0.0086) (Figure 12B), but none association with 13q and 17p deletion, gain of chromosome 1q and loss of chromosome 1p.

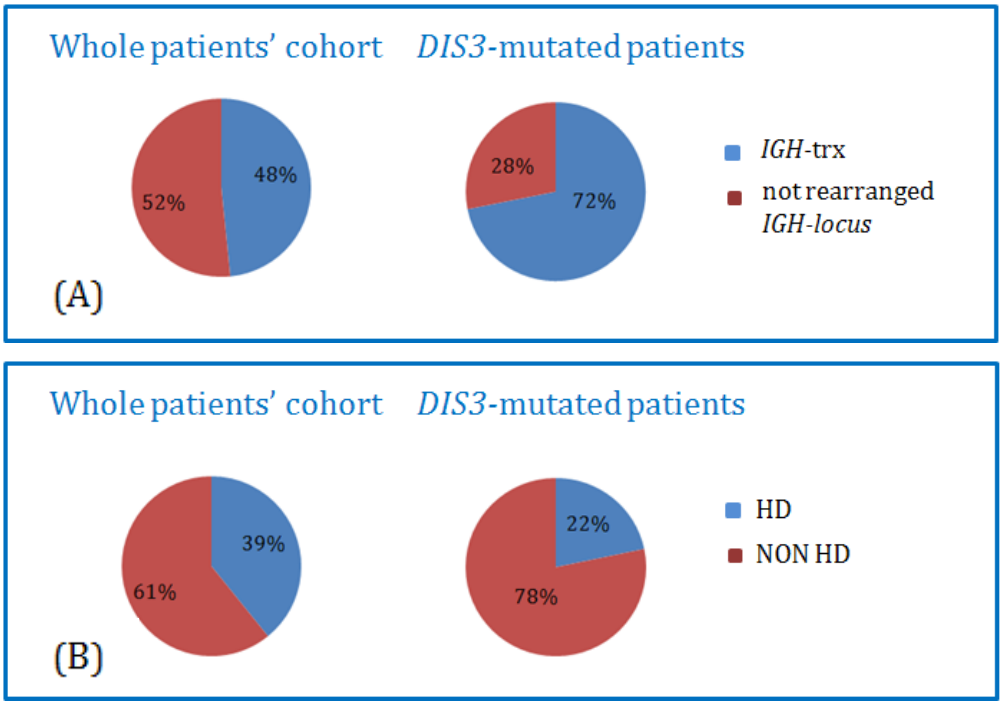

Figure 12. Pie chart of the proportional representation of *IGH*-translocation (A) and hyperdiploidy (HD) (B) in whole patients' cohort and *DIS3*-mutated cases.

Among our mutated cases with 13q deletion, mutant allele frequency was ≥90% in 10 out of 22 patients (5 MM and 5 PCL) while, in the remaining 12 patients (11 MM and 1 PCL), the percentage of variant reads was comprise between 3 and 82%. In the remaining 18 cases, disomic for chromosome 13, the percentage of variant reads of all but one samples (one relapsed MM patient harboring a mutation presumably present in both alleles at a frequency of 97.2%) was suggestive of a mutation present in heterozygosis or in a small tumor subclone (range between 0.38-47.56%).

Sequential analysis of *DIS3* mutations

To gain further insight into the state of *DIS3* mutation longitudinally, we analyzed samples of 20 patients for whom bone marrow specimens were collected at two different time points: in particular, fifteen MM and two pPCL cases at onset and at relapse, respectively; two patients with MM at onset and at leukemic transformation, and one patient in early and relapsed leukemic phase of MM (Figure 13).

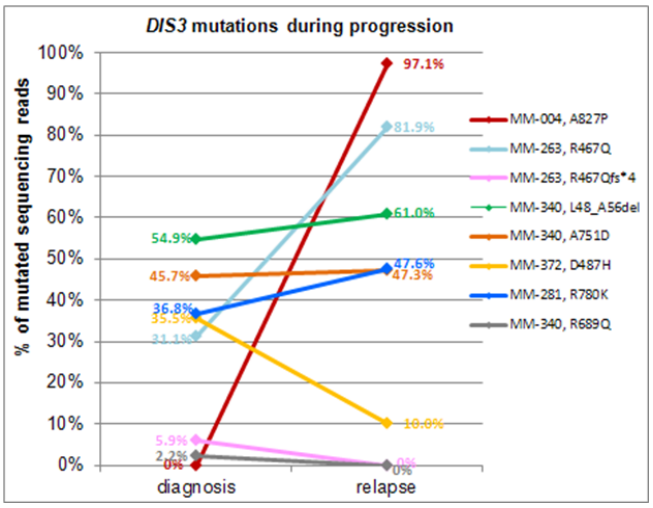

Figure 13. Grafic representation of *DIS3* mutations during disease progression.

Fifteen cases displayed a wild-type status for the gene at both time-points, while one patient that harbored the R780K variant showed a quite constant load during disease course (present in 36.8% of sequencing reads of sample obtained at diagnosis and in 47.56% during leukemic transformation). One wild type MM sample at onset acquired a SNV (A827P) detected in 97.1% of the sequencing reads at relapse: this patient showed two copies of chromosome 13 both at onset both at relapse, and also the other cytogenetic lesions remained stable during disease progression. A considerable increase in mutation burden was observed in only one MM case, with R467Q mutation occurring in 31.1% of the sequencing reads obtained at onset and 81.9% in the sample obtained at relapse. Additionally, this case carried a frameshift mutation at the same position at diagnosis that disappeared at relapse. The case MM-372 harbored the D487H variant at onset but showed a decrease in mutation burden during tumor progression (with an allele frequency of 35.48% at MM onset and 10% at post-therapy control). One case carried at onset three variants (del AA48-56, A751D and R689Q at 54.86%, 45.71% and 2.20% respectively) but at relapse only two (del AA48-56 and A751D) remained stable (with an allele frequency of 61% and 47.27% respectively) and the R689Q disappeared. The disappearance of a high frequency mutation was never observed in our cohort.

Expression of *DIS3* variants on cDNA

To verify if the mutations detected on genomic DNA are also expressed at transcriptomic level, we performed deep sequencing of cDNA of the *DIS3*-mutated cases. In particular, compatibly with the availability of material, 23 *DIS3* variants were evaluated. NGS results indicated that mutant allele frequencies detected on genomic DNA and on retrotranscribed total RNA are linearly correlated (Figure 14, showing the respective frequencies lying on the leading diagonal of the plot); in particular, a good level of correlation was found for all clonal variants, that were confirmed on cDNA generally at quite comparable allele frequency, while all but one subclonal mutations were not detectable at transcriptomic level (Figure 15).

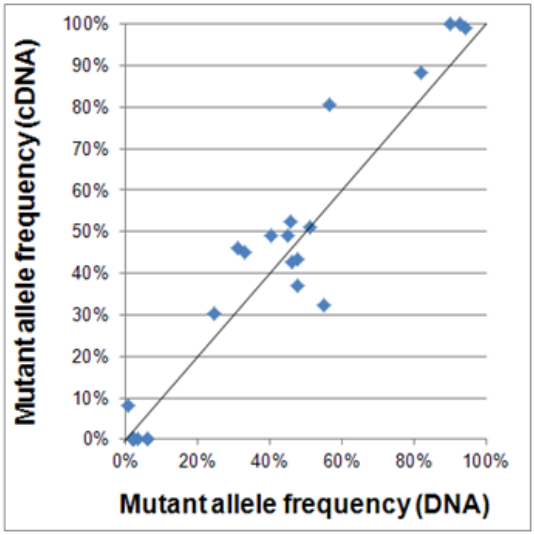

Figure 14. Mutant allele frequency of mutations observed on the cDNA (y-axis) plotted against mutant allele frequency observed on the genomic DNA (x-axis).

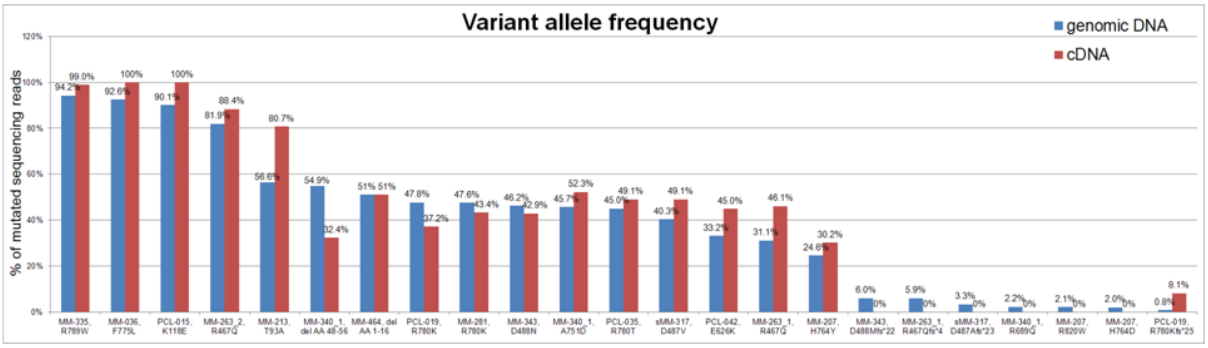

Figure 15. Variant allele frequencies indentified on genomic DNA and on cDNA.

Transcriptomic profile of *DIS3*-mutated patients

To identify transcriptomic profiles possibly related to *DIS3* mutations identified by NGS in MM patients, gene expression profiling was performed in a large part (n=102) of the samples analyzed by sequencing, including wild-type cases for the gene (n=89) and patients carrying mutations in *DIS3* (n=13), arbitrarily establishing 20% as the lower variant allele frequency cut-off, assuming that alterations present in a very limited number of MM cells might not appreciably affect gene expression.

Supervised analysis revealed 119 genes differentially expressed between *DIS3* mutated and *DIS3* wild-type cases; all this genes were up-regulated in *DIS3* mutated patients and 19 of these genes emerged at the highest stringency level. Notably, only seven of the 19 differentially expressed genes at the highest stringency level were protein coding, while ten genes encode for ncRNA, of which five were long non-coding RNAs and the others were not characterized; among protein coding genes, it is worth noting *APOBEC4*, belonging to a gene family involved in mRNA editing, somatic hypermutation and recombination of immunoglobulin genes.

Interestingly, functional enrichment analysis revealed the involvement of a statistically significant fraction of modulated transcripts in single-stranded RNA binding at molecular function level

(GO:0003727, q-value=1.96E-02), defense response to virus at biological process level (GO:0000165, q-value=1.96E-02) and interferon alpha/beta signaling at pathway level (187104BioSystems: REACTOME, q-value=8.23E-03).

Prognostic impact of *DIS3* mutations in pPCLs.

In the 16pPCL patients for whom the follow-up was available, the association of *DIS3* mutations with clinical outcome was tested, comparing wild-type patients (n=12) with those characterized by mutations in PIN or RNB functional domains of *DIS3* (n=4). Mutational events did not show any impact neither on PFS nor OS (Figure 16). Likewise, the response rate after a 4-cycle therapy with low-dose lenalidomide-dexamethasone (assessed according to International Uniform Criteria [155;156] and representing the primary endpoint of the RV-PCL-PI-350trial) did not associate with the occurrence of *DIS3* mutations (Fisher's exact test).

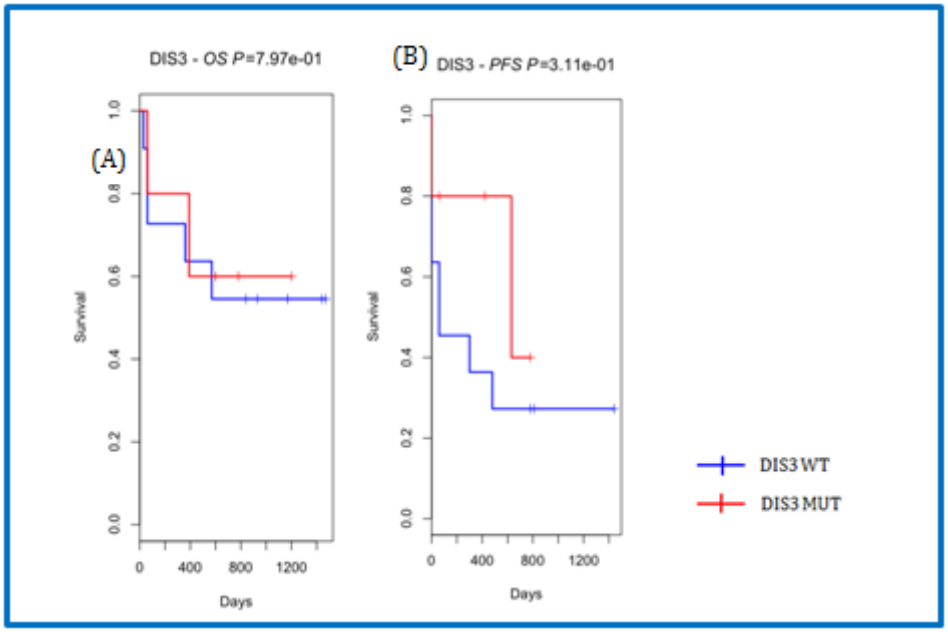

Figure 16. Correlation of genetic events with clinical outcome. Overall survival (A) and progression-free survival (B) of patients with *DIS3* WT compared with *DIS3* mut according to a univariate analysis.

Discussion

DIS3 is an active part of the human exosome complex that possesses both exo- and endonucleolytic activity and it is involved in regulating the processing and abundance of all RNA species [117]. Recently, the *DIS3* gene has emerged as recurrently mutated in MM patients from initial whole-genome and exome sequencing studies [43;61;65;136;137]). To characterized further the molecular spectrum of *DIS3* mutations in MMs, PCLs and HMCLs samples, we performed deep amplicon sequencing that covered the PIN (exons 1-4) and the RNB (exons 10-18) *DIS3* functional domains.

To the best of our knowledge, this is the first study assessing *DIS3* mutational status by means of next generation sequencing (NGS) technology in a large cohort of patients at different stages of plasma cell dyscrasia, evaluating it in the context of other clinical and biological features of the disease and determining a *DIS3* mutation-associated transcriptional signature in MM.

In our cohort of patients, *DIS3* mutations globally affect 26 MM at diagnosis (26/130, 20%), four at relapse (4/17, 23.5%)(of whom three mutated also at onset), six primary PCL (6/24, 25%), four secondary PCL (4/12, 33.3%) and three multiple myeloma cell lines (3/20, 15%).

The frequency of *DIS3* mutations in MM at onset is similar to that reported by Walker *et al* [136] that analyzed a cohort of patients exclusively composed by *IGH*-translocated samples. However, it is higher than that obtained in unselected patients' cohorts through less sensitive approaches by Chapman *et al* [43], Lohr *et al* [61] and Bolli *et al* [20], who reported a percentage of mutated MM samples of 10.5%, 11% and <3%, respectively. More recently Weissbach *et al* [137], despite yielding a high sequencing coverage comparable to that obtained by us, reported a mutation frequency of 11%.

Interestingly, the occurrence of *DIS3* mutations increased (although not at statistically significant level) in primary and secondary PCLs (25% and 33.3%, respectively), confirming the evidence that the complexity of genome aberrations increases with disease stage[136]. *DIS3* mutation status has never been thoroughly investigated before in these aggressive forms of plasma cell dyscrasia, and it would be worthy to confirm these findings in larger patients' cohorts, although difficult to collect due to the rarity especially of primary forms.

In line with what reported by Chapman *et al* [43], but in contrast with Weissbach *et al* [137], we found an accumulation of *DIS3* mutations in the RNB domain (30/41, 73.2%), while only nine variants (22%) were present in the PIN domain. The remaining two SNVs were found in a region downstream of the RNB domain. Interestingly, in our cohort of patients the majority of SNVs detected in RNB domain were located in highly conserved residues, suggesting their potential impact on the exonucleolytic activity of the enzyme[68]. Among these, one of the most frequently mutated residue was the conserved amino acid R780 and found altered in five patients of current dataset and in six patients of previously published reports (one by Chapman *et al* [43]; three by Lohr *et al*, [61]; one by Walker *et al* [136], and one Waissbach *et al* [137]). It is involved in substrate binding, and its mutation has been reported in yeast and

bacteria to cause a strong reduction of the exonucleolytic activity leading to a decreased degradation rate of RNA molecules belonging to different RNA classes[68]. Similar effects were caused by the D487 substitution [68;157], reported for the first time in our cohort and affecting two MM patients. D488 was another of the most frequently mutated residue: in our patients cohort four cases carried a mutation in this amino acid (three MMs and one PCL), while Lohr *et al* [61]and Weissbach *et al* [137]found this residue mutated in two and one patients, respectively. D488 is an aspartic acid residue involved in substrate catalysis essential for magnesium ion coordination in the catalytic centre of the protein and, because it lies next to D487, its substitution might have a similar effect to that displayed by D487 variants. Interestingly, these inactivating variants that affected conserved amino acids R780, D488 and D487 were always detected at allele frequency lower than 50% in our cohort of patients (in agreement with what found in the series of Weissbach *et al* [137] and Lohr *et al* [61], suggesting the hypothesis that the exclusive expression of such *DIS3* mutants in the absence of *DIS3* wild type protein might have deleterious cellular effects. Finally, two of these recurrently mutated amino acid residues in our dataset, i.e. D488 and R780, were recently proposed as mutational hot-spots by Weissbach *et al* [137].This is an important finding, because recurrent site-specific somatic SNVs in MM are rare [65].

As regard specific amino acid substitution, seven mutations of this dataset are recurrent, affecting two or more patients: of these, three SNVs were not described previously (T93A, R789W and R820W) while four (F775L, D488N, R780K and R780T) were already reported by other authors [43;61;136;137].

Notably, one MM sample of our dataset harbored a 9aa in-frame deletion in the PIN domain (Ala47_Pro55del) and a missense substitution in the RNB domain (A751D). In the yeast system, Tomecki *et al* [68] showed that MM-associated mutations in h*DIS3* RNB domain that inhibit its exonuclease activity are indeed synthetically lethal with inactivation of h*DIS3* endonucleolytic activity, suggesting that the PIN domain may be also a potential drug target for cancers bearing *DIS3* mutations.

With regard to the correlation of *DIS3* variants with cytogenetic alterations, in agreement with previous findings [43, 61, 136;137].) we observed a positive association between *DIS3* mutations and the occurrence of translocations involving IGH@ *locus* (particularly with the t(11;14)). With this regard it has been suggested that alteration of exosome complex (known to be involved in AID-mediated immunoglobulin hypermutations and class switches) due to *DIS3* mutations, may lead to chromosomal translocations [133].

Moreover, we showed a negative association between *DIS3* mutated cases and hyperdiploidy, but no association with 13q (in contrast to previous findings of Chapman *et al* [43], Lohr *et al* [61], Walker *et al* [136] and Weissbach *et al* [137]) and 17p deletion, gain of chromosome 1q and loss of chromosome 1p.

In a considerable fraction of patients, *DIS3* variants were identified at allele frequencies suggestive of the occurrence of the mutations in minor tumor subclones, confirming recently published data[137]. Besides a fraction of cases positive for del(13) and carrying mutations at allele frequency around 100%, and disomic patients with about 50% of mutated sequencing reads, we indeed detected *DIS3* substitutions at variant allele frequencies supporting the previously described and validated subclone

theory [20;27;136]. Notably, mutations at the lowest allele frequencies (undetectable by Sanger) exclusively affected MM patients at onset. Recently, Weissbach *et al* [137] reported that patients with *DIS3* mutations in a minor subclone showed a worse prognosis if compared with patients with mutations in a major subclone; unfortunately, we could not confirm this finding due to the unavailability of the clinical data for mutated patients at subclonal level.

Samples with *DIS3* mutations can also have deletion of 13q, including the *DIS3* locus in MM, and so deletion and mutation of *DIS3* may be collaborating lesions. The gene of interest at 13q14 has been debated for many years, and it remains unclear which region(s) of chromosome 13 is pathologically relevant; the discovery of another altered gene in this region will confound the situation further and only through analyses of large datasets, comprising copy number and mutational data, will the true nature of these abnormalities be discovered.

Although a number of studies have interrogated the role of mutational changes in cancer and in myeloma, the phenotypic impact of mutations has not been evaluated for the majority of these alterations [158]. The ultimate significance of these genetic changes will depend on whether the mutated allele is expressed, and whether the mutation affects expression, splicing, or function of the gene product. Recently, Rashid *et al* [159],for the first time in MM, used RNA-seq to examine the relationship between mutational status of a gene and its allelic expression, revealing that the majority of mutations are found in genes that have low or no detectable biological expression. Additionally, mutated genes often show differential allelic expression in multiple myeloma patient samples. In our study, we sequenced cDNA of the *DIS3*-mutated cases to verify if the mutations detected on genomic DNA are also expressed at transcriptomic level. In contrast with Rashid *et al* [159], we were able to confirm all clonal variants on cDNA, generally at quite comparable allele frequency, suggesting that these mutations may be functionally relevant. Furthermore, confirming the cellular expression of *DIS3* mutant proteins, our gene expression profiling analysis revealed 119 differentially expressed genes (all up-regulated in mutated cases) between *DIS3* mutated and *DIS3* wild-type cases. Functional enrichment analysis revealed the involvement of a statistically significant fraction of modulated transcripts in single-stranded RNA binding at molecular function level, consistent with the finding reported by Tomecki *et al* [68] that MM-associated *DIS3* mutations cause perturbations in cellular RNA metabolism. Furthermore, these data, although requiring confirmation in independent patients' series, support the hypothesis that the role of *DIS3* mutations in cancer development could be mediated by the alteration of gene expression, that could be a direct effect or a consequence of the overabundance (resulting from *DIS3* dysfunction and actually detected in our mutated samples) of unstable non-coding transcripts affecting chromatin structure and thus the expression of genes important for carcinogenesis [68].

Furthermore, looking at *DIS3* mutations during disease progression in serially analyzed patients, we observed a similar scenario to that found by Bolli *et al* [20], who exclusively reported clonal variants at both time points or acquired/increased variants in the late sample, consistent with the expected

positive selection of mutated subclones; only one MM case of the present series showed a decrease in mutation burden during tumor progression (with an allele frequency of 35.48% at MM onset and 10% at post-therapy control), and the disappearance of a high frequency mutation was never observed in our cohort.

The finding that the identified *DIS3* mutations did not have clinical impact was consistent with what reported by Weissbach *et al* [137], who found a trend towards a lower median overall survival (OS) but no effect on event-free survival (EFS) in patients with mutations in *DIS3* compared with those without. Moreover, no differences in response to therapy were observed [137]. In our study, we tested the association of *DIS3* mutations with clinical outcome in 16 pPCL patients for whom the follow-up was available. In agreement with Weissbach *et al* [137], mutational events did not show any impact neither on PFS nor OS. Likewise, the response rate after a 4-cycle therapy did not associate with the occurrence of *DIS3* mutations.

These findings, however, have to be treated with caution as, despite the robust size of the study cohort, the actual number of patients with *DIS3* mutations is quite small. Thus, clearly larger, multicentre efforts are required to better understand the oncogenic and clinical implications of *DIS3* mutations in MM.

Reference List

1. Gutierrez NC, Garcia-Sanz R, San Miguel JF. Molecular biology of myeloma. Clin.Transl.Oncol. 2007;9:618-624.
2. Kyle RA, Rajkumar SV. Multiple myeloma. N.Engl.J.Med. 2004;351:1860-1873.
3. Rajkumar SV, Kyle RA. Multiple myeloma: diagnosis and treatment. Mayo Clin.Proc. 2005;80:1371-1382.
4. Landgren O, Weiss BM. Patterns of monoclonal gammopathy of undetermined significance and multiple myeloma in various ethnic/racial groups: support for genetic factors in pathogenesis. Leukemia 2009;23:1691-1697.
5. Vij R, Siegel DS, Jagannath S et al. An open-label, single-arm, phase 2 study of single-agent carfilzomib in patients with relapsed and/or refractory multiple myeloma who have been previously treated with bortezomib. Br.J.Haematol. 2012;158:739-748.
6. Rajkumar SV. Multiple myeloma: 2011 update on diagnosis, risk-stratification, and management. Am.J.Hematol. 2011;86:57-65.
7. International Myeloma Working Group. Criteria for the classification of monoclonal gammopathies, multiple myeloma and related disorders: a report of the International Myeloma Working Group. Br.J.Haematol. 2003;121:749-757.
8. Kyle RA, Rajkumar SV. Criteria for diagnosis, staging, risk stratification and response assessment of multiple myeloma. Leukemia 2009;23:3-9.
9. Rajkumar SV. MGUS and Smoldering Multiple Myeloma: Update on Pathogenesis, Natural History, and Management. Hematology.Am.Soc.Hematol.Educ.Program. 2005340-345.
10. Podar K, Tai YT, Lin BK et al. Vascular endothelial growth factor-induced migration of multiple myeloma cells is associated with beta 1 integrin- and phosphatidylinositol 3-kinase-dependent PKC alpha activation. J.Biol.Chem. 2002;277:7875-7881.
11. Hideshima T, Mitsiades C, Tonon G, Richardson PG, Anderson KC. Understanding multiple myeloma pathogenesis in the bone marrow to identify new therapeutic targets. Nat.Rev.Cancer 2007;7:585-598.
12. Roodman GD. Pathogenesis of myeloma bone disease. Leukemia 2009;23:435-441.
13. Palumbo A, Anderson K. Multiple myeloma. N.Engl.J.Med. 2011;364:1046-1060.
14. Weiss BM, Abadie J, Verma P, Howard RS, Kuehl WM. A monoclonal gammopathy precedes multiple myeloma in most patients. Blood 2009;113:5418-5422.
15. Landgren O, Kyle RA, Pfeiffer RM et al. Monoclonal gammopathy of undetermined significance (MGUS) consistently precedes multiple myeloma: a prospective study. Blood 2009;113:5412-5417.
16. Kyle RA, Therneau TM, Rajkumar SV et al. Prevalence of monoclonal gammopathy of undetermined significance. N.Engl.J.Med. 2006;354:1362-1369.
17. Dispenzieri A, Katzmann JA, Kyle RA et al. Prevalence and risk of progression of light-chain monoclonal gammopathy of undetermined significance: a retrospective population-based cohort study. Lancet 2010;375:1721-1728.
18. Drexler HG, Matsuo Y. Malignant hematopoietic cell lines: in vitro models for the study of multiple myeloma and plasma cell leukemia. Leuk.Res. 2000;24:681-703.
19. Gabrea A, Martelli ML, Qi Y et al. Secondary genomic rearrangements involving immunoglobulin or MYC loci show similar prevalences in hyperdiploid and nonhyperdiploid myeloma tumors. Genes Chromosomes.Cancer 2008;47:573-590.
20. Bolli N, vet-Loiseau H, Wedge DC et al. Heterogeneity of genomic evolution and mutational profiles in multiple myeloma. Nat.Commun. 2014;5:2997.
21. Egan JB, Shi CX, Tembe W et al. Whole-genome sequencing of multiple myeloma from diagnosis to plasma cell leukemia reveals genomic initiating events, evolution, and clonal tides. Blood 2012;120:1060-1066.
22. Keats JJ, Chesi M, Egan JB et al. Clonal competition with alternating dominance in multiple myeloma. Blood 2012;120:1067-1076.
23. Ding L, Ellis MJ, Li S et al. Genome remodelling in a basal-like breast cancer metastasis and xenograft. Nature 2010;464:999-1005.
24. Yachida S, Jones S, Bozic I et al. Distant metastasis occurs late during the genetic evolution of pancreatic cancer. Nature 2010;467:1114-1117.
25. Ding L, Ley TJ, Larson DE et al. Clonal evolution in relapsed acute myeloid leukaemia revealed by whole-genome sequencing. Nature 2012;481:506-510.

26. Walker BA, Wardell CP, Melchor L et al. Intraclonal heterogeneity is a critical early event in the development of myeloma and precedes the development of clinical symptoms. Leukemia 2014;28:384-390.
27. Melchor L, Brioli A, Wardell CP et al. Single-cell genetic analysis reveals the composition of initiating clones and phylogenetic patterns of branching and parallel evolution in myeloma. Leukemia 2014;28:1705-1715.
28. Morgan GJ, Walker BA, Davies FE. The genetic architecture of multiple myeloma. Nat.Rev.Cancer 2012;12:335-348.
29. Brioli A, Melchor L, Cavo M, Morgan GJ. The impact of intra-clonal heterogeneity on the treatment of multiple myeloma. Br.J.Haematol. 2014;165:441-454.
30. Rollig, C., Knop, S., and Bomhauser, M. Multiple myeloma. Lancet Epub ahead of print. 2014. Ref Type: Generic
31. Bergsagel PL, Kuehl WM, Zhan F et al. Cyclin D dysregulation: an early and unifying pathogenic event in multiple myeloma. Blood 2005;106:296-303.
32. Hurt EM, Wiestner A, Rosenwald A 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;5:191-199.
33. Zingone A, Cultraro CM, Shin DM et al. Ectopic expression of wild-type FGFR3 cooperates with MYC to accelerate development of B-cell lineage neoplasms. Leukemia 2010;24:1171-1178.
34. Hideshima T, Bergsagel PL, Kuehl WM, Anderson KC. Advances in biology of multiple myeloma: clinical applications. Blood 2004;104:607-618.
35. Fonseca R, Debes-Marun CS, Picken EB et al. The recurrent IgH translocations are highly associated with nonhyperdiploid variant multiple myeloma. Blood 2003;102:2562-2567.
36. Kumar S, Fonseca R, Ketterling RP et al. Trisomies in multiple myeloma: impact on survival in patients with high-risk cytogenetics. Blood 2012;119:2100-2105.
37. Chang H, Qi X, Yeung J et al. Genetic aberrations including chromosome 1 abnormalities and clinical features of plasma cell leukemia. Leuk.Res. 2009;33:259-262.
38. Le BP, Leroux D, Dascalescu C et al. Novel evidence of a role for chromosome 1 pericentric heterochromatin in the pathogenesis of B-cell lymphoma and multiple myeloma. Genes Chromosomes.Cancer 2001;32:250-264.
39. Walker BA, Leone PE, Chiecchio L et al. A compendium of myeloma-associated chromosomal copy number abnormalities and their prognostic value. Blood 2010;116:e56-e65.
40. Qazilbash MH, Saliba RM, Ahmed B et al. Deletion of the short arm of chromosome 1 (del 1p) is a strong predictor of poor outcome in myeloma patients undergoing an autotransplant. Biol.Blood Marrow Transplant. 2007;13:1066-1072.
41. Shaughnessy JD, Jr., Qu P, Usmani S et al. Pharmacogenomics of bortezomib test-dosing identifies hyperexpression of proteasome genes, especially PSMD4, as novel high-risk feature in myeloma treated with Total Therapy 3. Blood 2011;118:3512-3524.
42. Zhan F, Colla S, Wu X et al. CKS1B, overexpressed in aggressive disease, regulates multiple myeloma growth and survival through. Blood 2007;109:4995-5001.
43. Chapman MA, Lawrence MS, Keats JJ et al. Initial genome sequencing and analysis of multiple myeloma. Nature 2011;471:467-472.
44. Boyd KD, Ross FM, Walker BA 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. Clin.Cancer Res. 2011;17:7776-7784.
45. Fonseca R, Oken MM, Harrington D et al. Deletions of chromosome 13 in multiple myeloma identified by interphase FISH usually denote large deletions of the q arm or monosomy. Leukemia 2001;15:981-986.
46. Avet-Louseau H, Daviet A, Sauner S, Bataille R. Chromosome 13 abnormalities in multiple myeloma are mostly monosomy 13. Br.J.Haematol. 2000;111:1116-1117.
47. Chiecchio L, Protheroe RK, Ibrahim AH et al. Deletion of chromosome 13 detected by conventional cytogenetics is a critical prognostic factor in myeloma. Leukemia 2006;20:1610-1617.
48. Keats JJ, Reiman T, Maxwell CA et al. In multiple myeloma, t(4;14)(p16;q32) is an adverse prognostic factor irrespective of FGFR3 expression. Blood 2003;101:1520-1529.
49. Fonseca R, Bergsagel PL, Drach J et al. International Myeloma Working Group molecular classification of multiple myeloma: spotlight review. Leukemia 2009;23:2210-2221.
50. Neri A, Baldini L, Trecca D et al. p53 gene mutations in multiple myeloma are associated with advanced forms of malignancy. Blood 1993;81:128-135.
51. Corradini P, Inghirami G, Astolfi M et al. Inactivation of tumor suppressor genes, p53 and Rb1, in plasma cell dyscrasias. Leukemia 1994;8:758-767.

52. Tiedemann RE, Gonzalez-Paz N, Kyle RA et al. Genetic aberrations and survival in plasma cell leukemia. Leukemia 2008;22:1044-1052.
53. Fonseca R, Blood E, Rue M et al. Clinical and biologic implications of recurrent genomic aberrations in myeloma. Blood 2003;101:4569-4575.
54. Avet-Loiseau H, Attal M, Moreau P et al. Genetic abnormalities and survival in multiple myeloma: the experience of the Intergroupe Francophone du Myelome. Blood 2007;109:3489-3495.
55. Annunziata CM, Davis RE, Demchenko Y et al. Frequent engagement of the classical and alternative NF-kappaB pathways by diverse genetic abnormalities in multiple myeloma. Cancer Cell 2007;12:115-130.
56. Keats JJ, Fonseca R, Chesi M et al. Promiscuous mutations activate the noncanonical NF-kappaB pathway in multiple myeloma. Cancer Cell 2007;12:131-144.
57. Bergsagel PL, Kuehl WM. Molecular pathogenesis and a consequent classification of multiple myeloma. J Clin.Oncol. 2005;23:6333-6338.
58. Agnelli L, Bicciato S, Mattioli M et al. Molecular classification of multiple myeloma: a distinct transcriptional profile characterizes patients expressing CCND1 and negative for 14q32 translocations. J.Clin.Oncol. 2005;23:7296-7306.
59. Shaughnessy JD, Jr., Zhan F, Burington BE 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;109:2276-2284.
60. Johnsen JM, Nickerson DA, Reiner AP. Massively parallel sequencing: the new frontier of hematologic genomics. Blood 2013;122:3268-3275.
61. Lohr JG, Stojanov P, Carter SL et al. Widespread genetic heterogeneity in multiple myeloma: implications for targeted therapy. Cancer Cell 2014;25:91-101.
62. Ley TJ, Mardis ER, Ding L et al. DNA sequencing of a cytogenetically normal acute myeloid leukaemia genome. Nature 2008;456:66-72.
63. Lee W, Jiang Z, Liu J et al. The mutation spectrum revealed by paired genome sequences from a lung cancer patient. Nature 2010;465:473-477.
64. Agnelli L, Neri A. Next generation sequencing in muliple myeloma: insights into the molecular heterogeneity of the disease. International Journal of Hematologic Oncology 2014;3:367-376.
65. Leich E, Weissbach S, Klein HU et al. Multiple myeloma is affected by multiple and heterogeneous somatic mutations in adhesion- and receptor tyrosine kinase signaling molecules. Blood Cancer J. 2013;3:e102.
66. Tiacci E, Trifonov V, Schiavoni G et al. BRAF mutations in hairy-cell leukemia. N.Engl.J.Med. 2011;364:2305-2315.
67. Hebraud B, Leleu X, Lauwers-Cances V 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;28:675-679.
68. Tomecki R, Drazkowska K, Kucinski I 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 Res. 2014;42:1270-1290.
69. Shaffer AL, Emre NC, Lamy L et al. IRF4 addiction in multiple myeloma. Nature 2008;454:226-231.
70. Fazzari MJ, Greally JM. Epigenomics: beyond CpG islands. Nat.Rev.Genet. 2004;5:446-455.
71. Chi P, Allis CD, Wang GG. Covalent histone modifications--miswritten, misinterpreted and mis-erased in human cancers. Nat.Rev.Cancer 2010;10:457-469.
72. Walker BA, Wardell CP, Chiecchio L et al. Aberrant global methylation patterns affect the molecular pathogenesis and prognosis of multiple myeloma. Blood 2011;117:553-562.
73. Kim JY, Kee HJ, Choe NW et al. Multiple-myeloma-related WHSC1/MMSET isoform RE-IIBP is a histone methyltransferase with transcriptional repression activity. Mol.Cell Biol. 2008;28:2023-2034.
74. Marango J, Shimoyama M, Nishio H et al. The MMSET protein is a histone methyltransferase with characteristics of a transcriptional corepressor. Blood 2008;111:3145-3154.
75. Nimura K, Ura K, Shiratori H et al. A histone H3 lysine 36 trimethyltransferase links Nkx2-5 to Wolf-Hirschhorn syndrome. Nature 2009;460:287-291.
76. Rouhi A, Mager DL, Humphries RK, Kuchenbauer F. MiRNAs, epigenetics, and cancer. Mamm.Genome 2008;19:517-525.
77. Lionetti M, Biasiolo M, Agnelli L et al. Identification of microRNA expression patterns and definition of a microRNA/mRNA regulatory network in distinct molecular groups of multiple myeloma. Blood 2009;114:e20-e26.
78. Gutierrez NC, Sarasquete ME, Misiewicz-Krzeminska I et al. Deregulation of microRNA expression in the different genetic subtypes of multiple myeloma and correlation with gene expression profiling. Leukemia 2010;24:629-637.

79. Pichiorri F, Suh SS, Ladetto M et al. MicroRNAs regulate critical genes associated with multiple myeloma pathogenesis. Proc.Natl.Acad.Sci.U.S.A 2008;105:12885-12890.
80. Pichiorri F, Suh SS, Rocci A et al. Downregulation of p53-inducible microRNAs 192, 194, and 215 Impairs the p53/MDM2 Autoregulatory Loop in Multiple Myeloma Development. Cancer Cell 2010;18:367-381.
81. Chang TC, Yu D, Lee YS et al. Widespread microRNA repression by Myc contributes to tumorigenesis. Nat.Genet. 2008;40:43-50.
82. Raab MS, Podar K, Breitkreutz I, Richardson PG, Anderson KC. Multiple myeloma. Lancet 2009;374:324-339.
83. Rajkumar SV. Multiple myeloma. Curr.Probl.Cancer 2009;33:7-64.
84. Singhal S, Mehta J, Desikan R et al. Antitumor activity of thalidomide in refractory multiple myeloma. N.Engl.J.Med. 1999;341:1565-1571.
85. Richardson PG, Blood E, Mitsiades CS et al. A randomized phase 2 study of lenalidomide therapy for patients with relapsed or relapsed and refractory multiple myeloma. Blood 2006;108:3458-3464.
86. Richardson PG, Barlogie B, Berenson J et al. A phase 2 study of bortezomib in relapsed, refractory myeloma. N.Engl.J.Med. 2003;348:2609-2617.
87. Richardson PG, Sonneveld P, Schuster MW et al. Bortezomib or high-dose dexamethasone for relapsed multiple myeloma. N.Engl.J.Med. 2005;352:2487-2498.
88. Albarracin F, Fonseca R. Plasma cell leukemia. Blood Rev. 2011;25:107-112.
89. Kyle RA, Maldonado JE, Bayrd ED. Plasma cell leukemia. Report on 17 cases. Arch.Intern.Med. 1974;133:813-818.
90. Dimopoulos MA, Palumbo A, Delasalle KB, Alexanian R. Primary plasma cell leukaemia. Br.J.Haematol. 1994;88:754-759.
91. Noel P, Kyle RA. Plasma cell leukemia: an evaluation of response to therapy. Am.J.Med. 1987;83:1062-1068.
92. Avet-Loiseau H, Daviet A, Brigaudeau C et al. Cytogenetic, interphase, and multicolor fluorescence in situ hybridization analyses in primary plasma cell leukemia: a study of 40 patients at diagnosis, on behalf of the Intergroupe Francophone du Myelome and the Groupe Francais de Cytogenetique Hematologique. Blood 2001;97:822-825.
93. Taniwaki M, Nishida K, Ueda Y, Takashima T. Non-random chromosomal rearrangements and their implications in clinical features and outcome of multiple myeloma and plasma cell leukemia. Leuk.Lymphoma 1996;21:25-30.
94. Garcia-Sanz R, Orfao A, Gonzalez M et al. Primary plasma cell leukemia: clinical, immunophenotypic, DNA ploidy, and cytogenetic characteristics. Blood 1999;93:1032-1037.
95. Fonseca R, Barlogie B, Bataille R et al. Genetics and cytogenetics of multiple myeloma: a workshop report. Cancer Res. 2004;64:1546-1558.
96. Chang H, Qi X, Trieu Y et al. Multiple myeloma patients with CKS1B gene amplification have a shorter progression-free survival post-autologous stem cell transplantation. Br.J.Haematol. 2006;135:486-491.
97. Chng WJ, Price-Troska T, Gonzalez-Paz N et al. Clinical significance of TP53 mutation in myeloma. Leukemia 2007;21:582-584.
98. Gertz MA, Lacy MQ, Dispenzieri A et al. Clinical implications of t(11;14)(q13;q32), t(4;14)(p16.3;q32), and -17p13 in myeloma patients treated with high-dose therapy. Blood 2005;106:2837-2840.
99. Bezieau S, Devilder MC, vet-Loiseau H et al. High incidence of N and K-Ras activating mutations in multiple myeloma and primary plasma cell leukemia at diagnosis. Hum.Mutat. 2001;18:212-224.
100. Ahuja HG, Foti A, Bar-Eli M, Cline MJ. The pattern of mutational involvement of RAS genes in human hematologic malignancies determined by DNA amplification and direct sequencing. Blood 1990;75:1684-1690.
101. Ortega MM, Faria RM, Shitara ES et al. N-RAS and K-RAS gene mutations in Brazilian patients with multiple myeloma. Leuk.Lymphoma 2006;47:285-289.
102. Ramsingh G, Mehan P, Luo J, Vij R, Morgensztern D. Primary plasma cell leukemia: a Surveillance, Epidemiology, and End Results database analysis between 1973 and 2004. Cancer 2009;115:5734-5739.
103. Mahindra A, Kalaycio ME, Vela-Ojeda J et al. Hematopoietic cell transplantation for primary plasma cell leukemia: results from the Center for International Blood and Marrow Transplant Research. Leukemia 2012;26:1091-1097.
104. Hayman SR, Fonseca R. Plasma cell leukemia. Curr.Treat.Options.Oncol. 2001;2:205-216.
105. Musto P, Rossini F, Gay F et al. Efficacy and safety of bortezomib in patients with plasma cell leukemia. Cancer 2007;109:2285-2290.
106. Benson DM, Jr., Smith MK. Effectiveness of lenalidomide (Revlimid) for the treatment of plasma cell leukemia. Leuk.Lymphoma 2007;48:1423-1425.
107. Chlebowski A, Lubas M, Jensen TH, Dziembowski A. RNA decay machines: the exosome. Biochim.Biophys.Acta 2013;1829:552-560.

108. Liu Q, Greimann JC, Lima CD. Reconstitution, activities, and structure of the eukaryotic RNA exosome. Cell 2006;127:1223-1237.
109. Lorentzen E, Dziembowski A, Lindner D, Seraphin B, Conti E. RNA channelling by the archaeal exosome. EMBO Rep. 2007;8:470-476.
110. Koonin EV, Wolf YI, Aravind L. Prediction of the archaeal exosome and its connections with the proteasome and the translation and transcription machineries by a comparative-genomic approach. Genome Res. 2001;11:240-252.
111. Symmons MF, Jones GH, Luisi BF. A duplicated fold is the structural basis for polynucleotide phosphorylase catalytic activity, processivity, and regulation. Structure. 2000;8:1215-1226.
112. Zuo Y, Deutscher MP. Exoribonuclease superfamilies: structural analysis and phylogenetic distribution. Nucleic Acids Res. 2001;29:1017-1026.
113. Schaeffer D, Reis FP, Johnson SJ, Arraiano CM, van HA. The CR3 motif of Rrp44p is important for interaction with the core exosome and exosome function. Nucleic Acids Res. 2012;40:9298-9307.
114. Lebreton A, Tomecki R, Dziembowski A, Seraphin B. Endonucleolytic RNA cleavage by a eukaryotic exosome. Nature 2008;456:993-996.
115. Schneider C, Leung E, Brown J, Tollervey D. The N-terminal PIN domain of the exosome subunit Rrp44 harbors endonuclease activity and tethers Rrp44 to the yeast core exosome. Nucleic Acids Res. 2009;37:1127-1140.
116. Schaeffer D, Tsanova B, Barbas A et al. The exosome contains domains with specific endoribonuclease, exoribonuclease and cytoplasmic mRNA decay activities. Nat.Struct.Mol.Biol. 2009;16:56-62.
117. Dziembowski A, Lorentzen E, Conti E, Seraphin B. A single subunit, Dis3, is essentially responsible for yeast exosome core activity. Nat.Struct.Mol.Biol. 2007;14:15-22.
118. 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. Mol.Cell 2008;29:717-728.
119. Tomecki R, Kristiansen MS, Lykke-Andersen S et al. The human core exosome interacts with differentially localized processive RNases: hDIS3 and hDIS3L. EMBO J. 2010;29:2342-2357.
120. Staals RH, Bronkhorst AW, Schilders G et al. Dis3-like 1: a novel exoribonuclease associated with the human exosome. EMBO J. 2010;29:2358-2367.
121. Chang HM, Triboulet R, Thornton JE, Gregory RI. A role for the Perlman syndrome exonuclease Dis3l2 in the Lin28-let-7 pathway. Nature 2013;497:244-248.
122. Lubas M, Damgaard CK, Tomecki R et al. Exonuclease hDIS3L2 specifies an exosome-independent 3'-5' degradation pathway of human cytoplasmic mRNA. EMBO J. 2013;32:1855-1868.
123. Astuti D, Morris MR, Cooper WN et al. Germline mutations in DIS3L2 cause the Perlman syndrome of overgrowth and Wilms tumor susceptibility. Nat.Genet. 2012;44:277-284.
124. Briggs MW, Burkard KT, Butler JS. Rrp6p, the yeast homologue of the human PM-Scl 100-kDa autoantigen, is essential for efficient 5.8 S rRNA 3' end formation. J.Biol.Chem. 1998;273:13255-13263.
125. Januszyk K, Liu Q, Lima CD. Activities of human RRP6 and structure of the human RRP6 catalytic domain. RNA. 2011;17:1566-1577.
126. Midtgaard SF, Assenholt J, Jonstrup AT et al. Structure of the nuclear exosome component Rrp6p reveals an interplay between the active site and the HRDC domain. Proc.Natl.Acad.Sci.U.S.A 2006;103:11898-11903.
127. Tomecki R, Drazkowska K, Dziembowski A. Mechanisms of RNA degradation by the eukaryotic exosome. Chembiochem. 2010;11:938-945.
128. Mukherjee D, Gao M, O'Connor JP et al. The mammalian exosome mediates the efficient degradation of mRNAs that contain AU-rich elements. EMBO J. 2002;21:165-174.
129. Gherzi R, Lee KY, Briata P et al. A KH domain RNA binding protein, KSRP, promotes ARE-directed mRNA turnover by recruiting the degradation machinery. Mol.Cell 2004;14:571-583.
130. Allmang C, Kufel J, Chanfreau G et al. Functions of the exosome in rRNA, snoRNA and snRNA synthesis. EMBO J. 1999;18:5399-5410.
131. Gudipati RK, Xu Z, Lebreton A et al. Extensive degradation of RNA precursors by the exosome in wild-type cells. Mol.Cell 2012;48:409-421.
132. Orban TI, Izaurralde E. Decay of mRNAs targeted by RISC requires XRN1, the Ski complex, and the exosome. RNA. 2005;11:459-469.
133. Basu U, Meng FL, Keim C et al. The RNA exosome targets the AID cytidine deaminase to both strands of transcribed duplex DNA substrates. Cell 2011;144:353-363.
134. Schneider C, Anderson JT, Tollervey D. The exosome subunit Rrp44 plays a direct role in RNA substrate recognition. Mol.Cell 2007;27:324-331.
135. Ibrahim H, Wilusz J, Wilusz CJ. RNA recognition by 3'-to-5' exonucleases: the substrate perspective. Biochim.Biophys.Acta 2008;1779:256-265.

136. Walker BA, Wardell CP, Melchor L et al. Intraclonal heterogeneity and distinct molecular mechanisms characterize the development of t(4;14) and t(11;14) myeloma. Blood 2012;120:1077-1086.
137. Weissbach S, Langer C, Puppe B et al. The molecular spectrum and clinical impact of DIS3 mutations in multiple myeloma. Br.J.Haematol. 2014
138. Ho AS, Kannan K, Roy DM et al. The mutational landscape of adenoid cystic carcinoma. Nat.Genet. 2013;45:791-798.
139. de Groen FL, Krijgsman O, Tijssen M et al. Gene-dosage dependent overexpression at the 13q amplicon identifies DIS3 as candidate oncogene in colorectal cancer progression. Genes Chromosomes.Cancer 2014;53:339-348.
140. Rose AE, Poliseno L, Wang J et al. Integrative genomics identifies molecular alterations that challenge the linear model of melanoma progression. Cancer Res. 2011;71:2561-2571.
141. Camps J, Pitt JJ, Emons G et al. Genetic amplification of the NOTCH modulator LNX2 upregulates the WNT/beta-catenin pathway in colorectal cancer. Cancer Res. 2013;73:2003-2013.
142. Pils D, Tong D, Hager 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.
143. Rehmsmeier M, Steffen P, Hochsmann M, Giegerich R. Fast and effective prediction of microRNA/target duplexes. RNA. 2004;10:1507-1517.
144. Todoerti K, Agnelli L, Fabris S et al. Transcriptional characterization of a prospective series of primary plasma cell leukemia revealed signatures associated with tumor progression and poorer outcome. Clin.Cancer Res. 2013;19:3247-3258.
145. Musto P, D'Auria F, Petrucci MT 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;118:2925.
146. 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.
147. Mattioli M, Agnelli L, Fabris S et al. Gene expression profiling of plasma cell dyscrasias reveals molecular patterns associated with distinct IGH translocations in multiple myeloma. Oncogene 2005;24:2461-2473.
148. Fabris S, Agnelli L, Mattioli M et al. Characterization of oncogene dysregulation in multiple myeloma by combined FISH and DNA microarray analyses. Genes Chromosomes.Cancer 2005;42:117-127.
149. Verdelli D, Nobili L, Todoerti K et al. Molecular events underlying interleukin-6 independence in a subclone of the CMA-03 multiple myeloma cell line. Genes Chromosomes.Cancer 2014;53:154-167.
150. Dai M, Wang P, Boyd AD et al. Evolving gene/transcript definitions significantly alter the interpretation of GeneChip data. Nucleic Acids Res. 2005;33:e175.
151. Tusher VG, Tibshirani R, Chu G. Significance analysis of microarrays applied to the ionizing radiation response. Proc.Natl.Acad.Sci.U.S.A 2001;98:5116-5121.
152. Agnelli L, Fabris S, Bicciato S et al. Upregulation of translational machinery and distinct genetic subgroups characterise hyperdiploidy in multiple myeloma. Br.J.Haematol. 2007;136:565-573.
153. Agnelli L, Mosca L, Fabris S 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;48:603-614.
154. Fabris S, Ronchetti D, Agnelli L et al. Transcriptional features of multiple myeloma patients with chromosome 1q gain. Leukemia 2007;21:1113-1116.
155. Bird JM, Owen RG, D'Sa S et al. Guidelines for the diagnosis and management of multiple myeloma 2011. Br.J.Haematol. 2011;154:32-75.
156. van de Donk NW, Lokhorst HM, Anderson KC, Richardson PG. How I treat plasma cell leukemia. Blood 2012;120:2376-2389.
157. Reis FP, Barbas A, Klauer-King AA et al. Modulating the RNA processing and decay by the exosome: altering Rrp44/Dis3 activity and end-product. PLoS.One. 2013;8:e76504.
158. Watson IR, Takahashi K, Futreal PA, Chin L. Emerging patterns of somatic mutations in cancer. Nat.Rev.Genet. 2013;14:703-718.
159. Rashid NU, Sperling AS, Bolli N et al. Differential and limited expression of mutant alleles in multiple myeloma. Blood 2014;124:3110-3117.