Integrative analysis of the genomic and transcriptomic landscape of double-refractory multiple myeloma

Bachisio Ziccheddu,[1,2] Giulia Biancon,[3] Filippo Bagnoli,[1,3] Chiara De Philippis,[1,3] Francesco Maura,[3,4] Even H. Rustad,[4] Matteo Dugo,[5] Andrea Devecchi,[5] Loris De Cecco,[5] Marialuisa Sensi,[5] Carolina Terragna,[6] Marina Martello,[7] Tina Bagratuni,[8] Efstathios Kastritis,[8] Meletios A. Dimopoulos,[8] Michele Cavo,[6] Cristiana Carniti,[1] Vittorio Montefusco,[1] Paolo Corradini,[1,3] and Niccolo Bolli[1,3]

[1]Department of Clinical Oncology and Hematology, Fondazione IRCCS Istituto Nazionale dei Tumori, Milan, Italy; [2]Department of Molecular Biotechnologies and Health Sciences, University of Turin, Turin, Italy; [3]Department of Oncology and Onco-Hematology, University of Milan, Milan, Italy; [4]Memorial Sloan Kettering Cancer Center, New York, NY; [5]Department of Applied Research and Technological Development, Fondazione IRCCS Istituto Nazionale dei Tumori, Milan, Italy; [6]Istituto di Ematologia "Seràgnoli," Azienda Ospedaliero-Universitaria Sant'Orsola-Malpighi, Bologna, Italy; [7]Seràgnoli Institute of Hematology, Bologna University School of Medicine, Bologna, Italy; and [8]Plasma Cell Dyscrasia Unit, Department of Clinical Therapeutics, School of Medicine, Alexandra Hospital, National and Kapodistrian University of Athens, Athens, Greece

Key Points

- Double-refractory samples showed high subclonal heterogeneity and a distinct spectrum of mutations, aneuploidies, and mutational signatures.

- Resistance may be polyclonal and dictated by a more complex genomic architecture rather than mutations or expression of drug-target genes.

In multiple myeloma, novel treatments with proteasome inhibitors (PIs) and immunomodulatory agents (IMiDs) have prolonged survival but the disease remains incurable. At relapse, next-generation sequencing has shown occasional mutations of drug targets but has failed to identify unifying features that underlie chemotherapy resistance. We studied 42 patients refractory to both PIs and IMiDs. Whole-exome sequencing was performed in 40 patients, and RNA sequencing (RNA-seq) was performed in 27. We found more mutations than were reported at diagnosis and more subclonal mutations, which implies ongoing evolution of the genome of myeloma cells during treatment. The mutational landscape was different from that described in published studies on samples taken at diagnosis. The TP53 pathway was the most frequently inactivated (in 45% of patients). Conversely, point mutations of genes associated with resistance to IMiDs were rare and were always subclonal. Refractory patients were uniquely characterized by having a mutational signature linked to exposure to alkylating agents, whose role in chemotherapy resistance and disease progression remains to be elucidated. RNA-seq analysis showed that treatment or mutations had no influence on clustering, which was instead influenced by karyotypic events. We describe a cluster with both amp(1q) and del(13) characterized by CCND2 upregulation and also overexpression of MCL1, which represents a novel target for experimental treatments. Overall, high-risk features were found in 65% of patients. However, only amp(1q) predicted survival. Gene mutations of IMiD and PI targets are not a preferred mode of drug resistance in myeloma. Chemotherapy resistance of the bulk tumor population is likely attained through differential, yet converging evolution of subclones that are overall variable from patient to patient and within the same patient.

Introduction

Over the last 2 decades, novel therapeutic approaches in multiple myeloma (MM), mainly represented by proteasome inhibitors (PIs) and immunomodulatory drugs (IMiDs), have produced deeper and more durable remissions.[1] However, relapses are inevitable, and patients will eventually become refractory to both classes of drugs, at which point survival is dismal.[2]

The revised international staging system has incorporated few translocations and copy-number abnormalities (CNAs) that have prognostic potential in risk stratification of newly diagnosed patients,[3] which confirms the utility of genotyping for clinical purposes. Next-generation sequencing (NGS) techniques have greatly expanded our view of the genomic makeup of newly diagnosed MM (NDMM).[4-7] But very little value is attributed to genomic lesions in predicting response to specific drugs, because even the few classic druggable mutations in MM (eg, BRAFV600E) are often found at a subclonal level.[5] Current risk definitions may need to be redefined because modern effective combination therapies seem to have the potential to abrogate the negative prognostic effect of some chromosomal abnormalities.[8] More recently, NGS-based studies have suggested that extended genotyping is feasible and can offer improved prognostication at diagnosis.[9-13]

In relapsed or refractory MM (RRMM), a small number of mutations implicated in resistance to PIs or IMiDs have been identified through custom targeted gene panels or in occasional patients.[14-16] These mutations have been validated in in vitro models, but their clinical relevance needs additional study. Expression levels and splice isoforms of the IMiD target CRBN have been proposed to mediate resistance to those drugs, but results are conflicting and are not ready for the clinic.[17] Furthermore, the few RNA sequencing (RNA-seq) studies in MM so far have found little correlation of genomic and transcriptomic findings.[18] Finally, in vitro genetic screens have suggested candidate targets that confer drug resistance.[19-21] In light of all this uncertainty, NGS studies have not provided any real clinical value to help make treatment decisions in the relapse setting.

We and others have shown that MM is characterized by the presence of subclonal heterogeneity and branching evolution from the very early preclinical stages all the way to relapse.[5,7,22-28] At diagnosis, treatment results in a selective pressure on the tumor population, which leads to the selection of resistant subclones. This suggests that chemotherapy resistance is carried by a fraction of myeloma cells that may be missed at diagnosis or that evolve during treatment. In the RRMM setting, broader sequencing approaches thus have the potential to better inform the catalog of genomic alterations that mediate drug resistance by the acquisition of the drug targets, by their downstream effectors, or by mutations of alternative oncogenic pathways.

In this study, we performed whole-exome sequencing (WES) in a highly selected cohort of patients refractory to both PIs and IMiDs and also performed RNA-seq in a subset of them, looking for genomic and phenotypic correlates of drug response.

Materials and methods

Sample selection

All patients enrolled in our study carried either refractory or relapsed-and-refractory (R/R) disease defined as documented disease progression during or within 60 days of completing the last therapy. All patients were refractory to both PIs and IMiDs and had R/R disease.

The samples were obtained at the Fondazione IRCCS Istituto Nazionale dei Tumori of Milan, Istituto di Ematologia "L.A. Seràgnoli" (Azienda Ospedaliero-Universitaria Sant'Orsola-Malpighi, Bologna, Italy) and National and Kapodistrian University of Athens in accordance with a protocol approved by the local ethics committee after informed consent has been obtained from each patient (Institutional Review Board INT 15/14).

We processed tumors from CD138$^+$ bone marrow plasma cells and matched germline samples from buccal swabs of DNA from 59 patients (118 total samples). Seventeen patients were excluded from the study because their tumor and normal samples were either mismatched or contaminated. Thus, our analysis was restricted 42 patients.

DNA extraction and library preparation

For plasma cells, 8 mL of a bone marrow aspirate were subjected to red cell lysis in NH_4Cl 0.15M, $KHCO_3$ 10 nM, and Na_4EDTA 1 nM (to pH 7.2). CD138$^+$ cells were then selected by immunomagnetic bead separation (Miltenyi Biotec). The CD138$^+$ fraction was frozen as viable cells in 90% fetal bovine serum and 10% dimethyl sulfoxide. For DNA extraction, cells were thawed and rinsed twice in phosphate-buffered saline; DNA was then extracted using the Nucleospin Tissue kit (Macherey-Nagel). A minimum of 200 ng of high-quality DNA was considered suitable for analysis. DNA amount was assessed by using a fluorimetric technique (Qubit 2.0, Thermo Fisher), and DNA integrity was assessed with an Agilent 2100 Bioanalyzer. For buccal swabs, DNA was obtained with the Nucleospin Tissue kit (Macherey-Nagel).

High-molecular-weight DNA was sheared to smaller fragments before preparing the library. DNA fragments were prepared using a Covaris M220 sonicator. After sonication, DNA integrity was checked again with the Agilent 2100 Bioanalyzer to confirm the desired molecular weight of 150-300 bp.

Sheared DNA was used to prepare indexed libraries according to Agilent SureSelect XT2 protocol (SureSelectXT2 Target Enrichment System for Illumina Paired-End Multiplexed Sequencing). Briefly, Illumina adapters and patient tags were ligated to the ends of the DNA fragments, which were then amplified with 7 cycles of polymerase chain reaction (PCR) before being captured. After indexing and amplification, libraries were pooled in equimolar amounts for a total of 1500 ng per pool. Each pool of 8 libraries was then subjected to target enrichment through hybridization with biotin-tagged 120-mer complementary RNA probes (Agilent Sure-Select Human All-Exon V6) at 65°C degrees for 24 hours. Pulldown was performed with streptavidin-coated magnetic beads. Nine cycles of postcapture PCR were performed, and DNA was then purified with AMPure XP beads.

Sequencing

We performed sequencing with the NextSeq500 sequencer (Illumina) and obtained an average of 15.63 Gb per patient on a 75-bp paired-end protocol. Raw fastq files were quality-controlled with FastQC.[29] Then, paired-end reads were aligned to the reference human genome (hg19) using a Burrows-Wheeler Aligner (BWA-MEM, v0.7.12).[30] Duplicate reads were removed with Picard software (http://broadinstitute.github.io/picard/), and unmapped reads were removed with samtools v1.3.1.[31] The reads were then post-processed according to Genome Analysis Toolkit (GATK) best practices 3.7,[32] which include left alignment of small insertions and deletions, indel realignment, and base quality score recalibration. Finally, we assessed the concordance and cross-sample contamination using the Conpair tool, which was developed

for detecting samples swaps and cross-individual contamination in paired samples. It can measure contamination levels as low as 0.1%, even in the presence of copy number changes.[33] Variant call format (VCF) files have been deposited in European Genome-phenome Archive (EGA), accession number EGAD00001005098. WES and RNA-seq data from NDMM patients were also gathered from the CoMMpass data set, which is part of the Multiple Myeloma Research Foundation Personalized Medicine Initiatives (https://research.themmrf.org and www.themmrf.org).

Detection of mutations and copy-number abnormalities

Somatic single nucleotide variants (SNVs) were called with 3 different variant callers: MuTect2 (v3.7),[34] CaVEMan (v1.13.2),[35] and MuSE (v1.0).[36] Small indels were called with the MuTect2 algorithm. To create a high-confidence variant list, we chose the variants called by at least 2 algorithms. Somatic variants (substitutions and indels) were annotated with Oncotator v1.9.9.0.[37] We then performed additional filtering to remove false positives and polymorphisms. In particular, we excluded variants based on at least 1 of the following filters: read depth was <30 in both normal and tumor samples; call was unidirectional; an alternative allele was present in matched normal if the variant was not listed in the Catalogue Of Somatic Mutations In Cancer (COSMIC) database; the C>A/G>T variants had a frequency <0.1 (oxoG artifacts); or variants were annotated in polymorphism databases (ExAC,[38] NCBI dbSNP,[39] 1000 Genome Project[40]) without a COSMIC annotation. Significantly mutated genes were identified by using the MutSigCV algorithm.[41] The analysis was performed using the Gene Pattern Web interface (http://genepattern.broadinstitute.org). $P < .005$ was used to create the list of the most frequently mutated genes.

CNAs were called with Excavator2 v1.1,[42] an algorithm that detects copy-number variations from WES data by integrating the analysis of on-target and off-target reads. Hyperdiploid samples were identified by the presence of 2 or more trisomies involving chromosomes 3, 5, 7, 9, 11, 15, 19, or 21. The genome regions that were significantly modified in our samples were identified by using GISTIC v2.0.[43] The analysis was executed using the Gene Pattern Web interface (http://genepattern.broadinstitute.org) and setting a q value threshold of 0.01.

Subclonal analysis using ABSOLUTE v2.0[44] was performed in each patient by using the cancer cell fraction (CCF) of each mutation, corrected by purity of the sample and ploidy of the locus to return more solid information compared with the analysis of raw allelic frequency. This computational method starts from previous information such as segmentation copy number data, cancer karyotype, and somatic point mutations data to infer absolute ploidy, tumor purity, and subclonal heterogeneity of each sample. The software computes possible solutions in each sample, characterized by a combination of purity, ploidy, and subclonal fraction ranked by probability. The user then manually chooses the most likely solution based on the software output and on the features of each patient. This is a more authoritative way to address subclonality.

Mutational signatures

Mutational signatures were investigated by using the implemented algorithm in MutationalPatterns R package,[45] as previously described.[46] This process exploits the non-negative matrix factorization approach for extracting known and unknown mutational signatures. The contribution of the shortlist of identified signatures to each case was then quantified by using the bespoke mmSig R code (https://github.com/evenrus/mmsig).

RNA-seq

RNA was extracted with the Qiagen RNeasy kit and quality controlled using the Agilent Bioanalyzer and TapeStation systems. RNA-seq libraries were produced using the Illumina TruSeq RNA Access Library Prep Kit (Illumina). Samples were sequenced with a plexing of 8 per high-output flowcell in the Illumina NextSeq500 system. Raw files generated from the sequencer were quality controlled with FastQC. Then, RNA-seq reads were mapped to the reference human genome (hg19) using the STAR aligner[47] and the parameters were set to count read numbers per gene while mapping. To analyze the gene expression profile, we applied the voom/limma pipeline.[48] First, the data set of raw counts was filtered to remove genes with <10 reads in >95% of samples. Then, we performed the trimmed mean of M-values (TMM) normalization.[49] This method estimates a scale factor used to reduce technical bias between samples that result from differences in library size. We then applied the voom transformation to convert the raw counts in log2-counts per million (log2-CPM) and calculated the respective observation-level weights to be used in differential expression analysis. P values were corrected for multiple testing by using the Benjamini-Hochberg false discovery rate method. We performed unsupervised hierarchical clustering with Euclidean distance and Ward's linkage. To compare expression levels in our cohort and those in the public CoMMpass database, we started from raw counts for both data sets. After TMM normalization and voom transformation in log2-CPM, we performed a z score normalization to allow the standardization of data across experiments and sequencing platforms for candidate gene analysis as in Cheadle et al.[50] Results were also confirmed by Singscore, an orthogonal method that uses a rank-based scoring system to score individual samples independently, which provides stable scores that are less likely to be affected by varying sample size and unwanted variations across samples.[51] Then we assessed the impact of mutations on expression of the same gene (cis analysis) or of other genes (trans analysis) with a hierarchical Bayesian approach implemented in the xseq R package.[52] This method works with 4 input files: a patient-gene expression matrix, a patient-gene mutation data set, patient CNV data, and a graph containing known interactions between genes. The outputs are the posterior probability that a mutation in a given gene in an individual patient influences gene expression within that patient (P[F]) and the posterior probability that mutations in a given gene influences gene expression across all patients (P[D]). The list of gene interactions used was downloaded from the Shah Lab algorithm Web page (https://shahlab.ca/projects/xseq/).

To call variants on RNA-seq data, including SNVs and indels, we used the GATK best practice workflow for RNA-seq 3.7 (https://software.broadinstitute.org/gatk/documentation). This includes removal of reads marked with the "N" symbol, hard-clipping of regions overhanging into introns, base recalibration, and variant calling with the HaplotypeCaller function. Called variants were annotated with Oncotator. We then applied the same filters as for variants obtained by WES and cross-referenced the results with the mutations identified in the previous WES analysis. We defined a gene as being not expressed or as being at a low expression level if it had a log2-CPM <3.

Table 1. Clinical features of cohort patients

Factor	n/N	%	Median	Range
Age at diagnosis, y			61	43-78
No. of lines of therapy before sampling			3	1-9
Time from diagnosis to sampling, mo			52.3	3-165
Follow-up after sampling, mo			14.6	
Best response after sampling				
VGPR	5/41	12.2		
PR	14/41	34.2		
MR	6/41	14.6		
SD	7/41	17.1		
PD	8/41	19.5		
Lost to follow-up	1/41	2.4		
Relapse	40/41	97.6		
Death	37/41	90.2		
Cause of death				
Progressive disease	34/37	91.9		
Treatment related mortality	2/37	5.4		

MR, minor response; PD, progressive disease; PR, partial response; SD, stable disease; VGPR, very good partial response.

Statistical analysis

The comparison tests have been performed with unpaired 2-tailed Student *t* tests. When the comparisons concerned paired samples, we used a paired 2-tailed Student *t* test. When the comparison included dichotomic variables, we used a Fisher's exact test. The P values were adjusted using the Bonferroni method. The exact number of hypotheses tested in each panel was used to correct the P value. An adjusted $P < .05$ was considered statistically significant. The correlations were computed using Pearson's correlation coefficient.

Association of categorical variables with both progression-free disease (PFS) and overall survival (OS) was performed in a univariable fashion using Kaplan-Meier curves and a log-rank test. P values were corrected for multiple testing using the Bonferroni method. All analyses were performed in R, the language and environment for statistical computing (R Core Team, 2018).

Results

Composition of the cohort and mutational spectrum

Our cohort included 42 heavily pretreated patients with a median age of 61 years. Patients were uniformly refractory to both PIs and IMiDs and received a median of 3 lines of treatment before sampling, which occurred a median of 52.3 months after diagnosis (Table 1). Three of 42 patients were primary refractory to both drugs. All patients were refractory to the last line of treatment, which consisted of a PI in 17, an IMiD in 15, both in 3, and other regimens (mostly alkylators) in 7 patients (supplemental Table 1).

Forty samples were sequenced by WES to an average depth of 97× (supplemental Figure 1A-C). We found a median of 77.5 somatic variants per sample (range, 22-333 somatic variants), and a total of 3683 variants (Figure 1A; supplemental Table 2). We then applied ABSOLUTE[44] to infer the subclonal composition of

samples: all samples showed evidence of subclonal mutations, and 37.5% of them had more subclonal than clonal mutations (Figure 1A). This was not influenced by purity or coverage of the samples (supplemental Figure 2A-B; supplemental Table 3). An increased number of mutations correlated positively with the number of subclonal mutations (Figure 1B), implying that hypermutation is a feature of late subclones that cause ongoing evolution, even in this highly selected cohort of double-refractory samples. Compared with NDMM patients from the CoMMpass data set, our RRMM samples showed an increased number of mutations (Figure 1C). The increase of mutations at relapse was also shown by paired analysis of pretreatment and relapse samples in the CoMMpass data set and in a recent publication by Bolli et al[5] in our group (supplemental Figure 2C-D), confirming recently published reports.[53] Perhaps not surprisingly, *TP53* emerged as the most commonly mutated gene with 10 mutations in 7 double-refractory patients (Figure 1D). The gene mutation spectrum was otherwise not different from those in previously published studies. We also found mutations previously implicated in resistance to IMiDs (ie, 2 in *CRBN*, 1 in *IKZF3*) in some patients in our cohort, but their overall number was very low when compared with the 100% IMiD refractory population in our study. Analysis of recurrently mutated genes did not show novel genes in RRMM compared with NDMM, and only *KRAS*, *TP53*, and *NRAS* showed significance (Figure 1D).

Genomic makeup of treatment-resistant samples

We combined fluorescence in situ hybridization (FISH) analysis and WES to obtain information on CNAs and translocations (Figure 2A). Analysis of significant areas of recurrent copy number changes (supplemental Figure 3) showed an excess in gain(1q) and del(17p) in our refractory cohort (45% and 35%, respectively). By combining mutational data with structural abnormalities, we found that deletions and amplifications of driver genes were frequent events. For tumor suppressors, we found frequent instances of biallelic inactivation through mutations and deletions that were particularly notable for *TP53* in which 6 of 7 patients showed both a mutated and a deleted allele (Figure 2A-B).[23,54] Combining CNAs and mutations, our selected cohort of double-resistant patients showed inactivation of the TP53 pathway through mutations and/or deletions in 45% of the patients (Figure 2C), more than what was previously reported in NDMM. In paired samples from the CoMMpass cohort, we also observed the same phenomenon (supplemental Figure 4A). Notably, we also found a significant increase in *TP53* deletions in our own patients when FISH data at diagnosis were compared with the *TP53* copy-number at the time of sampling in the double-refractory phase (supplemental Figure 4B). This likely represents a key mediator of the chemotherapy resistance observed in the cohort. By aggregating evidence from refractory patients and from in vitro screens, we created a shortlist of genes potentially implicated in IMiD and PI resistance (supplemental Table 4); altogether, mutations and deletions in IMiD resistance genes were found in 32.5% of patients, regardless of the last line of treatment. Mutations were rare. In particular, we found *IKZF3* mutations in 1 patient and *CRBN* mutations in 2, always at the subclonal level (Figure 2A). Conversely, deletions were more frequent, especially in *RBX1*, a component of the E3 ubiquiting ligase complex in chr22q, which was co-deleted with *XBP1* in 9 patients (Figure 2A). Interestingly, very few nonrecurring mutations or aneuploidies were found in proteasome subunit genes in our cohort, confirming their rarity.[16,55] We assessed high-risk genetic
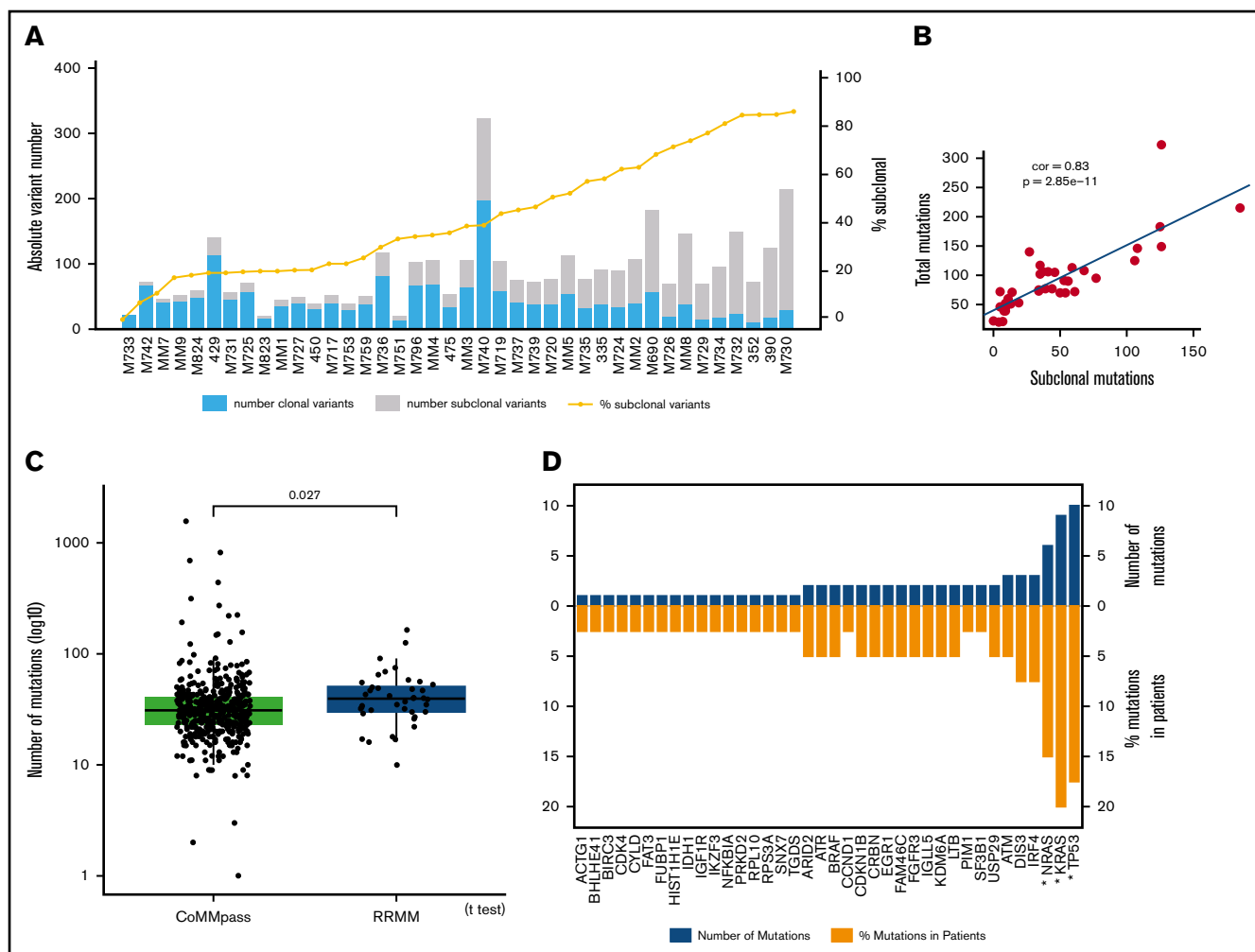
Figure 1. Mutations and subclonality in the cohort. (A) Stacked bar chart showing the number of variants per patient (left y-axis). Bars indicate the contributions of clonal variants and subclonal variants. Yellow dots represent the percentage of subclonal variants per patient (right y-axis). (B) Scatter plot representing correlation between the number of subclonal mutations and the total number of mutations. Each dot represents a patient. The correlation was computed by using Pearson's correlation coefficient. (C) Box plot representing the distribution of variant number in NDMM (samples pertain to the CoMMpass data set) and RRMM. First, second, and third quartiles are represented by horizontal bars, and the whiskers point to the 1.5 interquartile range of the upper quartile and lower quartile. To make this comparison, our cohort was re-analyzed with MuTect, Strelka, and Seurat to ensure accuracy of the analysis. (D) Waterfall plot showing the overall number of mutations for the most commonly mutated genes and their prevalence (% of mutations) in patients. Genes mutated at a statistically significant rate are indicated by an asterisk.

features (including those described in the Revised International Staging System [R-ISS][3]), double-hit events,[11] apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like (APOBEC) signature contribution,[13] TP53 mutations,[9] and CRBN pathway mutations[15] and found that at least 1 such event was present in 65% of patients (Figure 2D).

When the cohort was split into patients receiving IMiDs or PIs as the last treatment, the only significant difference in the mutational spectrum was in TP53 mutations, found only in patients receiving PIs as the last line of treatment (Figure 2E). We found no difference when we assessed primary refractory and R/R patients.

Mutational signatures

We next asked whether our treatment-resistant patients who showed an increased number of mutations and defective DNA

repair pathways also showed a distinct mutational signature profile. To this end, we performed a de novo extraction on the 96-class profile of mutations in the cohort (Figure 3A), followed by assignment using the COSMIC catalog of mutational signatures and finally fitting the identified signatures. This approach allows precise quantification of the contribution of each signature and the identification of novel ones.[46] We found that even at this late stage, the mutational catalog is mainly composed of mutational processes attributable to cell aging (signatures 1 and 5) and APOBEC (signatures 2 and 13)[56] (Figure 3B). Interestingly, we confirmed the activity of a novel signature (still not included in the COSMIC database) first reported in MM genomes as MM-1[28]; it was later believed to be linked to exposure to alkylating agents.[57] Not surprisingly, this signature was well represented in our heavily treated cohort, in which most patients were exposed to oral or intravenous alkylating agents at some point in the past. Its activity was not correlated with the cumulative dose of

Figure 2. Genomic makeup of RRMM. (A) Oncoplot columns showing genomic alterations for the patients in the study: cytogenetic events are in red in the top part (absence of information [NA] in white). Mutations are in the bottom part, color-coded by type. Subclonal mutations have a yellow square in the middle, and biallelic instances of mutations or deletions are outlined in black. (B) Circos plots summarizing the genomic makeup of each patient. Genomic coordinates on chromosomes are on the outer ring; then, working inward, copy number (CN) status (green, deletion or loss; red, amplification or gain), mutations (variants), and interchromosomal rearrangements. (C) Bar chart of

melphalan or cyclophosphamide administered to the patients when both oral and intravenous administration were considered (supplemental Figure 5A-B). Interestingly, it was significantly higher in people who underwent autologous transplantation before sampling and thus received a single high dose of melphalan (Figure 3C).

Transcriptomic profile of the cohort

In 27 of 42 patients (including 25 with whole-exome data available), RNA from CD138[+] cells was available and RNA-seq was performed. We obtained a median of 69.9 million reads per sample to extract data on the global transcriptome as well as expressed mutations.

We asked whether variation in gene expression across patients could somehow be correlated with the presence of driver mutations or cytogenetic abnormalities. To obtain an overview of the main patterns of expression changes, we first performed an analysis of principal components (Figure 4A). The first 2 principal components of the analysis accounted for 9.4% and 8.7% of the variability (supplemental Figure 6). Globally, expression changes were rarely correlated with genomic abnormalities, with the exception of immunoglobulin heavy chain translocations that clearly separated samples mostly along PC1. Second, we performed an unsupervised clustering of samples, in which the optimal number of clusters was determined to be 5 by consensus clustering analysis. Despite the limited value of clustering in our small cohort, we could observe distinct groups of patients segregating mainly on the basis of cytogenetic features (Figure 4B): in particular, the most distinctive group of patients was characterized by the simultaneous presence of del(13) and amp(1q) in the context of immunoglobulin heavy chain translocations. Most of these patients received lenalidomide as a last line of treatment, which failed. Although overall differences in gene expression were quite limited between clusters, patients with concomitant del(13) and amp(1q) showed a significant number of deregulated genes, with overexpression of CCND2 as the most prominent one. This cluster also showed EGR1 and CD79a downregulation and MCL1 overexpression. Gene set enrichment analysis (GSEA) on this cluster showed significant downregulation of genes involved in NF-kB signaling downstream of tumor necrosis factor (TNF) (supplemental Figure 7A; supplemental Table 5). Although MCL1 overexpression can be explained by chr(1q) amplification, upregulation of CCND2 was previously described in patients with MAF translocations.[58-60] Here, only 1 patient carried a t(14;16); therefore, other high-risk genetic features must underlie this observation. Interestingly, no significant gene dysregulation was observed when either CNA was considered in isolation, confirming that this is a fairly homogeneous subgroup characterized by a significant interaction between genetic lesions (supplemental Figure 7B).

Gene expression analysis could also carry predictive value and might correlate with drug response. For example, CRBN expression levels have been inconsistently linked to IMiD sensitivity.[61,62] In our data set, the type of last treatment did not correlate with any cluster (Figure 4B), in line with the high number of subclonal mutations

observed by WES and suggesting that treatment does not result in the selection of a transcriptionally homogeneous clone. Consistent with this observation, we did not observe significant differences in expression levels of genes associated with resistance to PIs or IMiDs among patients who showed refractoriness to either drug as a last line of treatment, including CRBN (supplemental Figure 8). Next, we compared z scores and rank scores of candidate genes at diagnosis and at the double-refractory stage by using data from the IA12a release of the CoMMpass study. This again confirmed no general downregulation of CRBN at the RNA level, whereas XBP1 expression levels were significantly lower in our RRMM stage (supplemental Figure 9).

Expressed mutations

We next asked to what extent RNA-seq data would show expression of the genomic spectrum of mutations observed. We found that an average of only 26.3% of mutations per sample were expressed, confirming previous data on a smaller data set[18] (Figure 5A). Expression levels of the mutation positively correlated with its CCF by WES (Figure 5B), and this was especially true for mutations with higher CCF. As expected, the probability of a mutation being expressed positively correlated with the overall expression levels of that gene (Figure 5C). Overall, mutations of driver genes such as KRAS, NRAS, and IRF4 were mostly expressed (Figure 5D). Notable exceptions were a TP53 mutation that caused nonsense-mediated decay in M725. Overall, mutations that were not expressed were the ones with the lowest cancer cell fraction in DNA sequencing, and this was independent of expression levels of the gene.

Finally, to systematically explore transcriptomic consequences of gene mutations, we applied the Bayesian xseq package.[52] In our small cohort, we found that the only mutated gene associated with a significant transcriptomic deregulation was TP53: in fact, patients with mutations often showed BCL2 upregulation and PTEN and BAX downregulation, suggesting that these mutations are linked with activation of other cancer-associated pathways and could be used as surrogates of linked gene expression levels (Figure 5E).

Survival analysis

Clinical data and adequate follow-up were available for 41 of 42 patients. Median follow-up after sampling was 14.6 months. In all, 97.6% of patients relapsed and 90.2% died during follow-up, with median PFS of 189 days after sampling and a median OS of 440 days after sampling, in line with the literature[63] (Figure 6A-B). Best response from sampling was very good partial response or better in 12.2%, partial response (PR) in 34.2%, minor response or stable disease in 31.7%, and progressive disease in 19.5%, highlighting the chemotherapy resistance of the cohort (Table 1). Transcriptomic clusters did not show distinct survival, whereas we found that patients who harbored chr(1q) amplification, but not a single chr(1q) gain, showed a poorer PFS and a trend toward worse OS (Figure 6C-D). Conversely, patients with TP53 mutations and chr(17p) deletions did not have a worse prognosis

---

Figure 2. (continued) the most commonly mutated pathways in the cohort. MAPK_pathway: KRAS, NRAS, BRAF, and FGFR3 mutations; NF-κB pathway: CYLD, BIRC2, BIR3, TRAF2, TRAF3, NFKBIA, and NFKBIE mutations and/or deletions; CRBN pathway: IKZF1 and IKZF3 mutations and CRBN, RBX1, DDB1, and CUL4B mutations and/or deletions; proteasome subunit: mutations in proteasome subunit genes; and TP53 pathway: TP53, ATM, and ATR mutations and/or deletions. (D) Breakdown of the patients based on the presence or absence of high-risk features; left-most bar, % contribution of each feature. (E) Frequency of mutations according to the last line of treatment.
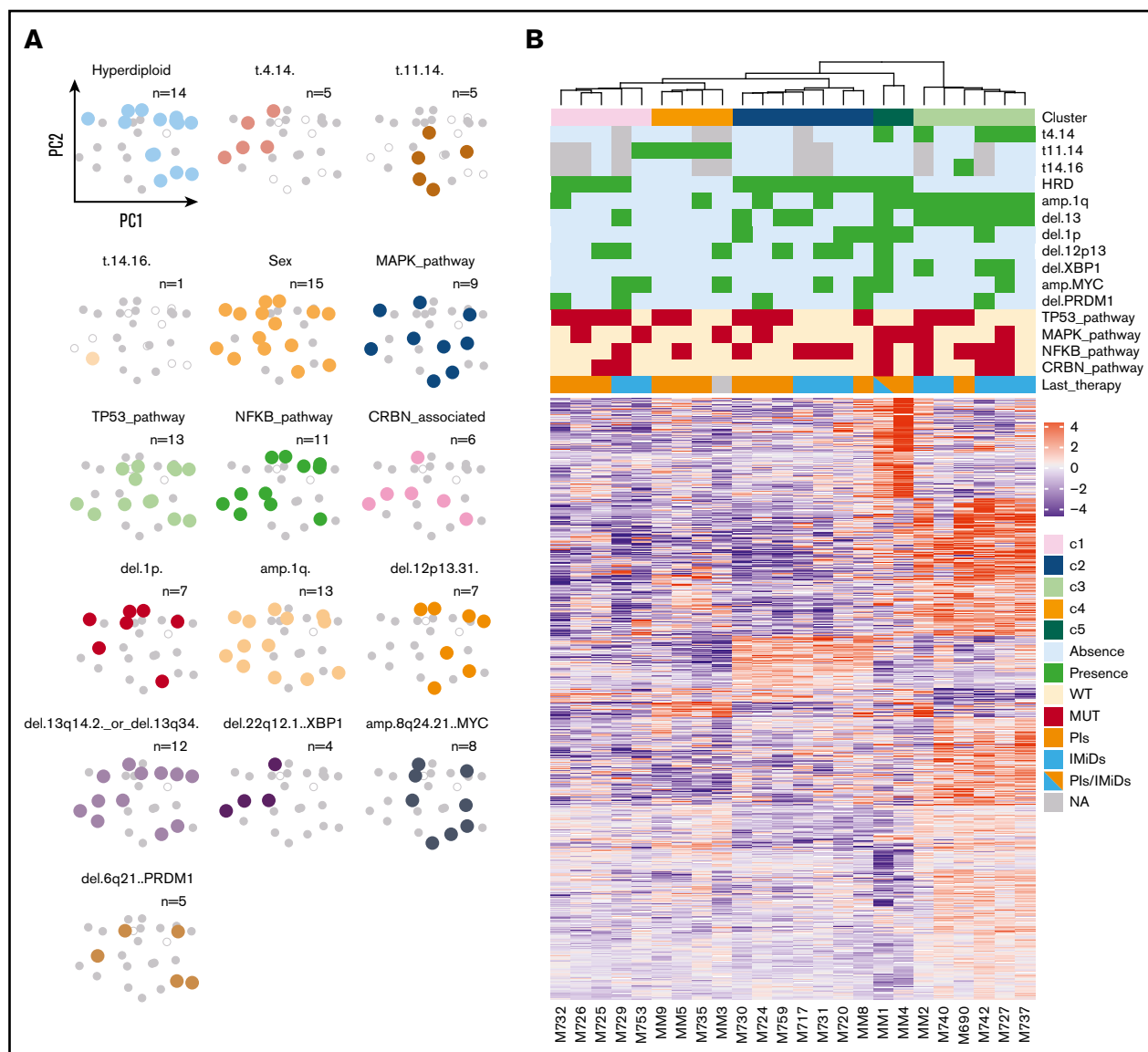
Figure 3. Mutational signatures in RRMM. (A) A 96-class plot of mutational frequencies in the pyrimidine context, color-coded by nucleotide change and further sub-divided on the basis of the 5′ and 3′ nucleotides. (B) De novo extraction of the mutational signatures contributing to the spectrum. (C) Proportion of mutations caused by signature MM1 in people who did not receive (blue) or did receive (yellow) high-dose melphalan before sampling.

Figure 4. Transcriptomic profile of RRMM patients. (A) Principal component analysis based on genetic features: samples are represented as dots in the space identified by the 2 principal components and are color-coded based on the presence or absence of the genetic feature listed in the top label: color: present; gray, absent; empty, information not available. (B) Unsupervised clustering of the samples based on their gene expression profile. Genomic features are annotated above the horizontal white line (green, cytogenetic events; red, mutations), as is last treatment (yellow, bortezomib; light blue, lenalidomide).

(Figure 6E-F). Gene mutations were found to have little prognostic significance in NDMM, and this was also the case at the RRMM stage. One exception was the *PRDM1* gene: when mutated or deleted, patients had a trend toward worse OS (Figure 6G-H). PRDM1 is a crucial transcription factor in plasma cells and its mutations-deletions have been observed as late events in MM development,[64] with a possible prognostic role, even at diagnosis.[9]

We had data on pomalidomide-based salvage treatment of 14 patients. Of these, 4 showed primary refractoriness, including 3 with loss of chr13 and del(17p). Interestingly, 1 patient with a *CRBN* mutation and 1 with an *IKZF3* mutation were treated with pomalidomide and showed sensitivity to the drug, with a PFS in line with that of the rest of the cohort.

Discussion

Here we performed WES and RNA-seq in a highly selected population that homogeneously showed refractoriness to both IMiDs and PIs. Contrary to previous studies, our approach allowed us to study the genomic and transcriptomic landscape of double-resistant myeloma in deeper detail.

We unexpectedly found that the subclonal composition is very rich in these late stages. Previous studies showed anecdotal evidence of 1 subclone taking over at disease relapse and accounting for the bulk of refractory disease.[65] Here, one-third of samples showed more subclonal than clonal mutations, implying ongoing acquisition of novel mutations and genomic heterogeneity, even at late stages.
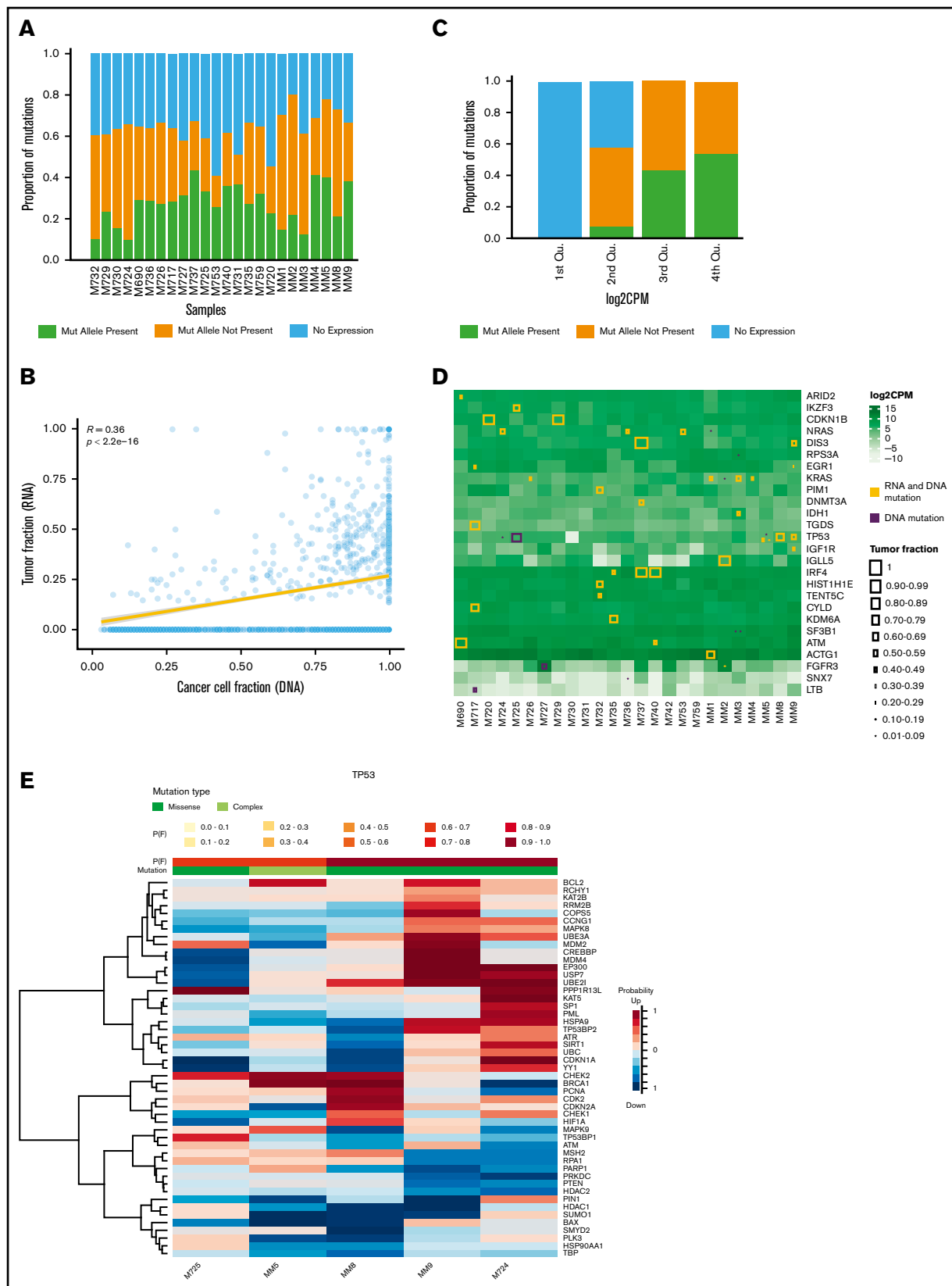
Figure 5. Expression of gene mutations. (A) Stacked bar chart showing the proportion of mutations for each patient that are expressed (present) or not because the mutated (mut) allele is not expressed (not present) or the gene is not expressed at all. (B) Scatter plot showing a correlation between the cancer cell fraction by DNA analysis and the tumor fraction by RNA analysis, the latter representing the ratio between mutant and wild-type transcripts. (C) Stacked bar chart showing the proportion of mutations

Indeed, this heterogeneity may explain the PRs observed to subsequent treatments, but almost half of our cohort did not even achieve a PR after sampling. Here, it is likely that different subclones attained drug resistance through convergent evolution, a well-described phenomenon in MM.[5,24,28] Supporting this hypothesis, known mutations conferring resistance to IMiDs were always found at the subclonal level, which therefore did not entirely explain resistance to treatment in the few patients carrying them. Consistently, the 2 patients with *CRBN* and *IKZF3* mutations were not primary refractory to pomalidomide. Furthermore, when stratified by last line of treatment, no common genomic pattern was found across patients, again reinforcing the hypothesis that resistance is not mediated by a major clone with a treatment-specific genotype that mediates resistance. Even in the context of spatial heterogeneity,[66] gene mutations are likely not a preferred mode of developing chemotherapy resistance, because recent studies of patients who relapsed during lenalidomide maintenance showed similar low rates of mutation in drug target genes.[53] In contrast, high-risk features were clearly enriched in our samples, including chr(1q) amplification,[11] biallelic inactivation of tumor suppressors,[23] and particularly *TP53* pathway inactivation,[67] which are likely responsible for the chemotherapy resistance we observed. Although genomic analyses are increasingly performed at relapse to determine whether high-risk features are present, we still lack clear correlates of subsequent drug response. Only larger studies with uniformly treated patients will have the potential to provide results that could be incorporated into clinical practice in the future.

WES analyses are limited in their ability to discover mutational signatures because of a limited view of the genome. However, we were able to confirm activation-induced cytidine deaminase and APOBEC activity in MM and to identify an additional mutational process active at this late stage. The signature of this process has been linked to exposure to alkylating agents,[57] and it was enriched in patients who underwent autologous transplantation. This indicates that the relapsing clone has been exposed to melphalan and therefore relapse arose from residual disease at the time of transplant rather than from contaminating clonal plasma cells in the graft. Likely, not only the dose but also the mode of exposure to alkylating agents is relevant in the activity of this mutational process. Clearly, a pulsed exposure of a few residual cells to high-dose melphalan is expected to result in more visible mutational activity compared with longer exposure of millions of cells to lower doses of drugs (eg, as in first-line triplet therapy with bortezomib, mephalan, and prednisone [VMP]), and this does not necessarily imply positive selection of mutagenized cells. Further studies with more patients will be required to confirm our findings and to clarify whether there is any pathogenetic role of this mutational process or whether it merely represents a sign of drug exposure.

The study of the whole transcriptome returned a fairly heterogeneous picture that paralleled that of genomic analysis. As in NDMM,

we found that patients clustered mostly on the basis of cytogenetic events rather than on gene mutations or treatment. This is likely explained by the high number of subclonal mutations we observed, which underlies ongoing evolution in genetically different cells that may average the RNA-seq signal and make it less distinct from patient to patient. This heterogeneity suggests that single-cell analysis will be needed in the future to functionally dissect heterogeneity and highlight therapeutic vulnerabilities at the subclonal level in RRMM. However, our RNA-seq study also highlighted several interesting observations. For example, our finding that RNA expression levels of the IMiD target *CRBN* are not modulated in drug-resistant stages adds evidence that this may not be a useful biomarker for IMiD sensitivity.[61,68,69] Additional studies will be needed to assess whether downregulation of *XBP1* (encoding a key member of the unfolded protein response linked to PI resistance[70-72]) can be a useful biomarker. In our cohort, the finding that amp(1q) correlated with increased *MCL-1* expression suggests that these patients may respond to MCL1 inhibitors. Finally, we found that most driver gene mutations are expressed, especially when found at a relatively high CCF, suggesting that future strategies of targeted treatment could be feasible in such patients.

To investigate whether prognostic factors used at diagnosis are also relevant at the double-refractory stage, we measured survival from the time of sampling. With the limitations of the small sample size and the different salvage treatments used, we found that not all adverse prognostic markers retained their effect. For example, biallelic inactivation of *TP53* did not have prognostic value. Median PFS for patients with *TP53* wild-type or single-hit mutations was less than 200 days, much less than what was found at diagnosis, likely because of the addition of other high-risk features, which changed the effect size of each. In the future, dynamic prognostic scores could help predict advanced stages. Of course, this would also depend on the subsequent treatment used.

Overall, clonal heterogeneity of MM is best addressed by the use of combination treatment, even in R/R stages. Not surprisingly, the best results in the RRMM setting have come from combination treatments, including monoclonal antibodies[73,74] that act regardless of the genomic makeup of the cancer cell and have pleiotropic effects that also target the microenvironment.[75] By comparing our data with those in studies on newly diagnosed and asymptomatic stages,[5,7,25,27,28,76] MM seems to bear an increasing genomic complexity as it advances from asymptomatic to R/R stages, which includes a higher number of mutations and aneuploidies, as well as the appearance of a new mutational process and a higher level of subclonal mutations compared with samples taken at diagnosis. Supposedly, treatment is more likely to be effective in earlier stages when the MM clonal structure is less complex and high-risk clones are less frequent. This is where the maximum effort should be directed to achieve long-term control of the disease. However,

Figure 5. (continued) that are expressed or not expressed for each quartile (Qu) of gene expression because the mutated allele is not expressed (not present) or the gene is not expressed at all. (D) Heat map for the most commonly mutated genes (in rows) showing their expression level in each patient (columns); whether a mutation is present in both RNA and DNA samples or in DNA samples only; and the cancer cell fraction of that mutation, which is proportional to the size of the square. (E) Heat map showing the probabilities of genes being dysregulated in patients with mutated *TP53*. P(F) denotes the probability that the mutation affects the expression in the same individual; mutation type can be either missense or complex (complex indicates that the patient has more than 1 mutation). The probability scale indicates the probability that the gene is upregulated (red) or downregulated (blue).
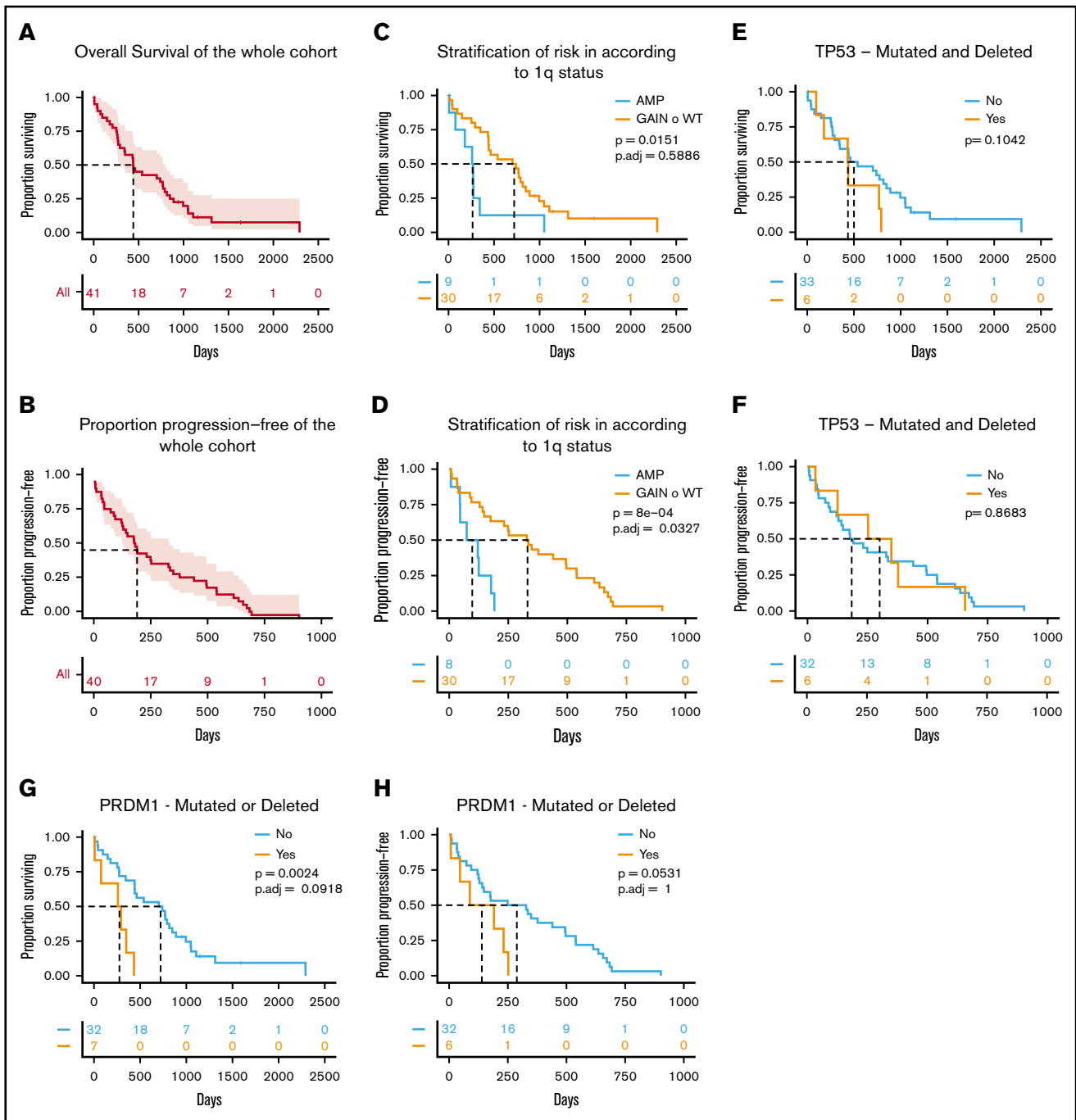
Figure 6. Survival analysis. (A-B) Kaplan-Meier plot of the OS and PFS of the cohort. (C-D) A significant negative effect is observed on PFS (but not on OS after *P* value correction) in patients harboring amplification of chr1q (ie, >3 copies). (E-F) No significant effect is observed in patients showing a double-hit of the TP53 locus. (G-H) A trend toward shorter survival is observed in patients with deletion or mutations of the PRDM1 gene in chr6q. For all plots, numbers of patients at risk are shown in tables below the graphs.

additional research is needed before personalized treatment can be applied to MM based on each patient's genomic and transcriptomic features.

Acknowledgments

This work was supported by the Associazione Italiana Contro le Leucemie-Linfomi e Mieloma, by the Società Italiana di Ematologia

Authorship

The current affiliation for G.B. is Section of Hematology, Department of Internal Medicine and Yale Comprehensive Cancer Center, Yale University School of Medicine, New Haven, CT.

ORCID profiles: C.D.P., 0000-0002-7738-997X; F.M., 0000-0002-5017-1620; M.M., 0000-0001-8196-8707; N.B., 0000-0002-1018-5139.

Correspondence: Niccolo Bolli, Department of Clinical Oncology and Hematology, Fondazione IRCCS Istituto Nazionale dei Tumori and Department of Oncology and Onco-Hematology, University of Milan, Via G. Venezian 1, 20133 Milan, Italy; e-mail: niccolo.bolli@unimi.it.

References

1. Kumar SK, Rajkumar SV, Dispenzieri A, et al. Improved survival in multiple myeloma and the impact of novel therapies. *Blood.* 2008;111(5):2516-2520.

2. Kumar SK, Lee JH, Lahuerta JJ, et al; International Myeloma Working Group. Risk of progression and survival in multiple myeloma relapsing after therapy with IMiDs and bortezomib: a multicenter international myeloma working group study. *Leukemia.* 2012;26(1):149-157.

3. Palumbo A, Avet-Loiseau H, Oliva S, et al. Revised international staging system for multiple myeloma: A report from International Myeloma Working Group. *J Clin Oncol.* 2015;33(26):2863-2869.

4. Chapman MA, Lawrence MS, Keats JJ, et al. Initial genome sequencing and analysis of multiple myeloma. *Nature.* 2011;471(7339):467-472.

5. Bolli N, Avet-Loiseau H, Wedge DC, et al. Heterogeneity of genomic evolution and mutational profiles in multiple myeloma. *Nat Commun.* 2014;5(1):2997.

6. 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(5):1077-1086.

7. Lohr JG, Stojanov P, Carter SL, et al; Multiple Myeloma Research Consortium. Widespread genetic heterogeneity in multiple myeloma: implications for targeted therapy. *Cancer Cell.* 2014;25(1):91-101.

8. Kazandjian D, Korde N, Mailankody S, et al. Remission and progression-free survival in patients with newly diagnosed multiple myeloma treated with carfilzomib, lenalidomide, and dexamethasone: Five-year follow-up of a phase 2 clinical trial. *JAMA Oncol.* 2018;4(12):1781-1783.

9. Bolli N, Biancon G, Moarii M, et al. Analysis of the genomic landscape of multiple myeloma highlights novel prognostic markers and disease subgroups. *Leukemia.* 2018;32(12):2604-2616.

10. Bolli N, Li Y, Sathiaseelan V, et al. A DNA target-enrichment approach to detect mutations, copy number changes and immunoglobulin translocations in multiple myeloma. *Blood Cancer J.* 2016;6(9):e467.

11. Walker BA, Mavrommatis K, Wardell CP, et al. A high-risk, double-hit, group of newly diagnosed myeloma identified by genomic analysis. *Leukemia.* 2019;33(1):159-170.

12. Walker BA, Boyle EM, Wardell CP, et al. Mutational spectrum, copy number changes, and outcome: Results of a sequencing study of patients with newly diagnosed myeloma. *J Clin Oncol.* 2015;33(33):3911-3920.

13. Maura F, Petljak M, Lionetti M, et al. Biological and prognostic impact of APOBEC-induced mutations in the spectrum of plasma cell dyscrasias and multiple myeloma cell lines. *Leukemia.* 2018;32(4):1044-1048.

14. Egan JB, Kortuem KM, Kurdoglu A, et al. Extramedullary myeloma whole genome sequencing reveals novel mutations in Cereblon, proteasome subunit G2 and the glucocorticoid receptor in multi drug resistant disease. *Br J Haematol.* 2013;161(5):748-751.

15. Kortüm KM, Mai EK, Hanafiah NH, et al. Targeted sequencing of refractory myeloma reveals a high incidence of mutations in CRBN and Ras pathway genes. *Blood.* 2016;128(9):1226-1233.

16. Barrio S, Stühmer T, Da-Viá M, et al. Spectrum and functional validation of PSMB5 mutations in multiple myeloma. *Leukemia.* 2019;33(2):447-457.

17. Lodé L, Amiot M, Maïga S, et al. Cereblon expression in multiple myeloma: not ready for prime time. *Br J Haematol.* 2013;163(2):282-284.

18. Rashid NU, Sperling AS, Bolli N, et al. Differential and limited expression of mutant alleles in multiple myeloma. *Blood.* 2014;124(20):3110-3117.

19. Tsvetkov P, Mendillo ML, Zhao J, et al. Compromising the 19S proteasome complex protects cells from reduced flux through the proteasome. *eLife.* 2015;4.

20. Acosta-Alvear D, Cho MY, Wild T, et al. Paradoxical resistance of multiple myeloma to proteasome inhibitors by decreased levels of 19S proteasomal subunits. *eLife.* 2015;4:e08153.

21. Sievers QL, Gasser JA, Cowley GS, Fischer ES, Ebert BL. Genome-wide screen identifies cullin-RING ligase machinery required for lenalidomide-dependent CRL4^{CRBN} activity. *Blood.* 2018;132(12):1293-1303.

22. Magrangeas F, Avet-Loiseau H, Gouraud W, et al. Minor clone provides a reservoir for relapse in multiple myeloma. *Leukemia.* 2013;27(2):473-481.

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

24. 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(8):1705-1715.

25. 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(2):384-390.

26. Raab MS, Lehners N, Xu J, et al. Spatially divergent clonal evolution in multiple myeloma: overcoming resistance to BRAF inhibition. *Blood*. 2016; 127(17):2155-2157.

27. Ledergor G, Weiner A, Zada M, et al. Single cell dissection of plasma cell heterogeneity in symptomatic and asymptomatic myeloma. *Nat Med*. 2018; 24(12):1867-1876.

28. Bolli N, Maura F, Minvielle S, et al. Genomic patterns of progression in smoldering multiple myeloma. *Nat Commun*. 2018;9(1):3363.

29. Babraham Bioinformatics. FASTQC. A quality control tool for high throughput sequence data. https://www.bioinformatics.babraham.ac.uk/projects/fastqc/.

30. Li H, Durbin R. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics*. 2010;26(5):589-595.

31. Li H, Handsaker B, Wysoker A, et al; 1000 Genome Project Data Processing Subgroup. The sequence alignment/map format and SAMtools. *Bioinformatics*. 2009;25(16):2078-2079.

32. McKenna A, Hanna M, Banks E, et al. The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res*. 2010;20(9):1297-1303.

33. Bergmann EA, Chen BJ, Arora K, Vacic V, Zody MC. Conpair: concordance and contamination estimator for matched tumor-normal pairs. *Bioinformatics*. 2016;32(20):3196-3198.

34. Cibulskis K, Lawrence MS, Carter SL, et al. Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. *Nat Biotechnol*. 2013;31(3):213-219.

35. Jones D, Raine KM, Davies H, et al. cgpCaVEManWrapper: Simple execution of CaVEMan in order to detect somatic single nucleotide variants in NGS data. *Curr Protoc Bioinformatics*. 2016;56:15.10.1-15.10.18.

36. Fan Y, Xi L, Hughes DS, et al. MuSE: accounting for tumor heterogeneity using a sample-specific error model improves sensitivity and specificity in mutation calling from sequencing data. *Genome Biol*. 2016;17(1):178.

37. Ramos AH, Lichtenstein L, Gupta M, et al. Oncotator: cancer variant annotation tool. *Hum Mutat*. 2015;36(4):E2423-E2429.

38. Lek M, Karczewski KJ, Minikel EV, et al; Exome Aggregation Consortium. Analysis of protein-coding genetic variation in 60,706 humans. *Nature*. 2016; 536(7616):285-291.

39. Sherry ST, Ward MH, Kholodov M, et al. dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res*. 2001;29(1):308-311.

40. 1000 Genomes Project Consortium, Abecasis GR, Auton A, et al. An integrated map of genetic variation from 1,092 human genomes. *Nature*. 2012; 491(7422):56-65.

41. Lawrence MS, Stojanov P, Polak P, et al. Mutational heterogeneity in cancer and the search for new cancer-associated genes. *Nature*. 2013;499(7457): 214-218.

42. Magi A, Tattini L, Cifola I, et al. EXCAVATOR: detecting copy number variants from whole-exome sequencing data. *Genome Biol*. 2013;14(10):R120.

43. Mermel CH, Schumacher SE, Hill B, Meyerson ML, Beroukhim R, Getz G. GISTIC2.0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers. *Genome Biol*. 2011;12(4):R41.

44. Carter SL, Cibulskis K, Helman E, et al. Absolute quantification of somatic DNA alterations in human cancer. *Nat Biotechnol*. 2012;30(5):413-421.

45. Blokzijl F, Janssen R, van Boxtel R, Cuppen E. MutationalPatterns: comprehensive genome-wide analysis of mutational processes. *Genome Med*. 2018; 10(1):33.

46. Maura F, Degasperi A, Nadeu F, et al. A practical guide for mutational signature analysis in hematological malignancies. *Nat Commun*. 2019;10(1):2969.

47. Dobin A, Davis CA, Schlesinger F, et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*. 2013;29(1):15-21.

48. Law CW, Chen Y, Shi W, Smyth GK. voom: Precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biol*. 2014;15(2): R29.

49. Robinson MD, Oshlack A. A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol*. 2010;11(3):R25-R29.

50. Cheadle C, Vawter MP, Freed WJ, Becker KG. Analysis of microarray data using Z score transformation. *J Mol Diagn*. 2003;5(2):73-81.

51. Foroutan M, Bhuva DD, Lyu R, et al. Single sample scoring of molecular phenotypes. *BMC Bioinformatics*. 2018;19(1):404.

52. Ding J, McConechy MK, Horlings HM, et al. Systematic analysis of somatic mutations impacting gene expression in 12 tumour types. *Nat Commun*. 2015;6(1):8554.

53. Jones JR, Weinhold N, Ashby C, et al; NCRI Haemato-Oncology CSG. Clonal evolution in myeloma: the impact of maintenance lenalidomide and depth of response on the genetics and sub-clonal structure of relapsed disease in uniformly treated newly diagnosed patients. *Haematologica*. 2019;104(7): 1440-1450.

54. Chavan SS, He J, Tytarenko R, et al. Bi-allelic inactivation is more prevalent at relapse in multiple myeloma, identifying RB1 as an independent prognostic marker. *Blood Cancer J*. 2017;7(2):e535.

55. Corre J, Cleynen A, Robiou du Pont S, et al. Multiple myeloma clonal evolution in homogeneously treated patients. *Leukemia*. 2018;32(12):2636-2647.

56. Alexandrov LB, Nik-Zainal S, Wedge DC, et al; ICGC PedBrain. Signatures of mutational processes in human cancer. *Nature*. 2013;500(7463): 415-421.

57. Rustad EH, Yellapantula V, Bolli N, et al. Timing the initiation of multiple myeloma. CANCER-CELL-D-19-00505. SSRN. 2019. https://ssrn.com/abstract=3409453 or http://dx.doi.org/10.2139/ssrn.3409453.

58. 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(15):2461-2473.

59. Bergsagel PL, Kuehl WM. Molecular pathogenesis and a consequent classification of multiple myeloma. *J Clin Oncol*. 2005;23(26):6333-6338.

60. Zhan F, Huang Y, Colla S, et al. The molecular classification of multiple myeloma. *Blood*. 2006;108(6):2020-2028.

61. Qian X, Dimopoulos MA, Amatangelo M, et al. Cereblon gene expression and correlation with clinical outcomes in patients with relapsed/refractory multiple myeloma treated with pomalidomide: an analysis of STRATUS. *Leuk Lymphoma*. 2019;60(2):462-470.

62. Heintel D, Rocci A, Ludwig H, et al. High expression of cereblon (CRBN) is associated with improved clinical response in patients with multiple myeloma treated with lenalidomide and dexamethasone. *Br J Haematol*. 2013;161(5):695-700.

63. Miguel JS, Weisel K, Moreau P, et al. Pomalidomide plus low-dose dexamethasone versus high-dose dexamethasone alone for patients with relapsed and refractory multiple myeloma (MM-003): a randomised, open-label, phase 3 trial. *Lancet Oncol*. 2013;14(11):1055-1066.

64. Maura F, Bolli N, Angelopoulos N, et al. Genomic landscape and chronological reconstruction of driver events in multiple myeloma. *Nat Commun*. 2019; 10(1):3835.

65. Keats JJ, Chesi M, Egan JB, et al. Clonal competition with alternating dominance in multiple myeloma. *Blood*. 2012;120(5):1067-1076.

66. Rasche L, Chavan SS, Stephens OW, et al. Spatial genomic heterogeneity in multiple myeloma revealed by multi-region sequencing. *Nat Commun*. 2017;8(1):268.

67. Lakshman A, Painuly U, Rajkumar SV, et al. Impact of acquired del(17p) in multiple myeloma. *Blood Adv*. 2019;3(13):1930-1938.

68. Dimopoulos K, Fibiger Munch-Petersen H, Winther Eskelund C, et al. Expression of CRBN, IKZF1, and IKZF3 does not predict lenalidomide sensitivity and mutations in the cereblon pathway are infrequent in multiple myeloma. *Leuk Lymphoma*. 2019;60(1):180-188.

69. Zhu YX, Braggio E, Shi CX, et al. Cereblon expression is required for the antimyeloma activity of lenalidomide and pomalidomide. *Blood*. 2011;118(18): 4771-4779.

70. Leung-Hagesteijn C, Erdmann N, Cheung G, et al. Xbp1s-negative tumor B cells and pre-plasmablasts mediate therapeutic proteasome inhibitor resistance in multiple myeloma. *Cancer Cell*. 2013;24(3):289-304.

71. Ling SCW, Lau EK, Al-Shabeeb A, et al. Response of myeloma to the proteasome inhibitor bortezomib is correlated with the unfolded protein response regulator XBP-1. *Haematologica*. 2012;97(1):64-72.

72. Bagratuni T, Wu P, Gonzalez de Castro D, et al. XBP1s levels are implicated in the biology and outcome of myeloma mediating different clinical outcomes to thalidomide-based treatments. *Blood*. 2010;116(2):250-253.

73. Palumbo A, Chanan-Khan A, Weisel K, et al; CASTOR Investigators. Daratumumab, bortezomib, and dexamethasone for multiple myeloma. *N Engl J Med*. 2016;375(8):754-766.

74. Dimopoulos MA, Oriol A, Nahi H, et al; POLLUX Investigators. Daratumumab, lenalidomide, and dexamethasone for multiple myeloma. *N Engl J Med*. 2016;375(14):1319-1331.

75. van de Donk NWCJ, Usmani SZ. CD38 antibodies in multiple myeloma: Mechanisms of action and modes of resistance. *Front Immunol*. 2018;9:2134.

76. Dutta AK, Fink JL, Grady JP, et al. Subclonal evolution in disease progression from MGUS/SMM to multiple myeloma is characterised by clonal stability. *Leukemia*. 2019;33(2):457-468.