

Article

A Network Analysis of Multiple Myeloma Related Gene Signatures

Yu Liu ¹, Haocheng Yu ¹, Seungyeul Yoo ² , Eunjee Lee ^{1,2}, Alessandro Laganà ³, Samir Parekh ³, Eric E. Schadt ^{1,2}, Li Wang ^{1,2} and Jun Zhu ^{1,2,3,*} 

¹ Sema4, a Mount Sinai Venture, 333 Ludlow St., Stamford, CT 06902, USA

² Department of Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai, 1 Gustave L. Levy Pl, New York, NY 10029, USA

³ The Tisch Cancer Institute, Icahn School of Medicine at Mount Sinai, 1 Gustave L. Levy Pl, New York, NY 10029, USA

* Correspondence: jun.zhu@mssm.edu

Received: 31 July 2019; Accepted: 20 September 2019; Published: 27 September 2019



Abstract: Multiple myeloma (MM) is the second most prevalent hematological cancer. MM is a complex and heterogeneous disease, and thus, it is essential to leverage omics data from large MM cohorts to understand the molecular mechanisms underlying MM tumorigenesis, progression, and drug responses, which may aid in the development of better treatments. In this study, we analyzed gene expression, copy number variation, and clinical data from the Multiple Myeloma Research Consortium (MMRC) dataset and constructed a multiple myeloma molecular causal network (M3CN). The M3CN was used to unify eight prognostic gene signatures in the literature that shared very few genes between them, resulting in a prognostic subnetwork of the M3CN, consisting of 178 genes that were enriched for genes involved in cell cycle (fold enrichment = 8.4, p value = 6.1×10^{-26}). The M3CN was further used to characterize immunomodulators and proteasome inhibitors for MM, demonstrating the pleiotropic effects of these drugs, with drug-response signature genes enriched across multiple M3CN subnetworks. Network analyses indicated potential links between these drug-response subnetworks and the prognostic subnetwork. To elucidate the structure of these important MM subnetworks, we identified putative key regulators predicted to modulate the state of these subnetworks. Finally, to assess the predictive power of our network-based models, we stratified MM patients in an independent cohort, the MMRF-CoMMpass study, based on the prognostic subnetwork, and compared the performance of this subnetwork against other signatures in the literature. We show that the M3CN-derived prognostic subnetwork achieved the best separation between different risk groups in terms of log-rank test p -values and hazard ratios. In summary, this work demonstrates the power of a probabilistic causal network approach to understanding molecular mechanisms underlying the different MM signatures.

Keywords: multiple myeloma; Bayesian network; gene signature; prognostic; treatment response

1. Introduction

Multiple myeloma (MM) is the second most common hematological malignancy that derives from the neoplastic transformation and proliferation of plasma cell in bone marrow [1]. Genetic aberrations, particularly genomics translocations involving immunoglobulins, are common [2,3] and are associated with prognosis [4]. These include t(4;14) (the translocation of IgH enhancer and the region of *FGFR3/MMSET* [5]), and the deletion of 17p or 17p13 [6,7]. Other putative prognostic aberrations in literature, such as t(14;16) [8] and 1q21 amplification [9,10], are mixed. More recently, patients with bi-allelic *TP53* inactivation, the amplification (>three copies) of *CKS1B*, and IgL-MYC

translocation, have been identified as high-risk for relapse in large scale cohorts [11,12]. However, these findings need further validation in independent studies.

Until 20 years ago, the treatment of MM was limited to steroids and alkylating agents. However, since the early 2000s, immunomodulatory agents and proteasome inhibitors have been introduced to treat MM. Other new agents against MM over the last decade, such as filanesib, dinaciclib, venetoclax, and daratumumab, have been developed and approved as well, or are still working their way through clinical trials [13–15]. As a result, the median survival has improved from 3 years to around 7 years [16–18]. Despite these improvements in the treatment of MM patients, roughly 20% of MM patients relapse and become drug resistant in 2 years after initial treatments [19,20]. Thus, while the therapeutic options available are increasing, it remains critical to identify high risk patients early and develop personalized treatment options for them to improve outcomes.

More than 20 different gene expression-based prognostic/response signatures have been reported for MM in last decade [21–42]. For example, Shaughnessy et al., applied a log-rank test coupled with stepwise multiple linear discriminant analysis to first derive a 70-gene signature (GEP70), and then selected a subset of these genes to predict prognosis of MM patients [38]. Later, it was found that five genes carried most of the discriminatory power of the 70-gene risk model [43]. Others, such as Kuiper et al., first employed univariate cox regression to select 1093 genes, then combined supervised principal component analysis with simulated annealing to reduce to a 92-gene signature (EMC-92) that was predictive of MM prognosis [31]. Only two genes overlap between GEP70 and EMC-92. This sparse overlap may be attributable to different aspects of myeloma biology. MM is a highly heterogeneous disease with diverse distinguished driver events, including simple mutations, genomic rearrangement, and other structure variants [4,14,44]. Multiple subgroups were proposed based on gene expression and/or genomic profiling [44,45]. Different MM subgroups might not share prognostic gene expression signatures. Recent efforts have been made to combine different gene expression signatures or combine gene expression signatures with clinical data to improve the prognostic stratification of MM patients using simulations [46–48].

In this study, we analyzed a large MM cohort (the Multiple Myeloma Research Consortium, MMRC study) [49], which contains both mRNA and copy number variation (CNV) profiles. We first conducted extensive QC to identify and correct sample-labeling errors (refer to the workflow in Figure 1). To distinguish potential associations and causal relationships among molecular features, we then constructed a multiple myeloma molecular causal network (M3CN) by integrating gene expression and CNV data using our previously described Bayesian network construction procedure [50]. Next, we collected prognostic and treatment response gene signatures from the literature, and then projected these signatures onto the M3CN to derive a robust prognostic MM subnetwork. Our network analysis of this subnetwork and of M3CN more generally, revealed several key regulators of the previously published prognostic and treatment response signatures. Even though these signatures had minimal overlap, most of them were significantly enriched in subnetworks that contained significant overlaps of gene signatures for prognosis and response. Pathway analysis showed that genes in the cell cycle and metabolic processes were enriched in these subnetworks. To assess the predictive power of the prognostic MM subnetwork we derived from M3CN, we stratified MM patients in an independent cohort, the MMRF-CoMMpass study, based on this subnetwork, and compared the accuracy of this stratification against the prognostic signatures from the literature. The M3CN-derived prognostic subnetwork was shown to achieve the best separation between different risk groups in terms of log-rank test *p*-values and hazard ratios. In summary, our results suggest that the network approach can reveal biological mechanisms unifying the diverse MM signatures.

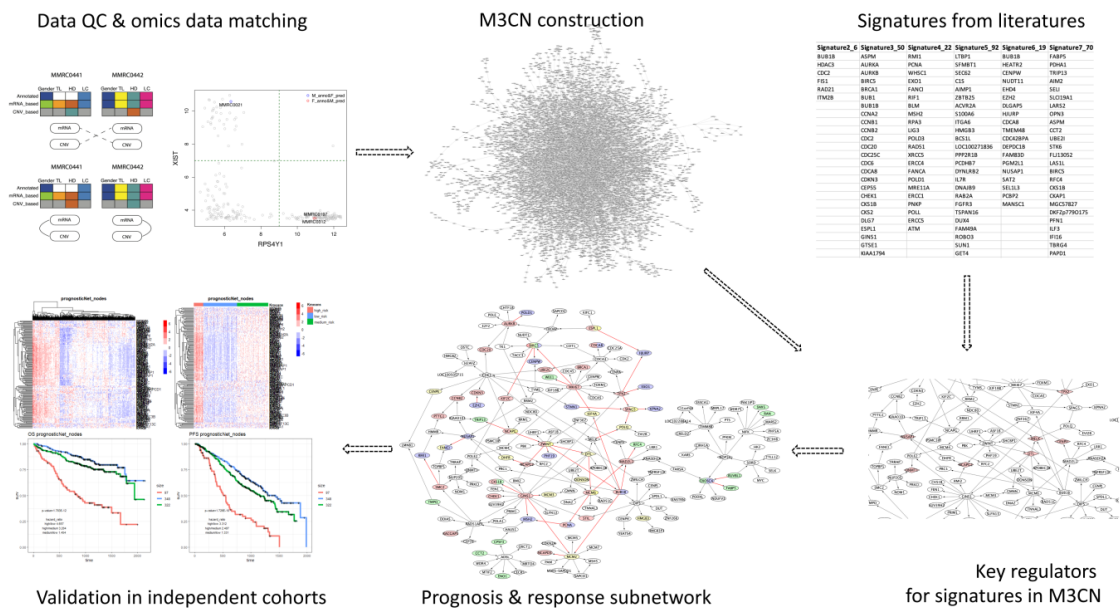


Figure 1. Workflow in this study. After extensive QC and omics data matching to correct sample-labeling errors, multiple myeloma molecular causal network (M3CN) was constructed, and multiple myeloma (MM) prognostic signature gene sets from the literature were projected onto it. Network analysis was subsequently performed to identify subnetwork associated with prognosis/response. The predictive power of subnetworks was assessed in independent cohorts.

2. Results

2.1. QC of the MMRC Dataset for Integrative Network Analysis

Gene expression profiles ($n = 304$, GSE26760), CNV profiles ($n = 254$, GSE26849), and associated clinical data for the MMRC study were downloaded from the GEO database [49]. Sample-labeling errors, including sample mislabeling, swapping, duplication, or contamination frequently occur in such multi-omics datasets [51,52]. Thus, it is critical to perform extensive QC to identify and correct such errors before integrating gene expression and CNV profiles for further analysis.

In MM, genomic alterations are common [49], and gene expression variations are strongly associated with such alterations [53]. In the MMRC dataset, the expression levels of 8182 genes were significantly associated with CNVs that contained the respective genes in cis form (cis-regulation), with a Benjamini–Hochberg multiple testing corrected p -value < 0.01 . Probabilistic multi-omics data matcher (proMODMatcher), a computational approach to identify and correct sample-labeling errors in multiple types of omics data [51,52], was applied to match mRNA and CNV profiles in the MMRC datasets. Among 246 pairs of gene expression and the CNV profiles of the common patient names, 10 profile pairs were not self-matched (i.e., the mRNA and CNV data annotated as having come from the same sample in these cases, were not correlated; see Supplementary Materials, Table S1a). Moreover, we detected six pairs of gene expression and CNV profiles that were cross-matched (i.e., the mRNA profile of one patient significantly correlated to the CNV profile of another patient; see Supplementary Materials, Table S1b). In total, 252 (246 self-matched and six cross-matched) pairs of gene expression and CNV profiles were used in the network reconstruction process.

To identify the source of these sample-labeling errors in the cross-matched mRNA–CNV profile pairs (whether the gene expression or CNV profiles were mis-labeled), the samples’ metadata (clinical annotations) and information inferred from mRNA and CNV profiles were compared. Based on the expression levels of the X-chromosome specific gene XIST and Y-chromosome specific gene RPS4Y1, MMRC samples were clustered into three groups to achieve a clear separation of samples into sex-specific groups (Male Sex, Female Sex, and No Call groups; left panel in Supplementary Materials,

Figure S1). There were inconsistencies between annotated sex/gender in the metadata and inferred sex based on the gene expression data for three samples: MMRC0021, MMRC0197, and MMRC0312 (Supplementary Materials, Table S1a). Sex was not inferred for samples in the third group containing profiles in which the levels of expression for both XIST and RPS4Y1 were too low due to X- or Y- chromosome loss in MM cells (left-bottom group in the left panel in Supplementary Figure S1). Immunoglobulin isotype can be inferred based on gene expression profiles. Comparison between the annotated and inferred isotypes revealed one inconsistent sample for light chain isotype (MMRC0207) and two for the heavy chain isotype (MMRC0039 and MMRC0312; Supplementary Materials, Table S1a). Analysis of hyperdiploidy estimated based on CNV profiles [54] and annotated hyperdiploidy yielded one inconsistency for sample MMRC0442 (Supplementary Materials, Table S1a). By comparing the clinically annotated and inferred metadata, the source of labeling errors was unambiguously identified for a few cross-matched profiles. For example, the mRNA and CNV profiles of MMRC0441 and MMRC0442 were cross-matched, and the labeling errors were likely due to a sample swap in CNV profiles based on the consistency with metadata. Similarly, mRNA and CNV profiles of MMRC0312 and MMRC0404 were cross-matched. The sex inconsistency between annotated and inferred sex of MMRC0312 indicated that a sample swap occurred in the mRNA expression profiles (Supplementary Materials, Figure S1).

After sample labeling error corrections, 304 more cis-associations (in total, 8486) were identified between gene expressions and the respective cis-CNVs at a multiple testing adjusted p -value < 0.01 . These results confirm that sample-labeling errors frequently occurred in large, complex datasets, and that proMODMatcher can efficiently identify and correct sample-labeling errors to improve power and accuracy in subsequent analyses.

2.2. Multiple Myeloma Molecular Causal Network (M3CN)

CNVs often occur in large blocks, with genes residing in these same blocks, likely sharing common CNV profiles across many samples. As a consequence of the shared common CNVs impacting the expression levels of genes in cis, the expression levels of these genes are likely to be correlated [55–57]. For example, the most significant cis-CNV regulated genes (of the 8394 genes identified at false discovery rate (FDR) value < 0.01) were enriched in chr1q (p -value = 2.9×10^{-54}), and these genes were co-expressed (Supplementary Materials, Figure S2). To distinguish gene co-expression due to biological regulations of core molecular and cellular functions from co-expression that is more an artifact of genomic co-localizations to common CNV regions, we included CNV data as nodes in the construction of the molecular causal MM networks, using our previously described Bayesian network reconstruction algorithm [58,59]. Serving as input into the reconstruction algorithm, were 7920 informative genes (mean expression levels > 4.8 and variance of expression levels > 0.4 ; see Supplementary Table S2), of which 3724 were cis-regulated by CNVs (Supplementary Materials, Table S3). In total 11,644 nodes (7920 nodes for gene expression and 3724 for CNVs) were included in the process of constructing a M3CN using RIMBANet [50] (detailed in Methods). The network reconstruction process searches for a structure G and associated parameters Θ that can best explain the given data $P(G, \Theta|D)$, which can be decomposed into a series of substructures (Methods). Given a potential regulation between nodes X and Y (i.e., X and Y are strongly associated), the joint probability $p(X, Y|D)$ can be represented as the structures $X \rightarrow Y$ $p(X, Y|D) = p(Y|X, D)p(X|D)$, $Y \rightarrow X$ $p(X, Y|D) = p(X|Y, D)p(Y|D)$, or the structure in which X and Y are both regulated by a third node Z (Figure 2A). Even though there is a directed edge between X and Y , the structures $X \rightarrow Y$ and $Y \rightarrow X$ are Markov equivalent (i.e., they have the same probability given the data D , so that they are statistically indistinguishable). However, when cis-CNV nodes are included, serving as a source of perturbation acting on X and Y , the structures $X \rightarrow Y$ $p(X, Y|D) = p(Y|X, CNV_y, D)p(X|D)$ and $Y \rightarrow X$ $p(X, Y|D) = p(X|Y, CNV_x, D)p(Y|D)$ (Figure 2B) are no longer equivalent, so that potential causal relationships between X and Y can be inferred unambiguously. In addition, when conditioning on a given CNV, gene expression correlations due to chromosome co-localization are able to be filtered out. For example, of the 140,283 pairs of

(X,Y) that were cis-regulated by CNVs and on the same chromosome associated at a multiple testing adjusted $p < 0.01$, after conditioning on CNV_x (or CNV_y), only 49% of the pairs (X,Y|CNV_x) were associated at the same multiple testing adjusted p -value < 0.01 , demonstrating the need to integrating CNV data in order to identify true biological regulation.

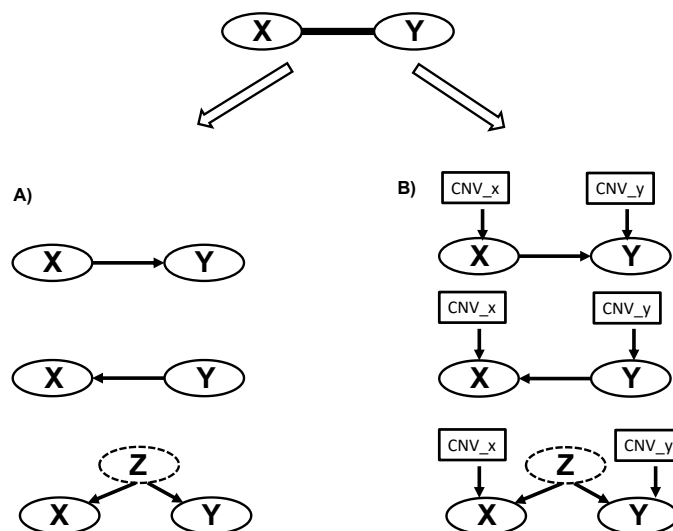


Figure 2. The causal structural inference schemes based on gene expression data, (A) without copy number variation (CNV) and (B) with CNV data integrated.

The resulting M3CN consisted of 9102 interactions between nodes representing variations in gene expression levels. To assess the quality of the constructed network, M3CN was compared with public network or pathway databases that are not MM specific. Nevertheless, a significant number of regulations in M3CN overlapped with interactions in pathway databases, supporting the biological relevance of the network. For example, over 30% of the regulations captured by the M3CN overlapped with regulatory relationships represented in the KEGG and Hallmark databases (Supplementary Materials, Table S4), indicating that M3CN was recapitulating known biological pathways. Causal links in M3CN were also compared with top correlated gene pairs derived from more basic co-expression analysis. Of the top 9102 gene pairs with the highest Pearson's correlation coefficients between their expression profiles, the correlations were more often being driven by genomic co-localization rather than coherent regulation of biological processes (Supplementary Materials, Figure S3).

The M3CN revealed many known features of MM. For example, the IgH gene [4] is one of the most frequently translocated genes in MM. The subnetwork associated with IgH is composed of mostly IgH family genes or genes located in the same region as IgH (14q32.33). However, this subnetwork is almost completely isolated from the rest of M3CN, with only a single node (the gene MIR8071-2) connecting this subnetwork to the rest of M3CN, (Supplementary Materials, Figure S4 This suggests that the pattern of gene expression variation driven by IgH and its corresponding subnetwork may have a limited impact on the rest of M3CN. Translocations between the immunoglobulin heavy chain locus and oncogene loci, including *CCND1*, *CCND3*, *MAF*, *FGFR3*, and *MMSET* (*WHSC1*) commonly occur in MM patients [60]. Further, MM patients can be divided into TC1-5 molecular subtypes based on translocations and *CCND1-3* expression levels [61]. Subtype-specific signatures were derived based on GSE13591 [62]. Putative key regulators were inferred for the TC subtype-specific signatures, including *CCND1* and *WHSC1* as key regulators for the TC1 and TC4 subtype-specific signatures, respectively. TC1-3 subtypes had one of the D-cyclin genes, *CCND1-3*, highly expressed, and the CD4/6-Rb pathway [63] activated. The subnetworks for the TC1-3 subtype-specific signatures all significantly overlapped with each other (e.g., the subnetwork for TC1 overlapped the TC2 and TC3 subnetworks, with TC1 2.2-fold enriched for TC2, p -values = 5.7×10^{-25} , and 6.2-fold enriched for TC3,

p -value = 1.6×10^{-39}), consistent with the observations that MM patients of TC1-3 subtypes shared similar survival patterns [64]. In contrast, *WHSC1* (*MMSET*) was upregulated in the TC4 subtype; *WHSC1* regulates the histone methylation of MM cells [65], which in turn regulates cell proliferation. The subnetwork for the TC4 specific signature was distinct from the subnetworks for the TC1-3 specific signatures (overlaps were not significant), consistent with the observations that MM patients of TC4 subtype had worse prognosis than the ones of TC1-3 subtypes [64].

At the global level, there were two highly connected genes, AGPS (Alkylglycerone Phosphate Synthase) and ATRX (Alpha Thalassemia/Mental Retardation Syndrome, X-Linked), regulated dozens of genes directly (41 and 32 respectively, Supplementary Figure S4). AGPS is a metabolic enzyme, a critical component in the synthesis of ether lipids, and is up-regulated across multiple types of aggressive human cancer cells and primary tumors [66]. Multiple studies show that lipid metabolism plays a critical role in MM tumorigenesis and progression [67]. Previous studies have also shown the potential of AGPS as a therapeutic target of cancer, and multiple AGPS inhibitors are in development [68]. ATRX is a chromatin remodeling protein whose main function is the deposition of the histone variant H3.3. A recent study showed that ATRX is a potential mutational driver in MM [69].

2.3. MM Prognostic Signature Genes in the M3CN

Eight large prognostic gene expression signatures were collected from the literature, with the number of genes across these eight signatures ranging from 15 to 92 (Table 1). The overlap of genes among these different signatures was limited (Supplementary Materials, Table S5). For example, only one gene, *BIRC5*, appeared in more than two signatures (Hose_50, Shaughnessy_70 and Kuiper_92 (EMC-92)). Some of the signatures were enriched in specific chromosomal locations (Supplementary Materials, Table S6). For instance, 14 genes from Shaughnessy_70 and 13 genes from Burlington_92 were located in chromosome 1 (Fisher's exact test $p = 5.54 \times 10^{-5}$ and 2.70×10^{-2} , respectively), consistent with the frequently described chromosome 1 aberration in MM [70].

Table 1. MM prognostic signature genes collected from literature.

Signature	PMID	Num. of sig. Genes	Num. of Genes in M3CN	Num. of Genes in Subnet	p -Values
Burlington_92	18676754	92	62	0	0.41
Decaux_15	18591550	15	6	1	0.13
Hose_50	20884712	50	30	28	2.2×10^{-16}
Kassambara_22	24809299	22	14	6	4.8×10^{-7}
Kuiper_92	22722715	92	55	16	2.1×10^{-13}
Reme_19	23493321	19	12	6	1.6×10^{-7}
Shaughnessy_70	17105813	70	33	14	1.5×10^{-14}
Zhan_52	17023574	52	31	0	1.0

These prognostic gene signatures were projected onto M3CN to infer the key regulators of each signature. It is of note that only a fraction of genes in each signature were included in the M3CN (Table 1), as only the most informative genes (e.g., those genes detectably expressed and that were observed to vary across samples) were included in the M3CN construction procedure (detailed in Methods). Our network analysis identified 15 potential key regulators for three of the eight prognostic signatures (Supplementary Materials, Table S7, detailed in Methods), including *NOPI6* and *CECR5* for Shaughnessy_70; *MELK*, *TPX2*, and *NCAPG2* for Kuiper_92; and *CDK1*, *DTL*, and eight other genes for the Hose_50 signature. *TPX2* is known to regulate *AURKA* and interacts with *RHAMM* [71], which is known to correlate with centrosome amplification and with poor prognosis in MM [72]; the *AURKA* inhibitor is a potential treatment for MM [73]. *CDK1* is one of the key regulators, and its inhibition has potential as an anti-cancer treatment [74]. Cyclin-dependent kinase inhibitors have been shown to induce cell cycle arrest and eventual apoptotic cell death in MM cells [75].

Using the 15 putative key regulators that we identified as regulating the prognostic gene signatures, we extracted a subnetwork consisting of 178 nodes and 218 interactions (Supplementary Materials, Tables S8 and S9) from M3CN (Figure 3, detailed in Methods) with the 15 key regulators as seeds, referred to as the prognostic subnetwork hereafter. Even though different prognostic signatures overlapped with each other sparsely at the individual gene level (Supplementary Materials, Table S5), five of the eight prognostic signatures we examined, were significantly enriched in this subnetwork (Table 1). For example, 28, 16, and 14 genes in Hose_50, Kuiper_92, and Shaughnessy_70, respectively, were in the prognostic subnetwork (corresponding to Fisher's exact test p -values: 2.20×10^{-16} , 2.10×10^{-13} , and 1.45×10^{-14} , respectively). Other smaller signatures, such as Kassambara_22 and Reme_19, were also enriched in this subnetwork ($p = 4.80 \times 10^{-7}$ and 1.55×10^{-7} , respectively).

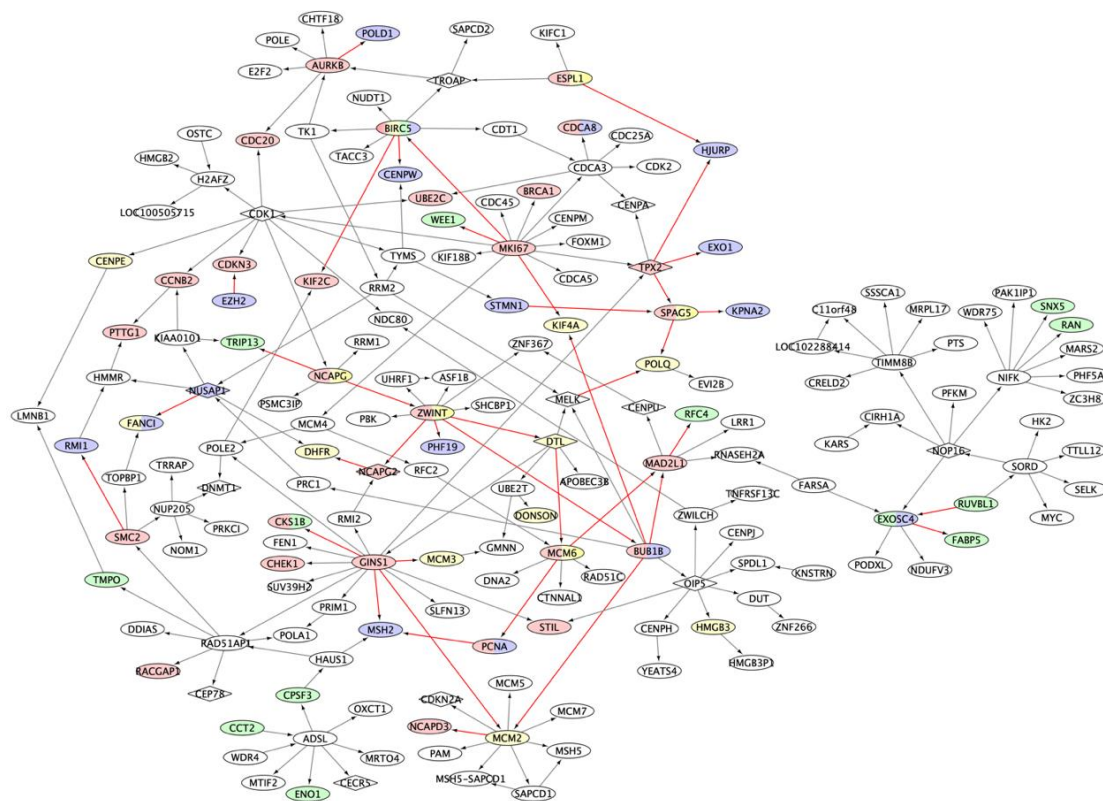


Figure 3. The prognostic subnetwork. The subnetwork was generated using key regulators for the Hose_40, Kuiper_92, and Shaughnessy_70 signatures as seeds. Nodes are color coded based on signatures: yellow nodes are genes from Kuiper_92, green nodes from Shaughnessy_70, and red nodes are Hose_40 genes; blue nodes are genes of other signatures. Edges in red indicate connections between genes of different signatures. Nodes of a diamond shape indicate key regulators for signatures.

The prognostic subnetwork was enriched for genes involved in the cell cycle and metabolic processes (Supplementary Materials, Table S10, detailed in Methods). More specifically, 42 out of 178 genes are involved in the cell cycle (fold enrichment = 8.4, FDR = 1.1×10^{-22}), and 25 in mitotic cell cycle process (FDR = 7.81×10^{-18}). Other top pathways include DNA replication (20 genes, FDR = 1.42×10^{-15}), cellular macromolecule biosynthetic process (20 genes, FDR: 8.8×10^{-5}), and cellular response to DNA damage stimulus (16 genes, FDR: 1.75×10^{-8}). The results are consistent with the essential roles cell cycle genes are known to play in MM progression, and the therapeutic implications [76].

To assess the prognostic values of the prognostic subnetwork or other prognostic signatures in the literature, we applied them to the Multiple Myeloma Research Foundation, MMRF's, CoMMpass cohort (Methods), which was not used in constructing M3CN or training other prognostic signatures.

In addition to the above prognostic signatures, we also included a four-gene signature [48] that was developed based on the MMRF-CoMMpass RNAseq dataset for comparison. Hierarchical clustering was applied to the CoMMpass RNAseq data based on genes in each signature, and the clustering result suggested that MM patients could be divided into three groups in general (Figure 4A and Supplementary Materials, Figure S5). Then, k-means clustering ($k = 3$) was applied to the data (Figure 4B and Supplementary Materials, Figure S5), and the three groups were compared in terms of overall survival (OS) (Figures 4C and 5) and progression free survival (PFS) (Figures 4D and 6). The patient groups, based on the prognostic subnetwork, had significantly different survival (log-rank test p -values = 1.8×10^{-12} and 1.7×10^{-10}), and hazard ratios (HRs) between high and low risk groups (4.7 and 3.3 for OS and PFS, respectively), were the best among all signatures tested (Table 2).

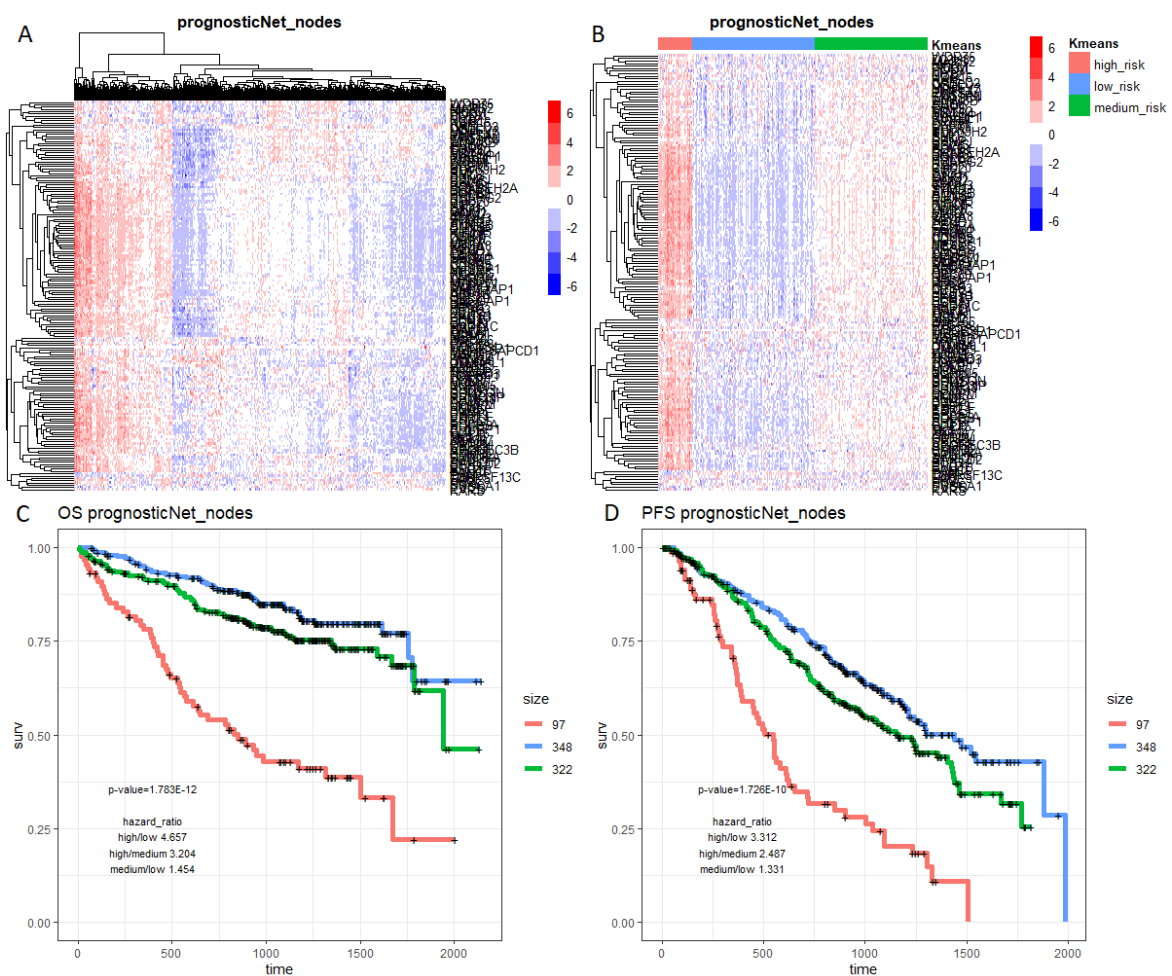


Figure 4. Heatmaps and K–M plots based on genes in the prognostic subnetwork and CoMMpass data. Rows in the heatmaps are genes and columns are samples. (A) The hierarchical clustering result and (B) the k-means ($k = 3$) clustering result were based on the z-scores of expression levels. K–M plots of three patient groups (based on k-means clusters) for (C) overall survival (OS) and (D) progress free survival (PFS).

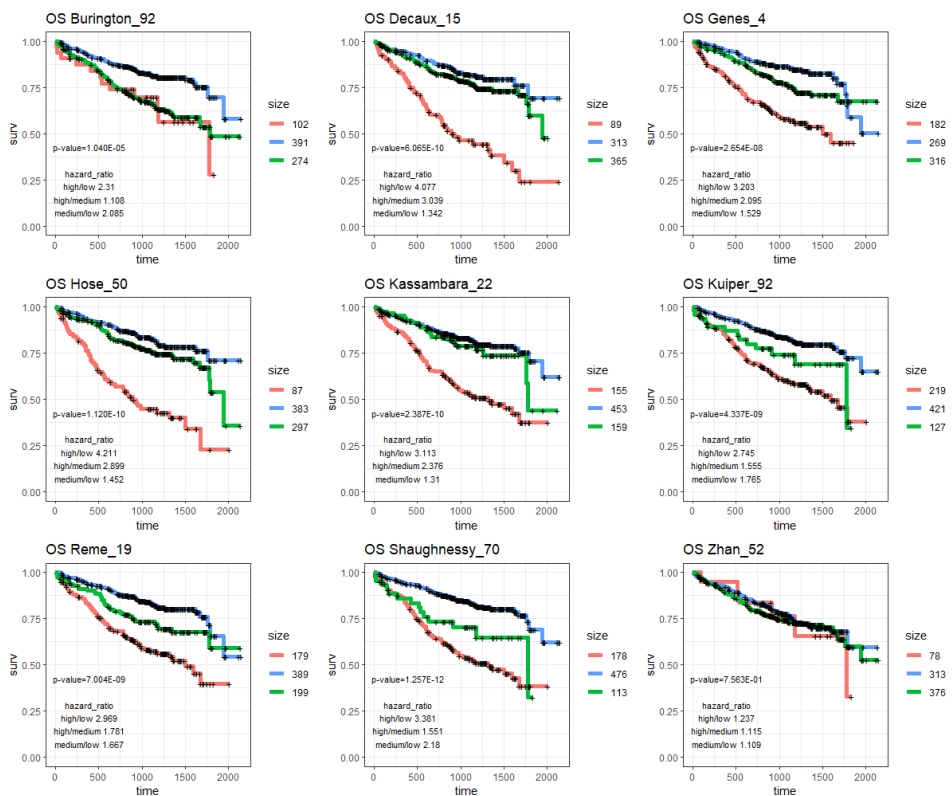


Figure 5. K-M plots of OS in patients in the CoMMpass study. Patients were partitioned into three groups based on genes in the nine different prognostic signatures in the literature (Supplementary Materials, Figure S5).

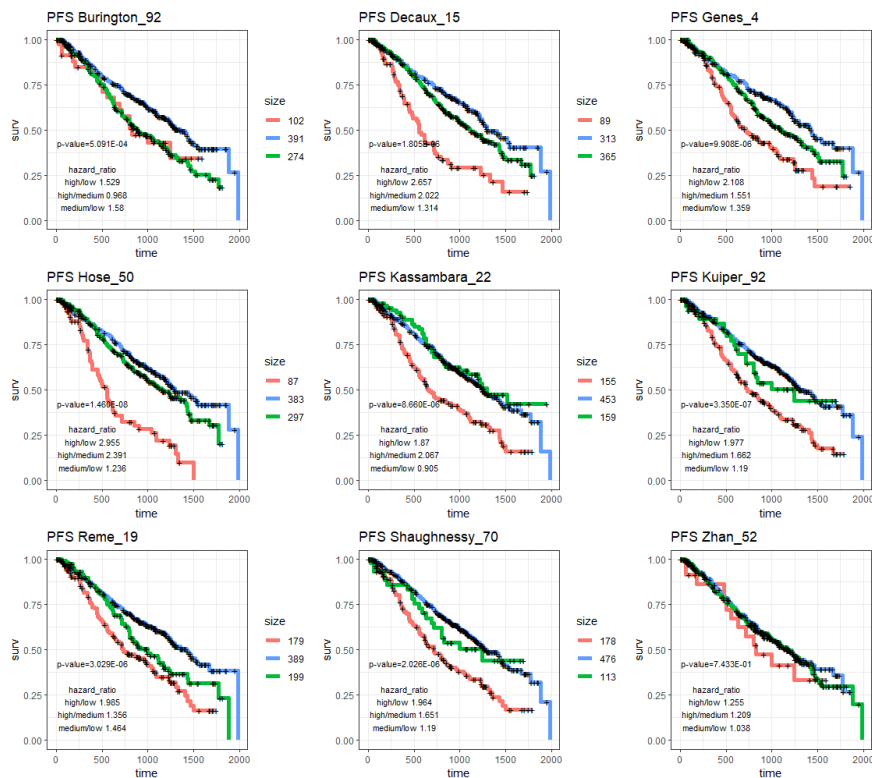


Figure 6. K-M plots of PFS in patients in the CoMMpass study. Patients were partitioned into three groups based on genes in the nine different prognostic signatures in the literature (Supplementary Figure S5).

Table 2. Summary of differences among MM patients in different risk groups, stratified based on genes in the prognostic subnetwork and the prognostic signatures in the literature. A total of 648 patients with survival information in the MMRF-CoMMpass study were included in the analysis. *p*-values were log-rank test *p*-values based on three risk groups. HR: hazard ratio.

Signature	OS				PFS			
	<i>p</i> -Value	HR High/Low	HR High/Med.	HR Med./Low	<i>p</i> -Value	HR High/Low	HR High/Med.	HR Med./Low
progNet	1.78×10^{-12}	4.657	3.204	1.454	1.73×10^{-10}	3.312	2.487	1.331
Burington_92	1.04×10^{-5}	2.31	1.108	2.085	5.09×10^{-4}	1.529	0.968	1.58
Decaux_15	6.07×10^{-10}	4.077	3.039	1.342	1.81×10^{-6}	2.657	2.022	1.314
Genes_4	2.65×10^{-8}	3.203	2.095	1.529	9.91×10^{-6}	2.108	1.551	1.359
Hose_50	1.12×10^{-10}	4.211	2.899	1.452	1.46×10^{-8}	2.955	2.391	1.236
Kassambara_22	2.39×10^{-10}	3.113	2.376	1.31	8.66×10^{-6}	1.87	2.067	0.905
Kuiper_92	4.34×10^{-9}	2.745	1.555	1.765	3.35×10^{-7}	1.977	1.662	1.19
Reme_19	7.00×10^{-9}	2.969	1.781	1.667	3.03×10^{-6}	1.985	1.356	1.464
Shaughnessy_70	1.26×10^{-12}	3.381	1.551	2.18	2.03×10^{-6}	1.964	1.651	1.19
Zhan_52	7.56×10^{-1}	1.237	1.115	1.109	7.43×10^{-1}	1.255	1.209	1.038

2.4. MM Treatment Response Signature Genes in the M3CN Network

To explore treatment response signatures in the context of the M3CN network, we identified five large signature gene sets (number of genes ≥ 15) related to various aspects of treatment in MM from the literature (Table 3). These signature genes encompass common types of treatments for MM. They also capture different aspects of treatments, including signature genes associated with patient survival after certain treatments (i.e., an after-treatment survival signature, such as shown in Mulligan_100), genes associated with drug resistance in MM cells (resistance signature, Mitra_42), genes whose expression changed after treatment (treatment effect signature, Bhutani_176), and genes identified through drug binding affinity screen (drug mechanism signature, Zhu_244). Similar to the prognostic gene signatures, the overlap among the drug-response signatures was limited: no genes were shared by three or more publications and only 13 genes were shared by two publications: *FLNA*, *SSX4*, *FAIM3*, *ITGB7*, *PPP1R16B*, *HLA-DOB*, *DEK*, *BLVRB*, *FKBP5*, *HCLS1*, *IDH1*, *DNAJA1*, and *XPO1*.

Table 3. MM drug-response signatures collected from the literature.

Signature	PMID	Num. of Sig. Genes	Num. Genes in M3CN	Treatment	Patients
Bhutani_176	28863804	176	132	IMiDs	Mixed
Zhu_244	24914135	244	143	IMiDs	Mixed
Mitra_42	28665416	42	29	PI	Mixed
Mulligan_100	17185464	100	30	PI	Mixed
Shaughnessy_80	21628408	80	40	TT2/TT3	NDMM

2.4.1. Immunomodulatory Drugs (IMiDs) Response Signatures

The mechanisms of IMiDs are complex, including apoptosis, immune responses, and anti-angiogenesis [77]. Two response signatures, which were derived from completely different approaches, were reported in the literature (Table 2) [21,42]. Zhu et al. reported 244 proteins (Zhu_244) that bind to E3 ligase protein cereblon (*CRBN*), which is the target of IMiDs [42], while Bhutani et al. report 176 genes (Bhutani_176) that were differently expressed when comparing bone marrow samples before and after IMiD treatment [21]. The two IMiDs' signatures were enriched for genes in different biological processes (Supplementary Materials, Table S11), indicating the two studies captured different aspects of IMiD treatment.

Zhu_244 was enriched for genes on Chromosome 4 (*p*-value = 0.007; Supplementary Materials, Table S12). Among the 244 genes in this signature, 143 were included in the M3CN, and 16 of them were in a subnetwork of 117 nodes (the largest connected subgraph and noted as subnetwork_Zhu_244, *p* = 2.3×10^{-10} ; Supplementary Materials, Figure S6). Subnetwork_Zhu_244 was enriched for the organic substance metabolic process (*p* = 4.75×10^{-5} ; Supplementary Materials, Table S13). One of the key

regulators of the subnetwork_Zhu_244 was *CCT2*, a chaperonin containing TCP1 subunit 2 (Supplementary Materials, Figure S6). *CCT2* is also in the prognostic signature Shaughnessy_70 [38]. In the M3CN, *CCT2* regulated *ADSL* (Supplementary Figure S6), with both of these genes also included in the prognostic subnetwork (Figure 3). The subnetwork_Zhu_244 and the prognostic subnetwork significantly overlap ($p = 2.9 \times 10^{-4}$), suggesting IMiDs bind *CRBN* and in turn suppressing cell proliferation [77].

On the other hand, 132 of 176 genes in *Bhutani_176* were included in the M3CN, and 19 were in a subnetwork comprised of 79 nodes (the largest connected subgraph and noted as subnetwork_Bhutani_176, $p = 2.82 \times 10^{-17}$; Supplementary Materials, Figure S7). *MX1*, an interferon-regulated resistance GTP-binding protein, was one of the key regulators of subnetwork_Bhutani_176 (Supplementary Materials, Figure S7). *MX1* is a key player of MM defined in proteomic studies [78]. Five genes (*MX1*, *STAT1*, *IFI27*, *IFI35*, and *OAS1*) in the *MX1*-regulated subnetwork overlapped with the 121 genes regulated by the DNA methylation inhibitor decitabine in MM cells; three of them (*IFI27*, *IFI35*, and *OAS1*) were also directly connected with *MX1* in subnetwork_Bhutani_176 [79]. *SOCS3*, suppressor of cytokine signaling 3, was also a key regulator of subnetwork_Bhutani_176 (Supplementary Materials, Figure S6). *SOCS3*, together with *STAT3* regulate the proliferation of MM cells [80,81]. The subnetwork_Bhutani_176 did not overlap with the prognostic subnetwork, and only shared one gene, *PRB1*, with subnetwork_Zhu_244 (Supplementary Materials, Figure S6), suggesting pleiotropic effects of IMiDs.

Furthermore, we generated a subnetwork by combining the *Zhu_244* and *Bhutani_176* signatures (Figure 7), which has 239 genes (Supplementary Materials, Table S14) and 244 edges (Supplementary Materials, Table S15), including 50 genes from *Zhu_244* and 12 genes from *Bhutani_176*. Two genes in the subnetwork, *XPO1* and *IDH1* were present in both signatures; seven edges connect two signatures. Pathway analysis showed that the subnetwork enriched genes as a defensive response to viruses ($p = 2.47 \times 10^{-9}$) and metabolic processes ($p = 2.76 \times 10^{-6}$) (Supplementary Materials, Table S16).

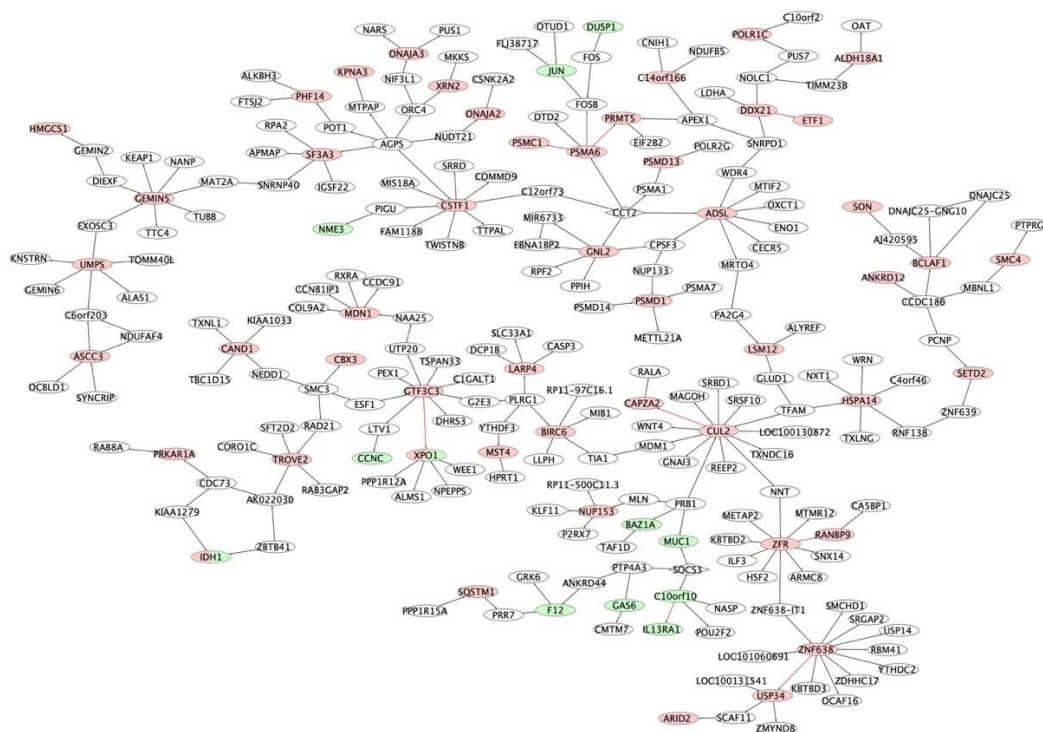


Figure 7. The IMiDs’ treatment response subnetwork. The subnetwork was generated using a combination of *Zhu_244* and *Bhutani_176* signatures. Nodes are color coded based on signatures: red nodes are genes from *Zhu_244*, and green nodes from *Bhutani_176* genes. Edges in red indicate connections between genes of signatures. Nodes which are diamond shaped indicate key regulators for signatures.

2.4.2. Proteasome Inhibitor (PI) Response Signatures

Mulligan et al. report a signature of 100 genes that associated with the survival of patients treated by proteasome inhibitor bortezomib [33]. Among them, 10 genes (including *GAGE4*, *MAGEA3*, *MAGEA6*, and *SSX2*) were on X chromosome (Supplementary Materials, Table S12; $p = 2.9 \times 10^{-5}$). Melanoma Antigen genes (MAGEs), more specifically *MAGEA3*, promote the survival of MM cells [82]. Cancer-testis antigen genes (*MAGEA3*, *MAGEA6*, and *GAGE4*) co-express with cell cycle genes (e.g., *CCNB2* and *MCM2*) in high proliferation MM patients [64], suggesting that PI drugs work better in highly proliferative MM cells [83]. Thirty-four genes in the signature were included in the M3CN, and the subnetworks of the signature (noted as subnetwork_Mulligan_100) consisted of 48 genes (Supplementary Materials, Figure S8), including nine genes in the signature ($p = 2.3 \times 10^{-13}$). *DEFA1*, defensin alpha 1, is significantly downregulated in MM [84], but upregulated in the Myeloid subgroup of MM patients [85]. MM patients with high myeloid signatures had much better survival [86]. Note that *MAGEA3*, *MAGEA6*, and *GAGE4* have positive weight in Mulligan_100, whereas *DEFA1* has a negative weight in the original publication [33].

Mitra_42 (Table 3) is derived by treating myeloma cell lines with multiple PIs [32]. This in vitro signature was not enriched in the M3CN and did not overlap with the Mulligan_100 signature, suggesting the in vitro signature does not reflect regulations in vivo.

2.4.3. Drug-Combination Response Signatures

The Shaughnessy_80 signature (Table 3) was derived in a two-step way: first, identifying differently expressed genes by comparing before and after thalidomide and bortezomib; then, differentially expressed genes (DEGs) are filtered based on association with survival [37]. Forty genes in the signatures were included in the M3CN. The subnetwork of Shaughnessy_80 consisted of 41 genes with eight genes in the signature ($p = 1.7 \times 10^{-11}$; Supplementary Materials, Figure S9). There were five genes in the overlap between subnetwork_Shaughnessy_80 and the IMiD response signature subnetwork, subnetwork_Zhu_244 ($p = 3.2 \times 10^{-4}$; Supplementary Materials Figure S6), but there was no overlap between subnetwork_Shaughnessy_80 and subnetwork_Bhutani_176, nor the former and subnetwork_Mulligan_100.

3. Discussion

Our network model integrated the multi-omics data available on MM patients. During data preprocessing for multi-omics data integration, our analyses suggested that samples of 12 patients in the MMRC dataset (5% of the total number of patients) were annotated incorrectly due to sample swaps and other unknown causes. Sample-labeling errors occur in clinics [87], clinical trials [88], and research databases [52]. Analyses based on error-containing datasets might decrease the power of the datasets, but may also lead to incorrect or contradictory scientific conclusions, or even harming patients [89,90]. Thus, it is critical to perform extensive QC to identify and correct potential sample-labeling errors in large-scale, multi-omics datasets prior to integrative data analysis, especially in the field of precision medicine.

Our network model aimed to identify molecular causal regulatory relationships between genes and among gene groups, with the ability to distinguish between these relationships and those co-expression patterns that may be driven by both biological regulation and genomic co-localization. A co-expression network of MM was constructed based on the MMRF-CoMMpass data by Lagana et al. [48]. The M9 module in Lagana et al.'s network was associated with prognosis ($p = 5 \times 10^{-3}$), and enriched for genes in the 13q deletion and 1q amplification blocks [48]. There were 972 unique genes in the M9 module, and 120 of them overlapped with our prognostic subnetwork (fold enrichment =14.1, $p = 7.9 \times 10^{-41}$). Even though they were significantly overlapping, genes in our prognostic subnetwork were not biased with respect to any chromosome location. The expressions of genes in our prognostic subnetwork were used to cluster patients in the MMRF-CoMMpass study into three risk groups

(Figure 4B); these groups were significantly different in terms of overall survival and progression free survival (log-rank test $p = 1.8 \times 10^{-12}$ and 1.7×10^{-10} , respectively), with hazard ratios (HRs) of 4.7 and 3.3, which were larger than the corresponding HRs based on the M9 module as a whole (HR = 1.75 for PFS), and the best four-gene combination used to predict prognosis (HR = 3 for PFS as reported in Lagana et al.). This is of particular note, given that the MMRF-CoMMpass study was not used in constructing the prognostic subnetwork, while Lagana et al.'s network was built on the MMRF-CoMMpass dataset. The results further suggest that our prognostic subnetwork is robust in identifying molecular mechanisms underlying high progression risk in MM patients.

With the increase of therapeutic options for MM patients, the development of an efficient prognostic and predictive signature becomes important, as personalized treatment plans can be tailored for patients with different risks, so that both under and over-treatment can be avoided. Molecular factors, such as genetic aberrations and gene expression profiles, have been investigated for their prognostic/predictive ability. Several genetic aberrations have been added to the Revised International Staging System (R-ISS), such as del(17p), t(4;14) and t(14;16), which are consistently associated with poor survival. On the other hand, studies of gene expression signatures have not led to consistent results, as shown by our analysis, where the number of genes in each signature varied significantly and the overlap among signatures was extremely limited. There are several potential reasons for those inconsistencies. First, the heterogeneity of MM may contribute to the inconsistencies. Though MM is still considered a single disease entity, its clinical presentation, response to treatment, and survival outcomes are highly heterogeneous. Based on cytogenetic abnormalities, MM can originate from two pathways: the hyperdiploidy of chromosomes 3, 5, 7, 9, 11, 15, 19 and 21, which is observed in 55% of patients; the translocation of *IGH* or *MYC*, observed in 40% to 50% of patients [4]. In most cases, hyperdiploidy and translocation are mutually exclusive in MM patients. The clustering analysis of genomic landscape and/or gene expression profiles reveal four to 10 subgroups using a large number of MM patients [44,69], indicating the underlying biology of these two types of MM is different. It is very likely that signatures will be different for those subgroups, but so far there are no signatures developed for subgroups separately. Second, variability in the methods applied to derive the signatures may result in differences among the signatures. Although almost all signatures were derived from microarray-based gene expression profiling, the methods are very different. For example, Kuiper_92 was developed using univariate Cox regression analysis, followed by a supervised principal component analysis in combination with simulated annealing [31], whereas Shaughnessy_70 was derived by a combination of log-rank tests with a stepwise multiple linear discriminant analysis [38], and then the method of Shaughnessy_80 used the expression profile after 48-h bortezomib test-dosing [37]. More recently, a simulation-based sophisticated machine learning approach was applied to derive a signature to predict the treatment benefits for MM patients [91]. Given the big differences among the applied methods, it is not surprise that the overlap among signatures was very limited.

To develop personalized treatments, it is critical to separate prognosis without a specific treatment and with the benefit due to the specific treatment. Even though the three prognostic signatures (Hose_50, Kuiper_92, and Shaughnessy_70) overlapped with each other sparsely, they were all enriched in the prognostic subnetwork (Figure 2), indicating that they were closely connected and regulated, and potentially in the same or related biological pathways; our pathway analysis confirmed that the prognostic subnetwork enriched genes of the cell cycle process (Supplementary Table S6). The prognostic subnetwork developed in this study reflects most prognostic signatures and has the potential to enable more robust estimation of baseline prognosis (Figures 4–6 and Table 2) when developing new treatments for MM.

Our network analysis also showed that MM drugs had pleiotropic effects with signature genes enriched in different subnetworks. However, some MM signatures were not enriched in any MM subnetwork, such as Mitra_42. The potential reasons for that disconnection may be due to (1) MM heterogeneity as discussed above, and (2) MM cells' interactions with the tumor microenvironment. Recent studies showed that bone marrow niche is critical for MM cell survival and proliferation [92].

Bone marrow adipocytes provide metabolites and metabolic signals for MM cell proliferation [67]. Thus, in vitro signatures may not faithfully reflect changes in vivo. As new high-throughput technology, such as single cell RNA sequencing [93], rapidly develops and is applied in studying diseases, including MM [94,95], more understanding of the MM microenvironment and treatment response is expected in the near future.

4. Materials and Methods

4.1. The Preprocess of Gene Expression, and CNV Data and Omics Data Matching

We downloaded the MMRC reference dataset consisting of 304 mRNA expression profiles (GSE26760), 254 CNV profiles (GSE26849), and associated metadata [49]. For the CNV profiles, the downloaded Circular binary segmentation (CBS) of the logR ratio values was mapped to gene levels' CNVs based on the coordinate information from Refgene annotations against hg19. The coordinate was downloaded from UCSC genome browser (<https://genome.ucsc.edu>). The sex prediction of MM patients was based on the expression levels of X chromosome specific gene *XIST* and Y chromosome specific gene *RPS4Y1* [96]. Immunoglobulin light chain and heavy chain isotypes of each MM sample were predicted by the corresponding Immunoglobulin genes. CNV data was used to estimate hyperdiploidy using a previous published method [97].

proMODMatcher was applied to detect sample labelling errors using mRNA and CNV profiles; the detail of the method can be found in [51]. Briefly, starting with 246 mRNA and CNV pairs that were matched based on their labels, we first identified a set of genes whose mRNA expression was highly correlated with their CNV profiles across all the 246 sample pairs. Out of 20,395 genes, 8404 genes were significantly correlated between mRNA and CNV profiles (FDR < 0.01). Then the profiles of selected genes were rank-transformed for non-parametric comparison among samples. Finally, the sample similarity score was calculated as the correlation coefficient between RNA expression of cis genes in sample i and the CNV values of cis genes in samples j , $S_{g_{cnv}}(i, j)$ for all $i, j = 1 \dots n$. We expected that (1) $S_{g_{cnv}}(i, i)$ was greater than $S_{g_{cnv}}(i, j)$ for $j \neq i$, and (2) $S_{g_{cnv}}(i, i)$ was greater than $S_{g_{cnv}}(k, i)$ for $k \neq i$, as one sample mRNA profile of cis genes should be the most similar to its own CNV profile and vice versa. If neither (1) nor (2) was satisfied for one sample, we tried to identify potential cross matches, such that $S_{g_{cnv}}(i, j)$ was the most significant among all possible permutations. With updated gene expression and CNV profiles pairs, cis gene expression-CNV pairs were refined and the procedure was iterated until there was no update between iterations.

4.2. Gene Signatures for MM

Twenty-two gene signatures for MM were collected from publications up to April 2019, 13 of them with numbers of genes > 15 were used for network analysis and were listed in Tables 1 and 2, along with the PMID of the publication. Probe IDs of microarrays were converted to gene symbols based on annotation of the relevant microarray platform. Duplicated gene symbols were removed, as were probes that failed to map to any annotated genes. For signatures that are not lists of genes, for example, the ratio of two genes, the resulting signature contained both genes.

4.3. M3CN Construction and Network Analysis

The M3CN was constructed by integrating gene expression and CNV profiles. GSE26863 (GSE26760 and GSE26849) for MM patients was downloaded from GEO, processed and normalized using R/Bioconductor. In total, 7920 genes with detectable expression levels and large variances across samples were selected to be included in the network reconstruction process. Among them, the expression of 3724 genes was cis-regulated by CNVs (FDR < 0.01) so that 3724 nodes for cis-CNVs were also included in the network reconstruction. The gene expression or CNV profile was discretized into three states: low, normal, and high level, guided by k-means clustering ($k = 3$) and biological meaningful cutoff values [98]. The gene expression and CNV nodes were then input into the software

suite, Reconstructing Integrative Molecular Bayesian Network (RIMBANet) [50,99], to construct a biological causal network given the data and priors, as previously described. Briefly, the network reconstruction process searches for a directed acyclic graph (DAG) structure G and associated parameters Θ that can best explain the given data D , $P(G, \Theta|D)$. If the structure G is a DAG, then $P(G, \Theta|D)$ can be decomposed into a series of sub-structures $P(G, \Theta|D) = \prod_i P(G^i, \Theta^i|D)$. With cis-CNV nodes included (Figure 2B), the structures $X \rightarrow Y$, given by $p(X \rightarrow Y|D) = p(Y|X, CNV_y, D)p(X|D)$, and $Y \rightarrow X$, given by $p(Y \rightarrow X|D) = p(X|Y, CNV_x, D)p(Y|D)$, are no longer equivalent, so that potential causal relationships between X and Y can be inferred unambiguously. To speed up the searching process, for each gene, the bottom 20% genes based on their mutual information were excluded as potential candidate regulators (sparse candidate search [100]). The network reconstruction process is a Monte Carlo Markov chain (MCMC). Given different random seeds, we might end up with different structures. Thus, we ran 1000 independent MCMC processes based on 1000 random seed numbers that resulted in 1000 candidate structures. Then, we selected consensus structure features with posterior probabilities >0.3 among candidate structures [98]. Finally, loops in the consensus network were removed by deleting the weakest link in the loops. The resulting network was visualized using Cytoscape 3.7 [101]. Given a set of seed nodes N_s , $N_s = \bigcup_i d(\text{node}, N_s^i) \leq l$ is the union of nodes that are within l steps from the seed node N_s^i , and the subnetwork for the seed nodes N_s is the set of connections among N_s .

4.4. The Identification of Key Regulators for Signature Genes

For the three signature genes, *Hose_50*, *Huiper_92*, and *Shaugnessy_70*, we further identified the putative key regulators [50,102,103] in M3CN. Briefly, for each gene i in the M3CN, a two-step subnetwork sub_i ($l = 2$) was extracted as described above. If sub_i was enriched for the input signature genes, then the gene i was a candidate regulator. After all candidate regulators were identified, they were sorted by corresponding enrichment p -values. The gene with the most significant enrichment p -value was a key regulator. Then, any candidate regulators in its two-step subnetwork were excluded as candidate key regulators. Moreover, the next candidate regulator in the sorted list was selected as a key regulator. The process was repeated throughout the sorted candidate list.

4.5. Pathway Analysis

PANTHER was used for pathway analysis based on overrepresentation test [104]. The gene list was submitted to PANTHER web server (version 14.1); Fisher's exact test was used for statistical hypothesis testing. FDR was used for the correction of the multiple-hypothesis testing problem, with $\text{FDR} < 0.05$ as the cutoff.

4.6. The MMRF-CoMMpass Cohort

The IA14 Release of CoMMpass data was downloaded from the MMRF researcher gateway portal (<https://research.themmr.org>), which consists of 767 RNA sequencing data of baseline CD138+ plasma cell bone marrow samples from newly diagnosed MM patients. Among them, overall survival or progression-free survival information was available for 648 patients which were included in our analyses.

5. Conclusions

We constructed a molecular causal network describing potential regulatory relationships in MM and analyzed different MM gene signatures using the network. We identified a prognostic subnetwork that shared among different MM prognostic signatures and showed that the prognostic subnetwork was more robust than individual signatures in terms predicting prognosis in an independent cohort. Our study demonstrated that MM was complex so that an integrative network is needed to reveal molecular mechanisms underlying different MM signatures.

Supplementary Materials: The following are available online at <http://www.mdpi.com/2072-6694/11/10/1452/s1>: Supplementary Table S1a. Sample error detected by analysis of sex, immunoglobulin, and hyperdiploidy. Supplementary Table S1b. Sample matches predicted by proMODMatcher between mRNA expression and CNV. Supplementary Table S2. Informative genes included in the construction of M3CN. Supplementary Table S3. Cis-CNV regulated genes among those that were informative, included in the construction of M3CN. Supplementary Table S4. Comparison of M3CN with public network/pathway databases. Supplementary Table S5. Comparison of prognostic signatures in the literature. Supplementary Table S6. Enrichment of prognostic signature genes in chromosomes. Supplementary Table S7. Key regulators identified in M3CN for the three large prognostic signatures. Supplementary Table S8. Nodes in the prognostic subnetwork. Supplementary Table S9. Regulations among nodes in the prognostic subnetwork. Supplementary Table S10. Pathway analysis of the prognostic subnetwork. Supplementary Table S11. Top biological processes enriched in the treatment response signatures. Supplementary Table S12. Enrichment of response signature genes in chromosomes. Supplementary Table S13. Top biological processes enriched in the MM drug-response subnetworks. Supplementary Table S14. Nodes in the IMiD response subnetwork. Supplementary Table S15. Regulations among nodes in the IMiD response subnetwork. Supplementary Table S16. Pathway analysis of the IMiD treatment response subnetwork. Supplementary Figure S1. Sample-labeling errors identified in the MMRC dataset. Supplementary Figure S2. Associations between CNV and mRNA. Supplementary Figure S3. Distributions of correlations between gene expressions, and correlations between CNVs. Supplementary Figure S4. M3CN, hub genes, and IGH subnet. Supplementary Figure S5. MM patients in the CoMMpass study were clustered based on genes in the different prognostic signatures. Supplementary Figure S6. subnetwork generated based on Zhu_244 (subnetwork_Zhu_244). Supplementary Figure S7 subnetwork generated based on Bhutani_176 (subnetwork_Bhutani_176). Supplementary Figure S8 subnetwork generated based on Mulligan_100 (subnetwork_Mulligan_100). Supplementary Figure S9 subnetwork generated based on Shaughnessy_80 (subnetwork_Shau-ghnessy_80).

Author Contributions: Conceptualization, Y.L., L.W., and J.Z.; methodology, Y.L., L.W., and J.Z.; formal analysis, Y.L., S.Y., E.L., L.W., and H.Y.; resources, Y.L. and H.Y.; data curation, Y.L., H.Y., E.L. and S.Y.; writing—original draft preparation, Y.L. and J.Z.; writing—review and editing, L.W., A.L., S.P., E.E.S., and J.Z.; visualization, Y.L. and H.Y.; supervision, L.W. and J.Z.; project administration, L.W.; funding acquisition, E.E.S. and J.Z.

Funding: This research was partially funded by the National Institute of Health, grant number U01 HG008451.

Acknowledgments: We would like to acknowledge the contributions of all our academic and technology partners who helped with the project, with special thanks to our colleagues in IT department, Kelly Farrar, Anatol Blass, and Shay Hassidim. The MMRF-CoMMpass data used in the study were generated as part of the Multiple Myeloma Research Foundation Personalized Medicine Initiatives (<https://research.themmr.org> and www.themmr.org).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Siegel, R.L.; Miller, K.D.; Jemal, A. Cancer statistics, 2018. *CA Cancer J. Clin.* **2018**, *68*, 7–30. [[CrossRef](#)] [[PubMed](#)]
2. Drach, J.; Schuster, J.; Nowotny, H.; Angerler, J.; Rosenthal, F.; Fiegl, M.; Rothermundt, C.; Gsur, A.; Jäger, U.; Heinz, R. Multiple myeloma: High incidence of chromosomal aneuploidy as detected by interphase fluorescence in situ hybridization. *Cancer Res.* **1995**, *55*, 3854–3859. [[PubMed](#)]
3. Kuehl, W.M.; Bergsagel, P.L. Early genetic events provide the basis for a clinical classification of multiple myeloma. *Hematol. Am. Soc. Hematol. Educ. Program* **2005**, *2005*, 346–352. [[CrossRef](#)] [[PubMed](#)]
4. Manier, S.; Salem, K.Z.; Park, J.; Landau, D.A.; Getz, G.; Ghobrial, I.M. Genomic complexity of multiple myeloma and its clinical implications. *Nat. Rev. Clin. Oncol.* **2017**, *14*, 100–113. [[CrossRef](#)] [[PubMed](#)]
5. Keats, J.J.; Reiman, T.; Maxwell, C.A.; Taylor, B.J.; Larratt, L.M.; Mant, M.J.; Belch, A.R.; Pilarski, L.M. In multiple myeloma, t(4;14)(p16;q32) is an adverse prognostic factor irrespective of FGFR3 expression. *Blood* **2003**, *101*, 1520–1529. [[CrossRef](#)] [[PubMed](#)]
6. Gaballa, S.; Saliba, R.M.; Srour, S.; Lu, G.; Brammer, J.E.; Shah, N.; Bashir, Q.; Patel, K.; Bock, F.; Parmar, S.; et al. Outcomes in patients with multiple myeloma with TP53 deletion after autologous hematopoietic stem cell transplant. *Am. J. Hematol.* **2016**, *91*, E442–E447. [[CrossRef](#)] [[PubMed](#)]
7. Thakurta, A.; Ortiz, M.; Blecua, P.; Towfic, F.; Corre, J.; Serbina, N.V.; Flynt, E.; Yu, Z.; Yang, Z.; Palumbo, A.; et al. High subclonal fraction of 17p deletion is associated with poor prognosis in multiple myeloma. *Blood* **2019**, *133*, 1217–1221. [[CrossRef](#)]
8. Avet-Loiseau, H.; Malard, F.; Campion, L.; Magrangeas, F.; Sebban, C.; Lioure, B.; Decaux, O.; Lamy, T.; Legros, L.; Fuzibet, J.G.; et al. Translocation t(14;16) and multiple myeloma: Is it really an independent prognostic factor? *Blood* **2011**, *117*, 2009–2011. [[CrossRef](#)]

9. Hanamura, I.; Stewart, J.P.; Huang, Y.; Zhan, F.; Santra, M.; Sawyer, J.R.; Hollmig, K.; Zangarri, M.; Pineda-Roman, M.; Van Rhee, F.; et al. Frequent gain of chromosome band 1q21 in plasma-cell dyscrasias detected by fluorescence in situ hybridization: Incidence increases from MGUS to relapsed myeloma and is related to prognosis and disease progression following tandem stem-cell transplantation. *Blood* **2006**, *108*, 1724–1732. [[CrossRef](#)]
10. Avet-Loiseau, H.; Attal, M.; Moreau, P.; Charbonnel, C.; Garban, F.; Hulin, C.; Leyvraz, S.; Michallet, M.; Yakoub-Agha, I.; Garderet, L.; et al. Genetic abnormalities and survival in multiple myeloma: The experience of the Intergroupe Francophone du Myélome. *Blood* **2007**, *109*, 3489–3495. [[CrossRef](#)]
11. Walker, B.A.; Mavrommatis, K.; Wardell, C.P.; Ashby, T.C.; Bauer, M.; Davies, F.; Rosenthal, A.; Wang, H.; Qu, P.; Hoering, A.; et al. A high-risk, Double-Hit, group of newly diagnosed myeloma identified by genomic analysis. *Leukemia* **2019**, *33*, 159–170. [[CrossRef](#)] [[PubMed](#)]
12. Barwick, B.G.; Neri, P.; Bahlis, N.J.; Nooka, A.K.; Dhodapkar, M.V.; Jaye, D.L.; Hofmeister, C.C.; Kaufman, J.L.; Gupta, V.A.; Auclair, D.; et al. Multiple myeloma immunoglobulin lambda translocations portend poor prognosis. *Nat. Commun.* **2019**, *10*, 1911. [[CrossRef](#)] [[PubMed](#)]
13. Kumar, S. Emerging options in multiple myeloma: Targeted, immune, and epigenetic therapies. *ASH Educ. Progr. Book* **2017**, *2017*, 518–524. [[CrossRef](#)] [[PubMed](#)]
14. Kumar, S.K.; Rajkumar, S.V. The multiple myelomas—Current concepts in cytogenetic classification and therapy. *Nat. Rev. Clin. Oncol.* **2018**, *15*, 409–421. [[CrossRef](#)] [[PubMed](#)]
15. Plesner, T.; Krejčík, J. Daratumumab for the Treatment of Multiple Myeloma. *Front. Immunol.* **2018**, *9*, 1228. [[CrossRef](#)] [[PubMed](#)]
16. Durie, B.G.; Hoering, A.; Abidi, M.H.; Rajkumar, S.V.; Epstein, J.; Kahanic, S.P.; Thakuri, M.; Reu, F.; Reynolds, C.M.; Sexton, R.; et al. Bortezomib with lenalidomide and dexamethasone versus lenalidomide and dexamethasone alone in patients with newly diagnosed myeloma without intent for immediate autologous stem-cell transplant (SWOG S0777): A randomised, open-label, phase 3 trial. *Lancet* **2017**, *389*, 519–527. [[CrossRef](#)]
17. Attal, M.; Lauwers-Cances, V.; Hulin, C.; Leleu, X.; Caillot, D.; Escoffre, M.; Arnulf, B.; Macro, M.; Belhadj, K.; Garderet, L.; et al. Lenalidomide, Bortezomib, and Dexamethasone with Transplantation for Myeloma. *N. Engl. J. Med.* **2017**, *376*, 1311–1320. [[CrossRef](#)] [[PubMed](#)]
18. Goldschmidt, H.; Lokhorst, H.M.; Mai, E.K.; van der Holt, B.; Blau, I.W.; Zweegman, S.; Weisel, K.C.; Vellenga, E.; Pfreundschuh, M.; Kersten, M.J.; et al. Bortezomib before and after high-dose therapy in myeloma: Long-term results from the phase III HOVON-65/GMMG-HD4 trial. *Leukemia* **2018**, *32*, 383–390. [[CrossRef](#)] [[PubMed](#)]
19. Kumar, S.K.; Therneau, T.M.; Gertz, M.A.; Lacy, M.Q.; Dispenzieri, A.; Rajkumar, S.V.; Fonseca, R.; Witzig, T.E.; Lust, J.A.; Larson, D.R.; et al. Clinical Course of Patients with Relapsed Multiple Myeloma. *Mayo Clin. Proc.* **2004**, *79*, 867–874. [[CrossRef](#)] [[PubMed](#)]
20. Sonneveld, P.; Broijl, A. Treatment of relapsed and refractory multiple myeloma. *Haematologica* **2016**, *101*, 396–406. [[CrossRef](#)]
21. Bhutani, M.; Zhang, Q.; Friend, R.; Voorhees, P.M.; Druhan, L.J.; Barlogie, B.; Sonneveld, P.; Morgan, G.J.; Symanowski, J.T.; Avalos, B.R.; et al. Investigation of a gene signature to predict response to immunomodulatory derivatives for patients with multiple myeloma: An exploratory, retrospective study using microarray datasets from prospective clinical trials. *Lancet Haematol.* **2017**, *4*, e443–e451. [[CrossRef](#)]
22. Burington, B.; Barlogie, B.; Zhan, F.; Crowley, J.; Shaughnessy, J.D. Tumor cell gene expression changes following short-term in vivo exposure to single agent chemotherapeutics are related to survival in multiple myeloma. *Clin. Cancer Res.* **2008**, *14*, 4821–4829. [[CrossRef](#)] [[PubMed](#)]
23. Chapman, M.A.; Sive, J.; Ambrose, J.; Roddie, C.; Counsell, N.; Lach, A.; Abbasian, M.; Popat, R.; Cavenagh, J.D.; Oakervee, H.; et al. RNA-seq of newly diagnosed patients in the PADIMAC study leads to a bortezomib/lenalidomide decision signature. *Blood* **2018**, *132*, 2154–2165. [[CrossRef](#)] [[PubMed](#)]
24. Chng, W.J.; Dispenzieri, A.; Chim, C.S.; Fonseca, R.; Goldschmidt, H.; Lentzsch, S.; Munshi, N.; Palumbo, A.; Miguel, J.S.; Sonneveld, P.; et al. IMWG consensus on risk stratification in multiple myeloma. *Leukemia* **2014**, *28*, 269–277. [[CrossRef](#)] [[PubMed](#)]

25. Decaux, O.; Lode, L.; Magrangeas, F.; Charbonnel, C.; Gouraud, W.; Jézéquel, P.; Attal, M.; Harousseau, J.L.; Moreau, P.; Bataille, R.; et al. Prediction of Survival in Multiple Myeloma Based on Gene Expression Profiles Reveals Cell Cycle and Chromosomal Instability Signatures in High-Risk Patients and Hyperdiploid Signatures in Low-Risk Patients: A Study of the Intergroupe Francophone du Myelome. *J. Clin. Oncol.* **2008**, *26*, 4798–4805. [[CrossRef](#)] [[PubMed](#)]
26. Dickens, N.J.; Walker, B.A.; Leone, P.E.; Johnson, D.C.; Brito, J.L.; Zeisig, A.; Jenner, M.W.; Boyd, K.D.; Gonzalez, D.; Gregory, W.M.; et al. Homozygous deletion mapping in myeloma samples identifies genes and an expression signature relevant to pathogenesis and outcome. *Clin. Cancer Res.* **2010**, *16*, 1856–1864. [[CrossRef](#)]
27. Hofman, I.J.F.; van Duin, M.; De Bruyne, E.; Fancello, L.; Mulligan, G.; Geerdens, E.; Garelli, E.; Mancini, C.; Lemmens, H.; Delforge, M.; et al. RPL5 on 1p22.1 is recurrently deleted in multiple myeloma and its expression is linked to bortezomib response. *Leukemia* **2017**, *31*, 1706–1714. [[CrossRef](#)] [[PubMed](#)]
28. Hose, D.; Reme, T.; Hielscher, T.; Moreaux, J.; Messner, T.; Seckinger, A.; Benner, A.; Shaughnessy, J.D., Jr.; Barlogie, B.; Zhou, Y.; et al. Proliferation is a central independent prognostic factor and target for personalized and risk-adapted treatment in multiple myeloma. *Haematologica* **2011**, *96*, 87–95. [[CrossRef](#)] [[PubMed](#)]
29. Kassambara, A.; Gourzones-Dmitriev, C.; Sahota, S.; Rème, T.; Moreaux, J.; Goldschmidt, H.; Constantinou, A.; Pasero, P.; Hose, D.; Klein, B. A DNA repair pathway score predicts survival in human multiple myeloma: The potential for therapeutic strategy. *Oncotarget* **2014**, *5*, 2487–2498. [[CrossRef](#)]
30. Kronke, J.; Kuchenbauer, F.; Kull, M.; Teleanu, V.; Bullinger, L.; Bunjes, D.; Greiner, A.; Kolmus, S.; Kopff, S.; Schreder, M.; et al. IKZF1 expression is a prognostic marker in newly diagnosed standard-risk multiple myeloma treated with lenalidomide and intensive chemotherapy: A study of the German Myeloma Study Group (DSMM). *Leukemia* **2017**, *31*, 1363–1367. [[CrossRef](#)]
31. Kuiper, R.; Broyl, A.; De Knegt, Y.; Van Vliet, M.H.; Van Beers, E.H.; Van Der Holt, B.; El Jarari, L.; Mulligan, G.; Gregory, W.; Morgan, G.; et al. A gene expression signature for high-risk multiple myeloma. *Leukemia* **2012**, *26*, 2406–2413. [[CrossRef](#)]
32. Mitra, A.K.; Harding, T.; Mukherjee, U.K.; Jang, J.S.; Li, Y.; HongZheng, R.; Jen, J.; Sonneveld, P.; Kumar, S.; Kuehl, W.M.; et al. A gene expression signature distinguishes innate response and resistance to proteasome inhibitors in multiple myeloma. *Blood Cancer J.* **2017**, *7*, e581. [[CrossRef](#)]
33. Mulligan, G.; Mitsiades, C.; Bryant, B.; Zhan, F.; Chng, W.J.; Roels, S.; Koenig, E.; Fergus, A.; Huang, Y.; Richardson, P.; et al. Gene expression profiling and correlation with outcome in clinical trials of the proteasome inhibitor bortezomib. *Blood* **2007**, *109*, 3177–3188. [[CrossRef](#)]
34. Narita, T.; Ri, M.; Masaki, A.; Mori, F.; Ito, A.; Kusumoto, S.; Ishida, T.; Komatsu, H.; Iida, S. Lower expression of activating transcription factors 3 and 4 correlates with shorter progression-free survival in multiple myeloma patients receiving bortezomib plus dexamethasone therapy. *Blood Cancer J.* **2015**, *5*, e373. [[CrossRef](#)]
35. Rème, T.; Hose, D.; Theillet, C.; Klein, B. Modeling risk stratification in human cancer. *Bioinformatics* **2013**, *29*, 1149–1157. [[CrossRef](#)]
36. Samur, M.K.; Minvielle, S.; Gullà, A.; Fulciniti, M.; Cleynen, A.; Samur, A.A.; Szalat, R.; Shammas, M.; Magrangeas, F.; Tai, Y.T.; et al. Long Intergenic Non-Coding RNAs have an Independent Impact on Survival in Multiple Myeloma. *Leukemia* **2018**, *32*, 2626–2635. [[CrossRef](#)]
37. Shaughnessy, J.D.; Qu, P.; Usmani, S.; Heuck, C.J.; Zhang, Q.; Zhou, Y.; Tian, E.; Hanamura, I.; Van Rhee, F.; Anaissie, E.; 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. [[CrossRef](#)]
38. Shaughnessy, J.D.; Zhan, F.; Burington, B.E.; Huang, Y.; Colla, S.; Hanamura, I.; Stewart, J.P.; Kordsmeier, B.; Randolph, C.; Williams, D.R.; 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. [[CrossRef](#)]
39. Terragna, C.; Remondini, D.; Martello, M.; Zamagni, E.; Pantani, L.; Patriarca, F.; Pezzi, A.; Levi, G.; Offidani, M.; Proserpio, I.; et al. The genetic and genomic background of multiple myeloma patients achieving complete response after induction therapy with bortezomib, thalidomide and dexamethasone (VTD). *Oncotarget* **2016**, *7*, 9666–9679. [[CrossRef](#)]

40. Wu, S.P.; Pfeiffer, R.M.; Ahn, I.E.; Mailankody, S.; Sonneveld, P.; van Duin, M.; Munshi, N.C.; Walker, B.A.; Morgan, G.; Landgren, O. Impact of Genes Highly Correlated with MMSET Myeloma on the Survival of Non-MMSET Myeloma Patients. *Clin. Cancer Res.* **2016**, *22*, 4039–4044. [[CrossRef](#)]
41. Zhan, F.; Barlogie, B.; Arzoumanian, V.; Huang, Y.; Williams, D.R.; Hollmig, K.; Pineda-Roman, M.; Tricot, G.; van Rhee, F.; Zangari, M.; et al. Gene-expression signature of benign monoclonal gammopathy evident in multiple myeloma is linked to good prognosis. *Blood* **2007**, *109*, 1692–1700. [[CrossRef](#)] [[PubMed](#)]
42. Zhu, Y.X.; Braggio, E.; Shi, C.X.; Kortuem, K.M.; Bruins, L.A.; Schmidt, J.E.; Chang, X.B.; Langlais, P.; Luo, M.; Jedlowski, P.; et al. Identification of cereblon-binding proteins and relationship with response and survival after IMiDs in multiple myeloma. *Blood* **2014**, *124*, 536–545. [[CrossRef](#)] [[PubMed](#)]
43. Heuck, C.J.; Qu, P.; Van Rhee, F.; Waheed, S.; Usmani, S.Z.; Epstein, J.; Zhang, Q.; Edmondson, R.; Hoering, A.; Crowley, J.; et al. Five gene probes carry most of the discriminatory power of the 70-gene risk model in multiple myeloma. *Leukemia* **2014**, *28*, 2410–2413. [[CrossRef](#)] [[PubMed](#)]
44. Bolli, N.; Biancon, G.; Moarii, M.; Gimondi, S.; Li, Y.; De Philippis, C.; Maura, F.; Sathiaselvan, V.; Tai, Y.T.; Mudie, L.; et al. Analysis of the genomic landscape of multiple myeloma highlights novel prognostic markers and disease subgroups. *Leukemia* **2018**, *32*, 2604–2616. [[CrossRef](#)] [[PubMed](#)]
45. Zhan, F.; Tian, E.; Bumm, K.; Smith, R.; Barlogie, B.; Shaughnessy, J., Jr. Gene expression profiling of human plasma cell differentiation and classification of multiple myeloma based on similarities to distinct stages of late-stage B-cell development. *Blood* **2003**, *101*, 1128–1140. [[CrossRef](#)] [[PubMed](#)]
46. Kuiper, R.; Van Duin, M.; Van Vliet, M.H.; Broijl, A.; Van Der Holt, B.; El Jarari, L.; Van Beers, E.H.; Mulligan, G.; Avet-Loiseau, H.; Gregory, W.M.; et al. Prediction of high- and low-risk multiple myeloma based on gene expression and the International Staging System. *Blood* **2015**, *126*, 1996–2004. [[CrossRef](#)] [[PubMed](#)]
47. Chng, W.J.; Chung, T.H.; Kumar, S.; Usmani, S.; Munshi, N.; Avet-Loiseau, H.; Goldschmidt, H.; Durie, B.; Sonneveld, P. Gene signature combinations improve prognostic stratification of multiple myeloma patients. *Leukemia* **2016**, *30*, 1071–1078. [[CrossRef](#)] [[PubMed](#)]
48. Lagana, A.; Perumal, D.; Melnekoff, D.; Readhead, B.; Kidd, B.A.; Leshchenko, V.; Kuo, P.Y.; Keats, J.; DeRome, M.; Yesil, J.; et al. Integrative network analysis identifies novel drivers of pathogenesis and progression in newly diagnosed multiple myeloma. *Leukemia* **2018**, *32*, 120–130. [[CrossRef](#)] [[PubMed](#)]
49. Chapman, M.A.; Lawrence, M.S.; Keats, J.J.; Cibulskis, K.; Sougnez, C.; Schinzel, A.C.; Harview, C.L.; Brunet, J.P.; Ahmann, G.J.; Adli, M.; et al. Initial genome sequencing and analysis of multiple myeloma. *Nature* **2011**, *471*, 467–472. [[CrossRef](#)] [[PubMed](#)]
50. Zhu, J.; Zhang, B.; Smith, E.N.; Drees, B.; Brem, R.B.; Kruglyak, L.; Bumgarner, R.E.; Schadt, E.E. Integrating large-scale functional genomic data to dissect the complexity of yeast regulatory networks. *Nat. Genet.* **2008**, *40*, 854–861. [[CrossRef](#)]
51. Yoo, S.; Huang, T.; Campbell, J.D.; Lee, E.; Tu, Z.; Geraci, M.W.; Powell, C.A.; Schadt, E.E.; Spira, A.; Zhu, J. MODMatcher: Multi-omics data matcher for integrative genomic analysis. *PLoS Comput. Biol.* **2014**, *10*, e1003790. [[CrossRef](#)] [[PubMed](#)]
52. Lee, E.; Yoo, S.; Wang, W.; Tu, Z.; Zhu, J. A probabilistic multi-omics data matching method for detecting sample errors in integrative analysis. *GigaScience* **2019**, *8*, giz080. [[CrossRef](#)] [[PubMed](#)]
53. Szalat, R.; Avet-Loiseau, H.; Munshi, N.C. Gene Expression Profiles in Myeloma: Ready for the Real World? *Clin. Cancer Res.* **2016**, *22*, 5434–5442. [[CrossRef](#)] [[PubMed](#)]
54. Smetana, J.; Fröhlich, J.; Zaoralova, R.; Vallová, V.; Greslikova, H.; Kupska, R.; Němec, P.; Mikulášová, A.; Almasi, M.; Pour, L.; et al. Genome-Wide Screening of Cytogenetic Abnormalities in Multiple Myeloma Patients Using Array-CGH Technique: A Czech Multicenter Experience. *BioMed Res. Int.* **2014**, *2014*, 1–9. [[CrossRef](#)] [[PubMed](#)]
55. Redon, R.; Ishikawa, S.; Fitch, K.R.; Feuk, L.; Perry, G.H.; Andrews, T.D.; Fiegler, H.; Shaperro, M.H.; Carson, A.R.; Chen, W.; et al. Global variation in copy number in the human genome. *Nature* **2006**, *444*, 444–454. [[CrossRef](#)] [[PubMed](#)]
56. Ohshima, K.; Hatakeyama, K.; Nagashima, T.; Watanabe, Y.; Kanto, K.; Doi, Y.; Ide, T.; Shimoda, Y.; Tanabe, T.; Ohnami, S.; et al. Integrated analysis of gene expression and copy number identified potential cancer driver genes with amplification-dependent overexpression in 1,454 solid tumors. *Sci. Rep.* **2017**, *7*, 1546. [[CrossRef](#)] [[PubMed](#)]

57. Zhou, C.; Zhang, W.; Chen, W.; Yin, Y.; Atyah, M.; Liu, S.; Guo, L.; Shi, Y.; Ye, Q.; Dong, Q.; et al. Integrated analysis of copy number variations and gene expression profiling in hepatocellular carcinoma. *Sci. Rep.* **2017**, *7*, 10570. [[CrossRef](#)] [[PubMed](#)]
58. Gong, Y.; Wang, L.; Chippada-Venkata, U.; Dai, X.; Oh, W.K.; Zhu, J. Constructing Bayesian networks by integrating gene expression and copy number data identifies NLGN4Y as a novel regulator of prostate cancer progression. *Oncotarget* **2016**, *7*, 68688–68707. [[CrossRef](#)] [[PubMed](#)]
59. Tran, L.M.; Zhang, B.; Zhang, Z.; Zhang, C.; Xie, T.; Lamb, J.R.; Dai, H.; Schadt, E.E.; Zhu, J. Inferring causal genomic alterations in breast cancer using gene expression data. *BMC Syst. Boil.* **2011**, *5*, 121. [[CrossRef](#)] [[PubMed](#)]
60. Kuehl, W.M.; Bergsagel, P.L. Multiple myeloma: Evolving genetic events and host interactions. *Nat. Rev. Cancer* **2002**, *2*, 175–187. [[CrossRef](#)] [[PubMed](#)]
61. Agnelli, L.; Bicciato, S.; Mattioli, M.; Fabris, S.; Intini, D.; Verdelli, D.; Baldini, L.; Morabito, F.; Callea, V.; Lombardi, L.; 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. [[CrossRef](#)]
62. Agnelli, L.; Mosca, L.; Fabris, S.; Lionetti, M.; Andronache, A.; Kwee, I.; Todoerti, K.; Verdelli, D.; Battaglia, C.; Bertoni, F.; 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 Chromosom. Cancer* **2009**, *48*, 603–614. [[CrossRef](#)]
63. Musgrove, E.A.; Caldon, C.E.; Barraclough, J.; Stone, A.; Sutherland, R.L. Cyclin D as a therapeutic target in cancer. *Nat. Rev. Cancer* **2011**, *11*, 558–572. [[CrossRef](#)]
64. Zhan, F.; Huang, Y.; Colla, S.; Stewart, J.P.; Hanamura, I.; Gupta, S.; Epstein, J.; Yaccoby, S.; Sawyer, J.; Burington, B.; et al. The molecular classification of multiple myeloma. *Blood* **2006**, *108*, 2020–2028. [[CrossRef](#)]
65. Kuo, A.J.; Cheung, P.; Chen, K.; Zee, B.M.; Kioi, M.; Lauring, J.; Xi, Y.; Park, B.H.; Shi, X.; Garcia, B.A.; et al. NSD2 links dimethylation of histone H3 at lysine 36 to oncogenic programming. *Mol. Cell* **2011**, *44*, 609–620. [[CrossRef](#)]
66. De Vet, E.C.; Hilkes, Y.H.; Fraaije, M.W.; van den Bosch, H. Alkyl-dihydroxyacetonephosphate synthase. Presence and role of flavin adenine dinucleotide. *J. Biol. Chem.* **2000**, *275*, 6276–6283. [[CrossRef](#)]
67. Falank, C.; Fairfield, H.; Reagan, M.R. Signaling interplay between bone marrow adipose tissue and multiple myeloma cells. *Front. Endocrinol.* **2016**, *7*, 57. [[CrossRef](#)]
68. Piano, V.; Benjamin, D.I.; Valente, S.; Nenci, S.; Mai, A.; Aliverti, A.; Nomura, D.K.; Mattevi, A. Discovery of Inhibitors for the Ether Lipid-Generating Enzyme AGPS as Anti-Cancer Agents. *ACS Chem. Boil.* **2015**, *10*, 2589–2597. [[CrossRef](#)]
69. Walker, B.A.; Mavrommatis, K.; Wardell, C.P.; Ashby, T.C.; Bauer, M.; Davies, F.E.; Rosenthal, A.; Wang, H.; Qu, P.; Hoering, A.; et al. Identification of novel mutational drivers reveals oncogene dependencies in multiple myeloma. *Blood* **2018**, *132*, 587–597. [[CrossRef](#)]
70. Marzin, Y.; Jamet, D.; Douet-Guilbert, N.; Morel, F.; Le Bris, M.J.; Morice, P.; Abgrall, J.F.; Berthou, C.; De Braekeleer, M. Chromosome 1 abnormalities in multiple myeloma. *Anticancer. Res.* **2006**, *26*, 953–959.
71. Evans, R.; Naber, C.; Steffler, T.; Checkland, T.; Keats, J.; Maxwell, C.; Perry, T.; Chau, H.; Belch, A.; Pilarski, L.; et al. Aurora A kinase RNAi and small molecule inhibition of Aurora kinases with VE-465 induce apoptotic death in multiple myeloma cells. *Leuk. Lymphoma* **2008**, *49*, 559–569. [[CrossRef](#)]
72. Maxwell, C.A.; Rasmussen, E.; Zhan, F.; Keats, J.J.; Adamia, S.; Strachan, E.; Crainie, M.; Walker, R.; Belch, A.R.; Pilarski, L.M.; et al. RHAMM expression and isoform balance predict aggressive disease and poor survival in multiple myeloma. *Blood* **2004**, *104*, 1151–1158. [[CrossRef](#)]
73. Santo, L.; Hideshima, T.; Cirstea, D.; Bandi, M.; Nelson, E.A.; Gorgun, G.; Rodig, S.; Vallet, S.; Pozzi, S.; Patel, K.; et al. Antimyeloma activity of a multitargeted kinase inhibitor, AT9283, via potent Aurora kinase and STAT3 inhibition either alone or in combination with lenalidomide. *Clin. Cancer Res.* **2011**, *17*, 3259–3271. [[CrossRef](#)]
74. Goga, A.; Yang, D.; Tward, A.D.; Morgan, D.O.; Bishop, J.M. Inhibition of CDK1 as a potential therapy for tumors over-expressing MYC. *Nat. Med.* **2007**, *13*, 820–827. [[CrossRef](#)]
75. McMillin, D.W.; Delmore, J.; Negri, J.; Buon, L.; Jacobs, H.M.; Laubach, J.; Jakubikova, J.; Ooi, M.; Hayden, P.; Schlossman, R.; et al. Molecular and cellular effects of multi-targeted cyclin-dependent kinase inhibition in myeloma: Biological and clinical implications. *Br. J. Haematol.* **2011**, *152*, 420–432. [[CrossRef](#)]

76. Pruneri, G.; Carboni, N.; Baldini, L.; Intini, D.; Colombi, M.; Bertolini, F.; Valentini, S.; Maisonneuve, P.; Viale, G.; Neri, A. Cell cycle regulators in multiple myeloma: Prognostic implications of p53 nuclear accumulation. *Hum. Pathol.* **2003**, *34*, 41–47. [[CrossRef](#)]
77. Quach, H.; Ritchie, D.; Stewart, A.K.; Neeson, P.; Harrison, S.; Smyth, M.J.; Prince, H.M. Mechanism of action of immunomodulatory drugs (IMiDs) in multiple myeloma. *Leukemia* **2010**, *24*, 22–32. [[CrossRef](#)]
78. Drach, J.; Slany, A.; Sagaster, V.; Gundacker, N.; Haudek, V.; Odelga, V.; Wimmer, H.; Zielinski, C.; Gerner, C. A cellular proteome map of human multiple myeloma. *Blood* **2007**, *110*, 111.
79. Moreaux, J.; Rème, T.; Leonard, W.; Veyrune, J.L.; Requirand, G.; Goldschmidt, H.; Hose, D.; Klein, B. Development of Gene Expression-Based Score to Predict Sensitivity of Multiple Myeloma Cells to DNA Methylation Inhibitors. *Mol. Cancer Ther.* **2012**, *11*, 2685–2692. [[CrossRef](#)]
80. Lin, L.; Cao, L.; Liu, Y.; Wang, K.; Zhang, X.; Qin, X.; Zhao, D.; Hao, J.; Chang, Y.; Huang, X.; et al. B7-H3 promotes multiple myeloma cell survival and proliferation by ROS-dependent activation of Src/STAT3 and c-Cbl-mediated degradation of SOCS3. *Leukemia* **2019**, *33*, 1475–1486. [[CrossRef](#)]
81. Wilop, S.; Van Gemmeren, T.B.; Lentjes, M.H.; Van Engeland, M.; Herman, J.G.; Brümmendorf, T.H.; Jost, E.; Galm, O. Methylation-associated dysregulation of the suppressor of cytokine signaling-3 gene in multiple myeloma. *Epigenetics* **2011**, *6*, 1047–1052. [[CrossRef](#)]
82. Nardiello, T.; Jungbluth, A.A.; Mei, A.; DiLiberto, M.; Huang, X.; Dabrowski, A.; Andrade, V.C.C.; Wasserstrum, R.; Ely, S.; Niesvizky, R.; et al. MAGE-A inhibits apoptosis in proliferating myeloma cells through repression of Bax and maintenance of survivin. *Clin. Cancer Res.* **2011**, *17*, 4309–4319. [[CrossRef](#)]
83. Chauhan, D.; Hideshima, T.; Anderson, K.C. PROTEASOME INHIBITION IN MULTIPLE MYELOMA: Therapeutic Implication. *Annu. Rev. Pharmacol. Toxicol.* **2005**, *45*, 465–476. [[CrossRef](#)]
84. Zhan, F.; Hardin, J.; Kordsmeier, B.; Bumm, K.; Zheng, M.; Tian, E.; Sanderson, R.; Yang, Y.; Wilson, C.; Zangari, M.; et al. Global gene expression profiling of multiple myeloma, monoclonal gammopathy of undetermined significance, and normal bone marrow plasma cells. *Blood* **2002**, *99*, 1745–1757. [[CrossRef](#)]
85. Broyl, A.; Hose, D.; Lokhorst, H.; De Knecht, Y.; Peeters, J.; Jauch, A.; Bertsch, U.; Buijs, A.; Stevens-Kroef, M.; Beverloo, H.B.; et al. Gene expression profiling for molecular classification of multiple myeloma in newly diagnosed patients. *Blood* **2010**, *116*, 2543–2553. [[CrossRef](#)]
86. Khoo, W.H.; Ledergor, G.; Weiner, A.; Roden, D.L.; Terry, R.L.; McDonald, M.M.; Chai, R.C.; De Veirman, K.; Owen, K.L.; Opperman, K.S.; et al. A niche-dependent myeloid transcriptome signature defines dormant myeloma cells. *Blood* **2019**, *134*, 30–43. [[CrossRef](#)]
87. Layfield, L.J.; Anderson, G.M. Specimen labeling errors in surgical pathology: An 18-month experience. *Am. J. Clin. Pathol.* **2010**, *134*, 466–470. [[CrossRef](#)]
88. Yelland, L.N.; Kahan, B.C.; Dent, E.; Lee, K.J.; Voysey, M.; Forbes, A.B.; Cook, J.A. Prevalence and reporting of recruitment, randomisation and treatment errors in clinical trials: A systematic review. *Clin. Trials* **2018**, *15*, 278–285. [[CrossRef](#)]
89. Suba, E.J.; Pfeifer, J.D.; Raab, S.S. Patient Identification Error Among Prostate Needle Core Biopsy Specimens—Are We Ready for a DNA Time-Out? *J. Urol.* **2007**, *178*, 1245–1248. [[CrossRef](#)]
90. Lippi, G.; Chiozza, L.; Mattiuzzi, C.; Plebani, M. Patient and Sample Identification. Out of the Maze? *J. Med. Biochem.* **2017**, *36*, 107–112. [[CrossRef](#)]
91. Ubels, J.; Sonneveld, P.; Van Beers, E.H.; Broijl, A.; Van Vliet, M.H.; De Ridder, J. Predicting treatment benefit in multiple myeloma through simulation of alternative treatment effects. *Nat. Commun.* **2018**, *9*, 2943. [[CrossRef](#)] [[PubMed](#)]
92. Basak, G.W.; Srivastava, A.S.; Malhotra, R.; Carrier, E. Multiple myeloma bone marrow niche. *Curr. Pharm. Biotechnol.* **2009**, *10*, 345–346. [[CrossRef](#)] [[PubMed](#)]
93. Ramsköld, D.; Luo, S.; Wang, Y.C.; Li, R.; Deng, Q.; Faridani, O.R.; Daniels, G.A.; Khrebtkova, I.; Loring, J.F.; Laurent, L.C.; et al. Full-length mRNA-Seq from single-cell levels of RNA and individual circulating tumor cells. *Nat. Biotechnol.* **2012**, *30*, 777–782. [[CrossRef](#)] [[PubMed](#)]
94. Ledergor, G.; Weiner, A.; Zada, M.; Wang, S.Y.; Cohen, Y.C.; Gatt, M.E.; Snir, N.; Magen, H.; Koren-Michowitz, M.; Herzog-Tzarfati, K.; et al. Single cell dissection of plasma cell heterogeneity in symptomatic and asymptomatic myeloma. *Nat. Med.* **2018**, *24*, 1867–1876. [[CrossRef](#)] [[PubMed](#)]

95. Jang, J.S.; Li, Y.; Mitra, A.K.; Bi, L.; Abyzov, A.; van Wijnen, A.J.; Baughn, L.B.; Van Ness, B.; Rajkumar, V.; Kumar, S.; et al. Molecular signatures of multiple myeloma progression through single cell RNA-Seq. *Blood Cancer J.* **2019**, *9*, 2. [[CrossRef](#)] [[PubMed](#)]
96. Staedtler, F.; Hartmann, N.; Letzkus, M.; Bongiovanni, S.; Scherer, A.; Marc, P.; Johnson, K.J.; Schumacher, M.M. Robust and tissue-independent gender-specific transcript biomarkers. *Biomarkers* **2013**, *18*, 436–445. [[CrossRef](#)] [[PubMed](#)]
97. Li, Y.; Wang, X.; Zheng, H.; Wang, C.; Minvielle, S.; Magrangeas, F.; Avet-Loiseau, H.; Shah, P.K.; Zhang, Y.; Munshi, N.C.; et al. Classify hyperdiploidy status of multiple myeloma patients using gene expression profiles. *PLoS ONE* **2013**, *8*, e58809. [[CrossRef](#)] [[PubMed](#)]
98. Zhu, J.; Wiener, M.C.; Zhang, C.; Fridman, A.; Minch, E.; Lum, P.Y.; Sachs, J.R.; Schadt, E.E. Increasing the power to detect causal associations by combining genotypic and expression data in segregating populations. *PLoS Comput. Biol.* **2007**, *3*, 69. [[CrossRef](#)] [[PubMed](#)]
99. Zhu, J.; Sova, P.; Xu, Q.; Dombek, K.M.; Xu, E.Y.; Vu, H.; Tu, Z.; Brem, R.B.; Bumgarner, R.E.; Schadt, E.E. Stitching together multiple data dimensions reveals interacting metabolomic and transcriptomic networks that modulate cell regulation. *PLoS Biol.* **2012**, *10*, e1001301. [[CrossRef](#)] [[PubMed](#)]
100. Friedman, N.; Linial, M.; Nachman, I.; Pe’Er, D. Using Bayesian networks to analyze expression data. *J. Comput. Biol.* **2000**, *7*, 127–135. [[CrossRef](#)]
101. Su, G.; Morris, J.H.; Demchak, B.; Bader, G.D. Biological network exploration with cytoscape 3. *Curr. Protoc. Bioinform.* **2014**, *47*, 8–13. [[CrossRef](#)] [[PubMed](#)]
102. Yang, X.; Peterson, L.; Thieringer, R.; Deignan, J.L.; Wang, X.; Zhu, J.; Wang, S.; Zhong, H.; Stepaniants, S.; Beaulaurier, J.; et al. Identification and validation of genes affecting aortic lesions in mice. *J. Clin. Investig.* **2010**, *120*, 2414–2422. [[CrossRef](#)] [[PubMed](#)]
103. Yang, X.; Tu, Z.; Zhu, J. Systems Biology Approaches to Studying Diet x Genome Interactions. In *Nutritional Genomics*; CRC Press: Boca Raton, FL, USA, 2011; pp. 63–76. [[CrossRef](#)]
104. Mi, H.; Muruganujan, A.; Ebert, D.; Huang, X.; Thomas, P.D. PANTHER version 14: More genomes, a new PANTHER GO-slim and improvements in enrichment analysis tools. *Nucleic Acids Res.* **2019**, *47*, D419–D426. [[CrossRef](#)] [[PubMed](#)]



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).