Whole-genome analysis identifies novel drivers and high-risk double-hit events in relapsed/refractory myeloma

Supplemental Methods and Data File

Patient characteristics

The Myeloma Genome Project is an ongoing initiative to assemble and analyse, in a uniform and innovative fashion, genetic data sets that have been generated on samples obtained from patients with myeloma who have been entered into clinical trials. The relapse/refractory multiple myeloma (rrMM) cohort comprised patients from the following clinical trials of immunomodulatoy agents (avadomide (CC-122), pomalidomide (CC-4047), or iberdomide (CC-220)) either alone or in combination): CC-4074-MM010 (STRATUS; NCT017127899)[1], CC-4047-MM-007 (OPTIMISMM; NCT0173492810)[2], CC-4047-MM013 (NCT02045017)[3], CC-4047-MM-014-B (NCT01946477)[4], CC220-MM001 (NCT0277303011)[5] and CC122-ST-001-MM2 (NCT01421524)[6]. The rrMM cohort was compared with 198 ndMM patients collected as part of the IFM/DFCI-2009 trial (NCT0119106012)[7]. Most tumor samples were collected at screening/study entry. All the rrMM trials included combination therapy with dexamethasone, the exception was CC122-ST-001-MM2 which was a cohort of MM patients in a trial investigating CC-122 monotherapy in cohorts of patients with advanced malignancies. Two of the pomalidomide (CC-4047) trials included other anti-MM agents, either bortezomib in CC-4047-MM-007 or daratumumab in CC-4047-MM-014B. Patients from the CC-220-MM-001 study were in dose escalation cohorts of either monotherapy or combination with dexamethasone. As most patients were from pomalidomide trials, which required either exposure or refractoriness to lenalidomide, the lenalidomide resistance (LENR) cohort was much larger than the pomalidomide resistance (POMR) cohort comprising samples from patients from the CC-122 or CC-220 trials, or patients in MM-010 with a sample at disease progression. The ndMM dataset comprised WGS data from patients in IFM2009 (n=198)[8].
Summary patient characteristics per clinical trial are described in Supplemental Table 1. All participants in the trials that have contributed to the Myeloma Genome Project gave consent for the use of their deidentified genetic and clinical data. Each trial received approval from the relevant ethical review committee.

Genomic material preparation

CD138+ plasma cells and matched germline controls (peripheral blood mononuclear cells) were stored in RLT buffer (Qiagen) and DNA was extracted using the Qiagen AllPrep DNA Mini Kit.

WGS was performed on tumor/normal pairs at an average of 60X/30X depth. All sequencing was performed on Illumina HiSeq2500 instruments.

Whole-genome sequence analysis, variant calling and annotation

All whole-genome sequence analyses included herein is based on GATK best practices (https://gatk.broadinstitute.org/hc/en-us/articles/360035894751-Pipeline-Index). A total of 418 relapse/refractory and 198 newly diagnosed tumors were sequenced at the whole-genome level along with paired normal sample. All FASTQ files were aligned to the human genome assembly hg38 using BWA-mem. Duplicate reads were removed using Picard MarkDuplicates (v1.119; https://broadinstitute.github.io/picard/) and base recalibration of alignments was performed using BQSR according to GATK best practices (https://gatk.broadinstitute.org/hc/en-us/articles/360035535912). Somatic variants in WGS data were called using MuTect2 *(*v4) under default parameters. Somatic SNVs and indels were annotated by ANNOVAR[9] (version May2018; http://annovar.openbioinformatics.org/) to predict functional consequence of variants. In addition, variants were identified based on dbSNP (v150), and population frequency of variants were reported based on the Exome Aggregation Consortium dataset (ExAC v0.3; http://exac.broadinstitute.org/) and the Genome Aggregation Database (gnomAD v2.1.1; https://gnomad.broadinstitute.org/).

Copy number aberration (CNA) analysis

Genome-wide copy number aberration analysis was undertaken in the entire dataset by using Battenberg (v2.2.8; https://github.com/Wedge-lab/battenberg), which has been described in detail previously[10]. To handle hg38-based assembled genomes, germline SNP data underwent liftover to hg19 (https://genome.ucsc.edu/cgi-bin/hgLiftOver) for compatibility purposes. In addition to calling clonal and subclonal allele-specific CNAs, it was also used to estimate purity and average ploidy of each tumor. To identify significantly enriched CNA regions, first, CNA of each type – i.e. Gain, loss of heterozygosity (LOH) and homozygous deletion (HD) – were aggregated across all rrMM samples along the chromosomes to obtain the frequency landscape of each CNA type based on all observed breakpoints. This was followed by a permutation test (n=1,000) which was then corrected for multiple testing (FDR<0.05). Regions that were significantly enriched above the random background rate for any CNA type were defined as enriched CNA[11]. Finally, the enriched regions that encompassed the HLA region (6p21), were specific to telomeric ends or present as a singleton were excluded. To estimate the overall LOH burden in each tumor, we calculated the proportion of genome altered (PGA) that could be assigned to LOH events (where nMinor is 0) at both clonal (frac1=1) and subclonal (frac1<1) levels.

Clonality of somatic variants

To estimate clonality, we calculated cancer cell fraction (CCF) of each variant by adjusting the variant allele frequency (VAF) for copy number status at the given locus, multiplicity of the variant and tumor purity[12]. To assign somatic mutations as clonal or subclonal, the CCF of each mutation was statistically assessed for clonal status. Briefly, the observed VAF was modelled using a binomial distribution (qbinom) by taking into account the total coverage at a given locus and VAFs representing the 95% interval boundaries (i.e. 2.5% and 97.5%) were used to generate the 95% confidence interval (CI) of the observed CCF by computing CCF for the VAF boundaries. Any variant with an upper CI above 1 was considered to not significantly deviate from a clonal state and a CCF value of 1 was thus assigned. Otherwise, variants were considered subclonal and the original estimated CCF value was retained. This allowed us to assess the clonality of all somatic mutations without the introduction of an arbitrary CCF cut-off.

Mutational Drivers

To identify potential drivers of relapse/refractory stage of MM, we used cDriver[13] (v0.4.2; https://github.com/hanasusak/cDriver), which not only relies on recurrence and functional consequence of variants, but also takes into account CCF of each mutation, which may lead to the identification of subclonal drivers. Genes mapping to immune hypermutated regions were excluded to avoid bias from mutation due to normal B cell development. Drivers not previously reported in newly diagnosed MM[14] were checked against the COSMIC v92 cancer gene census (CGC; https://cancer.sanger.ac.uk/census) to identify known cancer drivers. Assignment of identified drivers into tumor suppressor gene (TSG) and oncogene (ONC) categories was undertaken using the 20/20 principle[15]. We analyzed all drivers for gene essentiality as reported in DepMap[16]. A customised dataset of CRISPR effect scores was generated and downloaded from DepMap (https://depmap.org/) by limiting the genes to the drivers identified in this study and cell lines to multiple myeloma cell lines (N=21). The mutational landscape plot was generated using maftools[17] (v2.6.05; https://www.bioconductor.org/packages/release/bioc/html/maftools.html).

Network and pathway enrichment analysis

The interactome database STRING (Search Tool for the Retrieval of Interacting Genes/Proteins) v11.5 (http://string-db.org/)[18] was used to extract all known and predicted protein–protein interactions (PPIs) among driver genes. Interactions based on only textmining data were removed. This filtering

process was undertaken to remove possible spurious interactions that might result in misleading conclusions. Interactions with scores above the STRING default threshold (0.4) were analyzed. Enrichr[19] was used to identify enriched pathways among the novel promising candidate gene set (N=10). The pathways were filtered with an enrichment *P*-value < 0.05 after adjusting with FDR multiple testing correction.

Bi-allelic and double-hit analysis

To test for somatic interactions at the individual tumor level and assess likely tumor suppression activity, we combined the CNA and mutation data to assess the rate of bi-allelic inactivation of driver genes. Biallelic events were defined as the occurrence of two independent events inactivating both alleles at the same locus, which could result from a combination of a non-silent mutation and loss of heterozygosity (LOH) or by homozygous deletion (HD) in a given tumor. Drivers with a bi-allelic inactivation rate above 1% in rrMM were retained for subsequent differential analysis. In addition, we identified double-hit events in each individual tumor based on the co-occurrence of two or more high-risk genomic and clinical features[20] and those above 1% frequency in rrMM were used for enrichment analysis.

Mutational signature analysis

De-novo extraction of single base substitution (SBS) signatures was independently implemented based on a non-negative matrix factorization (NNMF) framework using SigProfilerExtractor[21] (v1.0.19, https://github.com/AlexandrovLab/SigProfilerExtractor) and signeR[22] which is an empirical Bayesian treatment of the NNMF model. Because de novo extraction of signatures is more accurate with a larger number of samples[23], the entire rrMM dataset was combined with the ndMM dataset. During mutational spectrum quality control screening, an ndMM tumor showed substantial SBS53 activity (likely artefact signature[21]) with high mutation burden and was removed from further analysis. The final input dataset for both SigProfilerExtractor and signeR thus comprised a total of 616 tumors. De novo signatures extracted by both methods were each assigned to a specific SBS signature in the recently updated COSMIC catalogue (https://cancer.sanger.ac.uk/cosmic/signatures/SBS/) by computing cosine similarities. The mutational signature profiles from both extraction methods were similar and identified the same set of COSMIC signatures in the dataset. To fit the mutational spectrum of each tumor with the identified COSMIC signatures, we used mmsig[24]. This was implemented to not only minimise mutational signature bleeding across samples[25], but to also accurately estimate the relative contribution of COSMIC signatures. Finally, at the fitting stage, we also included the novel multiple myeloma signature (SBS-MM1[26,27] to see whether its inclusion better reconstructs the observed mutational spectrum in

each tumor and evaluate its activity across this large MM dataset with both ndMM and rrMM tumors.

Detection of drivers based on non-coding variants

We partitioned each chromosome in the genome into discrete bins of 100 kb and undertook a genome-wide screening of variant recurrence in each of the non-overlapping bins[11]. Similar to a genome-wide association study construct, we compared the rrMM cohort with the ndMM cohort for differential prevalence of non-coding variant recurrence using pairwise Fisher's exact test followed by multiple testing correction to detect enriched non-coding mutation hotspots in or flanking coding genes in rrMM. This analysis was based only on SNVs since overall rate of SNV was not significantly different between the two groups. Therefore, no overall bias is present in the rate of SNVs and local over-representation signals are likely to be genuine relapse/refractory hotspot signals.

Kataegis

We followed previous studies in defining kataegis events[11,28]. KataegisPCF (https://github.com/nansari-pour/KataegisPCF) was used to accurately detect kataegis foci and visualize the events based on SNVs. A minimum of six consecutive SNVs with mean distance ≤1kb were required to call kataegis foci, which were identified systematically by applying piecewise constant fitting (PCF)[29] on inter-variant distance of all SNVs per chromosome across the genome.

Structural variant analysis

Structural variants (SV) from whole genome data were detected using MANTA *(v1.2.0)*[30] with default tumor-normal pair settings where tumor samples were compared against the germline control to detect somatic SVs. A custom script was then used to parse and summarise translocation (TL) events from the VCF output. All samples with TL events with a breakpoint within the immunoglobulin heavy (IgH) and light (IgL) chains were analysed further for molecular subtyping of MM.

Statistical analysis

All statistical calculations were implemented in R (v3.4.3; https://www.r-project.org/). For continuous data, we used Wilcoxon rank test (*wilcox.test*) or Student's t-test (*t.test*) wherever appropriate. For categorical data, we used Fisher's exact test (*fisher.test*). P-values were adjusted for multiple testing (*p.adjust*) based on the false discovery rate (FDR) proposed by

Benjamini and Hochberg with FDR<0.05 considered significant, unless stated otherwise. This was done to not only reduce type I error, but to also minimize type II error[31]. The two-sample Kolmogorov-Smirnov test (*ks.test*) was implemented to detect significant differences in the distribution of a variable across different groups. Due to the low frequencies of mutational drivers, we used Fisher's exact test as a non-parametric test to assess equality of proportions by using the fisher.test function in R. Such a test was undertaken when the mutational driver had a minimum frequency of 5% in at least one cohort. The Pearson correlation test (*cor.test*) was used to assess the magnitude and significance of pairwise correlation between mutation burden and mutational signature activity. Cox regression survival analysis (*coxph*) was undertaken to examine association between common rrMM-enriched events and progression free survival (PFS) and, in turn, identify aberrations with prognostic effect.

Supplemental Results

Kataegis

We also analyzed all tumors for kataegis[28] to see whether hypermutated genes can be identified at the individual tumor level, however, all events were mainly localized at immunoglobulin (IG) loci consistent with their likely link to the IG translocation events (Supplemental Figure 11).

Copy number analysis

We saw a slight enrichment of whole-genome duplication (WGD) in rrMM (7.8% vs 5.5%, OR=1.5) but was not statistically significant. We also assessed the proportion of the genome altered due to LOH (PGA_{LOH}) in rrMM and whether it varies between LENR and POMR cohorts. We observed no significant difference in the rate of PGA_{LOH} between ndMM and rrMM, nor among ndMM, LENR and POMR (Kolmogorov-Smirnov test, P>0.40; Supplemental Figure 12).

In a supervised analysis of the Cereblon locus from a subset of our rrMM dataset, we previously reported an increase in LOH of 3p26.3 (containing *CRBN*) in LENR and POMR patients compared to ndMM[32], however 3p26.3 LOH was not significant in this genome-wide unsupervised analysis. Similarly, we have performed supervised analysis of chromosome 2q, which contains 2 members of the COPS9 signalosome which are required for Cereblon function, but this region was also not one of the significant LOH events at the genome-wide

level[33]. This suggests that while 3p26.3 and 2q LOH play important roles in the acquisition of IMiD resistance, they are not among the most frequent copy number changes driving aggressive disease biology in rrMM.

References for Supplemental Methods and Results

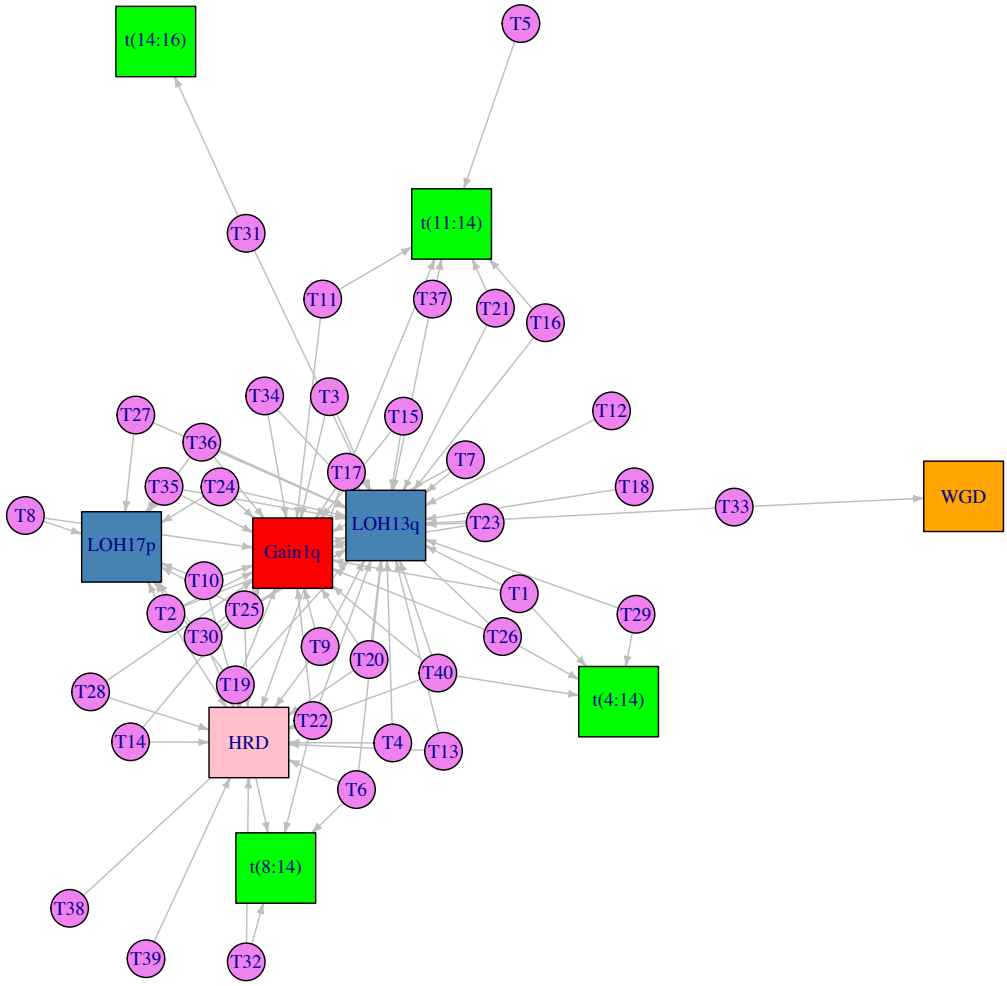
1. Dimopoulos MA, Palumbo A, Corradini P, et al. Safety and efficacy of pomalidomide plus low-dose dexamethasone in STRATUS (MM-010): a phase 3b study in refractory multiple myeloma. *Blood*. 2016;128(4):497-503.
2. Richardson PG, Oriol A, Beksac M, et al. Pomalidomide, bortezomib, and dexamethasone for patients with relapsed or refractory multiple myeloma previously treated with lenalidomide (OPTIMISMM): a randomised, open-label, phase 3 trial. *Lancet Oncol*. 2019;20(6):781-794.
3. Dimopoulos M, Weisel K, van de Donk N, et al. Pomalidomide Plus Low-Dose Dexamethasone in Patients With Relapsed/Refractory Multiple Myeloma and Renal Impairment: Results From a Phase II Trial. *J Clin Oncol*. 2018;36(20):2035-2043.
4. Pelligra CG, Parikh K, Guo S, et al. Cost-effectiveness of Pomalidomide, Carfilzomib, and Daratumumab for the Treatment of Patients with Heavily Pretreated Relapsed-refractory Multiple Myeloma in the United States. *Clin Ther*. 2017;39(10):1986-2005 e1985.
5. Lonial S, van de Donk NWCJ, Popat R, et al. First clinical (phase 1b/2a) study of iberdomide (CC-220; IBER), a CELMoD, in combination with dexamethasone (DEX) in patients (pts) with relapsed/refractory multiple myeloma (RRMM). *Journal of Clinical Oncology*. 2019;37(15_suppl):8006-8006.
6. Rasco DW, Papadopoulos KP, Pourdehnad M, et al. A First-in-Human Study of Novel Cereblon Modulator Avadomide (CC-122) in Advanced Malignancies. *Clin Cancer Res*. 2019;25(1):90-98.
7. Attal M, Lauwers-Cances V, Hulin C, et al. Lenalidomide, Bortezomib, and Dexamethasone with Transplantation for Myeloma. *N Engl J Med*. 2017;376(14):1311-1320.
8. Samur MK, Aktas Samur A, Fulciniti M, et al. Genome-Wide Somatic Alterations in Multiple Myeloma Reveal a Superior Outcome Group. *J Clin Oncol*. 2020;38(27):3107-3118.
9. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res*. 2010;38(16):e164.
10. Nik-Zainal S, Van Loo P, Wedge DC, et al. The life history of 21 breast cancers. *Cell*. 2012;149(5):994-1007.
11. Ansari-Pour N, Zheng Y, Yoshimatsu TF, et al. Whole-genome analysis of Nigerian patients with breast cancer reveals ethnic-driven somatic evolution and distinct genomic subtypes. *Nat Commun*. 2021;12(1):6946.
12. Rabbie R, Ansari-Pour N, Cast O, et al. Multi-site clonality analysis uncovers pervasive heterogeneity across melanoma metastases. *Nat Commun*. 2020;11(1):4306.
13. Zapata L, Susak H, Drechsel O, Friedlander MR, Estivill X, Ossowski S. Signatures of positive selection reveal a universal role of chromatin modifiers as cancer driver genes. *Sci Rep*. 2017;7(1):13124.
14. Walker BA, Mavrommatis K, Wardell CP, et al. Identification of novel mutational drivers reveals oncogene dependencies in multiple myeloma. *Blood*. 2018;132(6):587-597.
15. Vogelstein B, Papadopoulos N, Velculescu VE, Zhou S, Diaz LA, Jr., Kinzler KW. Cancer genome landscapes. *Science*. 2013;339(6127):1546-1558.

16. Tsherniak A, Vazquez F, Montgomery PG, et al. Defining a Cancer Dependency Map. *Cell*. 2017;170(3):564-576 e516.
17. Mayakonda A, Lin DC, Assenov Y, Plass C, Koeffler HP. Maftools: efficient and comprehensive analysis of somatic variants in cancer. *Genome Res*. 2018;28(11):1747-1756.
18. Szklarczyk D, Gable AL, Nastou KC, et al. The STRING database in 2021: customizable protein-protein networks, and functional characterization of user-uploaded gene/measurement sets. *Nucleic Acids Res*. 2021;49(D1):D605-D612.
19. Xie Z, Bailey A, Kuleshov MV, et al. Gene Set Knowledge Discovery with Enrichr. *Curr Protoc*. 2021;1(3):e90.
20. 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.
21. Alexandrov LB, Kim J, Haradhvala NJ, et al. The repertoire of mutational signatures in human cancer. *Nature*. 2020;578(7793):94-101.
22. Rosales RA, Drummond RD, Valieris R, Dias-Neto E, da Silva IT. signeR: an empirical Bayesian approach to mutational signature discovery. *Bioinformatics*. 2017;33(1):8-16.
23. Alexandrov LB, Nik-Zainal S, Wedge DC, et al. Signatures of mutational processes in human cancer. *Nature*. 2013;500(7463):415-421.
24. Rustad EH, Nadeu F, Angelopoulos N, et al. mmsig: a fitting approach to accurately identify somatic mutational signatures in hematological malignancies. *Commun Biol*. 2021;4(1):424.
25. Maura F, Degasperi A, Nadeu F, et al. A practical guide for mutational signature analysis in hematological malignancies. *Nat Commun*. 2019;10(1):2969.
26. 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.
27. Rustad EH, Yellapantula V, Leongamornlert D, et al. Timing the initiation of multiple myeloma. *Nat Commun*. 2020;11(1):1917.
28. Nik-Zainal S, Davies H, Staaf J, et al. Landscape of somatic mutations in 560 breast cancer whole-genome sequences. *Nature*. 2016;534(7605):47-54.
29. Nilsen G, Liestol K, Van Loo P, et al. Copynumber: Efficient algorithms for single- and multi-track copy number segmentation. *BMC Genomics*. 2012;13:591.
30. Chen X, Schulz-Trieglaff O, Shaw R, et al. Manta: rapid detection of structural variants and indels for germline and cancer sequencing applications. *Bioinformatics*. 2015;32(8):1220-1222.
31. Jafari M, Ansari-Pour N. Why, When and How to Adjust Your P Values? *Cell J*. 2019;20(4):604-607.
32. Gooding S, Ansari-Pour N, Towfic F, et al. Multiple cereblon genetic changes are associated with acquired resistance to lenalidomide or pomalidomide in multiple myeloma. *Blood*. 2021;137(2):232-237.
33. Gooding S, Ansari-Pour N, Kazeroun MH, et al. Loss Of COP9-Signalosome Genes At 2q37 Is Associated With IMiD Agent Resistance In Multiple Myeloma. *Blood*. 2022.

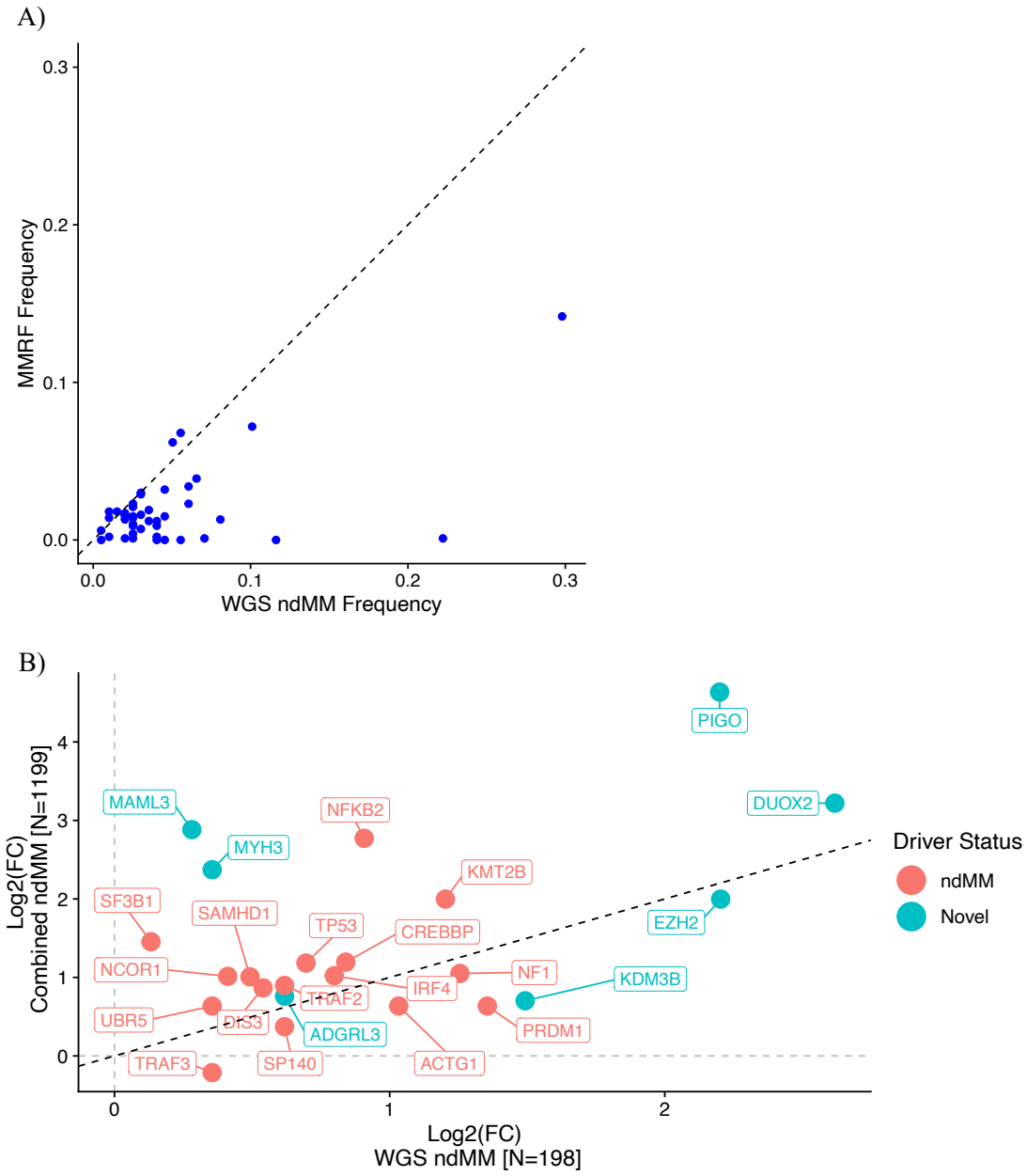
Data sharing statement

Proposals to access deidentified participant data from the rrMM cohort may be sent to the corresponding author or erin.flynt@bms.com.

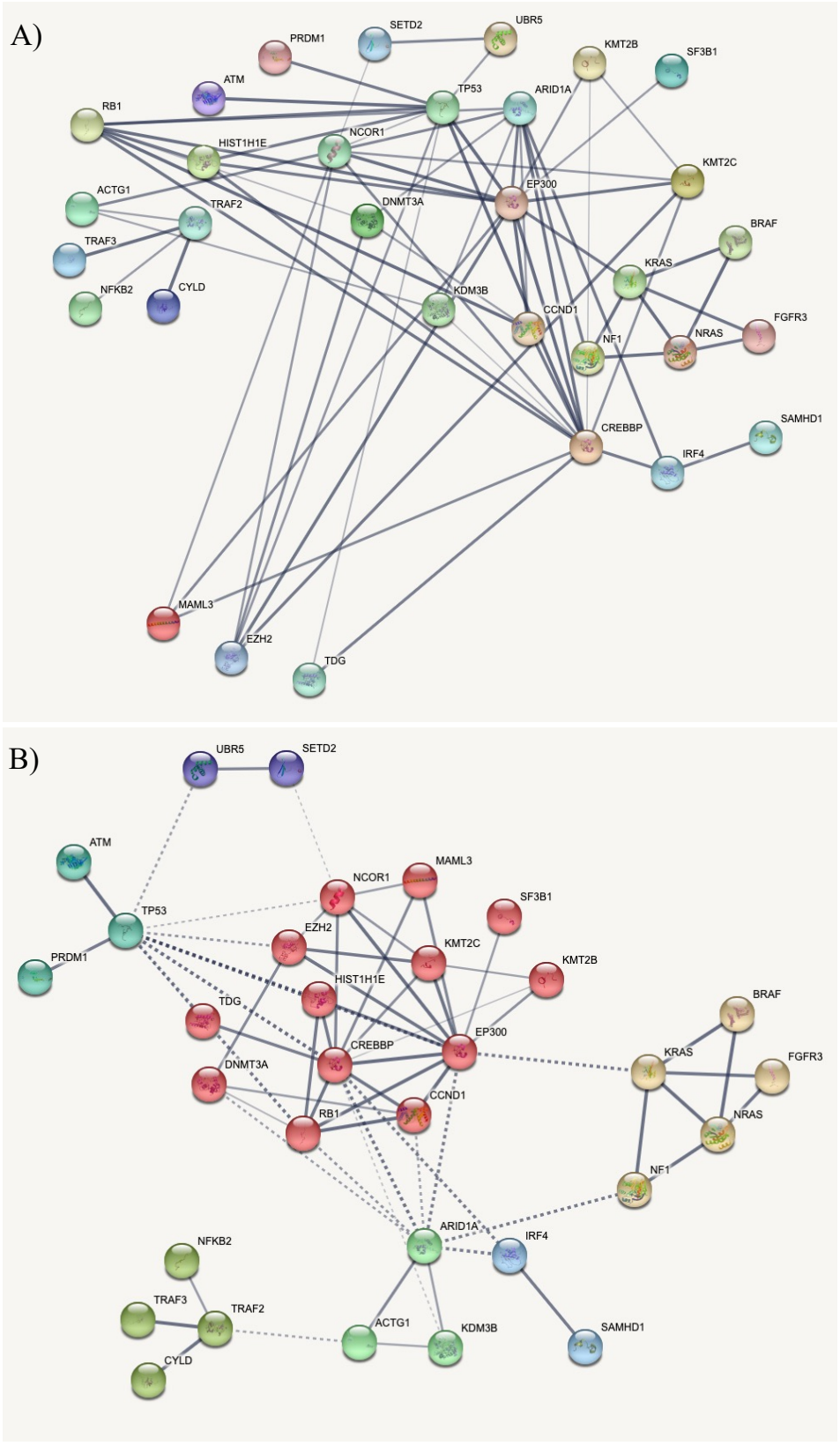
Supplemental Figures



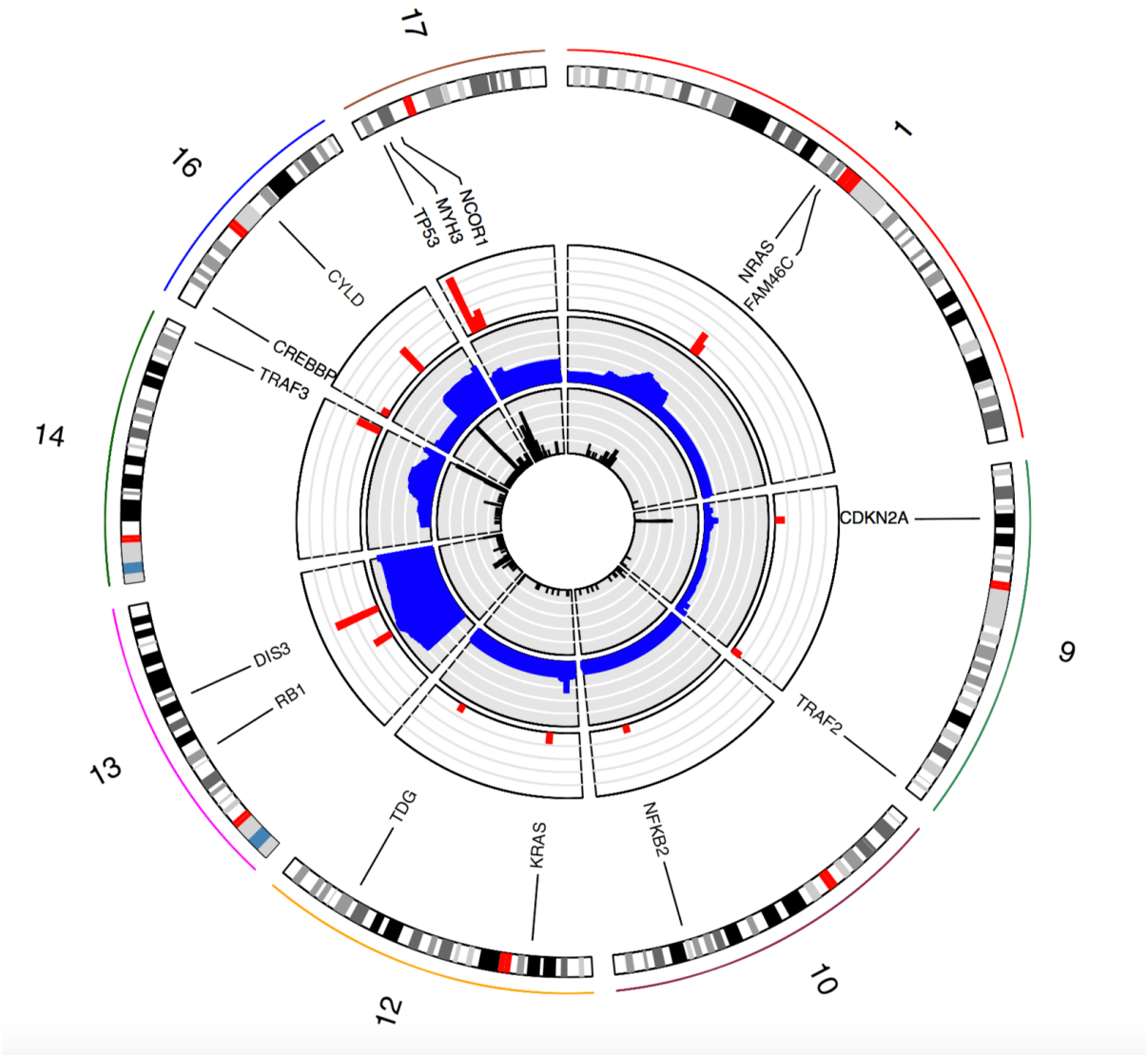
Supplemental Figure 1. Bipartite graph between tumors negative for all mutational drivers and major copy number- and translocation-based genomic aberrations. Tumors with at least one major event are shown. WGD: whole-genome duplication, HRD: hyperdiploidy.



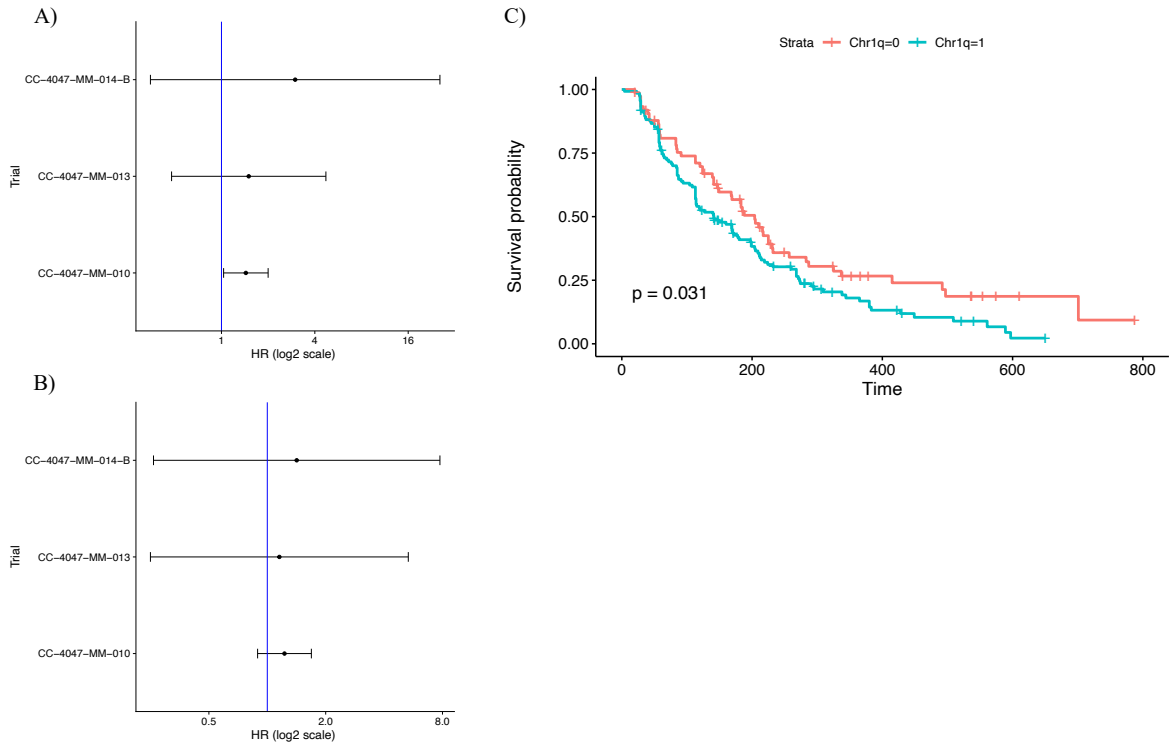
Supplemental Figure S2. MMRF CoMMpass-based driver frequency and enrichment comparison. A) Comparison of the cohort-level recurrence of mutational drivers in the WGS ndMM cohort and the WES MMRF CoMMpass cohort. **B)** Log-fold recurrence enrichment of drivers in rrMM compared between when the WGS ndMM cohort (x-axis) and the combined ndMM dataset (N=1199, y-axis) is used. Drivers with no enrichment against the WGS ndMM cohort are not shown. Genes previously detected in ndMM as drivers (Walker et al. 2018) are shown in blue and genes detected only in the rrMM cohort (labelled 'Novel') are shown in red. The dotted black line in both A) and B) represent y=x.



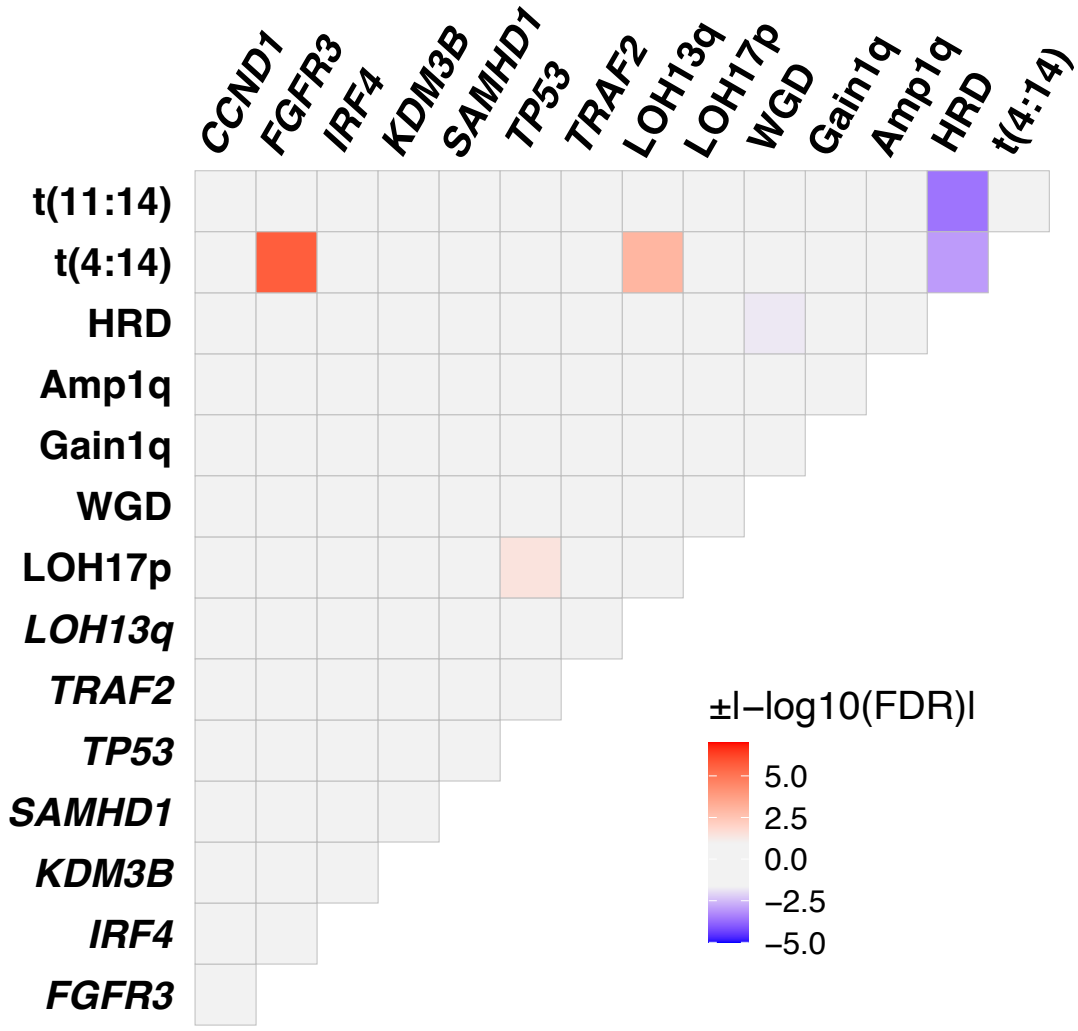
Supplemental Figure 3. Network analysis of novel promising candidate drivers in the context of known myeloma drivers. A) Functional interactions between novel and known drivers. **B)** MCL-based clustering of drivers based on the network topology.



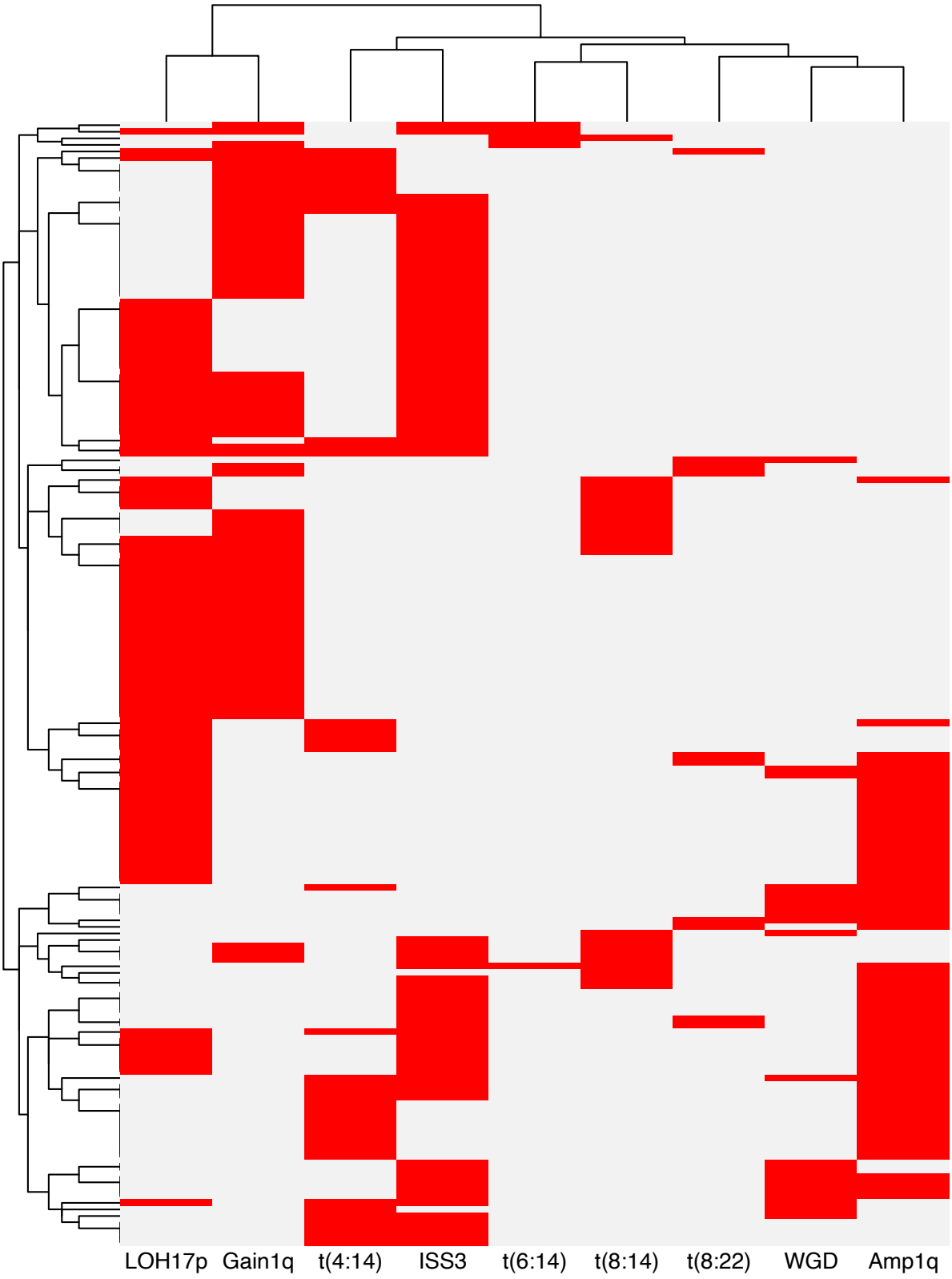
Supplemental Figure 4. Somatic association in individual tumors. Circos plot of double-hit drivers identified in individual tumors across the rrMM cohort. From inside, the first track represents the homozygous deletion (HD) frequency landscape (histogram height ranging from 0 to 2.5%), the second track represents the loss-of-heterozygosity (LOH) frequency landscape (histogram height ranging from 0 to 60%), and the third track represents the overall frequency of double-hit driver events through a combination of mutation and LOH and/or HD (histogram height ranging from 0 to 10%). Chromosome ideograms are represented and labelled on the outside.



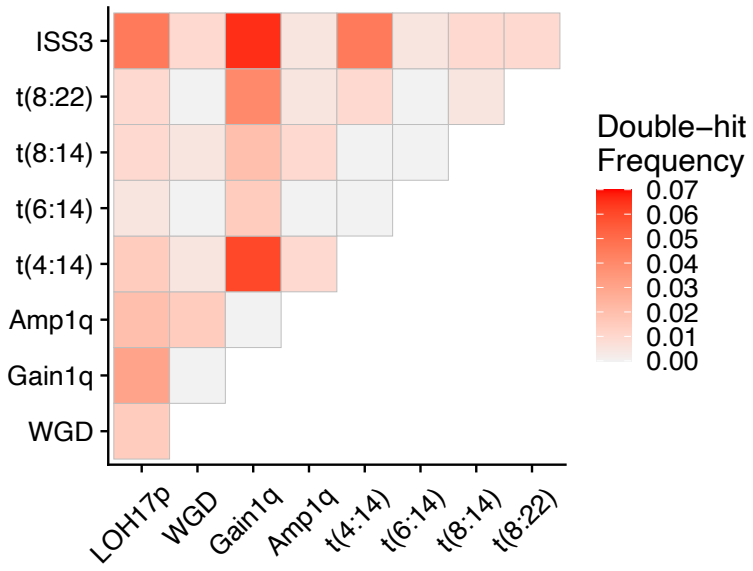
Supplemental Figure 5. Survival analysis of common rrMM-enriched CNA drivers with progression free survival (PFS). A) Cox regression analysis of 1q Gain in three trials with PFS data shows prognostic effect in the MM010 trial cohort. **B)** Cox regression analysis of 17p LOH in three trials with no significant association. **C)** Kaplan-Meier curve showing 1q Gain correlates with poor prognosis in the MM010 trial cohort.



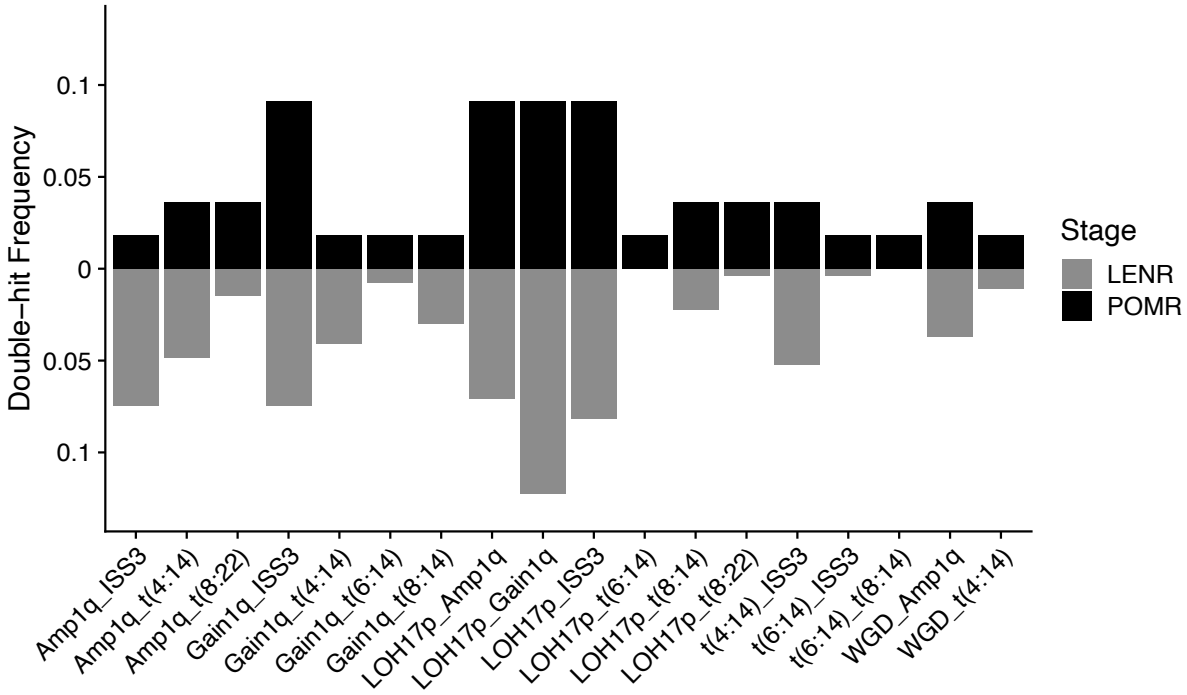
Supplemental Figure 6. Somatic interactions of genomic aberrations in the ndMM dataset. Genomic events with at least one significant pairwise association (FDR<0.1) in the rrMM dataset are shown.



Supplemental Figure 7. Double-hit of high risk events observed in the rrMM dataset. Each row represents an rrMM patient with at least two high risk events (N=171; 44.3%).

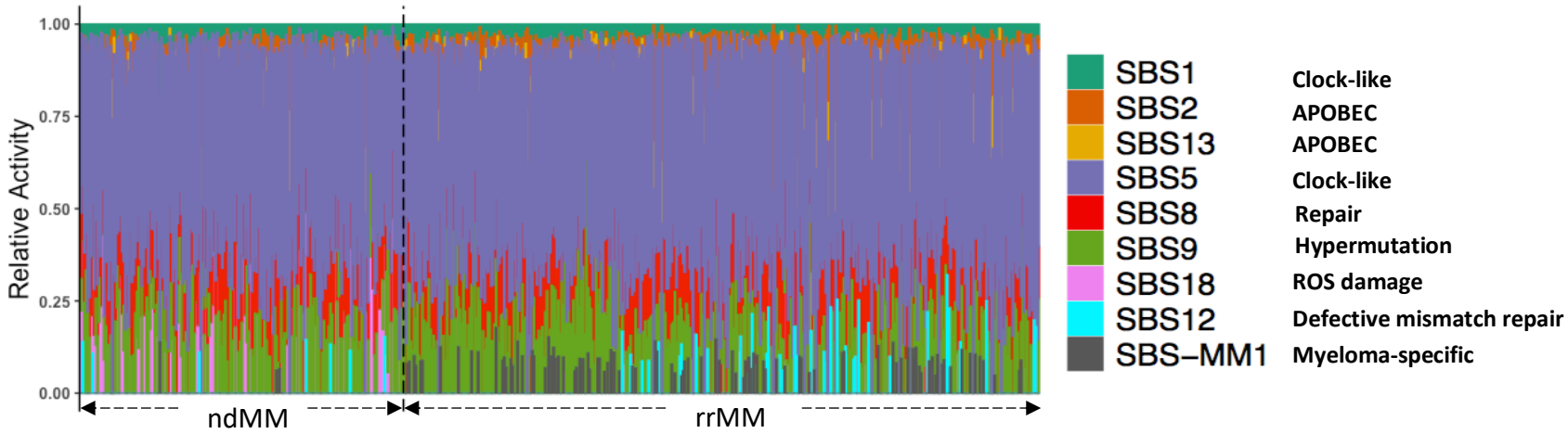


Supplemental Figure 8. Frequency of double-hit events of high-risk features observed in the ndMM dataset.

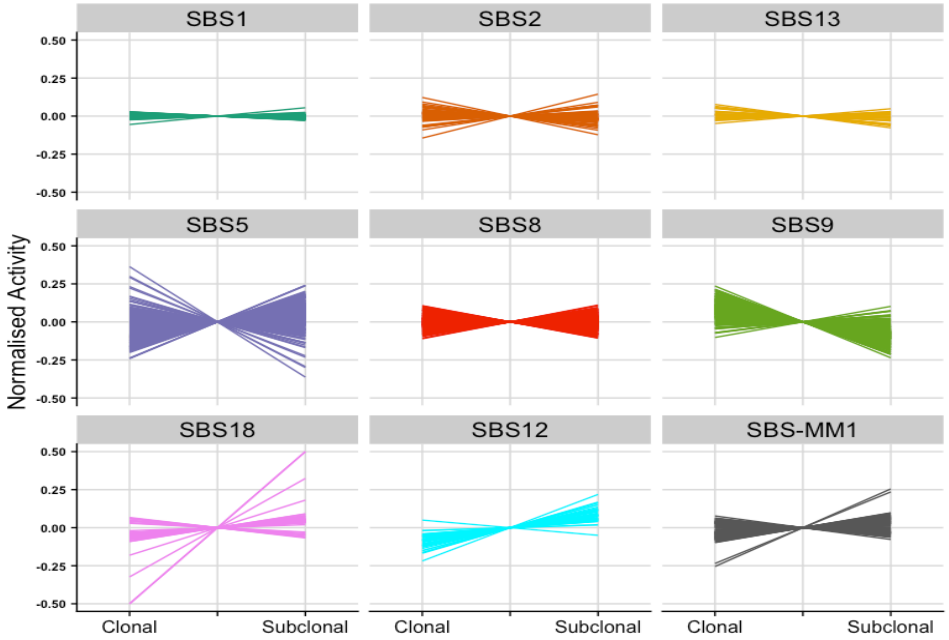


Supplemental Figure 9. Enrichment analysis of double-hit events in POMR compared with LENR. Frequency of each event was compared between POMR and LENR sub-groups using Fisher's exact test. None were significant after multiple-testing correction (FDR<0.05).

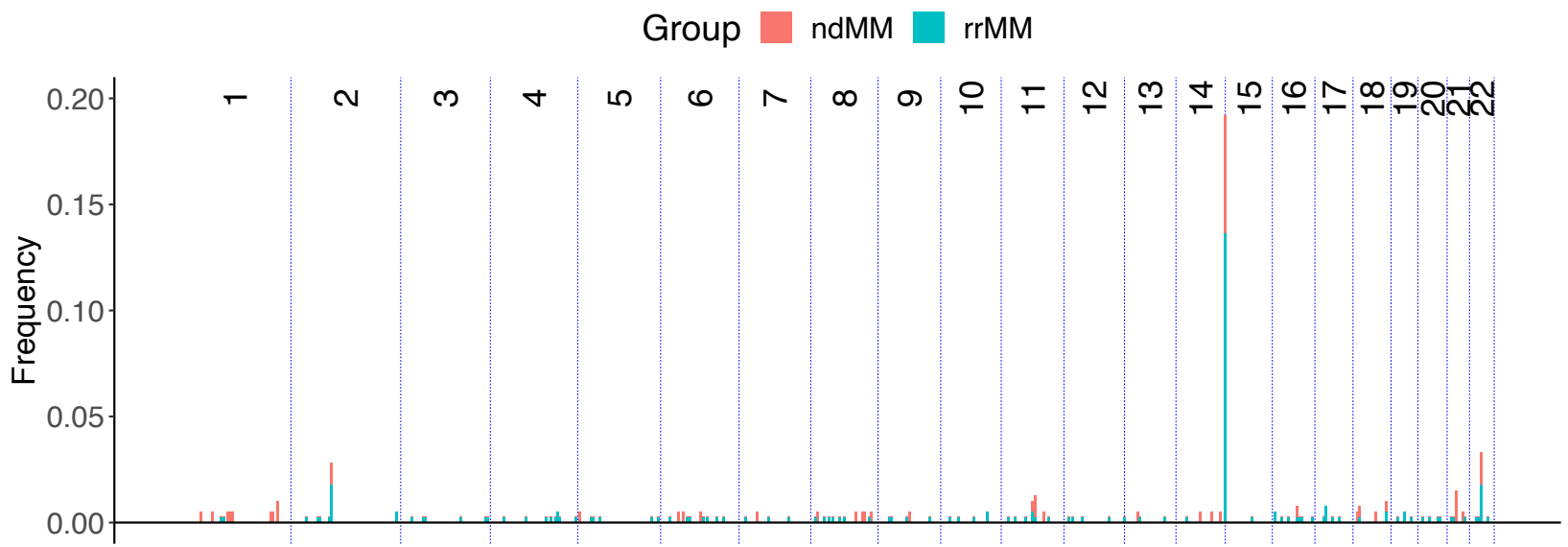
A



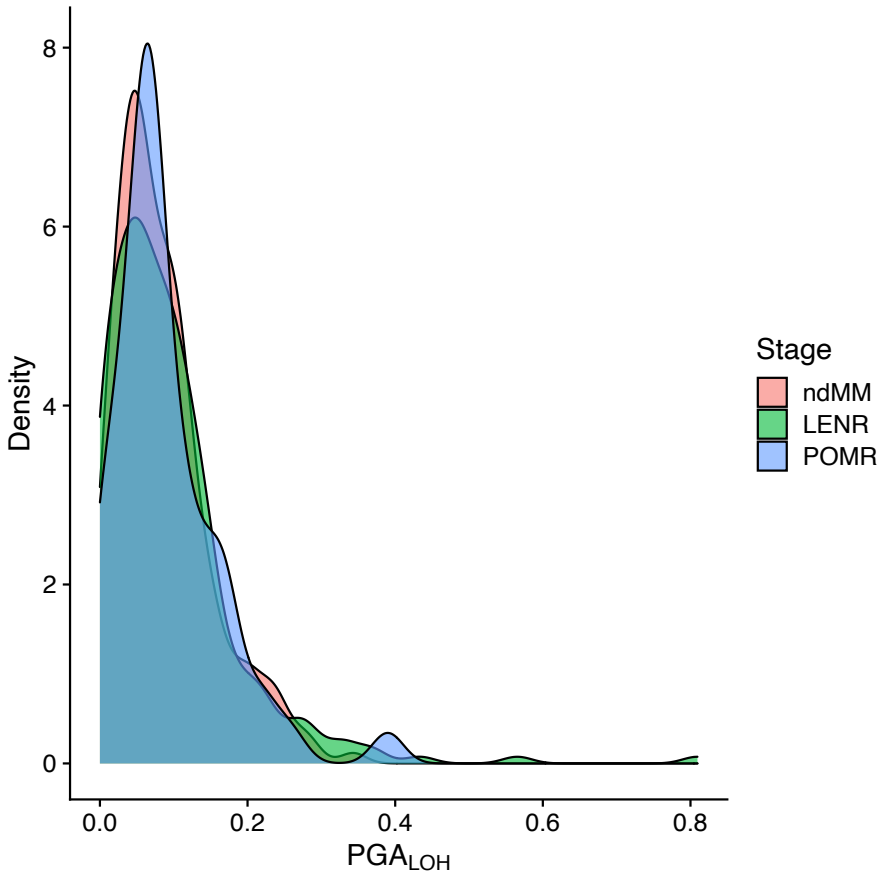
B



Supplemental Figure 10. Differential analysis of genome-wide mutational signatures. A) Frequency landscape of single base substitution (SBS) signatures in both ndMM and rrMM cohorts. **B)** Shift in mutational signature activity between clonal and subclonal mutations in individual samples of the rrMM cohort. Each line represents a single tumor where the activity of a given signature increases or decreases at the subclonal stage when compared with the clonal activity level. SBS12 shows a strong signal of subclonal increase while SBS9 shows a decrease across the majority of tumors.



Supplemental Figure 11. Genome-wide landscape of kataegis events in the rrMM and ndMM datasets. The most frequent events are present at the immunoglobulin loci (IgK: 2q, IgH: 14q and IgL: 22q).



Supplemental Figure 12. Distribution of proportion of genome altered due to loss-of-heterozygosity (PGA_{LOH}) at different therapy stages. ndMM: newly diagnosed multiple myeloma, LENR: lenalidomide resistant, POMR: pomalidomide resistant.

	ndMM	rrMM													
		TOTAL						LENR					POMR		
Cohort	**IFM-2009**	**CC-122-ST-MM**	**CC-220-MM-001**	**CC-4047-MM-007**	**CC-4047-MM-010**	**CC-4047-MM-013**	**CC-4047-MM-014-B**	**CC-122-ST-MM**	**CC-220-MM-001**	**CC-4047-MM-007**	**CC-4047-MM-010**	**CC-4047-MM-014-B**	**CC-122-ST-MM**	**CC-220-MM-001**	**CC-4047-MM-010**
Number of patients (Number of samples)	198 (198)	11 (11)	38 (51)	64 (64)	223 (242)	18 (18)	32 (32)	3 (3)	14 (17)	34 (34)	188 (189)	30 (30)	7 (7)	23 (33)	25 (43)
Mean study entry age (95% CI)	57.2 (56.3 - 58.2)	70.5 (67.4 - 73.6)	63.2 (60.5 - 65.9)	68.5 (66.8 - 70.2)	66.3 (65.3 - 67.3)	71.4 (67.6 - 75.2)	66.6 (64.2 - 69)	65.7 (62.4 - 69)	65.4 (61.2 - 69.5)	68.6 (66.4 - 70.7)	66.8 (65.7 - 67.9)	66.3 (63.8 - 68.8)	71.7 (67.6 - 75.8)	62 (58.3 - 65.7)	62.6 (60.1 - 65.1)
Mean time since diagnosis in years (95% CI)	-	6.2 (4.4 - 8.1)	7.1 (6.3 - 7.9)	6.0 (5.2 - 6.9)	5.9 (5.6 - 6.3)	5.4 (3.8 - 7.0)	4.0 (3.2 - 4.9)	8.3 (2.7 - 13.8)	7.6 (6.2 - 9.1)	5.9 (4.7 - 7.0)	5.8 (5.4 - 6.2)	4.0 (3.2 - 4.9)	6.1 (5.2 - 7.0)	6.7 (5.7 - 7.7)	7.1 (6.0 - 8.3)
Sex															
Male / total patients with data (%)	93/149 (62.4)	6/11 (54.5)	21/38 (55.3)	33/64 (51.6)	122/223 (54.7)	9/18 (50)	22/32 (68.8)	0/3 (0)	6/14 (42.9)	18/34 (52.9)	103/188 (54.8)	20/30 (66.7)	5/7 (71.4)	14/23 (60.9)	12/25 (48)
Female / total patients with data (%)	56/149 (37.6)	5/11 (45.5)	17/38 (44.7)	31/64 (48.4)	101/223 (45.3)	9/18 (50)	10/32 (31.2)	3/3 (100)	8/14 (57.1)	16/34 (47.1)	85/188 (45.2)	10/30 (33.3)	2/7 (28.6)	9/23 (39.1)	13/25 (52)
ISS															
I / total patients that had data(%)	55/156 (35.3)	0/11 (0)	13/38 (34.2)	20/48 (41.7)	59/174 (33.9)	0/18 (0)	5/26 (19.2)	0/3 (0)	7/14 (50)	11/34 (32.4)	54/150 (36)	5/25 (20)	0/7 (0)	6/23 (26.1)	3/16 (18.8)
II / total patients that had data(%)	66/156 (42.3)	4/11 (36.4)	17/38 (44.7)	17/48 (35.4)	58/174 (33.3)	3/18 (16.7)	14/26 (53.8)	1/3 (33.3)	4/14 (28.6)	14/34 (41.2)	49/150 (32.7)	14/25 (56)	3/7 (42.9)	12/23 (52.2)	9/16 (56.2)
III / total patients that had data(%)	35/156 (22.4)	7/11 (63.6)	8/38 (21.1)	11/48 (22.9)	57/174 (32.8)	15/18 (83.3)	7/26 (26.9)	2/3 (66.7)	3/14 (21.4)	9/34 (26.5)	47/150 (31.3)	6/25 (24)	4/7 (57.1)	5/23 (21.7)	4/16 (25)
Number of patients that received stem cell Transplants / total patients that had data (%)	87/187 (46.5)	8/11 (72.7)	32/38 (84.2)	29/48 (60.4)	143/223 (64.1)	6/18 (33.3)	14/14 (100)	3/3 (100)	12/14 (85.7)	18/34 (52.9)	118/188 (62.8)	14/14 (100)	5/7 (71.4)	19/23 (82.6)	19/25 (76)
BORT															
Exposed / total patients that had data (%)	-	10/10 (100)	38/38 (100)	60/60 (100)	212/212 (100)	-	-	3/3 (100)	14/14 (100)	34/34 (100)	188/188 (100)	-	7/7 (100)	23/23 (100)	24/24 (100)
Refractory / total patients that had data (%	-	-	14/38 (36.8)	5/60 (8.3)	212/212 (100)	-	-	-	4/14 (28.6)	3/34 (8.8)	188/188 (100)	-	-	9/23 (39.1)	24/24 (100)
DAR															
Exposed / total patients that had data (%)	-	-	38/38 (100)	-	-	-	-	-	14/14 (100)	-	-	-	-	23/23 (100)	-
Refractory / total patients that had data (%	-	-	21/38 (55.3)	-	-	-	-	-	4/14 (28.6)	-	-	-	-	17/23 (73.9)	-
Translocation															
t(4;14) / total patients that had data (%)	24/198 (12.1)	2/11 (18.2)	2/38 (5.3)	8/64 (12.5)	29/223 (13)	1/18 (5.6)	3/32 (9.4)	1/3 (33.3)	0/14 (0)	7/34 (20.6)	27/188 (14.4)	3/30 (10)	1/7 (14.3)	2/23 (8.7)	2/25 (8)
t(8;14) / total patients that had data(%)	12/198 (6.1)	2/11 (18.2)	3/38 (7.9%)	3/64 (4.7)	14/223 (6.3)	1/18 (5.6)	2/32 (6.3)	0/3 (0)	1/14 (7.1)	2/34 (5.9)	13/188 (6.9)	2/30 (6.7)	1/7 (14.3)	2/23 (8.7)	0/25 (0)
t(11;14) / total patients that had data(%)	41/198 (20.7)	3/11 (27.3)	5/38 (13.2)	12/64 (18.8)	47/223 (21.1)	6/18 (33.3)	7/32 (21.9)	0/3 (0)	0/14 (0)	6/34 (17.6)	39/188 (20.7)	7/30 (23.3)	2/7 (28.6)	5/23 (21.7)	7/25 (28)
t(14;16) / total patients that had data(%)	5/198 (2.5)	0/11 (0)	0/38 (0)	0/64 (0)	5/223 (2.2)	0/18 (0)	1/32 (3.1)	0/3 (0)	0/14 (0)	0/34 (0)	4/188 (2.1)	1/30 (3.3)	0/7 (0)	0/23 (0)	1/25 (4)
t(14;20) / total patients that had data(%)	0/198 (0)	0/11 (0)	0/38 (0)	0/64 (0)	0/223 (0)	0/18 (0)	0/32 (0)	0/3 (0)	0/14 (0)	0/34 (0)	0/188 (0)	0/30 (0)	0/7 (0)	0/23 (0)	0/25 (0)
t(6;14) / total patients that had data(%)	5/198 (2.5)	0/11 (0)	3/38 (7.9%)	1/64 (1.6)	2/223 (0.9)	0/18 (0)	0/32 (0)	0/3 (0)	1/14 (7.1)	0/34 (0)	1/188 (0.5)	0/30 (0)	0/7 (0)	2/23 (8.7)	0/25 (0)
t(8;22) / total patients that had data (%)	11/198 (5.5)	0/11 (0)	1/38 (2.6)	1/64 (1.6)	9/223 (4)	0/18 (0)	1/32 (3.1)	0/3 (0)	1/14 (7.1)	0/34 (0)	7/188 (3.7)	1/30 (3.3)	0/7 (0)	0/23 (0)	2/25 (8)

Supplemental Table S1. The demographics of the newly diagnosed and relapse/refractory patients per clinical trial.

Supplemental Table S2. Details of coding variants in the 44 driver genes identified in the rrMM cohort.
[this is provided as a separate file due to size]

ID	Term	P.value	Adjusted.P.value	Odds.Ratio	Combined.Score	Genes
BioPlanet_2019	Cleavage of a damaged pyrimidine	0.002497708	0.041251454	555.1666667	3326.770642	TDG
BioPlanet_2019	Resolution of AP sites via the single-nucleotide replacement pathway	0.005985094	0.041251454	201.8080808	1032.951295	TDG
BioPlanet_2019	NICD trafficking to the nucleus	0.006482395	0.041251454	184.9814815	932.0597464	MAML3
BioPlanet_2019	PRC2 complex long-term gene silencing through modification of histone tails	0.007476328	0.041251454	158.5396825	776.2124253	EZH2
BioPlanet_2019	Glycosylphosphatidylinositol (GPI) biosynthesis	0.008469366	0.041251454	138.7083333	661.8190143	PIGO
BioPlanet_2019	Resolution of abasic sites (AP sites)	0.009461511	0.041251454	123.2839506	574.5677164	TDG
BioPlanet_2019	Cell differentiation pathway	0.009957248	0.041251454	116.7894737	538.3357754	EZH2
BioPlanet_2019	Post-translational modification: biosynthesis of GPI-anchored proteins	0.014408847	0.045535866	79.21428571	335.8616723	PIGO
BioPlanet_2019	MicroRNAs in muscle cell differentiation	0.014902355	0.045535866	76.4789272	321.6884166	EZH2
BioPlanet_2019	Fluoropyrimidine activity	0.016381548	0.045535866	69.29861111	284.9281499	TDG
BioPlanet_2019	Base excision repair	0.017366565	0.045535866	65.21568627	264.3327714	TDG
BioPlanet_2019	Striated muscle contraction	0.018842427	0.045535866	59.91891892	237.9766259	MYH3
BioPlanet_2019	Regulation of transcription by NOTCH1 intracellular domain	0.023258053	0.050204947	48.17391304	181.1870902	MAML3
BioPlanet_2019	Muscle contraction	0.024236871	0.050204947	46.16203704	171.7172484	MYH3
BioPlanet_2019	Pre-NOTCH expression and processing	0.028143323	0.054410424	39.5515873	141.2167736	MAML3
BioPlanet_2019	Viral myocarditis	0.03494576	0.061265111	31.61904762	106.0489619	MYH3
BioPlanet_2019	Signaling by NOTCH1	0.035914031	0.061265111	30.73765432	102.2527178	MAML3
BioPlanet_2019	NOD signaling pathway	0.041705318	0.067191901	26.33068783	83.65592939	DUOX2
BioPlanet_2019	DNA repair	0.051287922	0.078281565	21.2457265	63.1061812	TDG
BioPlanet_2019	Signaling by NOTCH	0.057944341	0.081326797	18.71186441	53.29648635	MAML3
BioPlanet_2019	Notch signaling pathway	0.058891819	0.081326797	18.39814815	52.10453246	MAML3
BioPlanet_2019	Tight junction	0.064558698	0.085100102	16.71548822	45.8034535	MYH3
KEGG_2021_Human	Glycosylphosphatidylinositol (GPI)-anchor biosynthesis	0.012926986	0.070822859	88.73333333	385.8514207	PIGO
KEGG_2021_Human	Base excision repair	0.016381548	0.070822859	69.29861111	284.9281499	TDG
KEGG_2021_Human	Notch signaling pathway	0.029117734	0.070822859	38.18390805	135.0338735	MAML3
KEGG_2021_Human	Lysine degradation	0.031063919	0.070822859	35.71326165	123.9860269	EZH2
KEGG_2021_Human	Thyroid hormone synthesis	0.036881427	0.070822859	29.9039039	98.68429377	DUOX2
KEGG_2021_Human	B cell receptor signaling pathway	0.039778377	0.070822859	27.65277778	89.16449602	LILRA6
KEGG_2021_Human	Th1 and Th2 cell differentiation	0.045069092	0.070822859	24.2967033	75.30905521	MAML3
KEGG_2021_Human	Osteoclast differentiation	0.061729109	0.084877525	17.51675485	48.78415662	LILRA6
WikiPathway_2021_Human	Thyroid hormones production and their peripheral downstream signaling effects	0.000980033	0.024500814	53.48655914	370.550857	KDM3B;DUOX2
WikiPathway_2021_Human	Cytosine methylation	0.004491844	0.042325613	277.5277778	1500.174168	TDG
WikiPathway_2021_Human	Interactome of polycomb repressive complex 2 (PRC2)	0.007972959	0.042325613	147.962963	714.9125872	EZH2
WikiPathway_2021_Human	Cell Differentiation - Index expanded	0.009461511	0.042325613	123.2839506	574.5677164	EZH2
WikiPathway_2021_Human	GPCRs, Class B Secretin-like	0.011937965	0.042325613	96.4589372	427.1232267	ADGRL3
WikiPathway_2021_Human	Canonical and non-canonical Notch signaling	0.013421162	0.042325613	85.31623932	367.791702	MAML3
WikiPathway_2021_Human	Tumor suppressor activity of SMARCB1	0.015395642	0.042325613	73.92592593	308.5424793	EZH2
WikiPathway_2021_Human	Base Excision Repair	0.015395642	0.042325613	73.92592593	308.5424793	TDG
WikiPathway_2021_Human	Fluoropyrimidine Activity	0.016381548	0.042325613	69.29861111	284.9281499	TDG
WikiPathway_2021_Human	FBXL10 enhancement of MAP/ERK signaling in diffuse large B-cell lymphoma	0.017366565	0.042325613	65.21568627	264.3327714	EZH2
WikiPathway_2021_Human	Striated Muscle Contraction Pathway	0.018842427	0.042325613	59.91891892	237.9766259	MYH3
WikiPathway_2021_Human	Nucleotide-binding Oligomerization Domain (NOD) pathway	0.020316294	0.042325613	55.41666667	215.9217342	DUOX2
WikiPathway_2021_Human	Notch Signaling	0.022278351	0.042842983	50.36868687	191.6095305	MAML3
WikiPathway_2021_Human	Notch Signaling Pathway Netpath	0.030091266	0.051572801	36.90740741	129.3058522	MAML3
WikiPathway_2021_Human	Pathways affected in adenoid cystic carcinoma	0.032035694	0.051572801	34.59375	119.0337919	MAML3
WikiPathway_2021_Human	Histone Modifications	0.033006592	0.051572801	33.54208754	114.4136696	EZH2
WikiPathway_2021_Human	MECP2 and Associated Rett Syndrome	0.035430005	0.052102948	31.17214397	104.1210777	EZH2
WikiPathway_2021_Human	ncRNAs involved in Wnt signaling in hepatocellular carcinoma	0.042186509	0.057535271	26.01960784	82.36909635	EZH2
WikiPathway_2021_Human	GPCRs, Other	0.044589205	0.057535271	24.56790123	76.41264641	ADGRL3
WikiPathway_2021_Human	LncRNA involvement in canonical Wnt signaling and colorectal cancer	0.046028217	0.057535271	23.77180406	73.18151447	EZH2
WikiPathway_2021_Human	Thermogenesis	0.05271785	0.062759346	20.64693666	60.75982922	KDM3B
WikiPathway_2021_Human	DNA Repair Pathways Full Network	0.058418187	0.066384304	18.55368814	52.69484947	TDG
WikiPathway_2021_Human	Endoderm differentiation	0.068319537	0.074260367	15.75396825	42.27671122	EZH2
WikiPathway_2021_Human	Epithelial to mesenchymal transition in colorectal cancer	0.077197051	0.080413595	13.85814116	35.49615996	EZH2
WikiPathway_2021_Human	Circadian rhythm related genes	0.09609492	0.09609492	10.99444444	25.75359368	EZH2

Supplemental Table 3. Enriched pathways based on the ten novel rrMM drivers.