Genomic and transcriptional profiling stratifies VQ myeloma lines into two clusters with distinct risk signatures and drug responses

Evan Flietner[1,2]*, Mei Yu[3]*, Adhithi Rajagopalan[1], Yun Zhou[1], Yubin Feng[1], Anthony J. Veltri[3], Terra Lasho[4], Zhi Wen[6], Yuqian Sun[1], Mrinal M. Patnaik[4], Natalie S. Callander[5], Fotis Asimakopoulos[7], Demin Wang[3#], Jing Zhang[1#]

[1] McArdle Laboratory for Cancer Research, University of Wisconsin-Madison, Madison, WI 53705, USA
[2] Department of Pathology and Laboratory Medicine, University of Wisconsin School of Medicine and Public Health, University of Wisconsin-Madison, Madison, WI 53705, USA
[3] Versiti Blood Research Institute, Milwaukee, WI 53226
[4] Division of Hematology, Department of Internal Medicine, Mayo Clinic, Rochester, MN 55902
[5] Division of Hematology/Oncology, Department of Medicine, UW Comprehensive Cancer Center, University of Wisconsin-Madison, Madison, WI 53705, USA
[6] Center for Precision Medicine Research, Marshfield Clinic Research Institute, Marshfield, WI 54449
[7] Division of Blood and Marrow Transplantation, Department of Medicine and Moores Cancer Center, University of California-San Diego, La Jolla, CA 92093

*These authors contributed equally to this work.

Running Title: VQ myeloma clusters with distinct gene signature and drug response
Key words: multiple myeloma, VQ model, high-risk multiple myeloma, genomic and transcriptional profiling, drug-resistance

#Corresponding authors:
Demin Wang
Address: Blood Research Institute
              8727 Watertown Plank Rd
              Versiti
              Milwaukee, WI 53226
Telephone: (414) 937-3874
e-mail: dwang@versiti.org

Jing Zhang
Address: Wisconsin Institute for Medical Research II
              McArdle Lab for Cancer Research
              1111 Highland Avenue
              University of Wisconsin-Madison
              Madison, WI 53705
Telephone: (608)263-1147
Fax: (608)262-2824
e-mail: zhang@oncology.wisc.edu

Conflict of interest:
We declare that no conflict of interest exists.

Abstract

Multiple myeloma (MM) is a cancer of malignant plasma cells in the bone marrow and extramedullary sites. We previously characterized a VQ model for human high-risk MM. Different VQ lines display distinct disease phenotypes and survivals, suggesting significant intra-model variation. Here, we use whole exome sequencing and copy number variation (CNV) analysis coupled with RNA-Seq to stratify VQ lines into corresponding clusters: Cluster I VQ cells carried recurrent amplification of chromosome (chr) 3 and displayed upregulation of growth pathways and high-risk myeloma gene signatures, whereas Cluster II cells had monosomy chr5 and overexpressed genes and pathways associated with positive response to bortezomib (Btz) treatment in human MM patients. Consistently, in sharp contrast to Cluster II VQ cells that showed short-term response to Btz, Cluster I VQ cells were de novo resistant to Btz *in vivo*. Our study highlights Cluster I VQ lines as highly representative of human high-risk MM subset.

Introduction

Multiple myeloma (MM) is a malignancy of terminally differentiated plasma cells (PCs) that primarily grow in the bone marrow (BM) [1]. MM arises from the pre-malignant condition monoclonal gammopathy of undetermined significance (MGUS), in which the accumulation of chromosomal copy number variations (CNVs), primary translocations, and somatic mutations leads to the expansion of a clonal PC population [2]. MGUS and MM share a number of overlapping CNVs, including hyperdiploidy [3], as well as primary translocations such as t(4;14) [4] and t(11;14) [5]. The genetic heterogeneity of MM is further increased by acquiring secondary CNVs [6] and point mutations [7] and a greatly altered landscape of DNA methylation compared to healthy PCs [8].

Genetic events (i.e. CNVs and primary translocations) present at time of diagnosis play a significant role in patient prognosis [9] and in the stratification of high-risk multiple myeloma (hrMM) [10,11]. These events include two translocations involving the immunoglobulin heavy chain (IgH) locus: the t(4;14) translocation in which both fibroblast growth factor receptor 3 (FGFR3) and multiple myeloma SET domain (MMSET) are put under the control of the IgH promoter [12,13], and t(14;16) in which the transcription factor c-MAF is overexpressed [14]. Amplification (≥4 copies) of the long arm of chromosome 1 (amp(1q)) is an hrMM prognostic marker [15], while a gain of a single copy (gain(1q)) is considered high risk when combined with a second hrMM chromosomal abnormality [16]. At the level of gene expression, two independent, and largely non-overlapping gene signatures predicting increased relapse risk and poorer overall survival have been developed for use in diagnosing hrMM: a 70 gene signature developed by the University of Arkansas for Medical Sciences (UAMS-70) [17], and a 92 gene signature developed by the Erasmus University Medical Center (EMC-92 or SKY92) [18]. Diagnosis of patients with hrMM is particularly important as these patients are likely to have a poor response to current treatment options [19,20].

Murine models of MM play an essential role in dissecting mechanisms of disease growth [21] and as pre-clinical platforms for testing new anti-MM therapies [22]. The genetic basis for MM models ranges from the forced overexpression of a single MM-related oncogene such as the Eµ-MAF model [23], transgenic models in which oncogene overexpression is activated through plasma cell maturation as in the V_K*MYC model [24], or spontaneous MM development owing to genetic inbreeding as in the 5TMM family of myeloma models [25,26]. Due to the heterogeneous genetic origins of human MM and the downstream pathological and therapeutic consequences, genomic characterization of these murine myeloma cells is necessary to establish what sub-type of MM they best model. Genetic landscape characterization of the 5TMM models has shown notable genomic differences between different lines, with the 5T2 line showing CNVs syntenic for gain(1q) MM whereas the 5T33 and 5TGM1 lines show CNVs syntenic for del(13q) [27]. Within transgenic MM models some chromosomal variability is also observed: CNV analysis of both primary and transplantable V_K*MYC showed that about 50% of sequenced lines had monosomy 5, and within these mice half additionally had monosomy 14, including a region syntenic to del(13) in human MM [28]. Further analysis of monosomy 14 V_K*MYC lines identified loss of the region containing the cell-cycle regulating miRNA cluster MIR15A/16-1 as driving MM progression in the V_K*MYC model and as a potential mechanism to explain del(13) as an early initiating CNV in the transition from MGUS to MM [29].

Previously, we developed and characterized a mouse model of MM driven by both the V_K*MYC transgene and oncogenic NRas$^{LSL-Q61R}$ inducible via IgG1-Cre [30]. This so-called VQ model shows characteristics representative of hrMM, as it is highly proliferative, develops extramedullary disease, and is enriched for the UAMS-70 high-risk gene signature compared to control PCs [30]. In our initial study, five lines of VQ myeloma were isolated from primary mice (VQ-D1 through VQ-D5). In four of these VQ lines (VQ-D1, D2, D4, and D5), we established

survival of recipient mice transplanted with donor VQ cells and carried out bulk RNA-Seq analysis. Finally, we established two VQ cell lines from two VQ-D2 recipients, termed VQ 4935 and 4938. Further drug testing of the VQ model also showed *de novo* resistance to the BCL-2 inhibitor venetoclax in the VQ cell lines, as well as reduced responses to the proteasome-inhibitors (PIs) bortezomib and carfilzomib in VQ-D1 recipient mice *in vivo* [31]. Although all VQ cells are driven by MYC dysregulation and RAS hyperactivation, different lines display distinct patterns of extramedullary disease, antibody sub-type, and survival (Figure 1A). The genetic and/or transcriptional bases for these phenotypic differences remain elusive. Additionally, what subtypes of hrMM they may individually represent (if any) are unknown. In the current study, we aim to address these questions through the molecular characterization of five VQ lines using B cell receptor (BCR) repertoire sequencing, whole exome sequencing, CNV analysis, and RNA-Seq.

Materials and Methods

High-throughput sequencing of B cell receptor repertoire

We performed paired-end full-length high-throughput sequencing of the immunoglobulin heavey and light (IgH and IgL) chains using a unique molecular identifier (UMI)-based protocol detailed by Turchaninova et al [32]. In brief, we extracted mRNA, synthesized cDNA, and incorporated UMIs and a universal primer at the 5' end with a template switch reaction. To prepare library, cDNA equivalent to 5×10^4 B cells for each sample was amplified with two rounds of PCR. For immunoglobulin heavy chain amplification, the first PCR used a 5' end universal primer and pooled 3' end primers specific to mouse IgG or IgA constant region; for immunoglobulin light chain amplification, pooled 3' end primers specific to mouse Igκ and Igλ were used. The second PCR used a 5' and a 3' universal primer with integrated sample barcoding at both ends. The Illumina adaptors were ligated to each sample library to allow for multiple sample sequencing simultaneously. We prepared barcoded libraries for each sample and then pooled the libraries

to perform asymmetric 400 + 100 bp paired-end IgSeq in one run using Miseq Reagent Kit V3 (Illumina, Miseq Reagent Kit V3, 600-cycles) on Illumina MiSeq platform.

B cell receptor repertoire analysis

Sequences were firstly demultiplexed and unique molecular identifier (UMI) were extracted by MIGEC v1.2.9 as previously described [32]. Briefly, we demultiplexed fastq files using sample barcodes and estimated the MIG (molecular identifier groups, a set of reads tagged with the same UMI) size distribution. Sequences with the same UMI in the presence of at least 2 times were then assembled and paired sequences were merged from 5' end to 3' end reads. We then mapped MIG consensus sequences to mouse IgH or IgL germline references and assembled clonotypes using MiXCR v3.0.3 [33]. Somatic hypermutation frequencies were calculated using the SHazaM package v1.1.0 [34] and plotted using ggplot2 v3.3.5 in R v4.0.3.

Whole exome sequencing

Genomic DNA was extracted from CD138+ BM cells and tail tissues of moribund V_κ*Myc; $Nras^{LSL-Q61R/+;}$ IgG1-Cre mice using Gentra Puregene Cell Kit (Qiagen). Whole-exome targeted capture was carried out using the SureSelect XT Mouse Exome Kit, 49.6 Mb (Cat# 5190-4641; Agilent Technologies). Exome capture, exome library amplification, and data analysis was carried out as described previously [35].

Copy number variation analysis

Whole exome sequencing data was aligned to the mm10 reference genome using BWA MEM [36]. Copy number variation analysis was carried out using matched controls for each sample with the CNVkit Python library and command-line software toolkit to infer and visualize copy number from high-throughput DNA sequencing data as described [37]. Copy number calling

6

was performed with default parameters, with the exception of increasing stringency through the use of `-m clonal -purity 1` arguments.

RNA-Seq analysis

Demultiplexed RNA-Seq reads were aligned to Mus musculus mm10 reference genome and expression counts were generated using STAR pipeline on Basepair tech (https://www.basepairtech.com). Gene counts were then normalized, filtered and potential batch effect from tissue was removed using limma v3.46.0 [38] and edgeR v3.32.1 (39). TSNE plots were generated using Rtsne v0.15. Differential gene expression among groups was determined using DESeq2 v1.30.1 [40]. For gene set enrichment analyses (GSEA), pre-ranked gene lists were generated based on shrunken log2 fold changes (LFC) using "ashr" method [41]. GSEA tests were done in clusterProfiler v3.18.1 [42] using KEGG [43] and Reactome [44] database. For differential expression test and GSEA analysis, adjust p-values for false-discovery rate was calculated using the Benjamini-Hochberg method [45]. Plots were generated using ggplot2 v3.3.6.

High-risk multiple myeloma gene signature analysis

The method for the calculation of EMC-92 signature risk score was modified for use in murine samples from the initial study [18]. Briefly, gene counts of RNA-Seq data were imported as a DESeq object and then applied with DESeq normalization. Variance-stabilizing log2-like transformation of the data was applied using vst command in DESeq package. Human EMC-92 signature genes were firstly mapped to HGNC symbols (https://www.genenames.org) and then converted to *mus musculus* orthologues. Among all the EMC-92 signature genes, 84 genes were converted to mouse orthologues, in which "C1S", "FTL", "DONSON" were found with 2 different orthologues and "SUN1 / GET4" was converted to 2 individual mouse genes. In total, a gene set of 88 genes was used for the calculation. EMC-92 signature risk scores of

7

each sample were then calculated by multiplying the normalized, mean-variance standardized gene expression value with the weighting coefficient factors and the sums of the products (expression value x weighting coefficient) were determined as the EMC-92 signature risk scores.

Analysis of the UAMS-70 gene signature was done through gene set enrichment analysis (GSEA) of the mouse orthologs of the 51 human genes determined to be upregulated in high-risk myeloma patients as described previously [17].

Mice

CD45.1$^+$ transplant recipients were purchased from Jackson Laboratory (stock # 002014) and maintained at Biotron Animal Research Services facility, University of Wisconsin-Madison. Mice were 8-14 weeks old at time of transplant and male and female mice were used in approximately equal proportion. All animal experiments were conducted in accordance with the Guide for the Care and Use of Laboratory Animals and approved by an Animal Care and Use Committee at UW-Madison. The program is accredited by the Association for Assessment and Accreditation of Laboratory Animal Care. All animal experiments in this study are reported in accordance with ARRIVE guidelines (https://arriveguidelines.org).

Transplantation of myeloma cells

Total splenocytes from a moribund VQ-D2 MM bearing mouse were manually ground through a 70μm filter and washed with sterile PBS containing 2% FBS and antibiotic before being resuspended in 90% FBS/10% DMSO and stored in liquid nitrogen until needed. Frozen VQ-D2 donor cells were gently thawed at 37°C before being washed twice in PBS containing 2% mouse serum (Jackson ImmunoResearch, #015-000-120). Donor cells were then resuspended in 200 μl of PBS containing 2% mouse serum. Eight- to fourteen-week-old CD45.1+ recipient

8

mice were sub-lethally irradiated at 4.0 Gy using a CIX3 irradiator (Xstrahl) and transplanted with 2×10^5 donor cells via retro-orbital injection.

Serum protein electrophoresis (SPEP)

Mice were retro-orbitally bled with plain micro hematocrit tubes (Bris, ISO12772). Blood samples were spun in microtainer tubes (BD, 365967) at 4,000x g for 10 minutes to collect serum. Serum was loaded into Hydragel agarose gel (Sebia, 4140) and processed using the Hydrasys instrument (Sebia) following the manufacturer's instruction. The processed film was scanned, and pixel density of Albumin and γ-globulin bands were quantified using Adobe Photoshop.

Complete blood count

Peripheral blood samples were collected via retro-orbital bleeding and analyzed with a Hemavet 950FS (Drew Scientific).

Small compound treatment

For *in vivo* bortezomib treatment, bortezomib (Selleck) was dissolved in sterile PBS and administered at 0.5mg/kg twice a week via intra-peritoneal (IP) injection.

For *in vivo* trametinib treatment, trametinib (Chemietek) was dissolved in 0.5% hydroxypropylmethylcellulose (Sigma) and 0.2% Tween-80 (Sigma) in distilled water (pH 8.0) and administered at 0.25mg/kg via oral gavage daily.

Mice were allocated to treatment groups so that CBC parameters were statistically similar between each group prior to treatment (One-way Analysis of Variance with Tukey-Kramer test). Small compounds were not administered to animals in a blinded manner due to necessary daily preparation of working concentrations for treatment. Animal care staff were

blinded to experimental groups during animal assessment. Post-experiment data analysis was not blinded.

Statistics

For Kaplan–Meier survival curves, survival differences between groups were assessed with the log-rank test, assuming significance at $p<0.05$. Unpaired, two-way t Test was used to determine significant differences between two groups unless specified. One-way Analysis of Variance with Tukey-Kramer test was used to determine the significance between multiple data sets simultaneously unless specified, assuming significance at $p<0.05$. Statistical analysis was carried out using GraphPad Prism v9.1.

Data Availability

All data generated in this study are available upon reasonable request from the corresponding author.

Results

B cell receptor repertoire sequencing shows low rates of somatic hypermutation in VQ myeloma

We previously characterized 4-5 VQ myeloma lines, which have distinct sites of MM growth (including the bone marrow, spleen, or lymph node) and dramatically different survival (Figure 1A) [30]. In addition, isotyping of serum antibody from 5 primary VQ mice revealed ubiquitous kappa light chain secretion, with the VQ-D3 line also secreting lambda light chain but no heavy chain (HC) (Figure 1A). All other lines (VQ-D1, D2, D4, and D5) characterized secreted IgA or IgG class-switched HC (Figure 1A), consistent with a post-germinal center derivation [46]. Germinal center B cells undergo somatic hypermutation (SHM) as part of the antibody affinity maturation process [47]. To assess SHM rates in the VQ model, we carried out high-throughput

10

sequencing analysis of BCR repertoire [48] on four primary VQ donor lines (VQ-D1, D2, D4, and D5), as well as two cell lines derived from VQ-D2 recipient mice (4935 and 4938) [30]. Initial library preparation results showed BCR amplicons that were either clonal or oligoclonal with a dominant clone present in the four primary VQ lines (Figure 1B). Sequencing analysis showed monoclonal immunoglobulin heavy chain (IgH) sequences across all primary VQ lines as well as both cell lines (Figure 1C). In terms of immunoglobulin light chain (IgL) sequencing, VQ-D1, D2, and D4 cells showed near clonal IgL, with dominant IgL clones at a frequency of 95% or higher (Figure 1C). Meanwhile, VQ-D5 cells showed oligo-clonal IgL sequences, with three separate clones displaying a frequency of 5% or higher (Figure 1C). Although VQ 4935 and 4938 cell lines maintained clonal IgH similar to parental VQ-D2 cells, both cell lines developed a secondary IgL clone making up approximately 20-25% of sequences (Figure 1C). Interestingly, differences between IgL sequences in VQ cell lines and parental VQ-D2 cells showed clonal selection *in vitro*: the dominant IgL in the VQ 4935 cell line (clone 1) actually arose from a very minor clone in parental VQ-D2 cells with a frequency of only 0.7% (Figure 1C). This same IgL sequence was also present in VQ 4938 cells (clone 2) but did not become the dominant clone.

Subsequent BCR repertoire analysis of all VQ lines showed median SHM rates of IgH chains ≤3.0% (Figure 1D). This SHM frequency is significantly lower than studies in human MM patients, where median SHM rates of ~8-9% have been previously reported [49,50]. Corroborating initial library preparation, the VQ 4935 cell line showed a higher mean in SHM (mean = 7.1) compared to other samples, which all had mean SHM of < 4 (Figure 1D). Despite this, median SHM for 4935 cells remained low at 2.64%, while both median and mean SHM rates in VQ 4938 cells were similar as those in parental VQ-D2 cells (Figure 1D). Altogether, BCR repertoire analysis shows that VQ myeloma is characterized by secretion of dominant antibody clones that underwent relatively little affinity maturation prior to BM trafficking.

11

Identification of recurrently mutated genes and associated pathways in VQ myeloma cells

Malignant transformation of the VQ model is driven by two common genetic events in human MM: MYC dysregulation and *NRAS Q61R* mutation. However, the prolonged disease latency and phenotypic heterogeneity across different VQ lines suggests the accumulation of additional driver mutations in VQ myeloma cells (Figure 1A). To identify additional mutations, we carried out whole exome sequencing (WES) of CD138+ cells from five VQ mice, as well as corresponding tail samples as germline controls. True mutations were defined as base changes with variant allele frequency (VAF) of >10% in CD138+ cells and <2.5% in matched tail samples. Overall, 11 genes were identified as being recurrently mutated in 2 or more VQ myeloma lines (Figure 2A).

To determine if these genes were recurrently mutated in human MM, mutation status of human orthologs was determined using the Multiple Myeloma Research Foundation (MMRF) CoMMPass database. Of the 11 genes, only *SP140* and *FAT4* were found to be mutated in >3% of patients, Figure 2A), with 37 and 52 out of 1,171 patients, respectively. *SP140* encodes a member of the speckled protein family of chromatin reader proteins [51], which is known to play a role in suppressing Type I interferon signaling in both B cells [52] and macrophages [53]. Frequency of *SP140* mutation ranges from 2.5-7.5% [54,55,56,57] and is found in all stages of MM development from newly diagnosed [54,55] to drug-refractory [57] disease. *FAT4* encodes a member of the protocadherin family previously identified as putative tumor suppressor gene in the context of breast [58] and colorectal cancers [59]. However, a role of *FAT4* in MM pathogenesis has not been characterized. Frequency of *FAT4* mutation was found in 2-12% of newly diagnosed patients [54,56], but not significant in rrMM. Of note, *SP140* and *FAT4* mutations are not particularly enriched in MM patients with RAS/RAF pathway mutations as

12

44% of sequenced patients in the CoMMPass database display mutations in *NRAS, KRAS,* or *BRAF*.

Next, we sought to determine expression of recurrently mutated genes and if their expression levels are altered in myeloma vs control PCs. We generated an expanded RNA-Seq dataset, which includes RNA-Seq analysis of MM cells isolated from multiple tissues from multiple recipient mice in each donor line. Minimum tissue differences were observed in Fragments Per Kilobase of transcript per Million mapped reads (FPKM) from the same mouse. We found that *Erbb4* and *Fat4* were not expressed in PCs (FPKM < 0.1) and 6 of the 11 recurrently mutated genes were differentially expressed in BM VQ cells compared to control PCs (Figure 2A, far right column). Among them, *Map3k13*, *Memo1*, and *Dtwd2* were upregulated, while *Phactr4*, *Sp110*, and *Sp140* were downregulated in VQ myeloma cells. Interestingly, none of these genes have previously been characterized in the context of MM. To determine the potential pathways affected by their dysregulation in the VQ model, we used the Cytoscape network analysis package [60] to create a protein interaction map using human orthologs of the six recurrently mutated and dysregulated genes, along with their 50 closest interacting proteins as determined by Search Tool for the Retrieval of Interacting Genes/Proteins (STRING) analysis [61] (Figure 2B). We further integrated our RNA-Seq data with the protein-protein interaction network (see Figure 2B legend at bottom right). Intriguingly, we noticed that most of the genes in this interaction network were downregulated at the transcriptional level in VQ cells (Figure 2B). STRING functional enrichment of these proteins showed a significant association with both Type I and Type II Interferon signaling, as well as immune responses to viral and tuberculosis infection (Figure 2C). Corroborating these findings, gene set enrichment analysis (GSEA) of VQ MM cells showed downregulation of both Type I and Type II interferon signaling compared to control PCs (Figure 2D). Altogether, our data suggests that interferon signaling is downregulated in VQ myeloma cells.

Copy number variation analysis stratifies VQ myeloma lines into two clusters based on recurrent amplification of chromosome 3 and monosomy of chromosome 5

CNVs are frequent in MM and play a well-established role in predicting patient prognosis [62]. Correspondingly, genetic characterization of the 5T and V$_K$*MYC transgenic mouse models also show a prevalence of CNVs in murine myeloma cells [27,28]. We thus sought to carry out CNV analysis of the five VQ lines using our WES data. Amplification (amp) of chromosome (chr) 3 was observed in three of the five VQ lines (Cluster I: VQ-D2, VQ-D3, and VQ-D5) (Figure 3A), while monosomy of chr5 was observed in the other two lines (Cluster II: VQ-D1 and VQ-D4, Figure 3B). In addition, monosomy chr8, trisomy chr7, and trisomy chr15 were observed in VQ-D1, VQ-D3, and VQ-D4, respectively (Figure 3A-B). Similar CNVs have also been observed in other transgenic MM models: the 5T2 line of the 5T model shows both trisomy 3 and monosomy 5 [27], whereas monosomy 5 is observed in approximately 50% of primary V$_K$*MYC lines [28].

CNVs identified in the VQ model include several tumor suppressor genes and oncogenes that are highly relevant for MM pathogenesis (Figure 3A-B, highlighted in red). Notably, chr3 includes the proto-oncogene *Nras*, as well as *Fam46c* encoding a non-canonical poly(A) polymerase that is frequently mutated in MM and acts as a tumor suppressor for MM development [63,64]. Chr5 includes *Fgfr3* and *Mmset*, both of which are overexpressed in t(4;14) MM patients [12], and cyclin-dependent kinase 6 (*Cdk6*) that was recently found to be over-expressed in immunomodulatory drug-resistant MM cells [65]. Finally, chr15 includes the transcription factor *Myc*, frequently dysregulated in MM [66].

We next wanted to determine if CNVs of the affected genes correlate to their transcriptional changes in VQ vs control PCs. We previously reported moderate overexpression of Myc in all VQ MM lines due to the Vk*MYC transgene [30] and absence of Fgfr3 expression in both

control and VQ myeloma PCs [31]. Using the expanded RNA-Seq dataset, we observed an approximately two-fold increase in *Nras* expression in Cluster I (VQ-D2 and VQ-D5) cells with chr3 amplification compared to control and Cluster II VQ lines without chr3 amplification (Figure 3C). Despite chr3 amplification, expression of *Fam46c* was significantly lower in Cluster I cells compared to Cluster II and control PCs, suggesting an epigenetic mechanism involved in *Fam46c* downregulation (Figure 3D). In Cluster II (VQ-D1 and VQ-D4) cells with monosomy chr5, *Cdk6* was nearly undetectable (average FPKM < 0.5), whereas VQ-D5 cells had a 3-fold decrease compared to VQ-D2 cells and control PCs (Figure 3E). Despite monosomy chr5, *Mmset* expression in Cluster II MM cells was comparable to that in control PCs, whereas it was approximately two-fold higher in Cluster I cells (Figure 3F). Taken together, our data demonstrate that gene copy numbers do not necessarily correlate with mRNA levels in VQ cells, suggesting that additional epigenetic mechanisms play an important role in controlling gene expression.

We previously showed that daily treatment of trametinib (Tra), a potent MEK inhibitor, did not lower myeloma disease burden in VQ-D1 recipients as measured by the ratio of serum gamma-globulin to albumin (G/A) using serum protein electrophoresis (SPEP) (Figure 3G). However, it significantly prolonged the survival of diseased mice [30,31] (Figure 3H). Because Cluster I VQ cells show increased *Nras* copy number and mRNA expression, we wanted to determine if VQ-D2 cells were resistant to MEK inhibition. Here, we followed the same treatment schemes in VQ-D2 recipient mice as previously carried out in VQ-D1 mice (see Materials and Methods). Two weeks after VQ-D2 transplantation, recipients were divided into two groups with comparable CBC parameters and treated with vehicle (Veh) or Tra. Three weeks later, treatment efficacy was assessed via G/A ratio using SPEP assay. Despite the amplification of chr3 with the *Nras* locus in Cluster I myeloma lines, daily Tra treatment provided a survival

benefit to VQ-D2 recipients similar to what has been previously observed in VQ-D1 mice (Figure 3G and 3H).

Gene transcription-based clustering of VQ lines yields highly consistent result with the CNV study

Our CNV analysis separated VQ lines into two distinct clusters based on their characteristic genomic changes, which may cause consistent global transcriptional changes in VQ cells. To test this idea, we performed unsupervised hierarchical clustering analysis of all control and VQ samples based on the transcriptional profiles generated in RNA-Seq. Not surprisingly, these samples split into three groups: Cluster I with chr3 amplification (VQ-D2 and VQ-D5), Cluster II with monosomy chr5 (VQ-D1 and VQ-D4), and control PCs (Figure 4A). Further analysis of sample similarity via t-distributed stochastic neighbor embedding (tSNE) (Figure 4B) and principal component analysis (PCA) (Figure S1) confirmed that VQ myeloma lines cluster into two distinct transcriptional subtypes. This result was not affected by tissue origins of MM cells as similar clustering result was obtained when tissue effect was removed (Figure S2). More importantly, the transcriptional clusters mirror recurrent CNVs observed in these VQ lines.

Cluster I VQ cells display upregulation of growth pathways and high-risk myeloma gene signatures

We next sought to differentiate transcriptional activity between Cluster I and Cluster II VQ myeloma. GSEA of hallmark signaling pathways, oncogenic signatures, and relevant MM-related gene signatures showed numerous pathways significantly upregulated in Cluster I VQ cells, but relatively few in Cluster II (Figure 5A). Interestingly, we observed that essential cancer growth pathways were highly upregulated in Cluster I VQ cells (Figure 5A), including E2F targets (Figure 5B), G2M checkpoint (Figure 5C), and Myc target pathways (Figure 5D). Moreover, Cluster I VQ cells showed significant upregulation of the UAMS-70 and EMC-92 high-

risk MM gene signatures compared to Cluster II VQ lines (Figures 5E). Similar results were observed when we compared hrMM gene signatures of VQ clusters to the transplantable Vκ*Myc line t-Vk12653 [67]. While the UAMS-70 gene signature was enriched in Cluster I VQ compared to t-Vk12653 (Figure S3A), neither of the hrMM gene signatures was enriched in VQ Cluster II cells (Figure S3B). Because EMC92 gene signature includes both positive and negative risk genes with different weights, we calculated the EMC-92 risk scores in Cluster I and II as well as t-Vk12653 MM cells based on a modified version of the published algorithm [18]. The EMC-92 risk score of Cluster I MM cells was significantly higher than those of Cluster II and t-Vk12653 cells (Figure S3C). These analyses clearly defined VQ Cluster I myeloma lines as having high-risk gene expression in comparison to either VQ Cluster II or t-Vk12653 cells.

One of the most significantly upregulated genes in Cluster II VQ is the transcription factor Pbx1, which has recently been characterized as a driver oncogene in the context of chr1q-amp hrMM [68]. PBX1, along with its downstream target FOXM1, were found to drive proliferation and drug resistance in chr1q-amp MM patients [68]. Interestingly, despite loss of *Pbx1* in Cluster I VQ (Figure S4A), transcriptional levels of *Foxm1* in Cluster I VQ cells were significantly higher than that in Cluster II VQ cells or control PCs (Figure S4B). We further assessed the PBX1-FOXM1 axis in VQ myeloma using the recently characterized PBX1 gene signature [68] and a previously established FOXM1 pathway signature [69]. Interestingly, we found that while neither VQ-D2 nor VQ-D5 cells expressed detectable PBX1, Cluster I VQ was enriched for both the PBX1 (Figure S4C) and FOXM1 (Figure S4D) gene signatures compared to both Cluster II VQ and control PCs. In addition, despite high *Pbx1* mRNA expression in VQ-D1 and VQ-D4 cells, neither gene signature was significantly enriched in Cluster II VQ compared to control PCs. The enrichment of PBX1 and FOXM1 gene signatures in Cluster I VQ cells is consistent with their high-risk nature.

Cluster I and Cluster II VQ show distinct responses to bortezomib *in vivo*

When studying the pathways enriched in Cluster II vs Cluster I VQ MM cells, we found that both NF-κB and TNF-α signaling via NF-κB pathways were significantly upregulated in Cluster II VQ (Figure 6A and 6B). We also observed comparable expression of *Il2rg* and increased expression of *Cd74* and *Tnfaip3* in Cluster II vs Cluster I myeloma cells (Figure S5A-C). The expression levels of these three genes are positively correlated to bortezomib (Btz) response in human MM [70]. In addition, in comparison to control PCs, Cluster I myeloma cells overexpressed *Psmb1*, *Psmb2*, and *Psmb 5* (Figure 6E), which encode the potential Btz targeted proteasome subunits [71]. By contrast, Cluster II myeloma cells only overexpressed *Psmb1* and *Psmb 5* (Figure 6E). Together, our results suggest that Cluster I and II myeloma cells may respond differently to Btz treatment.

We previously established the single-agent efficacy of inhibiting the proteasome pathway in VQ-D1 via Btz [31]. Here, we followed the same Btz treatment scheme in VQ-D2 recipient mice as previously established in VQ-D1 (see Materials and Methods). Two weeks after VQ-D2 transplantation, recipients were divided into two groups with comparable CBC parameters and treated with vehicle (Veh) or Btz. Three weeks later, treatment efficacy was assessed via G/A ratio. In contrast to VQ-D1 mice, in which Btz lowered myeloma burden after three weeks of treatment and provided a moderate but significant survival boost [31], this same Btz treatment showed no efficacy in VQ-D2 mice and did not provide any survival benefit compared to Veh treated mice (Figure 6F). This *in vivo* study validated our molecular characterization, demonstrating that Cluster I VQ-D2 has *de novo* Btz resistance and can thus be considered higher risk than Cluster II VQ-D1.

Discussion

18

Previously, our group developed the Ras-driven VQ myeloma model and phenotypically characterized five lines derived from primary VQ mice (Figure 1A) [30]. In the current study, we combined BCR repertoire sequencing, WES, and CNV with RNA-Seq to characterize the genomic and transcriptomic landscapes of VQ MM lines. Both genomic and transcriptional analyses stratified VQ lines into two distinct clusters. Cluster I includes VQ-D2 and VQ-D5, which carry chr3 amplification and display both UAMS-70 and EMC-92 hrMM gene signatures vs Cluster II VQ and t-Vk12653 Vκ*Myc myeloma. By contrast, Cluster II includes VQ-D1 and VQ-D4, which harbor monosomy chr5 and express hrMM gene signatures comparable to t-Vk12653 Vκ*MYC myeloma. Consistent with their molecular classification, Cluster I myeloma cells showed de novo resistance to Btz treatment *in vivo*, while Cluster II myeloma cells exhibited a reduced response to Btz. Interestingly, both Cluster I and II VQ lines responded to Tra. Our molecular stratification of VQ lines provide a foundation to predict the potential outcomes of therapeutic regiments in different populations of human MM patients.

A negative association between the rate of SHM and patient prognosis has been previously established in CLL [72]. However, such an association is unclear in the context of MM, as SHM has been found to have a positive correlation with patient prognosis [49] or none at all [50]. With the exception of the VQ 4935 cell line, BCR repertoire analysis showed relatively low rates of SHM in VQ cells compared to human myeloma [49], with mutation frequencies $\leq 3.0\%$. Although median SHM frequency was slightly higher in Cluster I VQ (VQ-D2/D5) (Figure 1B), there was no significant correlation with median survival of different VQ lines (Figure 1A). Notably, SHM in VQ cells is comparable to levels in primary Vκ*MYC mice, where a median frequency of 2.6% in IgH was observed [24].

WES of five VQ lines identified 11 recurrently mutated genes (Figure 2). Human orthologs of two of these genes, *FAT4* and *SP140*, have been identified as recurrently mutated in numerous MM patient sequencing studies [36-40]. Combining mRNA expression and mutation frequencies during human MM development, we believe that *FAT4* and *SP140* mutations play distinct roles in MM development. *Fat4* expression was nearly undetectable at the mRNA level in both VQ and control PCs (Figure 2A). RNA-Seq data compiled by the MMRF's CoMMPass database shows similarly low expression, with 80% of patients having <2 transcripts per million reads for the *FAT4* gene (data not shown). Moreover, although *FAT4* mutations were detected in a fraction of newly diagnosed patients [54,55], they remained stable in a recent temporal sequencing study of 62 MM patients [73]. Together, these data suggest that *FAT4* mutations serve a passenger role in myeloma progression. By contrast, *Sp140* was found to be expressed in both control and VQ PCs and its mRNA level was significantly downregulated in VQ MM cells (Figure 2A). In MM patients, frequency of *SP140* mutation increases in more advanced stages of disease [54,55,56,57] and also increased over time in the same temporal sequencing study [73]. Altogether, these data suggest that *SP140* mutation plays an important role in MM development and progression.

SP140 encodes a member of the speckled protein family of chromatin reader proteins [51], and has been previously characterized for its role in downregulating interferon signaling in select immune cells [52, 53]. *Sp110*, encoding a homolog of Sp140 [51], was also found to be recurrently mutated in VQ cells and downregulated at the transcriptional level compared to control PCs (Figure 2A). STRING analysis showed that Sp110, along with Phosphatase and actin regulator 4 (Phactr4), whose gene was also recurrently mutated in VQ MM cells, share interactions with several proteins that were also downregulated at the mRNA level in VQ cells (Figure 2B). Pathway analysis of this group of proteins found that they were enriched in Type I and Type II interferon signaling (Figure 2C). Although Phactr4 has been found to act as a tumor

suppressor in hepatocellular carcinoma due to its inhibition of the IL-6/STAT3 pathway [74] and both Sp110 and Sp140 were previously implicated in interferon signaling in response to tuberculosis infection [75], none of these genes have previously been implicated in interferon signaling in the context of MM. Evolutionarily, attenuating Type I and Type II interferon signaling could be beneficial to MM immune evasion and survival. Prior to the development of current novel therapy regiments, IFN-α2b was used as a MM treatment [76] with limited long-term use due to its systemic toxicity [77]. More recent studies in which IFN-α2b is conjugated to antibodies targeting CD38 [78] or HLA-DR [79] on the myeloma cell surface have also shown both *in vitro* and *in vivo* efficacy. Therefore, the role of Sp110 and Sp140 in shaping the interferon response in VQ as well as in human MM is an interesting question worth future consideration.

Large-scale chromosomal changes, including translocations and aneuploidies, have significant impacts on treatment outcomes in MM patient prognosis. For example, use of the BCL-2 inhibitor venetoclax for t(11;14) MM [80] or the innate resistance of t(14;16) MM to PIs due to c-MAF overexpression [81,82] are well documented. Therefore, we performed CNV analysis in five VQ lines and identified two recurrent, mutually exclusive events: amplification of chr3, which was present in Cluster I (VQ-D2, D3, and D5; Figure 3A), and monosomy chr5, which was present in Cluster II (VQ-D1 and D4; Figure 3B). Neither of these CNVs are unique to the VQ model. Full gain(chr3) has been identified in the 5T2 line, while the 5TGM1 cells showed partial duplication [27]. Interestingly, monosomy 5 is a unifying CNV across all three murine MM models. Monosomy 5 was also observed in 5T2 cells, while only partially deleted in 5T33vv and 5TGM1 lines. Noticeably, monosomy 5 was present in 13/26 sequenced Vκ*MYC lines [28]. The high incidence of monosomy chr5 across multiple MM models warrants further investigation.

21

Consistent with our CNV analysis, gene transcriptional analyses of RNA-Seq data, including non-supervised hierarchal clustering, PCA, and tSNE, identified two distinct transcriptional clusters (Figure 4 and Figure S1) regardless of their discrete tissue origins (Figure S2). Extensive pathway analysis between these two clusters as well as in comparison to the t-Vk12653 Vk*MYC line revealed the hrMM order: Cluster I > Cluster II = t-Vk12653 (Figure 5 and S4), On the other hand, Cluster II (VQ-D1/D4) cells were enriched for the TNF-α/NF-κB signaling (Figure 6A and 6B) and increased expression of genes linked to Btz response in patients (Figure 6C). These transcriptional differences led us to speculate that Cluster I cells may be more resistant to PIs than Cluster II cells. In our initial study, treatment of VQ-D1 recipient mice with Btz following a dosing schedule found to be effective in the Vκ*MYC model (1.0mg/kg Btz treatment on days 1,4,8,11) [67] was not effective as a single agent, but did provide a survival benefit when combined with the MEK inhibitor selumetinib [30]. However, in a more recent study, we found that reducing Btz dosage to current clinical practices provided a modest but significant boost to VQ-D1 survival and significantly reduced G/A ratio after three weeks of treatment [31] (Figure 6F). By contrast, in our current study we found that the same Btz treatment was completely ineffective against VQ-D2 cells (Figure 6F), corroborating our molecular analysis.

We observed that despite chr3 amplification and Nras overexpression in Cluster I MM, both VQ-D1 and VQ-D2 MM cells responded to Tra treatment (Figure 3). By contrast, 5T2 MM cells showed increased resistance to Tra compared to other 5T lines, though this is complicated by the fact that the genomic region containing *Kras* is also duplicated in 5T2 cells [27]. Nonetheless, our results provide a strong rationale to develop Tra-based combination therapies in hrMM and rrMM, particularly in the context of RAS mutations.

Altogether, our data further elucidate the genomic and transcriptional landscapes of the VQ model. These molecular characterizations, along with functional validation via *in vivo* treatment experiments, supports an intra-model stratification where Cluster II VQ (VQ-D1/D4) models Ras-driven MM with some hrMM features comparable to non-Ras-driven MM (e.g. t-Vk*MYC), whereas Cluster I VQ (VQ-D2/D5) represents ultra-high risk multiple myeloma. Future studies using the VQ model as pre-clinical platform must necessarily take these intra-model differences into account when designing mechanistic studies or testing novel treatments.

Acknowledgments

References

1. Palumbo, A. & Anderson, K. Multiple myeloma. *N Engl J Med* **364**, 1046–1060 (2011).

2. Mouhieddine, T. H., Weeks, L. D. & Ghobrial, I. M. Monoclonal gammopathy of undetermined significance. *Blood* **133**, 2484–2494 (2019).

3. Brousseau, M. *et al.* Hyperdiploidy Is a Common Finding in Monoclonal Gammopathy of Undetermined Significance and Monosomy 13 Is Restricted to These Hyperdiploid Patients. *Clinical Cancer Research* **13**, 6026–6031 (2007).

4. Karlin, L. *et al.* Clinical and biological features of t(4;14) multiple myeloma: a prospective study. *Leukemia & Lymphoma* **52**, 238–246 (2011).

5. Fonseca, R. *et al.* Genomic abnormalities in monoclonal gammopathy of undetermined significance. *Blood* **100**, 1417–1424 (2002).

6. Avet-Loiseau, H. *et al.* Genetic abnormalities and survival in multiple myeloma: the experience of the Intergroupe Francophone du Myélome. *Blood* **109**, 3489–3495 (2007).

7. Vikova, V. *et al.* Comprehensive characterization of the mutational landscape in multiple myeloma cell lines reveals potential drivers and pathways associated with tumor progression and drug resistance. *Theranostics* **9**, 540–553 (2019).

8. Walker, B. A. *et al.* Aberrant global methylation patterns affect the molecular pathogenesis and prognosis of multiple myeloma. *Blood* **117**, 553–562 (2011).

9. Walker, B. A. *et al.* A compendium of myeloma-associated chromosomal copy number abnormalities and their prognostic value. *Blood* **116**, e56–e65 (2010).

10. Mikhael, J. R. *et al.* Management of Newly Diagnosed Symptomatic Multiple Myeloma: Updated Mayo Stratification of Myeloma and Risk-Adapted Therapy (mSMART) Consensus Guidelines 2013. *Mayo Clinic Proceedings* **88**, 360–376 (2013).

11. D'Agostino, M. *et al.* Second Revision of the International Staging System (R2-ISS) for Overall Survival in Multiple Myeloma: A European Myeloma Network (EMN) Report Within the HARMONY Project. *JCO* JCO.21.02614 (2022) doi:10.1200/JCO.21.02614.

12. Kalff, A. & Spencer, A. The t(4;14) translocation and FGFR3 overexpression in multiple myeloma: prognostic implications and current clinical strategies. *Blood Cancer Journal* **2**, e89–e89 (2012).

13. Mirabella, F. *et al.* MMSET is the key molecular target in t(4;14) myeloma. *Blood Cancer Journal* **3**, e114–e114 (2013).

14. Mina, R. *et al.* Clinical features and survival of multiple myeloma patients harboring t(14;16) in the era of novel agents. *Blood Cancer J.* **10**, 1–4 (2020).

15. Avet-Loiseau, H. *et al.* Long-Term Analysis of the IFM 99 Trials for Myeloma: Cytogenetic Abnormalities [t(4;14), del(17p), 1q gains] Play a Major Role in Defining Long-Term Survival. *JCO* **30**, 1949–1952 (2012).

16. Locher, M. *et al.* The prognostic value of additional copies of 1q21 in multiple myeloma depends on the primary genetic event. *Am J Hematol* **95**, 1562–1571 (2020).

17. Shaughnessy, J. D., Jr *et al.* A validated gene expression model of high-risk multiple myeloma is defined by deregulated expression of genes mapping to chromosome 1. *Blood* **109**, 2276–2284 (2006).

18. Kuiper, R. *et al.* A gene expression signature for high-risk multiple myeloma. *Leukemia* **26**, 2406–2413 (2012).

19. Caro, J. *et al.* How to Treat High-Risk Myeloma at Diagnosis and Relapse. *American Society of Clinical Oncology Educational Book* 291–309 (2021) doi:10.1200/EDBK_320105.

20. Solimando, A. G. *et al.* High-Risk Multiple Myeloma: Integrated Clinical and Omics Approach Dissects the Neoplastic Clone and the Tumor Microenvironment. *J Clin Med* **8**, (2019).

21. Rossi, M. *et al.* Mouse models of multiple myeloma: technologic platforms and perspectives. *Oncotarget* **9**, 20119–20133 (2018).

22. Cooke, R. E., Koldej, R. & Ritchie, D. Immunotherapeutics in Multiple Myeloma: How Can Translational Mouse Models Help? *J Oncol* **2019**, (2019).

23. Morito, N. *et al.* A Novel Transgenic Mouse Model of the Human Multiple Myeloma Chromosomal Translocation t(14;16)(q32;q23). *Cancer Research* **71**, 339–348 (2011).

24. Chesi, M. *et al.* AID-dependent activation of a MYC transgene induces multiple myeloma in a conditional mouse model of post-germinal center malignancies. *Cancer Cell* **13**, 167–180 (2008).

25. Radl, J., Glopper, E. D., Schuit, H. R. E. & Zurcher, C. Idiopathic Paraproteinemia: II. Transplantation of the Paraprotein-Producing Clone from Old to Young C57BL/KaLwRij Mice. *The Journal of Immunology* **122**, 609–613 (1979).

26. Garrett, I. R., Dallas, S., Radl, J. & Mundy, G. R. A murine model of human myeloma bone disease. *Bone* **20**, 515–520 (1997).

27. Maes, K. *et al.* The genetic landscape of 5T models for multiple myeloma. *Scientific Reports* **8**, 15030 (2018).

28. Chesi, M. *et al.* Monosomic Loss of MIR15A/MIR16-1 Is a Driver of Multiple Myeloma Proliferation and Disease Progression. 14 (2020).

29. Avet-Loiseau, H. *et al.* Monosomy 13 Is Associated With the Transition of Monoclonal Gammopathy of Undetermined Significance to Multiple Myeloma. *Blood* **94**, 2583–2589 (1999).

30. Wen, Z. *et al.* Expression of NrasQ61R and MYC transgene in germinal center B cells induces a highly malignant multiple myeloma in mice. *Blood* **137**, 61–74 (2021).

31. Flietner, E. *et al.* Ponatinib sensitizes myeloma cells to MEK inhibition in the high-risk VQ model. *Sci Rep* **12**, 10616 (2022).

32. Turchaninova, M. A. *et al.* High-quality full-length immunoglobulin profiling with unique molecular barcoding. *Nat Protoc* **11**, 1599–1616 (2016).

33. Bolotin, D. A. *et al.* MiXCR: software for comprehensive adaptive immunity profiling. *Nat Methods* **12**, 380–381 (2015).

34. Gupta, N. T. *et al.* Change-O: a toolkit for analyzing large-scale B cell immunoglobulin repertoire sequencing data. *Bioinformatics* **31**, 3356–3358 (2015).

35. You, X. *et al.* Asxl1loss cooperates with oncogenic Nras in mice to reprogram immune microenvironment and drive leukemic transformation. *Blood* blood.2021012519 (2021) doi:10.1182/blood.2021012519.

36. Li, H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. Preprint at https://doi.org/10.48550/arXiv.1303.3997 (2013).

37. Talevich, E., Shain, A. H., Botton, T. & Bastian, B. C. CNVkit: Genome-Wide Copy Number Detection and Visualization from Targeted DNA Sequencing. *PLOS Computational Biology* **12**, e1004873 (2016).

38. Ritchie, M. E. *et al.* limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Research* **43**, e47 (2015).

39. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2010).

40. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology* **15**, 550 (2014).

41. Stephens, M. False discovery rates: a new deal. *Biostatistics* **18**, 275–294 (2017).

42. Yu, G., Wang, L.-G., Han, Y. & He, Q.-Y. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* **16**, 284–287 (2012).

43. Kanehisa, M. & Goto, S. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Research* **28**, 27–30 (2000).

44. Fabregat, A. *et al.* The Reactome Pathway Knowledgebase. *Nucleic Acids Research* **46**, D649–D655 (2018).

45. Benjamini, Y. & Hochberg, Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society: Series B (Methodological)* **57**, 289–300 (1995).

46. González, D. *et al.* Immunoglobulin gene rearrangements and the pathogenesis of multiple myeloma. *Blood* **110**, 3112–3121 (2007).

47. Maul, R. W. & Gearhart, P. J. AID AND SOMATIC HYPERMUTATION. *Adv Immunol* **105**, 159–191 (2010).

48. Turchaninova, M. A. *et al.* High-quality full-length immunoglobulin profiling with unique molecular barcoding. *Nat Protoc* **11**, 1599–1616 (2016).

49. Medina, A. *et al.* Molecular profiling of immunoglobulin heavy-chain gene rearrangements unveils new potential prognostic markers for multiple myeloma patients. *Blood Cancer J.* **10**, 1–12 (2020).

50. Ferrero, S. *et al.* Multiple myeloma shows no intra-disease clustering of immunoglobulin heavy chain genes. *Haematologica* **97**, 849–853 (2012).

51. Fraschilla, I. & Jeffrey, K. L. The Speckled Protein (SP) Family: Immunity's Chromatin Readers. *Trends in Immunology* **41**, 572–585 (2020).

52. Karaky, M. *et al.* SP140 regulates the expression of immune-related genes associated with multiple sclerosis and other autoimmune diseases by NF-κB inhibition. *Human Molecular Genetics* **27**, 4012–4023 (2018).

53. Ji, D. X. *et al.* Role of the transcriptional regulator SP140 in resistance to bacterial infections via repression of type I interferons. *eLife* **10**, e67290 (2021).

54. Hu, Y., Chen, W. & Wang, J. Mutations In Thirty Hotspot Genes In Newly Diagnosed Chinese Multiple Myeloma Patients. *Onco Targets Ther* **12**, 9999–10010 (2019).

55. Walker, B. A. *et al.* Identification of novel mutational drivers reveals oncogene dependencies in multiple myeloma. *Blood* **132**, 587–597 (2018).

56. Bolli, N. *et al.* Heterogeneity of genomic evolution and mutational profiles in multiple myeloma. *Nat Commun* **5**, (2014).

57. Kortüm, K. M. *et al.* Targeted sequencing of refractory myeloma reveals a high incidence of mutations in CRBN and Ras pathway genes. *Blood* **128**, 1226–1233 (2016).

58. Qi, C., Zhu, Y. T., Hu, L. & Zhu, Y.-J. Identification of Fat4 as a candidate tumor suppressor gene in breast cancers. *Int J Cancer* **124**, 793–798 (2009).

59. Wei, R. *et al.* FAT4 regulates the EMT and autophagy in colorectal cancer cells in part via the PI3K-AKT signaling axis. *Journal of Experimental & Clinical Cancer Research* **38**, 112 (2019).

60. Franz, M. *et al.* Cytoscape.js: a graph theory library for visualisation and analysis. *Bioinformatics* **32**, 309–311 (2016).

61. Szklarczyk, D. *et al.* The STRING database in 2021: customizable protein-protein networks, and functional characterization of user-uploaded gene/measurement sets. *Nucleic Acids Res* **49**, D605–D612 (2021).

62. Avet-Loiseau, H. *et al.* Prognostic Significance of Copy-Number Alterations in Multiple Myeloma. *J Clin Oncol* **27**, 4585–4590 (2009).

63. Mroczek, S. *et al.* The non-canonical poly(A) polymerase FAM46C acts as an onco-suppressor in multiple myeloma. *Nat Commun* **8**, 619 (2017).

64. Zhu, Y. X. *et al.* Loss of FAM46C promotes cell survival in myeloma. *Cancer Res* **77**, 4317–4327 (2017).

65. Ng, Y. L. D. *et al.* Proteomic profiling reveals CDK6 upregulation as a targetable resistance mechanism for lenalidomide in multiple myeloma. *Nat Commun* **13**, 1009 (2022).

66. Jovanović, K. K. *et al.* Targeting MYC in multiple myeloma. *Leukemia* **32**, 1295–1306 (2018).

67. Chesi, M. *et al.* Drug response in a genetically engineered mouse model of multiple myeloma is predictive of clinical efficacy. *Blood* **120**, 376–385 (2012).

68. Trasanidis, N. *et al.* Systems medicine dissection of chr1q-amp reveals a novel PBX1-FOXM1 axis for targeted therapy in multiple myeloma. *Blood* **139**, 1939–1953 (2022).

69. Schaefer, C. F. *et al.* PID: the Pathway Interaction Database. *Nucleic Acids Res* **37**, D674-679 (2009).

70. Broyl, A. *et al.* Gene expression profiling for molecular classification of multiple myeloma in newly diagnosed patients. *Blood* **116**, 2543–2553 (2010).

71. Besse, A. *et al.* Proteasome Inhibition in Multiple Myeloma: Head-to-Head Comparison of Currently Available Proteasome Inhibitors. *Cell Chemical Biology* **26**, 340-351.e3 (2019).

72. Hamblin, T. J., Davis, Z., Gardiner, A., Oscier, D. G. & Stevenson, F. K. Unmutated Ig V(H) genes are associated with a more aggressive form of chronic lymphocytic leukemia. *Blood* **94**, 1848–1854 (1999).

73. Farswan, A. *et al.* Branching clonal evolution patterns predominate mutational landscape in multiple myeloma. *Am J Cancer Res* **11**, 5659–5679 (2021).

74. Cao, F., Liu, M., Zhang, Q.-Z. & Hao, R. PHACTR4 regulates proliferation, migration and invasion of human hepatocellular carcinoma by inhibiting IL-6/Stat3 pathway. *Eur Rev Med Pharmacol Sci* **20**, 3392–3399 (2016).

75. Pichugin, A. V., Yan, B.-S., Sloutsky, A., Kobzik, L. & Kramnik, I. Dominant Role of the sst1 Locus in Pathogenesis of Necrotizing Lung Granulomas during Chronic Tuberculosis Infection and Reactivation in Genetically Resistant Hosts. *Am J Pathol* **174**, 2190–2201 (2009).

76. Mellstedt, H. *et al.* Interferon therapy in myelomatosis. *Lancet* **1**, 245–247 (1979).

77. Fritz, E. & Ludwig, H. Interferon-alpha treatment in multiple myeloma: meta-analysis of 30 randomised trials among 3948 patients. *Ann Oncol* **11**, 1427–1436 (2000).

78. Pogue, S. L. *et al.* Targeting Attenuated Interferon-α to Myeloma Cells with a CD38 Antibody Induces Potent Tumor Regression with Reduced Off-Target Activity. *PLoS ONE* **11**, e0162472 (2016).

79. Rossi, E. A. *et al.* Preclinical studies on targeted delivery of multiple IFNα2b to HLA-DR in diverse hematologic cancers. *Blood* **118**, 1877–1884 (2011).

80. Kumar, S. *et al.* Efficacy of venetoclax as targeted therapy for relapsed/refractory t(11;14) multiple myeloma. *Blood* **130**, 2401–2409 (2017).

81. Qiang, Y.-W. *et al.* MAF protein mediates innate resistance to proteasome inhibition therapy in multiple myeloma. *Blood* **128**, 2919–2930 (2016).

82. Wei, G. *et al.* Clinical implications of c-maf expression in plasma cells from patients with multiple myeloma. *Experimental Hematology & Oncology* **6**, 16 (2017).

Figure Legends

Figure 1. B cell receptor repertoire analysis shows dominant clonality and low somatic hypermutation (SHM) rates in primary VQ cells and VQ cell lines. (A) Table summarizing previously established characteristics of VQ donor lines. (B-D) B cell receptor heavy-chain and light-chain repertoire analysis was carried out as described in Methods. (B) Tapestation image showing immunoglobulin heavy chain and light chain library amplification for samples in panels C and D. (C) Pie charts depicting clonal frequency for immunoglobulin heavy chain (IgH) sequences (top) and immunoglobulin light chain sequences (bottom) sequences from primary VQ cells (VQ-D1, D2, D4, and D5) and from VQ-D2 derived cell lines (4935 and 4938). (D) Violin plots showing mean (black dots) and distribution of the total somatic hypermutation frequencies across B-cell IgH sequences from primary VQ cells (VQ-D1, D2, D4, and D5) and from VQ-D2 derived cell lines (4935 and 4938).

Figure 2. Whole exome sequencing identifies recurrently mutated genes in VQ myeloma cells. Five paired tail DNA (non-leukemia control) and genomic DNA from VQ CD138$^+$ B220$^-$ plasma cells were subjected to whole exome sequencing as described in Materials and Methods. (A) Recurrently mutated genes (mutated in $\geq$ 2/4 mice) and their variant allele frequencies (VAF) in VQ cells, frequency of mutation in human orthologs as determined from a cohort of 1,171 MM patient samples from the CoMMPass database, and mRNA expression (as determined by Fragments Per Kilobase of gene per Million mapped reads [FPKM]) in VQ cells

are shown. (B) Cytoscape-generated STRING network of recurrently mutated genes with differential RNA expression, as well as 40 closest interacting proteins as determined by STRING analysis. Recurrently mutated genes are represented by diamonds. Interacting proteins are denoted by circles. Genes with differential expression are color-coded. (C) Select list of pathways enriched in genes highlighted in panel B. (D) GSEA of Reactome gene sets for interferon alpha/beta signaling (top) and interferon gamma signaling (bottom) between VQ and control plasma cells. NES, normalized enrichment score. FDR, false discovery rate.

Figure 3. Copy number variation (CNV) analysis stratifies VQ cells based on recurrent amplification of chromosome 3 and monosomy chromosome 5. (A-B) CNV analysis was performed using the whole exome sequencing data as described in Materials and Methods. Orange dots indicate significant changes in log2 copy ratio for a given call segment in plasma cells compared to non-leukemia control samples. Location and name of tumor suppressors and oncogenes related to myeloma pathogenesis are shown in red. CNV plots are grouped according to recurrent CNV status. (C-F) Transcript levels of genes highlighted in red in panels A and B are shown in CD138+ B220- cells from control and VQ recipient mice. FPKM, Fragments Per Kilobase of transcript per Million mapped reads. One-way analysis of variance with Tukey's post-test was performed. Tissue of origin for individual samples is denoted by legend. (G-H) CD45.1 recipient mice were sub-lethally irradiated and injected with bone marrow cells from moribund VQ-D1 donor mouse or splenocytes from moribund VQ-D2 donor mouse. Six weeks (VQ-D1) or two weeks (VQ-D2) post-transplant, mice were treated with vehicle or trametinib. (G) Serum protein electrophoresis was performed to quantify the γ-globulin/Albumin (G/A) ratios in VQ-D1 and VQ-D2 recipient mice at day 21 of treatment. Two-sided t-test was performed. (H) Kaplan-Meier survival curves were plotted against days after treatment. Log-rank

test was performed. Note: VQ-D1 results are combined from historical [30,31] and new data. ns, not significant. *, p<0.05. **, p<0.01. ***, p<0.001. ****, p<0.0001.

Figure 4. RNA-Seq analysis reveals two distinct transcriptional clusters of VQ myeloma. Bulk RNA-Seq analysis was performed using flow sorted CD138+ B220- CD45.2+ cells from bone marrow (BM) of control mice (n=3) and BM, spleen (SPL), lymph node (LN), or liver (Liv) of multiple VQ-D1, VQ-D2, VQ-D4, and VQ-D5 recipients. (A) Clustered heat map of RNA-seq gene count data. Samples are color-coded by tissue sites and VQ lines as indicated. (B) tSNE analysis of gene counts data of VQ myeloma samples. Samples are color-coded by VQ lines as in panel A. Tissue of origin for individual samples is denoted by legend.

Figure 5. VQ Cluster I myeloma cells have increased expression of Hallmark growth pathways and high-risk MM gene signatures. (A) Overview of gene set enrichment analysis between Cluster I (VQ-D2/D5) and Cluster II (VQ-D1/D4) VQ myeloma cells. Relevant pathways are highlighted in red. (B-E) GSEA plots comparing Cluster I to Cluster II for (B) Hallmark_E2F_Targets, (C) Hallmark_G2M Targets, (D) Hallmark_MYC_Targets_V1 (left) Hallmark_MYC_Targets_V2 (right), and (E) EMC-92 hrMM gene signature (left) and UAMS-70 hrMM gene signature (right). NES, normalized enrichment score; FDR, false discovery rate; p. adj., adjusted P-value.

Figure 6. Cluster I and Cluster II VQ cells display distinct responses to bortezomib *in vivo.* (A-C) Gene set enrichment analysis of proteasome inhibitor related pathways between Cluster I (VQ-D2/D5) and Cluster II (VQ-D1/D4) VQ myeloma cells. KEGG_ NF-kappa B signaling pathway (A), Hallmark_TNFA signaling via NFKB (B), and KEGG_Proteasome

pathway (C). (D) Transcript levels of NF-κB related genes *Tnfaip3*, *CD74*, and *Il2rg* in CD138+ B220- cells from Cluster I and II VQ myeloma mice. Two-sided t-Test was performed. VQ donor of origin is color-coded as indicated. Results are presented as mean + SD. FPKM, Fragments Per Kilobase of transcript per Million mapped reads. (E) Transcript levels of proteasome related genes *Psmb1*, *Psmb2*, and *Psmb5* in CD138+ B220- cells from control plasma cells (Control) and Cluster I and II VQ cells. Two-sided t-Test was performed. VQ donor of origin is color-coded as indicated. Results are presented as mean + SD. (F-G) CD45.1 recipient mice were sub-lethally irradiated and injected with bone marrow cells from moribund VQ-D1 donor mouse or splenocytes from moribund VQ-D2 donor mouse. Six weeks (VQ-D1) or two weeks (VQ-D2) posttransplant, mice were treated with vehicle or bortezomib as described in Materials and Methods. (F) Serum protein electrophoresis was performed to quantify the γ-globulin/Albumin (G/A) ratios in VQ-D2 recipient mice at day 21 of treatment. Two-sided t-Test was performed. (G) Kaplan-Meier survival curves were plotted against days after treatment. Log-rank test was performed. Note: VQ-D1 results are taken from historical data [31]. FDR, false discovery rate; NES, normalized enrichment score; ns, not significant; p. adj., adjusted P-value. *, p <0.05; **, p <0.01; ***, p <0.001; ****, p <0.0001.

Figure 1-Zhang

B

Mouse Ig Heavy Chain | Mouse Ig Light Chain

| [bp] | Ladder A0 (L) | VQ-D1 A1 | VQ-D2 B1 | VQ-D4 C1 | VQ-D5 D1 | 4935 E1 | 4938 F1 | VQ-D1 G1 | VQ-D2 H1 | VQ-D4 A2 | VQ-D5 B2 | 4935 C2 | 4938 D2 |

A

VQ Line	Gender	Dominant clone	Disease sites	Median Survival (days)
D1	M	IgG2κ	BM, SPL	80
D2	M	IgAκ	SPL, LN	21
D3	F	κ + λ only	LN	n/a
D4	M	IgAκ	BM, LN	60
D5	F	IgAκ	SPL, LN	40

C

D

	VQ-D1	VQ-D2	VQ-D4	VQ-D5	4935 cell line	4938 cell line
Mean	1.59	2.94	1.65	2.97	7.06	3.87
Median	1.37	2.56	0.63	2.68	2.64	2.74
StdDev	±1.46	±2.06	±2.87	±0.93	±8.56	±3.4
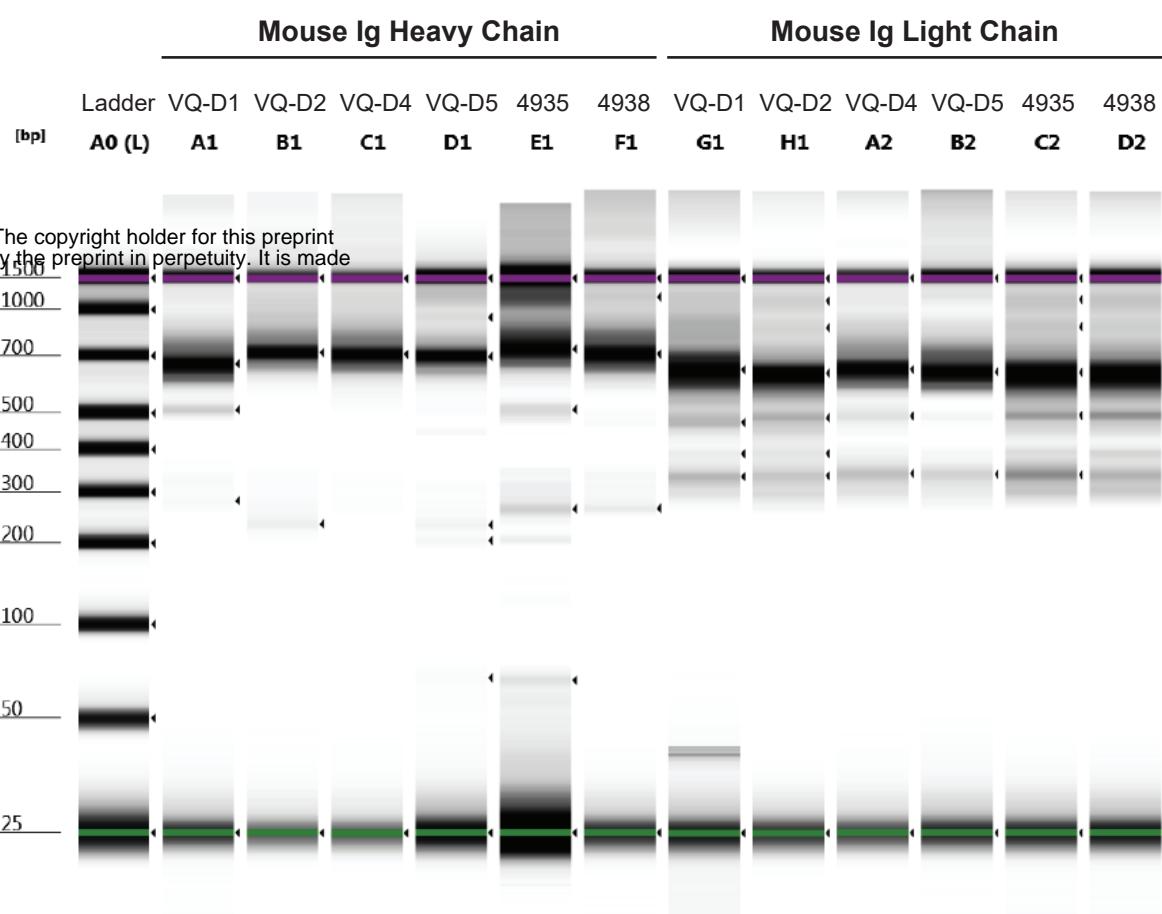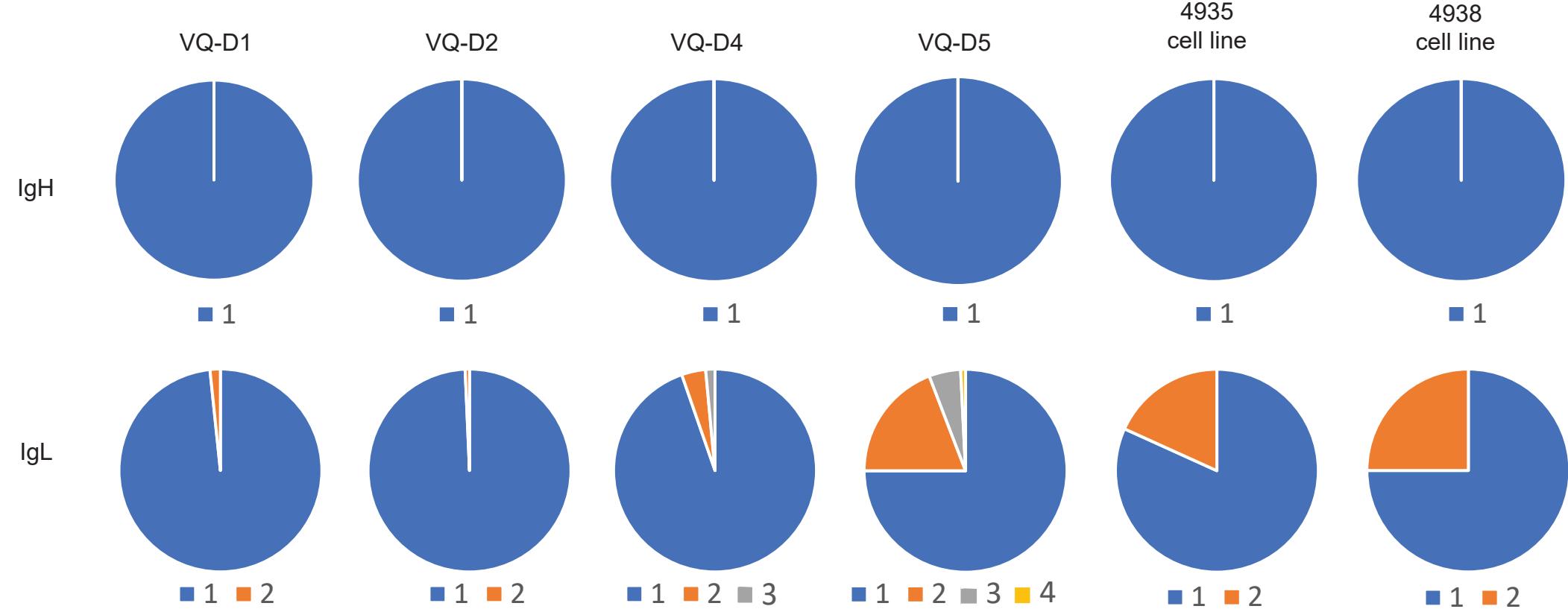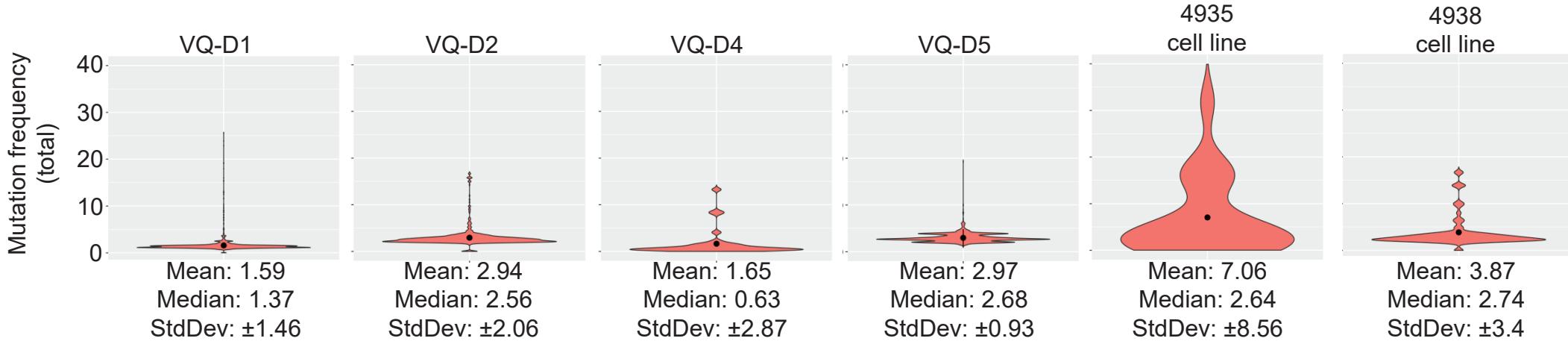
Figure 2-Zhang

A

Variant Allele Frequency for recurrent mutations in >=2 mice		VQ Donor #					CoMMPass database n=1171 (% reurrence with RAS/RAF)	VQ RNA expression (FPKM)
		D1	D2	D3	D4	D5		
Protein kinase	Erbb4	0.39	0	0	0.32	0.69	13 (31%)	0.00
	Map3k13	0	0	0	0.24	0.11	1 (0%)	4.31
Protein phosphatase	Trrap	0	0.15	0	0.45	0.44	14 (50%)	8.74
	Phactr4	0	0.17	0.73	0	0.40	4 (75%)	2.31
Transcription factor	Sp110	0	0	0.21	0.19	0	2 (0%)	0.61
	Sp140	0.23	0	0	0.15	0.33	37 (46%)	7.09
Cell Migration	Fat4	0	0	1.00	0.28	0.59	52 (56%)	0.02
	Memo1	0	0.23	0	0.50	0	2 (0%)	7.08
Other	Cdc5l	0	0.25	0	0.52	1.00	1 (0%)	20.12
	Dtwd2	0.31	0	0.80	0	0	1 (0%)	1.84
	Dnajb6	0	0.12	0.64	0	0	2 (0%)	3.14

Missense   Upregulated
Nonsense   Downregulated
Frameshift   No Change

B

STRING of human orthologs

Log2 fold change
-3.5   3.5

◆ Recurrently mutated gene
● Interacting protein (expressed)
● Interacting protein (not expressed)

Type I Interferon signaling

C

FDR

Type II interferon signaling (IFNG)	(1.31E-5)
Response to interferon-alpha	(6.56E-6)
SARS-CoV-2 innate immunity evasion and cell-specific immune response	(1.79E-6)
ISG15 antiviral mechanism	(6.64E-8)
Negative regulation of viral genome replication	(3.55E-8)
Immune response to tuberculosis	(1.1E-10)
Type I interferon induction and signaling during SARS-CoV-2 infection	(4.45E-12)
Type I interferon signaling pathway	(1.91E-25)

Number of proteins

D

REACTOME_INTEFERON_ALPHA_BETA_SIGNALING
NES: -1.79
FDR q-val:0.005
VQ   WT

REACTOME_INTEFERON_GAMMA_SIGNALING
NES: -1.92
FDR q-val:0.001
VQ   WT

Figure 3-Zhang

Figure 4-Zhang

A

B

Figure 5-Zhang

A

Cluster I
(VQ-D2/D5)

Cluster II
(VQ-D1/D4)

E2F_Targets

MYC_Targets_v1

G2M_Checkpoint

MYC_Targets_v2

TNFA_Signaling_via_NFKB

UAMS_70

EMC_92

NF-kappa B signaling pathway

−log10(p.adjust)

NES

B HALLMARK_E2F_TARGETS
NES = −2.26
FDR = 1.68E-09
VQ-D1/D4   VQ-D2/D5
Rank in Ordered Dataset

C HALLMARK_G2M_CHECKPOINT
NES = −2.01
FDR = 8.53E-08
VQ-D1/D4   VQ-D2/D5
Rank in Ordered Dataset

D HALLMARK_MYC_TARGETS_V1
NES = −2.16
FDR = 1.68E-09
VQ-D1/D4   VQ-D2/D5
Rank in Ordered Dataset

HALLMARK_MYC_TARGETS_V2
NES = −2.04
FDR = 2.41E-05
VQ-D1/D4   VQ-D2/D5
Rank in Ordered Dataset

E EMC_92_gene_signature
NES = −1.42
FDR = 0.037
VQ-D1/D4   VQ-D2/D5
Rank in Ordered Dataset

UAMS_70_gene_signature_UP
NES = −1.75
FDR = 0.0022
VQ-D1/D4   VQ-D2/D5
Rank in Ordered Dataset

Figure 6-Zhang