JOURNAL OF
EMERGING INVESTIGATORS

Article

Identification of potential therapeutic targets for multiple myeloma by gene expression analysis

Nathaniel J. Kochenderfer[1], James N. Kochenderfer[1]
[1] Sidwell Friends School, Washington, District of Columbia

SUMMARY

New treatments are needed for multiple myeloma (MM), an almost always incurable cancer of plasma cells. Cancer therapies might have fewer side effects if they are aimed at specific protein targets expressed by cancer cells. We hypothesized that differences in gene expression between multiple myeloma cells and normal cells could be used to identify new targets for MM therapies. RNA sequencing (RNAseq) is a technique for quantifying gene expression in cells. We analyzed RNAseq data from the Cancer Cell Line Encyclopedia to identify genes with high expression in MM cell lines versus cell lines derived from cancers other than MM. Next, the 200 genes with the highest expression in MM cell lines versus other cell lines were evaluated for RNA expression in primary MM samples from patients. Of these genes, 72 were expressed in at least 23 of 25 primary samples from patients. We evaluated the 72 genes expressed at high levels in MM cell lines and expressed in primary MM cells for expression in normal human tissues. We evaluated gene expression in normal human tissues by using publicly available RNAseq data. Genes with high expression in normal tissues were eliminated as potential targets to decrease the chance of side effects caused by therapies targeting proteins expressed by normal tissues. We identified 19 genes with high expression in MM and low expression in normal tissues. The proteins encoded by these genes are promising therapeutic targets for MM and include Prepronociceptin (PNOC) and Interferon Regulatory Factor 4 (IRF4).

INTRODUCTION

Nonspecific therapies for cancer, such as chemotherapy, do not specifically target cancer cells and, consequently, damage both cancer and normal cells (1–3). While nonspecific therapies can lead to cures for certain cancers, they can also be detrimental to the health of patients (1–3). Targeted therapies for cancer can be less toxic than chemotherapy because targeted therapies are specifically designed to target proteins expressed at higher levels in cancer cells versus normal cells (4). Targeted therapies are usually in the form of small-molecule drugs or monoclonal antibodies (4). Some immunotherapies such as chimeric antigen receptor-expressing (CAR) T-cells are also targeted to specific proteins on cancer cells (5).

Multiple myeloma (MM) is a cancer of plasma cells (6). The American Cancer Society estimates that 35,730 new cases of MM will occur in the United states in 2023 and 12,590 patients will die of MM in 2023 (7). MM is staged by the Revised International Staging System, which divides patients into three groups. For the lowest risk group, the estimated chance of overall survival five years after diagnosis was 82%, and the estimated chance of progression-free survival, which means MM did not progress after initial therapy, was 55% (8). For the highest risk group, the estimated chance of overall survival five years after diagnosis was 40%, and the estimated chance of progression-free survival was 24% (8). The median survival of patients with MM is highly variable, but, compared with people who have not been diagnosed with MM, survival has been shown to be decreased for MM patients of all ages (6, 9). MM is rarely curable with traditional therapies (10, 11).

MM can be treated by a variety of approaches including cytotoxic chemotherapy followed by autologous hematopoietic stem cell transplant, corticosteroids, immunomodulatory drugs, proteasome inhibitors, monoclonal antibodies, and CAR T cells (6, 12, 13). Although many therapies are available for MM, patients almost always relapse, even after treatment with newer monoclonal antibody or CAR T cell treatments (14). There are also significant toxicities with most currently available therapies such as neurologic toxicity with proteosome inhibitors and cytokine-release syndrome with CAR T cells (15, 16). Most established multiple myeloma therapies are nonspecific. New therapies are needed to increase the median survival time of patients with MM, and our work aims to aid in the development of new more specifically targeted therapies for multiple myeloma.

RNA sequencing (RNAseq) is a technique used to measure the expression of genes in cells (17, 18). Differential gene expression analysis can be used to compare the expression of genes across different samples of cells (17, 19). We used data from cell lines and primary cells in this work.

We hypothesized that differences in gene expression between MM cells and normal healthy cells could be used to identify new targets for MM therapies. Because expression of a therapeutic target in normal cell could lead to damage of normal cells by the therapy, a preferable target would have high levels of gene expression in MM cells but limited to no expression in normal cells. We tested this hypothesis by performing differential gene expression analysis with publicly available data on cell lines. Next, we evaluated RNAseq data from primary MM cells and assessed gene expression in normal human tissues. We identified 19 genes that were expressed in MM cells but had limited expression in normal human tissues. The proteins encoded by these 19 genes may be potential targets for MM therapies.

https://doi.org/10.59720/23-146

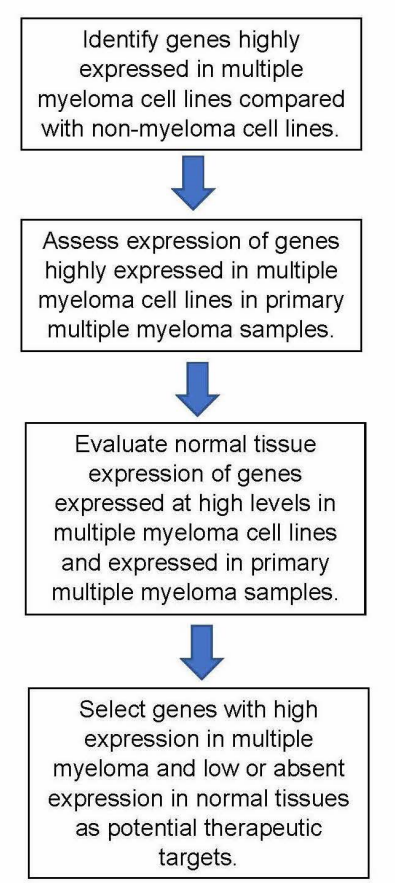

Figure 1. Flowchart of overall project. The flowchart shows the four main steps of the project. The aim of the project was to identify genes with high expression in MM and low or absent expression in normal tissues because such genes encode proteins that might be suitable targets for MM therapies. Genes were assessed by RNAseq performed on cell lines, primary MM cells, and normal human tissues.

RESULTS

Differential gene expression analysis

We tested our hypothesis by following a four-step process that involved analyzing publicly available RNAseq data from cell lines, primary MM cells, and normal human tissues (**Figure 1**). We performed differential gene expression analysis on RNAseq data from cell lines to identify genes with higher expression in MM cell lines versus other types of cancer cell lines. We downloaded cell line RNAseq gene level count data from the Cancer Cell Line Encyclopedia (CCLE) from the DepMap website (20, 21). The four MM cell lines used in this analysis were selected for similarity of gene expression with primary MM cells. This selection was based on a comparison of gene expression of MM cell lines to primary MM cells performed by Sarin et al. (22). MM cell lines selected were MM.1S, EJM, KMS34, and Molp2 (21, 22). Gene expression of the MM cell lines was compared with gene expression of four cell lines derived from cancers other than MM (non-myeloma cell lines). These four non-myeloma cell lines were from a variety of important organs. HepG2 was from liver cancer; A549 was from lung cancer; MOLT3 was from T-cell leukemia; U251MG was from brain cancer (21). By using the programming language R, differential gene expression analysis was performed with the DESeq2 program to compare gene expression in MM cell lines and non-myeloma cell lines.

DESEq2 output from the 20 genes most highly expressed in MM cell lines compared with non-myeloma cell lines is shown (**Table 1**). These results provide an example of the large differences in gene expression between multiple myeloma cell lines and non-myeloma cell lines. DESeq2-normalized gene counts of non-MM cell lines (HepG2, A549, MOLT3, U251MG) and MM cell lines (MM1S, EJM, KMS34, Molp2) are shown (**Table 1)**. The genes with the greatest difference in expression between MM cell lines and non-myeloma cell lines were the 250 genes with the highest expression in MM cell lines versus non-myeloma cell lines and the 250 genes with the lowest expression in MM cell lines versus non-myeloma cell lines. For these genes with the greatest difference in expression between MM cell lines and non-myeloma cell lines, we summarized the adjusted *p*-values and the log2 fold-change for the comparison of the normalized gene counts for MM cell lines and non-myeloma cell lines (**Figure 2**).

We determined the 250 genes with the highest expression in MM cell lines versus non-myeloma cell lines and the 250 genes with the highest expression in non-myeloma cell lines relative to MM cell lines (**Figure 3A, B**). We performed these analyses to identify genes with the largest fold-change in expression between MM cell lines and non-myeloma cell lines. Genes with the highest expression in MM cell lines relative to non-myeloma cell lines are of interest because this expression pattern might allow specific targeting of MM cells while sparing non-myeloma cells from damage.

Assessment of gene expression in primary MM cells

In the second step of the project, genes found in the DESeq2 analysis to be expressed at much higher levels in MM cell lines compared with non-myeloma cell lines were further evaluated for expression in primary human MM cells

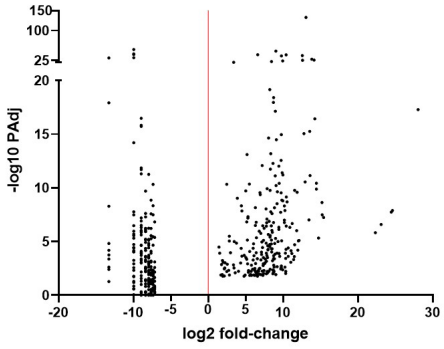

Figure 2. DESeq2 output depicting statistical parameters of the most differentially expressed genes between MM cell lines and non-myeloma cell lines. Log2 fold-change is on the x-axis, and -log10 *p*-value adjusted for multiple comparisons (PAdj) is on the y-axis. The group of dots to the left of the red line represents the 250 genes with the lowest expression in MM cell lines versus non-myeloma cell lines. The group of dots to the right of the red line represents the 250 genes with the highest expression in MM cell lines versus non-myeloma cell lines.

https://doi.org/10.59720/23-146

				Non-myeloma cell lines*				Multiple myeloma cell lines*			
Gene name	Fold Change	log2Fold-Change	PAdj	HepG2	A549	MOLT3	U251MG	MM.1S	EJM	KMS34	Molp2
CCL3L3	275318353	28	5.2E-18	0	0	0	0	10725	6000	0	54
AC097512.1	25784905	24.6	1.3E-08	0	0	0	0	0	1260	0	193
OR7E83P	23121148	24.5	1.8E-08	0	0	0	0	1189	106	0	0
AC234778.1	9004301	23.1	2.6E-07	0	0	0	0	0	449	0	39
AC093791.1	5275967	22.3	1.5E-06	0	0	0	0	0	72	0	740
IGLL5	43326	15.4	5.7E-08	0	0	5	0	2E+05	7	9630	57127
IGLJ1	38211	15.2	3.2E-08	1.3	0	63	1	3E+05	3678	714	2E+06
IGHG2	37505	15.2	2.3E-09	0	0	0	0	18	3606	7135	16401
NLRP7	27380	14.7	4.7E-06	0	0	0	0	7898	11370	2	559
TRDC	22354	14.4	1.3E-10	0	0	1	0	30	7203	4203	15255
IGHG4	22353	14.4	3.7E-11	0	0	0	0	69	3427	4008	8684
CCL3	19377	14.2	3.7E-17	1	0	1	0	28535	17186	478	6771
TNFRSF17	18059	14.1	1.5E-26	0	0	3	0	10735	31470	6848	10777
ITGB7	14472	13.8	1.1E-28	8	56	17	4	7E+05	3E+05	2E+05	12551
COX5BP7	12445	13.6	7.1E-12	0	0	0	0	747	5444	2593	229
JCHAIN	12086	13.6	5.5E-16	4	6	54	2	1E+05	2E+05	714	5E+05
TBC1D27P	11212	13.5	9.7E-08	0	0	0	0	6254	35	1085	746
IRF4	8544	13.1	1E-133	11	4	12	8	55087	86702	54714	1E+05
TNFRSF13B	7892	12.9	2.8E-11	0	0	1	1	11441	163	1720	1634
P2RY10	7047	12.8	8.9E-16	0	0	1	1	5508	564	1366	5985

Table 1. Example of raw DESeq2 output. As an example of DESeq2 output, the results of the 20 genes with the highest expression in MM cell lines versus non-myeloma cell lines are shown. Gene names are in the first column. Fold-change and log2 fold-change in gene counts between MM cell lines and non-myeloma cell lines are shown. The *p*-values adjusted for multiple comparisons (PAdj, Benjamini-Hochberg approach) are shown. The DESeq2-normalized gene counts for non-myeloma cell lines (HepG2, A549, MOLT3, U251MG) and MM cell lines (MM.1S, EJM, KMS34, Molp2) are shown. *The units in all columns under "Non-myeloma cell lines and "Multiple myeloma cell lines" are normalized gene counts.

from patients. This was done to confirm that genes highly expressed in MM cell lines are also expressed in primary MM cells from patients. We used primary MM cell RNAseq data derived from bone marrow samples of 25 patients (23). These data originated from a study by Alameda et al. in which different plasma cell malignancies, including multiple myeloma, were investigated (23). These publicly available data were from the National Center for Biotechnology Information Gene Expression Omnibus (GEO) website (24). By necessity, this evaluation of primary MM was limited to genes that were listed in both the cell line and primary MM data sets. Among genes included in both data sets, we assessed the 200 genes most highly expressed in MM cell lines versus non-myeloma cell lines for expression in primary MM patient samples. Since our goal was to determine if the genes were expressed in primary MM cells, we defined a gene as expressed in primary MM if it had a read count of 10 or greater in at least 23 of 25 samples assessed. Expression in at least 23 of 25 primary MM samples indicates that a gene is consistently expressed in MM cells of patients (**Figure 4**). Of the top 200 genes most highly expressed in MM cell lines relative to non-myeloma cell lines, 72 genes were also expressed in primary MM samples (**Table 2**). These genes have a preferable expression pattern for the target of a specific therapy for MM.

Assessment of gene expression in normal human tissues

In the third step of the project, we assessed expression of the 72 genes found to be highly expressed by MM in normal human tissues. For this analysis, we used publicly available human normal tissue RNAseq data from the Genotype-Tissue Expression Project (GTEx) website (25). We assessed gene expression in 53 normal tissues. For each normal tissue assessed, the GTEx data that we used were the median transcripts per million (TPM) of samples from multiple subjects. We considered a gene to have restricted normal tissue expression if the gene met two criteria. First, the gene had to be expressed at a median level of 30 TPM or greater in no more than two normal tissues. Second, the gene had to have a median of less than 5 TPM in kidney, brain, and heart tissues. The expression limit was lower for kidney, brain, and heart since these are particularly critical organs. We set these criteria arbitrarily but with an interest in selecting genes with a limited tissue distribution and with very low expression in the most critical organs.

Figure 3. Differences in gene expression by MM cell lines relative to non-myeloma cell lines. Heats maps represent DESeq2-normalized gene counts of four MM cell lines and four non-myeloma cell lines. From top to bottom, genes are ranked by log2 fold-change. The gene counts were normalized with GraphPad Prism for display on the heat maps. (A) Heat map shows the 250 genes most highly expressed in MM cell lines compared with non-myeloma cell lines. (B) Heat map shows the 250 genes most highly expressed in non-myeloma cell lines compared with MM cell lines. Color scales to the right of the heat maps indicate the level of gene expression. Yellow represents highly expressed genes while blue represents genes with low expression. Non-MM cell lines were HepG2, A549, MOLT3, and U251MG. MM cell lines were MM1S, EJM, KMS34, and Molp2.

Nineteen of the genes determined to be highly expressed in MM cells also had a restricted expression pattern in normal human tissues (**Table 3**). Expression of the tumor necrosis factor receptor superfamily member 13B (*TNFRSF13B*) gene in normal human tissues is shown as an example of a gene with a restricted normal tissue expression pattern (**Figure**

https://doi.org/10.59720/23-146

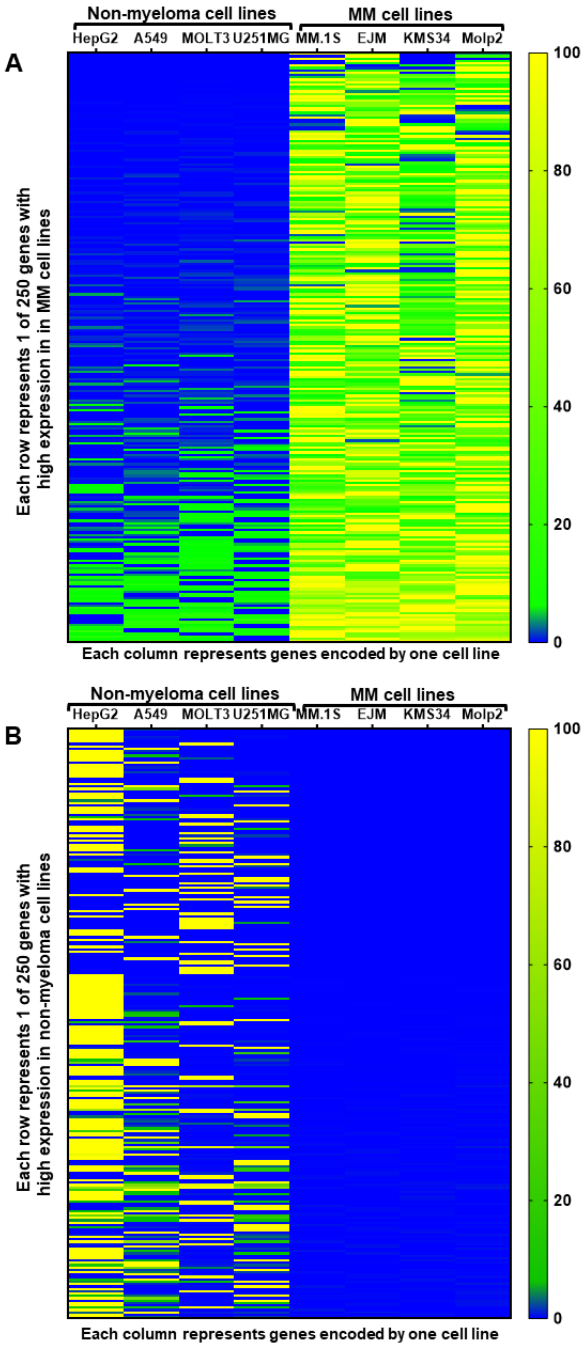

Figure 3. Differences in gene expression by MM cell lines relative to non-myeloma cell lines. Heats maps represent DESeq2-normalized gene counts of four MM cell lines and four non-myeloma cell lines. From top to bottom, genes are ranked by log2 fold-change. The gene counts were normalized with GraphPad Prism for display on the heat maps. (A) Heat map shows the 250 genes most highly expressed in MM cell lines compared with non-myeloma cell lines. (B) Heat map shows the 250 genes most highly expressed in non-myeloma cell lines compared with MM cell lines. Color scales to the right of the heat maps indicate the level of gene expression. Yellow represents highly expressed genes while blue represents genes with low expression. Non-MM cell lines were HepG2, A549, MOLT3, and U251MG. MM cell lines were MM1S, EJM, KMS34, and Molp2.

5A). *TNFRSF13B* was not expressed at a median level of 30 TPM or more in any normal tissue. The median TPM of the Copine 5 (*CPNE5*) gene in each normal tissue is shown as an example of a gene with widespread expression in important tissues (**Figure 5B**). *CPNE5* has expression above five TPM in the heart and brain, among other tissues; thus, we eliminated the protein encoded by this gene as a potential therapeutic target due to the gene expression in tissue samples from critical organs.

Selection of genes with a preferential gene expression pattern for targeted therapies

In the final step, we narrowed down the genes identified in previous steps to a pool of genes that fit all three criteria. First, the gene was expressed at high levels in MM cell lines while also being expressed at low levels in non-myeloma cell lines. Second, the gene was expressed in primary MM cells. Finally, the gene had limited expression in normal human tissues. We narrowed the data to a group of 19 genes fitting all criteria (**Table 3**). These 19 genes that are expressed in MM cells but expressed at low levels in normal human tissues might be appropriate targets for new specific MM therapies.

DISCUSSION

MM is a cancer that is almost always incurable with current therapies, so new therapies are needed for MM. Our hypothesis that differential gene expression analysis could be used to identify potential targets for MM therapies was shown to be supported by our analyses and results. We found, using differential gene expression analysis, 19 genes with a favorable expression pattern that could result in novel targeted therapies for MM.

Our approach identified two genes encoding proteins that are already targets of MM therapies approved by the United States Food and Drug Administration (FDA), which suggests that our approach was valid. One gene that we identified as a promising therapeutic target for MM was *SLAMF7*. *SLAMF7* encodes a protein that is a surface antigen and good marker of normal plasma cells and MM cells (26). The *SLAMF7*

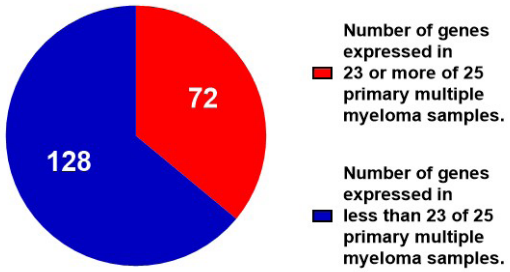

Figure 4. Assessment of gene expression in primary MM cells. The 200 genes most highly expressed in MM cell lines versus non-myeloma cell lines as ranked by log2 fold-change were assessed for expression in primary multiple myeloma samples. Primary MM cell RNAseq data were obtained from bone marrow MM samples of patients. A primary MM sample was counted as expressing a gene if the sample had a raw gene count of 10 or higher. Genes meeting this threshold for expression in at least 23/25 primary MM samples were counted as being expressed in primary MM cells.

https://doi.org/10.59720/23-146

Gene name	Abbreviation	Gene name	Abbreviation
Immunoglubulin heavy constant gamma 4	*IGHG4*	Tumor necrosis factor receptor superfamily 17	*TNFRSF17*
Immunoglubulin heavy constant gamma 1	*IGHG1*	Tumor necrosis factor receptor superfamily 13B	*TNFRSF13B*
Immunoglobulin kappa constant	*IGKC*	Integrin subunit beta 7	*ITGB 7*
Chemokine (C-C motif) ligand 3	*CCL3*	Regulator of G protein signaling 1	*RGS1*
Immunoglubulin heavy constant gamma 2	*IGHG2*	Prepronociceptin	*PNOC*
Interferon Regulatory Factor 4	*IRF4*	Potassium Calcium-Activated Channel Subfamily N Member 3	*KCNN3*
Signaling lymphocytic activation molecule 7	*SLAMF7*	Pleckstrin	*PLEK*
Cytohesion 1 interacting protein	*CYTIP*	Neutrophil Cytosolic Factor 4	*NCF$*
Fc receptor like 5	*FCRL5*	Dual Adaptor Of Phosphotyrosine And 3-Phosphoinositides 1	*DAPP1*
Copine 5	*CPNE5*	Wingless-type mouse mammary tumor virus integration site family member 10A	*WNT10A*
Chemokine (C-C motif) ligand 4	*CCL4*	Sodium Channel Epithelial 1 Subunit Beta	*SCNN1B*
Insulin like growth factor 1	*IGF1*	B cell linker	*BLNK*
Fc receptor like A	*FCRLA*	Colony Stimulating Factor 2 Receptor Subunit Beta	*CSF2RB*
Z-DNA binding protein 1	*ZBP1*	Bruton tyrosine kinase	*BTK*
Cluster of differentiation 79b antigen	*CD79b*	Translation Machinery Associated 16 Homolog	*TMA16*
Glycophorin C	*GYPC*	POU class 2 associating factor 1	*POU2AF1*
Spi-1 proto-oncogene	*SPI1*	Neutrophil Cytosolic Factor 1B Pseudogene	*NCF1B*
Arachidonate 5-Lipoxygenase	*ALOX5*	G Protein-Coupled Receptor Class C Group 5 Member D	*GPRC5D*
Translocase Of Outer Mitochondrial Membrane 7	*TOMM7*	ADAM Metallopeptidase Domain 8	*ADAM8*
Pim-2 Proto-Oncogene, Serine/Threonine Kinase	*PIM2*	POU Class 2 Homeobox 2	*POU2F2*
DEP Domain Containing MTOR Interacting Protein	*DEPTOR*	Docking Protein 3	*DOK3*
NLR Family Pyrin Domain Containing 1	*NLRP1*	Thymocyte Selection Associated Family Member 2	*THEMIS2*
Interleukin 15 Receptor Subunit Alpha	*IL15RA*	Transmembrane Protein 238	*TMEM238*
SH2B Adaptor Protein 2	*SH2B2*	Tumor Protein P53 Inducible Nuclear Protein 1	*TP53INP1*
Aquaporin 3 (Gill Blood Group)	*AQP3*	Vitamin D Receptor	*VDR*
Glucocorticoid Induced 1	*GLCCI1*	Thioredoxin Domain Containing 11	*TXNDC11*
Thioredoxin Domain Containing 5	*TXNDC5*	Beta-1,4-Galactosyltransferase 3	*B4GALT3*
BTG Anti-Proliferation Factor 2	*BTG2*	FKBP Prolyl Isomerase 11	*FKBP11*
Solute Carrier Family 25 Member 45	*SLC25A45*	Interferon Alpha And Beta Receptor Subunit 2	*IFNAR2*
ST6 N-Acetylgalactosaminide Alpha-2,6-Sialyltransferase 4	*ST6GALNAC 4*	CLPTM1 Like	*CLPTM1L*
Myocyte Enhancer Factor 2D	*MEF2D*	Torsin Family 3 Member A	*TOR3A*
Phosphomevalonate Kinase	*PMVK*	SUB1 Regulator Of Transcription	*SUB1*
Proteasome 20S Subunit Beta 8	*PSMB8*	Protein Phosphatase 3 Catalytic Subunit Gamma	*PPP3CC*
Signal Sequence Receptor Subunit 3	*SSR3*	Zinc Finger Protein 19	*ZNF19*
Vascular Endothelial Zinc Finger 1	*VEZF1*	Cysteine And Glycine Rich Protein	*CSRP1*
Leptin Receptor Overlapping Transcript Like 1	*LEPROTL1*	Ras Related Dexamethasone Induced 1	*RASD1*

Table 2. Genes with high expression in MM cell lines and primary MM cells. The names and abbreviations of 72 genes highly expressed in MM cell lines compared with non-myeloma cell lines and expressed in at least 23/25 primary MM samples are displayed.

https://doi.org/10.59720/23-146

Gene name	Abbreviation	Gene name	Abbreviation
C-C motif chemokine ligand 3	*CCL3*	Tumor necrosis factor receptor superfamily 17	*TNFRSF17*
Interferon Regulatory Factor 4	*IRF4*	Tumor necrosis factor receptor superfamily 13B	*TNFRSF13B*
Signaling lymphocytic activation molecule 7	*SLAMF7*	Integrin subunit beta 7	*ITGB 7*
Fc receptor like 5	*FCRL5*	Bruton tyrosine kinase	*BTK*
Insulin like growth factor 1	*IGF1*	Prepronociceptin	*PNOC*
Fc receptor like A	*FCRLA*	Dual Adaptor Of Phosphotyrosine And 3-Phosphoinositides 1	*DAPP1*
Z-DNA binding protein 1	*ZBP1*	Colony Stimulating Factor 2 Receptor Subunit Beta	*CSF2RB*
G Protein-Coupled Receptor Class C Group 5 Member D	*GPRC5D*	POU class 2 associating factor 1	*POU2AF1*
Neutrophil Cytosolic Factor 1B Pseudogene	*NCF1B*	Wingless-type family member 10A	*WNT10A*
Zinc Finger Protein 19	*ZNF19*		

Table 3. Genes with high expression in MM and low expression in normal tissues. All of the listed genes have high expression in MM and low expression in normal tissues. This is the preferred expression pattern for potential therapeutic targets.

protein is the target of an FDA-approved monoclonal antibody therapy (27). The second gene identified by our approach that encodes a protein already targeted by FDA-approved therapies is *TNFRSF17*, which encodes the B-cell maturation antigen protein (13, 28).

Integrin Beta 7 (*ITGB7*) encodes a cell-surface protein that might be a suitable target for monoclonal antibody or chimeric antigen receptor (CAR) T-cell therapies for MM (29). The Fc receptor like 5 (*FCRL5*) and *TNFRSF13B* genes also encode cell-surface proteins, which might be targeted by monoclonal antibodies or CAR T cells (30, 31). Expression of TNFRSF13B in the spleen is due to expression of this gene by normal B cells and plasma cells (**Figure 5A**) (32). Notably, these genes are under evaluation by other investigators as therapeutic targets for MM (29–31).

The wingless-type mouse mammary tumor virus integration site family member 10A (WNT10A) protein is a member of the WNT gene family (33). This gene is expressed in different forms of lymphoma and may be implicated in oncogenesis (33). WNT10A signaling might be inhibited by blocking post-translational modification or receptor binding (33).

Interferon regulatory factor 4 (IRF4) is a transcription factor that interacts with many different proteins (34). One important protein that IRF4 interacts with is transcription factor PU.1 (34). Interestingly, the Spi1 proto-oncogene (*SPI1*), which encodes PU.1, was one of the 72 genes that we found to be expressed at high levels in MM cells. This offers support for IRF4 and PU.1 being present in the same cells (**Table 2**). It is conceivable that a therapy could be designed to block the interaction between IRF4 and PU.1. Another gene with a favorable expression pattern was *POU2AF1*. The POU2AF1 protein is also known as BOB1 (35). The *POU2AF1* gene encodes a transcriptional cofactor that is important for expression of many genes critical for B-lineage cells (35).

We used cell line RNAseq data for differential gene expression analysis. It would have been preferable to use primary MM cells and primary normal human tissue RNAseq data for this step since the goal was to find targets for therapies that could eliminate primary MM cells while not harming normal human cells. However, we did not have access to primary human MM and normal human tissue RNAseq data that was generated by the same methods; therefore, performing differential gene expression analysis would not have been valid. We set arbitrary criteria of restricted normal tissue gene expression. These criteria were 1) no more than two tissues could have a median gene expression of greater than 30 TPM, and 2) gene expression of heart, kidney, or brain could be no more than a median of 5 TPM. These arbitrary criteria could be relaxed to increase the number of genes determined to be suitable therapeutic targets. Alternatively, the criteria could be tightened by lowering the allowed median TPM levels to decrease the number of genes determined to be suitable therapeutic targets.

Future steps for this work could include assessing more than 200 genes for expression in primary MM samples and normal human tissues. Assessing a larger number of genes would increase the chance of finding genes with the appropriate expression pattern. Another future project could include studying protein expression of promising genes in primary MM cells and normal human tissues. Finally, future steps for this work could involve projects aimed at developing methods of targeting the proteins encoded by the 19 genes with promising expression patterns. Such methods could be monoclonal antibodies for cell-surface proteins or small molecule drugs to target some intracellular proteins. Development of any therapy targeting one of the 19 genes that we identified would be a large project involving generation of a monoclonal antibody or small molecule drug followed by testing *in vitro* and possibly *in vivo* in animals such as mice.

Our approach to finding therapeutic targets could be applied to other types of cancer. This could be accomplished by finding cell lines in the CCLE derived from any type of cancer and then comparing gene expression of these cell lines to control cell lines derived from other cell types. Subsequently, candidate genes could be assessed for expression in the appropriate type of primary cancer cells and in normal human tissues.

Many cancer therapies cause toxicity by damaging normal tissues. We showed that gene expression analysis

https://doi.org/10.59720/23-146

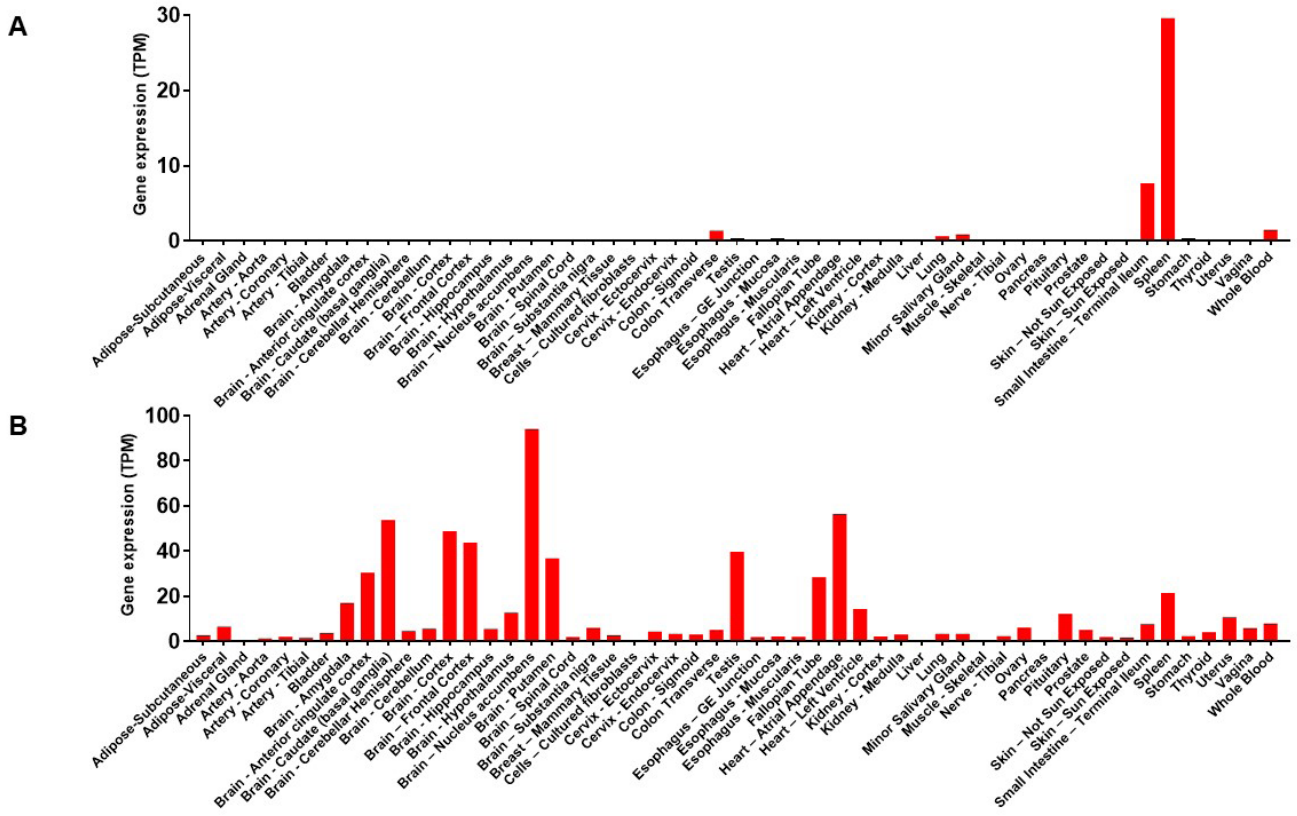

Figure 5. Normal tissue expression of genes with restricted and extensive patterns of expression. (A) *TNFRSF13B* gene expression in transcripts per million (TPM) is shown for major human tissues. *TNFRSF13B* had a restricted expression pattern. (B) *CPNE5* gene expression in TPM is shown for major human tissues. *CPNE5* had a widespread expression pattern. These genes were selected to show one gene with a restricted expression pattern (A) and a second gene with a widespread expression pattern (B). In Esophagus-GE Junction, GE stands for gastroesophageal junction.

can be used to identify genes encoding proteins that might be suitable targets for new cancer therapies that selectively target cancer cells but not normal cells. Our findings could potentially lead to new targeted therapies for MM.

MATERIALS AND METHODS

Differential gene expression analysis with DESeq2

By using RNASeq data from the CCLE, gene expression was compared in multiple myeloma cell lines versus cell lines derived from cancers other than MM (non-myeloma cell lines) (21). Cell line RNAseq gene-level count data from the CCLE was downloaded from the DepMap website (20,21). The specific data used was the 2022 Quarter 2 release of the CCLE RNAseq data. These data were downloaded in a file called CCLE_RNAseq_reads.csv (20). RNAseq was carried out by CCLE investigators (21). Gene-level read counts were calculated; details are available in Ghandi et al. (21). Gene-level RNAseq fragment counts were analyzed. A comma-separated values (csv) file that contained RNASeq gene counts of all genes included in the CCLE data was prepared. Data from eight cell lines: four MM cell lines and four non-myeloma cell lines was included. The four MM cell lines utilized were MM.1S, EJM, KMS34, and Molp2 (21). The four non-myeloma cell lines were HepG2 (liver cancer), A549 (lung cancer), MOLT-3 (T-cell leukemia), and U251MG (brain cancer) (21). To analyze the RNASeq data, R version 4.3.0 was run on RStudio. Instructions were followed from *RNA-SEQ by Example*, a biostar handbook (Updated December 11, 2021) (36). The DESeq2 package from Bioconductor was run to normalize the RNAseq gene count data and perform a statistical analysis. The statistical analysis compared the normalized gene counts of the four MM cell lines and the four non-myeloma cell lines. Statistical analyses performed by DESeq2 included a test of statistical significance with correction for multiple comparisons by the Benjamini-Hochberg approach and the log2 fold-change between the normalized gene counts of the MM cell lines and non-myeloma cell lines.

Evaluation of RNA transcript counts in primary MM samples

We obtained RNAseq data from primary MM samples from the GEO website (accession #GSE175384) (24). The primary MM data that we evaluated came from primary MM samples from 25 different patients. Plasma cells were purified from the 25 MM patients, and RNAseq was performed on mRNA from the plasma cells by Alameda and coworkers. Details are available at Alameda et al. (23). The data that we assessed was gene-level RNAseq data.

https://doi.org/10.59720/23-146

A list of genes that were present in both the MM cell line data set and the primary MM data set was selected. From the DESeq2 results comparing MM cell lines and non-myeloma cell lines, the 200 genes with the highest expression in MM cell lines versus non-myeloma cell lines as defined by log2 fold-change were selected. The name of each of these 200 genes was searched for in the primary MM data set. If a gene had a read count higher than 10 in a sample, that gene was considered expressed in that sample. If the gene was expressed in 23 out of 25 or more of the samples that we assessed, the gene was designated as expressed in primary MM.

Evaluation of normal tissue gene expression

Expression of genes in normal human tissues was evaluated using RNAseq data from the GTEx website (25). Sequencing was performed by investigators of GTEx, and RNAseq reads were aligned to the human genome (37). Gene-level quantification was performed, and results were presented as TPM. From the GTEx site, Expression then Search Gene/Transcript Expression were selected. This led to a text box where the gene abbreviations could be entered to obtain the RNAseq gene counts for the selected gene. RNAseq gene counts were expressed as TPM for each of 53 different tissues. For each tissue, median TPM were determined from the results of multiple subjects. TPM is a standard method for reporting RNAseq count data that normalizes for mapped read (transcript) frequency and gene length (38). The specific GTEx data was GTEx Analysis Release V8 (dbGaP Accession phs000424.v8p2). Gene counts were provided for all major human tissues.

Graphical representation

The volcano plot, heat maps, the pie chart, and bar graphs were made with Graph Pad Prism version 8.4.3. Values on the heat map were normalized by Prism with final values presented as percentages. By the normalization process, the smallest transcript count for each gene was defined as 0% and the largest transcript count for each gene was defined as 100%.

Received: July 8, 2023
Accepted: October 12, 2023
Published: April 26, 2024

REFERENCES

1. Krishnamurthi, S.S. et al. "Chemotherapy-associated diarrhea, constipation and intestinal perforation: Pathogenesis, risk factors, and clinical presentation." *UpToDate*, www.uptodate.com/contents/chemotherapy-associated-diarrhea-constipation-and-intestinal-perforation-pathogenesis-risk-factors-and-clinical-presentation?search=chemotherapy%-20toxicities&source=search_result&selectedTitle=8~150&usage_type=default&display_rank=8. Accessed 5 July 2023.
2. Maldonado, F. et al. "Pumonary toxicity associated with antineoplastic therapy: Cytotoxic agents." *UpToDate*, www.uptodate.com/contents/pulmonary-toxicity-associated-with-antineoplastic-therapy-cytotoxic-agents?search=chemotherapy%20toxicity%20&source=search_result&selectedTitle=6~150&usage_type=default&display_rank=6. Accessed 5 July 2023.
3. Kumar, S. "Multiple myeloma: Use of hematopoietic cell transplantation." *UpToDate*, www.uptodate.com/contents/multiple-myeloma-use-of-hematopoietic-cell-transplantation?search=autologous%20stem%20cell%20transplant%20toxicity%20&source=search_result&selectedTitle=1~150&usage_type=default&display_rank=1. Accessed 5 July 2023.
4. "Targeted therapy to treat cancer." *National Cancer Institute, National Institutes of Health*, www.cancer.gov/about-cancer/treatment/types/targeted-therapies. Accessed 5 July 2023.
5. June, C.H. et al. "Chimeric antigen receptor therapy." *New England Journal of Medicine*, vol. 379, no. 1, 2018, pp. 64-73, doi:10.1056/NEJMra1706169.
6. Albagoush, S.A. et al. "Multiple myeloma." *Statpearls, StatPearls Publishing*. Copyright © 2023, StatPearls Publishing LLC., 2023.
7. "Facts and figures 2023." American Cancer Society www.cancer.org/content/dam/cancer-org/research/cancer-facts-and-statistics/annual-cancer-facts-and-figures/2023/2023-
8. Laubach, J.P. " Multiple myeloma: Staging and prognostic studies." *UpToDate*, www.uptodate.com/contents/multiple-myeloma-staging-and-prognostic-studies?search=multiple%20myeloma%20prognosis&source=search_result&selectedTitle=1~150&usage_type=default&display_rank=1. Acessed 5 July 2023.
9. Binder, M. et al. "Mortality trends in multiple myeloma after the introduction of novel therapies in the united states." *Leukemia*, vol. 36, no. 3, 2022, pp. 801-808, doi:10.1038/s41375-021-01453-5.
10. Dimopoulos, M.A. et al. "Treatment options for patients with heavily pretreated relapsed and refractory multiple myeloma." *Clinical Lymphoma, Myeloma and Leukemia,* vol. 22, no. 7, 2022, pp. 460-473, doi:10.1016/j.clml.2022.01.011.
11. Stalker, M.E. et al. "Clinical management of triple-class refractory multiple myeloma: A review of current strategies and emerging therapies." *Current Oncology*, vol. 29, no. 7, 2022, pp. 4464-4477, doi:10.3390/curroncol29070355.
12. Rajkumar, S.V. et al. "Multiple myeloma current treatment algorithms." *Blood Cancer Journal*, vol. 10, no. 9, 2020, doi:10.1038/s41408-020-00359-2.
13. Berdeja, J.G. et al. "Ciltacabtagene autoleucel, a b-cell maturation antigen-directed chimeric antigen receptor t-cell therapy in patients with relapsed or refractory multiple myeloma (cartitude-1): A phase 1b/2 open-label study." *The Lancet*, vol. 398, no. 10297, 2021, pp. 314-324, doi:10.1016/S0140-6736(21)00933-8.
14. Kumar, S. et al. "Gaps and opportunities in the treatment of relapsed-refractory multiple myeloma: Consensus recommendations of the nci multiple myeloma steering committee." *Blood Cancer Journal*, vol. 12, no. 6, 2022, doi:10.1038/s41408-022-00695-5.
15. Markouli, M. et al. "Toxicity profile of chimeric antigen receptor t-cell and bispecific antibody therapies in multiple myeloma: Pathogenesis, prevention and management." *Current Oncology*, vol. 30, no. 7, 2023, pp. 6330-6352, doi:10.3390/curroncol30070467.
16. Was, H. et al. "Mechanisms of chemotherapy-induced

https://doi.org/10.59720/23-146

neurotoxicity." *Frontiers in Pharmacology*, vol. 13, 2022, doi:10.3389/fphar.2022.750507.
17. Marioni, J.C. et al. "Rna-seq: An assessment of technical reproducibility and comparison with gene expression arrays." *Genome Research*, vol. 18, no. 9, 2008, pp. 1509-1517, doi:10.1101/gr.079558.108.
18. Wang, Y. et al. "Changing technologies of rna sequencing and their applications in clinical oncology." *Frontiers in Oncology*, vol. 10, 2020, doi:10.3389/fonc.2020.00447.
19. Kukurba, K.R. et al. "Rna sequencing and analysis." *Cold Spring Harbor Protocols*, vol. 2015, no. 11, 2015, pp. 951-969, doi:10.1101/pdb.top084970.
20. "Depmap portal cancer cell line encyclopedia (ccle)." *Cancer Dependency Portal (DepMap)*. depmap.org/portal/download/all/. Accessed 8/15/2022.
21. Ghandi, M. et al. "Next-generation characterization of the cancer cell line encyclopedia." *Nature*, vol. 569, no. 7757, 2019, pp. 503-508, doi:10.1038/s41586-019-1186-3.
22. Sarin, V. et al. "Evaluating the efficacy of multiple myeloma cell lines as models for patient tumors via transcriptomic correlation analysis." *Leukemia*, vol. 34, no. 10, 2020, pp. 2754-2765, doi:10.1038/s41375-020-0785-1.
23. Alameda, D. et al. "Tumor cells in light-chain amyloidosis and myeloma show distinct transcriptional rewiring of normal plasma cell development." *Blood*, vol. 138, no. 17, 2021, pp. 1583-1589, doi:10.1182/blood.2020009754.
24. "Gene expression omnibus site of primary multiple myeloma data used in this study." *National Center for Biotechnology Information*. www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE175384. Accessed 8/15/2022.
25. "Gtex. Genotype-tissue expression (gtex) project." supported by the U.S. National Institutes of Health gtexportal.org/home. Accessed 5 July 2023.
26. Malaer, J.D. et al. "Cs1 (slamf7, cd319) is an effective immunotherapeutic target for multiple myeloma." *American Journal of Cancer Research*, vol. 7, no. 8, 2017, pp. 1637-1641.
27. Lonial, S. et al. "Elotuzumab therapy for relapsed or refractory multiple myeloma." *New England Journal of Medicine*, vol. 373, no. 7, 2015, pp. 621-631, doi:10.1056/NEJMoa1505654.
28. Granger, K. et al. "Newly approved and forthcoming t-cell-redirecting bispecific antibodies for the treatment of relapsed/refractory multiple myeloma." *Journal of Oncology Pharmacy Practice*, vol. 29, no. 3, 2023, pp. 722-726, doi:10.1177/10781552231154809.
29. Hosen, N. et al. "Expression of activated integrin β7 in multiple myeloma patients." *International Journal of Hematology*, vol. 114, no. 1, 2021, pp. 3-7, doi:10.1007/s12185-021-03162-2.
30. Polson, A.G. et al. "Expression pattern of the human fcrh/irta receptors in normal tissue and in b-chronic lymphocytic leukemia." *International Immunology*, vol. 18, no. 9, 2006, pp. 1363-1373, doi:10.1093/intimm/dxl069.
31. Kampa, M. et al. "The tnfsf members april and baff and their receptors taci, bcma, and baffr in oncology, with a special focus in breast cancer." *Frontiers in Oncology*, vol. 10, 2020, doi:10.3389/fonc.2020.00827.
32. Bossen, C. et al. "Baff, april and their receptors: Structure, function and signaling." *Seminars in Immunology*, vol. 18, no. 5, 2006, pp. 263-275, doi:10.1016/j.smim.2006.04.006.
33. Aguilera, K.Y. et al. "Wnt ligand dependencies in pancreatic cancer." *Frontiers in Cell and Developmental Biology*, vol. 9, 2021, doi:10.3389/fcell.2021.671022.
34. Wong, R.W.J. et al. "Irf4 as an oncogenic master transcription factor." *Cancers*, vol. 14, no. 17, 2022, doi:10.3390/cancers14174314.
35. Gstaiger, M. et al. "A b-cell coactivator of octamer-binding transcription factors." *Nature*, vol. 373, no. 6512, 1995, pp. 360-362, doi:10.1038/373360a0.
36. "Rna-seq by example." *The Biostar Handbook Collection*. www.biostarhandbook.com/.
37. Gtex laboratory methods." *Broad Insitute of MIT and Harvard*. gtexportal.org/home/methods. Accessed 7/7/2023.
38. Zhao, Y. et al. "Tpm, fpkm, or normalized counts? A comparative study of quantification measures for the analysis of rna-seq data from the nci patient-derived models repository." *Journal of Translational Medicine*, vol. 19, no. 1, 2021, doi:10.1186/s12967-021-02936-w.