# *PARP1* and *POLD2* as prognostic biomarkers for multiple myeloma in autologous stem cell transplant

Melissa Thomas,[1] Junan Li,[2,3] Kevan King,[1] Avinash K. Persaud,[2] Ernest Duah,[2] Zachary Vangundy,[2] Craig C. Hofmeister,[4] Jatinder K. Lamba,[5] Aik Choon Tan,[6] Brooke L. Fridley,[6] Ming J. Poi[2,3] and Nathan D. Seligson[1,3,7]

[1]Department of Pharmacotherapy and Translational Research, The University of Florida, Jacksonville, FL; [2]Division of Pharmacy Practice and Science, College of Pharmacy, The Ohio State University, Columbus, OH; [3]Comprehensive Cancer Center, The Ohio State University, Columbus, OH; [4]Department of Hematology and Medical Oncology, Winship Cancer Institute of Emory University, Atlanta, GA; [5]Department of Pharmacotherapy and Translational Research, The University of Florida, Gainesville, FL; [6]Department of Biostatistics and Bioinformatics, H. Lee Moffitt Cancer Center and Research Institute, Tampa, FL and [7]Center for Pharmacogenomics and Translational Research, Nemours Children's Health, Jacksonville, FL, USA

**Correspondence:** N.D. Seligson
nseligson@cop.ufl.edu

**Received:** November 9, 2022.
**Accepted:** February 23, 2023.
**Early view:** March 2, 2023.

**https://doi.org/10.3324/haematol.2022.282399**

**Supplementary Materials**

Supplementary Appendix

Supplementary Figures

- Supplementary Figure 1
- Supplementary Figure 2
- Supplementary Figure 3
- Supplementary Figure 4
- Supplementary Figure 5
- Supplementary Figure 6

Supplementary Tables

- Supplementary Table 1
- Supplementary Table 2
- Supplementary Table 3

**Supplementary Appendix**

*Publicly Available Datasets*

Publicly available data were collected from the Gene Expression Omnibus (GEO, https://www.ncbi.nlm.nih.gov/geo, downloaded November 13, 2021) and the Multiple Myeloma Research Foundation (MMRF, https://research.themmrf.org/, downloaded September 2020) as approved by the University of Florida institutional review board (#IRB202101136). Processed data was collected from GEO using the getGEO function from the R package GEOquery$_{v2.60.0}$. Data from the MMRF was manually downloaded. Data was collected from GEO series GSE13591[1], GSE23113[2], GSE6477[3], GSE5900[4], and GSE2658[5]. All samples, except for disease-free donors and relapsed disease, were collected at diagnosis. GSE13591 contains 158 gene expression profiles. Samples were collected from 5 disease-free donors, 11 patients with monoclonal gammopathy of undetermined significance (MGUS), 133 patients with MM and 9 patients with plasma cell leukemia (PCL). GSE23113 contains 52 gene expression profiles. Samples were collected from 7 patients with MGUS, 39 patients with MM and 6 patients with PCL patients. GSE6477 contains 162 gene expression profiles. Samples were collected from 15 disease-free donors, 22 patients with MGUS, 24 patients with smoldering MM (sMM), 73 patients with MM, and 28 patients with relapsed MM (rMM). GSE5900 contains 72 gene expression profiles. Samples were collected from with 22 disease-free donors, 44 patients with MGUS, and 12 patients with sMM. GSE2658 contains 559 samples collected from patients with newly diagnosed multiple myeloma subsequently treated with high dose therapy and ASCT. Plasma cells were isolated from bone marrow specimens. After RNA processing, samples were hybridized on Affymetrix GeneChip Human Genome U133A Arrays (GSE13591, GSE23113, GSE6477) or Affymetrix Human Genome U133A Plus 2.0 Arrays (GSE5900, GSE2658). Further processing included Affymetrix MicroArray Suite (MAS) 5.0. Robust Multi-array Analysis (RMA) quantile-quantile normalization log2 transformed (GSE13591), MAS normalization using global scaling procedure (GSE23113), MAS normalization using global scaling procedure to a target value of 500, log transformed, median centered, and analyzed using GeneSpring7 (GSE6477), and log-base 2 transformation (GSE5900, GSE2658). Data was collected from the Multiple Myeloma Research Foundation (MMRF) CoMMpass Study NCT01454297 (accessed September 2020). This study contains clinical and genomic data from 1,150 patients from 76 sites worldwide. This data represented real-world outcomes for patients including long-term follow-up. We collected complete clinical and transcriptomic data for 675 patients with newly diagnosed multiple myeloma. For transcriptomic data, plasma cells were isolated from bone

marrow aspirates. After RNA sequencing and processing with Cufflinks, gene transcription expression samples were quantified as fragments per kilobase of transcript per million mapped (FPKM). These data were generated as part of the Multiple Myeloma Research Foundation Personalized Medicine Initiatives (https://research.themmrf.org and www.themmrf.org). Data, including mRNA expression, was collected from the genomics of drug sensitivity in cancer database for cell lines noted as being MM (http://cancerrxgene.org, downloaded August 22, 2019)[6].

*MM Cell Lines*

MM cell lines MM1S and NCI-H929 were obtained from American Type Culture Collection (ATCC; Manassas, VA). Cells were maintained in advanced RPMI1640 (Life Technologies, Carlsbad, CA), supplemented with 5% fetal bovine serum (FBS; Life Technologies) and 1X antibiotic antimycotic solution (Sigma-Aldrich, St Louis, MO), and were grown at 37°C and 5% $CO_2$. These cell lines were regularly authenticated using short tandem repeat polymorphism (STRP) analysis as recommended by ATCC, were mycoplasma free, and used within 6 months of receipt from ATCC. All *in vitro* studies were conducted in at least triplicate and in at least three independent experiments.

*Cell Viability Assay*

MM cells were seeded at 2,000 cells per well for the NCI-H929 cell line and 2,500 cells per well for the MM1S cell line (96 well plates) in 100µL of RPMI1640-5% FBS media and incubated at 37°C and 5% $CO_2$ overnight. Subsequently, cells were treated with various concentrations of melphalan, olaparib, and talazoparib (Sigma-Aldrich) for another 72 hours. Cell viability was assayed using the MTT Cell proliferation Assay Kit (Roche, Indianapolis, IN) following the manufacturer's directions. Melphalan was dosed from 0.5-20µM, olaparib was dosed from 0.05-30µM, and talazoparib was dosed from 0.005-1µM.

*Apoptosis Assay*

MM cells were seeded at 200,000 cells per well (24-well plates) in 500 µL of advanced RPMI1640-5% FBS media and incubated at 37°C and 5% $CO_2$ overnight. Subsequently, cells

were treated with advanced RPMI1640-5% FBS media containing various concentrations of melphalan, olaparib, talazoparib or a combination of drugs (Sigma-Aldrich) for another 24 hours. Cells were washed once with cold FACS buffer and re-suspend in 250µL 1X binding buffer (10X binding buffer diluted in distilled water) to a concentration of 1-5x$10^5$ cells/250µL. Annexin V and PI were used to stain cells per the manufacture's recommendations. Five untreated controls were used: unstained control, Annexin V (no PI) control, PI (no Annexin V) control, dual stained untreated control, and heat-treated control. Cellular apoptosis was measured using the Attune NxT Flow Cytometer and the standard manufacture recommended protocol (ThermoFisher).

*In Vivo Experiments*

MM1S cells (4x$10^6$ cells/100µL) were injected in a PBS/matrigel (1:1) suspension of 100µL in both right and left flanks of 60 nude mice (22 females and 38 males; The Jackson Laboratory, Bar Harbor, ME). Mice were monitored daily until tumors reached 200$mm^2$. Mice were then randomized to receive vehicle control (18 mice, 28 tumors), talazoparib alone (12 mice, 16 tumors), melphalan alone (15 mice, 27 tumors), or melphalan in combination with talazoparib (15 mice, 28 tumors) in a 2:1:2:2 ratio. Melphalan was reconstituted in acid alcohol (5% HCl, 95% EtOH) and then diluted to a final concentration of 30µg/ µL (EtOH of <10%)[7,8]. Melphalan was dosed at 10mg/kg once on day 0 by intravenous injection into the lateral tail vein. Talazoparib was dissolved in 2% DMSO, 0.5% hydroxypropylmethylcellulose, and 0.2% tween 80. This stock solution was then diluted in distilled water to a concentration of 6µg in 100µL. Talazoparib was dosed at 0.2mg/kg delivered by oral gavage twice daily from days -3 to +3. Mice randomized to the control arm were doses with drug-free vehicles for both melphalan and talazoparib. Tumor size and mice weights were measured daily for the full duration of the study. Mice were removed from the study if their weight was <80% of baseline for two days, or if total tumor size per mouse reached >1,600$mm^2$. All animal studies were conducted following the Ohio State University Institutional Animal Care and Use Committee (IACUC) approval.

*Statistical Methods*

All data were analyzed in $R_{v.4.1.1}$ (The R Project for Statistical Computing, https://www.r-project.org) or Graphpad $Prism_{v.9.2.0}$ (GraphPad Software, San Diego, CA). Additional graphics were created with BioRender.com (BioRender, Toronto, Ontario). Members of the BER pathway

were defined by the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway database[9]. Single sample gene set enrichment analysis (ssGSEA) was conducted using the $GSVA_{v1.40.1}$ package using the ssgsea method without final normalization. ssGSEA is an extension of Gene Set Enrichment Analysis (GSEA), which calculates gene set enrichment scores for each sample. Each ssGSEA enrichment score represents the degree of gene set up or down regulated within a sample. Two-group analysis of continuous variables was conducted using the Mann Whitney Wilcoxon test while analysis of continuous variables across multiple groups was conducted using the Kruskal-Wallis test. Heatmaps were constructed using the $pheatmap_{v1.0.12}$ package. Uniform Manifold Approximation and Projection for Dimension Reduction (UMAP) plots were generated using the package $umap_{v0.2.7.0}$. Survival analysis was tested using Cox proportional hazard regression for continuous variables and log-rank tests for categorical variables. Stepwise Cox proportional hazard regression was conducted using the package $My.stepwise.coxph_{v0.1.0}$ using a significance level for entry and stay of 0.1. Survival graphs were created using the Kaplan–Meier estimator. Combination index was calculated using CompuSyn[10] (ComboSyn, Inc, Paramus, NJ) and Combenefit[11] (CRUK Cambridge Institute, Cambridge, UK) using the highest single agent (HSA) model. Changes in tumor growth over time were tested by ANOVA with repeated measures. Unless otherwise stated, two-sided p-values ≤0.05 were considered statistically significant. Adjustment of p-values was conducted using False Discovery Rate.

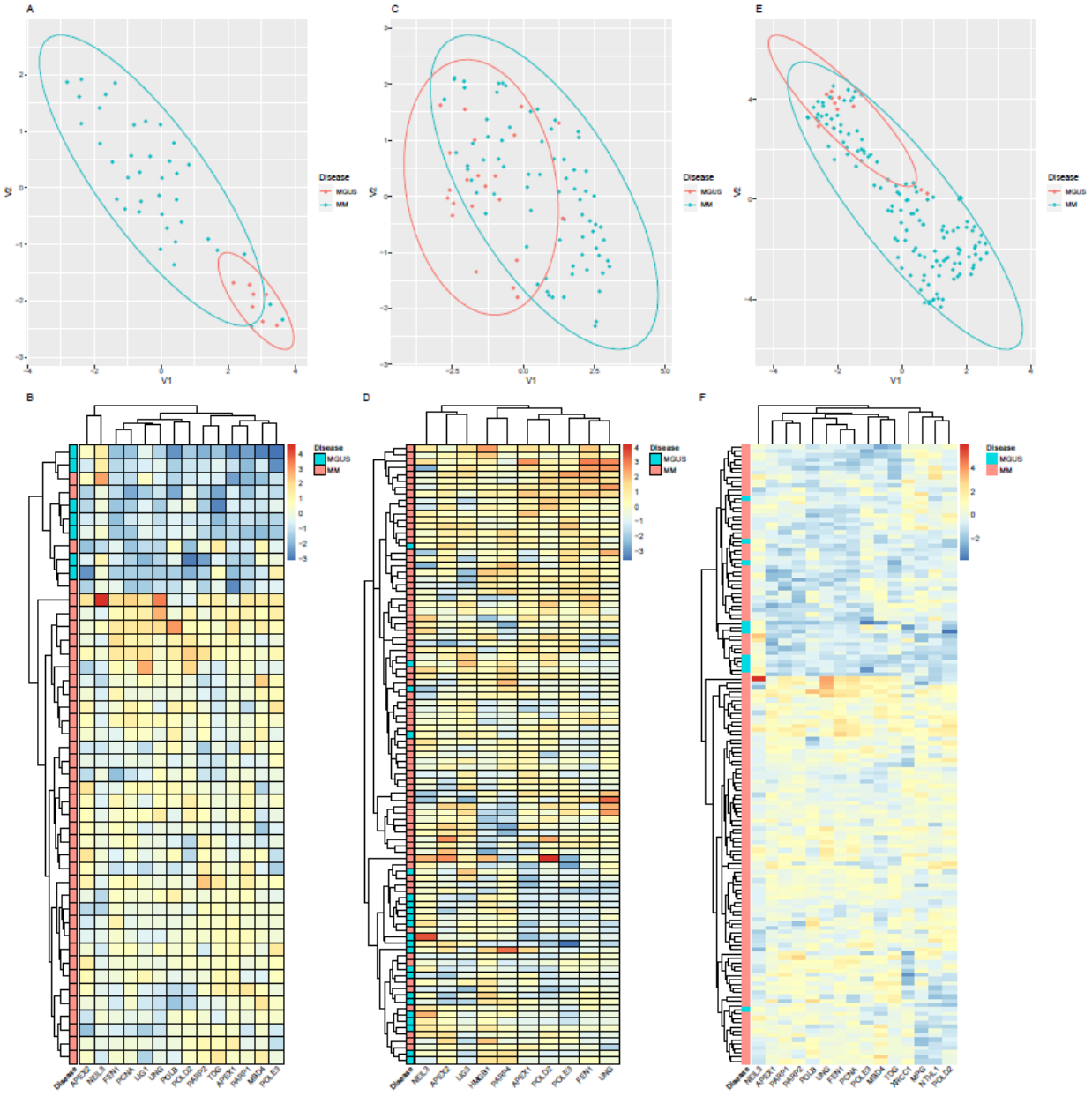
**Supplementary References**


1. Riehn, M., Klopocki, E., Molkentin, M., Reinhardt, R. & Burmeister, T. A BACH2-BCL2L1 Fusion Gene Resulting from a Lymphoma Cell Line BLUE-1. *Cancer* **396**, 389–396 (2011).
2. Mattioli, M. *et al.* Gene expression profiling of plasma cell dyscrasias reveals molecular patterns associated with distinct IGH translocations in multiple myeloma. *Oncogene* **24**, 2461–2473 (2005).
3. Chng, W. J. *et al.* Molecular dissection of hyperdiploid multiple myeloma by gene expression profiling. *Cancer Res* **67**, 2982–2989 (2007).
4. Zhan, F. *et al.* Gene-expression signature of benign monoclonal gammopathy evident in multiple myeloma is linked to good prognosis. *Blood* **109**, 1692–1700 (2007).
5. Zhan, F. *et al.* The molecular classification of multiple myeloma. *Blood* **108**, 2020–2028 (2006).
6. Yang, W. *et al.* Genomics of Drug Sensitivity in Cancer (GDSC): A resource for therapeutic biomarker discovery in cancer cells. *Nucleic Acids Res* **41**, (2013).
7. Nath, K. *et al.* Effects of hyperglycemia on lonidamine-induced acidification and de-energization of human melanoma xenografts and sensitization to melphalan. (2015) doi:10.1002/nbm.3260.
8. Sanchez, E. *et al.* The proteasome inhibitor CEP-18770 enhances the anti-myeloma activity of bortezomib and melphalan. doi:10.1111/j.1365-2141.2009.08008.x.
9. Kanehisa, M., Sato, Y., Kawashima, M., Furumichi, M. & Tanabe, M. KEGG as a reference resource for gene and protein annotation. *Nucleic Acids Res* **44**, D457–D462 (2016).

10. *Synergy of FD and PXT by compusyn quantification*. www.combosyn.com.
11. di Veroli, G. Y. *et al.* Combenefit: An interactive platform for the analysis and visualization of drug combinations. *Bioinformatics* **32**, 2866–2868 (2016).

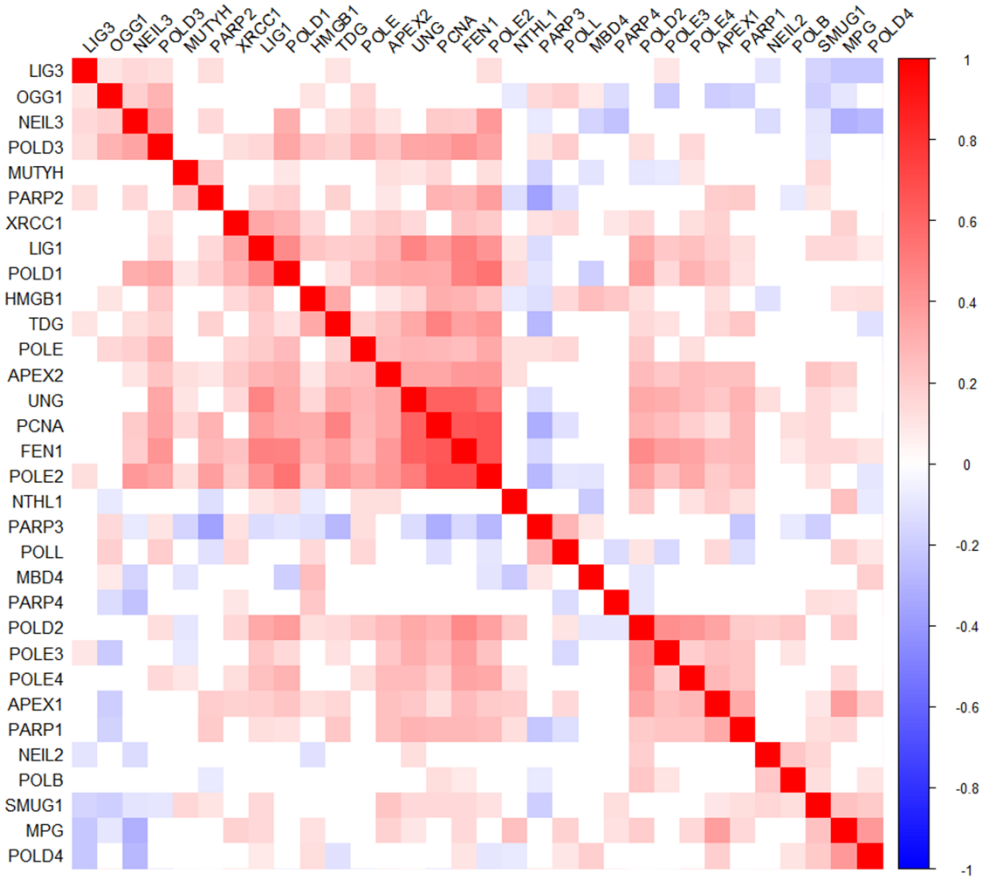
**Supplementary Figures**

Supplementary Figure 1



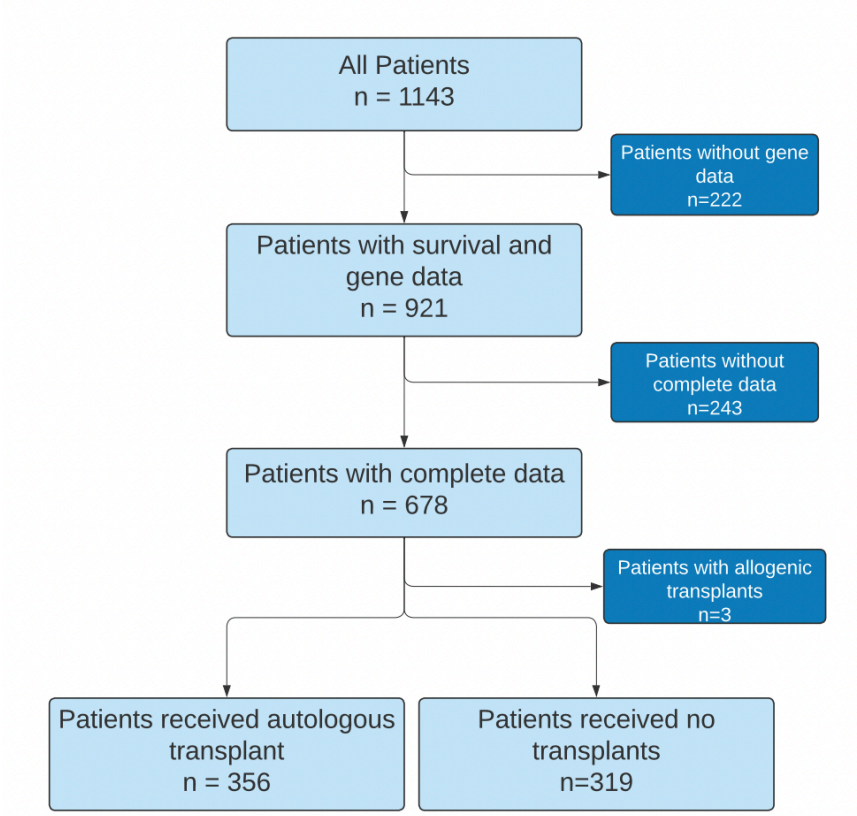

UMAP (GSE2113 (A), GSE6477 (C), GSE13591 (E)) and heatmap of significantly dysregulated BER genes (GSE2113 (B), GSE6477 (D), GSE13591 (F)) between MGUS and MM samples on BER genes alone demonstrate clear delineation between samples based on increased BER expression in MM.

Supplementary Figure 2



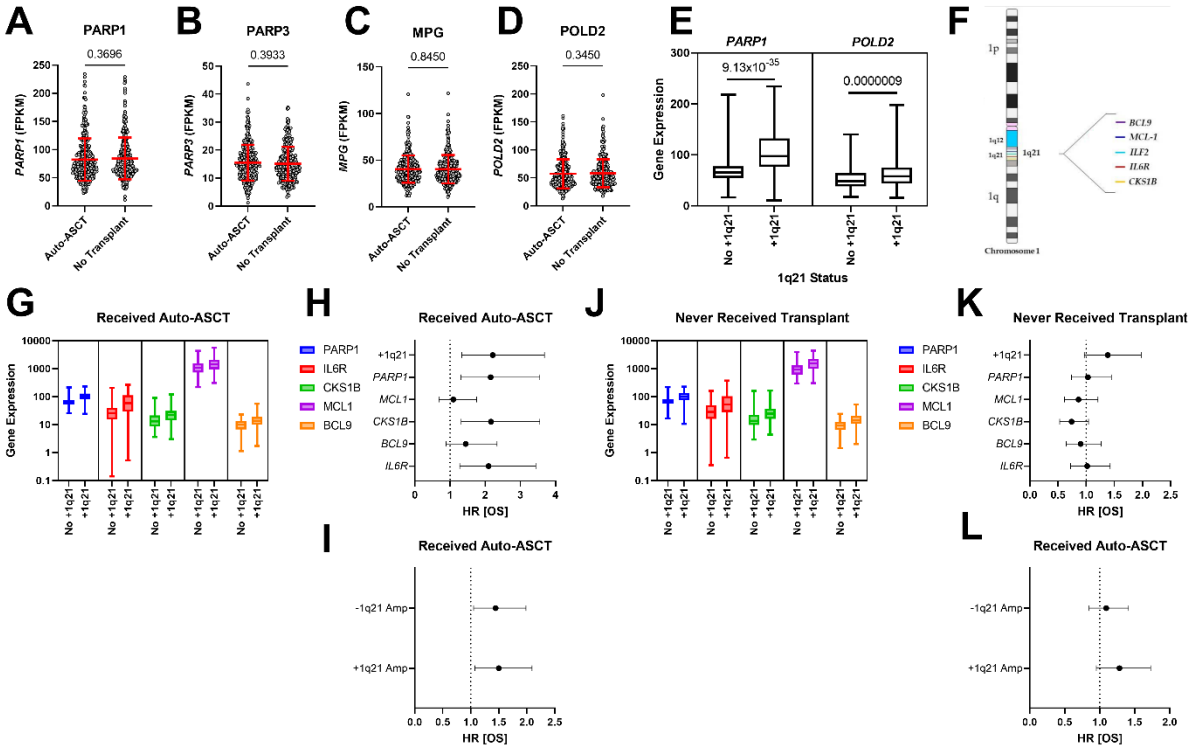

Co-expression of BER pathway genes in GSE2658 measured by Pearson's correlation coefficient.

Supplementary Figure 3



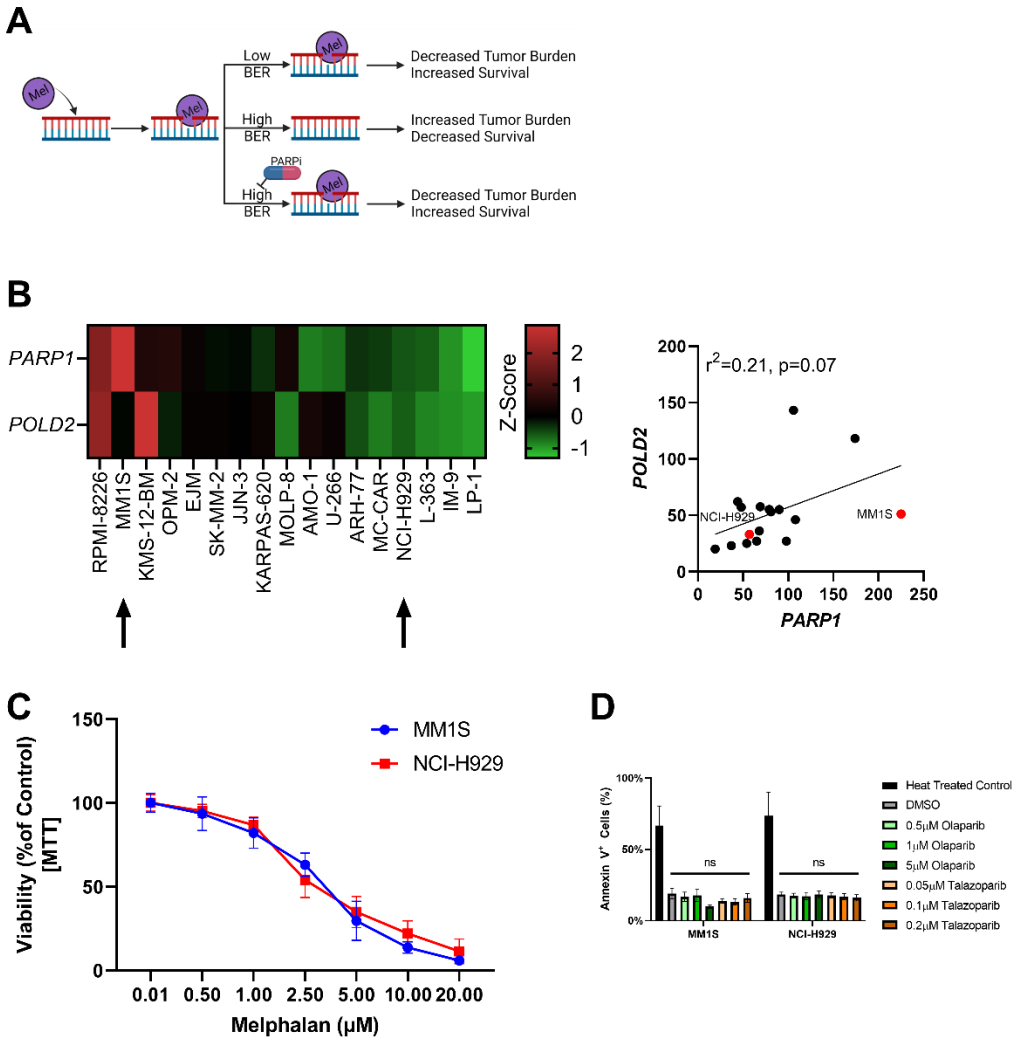

Gene expression and clinical data for 356 patients with MM who had received auto-ASCT and 319 patients with MM who had never received a transplant were collected from the MMRF.

Supplementary Figure 4



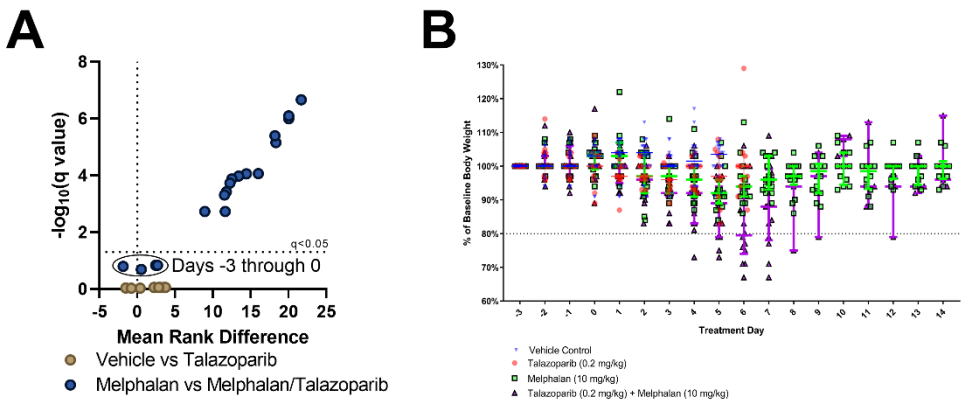

**A-D**) In the MMRF, expression of *PARP1* (**A**), *PARP3* (**B**), *MPG* (**C**), and *POLD2* (**D**) was consistent between patients who received auto-ASCT and those who never received a transplant. **E**) *PARP1* (1q42.12) and *POLD2* (7p13) were over expressed in samples with gain of 1q21 in at least 20% of cells identified by seqFISH. **F**) Genes located on 1q21 hypothesized to play a role in the mechanism of gain of 1q21 related clinical outcomes, including *CKS1B* (1q21.3), *IL6R* (1q21.3), *MCL1* (1q21.2), and *BCL9* (1q21.2). Adapted from Garcia et al[31]. **G-I**) Patients receiving auto-ASCT. **G**) 1q21 genes were increased in MM with gain of 1q21. **H**) In univariable analysis gain of 1q21, *CSK1B*, and *IL6R* were statistically associated with reduced OS. **I**) Increased *PARP1* expression (z-scores) were associated with poor OS in both 1q21 amplified and non-amplified disease. **J-L**) Patients receiving auto-ASCT. **J**). For patients who had never received a transplant, 1q21 genes were increased in MM with gain of 1q21. No 1q21 genes or 1q21 status were associated with OS. **L**) *PARP1* (z-scores) was not associated with OS regardless of 1q21 status.

Supplementary Figure 5




**A**) The proposed mechanism for the correlation between PARP1/POLD2 expression and survival in MM treated with ASCT. In MM cells with active BER, melphalan lesions can be repaired leading to an increased tumor burden, decreased efficacy of ASCT, and ultimately reduced overall survival. In MM cells with less active BER, melphalan lesions cannot be readily repaired leading to a reduced tumor burden, an increased efficacy of ASCT, and ultimately improved overall survival. Graphic created with BioRender.com. **B**) Data from the Genomics of Drug Sensitivity in Cancer database and the Broad Institute Cancer Cell Line Encyclopedia (CCLE), including DNA variants, mRNA expression, and drug sensitivity, were collected for 17 multiple myeloma cancer cell lines. The MM1S and NCI-H929 cell lines represent low and high expression, respectively, of both PARP1 and POLD2. **C**) Sensitivity to melphalan was not different between cell lines. **D**) PARP inhibitors olaparib or talazoparib had no cytotoxic effect on MM cell lines as measured by Annexin V positivity. A heat-treated positive control was created by water bath incubation for two minutes at 70°C

Supplementary Figure 6




**A**) For each treatment day, tumors burden was compared between vehicle control and talazoparib or melphalan and talazoparib with melphalan using the Mann Whitney Wilcoxon test for each comparison with p value adjustment. There was no difference in tumor size from day -3 to day 0 across all groups ($q>0.05$). Mice treated with melphalan and talazoparib demonstrated statistically smaller tumor volume from day +3 until the end of the study compared to melphalan alone ($q<0,05$). Mice receiving talazoparib alone demonstrated no significant difference in tumor burned compared to vehicle control ($q>0.05$). **B**) Mouse weight throughout the treatment period. The combination arm experienced severe adverse effects from their treatment with 40% mice in this arm being removed from study due to extreme weight loss. No other treatment arm experienced this effect.

**Supplementary Tables**

Supplementary Table 1: Transcriptomic Datasets

| Supplementary Table 1: Transcriptomic Datasets | | | | | | |
|---|---|---|---|---|---|---|
| Dataset | Normal Plasma | MGUS | sMM | MM | rMM | PCL |
| GSE13591 | 5 | 11 | - | 133 | - | 9 |
| GSE2113 | - | 7 | - | 39 | - | 6 |
| GSE6477 | 15 | 22 | 24 | 73 | 28 | - |
| GSE5900 | 22 | 44 | 12 | - | - | - |

MGUS: Monoclonal gammopathy of undetermined significance; sMM: smoldering multiple myeloma; MM: multiple myeloma; rMM: relapsed multiple myeloma; PCL: plasma cell leukemia

Supplementary Table 2: BER Gene Overall Survival Model

| Supplementary Table 2: BER Gene Overall Survival Model | | |
|---|---|---|
| Gene | Multivariate, Stepwise Cox-Proportional Hazard Regression | |
| | Hazard Ratio (95% Confidence Interval) | p Value |
| *MPG* | 0.9993 (0.9987-0.9998) | 0.009 |
| *PARP1* | 1.0001 (1-1.0002) | 0.03 |
| *PARP3* | 0.9989 (0.9981-0.9996) | 0.002 |
| *POLD1* | 1.0006 (0.9999-1.0013) | 0.07 |
| *POLD2* | 1.0005 (1.0001-1.0009) | 0.03 |

Supplementary Table 3: Multivariable Clinico-genomic model of Overall Survival

| Supplementary Table 3: Multivariable Clinico-genomic model of Overall Survival | | | | |
|---|---|---|---|---|
| Variable | Univariable | | Multivariable | |
| | HR (95%CI) | p Value | HR (95%CI) | p Value |
| *PARP1* - High | 2.15 (1.31-3.55) | 0.003 | 2.31 (1.15-4.67) | 0.02 |
| *PARP3* - High | 0.88 (0.55-1.41) | 0.58 | - | - |
| *MPG* - High | 0.79 (0.49-1.27) | 0.34 | - | - |
| *POLD2* - High | 1.67 (1.03-2.71) | 0.04 | 2.51 (1.24-5.06) | 0.01 |
| Race - White | 0.72 (0.42-1.23) | 0.23 | - | - |
| Age (Continuous) | 1.01 (0.98-1.04) | 0.49 | - | - |
| Age (>65 years) | 1.19 (0.74-1.94) | 0.47 | - | - |
| IMWG Risk Class (>0) | 2.10 (1.12-3.69) | 0.009 | 1.89 (0.95-3.77) | 0.07 |
| TP53 (No Functional Copies) | 2.99 (1.07-8.36) | 0.04 | 1.56 (0.51-4.79) | 0.44 |
| R-ISS (Stage 2-3) | 1.60 (0.78-3.29) | 0.2 | - | - |
| ECOG Performance Status (>0) | 1.88 (1.06-3.34) | 0.03 | 1.34 (0.66-2.72) | 0.42 |
| Received Maintenance Therapy | 0.66 (0.41-1.06) | 0.08 | - | - |
| Induction Regimen Contains: | | | | |
| Bortezomib | 0.59 (0.37-0.95) | 0.03 | 0.47 (0.24-0.91) | 0.03 |
| Cyclophosphamide | 1.26 (0.77-2.01) | 0.36 | - | - |
| Carfilzomib | 0.87 (0.37-2.00) | 0.74 | - | - |