Original Article

Genomic analysis of response to bacillus Calmette-Guérin (BCG) treatment in high-grade stage 1 bladder cancer patients

J. Alexa Sanders[1,2#], Connor Frasier[1,2#], Justin T. Matulay[3]^, Nury M. Steuerwald[4], Jason Zhu[5]^, Claud M. Grigg[5]^, James T. Kearns[3]^, Stephen B. Riggs[3], Kris E. Gaston[3], Cory R. Brouwer[1,2], R. Tucker Burks[6], Aaron L. Hartman[6], David M. Foureau[7], Earle F. Burgess[5]^, Peter E. Clark[3]

[1]Department of Bioinformatics & Genomics, University of North Carolina at Charlotte, Charlotte, NC, USA; [2]Bioinformatics Services Division, University of North Carolina at Charlotte, Kannapolis, NC, USA; [3]Department of Urology, Levine Cancer Institute/Atrium Health, Charlotte, NC, USA; [4]Molecular Biology and Microarray Core Facilities, Atrium Health, Charlotte, NC, USA; [5]Department of Medical Oncology, Levine Cancer Institute/Atrium Health, Charlotte, NC, USA; [6]Carolinas Pathology, Charlotte, NC, USA; [7]Immune Monitoring Core Laboratory, Levine Cancer Institute/Atrium Health, Charlotte, NC, USA

Contributions: (I) Conception and design: PE Clark, JA Sanders, C Frasier, JT Matulay; (II) Administrative support: None; (III) Provision of study materials or patients: PE Clark, SB Riggs, KE Gaston, RT Burks, AL Hartman; (IV) Collection and assembly of data: PE Clark, JA Sanders, C Frasier, JT Matulay; (V) Data analysis and interpretation: PE Clark, JA Sanders, C Frasier, JT Matulay, NM Steuerwald, CR Brouwer, JT Kearns, CM Grigg, J Zhu, DM Foureau, EF Burgess; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

[#]These authors contributed equally to this work.

Correspondence to: Peter E. Clark, MD. Professor and Chair, Department of Urology, Chair, Urologic Oncology, Specialty Medical Director, Urology, Levine Cancer Institute/Atrium Health, 1021 Morehead Medical Drive, Bldg One, Suite 5300, Charlotte, NC 28204, USA. Email: Peter.Clark@atriumhealth.org.

Background: Intravesical bacillus Calmette-Guérin (BCG) therapy is standard treatment for high-risk non-muscle invasive bladder cancer (NMIBC) but overall efficacy is low, and no reliable predictive biomarkers currently exist to refine patient selection. We performed genomic analysis on high-grade (HG) T1 NMIBCs to determine if response to therapy is predicted by certain mutational and/or expressional changes.

Methods: Patients with HG T1 NMIBC treated with induction BCG were stratified by response into durable and non-durable responders. Baseline tumor samples were subjected to targeted DNA sequencing and whole-exome RNAseq. Genomic variants differing significantly between response groups were analyzed using Ingenuity Pathway Analysis (IPA) software. Variant selection was refined to target potential biomarker candidates for responsiveness to BCG.

Results: Among 42 patients, the median follow-up was 51.7 months and 40.5% (n=17) were durable BCG responders. Deleterious mutations in the RNA sequence of *JCHAIN*, *S100A7*, *CLEC2B*, and *ANXA10* were more common in non-durable responders. Mutations in *MCL1* and *MSH6* detected on targeted sequencing were more commonly found in durable responders. Of all deleterious DNA and RNA mutations identified, only *MCL1* was significantly associated with longer recurrence free survival (RFS) (P=0.031).

Conclusions: Differences in the genomic profiles of HG T1 NMIBC tumors exist between those who show durable response to BCG and those who do not. Using pathway analysis, those differences imply upregulation of several interconnected inflammatory pathways among responders. Specific variants identified here, namely *MCL1*, are candidates for further study and, if clinically validated, may serve as useful biomarkers in the future.

Keywords: Non-muscle invasive bladder cancer (NMIBC); bacillus Calmette-Guérin (BCG); intravesical therapy; genomics

^ ORCID: Justin T. Matulay: 0000-0002-2467-5710; Jason Zhu: 0000-0003-0328-665X; Claud M. Grigg: 0000-0001-8847-5366; James T. Kearns: 0000-0002-5416-4671; Earle F. Burgess: 0000-0003-4293-4680.

Submitted Feb 24, 2021. Accepted for publication Jun 16, 2021.
doi: 10.21037/tau-21-158
View this article at: https://dx.doi.org/10.21037/tau-21-158

Introduction

The annual incidence of urothelial carcinoma of the bladder (UCB) in the United States is estimated to reach more than 83,000 cases in 2020, making it the 6th most common non-cutaneous malignancy (1). Non-muscle invasive bladder cancer (NMIBC) consists of carcinoma *in situ* (CIS), Ta, and T1 tumors and accounts for approximately 75% of new diagnoses (2). Overall, NMIBC carries excellent long-term prognosis with a treatment approach focusing on local resection with or without intravesical therapy (3). Tumors with the highest risk for progression (15% at 5 years) to muscle invasive bladder cancer (MIBC) are the subset including high-grade (HG) T1 disease, having demonstrated invasive biology by disrupting the lamina propria. Standard front-line therapy for high-risk NMIBC is visually complete transurethral resection (TUR) plus intravesical bacillus Calmette-Guérin (BCG) administration, but recurrence is common at 40% within 1 year of treatment and there remains a 5-year cancer specific mortality rate of 25% (4,5).

The mechanism of action for BCG therapy is believed to involve immune activation and destruction of local cancer cells but the phenomenon is poorly understood (6). There are no predictors for treatment response outside of classical clinicopathologic features, and the high treatment failure rate highlights the need for additional measures. In the current era of next generation sequencing (NGS) for UCB, much research effort to date has been focused on exploring the prognostic and predictive role for MIBC genomics as it pertains to treatment response, with relatively less attention paid to NMIBC genomics and BCG responsiveness on a *post hoc* basis (7-10).

There remains ample opportunity and need for NMIBC biomarkers associated with BCG response to improve patient selection for intravesical therapy versus upfront surgery. Furthermore, comparisons between responders and non-durable responders might lead to a better mechanistic understanding of BCG therapy. We used NGS (RNAseq and targeted DNA sequencing) to analyze a cohort of BCG treated NMIBC patients with the highest risk of disease progression (HG T1) stratified by treatment response to compare differences in expressional and mutational profiles in relation to BCG responsiveness.

We present the following article in accordance with the MDAR reporting checklist (available at https://dx.doi.org/10.21037/tau-21-158).

Methods

Patient characteristics

Recognizing the difficulties with clinical heterogeneity in NGS studies, we consciously elected to *a priori* restrict our analysis to patients with HG T1 UCB treated with complete induction BCG (≥5 of 6 weekly instillations) and compare those who responded to therapy versus those who did not. We retrospectively identified such patients at least 18 years or older at our institution between 2007 and 2019. Patients were excluded from further analysis if they had received any BCG within the previous 1 year. Response to BCG therapy was then assessed for each patient in order to create two discrete categories: BCG non-durable responders and BCG durable responders. Using date of first induction BCG instillation as the reference point, BCG non-durable responders were defined as patients with recurrence of UCB (any stage or grade) within 2 years, and BCG durable responders had no recurrences detected at any point during observable follow-up with a minimum disease-free interval of at least 2 years. BCG non-durable responders included those with a partial response, defined as any recurrence of lower grade or stage than the index lesion. Maintenance BCG was administered in all eligible patients in accordance with accepted guidelines (11). Formalin-fixed paraffin-embedded (FFPE) tissue blocks from the diagnostic TUR immediately preceding induction BCG were available for 45 patients. After microdissection and purification of DNA and RNA, a total of 42 samples were available for analysis.

DNA sequencing and variant calling

DNA samples were sequenced using the Trusight Tumor 170 gene panel (*TST170*, Illumina Inc., San Diego, CA, USA). Three hundred and fifty paired-end reads in fastq file format were first checked for quality using FastQC, a freely available software. The reads were then trimmed of adapter sequences and low-quality bases using Trimmomatic (12)

and checked for quality again following trimming. The reads were aligned to the human genome 38 (hg38) reference genome using the BWA MEM aligner (13). The alignment files were then condensed by lane to a single alignment per sample. These alignments were assessed for quality and depth of coverage using Picard HSMetrics tool (http://broadinstitute.github.io/picard/). The average depth of coverage was 580X, with a median depth of 375X. After being assessed for quality, the alignments were passed to the Genome Analysis Toolkit (GATK) for variant calling (14). A panel of normal references was created using genomes listed in the 1000 Genomes Project. The variants were also quality checked, filtered, and annotated using the GATK package. After variants were called, they were compared between groups (durable responders *vs.* non-durable responders) using a contingency table and a fisher's exact test. All variant P values listed are raw P values and any meeting a cutoff P value <0.05 were selected for downstream analysis.

RNA sequencing, gene expression, and variant calling

RNA samples were sequenced using Illumina's TruSeq RNA whole exome panel. Sequenced reads were quality checked and reads were aligned to the human genome hg38 reference set using STAR aligner (15). The aligned files were used for further analysis of gene expression and variant calling. Counts of mapped reads for the genes were calculated by featureCounts in the R package SubRead (16). These read counts were used as input for DESeq2 to identify significant differentially expressed genes between durable responders and BCG non-durable responder samples. After adjusting for multiple correction using the Benjamini-Hochberg method, those with an adjusted P value of less than 0.05 and log-fold change difference greater than two were considered statistically significant. These significantly differentially expressed genes were inputted into the Ingenuity Pathway Analysis (IPA) software (QIAGEN Inc., https://www.qiagenbioinformatics.com/products/ingenuity-pathway-analysis) to identify pathways that are potentially affected (17). Clinically annotated DNA sequencing and RNAseq data was uploaded to the National Center for Biotechnology Information Gene Expression Omnibus (https://www.ncbi.nlm.nih.gov/geo/).

Statistical analysis

Read counts were normalized and log transformed in R for input into classifyT1BC (Meeks classifier), consensusMIBC, and NMIBC classification (UROMOL classifier) to identify classification subtypes within our dataset (10,18,19). The classifyT1BC and consensusMIBC classifier tools were run using their respective R packages, while the NMIBC classification was run by uploading the log transformed normalized counts at http://nmibc-class.dk. For RNA-seq variant calling, aligned reads were merged per sample and duplicates marked with Picard tools. Sequence recalibration from GATK tools were used as input to Mutect2 for variant calling. Since matched normal samples were not available, a panel of normal references was created with MuTect2 to filter out potential germline variants. Raw variant calls for samples were then used for downstream analyses. After all variants were called, they were passed to bcftools to pull common variants from all samples. Only variants present in at least three samples were included. These common variants were then assessed for significance between groups (durable responders *vs.* non-durable responders) using fisher's exact test in R. Variants were filtered to only include those with a raw P value of <0.05 (none had an adjusted P value of <0.05). Further annotation was done on the common variants using the Annotate Variation (Annovar) software to predict amino acid changes (20). The significant common variants were further filtered to only nonsynonymous variants. These filtered variants were further filtered to those that were annotated to genes that were found to be differentially expressed. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). The study was approved by institutional/regional/national ethics/committee/ethics board of the Levine Cancer Institute of Atrium Health (No. LCI-GU-SPEC-PHA-001) and individual consent for this retrospective analysis was waived.

Results

Patient characteristics

A total of 42 patients were included in the final analysis, of which 17 (40.5%) were BCG durable responders and 25 (59.5%) were BCG non-durable responders. With a median overall follow-up of 51.7 months [interquartile range (IQR), 29.3–69.7 months] baseline characteristics of the cohort (*Table 1*) included a median age of 68.9 years old (IQR, 64.7–73.5 years old) comprised of 78.6% men (n=33), 78.6% current or former smokers (n=33), and 12.1% non-white ethnicities (n=5). A markedly higher proportion of females were observed among durable responders (n=7,

Table 1 Baseline demographics and outcomes in the final cohort of 42 patients with HG T1 UCB

Variables	Total		Durable responders		Non-durable responders		P value
	N	%	N	%	N	%	
Age in years, median (IQR)	68.9 (64.7–73.5)		66.1 (64.5–71.1)		69.4 (65.4–74.9)		0.303*
Sex							
Male	33	78.6	10	58.8	23	92.0	
Female	9	21.4	7	41.2	2	8.0	0.019^
BMI							
<25	13	31.0	5	29.4	8	32.0	
25–30	17	40.5	8	47.1	9	36.0	
>30	12	28.6	4	23.5	8	32.0	0.859†
CCI							
0–2	9	21.4	5	29.4	4	16.0	
3–9	33	78.6	12	70.6	21	84.0	0.122*
Race/ethnicity							
White/not Hispanic or Latino	28	66.7	14	82.4	14	56.0	
White/not specified	9	21.4	0	0.0	9	36.0	
African American	4	9.5	3	17.6	1	4.0	
Asian	1	2.4	0	0.0	1	4.0	0.020‡
Smoking status							
Current	6	14.3	0	0.0	6	24.0	
Former	27	64.3	14	82.4	13	52.0	
Never	9	21.4	3	17.6	6	24.0	0.057‡
Family history							
Yes	25	59.5	9	52.9	16	64.0	
No	17	40.5	8	47.1	9	36.0	0.534^
BCG responsive							
Yes	17	40.5	17	100.0	0	0.0	
No	25	59.5	0	0.0	25	100.0	N/A
Survival status							
Alive	33	78.6	15	88.2	18	72.0	
Dead	9	21.4	2	11.8	7	28.0	0.271^
Follow-up in months, median (IQR)	51.7 (29.3–69.7)		70.4 (56.3–79.7)		32.7 (25.3–51.6)		<0.001*

*, Welch two sample *t*-test; ^, Fisher's exact test; †, Wilcoxon rank sum test; ‡, chi-squared test. HG, high-grade; UCB, urothelial carcinoma of the bladder; IQR, interquartile range; BMI, body mass index; CCI, Charlson comorbidity index; BCG, bacillus Calmette-Guérin.

Table 2 Characteristics of recurrences among BCG non-durable responders

Variables	Non-durable responders	
	N	%
Grade at recurrence		
LG	3	12.0
HG	22	88.0
Stage at recurrence		
Ta	11	44.0
CIS	1	4.0
T1	10	40.0
T2 or greater	3	12.0
Partial response*		
Yes	12	48.0
No	13	52.0
Time to recurrence months, median (IQR)	7.7 (5.4–8.8)	

*, partial response defined as patients with T-stage < T1 at recurrence. BCG, bacillus Calmette-Guérin; LG, low grade; HG, high-grade; CIS, carcinoma *in situ*; IQR, interquartile range.

41.2%) than non-durable responders (n=2, 8.0%, P=0.019). Additionally, there were no active smokers among the durable responders, as compared to 24.0% (n=6, P=0.057) in the non-durable responder group.

Recurrences in the non-durable responders (*Table 2*) occurred at a median of 7.7 months (IQR, 5.4–8.8 months) from initiation of induction BCG and were predominantly HG (n=22, 88.0%). Partial response was observed in 48.0% (n=12). There were 3 patients (12.0%) who progressed to muscle invasion or higher during follow-up, including a single patient who developed distant metastasis in the setting of negative surveillance cystoscopy.

Gene expression

In total, 67 genes (Table S1) were found to have statistically significant (P<0.05) differential RNA expression on RNAseq when comparing durable responders to non-durable responders. The top genes found to be mutated across all samples, including both durable responders and non-durable responders, were *BRCA2*, *ATM*, and *EP300* at 96.73%, 93.55%, and 90.32% respectively. The results of unsupervised hierarchical clustering analysis are shown in *Figure 1* along with select clinicopathologic factors. Selected publicly available classification schemes for urothelial carcinoma were applied. To test whether these were involved in mediating inflammatory pathways, the differentially expressed genes were analyzed using the IPA software, using the expression levels for non-durable responders as the reference value. Three of the top pathways with statistically significant differences in gene expression (*Table 3*; *Figure 2*) were directly or indirectly associated with inflammation including macrophage migration inhibitory factor (MIF)-mediated glucocorticoid regulation (Figure S1), MIF regulation of innate immunity (Figure S2), and p38 mitogen activated protein family of kinases (MAPK) signaling (Figure S3). Furthermore, the predicted impact of changes in specific gene expression among durable responders was significant global up-regulation as well as activation of the key inflammatory mediator NF-κB when compared to non-durable responders (*Table 3*).

DNA and RNA variant analysis

In addition to gene expression, we analyzed the RNAseq dataset for variants that may be potentially useful biomarker candidates. The pool of RNA variants (Figure S4) was first filtered to variants that appeared in at least three samples (n=126,033 variants). These were further filtered to variants that were annotated to proteins as nonsynonymous mutations (n=44,798 variants), of which 856 variants were significantly enriched (raw P value <0.05). Finally, of these significant nonsynonymous variants, there were six variant genes identified (*Figure 3A*) from among the 228 genes previously identified as being differentially expressed based on BCG responsiveness. Notably, using these criteria, none of the differentially expressed genes from the pathway analysis were also RNA variants considered to be biomarker candidates. We found that *JCHAIN* variants were observed only in the non-durable responder samples (47.8% *vs.* 0%, P<0.001), while *S100A7* (39.1% *vs.* 5.9%, P=0.026), *CLEC2B* (43.5% *vs.* 11.8%, P=0.041), and *S100A7* (52.2% *vs.* 17.6%, P=0.046) were also significantly enriched in those patients (*Figure 3B*; Table S2).

Analysis of DNA variants was conducted with the *TST170* gene panel using similar iterative exclusion methodology as above (Figure S4), with 150 DNA variants initially identified in three or more tumor samples, of which 45 were annotated to proteins as nonsynonymous mutations. There was statistically significant (P value <0.05) enrichment of two DNA variants that also passed

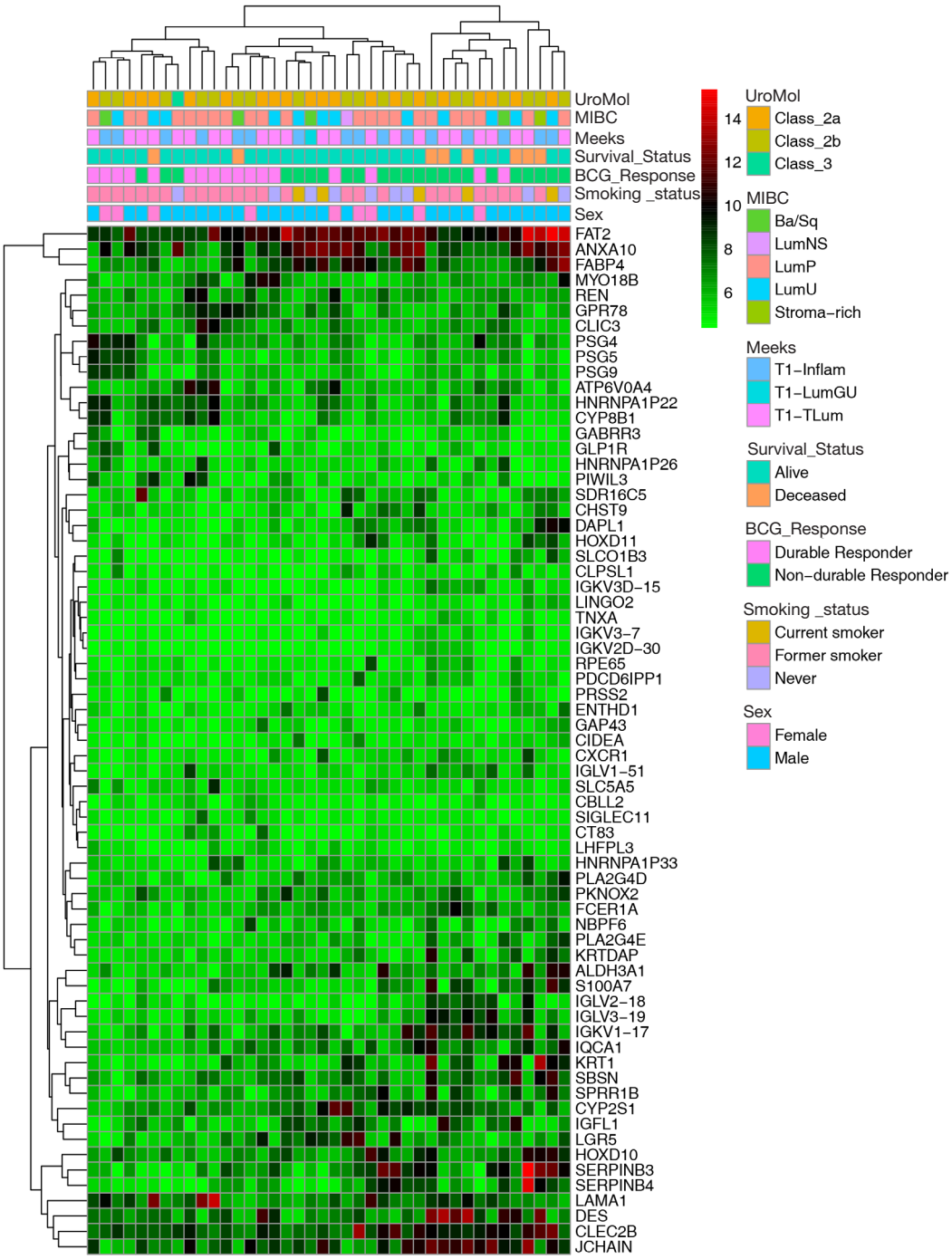


Figure 1 Gene expression heatmap of 67 statistically significant differentially expressed genes identified. Heatmap shows the count matrix data for 67 genes in all samples. Samples are colored based on their response group, shown in the legend. Additional clinical information is also provided. Publicly available classifiers from UroMol, consensusMIBC, and Meeks have been applied to the RNAseq data. MIBC, muscle invasive bladder cancer; BCG, bacillus Calmette-Guérin.

Table 3 Differentially expressed genes identified within the top 5 pathways as determined by IPA software

Pathway	Gene	Expr log ratio	Expr P value	Expected impact
MIF-mediated glucocorticoid regulation	*PLA2G4D*	–2.119	0.042	Up
	PLA2G4E	–2.338	0.027	Up
MIF regulation of innate immunity	*PLA2G4D*	–2.119	0.042	Up
	PLA2G4E	–2.338	0.027	Up
p38 MAPK signaling	*PLA2G4D*	–2.119	0.042	Up
	PLA2G4E	–2.338	0.027	Up

Expression values included in the table are representative of values among durable responders as compared to non-durable responders as baseline. The expected impact is generated by IPA software to indicate the predicted effect the change in expression in each particular gene would have on the associated pathway. IPA, Ingenuity Pathway Analysis; MIF, macrophage migration inhibitory factor; MAPK, mitogen activated protein family of kinases.

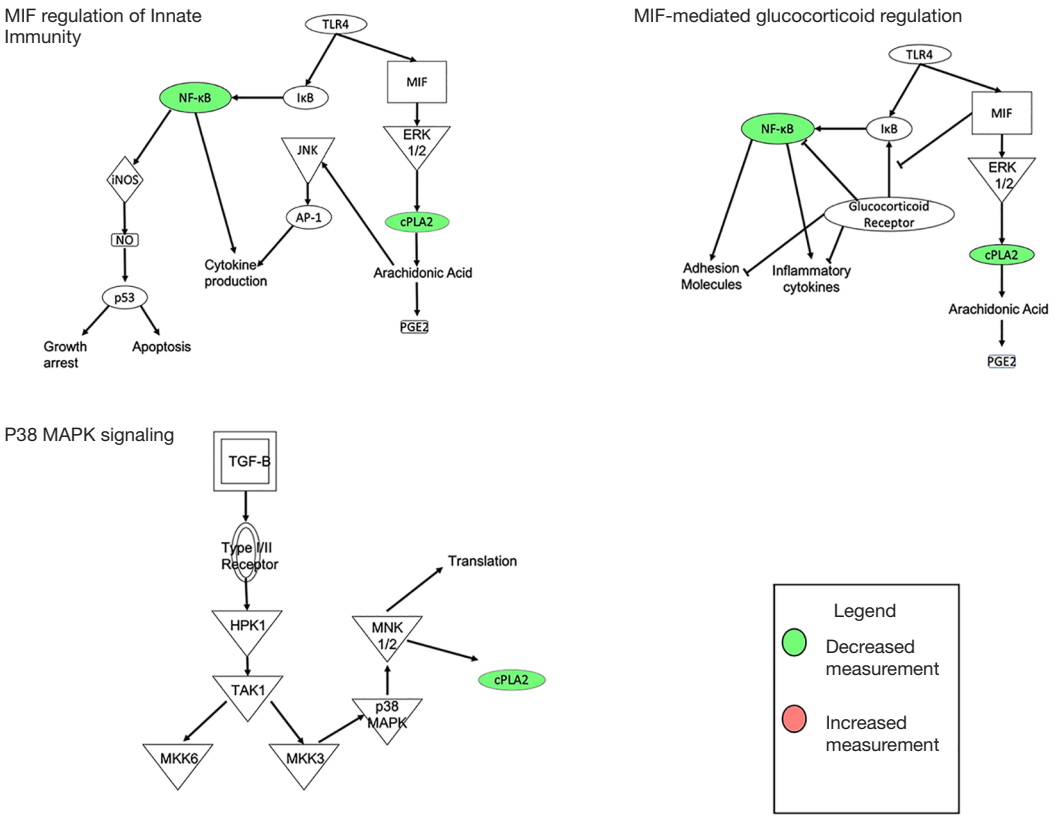
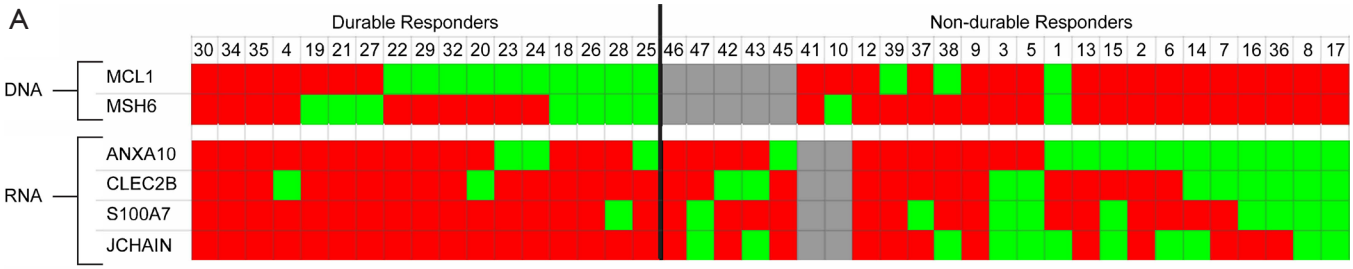


Figure 2 Simplified pathway analysis of genes with statistically significant differential expression between durable responders and non-durable responders. Pathway analysis was performed using IPA software which integrates gene expression results to visualize associated biological pathways. The top affected pathways are summarized here. IPA, Ingenuity Pathway Analysis.

nonsynonymous filtration (*Figure 3A*). Each of the identified DNA variants were observed in a higher proportion of durable responders (*Figure 3C*; Table S3): frameshift deletions in *MCL1* (58.8% *vs.* 15%, P=0.007) and *MSH6* (41.2% *vs.* 10.0%, P=0.034).

All DNA and RNA variants showing a statistically significant association with BCG response were then analyzed for impact in recurrence free survival (RFS) using



B

Gene	Expression (log2 fold change)	Ratio in Durable Responders	Ratio in Non-durable Responders	P-value
JCHAIN	-2.082	0/17	11/23	<0.001
S100A7	-3.534	1/17	9/23	0.026
CLEC2B	-2.116	2/17	10/23	0.041
ANXA10	-2.588	3/17	12/23	0.046

C

Gene	Ratio in Durable Responders	Ratio in Non-durable Responders	P-value
MCL1	10/17	3/20	0.007
MSH6	7/17	2/20	0.034

Figure 3 Analysis of genetic variation between durable BCG responders and non-durable responders. (A) The heatmap depicts the two statistically significant DNA variants and the post-filtering RNAseq variants. Red indicates the absence of the variant while green indicates the variant is present. Grey squares indicate that either DNA or RNA was unable to be analyzed for that particular sample. The tables numerically outline the presence of each RNA variant (B) and DNA variant (C) as a ratio for the responders and non-durable responders. The associated raw P values for Fisher's exact test are also listed. BCG, bacillus Calmette-Guérin.

the Kaplan-Meier survival method. A deleterious mutation in the *MCL1* gene (*Figure 4A*) was associated with significantly prolonged RFS as compared to wild type (2-year RFS 77.0% *vs.* 37.5%; log-rank test P=0.031). Though not reaching predefined statistical significance, mutation of the *MSH6* gene (*Figure 4B*) was also associated with longer 2-year RFS (77.8% *vs.* 42.9%; log-rank test P=0.073).

Discussion

Predicting BCG responsiveness in NMIBC has remained an elusive goal, even as NGS techniques have vastly increased availability of genomic data. We sought to identify potential biomarkers for predicting response using a cohort of patients which were *a priori* selected to have similar baseline pathologic features (HG T1 tumors) who were treated with "adequate" BCG and displayed either durable complete response (median follow-up 70.4 months; IQR, 56.3–79.7) or disease relapse before 2 years of follow-up (median time to recurrence 7.7 months; IQR, 5.4–8.8). In this cohort, deleterious mutations in the *MCL1* gene were associated with significantly longer RFS (median 9.3 months *vs.* not reached; P=0.031). Additional pathway analysis identified notable inflammatory pathways linked to durable responders with predicted increased pathway activation and activation of NF-κB.

MCL1 is an anti-apoptotic protein of the *BCL2* family and its down-regulation has been linked to tumor necrosis factor-related apoptosis-inducing ligand (TRAIL) sensitization (6,21). In prior work using immunohistochemical staining for *BCL2*, expression was shown to be inversely correlated with recurrence of NMIBC following BCG treatment with 47% prevalence at baseline in patients with tumor recurrences following BCG therapy as compared to 79% staining positive among non-recurrent cases (22). TRAIL expression has been shown to be up-regulated in granulocytes following induction of BCG and release of interferon signaling particles by cancer cells, and is therefore believed to be a key death ligand in the anti-neoplastic properties of BCG (23). The location of the frameshift deletion within *MCL1* we detected falls in the first exon near the N-terminus within a region believed to be involved in protein regulation (24). We speculate, therefore, that sensitization to TRAIL, induced by a loss of function mutation in *MCL1* that is seen more frequently in the durable responders and in association with longer RFS, provides a mechanism by which BCG responsiveness could be enhanced. The combination of a plausible mechanistic explanation for the association between *MCL1* and BCG responsiveness makes this a promising target for further

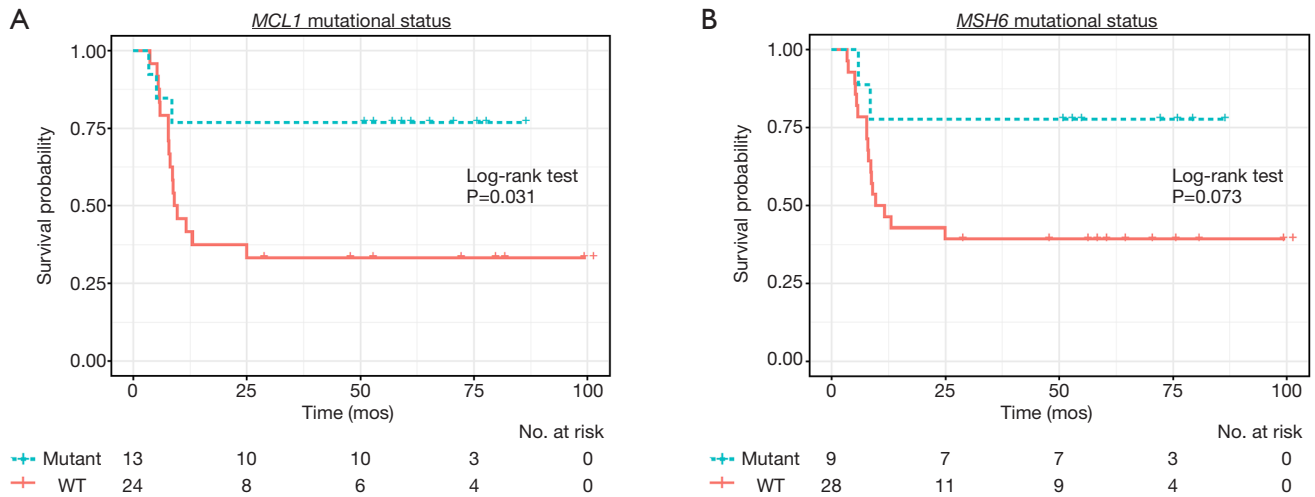


Figure 4 Kaplan-Meier survival plots. (A) Kaplan-Meier survival plot for RFS based on presence of deleterious mutations in *MCL1* (log-rank test P=0.031) on targeted DNA sequencing panel *TST170*. (B) Kaplan-Meier survival plot for RFS based on presence of deleterious mutations in *MSH6* (log-rank test P=0.073) on targeted DNA sequencing panel *TST170*. RFS, recurrence free survival.

biomarker validation.

Deleterious mutations in the DNA mismatch repair protein *MSH6* were also closely linked to longer RFS, though this did not meet our prespecified cutoff for statistical significance (P=0.073), a fact that may be explained by the small sample sizes. The *MSH6* gene is well known for its role in Lynch syndrome whereby germline mutations lead to accumulation of errors in regions of DNA called microsatellites. The classic genitourinary manifestation of Lynch syndrome is that of upper tract urothelial carcinoma which is related to UCB through a shared histologic origin, but which has significant differences in genomic signatures (25). Beyond the association with Lynch syndrome, *MSH6* has been linked specifically to UCB where methylation status of *MSH6* has shown an association with progression and recurrence among patients treated with intravesical BCG (26). The deactivating, hypermethylated state of *MSH6* corresponded to an increased time to progression and improved survival outcomes following BCG therapy in patients with HG NMIBC, and as such, should be considered for further investigation into its potential for a biomarker of response to therapy.

The pathways identified in our gene expression analysis show that inflammation and immune response are important triggers for BCG response, which we confirm in our formal pathway analysis using IPA software. MIF-mediated glucocorticoid regulation and MIF regulation of innate immunity, two of the top pathways from our analysis, are strongly involved in the immune system response. MIF is released during inflammatory reactions and has been shown to promote the expression of several pro-inflammatory cytokines (27). Upregulation of MIF-related pathways predicted by IPA software (Figures S1,S2) that we observed in our samples would be expected to result in a pro-inflammatory environment conducive to immune-mediated tumor destruction. Interestingly, there is a common linkage between all three of these pathways in the form of decreased expression in the *NFKB1* gene leading to upregulation of each associated pathway (*Table 3*). Functionally, the *NFKB1* gene is involved in regulating transcription for many pathways associated with inflammation, immunity, and apoptosis (28).

Other variants identified in our cohort, *CLEC2B* and *S100A7*, can also be linked to NF-κB activity but the correlation is less direct. *CLEC2B* is a member of the C-type lectin superfamily of proteins, carbohydrate binding proteins that can also self-bind and modulate many physiological activities (29). Increased TRAIL levels have been linked to C-type lectin-like molecule binding, but no specific work has shown direct correlation with expression of the *CLEC2B* gene. *S100A7* is a less well described pro-survival protein present in other cancer models (30). It has also been linked to NF-κB signaling through suppression of the Yes-associated protein (YAP) oncogene cascade (31) and can be evaluated with immunohistochemical staining, though a previous study failed to find statistically significant

association with BCG response (32). Though speculative until confirmed by further study, these observations lend additional support to support the hypothesis that alterations in *NF-κB* and its associated inflammatory regulatory networks may impact the response to BCG in patients with HG T1 UCB. *ANXA10* is from the annexin family of proteins that are calcium dependent membrane binding proteins (33). While *ANXA10*'s specific function is still unknown, high expression levels have previously been linked to lower rates of NMIBC progression and improved survival outcomes (33,34).

The present study is limited by the small sample size available for genomic analysis, though we believe a strength of our data lies in the homogeneity of primary tumor samples. By narrowing our criteria to only HG T1 tumors that had clearly either responded to BCG, with a median disease-free follow-up of nearly 6 years, versus patients with relapse at a median of less than 8 months, we have intentionally maximized the difference in clinical outcomes. The use of a targeted 170 gene panel (Illumina's *TST170*) is another potential limitation. While there may be additional interesting variants to identify through whole-exome sequencing, the targeted panel increases confidence as it provides higher sequencing coverage depth of the selected genes. Due to the retrospective nature of our dataset, we were not able to collect normal samples for matched analysis. Matching tumor/normal sample pairs are important in filtering out potential germline variants from the true somatic variants and represents an important future step in this work.

Conclusions

Intravesical BCG remains the preferred frontline treatment for most patients with HG T1 bladder cancer who have the highest risk of progression to muscle invasive disease. In this vulnerable patient population there is an ongoing need to identify reliable biomarkers that can predict patient responsiveness to BCG immunotherapy. We have identified several differentially mutated and expressed genes that vary based on responsiveness to therapy that can serve as candidate biomarkers to be further interrogated with validation studies, specifically *MCL1* and *MSH6*. Furthermore, based on our work we hypothesize that NF-κB signaling may be a common thread linking these variants and pathways we analyzed. This serves as the foundation for further investigation into these alterations as possible rational biomarkers for high-risk UCB treatment.

Acknowledgments

The authors would like to acknowledge Illumina, Inc. for providing TruSight Tumor 170 kits and associated consumables for the targeted DNA sequencing portion of this study.

Funding: Carolinas Bladder Cancer Fund; Don and Betty Anderson Fund for Bladder Cancer Research.

Footnote

Reporting Checklist: The authors have completed the MDAR reporting checklist. Available at https://dx.doi.org/10.21037/tau-21-158

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at https://dx.doi.org/10.21037/tau-21-158). The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). The study was approved by institutional/regional/national ethics/committee/ethics board of the Levine Cancer Institute of Atrium Health (No. LCI-GU-SPEC-PHA-001) and individual consent for this retrospective analysis was waived.

References

1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2020. CA Cancer J Clin 2020;70:7-30.
2. Burger M, Catto JW, Dalbagni G, et al. Epidemiology

and risk factors of urothelial bladder cancer. Eur Urol 2013;63:234-41.
3. Cambier S, Sylvester RJ, Collette L, et al. EORTC nomograms and risk groups for predicting recurrence, progression, and disease-specific and overall survival in non-muscle-invasive stage Ta-T1 urothelial bladder cancer patients treated with 1-3 years of maintenance bacillus Calmette-Guérin. Eur Urol 2016;69:60-9.
4. Chang SS, Bochner BH, Chou R, et al. Treatment of nonmetastatic muscle-invasive bladder cancer: American Urological Association/American Society of Clinical Oncology/American Society for Radiation Oncology/ Society of Urologic Oncology Clinical Practice Guideline Summary. J Oncol Pract 2017;13:621-5.
5. Dalbagni G, Vora K, Kaag M, et al. Clinical outcome in a contemporary series of restaged patients with clinical T1 bladder cancer. Eur Urol 2009;56:903-10.
6. Redelman-Sidi G, Glickman MS, Bochner BH. The mechanism of action of BCG therapy for bladder cancer--a current perspective. Nat Rev Urol 2014;11:153-62.
7. Hedegaard J, Lamy P, Nordentoft I, et al. Comprehensive transcriptional analysis of early-stage urothelial carcinoma. Cancer Cell 2016;30:27-42.
8. Pietzak EJ, Bagrodia A, Cha EK, et al. Next-generation sequencing of nonmuscle invasive bladder cancer reveals potential biomarkers and rational therapeutic targets. Eur Urol 2017;72:952-9.
9. van Kessel KEM, van der Keur KA, Dyrskjøt L, et al. Molecular markers increase precision of the European Association of Urology Non-Muscle-Invasive Bladder Cancer Progression Risk Groups. Clin Cancer Res 2018;24:1586-93.
10. Robertson AG, Groeneveld CS, Jordan B, et al. Identification of differential tumor subtypes of T1 bladder cancer. Eur Urol 2020;78:533-7.
11. Chang SS, Boorjian SA, Chou R, et al. Diagnosis and treatment of non-muscle invasive bladder cancer: AUA/ SUO guideline. J Urol 2016;196:1021-9.
12. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 2014;30:2114-20.
13. Li H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv, 2013:1303.3997.
14. Van der Auwera GA, Carneiro MO, Hartl C, et al. From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. Curr Protoc Bioinformatics 2013;43:11.10.1-11.10.33.
15. Dobin A, Davis CA, Schlesinger F, et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics 2013;29:15-21.
16. Liao Y, Smyth GK, Shi W. The R package Rsubread is easier, faster, cheaper and better for alignment and quantification of RNA sequencing reads. Nucleic Acids Res 2019;47:e47.
17. Krämer A, Green J, Pollard J Jr, et al. Causal analysis approaches in Ingenuity Pathway Analysis. Bioinformatics 2014;30:523-30.
18. Kamoun A, de Reyniès A, Allory Y, et al. A consensus molecular classification of muscle-invasive bladder cancer. Eur Urol 2020;77:420-33.
19. Lindskrog SV, Prip F, Lamy P, et al. An integrated multi-omics analysis identifies prognostic molecular subtypes of non-muscle-invasive bladder cancer. Nat Commun 2021;12:2301.
20. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res 2010;38:e164.
21. Kim SH, Ricci MS, El-Deiry WS. Mcl-1: a gateway to TRAIL sensitization. Cancer Res 2008;68:2062-4.
22. Okamura T, Akita H, Kawai N, et al. Immunohistochemical evaluation of p53, proliferating cell nuclear antigen (PCNA) and bcl-2 expression during bacillus Calmette-Guerin (BCG) intravesical instillation therapy for superficial bladder cancers. Urol Res 1998;26:161-4.
23. Ludwig AT, Moore JM, Luo Y, et al. Tumor necrosis factor-related apoptosis-inducing ligand: a novel mechanism for Bacillus Calmette-Guérin-induced antitumor activity. Cancer Res 2004;64:3386-90.
24. Germain M, Duronio V. The N terminus of the anti-apoptotic BCL-2 homologue MCL-1 regulates its localization and function. J Biol Chem 2007;282:32233-42.
25. Moss TJ, Qi Y, Xi L, et al. Comprehensive genomic characterization of upper tract urothelial carcinoma. Eur Urol 2017;72:641-9.
26. Agundez M, Grau L, Palou J, et al. Evaluation of the methylation status of tumour suppressor genes for predicting bacillus Calmette-Guérin response in patients with T1G3 high-risk bladder tumours. Eur Urol 2011;60:131-40.
27. Nobre CC, de Araújo JM, Fernandes TA, et al. Macrophage migration inhibitory factor (MIF): biological activities and relation with cancer. Pathol Oncol Res 2017;23:235-44.
28. Li P, Gu J, Yang X, et al. Functional promoter -94 ins/del ATTG polymorphism in NFKB1 gene is associated with bladder cancer risk in a Chinese population. PLoS One 2013;8:e71604.

29. Wiersma VR, de Bruyn M, Shi C, et al. C-type lectin-like molecule-1 (CLL1)-targeted TRAIL augments the tumoricidal activity of granulocytes and potentiates therapeutic antibody-dependent cell-mediated cytotoxicity. MAbs 2015;7:321-30.
30. Emberley ED, Niu Y, Curtis L, et al. The S100A7-c-Jun activation domain binding protein 1 pathway enhances prosurvival pathways in breast cancer. Cancer Res 2005;65:5696-702.
31. Li Y, Kong F, Shao Q, et al. YAP expression and activity are suppressed by S100A7 via p65/NFκB-mediated repression of ΔNp63. Mol Cancer Res 2017;15:1752-63.
32. Kitamura H, Torigoe T, Honma I, et al. Effect of human leukocyte antigen class I expression of tumor cells on outcome of intravesical instillation of bacillus calmette-guerin immunotherapy for bladder cancer. Clin Cancer Res 2006;12:4641-4.
33. Munksgaard PP, Mansilla F, Brems Eskildsen AS, et al. Low ANXA10 expression is associated with disease aggressiveness in bladder cancer. Br J Cancer 2011;105:1379-87.
34. Jin S, Chang IH, Kim JW, et al. Identification of downstream genes of the mtor pathway that predict recurrence and progression in non-muscle invasive high-grade urothelial carcinoma of the bladder. J Korean Med Sci 2017;32:1327-36.

Cite this article as: Sanders JA, Frasier C, Matulay JT, Steuerwald NM, Zhu J, Grigg CM, Kearns JT, Riggs SB, Gaston KE, Brouwer CR, Burks RT, Hartman AL, Foureau DM, Burgess EF, Clark PE. Genomic analysis of response to bacillus Calmette-Guérin (BCG) treatment in high-grade stage 1 bladder cancer patients. Transl Androl Urol 2021;10(7):2998-3009. doi: 10.21037/tau-21-158

Table S1 Complete list of all 67 differentially expressed genes among the entire cohort

Column 1	baseMean	log2FoldChange	lfcSE	stat	P value
RPE65	7.68733898	–2.571350011	0.70442542	–3.65028	0.00026195
NBPF6	25.90198965	–2.442938581	0.68712315	–3.555314	0.00037753
SPRR1B	157.3235125	–2.52973519	0.70307749	–3.5980887	0.00032056
S100A7	171.5242641	–3.533723408	0.95979549	–3.6817462	0.00023164
FCER1A	64.08876755	–2.0302491	0.53216142	–3.8151001	0.00013613
REN	108.3446647	2.35955338	0.48155691	4.89984322	9.59E–07
IGKV3-7	3.794564464	–4.158328543	0.98978012	–4.201265	2.65E–05
IGKV1-17	373.0407807	–2.969023821	0.73768259	–4.0247986	5.70E–05
IGKV2D-30	3.317110822	–4.486840264	1.07764392	–4.1635648	3.13E–05
IGKV3D-15	9.567786631	–3.427569731	0.82681503	–4.1455097	3.39E–05
DAPL1	51.5754466	–3.733826377	0.80840908	–4.6187338	3.86E–06
HOXD11	31.05264799	–3.60949324	0.8341999	–4.3268924	1.51E–05
HOXD10	225.067114	–2.082518063	0.52159398	–3.9926037	6.54E–05
CXCR1	14.51362623	–2.694638035	0.62956326	–4.2801704	1.87E–05
DES	1105.806282	–2.803220454	0.74075671	–3.784266	0.00015416
IQCA1	157.7611797	–2.095790107	0.57000182	–3.676813	0.00023617
CYP8B1	100.6244139	2.469417197	0.66801656	3.69664067	0.00021847
HNRNPA1P2	118.0411934	2.262877569	0.54673085	4.13892421	3.49E–05
GABRR3	13.83206586	2.790034517	0.78782189	3.5414534	0.00039793
GAP43	7.858817376	–2.4393895	0.61124018	–3.9908854	6.58E–05
GPR78	117.6919983	2.569996155	0.48736887	5.27320542	1.34E–07
JCHAIN	1005.933162	–2.081747521	0.60040368	–3.4672464	0.00052582
ANXA10	1751.90732	–2.588131704	0.57452764	–4.5047993	6.64E–06
FAT2	4586.23489	–2.057745311	0.52966203	–3.8850157	0.00010232
TNXA	3.164020447	–3.667106616	0.99169801	–3.6978057	0.00021747
CLPSL1	7.293100841	–3.464649696	0.90966568	–3.8087066	0.0001397
GLP1R	28.53075021	2.970804223	0.84526271	3.51465195	0.00044033
LHFPL3	6.707658831	2.234167912	0.63836499	3.49982841	0.00046556
ATP6V0A4	145.3407964	3.171395695	0.66013045	4.80419541	1.55E–06
PRSS2	6.5480637	–3.991690365	1.04962055	–3.8029842	0.00014296
CBLL2	5.102838811	2.46572491	0.61989647	3.97763984	6.96E–05
HNRNPA1P2	22.82667343	2.411431636	0.62060168	3.88563502	0.00010206
CT83	10.51343918	3.741888855	0.85245742	4.38953166	1.14E–05
SDR16C5	35.39316594	–2.619917391	0.6338022	–4.1336515	3.57E–05
FABP4	732.2650152	–2.615212461	0.63419368	–4.123681	3.73E–05
LINGO2	5.626760599	–2.754688192	0.7951548	–3.464342	0.00053153
CLIC3	106.6775328	2.528410334	0.52383202	4.82675791	1.39E–06
PKNOX2	67.3993649	–2.200218422	0.49988345	–4.4014628	1.08E–05
HNRNPA1P3	31.82147202	2.364252715	0.5995125	3.94362537	8.03E–05
CLEC2B	892.3273153	–2.116153228	0.4844572	–4.3680912	1.25E–05
SLCO1B3	10.89656062	–5.341511987	1.11677497	–4.7829797	1.73E–06
KRT1	231.4617811	–3.530174731	0.80998071	–4.3583442	1.31E–05
LGR5	215.3834438	–3.007485344	0.69966071	–4.2984911	1.72E–05
PDCD6IPP1	14.88562238	–2.198570803	0.60631082	–3.6261447	0.00028768
PLA2G4E	32.03878282	–2.338369072	0.62479961	–3.7425904	0.00018213
PLA2G4D	51.60568941	–2.118924598	0.59692566	–3.5497294	0.00038563
ALDH3A1	158.8963232	–2.536654695	0.681123	–3.7242241	0.00019592
LAMA1	496.0994292	2.619976536	0.56935057	4.60169301	4.19E–06
CIDEA	9.546056103	–2.830693268	0.77815274	–3.6377091	0.00027507
CHST9	44.2382589	–2.760563673	0.73014027	–3.7808676	0.00015628
SERPINB4	111.4501069	–2.903643906	0.76793457	–3.7811085	0.00015613
SERPINB3	590.7329976	–3.369561537	0.80818	–4.1693206	3.06E–05
SLC5A5	25.27003541	2.792910318	0.57500228	4.85721605	1.19E–06
KRTDAP	41.57913812	–2.384849081	0.64584581	–3.6925982	0.00022197
SBSN	255.5431969	–2.499487433	0.60155142	–4.1550686	3.25E–05
CYP2S1	214.9712457	–2.197729068	0.49353876	–4.453002	8.47E–06
PSG5	58.55825416	2.560115345	0.63566379	4.02746764	5.64E–05
PSG4	137.2964535	2.808715606	0.58875103	4.77063384	1.84E–06
PSG9	44.72608203	2.460799793	0.63693701	3.86349006	0.00011178
IGFL1	148.0335765	–3.287593779	0.68100987	–4.8275273	1.38E–06
SIGLEC11	4.632318214	2.031872984	0.57918796	3.5081409	0.00045125
IGLV1-51	22.07888685	–2.409881302	0.60716704	–3.9690582	7.22E–05
IGLV3-19	127.9767079	–3.04369407	0.75298639	–4.0421635	5.30E–05
IGLV2-18	56.67811333	–4.917737728	0.91828702	–5.3553384	8.54E–08
PIWIL3	53.13670612	3.216643423	0.90951327	3.53666463	0.00040521
MYO18B	116.2046349	2.841135925	0.58540301	4.85329912	1.21E–06
ENTHD1	16.42218903	–2.64700336	0.64662671	–4.0935571	4.25E–05

https://dx.doi.org/10.21037/tau-21-158

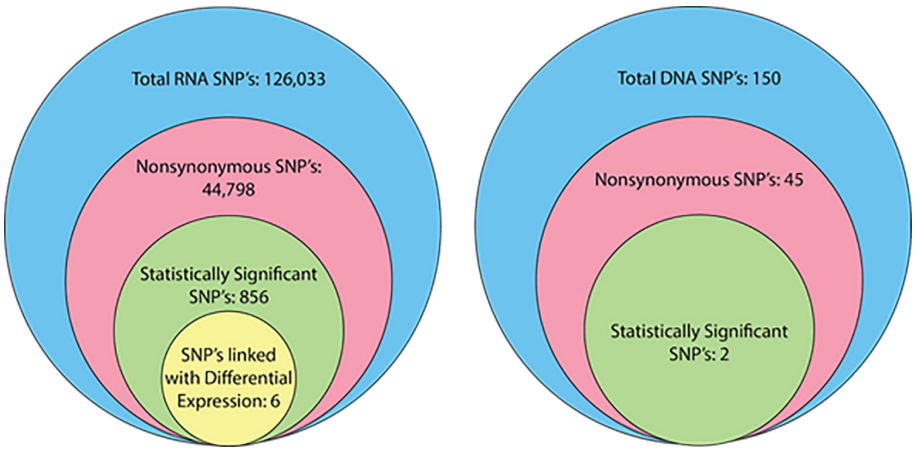


Figure S1 Pathway analysis of genes with statistically significant differential expression between responders and non-durable responder in the MIF regulation of innate immunity pathway. Pathway analysis was performed using IPA software. IPA allows for integration of gene expression results to identify and visualize associated biological pathways. Components highlighted in purple showed differential expression on RNAseq in our dataset. MIF, macrophage migration inhibitory factor; IPA, Ingenuity Pathway Analysis.

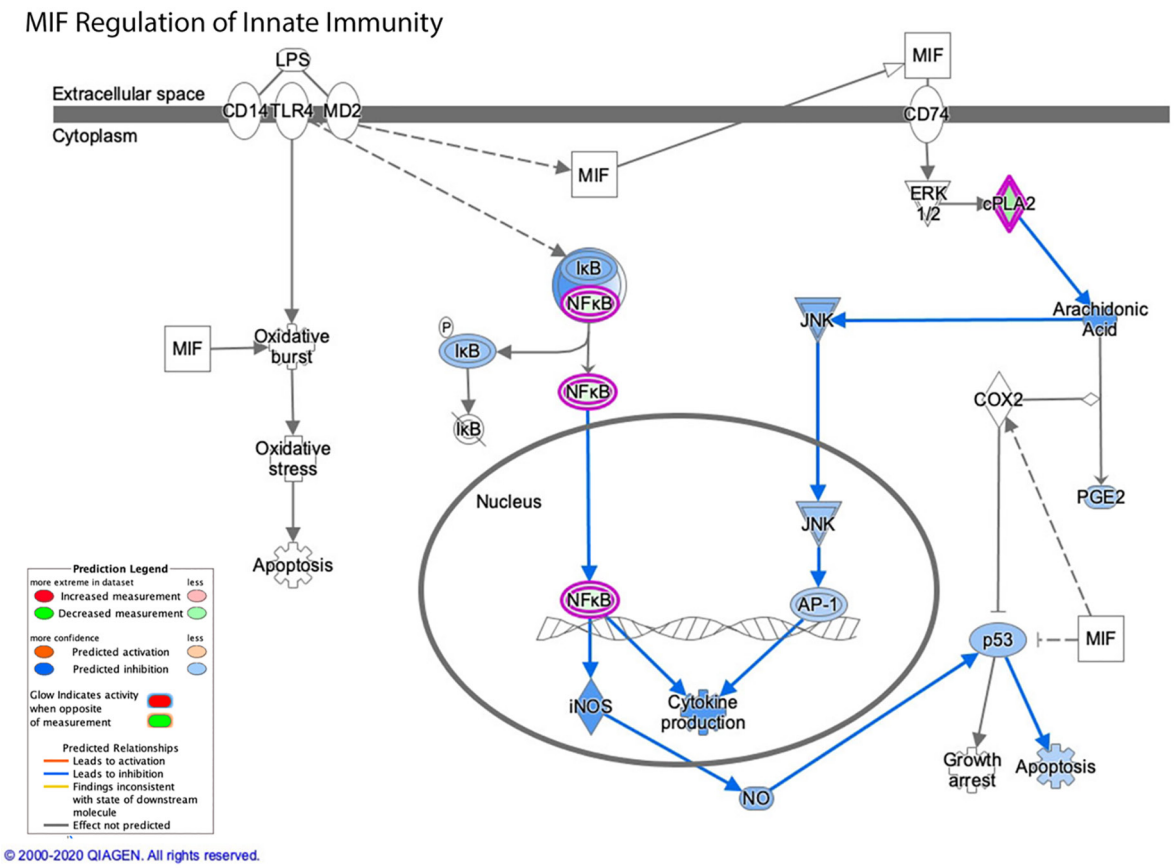


Figure S2 IPA for the MIF-mediated glucocorticoid regulation pathway. IPA, Ingenuity Pathway Analysis; MIF, macrophage migration inhibitory factor.

https://dx.doi.org/10.21037/tau-21-158

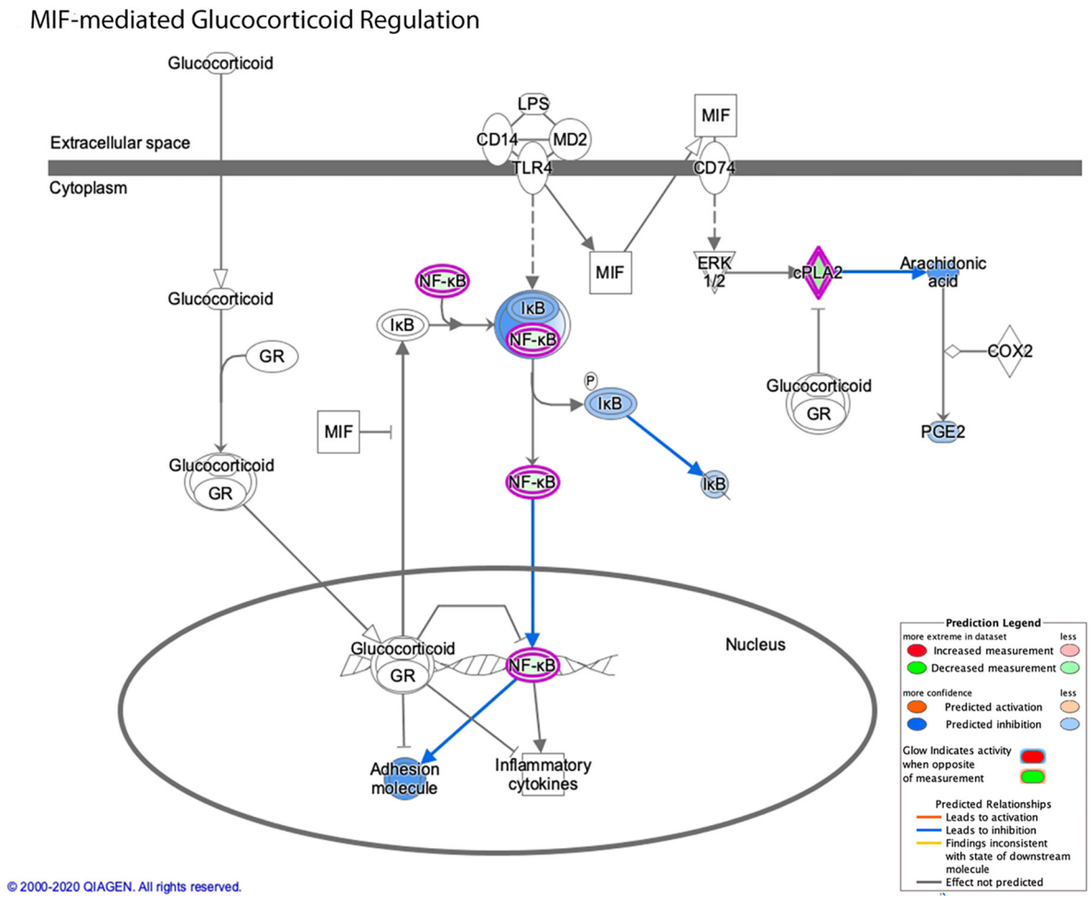


Figure S3 IPA for the p38 MAPK signaling pathway. IPA, Ingenuity Pathway Analysis.

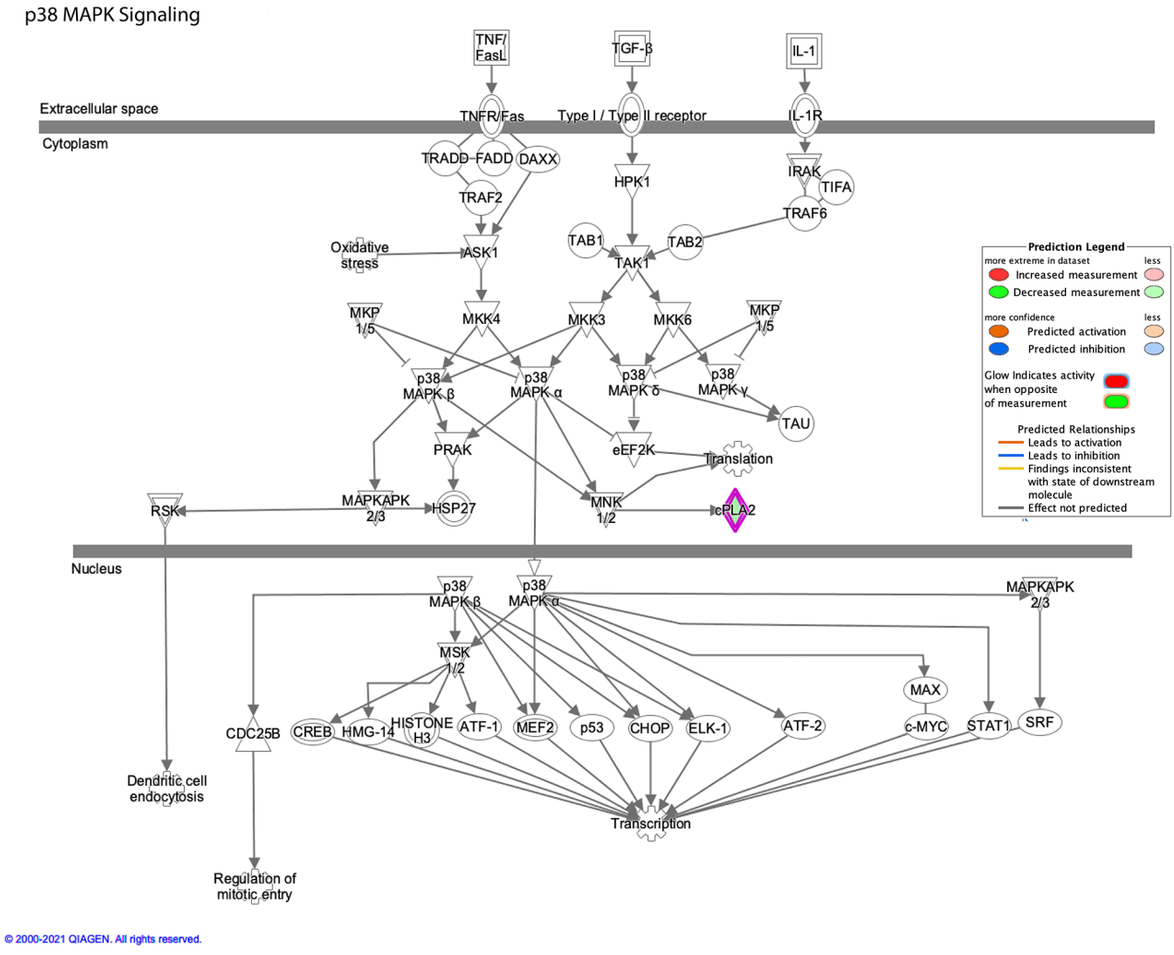


Figure S4 Diagram showing number of SNPs remaining after each filtering step. (A) The top blue circle shows the total number of raw RNA SNPs called, followed by numbers after filtering for nonsynonymous only SNPs, statistically significant SNPs, and then SNPs that were in a gene that was also considered differentially expressed between groups. (B) The same process is depicted for DNA SNPs. SNPs, single nucleotide polymorphisms.

https://dx.doi.org/10.21037/tau-21-158

Table S2 Specific RNA variants that can be linked to differentially expressed genes on RNAseq

Gene	Chromosome	Position	Reference	Alternate	Type	Expression (log2 fold change)	Ratio in responders	Ratio in non-responders	P value (raw)
JCHAIN	Chr 4	70656441	TACG	T	Nonframeshift substitution	–2.082	0/17	11/20	<0.001
S100A7	Chr 1	153458930	C	G	Nonframeshift substitution	–3.534	1/17	9/20	0.026
CLEC2B	Chr 2	9854422	A	AT	Frameshift subsitution	–2.116	2/17	10/20	0.041
ANXA10	Chr 4	168162543	A	C	Nonframeshift substitution	–2.588	3/17	12/20	0.046

Change in expression is represented as log2-fold change in comparison to the entire durable responder group. The ratio of each RNA variant in the durable responders versus non-durable responders is compared using Fisher's exact test.

Table S3 DNA variants identified using the *TST170* gene panel that differed between durable responders and non-durable responders

Gene	Chromosome	Position	Reference	Alternate	Type	Ratio in responders	Ratio in non-responders	P value (raw)
MCL1	Chr 1	150579476	C	–	Frameshift Deletion	10/17	3/20	0.007
Msh-6	Chr 2	47803501	C	–	Frameshift Deletion	7/17	2/20	0.034

Ratios were compared using Fisher's exact test.

https://dx.doi.org/10.21037/tau-21-158