AWARD NUMBER: W81XWH-17-1-0232

TITLE: Forward Genetic Screen to Identify Novel Therapeutic Entry Points of an Autism Spectrum Disorder

PRINCIPAL INVESTIGATOR: Jimmy L Holder, Jr., MD, PhD

CONTRACTING ORGANIZATION: Baylor College of Medicine, Houston, TX

REPORT DATE: November 2020

TYPE OF REPORT: Final

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
November 2020	Final	01Aug2017-31Jul2020

4. TITLE AND SUBTITLE

Forward Genetic Screen to Identify Novel Therapeutic Entry Points of an Autism Spectrum Disorder

5a. CONTRACT NUMBER
W81XWH-17-1-0232

5b. GRANT NUMBER
AR160091

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Jimmy L. Holder, Jr., MD, PhD

E-Mail: holder@bcm.edu

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Baylor College of Medicine
Department of Pediatric Neurology
Jan and Dan Duncan Neurological Research Institute
1250 Moursund St., Suite 925
Houston, TX 77030-3411

8. PERFORMING ORGANIZATION REPORT

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

During the life of this grant, we have successfully performed a cell based CRISPR/Cas9 screen to identify regulators of Shank3 protein stability from three sub-libraries, kinases/phosphatases, G-protein coupled receptors and ubiquitin-related proteins. We have performed a secondary screen with siRNA knockdown of the most promising hits from the primary screen. Several of these hits were validated in the secondary screen. For one of the kinase hits, a small molecule inhibitor of the hit was found to also increase Shank3 abundance. Thus, we have successfully proven the validity of the approach for identifying regulators of protein stability for Shank3. We will next take our most promising validated hits to a mouse in vivo system to confirm rescue of behavioral and molecular phenotypes.

15. SUBJECT TERMS
Autism, SHANK3, flow cytometry, primary neurons, druggable genome

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
					USAMRMC
a. REPORT	b. ABSTRACT	c. THIS PAGE			19b. TELEPHONE NUMBER *(include area code)*
Unclassified	Unclassified	Unclassified	Unclassified	13	

Table of Contents

Introduction

The goal of this work is to identify therapeutic entry points for the Autism Spectrum Disorder caused by haploinsufficiency of the *SHANK3* gene, also known as Phelan-McDermid Syndrome. The experimental system we are using is a cell-based high-throughput screen to identify genetic modifiers of SHANK3 protein stability. Utilizing engineered cells expressing a GFP-tagged SHANK3 and a second fluorescent reporter (DsRed) as an internal control, we disrupt the expression of genes of the druggable genome to identify those that when suppressed lead to increased SHANK3 protein abundance. Those candidate hits that can be confirmed by a secondary method will then be further tested in primary cortical neurons from mice for validation. The best candidates that are validated in our secondary screen will then ultimately be tested in a genetic interaction experiment for increasing Shank3 abundance in mice haploinsufficient for *Shank3*.

Keywords

Autism, SHANK3, flow cytometry, primary neurons, druggable genome

Accomplishments

The major goals of the project are

1. Perform a cell based screen for genetic modifiers of SHANK3
2. Bioinformatic prioritization of candidates and confirmation of direct interaction
3. *In vivo* confirmation of candidate genetic modifiers of SHANK3

Since the beginning of the project, significant progress has been made toward accomplishing the major goals. All sub-libraries of the druggable genome (kinase/phosphatase, G-protein coupled receptors and ubiquitin) have been screened against the DsRed-ires-EGFP:Shank3 cell line, and the cells with the top 10% and bottom 10% GFP to DsRed ratios sorted from the bulk cells. Four replicates of each sub-library were performed. We isolated genomic DNA from all of these 12 replicate experiments. The indexing PCRs were completed for each replicate. The PCR libraries were sent for next generation sequencing and bioinformatics analysis has been performed to identify genes which when depleted result in a significant increase in the ratio of GFP:DsRed (thus an increase in SHANK3 abundance after normalizing to transcriptional effects on the reporter transgene).

Quality control of the next generation sequencing was performed for each sub-library to confirm adequate sequencing coverage. Quality control included evaluating each PCR generated library for total reads, mapped reads and sequence coverage. Each PCR library was then compared to the others within a sub-library of the druggable genome guides by principal component analysis to determine if the PCR libraries grouped together. The sequencing quality control for each sub-library: kinases and phosphatases (Table I), G-protein coupled receptors (Table II) and Ubiquitin related proteins (Table III) are below. Overall, the sequence quality and coverage for the sorted cells was excellent.

Table I: Quality Control for Kinase/Phosphatase Library

sample	Mapped reads	Total reads	% mapped	Coverage
Base_1	11986149	14227759	84.24	2682
Base_2	11190500	13339664	83.89	2504
Base_3	9728278	11642744	83.56	2177
Base_4	11195734	13304223	84.15	2505
High_1	14104270	16823339	83.84	3156
High_2	12216546	14639230	83.45	2734
High_3	13013201	15551477	83.68	2912
High_4	12099660	14443863	83.77	2707
Low_1	11053443	13181320	83.86	2473
Low_2	12600435	15118313	83.35	2820
Low_3	12198485	14612748	83.48	2730
Low_4	11637377	13927533	83.56	2604

Table II: Quality Control for G-protein Coupled Receptor Library

sample	Mapped reads	Total reads	% mapped	Coverage
Base_1	4742114	5592357	84.80	2366
Base_2	4584389	5421864	84.55	2288
Base_3	3651316	4389271	83.19	1822
Base_4	5369547	6358685	84.44	2679
High_1	5561585	6596029	84.32	2775
High_2	4246326	5111826	83.07	2119
High_3	4889696	5781522	84.57	2440
High_4	4665773	5508976	84.69	2328
Low_1	5255169	6228055	84.38	2622
Low_2	4731541	5622638	84.15	2361
Low_3	5103544	6075788	84.00	2547
Low_4	4719625	5616096	84.04	2355

Table III: Quality Control for Ubiquitin-related Protein Library

sample	Mapped reads	Total reads	% mapped	Coverage
Base_1	19665837	27006838	72.82	2306
Base_2	21208925	25787011	82.25	2487
Base_3	22013777	26596578	82.77	2581
Base_4	22696792	27384821	82.88	2661
High_1	24866529	29905492	83.15	2916
High_2	25155756	30408705	82.73	2949
High_3	23454806	28232506	83.08	2750
High_4	21371641	25686490	83.20	2506
Low_1	22760355	27336630	83.26	2669
Low_2	23461675	28174157	83.27	2751
Low_3	21900556	26612701	82.29	2568
Low_4	22971341	27741335	82.81	2693

We next performed principal component analysis for each sub-library to determine degree of concordance among the low GFP/DsRed, high GFP/DsRed and bulk (base) libraries. The graphs for principal component analysis for each sub-library: kinases and phosphatases (Figure 1), G-protein couple receptors (Figure 2) and Ubiquitin related proteins (Figure 3) are below. Overall, there was excellent concordance among the replicates.

Figure 1: Principal Component Analysis for Kinase/Phosphatase Library

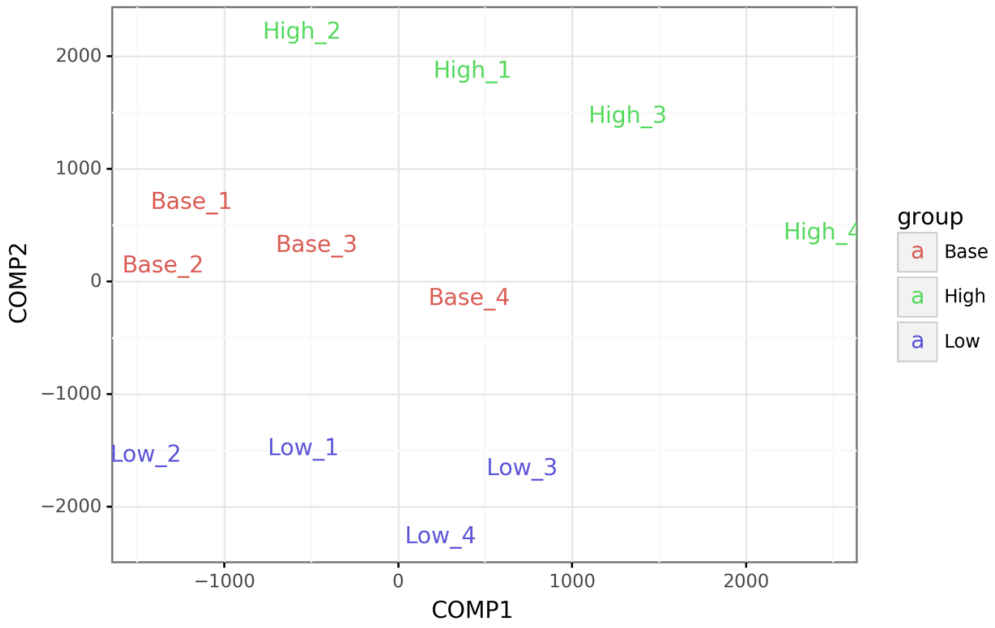


Figure 2: Principal Component Analysis for G-protein Coupled Receptor Library

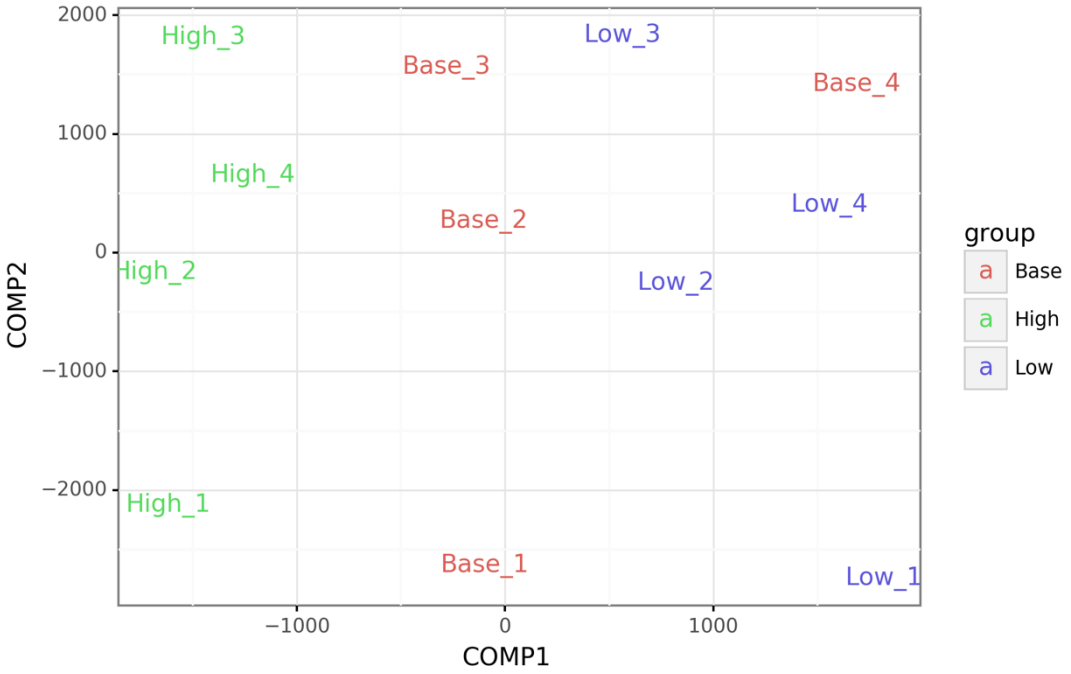
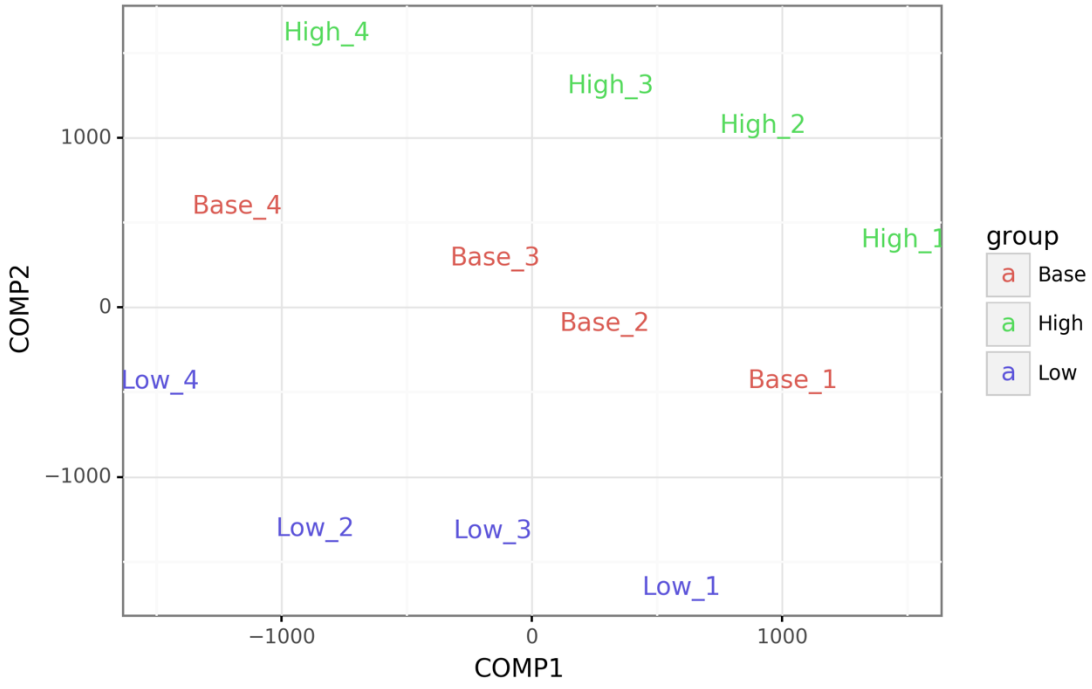


Figure 3: Principal Component Analysis for Ubiquitin-related Protein Library



We then analyzed the fold change in the ratio of GFP:Shank3/DsRed first guide by guide, then combining the guides to give a fold change in the ratio gene by gene. We have focused on those genes which when knocked down cause an increase in GFP:Shank3 abundance (High-Positive). We then ranked the genes beginning with those with the smallest false discovery rate (FDR) value. We chose to use FDR for our statistical analysis because it is better suited to the multiple comparisons we are making. Table IV lists the kinases/phosphatases with increased GFP:Shank3 expression and a FDR<0.01. Table V lists primary hits from the GPCR sub-library with an FDR < 0.05 and Table VI are the top 25 ubiquitin related genes from the primary screen with an FDR<<0.01

Table IV: Kinases and Phosphatases

Gene	log2FC:High	FDR_pos:High
CDK8	0.618811136	9.39E-11
NT5C3A	0.137664289	8.26E-07
NEK9	0.184628704	2.27E-05
CSNK1A1	0.398427106	0.00021823
PPP2R2A	0.577667901	0.000220517
FGFR3	0.043142468	0.000349231
SGPP1	0.080830387	0.000448681
DGKQ	0.135630316	0.000449597
CDK13	0.133347319	0.00154552
PI4K2B	0.044645108	0.001691471
PRKDC	0.097569413	0.004168655
STK31	-0.091174953	0.004214097
CSNK2B	0.479208179	0.005401772
DCLK2	0.00857118	0.00563236

Table V: GPCRs

Gene	log2FC:High	FDR_pos:High
LTB4R2	0.091149952	8.15E-07
TAPT1	0.180633718	8.15E-07
ADRB2	0.148524601	0.000916856
S1PR5	0.051316669	0.00599808
GPR78	0.050760479	0.019023492
MLNR	0.048208245	0.019023492
P2RY2	0.079345699	0.022389708
MTNR1A	0.056868486	0.036172028
PTGER3	0.105426924	0.042090944

Table VI: Ubiquitin-related Genes

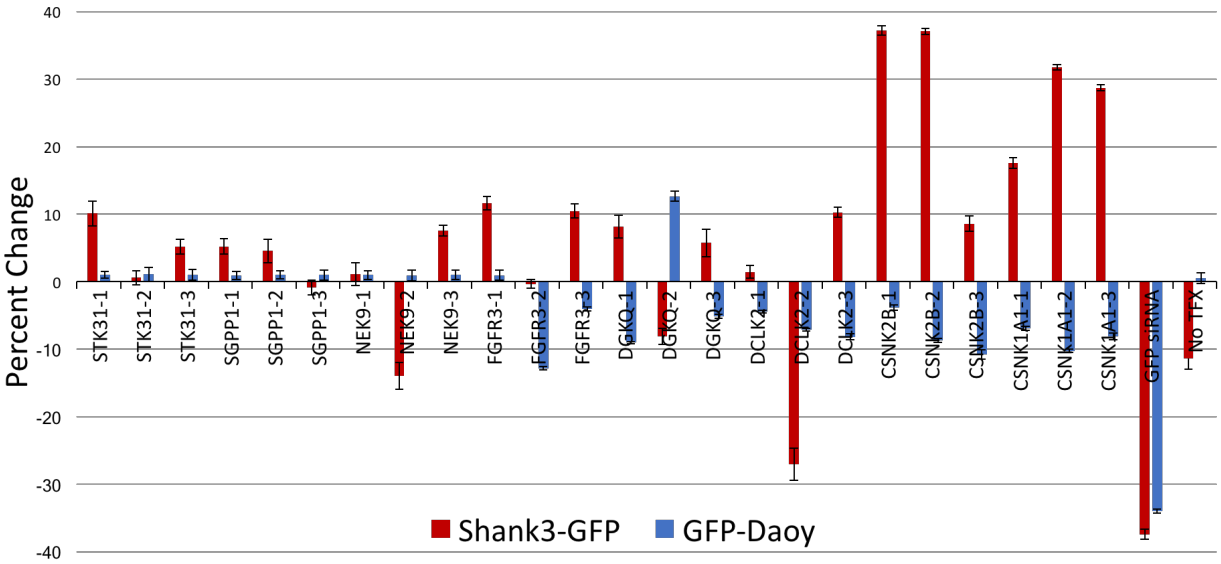
Gene	log2FC:High	FDR_pos:High
UBE4B	0.530172732	6.89E-17
FBXW11	0.83241468	1.47E-10
CUL1	0.579154746	4.67E-08
PCGF3	0.211035953	5.42E-08
DCAF12	0.305627556	1.08E-07
WDR59	0.415084536	1.08E-07
UBE3C	0.220693391	1.51E-07
PSMD4	0.735634822	1.72E-07
UBR5	0.349985311	5.69E-07
RFFL	0.231712543	8.97E-06
NUP43	0.31448088	1.92E-05
RNF213	0.128465774	1.96E-05
TRIM47	0.129991636	2.55E-05
PSMD2	0.317375073	2.97E-05
PCGF2	0.145772848	6.25E-05
HERC5	0.170592937	6.43E-05
KLHL11	0.273952797	0.000102183
HERC3	0.16637654	0.000144277
NSFL1C	0.162132849	0.000144277
GNB2	0.147867005	0.000172544
HERC6	0.225759259	0.000172544
BPTF	0.170874113	0.000181023
DCUN1D4	0.186038918	0.000181023
SMURF1	0.146698556	0.000184577
LNX1	0.087792985	0.00022974

We have started our validation of primary hits from the screen. For each primary hit gene, we obtain three independent siRNAs targeting the human gene. We then transfect these siRNAs back to the engineered cell line expressing DsRed-ires-EGFP:Shank3 and perform flow analysis to validate increased expression of EGFP:Shank3. To ensure the increase in EGFP:Shank3 abundance is not due to an effect on only EGFP, we also transfect the same siRNAs into a control cell line expressing DsRed-ires-EGFP (negative control). We have completed flow cytometry validation of genes encoding all three sub-libraries for which siRNAs are available. Secondary screening for GPCRs identified from the primary screen failed to identify any genes which when inhibited significantly increase Shank3 abundance. Thus, for the remainder of this report, we will focus on hits from the kinase/phosphatase and ubiquitin libraries.

Kinases/Phosphatases

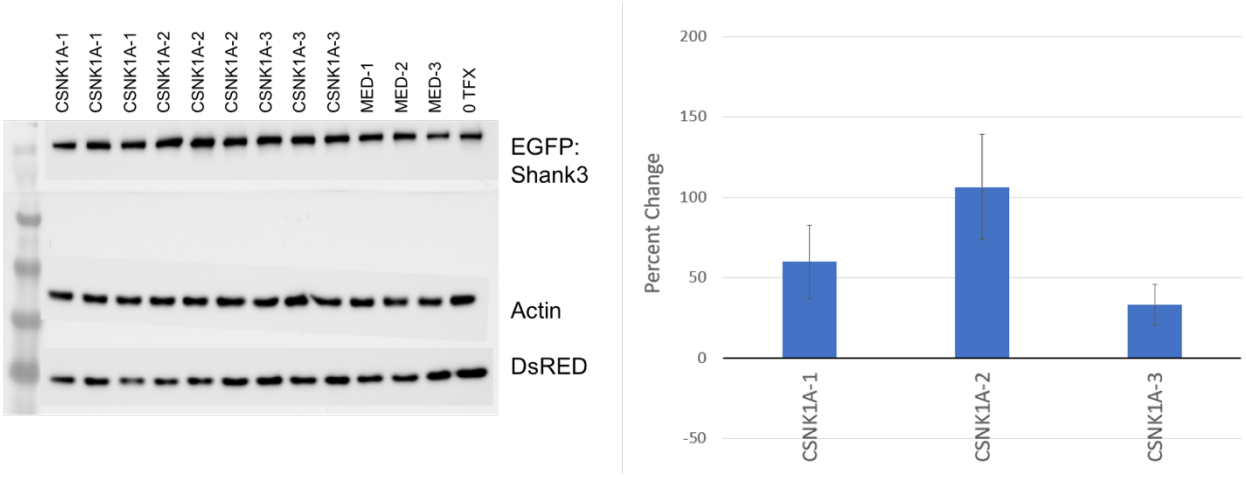
We began by validating primary hits from the kinase/phosphatase sub-library. From this experiment, the most consistently robust increase in EGFP:Shank3 abundance occurred with depletion of Casein Kinase I alpha (CSNK1A1) and Casein Kinase II beta (CSNK2B). The siRNAs targeting these genes did not result in a significant increase in EGFP expression in the negative control cell line (DsRed-ires-EGFP). Additionally, our positive control (GFP siRNA) effectively depleted both EGFP (negative control line) and EGFP:Shank3 (test line).

Figure 4: Flow cytometry validation of primary kinase and phosphatase hits



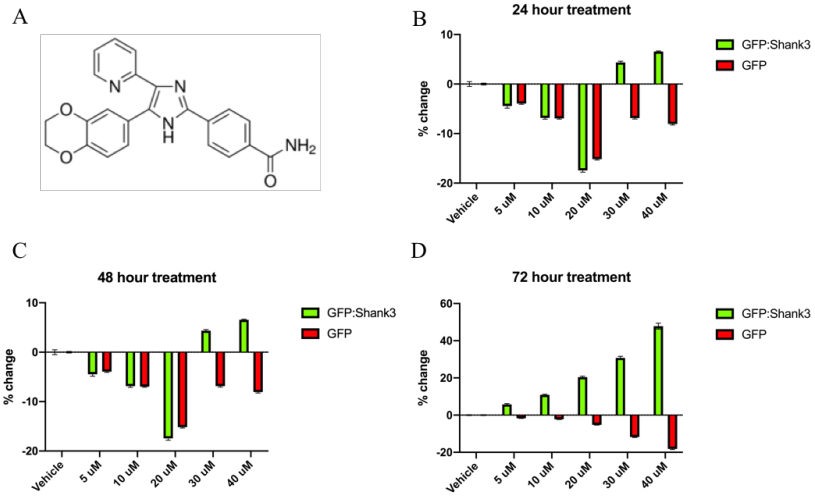
For those genes that validate by flow cytometry, we then perform Western blotting from the same cell line. We have completed evaluation of *CSNK1A1* depletion in EGFP:Shank3 cells by Western blotting (Figure 5). Similar to our flow cytometry results, we see a significant increase in the abundance of EGFP:Shank3 when *CSNK1A1* is depleted.

Figure 5: Western blot verification of CSNK1A



To further validate Casein Kinase I alpha as a regulator of Shank3 protein stability, we utilized a small molecule inhibitor of Casein Kinase I, D4476. D4476 is a cell-permeable, high affinity allosteric inhibitor of Casein Kinase I (Figure 6A). Previous studies have determined that in cells in culture, D4476 linearly inhibits CKI alpha from 5 to 40uM. We thus tested whether D4476 could increase Shank3 abundance in cultured cells. Daoy cells expressing either our reporter construct (DsRed-ires-EGFP:Shank3) or our negative control line (DsRed-ires-EGFP) were treated for 24 to 72 hours with 5 to 40 micromolar of D4476 or with vehicle. Cells then underwent flow cytometry as for siRNA knockdown of CK1α. We found that at 30 and 40 micromolar of D4476, Shank3 abundance was significantly increased as determined by measuring EGFP to DsRed ratios (Figures 6B-D). In contrast, for the negative control line, all concentrations of D4476 and all durations of treatment with D4476 resulting in significant decreases of EGFP to DsRed ratios. This confirms that the change in the EGFP to DsRed ratio in our reporter line is due to an effect on Shank3 and not on EGFP. Intriguingly, lower concentrations of D4476 (5-20 micromolar at 24 and 48 hours of treatment duration) cause a significant reduction in EGFP:DsRed by flow cytometry. The reason for this difference compared with longer duration of treatment is currently unclear but warrants further investigation.

Figure 6: D4476 increases Shank3 abundance

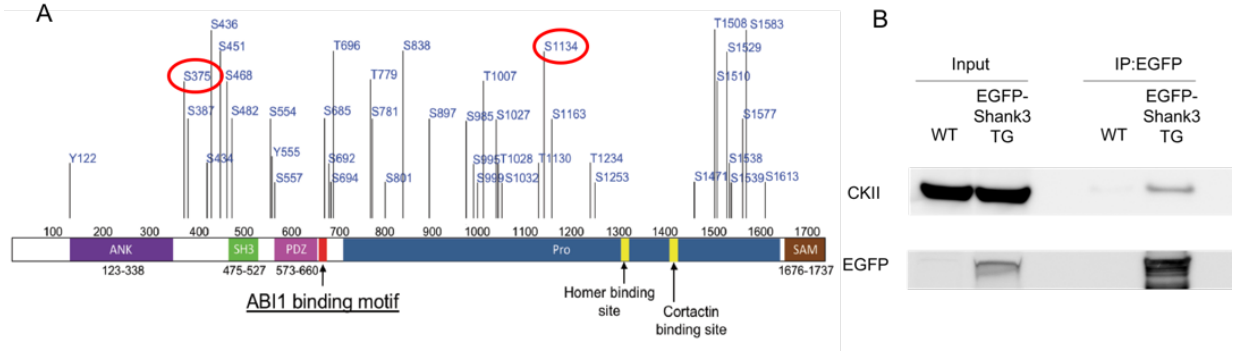


Next, we will deplete Casein Kinase I alpha in mouse cortical neurons to confirm its regulation of Shank3 *in vivo*. We have obtained three lentivirus plasmids (Transomic.com) that target *Csnk1A1* as well as a scrambled control. We generated lentivirus from these plasmids and infected mouse HT22 cells with these four plasmids. After selection for infected cells with puromycin, we isolated RNA from all four infections and measured abundance of *Csnk1A1* mRNA by quantitative Polymerase Chain Reaction (qPCR). We determined that all three significantly deplete *Csnk1A1* from 51-78% in this assay. We are now poised to confirm that depletion of *Csnk1A1* in mouse neurons leads to increased Shank3 levels.

In order to determine if Casein Kinase I and Casein Kinase II directly interact with and phosphorylate Shank3, we have performed biochemical phosphorylation and interaction experiments beginning with Casein Kinase II. First, we biochemically purified bacterially expressed fragments of Shank3 to use as a substrate for Casein Kinase II. We then performed *in vitro* phosphorylation assays with the recombinant Shank3 fragments and enzymatically active Casein Kinase II. The products of these assays then underwent mass spectrometry to identify which residues are phosphorylated by Casein Kinase II *in vitro*. Not all phosphorylation events that can occur *in vitro* occur *in vivo*. To determine which phosphorylation events are physiologically relevant, we mapped the *in vitro* phospho-sites back to the *in vivo* phospho-sites we previously identified from mouse brain (Figure 7A). We have identified S375 and S1134 as phosphorylation sites for Casein Kinase II.

In order to further confirm that Casein Kinase II directly interacts with Shank3, we have performed *in vivo* immunoprecipitation against EGFP from brains of mice expressing the transgenic EGFP-tagged Shank3. Following immunoprecipitation, we then blotted against a subunit of Casein Kinase II, Csnk2A1 that forms part of the tetramer to make the active enzyme by binding to Csnk2B. We found that by immunoprecipitation, Shank3 and Casein Kinase II do form a complex in mouse brain (Figure 7B).

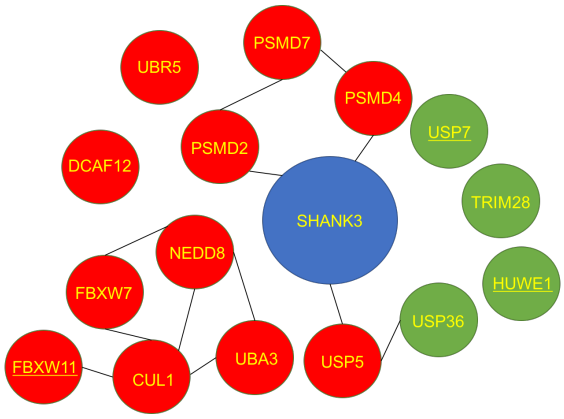
Figure 7: CSNK2A1 and Shank3 form a protein complex



Ubiquitin-related proteins

From our primary screen of the ubiquitin sub-library, we identified numerous ubiquitin-related proteins that raise Shank3 abundance. We investigated whether any of these ubiquitin-related proteins form a complex that could act in concert to regulate Shank3 protein stability (Figure 8). From the primary screen, we identified multiple core proteasomal subunits (PSMD 2, 4, 7) as potential negative regulators of Shank3 protein stability. Additionally, we identified a

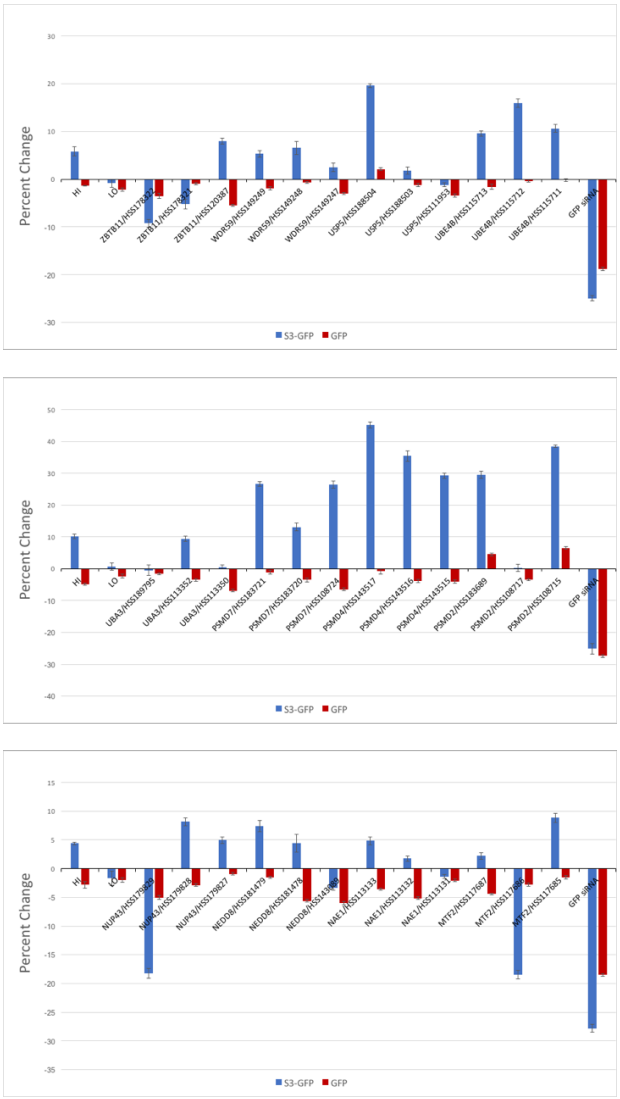
Figure 8: Ubiquitin-related proteins regulating Shank3



FBox/NEDD8/Cullin complex as potentially critical for Shank3 stability. In addition to these negative regulators of Shank3 stability (increased Shank3 abundance with their depletion), we also identified positive regulators of Shank3 stability (green in Figure 8). These discoveries could be critical in understanding the complex regulation of Shank3 protein stability.

Next, we performed a secondary validation of primary hits from the ubiquitin-related gRNA library by siRNA knockdown in our reporter (DsRed-ires-EGFP:Shank3) and negative control (DsRed-ires-EGFP) cell lines for those genes which commercial siRNAs are available (Figure 9). We were able to validate several of our hits including members of the core proteasome machinery (PSMD2, PSMD4 and PSMD7 which when inhibited resulted in the greatest increase in Shank3. For many of the ubiquitin-related proteins, we could not purchase pre-designed siRNAs and will need to design these ourselves.

Figure 9: SiRNA validation of Ubiquitin-related proteins regulating Shank3



In conclusion, we have successfully completed a CRISPR/Cas9 cell-based screen for genetic modifiers of Shank3 protein stability. From this primary screen, we have identified multiple candidate hits, genes which when inhibited increase Shank3 abundance. From these hits, we performed a secondary validation screen using siRNA to deplete the hits and again measure Shank3 abundance in our reporter cell line as well as a negative control cell line to control for hits that regulate EGFP and not Shank3. We found approximately 10% of our primary hits resulted in significant (>10%) increase in Shank3 abundance in our secondary screen.

We focused on two kinases for further validation, Casein Kinase I and Casein Kinase II. In particular, for Casein Kinase I, we found that a small molecule inhibitor, D4476, can also robustly increase Shank3 abundance. We are now poised to further investigate the impact of

inhibition of this kinase in cultured neurons and whole animals (mice). We have also identified a number of ubiquitin-related proteins that positively and negatively regulate Shank3 protein stability. Investigating the interplay between the kinases and ubiquitin-related proteins in regulated Shank3 stability is critical for understanding the mechanisms of Shank3 protein stability regulation.
In this grant period, we have focused exclusively on those proteins which when inhibited increase Shank3 abundance (negative regulators). Investigating positive regulators or those proteins which when inhibited lead to reduced Shank3 abundance will lead to better understanding of Shank3 regulatory mechanisms.

Disseminating results: The preliminary results from this study were presented at Neuroscience 2019 in Chicago, IL.

Impact
There is nothing to report yet.

Changes/Problems
No major changes to the main objectives or approaches have occurred.
The COVID-19 pandemic significantly impacted our progress in the last half of our final year of this award. BCM severely limited access to research laboratories from March until May to allow only for maintenance of equipment. Since May of 2020, BCM began slowly reopening laboratories for limited research activities until this summer when full access was again allowed; however, this was at the end of the project period. Because of this, we did not fully reach all of the ambitious milestones we had aimed to achieve. Nevertheless, we did succeed in completing the primary screen, validate multiple hits with a secondary screen, identify a small molecule inhibitor of one of our hits that increases Shank3 abundance and biochemically verify that one of the kinases that we identified as a regulator of Shank3 physically binds and phosphorylates it. Thus, this pilot award has provided us the opportunity to develop the preliminary data necessary to apply for a NIH grant to take our work to an *in vivo* system.

Products
Nothing to report.

Participants and other collaborating organizations
Jimmy Holder, MD/PhD: no change
Lunhui Lin, PhD: no change

Other organizations: Nothing to report.

Special Reporting Requirements
Nothing to report.

Appendices
Nothing to report.