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 77030

REPORT DATE: August 2019

TYPE OF REPORT: Annual

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 |
|---|---|---|
| August 2019 | Annual | 1 Aug 2018 - 31 Jul 2019 |

**4. TITLE AND SUBTITLE**

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

**5a. CONTRACT NUMBER**

**5b. GRANT NUMBER**
W81XWH-17-1-0232

**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 reporting period, the PCR libraries from the primary screen were sequenced. The sequencing data sets then underwent quality control. Bioinformatics of the sequencing data identified those genes with significantly increased abundance of EGFP:Shank3. The positive hits from the kinase/phosphatase sub-library were then tested in a secondary screen with siRNA depletion of those hits and measurement of EGFP:Shank3 abundance by flow cytometry. The two most robustly increased gene from the kinase/phosphatase sub-library are now undergoing validation by Western blotting. These same two candidate hits will be tested for Shank3 protein regulation by knockdown in primary mouse neurons. We have already discovered that the protein complex for one of these hits interacts with Shank3 in vivo.

**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 | 12 | |

**Standard Form 298 (Rev. 8-98)**
**Prescribed by ANSI Std. Z39.18**

**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 1), 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 |

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

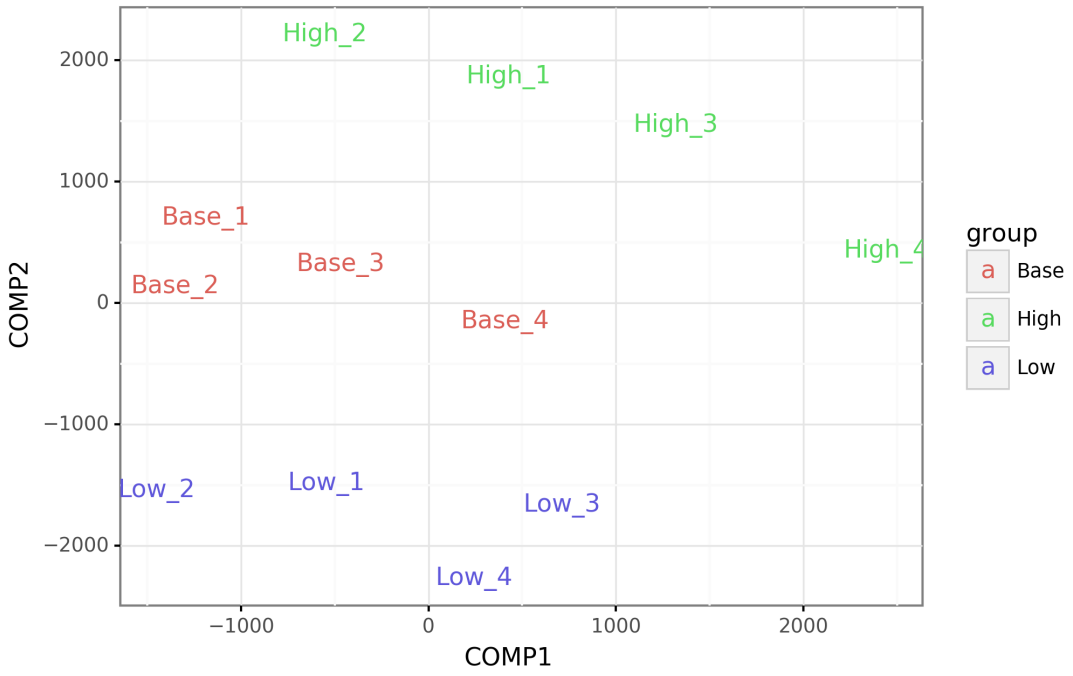


Figure 2: Principal Component Analysis for G-protein Coupled Receptor Library

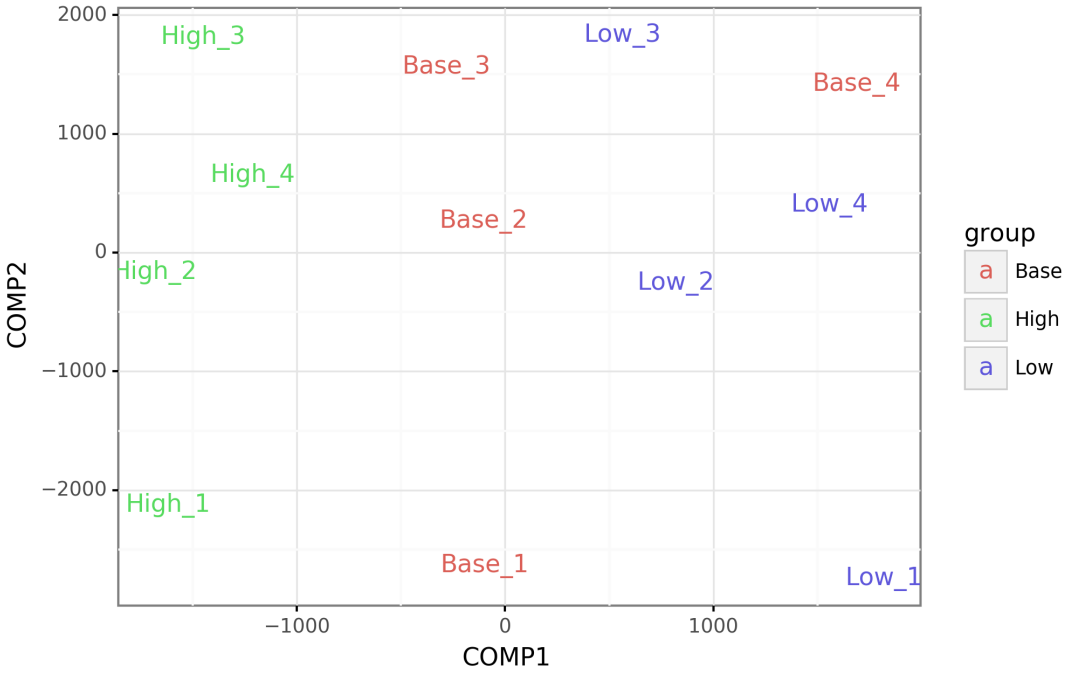
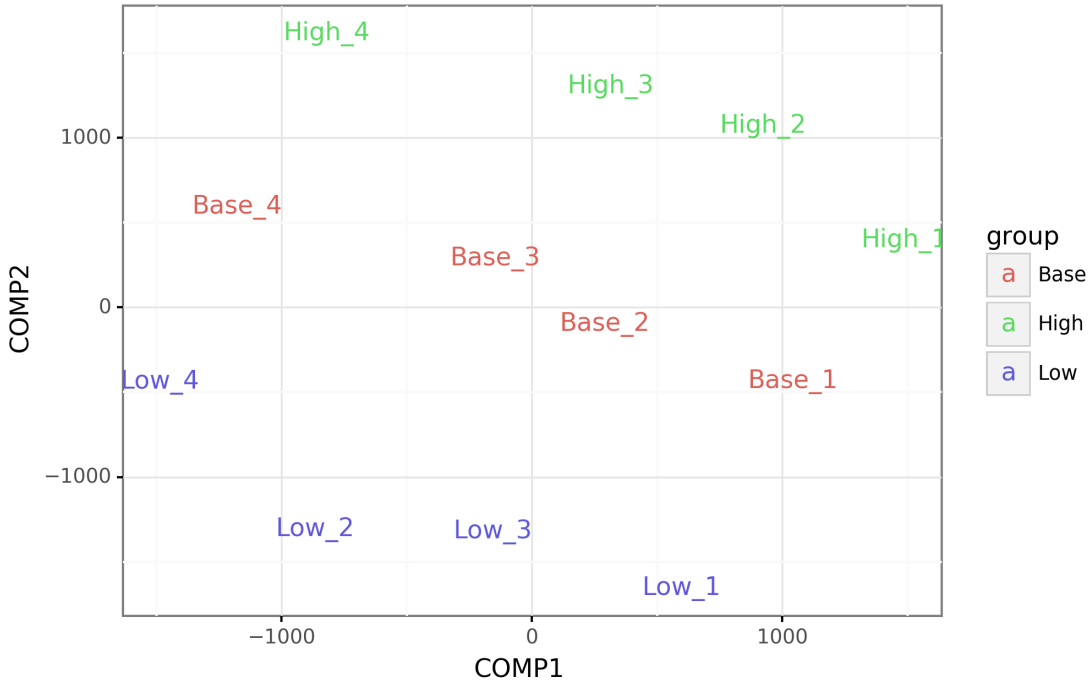



Figure 3: Principal Component Analysis for Ubiquitin-related Protein Library



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

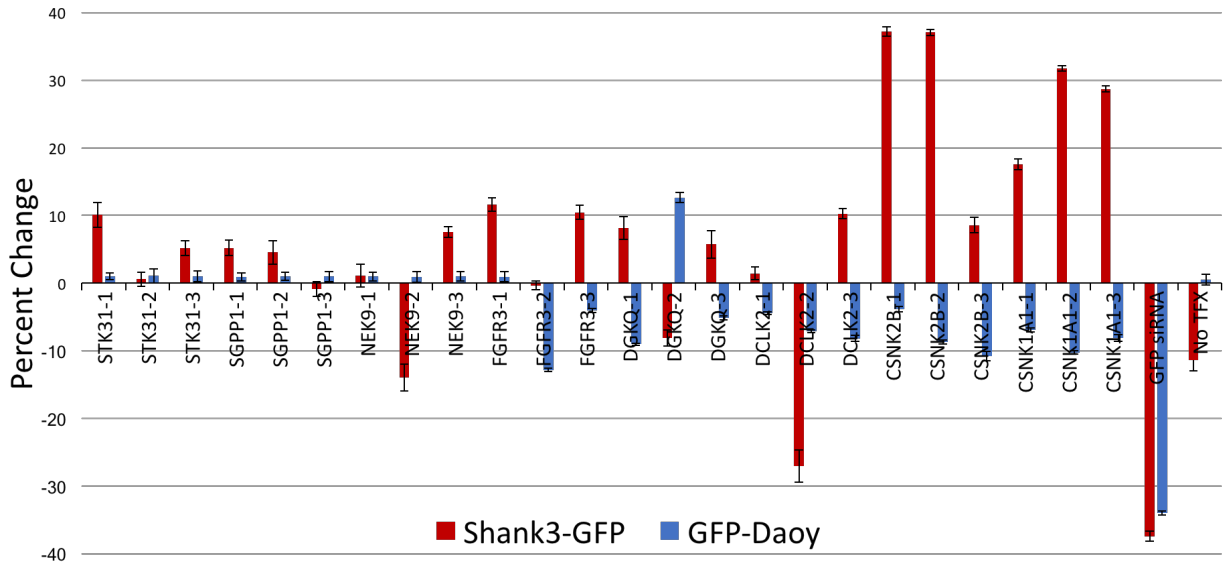
| Gene | log2FC:High | FDR_pos:High |
|---|---|---|
| LTB4R2 | 0.091149952 | 8.15E-07 |
| TAPT1 | 0.180633718 | 8.15E-07 |
| ADRB2 | 0.148524601 | 0.000916856 |
| S1PR5 | 0.051316669 | 0.00599808 |
| GPR78 | 0.050760479 | 0.019023492 |
| MLNR | 0.048208245 | 0.019023492 |
| P2RY2 | 0.079345699 | 0.022389708 |
| 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 DsRediresEGFP: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 DsRediresEGFP. We have completed flow cytometry validation of genes encoding kinases and phosphatases for which siRNAs are available.

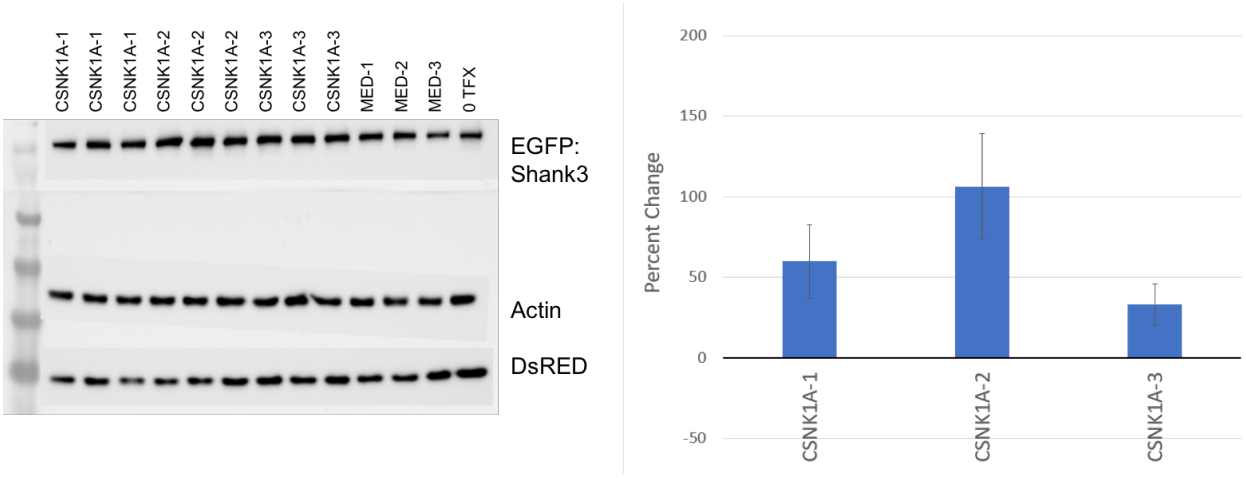
Figure 4: Flow cytometry validation of primary kinase and phosphatase hits




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 (DsRediresEGFP). Additionally, our positive control (GFP siRNA) effectively depleted both EGFP (negative control line) and EGFP:Shank3 (test line).

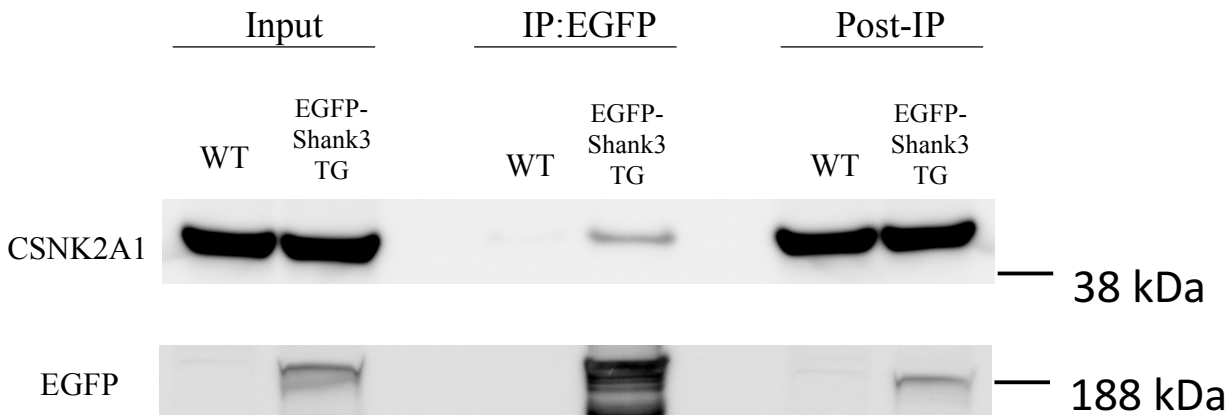
For those genes that validate by flow cytometry, we then perform Western blotting from the same cell line. We have completed evaluation of CSNK1A1 depletion in EGFP:Shank3 cells by Western blotting (Figure 5). Similar to our flow cytometry results, we see a significant increase in the abundance of EGFP:Shank3 when CSNK1A1 is depleted. We are currently validating depletion of CSNK2B as a regulator of Shank3 abundance by Western blotting.

Figure 5: Western blot verification of CSNK1A




Finally, in order to determine if direct interaction of the kinase with Shank3 is necessary to alter protein abundance (likely through direct phosphorylation), we have performed in vivo immunoprecipitation of EGFP:Shank3 from mouse brain and 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 6).

Figure 6: CSNK2A1 and Shank3 form a protein complex




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

For the next reporting period, we will complete our validation of hits for the GPCR and Ubiquitin sub-libraries by siRNA knockdown in our engineered cell line by flow cytometry and

Western blotting. If our experience with the kinase/phosphatase library holds that only approximately 10% of primary hits significantly and robustly regulate Shank3, we will move directly to testing for increased Shank3 abundance following knockdown in primary mouse neurons. We already have three lentiviral vectors for expressing shRNAs against *Csnk1a1*. We will test for knockdown efficiency in a mouse cell line then make lentivirus to knockdown the gene in cultured mouse neurons followed by Western blotting for endogenous Shank3 abundance. We will also test the efficacy of small molecule inhibitor D4476 (a Casein Kinase inhibitor) for increasing Shank3 abundance both in our engineered cell line and in primary mouse cortical neurons. If these experiments for inhibiting Casein kinase 1 or 2 are successful in increasing Shank3 abundance, we will move to evaluate for organism-level rescue as described in my original proposal.
Finally, those hits that are most promising from the tertiary screen will be testing for direct interaction with Shank3 using Bimolecular Fluorescent Complementation (BiFC) or in vivo immunoprecipitation and Western blotting as performed and reported above for Casein kinase 2. This will also be completed during the next reporting period.

**Impact**
There is nothing to report yet as we are now validating our primary screen.

**Changes/Problems**
No major changes to the main objectives or approaches have occurred.
We are currently approximately nine months behind the original proposed timeline. The delay is multifactorial, optimizing the cell sorting for developing the libraries took longer than anticipated. We also needed to re-optimize the flow cytometry validation process due to new personnel since our early preliminary work. However, we already have two promising and partially validated hits for boosters of Shank3 abundance that we are pursuing and anticipating testing at least one in vivo for behavioral rescue in mice in the next reporting period.

**Products**
Nothing to report.

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

Change in <u>other support</u> for PI:
The following grant was funded during the last reporting period. It has not altered the PI's effort on the current project.
19/1(Holder) 12/1/18-11/30/21 0 calendar
SYNGAP Research Fund
**DEVELOPMENT OF IPSC LINES FOR THE PURPOSE OF UNDERSTANDING HOW SYNGAP1 PATIENT MUTATIONS IMPACT SYNGAP PROTEIN LEVELS AND HUMAN NEURON FUNCTION**
Develop induced pluripotent stem cells from patients with a spectrum of variants in *SYNGAP1* to determine the phenotypic spectrum of abnormalities in induced neurons.
Overlap: none

Other organizations: Nothing to report.

**Special Reporting Requirements**
Nothing to report.

**Appendices**
Nothing to report.