

AWARD NUMBER: W81XWH-20-1-0028

TITLE: Systematic Analysis of Genetic Mosaicism in FTD/ALS Brains

PRINCIPAL INVESTIGATOR: Zinan Zhou

CONTRACTING ORGANIZATION: Boston Children's Hospital

REPORT DATE: MARCH 2022

TYPE OF REPORT: ANNUAL

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

DISTRIBUTION STATEMENT: Approved for Public Release;  
Distribution Unlimited

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.

# REPORT DOCUMENTATION PAGE

Form Approved  
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.

<b>1. REPORT DATE</b> MARCH 2022		<b>2. REPORT TYPE</b> ANNUAL		<b>3. DATES COVERED</b> 02/01/2021 – 01/31/2022	
<b>4. TITLE AND SUBTITLE</b>  Systematic Analysis of Genetic Mosaicism in FTD/ALS Brains				<b>5a. CONTRACT NUMBER</b>	
				<b>5b. GRANT NUMBER</b> W81XWH-20-1-0028	
				<b>5c. PROGRAM ELEMENT NUMBER</b>	
<b>6. AUTHOR(S)</b> Zinan Zhou  E-Mail: <a href="mailto:zinan.zhou@childrens.harvard.edu/">zinan.zhou@childrens.harvard.edu/</a>				<b>5d. PROJECT NUMBER</b>	
				<b>5e. TASK NUMBER</b>	
				<b>5f. WORK UNIT NUMBER</b>	
<b>7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)</b>  Boston Children's Hospital 300 Longwood Ave, Boston, MA 02115				<b>8. PERFORMING ORGANIZATION REPORT NUMBER</b>	
<b>9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)</b>  U.S. Army Medical Research and Development Command Fort Detrick, Maryland 21702-5012				<b>10. SPONSOR/MONITOR'S ACRONYM(S)</b>	
				<b>11. SPONSOR/MONITOR'S REPORT NUMBER(S)</b>	
<b>12. DISTRIBUTION / AVAILABILITY STATEMENT</b>  Approved for Public Release; Distribution Unlimited					
<b>13. SUPPLEMENTARY NOTES</b>					
<b>14. ABSTRACT</b> Frontotemporal dementia (FTD) and amyotrophic lateral sclerosis (ALS) are two closely related neurodegenerative diseases with TAR DNA-binding protein 43 (TDP-43) inclusions as a common pathological hallmark. They are characterized as age-related neurodegeneration, yet the mechanisms by which age and genetic risk interact, as well as the ultimate cause of neuronal loss, remains murky. Clinical features of FTD/ALS, including focal onset, stereotyped patterns of spread, and increased risk due to smoking, have suggested to some a role of somatic mutations as causative for some unexplained cases of FTD/ALS. Our major goal is to reveal the potential roles of genetic mosaicism in the etiology and progression of FTD/ALS. Our study identifies clonal and non-clonal somatic mutations in FTD/ALS brains using ultra-deep targeted sequencing and single-cell whole genome sequencing. In the last funding period, we have collected human postmortem brain tissues from several brain banks. We have also tested the targeted panel sequencing and single-cell L1 targeted sequencing approaches and gotten great performance from testing experiments. In the next funding period, we will apply these approaches to all collected human postmortem brain tissues to systematically examine the burden and pattern of various types of somatic mutations in FTD/ALS brains.					
<b>15. SUBJECT TERMS</b> amyotrophic lateral sclerosis, double-strand break, frontotemporal dementia, Indel, Molecular Inversion Probes, retrotransposon insertion, single-cell whole genome sequencing, somatic mutation, single-nucleotide variant, TDP-43					
<b>16. SECURITY CLASSIFICATION OF:</b> Unclassified			<b>17. LIMITATION OF ABSTRACT</b>  Unclassified	<b>18. NUMBER OF PAGES</b>  11	<b>19a. NAME OF RESPONSIBLE PERSON</b> USAMRDC
<b>a. REPORT</b>  Unclassified	<b>b. ABSTRACT</b>  Unclassified	<b>c. THIS PAGE</b>  Unclassified			<b>19b. TELEPHONE NUMBER</b> (include area code)

## TABLE OF CONTENTS

	<u>Page</u>
1. Introduction	4
2. Keywords	5
3. Accomplishments	5
4. Impact	9
5. Changes/Problems	10
6. Products	10
7. Participants & Other Collaborating Organizations	10
8. Special Reporting Requirements	11
9. Appendices	11

## 1. Introduction

Frontotemporal dementia (FTD), a neurodegenerative disorder characterized by behavioral, language, and memory dysfunction, is increasingly recognized to have clinical, genetic and pathological overlap with amyotrophic lateral sclerosis (ALS), a disease in which premature loss of upper and lower motor neurons leads to fatal paralysis, since the seminal discovery of TDP-43 inclusions as a common pathological hallmark in FTD and ALS. FTD and ALS are characterized as age-related neurodegeneration, yet the mechanisms by which age and genetic risk interact, as well as the ultimate cause of neuronal loss, remains murky. Clinical features of FTD/ALS, including focal onset, stereotyped patterns of spread, and increased risk due to smoking, have suggested to some a role of somatic mutation in FTD/ALS. Although many causative genes involved in various pathways have been identified to predispose to both familial and sporadic FTD/ALS over the past two decades, the possibility that clonal mutations of these genes—i.e., present initially in some cells but not all cells—may contribute to some cases of sporadic FTD/ALS has not been evaluated.

Whereas clonal mosaicism creates risk for some diseases, a recent study suggested that nonclonal mutations accumulate in postmitotic neurons with age, and may contribute to age-related degenerative conditions. For example, Cockayne Syndrome (CS) and Xeroderma Pigmentosum (XP), two human progeroid diseases characterized by accelerated aging and premature neuronal loss, exhibit faster accumulation of somatic single-nucleotide variant (sSNV), as well as distinct mutational signatures compared to age-matched controls through single-cell whole genome sequencing (scWGS) of post-mitotic neurons. These results suggest that accelerated accumulation of DNA mutations is a potential mechanism of other forms of neuronal degeneration and emerging evidence has implicated DNA damage in FTD/ALS.

TDP-43 is an aggregation-prone RNA- and DNA-binding protein identified as the major component of the cytoplasmic ubiquitinated inclusions formed in neurons and glial cells of almost all ALS and about 40% of FTD patients. In TDP-43 proteinopathies, the protein is detectable in the nuclei of unaffected neurons but partially cleared from nuclei in neurons containing cytoplasmic aggregations supporting that pathogenesis is driven, at least in part, by a loss of TDP-43 nuclear function. TDP-43 interacts with Fused in Sarcoma (FUS), a DNA/RNA binding protein shown to participate in homologous recombination (HR) and non-homologous end joining (NHEJ) mediated double-strand break (DSB) repair. Mutations of both *TDP-43* and *FUS* have been identified in FTD/ALS patients. TDP-43 and FUS both localize at sites of transcription-associated DNA damage and depletion of either TDP-43 or FUS leads to increased DNA breakage. TDP-43 has also been shown to be a critical component of the NHEJ-mediated DNA double-strand break (DSB) repair pathway. Hence TDP-43 would be normally suppressing transcription-associated DNA damage and DSB, both of which could otherwise increase somatic mutations in post-mitotic neurons and lead to cell death. In addition, *C9orf72* repeat expansion is the most common genetic cause of familial FTD/ALS, accounting for approximately 30% of familial cases. Recent studies have shown that *C9orf72* repeat expansion causes DNA breakage and leads to neuronal cell death. Taken together, these data suggest our hypothesis that clonal and non-clonal somatic mutations could be associated with FTD/ALS, which might arise from both development and postnatal DNA damage that go unrepaired.

To discover the presence of rare clonal mutations in FTD/ALS brains directly, the first goal of our project will be devoted to high-throughput screening of clonal mutations in a panel of 90 neurodegeneration related genes in the prefrontal cortex, motor cortex and cerebellum of a cohort of sporadic FTD/ALS and age-matched normal cases. Clonal mutations of previously identified FTD/ALS genes will be identified and evaluated based on their predicted pathogenicity. Potential crossover between FTD/ALS and other neurodegenerative diseases can also be evaluated because our targeted gene sequencing panel also contains genes that are related to Alzheimer's Disease, Parkinson's Disease and other rare types of dementia. Results of our project will reveal the frequency of pathogenic clonal mutation in FTD/ALS and may change our understanding of the genetics of the disease.

Our second goal is to compare the rate and pattern of non-clonal somatic mutations in neurons of FTD/ALS brains to neurons of normal brains. Various types of mutation, including SNV, Indel, retrotransposon insertion and double-strand break, will be evaluated through single-cell whole-genome and targeted sequencing. These approaches allow us to define ultra-rare mutations at single-cell level. An increased burden of somatic mutation will suggest that DNA damage exceeds the repair capacity. The sources of exogenous and endogenous DNA damage can be identified by analyzing the specific patterns of somatic mutations in FTD/ALS brains, which may potentially reveal the environmental risk factors of FTD/ALS.

## 2. Keywords

Amyotrophic Lateral Sclerosis  
double-strand break  
Frontotemporal Dementia  
Indel  
Molecular Inversion Probes  
retrotransposon insertion  
single-cell whole genome sequencing  
somatic mutation  
single-nucleotide variant  
TDP-43

## 3. Accomplishments

### What were the major goals of the project?

**Specific Aim 1:** Ultra-deep targeted sequencing of neurodegenerative genes in sporadic FTD/ALS brains using Molecular Inversion Probes (MIPs)

#### Major Task 1

Milestone at 3 months: HRPO Approval

We submitted HRPO forms for review in September 2020 and got approved in February 2021. This step was dramatically delayed due to the COVID-19 pandemic and also caused the delay of subsequent experiments.

Milestone at 6 months: Extracted gDNA from all brain samples

We finished gDNA extraction of all samples in February 2021.

Milestone at 10 months: Raw sequencing data

All samples were sequenced in August 2021.

Milestone at 14.5 months: Validated pathogenic germline and somatic variants

We have validated a set of somatic variants that were predicted to be pathogenic in December 2021. These validated variants are now under further investigation for their cell-type distributions.

**Specific Aim 2:** Genome-wide burden of sSNV and retrotransposon insertion in single neurons of FTD/ALS brains

#### Major Task 2: Determine the burden of sSNV through scWGS

Milestone at 3 month: Raw data of low-coverage sequencing

A new whole genome amplification method, primary template-directed amplification (PTA), became commercially available in 2020 and showed a much better quality of whole-genome amplification of single cells in our hands. Using this new method, we no longer needed to do low-coverage sequencing for QC analyses. All PTA amplified cells were directly used for high-coverage sequencing before data analyses.

Milestone at 6 month: Raw data of high-coverage sequencing

High-coverage sequencing were generated in September 2021.

Milestone at 11 month: Complete comparison of somatic SNV between FTD/ALS and normal brains

Somatic SNVs and Indels were called and compared in January 2022. We are performing further analyses to better understand potential mechanisms underlying the observed changes.

#### Major Task 3: Determine the rate of retrotransposon insertion through L1 targeted sequencing

Milestone at 6 months: Raw data of L1 targeted sequencing

Using the newly developed PTA-HATseq method, we finished the sequencing for FTD/ALS neurons in December 2021.

Milestone at 12 months: Complete comparison of somatic L1 insertion between FTD/ALS and normal brains

We have made the initial calls of L1 insertions. We are now validating the L1 insertions.

**Specific Aim 3:** Genome-wide burden of DSB in neurons of FTD/ALS brains

#### Major Task 4: Determine the burden of somatic Indels through scWGS

Milestone at 4 months: Complete comparison of somatic Indel between FTD/ALS and normal brains

The Indel calls were made in December 2021. We are doing further analyses for genomic distribution of Indels.

#### Major Task 5: Identify DSB using BLISS

Milestone at 6 months: Raw data of the BLISS sequencing

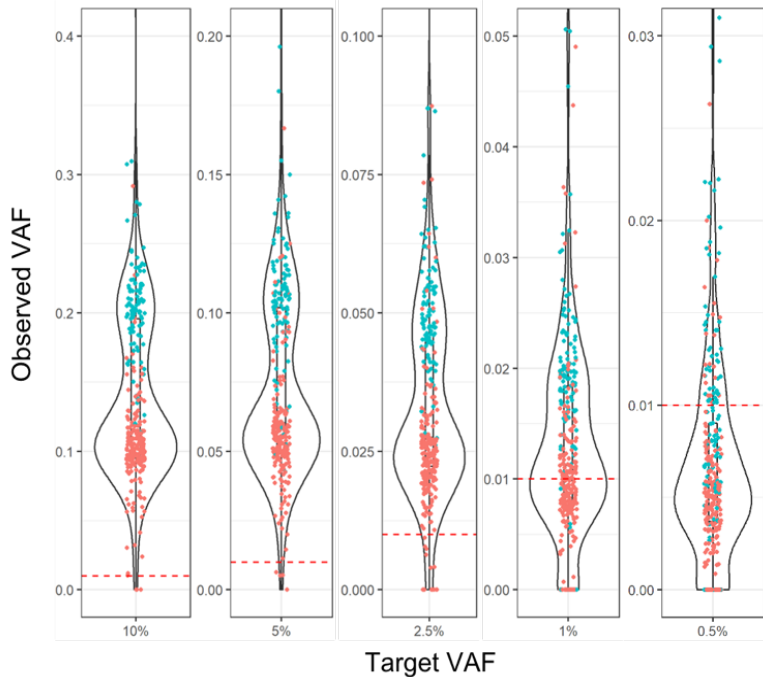
Due to the technical challenges of BLISS using postmortem tissues, we couldn't generate clear results of double-strand breaks in postmortem brains. We don't think this technical problem can be solved within the time period of this grant.

### What was accomplished under these goals?

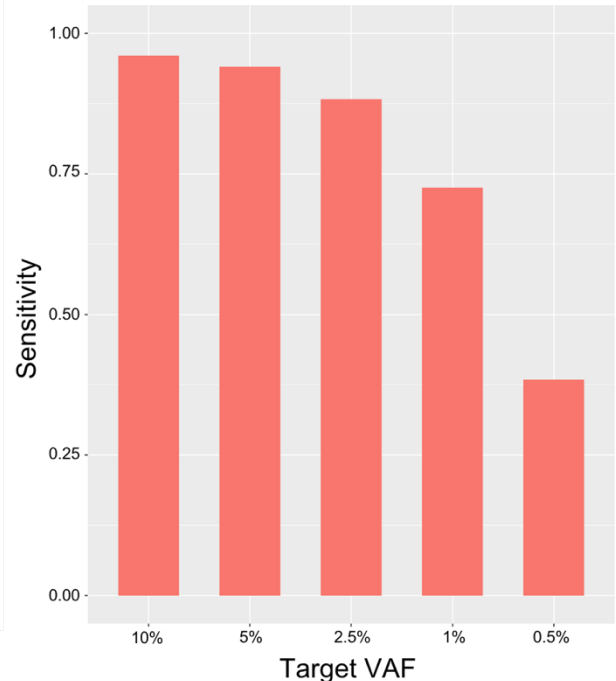
**Specific Aim 1:** Ultra-deep targeted sequencing of neurodegenerative genes in sporadic FTD/ALS brains using Molecular Inversion Probes (MIPs)

We have collected 1,706 human postmortem brain tissues of 533 individuals from brain banks and gotten the HPRO forms approved. gDNA from all these samples were extracted. We have designed a MIPs panel targeting the exons and exon-intron junctions of 88 neurodegeneration related genes, which includes 30 FTD/ALS genes, 8 AD genes,

27 PD genes as well as genes associated with other rare dementia. We performed a spike-in experiment to test the performance of the MIPs panel using commercially available human gDNA from the Genome in a Bottle Consortium (GIAB). The GIAB cases have been well characterized and the WGS data of them are available, which has been widely used for benchmarking variant calling. The spike-in test sequencing was done by mixing one normal individual's gDNA (GIAB12878) into another normal individual's gDNA (GIAB24695) at five different variant allele frequencies (VAFs), 10%, 5%, 2.5%, 1% and 0.5%. We developed and applied a customized pipeline for calling somatic SNVs from the spike-in test experiment with our MIPs panel. The observed VAFs of somatic SNVs were very well in line with the target VAFs at all five VAF ratios (Fig. 1). And the sensitivity of our customized pipeline for calling somatic variants from the spike-in MIPs data ranged from 40% to 95% (Fig. 2). The MIPs panel sequencing, together with our customized pipeline, allows us to confidently identify clonal somatic mutations even with a VAF of 0.5%.



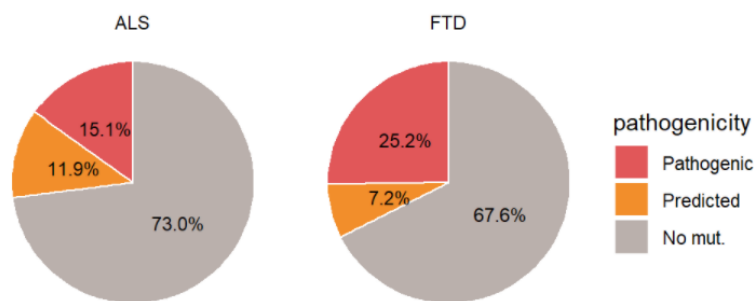
**Figure 1. VAFs of somatic variants called in the spike-in MIPs sequencing data.** The gDNA of GIAB12878 was mixed into the gDNA of GIAB24695 at the listed target VAFs for heterozygous germline variants of GIAB12878. Heterozygous germline variants (red) are clustered at the expected VAFs and homozygous germline variants (green) are clustered at VAFs with doubled values.



**Figure 2. Sensitivity of the customized pipeline in calling somatic variants from the spike-in MIPs sequencing data.** The sensitivity of our customized pipeline ranges from ~40% to 95% at different target VAFs.

We then sequenced all 1,706 samples using our MIPs panel and reached ~1800X average sequencing depth without variations among different batches and disease conditions. Germline mutations were called following the GATK best practice and annotated using our customized pipeline. In addition to the MIPs sequencing, repeat-primed PCR was also performed to genotype germline C9ORF72 repeat expansions in our cases. Overall, we found that 27% and 32.4% ALS and FTD cases respectively carry previously reported or predicted pathogenic germline mutations (Fig. 3).

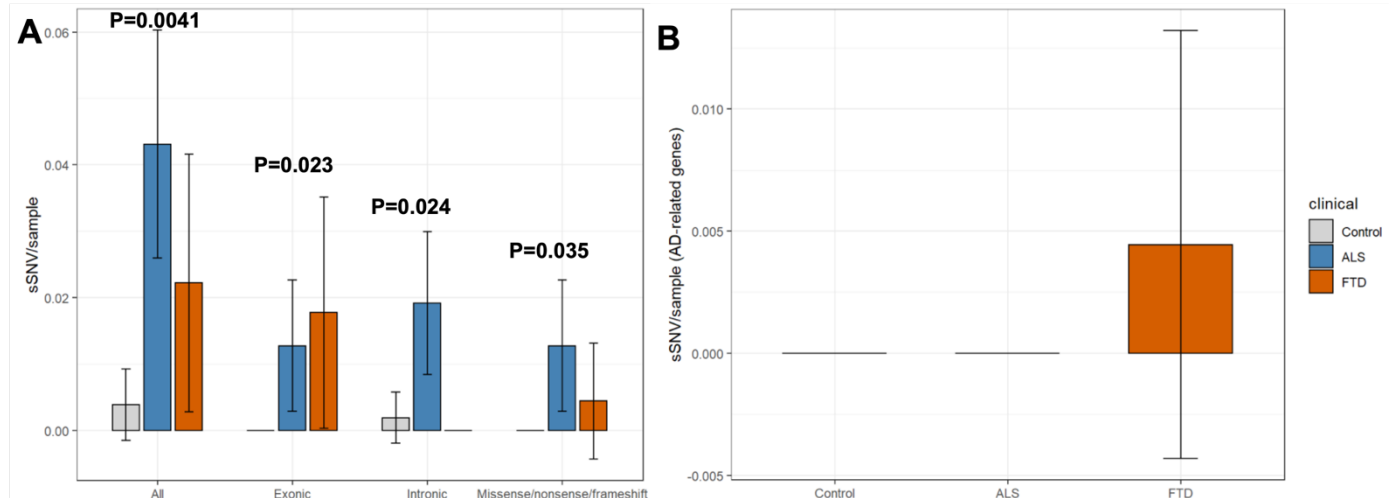
Using this customized pipeline, we identified 140 somatic SNVs from all the samples. We selected 27 SNVs for validation and validated 70.3% of them through Amplicon sequencing. After excluding cases with germline pathogenic and predicted pathogenic variants in FTD/ALS genes,



**Figure 3. Proportion of FTD/ALS cases with germline variants in FTD/ALS genes.** Pathogenic mutations were previously reported and predicted mutations were predicted by our customized annotation pipeline.

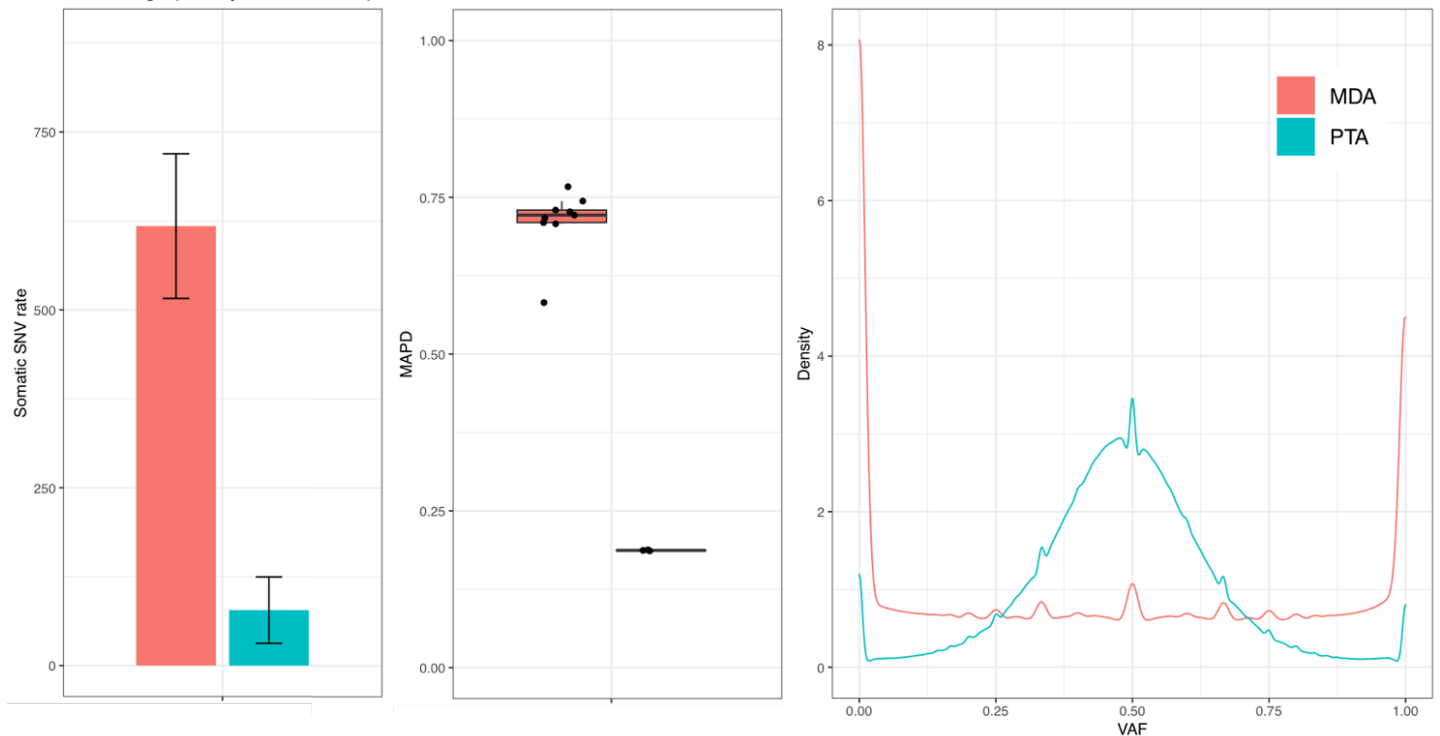
we were able to see enrichment of somatic SNVs in FTD/ALS genes in FTD/ALS cases. The SNVs were then further divided into exonic, intronic and functional (missense/nonsense/frameshift) mutations (Fig. 4A). Same enrichment was seen for exonic mutations in FTD/ALS cases. ALS cases also had enrichment of intronic and functional mutations. FTD cases only had a slight increase in functional mutations likely due to the fact that only half of FTD cases have TDP43 proteinopathy while the other half have Tau proteinopathy. Indeed, only FTD cases had functional somatic mutations in Tau proteinopathy genes (Fig. 4B). We then checked the distribution of functional mutations across brain regions. Interestingly, functional mutations were only detected in the prefrontal cortex (PFC) and primary motor cortex (BA4)

samples of FTD cases and BA4, premotor cortex (BA6) and spinal cord samples of ALS cases. No functional mutation was detected in these regions in control cases, suggesting these functional mutations were indeed pathogenic. We are currently performing cell-type analyses to figure out the cell-type distribution of these functional mutations.



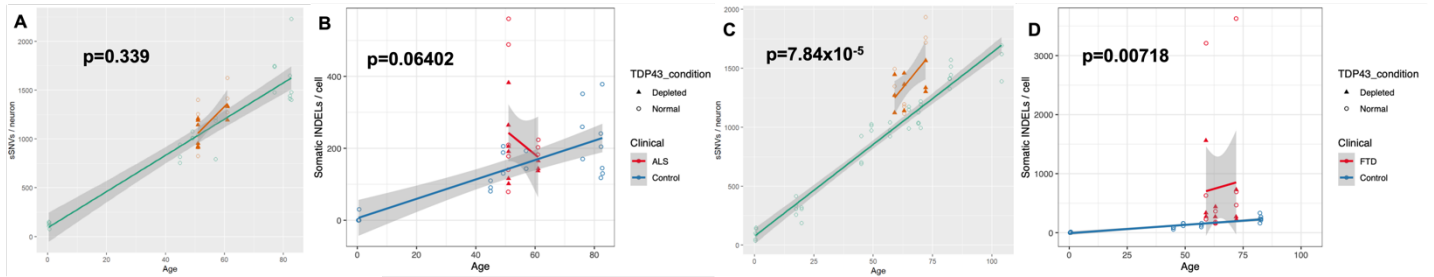
**Specific Aim 2: Genome-wide burden of sSNV and retrotransposon insertion in single neurons of FTD/ALS brains**

PTA, a new whole genome amplification (WGA) method, became commercially available in 2020. We compared scWGS data generated by PTA and MDA. PTA data had higher sensitivity, more even genome amplification, and fewer artifacts than other commonly used MDA data (Fig. 5). We decided to generate scWGS data for this project using PTA instead of MDA as we proposed originally in this grant. PTA not only improves the quality of scWGS data, but also saves time in doing quality control experiments.



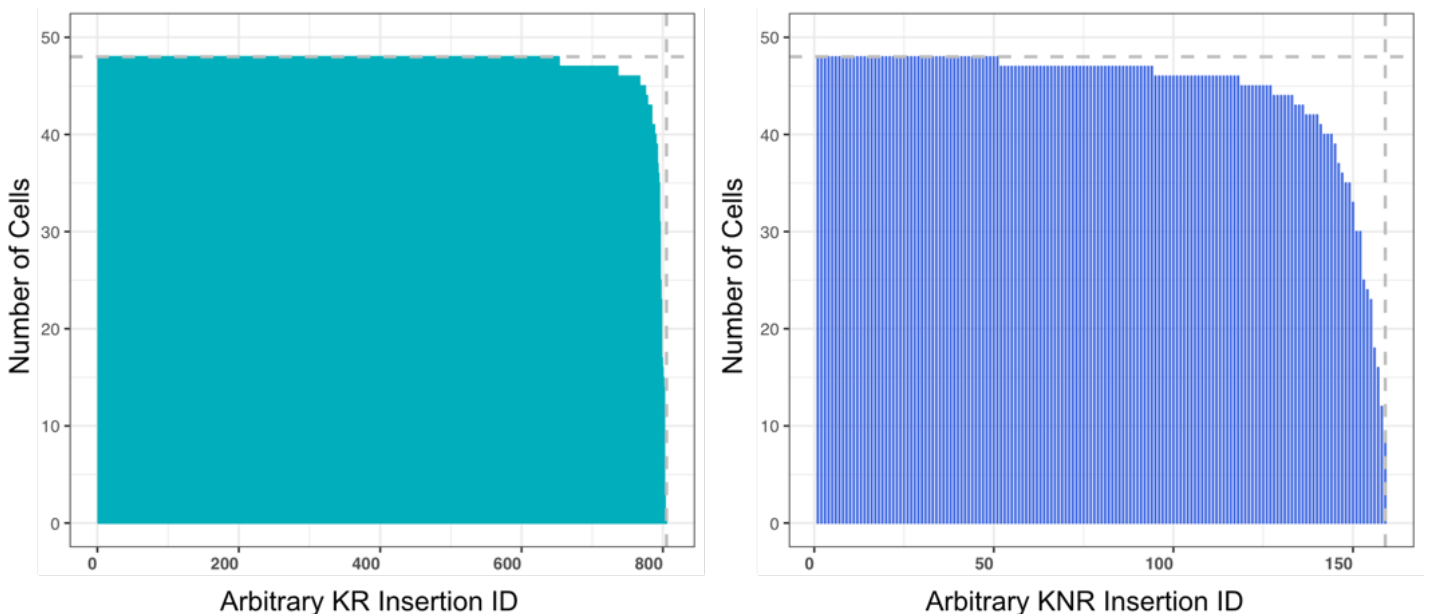
Single-cell whole genome sequencing (scWGS) was performed for 24 neurons from the BA6 region of 8 normal brains with ages ranging from 0 to 83 yo and established the base trendline of somatic mutation accumulation in neurons of BA6 region. The newly developed SCAN2 single-cell variant caller was used to identify single-cell private SNVs and Indels. BA6 neurons showed slightly higher sSNV rate than PFC neurons from the same normal individuals that our lab

generated for another project (Data not shown). On the other hand, somatic Indels (sIndels) accumulate similarly in BA6 and PFC neurons. scWGS was then performed for 3 normal and 3 TDP43 depleted neurons from each of the BA6 of 3 ALS brains and PFC of 3 FTD brains. BA6 neurons of ALS brains showed no increase of sSNVs and a slight increase in sIndels, although the p value was not significant (Fig. 6A, B). PFC neurons of FTD brains showed both significant increase of sSNVs and sIndels (Fig. 6C, D). Although normal PFC neurons showed higher sSNV and sIndel burdens than PFC neurons with TDP43 nuclear depletion, the p values didn't reach significance. More cells need to be studied to prove this potentially interesting observation which may suggest a cell-autonomous protecting role or non-cell-autonomous damaging role of TDP43 pathology. Mutational signature analyses are undergoing for these sSNVs and sIndels.



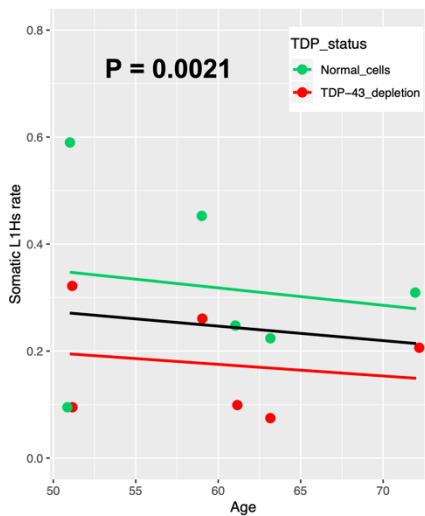
**Figure 6. sSNVs and sIndels in FTD/ALS brains. A-B.** sSNV burden (A) and sIndel burden (B) of BA6 neurons of ALS and control brains. Twenty-four control BA6 neurons are included. **C-D.** sSNV burden (C) and sIndel burden (D) of PFC neurons of FTD and control brains. Forty control PFC neurons are included. P value are calculated using a linear mixed model.

To identify somatic L1 retrotransposon insertions in single cells, we decided to combine human active transposon sequencing (HATseq) and PTA. Before applying it to FTD/ALS neurons, we first tested its performance. A test sequencing with the PTA-HATseq showed sensitivities of 98.3% and 93.1% in detecting known reference (KR) and known non-reference (KNR) insertions, respectively. This represents greater sensitivity than ever before reported for any single-cell method. The slightly lower sensitivity for the non-reference insertions is probably caused by the allele and locus dropout arising during single-cell WGA. Most non-reference insertions are heterozygous and more vulnerable to being missed due to allele and locus dropouts than reference insertions which are mostly present in two (homozygous) copies. This combination approach also has very subtle variations in the sensitivity across single cells and the bulk DNA sample (Fig. 7). Our pilot test sequencing results suggest that the combination of PTA and L1 targeted sequencing can recover >93% of L1 insertions genome-wide with an average depth of 200X by multiplex-sequencing of 96 single-neuron HAT-seq libraries in one lane of the Illumina HiSeq X sequencer. The estimated rate of somatic L1 insertions is 0.242 [95%CI: 0.108, 0.377] per neuron using this approach which is slightly higher than earlier reports and generally consistent with more recent studies. Therefore, it is a robust and sensitive approach to quantitatively study somatic L1 insertions in single cells.



**Figure 7. Recovery of germline KR and KNR insertions in single cells using PTA L1 targeted sequencing.** 48 neurons of a normal human postmortem brain show great evenness of the recovery of germline KR and KNR insertions. Each vertical bar represents the number of neurons having each KR or KNR insertion was called. Since KR and KNR are germline insertions, all 48 single-neurons from the same donor are expected to have each insertion called.

We then applied this method to single cells of both FTD/ALS brains. A bulk sample, 47 normal neurons and 48 neurons with TDP43 nuclear depletion were sequenced using PTA-HATseq for each of the BA6 of 3 ALS brains and the PFC of 3 FTD brains. Similar sequencing depth and sensitivity were achieved across all samples regardless of the disease and TDP43 conditions. A higher L1 insertion rate was observed in normal neurons than neurons with TDP43 nuclear depletion (Fig. 8), which again suggested a non-cell-autonomous damaging role of TDP43 related dysregulation of retrotransposons. The identified L1 insertions will be validated through a combination of 5' and 3' junction PCR validations.



**Figure 8. Somatic L1 insertions in FTD/ALS neurons.** Normal neurons show higher rate of L1 insertions than neurons with TDP43 nuclear depletion in both FTD and ALS brains. Black line indicates the trendline of average L1 insertion rate of both normal and TDP43 depleted neurons.

**Specific Aim 3: Genome-wide burden of DSB in neurons of FTD/ALS brains**

Due to the technical challenges of BLISS using postmortem tissues, we couldn't generate clear results of double-strand breaks in postmortem brains. Postmortem brains in general had very high background noise of double-strand breaks, likely due to the freeze and thaw cycles. In addition, the inter-sample variation was high. We have communicated with other labs that were also trying to use BLISS to study double-strand breaks in postmortem tissues and realized that it was a common problem. Although we are not able to determine DSBs directly, we expect that mutational signatures of the sIndels identified in Aim 2 should represent a mutational process of DSB and the associated DSB repair; e.g., homologous recombination, non-homologous end joining and microhomology-mediated end joining. The mutational signature analysis is currently undergoing.

**What opportunities for training and professional development has the project provided?**

Nothing to Report

**How were the results disseminated to communities of interest?**

Nothing to Report

**What do you plan to do during the next reporting period to accomplish the goals?**

We have identified clonal somatic mutations in FTD/ALS brain and spinal cord samples and validated some of the identified mutations. We have requested additional brain tissues and section slides of several cases carrying pathogenic somatic mutations for a more detailed regional distribution analysis. The regional distribution of pathogenic somatic mutations and TDP43 pathology through immunohistochemistry will also be compared for a potential correlation between the two. Cell-type analysis will be performed to see if the pathogenic somatic mutations have a preferential enrichment in neurons or glial cells. We have also requested the disease onset information of the FTD/ALS cases. The age of onset will be compared between cases carrying pathogenic germline mutations and somatic mutations.

The mutational signatures of sSNVs and sIndels from the scWGS data will be analyzed. Signatures will be made through both de novo signature decomposition and decomposition with COSMIC signatures. This should help figure out the potential mechanisms underlying the changes of sSNVs and sIndels in disease brains.

The candidate L1 retrotransposon insertions will be validated through both 5' and 3' junction PCR experiments. PTA-HATseq of neurons from aged-matched control brains will also be performed, which will be used as the baseline of L1 insertions.

Results generated from these studies will be wrapped up for two or three potential publications over the next year or so.

**4. Impact**

**What was the impact on the development of the principal discipline(s) of the project?**

Our study of clonal somatic mutations showed that a small fraction of FTD/ALS brains carry pathogenic somatic mutations that can explain the pathogenesis of the disease. Our findings will change the understanding of the pathogenesis of FTD/ALS and potentially change the diagnostic process of the diseases. Although our current results suggest that only 3-4% FTD/ALS cases can be potentially explained by pathogenic somatic mutations, the study could be limited by the detection sensitivity of ultra-low level somatic mutations and the limited number of brain regions involved in this study. Indeed, our preliminary data suggest that the identified somatic mutations are very focal with VAFs around 0.5-1%. With more advanced sequencing approaches in the future, it should allow us to detect somatic mutations with VAFs <0.5% and presumably increase the fraction of FTD/ALS cases carrying pathogenic somatic mutations.

Our scWGS data showed that normal neurons exhibit higher rate of sSNVs, sIndels and somatic L1 insertions than neurons with TDP43 nuclear depletion. These results raise an interesting hypothesis that TDP43 pathology may have a non-cell-autonomous damaging role which leads to elevated cellular stress and somatic mutations in adjacent normal neurons. Further analyses will be performed over the next year to address this hypothesis.

**What was the impact on other disciplines?**

Nothing to Report

**What was the impact on technology transfer?**

Nothing to Report

**What was the impact on society beyond science and technology?**

Nothing to Report

**5. Changes/Problems**

**Changes in approach and reasons for change**

For whole-genome amplification of single cells, we performed primary template-directed amplification (PTA) instead of multiple displacement amplification (MDA). PTA is an improved method for whole-genome amplification of single cells which is based on MDA. PTA became commercially available in 2020 and showed a much better quality of genome amplification in our hands. Therefore, PTA was used for this project.

**Actual or anticipated problems or delays and actions or plans to resolve them**

Shortly after the grant started, our institution was forced to shut down due to the COVID-19 pandemic. Our institution reopened in July 2020, but has been only allowing a 50% capacity of lab members who can work at the same time. Brain banks that we requested samples from were also shut down in the meantime. This significantly slowed down our speed of getting the necessary human brain tissue samples and the associated documents from those brain banks. We immediately resumed the process of getting the samples after the brain banks reopened and submitted the Human Research Protection Office (HRPO) form for review in September 2020. The HRPO approval was received in February 2021. The COVID-19 pandemic has delayed our process of getting the samples and the downstream experiments using them for about 6 months. We requested an one-year no-cost extension for this grant, which should give us enough time to finish the proposed studies. In addition, the originally proposed BLISS sequencing experiments turned out to be technically challenging with postmortem tissues. We don't think this issue can be resolved before the end of this grant. We will focus on the mutational signature analyses of indels, which can potentially link specific indel signatures to DSB and associated DSB repair processes.

**Changes that had a significant impact on expenditures**

Nothing to Report

**Significant changes in use or care of human subjects, vertebrate animals, biohazards, and/or select agents**

Nothing to Report

**Significant changes in use or care of human subjects**

Nothing to Report

**Significant changes in use or care of vertebrate animals**

Nothing to Report

**Significant changes in use of biohazards and/or select agents**

Nothing to Report

**6. Products**

**Publications, conference papers, and presentations**

Nothing to Report

**Website(s) or other Internet site(s)**

Nothing to Report

**Technologies or techniques**

Nothing to Report

**Inventions, patent applications, and/or licenses**

Nothing to Report

**Other Products**

Nothing to Report

**7. Participants & Other Collaborating Organizations**

**What individuals have worked on the project?**

Zinan Zhou: no change

Junho Kim:

Name:	<i>Junho Kim</i>
Project Role:	<i>Collaborator</i>
Researcher Identifier (e.g. ORCID ID):	0000-0003-3513-401X
Nearest person month worked:	3
Contribution to Project:	<i>Dr. Kim has performed computational analyses for all the studies.</i>
Funding Support:	<i>Assistant Professor startup package from the Sungkyunkwan University</i>

**Has there been a change in the active other support of the PD/PI(s) or senior/key personnel since the last reporting period?**

Zinan Zhou is partially supported by the postdoctoral fellowship of the Edward R. and Anne G. Lefler Center at Harvard Medical School from June 2021 to June 2022.

Junho Kim became an Assistant Professor at Sungkyunkwan University starting from July 2021. He is fully supported by his startup package since then.

**What other organizations were involved as partners?**

- **Organization Name:** Sungkyunkwan University
- **Location of Organization:** South Korea
- **Partner's contribution to the project**
  - Collaboration: Dr. Junho Kim is now working on these projects as a collaborator. He has been helping with the computational analyses of the studies including somatic detections in both bulk and single-cell samples.

**8. Special Reporting Requirements**

**Appendices**