Exploring the impact of mobile elements on Alzheimer's disease using targeted long-read sequencing

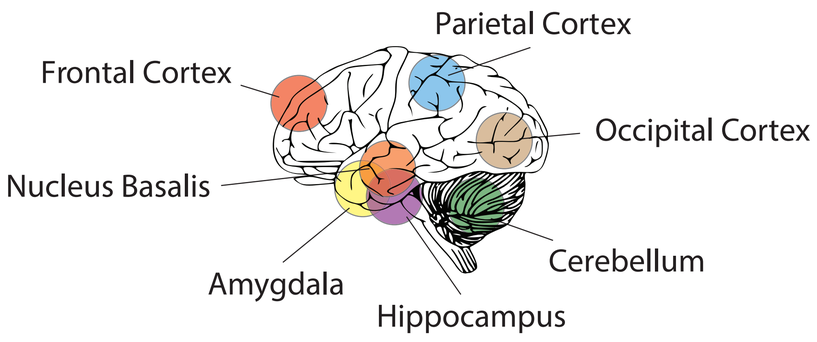
Camille Mumm[1,2], Torrin McDonald[2], Arthur Zhou[2], Jessica Switzenberg[2], Connor Maltby[3], Peter Todd[3], Ryan Mills[1,2], Alan P. Boyle[1,2]

[1]Department of Human Genetics. [2]Department of Computational Medicine and Bioinformatics. [3]Department of Neurology. University of Michigan, Ann Arbor, MI

Targeted sequencing across brain regions

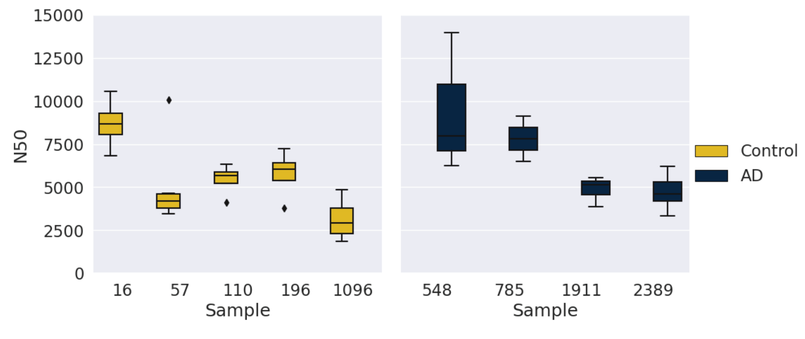
Transposable element-derived sequences make up more than 45% of the human genome. These sequences are not only an important source of variation but, mobile element insertions (MEIs) can also play a role in development and disease.

Here we aim to characterize MEIs both within and between individuals. To do this we used nanopore Cas9-targeted sequencing (nCATS[1]) to enrich for and sequence classes of mobile elements active in the human genome; L1HS, AluYa5, AluYb8. We performed these enrichments across multiple brain regions in postmortem tissue from individuals with a high-likelihood of having Alzheimer's disease (AD) and control individuals to identify non-reference and potentially somatic MEIs.

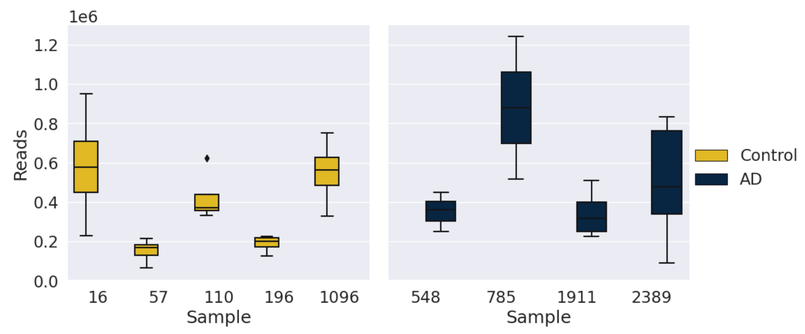


Sequencing Results

N50 by sample

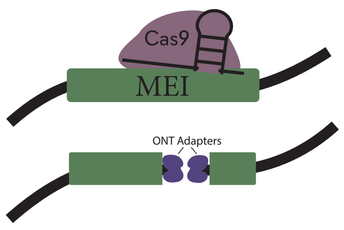


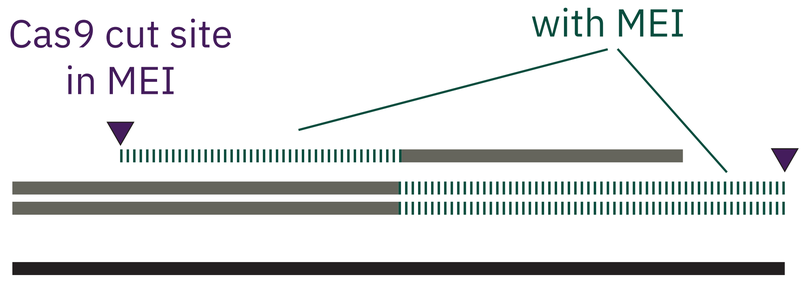
Reads by sample



nCATS enrichment for MEIs

sgRNA targets Cas9 to cut L1HS, AluYa5, AluYb8





Previous work from our lab adapted nCATs to target MEIs in control cell lines[2]. Using this technique we are able to efficiently capture elements across the genome on Minion flow cells. We then use NanoPal[2] identify and characterize events.

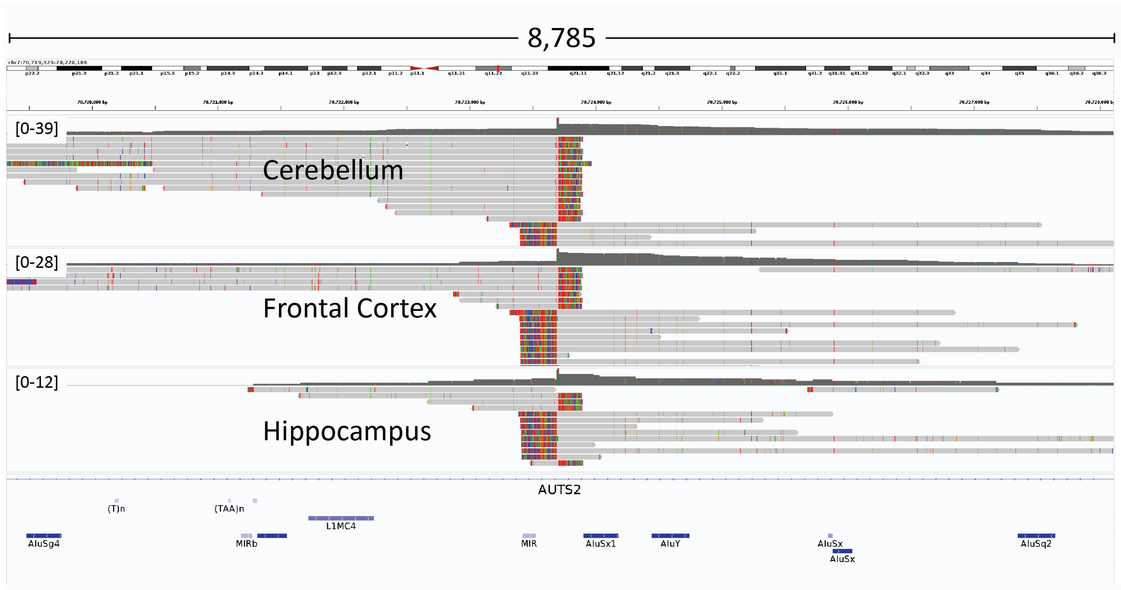
Preliminary Results

	Control	High Likelihood AD
Samples	5	4
Number Regions	4-6	2-7
Mean number reads	359337	469589
Mean N50	5566	5915

We sequenced multiple brain regions from control and AD samples. Using nCATs, we can obtain more than 60% on target rate (reads with MEI signal / total reads) and capture both reference and non-reference L1HS, AluYa5, and AluYb8.

A

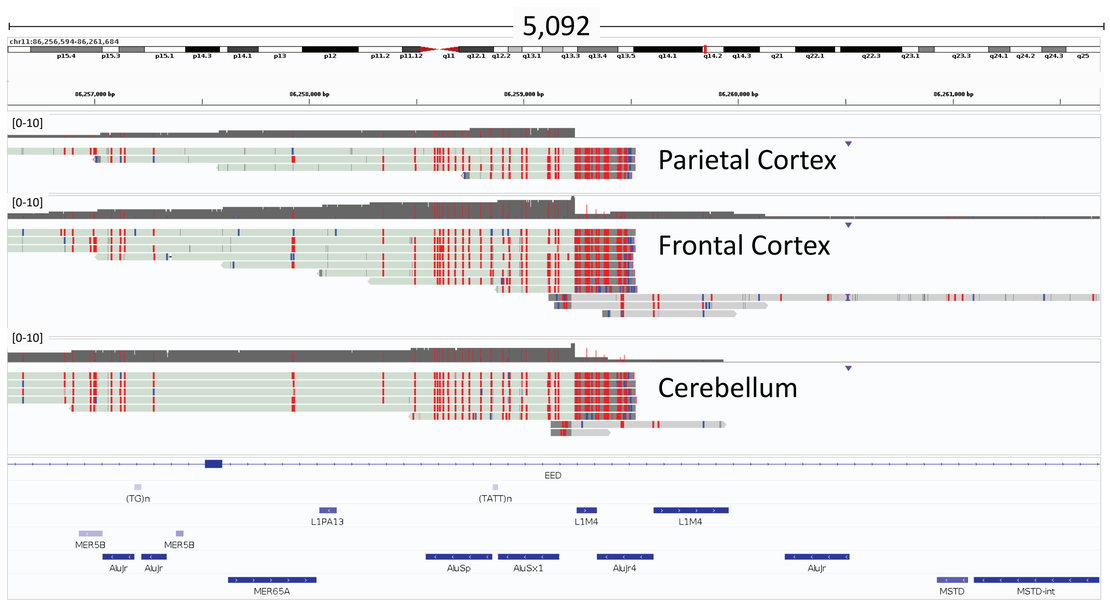
Non-reference L1HS insertion



A.) A ~480bp non-reference L1HS insertion on chromosome 7 in an intron of AUTS2. This event is shared across all brain regions from this individual. The dark gray tracks show the coverage at this region which varies between region and highlights the biased bi-directional sequencing from nCATS. The gray track shows the reads, with the multicolored portion showing the clipped sequence containing the truncated L1HS. The reference gene and RepeatMasker tracks are shown in blue.

B

CpG Methylation at non-reference AluYa5 insertion



B.) A ~320bp non-reference AluYa5 insertion in an intron of EED in 3 brain regions. The dark gray tracks show the coverage at this region which varies between region and highlights the biased bi-directional sequencing from nCATS. The light gray track shows the reads, which have been shaded by CpG methylation. Red bases show a methylated CpG called by guppy and blue bases are unmethylated CpGs. The reference gene and RepeatMasker tracks are shown in blue.

Project Goals

Here we show preliminary data from ongoing work characterizing MEIs in brain tissues from AD and control samples. Using nCATs we are able to capture potential individual specific non-reference mobile element insertions. We aim to compare MEIs both between individuals and also between regions from the same brain to characterize potential mosaic insertions. The potential role of somatic variation in AD is not yet fully understood, but using long reads we will be able to investigate differences between AD and control groups.

Using nanopore reads we are also able to directly obtain methylation data at mobile elements and can examine both global and locus-specific methylation patterns of both reference and non-reference repetitive elements across brain regions and samples.

In addition to transposable elements, we also aim to examine other kinds of variation between individuals, patient groups, and brain regions. In particular we will examine other types of repetitive variation such as tandem repeats.

Future Directions

- Analyze duplex reads from Kit 114 data
- Examine methylation patterns between reference and non-reference events
- Further quantify the number of non-reference elements per sample / per region

Acknowledgments

- Members of the Boyle, Mills, and Todd labs who have provided samples, designed MEI analysis, and performed enrichment experiments on this project
- Genome Science Training Program
- 5R21HG011493-02

References

1. Gilpatrick et al., 2020, Nat Biotechnol
2. McDonald et al., 2021, Nat Commun

the BOYLE LAB