Investigating the Effects of Allele-specific Alternative Polyadenylation on Alzheimer's Disease

Ryan Barney[1,4], Giovanni Quinones-Valdez[1,2,3], Xinshu Xiao[1,2,3,4]

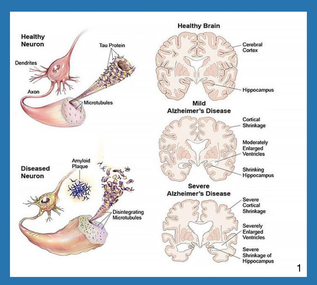
BIG Summer Program - Institute for Quantitative and Computational Biosciences[1], Department of Bioengineering[2], Department of Integrative Biology and Physiology[3], Bioinformatics Interdepartmental Master's Program[4]

Abstract

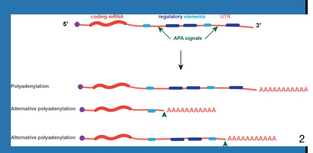
Alzheimer's Disease (AD) is the most common form of dementia, impacting millions of Americans across the country. Single nucleotide polymorphisms (SNPs) could contribute to the development or progression of Alzheimer's Disease pathology. One type of functional mechanism of SNPs is by altering RNA processing. This has the potential to drastically alter cellular pathways. Discovering these genetically regulated alternative RNA processing events is critical in our effort to comprehend and combat AD. We seek to understand how allele-specific alternative polyadenylation (asAPA) presents powerful insight into the functional consequences of SNPs in AD patients. Using the allele-specific alternative mRNA processing pipeline (ASARP) developed by our lab, we identified SNPs in RNA-seq data from 364 human brains from the Mount Sinai Brain Bank (MSBB). We reveal how asAPA events are found in genes that demonstrate AD relevance, and they also demonstrate downstream effects on processes such as RNA binding protein (RBP) binding. Taken together, we illustrate how asAPA has strong evidence for being a contributor of AD.

Introduction

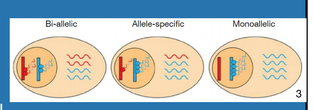
- Understanding Alzheimer's Disease (AD) can lead to
  - identifying risk factors and causes of neurodegeneration
  - discovering novel drug targets for treatment



- Alternative Polyadenylation (APA) is when the polyA tail is appended at a non-canonical location in a transcript
  - This has been shown to be AD relevant in previous literature

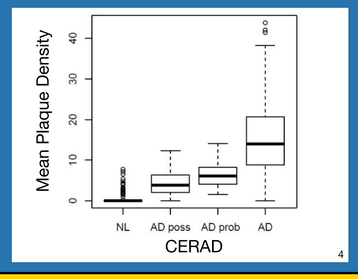


- Allele-specific expression (ASE) is the preferential expression of one allele
- SNPs may drive allele-specific APA, manifested as ASE in local regions of 3'UTRs - we call this asAPA



Materials

- Data came from the Mount Sinai Brain Bank
  - 1004 RNA-seq samples from 364 human brains
  - 4 Brain Regions of Interest:
    - Frontal Pole (FP)
    - Inferior Frontal Gyrus (IFG)
    - Parahippocampal Gyrus (PG)
    - Superior Temporal Gyrus (STG)



- NL = brain does not have AD
- AD poss = brain has a chance of having AD
- AD prob = brain likely has AD
- AD = brain does have AD

Methods

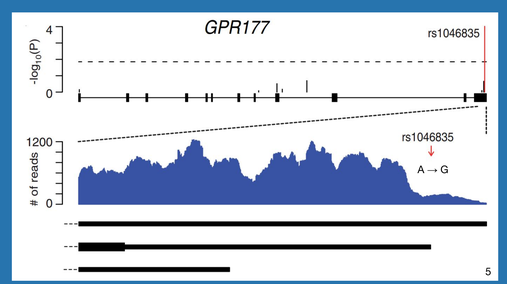
1. Map reads to a reference genome
We used STAR to align reads with the human genome

2. Remove duplicate reads
Duplicate reads can cause bias towards certain events

3. Calculate the normalized expression value
Finds inclusion of 3'UTR regions, or where APA occurs

4 Find RNA-DNA differences
Identify SNP locations, along with their allelic ratios

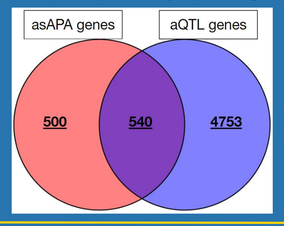
5. Identify asAPA events
Test whether an allele is associated with APA



Results

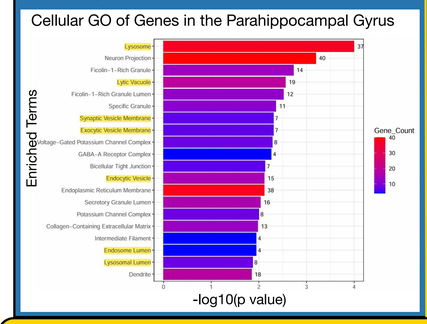
Overview and Validation

- Our method identifies 9318 asAPA events in 1040 unique genes
- Alternative polyadenylation quantitative trait loci (aQTLs) are known to alter polyA motifs
  - Over 50% of our asAPA-containing genes overlap with previously reported genes[6]



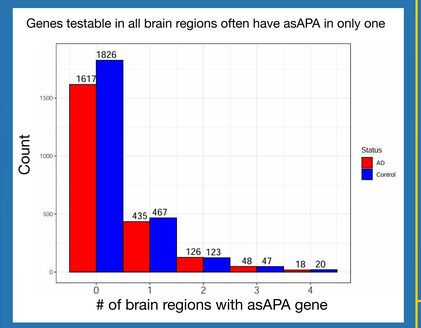
Gene Ontology finds genes relevant in vesicle trafficking

- Gene Ontology identifies the pathways our genes are present in
- We found terms related to vesicle transport, traffic, and formation

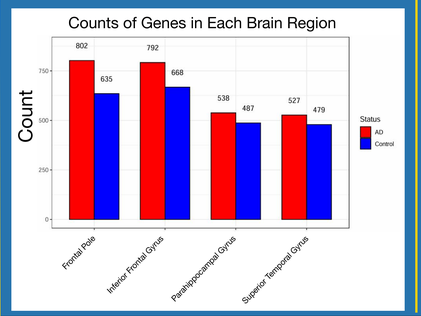


asAPA genes are brain-region specific

- Genes are testable if they have an imbalanced SNP in the alternate 3'UTR region
- For the set of genes that could be tested in all brain regions, they often have asAPA in one brain region or none at all

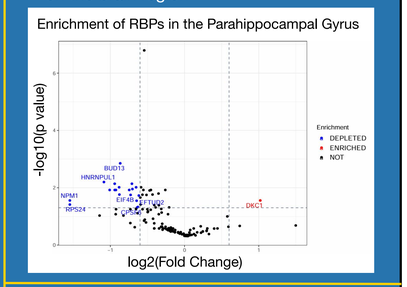


- This means asAPA genes demonstrate brain region specificity
- Of the brain region specific genes, most are in the FP or IFG



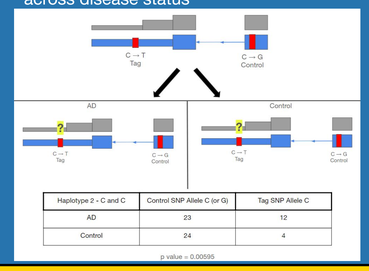
RNA Binding Proteins have differential enrichment

- Enhanced crosslinking and immunoprecipitation (eCLIP) data maps the binding site of RBPs
- RNA binding proteins overlapping our asAPA events can be enriched or depleted based on the brain region



Differences between AD and Control were detected

- The read counts for asAPA events across AD and Control brains can be compared
- 20 unique genes have significant differences across disease status



Haplotype 2 - C and C	Control SNP Allele C (or G)	Tag SNP Allele C
AD	23	12
Control	24	4

p value = 0.00595

Conclusions

1. asAPA demonstrates a difference in prevalence based on brain region, with most genes being significant in only one brain region
   a. Certain brain regions may have greater relevance in disease
2. asAPA-containing genes are important in vesicle pathways
   a. The clearance of amyloid-beta plaques characteristic of AD requires proper vesicle function
3. RNA binding proteins are altered in the presence of APA
   a. This may be a functional mechanism by which asAPA acts to contribute to AD
4. A subset of asAPA genes have significantly different allele usage across AD and Control
   a. These significant genes may potentially contribute to AD

Future Directions

- Develop a model to pinpoint specific RBP or miRNA motifs that are disrupted by the disease
  - We can pinpoint particular locations altered in AD
  - Depending on the functional mechanism at that location, we may identify risk factors or drug targets

Acknowledgements

We thank all the members of the Xiao laboratory for their helpful feedback. Their continued support and assistance make this research possible.
We also thank the BIG Summer program for the opportunity to present this work.

1. Morreale, *The Progression of Alzheimer's Disease*. 2015
2. Patel et al., *Alternative cleavage and polyadenylation of genes associated with protein turnover and mitochondrial function are deregulated in parkinson's, alzheimer's, and ALS*. 2019
3. Kukurba et al., *Allelic Expression of Deleterious Protein-Coding Variants across Human Tissues. PLoS genetics*. 2014
4. Wang et al., *The Mount Sinai Cohort of large-scale genomic, transcriptomic and proteomic data in alzheimer's disease*. 2018
5. Li et al., *Identification of allele-specific alternative mRNA processing via transcriptome sequencing*. Nucleic acids research. 2012
6. Li et al., An atlas of alternative polyadenylation quantitative trait loci contributing to complex trait and disease heritability. *Nature genetics*. 2021