Aberrant splicing in AD

Leafcutter analysis has shown aberrant splicing to be a component part of AD pathology.

# QTL

### Mendelian randomization

Mendelian randomization

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3981897/

### MultiXcan

### sQTL

## Causality

#### Potential way to assess the impact of SNPs

Functional annotation by the Roadmap Epigenomics and ENCODE projects [30] indicates many SNPs in this region have an impact on chromatin state or transcription factor binding in various tissues (Fig.3c), but only rs906482 alters DNA-protein binding.

https://www.nature.com/articles/s41380-018-0207-1#Sec11

### Epigenetics

Over the past decades, genome-wide association studies (GWAS) have identified thousands of phenotype-associated DNA sequence variants for potential explanations of inter-individual phenotypic differences and disease susceptibility. However, it remains a challenge for translating the associations into causative mechanisms for complex diseases, partially due to the involved variants in the noncoding regions and the inconvenience of functional studies in human population samples. So far, accumulating evidence has suggested a complex crosstalk among genetic variants, allele-specific binding of transcription factors (ABTF), and allele-specific DNA methylation patterns (ASM), as well as environmental factors for disease risk. This review aims to summarize the current studies regarding the interactions of the aforementioned factors with a focus on epigenetic insights. We present two scenarios of single nucleotide polymorphisms (SNPs) in coding regions and non-coding regions for disease risk, via potentially impacting epigenetic patterns. While a SNP in a coding region may confer disease risk via altering protein functions, a SNP in non-coding region may cause diseases, via SNP-altering ABTF, ASM, and allele-specific gene expression (ASE). The allelic increases or decreases of gene expression are key for disease risk during development. Such ASE can be achieved via either a SNP-introduced ABTF to ASM or a SNP-introduced ASM to ABTF. Together with our additional in-depth review on insulator CTCF, we are convinced to propose a working model that the small effect of a SNP acts through altered ABTF and/or ASM, for ASE and eventual disease outcome (named as a SNP intensifier model). In summary, the significance of complex crosstalk among genetic factors, epigenetic patterns, and environmental factors requires further investigations for disease susceptibility.

https://www.frontiersin.org/articles/10.3389/fgene.2018.00695/full

##### faulty CTCF binding implicated in neuronal disease

demonstrated that a nearby SNP, rs906482, strongly affects binding by the transcription factor, CTCF, and that the high-affinity allele usually occurs on haplotypes carrying the rs9834970 risk allele. Decreased expression of TRANK1 perturbed expression of many genes involved in neural development and differentiation.

https://www.nature.com/articles/s41380-018-0207-1#Sec11

#### mCpG + ATAC vs neuronal disease

Finally, both regions of differential methylation and those of differential accessibility showed a surprising >10-fold enrichment of explained heritability associated with addictive behavior, as well as schizophrenia- and neuroticism-associated regions, suggesting that common psychiatric illness is mediated through brain region-specific epigenetic marks.

https://www.biorxiv.org/content/10.1101/120386v1

##### finding DNMs in ATAC peaks

source: https://www.biorxiv.org/content/10.1101/2019.12.30.891549v2.full.pdfintersected our cell type-specific peak sets with de novo non-coding mutations (DNMs) identified from ASD and145 neurodevelopmental delay (NDD) cases and found significant enrichment of DNMs in 19 of 27 cell-type specific146 peak sets, compared to a merged background peak set (Extended Data Figure 5).

###### scATAC from Kellis et al.

These data could be used to mark regions of interest (bigwig files)- scATACseq data would enable us to identify the target cell types of functional noncoding SNPs

https://www.nature.com/articles/s41588-020-00721-x#Sec2

bw->peaks

https://www.biostars.org/p/106094/

#### DNA accessibility

footprints revealed by DNA accessibility are highly unbiased and capture variation in diverse regulatory elements, including promoters, enhancers, silencers, insulators, and locus-control regions.

https://www.nature.com/articles/ng.3969#Bib1

###### Allele specific OCRs

open chromatin regions (OCRs) are enriched for single-nucleotide polymorphisms (SNPs) associated with gene expression and complex disorders or traits (1, 2). For neuropsychiatric disorders, disease-associated variants are also enriched in OCRs found not only in adult brains but also in developing brains and induced pluripotent stem cell (iPSC) derived neuronal (iN) cells (36), suggesting the importance of neurodevelopmental chromatin regulation. However, as not all OCR variants are functional, it remains largely unknown whether neuropsychiatric disease risk variants affect chromatin accessibility during neurodevelopment.

https://science.sciencemag.org/content/369/6503/561

###### index of accessible regions

empty

https://www.nature.com/articles/s41586-020-2559-3#Abs1

###### DNAse vs TF access

empty

https://www.nature.com/articles/s41586-020-2528-x#Sec8

#### TFs

http://compbio.mit.edu/publications/185P\_Quon\_bioRxiv\_18.pdfwhat TF binding sites are proximal to SNP

###### Background investigation of the literature

Genomic Environment of Regulatory VariantsRVE-score = C-score + F-score. This RVE-score represents the accumulated evidence for a causative role of a regulatory variant in a disease, thus going beyond the molecular impact of the variant itself. Text S2 in the supplemental information online describes the classification scheme in more detail.

https://www.cell.com/trends/genetics/fulltext/S0168-9525(20)30088-3#secst0015

###### Riviera

#### histone marks (specifically H3K27ac)

(H3K27ac) were shown to be the most variable across tissues and cell types33,34, the most variable across individuals35, and the most highly enriched for disease-associated genetic variants34. Thus, to characterize the protein DNA interactions through which gene regulation is mediated, eGTEx

https://www.nature.com/articles/ng.3969#Bib1

###### graph genomes chip-seq peaks

Even if only a fraction of peaks are observed to be altered, these regions will correspond to biochemically active regions that are more likely to differ between individuals and, as such, could be relevant in the study of various human phenotypes.

https://genomebiology.biomedcentral.com/articles/10.1186/s13059-020-02038-8

###### Roadmap epigenomics 15 states

Core 15-state model

https://egg2.wustl.edu/roadmap/web\_portal/chr\_state\_learning.html#core\_15state

###### ML- relation epigenetic marks vs splicing

H3K36me3 gives rise to greater exon inclusionrepressive marks give rise to exon skipping

https://www.biorxiv.org/content/10.1101/2020.02.03.932251v2

#### EP pairs

empty

###### CT-FOCS

CT-FOCS

https://www.biorxiv.org/content/10.1101/707158v2

### Knowing implicated genes in advance

##### PLGC2

##### CD33

### caQTL

caQTL

https://www.biorxiv.org/content/10.1101/2020.01.13.904862v1.abstract

### eQTL

# AD

Alzheimers' Disease is a neurodegenerative disorder. Many different molecular pathways have been implicated in this disease- all of which seem to relate to oneanother.

### Inflammation

### Microglia

### Affected cellular pathways

S-predixcan of eQTLs across all tissues

https://alzres.biomedcentral.com/articles/10.1186/s13195-020-00611-8

### Aberrant histone marks

## Other neurobio disorders

Transcriptome-wide isoform-level dysregulation in ASD, schizophrenia, and bipolar disorder

https://science.sciencemag.org/content/362/6420/eaat8127.abstract?casa\_token=5u8UPD1sssUAAAAA:QKccAPbIw-PuG36TUQOyNt1nYUKGxyWl42JkPv7Xv4fSk0PGb8\_woeBT6--aQZNzK5G5Om6kbkhlv9s

# Sources of data

### GTEx

### ROSMAP

### HACER

http://bioinfo.vanderbilt.edu/AE/HACER/

## EpiMap

empty

http://compbio.mit.edu/epimap/

#### Aggregate ChromHMM regions across tissues

https://personal.broadinstitute.org/cboix/epimap/ChromHMM/observed\_aux\_18\_hg38/CALLS/

### ATAC seq

ATAC-seq for different types of cells in BW format

http://epigenomegateway.wustl.edu/legacy/?genome=hg38&session=drS3o1n4kJ

### RNAseq pseudotime

Could this pseudotime paper's data be complimented with leafcutter analysis to highlight the differential splicing across time

https://www.nature.com/articles/s41467-020-19622-y#data-availability

### H3K27ac differential levels in AD