# Spatiotempotal properties of NDD genes during brain development for biomarker discovery

Urwah Nawaz

February 12, 2020

# 1 Aims of the project

- To comprehensively characterize the expression properties of ID and CP enes in the normal human brain
  - To determine whether ID and CP genes are expressed in a cell-type speciic mannner using single-cell RNA-seq data
  - To characterize the developmental trajectory of gene expression for ID and CP genes and assess their expression during cellular maturation in brain organiods and assess their expression across multiple fetail developmental periods
  - To characterize the spatial and temporal properties of ID and CP gene expression in the adult brain, by assessing age-dependent changes brain-region and cortical layer-specificity
- To determine whether convergent gene expression changes are observed in ID and CP in patient-derived cells with heterogenous mutations

# 2 Introduction

Human brain development is a complex and a tightly regulated process during which changes occur at both anatomical and functional levels. The processes of brain development are highly dependent on the appropriate expression of RNA and proteins, Mutations that result in altered expression or function of these gene products can cause or contribute to neurodevelopmental disorders (NDDs).

# 3 Neurodevelopmental disorders

Neurodevelopmental disorders (NDDs) are a group of early onset neurological disorders that affect an estimated 10% to 15% of the poluation with revalence rates increasing worldwide. Common NDDs include autism spectrum disroder, intellectual disability, epilepsy and motor/tic movement disorders, and are characterized by strong clinical co-morbidity which suggests common genetic etiology.

- ASD and how many genes have been identified
- ID and how many genes have been identified
- Other NDDs and how many genes have been identified

The genetic heterogeneity and overlap observed in NDDs make it difficult to identify the genetic causes of specific clinical symptoms

**Autism**

Autism spectrum disorders represent genetically heterogeneous group of neurodevelopmental syndromes with high prevalence that has a wide range of phenotype. While there is no unifying hypothesis about the molecular pathology of autism, it is clear that the disorder is highly hertiable and results from the combination of genetic, neurologic , immunologic and environmental factors.

Recent advnces in sequencing technologies have made it possible to gain insight into the molecular aspects of ASD. Microarray technologies and next-generation sequencing have enabled high-throughput discovery of genes likely to be involved in the molecular pathology of autism.5, 6, 7, 8 However, as the success in discovery has risen, the number of candidate genes with associated risk for ASD has also stretched well into the hundreds.9, 10 As of December 2014, 667 genes have been implicated in autism. Despite the large amounts of data now avail-

able, the general lack of replication across studies suggests that more data will be needed to fully characterize the genetic models responsible for the various forms of autism.

### 3.1 Heterogeneity of NDDs

## 4 Expression studies

### 4.1 RNA: bulk vs single-cell

### 4.2 Co-expression network analysis

## 5 Characterisation of brain diversity using single-cell

### 5.1 Studies that have characterized the single-cell transcriptome in the brain

what to review

- human brain single cell rna-seq studies
- reread papers used for PhD proposal
- number and types of samples
- method used for sequencing
- data availablity

| Year | Cells reported | Method | Technique | Species | Cell isolation | Brain region | Developmental stag |
|---|---|---|---|---|---|---|---|
| 2014 | | | | | | | |

- main conclusions of the studies
- co-expression networks and their utility in spatiotemporal studes
- this kind of analysis in other areas (particularly cancer)
- Use of single-cell analysis vs bulk rna-seq
- what are the datasets available
- where am i going to get the genelists from
- what gene ontologies am I getting with my genes
- current co-expression networks that are available
- lack of ID gene networks
- benchmark potentially?
- Issues with bulk rna?