Interrogation of 16p11.2 Region Gene Perturbation

LISA DUNCAN[1], Alex Dai[2], Michael F. Wells[2]

1. BIG Summer Program, Institute for Quantitative and Computational Biosciences, UCLA;
2. Department of Human Genetics, David Geffen School of Medicine, Los Angeles, CA, USA;

Abstract

Copy number variation on chromosomal segment 16p11.2 is the greatest known contributing risk factor for autism spectrum disorder (ASD) and is correlated with other neuropsychiatric disorders, such as bipolar disorder and schizophrenia. Because the effect of 16p11.2 deletion on brain development is not well understood, we explored individual gene function by performing a CRISPR interference screen on 16p11.2 genes in stem cell-derived neural progenitor-like cells (NPCs). We performed Perturb-seq – which captures single guide RNA sequences and single cell gene expression – and analyzed sequencing data using the 10x platform and Cell Ranger pipeline. We identified that experimental improvements will be necessary to optimize the determination of 16p11.2 gene function in NPCs. The results of our analyses will inform future experiments where we aim to identify dysregulated pathways and driver genes of the 16p11.2 region that contribute to ASD and other neuropsychiatric disorders.

Background

- **Autism Spectrum Disorder** (ASD) impacts a large and growing percentage of the population (1 in 36 children)
- 1% of ASD patients have 16p11.2 duplication or deletion syndrome

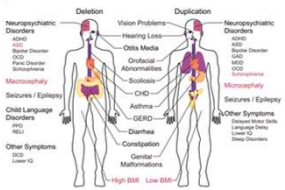

Figure 1: **Shared** and **reciprocal phenotypes** associated with duplication and deletion syndromes. (Roth 2020)

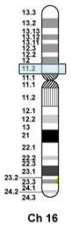

Figure 2: Chromosome 16p11.2 (adapted from ufl.edu). We focused on **deletion** syndrome.

Experimental Goals and Approach

- Interrogate **gene expression**, gene modules and biological pathways impacted by each gene in 16p11.2 deletion in Stem-cell-derived Neurogenin-2-accelerated Progenitors (SNaPs)
- Clustered Regularly Interspersed Short Palindromic Repeats interference (**CRISPRi**) **screen** of 25 genes of 16p11.2 expressed in NPCs

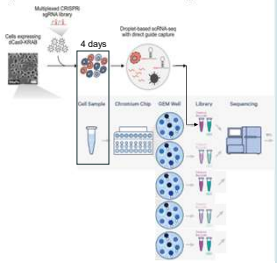

Figure 3: Experimental workflow (adapted from Wells et al 2023, Replogle et al 2022, 10xGenomics.com).

Analysis Goals and Methods

- Pilot a **data analysis pipeline** to optimize experimental workflow
- Perform preliminary identification of **gene perturbation effects**
- Validate expression of pluripotency and **neural progenitor markers** to confirm neural progenitor cell identity

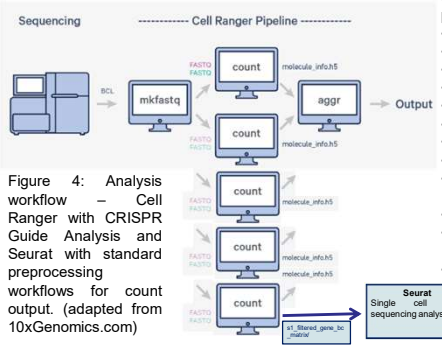

Figure 4: Analysis workflow – Cell Ranger with CRISPR Guide Analysis and Seurat with standard preprocessing workflows for count output. (adapted from 10xGenomics.com)

Standard **Seurat** preprocessing:

- Load 10x data
- Create Seurat object
- Gene Expression
- Quality Control
- Filter mitochondrial RNA
- Normalize data
- Find variable features
- Scale data
- Perform linear dimensional reduction
- Cluster cells
- Run non-linear dimensional reduction
- Plot UMAP

Results – Aggregated Analysis – Cell Ranger

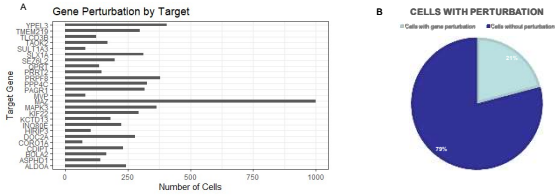

Figure 5 A: Numbers of cells expressing gene perturbation by each targeted gene. B: Percentages of cells with perturbation compared to cells without perturbation.

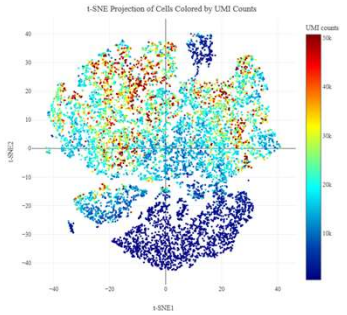

Figure 6: tSNE projections from aggregated summary, identified by UMI counts per cell.

Results – Example Analysis – Seurat

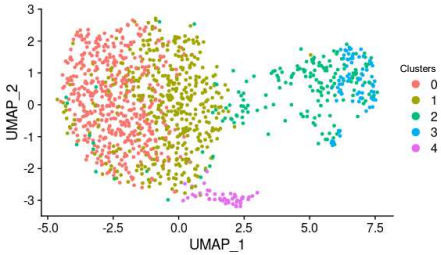

Figure 7: Uniform Manifold Approximation and Projection (UMAP) for sample 1 gene expression.

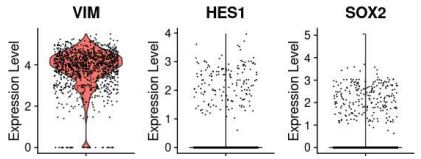

Figure 8: Sample 1 RNA expression of selected markers for SNAPs (neural progenitor cells)

Conclusion and Future Directions

- Proportion of cells with CRISPRi perturbations quite low
- Procedural improvements (experimental methods) to optimize number of cells expressing guides include:
- Increase time between guide transduction and 10x harvest to allow for expression of sgRNA, gene perturbation
- Improve selection for cells with sgRNA – time, fluorescence assisted cell sorting (FACS)
- Further explore perturb-seq-specific Seurat analysis
- Investigate Gene Ontology and gene modules, and pathway analysis
- Repeat 10x with single cell line and cell village.

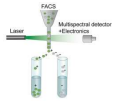

Figure 9: FACS (Hwang et al 2018)

Figure 10: Cell village model (Wells et al 2023)

Acknowledgements

Research reported in this publication was supported by the National Institute Of Dental & Craniofacial Research of the National Institutes of Health under Award Number R25DE030117. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

This work used computational and storage services associated with the Hoffman2 Shared Cluster provided by UCLA Institute for Digital Research and Education's Research Technology Group.

References

Chromosome Ideograms » Division of Genetics and Metabolism » College of Medicine » University of Florida (ufl.edu) accessed 8/5/2023

Data & Statistics on Autism Spectrum Disorder | CDC.gov accessed on 8/10/2023

Hwang, B., Lee, J.H. & Bang, D. Single-cell RNA sequencing technologies and bioinformatics pipelines. *Exp Mol Med* **50**, 1–14 (2018). https://doi.org/10.1038/s12276-018-0071-8

Replogle, J.M., Norman, T.M., Xu, A. et al. Combinatorial single-cell CRISPR screens by direct guide RNA capture and targeted sequencing. Nat Biotechnol 38, 954–961 (2020). https://doi.org/10.1038/s41587-020-0470-y

Roth JG., et al. 16p11.2 microdeletion imparts transcriptional alterations in human iPSC-derived models of early neural development. *Elife*. 9:e58178 (2020). https://doi.org/10.7554/eLife.58178

Wells, M.F., et al. Natural variation in gene expression and viral susceptibility revealed by neural progenitor cell villages. *Cell Stem Cell* **30**, 312-332 (2023). https://doi.org/10.1016/j.stem.2023.01.010

What is Cell Ranger? -Software -Single Cell Gene Expression -Official 10x Genomics Support (10xGenomics.com) accessed 8/5/2023