Data in Brief 35 (2021) 106897

Contents lists available at ScienceDirect

Data in Brief

journal homepage: www.elsevier.com/locate/dib

Data Article

Transcriptomic data of Clozapine-treated Ngn2-induced Human Excitatory Neurons

Anh-Thu N. Lam[a,#], Xi Peng[b,#], Debamitra Das[a,#], Joel S. Bader[a], Dimitrios Avramopoulos[a,c,*]

[a] Department of Genetic Medicine, Johns Hopkins University School of Medicine, United States of America
[b] Department of Biomedical Engineering, Whiting School of Engineering and School of Medicine, Johns Hopkins University, United States of America
[c] Department of Psychiatry, Johns Hopkins University School of Medicine, United States of America

ARTICLE INFO

Article history:
Received 4 February 2021
Revised 12 February 2021
Accepted 16 February 2021
Available online 19 February 2021

Keywords:
Schizophrenia
Clozapine
Cholesterol metabolism
Antipsychotics
Gene expression
Transcriptome

ABSTRACT

We generated human excitatory neurons using a protocol for rapid 21-day induction using neurogenin-2 overexpression (Zhang et al., 2013) in a publicly available control iPSC line. We validated the glutamatergic neuronal identity of the neurons by immunofluorescence and transcriptomics. We exposed 6 of the 12 replicate neuron cultures to therapeutic plasma levels of clozapine (300 ng/mL) for the last 3 days of culture, and the remaining 6 to replicates to the clozapine solvent alone (methanol) to be used as controls. We harvested the cultures and extracted total RNA, depleted ribosomal RNA and subjected them to RNA sequencing. Of the 6 control replicates 2 failed RNA quality control, and thus a total of 6 exposed and 4 control cultures were used for further analysis. Here, we provide that raw sequencing data as well as a list of all of the genes and their expression levels resulting from the RNA-sequencing. This dataset can be used as a reference data for future studies that access additional neuronal cell types, clozapine exposure conditions, and other antipsychotic medication.

DOI of original article: 10.1016/j.schres.2020.12.041
* Corresponding author at: 733N. Broadway, Miller Research Building 509, Baltimore, MD 21205.
E-mail address: adimitr1@jhmi.edu (D. Avramopoulos).
Social media: (A.N. Lam)
# The first three authors have contributed equally

https://doi.org/10.1016/j.dib.2021.106897

Related Research Article: Das, D., Peng, X., Lam, A.N., Bader, J.S., Avramopoulos, D., 2021. Transcriptome analysis of human induced excitatory neurons supports a strong effect of clozapine on cholesterol biosynthesis. Schizophr Res 228, 324–326. (Das et al., 2021).

Specifications Table

Subject	Biological Sciences
Specific subject area	Genetics: Human Neuroscience: Cellular and Molecular Neurons induced by Ngn2 expression from IPSCs Exposure to clozapine (300 ng/mL for three days) Transcriptomics: rRNA depletion
Type of data	Raw data, RNA Sequencing reads (Fastq) Processed data: All genes and their expression levels for all samples (.csv file)
How data were acquired	RNA Extraction was performed using the Zymo Quick RNA Miniprep kit. RNA-sequencing was performed by Novogene Corporation Inc. (Sacramento, CA) for rRNA depletion, library generation and RNA-sequencing.
Data format	Fastq files
Parameters for data collection	Using a rapid 21-day Ngn2-induction method, a publicly available control iPSC line was induced into excitatory cortical neurons [1]. At differentiation Day 18, we exposed 6 replicate excitatory neurons cultures to clozapine (300 ng/mL), and the remaining 6 to the methanol. Immunofluorescence and transcriptomics validated their glutamatergic neuronal identity.
Description of data collection	Six clozapine-exposed samples and four CH_3OH control samples (the solvent of clozapine) were included for RNA-sequencing after removing two CH_3OH control for low library concentration or high degradation. We harvested the two sets of culture, extracted total RNA using the depleted ribosomal RNA method and measured their transcriptomes by RNA sequencing.
Data source location	Institution: Johns Hopkins University School of Medicine City/Town/Region: Baltimore, Maryland Country: United States of America Latitude and longitude for collected samples/data: 39.2992° N, −76.5934° W
Data accessibility	Repository name: Gene Expression Omnibus (GEO) Data identification number: GSE163161 Direct URL to data: https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE163161 Instructions for accessing these data: Click on the link above. It will lead you to the GEO accession.
Related research article	Co-submission: Das, D., Peng, X., Lam, A.N., Bader, J.S., Avramopoulos, D., 2021. Transcriptome analysis of human induced excitatory neurons supports a strong effect of clozapine on cholesterol biosynthesis. Schizophr Res 228, 324–326. doi:10.1016/j.schres.2020.12.041

Value of the Data

- This data contributes to the understanding of the effect of clozapine on neuronal cells, the first analysis of its kind in human induced neurons.
- This data provides a method to elucidate possible mechanism(s) by which clozapine may alter disease state, an important step for future tailed treatments guided by precision medicine.

- Data can be used as baseline data for future studies that access additional neuronal cell types, clozapine exposure conditions, and other antipsychotic medication.

1. Data Description

1.1. This article presents data on the transcriptional profile of neuron cells treated with clozapine

This dataset consists of raw RNA-sequencing data from six clozapine-exposed samples and four CH_3OH control samples after removing two CH_3OH control for low library concentration or high degradation. We also provide processed data on the expression levels of all genes in .csv format. This file has a first column with gene names followed by columns 2–7 for each of the samples exposed to clozapine and columns 8–11 for each of the samples exposed only to the clozapine solvent (methanol).

The raw data are deposited to the Gene Expression Omnibus (GEO) database under the accession number GSE163161.

The proceeded data table (GSE163161_Processed_data.csv) can also be found under the same accession number.

Please use the direct URL to data:
https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE163161

2. Experimental Design, Materials and Methods

2.1. Induced pluripotent stem cell (iPSC) acquisition, culturing and maintenance

We obtained human iPSC lines from individuals without a psychiatric diagnosis from the National Institute of Mental Health (NIMH) at Rutgers University. For this study, we used the cell line MH0180967, a male of European descent, designated internally as ID RU01. We used Laminin (Biolamina)-coated 6-well cell culture plates to grow the cells and StemPro Accutase (Gibco) to dissociate them and seed them at the required density for experimentation. To enhance survival of dissociated cells, we added Rock inhibitor Y-27,632 dihydrochloride (Tocris) at 10 uM on the first day of passage to prevent dissociation-induced apoptosis. Finally, we tested cultured for mycoplasma contamination.

2.2. Ngn2 neuronal induction

We acquired Ngn2/rTTA expressing lentiviruses from the Core facility at the University of Pennsylvania (Lot# LV242 and LV243) and followed the protocol published by Zhang et al. [1]. This protocol yields high rates of converting iPSCs into functional neurons in under 2 weeks. We plated RU01 iPSCs at 250k per well on laminin-coated a 6-well cell culture plate in Stem Flex media with Rock inhibitor. After 24 h, we infected with Ngn2 lentivirus using Polybrene (Santa Cruz). After optimization for optimal differentiation, we transduced RU01 cells with 10 uL of Ngn2/rTTA.

2.3. Differentiation of Ngn2-transduced cells into neurons

On Day −2, we plated 250k Ngn2-transduced RU01 cells per well in laminin-located 6-well cell culture plates. On **Day −1**, we fed cultured cells with fresh Stem Flex media. On **Day 0**, we added Doxycycline to induce Ngn2 expression in induction media (DMEM/F12 (Thermo

Fisher), N2 (Thermo Fisher), D-Glucose (Thermo Fisher), 2-βME (Life technologies), Primocin (Invivogen), BDNF (10 ng/mL, Peprotech), NT3 (10 ng/mL, Peprotech), Laminin (200 ng/mL, Millipore Sigma) and Doxycycline (2 ug/mL, Sigma)). **On Day 1,** the induction media was supplemented with puromycin (2.5 ug/mL) to select for transduced cells. **On Day 2**, we harvested survivors and plated at 100k cells per well on matrigel-coated 24-well plates for immunocytochemistry or 1×10^6 cells per well in a 6-well plate for RNA sequencing in neural differentiation media (Neurobasal media (Thermo Fisher) Glutamax (Thermo Fisher), B27 (Thermo Fisher), D-Glucose (Thermo Fisher), Penn/Strep (Thermo Fisher), BDNF 10 ng/mL, Laminin 200 ng/mL, NT3 10 ng/mL and Doxycycline 2 ug/mL). Feeding with neural differentiation media continued every other day until **Day 12**. On **Day 4,** we treated the cells with 2 uM Cytosine β-D-arabinofuranoside hydrochloride (Ara-C) to inhibit proliferation of non-neuronal cells. Doxycycline induction continued until Day 12. Thereafter, cells were fed every two days until **Day 21** with Neural maturation media (Neurobasal media A (Thermo Fisher), B27, Glutamax, Penn/Strep, Glucose Pyruvate mix (1:100, final concentration of 5 mM glucose and 10 mM sodium pyruvate), BDNF (10 ng/mL), NT3 (10 ng/mL), Laminin (200 ng/mL) and Doxycycline (2 ug/mL). According to the protocol and supported by our immunohistochemistry and RNA-seq data, neurons were mature and ready for harvesting on **Day 21**.

2.4. Exposure of cells to clozapine and RNA extraction

On **Day 18**, we added clozapine to the medium at a concentration of 300 ng/mL, matching therapeutic plasma levels. **On Day 21**, neurons were fixed with PFA for ICC analysis or harvested using Accutase for RNA extraction. Six clones were exposed to clozapine and another six to methanol (CH_3OH), the solvent in which clozapine was dissolved in, to be used as controls. RNA extraction was performed on **Day 21** neurons (3 days post clozapine exposure) using the Zymo Quick RNA Miniprep kit. Novogene Corporation Inc. (Sacramento, CA) performed rRNA depletion, library generation and RNA-sequencing.

2.5. RNA-sequencing

We dropped two control lines due to quality control failure (low library concentration or high degradation). We received on average 28.2 million reads per sample with a range between 23.3 and 38 million. We used Hisat2 version 2.1.0 (Kim et al., 2019) to align the 150 bp paired-end reads to the human genome (GRCh38). We used Samtools 1.9 [2] to generate BAM files. After removing genes with low read counts (< 10) in order to decrease noise 19,296 genes were left for analysis. All 10 sequenced libraries passed Novogene's post-sequencing quality control. We used Stringtie 2.0.3 [3] to assemble transcripts and estimate abundance based on GRCh38 gene annotations [4]. We used the Bioconductor package tximport to compute raw counts by reversing the coverage formula used by Stringtie with the input of read length [5]. We then imported the output to DESeq2 (Bioconductor) for differential gene expression analysis [6]. We used the R package pheatmap for visualization. We performed principal component analysis using the Python package scikit-learn 0.22.1 (https://scikit-learn.org/stable/).

2.6. Immunofluorescence

We fixed neuronal cells grown in 24-well laminin-coated cell culture plates on freshly prepared 4% paraformaldehyde in PBS (pH 7.4) for 15 min after washing in Phosphate-Buffered Saline (PBS). We then permeabilized the cells with PBS plus 0.1% Triton X-100 for 15 min, blocked in 10% goat serum for 1 hr and stained with primary antibody at 4 °C overnight. After 24 h, we washed the cells 3 times with PBS +0.1% Tween20 (PBST) and incubated with

secondary antibodies for 1 hr. We then washed another three times with PBST and counterstained with DAPI diluted 1:1000 (Roche cat#10,236,276,001). The primary antibodies we used were for neuronal markers MAP2 diluted 1:250 (mouse, Sigma cat# M1406), Synapsin diluted 1:250 (Rabbit, Sigma cat# S193), and TUJ1 diluted 1:300 (Mouse, Biolegend cat#801,202).

2.7. Neuronal identity of IPSC-derived excitatory neurons

Twenty-one days after doxycycline induction, Ngn2/rTTA-transduced human iPSC cells showed morphological progression to neurons in Bright field images and the presence of cortical neuronal markers MAP2, Synapsin, and TUJ1 was confirmed by immunofluorescence staining [7]. RNA sequencing confirmed the expression of a series of genes expressing neuronal and glutamatergic markers such an NeuN, MAP2, TUJ1, VGlut1 and 2 and Glutaminase [7].

Ethics Statement

This study was approved by the Johns Hopkins University Institutional review board (protocol IRB00122135).

CRediT Author Statement

Debamitra Das: Laboratory Experimentation, manuscript reviewing; **Xi Peng:** Data curation guided by **Joel S. Bader** and Dimitrios Avramopoulos; **Anh-Thu N. Lam:** Writing original draft preparation, reviewing, editing; **Dimitrios Avramopoulos** and **Debamitra Das:** designed the project; **Dimitrios Avramopoulos:** provided funding for this project.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships which have or could be perceived to have influenced the work reported in this article.

Acknowledgments

This work was supported in part by NIMH grants P50 MH094268, R01 MH113215 and RF1 MH122936 to DA.

Bio-samples and/or data for this publication were obtained from NIMH Repository & Genomics Resource, a centralized national biorepository for genetic studies of psychiatric disorders.

References

[1] Y. Zhang, C. Pak, Y. Han, H. Ahlenius, Z. Zhang, S. Chanda, S. Marro, C. Patzke, C. Acuna, J. Covy, W. Xu, N. Yang, T. Danko, L. Chen, M. Wernig, T.C. Sudhof, Rapid single-step induction of functional neurons from human pluripotent stem cells, Neuron 78 (5) (2013) 785–798.

[2] H. Li, B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, R. Durbin, S. Genome Project Data Processing, The sequence alignment/Map format and SAMtools, Bioinformatics 25 (16) (2009) 2078–2079.

[3] M. Pertea, G.M. Pertea, C.M. Antonescu, T.C. Chang, J.T. Mendell, S.L. Salzberg, StringTie enables improved reconstruction of a transcriptome from RNA-seq reads, Nat. Biotechnol. 33 (3) (2015) 290–295.

[4] M. Pertea, D. Kim, G.M. Pertea, J.T. Leek, S.L. Salzberg, Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown, Nat. Protoc. 11 (9) (2016) 1650–1667.

[5] C. Soneson, M.I. Love, M.D. Robinson, Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences, F1000Res 4 (2015) 1521.
[6] M.I. Love, W. Huber, S. Anders, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2, Genome Biol. 15 (12) (2014) 550.
[7] D. Das, X. Peng, A.N. Lam, J.S. Bader, D. Avramopoulos, Transcriptome analysis of human induced excitatory neurons supports a strong effect of clozapine on cholesterol biosynthesis, Schizophr. Res. 228 (2021) 324–326.