Proceedings of ICMMGH 2026 Symposium: Cell and Molecular Biology: Understanding Immune Cell Functions
DOI: 10.54254/2753-8818/2026.PJ31204

Cell Type-specific Gene Expression Dysregulation Induced by CHD8 Loss in Mice: A Perspective from Neural Plasticity

He Yin

School of Engineering, The Hong Kong University of Science and Technology, Hong Kong, China
hyinal@connect.ust.hk

Abstract. In China and worldwide, neuroplasticity has been an active topic of research in autism studies, and since the human nervous system reacts to the external environment. In this project, CHD8+/- murine model was employed as a specific system to search for the molecular underpinnings of autism spectrum disorder, which means loss-of-function mutations in the transcriptional regulator CHD8 are responsible for causing disease. The strongest genetic risk factors for ASD are identified, RNA sequencing data from functionally impaired Chd8 KO mice displaying strong resemblance to human ASD condition showed dysregulation of networks involved in autism' s pathophysiology. By using the CHD8 mouse model and verifying the significant transcriptional changes, the aim is to conduct some progressive research on the potential mechanisms of autism spectrum disorders.

Keywords: Neuroplasticity, Autism, CHD8, Data Visualization

1. Introduction

1.1. Autism

Autism spectrum disorder (ASD) is a complex neurodevelopmental condition characterized by significant genetic and clinical heterogeneity. The most important features of the disorder include persistent challenges in social discussion and interchange, often co-occurring with restricted patterns of interest and repetitive behaviors [1]. Synaptic dysfunction represents a core pathological component of autism spectrum disorder, a finding consistently recapitulated across multiple established mouse models of the condition [2-4]. People with autism show significant impairments in social interaction and language communication, manifested by limited daily interests, repetitive and stereotyped movements, and strong preferences for specific foods. This neuronal activation, in turn, regulates synaptic development and function in the neural network by inducing the expression of multiple genes [5,6].

1.2. Influence of genetic variants on the risk of disease

Remarkably, protein-truncating variants (PTVs)—including nonsense, frameshift, and essential splice site mutations—exhibit a greater burden difference between ASD cases and controls than missense variants, indicating that their average impact on risk is likely to be higher [7]. Functional

DOI: 10.54254/2753-8818/2026.PJ31204

severity metrics, which evaluate the evolutionary constraint on deleterious variants—such as the probability of loss-of-function intolerance (pLI) score [8] and the integrated missense badness, PolyPhen-2, and constraint (MPC) score [9]—can further distinguish variant types with higher disease risk. Sequencing studies have shown that in unrelated individuals with ASD, genes harboring multiple de novo loss-of-function mutations are likely contributors to disease risk. Current research has established a core set of nine genes with the strongest evidence for conferring risk to autism spectrum disorder. These high-confidence risk loci encompass multiple genes, such as ANK2, CHD8, CUL3, DYRK1A, GRIN2B, KATNAL2, POGZ, SCN2A, and TBR1 [10]. The proteins encoded by these genes are involved in crucial biological processes, including chromatin remodeling and transcriptional regulation - this suggests that the pathogenesis of autism spectrum disorder (ASD) may be related to molecular dysfunction. Among them, the CHD8 gene demonstrates its crucial role in the genetic susceptibility of ASD, as inactivating mutations are most common in individuals with ASD.

1.3. CHD8

CHD8—a chromatin helicase—has emerged as one of the most prominent genetic risk factors implicated in autism spectrum disorder (ASD) etiology. Acting as an ATP-dependent chromatin remodeler, it binds to the promoter regions of a diverse set of genes, including multiple ASD-associated loci, and modulates their transcriptional output [11-13]. As a member of the SNF2 family of chromatin remodelers, CHD8 harbors loss-of-function mutations that trigger profound genome-wide alterations in chromatin accessibility, and it has emerged as one of the strongest candidate genes implicated in autism spectrum disorder (ASD) pathogenesis—garnering substantial attention as a leading focus of ASD research—with homozygous deletion of Chd8 in mice resulting in early embryonic lethality, a phenotype well-documented in preclinical studies [11,14,15]. Consequently, CHD8 heterozygous mice have been employed, which exhibit macrocephaly, motor delays, hypertelorism, pronounced hypoactivity, and aberrant responses to social stimuli [16]. It's obvious that CHD8's deficiency makes excitation neuronal progenitors in the forebrain and granule cell progenitors in the cerebellum fail to proliferate and differentiate, resulting in defective cerebral cortex and cerebellum development, and also playing key roles in the formation of special glial cell populations and the process of myelin production, inducing similar abnormalities to those shown by conditional mutant of CHD8 in the oligodendrocyte progenitor cell line [17]. Although these findings suggest that CHD8 serves as a central regulator of progenitor cell proliferation and differentiation in the brain, whether CHD8 similarly plays a pivotal role in neuronal development and overall brain maturation remains unresolved [18].

1.4. Research progress

Increasing evidence underlines the fact that autism susceptibility genes become coactivated within functional brain networks during early neurodevelopment; this notion was further elucidated by Willsey et al., who applied a detailed spatiotemporal transcriptomic atlas spanning human neurodevelopment from gestation through early adulthood. In their systematic coexpression analysis, they uncovered distinct transcriptional modules in each individual developmental window and brain region of interest, where nine well-established ASD-risk genes were enriched in networks related to similar functions within a specific module [10]. Another parallel study also confirmed that ASD risk genes showed a significant enrichment trend in the co-expression network in this developmental stage and brain region [19]. ASD analysis suggests that genes associated with the risk

DOI: 10.54254/2753-8818/2026.PJ31204

of autism spectrum disorder work together and may participate in gene regulation network formation related to autism spectrum disorder, which illustrates the collaborative nature of these genes. For the genes that were examined in this study, one with distinctive expression patterns during different developmental stages of the fetus was CHD8, which plays an important role in chromatin remodeling and directly interacted with a number of key transcriptional regulators. The specific expression pattern and molecular function ascribed to CHD8 make it a likely coordinator of gene regulatory networks described above [20]. Studies using both mouse and human embryonic stem cells (ESCs) have shown that when these cells differentiate into neural progenitor cells (NPCs) with reduced CHD8 function, the expression of genes critical for proper neural development becomes dysregulated [16,21,22]. A recent study has revealed that there is genome-wide transcription regulation disorder due to the chromatin remodeling dysfunction of CHD8. This mechanism has also been demonstrated for similar chromatin regulators such as CHD3 and CHD7: missense mutations associated with neurodevelopmental disorders (e.g., intellectual disability, autism spectrum disorder) and CHARGE syndrome have been shown to significantly inhibit ATP-dependent chromatin ununing activity of these proteins in vitro biochemically [23,24]. Breakpoints in the CHD8 gene were detected in three individuals, two of whom showed no signs of intellectual impairment. Although these findings suggest that higher CHD8 expression may be linked to the clinical features in these patients, it remains unclear whether CHD8 duplication directly causes neurodevelopmental disorders or how it might do so [25]. As of now, the only study that reported a behavioral phenotype similar to autism spectrum disorder associated with Chd8 deficiency was through the use of selective intrauterine electrical perforation technology to treat neural progenitor cells, reducing the expression of Chd8 in these cells to 20% of the wild-type level, thereby promoting the growth of upper neurons [26].

1.5. Research significance and objective

It remains unclear whether the molecular dysfunction of CHD8 directly contributes to the onset of Autism Spectrum Disorder. Following CHD8 knockdown, ASD-associated genes are downregulated in neural progenitor cells. In neural precursor cells, CHD8 is typically recruited to promoters enriched with permissive chromatin marks, suggesting its role in transcriptional activation [12,13]. Gene expression changes in Chd8+/- mice also appear to be strongly influenced by developmental stage, with more extensive transcriptional alterations observed during the perinatal and postnatal periods [16].This study uses the RStudio platform to build a standardized analysis workflow for RNA-seq differential expression. It applies two-way ANOVA for statistical testing, data processing, and result visualization. High-throughput sequencing results from ASD mouse models will be subjected to expression profiling to further elucidate the underlying molecular mechanisms.

2. Research method

2.1. Data sources

The study was conducted with the use of data from the National Center for Biotechnology Information (NCBI), an open, shared database, and the Gene Expression Omnibus (Gene Expression Omnibus) database. Chd8+/ -mouse data set retrieved from GEO (Small Extracellular RNA Sequencing, seRNA-seq) [27].

DOI: 10.54254/2753-8818/2026.PJ31204

2.2. Data processing method

The single-cell RNA sequencing (scRNA-seq) data from eight samples (GSE273765) were processed using the Seurat package in R. Raw gene expression matrices (in 10X Genomics format) were imported to create individual Seurat objects. Metadata annotations were applied, distinguishing four wild-type (WT) control and four autism model (Het) samples, with balanced sex distribution (four male and four female samples). Mitochondrial gene contamination was quantified using PercentageFeatureSet (pattern: ^mt-). Initial quality control was performed by visually screening three key metrics via violin plots: number of genes per cell (nFeature_RNA), UMI counts per cell (nCount_RNA), and mitochondrial gene percentage (percent.mt). Expression data were log-normalized (LogNormalize, scale factor = 10,000), followed by screening for highly variable genes (FindVariableFeatures, selection method = vst, 2,000 features). After merging all samples into a unified Seurat object, batch effect correction was performed using Harmony integration (dims = 1:30). Notably, principal component analysis (PCA) was conducted both before and after integration. Cell type annotation was performed using SingleR with the MouseRNAseqData reference atlas. Following data layer integration (JoinLayers), label transfer was executed based on RNA assay data. Differential gene expression (DEG) analysis between autism (Het) and control (WT) groups was conducted using FindMarkers, with the following thresholds: minimum expression cell proportion: 10% (min.pct = 0.1), absolute log2 fold-change threshold: > 0.25 (logfc.threshold = 0.25). Significantly differentially expressed genes (adjusted $P < 0.05$, $|\log2FC| > 0.5$) were visualized via volcano plots (EnhancedVolcano). Gene Ontology (GO) enrichment analysis was carried out with the clusterProfiler package and the org.Mm.eg.db database, covering all three categories: biological processes (BP), molecular functions (MF), and cellular components (CC).

2.3. Data processing step

• Data preprocessing (Quality Control & Adapter Trimming) was performed in RStudio, followed by quantification analysis (Quantification) using featureCounts, data normalization, differential expression analysis (Differential Expression, DE), functional enrichment analysis (Functional Enrichment), and visualization. In R, Seurat was employed for data integration, normalization (SCTransform), and dimensionality reduction (PCA + UMAP), converting raw single-cell expression data (sample.data) into Seurat objects (sample.obj).

• Sample-level metadata (metadata) was added to the Seurat objects in R, specifically assigning experimental condition (condition) and sex (sex) information to each sample (sample1.obj to sample8.obj).

• The percentage of mitochondrial gene expression (percent.mt) per cell was calculated and stored, and violin plots (Violin Plot) were generated to visualize key quality control (QC) metrics of the single-cell data.

• In R, NormalizeData was used to correct for sequencing depth differences, outputting log-transformed values. Log normalization eliminated the impact of sequencing depth, enabling cross-cell gene expression comparability. Subsequently, the VST method was applied to screen 2,000 highly variable genes (HVGs) exhibiting expression variation significantly higher than technical noise.

• To mitigate batch effects, IntegrateData() was applied after merge(), with batch correction effectiveness verified (Harmony + UMAP). After defining experimental groups (control vs. autism), automated cell type annotation (SingleR) was performed.

DOI: 10.54254/2753-8818/2026.PJ31204

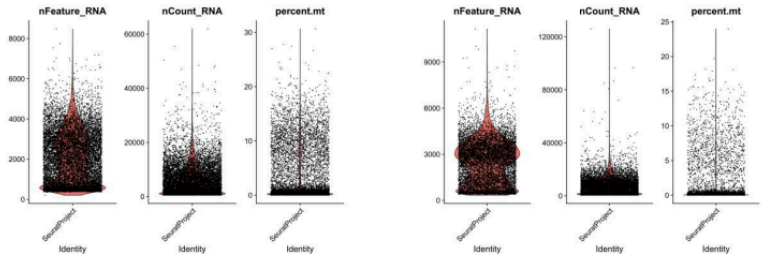
• Post-integration differential expression analysis (DEGs) identified potential autism biomarkers. Visualization (volcano plot) intuitively displayed the statistical significance of differential genes, and functional enrichment analysis (GO) was conducted to elucidate biological differences between the autism and control groups.

3. Research and discussion

3.1. Summary of single-cell RNA sequencing quality control results

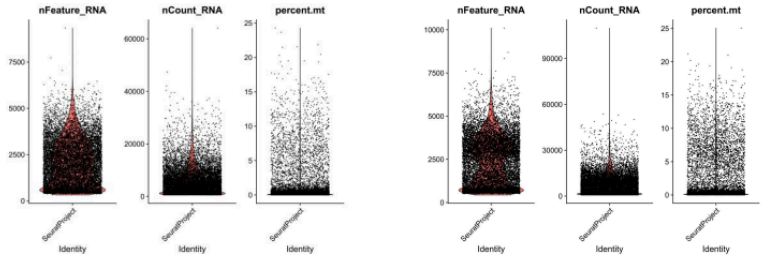
Quality control analysis confirmed that all eight samples showed high quality control data (Figure 1). Median gene count per cell (nFeature_RNA) ranged from about 2000 to 5000 for all samples, and the UMI counts (nCount_RNA) were primarily distributed between 10,000 and 40,000. Mitochondrial gene expression (percent. mt) was kept below 10%, except for sample S3 and S8, in which the latter displayed elevated mitochondrial levels, indicating cellular stress. Some cells had up to 15%, but they were within the permissible limit. Regarding the gene count distribution, both the Control (S1 - S4) and the Autism groups (S5 - S8) showed similar distributions. Yet there is an exception: two sample numbers from the Autism group (S6 and S8) having about a broad distribution with cells that exceeded 10,000 detected genes, signifying possible presence of doublets or possibly a large number of highly transcribed cells. For the female group (S3, S4, S7, S8), S8 had a much larger nFeature_RNA value than other samples, and thus should be given proper notice when conducting further analyses. For total UMI counts, most cells fell within the 5000–40000 range, but a subset of cells in sample S6 had nCount_RNA values above 150,000, indicating possible high-expression cell populations or doublets.

The part (a)-(d), quality control metrics for single-cell RNA-seq samples. Violin plots showing distribution of (A) number of genes detected per cell (nFeature_RNA), (B) total UMI counts per cell (nCount_RNA), and (C) percentage of mitochondrial gene expression (percent. mt) for samples S1–S8. Dashed lines indicate filtering thresholds. n = approximately 3000- 6000 cells per sample.



(a) This is the first part

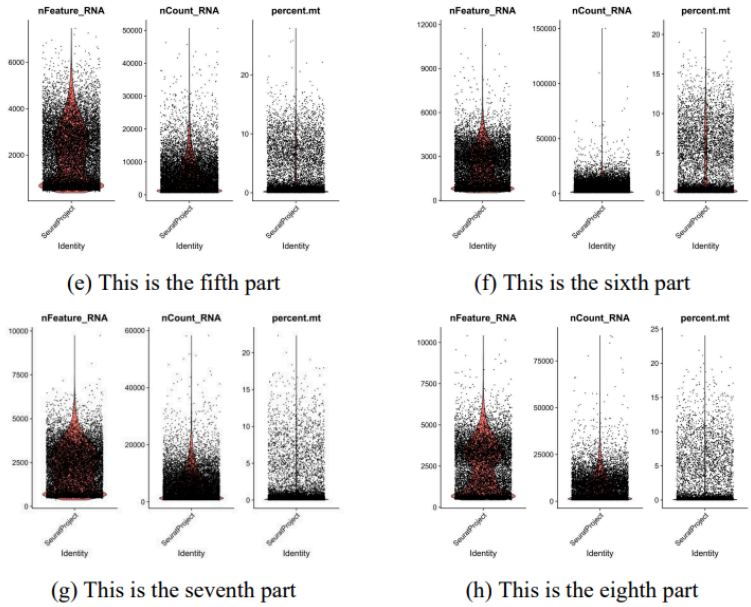
(b) This is the second part



(c) This is the third part

(d) This is the forth part

DOI: 10.54254/2753-8818/2026.PJ31204



(e) This is the fifth part (f) This is the sixth part

(g) This is the seventh part (h) This is the eighth part

Figure 1. Quality control metrics of eight single-cell RNA-seq samples based on gene expression UMI counts, and mitochondrial content

(a-h) Violin plots showing the distribution of three quality metrics—number of detected genes per cell (nFeature_RNA), total UMI counts per cell (nCount_RNA), and percentage of mitochondrial gene expression (percent.mt)—for individual samples. Panels (a-h) correspond to Sample 1 through Sample 8, respectively.Each point represents a single cell (n = total cells per sample, ranging from 2,000 to 5,000).The y-axis represents gene counts (nFeature_RNA, unitless), UMI counts (nCount_RNA, unitless), and percent.mt (percentage, %). All samples were untreated controls collected under identical conditions. Cells were filtered with standard thresholds (200 < nFeature_RNA < 7,500; percent.mt < 20%) prior to downstream analysis.All violin plots display the full distribution of values with embedded boxplots; the central line denotes the median, and the upper/lower limits indicate the interquartile range (IQR).

3.2. To remove batch effect among different samples, we integrated multiple samples

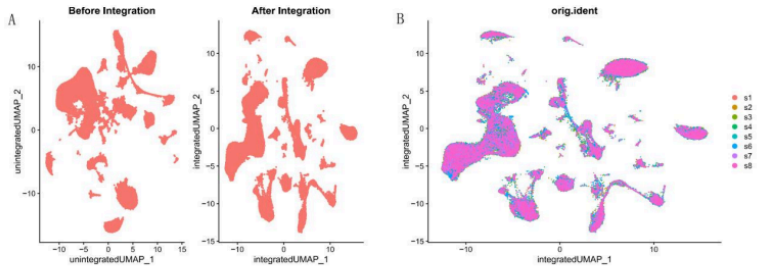
(S1-S8) using Seurat's built-in integration method (CCA) and Harmony method, respectively.

3.2.1. Clustering analysis was performed using the Louvain algorithm implemented in Seurat (v5.2.1)

A shared nearest neighbor(SNN) of the total of 97,772 cells, all cells were analyzed and a SNN graph consisting of about 3.1 million edges was constructed. The clustering resolution was set to 2.0 in order to get fine grained cell subpopulation clusters. Louvain algorithm was performed 10 times with random initialization and gave out a maximum modularity of 0.8935. 66 clusters have been identified through the initial analysis, including 4 clusters consisted of a single cell only the end result is 62 separate cell clusters. Comparisons of data from before and after Harmony integration is done via UMAP visualization as shown in Fig. 2. In the left panel of the figure is a UMAP projection of single-cell transcriptomes before batch effect corrections are made, where poorly mixed data points and poorly defined projection ranges show the presence of batch effects. After

DOI: 10.54254/2753-8818/2026.PJ31204

applying Harmony integration, the cells cluster better in the lower dimensional space, and the removal of technical variability across different samples can be observed (right panel). The integrated data (colored by sample s1-s8) exhibits well-mixed cell distribution across all clusters, confirming successful data integration.



(a) This is the first part (b) This is the second part

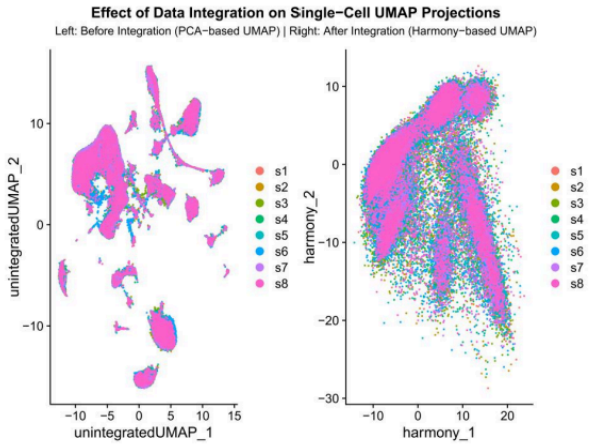
Figure 2. Impact of data integration on single-cell UMAP representations

(a) UMAP visualization of single-cell transcriptomic data before (left) and after (right) data integration.Prior to integration, cells exhibit a dispersed and unstructured distribution in the unintegrated UMAP space, with coordinates spanning approximately −10 to 15 along UMAP_1 and −10 to 10 along UMAP_2.Following integration, the cell populations are more compactly organized within a shifted coordinate space (UMAP_1: −15 to 10; UMAP_2: −10 to 10), indicating improved alignment of shared cellular structures across batches and reduced technical variation. (b) UMAP projection of the integrated dataset, colored by sample identity (s1–s8).Although each sample maintains distinguishable clustering patterns, considerable overlap is observed across samples, reflecting successful integration of biological replicates. The preservation of sample-specific characteristics alongside shared cell populations suggests effective harmonization of inter-sample variability.

3.2.2. Harmony algorithm was used to integrate scRNA-seq data

Before integration (Figure.3 left), PCA principal component was used for UMAP dimensionality reduction. The results showed that cells from different samples (S1-S8) clustered into multiple independent clusters, indicating that there was an obvious batch effect. After integration (right of Figure 3), UMAP dimension reduction using the Harmony corrected principal components can be observed that cells from different samples are fully mixed in the low-dimensional space, and batch specificity disappears, which validates that the Harmony algorithm effectively eliminates technical differences while preserving biological differences. In contrast, the Harmony integration system showed more continuous cell distribution, more full color mixing of each sample, and better batchless effect ability.

DOI: 10.54254/2753-8818/2026.PJ31204



(a) This is the first part (b) This is the second part

Figure 3. Harmony-based integration mitigates batch effects in multi-sample single-cell UMAP projections

(a) UMAP visualization of single-cell RNA-seq profiles before data integration by PCA dimensionality reduction. This panel shows the eight samples (s1–s8) before batch correction. Dimensionality reduction was performed using principal component analysis (PCA), followed by UMAP embedding. The x- and y-axes represent unintegratedUMAP_1 and unintegratedUMAP_2, respectively; both axes are unitless and reflect relative transcriptomic distances in the reduced feature space. Each dot denotes a single cell and is color-coded by its sample of origin. Cells from different samples form distinct, non-overlapping clusters (e.g., s1 in pink, s2 in orange), suggesting strong batch effects likely driven by technical variation rather than true biological differences. No statistical tests were applied in this panel; the visualization provides a qualitative baseline. A total of 97,772 cells are included across all samples. (b) UMAP projection after data integration using Harmony. This panel shows the same dataset following batch correction with the Harmony algorithm, applied across the top 30 principal components. The x- and y-axes represent harmony_1 and harmony_2, respectively, both of which are unitless. Cells remain color-coded by sample identity. After integration, cells from different samples appear more intermixed, and previously distinct sample-specific clusters are no longer clearly separated (e.g., s1 cells now overlap with s2 and other samples). This improved alignment indicates that Harmony effectively reduces batch-specific effects while preserving biological structure. No experimental treatments or time points were applied to the samples. Although no formal statistical tests were conducted, the increased overlap and continuity among samples serve as qualitative evidence of successful integration. The total number of analyzed cells remains 97,772, consistent with Panel (a) .

3.3. UMAP visualization of SingleR-annotated cell types from single-cell RNA-seq data (Figure 4)

Each dot represents a cell, and the different colors indicate the different cell types annotated by SingleR based on the MouseRNAseq reference dataset. To identify the cell type in the sample, The SingleR tool was used in combination with MouseRNAseq Single cell data were automatically annotated by the reference data set, and the annotated cell types formed obvious clusters in two-dimensional space.The main identified cell types included neurons, oligodendrocytes, astrocytes, microglia, monocytes, epithelial cells, endothelial cells, fibroblasts, and red blood cells.

DOI: 10.54254/2753-8818/2026.PJ31204

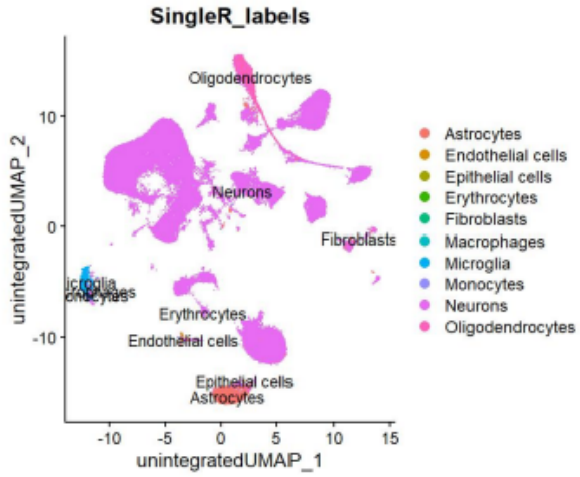


Figure 4. Cell type annotation of unintegrated single-cell RNA-seq data based on SingleR and MouseRNAseq reference

The UMAP graph visually presents the cell type annotation of the single-cell RNA sequencing data before batch integration. Each dot represents an individual cell(n = 97, 772), and colors indicate cell types automatically assigned using the SingleR tool, referencing the MouseRNAseq dataset.The x-and y-axes(unintegratedUMAP_1 and unintegratedUMAP_2) are unitless dimensions that reflect relative transcriptomicdistances in two-dimensional space. SingleR-based annotation reveals clearly clustered populations corresponding to distinct cell types, including neurons, oligodendrocytes, astrocytes, microglia, monocytes, epithelial cells, endothelial cells, fibroblasts, and erythrocytes. The observed clustering suggests that transcriptomic profiles alone are sufficient to distinguish major neural and non-neural cell types within the dataset. This figure provides a baseline map of cellular identities for subsequent integration and downstream analyses. No statistical tests or experimental treatments were applied across samples in this panel.

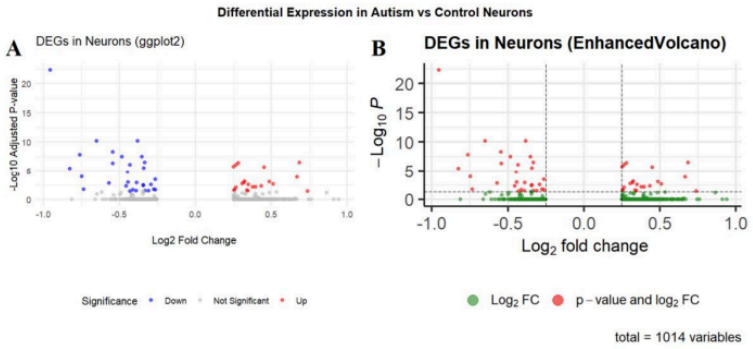
3.4. Enrichment analysis

3.4.1. Comparison of volcano plots

To pinpoint the transcriptional changes linked to autism within neurons, we performed a differential gene expression analysis between ASD samples and controls. The left panel of Figure 5 shows the results from the Seurat pipeline using a fairly permissive threshold (log2FC > 0.25 and adjusted p-value < 0.05). What we see is that ASD neurons have a lot more significantly upregulated (red) or downregulated (blue) genes compared to controls, with the downregulated ones being noticeably more dominant. This pattern hints that the overall gene expression program in ASD neurons might be under some form of suppression. For comparison, the right panel shows what we got using ShinyGO with much stricter filters (|log2FC| > 1 and FDR < 0.01). This time, we ended up with only 193 differential genes. What's interesting is that the fold changes for most of these genes aren't actually huge, but their p-values are all very strong. Among them, the gene 4930415C11Rik is the one that was knocked down the most. This makes it a strong candidate for directly playing a role in the neuronal dysfunction seen in ASD. Putting these two views side by side, we can draw a practical takeaway: using a looser cutoff with Seurat helps us see the broader shifts in the transcriptome,

DOI: 10.54254/2753-8818/2026.PJ31204

while tightening the criteria lets us zoom in on the most reliable and potentially central gene changes.



(a) This is the first part (b) This is the second part

Figure 5. Differential gene expression analysis in neurons from Autism versus Control mice

Part (a) Volcano plot showing differentially expressed genes (DEGs) between Autism and Control groups in neuronal cells, based on single-cell RNA-seq data. The x-axis indicates the Log_2 fold change (Log_2FC), and the y-axis represents the $-Log_{10}$ adjusted P-value (Benjamini-Hochberg correction). Genes significantly upregulated in Autism are shown in red (Up), downregulated in blue (Down), and non-significant genes in gray. Part (b) Volcano plot highlighting DEGs under stricter thresholds (FDR < 0.01, $|Log_2FC| > 1$). Each point represents one gene (total = 193 variables). Genes satisfying both criteria (FDR and fold change) are colored red; those meeting only the fold change threshold are in green. One key gene, 4930415C11Rik, is annotated as a representative significant DEG. Statistical significance is determined using the Wilcoxon rank-sum test, followed by FDR correction. No error bars are shown as data represent population-wide differential expression rather than mean-based measurements. Axes are dimensionless (Log-transformed values).

3.4.2. Oxidative phosphorylation pathway

KEGG pathway enrichment analysis of differentially expressed genes (DEGs) showed that genes involved in oxidative phosphorylation (OXPHOS) were significantly enriched, as shown in Figure 6, mitochondrial respiratory complex I (NADH dehydrogenase), III (cytochrome bc1 complex) in autism samples, Multiple components of IV (cytochrome c oxidase) and V (ATP synthase) were downregulated, and in complex I, including NDUFS1, NDUFS4, NDUFB4, and NDUFA1, which are essential for electron transfer and NADH oxidation, were observed to be downregulated. Similarly, complex III subunits such as UQCR10, UQCRQ and the Complex IV subunit COX4I1 are significantly reduced, possibly affecting the efficiency of the electron transport chain (ETC). ATP synthase (Complex V) subunits such as ATP5E and ATP5O are also inhibited, implying that autistic neurons have reduced ATP production capacity. The widespread downregulation of OXPHOS complexes suggests that mitochondrial dysfunction is a recurring theme in the pathophysiology of autism spectrum disorder (ASD). Impaired mitochondrial bioenergetics may result in inadequate energy supply to neurons, altered synaptic transmission, and increased oxidative stress.

DOI: 10.54254/2753-8818/2026.PJ31204

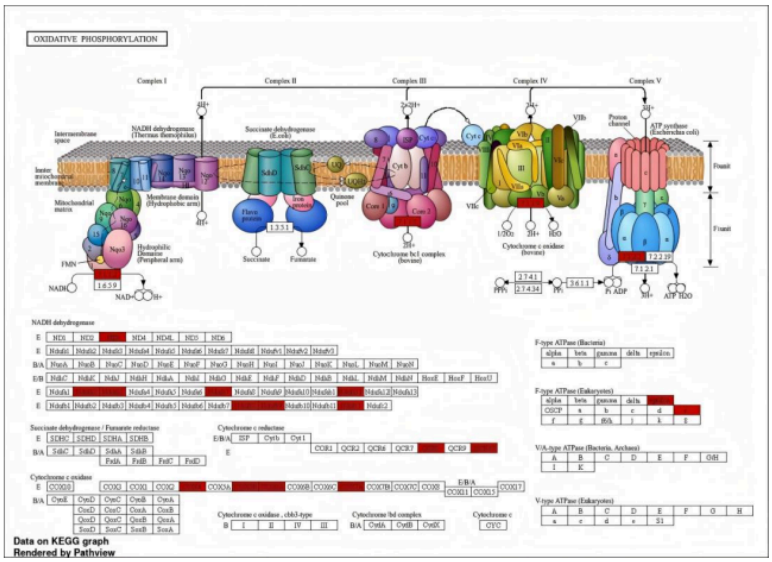


Figure 6. Schematic overview of the oxidative phosphorylation pathway and its core protein complexes

There are 5 key protein complexes (I to V), which are located inside the inner mitochondrial membrane. They work as one to carry out oxidative phosphorylation. NADH and succinate pass electrons to ubiquinone (UQ) via Complex I (NADH oxidase) and Complex II (succinate oxidase). After that, electrons pass across complex III (cytochrome bc1 complex) and cytochrome C (Cyt c), then arriving at complex IV (cytochrome c oxidase). During this process, molecular oxygen is converted to water, and protons (H^+) are pumped across the inner mitochondrial membrane into the intermembrane space, creating an electrochemical gradient to make an electrochemical potential across the membrane through electron transfer processes. The energy then catalyzes the synthesis of ATP from Complex V (ATP synthase). The highlighted subunits present in the figure represent altered genes or protein expression (upregulated) as per KEGG annotation. The labels include KEGG homologues, such as ND1–ND6, NDUFA–NDUFS (complex I), SDHA–SDHD (complex II), COR/ISP/Cytb (complex III), COX1–COX6 (complex IV), and ATP5 subunits (complex V). The figure is intended only to illustrate the pathway and is based on a KEGG pathway map produced using Pathview; thus, units, statistical analyses, or sample sizes from experiments are not included.

3.5. Discussion and prospect

The present study, which conducted a comprehensive transcriptome differential analysis based on single-cell RNA sequencing (scRNA-seq), found differences between ASD cases and controls with respect to the transcriptomes of different cell types. When performing the integrated single-cell transcriptome analyses, we found characteristic molecular signatures for ASD. By systematically processing eight single-cell RNA-seq samples (four ASD models [Het] and four sex-balanced wild-type [WT] samples), we used Harmony-based batch correction methods. The experiment constructed a complete cell map consisting of 97,772 cells and 32,285 features. Using 30 harmonious dimensions (Louvain algorithm, resolution = 2) for unsupervised clustering, different transcriptional cell clusters were identified. The reference maps of these cell clusters were annotated onto the MouseRNAseqData map. In the differential gene expression analysis, a total of 78 genes (adjusted p-value < 0.05, $|\log_2 FC| > 0.5$) were statistically significant, including the

DOI: 10.54254/2753-8818/2026.PJ31204

neurodevelopment-related factors Maml3, Pou6f2, Nr4a3, Fabp7, and Htr2c. It can be speculated that the oxidative phosphorylation disorder may be one of the molecular bases of the neurological dysfunction in autism. The above differential expression analysis revealed moderate but statistically significant transcriptional changes between neurons from autistic patients and control samples.

Both ggplot2 and EnhancedVolcano visualizations highlighted a consistent group of differentially expressed genes (DEGs), reinforcing the robustness of the analysis. Although no genes exhibited large fold changes ($|\log_2 FC| > 1$), the reproducible nature of these subtle alterations suggests potential biological relevance. These mild but consistent dysregulations may reflect the intricate transcriptional landscapes underlying autism, rather than extensive reprogramming. Several of the identified genes have been previously implicated in neurodevelopmental regulation and synaptic signaling. For instance, Maml3 (Mastermind-like transcriptional coactivator 3) has been shown to modulate Notch signaling, a pathway critical for cortical neurogenesis and neuronal differentiation. Abnormal functioning of the Notch signal pathway may lead to distorted brain structure and abnormal behavior in ASD models. Pou6f2 is also a POU domain transcription factor that regulates the formation of neuronal subtypes. It also participates in cortical identity and neuronal plasticity regulation. Additionally, Nr4a3 (also known as NOR-1) is an important nucleus-receptor that functions in the activity-dependent expression of genes and the excitability of neurons. The increase of Nr4a family genes are correlated with imbalance between excitation and inhibition, which is one of the key physiological aspects of ASD. Fabp7, an astrocyte-expressed fatty acid-binding protein, has been implicated in early brain development and neurogenesis. Elevated Fabp7 expression has been reported in human ASD brain tissues and may reflect astrocyte reactivity or altered lipid metabolism.Additionally, Htr2c, a serotonin receptor subtype expressed in various brain regions, plays a critical role in synaptic transmission and neuropsychiatric behaviors. Altered serotonergic signaling has long been implicated in ASD, and modulation of Htr2c activity is considered a potential therapeutic avenue. The over-representation of differentially expressed genes from the OXPHOS pathway indicates that mitochondrial dysfunction is a common cause of all three phenotypes. Previous findings have shown decreased mitochondrial function and energy metabolism in people with ASDs. Although missense variants represent the most frequent type of mutation identified in the CHD8 gene among individuals with autism spectrum disorder, their functional consequences on CHD8 activity and subsequent behavioral outcomes are still poorly characterized [28]. Based on evidence currently available, we found potential complicated connection among genetic regulation, cellular metabolism and neurodevelopmental pathway which may jointly lead to ASD. This discovery provides new direction for research to understand multifactorial etiology of ASD. It also stimulates us to pay close attention to potential interactions between metabolism and neurodevelopment; likewise epigenetic regulation may contribute to this process, but more experimental evidence is still needed.

4. Conclusion

This study constructed a high-resolution single-cell transcriptome map specifically for autism spectrum disorder (ASD); the data analysis revealed that DEGs in the oxidative phosphorylation pathway were abundantly enriched in this pathway. Therefore, mitochondrial dysfunction of neurons and basic energy deficiency may well be one of the possible core mechanisms to explain autism. it is worth noting that this model's specificity. These insights require verification from other systems; while CHD8 error mutation is quite pronounced in our study data, we still need to confirm its exact function through experiments. Currently, we should carry out functional tests for all of these DEGs and interfere with them directly to observe if this can destroy neuron energy metabolism and cause

DOI: 10.54254/2753-8818/2026.PJ31204

dysfunction. We want to examine whether this metabolic hypothesis can not only explain the pathogenesis of autism in a more detailed mechanism, but also reveal new therapeutic methods.

References

[1] Lord, C. et al. Autism spectrum disorder. Nat. Rev. Dis. Prim. 6, 5 (2020).

[2] Tabuchi K, Blundell J, Etherton MR, et al. A neuroligin-3 mutation implicated in autism increases inhibitory synaptic transmission in mice. Science. 2007; 318(5847): 71-76. doi: 10.1126/science.1146221

[3] Peça, J., Feliciano, C., Ting, J. et al. Shank3 mutant mice display autistic-like behaviours and striatal dysfunction. Nature 472, 437–442 (2011). https: //doi.org/10.1038/nature09965

[4] Bourgeron T. From the genetic architecture to synaptic plasticity in autism spectrum disorder. Nat Rev Neurosci. 2015; 16(9): 551-563. doi: 10.1038/nrn3992

[5] Ebert DH, Greenberg ME. Activity-dependent neuronal signalling and autism spectrum disorder. Nature. 2013; 493(7432): 327-337. doi: 10.1038/nature11860

[6] Minatohara, K., Akiyoshi, M. & Okuno, H. Role of immediate-early genes in synaptic plasticity and neuronal ensembles underlying the memory trace. Front. Mol. Neurosci. 8, 78 (2015).

[7] He X, Sanders S J, Liu L, et al. Integrated model of de novo and inherited genetic variants yields greater power to identify risk genes [J]. PLoS genetics, 2013, 9(8): e1003671.

[8] Kosmicki J A, Samocha K E, Howrigan D P, et al. Refining the role of de novo protein-truncating variants in neurodevelopmental disorders by using population reference samples [J]. Nature genetics, 2017, 49(4): 504-510.

[9] Samocha K E, Kosmicki J A, Karczewski K J, et al. Regional missense constraint improves variant deleteriousness prediction [J]. BioRxiv, 2017: 148353.

[10] Willsey AJ, Sanders SJ, Li M, et al. Coexpression networks implicate human midfetal deep cortical projection neurons in the pathogenesis of autism. Cell. 2013; 155(5): 997-1007. doi: 10.1016/j.cell.2013.10.020

[11] Nishiyama M, Oshikawa K, Tsukada Y, et al. CHD8 suppresses p53-mediated apoptosis through histone H1 recruitment during early embryogenesis. Nat Cell Biol. 2009; 11(2): 172-182. doi: 10.1038/ncb1831

[12] Sugathan A, Biagioli M, Golzio C, et al. CHD8 regulates neurodevelopmental pathways associated with autism spectrum disorder in neural progenitors. Proc Natl Acad Sci U S A. 2014; 111(42): E4468-E4477. doi: 10.1073/pnas.1405266111

[13] Cotney J, Muhle RA, Sanders SJ, et al. The autism-associated chromatin modifier CHD8 regulates other autism risk genes during human neurodevelopment. Nat Commun. 2015; 6: 6404. Published 2015 Mar 10. doi: 10.1038/ncomms7404

[14] Nishiyama M, Nakayama K, Tsunematsu R, Tsukiyama T, Kikuchi A, Nakayama KI. Early embryonic death in mice lacking the beta-catenin-binding protein Duplin. Mol Cell Biol. 2004; 24(19): 8386-8394. doi: 10.1128/MCB.24.19.8386-8394.2004

[15] Katayama Y, Nishiyama M, Shoji H, et al. CHD8 haploinsufficiency results in autistic-like phenotypes in mice. Nature. 2016; 537(7622): 675-679. doi: 10.1038/nature19357

[16] Suetterlin P, Hurley S, Mohan C, et al. Altered Neocortical Gene Expression, Brain Overgrowth and Functional Over-Connectivity in Chd8 Haploinsufficient Mice. Cereb Cortex. 2018; 28(6): 2192-2206. doi: 10.1093/cercor/bhy058

[17] Kawamura, A. et al. Oligodendrocyte dysfunction due to Chd8 mutation gives rise to behavioral deficits in mice. Hum. Mol. Genet. 29, 1274–1291 (2020).

[18] Kawamura, A., Nishiyama, M. Deletion of the autism-related gene Chd8 alters activity-dependent transcriptional responses in mouse postmitotic neurons. Commun Biol 6, 593 (2023).

[19] Parikshak, N. N. et al. Integrative functional genomic analyses implicate specific molecular pathways and circuits in autism. Cell 155, 1008–1021 (2013) .

[20] Bernier, R. et al. Disruptive CHD8 mutations define a subtype of autism early in development. Cell 158, 263–276 (2014) .

[21] Sood S, Weber CM, Hodges HC, Krokhotin A, Shalizi A, Crabtree GR. CHD8 dosage regulates transcription in pluripotency and early murine neural differentiation. Proc Natl Acad Sci U S A. 2020; 117(36): 22331-22340. doi: 10.1073/pnas.1921963117

[22] Ding S, Lan X, Meng Y, et al. CHD8 safeguards early neuroectoderm differentiation in human ESCs and protects from apoptosis during neurogenesis. Cell Death Dis. 2021; 12(11): 981. Published 2021 Oct 22. doi: 10.1038/s41419-021-04292-5

DOI: 10.54254/2753-8818/2026.PJ31204

[23] Snijders Blok L, Rousseau J, Twist J, et al. CHD3 helicase domain mutations cause a neurodevelopmental syndrome with macrocephaly and impaired speech and language. Nat Commun. 2018; 9(1): 4619. Published 2018 Nov 5. doi: 10.1038/s41467-018-06014-6

[24] Bouazoune K, Kingston RE. Chromatin remodeling by the CHD7 protein is impaired by mutations that cause human developmental disorders. Proc Natl Acad Sci U S A. 2012; 109(47): 19238-19243. doi: 10.1073/pnas.1213825109

[25] Kawamura A, Fujii K, Tamada K, et al. Duplication of the autism-related gene Chd8 leads to behavioral hyperactivity and neurodevelopmental defects in mice. Nat Commun. 2025; 16(1): 4641. Published 2025 May 26. doi: 10.1038/s41467-025-59853-5

[26] Durak O, Gao F, Kaeser-Woo YJ, et al. Chd8 mediates cortical neurogenesis via transcriptional regulation of cell cycle and Wnt signaling. Nat Neurosci. 2016; 19(11): 1477-1488. doi: 10.1038/nn.4400

[27] Yim KM, Baumgartner M, Krenzer M, et al. Cell type-specific dysregulation of gene expression due to Chd8 haploinsufficiency during mouse cortical development. Preprint. bioRxiv. 2024; 2024.08.14.608000. Published 2024 Aug 15. doi: 10.1101/2024.08.14.608000

[28] Shiraishi T, Katayama Y, Nishiyama M, et al. The complex etiology of autism spectrum disorder due to missense mutations of CHD8. Mol Psychiatry. 2024; 29(7): 2145-2160. doi: 10.1038/s41380-024-02491-y