Research Article

Analyzing the Expression and Function of Autism-associated Genes Throughout the Brain

Apoorva Siragavarapu

Central Bucks High School South, 1100 Folly Road, Warrington, PA 18976, United States

ABSTRACT

Autism Spectrum Disorder (ASD) is one of the most prevalent neurodevelopmental conditions, affecting approximately 1 in 54 children. Due to the complexity of the disorder, the pathogenesis of ASD is multifaceted. Pre-existing research has linked many genes to ASD; however, the mechanisms by which these genes contribute to ASD are not always clear. This study aims to identify similarities between ASD related genes. By analyzing their similarities in expression in the cerebellum, hippocampus, and prefrontal cortex, this study identifies a brain region with the highest expression of autism-associated genes. TSC1, SHANK2, TAOK1, and CHD8 are genes that were chosen to analyze as they are high-confidence ASD genes. Phase 1 includes gathering information from published literature on how the genes cause ASD. Phase 2 and 3 use RNA sequencing databases to analyze the percent expression of the genes in the cerebellum, hippocampus, and prefrontal cortex. It was found that cerebellar purkinje neurons have the highest expression (84.24%) of the selected autism-associated genes. Additional research is needed to better understand how autism-associated genes affect brain functions, so identifying regions like this may be promising targets which can lead to the treatment of ASD.

Keywords: Autism Spectrum Disorder; Purkinje Neurons; Cerebellum; Hippocampus; Prefrontal Cortex

INTRODUCTION

Autism Spectrum Disorder (ASD) is one of the most common neurodevelopmental conditions in children, acting as a heterogenous group of disorders (1). It affects learning, communication, memory, and behavior (2). Although people diagnosed with ASD share similar symptoms, their symptoms and intellectual functions significantly differ (3). Moreover, ASD co-exists with other conditions: motor abnormalities, gastrointestinal problems, epilepsy, intellectual disability, and sleep disorders (4).

ASD is characterized by three main clinical features: challenges in social interactions, deficits in verbal and nonverbal communication, and restricted interests (5). ASD presents with varying cognitive and language abilities. Some individuals are severely impaired, unable to speak, and display repetitive motor behaviors or self-injury (2). In contrast, others with high-functioning autism may be highly skilled, fluent in specialized topics, and engage in one-sided conversations (2). ASD is categorized into three levels: Level 1 (high functioning with minimal speech difficulties), Level 2 (moderate

Corresponding author: Apoorva Siragavarapu, E-mail: apoorvasirag@gmail.com.

Accepted August 27, 2025
https://doi.org/10.70251/HYJR2348.35916

challenges), and Level 3 (severe impairments with little to no verbal communication) (2).

The pathogenesis of ASD is multifaceted. Genetic studies have linked hundreds of genes to ASD; however, a better understanding of how these genes interact with each other is needed. This presents a gap in knowledge of how specific autism-associated genes interact with each other in different parts of the brain to cause ASD. This study analyzes autism-associated genes in different brain regions: cerebellum, hippocampus, and prefrontal cortex. As primary regions of the brain that contribute to controlling social cognition, these regions are suitable for the analysis of autism-associated genes. By analyzing the expression of these genes in the cerebellum, hippocampus, and prefrontal cortex, this study aims to discover a region in the brain with the highest expression of autism-associated genes.

The cerebellum is an important region of the human brain as it controls motor movement regulation and balance control (6). It is also involved in the coordination of posture, balance, voluntary movements, and muscle tone. The cerebellum is mainly known for these functions, but it is also important for cognitive and social behaviors. Reduced numbers of purkinje cells, abnormal connectivity, and genetic factors, occur in the cerebellum in ASD affected individuals (7). Mutations in autism-associated genes can affect the cerebellum, and its ability to connect to other brain regions (8).

The hippocampus is responsible for memory formation, memory transfer, and social navigation (6). The hippocampus organizes, stores, and retrieves memories within the brain and is responsible for cognitive functions such as learning, short and long-term memory, visual-spatial memory, verbal memory, and declarative memory (6). It also works with the amygdala to connect memories to emotions to create an emotional response. The hippocampus is part of the limbic system and is connected to the hypothalamus and amygdala to help regulate many bodily functions (6). Individuals with ASD usually have impairments in learning, memory, and cognitive map creation which are all hippocampal functions (6). In ASD the hippocampus is found to be enlarged which results in deficits in memory for person and emotion-related stimuli (9).

The prefrontal cortex located in the forefront of the brain serves as the brain's control center, regulating behavior and controlling cognitive processes (10). It is a large and complex region in the frontal lobe which has extensive connections with other brain regions (12). This network allows the prefrontal cortex to combine information and coordinate cognitive functions (10). It receives incoming information from the social world and adapts to the situational surrounding accordingly (10). Moreover, it controls reactions based on a person's perception of the situation (10). In ASD, the prefrontal cortex is observed to have structural differences which could reflect an unusual neurodevelopmental process and could contribute to the behavioral effects of ASD (10).

Overall, ASD is a highly complex condition characterized by a wide range of behaviors and an extensive genetic basis. Due to this complexity, it remains unclear whether autism-associated genes exhibit consistent functions across individuals or if their roles vary. Researchers aim to determine whether these genes contribute to ASD through similar mechanisms and to understand the associated processes. This study involves gathering information about autism-associated genes from published literature and using this information to analyze the RNA sequencing data to identify specific brain regions with the highest expression of autism-associated genes. By focusing on these regions, this study aims to provide insights into the genetic contributions to ASD development, potentially advancing our understanding of its biological mechanisms.

METHODS AND MATERIALS

This study was conducted in three sequential phases: (1) gathering gene information, (2) accessing RNA sequencing data and (3) statistically analyzing data.

Phase 1 – Gathering Gene Information

Four genes commonly associated with ASD were selected for analysis: TSC1, SHANK2, TAOK1, and CHD8. Research papers related to these genes and ASD were selected and analyzed. Particular attention was given to findings related to three brain regions frequently implicated in ASD: the cerebellum, hippocampus, and prefrontal cortex. Notes were compiled on each gene's functional role, mechanisms contributing to ASD, and region-specific involvement. Key insights from the research papers were synthesized to establish a foundational understanding of gene-brain region relationships.

Phase 2 – Accessing RNA Sequencing Data

Single-cell RNA sequencing data were obtained from the Brain Cell Data Viewer, a publicly available resource providing transcriptional profiles of cell types in the mouse brain (12). To create a comprehensive atlas of

cell types in each brain structure, the researchers paired high-throughput single-nucleus RNA sequencing with Slide-seq (a spatial transcriptomics method with near-cellular resolution) across the entire mouse brain.

The RNA sequencing data that was used was created by researchers who got gene information from neurons, and then clustered cells based on shared expression patterns using machine learning (12). Then, they further grouped the clusters into metaclusters which generally correspond to predefined neuronal cell types. Metaclusters help researchers identify broader biological trends by grouping related genes together based on common characteristics.

The four target genes (TSC1, SHANK2, TAOK1, CHD8) were selected within the Single Cell Data Viewer interface. Cell classes were filtered to include only neuronal cell types which would limit the data to the brain. Brain regions corresponding to the cerebellum, hippocampus, and prefrontal cortex were selected by choosing the following anatomical structures cerebellum, hippocampal formation, prelimbic, infralimbic, and anterior cingulate areas which are functionally analogous to the mouse prefrontal cortex. Expression data for each gene within these brain regions were exported in CSV format for analysis.

Phase 3 – Statistically Analyzing Data

Exported RNA sequencing data were processed and analyzed using Microsoft Excel. For each brain region, the average percent expression of each gene in each metacluster was calculated. The percent expression threshold for the metaclusters was 60%. The cerebellum and hippocampus each had three metaclusters with the highest average expression, while the prefrontal cortex had only two that met the threshold. This was done to isolate the cell types that likely cause differences in brain regions related to the ASD variant. The data was then graphed in Excel using bar graphs.

RESULTS

Phase 1 – Gathering Gene Information

In this phase, six published pieces of literature on TSC1, SHANK2, TAOK1, and CHD8 in ASD were reviewed to gain insight on the genes' role and function in ASD and in the cerebellum, hippocampus, and prefrontal cortex.

TSC1

Tuberous Sclerosis 1 (TSC1) is a tumor suppressor gene which encodes a growth inhibitory protein (13). TSC1 is a well-studied gene due to its role in Tuberous Sclerosis Complex (TSC), a condition where up to 60% of individuals also have ASD (13). It is widely used in animal models and listed in major ASD gene databases like SFARI (13).

Mutations of TSC1 and TSC2 causes tuberous sclerosis which is a combination of ASD, intellectual disability, and epilepsy (13). TSC1 and TSC2 mutation causes an upregulation of mTORC1 signaling, resulting in an increased phosphorylation of ribosomal proteins (13). The mutation reduces the synaptic inhibition of pyramidal cells and interneurons in the cerebellum and when the mTORC1 is dysregulated, it results in abnormal protein transformation which leads to synaptic dysfunction in cerebellar inhibitory cells (13).

In the cerebellum, TSC1 deletion causes a decrease in purkinje cell numbers and an increase in soma (cell body) size (13). In the hippocampus it decreases inhibitory synaptic transmission and impairs working memory and fear discrimination (13). The hippocampal pyramidal cells show a decrease in synaptic inhibition and an increase in synaptic excitation because of the TSC1 mutation (13).

Overall, TSC1 is needed to regulate mTORC1, a protein complex which phosphorylates proteins (13). A lack of TSC impairs synaptic inhibition in interneurons (13). The deletion of TSC1 causes an upregulation of mTORC1 signaling in interneurons which is related to the reduction of synaptic inhibition of hippocampal pyramidal cells (13).

SHANK2

SHANK2 is known for its role in synapse development and function. It is widely studied in mouse models and consistently appears in ASD gene databases.

SHANK2 is a scaffolding protein in post-synaptic dendrites (14). This protein helps with the structural and functional coordination of proteins (15). There are 3 SHANK proteins: SHANK1, SHANK2, and SHANK3 (16). All three of these proteins are implicated in ASD; however, SHANK2 and SHANK3 are only implicated in intellectual disability (16). SHANK proteins regulate the excitatory postsynaptic proteins in cerebellar purkinje and granular layers (15). The deletion of SHANK2 in the cerebellar purkinje layer causes dysregulation of excitatory transmission (15). SHANK2 mutations cause anxiety and abnormal social behaviors (15). The human brain stem and thalamus have an increased expression of SHANK2, while the human amygdala has the lowest

expression of SHANK2 (14).

SHANK2A is an isoform of SHANK2 that affects daily activities, anxiety and social behaviors (16). It also helps in plasticity of hippocampal synapses and is involved in extra synaptic function and neuronal network of anterior olfactory nucleus (16). It modulates receptors on a subcellular level (16).

SHANK2 is associated with ASD through an increase in presynaptic expression (14). In addition, the duplication and triplication of the gene is present in the mutation (16). SHANK2 mutations cause decreased excitatory synaptic transmission mediated by a dysregulated NMDA receptor which impairs cognitive ability (15). This mutation also decreased excitatory postsynaptic frequency, density, and increased mismatched excitatory synapses which causes impaired motor coordination (15).

In the cerebellum, the SHANK2 mutation causes increased repetitive behaviors, decreased sociability, altered plasticity, and reduced odor recognition (16). The mutation reduces excitatory current frequency, post-synaptic density adhesion in purkinje cells and reduces excitatory synaptic membrane proteins (15). SHANK2 regulates parallel fiber and purkinje synapse through glutamate D2 protein modulation (15).

In the hippocampus, pre and post synaptic receptors are affected by SHANK2 mutations (16). This mutation also suppresses motor coordination and causes dysregulation of synaptic plasticity in the hippocampus (15). Pre and post synaptic calcium permeable receptors are modulated by SHANK2 in the apical and basal hippocampal dendrites (16).

In the prefrontal cortex, SHANK2 has been localized mostly in post synaptic regions in the soma of the neurons, but not in their nuclei (14).

CHD8

Chromatin Helicase DNA-binding (CHD8) encodes chromatin remodelers that remodel chromatin by methylating DNA (17). CHD family has 8 different isotypes CHD1-8 all of which have same ATPase domain (17). CHD8 phosphorylates the histone tail of DNA and remodels nucleosomes which in turn control gene expression (17).

CHD8 is one of the most well-known genes associated with ASD (17). It's a top-ranked gene in databases like SFARI and is widely used in human and animal model studies to investigate early brain development and gene regulation (17).

Mutations of CHD8 cause ASD, macrocephaly, and facial dysmorphism (17). CHD8 mutation disrupts the modification of DNA and chromatin interaction with transcription factors (17). This mutation causes a decrease in synaptic proliferation in the prefrontal cortex and a decrease in inhibitory signaling in nucleus acumbens. It also increases the brain size in the prefrontal cortex, hippocampus, and cerebellum.

The molecular basis by which this mutation causes ASD could be as a result of decreased pre-synaptic endocytosis and glutamate transmission through endosome mediated synaptic vesicle cycling (17). Neurophysiological studies have shown that the excitatory/inhibitory ratios are affected leading to decreased neuronal plasticity and increased tactile hyper-sensitivity (17).

TAOK1

TAOK1 is a triple helix kinase with 1000 amino acids (18). Thousand and one amino acid kinase 1 (TAOK1) remodels the plasma membrane using phosphoinositides. TAOK1 is necessary for plasma membrane remodeling (18). In its active form, TAOK1 is in the cytoplasm, but when there's a mutation its access to the plasma membrane is blocked by auto phosphorylation (18).

TAOK1 is a less well-known ASD gene, involved in neuronal development and signaling pathways (18). Like the other genes, TAOK1 is included in gene databases, and is gaining attention through recent genetic studies (18).

In ASD mutation of TAOK1 causes auto phosphorylation of threonine 440 and threonine 443 in the triple helix which blocks the plasma membrane access of TAOK1 (18). This impairs the membrane remodeling and causes a disruption of kinase dependent membrane modeling and results in several plasma membrane protrusions in the neurons (18).

TAOK1 mutation causes a decrease in dendritic length which decreases the conduction of the neuronal impulse (18). In the hippocampus, mutations in TAOK1 cause aberrant neuronal membrane protrusions, abnormal growth of dendritic fibers, decreased dendritic length, and aberrant neuronal projections (18).

Overall, TAOK1 inhibits the membrane binding of the triple helix by phosphorylation (18). It induces conformational changes in its helical membrane through autophosphorylation (18).

Phase 2 & Phase 3 – Accessing RNA Sequencing Data & Statistically Analyzing Data

In phase two and three, the average percent expression of the genes in each metacluster in each brain

region was calculated. Then, the highest average percent expressing metaclusters were compiled (See Figure 1). The highest expressing metaclusters in the cerebellum were the pontine and medullary neurons, the cerebellar purkinje neurons, and the deep cerebellar nuclei neurons (See Figure 2). The highest expressing metaclusters in the hippocampus were the hippocampal CA1 neurons, the upper layer excitatory neurons, and the entorhinal, insular, and piriform excitatory neurons (See Figure 3). The highest expressing metaclusters in the prefrontal cortex were the cortical neurons and the isocortical neurons (See Figure 4). In the prefrontal cortex, only two metaclusters met the high-expression threshold. SHANK2 was found to be the highest expressing gene in the highest expressing metacluster (See Figure 5).

DISCUSSION

This study analyzed autism-associated genes and their similarities in expression in the cerebellum, hippocampus, and prefrontal cortex, identifying a brain region with the highest expression of autism-associated genes.

After comparing the average percent gene expressions in the highest expressing metaclusters in the brain regions, it was found that the cerebellar purkinje neurons in the cerebellum had the highest percent expression with 84.24% of the metacluster expressing autism-associated genes. By averaging the percent expression of each gene in the cerebellar purkinje neuron region, it was found that SHANK2 is the highest expressing gene in this region.

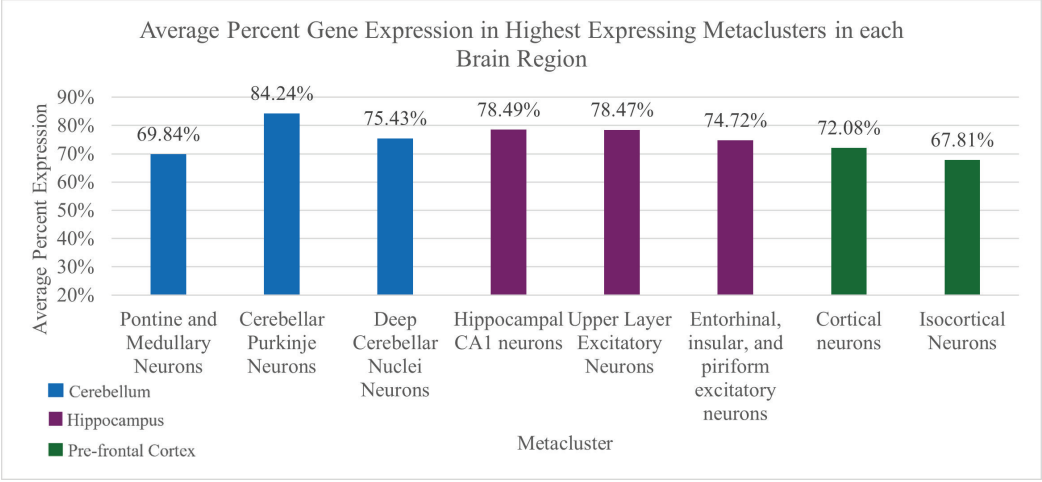

Figure 1. Top three highest expressing metaclusters in each brain region. The cerebellar purkinje neuron metacluster in the cerebellum has the highest average expression compared to the other metaclusters.

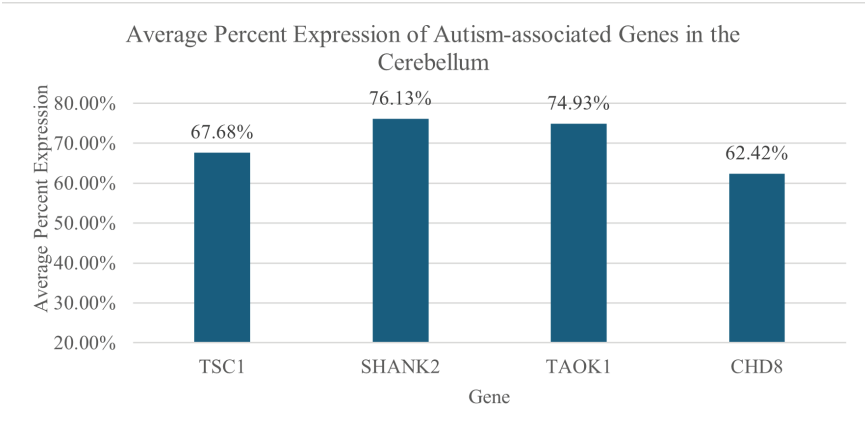

Figure 2. The average percent expression of autism-associated genes in the cerebellum.

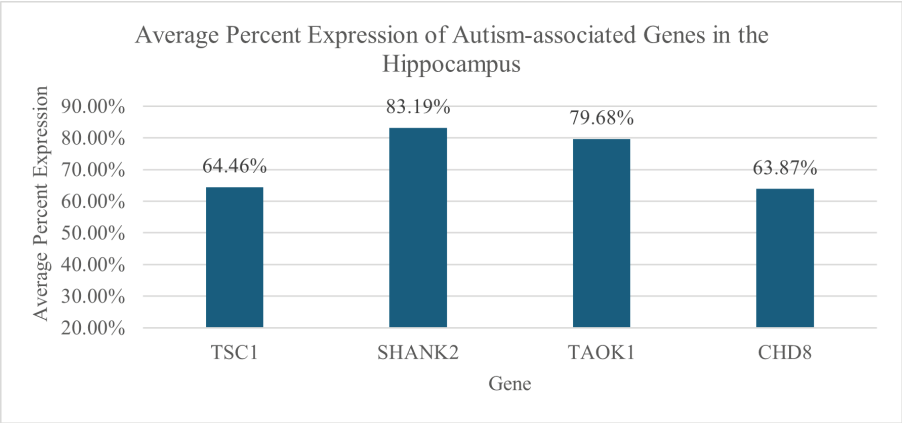

Figure 3. The average percent expression of autism-associated genes in the hippocampus.

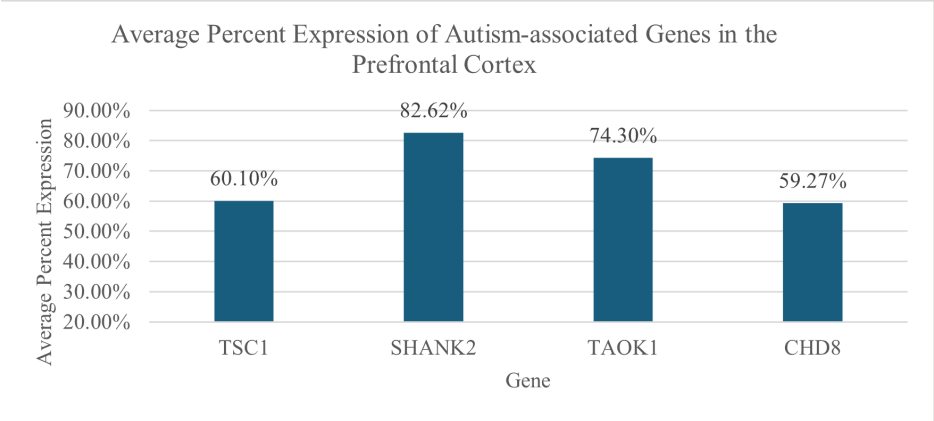

Figure 4. The average percent expression of autism-associated genes in the prefrontal cortex.

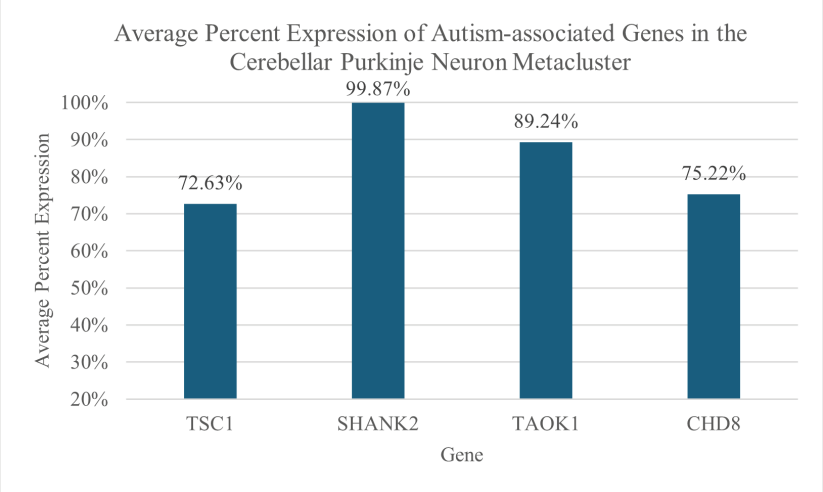

Figure 5. The average percent expression of autism-associated genes in the highest expressing metacluster, the cerebellar purkinje neurons. SHANK2 is the highest expressed gene in the highest expressed metacluster.

This finding aligns with the information collected from the published literature as SHANK2 is mostly only found in purkinje cells. These cells are large neurons in the cerebellum that help coordinate movement by sending signals to other brain regions. The cerebellum is known to be involved in a range of behaviors from motor coordination and balance to social and executive functions. Since purkinje cells are the sole output of the cerebellar cortex, they are critical players in all cerebellum dependent functions and behaviors. Therefore, disruptions in purkinje cell activity can potentially impair autism-associated behavior.

In this study, published literature was reviewed to gain insight into autism-associated genes. However, only a few relevant pieces of literature were identified which limited the amount of information about genes. Thus, using more pieces of literature can allow for a broader understanding of the autism-associated genes and their role in ASD. Moreover, this study only measured the amount of cells expressing in each metacluster. Using more metrics such as the amount of RNA expression per cell and analyzing the subclusters within the cell types could provide more accurate results.

Cerebellar purkinje neurons being the highest expressing region for autism-associated genes suggests that this region is important for future study. ASD is a complex disorder in which researchers are unsure of how autism-associated genes interact with each other to cause ASD. So, identifying the region with the highest expression of autism-associated genes, cerebellar purkinje neurons, can help researchers identify a potential target for future research. This finding suggests that purkinje neurons may be a crucial cell type with altered activity from variance in multiple ASD associated genes.

DECLARATION OF CONFLICT OF INTERESTS

The author declares that there are no conflicts of interest regarding the publication of this article.

ACKNOWLEDGEMENTS

I would like to thank my mentor whose guidance was fundamental to the success of this project. I would also like to extend my gratitude to my friends and family for supporting me throughout this whole study.

REFERENCES

1. Tsai L, Chen R & Wang M. Prevalence and heterogeneity of autism spectrum disorder in children. *Journal of Child Neurodevelopment*. 2012; 5 (2).
2. American Psychiatric Association. *Diagnostic and statistical manual of mental disorders* (5th ed.). 2013. American Psychiatric Publishing. https://doi.org/10.1176/appi.books.9780890425596
3. Jeste S, Geschwing D & State M. Genetic and phenotypic heterogeneity in autism spectrum disorders. *Neuron*. 2014; 82 (1): 1-23. https://doi.org/10.1016/j.neuron.2014.02.010
4. Lasheras I, Rivas P & Martinez-Gonalez M. Gastrointestinal, neurological, and sleep comorbidities in autism spectrum disorder. *European Journal of Paediatric Neurology*. 2020; 24: 14-25. https://doi.org/10.1016/j.ejpn.2020.04.003
5. Genovese S & Butler M. Core diagnostic features of autism spectrum disorder: A molecular and clinical review. *International Journal of Molecular Sciences*. 2020; 21 (13): https://doi.org/10.3390/ijms21134726
6. Cleveland Clinic. (2024, May 14). *Hippocampus: Function, location & damage*. Cleveland Clinic. Retrieved July 20, 2025, from https://my.clevelandclinic.org/health/body/hippocampus
7. Wang S, Kloth A & Badura A. The cerebellum, sensitive periods, and autism. *Neuron*. 20124; 83 (3): 518-532. https://doi.org/10.1016/j.neuron.2014.07.016
8. Hampson D & Blatt G. Autism spectrum disorders and neuropathology of the cerebellum. *Frontiers in Neuroscience*. 2015; 9. https://doi.org/10.3389/fnins.2015.00420
9. Long J, Li H, Liu Y, Liao X, et al. Insights into the structure and function of the hippocampus: Implications for the pathophysiology and treatment of autism spectrum disorder. *Frontiers in Psychiatry*. 2024; 15. https://doi.org/10.3389/fpsyt.2024.1364858
10. NeuroLanch Editorial Team. (2024, August 11). *Prefrontal cortex and autism: Exploring the neural link*. NeuroLanch. Retrieved July 20, 2025, from https://neurolaunch.com/prefrontal-cortex-autism/
11. Bandoim L. (2023, September 13). *Prefrontal cortex: Anatomy, function, and conditions*. Verywell Health. Retrieved July 20, 2025, from https://www.verywellhealth.com/prefrontal-cortex-5220699
12. Macosko E. (2023, December 13). *Brain Cell Data Viewer*. Retrieved August 9, 2025, from https://braincelldata.org/
13. Haji N, Riebe I, Aguilar-Valles A, Artinian J, et al. Tsc1 haploinsufficiency in Nkx2.1 cells upregulates hippocampal interneuron mTORC1 activity, impairs pyramidal cell synaptic inhibition, and alters contextual fear discrimination and spatial working memory in mice. *Molecular autism*. 2020; 11 (1): 29. https://doi.org/10.1186/s13229-020-00340-7
14. Woelfle S, Pedro MT, Wagner J, Schön M & Boeckers TM. Expression profiles of the autism-related SHANK proteins in the human brain. *BMC biology*. 2023; 21 (1): 254. https://

doi.org/10.1186/s12915-023-01712-0

15. Ha S, Lee D, Cho YS, Chung C, et al. Cerebellar Shank2 Regulates Excitatory Synapse Density, Motor Coordination, and Specific Repetitive and Anxiety-Like Behaviors. *The Journal of neuroscience: the official journal of the Society for Neuroscience*. 2016; 36 (48): 12129–12143. https://doi.org/10.1523/JNEUROSCI.1849-16.2016
16. Eltokhi A, Gonzalez-Lozano MA, Oettl LL, *et al*. Imbalanced post- and extrasynaptic SHANK2A functions during development affect social behavior in SHANK2-mediated neuropsychiatric disorders. *Mol Psychiatry*. 2021; 26: 6482–6504. https://doi.org/10.1038/s41380-021-01140-y
17. Hoffmann A & Spengler D. Chromatin Remodeler CHD8 in Autism and Brain Development. *Journal of Clinical Medicine*. 2021; 10 (2): 366. https://doi.org/10.3390/jcm10020366
18. Beeman N, Sapre T, Ong SE & Yadav S. Neurodevelopmental disorder-associated mutations in TAOK1 reveal its function as a plasma membrane remodeling kinase. *Science signaling*. 2023; 16 (766): eadd3269. https://doi.org/10.1126/scisignal.add3269