

## Research Article

# Clustering, Pathway Enrichment, and Protein-Protein Interaction Analysis of Gene Expression in Neurodevelopmental Disorders

Ruchi Yadav  and Prachi Srivastava 

Amity Institute of Biotechnology, Amity University Uttar Pradesh, Lucknow Campus, Lucknow 226028, UP, India

Correspondence should be addressed to Prachi Srivastava; [psrivastava@amity.edu](mailto:psrivastava@amity.edu)

Received 9 August 2018; Accepted 30 October 2018; Published 27 November 2018

Guest Editor: Azhar Rasul

Copyright © 2018 Ruchi Yadav and Prachi Srivastava. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Neuronal developmental disorder is a class of diseases in which there is impairment of the central nervous system and brain function. The brain in its developmental phase undergoes tremendous changes depending upon the stage and environmental factors. Neurodevelopmental disorders include abnormalities associated with cognitive, speech, reading, writing, linguistic, communication, and growth disorders with lifetime effects. Computational methods provide great potential for betterment of research and insight into the molecular mechanism of diseases. In this study, we have used four samples of microarray neuronal developmental data: control, RV (resveratrol), NGF (nerve growth factor), and RV + NGF. By using computational methods, we have identified genes that are expressed in the early stage of neuronal development and also involved in neuronal diseases. We have used MeV application to cluster the raw data using distance metric Pearson correlation coefficient. Finally, 60 genes were selected on the basis of coexpression analysis. Further pathway analysis was done using the Metascape tool, and the biological process was studied using gene ontology database. A total of 13 genes AKT1, BAD, BAX, BCL2, BDNF, CASP3, CASP8, CASP9, MYC, PIK3CD, MAPK1, MAPK10, and CYCS were identified that are common in all clusters. These genes are involved in neuronal developmental disorders and cancers like colorectal cancer, apoptosis, tuberculosis, amyotrophic lateral sclerosis (ALS), neuron death, and prostate cancer pathway. A protein-protein interaction study was done to identify proteins that belong to the same pathway. These genes can be used to design potential inhibitors against neurological disorders at the early stage of neuronal development. The microarray samples discussed in this publication are part of the data deposited in NCBI's Gene Expression Omnibus (Yadav et al., 2018) and are accessible through GEO Series (accession number GSE121261).

## 1. Introduction

**1.1. Neuronal Development Disorder.** Neurogenesis is a process of generating new and functional neurons from neuronal precursors known as NSC (neuronal stem cells) [1, 2]. Functional neurons are generated at the embryonic stage at different stages of development throughout life [3, 4]. With rapid advancement in techniques and curiosity to understand neuronal diseases at the development stage, researchers have explored a wide area of neuronal development diseases and their causes [5–8]. Neuronal stem cells have two major features that are regeneration capacity, that is, ability of self-renewal by process of cell division, and differentiation capacity, that is, process of generating new

and specialized cell types [9]. Developed neurons do not carry dendrites and axons, but they play an important role to receive and send signals to other neurons [10]. Significant development has been made to identify genes that are involved in neuronal diseases at the developmental stage [11]. It is important to study different stages of nervous system development and to identify abnormalities that can arise from improper development of brain at its early stage [12]. Significant contribution has been made by scientists to identify neuronal disorders that occur at the early stage of development [13]. Neuronal disorders include abnormalities associated with intellectual disability, attention deficit hyperactivity disorder (ADHD), and cognitive skills disorders, like dyslexia and dysgraphia, and language development

disorders like expression disorder [14–18]. Scientific evidence shows that neurological disorders can be identified at the early stage by the first week or month of a lifecycle [19–21]. It is important to identify which genes are crucial and result in neurological disorders.

We have used high-throughput microarray experiment to identify genes that are involved in the early stage of neurodevelopment. Our aim was to identify genes that were expressed when stem cells were exposed to MCP (monocrotophos), a neurotoxicant, and to evaluate the effective role of resveratrol (RV) and nerve growth factor (NGF) as the neuroprotectant.

*1.2. Resveratrol Clinical Perspectives.* Resveratrol is a natural phenol and phytoalexin produced naturally by several plants in response to injury [22]. There is exponential evidence since 1939 in the literature that resveratrol is a promising natural compound for prevention and treatment of a wide range of human diseases [23]. Resveratrol is also reported to be effective against neuronal cell dysfunction and cell death, Huntington's disease and Alzheimer's disease [24–27]. Molecular studies show that resveratrol is associated with an induction of genes for oxidative phosphorylation and mitochondrial biogenesis [28]. Effect of resveratrol is known to extend lifespan, and it impacts mitochondrial function and metabolic homeostasis [29]. In the current work, we have mapped effectiveness of resveratrol against injured neurodevelopment samples. In this study, four samples were prepared (control, resveratrol, NGF, and RV + NGF). Datasets of prepared samples were taken to investigate the neuroprotective role of resveratrol against exposure of monocrotophos. In silico expression analysis of different datasets is done to identify genes that are coexpressed.

*1.3. Microarray Data Analysis.* Microarrays provide a rich source of data on the molecular mechanism of cell function. Each microarray reports expression of thousands of mRNAs [30]. Virtually, almost every human disease is being studied using microarrays experiment, with the aim of finding the novel genes involved in diseases and disease markers and to identify drug targets [31]. Bioinformatics analysis plays an important part of processing the information, embedded in large-scale expression profiling studies, and for laying the biological interpretation of high throughput microarray data [32]. A basic yet challenging task in the analysis of microarray gene expression data is the identification of changes in gene expression that are associated with particular biological conditions [33, 34]. Careful statistical design and analysis are essential to identify genes involved in each biological condition.

A standard workflow is required to utilize computational tools at various steps of microarray analysis. This paper also describes use of different bioinformatics tools for quality control, normalization, coexpression, annotation, and pathway and protein-protein interaction analysis.

*1.4. Clustering and Coexpression Analysis.* Clustering is a method to identify genes that are coexpressed in each biological condition [35]. Clustering methods uses a distance

measure (e.g., Euclidean metric) to compare expression values of pairs of genes for each experiment [36]. When the distance between a pair of genes is small, then the two genes might be clustered. Clusters are analyzed to identify genes that are coexpressed and coregulated.

*1.5. Biological Annotation and Interpretation.* After extensive analysis of microarray data, one needs to annotate Affymetrix IDs for its significance. Annotation reveals the biological significance of genes like its molecular pathway, diseases involved, gene ontology, and so on [37]. Careful exploration is required to identify genes that are expressed in each condition of microarray experiment. Pathway and process enrichment is a crucial part of annotation, as it leads to the identification of set of genes that are involved in the same pathways [38]. Pathways analysis also highlights the set of proteins that interact with each other; this information is used to categorize protein interaction partners and to study protein-protein interaction network [39].

## 2. Materials and Methods

*2.1. Microarray Data.* The MSCs (mesenchymal stem cells) were used to study effect of monocrotophos (MCP) and repairing capability of resveratrol and nerve growth factor. MSCs were exposed to RV, NGF, and RV + NGF, respectively. In total, four samples were generated to identify genes that were coexpressed at the neuronal development stage. Affymetrix gene chip platform (Prime view.CDF) was used to identify gene expression using four samples as described in Table 1.

*2.2. Microarray Data Analysis and Annotation.* Computational software and tools were used to identify genes that are coexpressed. Figure 1 shows the workflow used for microarray data analysis and annotation. Raw files were used, i.e., chip electronic file (CEL) and chip description file (CDF) for quality control analysis. R and Bioconductor, Affy package, was used for data normalization and data transformation. Gene expression matrix was generated from Affy package, using RMA (robust multiarray average).

Significant analysis of microarray (SAM) [40] and clustering were done using MeV application [41]. The clustering method was used to cluster significant genes obtained from the SAM method. For clustering, the distance metric Pearson correlation coefficient was used, using parameter of k-means algorithm, number of cluster 10, and number of iteration 50. Coexpressed genes were identified by analyzing each ten clusters.

*2.3. Pathway Enrichment and Protein-Protein Interaction Analysis.* Coexpressed genes identified from clustering analysis were further annotated for biological intervention and pathway analysis. The list of coexpressed genes was searched against pathway and GO database using the Metascape tool (<http://metascape.org>) [42]. Each gene was studied for its pathway and process enrichment score for

TABLE 1: Four samples that were used for microarray gene expression analysis.

S.no	Samples	Description	Raw files
1	Control	Control sample	Control.CEL
2	RV	MSCs exposed to RV	RV.CEL
3	NGF	MSCs exposed to NGF	NGF.CEL
4	RV + NGF	MSCs exposed to combined RV + NGF	RV + NGF.CEL

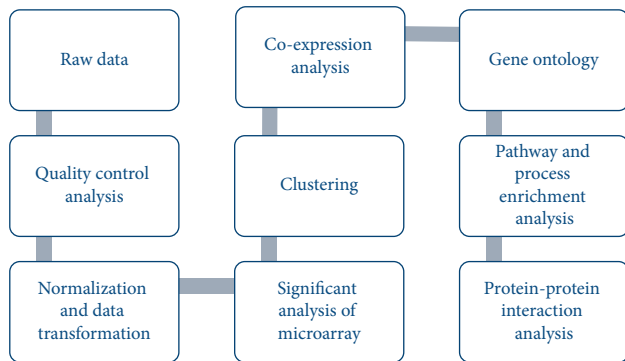


FIGURE 1: Workflow used for microarray data analysis and annotation.

statistical significance of genes in each biological process. Genes were also clustered according to their pathways. Relationship between genes was also identified by using the network map generated from the Metascape tool and visualized in the Cytoscape tool [43].

Protein-protein interaction analysis is carried out by different protein interaction databases like BioGrid, InWeb\_IM, and OmniPath by using the Metascape tool. Molecular complex detection (MCODE) algorithm was used by the Metascape tool to identify densely connected network of protein-protein interaction [44].

### 3. Results

**3.1. Quality Control Analysis.** Quality control (QC) is an important aspect of examining microarray data, before any statistical analysis to be done. QC analysis was done using Affy package of R and Bioconductor [45]. Figure 2 shows the chip image of four samples: (a) control sample, (b) NGF sample, (c) RV sample, and (d) RV + NGF sample; it signifies there is no error in chips of all four samples, and they can be used for data exploration and analysis.

**3.2. Data Normalization and Transformation.** Finding biologically relevant answer from microarray experiment is a primary need of any microarray experiment. Variations in gene expression should be biologically not from any source of errors like biasness in dyes, lasers, samples, and chip spotting during microarray experiment [46]. To analyze microarray data, one needs to remove these biasness and errors in microarray experiment. Normalization is a method to remove these systematic errors that affect gene expression measures [47]. After QC analysis, normalization was done

using Affy package of R and Bioconductor. We have used the RMA method of normalization. Figure 3 shows the box plot of four samples after data normalization. Box plot shows statistical values like mean or median and variations between samples [48]. Figure 3 shows that means of all four samples are at position. Data were transformed to logarithm base 2 value of the expression ratio, and expression matrix was written, for further statistical analysis and comparisons.

**3.3. Clustering and Coexpression Analysis.** Significant analysis of microarray (SAM) was done to identify the number of genes that were statistically significant. Out of 49,495 genes, 49,022 genes were found to be insignificant and 473 genes were statistically significant. K-means clustering was done on significant genes with parameter of  $k = 10$ . Ten clusters were generated and studied for coexpression of genes. Figure 4 shows cluster 1 (only one cluster is shown but all ten clusters were studied for coexpression analysis) which shows coexpression of Tp53 and B-cell cll/lymphoma2 gene. In addition, Caspase-8, Caspase-10, and dopamine receptor are also coregulated.

Analysis of all ten clusters results in identification of coexpressed genes. Rigorous analysis of clustering shows that 60 genes were coexpressed (AKT1, BAD, BAX, BCL2, BDNF, CASP3, CASP8, CASP9, MYC, PIK3CD, MAPK1, MAPK10, and CYCS). These genes were used for gene ontology, biological function, and pathway analysis. Descriptions including function of these 60 genes were shown in pathway and process enrichment analysis.

**3.4. Pathway and Process Enrichment Analysis.** Coexpressed genes that were clustered in the clustering step were used for biological annotation and interpretation. The Metascape tool was used to study pathway and process of these 60 genes. Protein-protein interaction network was constructed to identify more proteins that have similar function and belong to same pathway. 60 genes were further clustered into 20 groups on the basis of their enrichment score (enrichment score is the score between observed count and expected count by chance) [49].

In each cluster, one term represents the cluster that is most statistically significant [42]. Figure 5 shows the heat map of enriched terms colored by the  $p$  value. Pathway enrichment analysis shows that most of the genes were involved in colorectal cancer, neurotrophin signaling pathway, neuron death, and thyroid hormone signaling pathway. Other clusters indicate genes that were involved in cellular response to organonitrogen compounds, response to nicotine, and head development. Genes that belong to these clusters were further studied in detail for function and pathway analysis.

Top 5 clusters are shown in Table 2; count is the number of genes in each cluster; percentage is the total gene ontology provided in list of genes;  $\text{Log}_{10}(P)$  is the log base 10 value; and  $\text{Log}_{10}(q)$  is the log base 10 adjusted  $p$  value [42].

Pathway enrichment shows that neuronal development genes are involved in colorectal cancer, neuron death, and other diseases like leukemia and sclerosis [50]. Genes AKT1, BAD, BAX, BCL2, CASP3, CASP8, CASP9, MYC, PIK3CD,

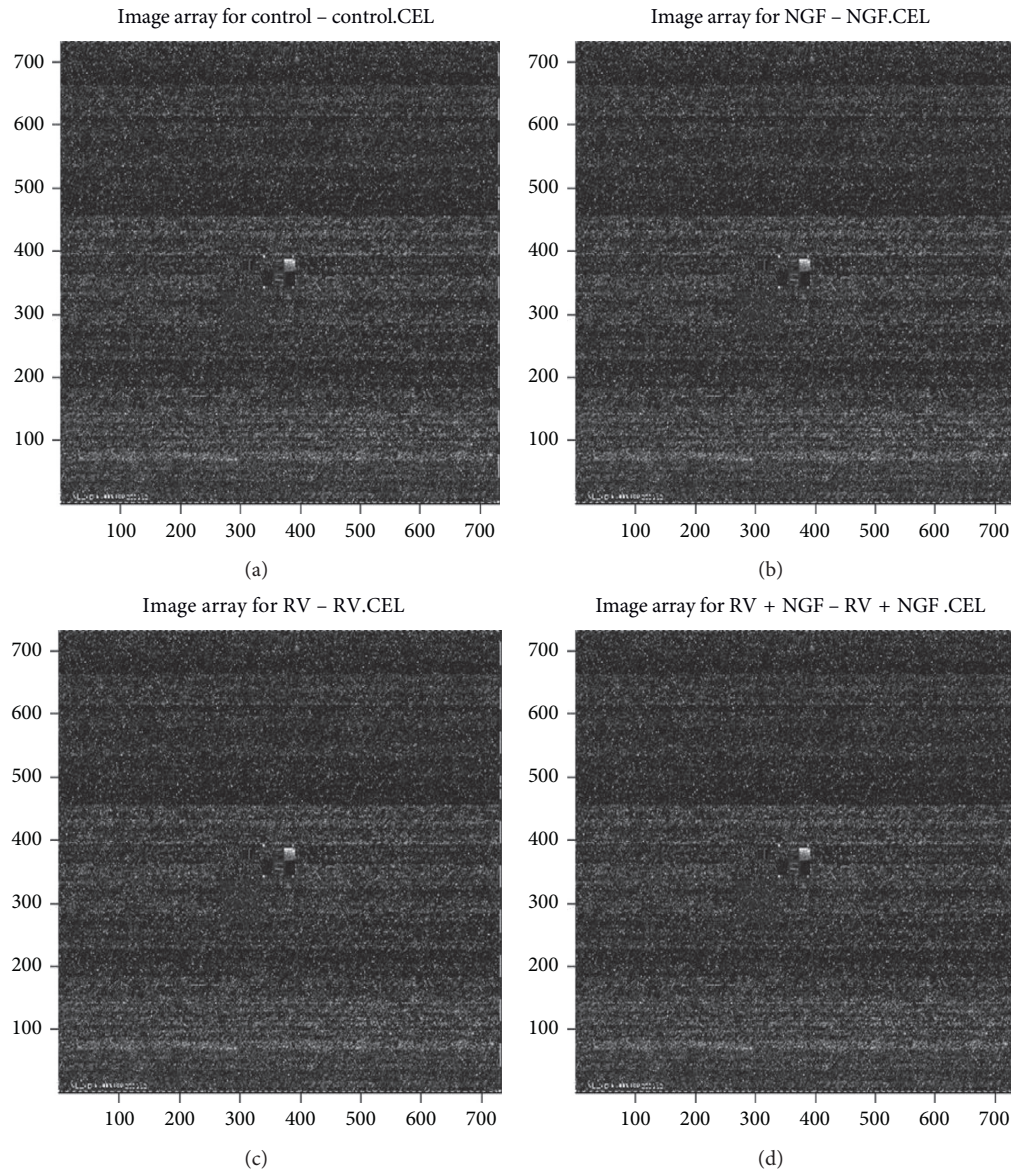


FIGURE 2: Affymetrix chip image: (a) control sample, (b) NGF sample, (c) RV sample, and (d) RV + NGF sample.

MAPK1, MAPK10, and CYCS are commonly expressed in the cluster of colorectal cancer, neuronal signaling pathway, neuronal death, amyotrophic lateral sclerosis, and tuberculosis [51]. Further proteins are identified that show interaction with these proteins on the basis of protein-protein interaction study.

**3.5. Protein-Protein Interaction Enrichment Analysis.** Protein-protein interaction (PPI) enrichment was done among the list of genes that were clustered in pathway and process enrichment analysis. The Metascape tool predicts PPI network by comparing it with protein interaction databases (BioGrid, InWeb\_IM, and OmniPath) [42]. PPI is made between proteins having physical interactions, and PPI network is further subclustered on the basis of the  $p$  value score. Figure 6 shows the PPI map between the set of input genes. Three best scoring genes by the  $p$  value

are identified; these proteins define the functionality of PPI network. Best scoring genes belong to apoptosis (hsa04210) [52], colorectal cancer (hsa05210) [53], and hepatitis B (hsa05161) [54]. PPI network represents involvement of neuronal development genes in diseases like cancer.

The molecular complex detection (MCODE) method was applied to identify closely related protein from PPI network. The MCODE algorithm subclustered PPI network into 3 subclusters. Figure 7 shows MCODE components (red, blue, and green as MCODE 1, 2, and 3). Three dense PPI were made and detail of each cluster is given in Table 3. MCODE prediction validates the results of clustering as previously shown in Figure 4. The same set of proteins was identified by MCODE algorithm as predicted by clustering using the MeV tool. These proteins have the same GO and pathway.

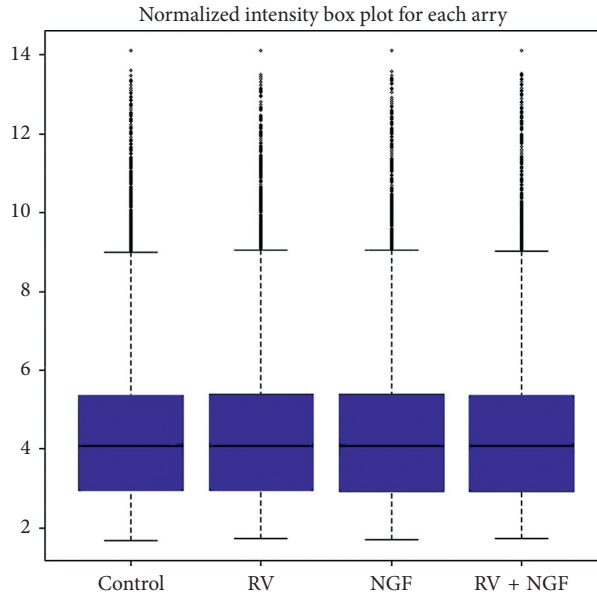


FIGURE 3: Box plot of microarray samples after normalization.

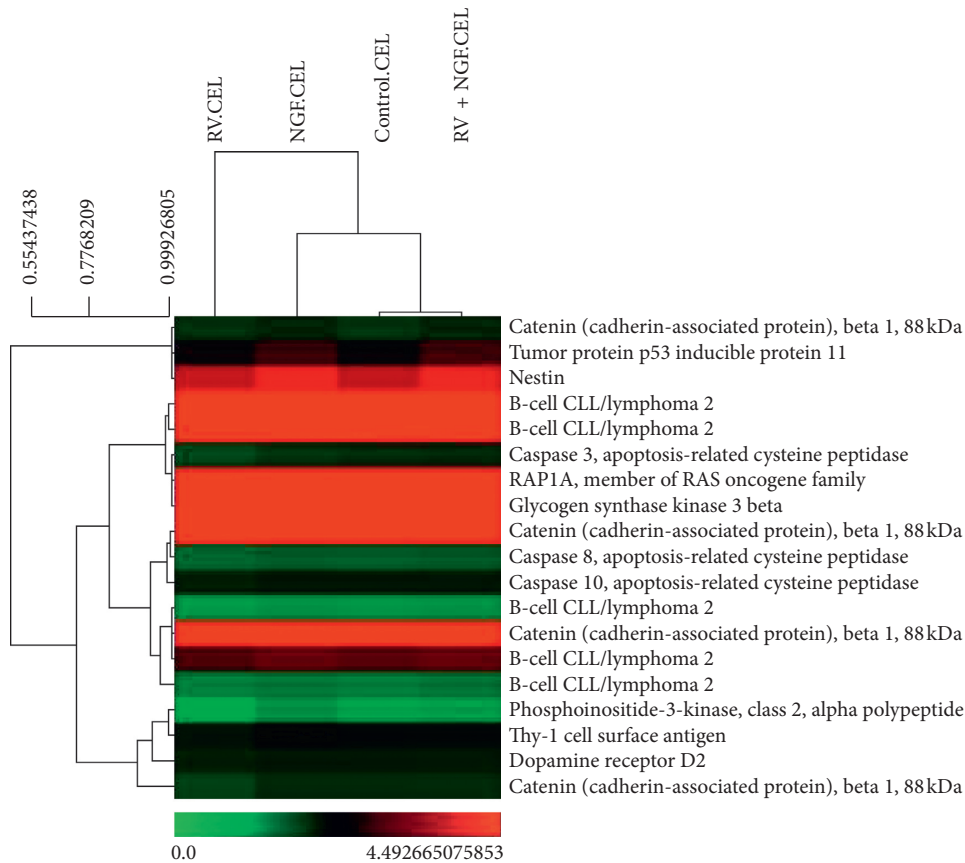


FIGURE 4: Clustering result of significant genes.

Cluster analysis of MCODE components is done, and details of proteins involved in each cluster and their corresponding pathways are shown in Table 3. Cluster 1 includes proteins CASP3, CASP9, BAX, TP53, BAD, GSK3B, POU5F1, MAPK14, CREB1, SOX2, and KLF4. Gene ontology data show

that these proteins are associated with amyotrophic lateral sclerosis (hsa05014) [55], colorectal cancer (hsa05210) [56], and positive regulation of neuron death (GO: 1901216) [57].

Cluster 2 genes are mentioned in Table 3. GO analysis shows that these proteins belong to thyroid hormone

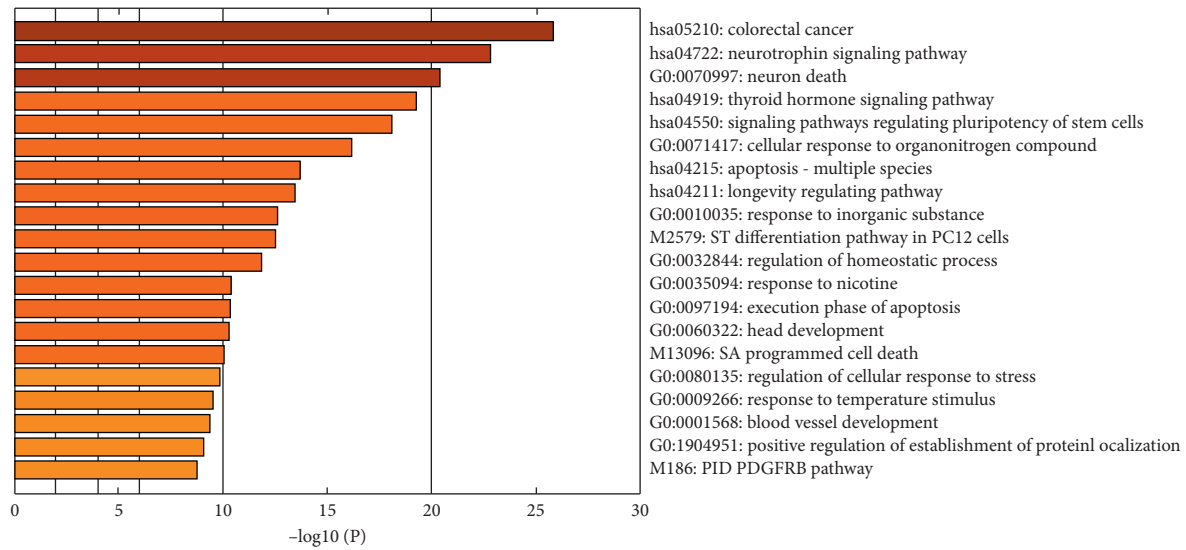


FIGURE 5: Heat map of enriched terms across input gene lists, colored by  $p$  values.

TABLE 2: Pathway and process enrichment analysis.

S.no	GO	Category	Description	Count	%	Log10 (P)	Log10 (q)
1	Hsa05210	KEGG pathway	Colorectal cancer	15	25	-25.7	-21.4
2	Hsa04722	KEGG pathway	Neurotrophin signaling pathway	16	26.6	-22.8	-19.1
3	Go: 0070997	Go biological processes	Neuron death	19	31.6	-20.3	-16.9
4	Hsa04919	KEGG pathway	Thyroid hormone signaling pathway	14	23.3	-19.2	-15.9
5	Hsa04550	KEGG pathway	Signaling pathways regulating pluripotency of stem cells	14	23.3	-18.0	-14.9

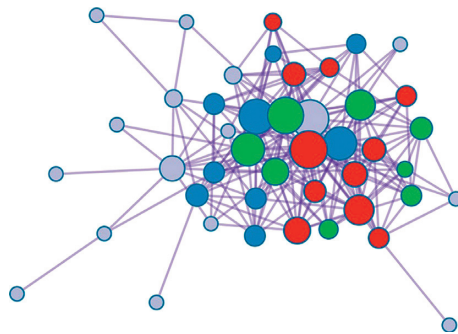


FIGURE 6: Protein-protein interaction network.

signaling pathway (hsa04919) [55], diseases of signal transduction (R-HSA-5663202) [56], and pathways in cancer (hsa05200) [57]. Cluster 3 proteins are involved in pathways of toxoplasmosis (hsa05145) [55], tuberculosis (hsa05152) [56], and fluid shear stress and atherosclerosis (hsa05418) [55].

MCODE cluster and cluster of MeV software show that some genes were commonly expressed and were coregulated. AKT1, BAD, BAX, BCL2, BDNF, CASP3, CASP8, CASP9, MYC, PIK3CD, MAPK1, MAPK10, and CYCS genes are coregulated. PPI analysis identifies other proteins that have interaction with abovementioned proteins. These proteins are important in neuronal differentiation, and regeneration proteins like ACTB, GSK3B, CREB1, and

CTNNB1 have physical interaction with coexpressed proteins [58]. Table 3 also gives the information about proteins and association with diseases. Analysis of disease associated with proteins highlights that some proteins belong to different classes of cancers. 12 proteins (CASP3, CASP9, BAX, TP53, BAD, GSK3B, MTOR, BCL2L11, SIRT1, CASP8, AKT1, and CTNNB1 proteins) are involved in diverse types of cancers like lung cancer, breast cancer, ovarian cancer, colorectal cancer, and leukemia [59].

While other proteins (GSK3B, POU5F1, MAPK14, CREB1, SOX2, KLF4, PRKACA, MAPK10, STAT1, ACTB, TUBB3, MYC, GAPDH, AKT1, and CTNNB1) are related with process of aging, neuronal diseases, cardiovascular diseases, abnormal brain development, mental retardation,

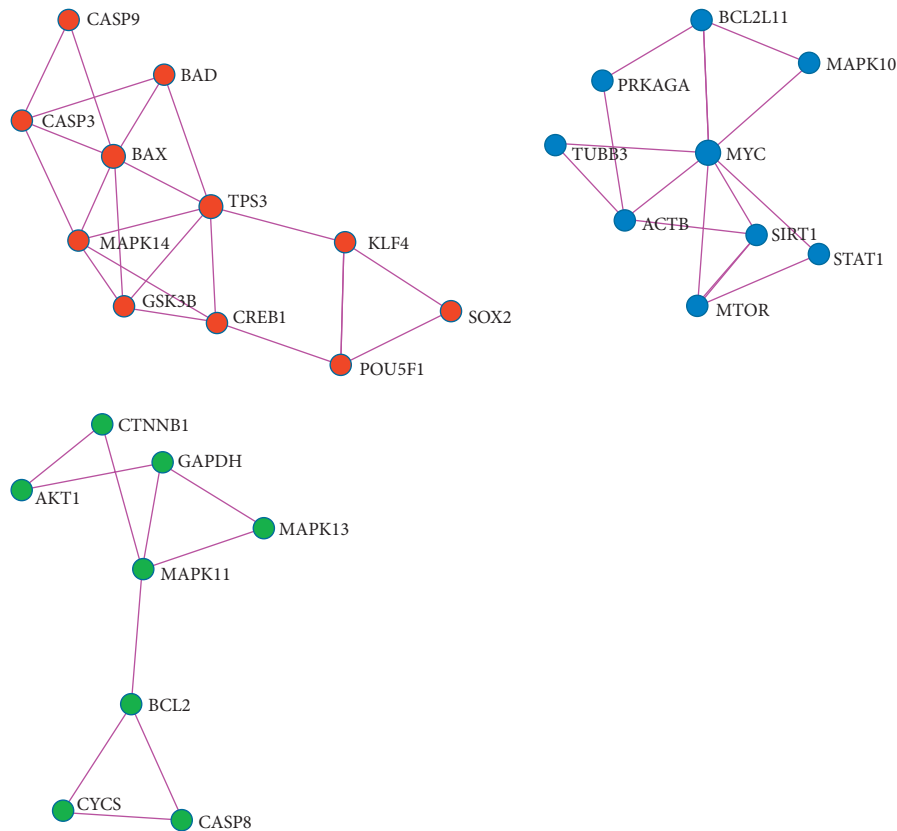


FIGURE 7: MCODE protein-protein interaction network. Colors show the different components of MCODE (red color: MCODE 1, blue color: MCODE2, and green color: MCODE 3).

schizophrenia, and mycobacterial and viral infections [60–62].

Key findings of the pathway and disease association study are identification of proteins involved in neurological diseases and also expressed at the early stage of neuronal development. SOX2 protein was expressed in optic nerve hypoplasia and abnormalities of the central nervous system [63], STAT1 was expressed during mycobacterial and viral infections [64], TUBB3 was related with fibrosis and cortical dysplasia and brain deformities, AKT1 was expressed in breast cancer, colorectal cancer, ovarian cancer, and schizophrenia [65] and CTNNB1 was expressed in colorectal cancer, hepatocellular carcinoma, ovarian cancer, and mental retardation [66]. The study shows that proteins (SOX2, STAT1, AKT1, and CTNNB1) can be used as markers for neurological disease at the early stage of neuronal development, and they can be potential drug targets for therapeutic development.

#### 4. Conclusion and Discussion

Microarray experiment is designed to investigate the genes that are expressed at the early stage of neuronal development. Neurodevelopmental microarray gene expression data are used to identify genes that are expressed in neuronal disorders, at its initial stage of progress [67]. Four samples were prepared, viz, control, resveratrol, nerve growth factor, and RV + NGF and hybridized to Affymetrix chip (Prime

view). Gene expression matrix was constructed, and computational analysis was done. The protocol is designed to study biologically significant genes. Microarray data analysis workflow includes quality control, data normalization, clustering, pathways enrichment, and PPI study. Clustering analysis identifies genes that are coexpressed. These sets of coexpressed genes are used for pathway and process enrichment analysis. Gene ontology and pathway study reveal proteins that share common pathways and function. Further protein-protein interaction network is constructed to identify more number of proteins, which have physical interaction with coexpressed proteins. PPI network is subclustered to predict closely related proteins. Gene ontology information of these proteins is used to identify function and disease associated with proteins. 12 proteins CASP3, CASP9, BAX, TP53, BAD, GSK3B, MTOR, BCL2L11, SIRT1, CASP8, AKT1, and CTNNB1 proteins are predicted that are involved in various types of cancers like lung cancer, breast cancer, ovarian cancer, colorectal cancer, and leukemia [60, 61, 62]. Some proteins like SOX2, STAT1, AKT1, and CTNNB1 proteins are associated with neurological disease like abnormal brain development, mental retardation, schizophrenia, and mycobacterial and viral infections [63–66]. These genes can be used as markers for neurological disease, for detection of abnormalities at the early stage of neuronal development [67]. Predicted proteins can also act as potential drug targets for the drug development process. Further work is required for wet lab

TABLE 3: Cluster details of MCODE protein-protein interaction and pathways.

Cluster	Symbol	Pathways	Diseases
1	CASP3	Caspase-3	Cancer
	CASP9	Caspase-9	Aging, cancer, chemdependency
	BAX	BCL2-associated X and apoptosis regulator	Colorectal cancer, somatic, T-cell acute lymphoblastic leukemia
	TP53	Tumor protein p53	Breast cancer, colorectal cancer, hepatocellular carcinoma, etc.
	BAD	BCL2-associated agonist of cell death	Cancer, hematological, infection
	GSK3B	Glycogen synthase kinase 3 beta	Cancer
	POU5F1	POU class 5 homeobox 1	Chemdependency
	MAPK14	Mitogen-activated protein kinase 14	Chemdependency
	CREB1	cAMP responsive element-binding protein 1	Histiocytoma
	SOX2	SRY-box 2	Optic nerve hypoplasia and abnormalities of the central nervous system
KLF4	Kruppel like factor 4	Cardiovascular and metabolic diseases	
2	MTOR	Mechanistic target of rapamycin	Aging, cancer, chemdependency
	PRKACA	Protein kinase cAMP-activated catalytic subunit alpha	Cushing syndrome
	MAPK10	Mitogen-activated protein kinase 10	Metabolic diseases
	STAT1	Signal transducer and activator of transcription 1	Mycobacterial and viral infections
	ACTB	Actin beta	Baraitser-winter syndrome 1 and dystonia
	BCL2L11	BCL2 like 11	Cancer, cardiovascular diseases
	SIRT1	Sirtuin 1	Aging, cancer, cardiovascular diseases
	TUBB3	Tubulin beta 3 class III	Fibrosis, cortical dysplasia, brain malformations
	MYC	v-myc avian myelocytomatosis viral oncogene homolog	Burkitt lymphoma
	GAPDH	Glyceraldehyde-3-phosphate dehydrogenase	Aging
3	CASP8	Caspase-8	Breast cancer, lung cancer
	AKT1	AKT serine/threonine kinase 1	Breast cancer, somatic, colorectal cancer, ovarian cancer, schizophrenia
	BCL2	BCL2, apoptosis regulator	Leukemia/lymphoma
	CYCS	Cytochrome c, somatic	Thrombocytopenia
	MAPK11	Mitogen-activated protein kinase 11	Unknown
	MAPK13	Mitogen-activated protein kinase 13	Unknown
	CTNNB1	Catenin beta 1	Colorectal cancer, hepatocellular carcinoma, ovarian cancer and mental retardation

verification of predicted genes that are expressed in neurological disorders and express at the developmental stage. More research is required in the field of neurodevelopmental biology to identify neurological abnormalities at its budding stage. This paper also highlights the importance of microarray experiment in understanding the neurological diseases and methodology to study various outcomes of gene expression data, like coexpression analysis, pathway and process identification, and protein-protein interaction network study.

### Data Availability

The microarray data used to support the findings of this study are included within the supplementary information file.

### Conflicts of Interest

The authors declare that they have no conflicts of interest.

### Supplementary Materials

Gene expression matrix microarray raw files (CEL and CDF files) were used to make gene expression matrix using Affy package and RMA (robust multiarray average) method. This

matrix file was used for further microarray analysis like clustering, pathway, and protein-protein interaction analysis. Pathway and protein-protein interaction result as generated from the Metascape tool is given. This file includes information about clustered formed within genes and network details with scores. Annotation and enrichment annotation file and pathway enrichment result as generated from the Metascape tool is given. Annotation file includes information about genes like gene symbol, gene description, GO biological process, protein function, and so on. (*Supplementary Materials*)

### References

- [1] P. S. Eriksson, E. Perfilieva, T. B. Eriksson et al., "Neurogenesis in the adult human hippocampus," *Nature Medicine*, vol. 4, no. 11, p. 1313, 1998.
- [2] H. Van Praag, G. Kempermann, and F. H. Gage, "Running increases cell proliferation and neurogenesis in the adult mouse dentate gyrus," *Nature Neuroscience*, vol. 2, no. 3, p. 266, 1999.
- [3] H. Van Praag, A. F. Schinder, B. R. Christie, N. Toni, T. D. Palmer, and F. H. Gage, "Functional neurogenesis in the adult hippocampus," *Nature*, vol. 415, no. 6875, p. 1030, 2002.



- [4] J. Wegiel, I. Kuchna, K. Nowicki et al., "The neuropathology of autism: defects of neurogenesis and neuronal migration, and dysplastic changes," *Acta Neuropathologica*, vol. 119, no. 6, pp. 755–770, 2010.
- [5] M. V. Johnston, M. E. Blue, and N. Sakkubai, "Rett syndrome and neuronal development," *Journal of Child Neurology*, vol. 20, no. 8, pp. 759–763, 2005.
- [6] A. M. Galaburda, J. LoTurco, F. Ramus, R. H. Fitch, and G. D. Rosen, "From genes to behavior in developmental dyslexia," *Nature Neuroscience*, vol. 9, no. 10, pp. 1213–1217, 2006.
- [7] D. A. Lewis and P. Levitt, "Schizophrenia as a disorder of neurodevelopment," *Annual Review of Neuroscience*, vol. 25, no. 1, pp. 409–432, 2002.
- [8] B. D. Trapp and K.-A. Nave, "Multiple sclerosis: an immune or neurodegenerative disorder?," *Annual Review of Neuroscience*, vol. 31, pp. 247–269, 2008.
- [9] G. Muñoz-Eliás, W. Dale, and I. B. Black, "Marrow stromal cells, mitosis, and neuronal differentiation: stem cell and precursor functions," *Stem Cells*, vol. 21, no. 4, pp. 437–448, 2003.
- [10] M. Schuldiner, R. Eiges, A. Eden et al., "Induced neuronal differentiation of human embryonic stem cells," *Brain Research*, vol. 913, no. 2, pp. 201–205, 2001.
- [11] D. P. Purpura, "Normal and aberrant neuronal development in the cerebral cortex of human fetus and young infant," in *Brain Mechanisms in Mental Retardation*, pp. 141–169, Elsevier Inc., Amsterdam, Netherlands, 1975.
- [12] A. E. West and M. E. Greenberg, "Neuronal activity-regulated gene transcription in synapse development and cognitive function," *Cold Spring Harbor Perspectives in Biology*, vol. 3, no. 6, article a005744, 2011.
- [13] M. P. Whitaker-Azmitia, "Serotonin and brain development: role in human developmental diseases," *Brain Research Bulletin*, vol. 56, no. 5, pp. 479–485, 2001.
- [14] T. L. Bale, T. Z. Baram, A. S. Brown et al., "Early life programming and neurodevelopmental disorders," *Biological Psychiatry*, vol. 68, no. 4, pp. 314–319, 2010.
- [15] H. Y. Zoghbi, "Postnatal neurodevelopmental disorders: meeting at the synapse?," *Science*, vol. 302, no. 5646, pp. 826–830, 2003.
- [16] M. Dennis, D. J. Francis, P. T. Cirino, R. Schachar, M. A. Barnes, and J. M. Fletcher, "Why IQ is not a covariate in cognitive studies of neurodevelopmental disorders," *Journal of the International Neuropsychological Society*, vol. 15, no. 3, pp. 331–343, 2009.
- [17] P. Krakowiak, C. K. Walker, A. A. Bremer et al., "Maternal metabolic conditions and risk for autism and other neurodevelopmental disorders," *Pediatrics*, vol. 129, no. 5, pp. e1121–e1128, 2012.
- [18] P. Levitt, K. L. Eagleson, and E. M. Powell, "Regulation of neocortical interneuron development and the implications for neurodevelopmental disorders," *Trends in Neurosciences*, vol. 27, no. 7, pp. 400–406, 2004.
- [19] I. C. Gillberg and C. Gillberg, "Children with preschool minor neurodevelopmental disorders. IV: behaviour and school achievement at age 13," *Developmental Medicine and Child Neurology*, vol. 31, no. 1, pp. 3–13, 1989.
- [20] M. J. Millan, "An epigenetic framework for neurodevelopmental disorders: from pathogenesis to potential therapy," *Neuropharmacology*, vol. 68, pp. 2–82, 2013.
- [21] J. A. Baur and D. A. Sinclair, "Therapeutic potential of resveratrol: the in vivo evidence," *Nature Reviews Drug Discovery*, vol. 5, no. 6, pp. 493, 2006.
- [22] L. Frémont, "Biological effects of resveratrol," *Life Sciences*, vol. 66, no. 8, pp. 663–673, 2000.
- [23] M. Lagouge, C. Argmann, Z. Gerhart-Hines et al., "Resveratrol improves mitochondrial function and protects against metabolic disease by activating SIRT1 and PGC-1 $\alpha$ ," *Cell*, vol. 127, no. 6, pp. 1109–1122, 2006.
- [24] M. H. Aziz, S. Reagan-Shaw, J. Wu, B. J. Longley, and N. Ahmad, "Chemoprevention of skin cancer by grape constituent resveratrol: relevance to human disease?," *FASEB Journal*, vol. 19, no. 9, pp. 1193–1195, 2005.
- [25] K. Magyar, R. Halmosi, A. Palfi et al., "Cardioprotection by resveratrol: a human clinical trial in patients with stable coronary artery disease," *Clinical Hemorheology and Microcirculation*, vol. 50, no. 3, pp. 179–187, 2012.
- [26] V. Vingtdeux, U. Dreses-Werringloer, H. Zhao, P. Davies, and P. Marambaud, "Therapeutic potential of resveratrol in Alzheimer's disease," *BMC Neuroscience*, vol. 9, no. 2, p. S6, 2008.
- [27] E. N. Frankel, A. L. Waterhouse, and J. E. Kinsella, "Inhibition of human LDL oxidation by resveratrol," *The Lancet*, vol. 341, no. 8852, pp. 1103–1104, 1993.
- [28] G. López-Lluch, P. M. Irusta, P. Navas, and R. de Cabo, "Mitochondrial biogenesis and healthy aging," *Experimental Gerontology*, vol. 43, no. 9, pp. 813–819, 2008.
- [29] A. Biala, E. Tauriainen, A. Siltanen et al., "Resveratrol induces mitochondrial biogenesis and ameliorates Ang II-induced cardiac remodeling in transgenic rats harboring human renin and angiotensinogen genes," *Blood Pressure*, vol. 19, no. 3, pp. 196–205, 2010.
- [30] de M. X. Renée, J. M. Boer, and H. C. van Houwelingen, "Microarray data analysis," *Applied bioinformatics*, vol. 3, no. 4, pp. 229–235, 2004.
- [31] A. T. Weeraratna and D. D. Taub, *Microarray Data Analysis*, Humana Press, New York City, NY, USA, 2007.
- [32] P. Behzadi, E. Behzadi, and R. Ranjbar, "Microarray data analysis," *Challenge*, vol. 7, p. 8, 2014.
- [33] G. K. Smyth, Y. H. Yang, and T. Speed, "Statistical issues in cDNA microarray data analysis," *Functional Genomics*, pp. 111–136, Humana Press, New York City, NY, USA, 2003.
- [34] T. Jirapech-Umpai and S. Aitken, "Feature selection and classification for microarray data analysis: evolutionary methods for identifying predictive genes," *BMC bioinformatics*, vol. 6, no. 1, p. 148, 2005.
- [35] G. Getz, E. Levine, and E. Domany, "Coupled two-way clustering analysis of gene microarray data," *Proceedings of the National Academy of Sciences*, vol. 97, no. 22, pp. 12079–12084, 2000.
- [36] R. K. Curtis, M. Orešič, and A. Vidal-Puig, "Pathways to the analysis of microarray data," *TRENDS in Biotechnology*, vol. 23, no. 8, pp. 429–435, 2005.
- [37] T. Werner, "Bioinformatics applications for pathway analysis of microarray data," *Current Opinion in Biotechnology*, vol. 19, no. 1, pp. 50–54, 2008.
- [38] G. Wu, X. Feng, and L. Stein, "A human functional protein interaction network and its application to cancer data analysis," *Genome Biology*, vol. 11, no. 5, p. R53, 2010.
- [39] S. Dudoit, R. C. Gentleman, and J. Quackenbush, "Open source software for the analysis of microarray data," *Biotechniques*, vol. 34, no. 13, pp. 45–51, 2003.
- [40] E. Howe, K. Holton, S. Nair, D. Schlauch, R. Sinha, and J. Quackenbush, "Mev: Multiexperiment Viewer," *Biomedical Informatics for Cancer Research*, pp. 267–277, Springer, Boston, MA, USA, 2010.
- [41] S. Tripathi, M. O. Pohl, Y. Zhou et al., "Meta-and orthogonal integration of influenza OMICs data defines a role for UBR4

- in virus budding,” *Cell Host and Microbe*, vol. 18, no. 6, pp. 723–735, 2015.
- [42] L. Shuaichen and G. Wang, “Bioinformatic analysis reveals CYP2C9 as a potential prognostic marker for HCC and liver cancer cell lines suitable for its mechanism study,” *Cellular and Molecular Biology*, vol. 64, no. 7, pp. 70–74, 2018.
- [43] A. Kuno, K. Nishimura, and S. Takahashi, “Time-course transcriptome analysis of human cellular reprogramming from multiple cell types reveals the drastic change occurs between the mid phase and the late phase,” *BMC Genomics*, vol. 19, no. 1, p. 9, 2018.
- [44] C. L. Wilson and C. J. Miller, “Simpleaffy: a BioConductor package for Affymetrix quality control and data analysis,” *Bioinformatics*, vol. 21, no. 18, pp. 3683–3685, 2005.
- [45] R. A. Irizarry, “Exploration, normalization, and summaries of high density oligonucleotide array probe level data,” *Bio-statistics*, vol. 4, no. 2, pp. 249–264, 2003.
- [46] D. P. Kreil, N. A. Karp, and K. S. Lilley, “DNA microarray normalization methods can remove bias from differential protein expression analysis of 2D difference gel electrophoresis results,” *Bioinformatics*, vol. 20, no. 13, pp. 2026–2034, 2004.
- [47] Y. H. Yang, “Normalization for cDNA microarray data: a robust composite method addressing single and multiple slide systematic variation,” *Nucleic Acids Research*, vol. 30, no. 4, p. e15, 2002.
- [48] Da W. Huang, B. T. Sherman, and R. A. Lempicki, “Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists,” *Nucleic Acids Research*, vol. 37, no. 1, pp. 1–13, 2008.
- [49] S. E. Soden, C. J. Saunders, L. K. Willig et al., “Effectiveness of exome and genome sequencing guided by acuity of illness for diagnosis of neurodevelopmental disorders,” *Science Translational Medicine*, vol. 6, no. 265, article 265ra168, 2014.
- [50] S. Wang, Z. Z. Chong, Y. C. Shang, and K. Maiese, “Wnt1 inducible signaling pathway protein 1 (WISP1) blocks neurodegeneration through phosphoinositide 3 kinase/Akt1 and apoptotic mitochondrial signaling involving Bad, Bax, Bim, and Bcl-xL,” *Current Neurovascular Research*, vol. 9, no. 1, pp. 20–31, 2012.
- [51] Y. S. Lee, S. G. Hwang, J. K. Kim et al., “Identification of novel therapeutic target genes in acquired lapatinib-resistant breast cancer by integrative meta-analysis,” *Tumor Biology*, vol. 37, no. 2, pp. 2285–2297, 2016.
- [52] A. L. Tarca, S. Draghici, G. Bhatti, and R. Romero, “Down-weighting overlapping genes improves gene set analysis,” *BMC Bioinformatics*, vol. 13, no. 1, p. 136, 2012.
- [53] S.-L. Chen, Z.-M. Wang, Z.-Y. Hu, and B. Li, “Genome-wide analysis of differentially expressed long noncoding RNAs induced by low shear stress in human umbilical vein endothelial cells,” *Integrative Molecular Medicine*, vol. 2, pp. 276–289, 2015.
- [54] Z.-P. Liu, Y. Wang, X.-S. Zhang, and L. Chen, “Identifying dysfunctional crosstalk of pathways in various regions of Alzheimer’s disease brains,” *BMC Systems Biology*, vol. 4, no. 2, 2010.
- [55] R. Wang, J. Wei, Z. Li, Y. Tian, and C. Du, “Bioinformatical analysis of gene expression signatures of different glioma subtypes,” *Oncology Letters*, vol. 15, no. 3, pp. 2807–2814, 2018.
- [56] H.-Q. Wang, X.-P. Xie, and C.-H. Zheng, “A pathway-based classification method that can improve microarray-based colorectal cancer diagnosis,” in *Proceedings of International Conference on Intelligent Computing*, August 2011.
- [57] J. K. Choudhari, B. P. Sahariah, J. K. Choubey, A. Patel, and M. K. Verma, “Identification of potential transcription factor and protein kinases for regulation of differentially expressed genes for fluoride exposure in human using Expression2Kinases (X2K) approach,” *Network Modeling Analysis in Health Informatics and Bioinformatics*, vol. 6, no. 1, p. 7, 2017.
- [58] V. Uversky, I. Na, K. Landau, and R. Schenck, “Highly disordered proteins in prostate cancer,” *Current Protein and Peptide Science*, vol. 18, no. 5, pp. 453–481, 2017.
- [59] E. Weinstein, X. Cui, and P. Simmons, *Genomic Editing of Neurodevelopmental Genes in Animals*, U.S. Patent Application No. 12/842.
- [60] D. Nosome, “Identification of genetic risk factors for cerebellar mutism in pediatric brain tumor patients,” Dissertation, The University of Texas School of Public Health, Houston, TX, USA, 2012.
- [61] C. R. Sullivan, C. Mielnik, S. M. O’Donovan et al., *Connectivity Analyses of Bioenergetic Changes in Schizophrenia: Identification of Novel Treatments*, bioRxiv, 2018.
- [62] A. L. M. Ferri, “Sox2 deficiency causes neurodegeneration and impaired neurogenesis in the adult mouse brain,” *Development*, vol. 131, no. 15, pp. 3805–3819, 2004.
- [63] T. Nishibori, Y. Tanabe, L. Su, and M. David, “Impaired development of CD4+CD25+ regulatory T cells in the absence of STAT1: increased susceptibility to autoimmune disease,” *Journal of Experimental Medicine*, vol. 199, no. 1, pp. 25–34, 2004.
- [64] G. Xiromerisiou, G. M. Hadjigeorgiou, A. Papadimitriou, E. Katsarogiannis, V. Gourbali, and A. B. Singleton, “Association between AKT1 gene and Parkinson’s disease: a protective haplotype,” *Neuroscience Letters*, vol. 436, no. 2, pp. 232–234, 2008.
- [65] T. Morikawa, “Association of CTNNB1 ( $\beta$ -catenin) alterations, body mass index, and physical activity with survival in patients with colorectal cancer,” *JAMA*, vol. 305, no. 16, pp. 1685–1694, 2011.
- [66] N. N. Parikhshak, M. J. Gandal, and D. H. Geschwind, “Systems biology and gene networks in neurodevelopmental and neurodegenerative disorders,” *Nature Reviews Genetics*, vol. 16, no. 8, p. 441, 2015.
- [67] P. B. Crino, J. Q. Trojanowski, M. A. Dichter, and J. Eberwine, “Embryonic neuronal markers in tuberous sclerosis: single-cell molecular pathology,” *Proceedings of the National Academy of Sciences*, vol. 93, no. 24, pp. 14152–14157, 1996.