Online Available at: https://www.sciplatform.com/index.php/wjb/article/view/1161

ISSN (Online) = 2522-6754 ISSN (Print) = 2522-6746

WORLD JOURNAL OF BIOLOGY AND BIOTECHNOLOGY

Research Manuscript www.sciplatform.com

Integrative bioinformatics analysis of transcriptomic data sheds light on the molecular complexity of parkinson's disease and potential therapeutic targets

Qudsia Yousafi*, Sawaira Asghar, Misbah Shoukat

[a] **Department of BioSciences, COMSATS University Islamabad, Sahiwal Campus, Pakistan.**

Authors' Contribution | **Yousafi, Q.,** conceptualized the idea, designed methodology interpreted results and wrote original manuscript, **S. Asghar & M. Shoukat** executed the methodology and result generation.

***Corresponding Author's Email Address**: qudsia@cuisahiwal.edu.pk

Review Process: Peer review

ABSTRACT

Parkinson's disease (PD) is a progressive neurodegenerative disorder that affects millions of individuals worldwide, with an increasing prevalence in aging populations. In this study, we retrieved the mRNA expression dataset, GSE165082, for PD through GEOmnibus. Total of 220 downregulated and 354 upregulated genes were identified after data normalization. Functional annotation carried out by DAVID tools, revealed that these DEGs were mainly enriched in biological processes i.e., cell division and protein phosphorylation, and they were localized mostly cytoplasm and nucleus. Two molecular function protein binding and ATP binding were predominant. Additionally KEGG pathway analysis highlighted their involvement in neurodegenerative, cancer, alzheimer's and coronavirus diseases. Armadillo-type fold and Armadillo-like helical domains were found by INTERPRO while TKc domain by SMART. Transcription factors IRF1 was predicted by FunRich tool. Upregulated genes were found expressed in 6 sites i.e., Palate, Ventral striatum, Pluripotent stem cells, Ganglia, Curtilage and Ciliary muscle. A protein-protein interaction network was constructed by using Cytoscape v 6.0. Ten hub genes EFI3A, RPL28, SMG8. UPF2, XAF1, IFITM1, IFIT3, LY63, IFI3 and LY6B were identified by Cytohubba. The expression patterns of hub genes across different organs and immune response cells using a heatmap and expression of EIF3 was found in almost all organs except liver. MicroRNA for were predicted by FunRich tool. Finally, we predicted microRNAs for RPL28, SMG8, UPF2 and EIF3a that could potentially regulate these hub genes, providing insights into post-transcriptional gene regulation. This comprehensive analysis contributes to our understanding of the molecular mechanisms underlying Parkinson's disease and provides a foundation for future research and therapeutic development in this complex and challenging condition.

Keywords: miRNA, gene ontology, hub genes, heat map, IRF1.

INTRODUCTION: Parkinson's disease is a neurodegenerative disorder that primarily affects movement and is characterized by a range of motor and non-motor symptoms (De Lau and Breteler, 2006). It is named after Dr. James Parkinson, who first described the condition in his groundbreaking essay, "An Essay on the Shaking Palsy," published in 1817 (Parkinson, 1817). This landmark publication laid the foundation for our understanding of the disease. Parkinson's disease is estimated to affect millions of individuals worldwide, with an increasing prevalence in aging populations. The condition results from the progressive degeneration of dopaminergic neurons in a specific region of the brain called the substantia nigra (Gao et al., 2002). This neuronal loss leads to a shortage of a neurotransmitter crucial for regulating movement i.e., dopamine. The cardinal motor symptoms of Parkinson's disease include resting tremors, bradykinesia (slowness of movement), rigidity, and postural instability. Additionally, non-motor symptoms such as depression, anxiety, cognitive impairment and autonomic dysfunction often coexist and can significantly impact a patient's quality of life (Jankovic, 2008). Several treatment options are available to manage its symptoms and improve the patient's overall well-being (Armstrong and Okun, 2020). These treatments include medication, physical therapy, and in some cases, surgical interventions like deep brain stimulation. Ongoing research efforts continue to explore the underlying mechanisms of Parkinson's disease and seek innovative therapies that may one day slow or halt its progression. Parkinson's disease remains a significant medical challenge, highlighting the importance of continued research and support for affected individuals and their families. With the advancement of genomic technologies, The investigation of gene expression patterns to understand the molecular causes of diseases and finding disease-specific biomarkers has grown popularly (Can, 2014). Differential gene expression analysis is a highly effective tool for analyzing the molecular processes underpinning genome regulation and identifying quantitative differences in expression levels between experimental groups and control groups (San Segundo-Val and Sanz-Lozano, 2016). These variations in gene expression may help identify potential biomarkers for a particular disease. Gene ontology and enrichment analysis are useful technique for comprehending gene function and gene connection from genome-wide expression (Langfelder and Horvath, 2008). These can be used to find relevant modules linked to clinical features and co-expression modules of highly correlated genes (Zhang and Horvath, 2005). providing great insight into predicting the functions of co-expression genes and finding genes that play key roles in human diseases (Li et al., 2018; Yang et al., 2014). In this study, differential co-expression genes were identified using expression the mRNA expression data of Parkinson's disease GEO databases. We have identified important genes involved in disease prognosis by performing functional enrichment, protein-protein interaction (PPI) analysis. These genes can be used as therapeutic targets for disease treatment therapies. The global Parkinson's disease (PD) population was estimated to be 9.4 million in 2020, surpassing the previously reported figure of 6 million cases in 2016 (Dorsey et al., 2018; Yang et al., 2020). This increasing trend in PD prevalence depicts the urgent requirement for initiatives to address this issue and design effective therapies to manage this complex challenging disease.

OBJECTIVES: The objectives of this study were as follows: (1) to determine the functionality and localization of DEGs (2) Identification of hub genes. (3) Interaction network analysis of hub genes and other genes in human being. (4) Identification miRNA for the hub genes regulation.

MATERIAL AND METHODS: Data retrieval, normalization and statistical analysis: Dataset was retrieved from Gene Expression Omnibus (GEO) database by using GSE165082 accession query. The data set was taken from an experiment of DNA methylation and expression profiles of whole blood in Parkinson's disease done by expression profiling by high throughput screening (Henderson et al., 2021). Total number of samples in dataset was 26 (14 were healthy/control and 12 were unhealthy/diseased). GSE165082_PD-CC.counts.txt.gz ftp file was downloaded for analysis.

The mRNA expression data was statistically analyzed using "BiocManager" Packages of R. The normalization and statistical analyses of data was done by installing "DeSeq2" library of "BiocManager" in R. These normalization techniques help us to handle numerical variables of varying units and scales, thus improving the quality of data. The purpose of normalization is to remove systematic variation in a microarray experiment which affects the measured gene expression levels (Park et al., 2003). The output expression file was saved in .csv format. The Gene Ensemble IDs in the output file were converted to Gene IDs by submitting to Ensemble BioMart. Differentially expressed genes (DEGs) were identified on the basis of Log fold Change (LogFC2) value. The threshold for LogFC2 was set as >1 and p value< 0.05.

Functional annotation analysis of DEGs: For Gene Ontology Database for Annotation, Visualization and Integrated Discovery (DAVID) was used in the present study. It provides a comprehensive set of functional annotation tools for investigators to understand biological meaning behind large list of genes (Dennis et al., 2003). The list of DEGs was submitted to DAVID for functional annotation analysis.

Protein domain and transcription factor prediction: The enrichment of DEGs in protein domains was predicted by two databases SMART and INTERPRO. The results were selected on the basis of threshold p-value. Transcription factor (TF) (or sequence-specific DNA-binding factor) is a protein that controls the rate of

transcription of genetic information from DNA to messenger RNA, by binding to a specific DNA sequence (Latchman, 1997). The transcription factors for DEGs were predicted by FunRich Gene Enrichment module.

Site of Expression of DEGs: The genes are either upregulated or downregulated in their expression. The sites for upregulated and downregulated gene expression were investigated by using FunRich gene expression module.

Protein-Protein Network and Identification of Hub genes: The list of DEGs imported to network construction and visualization tool Cytoscape v.2.3 (Cline et al., 2007). Protein - Protein Interaction of gene submitted list was constructed by using inbuilt plugin STRING. The node colors were assigned as per their expressions by using option from style menu. The preferred string type layout was selected for network display. The hub genes were identified from the main network by using Cytohubba tool. This tool identifies the hub genes based on Nodes score. Twelve different statistical approaches are used for hub genes score calculation & identification by this tool. The list of hub genes for each method was downloaded & saved. Consensus genes from more than 5 methods out of 12 were selected as hub genes of the Parkinson's disease datasets GSE165082.

Protein-protein Network construction, heat map plotting and micro RNA (miRNA) prediction for hub gene: The list of selected genes was imported to Cytoscape and plugin GeneMANIA was selected for interaction network of query hub genes with other human (*Homo sapiens*) genes.

A common method of visualizing gene expression data is to display it as a heatmap. In heat maps the data is displayed in a grid form where each row represents a gene and each column represents a sample. The color and intensity of the boxes is used to represent changes (not absolute values) of gene expression (Haarman et al., 2015). The heat map for hub genes was plotted by using FunRch heatmap generation module.

MicroRNAs (miRNAs) are small non-coding nucleic acids. These are involved in regulating post-transcriptional gene expression by binding to complementary sequences of target messenger RNA (mRNA) (Lukiw and Alexandrov, 2012). The miRNA for the hub genes were predicted by FunRich miRNA prediction module.

RESULTS: The normalized dataset, GSE165082, consisted of 220 downregulated and 354 upregulated genes, was used for further analysis

Functional annotation analysis of DEGs: The highest number of genes (20) were found to be enriched in two biological processes *i.e.*, cell division and protein phosphorylation (figure 1a). The DEGs were found to be localized in seventeen cellular location with highest number 166 in cytoplasm followed by 163 genes were found in nucleus (figure 1b). Twelve molecular functions were predicted for query genes (figure 1c). Less than twenty genes were found to be involved in each of ten molecular functions. Highest number of genes (271) were found for protein binding followed by those for ATP binding (57 genes). Similarly, KEGG pathway enrichment analysis showed that the DEGs mainly participated in neurodegenerative disease (24), cancer (23), alzheimer disease (20) and corona virus disease (16) pathways (figure 1d).

Protein domain and transcription factor prediction: The results from two domain databases SMART and INTERPRO were selected based on threshold p-value≤0.05. INTERPRO predicted 14 domains with highest gene count (22) for Armadillo-type fold domain and 15 genes were found to be involved in Armadillo-like helical domain (figure 2a). Only 3 domains were predicted by SMART with highest gene count (17) for S-TKc domain (figure 2b). The Transcription factor for whole data set was IRF1 representing 50% of genes and lowest p-value 0.098 (figure 3). In case of hub genes although IRF1represent small percentage of genes (9.47%) but lowest p-value 0.047 make it potential transcription factor with 94.9% confidence. Other transcription factors were SP4, KLF7, SP1, NRF1 and GABPA.

Site of Expression of DEGs: In current study the genes were found downregulated in six expression sites with -100 Fold Change (FC) value *i.e.*, Peyer's patches, Synovial cell, Transverse mesocolon, Lung epithelium, Articular cartilage and Subthalamus (figure 4). These were found to upregulated in six sites i.e., Palate, Ventral striatum, Pluripotent stem cells (FC 19.4) Ganglia, Curtilage and Ciliary muscle (FC 38.7).

Protein-Protein Interaction Network and Identification of Hub Genes: Main network containing 572 genes (nodes) was constructed (figure 5) by using Cytoscape. Red nodes represent upregulated genes while the green ones were for downregulated. The Interaction between the nodes shows by the edges. Ten hub genes were identified by each of 12 methods used by CytoHubba (table 1).

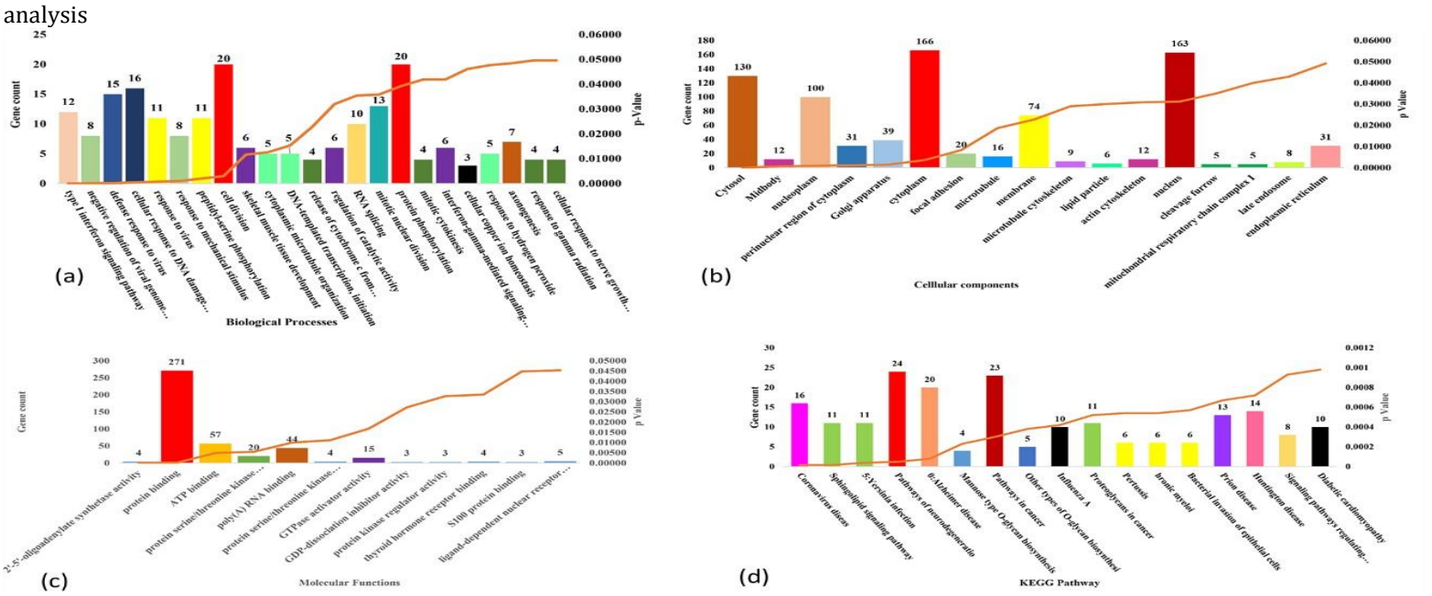

Figure 1: Number of genes of Parkinson's disease (GSE165082) involved in different Gene Ontology events. (a) Biological processes (b) cellular components (c) molecular functions (d) KEGG pathways.

	Statistical methods used for mapping hub genes													
NO	BETWeenness	Botteleneck	Closeness	ClusCoefficit	Degree	DMNC	EcCentricity	EPC	MCC	MNC	Radiality	Stress	Frequency	Consensus Hub genes
1	RPL28	RPL28	IFIT3	SMG8	IFIT3	XAF1	IFIT3	IFITM1	ISG15	ISG15	RPL28	RPL28	10	RPL28
2	IFIT3	IFIT3	IFI6	LY6E	IFI6	IFI35	IFI6	IFI6	RSAD2	SMC3	SMG8	IFIT3	11	IFIT3
3	IFI6	IFI6	XAF1	IFI35	XAF1	IFI6	XAF1	LY6E	OAS2	RPS5	UPF2	IFI6	11	IFI6
4	SMG8	SMG8	IFITM1	IFITM1	IFITM1	IFITM1	IFITM1	XAF1	OAS1	RPS24	EIF3A	SMG8	11	SMG8
5	UPF2	UPF2	IFI35	IFIT3	IFI35	IFIT3	IFI35	IFIT3	IFIT3	RSAD2	IFIT3	UPF2	9	UPF2
6	XAF1	XAF1	LY6E	XAF1	LY6E	EIF3A	LY6E	IFI35	XAF1	PTPRC	IFI6	XAF1	11	XAF1
7	IFITM1	IFITM1	RPL28	UPF2	RPL28	UPF2	RPL28	RPL28	IFI35	RAD21	XAF1	IFITM1	11	IFITM1
8	EIF3A	EIF3A	SMG8	IFI6	SMG8	SMG8	SMG8	UPF2	IFI6	RPL26	IFITM1	EIF3A	10	EIF3A
9	IFI35	IFI35	UPF2	RPL28	UPF2	LY6E	UPF2	SMG8	IFITM1	USP18	IFI35	IFI35	11	IFI35
10	LY6E	LY6E	EIF3A	EIF3A	EIF3A	RPL28	EIF3A	EIF3A	USP18	SOCS1	LY6E	LY6E	10	LY6E

Table 1: Hub genes of Parkinson's disease dataset GSE165082 predicted by Cytohubba.

 Digital Object Identifier: https://doi.org/10.33865/wjb.009.01.1161

The top ten genes appeared in more than 5 methods were selected as consensus hub genes data set GSE-165082.

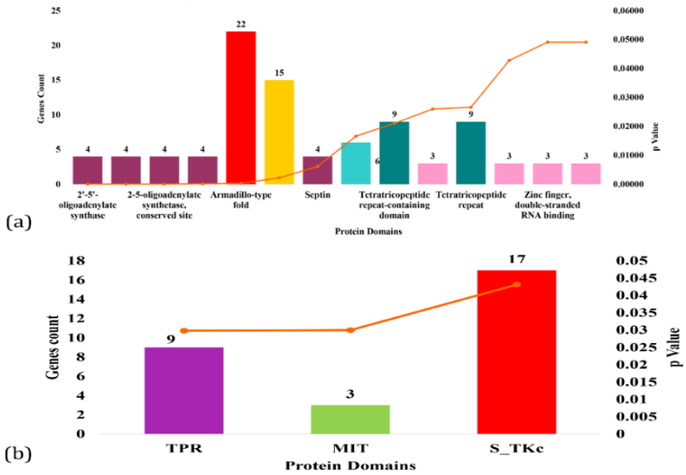

Figure 2: The number of genes of Parkinson disease (GSE165082) involved in different protein domains predicted by (a) INTERPRO (b) SMART.

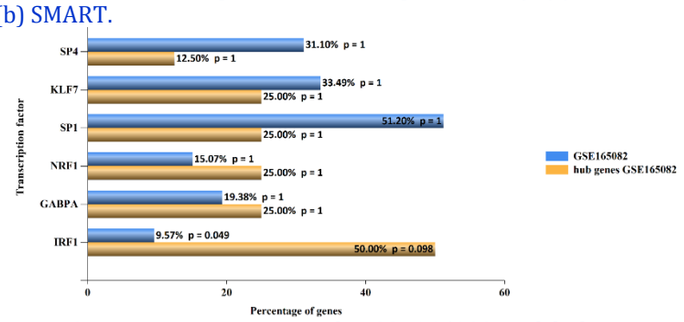

Figure 3: Transcription Factor for GSE165082 and hub genes of GSE165082.

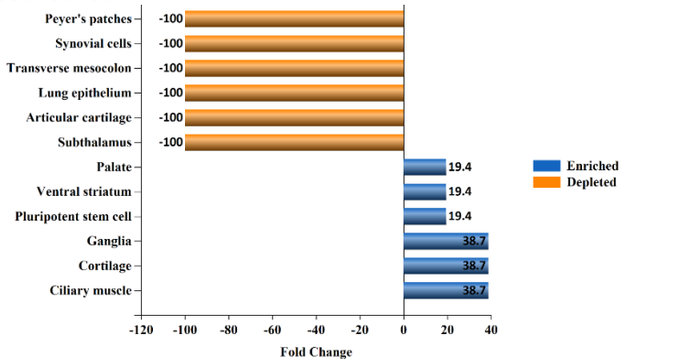

Figure 4: : Site of expression for genes in data set GSE165082 for Parkinson disease.

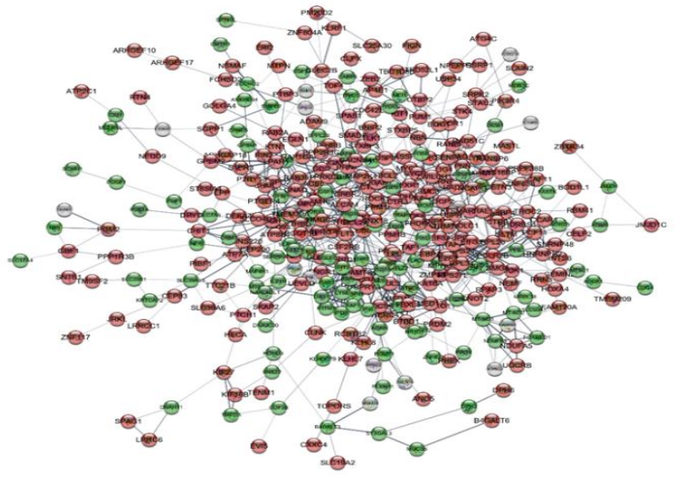

Figure 5: Protein-protein Interaction network of genes involved in Parkinson's disease (GSE165082). *Red nodes= down regulated genes, Green node= upregulated genes.

Protein-protein network construction, heat map plotting and micro RNA (miRNA) prediction for hub gene: The hub genes showed interaction with each other & nineteen other genes (figure 6a). The Black color nodes (EFI3A, RPL28, SMG8. UPF2, XAF1, IFITM1, IFIT3,LY63, IFI35) indicates the query hub gene while the grey color nodes (BST2, RSAD2, ADAR, STAT2, ISG20, RSAD2, RNASEL, IFITM2, GBP2, IRF9, ISG15, IFIT2, IFITM3, IFI27, IF144L, IP6K2, EGR1, MX1, QAS1, QAS2) represent interacting other genes.

The heat map shows the expression of hub genes in different adult and fetal organs and in immune response cell (figure 6b).

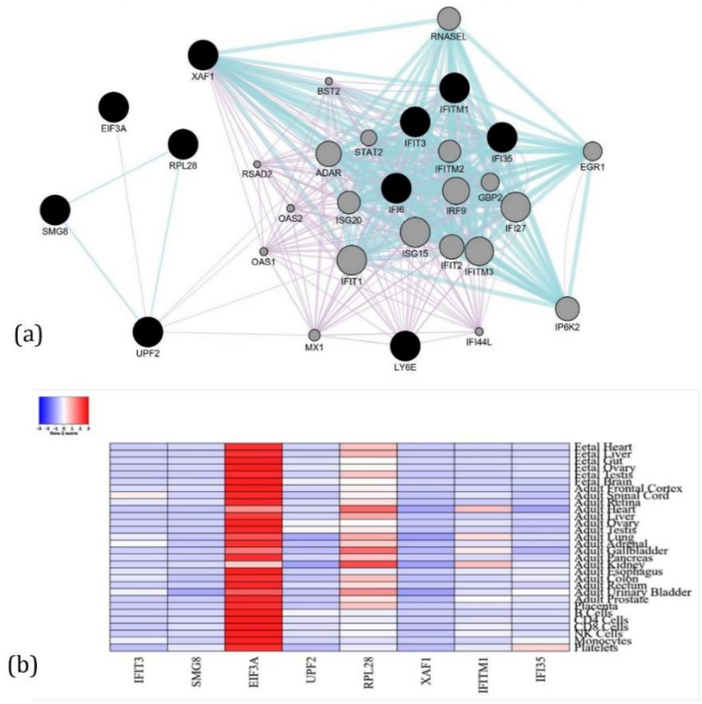

Figure 6: Presentation of interactions and enrichment of hub genes of data set GSE165082. (a) Protein-protein network with other genes in human (Black nodes=query genes, Grey nodes= interacting human genes). (b) Heat map for expression of hub genes.

The expression of a gene in the cell is represented by its *z*-score (legend). The value of *z*-score is directly proportional to its color. Only one gene, EIF3A, showed sharpest red color for almost all the categories except for adult kidney. This shows high expression of this gene in almost all the mentioned localization. MicroRNA (miRNA) have been predicted in the case of four hub genes, RPL28, SMG8, UPF2 and EIF3a (table 2). Eight miRNAs were predicted against UPF2 while only three for SMG8.

DISCUSSIONS: Differentially expressed genes (DEGs) from RNAseq dataset of Parkinson's disease were obtained from GEOomnibus database. We used the freely accessible, Database Visualization and Integrated Discovery (DAVID, http://david.niaid.nih.gov) for gene annotation analysis (Huang et al., 2009). These biological annotation of genes was done in terms of biological processes, cellular localization and molecular functions (Garcia-Moreno et al., 2022). Most of the genes were found be involved in protein phosphorylation and cell division. It has been reported that abnormal protein phosphorylation plays a key role in the development and progression of Alzheimer disease (AD). According to a research, changes in the pattern of protein phosphorylation of various brain regions may enhances AD progress from a pre-symptomatic to a symptomatic condition in response to the buildup of amyloid β-peptide (Perluigi et al., 2016). Amyloid-beta and tau

NO	Target Gene	miRNA
1	RPL28	hsa-miR-137; hsa-miR-652-3p; hsa-miR-3064-5p; hsa-miR-6504-5p
2	SMG8	hsa-miR-200b-3p; hsa-miR-200c-3p; hsa-miR-429
3	UPF2	hsa-miR-103a-3p; hsa-miR-107; hsa-miR-219a-5p; hsa-miR-138-5p; hsa-miR-143-3p; hsa-miR-493-3p; hsa-miR-4782-3p; hsa-miR-6766-3p
4	EIF3A	hsa-miR-182-5p; hsa-miR-424-5p; hsa-miR-488-3p; hsa-miR-497-5p; hsa-miR-500b-5p

Table 2: miRNA predicted for hub gene of data set GSE165082 of Parkinson's disease.

proteins may interact with each other to enhance the nucleation and propagation of different neurological diseases (Thompson et al., 2020). Aberrant phosphorylation of several proteins occurs in the brain of AD patients and also in its prodromal phase, amnestic mild cognitive impairment (MCI) (Chung, 2009; Thomas et al., 2009). The neuronal cells undergo mitotic catastrophe and endoreduplication prior to cell death in Parkinson's disease (Wang et al., 2014). Cell cycle activation and cell cycle re-entry are thought to be critical causes of neuronal death resulting in neurodegenerative disorders (Ogawa et al., 2003).

 Digital Object Identifier: https://doi.org/10.33865/wjb.009.01.1161

The highest number of DEGs were found to be involved in two molecular functions i.e., protein binding and ATP binding. The important landmark of neurodegenerative diseases is accumulation of protein inclusions, might be due to some mutation in genes encoding for aggregation-prone proteins (Calabrese et al., 2022). ATP binding ABC transporters are predominantly expressed in the brain (Katzeff and Kim, 2021) and dysregulation thought be linked with progression of neurodegenerative diseases (Katzeff et al., 2022). The KEGG pathway analysis predicts the genes involved in different biological pathways (Ogata et al., 1998). The highest number of genes from our dataset were found to be involved in neurodegenerative disease (24), cancer (23), alzheimer disease (20) and corona virus disease (16). It was reported that COVID increase the chances for neurodegenerative disease (Li et al., 2022). The common genes involved in COVID 19 and neurodegenerative diseases have been used as common drug targets (Deng et al., 2023). Many progeria syndromes are connected to an increased risk of developing both cancer and central nervous system (CNS) diseases (Navarro et al., 2006).

Highest gene count (22) for Armadillo-type fold domain were predicted by INTERPRO. The highest number of genes (17) were found to be part of S-TKc domain predicted by SMART. S-TKc is a conserved protein kinase domain which is involved in phosphorylation (Hanks et al., 1988) this is the major molecular function for DEGs has been predicted in current study. The armadillo repeats, ankyrin repeats and leucine-rich repeats together form an extended N-terminal flexible 'solenoid'-like structure composed of tandem repeat modules likely to be important in anchoring to the membrane and cytoskeletal structures as well as binding to other protein ligands (Mills et al., 2012). In the current study protein-protein binding has been predicted as a major molecular function of DEGs related to Parkinson disease.

Parkinson's disease (PD) is a condition associated with the degeneration of dopaminergic neurons in the basal ganglia and associated areas of the brain (Rusz et al., 2016). The Parkinson disease affects the motion and speech of a person (Vandana et al., 2021). The results of current study showed DEGs found to be upregulated in Palate, Ventral striatum, Pluripotent stem cells, Ganglia, Curtilage and Ciliary muscle. All these sites are linked with nervous system and oral functions

Interferon regulatory factor 1 (IRF1), a major transcription factor predicted for DEGs in current dataset. It is an important protein encoded by the *IRF1* gene in human (Itoh et al., 1991). In the current study hub genes were found to be interacting with other genes involved in interferon activities. A major factor in CNS disorders and neuro-inflammation is microglial activation. The transcription factor interferon regulatory factor 1 (IRF1) plays crucial roles in microglial activation and retinal inflammation through controlling the expression of pro- and anti-inflammatory genes (Yang et al., 2022). IRF1 is also main transcription factor for IFN-1 (Kano et al., 2008). Type I interferon (IFN-I) is an innate cytokine family produced in response to viral infections as a first line of defense for the host (Roy and Cao, 2022). It's strictly regulated production is necessary for normal functioning. Persistently elevated IFN-I levels result in auto-inflammatory diseases in multiple organs including brain (Crow and Stetson, 2022).

Hub genes CytoHubba, evaluates the significance of the nodes in a biological network as well as to select the key genes in protein-protein network by using 12 different statistical approaches i.e., Degree, Edge Percolated Component (EPC), Maximum Neighborhood Component (MNC), Density of Maximum Neighborhood Component (DMNC), Maximal Clique Centrality (MCC), Bottleneck, EcCentricity, Closeness, Radiality, Betweenness, Stress and Clustering Coefficient (Chin et al., 2014a). These methods are categorized into 2 main groups: i.e., local and global methods. A local method just inspect the link between the node and its neighbors for calculating the node score with in a network, on the other hand, the link between the whole network and its nodes can be scrutinize by global method. Degree, MNC, DMNC, MCC are local methods and Closeness, EcCentricity, Clustering Coefficient, Radiality, Stress, Betweenness, EPC and BottleNeck are global methods (Chin et al., 2014b). We have identified 10 hub genes, RPL28, IFIT3, IFI6, SMG8, UPF2, XAF1, IFITM1, EIF3A, IFI35, LY6E, by using Cytohubba. Six hub genes are related to interferon activity IFIT3, IFI6, XAF1, IFITM1, EIF3A, IFI35. Interferon factors are important to maintain the homeostasis of central nervous system (Biernacki et al., 2005). Some IFNs plays a crucial role by promoting the secretion of nerve growth factor (Tan et al., 2022).

The interaction network of hub genes with other human genes was created by GeneMANIA. The node size of interacting genes vary according to the confidence score of prediction. The edge thickness shows strength & weight of interaction (Mostafavi et al., 2008). The edge colors shows type of interaction study purple edges represent co-expression & blue represents genes in common pathways (Zuberi et al., 2013). Six genes ISG15 (4.52), IFTI27 (4.49), IFIT1 (4.49), IFITM3 (4.38), IRF9 (4.32) and ADAR (4.06) represented by larger nodes showed higher confidence of prediction. In our datasets highest scored four genes have interferon related activity *i.e.*, IFTI27 is interferon alpha inducible protein, IFIT1interferace induced protein with tetratricoprotien repeat, IFITM3 interferon induce transmembrane induce protein and IRF9 interferon regulatory factor. These are very important factors in nervous system processes Downregulation of interferon signaling activity may increase the probability of neurodegenerative illness progression and hence serve as a biomarker for disease prognosis (Song et al., 2022).

Translational control is important in regulating gene expression and takes place predominantly during the initiation step, which involves numerous eukaryotic translation initiation proteins (EIFs) (Dong et al., 2009; Mathews et al., 2000). EIF3 complex, is the largest initiation complex in human and critically involved in mRNA translation for cell proliferation (Dong et al., 2009; Dong and Zhang, 2006). In the current studies highest expression level of EIF3A in all the categories has been supported by the fact that these genes are components of eukaryotic translation complex.

The miRNA have been predicted in case of four hub genes, RPL28, SMG8, UPF2 and EIF3a with highest numbers (8) for UPF2. UPF1 (Up-frameshift protein 1) plays very important role in neuroprotective disease by inhibiting the accumulation of misfolded proteins in the cell (Staszewski et al., 2023). The microRNA (miRNA) is a type of single-stranded, 18-25nt long, RNA. They can regulate the expression of other protein-coding genes so, served as medically important biomarkers (Ying et al., 2008). Micro RNA (miRNA) can be used alternate targets (Ardekani and Naeini, 2010) as they can remain more stable than mRNA in the cell (Angelucci et al., 2019; Sun et al., 2018).

CONCLUSIONS: The integrative bioinformatics analysis of differentially expressed genes in Parkinson disease from dataset GSE165082 revealed: The DEGs were found to be localized mainly in cytoplasm and nucleus and significantly involved in cell division (biological processes) and molecular functions like protein binding and ATP binding. These were also predicted to be involved in some disease pathways i.e., neurodegenerative diseases, cancer, alzheimer's disease, and coronavirus disease. These genes were also reported to be enriched in important protein domains with highest prevalence in Armadillo-type fold domains. The IRF1 was identified as the most significant transcription factor for the IFN related DEGs in dataset. The gene were found upregulated in Palate, Ventral striatum, Pluripotent stem cells, Ganglia, Curtilage and Ciliary muscle. Ten hub genes, EFI3A, RPL28, SMG8. UPF2, XAF1, IFITM1, IFIT3, LY63, IFI3 and LY6B, were identified, by using 12 statistical methods. The miRNA prediction for hub genes resulted in eight miRNAs against UPF2 while only three for SMG8. In the heat map EIF3A showed sharpest red color for almost all the categories except for adult kidney. Overall, this analysis provides valuable insights into the molecular mechanisms, potential disease associations, and regulatory factors governing the expression of DEGs in the GSE165082 dataset. These findings can be served as a baseline information in understanding the biological significance of differentially expressed genes. These findings also provides an insight into the important biomarkers for therapeutic strategies against this disease.

CONFLICT OF INTEREST: The authors declared no conflict of interest.

REFERENCES: Angelucci, F., Cechova, K., Valis, M., Kuca, K., Zhang, B., Hort, J., 2019. MicroRNAs in Alzheimer's disease: diagnostic markers or therapeutic agents? Frontiers in pharmacology 10, 665.

Ardekani, A.M., Naeini, M.M., 2010. The role of microRNAs in human diseases. Avicenna journal of medical biotechnology 2, 161.

Armstrong, M.J., Okun, M.S., 2020. Diagnosis and treatment of Parkinson disease: a review. Jama 323, 548–560.

Biernacki, K., Antel, J.P., Blain, M., Narayanan, S., Arnold, D.L., Prat, A.

,2005. Interferon beta promotes nerve growth factor secretion early in the course of multiple sclerosis. Archives of neurology 62, 563-568.

Calabrese, G., Molzahn, C., Mayor, T., 2022. Protein interaction networks in neurodegenerative diseases: From physiological function to aggregation. Journal of biological chemistry 298.

Can, T., 2014. Introduction to Bioinformatics, in: Yousef, M., Allmer, J. (Eds.), miRNomics: MicroRNA Biology and Computational Analysis, Methods in Molecular Biology. Humana Press, Totowa, NJ, pp. 51-71.

Chin, C.-H., Chen, S.-H., Wu, H.-H., Ho, C.-W., Ko, M.-T., Lin, C.-Y., 2014a. cytoHubba: identifying hub objects and sub-networks from complex interactome. BMC systems biology 8, 1–7.

Chin, C.-H., Chen, S.-H., Wu, H.-H., Ho, C.-W., Ko, M.-T., Lin, C.-Y., 2014b. cytoHubba: identifying hub objects and sub-networks from complex interactome. BMC systems biology 8, 1–7.

Chung, S.-H., 2009. Aberrant phosphorylation in the pathogenesis of Alzheimer's disease. BMB reports 42, 467–474.

Cline, M.S., Smoot, M., Cerami, E., Kuchinsky, A., Landys, N., Workman, C., Christmas, R., Avila-Campilo, I., Creech, M., Gross, B., 2007. Integration of biological networks and gene expression data using Cytoscape. Nature protocols 2, 2366–2382.

Crow, Y.J., Stetson, D.B., 2022. The type I interferonopathies: 10 years on. Nature Reviews Immunology 22, 471–483.

De Lau, L.M., Breteler, M.M., 2006. Epidemiology of Parkinson's disease. The Lancet Neurology 5, 525–535.

Deng, L., Ding, L., Duan, X., Peng, Y., 2023. Shared molecular signatures between coronavirus infection and neurodegenerative diseases provide targets for broad-spectrum drug development. Scientific reports 13, 5457.

Dong, Z., Liu, Z., Cui, P., Pincheira, R., Yang, Y., Liu, J., Zhang, J.-T., 2009. Role of eIF3a in regulating cell cycle progression. Experimental cell research 315, 1889–1894.

Dong, Z., Zhang, J.-T., 2006. Initiation factor eIF3 and regulation of mRNA translation, cell growth, and cancer. Critical reviews in oncology/hematology 59, 169–180.

Dorsey, E.R., Elbaz, A., Nichols, E., Abbasi, N., Abd-Allah, F., Abdelalim, A., Adsuar, J.C., Ansha, M.G., Brayne, C., Choi, J.-Y.J., 2018. Global, regional, and national burden of Parkinson's disease, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. The Lancet Neurology 17, 939–953.

Gao, H.-M., Jiang, J., Wilson, B., Zhang, W., Hong, J.-S., Liu, B., 2002. Microglial activation-mediated delayed and progressive degeneration of rat nigral dopaminergic neurons: relevance to Parkinson's disease: Microglia-mediated dopaminergic neurodegeneration. Journal of Neurochemistry 81, 1285–1297.

Garcia-Moreno, A., López-Domínguez, R., Villatoro-García, J.A., Ramirez-Mena, A., Aparicio-Puerta, E., Hackenberg, M., Pascual-Montano, A., Carmona-Saez, P., 2022. Functional enrichment analysis of regulatory elements. Biomedicines 10, 590.

Haarman, B.C.B., Riemersma-Van der Lek, R.F., Nolen, W.A., Mendes, R., Drexhage, H.A., Burger, H., 2015. Feature-expression heat maps–A new visual method to explore complex associations between two variable sets. Journal of biomedical informatics 53, 156-161.

Hanks, S.K., Quinn, A.M., Hunter, T., 1988. The Protein Kinase Family: Conserved Features and Deduced Phylogeny of the Catalytic Domains. Science 241, 42–52.

Henderson, A.R., Wang, Q., Meechoovet, B., Siniard, A.L., Naymik, M., De Both, M., Huentelman, M.J., Caselli, R.J., Driver-Dunckley, E., Dunckley, T., 2021. DNA methylation and expression profiles of whole blood in Parkinson's disease. Frontiers in Genetics 12, 640266.

Huang, D.W., Sherman, B.T., Lempicki, R.A., 2009. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nature Protocols 4, 44-57.

Itoh, S., Harada, H., Nakamura, Y., White, R., Taniguchi, T., 1991. Assignment of the human interferon regulatory factor-1 (IRF1) gene to chromosome 5q23–q31. Genomics 10, 1097-1099.

Jankovic, J., 2008. Parkinson's disease: clinical features and diagnosis. Journal of neurology, neurosurgery & psychiatry 79, 368–376.

Kano, S., Sato, K., Morishita, Y., Vollstedt, S., Kim, S., Bishop, K., Honda, K., Kubo, M., Taniguchi, T., 2008. The contribution of transcription factor IRF1 to the interferon-γ–interleukin 12 signaling axis and TH1 versus TH-17 differentiation of CD4+ T cells. Nature immunology 9, 34-41.

Katzeff, J.S., Kim, W.S., 2021. ATP-binding cassette transporters and neurodegenerative diseases. Essays in biochemistry 65, 1013-1024.

Katzeff, J.S., Lok, H.C., Bhatia, S., Fu, Y., Halliday, G.M., Kim, W.S., 2022. ATP-binding cassette transporter expression is widely dysregulated in frontotemporal dementia with TDP-43 inclusions. Frontiers in molecular neuroscience 15, 1043127.

Langfelder, P., Horvath, S., 2008. WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics 9, 559.

Latchman, D.S., 1997. Transcription factors: an overview. The international journal of biochemistry & cell biology 29, 1305-1312.

Li, C., Liu, J., Lin, J., Shang, H., 2022. COVID-19 and risk of neurodegenerative disorders: A Mendelian randomization study. Translational psychiatry 12, 283.

Li, J., Zhou, D., Qiu, W., Shi, Y., Yang, J.-J., Chen, S., Wang, Q., Pan, H., 2018. Application of weighted gene co-expression network analysis for data from paired design. Scientific reports 8, 622.

Lukiw, W.J., Alexandrov, P.N., 2012. Regulation of Complement Factor H (CFH) by Multiple miRNAs in Alzheimer's Disease (AD) Brain. Molecular Neurobiology 46, 11–19.

Mathews, M.B., Sonenberg, N., Hershey, J.W., 2000. Origins and principles of translational control. Cold spring harbor monograph series 39, 1-32.

Mills, R.D., Mulhern, T.D., Cheng, H.-C., Culvenor, J.G., 2012. Analysis of LRRK2 accessory repeat domains: prediction of repeat length, number and sites of Parkinson's disease mutations. Biochemical society transactions 40, 1086–1089.

Mostafavi, S., Ray, D., Warde-Farley, D., Grouios, C., Morris, Q., 2008. GeneMANIA: A real-time multiple association network integration algorithm for predicting gene function. Genome Biology 9.

Navarro, C.L., Cau, P., Lévy, N., 2006. Molecular bases of progeroid syndromes. Human molecular genetics 15, R151–R161.

Ogata, H., Goto, S., Fujibuchi, W., Kanehisa, M., 1998. Computation with the KEGG pathway database. Biosystems 47, 119–128.

Ogawa, O., Lee, H., Zhu, X., Raina, A., Harris, P.L.R., Castellani, R.J., Perry, G., Smith, M.A., 2003. Increased p27, an essential component of cell cycle control, in Alzheimer's disease. Aging Cell 2, 105-110.

Park, T., Yi, S.-G., Kang, S.-H., Lee, S., Lee, Y.-S., Simon, R., 2003. Evaluation of normalization methods for microarray data. BMC bioinformatics 4, 33.

Parkinson, J., 1817. An essay on the shaking palsy: London: Whittingham and Rowland for Sherwood. Neely and jones.

Perluigi, M., Barone, E., Di Domenico, F., Butterfield, D.A., 2016. Aberrant protein phosphorylation in Alzheimer disease brain disturbs pro-survival and cell death pathways. Biochimica et Biophysica Acta (BBA)-Molecular Basis of Disease 1862, 1871–1882.

Roy, E., Cao, W., 2022. Glial interference: impact of type I interferon in neurodegenerative diseases. Molecular neurodegeneration 17, 78.

Rusz, J., Hlavnička, J., Tykalová, T., Bušková, J., Ulmanová, O., R\uužička, E., Šonka, K., 2016. Quantitative assessment of motor speech abnormalities in idiopathic rapid eye movement sleep behaviour disorder. Sleep medicine 19, 141–147.

San Segundo-Val, I., Sanz-Lozano, C.S., 2016. Introduction to the Gene Expression Analysis, in: Isidoro García, M. (Ed.), Molecular Genetics of Asthma, Methods in Molecular Biology. Springer New York, New York, NY, pp. 29-43.

Song, L., Chen, J., Lo, C.-Y.Z., Guo, Q., Feng, J., Zhao, X.-M., 2022. Impaired type I interferon signaling activity implicated in the peripheral blood transcriptome of preclinical Alzheimer's disease. EBioMedicine 82.

Staszewski, J., Lazarewicz, N., Konczak, J., Migdal, I., Maciaszczyk-Dziubinska, E., 2023. UPF1-From mRNA degradation to human disorders. Cells 12, 419.

Sun, Z., Shi, K., Yang, S., Liu, J., Zhou, Q., Wang, G., Song, J., Li, Z., Zhang, Z., Yuan, W., 2018. Effect of exosomal miRNA on cancer biology and clinical applications. Molecular cancer 17, 147.

Tan, P.-H., Ji, J., Hsing, C.-H., Tan, R., Ji, R.-R., 2022. Emerging roles of type-I Interferons in Neuroinflammation, neurological diseases, and long-haul COVID. International journal of molecular sciences 23, 14394.

Thomas, S.N., Cripps, D., Yang, A.J., 2009. Proteomic Analysis of protein phosphorylation and ubiquitination in alzheimer's

disease, in: Ottens, A.K., Wang, K.K.W. (Eds.), Neuroproteomics, methods in molecular biology. Humana press, Totowa, NJ, pp. 109–121.

Thompson, T.B., Chaggar, P., Kuhl, E., Goriely, A., Initiative, A.D.N., 2020. Protein-protein interactions in neurodegenerative diseases: A conspiracy theory. PLoS computational biology 16, e1008267.

Vandana, V.P., Darshini, J.K., Vikram, V.H., Nitish, K., Pal, P.K., Ravi, Y., 2021. Speech characteristics of patients with Parkinson's disease-does dopaminergic medications have a role? JNRP 12, 673–679.

Wang, H., Chen, Y., Chen, J., Zhang, Z., Lao, W., Li, X., Huang, J., Wang, T., 2014. Cell cycle regulation of DNA polymerase beta in rotenone-based Parkinson's disease models. PLoS one 9, e109697.

Yang, W., Hamilton, J.L., Kopil, C., Beck, J.C., Tanner, C.M., Albin, R.L., Ray Dorsey, E., Dahodwala, N., Cintina, I., Hogan, P., 2020. Current and projected future economic burden of Parkinson's disease in the US. NPJ Parkinson's disease 6, 15.

Yang, X., Diaz, V., Huang, H., 2022. The Role of Interferon Regulatory Factor 1 in Regulating Microglial Activation and Retinal Inflammation. International Journal of molecular sciences 23, 14664.

Yang, Y., Han, L., Yuan, Y., Li, J., Hei, N., Liang, H., 2014. Gene co-expression network analysis reveals common system-level properties of prognostic genes across cancer types. Nature communications 5, 3231.

Ying, S.-Y., Chang, D.C., Lin, S.-L., 2008. The MicroRNA (miRNA): Overview of the RNA genes that modulate gene function. Molecular biotechnology 38, 257-268.

Zhang, B., Horvath, S., 2005. A General Framework for Weighted Gene Co-Expression Network Analysis. Statistical applications in genetics and molecular biology 4.

Zuberi, K., Franz, M., Rodriguez, H., Montojo, J., Lopes, C.T., Bader, G.D., Morris, Q., 2013. GeneMANIA prediction server 2013 update. Nucleic acids research 41, W115-W122.