"***In silico approach towards Parkinson's disease pathophysiology, drug repurposing and post translational modifications***"

A DISSERTATION

SUBMITTED IN THE PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE AWARD OF THE DEGREE OF

MASTER OF TECHNOLOGY

IN

BIOINFORMATICS

Submitted by:

Devesh Srivastava

2K19/BIO/06

Under the supervision of

Prof. Pravir Kumar

Department of Biotechnology

Delhi Technological University

(Formerly Delhi College of Engineering)

Bawana Road, Delhi-110042

Delhi Technological University

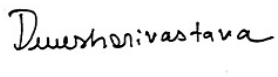
(Formerly Delhi College of Engineering)

Bawana Road, Delhi-110042

CANDIDATE'S DECLARATION

I, Devesh Srivastava, Roll No. - 2K19/BIO/06, student of M.Tech Bioinformatics, hereby declare that the project thesis titled "**In silico approach towards Parkinson's disease pathophysiology, drug repurposing and post translational modifications**" which is submitted by me to the Department of Biotechnology, Delhi Technological University, Delhi, in partial fulfilment of the requirement for the award of the degree of Master of Technology, is original and not copied from any source without paper citation. The work has not previously formed the basis for the award of any Degree, Diploma Associateship, Fellowship or any other similar title or recognition.

Place: Lucknow



Date: 3rd July 2021

Devesh Srivastava

2K19/BIO/06

M.Tech Bioinformatics

Department of Biotechnology
DELHI TECHNOLOGICAL UNIVERSITY
(Formerly Delhi College of Engineering)
Bawana Road, Delhi-110042

CERTIFICATE

I hereby certify that the Project Dissertation titled "**In silico approach towards Parkinson's disease pathophysiology, drug repurposing and post translational modifications**" which is submitted by Devesh Srivastava, 2K19/BIO/06, Department of Biotechnology, Delhi Technological University, Delhi in partial fulfilment of the requirement for the award of the degree of Master of Technology, is a record of the project work carried out by the student under my supervision. To the best of my knowledge this work has not been submitted in part or full for any Degree or Diploma to this University or elsewhere.

Place: DTU, Delhi

Date: 03.07.2021

Prof. Pravir Kumar
Guide and Head of Department
Department of Biotechnology
Delhi Technological University

ACKNOWLEDGEMENT

First of all, like I would like to thank Omnipotent and Almighty God for showing me the right way and giving me strength to complete this project. I express my deepest and sincere regards to Dr. Pravir Kumar, Professor and Head, Department of Biotechnology, Delhi Technological University, who selected me to work under him in his laboratory. He always kept guiding and motivating me, and also shared valuable life lessons throughout this whole journey, which I will remember my whole life. I wouldn't have been able to finish the project without his support.

I would also like to extend my deep regards and gratitude to Dr. Rashmi K. Ambasta for her valuable inputs in various scientific papers.

I would also give a special mention to my colleagues Rohan Gupta, Mehar Sahu and Swati Tiwari for being perfect team mates and helping each other. Specially, PhD scholar, Rohan Gupta, who was always there to help me and taught me new techniques related to the project and scientific publications. Lastly, I would thank my parents who kept cheering me up through Ups and downs of this journey.

Devesh Srivastava

2K19/BIO/06

Abstract

Parkinson's disease (PD) is an enervating, debilitating and lethal neurodegenerative disorder marked by deterioration of neurons which produces dopamine in the central nervous system. PD is accompanied by a constellation of lethal motor and non-motor symptoms which are observed when the disease is progressed at an advanced stage. Hence, there is a great necessity of novel blood-based biomarkers which can help in early detection of the disease and can serve as new therapeutic targets to impede the progression of disease. Herein, we performed blood-based differential gene expression analysis of Parkinson's samples and healthy samples to look for novel blood-based gene biomarkers and their target drugs. Herein firstly, we downloaded blood-based microarray gene expression omnibus (GEO) dataset to explore differentially expressed genes (DEGs) in PD samples compared to healthy control samples. We found 18 DEGs between PD and healthy samples, most of which were novel genes for PD. Further, we validated these DEGs via machine learning algorithms using their expression signature as input features. Validation with algorithms such as Artificial neural networks, Decision trees, Random Forest, Linear discriminant analysis and kernel principal component analysis (PCA) models resulted in accuracy of 92.8%, 78.5%, 92.8%, 100%, 92.8% respectively. Moreover, using hierarchical clustering based unsupervised machine learning we found that PD and healthy samples were well differentiated in two separate clusters. Furthermore, we used LINCS L1000 based drug repurposing search engine L1000CDS^2, and CoDReS tool to look for exemplar repurposed drugs which can reverse the expression of our obtained genes, thereby we obtained several drugs with neuroprotective properties. In addition, we looked for novel transcription factors regulating the dysregulation of genes targeted by the shortlisted drugs. Further, using in silico tools we found various post translational modifications involved in drug-gene pathway. Lastly, we searched for common drugs with can target both PD pathogenesis and ageing.

Keywords: Parkinson Disease, Differentially Expressed Genes, Machine Learning, Drug Repurposing, Therapeutic Targets

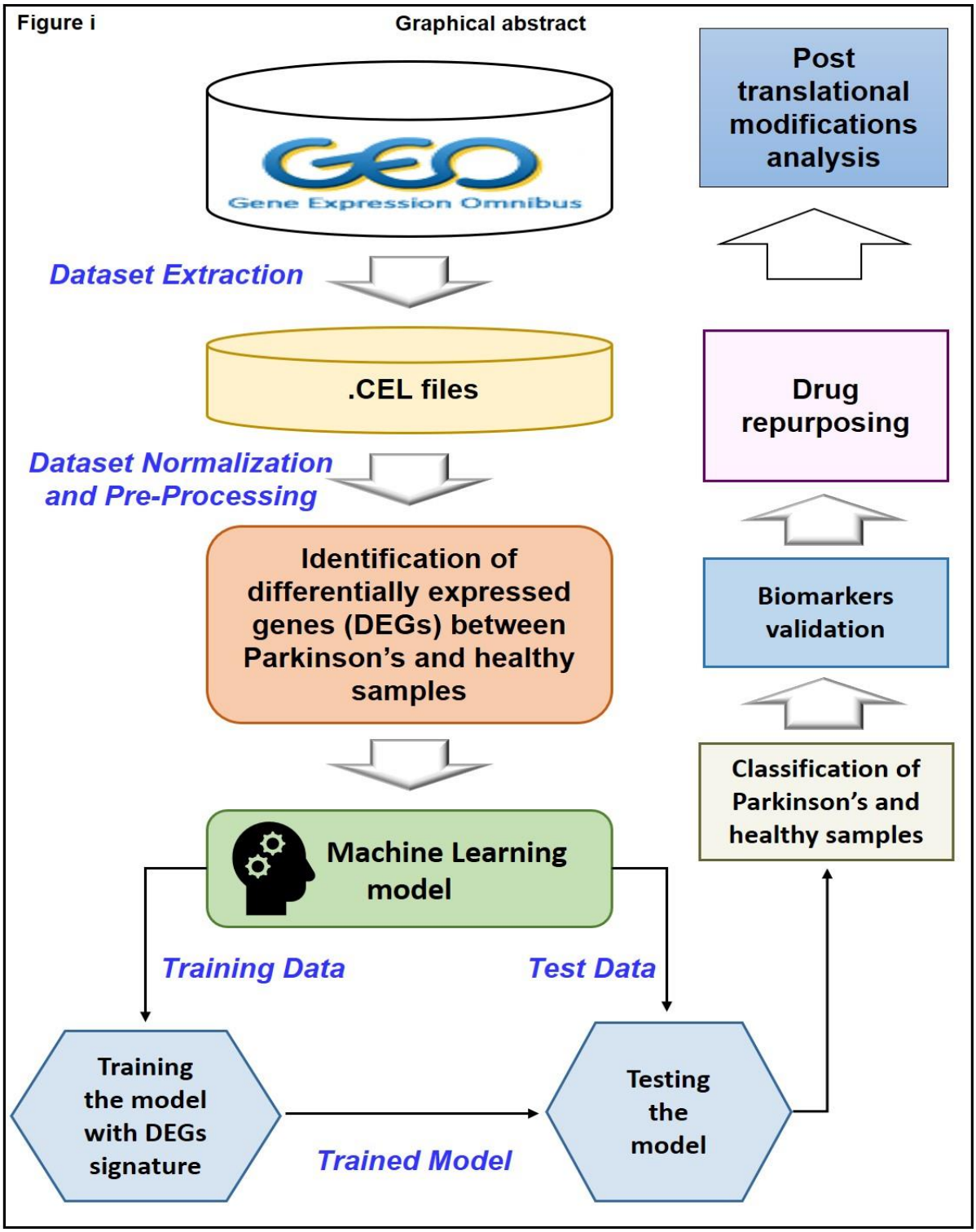


Figure i: Graphical abstract- Herein, firstly we acquired dataset from NCBI GEO, and then found out differentially expressed genes (DEGs) between Parkinson's and healthy control samples. Further, we created myriad of machine learning models to see if these DEGs can classify between Parkinson's and healthy samples. Thereafter, we used drug repositioning approach to find out drugs whichcan revert the expression levels of these DEGs, and looked for various post translational modifications in Drug-gene pathway.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVATIONS

PD: Parkinson's Disease

GEO: Gene expression Omnibus

DEG: Differentially Expressed Genes

PCA: Principal Component Ananlysis

LINCS: Library of Integrated Network Based Cellular Signature

L1000CDS^2: L1000 Characteristic Direction Signature Search Engine

RNA: Ribonucleic acid

NCBI: National Center for Biotechnology Information

Limma: Linear Model for Microarray data

ANN: Artificial Neural Network

LDA: Linear Discriminant Analysis

ROC: Receiver Operating Characteristic

ADAMTS2: ADAM metallopeptidase with thrombospondin type 1 motif 2

EHBP1: EH binding protein 1

CBX5: chromobox 5

PKD1: Polycystin 1

RIBC2: RIB43A Domain with Coiled-Coils 2

CCR9: C-C motif chemokine receptor 9

PCDHGA3: Protocadherin gamma subfamily A, 3

ZNF551: Zinc finger protein 551

PRDM12: PR/SET domain 12

FGF9: Fibroblast growth factor 9

HLA-DOA: Major histocompatibility complex, Class II, DOA

ADRA1D: Adrenoceptor alpha 1D

KLHL22: Kelch like family member 22

TMEM19: Transmembrane protein 19

NUBPL: Nucleotide binding protein

mRNA: messenger RNA

MPP+: 1-methyl-4phenylpyridinium ion

HDAC: Histone deacetylases

PTEN: phosphatase and tensin homolog

MPTP: 1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine

CDK: Cyclin dependent kinase

PI3K: Phosphoinositide 3-kinase

ROS: reactive oxygen species

SIN3B: SIN3 Transcription Regulator Family Member B

BMI1: BMI1 Polycomb Ring Finger Oncogene

SPI1: SPI1 Proto-Oncogene

POU5F1: POU Domain, Class 5, Transcription Factor 1

MTF2: Metal-Response Element-Binding Transcription Factor 2

HNF4A: Hepatocyte Nuclear Factor 4 Alpha

SOX2: SRY-Box Transcription Factor 2

FLI1: FLI1 Proto-Oncogene

SBDSP1: SBDS Pseudogene 1

SYMPK: Symplekin

CACNG2: Calcium Voltage-Gated Channel Auxiliary Subunit Gamma 2

TLK1: Tousled Like Kinase 1

1. Introduction

Parkinson's disease (PD) is very multifaceted, pernicious, debilitating neurodegenerative disorder (NDD) and a big burden globally. It is a result of progressive and chronic deterioration of dopamine producing neurons in the midbrain region of substantia nigra pars compacta [1]. In addition, pathogenic, toxic, intraneuronal aggregates known as Lewy bodies, mainly composed of α- synuclein proteins, are major histopathological trademark features of PD. PD is accompanied by a myriad of enervating motor symptoms like bradykinesia, tremors at rest, rigidity of muscles, postural instability which are a direct result of dopaminergic neurons dysfunction [2]. In addition, plethora of non-motor symptoms like sleeping problems, cognitive decline, depression, urinary disorder, hallucinations, obsessive compulsive disorders are also witnessed in PD [3,4]. The major issue with a NDD like PD is that it is diagnosed and its symptoms are visible only when the disease is already at an advanced stage, where around 70-80% dopamine producing neurons have already been degenerated, with symptoms gradually exacerbating with time [5]. Therefore, there is a great desire and need to look for novel blood-based biomarkers which can help in early detection and can also serve as potential novel therapeutic targets. The advantage with blood-based transcriptome is that it is non-invasive which can be easily obtained in a clinical setting and blood might contain a disease signature which can help in early detection. With the advent of microarray, RNA-SEQ technologies and with increasing genomic databases various gene expression-based biomarkers analysis is being performed nowadays. In addition, machine learning and other statistical tools are widely being used now for disease characterization as these tools can find out hidden patterns in the genomic data. For instance, one study used blood based gene expression data and performed machine learning based analysis for Alzheimer's disease (AD) pathophysiology [6]. Likewise, D.H. Oh, et al. performed blood transcriptome analysis of young adults with autism spectrum disorder then used machine learning to find biomarkers related to autism spectrum disorder [7]. Furthermore, it is imperative to find new target drugs for PD which can thwart the progression of disease and alleviate the symptoms. Nowadays, drug repositioning is very popular cost and time efficient drug discovery strategy, where existing approved, experimental drugs can be used for novel purposes [8]. For instance, one study in PD used genome wide association studies (GWAS) data and performed drug repositioning analysis to report some new drugs for PD [9]. Herein, we used microarray-based blood transcriptome dataset, GSE72267, from the national center for biotechnology information (NCBI) gene expression omnibus (GEO) and found differentially expressed genes (DEGs) between PD samples and healthy control samples.

Furthermore, we used a variety of machine learning (ML) algorithms to see if these DEGs alongwith their expression signature can classify between PD samples and healthy control samples. We used artificial neural network (ANN), decision tree, random forest, kernel principal component analysis (PCA), linear discriminant analysis (LDA) to find a gene set which can differentiate between Parkinson's sample and healthy samples, thus acting as biomarkers. Additionally, we used an unsupervised machine learning technique as well to see if PD and healthy control samples can form separate clusters based on their gene expression signatures. Furthermore, we used LINCS L1000 based drug repurposing search engine and CoDReS tool, to look for existing drugs which can revert the signature of these DEGs, thus serving as new target drugs for PD. Drug repurposing was followed by ADME analysis by SwissADME tool, to verify potential druggability of the shortlisted drugs. In addition, we used ChEA tool to identify novel transcription factors contributing to altered expression of genes targeted by the drugs in order to determine drug-gene pathway. Further, we used various in silico tools to find post translational modifications involved in drug-gene pathway. Lastly, we explored common drugs involved in both ageing and PD pathogenesis and we used molecular docking approach to look for binding sites of a shortlisted drug with HDAC6, as HDAC6 is known to promote protein aggregation in PD.

2. Literature review

2.1 Neurodegenerative disorders: an overview

Neurodegenerative disorders (NDDs) are disorders of the central nervous system which ultimately culminate in neuronal death, which can further result in death of the patient suffering from these disorders. Abnormal protein aggregation both intracellularly and extracellular in various regions of the central nervous system is one of the main causative factors of neurodegenerative disorders [10]. An abnormally folded protein has its hydrophobic core interacting with its outer hydrophilic milieu, which can lead to oligomer and fibril formations further culminating in bulk protein aggregates. Mutations, oxidative stress, post translational modifications, ubiquitin proteasome dysfunction, autophagy dysfunction are some of the causative factors which promote protein aggregation [11–13]. Interestingly, reactive oxygen species like superoxide anion also promotes formation of aggregated proteins [14]. This abnormal protein aggregation further activates immune response by microglia in central nervous leading to neuroinflammation due to cytokines release by microglia, culmination in neuronal death. Strikingly, it has been observed that these toxic aggregates can be travel from one neuron to another, hence, spreading their pathogenicity [15]. Alzheimer's disease, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis (also known as Lou Gehrig's disease), frontotemporal dementia, multiple sclerosis, Lafora disease and spinocerebellar ataxias are some of the commonly observed NDDs globally. Till date, no medication has been developed which can thwart the progress of these brain disorders, most of the medicines can relieve the symptoms only.

Alzheimer's disease (AD) is the most widely prevailing NDD and a major reason of death globally. Short term memory loss, dementia, cognitive dysfunction, speaking problems are main symptoms of AD [16]. Extracellular aggregation of amyloid beta proteins forming plaques and intraneuronal aggregates of tau proteins forming neurofibrillary tangles are considered to be main causative factors leading to AD pathogenesis [17]. Mutations in PSEN1 and PSEN2 genes promote amyloid beta aggregation. Further, abnormal cleavage of amyloid precursor protein: APP driven by β and γ secretase is considered to be the main reason for oligomer and fibril formation of amyloid beta proteins [18]. Further, phosphorylation and hyperphosphorylation of tau proteins is one of the primary reasons of tau protein aggregation which further gives rise to neurofibrillary tangles.

Further, Parkinson's disease (PD) is considered to be the second most commonly observed NDD, mainly affecting older population. Tremors during rest, slow movement or bradykinesia, postural instability, sleep disorder, cognitive decline are most commonly visible symptoms of PD [19]. Intraneuronal aggregates of lewy bodies composed and made up of α-synuclein protein is the main cause of neuronal death during PD [20]. During PD, dopamine engendering neurons in substantia nigra region of basal ganglia are damaged due to aggregated proteins. Hence, dopamine is not produced in sufficient amount during PD [21]. Dopamine is a neurotransmitter which controls voluntary movements, due to dopamine loss, there is no control over voluntary movements during PD. Mutations in SNCA, LRRK2, Parkin, PINK1 genes are known to promote protein aggregation resulting in PD [22]. Levodopa is the most widely used medication and treatment for PD. Levodopa is a precursor of dopamine which easily penetrates the blood-brain barrier and can be further transformed into dopamine. Levodopa is usually given in conjugation with carbidopa so that levodopa is not converted into dopamine before it reaches the desired brain region [23]. Dopamine agonists are also used in some cases to treat PD patients. Monoamine oxidase inhibitors are another type of drugs given to PD patients [24]. In addition, deep brain stimulation is an invasive procedure which is used sometimes to relieve motor symptoms of PD.

Further, amyotrophic lateral sclerosis (ALS) is another lethal NDD, famous scientist Stephen hawking died due to ALS. Unlike other NDDs, ALS is very fast progressing which can lead to death of the individual within 2-5 years of onset. ALS results due to degeneration and damage of both upper and lower motor neurons. Muscle atrophy and damage, paralysis, stiffness are the main symptoms of ALS [25]. Unlike other NDDs there is no cognitive decline in ALS. Aggregation of TDP-43 proteins in central nervous system is considered to be main causative factor of ALS [26]. Mutations in SOD1, and FUS genes are also known to be causative factors of ALS.

Further, Huntington's disease is another NDDs. Involuntary movements, abnormal walking, cognitive dysfunction are the main symptoms of Huntington's disease [27]. Aggregation of mutant huntingtin is main causative factor responsible for Huntington's disease. Expansion of CAG repeats coding for glutamine also causes Huntington's disease. Spinocerebellar ataxia is another disease which results due to polyglutamine expansion [28]. Lafora disease is another NDD which results due to aggregation of glycogen in the brain [29]. In addition, Frontotemporal dementia is another NDD which is a result of TDP-43 aggregation.

2.2 Protein aggregation and Post Translational Modifications

Mutations, oxidative stress, dysfunction of protein quality control system like ubiquitin-proteasome system, molecular chaperones, autophagy-lysosome pathway and various post translational modifications are major factors responsible of intracellular and extracellular protein aggregation. Post translational modifications are biochemical enzymatic modifications of proteins with different moieties after they have been synthesized [30]. Further various post translational modifications (PTMs) also promote protein aggregation contributing to pathogenesis of NDDs. For instance, phosphorylation is adduction of phosphate group by various kinases on serine, threonine, tyrosine residues. S129 phosphorylation of α-synuclein potentiates aggregation of α-syunclein. Likewise, tau phosphorylation medicated by GSK3β at serine and threonine residues also potentiates tau towards aggregation [31]. Similarly, acetylation is attachment of acetyl moiety on lysine residues and, acetylation at K145 and K192 in RRM domain site of TDP-43 protein increases susceptibility of TDP-43 towards aggregation [32]. GlcNAcylation and glycation of tau proteins also make them more susceptible towards aggregation in AD. Likewise, palmitoylation is covalent modification with 16 carbon fatty acids on cysteine residues and it has been observed that palmitoylation at C186 of amyloid precursor protein stimulated amyloid beta aggregation in AD [33]. Further, ADP ribosylation is covalent adduction of ADP-ribose group and ribosylation of TDP-43 also potentiates it towards aggregation leading to ALS [34]. Ubiquitination is covalent adduction of ubiquitin on lysine amino acid residues and ubiquitinated proteins are also found in aggregated lewy bodies, suggesting dysfunction of ubiquitin proteasome system resulting in aggregated proteins. Moreover, SUMOylation is another covalent adduction on lysine sites with small ubiquitin like modifier (SUMO) and it has been reported that SUMOylation of K75 residue of SOD1 also increases its propensity towards aggregation [35]. Further, nitration is addition of nitrate group and it has been reported that Y39 nitration of α-syunclein also increases its aggregation propensity [36]. Importantly, glycation is covalent addition of sugar moieties and it has been observed that glycation of Aβ amplifies its neural toxicity and increases GSK3β activity which is known to promote tau hyperphosphorylation [37]. Moreover, nitrosylation is adduction of nitric oxide group and nitrosylation of PINK1 has been observed to impair its mitophagy activity promoting neuronal cell death [38]. Further, carbamylation is addition of isocyanic acid and it is another PTM which is known to potentiate tau aggregation leading to neuronal toxicity in AD [39]. Further, Glutathionylation is addition of glutathione on cysteine sites and one study observed that glutathionylation of C111 residues of SOD1 increases its susceptibility

to aggregate [40]. Additionally, glycosylation is covalent modification with carbohydrate group and glycosylation also promotes tau hyperphosphorylation leading to its aggregation [41].

2.3 Signal transduction pathways involved in neuroprotection and neurodegeneration

Various different signal transduction pathways also regulate the process of PTMs resulting to neuroprotection or further promoting neurodegeneration. Different transduction pathways like Akt, Wnt, AMPK, MAPK are involved in this process. For instance, activated AKT pathway inhibits GSK3β. Inhibition of GSK3β can reduce tau phosphorylation as it is one of the main kinases promoting tau phosphorylation. Sulfhydration of AKT at C77 inhibits its activity, due to which AKT can't thwart GSK3β through phosphorylation, thus promoting tau hyperphosphorylation [42]. Further, ataxin-1 phosphorylation at S776 by AKT promotes docking of ataxin-1 with the protein named 14-3-3, and the interalinkage of 14-3-3 protein with ataxin, increases aggregation of ataxin in Spinocerebellar ataxia [43]. Additionally, it has been reported that overexpressing DJ-1 protein led to phosphorylation on T308 residue of AKT, which prevented mitochondrial dysfunction in PD model [44]. Further, activation of β catenin in wnt pathway, also inhibits GSK3β activity, thus preventing tau hyperphosphorylation in AD. Another study reported that inhibition of β catenin increased the activity of GSK3β, which in turn promoted tau hyperphosphorylation [45]. In addition, it was observed that DKK1 protein inhibits wnt signal cascade, resulting in no β catenin activation, culminating in tau hyperphosphorylation [46]. Further, it was reported that AMPK pathway reduced tau acetylation by increasing the expression of deacetylates SIRT1 [47]. Further, MAPK pathways are also involved in NDDs. For instance, Moreover, a study pointed out that α-synuclein promotes activation of p38 MAPK cascade, which is further involved in phosphorylation of PARKIN at residue S131, this contributes to mitochondria dysfunction in PD [48]. In addition, ER stress can activate JNK3, and it has been observed that phosphorylation of APP by JNK3 at T668 stimulates Aβ aggregation [49].

2.4 Crosstalk between ageing and neurodegeneration

Further, ageing is considered another major factor which increases risk towards NDDs. Mitochondrial dysfunction, oxidative stress, telomere shortening, cellular senescence or cell cycle arrest, epigenetic changes like DNA methylation, demethylation and histone acetylation, deacetylation are some of the common processes involved in both aging and neurodegeneration. Thus, linking both aging and neurodegeneration. There are various drugs

Further, various post translational modifications are also observed regulating the process of aging. One study reported that protein's O-GlcNAcylation decreased in the hippocampus of aging mouse which co-related with cognitive decline and increase protein O-GlcNAcylation restored cognitive functions [50]. Similarly, increase in phosphorylation, nitration of α-synuclein was observed in basal ganglia-substantia niagra of aging squirrel monkeys [51]. Further, S-Nitrosylation of Parkin was observed with aging, which impairs its mitophagy activity [52]. Additionally, increased succinylation of mitochondrial proteins with age, has been observed in Drosophila melanogaster and C.elegans [53]. Further, sialylation of proteins also decreases with age in hippocampus with potentiates neuronal dysfunction and cognitive decline [54]. Further, one study pointed out that S-Sulfhydration of proteins also decreases with age in rat model [55]. Moreover, it was reported that palmitoylation of NMDAR proteins increases with age in the frontal cortex of mice model which correlated with synaptic and cognitive defects [56]. Interestingly, age related protein carbonylation due to increased oxidative stress has also been observed in aged mice model. It has to be noted that protein carbonylation is a precursor of various protein aggregates [57]. Moreover, Advanced glycation and glycosylation end products are also observed in skin with increasing age [58].

2.5 Artificial Intelligence and Machine learning: an overview

ML is the science of teaching computers how to learn, act and give output without being explicitly programmed. In this era of digitalization various Machine learning is widely being used now in the area of healthcare, drug design and development, genomics, proteomics. ML comes under the branch of artificial intelligence. Artificial intelligence (AI) is science which allows computers to think, act and behave like humans do. All machine learning algorithms are considered to be subset of AI [59]. Alan Turing is considered as the father of AI. Deep learning is another term which is widely being used now. Deep learning algorithms try to emulate the behaviour and functioning of human brain [60]. A myriad of AI, ML and deep learning algorithms are widely being used now in the area of drug design, healthcare, medical informatics etc. Graphics processing units (GPUs) are best suited to handle larger AI and/or ML algorithms. Nvidia's CUDA framework allows user to use AI algorithms through GPUs. Only issue with machine learning algorithms is that sometimes they might suffer with overfitting of data when a large amount of data is being used in order to train them.

Supervised learning, unsupervised learning are the two main types of ML models. In supervised ML technique we feed input to the system and its desirable output, in order to teach and

supervise the model. The model then analyses the input-output relationship, so that whenever a fresh unknown input is applied to the system, the system can give the correct output [61]. Support vector machines, decision trees, artificial neural network, logistic regression, naïve bayes are some of the most popular supervised learning techniques. Whereas in unsupervised learning we only give an input to the system and don't give its expected output. The system is allowed to itself analyse the hidden patterns, structures in the data and the come at a conclusion [62]. Amazon, Netflix, stock market predictions use unsupervised learning techniques. K - means clustering and hierarchical clustering, are two types of unsupervised learning techniques. Artificial neural network is a simple ML algorithm which emulates the function of human brain. There is an input layer followed by a hidden layer followed by an output layer. Moreover, deep learning on the other hand uses a combination of neural networks for higher efficiency. In reinforcement learning there is an agent and a reward, and agent improves its performance in order to achieve maximum reward [63]. Python and R programming are the two main programming techniques use to write machine learning scripts and create machine learning models.

2.6 Big data in drug design and development

Big data means datasets that are so big and gigantic that they cannot be analysed, processed by conventional, traditional software tools and devices [64]. Due to increasing use of different types of sequencing, RNA-seq, microarray and other in silico, in vitro techniques a huge amount of biological data is being produces nowadays. Thus, increasing the amount of big data related to genomics, proteomics, transcriptomics. RNA-seq is used to obtain mRNA expression, this technique can give us dysregulated genes in different conditions. Likewise, microarray technique can also give us differentially expressed genes between diseased state and a normal healthy state. NCBI GEO is a databank which contains data obtained from plethora of microarray and RNA-seq experiments [65]. R programming is one the most widely used tool to analyse expression levels from NCBI GEO datasets. ArrayExpress is another big source of gene expression data obtained from microarray experiments [66]. In addition, TCGA is a big databank of gene expression data related to a myriad varied of cancer [67]. Further, analysis genome wide association studies (GWAS) datasets can give us target genes in a plethora of diseased conditions. NHGRI-EBI GWAS Catalog is a major databank containing GWAS data [68]. GWAS central is another big data repository of GWAS data [69].

Further, sequencing data can also be used to obtain target genes implicated in different diseased conditions. Sequence read archives is a databank which contains data endangered from various sequencing experiments [70]. Many times, literature survey and data mining can also help in obtaining target genes or proteins and PubMed is the biggest source of different biological published scientific literatures [71].

Further, protein data bank (PDB) consists of data of different 3D protein structures. PDB also contains DNA and RNA structural data [72]. PDB can be used to analyse binding of ligand, small molecule with target genes/proteins. Covid 3CL protease on PDB has widely been used to analyse various drugs related to covid [73]. PDB is the most popular databank for molecular docking in conjugation with different machine learning algorithms. Swiss-prot and uniprot are databases which contains all the information related to different protein structures

Further, there are various chemical databases available in order to obtain different small molecules and ligands related to target gene, protein of interest. Pubchem, ChEMBL, DrugBank, BindingDB, zinc database, LINCS L1000CDS^2, are some big chemical databases available to public use in order to obtain small molecules and ligands related to target protein [74–77]. These databases were very widely used during current covid situation in order to ascertain the viability of different drugs for covid treatment. There are various post translational modifications database available for public use. DEEP- PLA is a database which contains acetylation data on lysine sites by various acetylases and HDACs [78]. dbPTM also contains a lot of PTM data on different proteins [79]. PMLD database is also used to get the information of PTMs on different lysine residues.

2.7 Transformation of traditional drug development using artificial intelligence based tools

AI-ML based algorithms can indeed transform the whole landscape of drug design and development. Various algorithms and tools have been devised which has increased the efficiency of drug design and development. For instance, AlphaFold is an AI based tool which can ascertain 3D structure of different protein based on their input amino acid sequences [80]. Likewise, RGN: recurrent geometric network, is another deep learning based tool which can ascertain and predict protein's three dimensional structure using amino acid sequence as an input [81]. Further, SchNOrb is a deep learning driven tool which can predict arrangement of atoms and orbitals in a molecule, which can indeed help in the process of drug design [82]. Further, DriverML is a machine learning based tool which can help in obtaining genes

implicated in pathogenesis of different types of cancer [83]. Further, AI and ML based tools have accelerated the process of denovo drug design as well. For instance, MolAIcal is an AI driven tool to design 3D drugs in protein binding pockets [84]. Moreover, ReLeaSE is a deep reinforcement learning driven tool for de novo synthesis of drugs

Further, various text mining and natural language processing driven tools are also being used in the area of biological sciences. For instance, STRING is a text mining based tool which is used in order to obtain protein-protein interactions [85]. STITCH is another text mining based databank which gives us a relationship between proteins and different chemicals [86]. DisGeNET also uses text mining and gives us information related to gene and disease association [87]. Moreover, AI/ML based tools are also being used for synthesis planning of drugs as well. Further, synthesis planning determines the best synthesis route for the novel drug. Like, Chematica uses decision trees in order to determine best retrosynthesis pathway [88]. Further, AiZynthFinder uses neural networks for synthesis planning [89]. ICSYNTH also uses AI based algorithms in order to determine most optimum synthesis planning route. Further, AI based algorithms have also accelerated the process of docking based virtual screening as well.

Moreover, various AI based algorithms have been devised to ascertain the optimum drug dosage. For instance, CURATE.AI uses AI based algorithms to determine optimum drug dosage. In addition, comboFM uses factorization machines in order to determine the best drug dosage. In addition, AI-ML based algorithms are also being used in order to predict bioactivity as well. For instance, WDL-RF is a random forest drive approach in order to find out bioactive ligands for G-protein coupled receptors (GPCRs) [90]. In addition, pairwiseMKL is a kernel based approach in order to determine bioactivity of ligands [91]. Likewise, DeepMalaria is another deep learning based tool in order to find out compounds with inhibitory activity towards plasmodium falciparum responsible for causing malaria [92].

2.8 Drug development and drug repurposing

The biggest and the utmost challenge facing the field of NDDs is their drug design and development. As till date, no medicine has been found out which can revert and slow down the progression of these NDDs. Most of the drugs which are being used can only temporarily relieve the symptoms. The prolonged use to these drugs gives rise to debilitating side effects. Hence, there is always a need for new class of drugs which can help in the treatment of these disorders. Drug repurposing approach is one the most popular approaches for exploring new

drugs as in drug repurposing approach existing approved, experimental drugs for one disease can be used for another disease [93]. Designing and developing novel drugs from scratch is a time and cost consuming process. Most importantly, a novel drug has to follow a complex process of animal model testing, clinical trial stage 1, clinical trial stage 2, clinical 3 stage then lastly FDA/drug controller approval. So many drugs fail in clinical trials itself whereas many drugs fail to get approval to hit the market. Thus, resulting in wastage of time and money. In this regard, drug repurposing approach can aid and augment the process of drug development as in drug repurposing many drugs are already under experiment or have hit the market. Drug repurposing approach has been used in the field of NDDs as well, especially in the context of PTMs. For instance, a study reported that telmisartan, a drug used for the treatment of hypertension, can reduce tau phosphorylation [94]. Another study reported that, metformin, a drug for treating diabetes, can protect against advanced glycation end products induced neurotoxicity, by activating AMPK pathway [95]. In addition, curcumin, the major component of turmeric which is a popular Indian household spice, was observed to reduce tau phosphorylation in an AD model [96]. Additionally, resveratrol, an anti-oxidant has been reported to inhibit tau phosphorylation and toxicity [97]. Resveratrol has been shown to protect against advanced glycation end products toxicity as well [98]. Hence, we can see that drug repurposing can indeed aid and augment the whole process of drug development. **Figure 1** on page 12, shows different novel and repurposed drugs which can target PTMs driven neurodegeneration.

Further, various AI and ML tools are also being employed for in silico drug repurposing. For instance, deepDR is a deep learning driven tool for in silico drug repurposing. It has been used to find repurposed drugs for NDDs [99]. Likewise, iDrug is another AI algorithm driven tool for computational drug repurposing [100]. In addition, BiFusion is a convolutional neural network driven tool for computational drug repositioning. Moreover, Drugbank, PubChem, ChEMBL database driven tools are being used for drug repurposing with the help of virtual screening in conjugation with machine learning. Moreover, many researchers are using in silico drug repurposing approach in the field of NDDs as well. A study used scientific literature data mining approach in conjugation with ML algorithms in order to obtain repurposed drugs for PD [101]. Likewise, P. Chatterjee et al. used epigenetics driven protein- protein interaction networks in order to find out epigenetics drugs associated with PD [102].

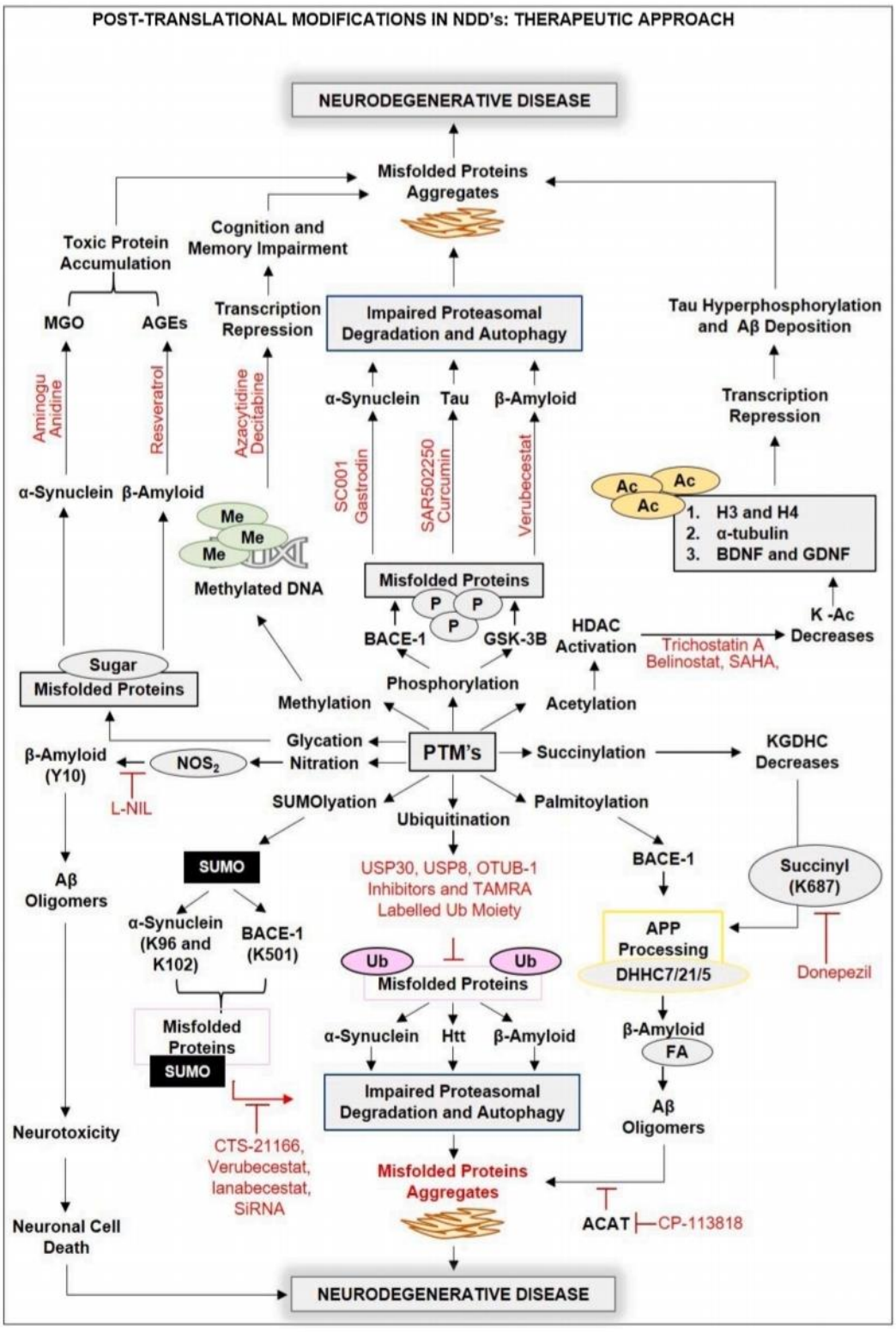


Figure 1: Different novel and repurposed drugs being used to target post translational modifications induced neurotoxicity. Figure taken from our published paper in Elseiver Ageing Research Reviews, titled "Post-translational modifications: Regulators of neurodegenerative proteinopathies".

3. Methodology

3.1. Data acquisition

In this study, we used and downloaded publicly available blood-based microarray dataset, GSE72267, deposited by Calligaris et al., from NCBI GEO. This microarray study was performed on the Affymetrix human genome array. Herein, we used 21 PD patient samples and 19 matched healthy control samples from GSE72267, to avoid class imbalance and make sure that our analysis doesn't get biased with PD data in the dataset.

3.2. Differentially expressed Genes

We used Limma package in R programming language to obtain DEGs between PD samples and healthy control samples, as it is very efficient package for differential gene expression analysis [103]. Firstly, we downloaded .CEL files from GSE72267 containing desired PD and control samples. Then, .CEL files were imported in R studio using read.celfiles function. Furthermore, exprs function was used to find expression values of the probes and normalizeQuantiles function was used to normalize the data (**figure 2**). Then finally, lmFit, eBayes, toptable functions were used to obtained DEGs between PD samples and healthy control samples. Moreover, to obtain a set of significant DEGs, we shortlisted obtained DEGs on the basis of adj.p.value < 0.05, to lower the false positives.

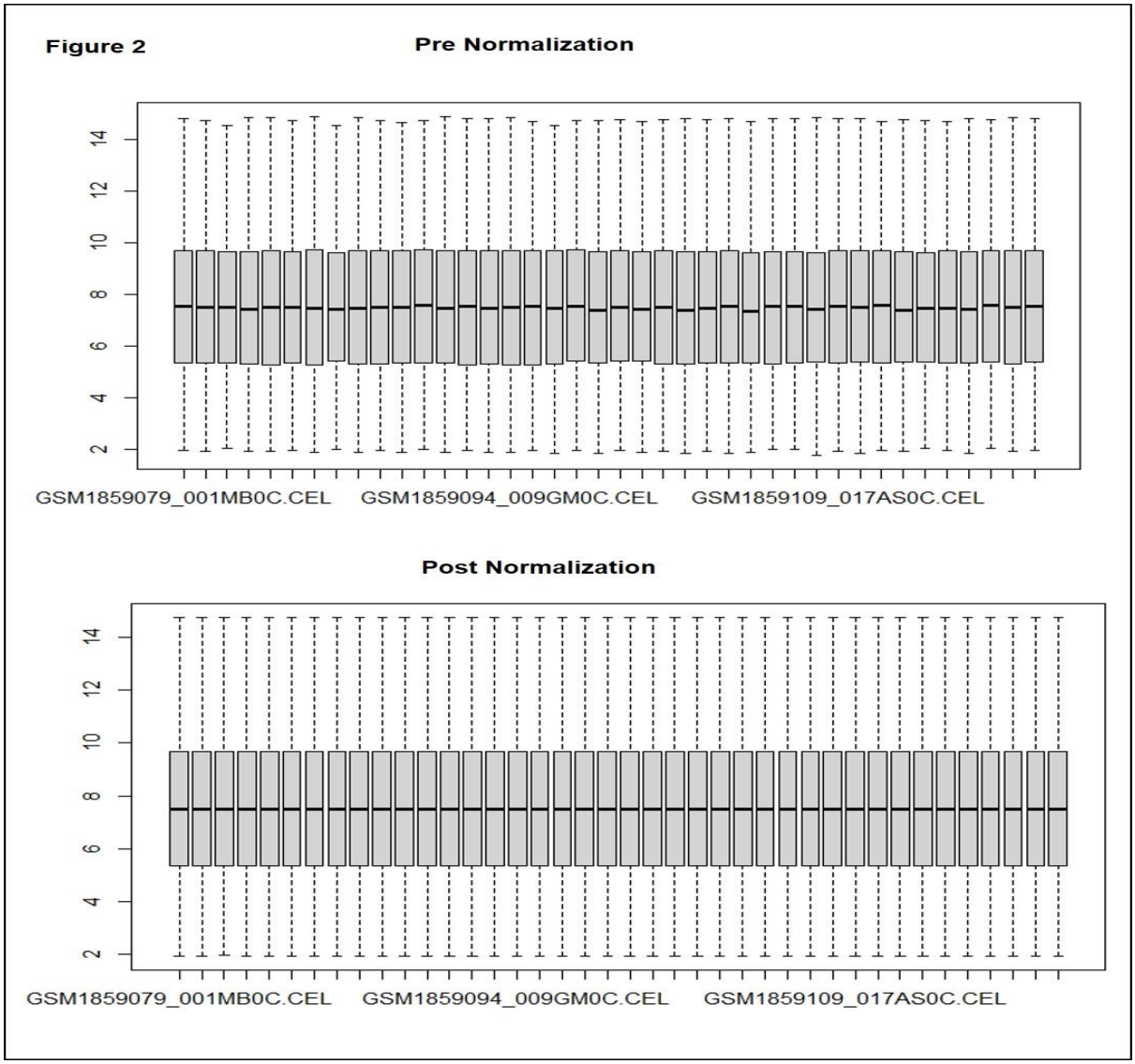


Figure 2: Pre and post normalization of Parkinson's and control samples obtained from GSE72267. normalizeQuantiles function in R studio was used to normalize the data

3.3. Machine learning

Next our aim was to employ different ML algorithms to look for a transcriptomic signature which can differentiate between PD samples and control samples, thus acting as biomarkers. Herein, we used both supervised and unsupervised machine learning techniques which were created using R studio. For supervised machine learning analysis firstly, we created a dataframe of 18 DEGs alongwith their expression signatures for PD and control samples (**figure 3**). Then using caTools and sample.split function we randomly partitioned our dataset into training data and testing data. Out of 40 PD and control samples, 26 samples randomly went into training set and 14 samples were randomly assigned into testset. Training dataset labelled with 18 DEGs

signature was employed to train different ML models. Herein, for supervised ML models we used ANN, random forest, decision tree, LDA and kernel PCA.

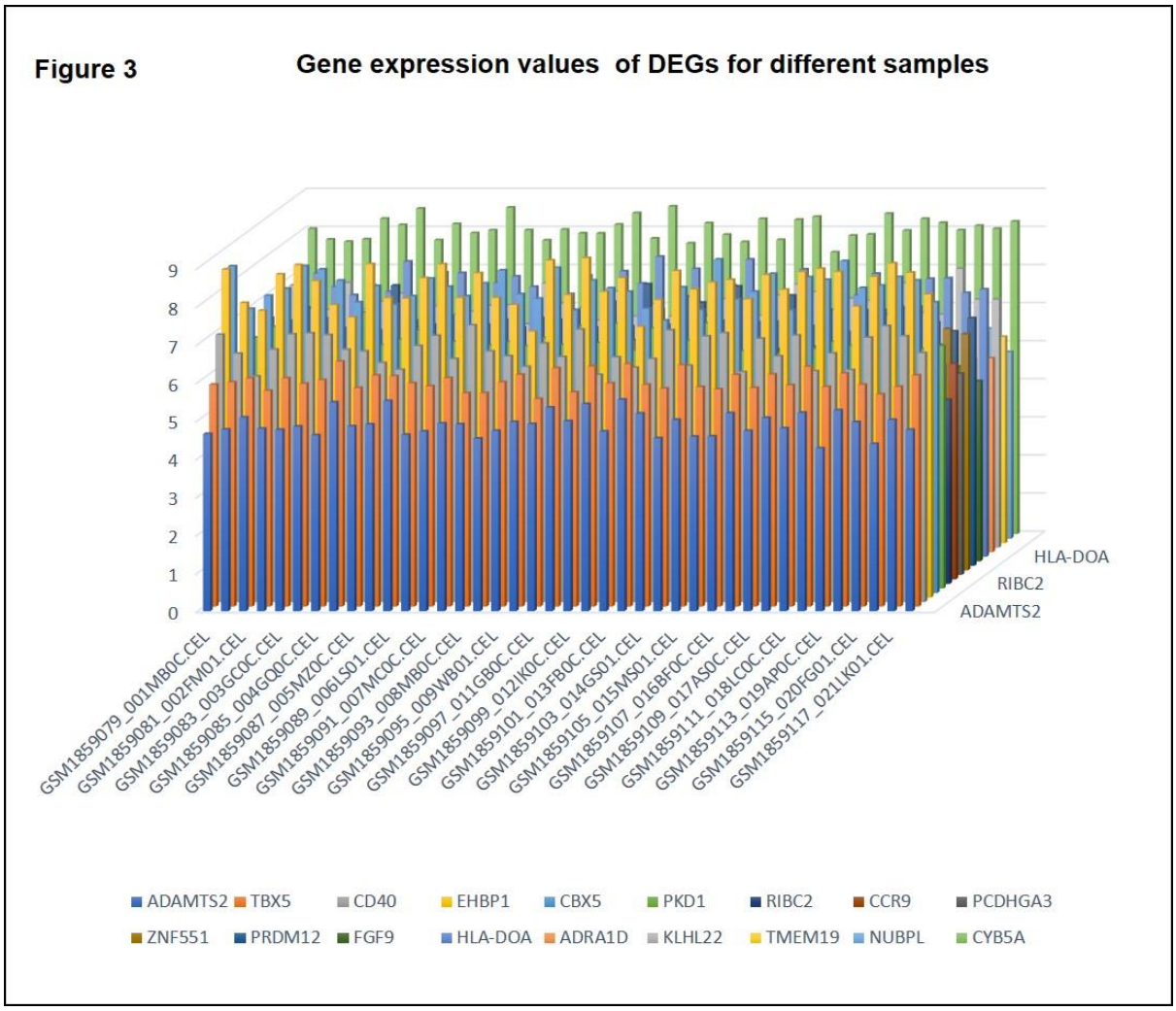


Figure 3: Gene expression values of differentially expressed genes (DEGs) for Parkinson's and control samples. We used exprs function in R studio to obtain these values embedded in the dataset. Y-axis represents their expression values.

Further, ANN, is an ML technique which mimics the neural network of human brain, where billions of neurons work in tandem with each other to process different types of information. Basically, ANN is composed of three layers of neurons working in tandem with each other. The input layer which takes the input, followed by multiple hidden layers where input data is processed, which is fed to the final output layer which gives classification result [104]. Here, we created neural network by using neuralnet function from neuralnet library [105].

Afterwards, we used decision tree machine learning algorithm. Decision tress are flowchart type algorithms where data is repeatedly partitioned by the algorithm based on different frameworks and conditions. Decision trees are made up of two most important entities decision

nodes and leaf nodes. Decision nodes, which have many branches, are used to split data based on certain set of guidelines and leaf nodes are the final output or result [106]. Here, we created decision tree model by using rpart function from the rpart library [107]. Further, we used random forest algorithms for classification. Random forest is an upgraded version of decision trees and it is essentially made up of myriad of decision trees, based on the idea that all these different decision trees when combined, can give more accurate classification result compared to a single decision tree. In random forest different decision trees are trained in different ways and the final classification output is selected by determining most common consensus output of all the decision trees in the algorithm [108]. Herein, we created random forest model by using randomforest function from the randomforest library in R studio.

In addition, we used dimensionality reduction techniques like LDA and Kernel PCA also. LDA and kernel PCA are supervised learning-based dimensionality reduction techniques. The main aim of dimensional reduction techniques is to extract the most optimum features from a dataset and exclude all the redundant features without any loss of information. We created LDA model by using lda function from MASS library and Kernel PCA model by using kpca function from kernlab library [109]. After creating all these models from 18 DEG expression signature, test dataset was used to validate the machine learning models to see if these DEGs expression signature can successfully classify between PD samples and healthy samples.

Furthermore, we used unsupervised machine learning model as well for the analysis of obtained DEGs. In unsupervised machine learning we don't train the model with labelled training dataset instead here the machine learning model works on its own to search for previously unmapped hidden structures and patterns in the given input dataset [62] Herein, for unsupervised machine learning, we used 18 DEGs expression signatures of all 40 PD and healthy control samples as input data, to look for patterns in this dataset. We used hierarchical clustering based unsupervised machine learning to find patterns. Hierarchical clustering model was created by using hclust function then cutree function was used to create clusters.

3.4. Drug repurposing

We decided to use LINCS L1000 based search engine known as L1000CDS^2 (https://maayanlab.cloud/L1000CDS2/#/index), to look for drugs which can reverse the expression of obtained DEGs [76]. Our aim was to shortlist mainly those drugs which have been reported for some neuroprotective properties, using literature survey. L1000CDS^2, is a LINCS 1000 based publicly available search engine which takes upregulated and

downregulated genes as input then look for drugs which can reverse the signature of these genes, based on data collected from various cell line experiments. In addition, we used DrugBank to look for antagonists for those genes which didn't get any significant result from L1000CDS^2 search engine. Further we used publicly available web based tool CoDReS (http://bioinformatics.cing.ac.cy/codres/) to find structurally and functionally active drugs from the list of shortlisted drugs. CoDReS, which stands for composite drug reranking scoring, is publicly available tool to rank drugs, it uses clustering algorithms to suggest most structurally and functionally promising drugs from an input list of drugs, for a particular disease, based on data collected from different sources like DrugBank, BindingDB, ChEMBL among others [110].

3.5 ADME analysis

The shortlisted drugs were further put through ADME analysis using publicly available online tool swissADME (http://www.swissadme.ch/), to ascertain their drug-likeness, lipophilicity, solubility and other physio-chemical properties [111]. ADME stands for adsorption, distribution, metabolism, excretion and it is performed in early stages of drug design and development, with an aim to eliminate those drugs which do not exhibit drug like properties.

3.6 Transcription factor regulatory network

Transcription factors (TFs) are a class of proteins which regulate gene expression by binding and/or unbinding near coding regions of DNA. Hence it is important to determine TFs which are responsible for altered expression during diseased conditions. Herein, we decided to look for TFs responsible for dysregulation genes targeted by shortlisted drugs, especially those drugs which were targeting PTM ezymes, in order to determine gene-drug pathway. Herein, we used NetworkAnalyst (https://www.networkanalyst.ca/) tool in order to determine TF-gene regulatory network. We shortlisted only those TFs having a degree of 3, that is, at least 3 genes were regulated by them.

3.7 Post-translational modification analysis

Further, we decided to explore different post translational modifications (PTMs) on the shortlisted TFs, as some of the shortlisted drugs are inhibitors of various PTM enzymes. The shortlisted drugs can act on those PTM enzymes. We used different computational PTM tools like Deep-PLA (http://deeppla.cancerbio.info/) [112], PLMD (http://plmd.biocuckoo.org/) , and GPS (http://gps.biocuckoo.org/online.php) . We shortlisted PTMs on only those enzymes who showed good high threshold score and low false discovery rate.

3.8 Common drug involved in Parkinson's and ageing

We decided to look for common drugs which can be used to target both PD and ageing. We downloaded GSE106940 dataset from NCBI GEO. This dataset contains gene expression data obtained from tissues of 5 young and 5 old samples. Then we decided to look for differentially expressed genes between young and old samples, to see which genes are dysregulated during ageing. Lastly, our aim was to look for drugs, with the help of LINCS L1000CDS[2], which can target these dysregulated genes, and see if there is any common drug which can be used for both PD and ageing.

3.9 Molecular docking

Lastly, we looked for docking between shortlisted HDAC inhibitor and HDAC6, as HDAC6 has recently been implicated in promoting protein aggregation in PD model, and see their binding affinity. Molecular docking was performed using Pymol and online available tool CB-Dock (http://clab.labshare.cn/cb-dock/php/).

4. Results

4.1 Differential gene expression analysis

Using Limma package in R programming we obtained 18 significant DEGs, with adj.p.value < 0.05. From these 18 DEGs 8 were upregulated and 10 were downregulated in PD conditions. Further, volcanoplot function was used to obtain volcano plot of the shortlisted DEGs (**figure 4**). Further, we used DAVID (https://david.ncifcrf.gov/conversion.jsp) tool to annotate our raw affy gene IDs, with their official gene name and function (**Table 1**). Out of these 18 DEGs, FGF9 and NUBPL have previously been implicated in PD.

Table 1: Obtained differentially expressed genes. We got 8 upregulated and 10 downregulated genes

Affymetrix ID	adj.P.Val	logFC	Gene id	Gene Name
214454_at	0.008264	0.425736	ADAMTS2	ADAM metallopeptidase with trombospondin type 2 motif 1
211886_s_at	0.008264	0.338014	TBX5	T-Box Transcription factor 5
222292_at	0.008264	-0.52801	CD40	CD40 molecule
212650_at	0.013411	-0.62785	EHBP1	EH domain binding protein 1
209715_at	0.023147	-0.50568	CBX5	chromobox 5
216949_s_at	0.03	0.337489	PKD1	Polycystin 1
206526_at	0.03	0.398937	RIBC2	RIB43A domain with coiled coils 2
207445_s_at	0.034967	-0.68645	CCR9	C-C motif chemokine receptor 9
216352_x_at	0.034967	0.358624	PCDHGA3	protocadherin gamma subfamily A, 3
211721_s_at	0.034967	-0.66767	ZNF551	Zinc finger protein 551
220894_x_at	0.034967	0.413379	PRDM12	PR/SET domain 12
206404_at	0.034967	-0.47345	FGF9	Fibroblast growth factor 9
211142_x_at	0.037017	0.34131	HLA-DOA	Major histocompatibility complex, Class II, DOA
210961_s_at	0.041502	0.337381	ADRA1D	Adrenoceptor alpha 1D
222141_at	0.044918	-0.4025	KLHL22	Kelch like family member 22
219941_at	0.044918	-0.50213	TMEM19	Transmembrane protein 19
220176_at	0.045747	-0.3836	NUBPL	Nucleotide binding protein
215726_s_at	0.049254	-0.3402	CYB5A	Cytochrome b5 type A

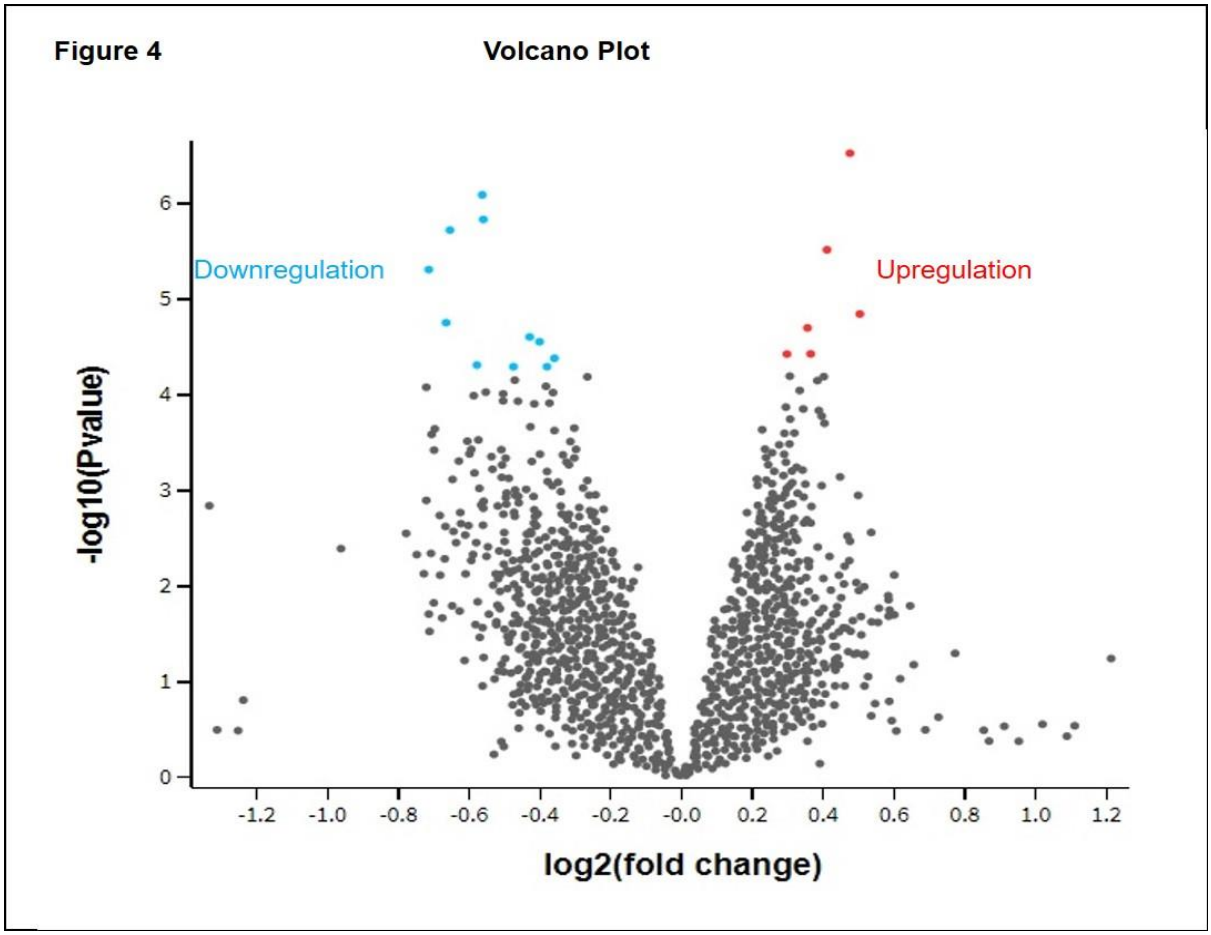


Figure 4: Volcano plot of DEGs, obtained through volcanoplot function in R programming, blue color signifies downregulated genes whereas genes in red color signifies upregulated genes in PD conditions.

4.2 Biomarker validation by Machine learning analysis

Using expression signatures of 18 DEGs we trained and created machine learning models. Further, test data was used to check and validate whether differential gene expression signature can classify between PD samples and healthy samples. Validation with test data set gave us very good results, our machine learning models successfully differentiated between PD samples and healthy samples, proving that gene expression signatures can be used as biomarkers and these genes can serve as therapeutic targets.

Using ANN model **(figure 5**) on test dataset, the model successfully classified PD and healthy samples in the test dataset with 92.8% accuracy. Moreover, with ANN model with achieved sensitivity of 100% , specificity of 85.7%. Furthermore, we used PRROC package to plot AUC curve for ANN model(**figure 6**) [113].

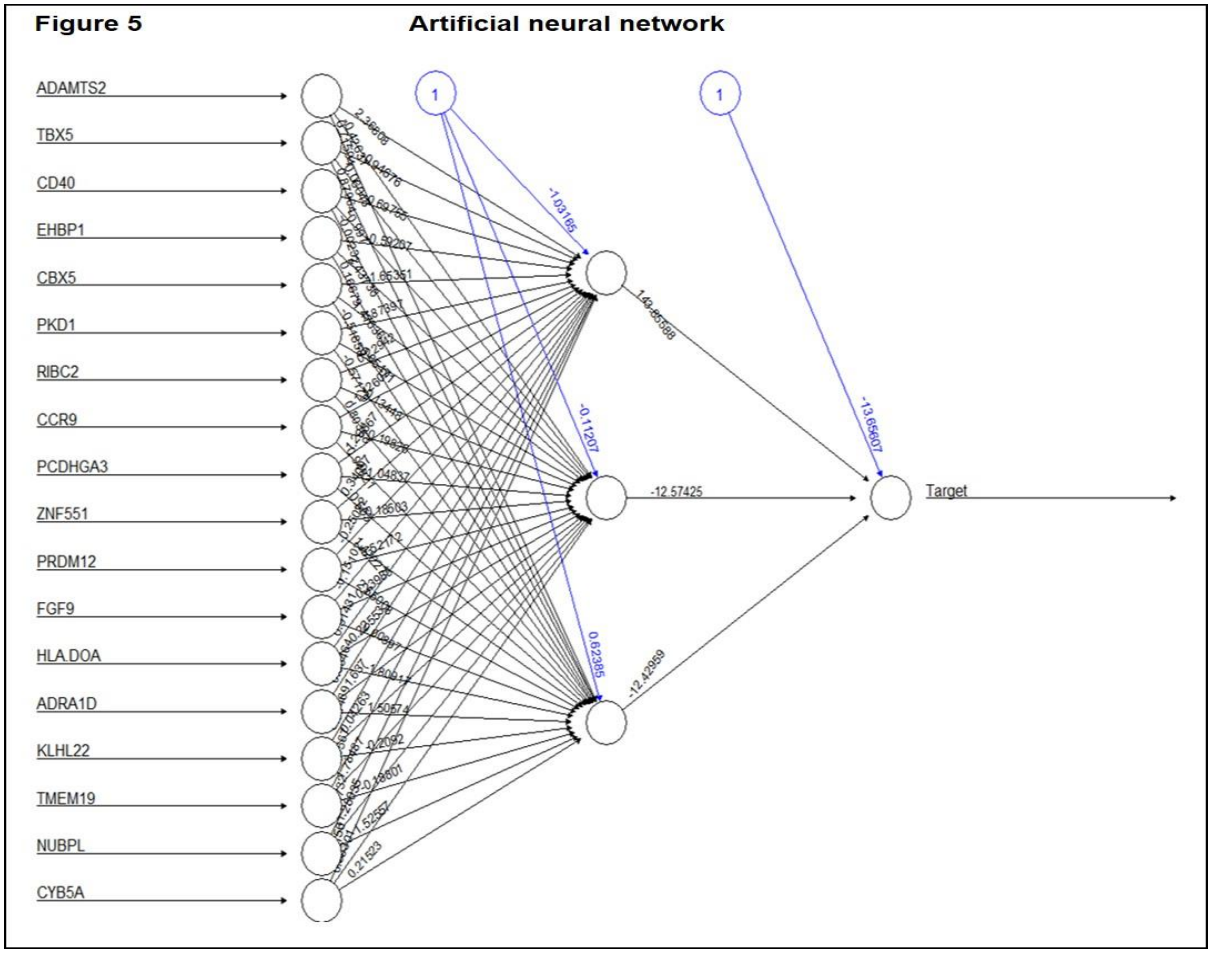


Figure 5: Artificial neural network (ANN) created through neuralnet function in R studio, using DEGs expression signature as input feature

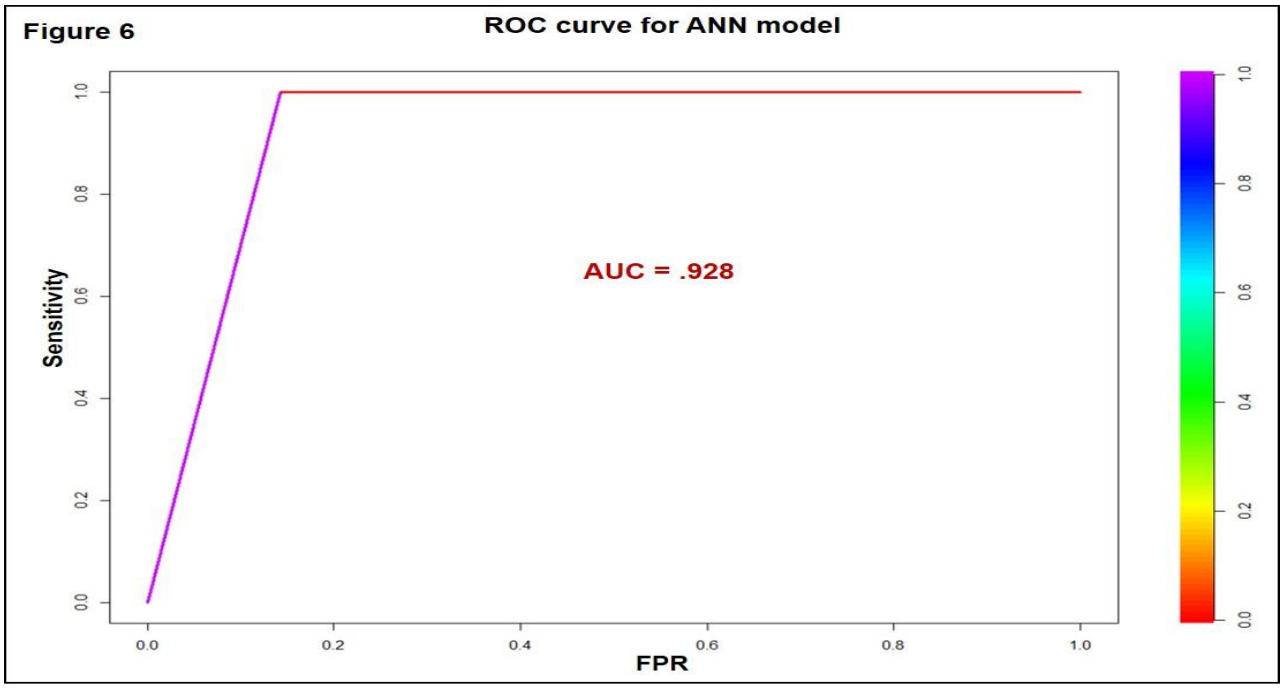


Figure 6: ROC curve for ANN model created using PRROC package in R studio. With ANN model we got

Area Under ROC curve (AUC) of .928, which signifies that our accuracy obtained was 92.8% with ANN model. Further with ANN model we got sensitivity of 100%, specificity of 85.7%.

Further, using decision tree model on test dataset, we got accuracy of 78.5%, sensitivity of 71.4%, specificity of 85.7%. In addition, random forest model which is an upgraded version of decision trees gave us accuracy of 92.8%, specificity of 85.7%, sensitivity of 100%. ROCit package was used to get ROC curve of decision tree and random forest model (**figure 7A and figure 7B respectively**). Moreover, we also used dimensionality reduction techniques like LDA and logarithmic regression-based Kernel PCA. Interestingly, using test dataset on LDA model we got accuracy of 97%, specificity of 100% and even sensitivity of 100%. Whereas with kernel PCA we again got accuracy of 92.8%, specificity of 85.7%, sensitivity of 100%. ROCit package was used to plot ROC curve for LDA model and Kernal PCA model as well. **Table 2** shows accuracy, sensitivity and specificity for different machine learning model we have used.

We created unsupervised machine learning using hierarchical clustering. Herein, 18 differential gene expression signatures were used to create clusters of all 40 samples. Using hierarchical cluster model on gene expression signatures, we found that PD samples and healthy samples were well discriminated and formed well separated clusters. We got clusters of PD and healthy controls with accuracy of 90%, thus providing further evidence that these gene expression signatures can indeed differentiate PD and healthy samples. In addition, we used clusplot function in R studio, to visualize the clusters of PD and healthy samples (**figure 8**)

Figure 7a **ROC curve for decision tree model**

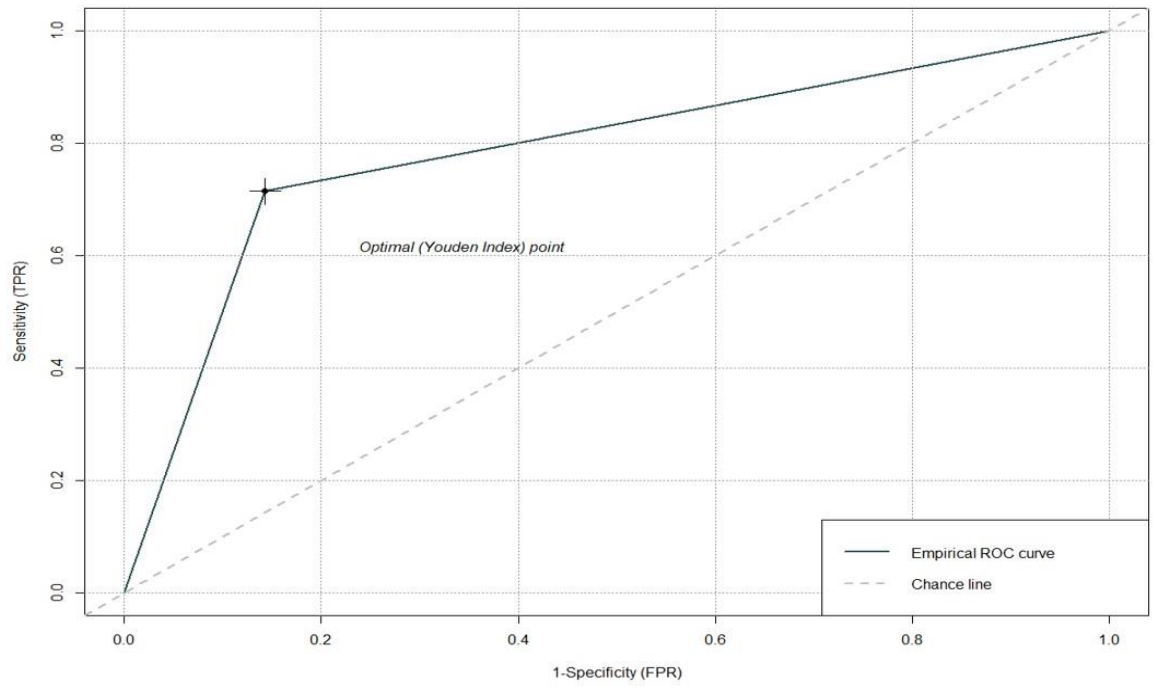


Figure 7b **ROC curve for random forest model**

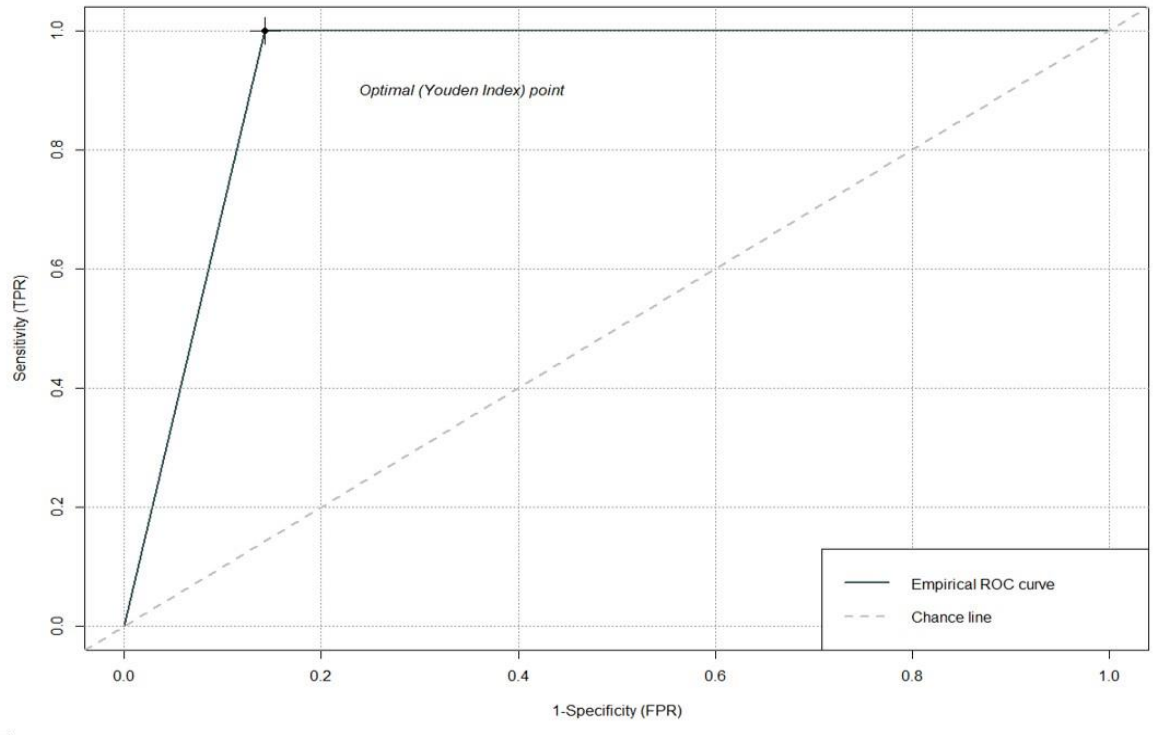


Figure 7a: ROC curve for decision tree model using ROCit package in R studio. In decision tree model we got 78.5% accuracy, sensitivity of 71.4%, specificity of 85.7%. Figure 7b: ROC curve for random forest model obtained using ROCit package in R studio. With random forest model we got accuracy of 92.8%, specificity of 85.7%, sensitivity of 100%.

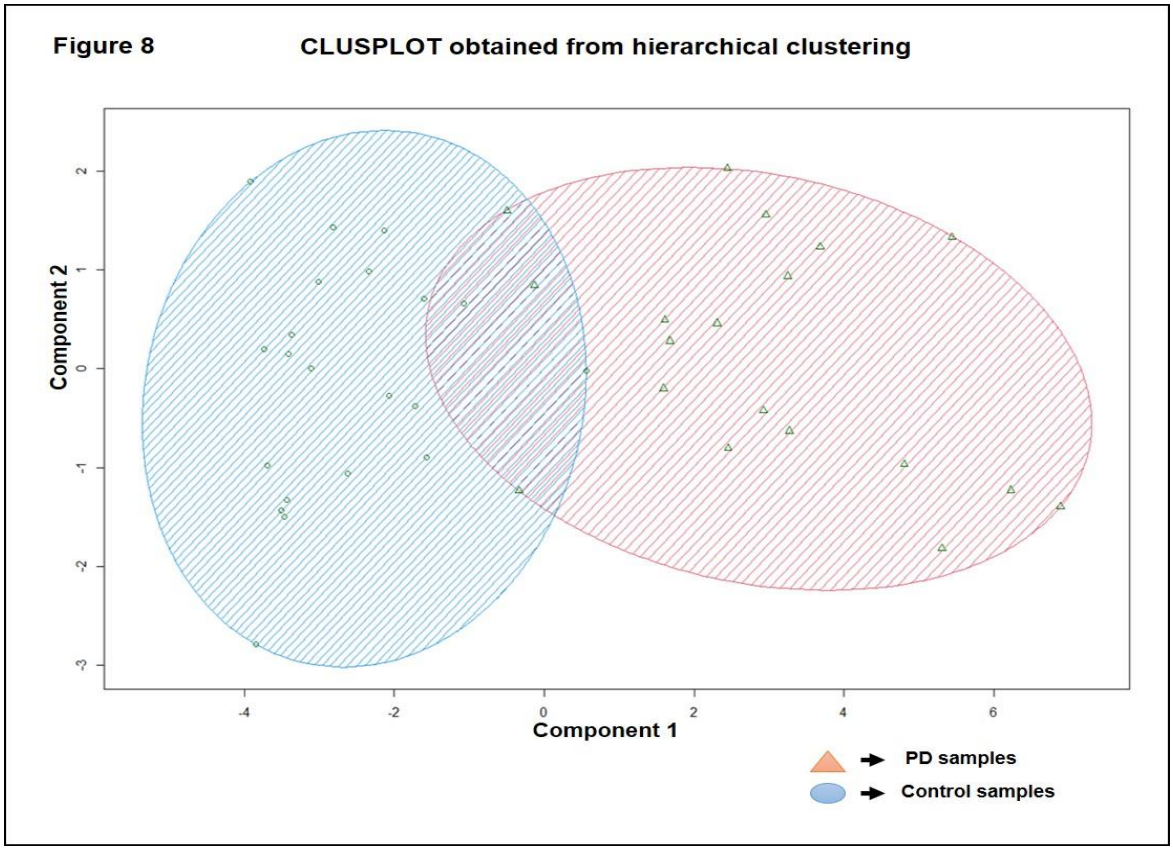


Figure 8: Unsupervised machine learning clusters of PD and controls samples using hierarchical clustering. Clusplot function in R studio was used to create these clusters. Herein, gene expression signatures of 18 DEGs were used to analyze 40 PD and control samples through hierarchical clustering. PD and control samples were clustered in two separate clusters with 90% accuracy

Table 2: Machine learning analysis results for different machine learning models

Machine Learning model	Accuracy	Sensitivity	Specificity
Artificial Neural Network	92.8%	100%	85.7%
Random Forest	92.8%	100%	85.7%
Decision trees	78.5%	71.4%	85.7%
Linear Discriminant Analysis	97%	100%	100%
Kernel PCA	92.8%	100%	85.7%

4.3 Novel drug hits through drug repurposing

Using L1000CDS[2] search engine we found those drugs which can reverse the expression of these obtained DEGs. Further using literature survey, we shortlisted those drugs which have been reported to exhibit some neuroprotective properties (**table 3**). ADRA1D was one of the upregulated genes in our result, but it didn't show any significant result from L1000CDS[2]

search engine, so we used DrugBank to look for its antagonist and selected those antagonists with some neuroprotective properties. Using L1000CDS^2 search engine we shortlisted several immunosupressants, HDAC inhibitors, CDK inhibitors, anti-inflammatory corticosteroids, tyrosine kinase inhibitor, calcium blockers, alpha blocker, K^+ATP channel opening vasodilator with neuroprotective properties.

Table 3: Shortlisted drugs obtained from L1000CDS^2 tool

Drug Name	Target gene	Drug function
Sirolimus	FGF9	Immunosupressant
Mitoxantrone	FGF9	Immunosupressant
Mycophenolate Mofetil	FGF9	immunosupressant
Tamoxifen Citrate	FGF9	Estrogen receptor modulator
Roscovitine	CYB5A	CDK inhibitor
Alvocidib	EHBP1	CDK inhibitor
Vorinostat	FGF9, CYB5A	HDAC inhibitor
Scriptaid	FGF9, CYB5A	HDAC inhibitor
HDAC6 inhibitor ISOX	FGF9	HDAC6 inhibitor
Etinostat	EHBP1	Benzamide HDAC inhibitor
Pracinostat	EHBP1	HDAC inhibitor
Cilnidipine	CD40	Calcium channel blocker
Nicardipine	ADRA1D	Calcium channel blocker
Afatinib	ZNF551	Tyrosine kinase inhibitor
Rosuvastatin	CBX5	HMG-CoA reductase inhibitor
Dexamethasone	CYB5A	Anti-inflammatory corticosteroid
Betamethasone	CYB5A	Anti-inflammatory corticosteroid
Gugglesterone	CYB5A	Phytosterol
Estradiol Valrate	CD40	Estrogen receptor agonist
Terazosin	ADRA1D	Alpha blocker
Levcromakalim	CD40	K^+ATP channel opening vasodilator

Further, we used CoDReS tool and uploaded the shortlisted drugs to get structurally and functionally favorable drugs for PD treatment. Using its clustering algorithm CoDRes tool gave dexamethasone, sirolimus, afatinib, nicardipine, scriptaid, rosuvastatin as the most promising drugs which are worth further investigation for PD treatment. Previously, it has been shown that dexamethasone protects against neuroinflammation by inhibiting nox-2 dependent overproduction of reactive oxygen species (ROS) [114]. Likewise, it has been shown that sirolimus can protect against cognitive dysfunction in mice model of PD, by inhibition of mTORC1 [115]. Further afatinib is a tyrosine kinase inhibitor used of treatment of lung carcinoma. Interestingly, it has been reported that afatinib can protect against neuroinflammation by thwarting oxygen/glucose deprivation (OGD) driven activation of

astrocytes and inflammasome [116]. Additionally, it has been observed that Rosuvastatin , a statin drug targeting CBX5 gene, protects against rotenone induced toxicity in PD models, by enhancing autophagy [117]. Scriptaid is an HDAC inhibitor, which was observed to have a protective effect in traumatic brain injury by promoting AKT pathway activation and thwarting AKT inhibition by PTEN [118]. Nicardipine is antagonist of ADRA1D and it has been reported that nicarpidine can protect brain during hypertension induced cognitive damage [119]. Hence, we can see that all these six drugs exhibit neuroprotective properties and are worth further investigation for PD treatment.

4.4 ADME analysis of shortlisted drugs

The proposed drug from CoDReS tool, dexamethasone, sirolimus, afatinib, nicardipine, scriptaid, rosuvastatin, vorinostat were analyzed through swissADME in order to ascertain their physiochemical properties, pharmacokinetics and drug likeness. Of all the shortlisted drugs only sirolimus showed one violation of Lipinski rule for druglikeness, which was its molecular weight > 500 g/mol. All other drugs didn't show any violation for druglikeness. The result obtained is shown in the **table 4**.

Table 4: ADME analysis of exemplar drugs suggested by CoDReS tool

Drug	Molecular weight	Druglikeness (Lipinski Rule)	GI absorption	CYP1A2 inhibitor	CYP2C19 inhibitor	CYP2C9 inhibitor	CYP2D6 inhibitor
Dexamethasone	392.46 g/mol	Yes	High	No	No	No	No
Sirolimus	914.17 g/mol	No	Low	No	No	No	No
Afatinib	485.94 g/mol	Yes	High	No	Yes	Yes	Yes
Nicardipine	479.52 g/mol	Yes	High	No	Yes	Yes	Yes
Scriptaid	326.35 g/mol	Yes	High	No	No	No	No
Rosuvastatin	481.54 g/mol	Yes	Low	No	No	No	No
Vorinostat	264.32 g/mol	Yes	High	No	No	No	No

4.5 Transcription factors- gene regulatory network

We decided to look for TFs responsible for dysregulation of genes targeted by the shortlisted drugs, in order to determine drug-gene pathway. Only TFs targeting regulating atleast 3 genes were shortlisted. We used Chea TF database present in NetworkAnalyst tool to obtain the TF-

gene network (**figure 9**). Obtained TFs is shown in **table 5**. SIN3B, SPI1, BMI1 were TFs regulating both FGF9, CYB5A. Whereas HNF4A, FLI1, Sox2 were TFs regulating ZNF551. In addition, POU5F1 and MTF2 were TFs regulating FGF9 only.

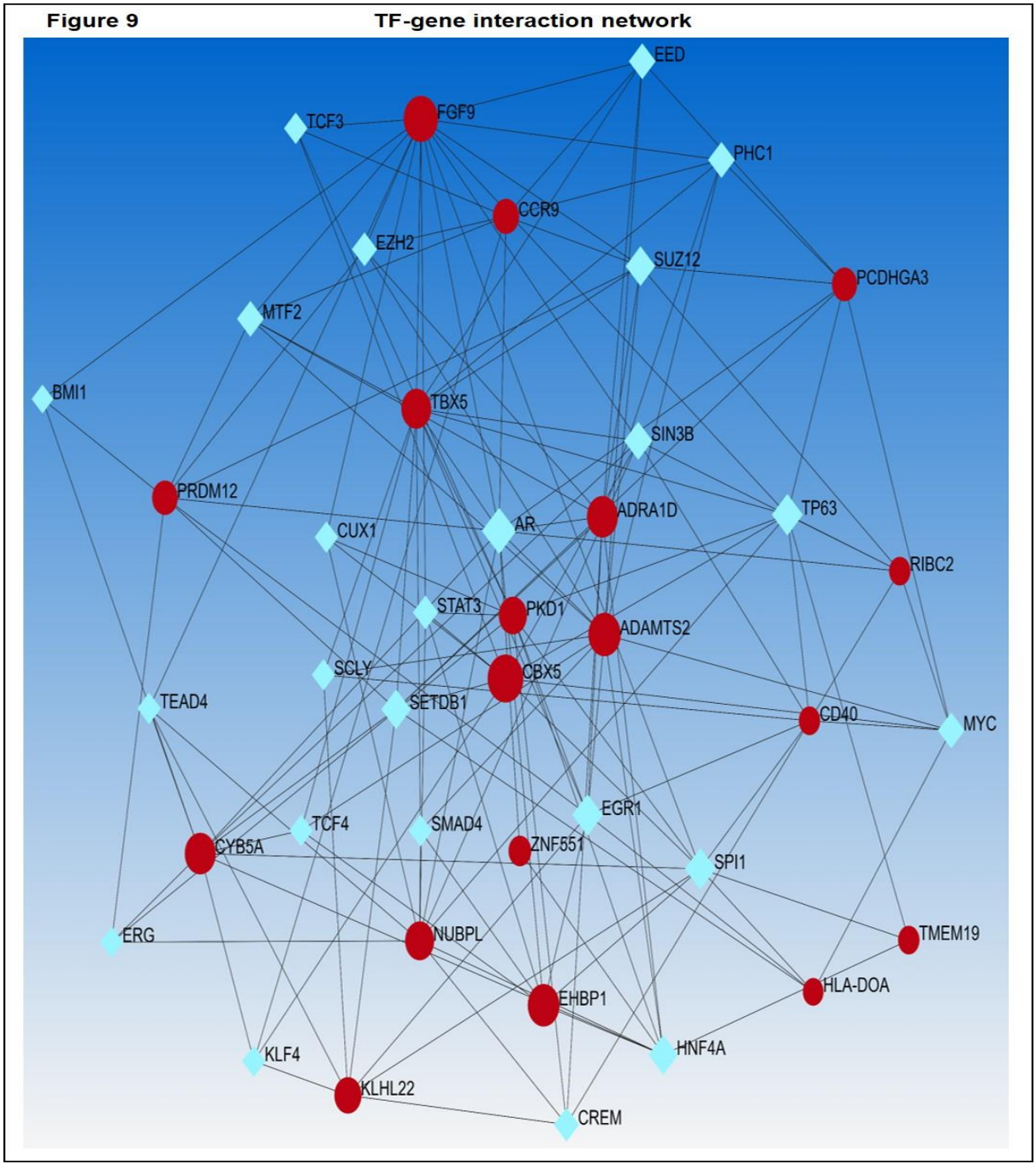


Figure 9: Transcription factor-gene interaction network obtained from chea database in NetworkAnalyst tool

Table 5: Transcription factors regulating genes targeted by shortlisted drugs, along with their description and implicated biological process

Transcription Factor	Overlapping genes	Description
SIN3B	FGF9, CYB5A	SIN3 Transcription Regulator Family Member B
BMI-1	FGF9, CYB5A	BMI1 Polycomb Ring Finger Oncogene
SPI-1	FGF9, CYB5A	Spi-1 Proto-Oncogene
POU5F1	FGF9	POU Domain, Class 5, Transcription Factor 1
MTF2	FGF9	Metal-Response Element-Binding Transcription Factor 2
HNF4A	ZNF551	Hepatocyte Nuclear Factor 4 Alpha
Sox2	ZNF551	SRY-Box Transcription Factor 2
FLI1	ZNF551	Fli-1 Proto-Oncogene

4.6 PTMs on Transcription factors

Further, we decided to look for PTMs on shortlisted TFs. As most of our shortlisted drugs are HDAC and kinase inhibitors, we decided to look for HDAC sites on lysine residues and phosphorylation sites on serine/threonine, tyrosine residues. Vorinostat and Scriptaid are HDAC inhibitors upregulating FGF9, CYB5A expression. SIN3B, BMI-1, SPI-1 were the TFs regulating both FGF9, CYB5A. We used Deep-PLA tool to look for HDAC sites on these TFs. We shortlisted only those sites which showed high threshold and FPR less than 1%. Only SIN3B and BMI-1 showed HDAC active sites with high threshold and low FDR. HDAC1 interacts with SIN3B on K797 site (**figure 10a**), whereas HDAC6 interacts with BMI-1 on K314 (**figure 10b**). Further, Sirolimus is an mTOR inhibitor which is a serine/threonine kinase. Sirolimus upregulates FGF9 gene, and POU5F1, MTF2 are TFs regulating FGF9 only. Using GPS tool for phosphorylation, we found that MTF2 showed good score for phosphorylation by mTOR at T24. Likewise, afatinib is a tyrosine kinase upregulating ZNF551. Using, same GPS tool, TF HNF4A showed good score for getting phosphorylated by Janus kinase: JAK, at Y286. **Figure 11** and **12** shows the proposed Drug-gene pathway.

Figure 10a **SIN3B PTM sites**

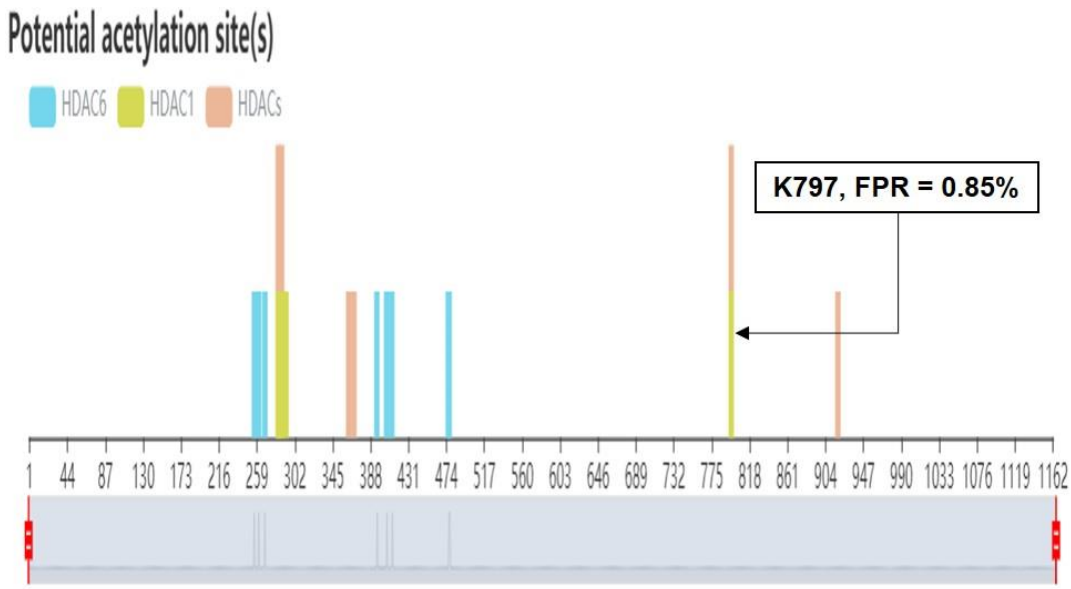


Figure 10b **BMI-1 PTM sites**

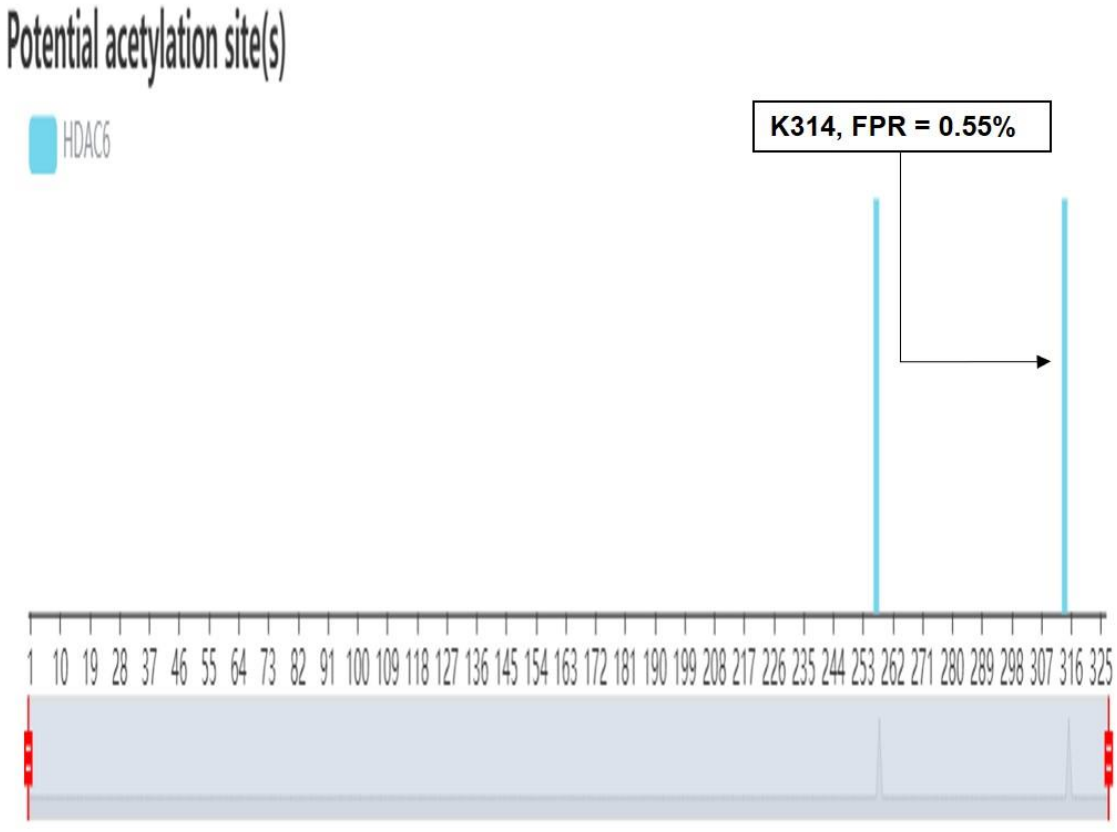


Figure 10a: HDAC PTM sites on SIN3B found using Deep-Pla tool. Figure 10b: HDAC PTM sites on BMI-1 found using Deep-pla tool.

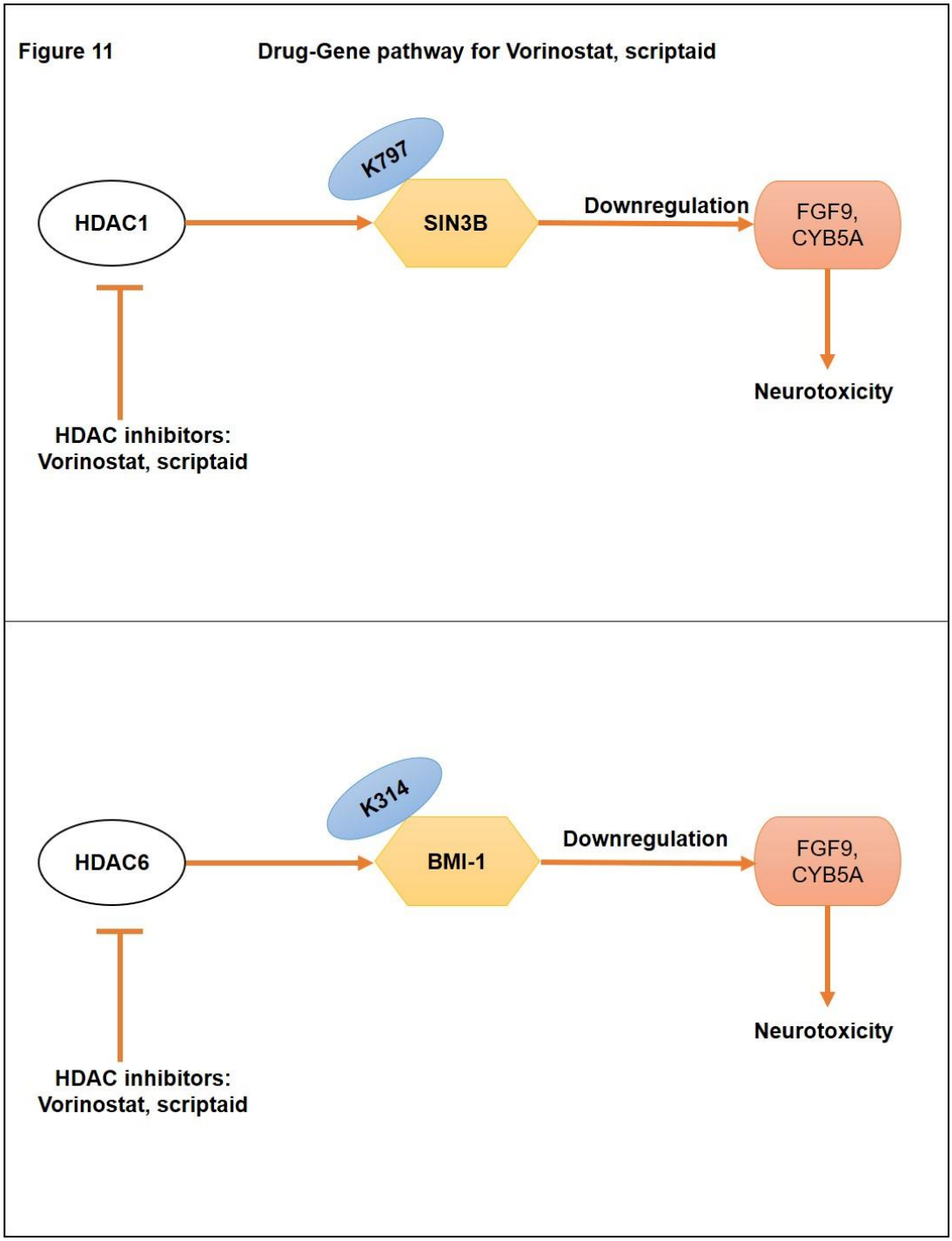


Figure 11: Drug gene pathway for vorinostat and scriptaid

Figure 12a Drug gene pathway for sirolimus

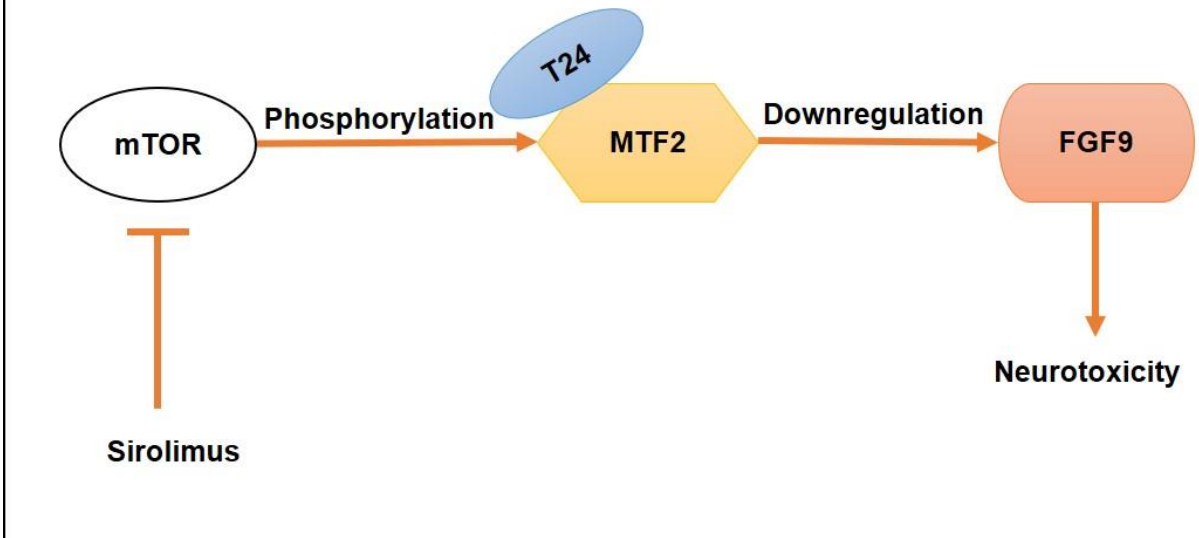


Figure 12b Drug gene pathway for afatinib

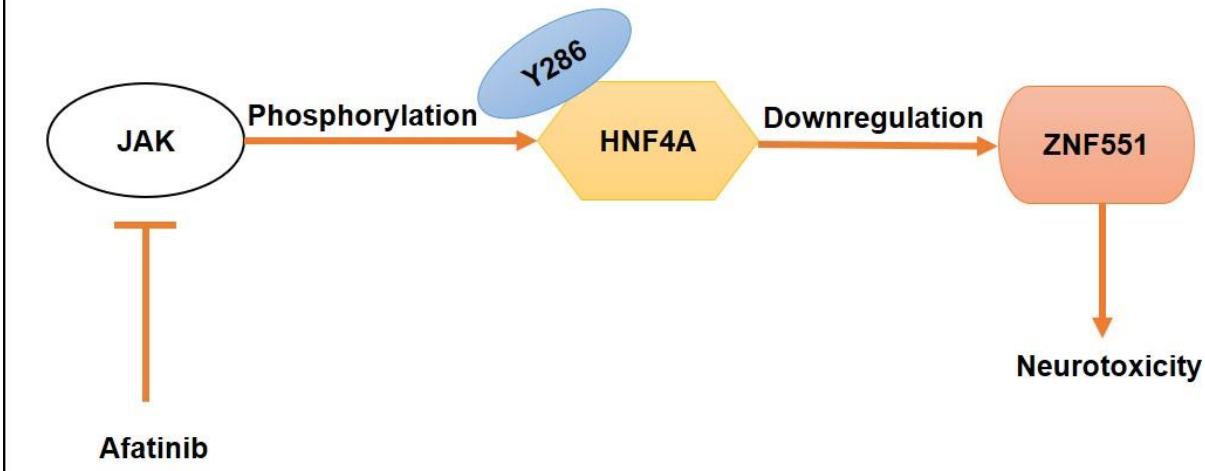


Figure 12a: Drug gene pathway for sirolimus. Figure 12b: Drug gene pathway for afatinib

4.7 Common drug involved in ageing and neurodegeneration

We found the DEGs between old and young samples, with the aim of finding any common drug from the above shortlisted drug. DEGs were shortlisted based on adjusted p value being less 0.05 (**table 6)**. We looked for the functions of shortlisted genes. Then we used LINCS L1000CDS2 to look for the drug targeting TLK1 as it was involved in telomere shortening, very characteristic feature of ageing and neurodegeneration. Strikingly, we found out that vorinistat is a drug which ameliorates the dysregulation of TLK1. Vorinostat is also one of the shortlisted drugs. Hence, Vorinostat is a drug which can ameliorate both PD and ageing.

Table 6: Differentially expressed genes between ageing and young samples

ID	adj.P.Val	P.Value	logFC	Symbol	Function
1554089_s_at	0.048	1.97E-06	-0.848	SBDSP1	Long non-coding RNA
1554595_at	0.048	4.69E-06	0.93	SYMPK	mRNA splicing
214495_at	0.048	6.67E-06	2.452	CACNG2	Voltage-gated ion channel activity
222966_at	0.048	6.55E-06	2.74	TLK1	Chromatin remodelling

4.8 Docking of vorinostat with HDAC6 to prevent α-synuclein aggregation

Recently it researchers found out that HDAC6 promotes α-synuclein toxicity and aggregation in Parkinson's rat model and inhibition of HDAC6 thwarts α-synuclein aggregation by increasing its acetylation status and promoting its recognition by proteasome system. Hence, we decided to check whether vorinostat shows good binding with HDAC6 via molecular docking approach. Docking result through CB Dock showed good binding score. A10, V9, P63, L8 residues take part in binding of receptor with ligand as shown in the **figure 13.**

Figure 13 Docking of vorinostat with HDAC6

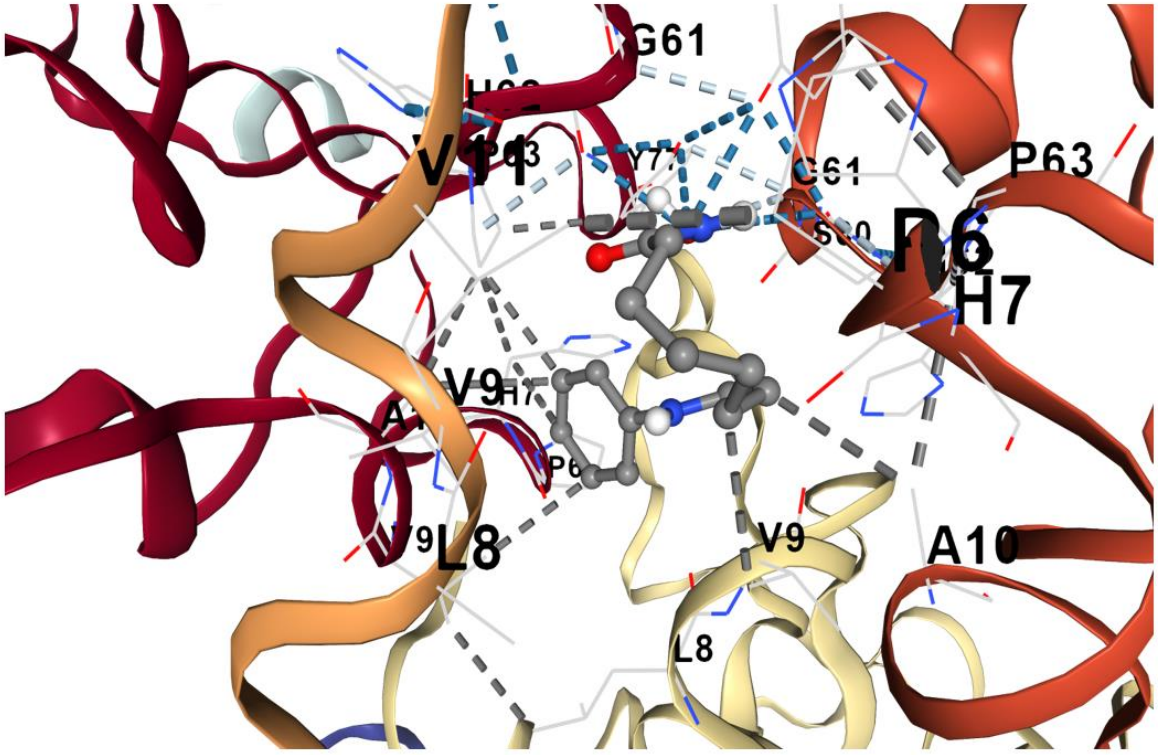


Figure 13: Docking result of vorinostat with HDAC6 obtained with the help of PyMol and CB Dock tool. A10, V9, P63, L8 residues are involved in binding of receptor with ligand as shown.

5. Discussion and conclusion

NDDs are a big burden on society, especially on ageing population. Most of the drug being used for NDD treatment can only relieve the symptoms but can't stop their progress. There is a big necessity to explore drugs which can somewhat impede and thwart the progress of the disease. In this regard, AI and/or ML driven techniques are widely being used now for disease characterization because of their extraordinary ability to find hidden and convoluted relationship among different biomedical datasets. Microarray techniques are able to find out dysregulated mRNAs within the genome in different diseased conditions. Herein, our main aim was to find a blood transcriptomic gene set which can identify between PD samples and healthy samples, using different ML algorithms. Herein, we downloaded microarray blood transcriptome dataset from NCBI GEO and used it to find differentially expressed genes between PD patient samples and healthy control samples. Further using supervised, unsupervised ML models we evaluated whether these DEGs and their expression level can differentiate PD samples from healthy samples. Moreover, using different ML algorithms on test dataset, our DEGs successfully differentiated PD samples from healthy samples. Hence, according to our analysis the expression level of these 18 DEGs can act as biomarker and help in identification of PD patients. Furthermore, these DEGs can even serve as potential therapeutic target in PD. Out of 18 DEGs, two DEGs have previously been reported for PD pathogenesis, like, FGF9 has been observed to be downregulated during MPP^+ induced neurotoxicity in PD model and it was observed here that upregulation of FGF9 can protect against neuronal cell death [120]. Likewise, loss of function mutation of NUBPL has been linked to PD [121].

Further, using L1NC data based $L1000CDS^2$ search engine we found drugs which can turn-around the expression of obtained differentially expressed genes. Afterwards, using published literature survey, we shortlisted only those drugs which have been reported for their neuroprotective properties. The shortlisted drugs were submitted CoDReS tool to obtain most promising set of drugs which are worth further investigation for PD treatment. Dexamethasone, sirolimus, afatinib, scriptaid, rosuvastatin, nicardipine, vorinostat were the most biologically promising drug suggested by CoDReS tool. All these drugs have been reported for their neuroprotective properties. Only dexamethasone showed one violation of lipnski rule for druglikeness during ADME analysis, remaining drugs didn't show any violation for druglikeness. Further, we used ChEA tool to obtained top TFs which regulate the expression of these DEGs targeted by the drugs. Most of the obtained TFs were implicated in nervous

system related biological processes. In addition, we found PTMs on carious sites of these TFs, with potential involvement in drug-gene pathway. We also found common drug which can be used for treating both ageing and Parkinson's symptoms.

In summary, our analysis provides further evidence dysregulation of mRNA in blood can lead to diseased conditions in the brain. We identified a blood transcriptomic set capable of identifying Pd patients from healthy controls. In addition, using drug repurposing approach we found some new drugs which can deaccelerate progress of PD. Moreover, ML analysis will further benefit from a bigger dataset and overall results can be confirmed with wet-lab experiments. Although in silico tools have given us a plethora of target genes, proteins and drugs which can interact with those targets, very few drugs have actually made it for public use during different disorders. In silico tools have big potential and they are expected to get more accurate in near future, but their validity remains a big issue as of now, which will have to be worked upon.

6. References

[1] J.D. Guo, X. Zhao, Y. Li, G.R. Li, X.L. Liu, Damage to dopaminergic neurons by oxidative stress in Parkinson's disease (Review), Int. J. Mol. Med. 41 (2018) 1817–1825. https://doi.org/10.3892/ijmm.2018.3406.

[2] E. Gökçal, V.E. Gür, R. Selvitop, G. Babacan Yildiz, T. Asil, Motor and non-motor symptoms in parkinson's disease: Effects on quality of life, Noropsikiyatri Ars. 54 (2017) 143–148. https://doi.org/10.5152/npa.2016.12758.

[3] A.Q. Rana, U.S. Ahmed, Z.M. Chaudry, S. Vasan, Parkinson's disease: A review of non-motor symptoms, Expert Rev. Neurother. 15 (2015) 549–562. https://doi.org/10.1586/14737175.2015.1038244.

[4] W. Poewe, Non-motor symptoms in Parkinson's disease, Eur. J. Neurol. 15 (2008) 14–20. https://doi.org/10.1111/j.1468-1331.2008.02056.x.

[5] Y. Wu, W. Le, J. Jankovic, Preclinical biomarkers of parkinson disease, Arch. Neurol. 68 (2011) 22–30. https://doi.org/10.1001/archneurol.2010.321.

[6] T. Lee, H. Lee, Prediction of Alzheimer's disease using blood gene expression data, Sci. Rep. 10 (2020) 1–13. https://doi.org/10.1038/s41598-020-60595-1.

[7] S.H. Kim, I.B. Kim, D.H. Oh, D.H. Ahn, Predicting autism spectrum disorder using blood-based gene expression signatures and machine learning, Eur. Neuropsychopharmacol. 27 (2017) S1090. https://doi.org/10.1016/s0924-977x(17)31894-1.

[8] A. Lau, H.C. So, Turning genome-wide association study findings into opportunities for drug repositioning, Comput. Struct. Biotechnol. J. 18 (2020) 1639–1650. https://doi.org/10.1016/j.csbj.2020.06.015.

[9] T. Uenaka, W. Satake, P.C. Cha, H. Hayakawa, K. Baba, S. Jiang, K. Kobayashi, M. Kanagawa, Y. Okada, H. Mochizuki, T. Toda, In silico drug screening by using genome-wide association study data repurposed dabrafenib, an anti-melanoma drug, for Parkinson's disease, Hum. Mol. Genet. 27 (2018) 3974–3985. https://doi.org/10.1093/hmg/ddy279.

[10] D.M. Skovronsky, V.M.Y. Lee, J.Q. Trojanowski, Neurodegenerative diseases: New concepts of pathogenesis and their therapeutic implications, Annu. Rev. Pathol. 1

(2006) 151–170. https://doi.org/10.1146/annurev.pathol.1.110304.100113.

[11] D. Chhangani, A. Mishra, Protein quality control system in neurodegeneration: A healing company hard to beat but failure is fatal, Mol. Neurobiol. 48 (2013) 141–156. https://doi.org/10.1007/s12035-013-8411-0.

[12] S. Vijayakumaran, D.L. Pountney, SUMOylation, aging and autophagy in neurodegeneration, Neurotoxicology. (2018). https://doi.org/10.1016/j.neuro.2018.02.015.

[13] U. Bandhyopadhyay, A.M. Cuervo, Chaperone-mediated autophagy in aging and neurodegeneration: Lessons from α-synuclein, Exp. Gerontol. 42 (2007) 120–128. https://doi.org/10.1016/j.exger.2006.05.019.

[14] A. Grimm, A. Eckert, Brain aging and neurodegeneration: from a mitochondrial point of view, J. Neurochem. 143 (2017) 418–431. https://doi.org/10.1111/jnc.14037.

[15] C. Soto, S. Pritzkow, Protein misfolding, aggregation, and conformational strains in neurodegenerative diseases, Nat. Neurosci. 21 (2018) 1332–1340. https://doi.org/10.1038/s41593-018-0235-9.

[16] J. Weller, A. Budson, Current understanding of Alzheimer's disease diagnosis and treatment, F1000Research. 7 (2018) 1–9. https://doi.org/10.12688/f1000research.14506.1.

[17] R. Gupta, M. Sahu, D. Srivastava, S. Tiwari, R.K. Ambasta, P. Kumar, Post-translational modifications: Regulators of neurodegenerative proteinopathies, Ageing Res. Rev. 68 (2021) 101336. https://doi.org/10.1016/j.arr.2021.101336.

[18] T. Pan, S. Kondo, W. Le, J. Jankovic, The role of autophagy-lysosome pathway in neurodegeneration associated with Parkinson's disease, Brain. 131 (2008) 1969–1978. https://doi.org/10.1093/brain/awm318.

[19] R. Balestrino, A.H.V. Schapira, Parkinson disease, Eur. J. Neurol. 27 (2020) 27–42. https://doi.org/10.1111/ene.14108.

[20] J.L. Webb, B. Ravikumar, J. Atkins, J.N. Skepper, D.C. Rubinsztein, α-synuclein Is Degraded by Both Autophagy and the Proteasome, J. Biol. Chem. 278 (2003) 25009–25013. https://doi.org/10.1074/jbc.M300227200.

[21] A.A. Cooper, A.D. Gitler, A. Cashikar, C.M. Haynes, K.J. Hill, B. Bhullar, K. Liu, K. Xu, K.E. Strathearn, F. Liu, S. Cao, K.A. Caldwell, G.A. Caldwell, G. Marsischky, R.D. Kolodner, J. LaBaer, J.C. Rochet, N.M. Bonini, S. Lindquist, α-synuclein blocks ER-Golgi traffic and Rab1 rescues neuron loss in Parkinson's models, Science (80-.). 313 (2006) 324–328. https://doi.org/10.1126/science.1129462.

[22] K. Senkevich, Z. Gan-Or, Autophagy lysosomal pathway dysfunction in Parkinson's disease; evidence from human genetics, Park. Relat. Disord. (2019). https://doi.org/10.1016/j.parkreldis.2019.11.015.

[23] N. Tambasco, M. Romoli, P. Calabresi, Levodopa in Parkinson's Disease: Current Status and Future Developments, Curr. Neuropharmacol. 16 (2017) 1239–1252. https://doi.org/10.2174/1570159x15666170510143821.

[24] P. Riederer, G. Laux, MAO-inhibitors in Parkinson's Disease, Exp. Neurobiol. 20 (2011) 1–17. https://doi.org/10.5607/en.2011.20.1.1.

[25] A. Prasad, V. Bharathi, V. Sivalingam, A. Girdhar, B.K. Patel, Molecular mechanisms of TDP-43 misfolding and pathology in amyotrophic lateral sclerosis, Front. Mol. Neurosci. 12 (2019). https://doi.org/10.3389/fnmol.2019.00025.

[26] A. Girdhar, V. Bharathi, V.R. Tiwari, S. Abhishek, W. Deeksha, U.S. Mahawar, G. Raju, S.K. Singh, G. Prabusankar, E. Rajakumara, B.K. Patel, Computational insights into mechanism of AIM4-mediated inhibition of aggregation of TDP-43 protein implicated in ALS and evidence for in vitro inhibition of liquid-liquid phase separation (LLPS) of TDP-432C-A315T by AIM4, Int. J. Biol. Macromol. 147 (2020) 117–130. https://doi.org/10.1016/j.ijbiomac.2020.01.032.

[27] R.A.C. Roos, Huntington's disease: A clinical review, Orphanet J. Rare Dis. 5 (2010) 2–9. https://doi.org/10.1186/1750-1172-5-40.

[28] L. Wan, K. Xu, Z. Chen, B. Tang, H. Jiang, Roles of Post-translational Modifications in Spinocerebellar Ataxias, Front. Cell. Neurosci. 12 (2018). https://doi.org/10.3389/fncel.2018.00290.

[29] B. Verhalen, S. Arnold, B.A. Minassian, Lafora Disease: A Review of Molecular Mechanisms and Pathology, Neuropediatrics. 49 (2018) 357–362. https://doi.org/10.1055/s-0038-1675238.

[30] N.N. Nalivaeva, A.J. Turner, Post-translational modifications of proteins: Acetylcholinesterase as a model system, Proteomics. 1 (2001) 735–747. https://doi.org/10.1002/1615-9861(200106)1:6<735::AID-PROT735>3.0.CO;2-8.

[31] J. Avila, Tau phosphorylation and aggregation in Alzheimer's disease pathology, FEBS Lett. 580 (2006) 2922–2927. https://doi.org/10.1016/j.febslet.2006.02.067.

[32] T.J. Cohen, A.W. Hwang, C.R. Restrepo, C.X. Yuan, J.Q. Trojanowski, V.M.Y. Lee, An acetylation switch controls TDP-43 function and aggregation propensity, Nat. Commun. 6 (2015) 1–13. https://doi.org/10.1038/ncomms6845.

[33] R. Bhattacharyya, C. Barren, D.M. Kovacs, Palmitoylation of amyloid precursor protein regulates amyloidogenic processing in lipid rafts, J. Neurosci. (2013). https://doi.org/10.1523/JNEUROSCI.4704-12.2013.

[34] L. McGurk, O.M. Rifai, N.M. Bonini, Poly(ADP-Ribosylation) in Age-Related Neurological Disease, Trends Genet. (2019). https://doi.org/10.1016/j.tig.2019.05.004.

[35] E. Foran, L. Rosenblum, A.I. Bogush, D. Trotti, Sumoylation of critical proteins in amyotrophic lateral sclerosis: Emerging pathways of pathogenesis, NeuroMolecular Med. 15 (2013) 760–770. https://doi.org/10.1007/s12017-013-8262-x.

[36] R. Burai, N. Ait-Bouziad, A. Chiki, H.A. Lashuel, Elucidating the role of site-specific nitration of α-synuclein in the pathogenesis of Parkinson's disease via protein semisynthesis and mutagenesis, J. Am. Chem. Soc. 137 (2015) 5041–5052. https://doi.org/10.1021/ja5131726.

[37] X.H. Li, L.L. Du, X.S. Cheng, X. Jiang, Y. Zhang, B.L. Lv, R. Liu, J.Z. Wang, X.W. Zhou, Glycation exacerbates the neuronal toxicity of β-amyloid, Cell Death Dis. 4 (2013) 1–9. https://doi.org/10.1038/cddis.2013.180.

[38] C.K. Oh, A. Sultan, J. Platzer, N. Dolatabadi, F. Soldner, D.B. McClatchy, J.K. Diedrich, J.R. Yates, R. Ambasudhan, T. Nakamura, R. Jaenisch, S.A. Lipton, S-Nitrosylation of PINK1 Attenuates PINK1/Parkin-Dependent Mitophagy in hiPSC-Based Parkinson's Disease Models, Cell Rep. 21 (2017) 2171–2182. https://doi.org/10.1016/j.celrep.2017.10.068.

[39] V. Guru KrishnaKumar, L. Baweja, K. Ralhan, S. Gupta, Carbamylation promotes amyloidogenesis and induces structural changes in Tau-core hexapeptide fibrils,

Biochim. Biophys. Acta - Gen. Subj. 1862 (2018) 2590–2604. https://doi.org/10.1016/j.bbagen.2018.07.030.

[40] K.C. Wilcox, L. Zhou, J.K. Jordon, Y. Huang, Y. Yu, R.L. Redler, X. Chen, M. Caplow, N. V. Dokholyan, Modifications of superoxide dismutase (SOD1) in human erythrocytes: A possible role in amyotrophic lateral sclerosis, J. Biol. Chem. 284 (2009) 13940–13947. https://doi.org/10.1074/jbc.M809687200.

[41] F. Liu, T. Zaidi, K. Iqbal, I. Grundke-Iqbal, R.K. Merkle, C.X. Gong, Role of glycosylation in hyperphosphorylation of tau in Alzheimer's disease, FEBS Lett. 512 (2002) 101–106. https://doi.org/10.1016/S0014-5793(02)02228-7.

[42] T. Sen, P. Saha, T. Jiang, N. Sen, Sulfhydration of AKT triggers Tau-phosphorylation by activating glycogen synthase kinase 3β in Alzheimer's disease, Proc. Natl. Acad. Sci. U. S. A. 117 (2020) 4418–4427. https://doi.org/10.1073/pnas.1916895117.

[43] H.K. Chen, P. Fernandez-Funez, S.F. Acevedo, Y.C. Lam, M.D. Kaytor, M.H. Fernandez, A. Aitken, E.M.C. Skoulakis, H.T. Orr, J. Botas, H.Y. Zoghbi, Interaction of Akt-phosphorylated ataxin-1 with 14-3-3 mediates neurodegeneration in spinocerebellar ataxia type 1, Cell. 113 (2003) 457–468. https://doi.org/10.1016/S0092-8674(03)00349-0.

[44] Y. Zhang, X.G. Gong, Z.Z. Wang, H.M. Sun, Z.Y. Guo, J.H. Hu, L. Ma, P. Li, N.H. Chen, Overexpression of DJ-1/PARK7, the Parkinson's disease-related protein, improves mitochondrial function via Akt phosphorylation on threonine 308 in dopaminergic neuron-like cells, Eur. J. Neurosci. 43 (2016) 1379–1388. https://doi.org/10.1111/ejn.13216.

[45] O. Mercado-Gómez, K. Hernández-Fonseca, A. Villavicencio-Queijeiro, L. Massieu, J. Chimal-Monroy, C. Arias, Inhibition of Wnt and PI3K signaling modulates GSK-3β activity and induces morphological changes in cortical neurons: Role of tau phosphorylation, Neurochem. Res. 33 (2008) 1599–1609. https://doi.org/10.1007/s11064-008-9714-9.

[46] A. Caricasole, A. Copani, F. Caraci, E. Aronica, A.J. Rozemuller, A. Caruso, M. Storto, G. Gaviraghi, G.C. Terstappen, F. Nicoletti, Induction of Dickkopf-1, a negative modulator of the Wnt pathway, is associated with neuronal degeneration in Alzheimer's brain, J. Neurosci. 24 (2004) 6021–6027.

https://doi.org/10.1523/JNEUROSCI.1381-04.2004.

[47] L. Wang, F.X. Shi, N. Li, Y. Cao, Y. Lei, J.Z. Wang, Q. Tian, X.W. Zhou, AMPK Ameliorates Tau Acetylation and Memory Impairment Through Sirt1, Mol. Neurobiol. (2020). https://doi.org/10.1007/s12035-020-02079-x.

[48] J. Chen, Y. Ren, C. Gui, M. Zhao, X. Wu, K. Mao, W. Li, F. Zou, Phosphorylation of Parkin at serine 131 by p38 MAPK promotes mitochondrial dysfunction and neuronal death in mutant A53T α-synuclein model of Parkinson's disease, Cell Death Dis. 9 (2018). https://doi.org/10.1038/s41419-018-0722-7.

[49] S.O. Yoon, D.J. Park, J.C. Ryu, H.G. Ozer, C. Tep, Y.J. Shin, T.H. Lim, L. Pastorino, A.J. Kunwar, J.C. Walton, A.H. Nagahara, K.P. Lu, R.J. Nelson, M.H. Tuszynski, K. Huang, JNK3 Perpetuates Metabolic Stress Induced by Aβ Peptides, Neuron. 75 (2012) 824–837. https://doi.org/10.1016/j.neuron.2012.06.024.

[50] E.G. Wheatley, E. Albarran, C.W. White, G. Bieri, C. Sanchez-Diaz, K. Pratt, C.E. Snethlage, J.B. Ding, S.A. Villeda, Neuronal O-GlcNAcylation Improves Cognitive Function in the Aged Mouse Brain, Curr. Biol. 29 (2019) 3359-3369.e4. https://doi.org/10.1016/j.cub.2019.08.003.

[51] A.L. McCormack, S.K. Mak, D.A. Di Monte, Increased a-synuclein phosphorylation and nitration in the aging primate substantia nigra, Cell Death Dis. 3 (2012) e315-9. https://doi.org/10.1038/cddis.2012.50.

[52] S. Rizza, S. Cardaci, C. Montagna, G. Di Giacomo, D. De Zio, M. Bordi, E. Maiani, S. Campello, A. Borreca, A.A. Puca, J.S. Stamler, F. Cecconi, G. Filomeni, S-nitrosylation drives cell senescence and aging in mammals by controlling mitochondrial dynamics and mitophagy, Proc. Natl. Acad. Sci. U. S. A. 115 (2018) E3388–E3397. https://doi.org/10.1073/pnas.1722452115.

[53] S.Y. Hong, L.T. Ng, L.F. Ng, T. Inoue, N.S. Tolwinski, T. Hagen, J. Gruber, The role of mitochondrial non-enzymatic protein acylation in ageing, PLoS One. 11 (2016) 12–14. https://doi.org/10.1371/journal.pone.0168752.

[54] C. Klaus, J.N. Hansen, A. Ginolhac, D. Gérard, V.S. Gnanapragassam, R. Horstkorte, C. Rossdam, F.F.R. Buettner, T. Sauter, L. Sinkkonen, H. Neumann, B. Linnartz-Gerlach, Reduced sialylation triggers homeostatic synapse and neuronal loss in

middle-aged mice, Neurobiol. Aging. 88 (2020) 91–107. https://doi.org/10.1016/j.neurobiolaging.2020.01.008.

[55] J. Zivanovic, E. Kouroussis, J.B. Kohl, B. Adhikari, B. Bursac, S. Schott-Roux, D. Petrovic, J.L. Miljkovic, D. Thomas-Lopez, Y. Jung, M. Miler, S. Mitchell, V. Milosevic, J.E. Gomes, M. Benhar, B. Gonzales-Zorn, I. Ivanovic-Burmazovic, R. Torregrossa, J.R. Mitchell, M. Whiteman, G. Schwarz, S.H. Snyder, B.D. Paul, K.S. Carroll, M.R. Filipovic, Selective Persulfide Detection Reveals Evolutionarily Conserved Antiaging Effects of S-Sulfhydration, Cell Metab. 30 (2019) 1152-1170.e13. https://doi.org/10.1016/j.cmet.2019.10.007.

[56] D.R. Zamzow, V. Elias, V.A. Acosta, E. Escobedo, K.R. Magnusson, Higher levels of protein palmitoylation in the frontal cortex across aging were associated with reference memory and executive function declines, ENeuro. 6 (2019). https://doi.org/10.1523/ENEURO.0310-18.2019.

[57] M. Tanase, A.M. Urbanska, V. Zolla, C.C. Clement, L. Huang, K. Morozova, C. Follo, M. Goldberg, B. Roda, P. Reschiglian, L. Santambrogio, Role of carbonyl modifications on aging-associated protein aggregation, Sci. Rep. 6 (2016) 1–14. https://doi.org/10.1038/srep19311.

[58] C.-S. Kim, S. Park, J. Kim, The role of glycation in the pathogenesis of aging and its prevention through herbal products and physical exercise, J. Exerc. Nutr. Biochem. 21 (2017) 55–61. https://doi.org/10.20463/jenb.2017.0027.

[59] T.S. Toh, F. Dondelinger, D. Wang, Looking beyond the hype: Applied AI and machine learning in translational medicine, EBioMedicine. (2019). https://doi.org/10.1016/j.ebiom.2019.08.027.

[60] C. Cao, F. Liu, H. Tan, D. Song, W. Shu, W. Li, Y. Zhou, X. Bo, Z. Xie, Deep Learning and Its Applications in Biomedicine, The Authors, 2018. https://doi.org/10.1016/j.gpb.2017.07.003.

[61] R. Gupta, D. Srivastava, M. Sahu, S. Tiwari, R.K. Ambasta, P. Kumar, Artificial intelligence to deep learning : machine intelligence approach for drug discovery, Springer International Publishing, 2021. https://doi.org/10.1007/s11030-021-10217-3.

[62] W.L. Martinez, A.R. Martinez, Unsupervised Learning, 2020.

https://doi.org/10.1201/b19035-16.

[63] M. Mahmud, S. Member, M. Shamim Kaiser, A. Hussain, S. Vassanelli, IEEE TRANSACTIONS ON NEURAL NETWORKS AND LEARNING SYSTEMS 1 Applications of Deep Learning and Reinforcement Learning to Biological Data, Ieee Trans. Neural Networks Learn. Syst. 1 (2018) 1–17.

[64] A. Gandomi, M. Haider, Beyond the hype: Big data concepts, methods, and analytics, Int. J. Inf. Manage. 35 (2015) 137–144. https://doi.org/10.1016/j.ijinfomgt.2014.10.007.

[65] R. Edgar, M. Domrachev, A.E. Lash, Gene Expression Omnibus: NCBI gene expression and hybridization array data repository, Nucleic Acids Res. (2002). https://doi.org/10.1093/nar/30.1.207.

[66] H. Parkinson, M. Kapushesky, M. Shojatalab, N. Abeygunawardena, R. Coulson, A. Farne, E. Holloway, N. Kolesnykov, P. Lilja, M. Lukk, R. Mani, T. Rayner, A. Sharma, E. William, U. Sarkans, A. Brazma, ArrayExpress - A public database of microarray experiments and gene expression profiles, Nucleic Acids Res. (2007). https://doi.org/10.1093/nar/gkl995.

[67] Z. Wang, M.A. Jensen, J.C. Zenklusen, A practical guide to The Cancer Genome Atlas (TCGA), in: Methods Mol. Biol., 2016. https://doi.org/10.1007/978-1-4939-3578-9_6.

[68] A. Buniello, J.A.L. Macarthur, M. Cerezo, L.W. Harris, J. Hayhurst, C. Malangone, A. McMahon, J. Morales, E. Mountjoy, E. Sollis, D. Suveges, O. Vrousgou, P.L. Whetzel, R. Amode, J.A. Guillen, H.S. Riat, S.J. Trevanion, P. Hall, H. Junkins, P. Flicek, T. Burdett, L.A. Hindorff, F. Cunningham, H. Parkinson, The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019, Nucleic Acids Res. (2019). https://doi.org/10.1093/nar/gky1120.

[69] T. Beck, R.K. Hastings, S. Gollapudi, R.C. Free, A.J. Brookes, GWAS Central: A comprehensive resource for the comparison and interrogation of genome-wide association studies, Eur. J. Hum. Genet. (2014). https://doi.org/10.1038/ejhg.2013.274.

[70] R. Leinonen, H. Sugawara, M. Shumway, The sequence read archive, Nucleic Acids Res. (2011). https://doi.org/10.1093/nar/gkq1019.

[71] K. Canese, S. Weis, PubMed: The bibliographic database, NCBI Handb. (2013).

[72] S.K. Burley, H.M. Berman, C. Bhikadiya, C. Bi, L. Chen, L. Di Costanzo, C. Christie, K. Dalenberg, J.M. Duarte, S. Dutta, Z. Feng, S. Ghosh, D.S. Goodsell, R.K. Green, V. Guranović, D. Guzenko, B.P. Hudson, T. Kalro, Y. Liang, R. Lowe, H. Namkoong, E. Peisach, I. Periskova, A. Prlić, C. Randle, A. Rose, P. Rose, R. Sala, M. Sekharan, C. Shao, L. Tan, Y.P. Tao, Y. Valasatava, M. Voigt, J. Westbrook, J. Woo, H. Yang, J. Young, M. Zhuravleva, C. Zardecki, RCSB Protein Data Bank: Biological macromolecular structures enabling research and education in fundamental biology, biomedicine, biotechnology and energy, Nucleic Acids Res. 47 (2019) D464–D474. https://doi.org/10.1093/nar/gky1004.

[73] Z. Xu, L. Yang, X. Zhang, Q. Zhang, Z. Yang, Y. Liu, S. Wei, W. Liu, Discovery of Potential Flavonoid Inhibitors Against COVID-19 3CL Proteinase Based on Virtual Screening Strategy, Front. Mol. Biosci. 7 (2020) 1–8. https://doi.org/10.3389/fmolb.2020.556481.

[74] S. Kim, J. Chen, T. Cheng, A. Gindulyte, J. He, S. He, Q. Li, B.A. Shoemaker, P.A. Thiessen, B. Yu, L. Zaslavsky, J. Zhang, E.E. Bolton, PubChem 2019 update: Improved access to chemical data, Nucleic Acids Res. (2019). https://doi.org/10.1093/nar/gky1033.

[75] A.P. Bento, A. Gaulton, A. Hersey, L.J. Bellis, J. Chambers, M. Davies, F.A. Krüger, Y. Light, L. Mak, S. McGlinchey, M. Nowotka, G. Papadatos, R. Santos, J.P. Overington, The ChEMBL bioactivity database: An update, Nucleic Acids Res. 42 (2014). https://doi.org/10.1093/nar/gkt1031.

[76] Q. Duan, S.P. Reid, N.R. Clark, Z. Wang, N.F. Fernandez, A.D. Rouillard, B. Readhead, S.R. Tritsch, R. Hodos, M. Hafner, M. Niepel, P.K. Sorger, J.T. Dudley, S. Bavari, R.G. Panchal, A. Ma'Ayan, L1000CDS2: LINCS L1000 characteristic direction signatures search engine, Npj Syst. Biol. Appl. 2 (2016) 1–12. https://doi.org/10.1038/npjsba.2016.15.

[77] D.S. Wishart, Y.D. Feunang, A.C. Guo, E.J. Lo, A. Marcu, J.R. Grant, T. Sajed, D. Johnson, C. Li, Z. Sayeeda, N. Assempour, I. Iynkkaran, Y. Liu, A. MacIejewski, N. Gale, A. Wilson, L. Chin, R. Cummings, Di. Le, A. Pon, C. Knox, M. Wilson, DrugBank 5.0: A major update to the DrugBank database for 2018, Nucleic Acids Res.

46 (2018) D1074–D1082. https://doi.org/10.1093/nar/gkx1037.

[78] K. Yu, Q. Zhang, Z. Liu, Y. Du, X. Gao, Q. Zhao, H. Cheng, X. Li, Z.X. Liu, Deep learning based prediction of reversible HAT/HDAC-specific lysine acetylation, Brief. Bioinform. 21 (2020) 1798–1805. https://doi.org/10.1093/bib/bbz107.

[79] T.Y. Lee, H. Da Huang, J.H. Hung, H.Y. Huang, Y.S. Yang, T.H. Wang, dbPTM: an information repository of protein post-translational modification., Nucleic Acids Res. 34 (2006) 622–627. https://doi.org/10.1093/nar/gkj083.

[80] A.W. Senior, R. Evans, J. Jumper, J. Kirkpatrick, L. Sifre, T. Green, C. Qin, A. Žídek, A.W.R. Nelson, A. Bridgland, H. Penedones, S. Petersen, K. Simonyan, S. Crossan, P. Kohli, D.T. Jones, D. Silver, K. Kavukcuoglu, D. Hassabis, Improved protein structure prediction using potentials from deep learning, Nature. 577 (2020) 706–710. https://doi.org/10.1038/s41586-019-1923-7.

[81] M. AlQuraishi, End-to-End Differentiable Learning of Protein Structure, Cell Syst. 8 (2019) 292-301.e3. https://doi.org/10.1016/j.cels.2019.03.006.

[82] M. Gastegger, A. McSloy, M. Luya, K.T. Schütt, R.J. Maurer, A deep neural network for molecular wave functions in quasi-atomic minimal basis representation, J. Chem. Phys. (2020). https://doi.org/10.1063/5.0012911.

[83] Y. Han, J. Yang, X. Qian, W.C. Cheng, S.H. Liu, X. Hua, L. Zhou, Y. Yang, Q. Wu, P. Liu, Y. Lu, DriverML: A machine learning algorithm for identifying driver genes in cancer sequencing studies, Nucleic Acids Res. 47 (2019). https://doi.org/10.1093/nar/gkz096.

[84] Q. Bai, S. Tan, T. Xu, H. Liu, J. Huang, X. Yao, MolAICal: a soft tool for 3D drug design of protein targets by artificial intelligence and classical algorithm, Brief. Bioinform. 00 (2020) 1–12. https://doi.org/10.1093/bib/bbaa161.

[85] D. Szklarczyk, A.L. Gable, D. Lyon, A. Junge, S. Wyder, J. Huerta-Cepas, M. Simonovic, N.T. Doncheva, J.H. Morris, P. Bork, L.J. Jensen, C. Von Mering, STRING v11: Protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets, Nucleic Acids Res. 47 (2019) D607–D613. https://doi.org/10.1093/nar/gky1131.

[86] D. Szklarczyk, A. Santos, C. Von Mering, L.J. Jensen, P. Bork, M. Kuhn, STITCH 5:

Augmenting protein-chemical interaction networks with tissue and affinity data, Nucleic Acids Res. 44 (2016) D380–D384. https://doi.org/10.1093/nar/gkv1277.

[87] J. Piñero, Á. Bravo, N. Queralt-Rosinach, A. Gutiérrez-Sacristán, J. Deu-Pons, E. Centeno, J. García-García, F. Sanz, L.I. Furlong, DisGeNET: A comprehensive platform integrating information on human disease-associated genes and variants, Nucleic Acids Res. 45 (2017) D833–D839. https://doi.org/10.1093/nar/gkw943.

[88] B.A. Grzybowski, S. Szymkuć, E.P. Gajewska, K. Molga, P. Dittwald, A. Wołos, T. Klucznik, Chematica: A Story of Computer Code That Started to Think like a Chemist, Chem. 4 (2018) 390–398. https://doi.org/10.1016/j.chempr.2018.02.024.

[89] S. Genheden, A. Thakkar, V. Chadimová, J.L. Reymond, O. Engkvist, E. Bjerrum, AiZynthFinder: a fast, robust and flexible open-source software for retrosynthetic planning, J. Cheminform. 12 (2020) 1–9. https://doi.org/10.1186/s13321-020-00472-1.

[90] J. Wu, Q. Zhang, W. Wu, T. Pang, H. Hu, W.K.B. Chan, X. Ke, Y. Zhang, WDL-RF: Predicting bioactivities of ligand molecules acting with G protein-coupled receptors by combining weighted deep learning and random forest, Bioinformatics. 34 (2018) 2271–2282. https://doi.org/10.1093/bioinformatics/bty070.

[91] A. Cichonska, T. Pahikkala, S. Szedmak, H. Julkunen, A. Airola, M. Heinonen, T. Aittokallio, J. Rousu, Learning with multiple pairwise kernels for drug bioactivity prediction, Bioinformatics. 34 (2018) i509–i518. https://doi.org/10.1093/bioinformatics/bty277.

[92] A.K. Arshadi, M. Salem, J. Collins, J.S. Yuan, D. Chakrabarti, Deepmalaria: Artificial intelligence driven discovery of potent antiplasmodials, Front. Pharmacol. 10 (2020). https://doi.org/10.3389/fphar.2019.01526.

[93] V. Parvathaneni, N.S. Kulkarni, A. Muth, V. Gupta, Drug repurposing: a promising tool to accelerate the drug discovery process, Drug Discov. Today. 24 (2019) 2076–2085. https://doi.org/10.1016/j.drudis.2019.06.014.

[94] T. Kurata, V. Lukic, M. Kozuki, D. Wada, K. Miyazaki, N. Morimoto, Y. Ohta, K. Deguchi, Y. Ikeda, T. Kamiya, K. Abe, Telmisartan reduces progressive accumulation of cellular amyloid beta and phosphorylated tau with inflammatory responses in aged spontaneously hypertensive stroke resistant rat, J. Stroke Cerebrovasc. Dis. 23 (2014)

2580–2590. https://doi.org/10.1016/j.jstrokecerebrovasdis.2014.05.023.

[95] M.M. Chung, Y.L. Chen, D. Pei, Y.C. Cheng, B. Sun, C.J. Nicol, C.H. Yen, H.M. Chen, Y.J. Liang, M.C. Chiang, The neuroprotective role of metformin in advanced glycation end product treated human neural stem cells is AMPK-dependent, Biochim. Biophys. Acta - Mol. Basis Dis. 1852 (2015) 720–731. https://doi.org/10.1016/j.bbadis.2015.01.006.

[96] J. Sun, X. Zhang, C. Wang, Z. Teng, Y. Li, Curcumin Decreases Hyperphosphorylation of Tau by Down-Regulating Caveolin-1/GSK-3β in N2a/APP695swe Cells and APP/PS1 Double Transgenic Alzheimer's Disease Mice, Am. J. Chin. Med. 45 (2017) 1667–1682. https://doi.org/10.1142/S0192415X17500902.

[97] S. Schweiger, F. Matthes, K. Posey, E. Kickstein, S. Weber, M.M. Hettich, S. Pfurtscheller, D. Ehninger, R. Schneider, S. Krauß, Resveratrol induces dephosphorylation of Tau by interfering with the MID1-PP2A complex, Sci. Rep. 7 (2017) 1–13. https://doi.org/10.1038/s41598-017-12974-4.

[98] Z. Yılmaz, E.B. Kalaz, A.F. Aydın, V. Olgaç, S. Doğru-Abbasoğlu, M. Uysal, N. Koçak-Toker, The effect of resveratrol on glycation and oxidation products in plasma and liver of chronic methylglyoxal-treated rats, Pharmacol. Reports. 70 (2018) 584–590. https://doi.org/10.1016/j.pharep.2017.12.005.

[99] X. Zeng, S. Zhu, X. Liu, Y. Zhou, R. Nussinov, F. Cheng, DeepDR: A network-based deep learning approach to in silico drug repositioning, Bioinformatics. 35 (2019) 5191–5198. https://doi.org/10.1093/bioinformatics/btz418.

[100] H. Chen, F. Cheng, J. Li, IDrug: Integration of drug repositioning and drug-target prediction via cross-network embedding, PLoS Comput. Biol. (2020). https://doi.org/10.1371/journal.pcbi.1008040.

[101] Y. Zhu, W. Jung, F. Wang, C. Che, Drug repurposing against Parkinson's disease by text mining the scientific literature, Libr. Hi Tech. 38 (2020) 741–750. https://doi.org/10.1108/LHT-08-2019-0170.

[102] P. Chatterjee, D. Roy, N. Rathi, Epigenetic Drug Repositioning for Alzheimer's Disease Based on Epigenetic Targets in Human Interactome, J. Alzheimer's Dis. 61

(2018) 53–65. https://doi.org/10.3233/JAD-161104.

[103] M.E. Ritchie, B. Phipson, D. Wu, Y. Hu, C.W. Law, W. Shi, G.K. Smyth, Limma powers differential expression analyses for RNA-sequencing and microarray studies, Nucleic Acids Res. 43 (2015) e47. https://doi.org/10.1093/nar/gkv007.

[104] J. Zou, Y. Han, S.S. So, Overview of artificial neural networks, Methods Mol. Biol. 458 (2008) 15–23. https://doi.org/10.1007/978-1-60327-101-1_2.

[105] F. Günther, S. Fritsch, Neuralnet: Training of neural networks, R J. 2 (2010) 30–38. https://doi.org/10.32614/rj-2010-006.

[106] S.B. Kotsiantis, Decision trees: A recent overview, Artif. Intell. Rev. 39 (2013) 261–283. https://doi.org/10.1007/s10462-011-9272-4.

[107] T. Therneau, B. Atkinson, B. Ripley, rpart: Recursive partitioning for classification, regression and survival trees. R package version 4.1-15, (2019). https://cran.r-project.org/package=rpart.

[108] A. Cutler, D.R. Cutler, J.R. Stevens, Ensemble Machine Learning, Ensemble Mach. Learn. (2012). https://doi.org/10.1007/978-1-4419-9326-7.

[109] A. Karatzoglou, K. Hornik, A. Smola, A. Zeileis, kernlab - An S4 package for kernel methods in R, J. Stat. Softw. 11 (2004) 1–20. https://doi.org/10.18637/jss.v011.i09.

[110] E. Karatzas, G. Minadakis, G. Kolios, A. Delis, G.M. Spyrou, A Web Tool for Ranking Candidate Drugs Against a Selected Disease Based on a Combination of Functional and Structural Criteria, Comput. Struct. Biotechnol. J. 17 (2019) 939–945. https://doi.org/10.1016/j.csbj.2019.05.010.

[111] A. Daina, O. Michielin, V. Zoete, SwissADME: A free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules, Sci. Rep. (2017). https://doi.org/10.1038/srep42717.

[112] K. Yu, Q. Zhang, Z. Liu, Y. Du, X. Gao, Q. Zhao, H. Cheng, X. Li, Z.-X. Liu, Deep learning based prediction of reversible HAT/HDAC-specific lysine acetylation., Brief. Bioinform. 21 (2020) 1798–1805. https://doi.org/10.1093/bib/bbz107.

[113] J. Grau, I. Grosse, J. Keilwagen, PRROC: Computing and visualizing Precision-recall and receiver operating characteristic curves in R, Bioinformatics. 31 (2015) 2595–

2597. https://doi.org/10.1093/bioinformatics/btv153.

[114] Y. Huo, P. Rangarajan, E.A. Ling, S.T. Dheen, Dexamethasone inhibits the Nox-dependent ROS production via suppression of MKP-1-dependent MAPK pathways in activated microglia, BMC Neurosci. 12 (2011). https://doi.org/10.1186/1471-2202-12-49.

[115] D. Masini, A. Bonito-Oliva, M. Bertho, G. Fisone, Inhibition of mTORC1 signaling reverts cognitive and affective deficits in a mouse model of Parkinson's disease, Front. Neurol. 9 (2018) 1–8. https://doi.org/10.3389/fneur.2018.00208.

[116] Y.J. Chen, C.C. Hsu, Y.J. Shiao, H.T. Wang, Y.L. Lo, A.M.Y. Lin, Anti-inflammatory effect of afatinib (an EGFR-TKI) on OGD-induced neuroinflammation, Sci. Rep. 9 (2019) 1–10. https://doi.org/10.1038/s41598-019-38676-7.

[117] S.Y. Kang, S. Bin Lee, H.J. Kim, H.T. Kim, H.O. Yang, W. Jang, Autophagic modulation by rosuvastatin prevents rotenone-induced neurotoxicity in an in vitro model of Parkinson's disease, Neurosci. Lett. 642 (2017) 20–26. https://doi.org/10.1016/j.neulet.2017.01.063.

[118] G. Wang, X. Jiang, H. Pu, W. Zhang, C. An, X. Hu, A.K.F. Liou, R.K. Leak, Y. Gao, J. Chen, Scriptaid, a Novel Histone Deacetylase Inhibitor, Protects Against Traumatic Brain Injury via Modulation of PTEN and AKT Pathway: Scriptaid Protects Against TBI via AKT, Neurotherapeutics. 10 (2013) 124–142. https://doi.org/10.1007/s13311-012-0157-2.

[119] F. Amenta, D. Tomassoni, Treatment with nicardipine protects brain in an animal model of hypertension-induced damage, Clin. Exp. Hypertens. 26 (2004) 351–361. https://doi.org/10.1081/CEH-120034139.

[120] J.Y. Huang, Y.T. Hong, J.I. Chuang, Fibroblast growth factor 9 prevents MPP+-induced death of dopaminergic neurons and is involved in melatonin neuroprotection in vivo and in vitro, J. Neurochem. 109 (2009) 1400–1412. https://doi.org/10.1111/j.1471-4159.2009.06061.x.

[121] P.S. Eis, N. Huang, J.W. Langston, E. Hatchwell, B. Schüle, Loss-of-Function NUBPL Mutation May Link Parkinson's Disease to Recessive Complex I Deficiency, Front. Neurol. 11 (2020) 1–5. https://doi.org/10.3389/fneur.2020.555961.

Journal publication 1-:

Title: "Post-translational modifications: Regulators of neurodegenerative proteinopathies"**, Journal:** Elsevier's Ageing Research Reviews**, Impact Factor:** 10.616, **Authorship:** Joint first author

Ageing Research Reviews 68 (2021) 101336

Contents lists available at ScienceDirect

Ageing Research Reviews

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

Review

Post-translational modifications: Regulators of neurodegenerative proteinopathies

Rohan Gupta [1], Mehar Sahu [1] Devesh Srivastava [1] Swati Tiwari [1], Rashmi K. Ambasta, Pravir Kumar *

Molecular Neuroscience and Functional Genomics Laboratory, Department of Biotechnology, Delhi Technological University (Formerly DCE), Delhi 110042, India

ARTICLE INFO

Keywords:
Post-translational modifications
Neurodegenerative disease
Proteinopathies
Protein aggregation

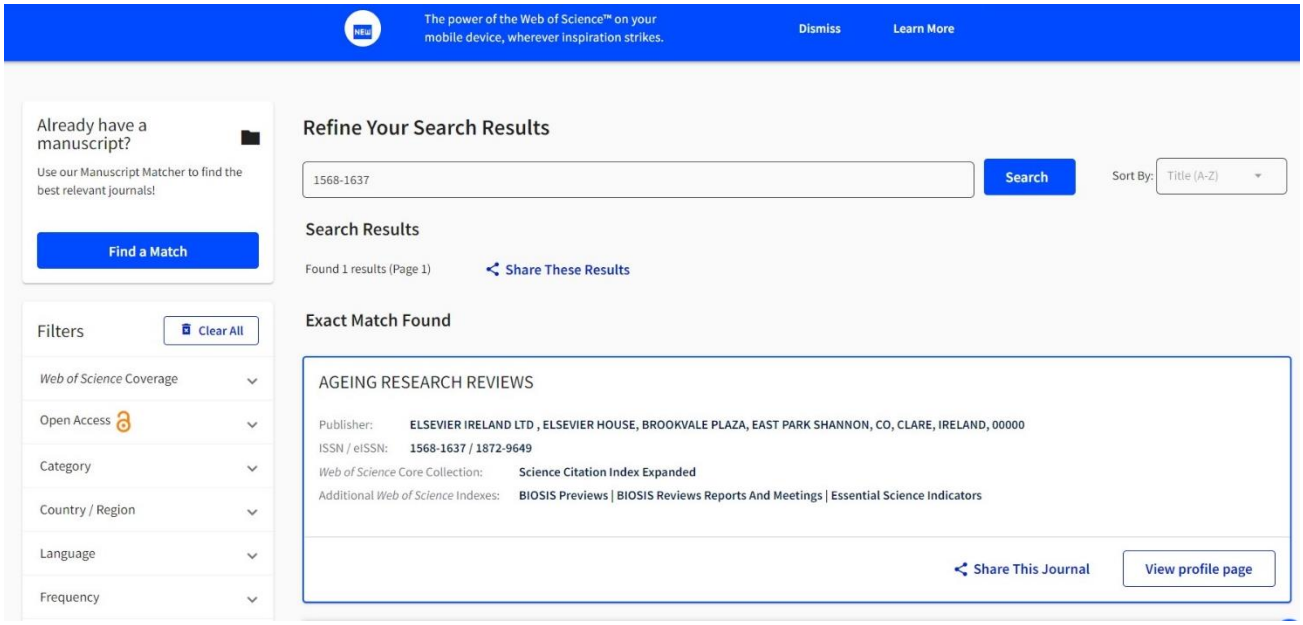
ABSTRACT

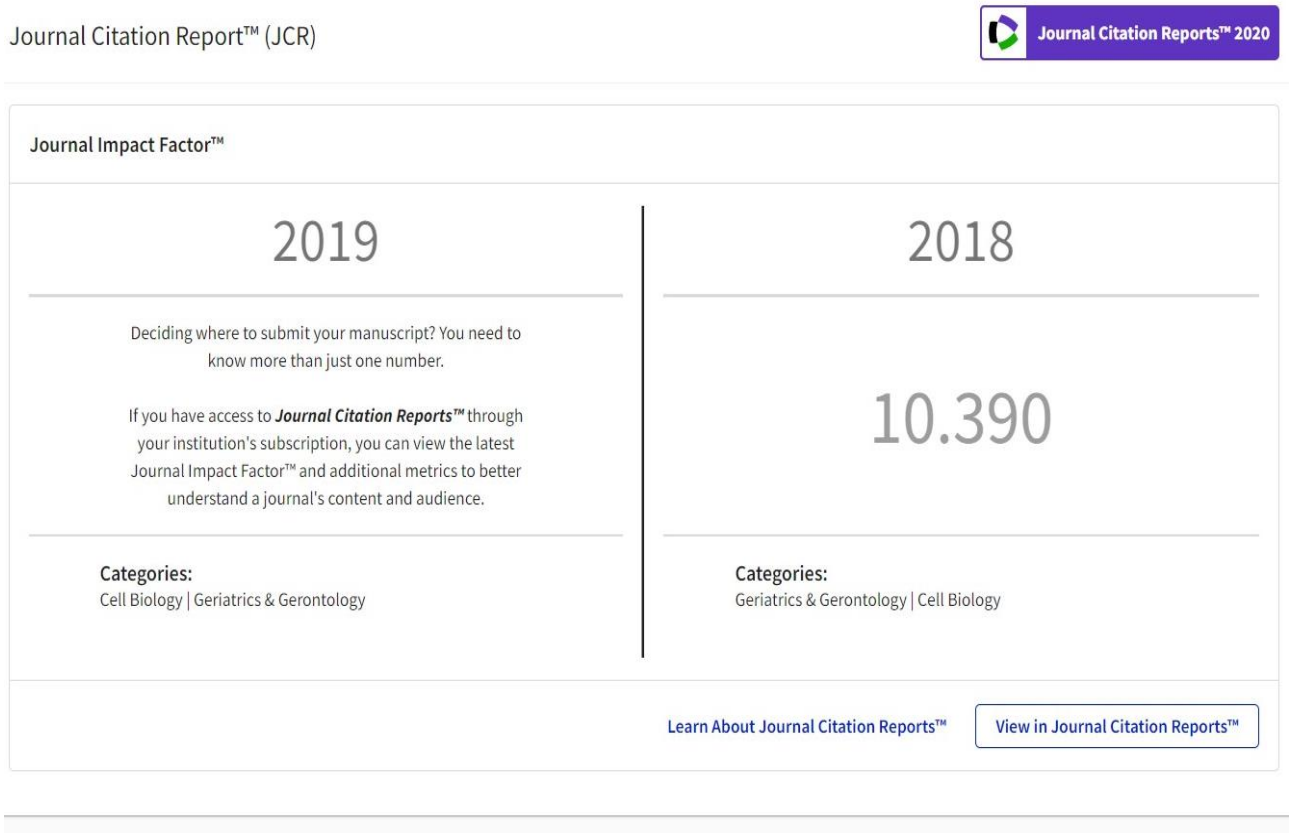
One of the hallmark features in the neurodegenerative disorders (NDDs) is the accumulation of aggregated and/or non-functional protein in the cellular milieu. Post-translational modifications (PTMs) are an essential regulator of non-functional protein aggregation in the pathogenesis of NDDs. Any alteration in the post-translational mechanism and the protein quality control system, for instance, molecular chaperone, ubiquitin-proteasome system, autophagy-lysosomal degradation pathway, enhances the accumulation of misfolded protein, which

Abbreviations: PTMs, Post-translational modifications; NDDs, Neurodegenerative diseases; AD, Alzheimer's disease; PD, Parkinson's disease; ALS, Amyotrophic lateral sclerosis; HD, Huntington's disease; TDP-43, Transactivation response DNA binding protein-43; Aβ, β-amyloid; NFTs, Neurofibrils tangles; SNpc, Substantia nigra pars compacta; polyQ, Polyglutamine; LBs, Lewy bodies; htt, Huntingtin protein; SOD1, Superoxide dismutase 1; UPS, Ubiquitin-proteasome system; CMA, Chaperone mediated autophagy; HSPs, Heat shock proteins; PSEN2, Presenilin-2; IT15, Interesting transcript 15; TARDBP, TAR DNA Binding Protein; NO, Nitric oxide; CK1, Casein kinase 1; GSK-3β, Glycogen synthase kinase 3β; PKA, Protein kinase A; CDK5, Cyclin-dependent kinase 5; DYRK1A, Dual Specificity Tyrosine Phosphorylation Regulated Kinase 1A; REP, Repressor element of PARKIN; MTS, Mitochondrial targeting sequence; TM, Transmembrane; FUS, Fused in sarcoma; NLS, Nuclear localization sequence; RRMs, RNA recognition motif; NES, Nuclear export sequence; ER, Endoplasmic reticulum; UPR, Unfolded protein response; IRE1α, Inositol-requiring enzyme 1 α; PERK, Protein kinase R like endoplasmic reticulum kinase; ATF6α, Activating transcription factor 6 α; DR5, Death receptor 5; eIF2α, Eukaryotic initiation factor 2 α; XBP1, X-box binding protein 1; ASK1-JNK, apoptosis signal-regulating kinase 1/ c-Jun N-terminal kinases; TRAF2, TNF receptor-associated factor 2; PAD4, Protein arginase deaminase 4; SP1, Specificity protein 1; SP2, Specificity protein 2; PARP16, poly ADP ribose polymerase 16; Umf1, ubiquitin fold modifier 1; CHOP, C/EBP homologous protein; ERAD, endoplasmic reticulum-associated degradation; APP, Amyloid precursor protein; PSEN1, Presenilin-1; BACE1, Beta-secretase 1; BiP, Binding immunoglobulin protein; PARKIN, E3 ubiquitin-protein ligase parkin; PINK1, PTEN-induced kinase 1; PDI, Protein disulfide isomerase; GADD34, Growth arrest and DNA damage-inducible protein; DJ1, Protein deglycase; ATFS1, Cyclic AMP-dependent transcription factor; PDR1, Pleiotropic drug resistance 1; CSMNs, Corticospinal motor neurons; FOXO1, Forkhead box protein O1; FTLD, frontotemporal lobar degeneration; ASK1, Signal-regulating kinase 1; PI3K, Phosphatidylinositol 3-kinase; PIP2, Phosphatidylinositol 4,5-bisphosphate; PIP3, Phosphatidylinositol (3,4,5)-trisphosphate; PH domain, Pleckstrin homology domain; PDK1, Phosphoinositide dependent kinase 1; mTORC2, Mammalian target of rapamycin complex 2; mTORC1, Mammalian target of rapamycin complex 1; IGF-1, Insulin like growth factor-1; AGEs, Advanced glycation end-product; RAGE, Receptor for advanced glycation end products; MAPK, Mitogen activated protein kinase; ERK ½, Extracellular signal regulated kinase ½; AMPK, Adenosine monophosphate activated protein kinase; CBS, cystathionine-beta-synthase; AMP, Adenosine monophosphate; ADP, Adenosine diphosphate; ATP, Adenosine triphosphate; LKB1, Liver kinase B1; TAK1, Transforming growth factor beta activated kinase 1; CaMKKβ, Calmodulin dependent protein kinase kinase-β; ACC, Acetyl CoA carboxylase; ULK1, Unc-51 like autophagy activating kinase 1; TSC 1/2, Tuberous sclerosis complex ½; SREBP, Sterol regulatory element binding protein; NAD^+, Nicotinamide adenine dinucleotide; PGC-1α, Peroxisome proliferator activated receptor gamma coactivator-1α; BCL-2, B-cell lymphoma 2; APC, Adenomatous polyposis coli; β-TrCP, β-transducin repeats containing proteins; LRP5/6, Low density lipoprotein receptor related protein 5/6; Dsh, Dishevelled; TCF/LEF, T-cell factor/lymphoid enhancer factor; DKK1, Dickkopf related protein 1; AVs, Autophagic vacuoles; PAR, UBE3A: Ubiquitin-protein ligase E3A; TRAF6, Tumor necrosis factor receptor associated factor 6; UBE2D2, Ubiquitin-conjugating enzyme E2 D2; UCH-L1, Ubiquitin carboxy-terminal hydrolase L1; USP14, Ubiquitin carboxyl-terminal hydrolase 14; BCL-xL, B-cell lymphoma-extra-large; LAMP2A, Lysosome-associated membrane protein 2A; LRRK2, Leucine-rich repeat kinase; CNS, Central nervous system; COX2, cyclooxygenase-2; iNOS, Inducible nitric oxide synthase; NF-κβ, Nuclear factor kappa-light-chain-enhancer of activated B cells; IKK, IκB kinase; Cdks, cyclin-dependent kinases; NGF, Nerve growth factor; PPIase, Peptidyl-prolyl cis/trans isomerase; Prx2, Peroxiredoxin-2; 3-NPA, 3-Nitropropionic acid; PLK2, Polo like kinase 2; VDAC1, Voltage-dependent anion-selective channel 1; DLP-1, Dynamin-like protein 1; Drp1, Dynamin-related protein 1; E-2609, Elenbacestat; ACAT, Sterol O-acyltransferase; HDAC, Histone deacetylase; ROS, Reactive oxygen species; NEDD4, Neural precursor cell expressed developmentally down-regulated protein 4; PIA2, Phytochrome-interacting ankyrin-repeat protein 2; SNX33, Sorting Nexin 33; DNMTs, DNA methyltransferases; TET-1, Ten-eleven translocation methylcytosine dioxygenase 1.

* Corresponding author.
E-mail address: pravirkumar@dtu.ac.in (P. Kumar).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.arr.2021.101336
Received 28 October 2020; Received in revised form 10 March 2021; Accepted 22 March 2021
Available online 26 March 2021


The power of the Web of Science™ on your mobile device, wherever inspiration strikes.
Dismiss
Learn More
Already have a manuscript?
Use our Manuscript Matcher to find the best relevant journals!
Find a Match
Refine Your Search Results
1568-1637
Search
Sort By: Title (A-Z)
Search Results
Found 1 results (Page 1)
Share These Results
Exact Match Found
Filters
Clear All
Web of Science Coverage
Open Access
Category
Country / Region
Language
Frequency
AGEING RESEARCH REVIEWS
Publisher: ELSEVIER IRELAND LTD , ELSEVIER HOUSE, BROOKVALE PLAZA, EAST PARK SHANNON, CO, CLARE, IRELAND, 00000
ISSN / eISSN: 1568-1637 / 1872-9649
Web of Science Core Collection: Science Citation Index Expanded
Additional Web of Science Indexes: BIOSIS Previews | BIOSIS Reviews Reports And Meetings | Essential Science Indicators
Share This Journal
View profile page


Journal Citation Report™ (JCR)
Journal Citation Reports™ 2020
Journal Impact Factor™
2019
Deciding where to submit your manuscript? You need to know more than just one number.
If you have access to Journal Citation Reports™ through your institution's subscription, you can view the latest Journal Impact Factor™ and additional metrics to better understand a journal's content and audience.
Categories:
Cell Biology | Geriatrics & Gerontology
2018
10.390
Categories:
Geriatrics & Gerontology | Cell Biology
Learn About Journal Citation Reports™
View in Journal Citation Reports™

Journal publication 2-:

Title: "Artificial intelligence to deep learning: machine intelligence approach for drug discovery"**, Journal:** Springer's Molecular Diversity**, Impact factor:** 2.013**, Authorship:** Joint first author

Molecular Diversity
https://doi.org/10.1007/s11030-021-10217-3

Artificial intelligence to deep learning: machine intelligence approach for drug discovery

Rohan Gupta[1] · Devesh Srivastava[1] · Mehar Sahu[1] · Swati Tiwari[1] · Rashmi K. Ambasta[1] · Pravir Kumar[1]

Received: 29 January 2021 / Accepted: 22 March 2021

Abstract
Drug designing and development is an important area of research for pharmaceutical companies and chemical scientists. However, low efficacy, off-target delivery, time consumption, and high cost impose a hurdle and challenges that impact drug design and discovery. Further, complex and big data from genomics, proteomics, microarray data, and clinical trials also impose an obstacle in the drug discovery pipeline. Artificial intelligence and machine learning technology play a crucial role in drug discovery and development. In other words, artificial neural networks and deep learning algorithms have modernized the area. Machine learning and deep learning algorithms have been implemented in several drug discovery processes such as peptide synthesis, structure-based virtual screening, ligand-based virtual screening, toxicity prediction, drug monitoring and release, pharmacophore modeling, quantitative structure–activity relationship, drug repositioning, polypharmacology, and physiochemical activity. Evidence from the past strengthens the implementation of artificial intelligence and deep learning in this field. Moreover, novel data mining, curation, and management techniques provided critical support to recently developed modeling algorithms. In summary, artificial intelligence and deep learning advancements provide an excellent opportunity for rational drug design and discovery process, which will eventually impact mankind.

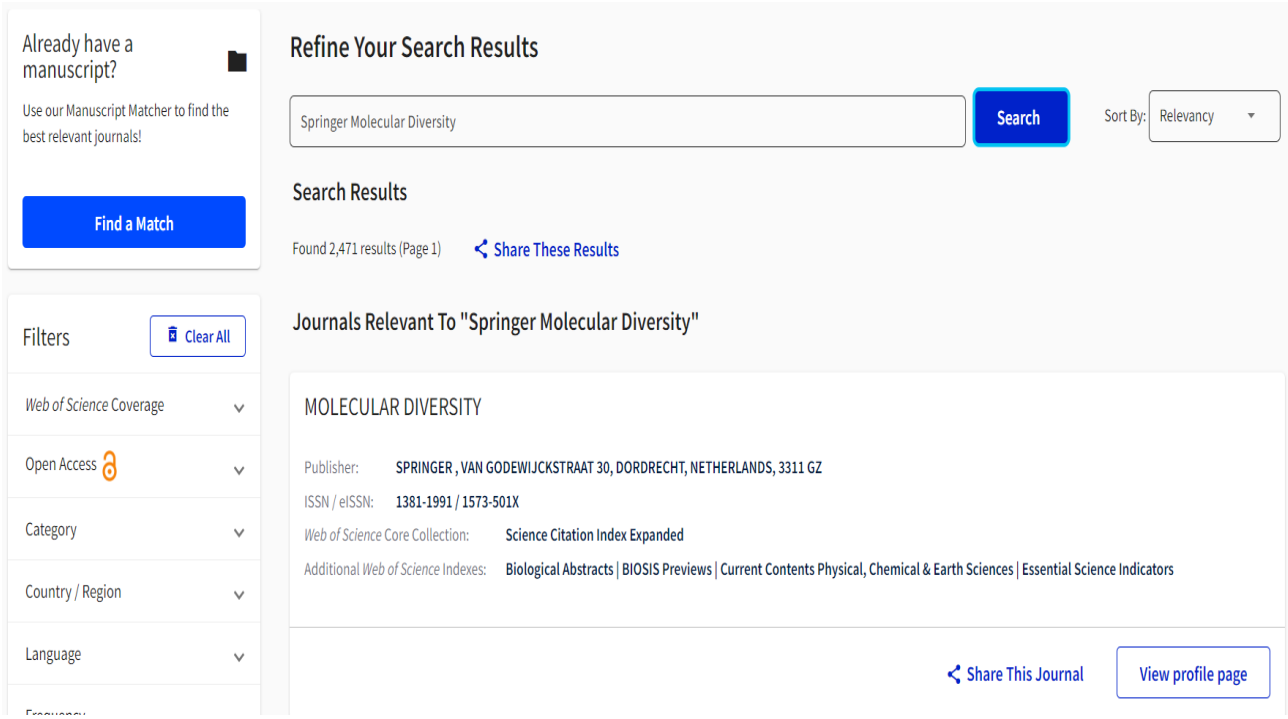
Graphic abstract
The primary concern associated with drug design and development is time consumption and production cost. Further, inefficiency, inaccurate target delivery, and inappropriate dosage are other hurdles that inhibit the process of drug delivery and development. With advancements in technology, computer-aided drug design integrating artificial intelligence algorithms can eliminate the challenges and hurdles of traditional drug design and development. Artificial intelligence is referred to as superset comprising machine learning, whereas machine learning comprises supervised learning, unsupervised learning, and reinforcement learning. Further, deep learning, a subset of machine learning, has been extensively implemented in drug design and development. The artificial neural network, deep neural network, support vector machines, classification and regression, generative adversarial networks, symbolic learning, and meta-learning are examples of the algorithms applied to the drug design and discovery process. Artificial intelligence has been applied to different areas of drug design and development process, such as from peptide synthesis to molecule design, virtual screening to molecular docking, quantitative structure–activity relationship to drug repositioning, protein misfolding to protein–protein interactions, and molecular pathway identification to polypharmacology. Artificial intelligence principles have been applied to the classification of active and inactive, monitoring drug release, pre-clinical and clinical development, primary and secondary drug screening, biomarker

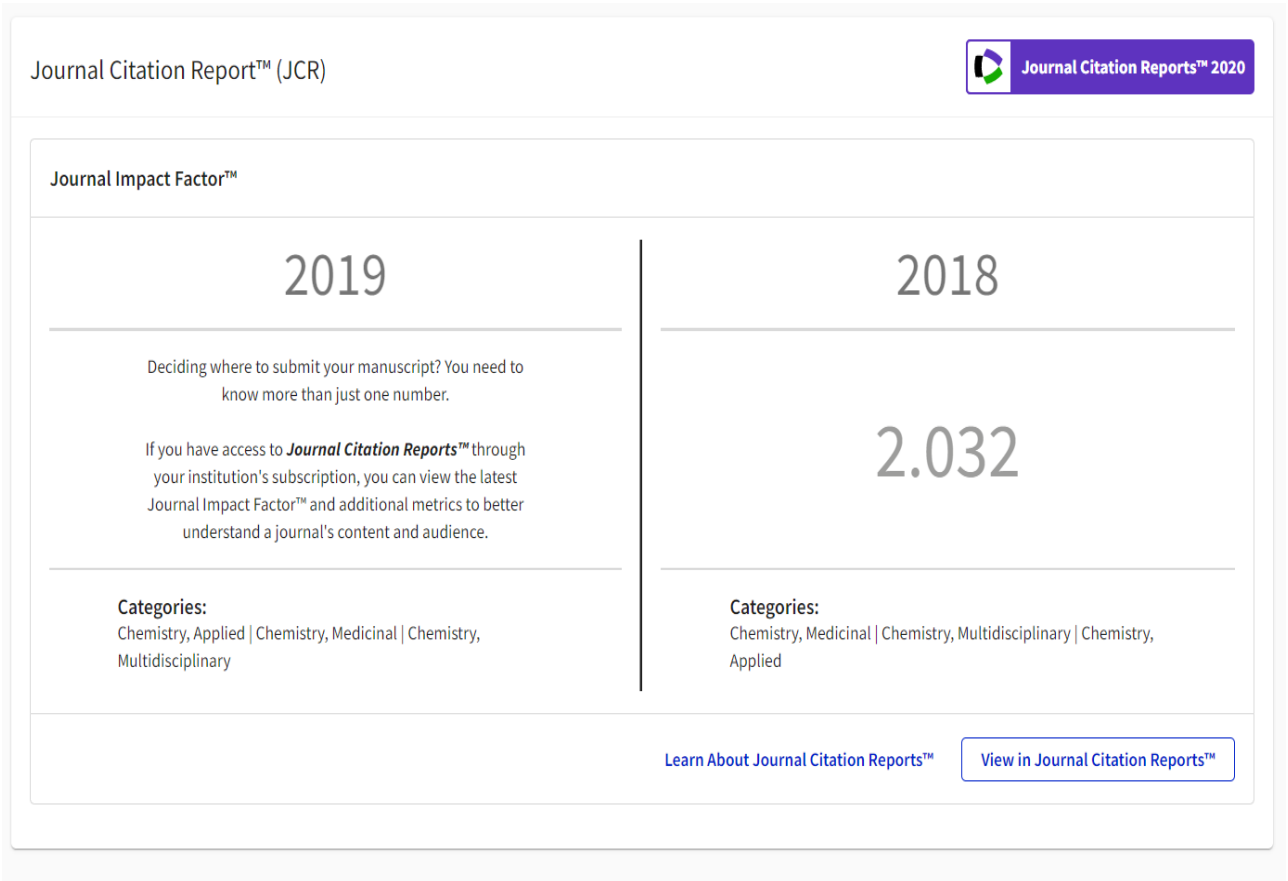
Rohan Gupta, Devesh Srivastava, Mehar Sahu, and Swati Tiwari contributed equally to this work.

✉ Pravir Kumar
pravirkumar@dtu.ac.in; kpravir@gmail.com

[1] Molecular Neuroscience and Functional Genomics Laboratory, Department of Biotechnology, Delhi Technological University (Formerly DCE), Shahbad Daulatpur, Bawana Road, Delhi 110042, India

Published online: 12 April 2021


Already have a manuscript?
Use our Manuscript Matcher to find the best relevant journals!
Find a Match
Refine Your Search Results
Springer Molecular Diversity
Search
Sort By: Relevancy
Search Results
Found 2,471 results (Page 1)
Share These Results
Filters
Clear All
Web of Science Coverage
Open Access
Category
Country / Region
Language
Journals Relevant To "Springer Molecular Diversity"
MOLECULAR DIVERSITY
Publisher: SPRINGER , VAN GODEWIJCKSTRAAT 30, DORDRECHT, NETHERLANDS, 3311 GZ
ISSN / eISSN: 1381-1991 / 1573-501X
Web of Science Core Collection: Science Citation Index Expanded
Additional Web of Science Indexes: Biological Abstracts | BIOSIS Previews | Current Contents Physical, Chemical & Earth Sciences | Essential Science Indicators
Share This Journal
View profile page


Journal Citation Report™ (JCR)
Journal Citation Reports™ 2020
Journal Impact Factor™
2019
Deciding where to submit your manuscript? You need to know more than just one number.
If you have access to Journal Citation Reports™ through your institution's subscription, you can view the latest Journal Impact Factor™ and additional metrics to better understand a journal's content and audience.
Categories:
Chemistry, Applied | Chemistry, Medicinal | Chemistry, Multidisciplinary
2018
2.032
Categories:
Chemistry, Medicinal | Chemistry, Multidisciplinary | Chemistry, Applied
Learn About Journal Citation Reports™
View in Journal Citation Reports™

29th June 2021

Letter Of Acceptance

Abstract id: **ICETSEM_2806827169**

Abstract Title: **In silico approach towards Parkinson's disease pathophysiology, drug repurposing and post translational modifications**

Author: **Devesh Srivastava**

Co-Authors: **Pravir Kumar**

Dear Mr/Ms Devesh Srivastava,

Congratulations!!

The scientific research paper reviewing committee of **" 2nd International Conference on Emerging Trends in Science, Engineering and Management-2021**" (ICETSEM-2021) scheduled to take place on the **15th and 16th July 2021** organized by GM Institute of Technology, Davangere, Karnataka & IFERP is pleased to inform your research paper titled "**In silico approach towards Parkinson's disease pathophysiology, drug repurposing and post translational modifications"** has been accepted after our double-blind peer review process for presenting paper at ICETSEM 2021. Authors and speakers are recommended to proceed for registration to confirm their slots in relevant scientific sessions by following the link: https://icetsem.net/conference-registration.php

Acceptance Status: Yes

For further details and other affiliated journals feel free to contact us,

***Additional Charges Applicable for publications in Scopus and Web of Science indexed Journals**

Registration Guidelines: https://icetsem.net/conference-registration-guidelines.php

No of authors per paper is limited, only 5 authors per paper. If Exceeds have to pay Rs.200 per head

Share your payment document after registering at https://icetsem.net/conference-registration.php and email the payment document to info@icetsem.net

Regards,
Dr.Praveen.J.,
IQAC Director, Professor & Head(ECE),
Convener, ICETSEM 2021,
GM Institute of Technology, Davnagere,Karnataka

K.Gomathi,
Conference Coordinator, IFERP,
+918122268465.