2013

A Network View on Neurodegenerative Disorders

Sreedevi Chandrasekaran
Virginia Commonwealth University

School of Life Sciences
Virginia Commonwealth University

This is to certify that the dissertation prepared by Sreedevi Chandrasekaran entitled "A Network View on Neurodegenerative Disorders" has been approved by his or her committee as satisfactory completion of the dissertation requirement for the degree of Doctor of Philosophy.

Danail G. Bonchev, Ph.D., D. Sc., School of Life Sciences

Tarynn M. Witten, Ph.D., LCSW, FGSA, School of Life Sciences

Vladimir I. Vladimirov, M.D., Ph.D., Department of Psychiatry, School of Medicine

---

Peter Uetz, Ph.D., School of Life Sciences

William B. Eggleston Jr., Ph.D., ILS Director, School of Life Sciences

---

Thomas F. Huff, Ph.D., Vice Provost, School of Life Sciences

F. Douglas Boudinot, Ph.D., Dean, School of Graduate Studies

May 1, 2013

A NETWORK VIEW ON NEURODEGENERATIVE DISORDERS

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at Virginia Commonwealth University.

by

SREEDEVI CHANDRASEKARAN
Master of Science, University of Oklahoma, 2006

Advisor: Danail G. Bonchev, Ph.D., D. Sc.
Professor, School of Life Sciences

Virginia Commonwealth University
Richmond, Virginia
May 2013

Acknowledgments

First and foremost, I like to thank my thesis advisor Dr. Danail Bonchev, who introduced me to Network biology, and whose enthusiasm for “molecular structures” had lasting effect. Without his guidance and persistent help this thesis would not have been possible. I am very grateful for all his valuable advice and for giving me a nurturing environment to grow and think freely. His calm and positive approach to various problems encountered along the way made my doctoral work far more enjoyable than otherwise. I am most grateful to have benefited from the supervision of such a distinguished academic and acknowledged expert in the field. I owe my deepest gratitude and respect to him for his continued support and direction.

I like to extend my thankfulness to my thesis committee members Dr. Tarynn Witten, Dr. Peter Uetz and Dr. Vladimir Vladimirov for their valuable expertise and insightful comments of my work. I wish I had more time to collaborate with them. Their guidance has directed me to think more critically and clearly. Also many thanks to Dr. Robert Tombes, Dr. William Eggleston and all administration personnel at Life Sciences department for their support and guidance.

I also thank most sincerely Late Dr. Anthony Marmarou, Dr. Harold Young and Dr. Shumei Sun for supporting me and my education. I am most thankful to my colleagues Jim, Brian, Lori, Paul and others at ABIC and IEC for their help and cooperation. I am very grateful for the personal encouragement and academic insights

of my good friends and colleagues Drs. Juan Lu and Gunes Aygok. Heartfelt thanks to Joy and Frank for their love and support. I am indebted to Christiane and James and also Malathi and Rangarajan for offering me a second home on so many occasions, and for supporting me in so many ways. I am thankful and happy to have so many encouraging and cherishing friends Singaravelan, Gayathri, Vidya, Bhuvana, Ramya, Anita, Lavanya and Lalitha.

Finally, my deepest thanks to my parents, Indira and Chandrasekaran, for inspiring me at every walk of my life. Without them and their unconditional love none my dreams would have come true. I owe my heartfelt gratitude to them for allowing me to be ambitious and adventurous. My brothers, Diwakar and Narendar have been motivating and instrumental in shaping my life. Thanks to their loving spouses and wonderful children. I also thank my parents-in-laws, Vijayalakshmi and Ganesan and my extended family members for their support and love. Finally, thanks to my loving husband Ayyappan, for sharing my dreams and always being there for me.

Dedication

To my beloved mother and father

Table of Contents

List of Tables

List of Figures

List of abbreviations

A-beta . . . . . amyloid-beta

AD . . . . . Alzheimer's disease

CSPNW . . . . . compact version of shortest-path network

DA . . . . . dopaminergic

DAVID . . . . . Database for Annotation, Visualization and Integrated Discovery

DI . . . . . Direct interaction

EBI . . . . . European Bioinformatics Institute's

GEO . . . . . Gene Expression Omnibus

GO . . . . . Gene Ontology

HD . . . . . Huntington's disease

HTT . . . . . Huntingtin

KEGG . . . . . Kyoto Encyclopedia of Genes and Genomes

LBs . . . . . Lewy bodies

miRNAs . . . . microRNAs

MRN . . . . . MicroRNA regulatory network

NCBI . . . . . National Center for Biotechnology Information's

NDDs . . . . . Neurodegenerative disorders

NFTs . . . . . Neurofibrillary tangles

OMIM . . . . . Online Mendelian Inheritance in Man database

PD . . . . . Parkinson's disease

PPI . . . . . protein-protein interaction

RMA . . . . . Robust multi-array average

SDEGs . . . . . significantly differentially expressed genes

SN . . . . . Substania nigra

SNCA . . . . . alpha-synuclein

SPNW . . . . . Shortest-path network

Abstract

A NETWORK VIEW ON NEURODEGENERATIVE DISORDERS

By Sreedevi Chandrasekaran, M.S.

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at Virginia Commonwealth University.

Virginia Commonwealth University, 2013

Advisor: Danail G. Bonchev, Ph.D., D. Sc.
Professor, School of Life Sciences

Neurodegeneration is a chronic, progressive and debilitating condition that affects majority of the World's elderly population who are at greater risk. Numerous scientific studies suggest that there could be a common underlying molecular mechanism that promotes the degeneration and the subsequent neuronal loss, however so far the progress in this direction is rather limited. Abnormal protein misfoldings, as well as protein plaque formations in the brain, are some of the hallmark characteristic features of neurodegenerative disorders (NDDs). Genetic and environmental factors, oxidative stress, excessive reactive oxygen species formation, mitochondrial dysfunction, energy depletion and autophagy disruption etc. are some of the widely suspected mechanisms that manifest the cognitive, motor and emotional symptoms

of these NDDs. Motivated by some molecular traits found in common in several NDDs, network-based systems biology tools and techniques were used in this study to identify critical molecular players and underlying biological processes that are common for Parkinson's, Alzheimer's and Huntington's disease. Utilizing multiple microarray gene expression datasets, several biomolecular networks such as direct interaction, shortest path, and microRNA regulatory networks were constructed and analyzed for each of the disease conditions. The network-based analysis revealed 26 genes of potential interest in Parkinson's, 16 in Alzheimer's and 30 in Huntington's disease. Many new microRNA-target regulatory interactions were identified. For each disorder, several routes for possible disease initiation and protection scenarios were uncovered. A unified neurodegeneration mechanism network was constructed by utilizing the significantly differentially expressed genes found in common in Parkinson's, Alzheimer's and Huntington's microarray datasets. In this integrated network many key molecular partakers and several biological processes that were significantly affected in all three NDDs were uncovered. The integrated network also revealed complex dual-level interactions that occur between disease contributing and protecting entities. Possibilities of microRNA-target interactions were explored and many such pairs of potential interest in NDDs were suggested. Investigating the integrated network mechanism, we have identified several routes for disease initiating, as well as alleviating ones that could be utilized in common for Parkinson's, Alzheimer's and Huntington's disease. Finding such crucial and universal molecular players in addition to maintaining a delicate balance between neurodegeneration promoters and protectors is vital for restoring the homeostasis in the three NDDs.

Chapter 1

Introduction

Neurodegeneration is a collection of neurological diseases or medical conditions that occur when nerve cells die or lose their capacity to function normally. Diseases like Alzheimer's, Parkinson's, Huntington's, Amyotrophic lateral sclerosis etc., exhibit loss of neurons that affects multiple facets of basic daily living functions including memory loss, cognitive decline, mood disorders, difficulty in swallowing, walking and speech. Currently there are medications and surgical treatments available to mitigate the symptoms but very little intervention is provided towards halting the disease progression. Neurodegeneration is a chronic, progressive and debilitating condition that ultimately is fatal.

As a result of advances in medicine and improvements in quality of life, it is quite impressive to note that the average human lifespan has increased considerably in the past century. It would be ideal to spend this long life healthy and independent. But the aging population is at the greatest risk for neurodegenerative diseases. As of 2012, one in eight older (age ≥ 65) American has been diagnosed with Alzheimer's disease (AD). It has been estimated that by 2050, there would be one new case of AD in every 33 seconds (Bleiler and Laura, 2012). There is an urgent need to

understand these complex neurodegeneration molecular processes so as to identify effective therapeutic measures to save our growing elderly population.

Even though each neurodegenerative disorder has distinct clinical phenotypical presentations, there is a growing interest toward searching for a unified underlying mechanism of degeneration. Such a mechanism is suggested to include dysfunction in protein folding and aggregation, oxidative stress, free radical formation, and mitochondrial function etc.(Chiti and Dobson, 2006; Emerit et al., 2004; Schon and Manfredi, 2003; Jellinger, 2010). Various genetic, environmental and endogenous factors were suspected to contribute to the deregulation of these biological processes that eventually manifests a fatal neurodengeneration. Discovering these kinds of common mechanisms offers hope for simultaneous therapeutic advancement for these diseases.

In this study, we aim to approach the problem for the supposed unified mechanism of neurodegenerative diseases proceeding from a systems biology approach, which analyzes the underlying processes in their entirety and interdependence. Molecular biology networks are the ideal tool for such global approach. Network analysis offers realistic chances to identify the most relevant of molecular pathways involved in neurodegeneration, as well as the critical molecular players in these pathways. An important result of such an analysis is finding potential candidates for new drugs, which would suppress the expression of disease causing genes.

We selected for our network analysis three widely spread and well-studied neurodegenerative diseases: Parkinson's, Alzheimer's and Huntington's disease. In the following few paragraphs, we summarize the available information on clinical symptoms, genes implicated in the disease pathogenesis, known disease signaling pathways, current therapeutic measures, etc. for these three neurodegenerative disorders. This information resulted from a broad literature search, in which the Online

Mendelian Inheritance in Man database (OMIM) catalog of human genes and genetic disorders (http://www.omim.org/) was the basic source about the genes already implicated in each of these neurodegenerative diseases.

1.1 Parkinson's disease (PD)

1.1.1 Introductory notes

James Parkinson has been the first to observe this disease in adults in the year 1817. In his essay entitled "An Essay of the Shaking Palsy" he described this disease as initiated with slow, progressive involuntary tremors, followed by difficulty in walking, swallowing and speech (Parkinson, 2002). Apart from motor symptoms, Parkinson's disease patients experienced significant non-motor symptoms including mood and cognition decline, sleep disturbances, and other autonomic dysfunctions (Micieli et al., 2003).

With the help of modern-day molecular and cellular research advancement, progressive degeneration of the dopaminergic (DA) neurons of the Substania nigra (SN) brain region were found in Parkinson's disease (PD) brains (Jellinger, 2009), in addition to the accumulation of misfolded protein aggregates. Both environmental factors and genetic mutations were suspected to cause PD (Di Monte et al., 2002; Lesage and Brice, 2009). One of the distinctive features of Parkinson's disease is severe damage to the nigrostriatal dopaminergic system. Neurotoxic agents such as manganese and 1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine (MPTP) were suspected for this type of neuronal damage. MPTP induced Parkinson's disease animal models were extensively used to study the neurodegeneration process as well as to identify potential therapeutic drug targets (Sedelis et al., 2001). Soluble fractalkine

(CX3CL1, chemokine ligand 1) isoform was shown to reduce impairment of motor coordination, decrease dopaminergic neuron loss, and ameliorate microglial (macrophages of brain) activation and proinflammatory cytokine release resulting from MPTP exposure (Morganti et al., 2012).

1.1.2 PD etiology

Long time belief was that Parkinson's disease etiology is sporadic (not genetically inherited) in nature. However, a small percentage of the PD patients were now known to inherit gene mutations. Genes including ATP13A2, DJ-1, GIGYF2, HTRA2, LRRK2, PARK2 (parkin), PINK1, SNCA and UCHL1 were associated with either autosomal dominant or recessive form of Parkinson's disease (Lesage and Brice, 2009). From the listed genes alpha-synuclein (SNCA) is critical to the pathogenesis in the early-onset of the rare familial form of PD. Insoluble form of α-syn fibrils were discovered in the protein aggregates called Lewy bodies (LBs), the hallmark pathological characteristics of Parkinson's disease. The aggregation and accumulation of abnormal α-syn in dopaminergic neurons have been postulated to be responsible for the neurodegeneration that ultimately leading to cell death (George, 2002; Recchia et al., 2004). Synucleins were also found in the amyloid-plaques in Alzheimer's disease brains.

In general, alpha-synuclein is highly expressed in brain at presynaptic terminals, particularly in the neocortex, hippocampus, striatum, thalamus, and cerebellum components. They function as molecular chaperones and interact with many proteins thus modifying their cellular activity. Due to its versatile interacting behavior, mutant alpha-synuclein has been implicated in the deregulation of many biological processes including oxidation, neuroinflammation, mitochondrial func-

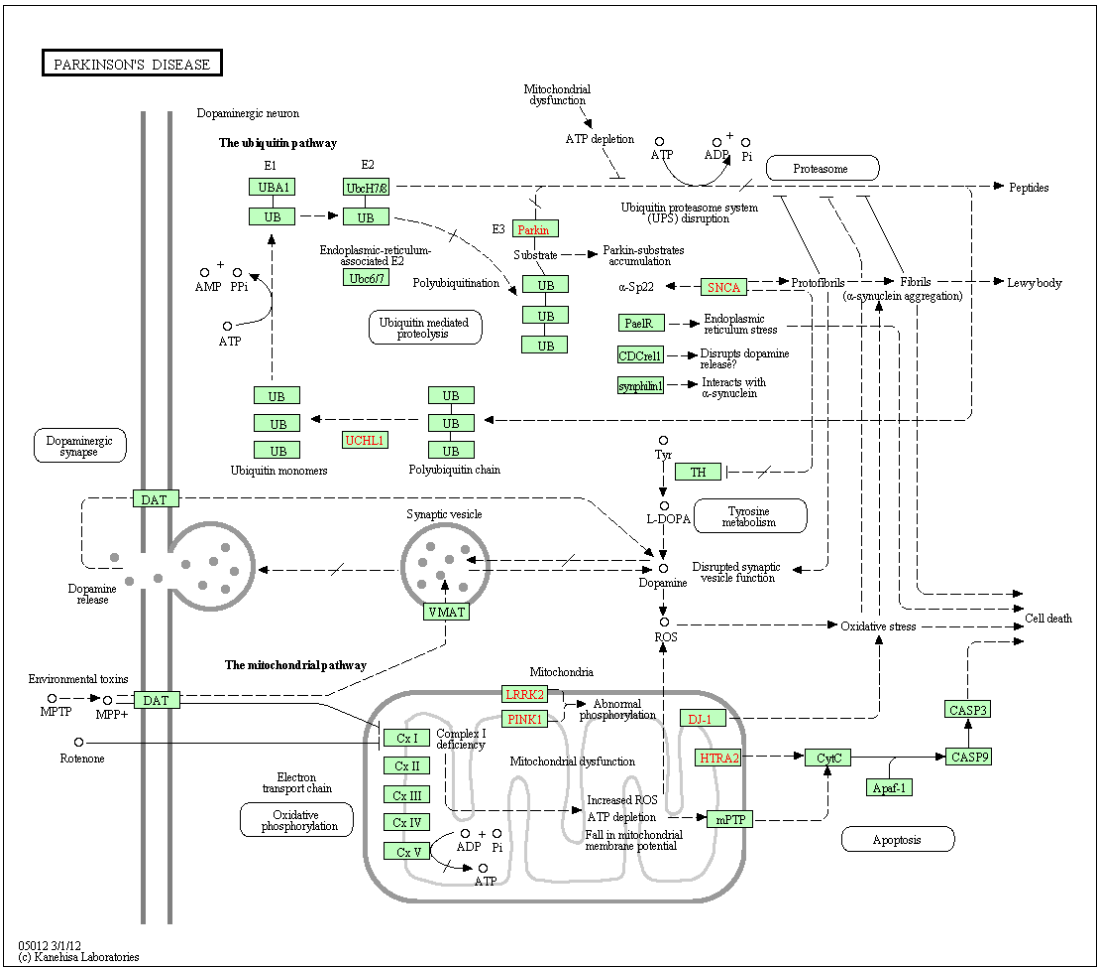


Figure 1.1. Parkinson's disease pathway from KEGG database. Biological processes and genes implicated in the Parkinson's disease. Courtesy: Parkinson's disease pathway from KEGG database, available at `http://www.genome.jp/kegg/pathway/hsa/hsa05012.html` retrieved on Apr 3, 2013.

tion, ubiquitination etc. (Jellinger, 2009, 2010; Hsu et al., 2000; Polymeropoulos et al., 1997). Figure 1.1 depicts the various genes already implicated in Parkinson's disease along with different deregulated biological processes caused by the several abnormal protein activities.

To date, several genetic modifiers of PD have been described. Information of many such genes and their role in PD pathogenesis can be found in literature (Klein and Schlossmacher, 2006; Gasser, 2009; Shulman et al., 2011; Farrer, 2006; von

Bohlen und Halbach et al., 2004). Some of these PD-known genes relate to neuronal growth and neuroprotective mechanisms in Parkinson's disease (refer Table 1.1).

Recent genome-wide studies have found that mutations in at least 13 PARK loci and related genes increase both early- and late-onset PD susceptibility (Hicks et al., 2002; Chung et al., 2011; Klein and Schlossmacher, 2006). Earlier study by Galvin et al. (1999) had shown that β- and γ-synuclein are associated with hippocampal axon pathology in Parkinson's disease and dementia with Lewy bodies. Genome-wide approaches were also used to identify microRNAs-target mRNA interactions in PD domain. MicroRNAs (miRNAs) are a class of small RNAs (22 nucleotides) that act as post-transcriptional regulators of gene expression by binding to the complementary sequences in target mRNAs. In recent years, miRNAs have emerged as potential drug targets in a variety of diseases including infections, metabolism and inflammation etc (van Rooij et al., 2012). A recent genome-wide miRNA profiling study for Parkinson's disease has reported several miRNAs to be differentially expressed in PD blood samples. Hundreds of genes were reported as targets of these miRNAs. The predicted target genes belonged to various biological pathways including synaptic long-term potentiation, semaphorin signaling in neurons and protein ubiquination pathway etc. many of which were previously found deregulated in Parkinson's disease mechanism (Martins et al., 2011).

1.1.3 PD treatment options

Even though there were some new treatment options available to PD patients, oral administration of levodopa (precursor of dopamine) has been the gold standard medication for Parkinson's disease. But prolonged use of levodopa increases the risk of developing levodopa-induced dyskinesias (involuntary movement) (Rascol

Table 1.1. Some of the well-known Parkinson's disease genes and their molecular functions.

Known PD genes	Functions (references)
CX3CL1 (fractalkine)	Produced by neurons, suppresses the activation of microglia and plays a neuroprotective role in 6-OHDA-induced (synthetic neurotoxic compound) dopaminergic lesions (Pabon et al., 2011).
FGFs (fibroblast growth factors)	Exhibits potent neurotrophic properties for dopaminergic (DA) neurons. Promote DA neuron's development and neurite outgrowth, rescue damaged DA neurons after toxic insults, and prevent apoptosis (Walker et al., 1998).
L1CAM (L1 cell adhesion molecule)	Enhances the survival of imperiled endogenous dopaminergic neurons in Substantia nigra (SN) (Cui et al., 2010).
MAPK signaling pathways	Contributes to neuroinflammatory responses and neuronal death which is triggered by α-syn aggregates or functional deficiencies in parkin or DJ-1 genes (Kim and Choi, 2010).
MT1F, MT2A (Metallothioneins 1 and 2)	Scavenges reactive oxygen species and free radicals in central nervous system (Michael et al., 2011).
RAB3A (member of RAS oncogene family)	Suppresses α-syn toxicity in neuronal models of PD (Gitler et al., 2008).
RNF11 (ring finger protein 11)	Found highly enriched in SN dopamergic neurons as well as its co-localizes with Lewy bodies (abnormal aggregates of protein) in PD brains (Anderson et al., 2007).

et al., 2011; Thanvi et al., 2007). Recently, deep brain stimulation (DBS) has been offered as a secondary treatment option in Parkinson's disease where the benefits of medication have failed/diminished. DBS therapy has been shown to increase the neuron firing rate, blood flow and to promote neurotransmitter release as wells as to stimulate neurogenesis. Although deep brain stimulation improves the motor symptoms of Parkinson's disease, it is a serious surgical intervention with major side effects of infection and intracranial hemorrhage including the risk of death (Okun, 2012). With current advancement of different "omics" technologies along with effective in-silico testing options, finding successful molecular therapeutic targets for Parkinson's disease seems much more possible than before.

1.2 Alzheimer's disease (AD)

According to a recent US Alzheimer's association's report, 5.4 million Americans of all ages had Alzheimer's disease in 2012 and the numbers are estimated to increase in the years to come. Alzheimer's is the most common form of dementia (serious progressive loss of cognitive functions). Aging, diet, lifestyle, heart problems, family history and gene mutations etc. have been suggested to increase the risk for developing the disease (Bleiler and Laura, 2012).

1.2.1 AD etiology

Unlike Parkinson's disease, memory loss is one of the first symptoms in Alzheimer's. At first, patients with AD start to forget names, things and words etc. which then manifests into confusion and disorientation followed by slow and progressive decline in standard of daily living. AD is ultimately fatal. In general, AD etiology is spo-

radic but a small percentage of early-onset familial form is also possible. Mutations in APP, APOE, PSEN1, PSEN2, and MAPT genes were found to cause Alzheimer's disease pathogenesis. Histopathologically, AD is characterized by the presence of amyloid-plaques (which consist of amyloid-β peptides) and Neurofibrillary tangles (NFTs) in the brain. APP encodes for the amyloid-β precursor protein, and the presenilin genes (PSEN1 and PSEN2) encodes for the proteolytic enzymes that cleave APP into amyloid-β and other fragments (Wolfe et al., 1999). PSEN1 mutations were attributed to more than 50% of early onset of familial form of AD (Raux et al., 2005). Neurofibrillary tangles consist of aggregations of hyperphosphorylated tau proteins (MAPT) (Selkoe, 2001; Bekris et al., 2010). Accumulation of amyloid-beta (A-beta) mediates migration of inflammatory molecules across blood brain barrier (BBB) and it is enhanced by PECAM1 adhesion, thus promoting the disease progression (Kalinowska and Losy, 2006).

Some of the other genes that were suspected to contribute the disease pathogenesis are shown in Table 1.2.

In addition, TGFβ1 has both protective and deleterious effect in AD. Overexpression of TGF-beta may initiate or promote amyloidogenesis in Alzheimer's disease. TGFβ1 has also been suggested to have an anti-amyloidogenic role; it decreases the Aβ load and the formation of neuritic plaques in brain (Tichauer and von Bernhardi, 2012). This neuroprotective effect of TGFβ1 appears to be at least partially mediated by the SMAD pathway (Ueberham et al., 2006). Another neuroprotective mechanism has been suggested to take place via Humanin-mediated JAK2/STAT3 signaling pathway. Activation of this pathway has reduced the memory and cognitive impairment found in Alzheimer's disease (Chiba et al., 2009; Niikura et al., 2004). Even after extensive biochemical research, little is known about how these protein aggregations trigger gradual memory loss and other symptoms in

Table 1.2. Some of the well-known Alzheimer's disease genes and their molecular functions.

Known AD genes	Functions (references)
CDK5 (cyclin-dependent kinase 5)	Along with GSK3B (glycogen synthase kinase-3beta) leads to accumulation of aberrantly phosphorylated forms of the microtubule-associated protein tau (Mateo et al., 2009).
CDK5R1 (cyclin-dependent kinase 5, regulatory subunit 1)	Activator of CDK5 which contributes to the hyperphosphorylated tau protein aggregate formations in AD brains (Mateo et al., 2009).
LMO4 (LIM domain only 4)	Play a secondary role in AD by increasing the complexity and the severity of the disease pathogenesis (Leuba et al., 2004).
PTEN (phosphatase and tensin homolog)	Found co-localized with abnormal tau and phosphorylated neurofilament proteins in neurons of AD brains (Sonoda et al., 2010).
APOE, A2M, LRP1, MPO, JUN, NOS1, SIRT1 ...	Found to increase the abnormal protein aggregation and other symptoms of Alzheimer's disease (Serretti et al., 2005).

Alzheimer's disease (Selkoe, 2001; Bekris et al., 2010). Apart from genes, the role of miRNA regulatory mechanisms was also studied in AD. A recent genome-wide miRNA profiling study has found a substantial number of differentially expressed miRNAs in the cortical region of AD brains. Many of these miRNAs and their predicted mRNA target pairs were part of several biological processes that were previously reported dysfunctional in Alzheimer's disease mechanism (Nunez-Iglesias et al., 2010).

1.2.2 AD deregulated biological processes

Many biological processes including neuroinflammation, oxidative stress, dysfunction of lysosomal/proteasomal degradation, mitochondrial dysfunction etc. have been associated with Alzheimer's disease. Figure 1.2 illustrates the different mutant

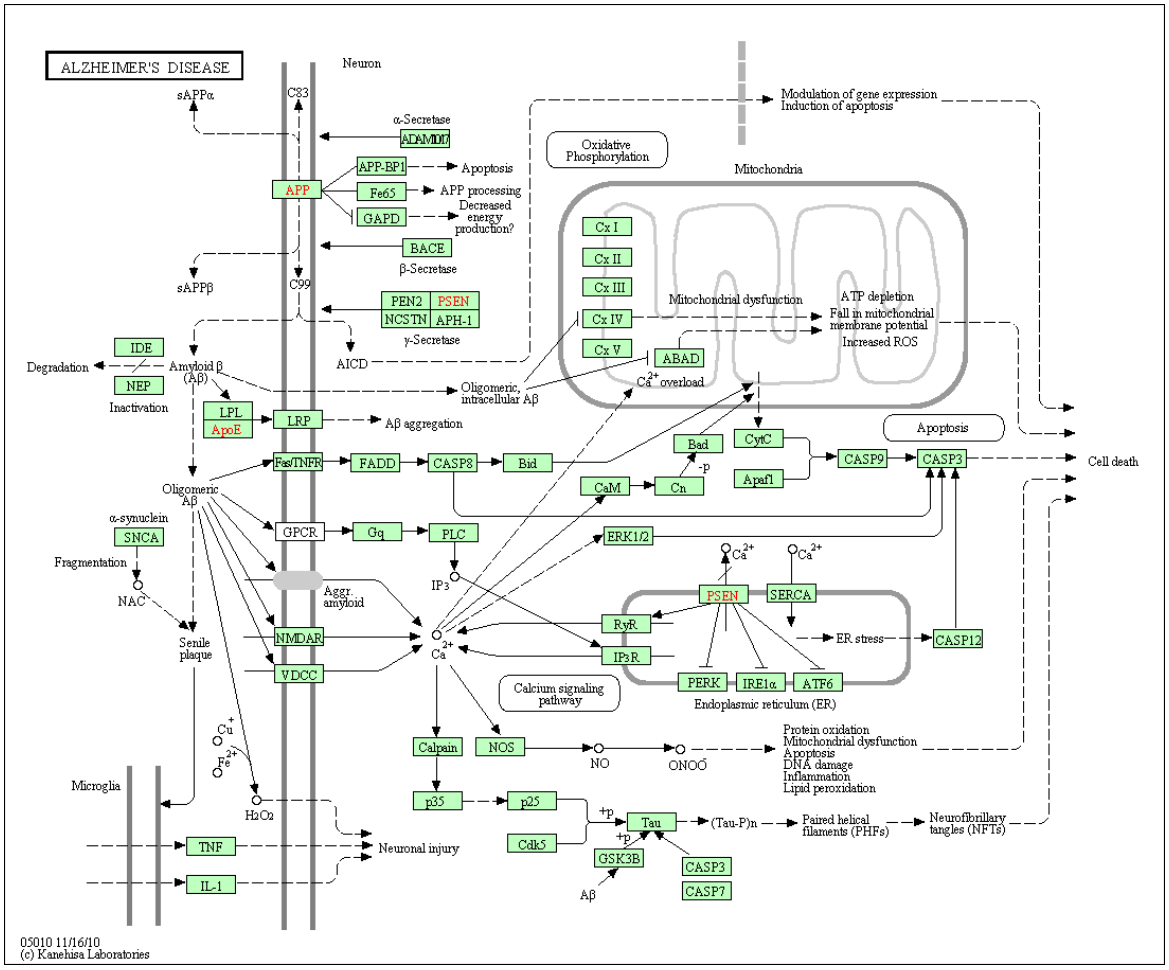


Figure 1.2. Alzheimer's disease pathway from KEGG database. Biological processes and genes implicated in the Alzheimer's disease. Courtesy: Alzheimer's disease pathway from KEGG database, available at http://www.genome.jp/kegg/pathway/hsa/hsa05010.html retrieved on Apr 3, 2013.

genes and its cellular localization along with the biological processes that are deregulated in Alzheimer's disease pathogenesis. Aggregation of amyloid-plaques and tau proteins were suggested to be the cause for these deregulations (Crews and Masliah, 2010; Doyle et al., 2011; Jomova et al., 2010; Selkoe, 2001; Tuppo and Arias, 2005; Jellinger, 2010). Blood and inflammatory markers, oxidative stress indicators, along with clinical examination, brain-imaging techniques, and CSF biochemical markers show promising steps towards a combinatorial biomarkers approach for early and accurate Alzheimer's disease diagnosis (Flirski and Sobow, 2005).

1.2.3 AD treatment options

Current thoughts of therapeutic measures in AD were along the lines of reducing APP production or by increasing the amyloid-beta clearance. These were suggested in order to protect the neurons from cell death and to promote synaptic formation and neurogenesis. Medications such as Donepezil and Memantine, mitochondrial-targeted antioxidants such as MitoQ, SS31 and neurotrophic factors like BDNF have been shown to reduce amyloid-β toxicity as well as reverse neuronal atrophy and improve age-related cognitive impairment (Manczak et al., 2010; Nagahara et al., 2009; Howard et al., 2012).

1.3 Huntington's disease (HD)

1.3.1 Introductory notes

Huntington's is an inherited autosomal dominant motor disorder. Expansion of 36 or more CAG trinucleotide (polyQ) repeats in Huntingtin (HTT) gene is the hallmark characteristic of the disease. PolyQ expanded HTT is considered as a trigger

of the neurodegeneration that eventually cause all the Huntington's disease (HD) symptoms. Offspring of an individual with a mutant allele have a 50% chance of inheriting the disease. Huntington's disease is described by progressive motor, cognitive, and emotional disturbances. The motor symptoms include chorea, dystonia, rigidity, postural instability etc. Depression and personality changes are the major emotional disturbances part of the disorder. Like Alzheimer's disease, short-term memory loss, confusion and disorientation are some of the cognitive issues found in HD patients. In HD, neurodegeneration is selective for striatal GABAergic medium-sized spiny neurons. These neurons project to substania nigra and globus pallidus parts of the brain affecting primarily the motor coordination (Bates, 2003; Reiner et al., 2011; Albin et al., 1990).

Figure 1.3 depicts HTT gene along with its interacting partners which trigger the striatal neuronal loss that eventually manifests all the Huntington's disease symptoms. Like the other neurodegenerative disorders, HD also showed protein misfolding, ubiquitin proteasome system deregulation, autophagy dysfunction, metabolic and mitochondrial dysfunction as well as oxidative stress, which over the years culminates into motor and cognitive disorders (Martinez-Vicente et al., 2010; Gil and Rego, 2008; Davies et al., 2007; Jellinger, 2010).

Apart from HTT, mutations in HDL3, JPH3 and PRNP genes were also related to Huntington's disease pathogenesis (OMIM database, retrieved on Dec 17, 2012). Some of the known HD related genes are shown in Table 1.3. 25 or more genes that interact with HTT and are part of Huntington's disease pathogenesis are listed in literature (`http://en.wikipedia.org/wiki/Huntingtin#cite_ref-pmid11532990_10-0`, retrieved on Apr 8, 2013).

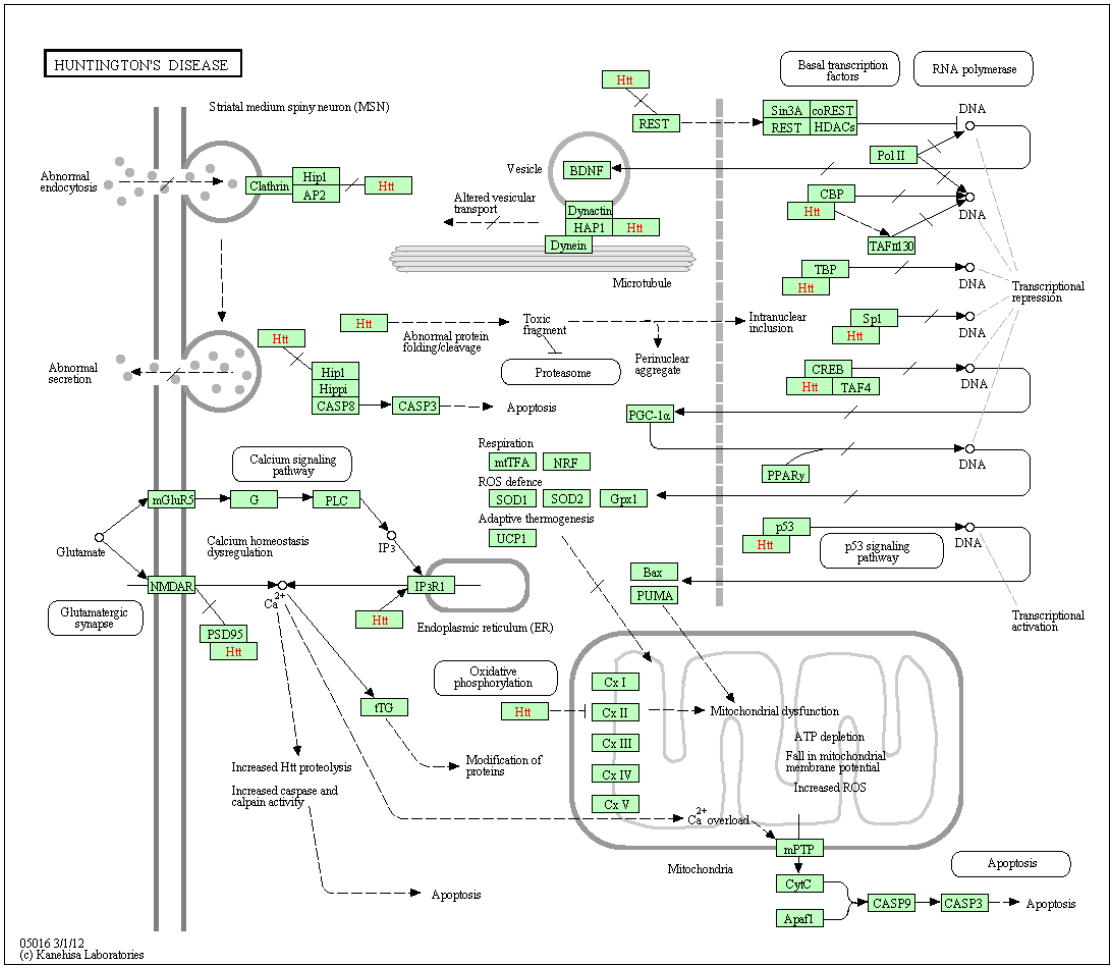


Figure 1.3. Huntington's disease pathway from KEGG database. Biological processes and genes implicated in the Huntington's disease. Courtesy: Huntington's disease pathway from KEGG database, available at http://www.genome.jp/kegg/pathway/hsa/hsa05016.html retrieved on Apr 3, 2013.

Table 1.3. Some of the well-known Huntington's disease genes and their molecular functions.

Known HD genes	Functions (references)
BCL2 (B-cell CLL/lymphoma 2)	Overexpression of BCL2 slows down the Huntington's disease progression (Mattson, 2000; Zhang et al., 2003).
CCKBR (cholecystokinin B receptor)	Encodes a G-protein coupled receptor for gastrin and cholecystokinin (CCK), the regulatory peptides of the brain and gastrointestinal tract. It was reported that there is selective loss of CCK receptor-containing neurons in cerebral cortex of Huntington's patients (Hays et al., 1981).
cytochrome c	It's release from the mitochondria triggers the downstream caspase activation leading to apoptotic neuronal death in many neurodegenerative diseases. This kind of neuronal death plays a greater role at the end stage of HD (Kiechle et al., 2002). An animal model of chronic HD neurodegeneration study has shown that methazolamide drug could inhibit cytochorome c release from mitochondria and thus acts as a neuroprotective agent (Wang et al., 2008).
FGF2 (fibroblast growth factor 2)	Improves motor performance and extends the lifespan by 20% by reducing the accumulation of polyglutamine aggregates in the brain (Jin et al., 2005; La Spada, 2005).
GAPDH (glyceraldehyde-3-phosphate dehydrogenase)	Due to its selective binding to the CAG repeats in huntingtin gene, GAPDH activity was found reduced in HD brains there by reducing the cellular energy production (Burke et al., 1996; Mazzola and Sirover, 2001).
GPCRs (G-protein coupled receptors)	Due to their abundant presence in central nervous system, as well as their complex interactions with many downstream targets, have made GPCRs as potential drug targets in many neurological diseases including Huntington's disease ((Dowie et al., 2010).
IRS2 (insulin receptor substrate 2)	Decreasing IRS2 signaling could be part of a therapeutic approach to slow down the progression of HD (Sadagurski et al., 2011).

1.3.2 HD potential therapeutic measures

Unlike Parkinson's and Alzheimer's neurodegenerative disorders, Huntington's disease main cause is a genetic defect. This has opened broad venues for developing effective animal models to study and understand the disease pathogenesis. With moderate success, these animal models have replicated HD related neuronal loss and the subsequent motor and cognitive symptoms of the disease. The animal research work have expanded the understanding of HD pathogenesis and thus suggested some valuable therapeutic measures to alleviate the disease symptoms. Some of the suggested beneficial pathways were by up-regulating the innate autophagy process and via increasing BDNF gene expression. Autophagy process protects against the toxic insults of mutant huntingtin proteins by enhancing its clearance from the cell.

Brain-derived neurotrophic factor (BDNF) is necessary for survival of striatal neurons in the brain and it promotes synaptic plasticity in addition to memory formation. It can also act as a neuromodulator affecting the pre-synaptic release of neurotransmitters in central nervous system. When administered systemically or delivered via genetically-grafted cells, BDNF has shown to prevent striatal neurons from cell death in HD animal models. BDNF has also been suggested to reduce amyloid-beta neurotoxicity in Alzheimer's disease (Ross and Tabrizi, 2011; Zuccato et al., 2010; Tapia-Arancibia et al., 2008; Ferrer et al., 2000). Currently, blood markers, brain imaging are used along with qualitative clinical measures as potential biomarkers to test for early-onset and/or the progression of Huntington's disease (Walker, 2007).

1.4 Network biology

Traditionally, medical practices have followed reductionist approach in treating diseases, based on the assumption that information about individual body parts is sufficient to explain the whole system. However, unpredicted behaviors evolve as a result of complex dynamic interactions between the parts, which cannot be explained by the parts alone. One needs to know the interactions of various parts in order to fully understand the system. Systems biology is a holistic approach to understand biological systems as a whole including its individual components and their interactions.

In an abstract level, behavior of complex systems (e.g., biological cell) can be represented as a network where the individual components (e.g., DNA, proteins, metabolites) of the system are the "nodes" of the network and their interactions (e.g., protein-DNA, protein-protein, and protein-metabolite) are the "edges/links" that connect the individual components. Network biology is a systems approach to study the internal workings of a biological system by representing its complexity in a network form. One such example is human protein-protein interaction (PPI) network shown in Figure 1.4. The variety of techniques developed in network biology makes possible to analyze and understand the internal molecular architecture and interactions of a biological cell (Barabási and Oltvai, 2004).

As shown in the previous subsections, not only individual genes but many biological processes are affected in each of the neurogenerative diseases. Neurodegeneration process is a complex system with many players and multiple interactions among them. Such systems can be analyzed in a network form to understand the underlying molecular mechanisms. Network analysis approach has been utilized to study complex diseases like cancer, virus-host interaction and others. (Cerami et al., 2010; Taylor et al., 2009; Uetz et al., 2006; Calderwood et al., 2007). With the ad-

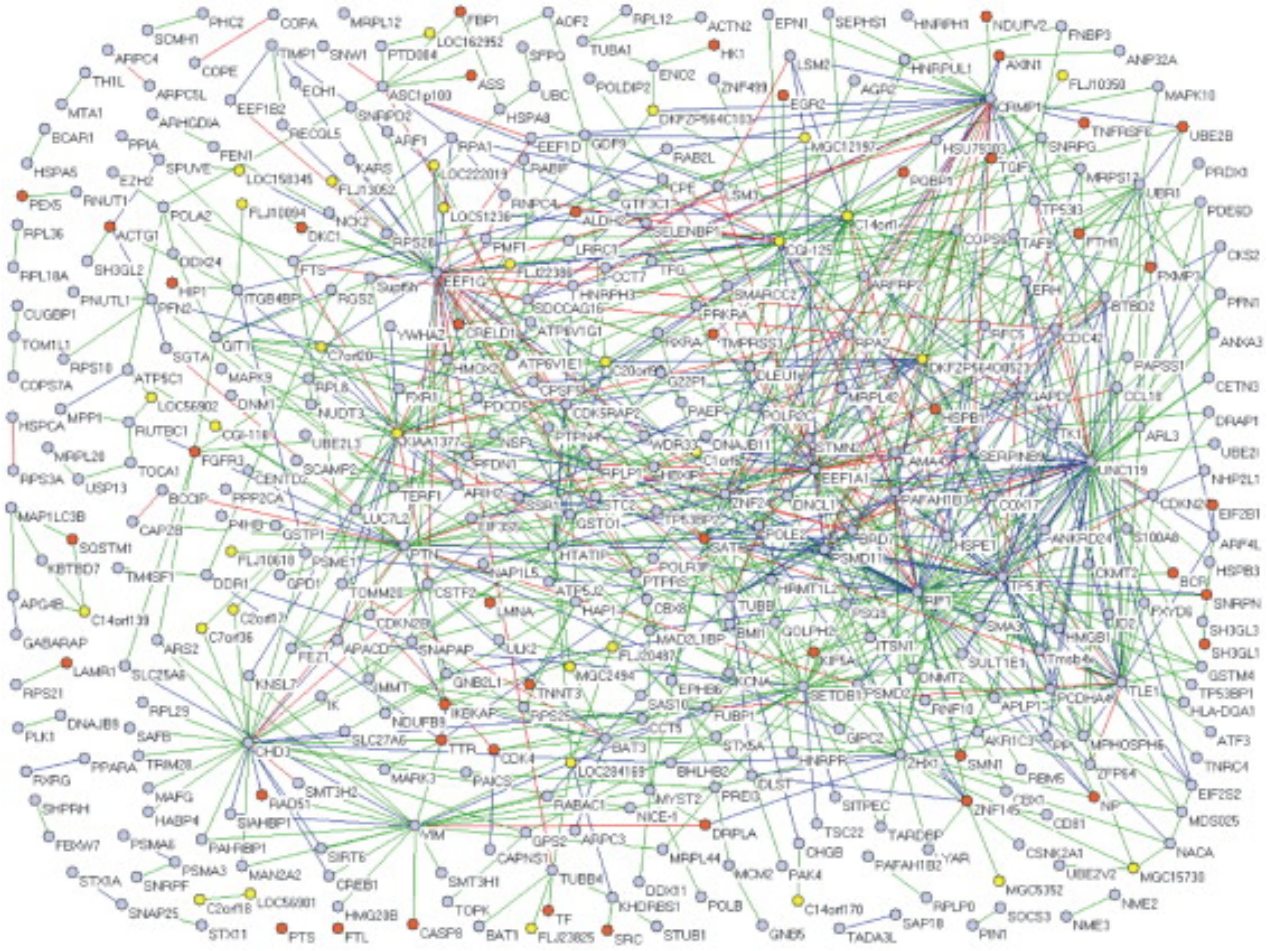


Figure 1.4. Network example. Shown here is the human protein-protein interaction network. Courtesy: A human protein-protein interaction network: a resource for annotating the proteome (Stelzl et al., 2005). Available at `http://www.sciencedirect.com/science/article/pii/S0092867405008664` retrieved on Apr 5, 2013.

vancement of high-throughput technologies like microarrays, protein and exon arrays we are now able to produce information about thousands of genes and proteins simultaneously. This has inspired a new wave of studies to understand complex disease mechanisms via network approach that involves all those genes/proteins along with their interactions. Network-based approaches have been suggested to provide a new and useful framework for classifying diseases and predicting their outcome, as well as for identifying therapeutic strategies (Loscalzo et al., 2007; Goh et al., 2007; Barabási et al., 2011).

1.4.1 Critical network measures

Network topology of biological systems not only provides means to visualize the complex interactions between the components, it also helps to mathematically derive specific scores/measures to identify key players within the network. Some of the important network descriptors are as follows. First is node *degree*. It is one of the basic characteristics of a network and it tells us how many links the given node has with other nodes. Nodes with highest degree are called "hubs". Studies have shown that highly connected nodes in a biological network tend be both essential and conserved (Bergmann et al., 2004). Next is network *distance* (or *shortest-path*), the minimal number of links we need to pass through to travel between two nodes. It describes the navigability of the network (Barabási and Oltvai, 2004). Another one is *betweenness centrality* which quantifies the number of times a node acts as a bridge along the shortest-path between any two nodes in a network. Modules within the network are often thought to communicate with each other via nodes with high betweenness centrality score. They are the "traffic-influential" nodes of a network. Finally, *closeness centrality* is a network measure which tells how fast it will take to

spread information from the given node to the entire network. These nodes could be described as the "monitors" since it is in an excellent position to watch the flow of information within the network (Estrada, 2011). Identifying as well as targeting these central nodes is beneficial from a therapeutic stand point. Thus, network analysis could offer the vital information about the critical players of any complex system. Once the important elements of a network are known, further downstream experimental investigations can be much more focused and less time consuming. In addition, network analysis could help in creating new study hypotheses.

1.4.2 Some previous network-based research work

Previously, some neurodegenerative disease studies have used network-based analysis to explore and expand the knowledge-base of disease related genes/proteins. One such network-based study was performed using the Parkinson's disease microarray gene expression data. In this study, using protein-protein interactions, Moran and Graeber et al., (2008) built a network of known genes/proteins involved in Lewy body formations. Expanding their network by including "candidate genes", they have been able to reveal more information about the underlying molecular mechanism of Lewy body formation in Parkinson's disease brains.

Another network-based neurodegeneration study was by Goehler et al. in 2004. They constructed a protein-protein interaction network around Huntington's disease gene HTT in order to identify the genes/proteins that directly interact with it. This immediate neighborhood network helped them to identify a new enhancer of HTT aggregation, called GIT1 (a G protein-coupled receptor kinase), and additional tests verified its role in Huntington's disease pathogenesis. Even though these studies were network-based, it only included limited use of techniques. A more of

network-based analysis would be to incorporate different network descriptors such as degrees and centrality scores of the PPI network to identify new important players, and to expand the knowledge-base of the underlying molecular mechanisms. With the advent of many high-throughput technologies, network-based analysis is an invaluable tool in studying biological and biomedical systems including viral-host interactions, and complex diseases like cancer, infections and neurodegeneration, as well as in drug discovery (Ideker and Sharan, 2008).

With encouragement from valuable previous research work (Witten and Bonchev, 2007; Chandrasekaran and Bonchev, 2012; Vladimir Kuznetsov and Bonchev, 2008), we ventured out to utilize a comprehensive network-based approach to study the neurodegeneration process. With the "Brain activity map" project on the horizon, it is the right time to apply such comprehensive holistic methods to study the complex nature of the neurodegenerative diseases as well as to find effective therapeutic measures.

1.5 Study proposal

As a result of vast number of biochemical, animal and human research work, many critical genes and molecular mechanisms of neurodegeneration process were found. Primarily, both genetic and environmental factors were suspected to contribute for the slow, progressive and irreversible dysfunction as well as specific neuronal loss in all neurodegeneration conditions. The latter always exhibit protein misfolding and abnormal aggregate formation in select brain regions. In addition, biological processes including inflammation, oxidative stress, synapse formation, mitochondrial function, microglial activation, reactive oxidative species and free radicals formation etc. were repeatedly found deregulated in such disease states. All these man-

ifestations indicate a possibility for the existence of a common unified underlying molecular mechanism in all neurodegenerative disorders. Network-based analysis provides valuable tools to study such complex systems which involve many players together with their intricate web of interactions.

Working hypothesis

"Applying network-based methods to build various types of interaction networks, one can identify the common underlying molecular mechanisms and the critical players in all neurodegenerative disorders. In doing so, we anticipate finding new neurodegenerative disease genes, pathways and/or drug targets."

We plan to accomplish this study goal by following our specific study proposals:

1. Identify key underlying molecular players of neurodegenerative disorders such as Parkinson's, Alzheimer's and Huntington's at both gene/protein level and at pathway level using microarray gene expression data.

2. Find common key genes/proteins and pathways involved in neurodegenerative disorders.

3. Expand the knowledge base of the molecular mechanisms of neurodegenerative disorders with the help of regulatory gene/protein networks and predict key genes/proteins and pathways involved with considerably high probability in the disease regulatory mechanisms.

4. Elucidate miRNAs regulatory functions in these diseases.

1.6 Thesis organization

Following our study proposal, we have dedicated individual chapters detailing the network-based analysis and results of the three neurogenerative disorders namely

Parkinson's, Alzheimer's and Huntington's disease. The study methodology chapter explains the study data, statistical analysis and the use of different network descriptors to summarize our research findings. The final chapter is devoted exclusively to our analysis and results obtained for the common molecular mechanisms found in all three neurogenerative disorders. The Appendix includes the R program (written in-house), different graphs/plots used for data quality assessment, and Venn-diagrams used in the study.

Chapter 2

Methods and Data

For a research work to be successful, it should include a well-thought out methods section which would explain the study design, the data, tools and different tasks needed to accomplish the proposed research goal. In this section, we explained in detail the various steps we followed to streamline the network analysis of Parkinson's, Alzheimer's and Huntington's disease. Figure 2.1 is an illustration of our research work flow

2.1 Step 1: Microarray gene expression data

In a simple form, the central dogma of molecular biology could be stated like DNA makes RNA, which in turn makes protein. The more active a gene is the more mRNA it will produce. By measuring the mRNA level, we could examine the gene expression pattern of the cell. DNA microarray is a powerful technology that provides a high throughput and detailed view of the entire genome and transcriptome of an organism by measuring the relative mRNA abundance intensity. Due to their ready availability, high volume capacity and parallel testing, microarrays have dramati-

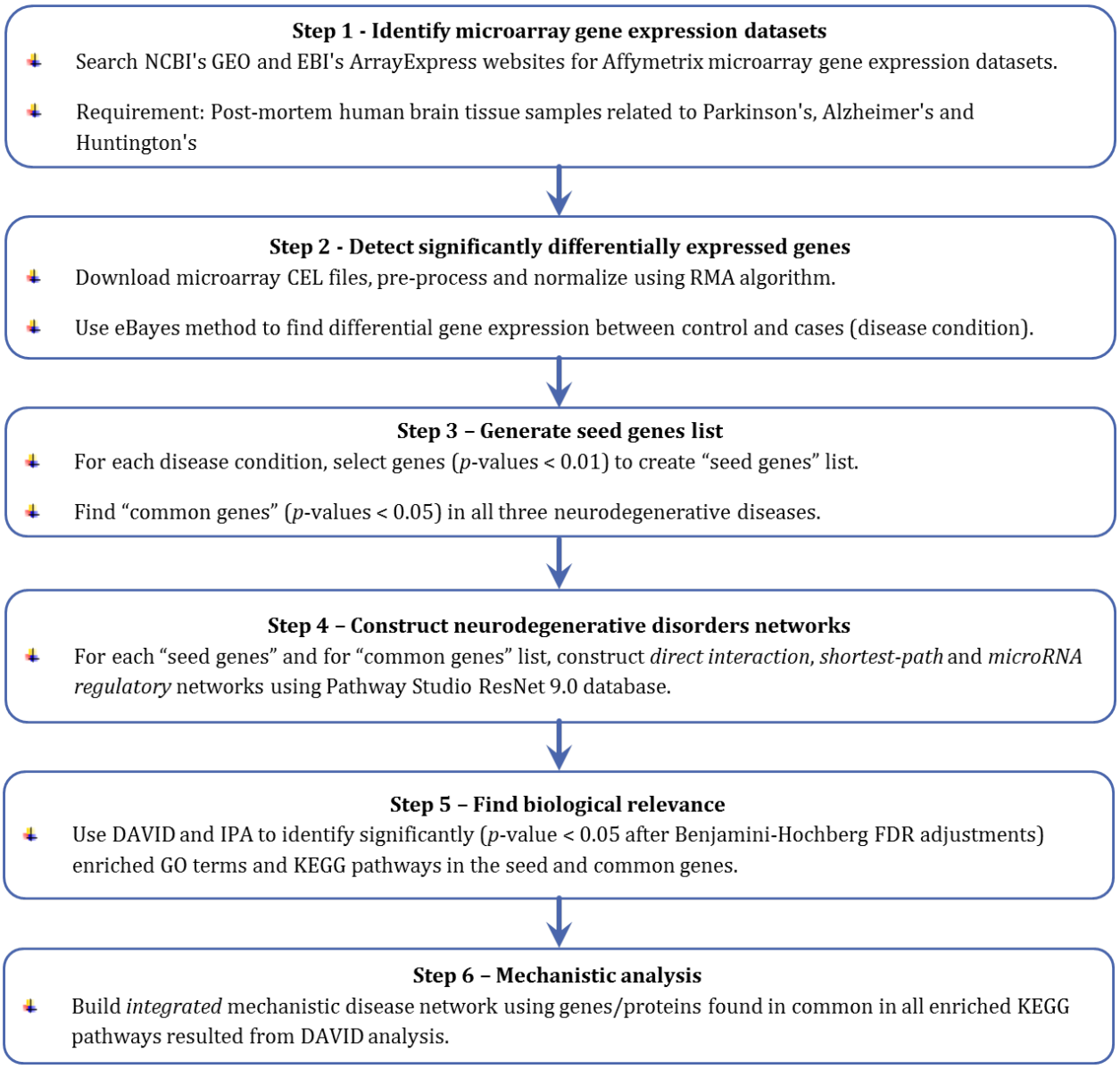


Figure 2.1. Methods and Data workflow. Workflow for finding significantly differentially expressed genes (SDEGs) and pathways in each and in common in the selected neurodegenerative diseases: Parkinson's, Alzheimer's and Huntington's.

cally accelerated many types of molecular biology investigation. They can be used to find genomic expression patterns that enable advancing new pathophysiological hypotheses. Such an approach has already yielded interesting new insights in the study of various cancers, neurological diseases, etc. The known limitation of this approach is that mRNA level does not necessarily correlates with its functional protein level in the cell. Also, post-translational modifications essential for determining protein function are not present on DNA microarray. However, these limitations could be partially overcome by careful handling of arrays, probe selections and repeat experiments. Moreover, microarray assays are inexpensive and less-time consuming when compared with proteomics experiments. Better results in understanding the underlying biological mechanisms are yielded by integrating gene expression and proteomics approaches (Tian et al., 2004; Ideker et al., 2001).

Neurological disorders such as Parkinson's, Alzheimer's and Huntington's are chosen as disease models to study the human neurodegenerative process at molecular level. For this study, DNA microarray gene expression data are used to identify genes involved in neurodegenerative disease conditions. Such microarray datasets are found in public data repositories such as National Center for Biotechnology Information's (NCBI) Gene Expression Omnibus (GEO) available at `www.ncbi.nlm.nih.gov/geo/` and European Bioinformatics Institute's (EBI) ArrayExpress database available at `www.ebi.ac.uk/arrayexpress/`.

Even though there is progress made in clinical diagnosis of neurodegenerative diseases, a definite diagnosis is still possible only by neuropathological examination of the brain. For this study, microarray gene expression of post-mortem brain tissue samples from control and diseased conditions (called cases) are used. All samples are age and sex matched along with various criteria including pH, PMI (post-mortem interval), disease duration etc. (please follow the link listed in Table

2.1 for more details about each dataset). The original authors of the selected microarray datasets had taken measures to include case subject samples' neuropsychological and/or neuropathological data demonstrating the neurodegeneration manifestation clinical diagnosis and symptoms. The control subjects were those with no known neurodegenerative disease history. For cross disease comparison and validation, care has been taken to include datasets that were performed using same DNA microarray technology such as Affymetrix GeneChips with either Human genome U133A/B or U133 Plus 2.0 Arrays. When searching for Parkinson's, Alzheimer's and Huntington's microarray gene expression datasets in GEO and ArrayExpress, Affymetrix datasets were found predominantly. Table 2.1 lists the microarray platforms and gene expression datasets utilized in this study. Unless stated otherwise any subsequent mention of *microarrays* in this text is referring specifically to the Affymetrix GeneChips.

Due to the unavailability, only one post-mortem human brain tissue microarray gene expression dataset was found and used for Huntington's disease study. More detailed information about each microarray dataset like sample's demographics, brain regions, disease stage, and post-mortem delay etc. can be obtained by following the links in Table 2.1 footnotes or by querying the dataset names (like GSE8397) in NCBI's GEO and also in EBI's ArrayExpress websites. Table 2.2 lists brief information about the different microarray gene expression datasets used in this study. For complete information about the entire sample parameters (such as age, sex, pH, brain regions, disease duration etc.) collected for each of the microarray dataset can be obtained from NCBI's GEO and/or EBI's ArrayExpress websites. Please use the website link provided in Table 2.1 to access such information.

Table 2.1. Microarray gene expression datasets and the platforms used.

Disease condition	Microarray dataset	Microarray platform
Parkinson's	GSE8397 [a]	GPL96/97 - Affymetrix HG[b]-U133A/B Array
	GSE20295 [c]	GPL96 - Affymetrix HG-U133A Array
Alzheimer's	GSE4757 [d]	GPL570 - Affymetrix HG-U133 Plus 2.0 Array
	GSE28146 [e]	GPL570 - Affymetrix HG-U133 Plus 2.0 Array
Huntington's	GSE3790 [f]	GPL96/97 - Affymetrix HG-U133A/B Array

[a] http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE8397 or http://www.ebi.ac.uk/arrayexpress/experiments/E-GEOD-8397/?query=gse8397

[b] Human Genome

[c] http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE20295 or http://www.ebi.ac.uk/arrayexpress/experiments/E-GEOD-20295/?query=gse20295

[d] http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE4757 or http://www.ebi.ac.uk/arrayexpress/experiments/E-GEOD-4757/?query=gse4757

[e] http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE28146 or http://www.ebi.ac.uk/arrayexpress/experiments/E-GEOD-28145/?query=gse28145

[f] http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE3790 or http://www.ebi.ac.uk/arrayexpress/experiments/E-GEOD-3790/?query=gse3790

Table 2.2. Brief information about each microarray gene expression datasets used for analysis.

Disease condition	Microarray dataset	Number of samples	Ratio Male to Female	Mean age ($\pm SD$)	Brain tissues (or) regions
Parkinson's	GSE8397	Cases: 15	9:6	80 (5.7)	SFG, MSN and LSN
		Controls: 8	6:2	70.6 (12.5)	
	GSE20295	Cases: 15	9:6	76.7 (6.2)	BA9, PT and SN
		Controls: 15	10:5	71.2 (11.1)	
Alzheimer's	GSE4757	Cases: 19	9:10	84.1 (7.5)	EC
		Controls: 15	9:6	80.1(7.9)	
	GSE28146	Cases: 22 Controls: 8	11:19[a]	86.3 (1.4)	EC
Huntington's	GSE3790	Cases: 39	23:16	58.3 (15.6)	CE, CN and FL
		Controls: 33	23:11	57.2 (17)	

SFG - superior frontal gyrus; MSN - medial Substania nigra; LSN - lateral Substania nigra; BA9 - Brodmann area 9; PT - Putamen; SN - Substania nigra; EC - entorhinal cortex; CE - cerebellum; CN - caudate nucleus; FL - frontal lobe.

[a]No separate information about sample's age and sex details were found.

2.2 Step 2. Detection of significantly differentially expressed genes

The statistical tools and methods used in this study were chosen following their wide use in the community and they are considered as standard procedures for microarray gene expression analysis. Goal of analyzing microarray data is to determine which genes are differentially expressed across different tissues and clinical conditions. In order to accomplish our study goal of identifying common genes and pathways among the different neurodegenerative diseases, we maintained consistency by following same techniques for pre-processing, normalizing and post-normalizing across all the microarray gene expression datasets. Several commercial and free tools like dCHIP, TM4, SAM and software like GeneSpring, Pathway Studio are available for microarray data analysis (Mehta and Rani, 2011).

For this study, we preferred to use Bioconductor package for analysis because it is freely available, widely used and relatively easy to write customized programs. "Bioconductor is an open source and open development software project for the analysis and comprehension of genomic data" (Gentleman et al., 2004) (http://www.bioconductor.org/). Bioconductor is based on the R programming language. All the microarray processing is implemented in R (written in-house) using different packages from Bioconductor software.

2.2.1 Data pre-processing

During the pre-processing step, the appropriate gene expression datasets were identified as listed in Table 2.1, the raw microarray CEL files were downloaded from the GEO/ArrayExpress databases, and the microarray chip quality was assessed.

It was evaluated using a Bioconductor package called *arrayQualityMetrics* (Kauffmann et al., 2009). "*arrayQualityMetrics* provides powerful, auto- mated, objective and comprehensive instruments to assess GeneChip reproducibility, identify apparent outlier arrays and compute measures of signal-to-noise ratio". Each and every microarray dataset that we used in our study were subjected to *arrayQualityMetrics* assessment and no extreme outliers were detected. *arrayQualityMetrics* also offered variety of graphs/plots options like MA plot, density plot, boxplot, heatmaps etc. to readily visualize the quality of the arrays. However, there is no one specific quality control tool that has been widely used. Other such tools *arrayQuality, arrayMvout, ArrayTool* etc. are also available with varying degree of quality control report options. Relevant quality assessment plots/graphs for one such microarray dataset are provided in the Appendix A.

2.2.2 Data normalization

Normalization aims to correct for systematic differences (due to sample preparation, batch processing etc.) between genes or arrays. Unless arrays are appropriately normalized, comparing data from different arrays can lead to misleading results. All microarray expression datasets were normalized using Robust multi-array average (RMA) expression measure (Irizarry et al., 2003), which consists of three steps: background correction, quantile normalization (each performed at the individual probe level), and robust linear model fit using median polish (log-transformed intensities at the probeset level). Although the RMA algorithm has been extended to account for background correction for GC content of the oligonucleotides (Wu et al., 2004) providing improvements in accuracy, particularly for weakly expressed genes, we chose to use the standard RMA approach for the ease of comparison with other

similar research results.

The differential gene expression changes were statistically evaluated by the empirical Bayes (eBayes) method (Smyth, 2004) from the *limma* Bioconductor package. Empirical Bayesian methods combine information across genes which provide more statistical power to identify significant changes. Studies have shown that empirical Bayes t-statistic performed better than other test statistics (such as t-test, ANOVA, SAM, Welch t-statistics etc.) especially when the sample size is small (Murie et al., 2009; Jeffery et al., 2006; Kooperberg et al., 2005; Jeanmougin et al., 2010). To account for potential multiple-testing problems, couple of standard procedures was implemented but with no productive results. Due to the nature of the samples (post-mortem tissues) used in this study; it was not possible to find many significantly changed genes using Bonferroni method, the most conservative procedure used for multiple-testing correction. Hence, the next best alternative to correct for multiple-testing problem, the Benjamini-Hochberg adjustment (Benjamini and Hochberg, 1995) was tried, however not all differentially expressed genesets were significant enough. For this study, we were investigating the key players and molecular mechanisms that are common in all three neurodegenerative disorders. In order to accomplish our study goal, we needed sufficient number of significantly differentially expressed genes (SDEGs) in each disease conditions as well as a reasonable number of overlapping genes and yet differentially expressed between different disorders. Hence, no multiple-corrections were utilized and the probe-sets with t-test p-values < 0.05 were considered to be SDEGs in a given disease condition. In this manner, considerable number of overlapping genes was found and, to partially compensate for not accounting for the multiple correlation, only those SDEGs with the p-value lower than 0.01 were included in network evaluation.

2.3 Step 3. Generation of *"seed genes"* for each neurodegenerative disorders

All seven microarray dataset listed in Table 2.1 were subjected to the same statistical analysis explained above and significantly differentially expressed genes lists called "seed genes" were generated for each dataset. To explain this process, we used Parkinson's disease microarray datasets as an example here. The final set of Parkinson's disease SDEGs/seed genes were derived by overlapping the individual seed genes of the three microarray datasets namely GSE8397 HG-U133A & B and GSE20295 HG-U133A. As listed in Table 2.2, GSE8397 dataset had gene expression data from three types of brain tissue samples namely superior frontal gyrus (SFG), medial and lateral Substantia nigra (M/LSN). For this dataset, four sets of differential gene expression changes were evaluated. They were

1. Differential gene expression changes found between control and cases (PD), irrespective of tissue types. We refer this set as *"Diagnosis"*.

2. Differential gene expression changes found between control and cases (PD) in SFG tissue only. We refer this as *"SFG"*.

3. Differential gene expression changes found between control and cases (PD) in MSN tissue only. We refer this as *"MSN"*.

4. Differential gene expression changes found between control and cases (PD) in LSN tissue only. We refer this as *"LSN"*.

On completion of evaluation of the four sets of gene expression changes (diagnosis, SFG, MSN and LSN) in GSE8397 HG-U133A, an overlap of 414 seed genes were found (see Figure 2.2). In a similar way, an overlap of 225 seed genes was

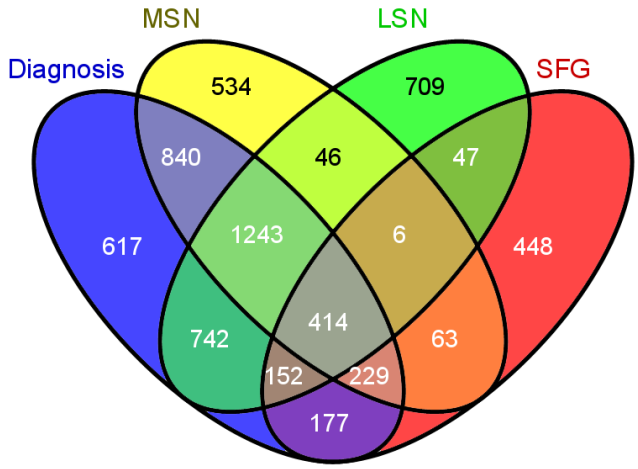


Figure 2.2. Four-set Venn diagram illustration of the overlap of SDEGs in GSE8397 HG-U133A gene expression dataset for Parkinson's disease. Courtesy: VENNY

found in the GSE8397 HG-U133B microarray gene expression dataset (see Figure 2.3). Together, $414 + 225 = 631$ seed genes (after removing duplicates) were found in GSE8397 U133A and U133B datasets.

Correspondingly using GSE20295 HG-U133A microarray dataset, four sets of seed genes namely diagnosis, BA9, PT and SN (tissue samples used) were generated and an overlap of 110 genes were considered SDEGs (p-values < 0.05). Finally, combining the two Parkinson's microarray datasets (GSE8397 and GSE20295) we found 719 (p-values < 0.05) genes to be significantly differentially expressed.

Until now we explained how we generated the "seed genes" for Parkinson's disease using GSE8397 and GSE20295 Affymetrix microarray dataset. Continuing in the similar manner, we evaluated the respective microarray gene expression datasets (see Table 2.1) and found 1205 "seed genes" for Alzheimer's and 925 for Hunting-

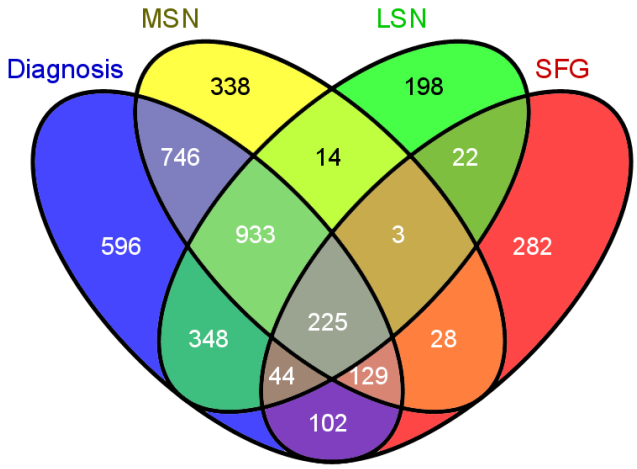


Figure 2.3. Four-set Venn diagram illustration of the overlap of SDEGs in GSE8397 HG-U133B gene expression dataset for Parkinson's disease. Courtesy: VENNY

ton's disease. The overlap numbers and the Venn diagram illustrations for each gene expression datasets are given in detail in Appendix B.

Following the p-values < 0.01 cut-off criteria, there were 267, 214 and 531 "seed genes" for Parkinson's, Alzheimer's and Huntington's diseases respectively. These seed gene list was used for further network analysis. Concluding this part, 22 genes were found in "common" in all three neurodegenerative diseases (see Figure 2.4). The genes were used as "seed" ones in constructing a larger network to explore and to understand the common molecular mechanisms and the network characteristics of the three neurodegenerative diseases.

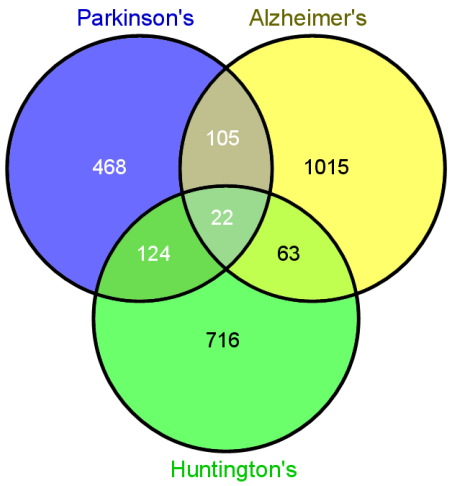


Figure 2.4. Number of common genes in all three neurodegenerative disorders. Courtesy: VENNY

2.4 Step 4. Construction of various types of neurodegenerative disorder networks

We constructed various kinds of networks to understand the interactions between seed genes and investigate the common molecular mechanisms in neurodegeneration. In order to partially compensate for the non-effectiveness of Bonferroni-based selection of SDEGs and to have a more concise and yet informative networks, we restricted our "seed list" (list of significantly differentially expressed genes) to include those genes/proteins with p-values < 0.01.

2.4.1 Pathway Studio software

A popular tool to analyze and construct biological networks and path- ways is Ariadne Genomics Pathway Studio (`www.ariadnegenomics.com`) database software (Nikitin et al., 2003). It helps to interpret gene expression and other high through-

put data. It provides tools to build networks as well as analyze pathways. Pathway Studio software offers many options to construct various kinds of networks such as direct interaction, shortest-path, common targets and regulators of pairs or multiple genes. The protein-protein interaction data are taken from the proprietary database ResNet 9.0 of Ariadne Genomics, provided jointly with the software. ResNet version 9.0 was released October 15, 2011. It covers human, mouse and rat proteins. Protein definitions are from Entrez Gene database as of April 7, 2011. ResNet database is compiled by Ariadne Genomics using MedScan technology from over 20 million NCBI's PubMed abstracts and 880,967 full-text articles as of May 27, 2011. Currently the database covers 125,342 entities such as cell process, complex, disease, functional class, treatment and small molecules including 112,096 proteins. It offers 1,160,524 protein-protein interactions of types binding, chemical reaction, direct regulation, expression, microRNA regulation, molecular synthesis, molecular transport, promoter binding, protein modification and regulations. ResNet 9.0 database also includes information about 5581 custom built Ariadne Genomics cell-process, metabolic and signaling pathways.

For this project, we used direct interaction, shortest-path, and common regulators network options in Pathway Studio software to identify various interactions between the SDEGs and connecting genes/proteins that could be of additional interest in neurodegeneration process. In the following paragraphs we explain the types of networks we constructed and their importance in our analysis.

2.4.2 *Direct interaction* (DI) network

Direct interactions are the simplest form of network representation. Using our list of "seed genes". Direct interaction (DI) network reveals the direct interactions be-

tween those genes. The seed gene lists of each of the chosen neurodegenerative disorders were subjected to direct interaction network construction. Interactions between the genes/proteins are of the following types: regulation, protein modification, promoter binding, direct regulation and microRNA regulations. Due to insufficient information in the databases, as well as due to their nature, not all genes/proteins would directly interact with each other; many genes/proteins remain isolated (unconnected) nodes in the DI network. The information regarding the unconnected nodes is maintained separately in Excel files but these nodes are not included in the later network analysis.

2.4.3 *Common regulator* network

Another network option in Pathway Studio 9.0 is common regulator network. We constructed this network with each of the following interaction types: regulation, protein modification, promoter binding, direct regulation and microRNA regulations. These individual networks revealed a subset of seed genes to be common regulators of pairs or multiples of genes; these are the genes/proteins that control the gene expression pattern of other genes/proteins. Based on their degree of local connectivity (node degree ≥ 5) the common regulators are classified as "Hubs" - the genes that represent the most highly connected nodes in the molecular network. The cut-off value for the hub node degree was selected in order to classify a sufficient number of significantly differentially expressed genes as network hubs. A node with higher connectivity usually means a higher functional role for the corresponding gene/protein in the network. The hub genes may represent key points of vulnerability in underlying signaling pathways responding to neurodegeneration. We therefore categorized these genes as either *"Genes of interest in SDEGs"* or *"Al-*

ready known genes in SDEGs" based on whether they could be of interest or have already been implicated in neurodegenerative disorders. In later chapters, for easy understanding the network nodes in the figures are highlighted in either *blue* or *green* color based on their categories respectively.

2.4.4 Sources used to search for gene information

Online Mendelian Inheritance in Man (OMIM) database (McKusick-Nathans In- stitute of Genetic Medicine, Johns Hopkins University (Baltimore, MD), Oct 5, 2012, World Wide Web URL: http://omim.org) was used as the primary source of information on genes that were already implicated in each of the three neurodegenerative disorders chosen for this study. For example, we compiled a list of Parkinson's disease associated genes found in OMIM by searching for the word "Parkinson's disease" in OMIM Gene Map query tool using "Phenotype only Entries" option. This query in OMIM resulted in a list of previously known PD related genes.

Additionally, Google search engine (available at https://www.google.com/) was used to find information about the SDEGs and their implication in neurodegeneration process, if any. NCBI's PubMed database (available at http://www.ncbi.nlm.nih.gov/pumed) was primarily used as a resource to download the corresponding journal articles. At the end, we used the information from OMIM, Google and NCBI's PubMed to categorize the SDEGs into either already-known or could be of potential interest in a given neurodegenerative disorder.

2.4.5 *Shortest-path* (SP) network

The Shortest-path network (SPNW) enables us to find indirect relationship between two or more entities through intermediary nodes in the absence of direct relation-

ship. By constructing networks that include only the shortest-path interactions among the genes/proteins, the network size is reduced considerably and also true-positive interactions are likely to be enriched (Managbanag et al., 2008).

By applying the *shortest-path* network strategy to the list of SDEGs, we could identify genes/proteins that might contribute to the neurodegenerative process but have not been related so far to it. This approach is based on the inference that when genes/proteins with well-defined biological functions (e.g., oxidative stress) interact with other genes/proteins, the latter have a higher probability to share that function, as compared to those selected at random (Schwikowski et al., 2000; Witten and Bonchev, 2007). By exploiting this known biological network property (guilt-by-association), we could find novel genes/proteins of importance for neurodegenerative process and expand the knowledge base of the molecular mechanisms of these disease conditions.

In Pathway Studio, shortest-path network option with protein modification, promoter binding, and direct regulation interaction types were used to construct a *shortest-path* network for each of the neurodegenerative disorder's seed list. This type network could introduce many connecting genes/proteins without which some of the seed genes would remain unconnected nodes in the network. One limitation of the shortest-path network is that sometimes it could bring in more intermediary nodes in order to have a unified network. For example, beginning with 531 seed genes (with p-values <0.01) from the Huntington's disease microarray dataset, the *shortest-path* network grew into a network with 937 nodes. Such huge networks are not only impractical to do further analysis it will also diminish the importance of the seed genes in a given scenario. Thus, care has been taken to reduce the number of connecting nodes in the shortest-path network.

We accomplished the reduction in number of connecting nodes by setting up

a cut-off rule, to include only those seed genes which had $\geq$ 25 neighbors in the Pathway Studio ResNet 9.0 database. The specific threshold value of 25 reflects the very high average number of neighbors per node in this database. The highly connected nodes we selected in this manner have a better chance to be directly connected to the genes known to be related to the diseases under study. The 531 seed Huntingon's genes were thus reduced to 258 genes. A lower ratio of 1.5 to 2 connecting vs seed genes/proteins was maintained for all datasets. As a rule, unconnected nodes were removed from the final network.

2.4.6 Construction of *compact* shortest-path network

The connecting genes/proteins of the shortest-path networks include both essential and non-essential genes. Essential genes are part of the normal cell functions (e.g. MAPK is involved in a wide variety of cellular processes such as proliferation, differentiation, transcription regulation and development), while non-essential genes could be disease causing ones. In order to determine whether connecting genes/proteins were already implicated or could be of interest in any neurodegenerative process or disorders, we cross-referenced (queried) each such gene/protein in NCBI's PubMed database. OMIM database was also searched for any such associations. The disease-related genes/proteins were categorized as *"Already known genes from the connecting nodes of SPNW"*. Based on their biological/molecular role if the genes are of interest in neurodegenerative disorders then those genes were categorized as *"Genes of interest from the connecting nodes of SPNW"*. In later chapters, for easy distinguishing the network nodes in the figures are highlighted in either *red* or *orange* color for these two categories respectively.

Next, a compact version of shortest-path network (CSPNW) was constructed

using the genes/proteins listed in all four categories discussed above. Few more generic genes were added without which some of the genes of interest would remain unconnected (not all genes of interest directly interact with previously known disease related genes). Some genes could be second- or third-level interacting partners to the known genes via intermediary nodes and yet be potentially important to the disease pathogenesis. The resulted *compact* shortest-path networks were concise and yet meaningful to carry out our network analysis of neurodegenerative disorders.

2.4.7 Network topology descriptors

Another criterion for selecting genes/proteins potentially important for neurodegenerative process is their central location in the network built. Pajek is freely available non-commercial software used for analysis and visualization of large networks such as Internets, collaboration networks, protein-protein interaction networks etc. (Batagelj and Mrvar, 1998). The *compact* shortest-path networks were subjected to quantitative analysis using Pajek. Network measures such as *local connectivity* (i.e., node degree), *closeness* (network monitors) and *betweeness* (traffic-influential) *centrality* scores were calculated (Estrada, 2011). The mathematical formulae for various centrality measures are given below.

Node degree k_i is defined as the number of connections of node i. It is a basic network connectivity measure, which for an undirected network $G = (V, E)$ is described as,

$$k_i = \sum_{j=1}^{V} a_{ij}$$

where V is the total number of *vertices* (nodes), E is the total number of *edges*, j represents all other nodes, and the *adjacency matrix* element a_{ij} is defined as 1 if

node i is connected to node j, and 0 otherwise.

Closeness centrality CC_i is defined as the inverse of the average *vertex distance*(d). It can be regarded as a measure of how fast it will take to spread information from node i to all other nodes in a network sequentially. It can be mathematically expressed as

$$CC_i = \frac{V-1}{d_i}$$

where V is the total number of *vertices* (nodes).

Betweeness centrality BC_i is defined by the fraction of all shortest-paths passing through the vertex. It is a measure of the influence a node has over the spread of information through the network. It can be described as

$$BC_i = \sum_i \sum_j \frac{\rho(i,k,j)}{\rho(i,j)}$$

where $\rho(i,j)$ is the number of shortest-paths from node i to j, and $\rho(i,k,j)$ is the number of the shortest-paths that pass through node k in the network.

2.4.8 *MicroRNA regulatory* network (MRN)

A critical moment in understanding complex biological systems (such as diseases) is to explore underlying molecular mechanisms. Regulatory network is one such way to describe potential pathways to regulate global gene expression programs in a biological system. Differential gene expression is achieved through complex regulatory networks that are controlled in part by two types of regulators: transcription factors (TFs) and microRNAs (miRNAs). TFs and miRNAs are critical modulators of gene expression and signaling pathways and provide potential novel targets for understanding biological behavior and for therapeutic applications. The main function of

transcription factors is transcriptional regulation of target genes, whereas miRNAs are post-transcriptional regulators. Few miRNAs (e.g., miR-133b, miR-7 in Parkinson's) have already been found to be involved in neurodegeneration. This kindled our interest to explore for microRNA regulation in our seed genes.

First step we ventured out to find miRNAs that would target our seed genes. In order to identify the microRNAs by employing seed genes we constructed *shortest-path* network with only miRNA regulation type interactions using Pathway Studio's ResNet 9.0 database. This network revealed the miRNAs that could be of interest in a given disease type. Secondly, in order to construct regulatory network, we used the *direct* interaction network option in Pathway studio utilizing the seed genes and the corresponding miRNAs (identified in the earlier step). We indeed identified many microRNA regulations of our seed genes which will be discussed in detail in the following chapters. The MicroRNA regulatory network (MRN) also revealed a probable integrated regulation by both transcription factors and microRNAs in neurodegeneration process. However, our microRNA regulatory analysis should be offered with some caution, because currently a high percentage of miRNA-targets interactions in Pathway Studio ResNet 9.0 database are based on predictions instead of experimental validation, as verified from the references given in this database.

2.5 Step 5. Finding biological relevance in neurodegenerative disorder networks

Gene Ontology (GO), an expert-curated database, assigns a list of genes into various biologically meaningful categories such as biological processes, molecular functions, and cellular components. p-values are used to rank the significantly modulated

genes into different GO categories. The Database for Annotation, Visualization and Integrated Discovery (DAVID) (Glynn Dennis et al., 2003; Huang et al., 2009a,b) is a widely used web-based application focusing on GO classification. It provides biological functional interpretation of large lists of genes derived from genomic studies such as microarray, proteomics experiments.

Next step is to evaluate whether the listed genes were part of already well- established biological pathways. Tools like Core analysis in Ingenuity's IPA (Ingenuity Systems, www.ingenuity.com) and Pathway Enrichment Analysis in Pathway Studio are widely used to identify enriched canonical pathways in a given list of genes. The list of seed genes were subjected to both tools in the search for known pathways to be affected in neurogenerative disorders.

2.6 Step 6. Mechanistic Analysis

The results from DAVID analysis (obtained in Step 5) were examined in an attempt to characterize the integrated molecular mechanisms involved in neurodegeneration process. In general, the output of DAVID analysis includes those GO categories (such as biological process, cellular component and molecular functions) along with KEGG pathways that are enriched in a given list of genes. Kyoto Encyclopedia of Genes and Genomes (KEGG) "is a database resource for understanding high-level functions and utilities of the biological system, such as the cell, the organism and the ecosystem, from molecular-level information, especially large-scale molecular datasets generated by genome sequencing and other high-throughput experimental technologies" (available at http://www.genome.jp/kegg/) (Kanehisa and Goto, 2000). The KEGG pathways that were significantly enriched (p-value ≤ 0.05 after Benjamini-Hochberg FDR adjustments) and previously known in neurodegenerative

disorders under study were identified and further investigated. Google and NCBI's PubMed databases were used to search for such previously known biological pathways in neurodegenerative disorders. Later, all the genes from the enriched KEGG pathways were combined into a list of genes called "mechanism genes". For each disease condition, such a "mechanism genes" list was prepared.

Based on their molecular functions we further classified these "mechanism genes" as either disease causing (leading to neuronal loss/death) or disease alleviating (helps in neuronal survival) agents. Once again Google and NCBI's PubMed databases were used to identify such previous implications. For easy understanding, the *loss* versus *survival* classification is represented by highlighting the "mechanism genes" in *purple* or *yellow*, respectively. Using the "mechanism genes" *direct* interaction network was constructed as well as investigated for integrated disease mechanism. As will be shown in the chapters to follow in each of the three diseases investigated we could outline three possible mechanisms for initiating the disease from an extracellular signaling.

Chapter 3

Parkinson's Disease Network Analysis

3.1 Introduction

As mentioned in Methods and Data chapter, for our Parkinson's disease (PD) network analysis we used GSE8397 and GSE20295 microarray gene expression datasets from NCBI's GEO database. The GSE8397 microarray dataset was published in 2006 by Moran et al. The data contain 15 cases of neuropathologically confirmed PD and eight controls. This was the first whole genome expression analysis of the Substantia nigra (SN). Some of the important findings of this research work were remarkable quantitative gene expression difference between the Parkinsonian Substantia nigra and control tissue, reported several new candidate genes which map to PARK loci, identified 570 "priority genes" after the Benjamini-Hochberg FDR correction. Two years later, (Moran and Graeber, 2008) published a network-based analysis using the then Pathway Studio's ResNet database version 5.0. No effect for sex difference in Parkinson's disease has been found in the expression levels and clustering of probes. Several *direct* interaction networks has been constructed for the interactions between their priority genes and known-PD genes, as well as for genes related

to Lewy body formations, with nearly 30% of the priority genes remaining unconnected. Cancer, diabetes and inflammation disease conditions have been associated with the top up-regulated priority genes. Drugs like clozapine, cocaine and haloperidol, which are used in the treatment of PD and cause side effects, have been found to interact with a large number of PD priority genes.

The second microarray dataset GSE20295 was published by Zhang et al. in 2005. That research group had highlighted some of the deregulated genes responsible for either disease aggravation (MKNK2) or neuroprotection (HSBP1, SMA5, and FGF13). Deregulation was noticed in various genes belonging to metallothionein group and the heat shock protein group. These patterns of multiple molecular process deregulations have been found across different brain regions studied. Another expression pattern discovered supports the hypothesis for ubiquitin/proteasome system (UPS) dysfunction in Parkinson's disease. A decrease in Complex I activity has also being found to reinforce the suspected mitochondrial deregulation in PD.

To the best of our knowledge, we present the first molecular level study with comprehensive network-based analysis of Parkinson's disease proceeding from the above mentioned microarray datasets. Helped by the rapidly accumulated new biological information, advanced ResNet 9.0 database, and a variety of specific network reconstructions and analyses, we were able to expand upon the previous author's analysis of this disease paradigm, and to explore the underlying cellular mechanisms and molecular players of this disease domain.

3.2 Parkinson's disease *direct* interaction network

We initiate our Parkinson's disease network analysis using the 267 "seed genes", selected as explained in Methods and Data. Out of the 267 significantly differentially

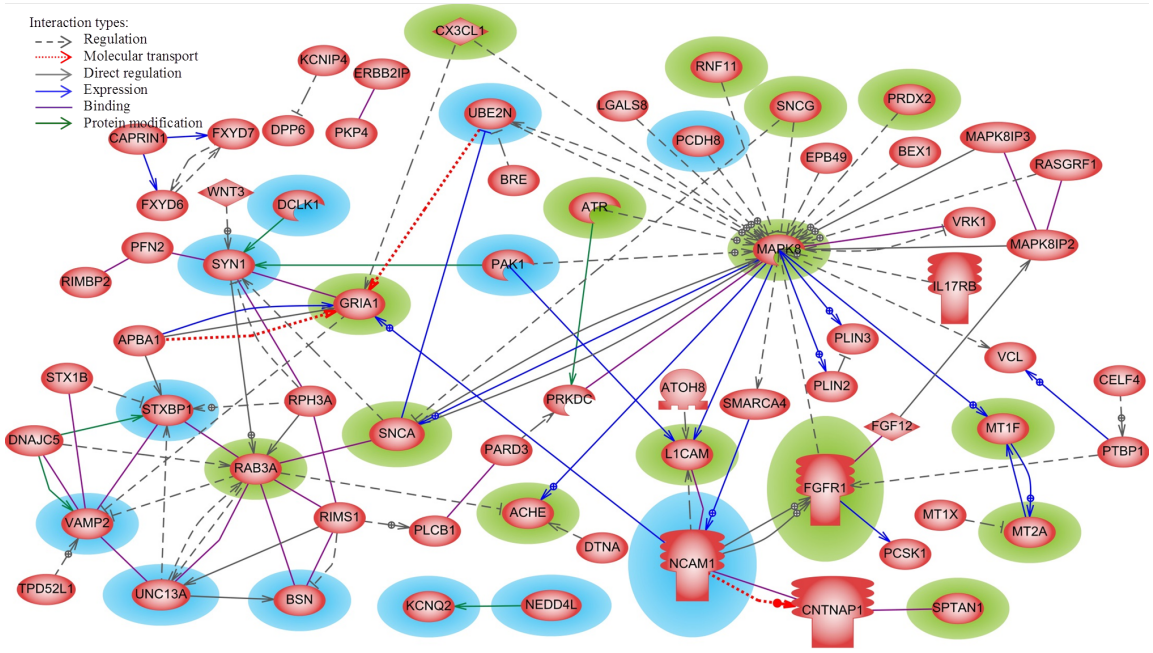


Figure 3.1. Parkinson's disease *direct interaction* network. The 15 genes/proteins implicated previously in PD pathology are highlighted in green and the twelve genes/proteins of potential interest for that disease are highlighted in blue.

expressed genes (SDEGs), 67 genes were directly connected to each other based on the different interaction types like regulations, promoter binding, direct regulation, protein modification and microRNA's regulation. This interaction network (Figure 3.1) has a relatively low average node degree of 2.84. Genes like MAPK8, RAB3A, STXBP1, SYN1 and VAMP2 are the top five hubs (high connectivity nodes) with node degree ≥ 7. One of the well-known Parkinson's gene SNCA (α-synuclein) was among the top five most influential (betweenness centrality) and highest accessible (closeness centrality) nodes in the network.

Based on their physiological/molecular role, 15 out of the 67 genes (ACHE, ATR, CX3CL1, FGFR1, GRIA1, L1CAM, MAPK8, MT1F, MT2A, PRDX2, RAB3A, RNF11, SNCA, SNCG and SPTAN1), have already been implicated in Parkinson's disease paradigm either as neuroprotective and therapeutic agents or as disease aggravat-

ing ones (see Chapter 1.). In Figure 3.1, these previously PD-known genes are highlighted in *green*.

3.2.1 Molecular functions of proposed candidate genes

Useful information about their characteristic physiological roles has also been found in the literature, as a preliminary basis for classifying twelve genes (BSN, DCLK1, KCNQ2, NCAM1, NEDD4L, PAK1, PCDH8, STXBP1, SYN1, UBE2N, UNC13A and VAMP2) colored in *blue* in Figure 3.1 as potentially involved in Parkinson's disease. The molecular functions of some of these candidate genes are summarized here. NCAM1 (neural cell adhesion molecule 1) is important in cognitive processes such as learning and memory. It plays a major role in both body and brain immune surveillance system (Mirnics et al., 2005). NCAM1 facilitates the release, repositioning, and/or expansion of the synaptic complex. BSN (bassoon presynaptic cytomatrix protein), is a scaffolding protein involved in organizing the presynaptic cytoskeleton, the specialized sites where neurotransmitters are released from the synaptic vesicles. (NCBI — Gene website. Retrieved on 25-Feb-2013, from `http://www.ncbi.nlm.nih.gov/gene/8927`; updated on 24-Feb-2013). Campbell et al. (2012) have shown that STXBP1 (syntaxin binding protein 1) has a vital part in the process of calcium ion—dependent exocytosis in neurons, as well as in neuroendocrine cells. It facilitates membrane fusion and neurotransmitter release.

Additionally, SYN1 (synapsin I) was found to be a key player in synapse formation and plasticity (Cesca et al., 2010). During an action potential (an important part of the neuron firing process), synapsins are phosphorylated by PKA (cAMP dependent protein kinase), releasing the synaptic vesicles and allowing them to move to the membrane and release their neurotransmitter. VAMP2 (vesicle-associated mem-

brane protein 2), gene is thought to participate in neurotransmitter release at a step between docking and fusion. A recent study (Diekstra et al., 2012) has shown that single nucleotide polymorphisms in UNC13A (unc-13 homolog A) gene may be associated with sporadic amyotrophic lateral sclerosis (ALS). It regulates neurotransmitter release at synapses, including at neuromuscular junctions. α-synuclein were shown to promote disruption of ubiquitin proteasome system (Betarbet et al., 2005). UBE2N (ubiquitin-conjugating enzyme E2N) targets proteins for degradation via the proteasome. In recent years, synaptic vesicle trafficking defects have been increasingly implicated as an important factor in many PD models, either via direct interactions with the synaptic vesicle (SV) cycling machinery or via indirect effects caused by mitochondrial dysfunction (Esposito et al., 2012). Even though genes BSN, NCAM1, STXBP1, SYN1, VAMP2 and UNC13A are not shown to be directly related to PD, they all seems to play an important role in the regulation as well as the release of neurotransmitters and synaptic vesicles during the SV cycle process.

3.2.2 Network attributes of proposed candidate genes

Additional arguments for considering the above mentioned genes as associated with Parkinson's disease are provided from network perspective. Figure 3.1 reveals the *direct* interacting partners for each candidate gene. For example, BSN, STXBP1, SYN1, VAMP2, and UNC13A directly interact with RAB3A, a gene well-known in PD, where RAB3A is able to provide substantial rescue against α-synuclein-induced degeneration of dopaminergic neurons. Besides with RAB3A, SYN1 is also directly connected to GRIA1 and SNCA, two known PD genes. Studies have suggested glutamate receptor (GRIA1) antagonists as potential treatment agent for Parkinson's

disease (Lai et al., 2003).

In the *direct* interaction network, potential candidate genes like PAK1 and UBE2N are among the top five nodes with high closeness (visibility) centrality score. Another of the proposed candidate genes SYN1, was among the top five hub nodes as well as among the top five nodes with highest betweenness (traffic-influential) centrality score. Being a first-level direct interacting neighbors of a known gene (guilt-by-association), makes also BSN, NCAM1, PAK1, PCDH8, STXBP1, SYN1, UBE2N, UNC13A or VAMP2 genes of potential interest in Parkinson's disease. The physiological role these genes play in synaptic vesicle trafficking, neurotransmitter release, and ubiquitination, as well as their other network attributes like being hubs, network traffic-influential and/or monitoring nodes, increases the chance of these genes to be involved in the PD pathology, which reinforces the arguments in favor of their experimental validation.

3.3 Parkinson's disease *shortest-path* network (SPNW)

Based on their number of neighbors (≥ 25) in the Pathway Studio's ResNet 9.0 database we selected 105 out of the 267 SDEGs for constructing a more *compact* shortest-path network (SPNW), the nodes of which would have a higher chance to be connected to some of the known PD genes. Interaction types included promoter binding, protein modification and direct regulation. 193 genes were software-added to connect the 105 seed genes along the shortest-paths between any pair of those. The connecting genes were examined in sources like OMIM, and NCBI's PubMed databases, along with Google search engines to verify whether they have already been or not implicated in PD. In the second case, whether they could be of potential interest in PD diagnosis was decided based on the gene's physiology/molecular

characteristics and network location (guilt-by-association). In addition to that, we also analyzed the most highly connected *common regulators* of SDEGs (discussed in detail in Chapter 2.) and categorized them as already known and not known but of potential interest in PD diagnosis.

Table 3.1. Summary of the genes of interest and genes already known in Parkinson's disease.

Different categories	Number of genes	Node color in figure
Genes of interest from SDEGs	16	blue
Known PD genes from SDEGs	15	green
Genes of interest in SPNW connecting nodes	35	orange
Known PD genes in SPNW connecting nodes	21	red

A more *compact* version of this 267-genes shortest-path type network was constructed using the four categories of genes from Table 3.1 and few generic genes without which some of the genes of interest would remain unconnected. The *compact* SPNW (see Figure 3.2) is considerably better connected (average node degree 6.79) than the one based on *direct* interactions. Many of the known PD genes, such as AKT1, CASP3, CDK5, MAPK1, MAPT and SNCA are highly connected in this *compacted* network version. CDK5 and MAPK1 are among the ten hub genes (AKT1, CASP3, CDK5, CREB1, CTNNB1, EGFR, MAPK1, SP1, SRC and TP53) with node degree > 15. In biological networks, highly connected nodes tend to be part of critical functions or pathways and some of the founded hubs like TP53, MAPK1, AKT1 and CASP3 being a typical example.

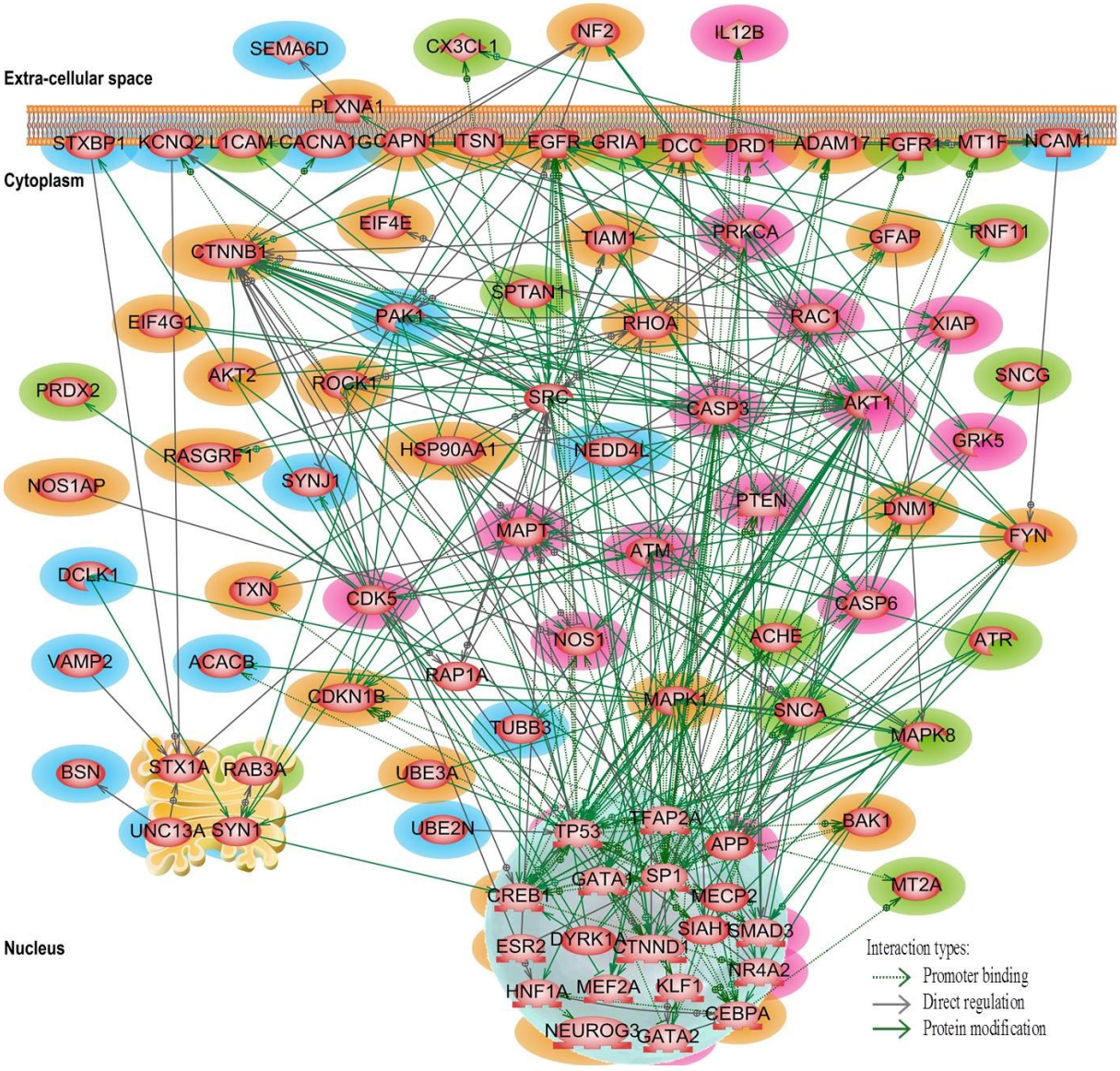


Figure 3.2. Parkinson's disease *compact* shortest path network.The genes/proteins implicated in PD pathology are highlighted in green and red. The genes/proteins of potential interest are highlighted in blue and orange. (see Table 3.1. for gene highlighting details).

3.3.1 DAVID enrichment analysis

The nodes included in the network were then subjected to enrichment analysis using DAVID software tool which systematically maps the given gene list to the associated biological annotation terms (e.g., GO terms or Pathways). Different statistically significantly enriched Gene Ontology categories/subcategories and pathways related to brain and nervous system (with Benjamini-Hochberg multiple correction) are presented in Table 3.2.

DAVID analysis identified several clusters of genes involved in neuron development, differentiation, projection and apoptosis, synaptic transmission, vesicle transport and regulation as biological processes affected by Parkinson's disease. Indeed, many of the enriched genes like CDK5, FGFR1, L1CAM, NR4A2, PRKCA, RAB3A, RAC1 and SNCA have already been studied as mediators, suppressors or regulators of neurodegeneration. Pathways such as ErbB signaling and Neurotrophin signaling are enriched in this PD related gene list. Both these pathways were considered as major players to promote survival of dopaminergic neurons (Sharif and Prevot, 2010). Synaptosomes, axons, and membrane-bounded vesicles are some of the cellular components that are found affected by PD.

3.3.2 IPA analysis

Similar analysis was also carried out using Ingenuity's IPA software to examine pathways that were enriched in our gene list. Many signaling pathways (see Figure 3.3) including 14-3-3 mediated, neuregulin, semaphorin, ephrin, gap-junction, axonal guidance, Huntington's disease, as well as different growth factor signaling like EGF, FGF, and NGF, were found to overlap and to be enriched in Parkinson's disease pathology. This finding extends over the recent report of (Matigian et al., 2010).

Table 3.2. Gene set DAVID enrichment analysis of Parkinson's disease *compact* shortest-path network.

Category	Term	Gene count	Fold Enrichment	Benjamini
GOTERM_BP_FAT	GO:0031175 neuron projection development	16	9.61	1.37E-08
GOTERM_BP_FAT	GO:0030182 neuron differentiation	19	6.67	2.88E-08
GOTERM_BP_FAT	GO:0048666 neuron development	17	7.71	4.11E-08
GOTERM_BP_FAT	GO:0048489 synaptic vesicle transport	8	37.27	1.11E-07
GOTERM_CC_FAT	GO:0043005 neuron projection	16	7.38	2.10E-07
GOTERM_BP_FAT	GO:0007268 synaptic transmission	15	7.74	3.66E-07
GOTERM_BP_FAT	GO:0060627 regulation of vesicle-mediated transport	10	16.01	5.66E-07
GOTERM_BP_FAT	GO:0048812 neuron projection morphogenesis	13	9.38	5.81E-07
GOTERM_BP_FAT	GO:0048488 synaptic vesicle endocytosis	6	65.88	9.59E-07
GOTERM_BP_FAT	GO:0007409 axonogenesis	12	9.56	1.88E-06
GOTERM_CC_FAT	GO:0019717 synaptosome	9	16.71	2.17E-06
GOTERM_CC_FAT	GO:0030424 axon	11	10.92	2.52E-06
GOTERM_CC_FAT	GO:0045202 synapse	14	6.22	7.42E-06
GOTERM_BP_FAT	GO:0050804 regulation of synaptic transmission	10	11.30	7.83E-06
GOTERM_BP_FAT	GO:0007611 learning or memory	9	12.46	1.62E-05
GOTERM_BP_FAT	GO:0051588 regulation of neurotransmitter transport	6	35.48	1.64E-05
GOTERM_BP_FAT	GO:0016192 vesicle-mediated transport	17	4.54	1.83E-05
GOTERM_BP_FAT	GO:0046928 regulation of neurotransmitter secretion	5	36.60	1.80E-04
GOTERM_BP_FAT	GO:0001764 neuron migration	6	14.41	9.20E-04
GOTERM_BP_FAT	GO:0021955 central nervous system neuron axonogenesis	4	40.99	0.002
GOTERM_CC_FAT	GO:0031982 vesicle	14	3.30	0.004
GOTERM_CC_FAT	GO:0030665 clathrin coated vesicle membrane	5	14.89	0.004
GOTERM_BP_FAT	GO:0007411 axon guidance	6	8.62	0.007
GOTERM_CC_FAT	GO:0030136 clathrin-coated vesicle	6	7.17	0.014
GOTERM_BP_FAT	GO:0001963 synaptic transmission, dopaminergic	3	51.24	0.014
GOTERM_BP_FAT	GO:0016358 dendrite development	4	17.57	0.014
GOTERM_BP_FAT	GO:0021952 central nervous system projection neuron axonogenesis	3	38.43	0.023
GOTERM_CC_FAT	GO:0012506 vesicle membrane	6	6.27	0.023
GOTERM_CC_FAT	GO:0030425 dendrite	6	5.81	0.028
GOTERM_CC_FAT	GO:0030426 growth cone	4	11.69	0.035

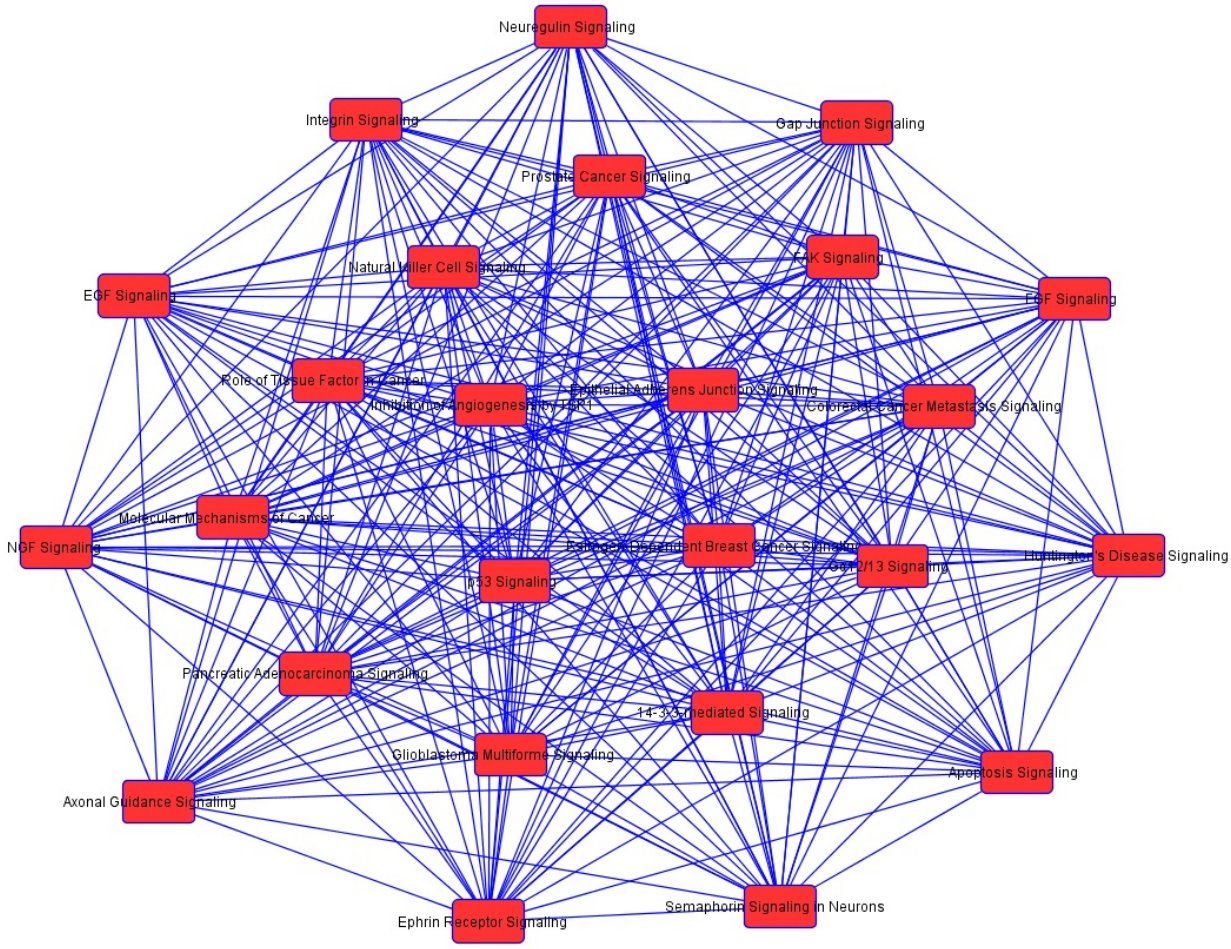


Figure 3.3. Enriched IPA canonical pathways in Parkinson's disease.

Neuregulins along with epidermal growth factors play a diverse role in neuronal development and as well as differentiation. Systemic administration of neuregulin-1β1 protects dopaminergic neurons in a mouse model of Parkinson's disease. (Carlsson et al., 2011). Semaphorins and ephrins are prominent families of axon guidance cues during normal nerve growth and also after injury. Previous studies have shown binding interactions between 14-3-3 proteins, synuclein-alpha and LRRK2 (leucine-rich repeat protein kinase 2), the genes linked to sporadic and familial form of PD (Steinacker et al., 2011). It was symptomatic to find out major neurodegenerative conditions like Alzheimer's disease, Amyotrophic lateral sclerosis (ALS) and Huntington's disease signaling, to be enriched in Parkinson's disease conditions as well. Discovering these overlapping pathways will help to better understand the complex neurodegenerative diseases mechanism and to search for therapeutic agents common for the entire family of these diseases.

Table 3.3. Genes of interest for Parkinson's disease identified by guilt-by-association with the known PD-related genes.

Genes of Interest	Interacts with no. of known PD genes
CTNNB1	10
EGFR	6
PAK1	5
CEBPA, CTNND1, ADAM17	4
CDKN1B, KLF1, ROCK1, SYN1	3
AKT2, BAK1, DNM1, DYRK1A, NF2, TUBB3	2
EIF4E, ITSN1, MECP2, NCAM1, NEDD4L, NOS1AP, RASGRF1, RHOA, STX1A, STXBP1, SYNJ1, UBE2N	1

3.3.3 Guilt-by-association analysis

Table 3.3 shows the genes of potential interest as determined by being neighbors of the known PD genes (guilt-by-association). CTNNB1 (catenin, beta 1) has the record environment of ten (!) nearest neighbors in the *compacted* shortest-path network (CSPNW, Figure 3.2) all of which known to be involved in Parkinson's disease (AKT1, CASP3, CASP6, CDK5, CREB1, MAPK8, NR4A2, PTEN, RAC1 and SMAD3). This makes CTNNB1 number one candidate gene of interest. This gene, along with Wnt1 and Fzd-1 critically contributes to the survival and protection of adult midbrain DA neurons (L'episcopo et al., 2011). In addition, it has a high betweenness centrality which increases its global influence in the network.

The next strongest candidate for implication with Parkinson's disease is EGFR (epidermal growth factor receptor) gene having six PD-related neighbors (CASP3, CDK5, PRKCA, RNF11 and TP53). It is one of the top ten nodes with highest node degree, closeness as well as betweenness centrality scores. This greatly contributes EGFR to be one of the critical positions in the *compact* shortest-path network with greater visibility and traffic-control. In general, many studies have shown that EFGR signaling play a major role in neurogenesis, neuron survival and maintenance (Chen et al., 2007; Dumstrei et al., 1998; Lillien and Raphael, 2000; Ayuso-Sacido et al., 2010). In a recent study, EGFR has been suggested as a preferred target treating amyloid-beta induced memory loss in Alzheimer's disease (Wang et al., 2012).

Third interesting PD candidate is PAK1 (p21 protein (Cdc42/Rac)-activated kinase 1) gene having five PD-related neighbors (AKT1, CASP3, CDK5, RAC1, and TP53). PAK1 regulates neuronal polarity, morphology, migration and synaptic function (Nikolić, 2008). There are also strong indications that PAK1 is required for normal cognitive function. The gallery of Parkinson's disease potentially related genes

includes also CEBPA (CCAAT/enhancer binding protein (C/EBP), alpha), which interacts with four known PD genes (GATA2, IL12B, MT2A, and TP53). CEBPA has been shown to bind to the promoter and modulate the expression of leptin, a hormone having easy accessibility to the brain. It is important to note that leptin receptors are expressed in neurons and other brain regions and are known to regulate neural development and neuroendocrine functions. Leptin could be a potential drug candidate for neurodegeneration (Tang, 2008).

3.3.4 Molecular functions and network attributes of some of the critical connecting genes

The *compact* shortest-path network included many noteworthy connecting proteins like APP, CREB1, HSP90AA1, MAPT and PTEN which were previously implicated to play critical roles in many neurodegeneration disease pathogenesis and couple of them were indicated to have neuroprotective mechanism. APP (syncluein, alpha), is the major component of the filamentous inclusions found in the Lewy bodies and Lewy neuritis, the characteristic hallmark features of many neurodegenerative diseases including Parkinson's, Alzheimer's, dementia with Lewy bodies and multiple system atrophy (MSA). Neurodegenerative diseases caused by these abnormal aggregations of alpha-syncluein proteins are specially classified as alpha-synucleinopathies (Galvin et al., 2001; Spillantini and Goedert, 2000; Stefanis, 2012; O'Brien and Wong, 2011). Similarly, tauopathies are a class of neurodegenerative diseases that are associated with the aberrant accumulations of tau proteins (MAPT) in the brain. Hyper phosphorylated tau proteins are the main component of neurofibrillary tangles (NFTs), another typical pathological feature of neurodegeneration. Tau proteins deformation are found in both genetic and sporadic forms

of Parkinson's and Alzheimer's diseases in addition to other neurodegenerative diseases such as progressive supranuclear palsy (PSP), Down's syndrome, Pick's disease (Spillantini and Goedert, 2001; Pittman et al., 2006; Lei et al., 2010; Churcher, 2006). Many molecular evidence suggests potential interaction between alpha-synclein and tau proteins (Galpern and Lang, 2006). PTEN gene mutations also contribute to the NFT formations and the deregulation of tau phosphorylation (Sonoda et al., 2010; Zhang et al., 2006; Kerr et al., 2006). Detailed biochemical and genetic studies about APP, MAPT and PTEN molecular processing will be crucial to the development of therapeutic targets to treat many neurodegenerative diseases.

CREB1 and HSPs were suggested for such therapeutic measures in neurodegenerative disorders. In a mice model study (Mantamadiotis et al., 2002), it is indicated that postnatal disruption of CREB1 along with CREM showed progressive neurodegeneration in the hippocampus and in the dorsolateral striatum. This evidences that both CREB1 and CREM can promote nerve cell survival globally in developing brain while more selectively in adult brain. Earlier studies have demonstrated that increase in the expression of HSPs especially HSP70 by gene transfer or HSPs inducers can reduce the aberrant protein misfolding and inhibit the pro-apoptotic pathway to attenuate dopaminergic neuron degeneration (Meriin and Sherman, 2005).

Besides being major contributors of neurodegeneration process, APP, CREB1, HSP90AA1, MAPT and PTEN have varying degree of interactions with many known Parkinson's disease genes (see CSPNW, Figure 3.2). Among these five genes, CREB1 appears to have major network advantage as being one of the top ten nodes with highest local connectivity, visibility and traffic-influential node in the *compact* shortest-path network. In addition, genes like APP, MAPT and HSP90AA1 are among the top 25 nodes with highest connectivity and higher accessibility to all other nodes as measured from their node degree and closeness centrality score. Other genes from

Table 3.3 might also be investigated for possible relations to Parkinson's disease, which are also interacting with many known PD genes.

3.4 *Integrated* Parkinson's disease mechanism

Apart from finding important biological players of Parkinson's disease process, we continued with our major study goal to trace down the underlying common molecular mechanisms of this disease. In our quest we proceeded from the significantly enriched KEGG pathways (p-value < 0.05 after Benjamini-Hochberg FDR adjustments) found in DAVID analysis. As described earlier, the genes used to construct the *compact* shortest-path network were subjected to DAVID gene set enrichment analysis which includes KEGG pathway classifications. Kyoto Encyclopedia of Genes and Genomes (KEGG) "is a database resource for understanding high-level functions and utilities of the biological system, such as the cell, the organism and the ecosystem, from molecular-level information, especially large-scale molecular datasets generated by genome sequencing and other high-throughput experimental technologies" (available at `http://www.genome.jp/kegg/`) (Kanehisa and Goto, 2000).

3.4.1 KEGG analysis

DAVID analysis revealed many KEGG pathways that were significantly affected in our PD gene list. After reviewing previous Parkinson's disease research literature, we selected 16 out of 34 such KEGG pathways (see Table 3.4 for details) to identify common molecular mechanisms of PD. The selected pathways belong to KEGG's categories of signal transduction, cell motility, cell communication, immune system, nervous system and neurodegenerative diseases.

Table 3.4. Enriched KEGG pathways in Parkinson's disease resulted from DAVID analysis.

KEGG Pathways	Gene count	FE [a]	Benjamini	Genes
Focal adhesion	17	6.62	1.49E-07	PRKCA, EGFR, ROCK1, XIAP, PTEN, SRC, CTNNB1, AKT1, MAPK1, FYN, RASGRF1, RAC1, RHOA, RAP1A, MAPK8, PAK1, AKT2
Adherens junction	10	10.16	6.08E-06	EGFR, FGFR1, MAPK1, FYN, RAC1, RHOA, SMAD3, CTNND1, SRC, CTNNB1
MAPK signaling pathway	15	4.39	4.57E-05	PRKCA, EGFR, FGFR1, TP53, AKT1, MAPK1, CASP3, RASGRF1, MAPT, RAC1, CACNA1G, RAP1A, MAPK8, PAK1, AKT2
Axon guidance	11	6.67	5.10E-05	DCC, MAPK1, ROCK1, PLXNA1, SEMA6D, FYN, RAC1, RHOA, L1CAM, PAK1, CDK5
ErbB signaling pathway	9	8.09	1.12E-04	PRKCA, EGFR, AKT1, MAPK1, CDKN1B, MAPK8, PAK1, SRC, AKT2
Chemokine signaling pathway	11	4.60	6.61E-04	AKT1, MAPK1, ROCK1, TIAM1, RAC1, RHOA, RAP1A, CX3CL1, GRK5, PAK1, AKT2
Wnt signaling pathway	9	4.66	0.002	PRKCA, ROCK1, RAC1, RHOA, TP53, SMAD3, SIAH1, MAPK8, CTNNB1
Neurotrophin signaling pathway	8	5.05	0.004	AKT1, MAPK1, RAC1, RHOA, TP53, RAP1A, MAPK8, AKT2
Alzheimer's disease	9	4.32	0.004	MAPK1, APP, CASP3, NOS1, MAPT, SNCA, ADAM17, CDK5, CAPN1
Tight junction	8	4.67	0.005	PRKCA, AKT1, RHOA, PTEN, SRC, SPTAN1, CTNNB1, AKT2
p53 signaling pathway	6	6.90	0.005	CASP3, TP53, SIAH1, ATR, PTEN, ATM
VEGF signaling pathway	6	6.26	0.008	PRKCA, AKT1, MAPK1, RAC1, SRC, AKT2
Amyotrophic lateral sclerosis (ALS)	5	7.38	0.012	CASP3, NOS1, GRIA1, RAC1, TP53
Gap junction	6	5.27	0.014	PRKCA, EGFR, MAPK1, DRD1, SRC, TUBB3
Toll-like receptor signaling pathway	6	4.65	0.022	AKT1, MAPK1, RAC1, MAPK8, IL12B, AKT2
Regulation of actin cytoskeleton	8	2.91	0.044	EGFR, FGFR1, MAPK1, ROCK1, TIAM1, RAC1, RHOA, PAK1

[a]Fold Enrichment

Literature review showed less interest to pathways involved in Parkinson's disease but instead revealed many genes/proteins of the individual pathways to be involved in PD mechanism. Next few lines will explain the molecular functioning of couple of the genes/proteins of different pathways (listed in Table 3.4) and how they are related to Parkinson's disease manifestation. FYN (oncogene related to SRC, FGR, YES) protein-tyrosine kinase oncogene belongs to focal adhesion pathway. Previous study (Nakamura et al., 2001) has shown that alpha-synuclein was phosphorylated by activation of FYN-mediated signaling which might influence the mobility of synaptic vesicles, neurotransmitter release and axonal transport. VEGF (vascular endothelial growth factor) molecule, part of the same named signaling network, is known to promote microglial proliferation, neurogenesis and angiogenesis. VEGF was shown to provide neuroprotective effects via both direct and indirect mechanisms with other players of VEGF signaling pathway (Yasuhara et al., 2004).

Figure 3.4 is our modest attempt to show the essential part of the integrated Parkinson's disease mechanism using the 46 genes/proteins found in common in all the 16 KEGG pathways listed in Table 3.4 Following earlier convention, in Figure 3.4 we have highlighted in *green* those genes that were already implicated in PD and in *blue* the genes that could be of potential interest to Parkinson's disease. Based on their molecular functions we further classified the 46 genes as either disease causing (leading to neuronal loss/death) or disease alleviating (helps in neuronal survival) agents. This loss versus survival classification is expressed in Figure 3.4 via highlighting the 18 loss causing genes in *purple* and those 26 genes that helps in neuron survival in *yellow*. Due to the high interconnectedness of our biomolecular network we were unable to see any clear separation between the loss or survival genes, they are part of a single integrated system. There was a considerable overlap between the pathways these genes are part of. Visual inspection of the 16 pathways

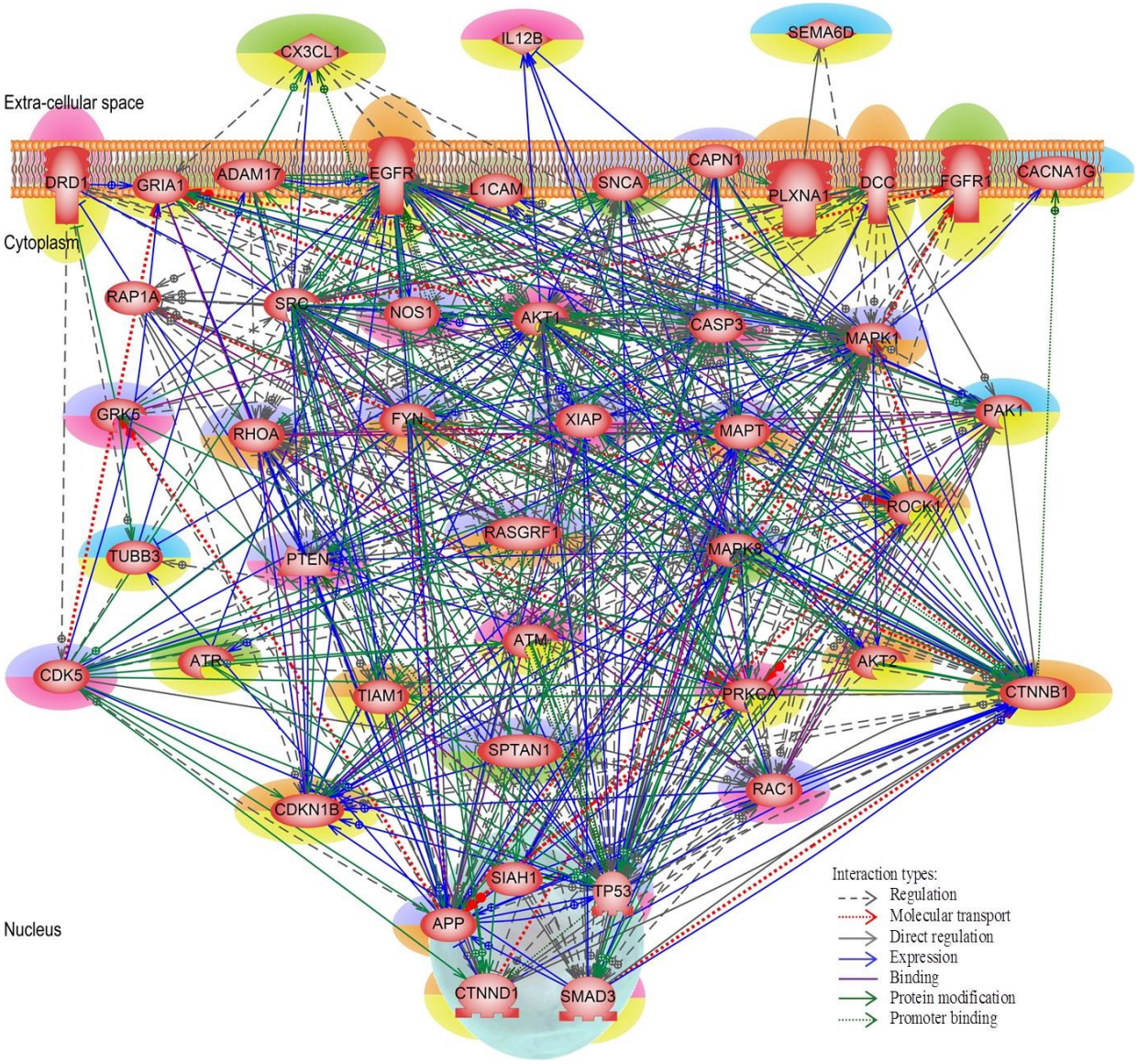


Figure 3.4. ***Integrated*** **Parkinson's disease mechanism.** The 46 genes/proteins are found in common in all 16 enriched KEGG pathways. Genes/proteins implicated in PD pathology are highlighted in green/red and the genes/proteins of potential interest are highlighted in blue/orange (See Table 3.1.) for details.

in KEGG database also revealed that there is no definite compartmentalization of different processes within a biological cell. One process/pathway feed into another or multiple pathways, e.g., WnT signaling pathway includes players from MAPK, focal adhesion, adherens junction, and Alzheimer's disease pathways.

3.4.2 Proposed Parkinson's disease initiation routes

On examining the integrated mechanism network we suspect that the Parkinson's disease mechanism could be initiated via one of the three routes triggered by the extra-cellular ligands namely, CX3CL1, IL12B and SEMA6D. First route is through fractalkine (CX3CL1) together with DRD1 (dopamine receptor D1) modulates GRIA1 gene expression. There is also interaction between CX3CL1, ADAM17 (metallopeptidase domain 17), and LCAM1 which then follow a downstream path into cytoplasm and finally to the nucleus for further regulation of gene expression. CX3CL1 lies upstream to and its gene expression is positively activated by ADAM17 and TP53 genes/proteins. As mentioned earlier, fractalkine suppresses microglial activation which contributes to neuronal survival. Studies have shown that ADAM17 mediated fractalkine cleavage would ultimately limit activation of microglia and support neuronal survival (Zujovic et al., 2000). There is two-way opposing gene expression modulation between CX3CL1 and SRC. Inside the cytoplasm AKTI, CASP3, MAPK1, MAPK8 genes/proteins are the direct (first-level) downstream targets of CX3CL1. Except GRIA1 and CASP3, all other downstream target genes of CX3CL1 are positively activated by CX3CL1. Many animal PD model studies have shown that ADAM17, CX3CL1, DRD1, GRIA1, and LCAM1 genes could be therapeutic targets for Parkinson's disease (Benarroch, 2012; Cui et al., 2010; Johnson et al., 2009; Pabon et al., 2011).

Second route is initiated via SEMA6D and its receptor PLXNA1 (plexin A1) which in turn regulates RHOA and AKT1 gene expression inside the cytoplasm. SEMA6D, the upstream target of PLXNA1 was negatively modulated. Inside cytoplasm MAPK1 gene expression was negatively modulated by SEMA6D thereby controlling the downstream activity of MAPK1. Apart from SEMA6D modulation, CAPN1 (calpain 1, (mu/I) large subunit) negatively regulates the expression of both PLXNA1 and SNCA in so doing CAPN1 can modulate PLXNA1 and SNCA's downstream actions inside the cytoplasm. Semaphorins, secreted proteins involved in the guidance of neuronal and nonneuronal cells, interact with receptor complexes formed by plexins and neuropilins. Earlier studies have shown semaphorins along with their receptors to promote or guide neuronal axon projection as therapeutic approaches for treatment of Parkinson's disease (Majed et al., 2006; Tamariz et al., 2010). Studies in rodent and cell culture models of PD suggest that treatment with calpain inhibitors can prevent neuronal death and restore functions thus suggesting that calpain inhibition could be a therapeutic strategy in PD (Samantaray et al., 2008).

Third route of the proposed integrated Parkisnon's disease mechanism takes place via another extra-cellular ligand IL12B (interleukin 12B) which lies upstream to MAP kinases, RAC1 and AKT1 that are present inside the cytoplasm, and all these genes negatively regulate the gene expression of IL12B. Many studies have suggested that neuroinflammation and activated microglia contribute to neurodegenerative processes. Studies have indicated that interleukins help in differentiation and survival of neuronal cells that were stressed out by activated microglial actions (Müller et al., 1998; Rentzos et al., 2009).

Thus, from the integrated disease mechanism network we propose three routes to enhance the survival of the dopaminergic neurons, which could be a source for potential therapeutic targets in Parkinson's disease.

3.5 Parkinson's disease *microRNA* regulatory network

3.5.1 Construction of PD microRNA network

A *shortest-path* network (SPNW) was constructed using all the 267 seed genes/proteins with only microRNA (miRNA) interactions from the ResNet 9.0 database in Pathway Studio software. This allowed identifying 71 regulatory miRNAs, which were used to create a smaller size regulatory network having only *direct* microRNA-target interactions (see Figure 3.5).

Table 3.5. Genes of interest determined from Parkinson's disease microRNA regulatory network

Genes of Interest	Targeted by no. of miRNAs
RIMS3, SEMA6D, SYNJ1	7
PCDH8	6
AQP11, VAMP2	5
DCLK1, PAK1	4
BSN, NCAMP1, STXBP1, UBE2N	3
CACNA1G, GLS, NEDD4L	2
ACACB, KCNQ2	1

Table 3.5 shows the genes of interest in the MicroRNA Regulatory Network (MRN) and how many miRNAs are targeting each gene's mRNA. In the regulatory network (Figure 3.5), miR-218-1 is found to be the top player regulating the expression of 16 genes of which three (PCDH8, RIMS3 and STXBP1) are potential genes of interest to Parkinson's disease. In animal model study, it was shown that miR-218-1 is expressed in hippocampus (Bak et al., 2008) and also through a volumetric MRI imaging study it was implicated that there is a progressive hippocampal volume loss in PD human subjects (Camicioli et al., 2003).

Other microRNAs like miR-29a, miR-132, miR-133a1, miR-182, and miR-330

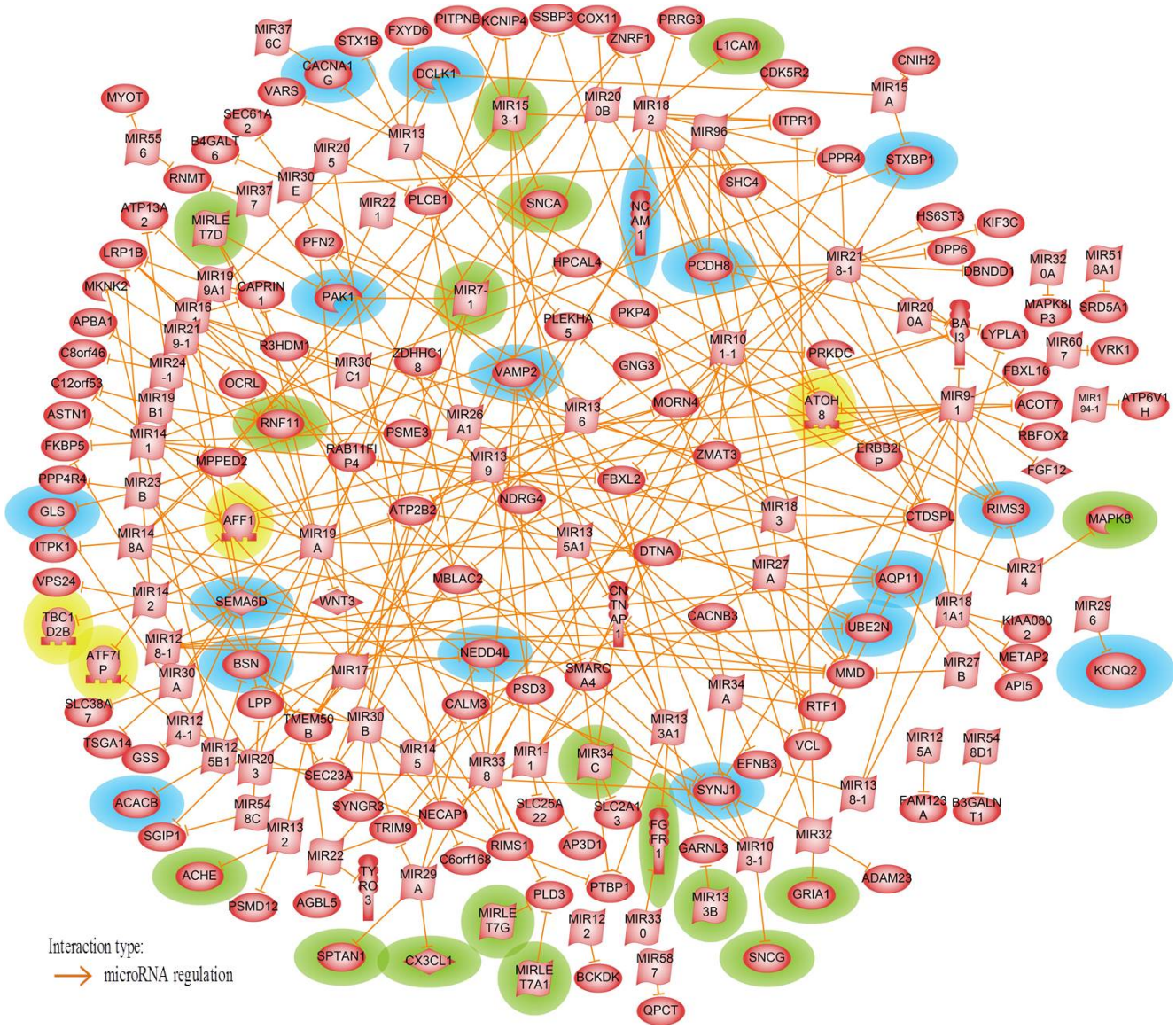


Figure 3.5. Parkinson's disease *miRNA regulatory* network.The genes/proteins and microRNAs implicated in PD pathology are highlighted in green and the genes/proteins of potential interest are highlighted in blue. Genes/proteins that code for transcription factors (TFs) are highlighted in yellow.

were found to regulate the expression of the known Parkinson's related genes ACHE, CX3CL1, FGFR1, L1CAM, and SPTAN1. Being direct interacting partners with known PD-related genes these miRNAs could make them potential regulatory targets in understanding Parkinson's disease mechanism.

3.5.2 Some of the well-known miRNAs in PD

The microRNA regulatory network also incorporates some of the already implicated miRNA's (miR-133b, miR-153, miR-34c, miR-7 and miR-let-7) mediated translation regulations that are relevant to PD pathophysiology. Studies (Harraz et al., 2011; Miñones Moyano et al., 2011; Mouradian, 2012) have shown that the gene expression of DJ1, PARKIN, PITX3, and SNCA were targeted by these miRNAs. These genes were shown to regulate dopaminergic neuron differentiation and activity, oxidative stress mediated cell death, and mitochondrial energy production dynamics, all of which are widely accepted as part of the PD molecular mechanisms.

3.5.3 Highly targeted genes of interest in PD

On further examination, microRNA regulatory network revealed that the expression of candidate genes like RIMS3, SEMA6D and SYNJ1 was tightly regulated by multiple miRNAs. RIMS3 (regulating synaptic membrane exocytosis 3) and other RIM family members are generally believed to be RAB3 isoforms (RAB3A/B/C/D)-specific effectors that regulate synaptic vesicle exocytosis in neurons and in some endocrine cells (Simunovic et al., 2009). Release and re-uptake of neurotransmitters in the synaptic junction is a highly coordinated process and RIMS3, and RAB3A along with other proteins play an important role during neurotransmitter release.

Extra-cellular ligand SEMA6D was proposed as one of three initiators involved

in integrated Parkinson's disease mechanism network (see Figure 3.4). In general, semaphorins act as axonal growth cone guidance molecules and the miRNA regulatory network revealed that SEMA6D gene expression was regulated by seven miRNAs (miR-124-1, miR-128-1, miR-16-1, miR-19a, miR-23b, miR-30a and miR-9). Previous studies have shown various miRNAs in the development and propagation of Alzheimer's disease neuropathology. Especially miR-124-1, miR-128-1 and miR-9 were found altered in abundance in fetal, adult and Alzheimer hippocampus where miR-9 and miR-124-1 are brain-specific and miR-128 is brain-enriched microRNAs (Lukiw, 2007). It is of interest to mention that many of the microRNAs that were listed above to regulate SEMA6D in the miRNA regulatory network have been previously implicated in Alzheimer's, which show protein accumulation neuropathology similar to Parkinson's disease. This is one more sign for the existence of common regulatory mechanisms in neurodegenerative diseases.

Another highly microRNA-regulated gene is SYNJ1 (synaptojanin 1) (see Table 3.5), a polyphosphoinositide phosphatase found enriched in the brain and located at nerve terminals, as well as associated with synaptic vesicles and coated endocytic intermediates. Synaptojanins were suggested to accelerate the synaptic vesicle recovery/trafficking process at the synapse (Harris et al., 2000). Dysfunction of synaptic transmission and membrane trafficking are implicated in PD. Based on its molecular function, SYNJ1 could be of potential interest in Parkinson's disease molecular mechanism.

In addition to miRNA mediated regulation, the network also included four genes AFF1, ATF7IP, ATOH8 and TBC1D2B that encode for transcription factors (TFs). These significantly differentially expressed TFs indicate a possible integrated regulation of the transcription of Parkinson's related genes. The microRNA analysis in this section should be taken with some caution, because currently a considerable

percentage of miRNA-target interactions in ResNet 9.0 database are based on predictions but not on experimental validation.

3.6 Summary

The microarray expression data used in our study were a combination of data produced and interpreted by different authors (Moran et al., 2006; Zhang et al., 2005) and referring to different regions of brain. Aiming at a search for a common molecular mechanism for neurodegenerative diseases, we renormalized the data for a better comparability. Then, a number of specific biomolecular networks were built and analyzed in a variety of ways. As a result, while confirming some of the previous finding, including part of the novel predicted Parkinson's genes, more such PD-related genes were proposed in this chapter based on guilt by association analysis and accounting for the importance of certain nodes in network topology.

As well known, the guilt-by-association approach is based on analysis of the nearest network neighborhood of genes with proved function in the search of interest. Many Parkinson's disease genes were found in the OMIM database, such as LRRK2, PARK2, PARK7, PLA2G6, PINK1, SNCA, and UCHL1. However, our list of significantly differentially expressed genes did not include any of these genes except SNCA. We found that the log fold-change of PARK2, PARK7 and PLA2G6 was only around 0.03, which was not significant enough to detect changes in gene expression. Even though PINK1 and UCHL1 were significantly differentially expressed in some brain tissue types (both in medial Substantia nigra, while UCHL1 also in lateral Substantia nigra) they did not qualify for our "seed genes" list used to construct the PD networks since they were not found significantly differentially expressed in *all* Parkinson's disease datasets. Affymetrix HG-U133A GeneChip did not have probe

for LRRK2 gene but instead included LRRK1 gene probe. Again, LRRK1 did not meet the criteria for "seed genes" list since it did not show strong differential gene expression and its log fold-change was also only around 0.03. While the lack of statistically significant presence of the above mentioned PD-related genes could be attributed to the loss of expression intensity in the post-mortem brain samples compared to a functioning brain, in this study we focused our attention mainly to the genes showing considerable change in the Parkinson's disease samples.

Despite the reduced base of known PD-genes needed for the guilt-by-association predictions we were able to identify from our *direct* interaction network BSN, NCAM1, PAK1, PCHD8, STXBP1, SYN1, UBE2N, UNC13A and VAMP2 as novel Parkinson's disease candidate genes. Second-level interacting partners generally have much lesser chance to be included in the list of candidate genes. However, this chance increases when the gene is known to show certain functions that may be related to the disease of interest. Such is the case with DCLK1 gene via its role in synaptic plasticity and neurodevelopment. Another group of novel PD gene candidates was found from similar analysis of the *shortest path* network. Such is the case with NEDD4L, SYNJ1, TUBB3 as direct partners and ACACB, CACNA1G, KCNQ2, and SEMA6D as second-level partners to already known PD genes. All 17 genes listed here are significantly differentially expressed in PD.

Our network analysis indicated that apart from the strongly differentially expressed genes some *connecting* genes/proteins from the *shortest path* networks could be of similar importance in the deregulation of the disease mechanisms. Considering such *connecting* genes/proteins via their guilt by associations to already known PD genes we concluded that CTNNB1, EGFR, ADAM17, CEBPA, CTNND1, CDKN1B, KLF1, ROCK1 and TIAM1 could also be genes of potential interest in Parkinson's disease realm. Some of the genes of this list were found to play an important role

in network topology. Thus, CTNNB1 and EGFR are among the top ten highly connected nodes (with degree > 15), among the top ten nodes with higher accessibility to all other nodes as assessed by the closeness centrality, and among the top ten traffic influential nodes in the network as judged by their betweenness centrality. Genes like ADAM17, CEBPA and CTNND1 are among the top 25 high connectivity nodes (with degree ≥ 8) and also among the top 25 traffic-influential nodes in the network. Besides helping in identifying novel PD-related genes, the same line of network analysis has shown that APP, MAPT and PTEN, well-known contributors of many other neurodegenerative diseases including Alzheimer's, MSA, Pick's, PSPs etc., are important *connecting* genes/proteins in the Parkinson's *shortest path* network. Finding such genes with common role in neurodegeneration process reinforces our study goal.

In addition to the numerous miRNAs already known to affect the expression of PD-relevant genes (Harraz et al., 2011; Miñones Moyano et al., 2011; Mouradian, 2012) we added another seven. With caution because some of their regulatory interactions are not yet validated, we predict that miR-132, miR-133a1, miR-181-1, miR-182, miR-218-1, miR-29a, and miR-330 could be potential regulators in Parkinson's disease mechanisms, due to their *direct* interaction with known PD related genes. Further investigation of the above mentioned miRNA-related regulatory interactions of candidate and known PD-genes would deepen our understanding of the molecular mechanisms of complex diseases like Parkinson's. Examining the microRNA regulatory network, one may conclude that disease pathogenesis is complex enough to be explained not only from protein-coding genes, but also via regulatory mechanisms mediated by the small non-coding microRNAs.

All genes listed in this summary were shown through gene set enrichment analysis to be key players in various cellular mechanisms like neuron development and

differentiation, synaptic transmission, vesicle transport and endocytosis, apoptosis, and memory/learning, which are altered in the underlying Parkinson's pathophysiology and potential compensatory responses. Moreover, enrichment of Alzheimer's, ALS and Huntington's disease signaling pathway was found to be also affected in PD brains. This supports our hypothesis for the presence of an underlying common mechanism for all neurodegenerative diseases. More evidence in favor of our hypothesis along this line is presented in the next chapters.

In the final stage of our systems biology approach to Parkinson's disease we used the 16 KEGG pathways found enriched by DAVID analysis to build an integrated mechanistic Parkinson's disease network containing 46 genes. We propose on this basis three routes of PD molecular mechanisms proceeding from signaling initiated via the extra-cellular ligands CX3CL1, SEMA6D and IL12B. Further analysis of these routes could reveal novel therapeutic targets for Parkinson's disease. However, the above findings could be considered as the tip of the iceberg in understanding the intertwined nature of these complex neurodegenerative diseases.

Chapter 4

Alzheimer's Disease Network Analysis

4.1 Introduction

Brain tissue samples usually consist of diverse cell types. We need to isolate individual cell types from a heterogeneous tissue sample in order to understand its specific role in a disease mechanism. One such technique is laser capture microdissection (LCM). Neurofibrillary tangles (NFTs) constitute one of the cardinal histopathological features of Alzheimer's disease. Earlier research work by Dunckley et al. in 2006 (GEO dataset GSE4757) claims that LCM can be utilized in a wide range of neurodegenerative diseases to isolate and characterize specific pathologically affected cell types. The method provides valuable new information about the gene dysregulation that occurs during disease pathogenesis. Applying LCM, Dunkley et al. proposed three AD candidate genes APOJ, TIMP3 and IRAK1 which were up-regulated and co-localized to NFTs. All three genes were shown to promote NFT formation and neuronal cell death in AD.

In our study we used the above mentioned gene expression dataset along with GSE28146 published by Blalock et al. in 2011. The latter study provided strong

reciprocal validation for two technically dissimilar microarray analyses (using LCM and heterogeneous tissue sample) of Alzheimer's disease. The authors report that LCM had uniquely identified previously undetected alterations in AD, including up-regulated vasculature development, as well as down-regulation of genes important for stabilization of ryanodine receptor-related intracellular Ca^{2+} release. Through DAVID analysis of the GSE28146 dataset we also observed vasculature development pathway to be affected in Alzheimer's disease.

For Alzheimer's disease (AD) network analysis we used 214 "seed genes" from GSE4757 and GSE28146 microarray gene expression datasets (see Methods and Data chapter for details). To the extent of our knowledge, this is the first study to implement network-based analysis of Alzheimer's disease using the above mentioned microarray datasets. We constructed various types of network to identify critical molecular players and mechanisms involved in AD.

4.2 Alzheimer's disease *direct* interaction network

We found 214 significantly differentially expressed genes (SDEGs), of which 48 genes were directly connected to each other based on the different interaction types like regulations, promoter binding, direct regulation, protein modification and microRNA's regulation. This interaction network (Figure 4.1) has an average node degree of 2.20 with BCL6, DCN, JAK2, PECAM1 and SMAD3 as top five hub genes. Some of these hub genes (BCL6, DCN and SMAD3) along with C1QA and PSEN1 (a well-known gene to cause familial early-onset of AD) have high betweenness centrality score which position them among the top influential nodes in the network. 14 out of 48 SDEGs (CDK5R1, DCN, GRIN2A, HSPB2, ICAM2, JAK2, LMO4, NEFL, NEFM, PECAM1, PSEN1, SMAD3, SYN1 and TGFBI) are already associated

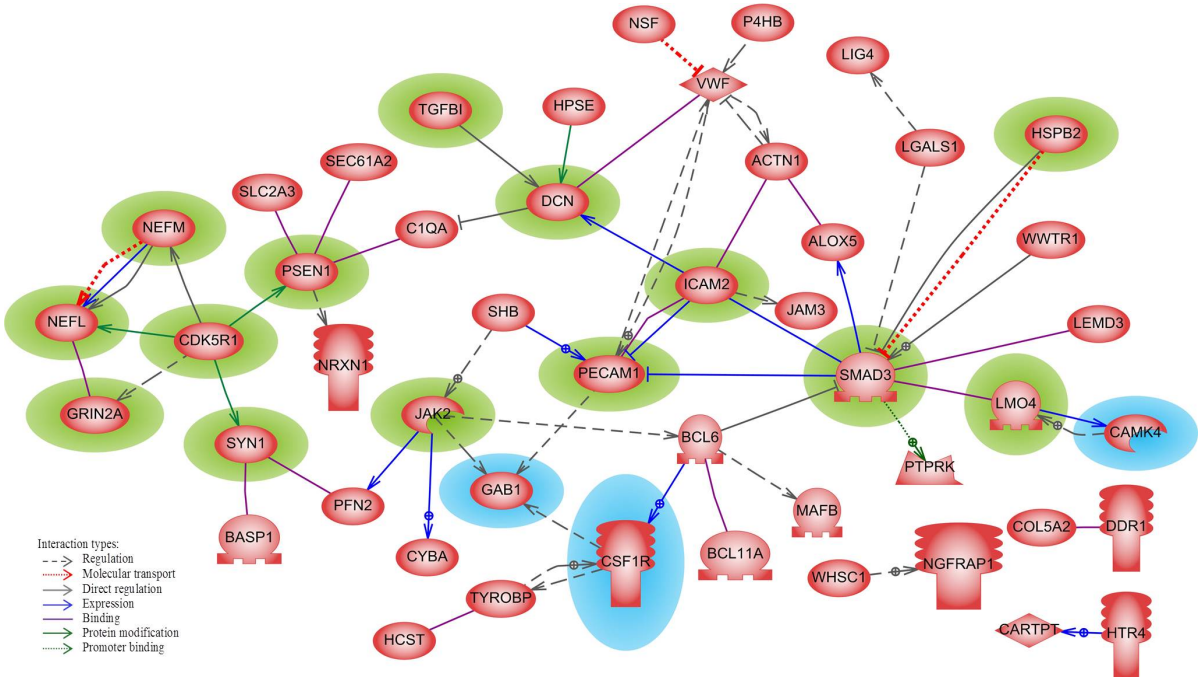


Figure 4.1. Alzheimer's disease *direct interaction* network. The 14 genes/proteins implicated in AD pathology are highlighted in green and the three genes/proteins of potential interest for that disease are highlighted in blue.

in Alzheimer's disease pathology (see Chapter 1). On the basis of their biological functions an additional three genes (CAMK4, CSF1R and GAB1) from our SDEG set could be considered as potentially related to AD mechanism. See Figure 4.1 for details about known and potential novel genes/proteins related to AD.

4.2.1 Molecular functions of some well-known and candidate genes

Figure 4.1 illustrates the well-connected *direct* interactions among many genes that contribute to the neurodegeneration process in Alzheimer's disease. The *direct* association of GRIN2A, NEFM, NEFL and SYN1 with CDK5R1, an activator of CDK5 that leads to the accumulation of aberrantly phosphorylated protein tau, contributes to

Alzheimer's disease. CDK5R1 is in turn has a first level network interaction with a well-known early onset familial AD mutant gene PSEN1. Regional alteration and reduced expression of GRIN2A, NEFL, NEFM and SYN1 were thought to influence to the severity of Alzheimer's disease (Tsang et al., 2008; Qin et al., 2004; Pollak et al., 2003). In addition, CAMK4 and GAB1 could be considered as candidate genes, due to their *direct* associations with JAK2, PECAM1 and LMO4, already known AD genes. Besides, the native biomolecular functions of these genes are noteworthy. They could be potential neuroprotective agents in Alzheimer's disease domain. CAMK4 phosphorylation of CREB/CRE and the subsequent CRE-mediated gene expression in the neuronal nucleus plays an important role in the consolidation/retention of hippocampus-dependent long-term memory formation. Disruption of this balanced pathway was suspected to cause impairments in both long-term potentiation and memory formation (Kang et al., 2001). Moreover, CAMK4 being part of Wnt signaling pathway, could have a pertinent role in Wnt-mediated neuroprotection against amyloid-beta accumulation in AD (Inestrosa and Toledo, 2008).

There is growing evidence that oxidative stress is involved in the pathogenesis of most of the neurodegenerative disorders, including AD (Christen, 2000; Emerit et al., 2004; Jellinger, 2010; Gil and Rego, 2008; Jellinger, 2009; Klein and Schlossmacher, 2006). GAB1 has been identified as an important component in oxidative stress signaling and influencing the cell survival (Holgado-Madruga and Wong, 2003). Although CSFIR (colony stimulating factor 1 receptor) is a second-level interacting partner to AD known gene, based on its intrinsic neuroprotective function, it could also be a potential gene of interest in Alzheimer's disease condition. In a recent study, CSF1R depleted mice show increased neuronal loss after injury and vice versa. Its ligands CSF1 and IL34 exhibit neuroprotection against excitotoxicity induced by neurodegeneration (Luo et al., 2013). Up-regulation of CSF1R could

play potential role in neuronal survival after injury and degeneration.

With their innate physiological roles as well as being *direct* interacting partners (guilt-by-association) to already known AD genes increases the chances of the three genes as novel candidate genes for Alzheimer's disease.

4.3 Alzheimer's disease *shortest-path* network

Using Ariadne's Pathway Studio software a shortest-path network (SPNW) was built with 96 Alzheimer's disease SDEGs. Interaction types included in this network were promoter binding, protein modification and direct regulation. These 96 out of 214 SDEGs were selected on the basis of their number of neighbors (≥ 25) in the Pathway Studio's ResNet 9.0 database. 131 genes were software-added to connect the 96 seed genes along the shortest-paths between any pair of those. Following our categorization technique (discussed in detail in Chapter 2.), the 96 SDEGs along with the 131 software-added connecting genes were classified into the genes that were already related to AD and the genes that could be of potential interest in AD diagnosis and medical treatment. Table 4.1 summarizes the different categories and the corresponding numbers of genes in each.

Table 4.1. Summary of the genes of interest and genes already known in Alzheimer's disease.

Different categories	Number of genes	Node color in figure
Genes of interest from SDEGs	5	blue
Known AD genes from SDEGs	17	green
Genes of interest in SPNW connecting nodes	24	orange
Known AD genes in SPNW connecting nodes	26	red

Utilizing all genes in Table 4.1, along with few additional genes without which

some of the genes of interest would remain unconnected, a more *compact* version of the 214-genes shortest-path network was constructed. Compared with *direct* interaction network this *compact* SPNW was better connected with an average node degree of 8.81 and genes like CREB1, MAPK1, JUN, SP1 and TP53 as highly connected nodes (hubs). In this network, we found 24 such genes with node degree > 10. Many of these hub genes were also the nodes with high closeness (better accessibility) and betweenness (traffic-influential) centrality scores in this *compact* SPNW. Figure 4.2 illustrates the interactions between the known and the genes of interest in the AD *compact* SPNW.

Table 4.2. Genes of interest for Alzheimer's disease identified by guilt-by-association with the known AD-related genes.

Genes of Interest	Interacts with no. of known-AD genes
SRC	**13**
MDM2	**11**
NRF1	5
CAMK4, ESRRA, PTK2B, SRF	4
CR2, NOX4	2
CHRM3, CSF1R, HEY2, IL3	1

4.3.1 Guilt-by-association analysis

Here we show that some genes listed in Table 4.2 could be of potential interest in Alzheimer's disease due to their immediate proximity (guilt-by-association) to some of the already known AD-genes.

Even though the tyrosine kinase SRC (v-src sarcoma (Schmidt-Ruppin A-2) viral oncogene homolog (avian)) is a very generic gene/protein, which controls a wide variety of processes, pathways, and actions, and is responsible for key events

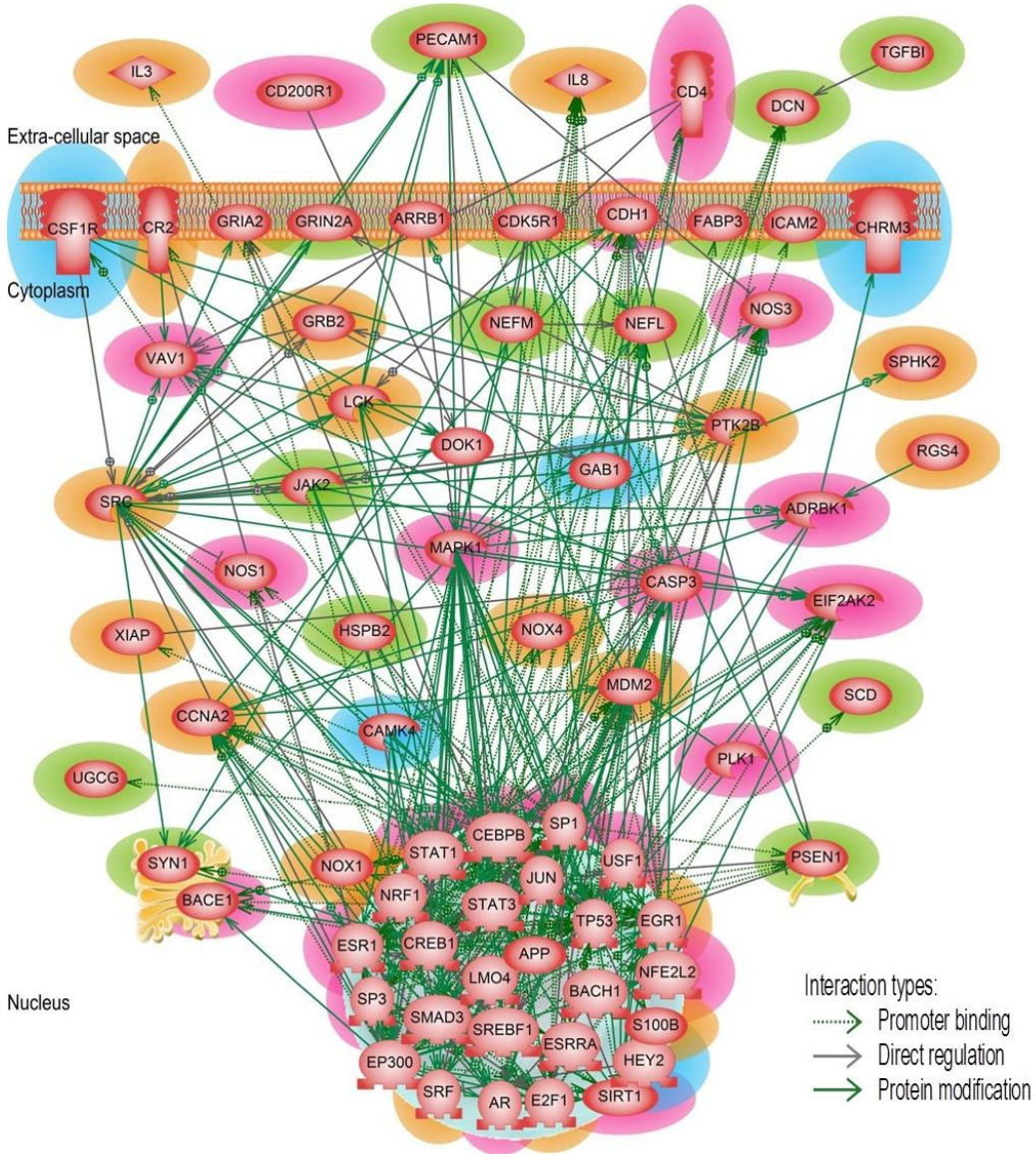


Figure 4.2. Alzheimer's disease *compact* shortest path network.The genes/proteins implicated in AD pathology are highlighted in green and red. The genes/proteins of potential interest are highlighted in blue and orange. (see Table 4.1. for gene highlighting details).

in the body, we found it to be the first-level neighbor to 13 (!) already known Alzheimer's disease genes (ADRBK1, AR, CASP3, CDH1, JAK2, JUN, NFE2L2, NOS1, PECAM1, SP1, STAT1, STAT3 and VAV1). This association with many known AD genes makes SRC a strong candidate for a role in Alzheimer's pathology. Moreover, it has been demonstrated that accumulation of amyloid-beta stimulates microglial (macrophages of brain) activation via SRC tyrosine kinase pathway. Murine AD model study has shown that a cancer drug named Dasatinib could inhibit SRC tyrosine kinase activity of microgliosis in the AD diseased brains. Decreased microglial activity was correlated with improved cognitive performance in mouse (Dhawan and Combs, 2012). Cell culture study also suggests that SRC tyrosine kinase inhibition shows promising route in mitigating the effects of amyloid-beta accumulation in brains which could be tested in human Alzheimer's disease brains (Zambrano et al., 2004).

Based on its high proximity to AD related genes, our next candidate gene of interest is MDM2; it is directly connected to eleven already known Alzheimer's disease genes. MDM2 (p53 E3 ubiquitin protein ligase homolog (mouse)) has E3 ubiquitin ligase activity, which targets tumor suppressor protein p53 for proteasomal degradation. It has been shown that amyloid-β signaling pathway disrupts biochemical pathways involving lipid metabolism, ultimately affecting tau phosphorylation which contributes to AD (Di Paolo and Kim, 2011). Perturbation of lipid metabolism, especially of cholesterol homeostasis, has been illustrated in a another severe progressive neurodegeneration called Niemann-Pick Type C (NPC) disease (Qin et al., 2010). NPC has often been used as a model system to study AD pathology, since both share several features, including neurofibrillary tangles, autophagic/lysosomal dysfunction, inflammation, and cholesterol metabolism abnormalities. In a mouse NPC model, it has been demonstrated that cholesterol perturbation-induced axonal

growth cone collapse and decrease in phosphorylated p53 were reduced by inhibition of MAPK and MDM2 E3 ligase. Due to its inhibitory role in neurodegeneration and also neighboring with many previously implicated AD genes marks MDM2 as a potential gene of considerable interest.

Another proposed candidate gene is NRF1; it neighbors with five (CREB1, ESR1, NFE2L2, NOS1 and TP53) genes that are related to Alzheimer's disease mechanism. One of the distinctive characteristic features of several neurodegenerative diseases is the aberrant accumulation of protein (like α-synuclein, amyloid-β) aggregations. This protein buildup leads to mitochondrial dysfunction which eventually increases the ROS production in the brain (Jomova et al., 2010; Emerit et al., 2004). Both NRF1 and NRF2 are transcription factors that play a pivotal role in the protection against the toxic effects of reactive oxygen species (ROS). Studies have suggested that NRF1 and NRF2 promote mitochondrial biogenesis as well as play crucial role in neuronal survival after acute brain injury. (Dhar et al., 2008; Hertel et al., 2002).

4.3.2 Molecular functions and network attributes of some of the critical connecting genes

The *compact* shortest-path network included many significant *connecting* proteins like APP, BACE1, CCNA2, CREB1, E2F1 and EGR1 which were previously implicated to play critical roles in many neurodegeneration disease pathogenesis including Alzheimer's. Abnormal proliferation of APP is the major contributor of the neurodegeneration process in AD. Pathogenic alterations of BACE1, CCNA2, CREB1, E2F1 and EGR1 are evident in the disease progression. BACE1 initiates the amyloid-beta formation (Cole and Vassar, 2007). In general, cyclins (CCNA2) were shown to increase the severity of Alzheimer-related pathology and other types of dementia

(Smith et al., 1999). In addition, cyclins are regulators of CDK kinases. It has been indicated that APP and tau, the major two proteins implicated in AD, are the physiological substrates for cyclin-dependent kinase especially by a neuron-specific cyclin kinase called CDK5 (Iijima et al., 2000). Similarly, E2F1 were found co-localized with amyloid-beta plaques. It may ultimately contribute to neuronal cell death in AD (Jordan-Sciutto et al., 2002). CREB1 along with CREM were suggested to play major therapeutic role in neurodegeneration (Mantamadiotis et al., 2002).

Apart from having pivotal in neurodegeneration mechanism, the above mentioned connecting genes interact with many previously known-AD genes. Except BACE1, all other genes are among the top 30 nodes with highest connectivity (degree > 8) in the *compact* shortest-path network. CREB1 turns out to be one of the critical nodes in the network. It has many advantages being one of the top five highly connected nodes, having better visibility as measured the closeness centrality as well as by its capacity to control the flow of traffic within the network as assessed by the betweenness centrality. Thus, due to their favorable network centrality and neighborhood connectivity, as well as by their physiological roles, the genes listed in Table 4.2 are our suggestions for candidate genes in Alzheimer's disease, which would require experimental validation.

4.3.3 DAVID enrichment analysis

Apart from *individual gene* assessment, we also carried out gene set enrichment analysis using DAVID software tool to identify different biological processes and/or pathways affected in Alzheimer's disease. Table 4.3 lists some of the Gene Ontology categories/subcategories related to nervous system and functions that were statistically significantly enriched in AD (with Benjamini-Hochberg multiple correction).

Table 4.3. Gene set DAVID enrichment analysis of Alzheimer's disease *compact* shortest-path network.

Category	Term	Gene count	Fold Enrichment	Benjamini
GOTERM_CC_FAT	GO:0043005 neuron projection	15	8.37	1.58E-07
GOTERM_BP_FAT	GO:0007568 aging	8	13.48	8.02E-05
GOTERM_BP_FAT	GO:0043523 regulation of neuron apoptosis	7	14.41	2.76E-04
GOTERM_CC_FAT	GO:0030425 dendrite	8	9.36	3.87E-04
GOTERM_BP_FAT	GO:0030182 neuron differentiation	12	5.08	5.43E-04
GOTERM_BP_FAT	GO:0007169 transmembrane receptor protein tyrosine kinase signaling pathway	9	7.45	6.43E-04
GOTERM_BP_FAT	GO:0006979 response to oxidative stress	8	9.04	6.66E-04
GOTERM_BP_FAT	GO:0042325 regulation of phosphorylation	12	4.77	8.15E-04
GOTERM_BP_FAT	GO:0001944 vasculature development	9	6.64	0.001
GOTERM_CC_FAT	GO:0045202 synapse	10	5.37	0.001
GOTERM_BP_FAT	GO:0048666 neuron development	10	5.47	0.002
GOTERM_CC_FAT	GO:0031594 neuromuscular junction	4	34.69	0.002
GOTERM_BP_FAT	GO:0007611 learning or memory	6	10.02	0.005
GOTERM_BP_FAT	GO:0031175 neuron projection development	8	5.79	0.006
GOTERM_BP_FAT	GO:0008637 apoptotic mitochondrial changes	4	23.91	0.008
GOTERM_BP_FAT	GO:0000302 response to reactive oxygen species	5	12.35	0.009
GOTERM_BP_FAT	GO:0007005 mitochondrion organization	6	8.06	0.010
GOTERM_BP_FAT	GO:0035235 ionotropic glutamate receptor signaling pathway	3	61.77	0.011
GOTERM_BP_FAT	GO:0007268 synaptic transmission	8	4.97	0.011
GOTERM_BP_FAT	GO:0007259 JAK-STAT cascade	4	19.01	0.013
GOTERM_BP_FAT	GO:0019226 transmission of nerve impulse	8	4.24	0.024
GOTERM_BP_FAT	GO:0007409 axonogenesis	6	5.76	0.032
GOTERM_BP_FAT	GO:0008088 axon cargo transport	3	30.89	0.035
GOTERM_BP_FAT	GO:0050877 neurological system process	15	2.30	0.036
GOTERM_BP_FAT	GO:0051402 neuron apoptosis	3	29.26	0.038
GOTERM_BP_FAT	GO:0048667 cell morphogenesis involved in neuron differentiation	6	5.32	0.042
GOTERM_BP_FAT	GO:0048812 neuron projection morphogenesis	6	5.22	0.045

DAVID analysis revealed (see Table 4.3) some of the biological processes like oxidative stress, neuron apoptosis, tyrosine kinase signaling pathway, regulation of phosphorylation and aging affected in Alzheimer's disease, which were also discussed above as part of AD mechanisms. It is valuable to note that the ALS pathway, characterizing another debilitating neurodegenerative condition called Amyotrophic lateral sclerosis, was also enriched among Alzheimer's disease genes. Enrichment of ALS disease pathway along with Huntington's disease pathway was reported by Ingenuity's IPA analysis. One more promising step towards our study goal of identifying and reporting common genes and pathways in major neurodegenerative disorders is achieved.

4.4 *Integrated* Alzheimer's disease mechanism

The results of DAVID analysis were further evaluated in an attempt to elucidate details of the underlying molecular mechanism of Alzheimer's disease. DAVID analysis revealed twelve KEGG pathways that were significantly (p-value <0.05 with Benjamini-Hochberg multiple correction) affected in AD. The pathways were found to take part in signal transduction, immune system, cell communications, nervous system and in general neurodegenerative diseases categories. The pathways classification was helped by previous Alzheimer's disease research literature reviews (Agadjanyan et al., 2005; Arikkath and Reichardt, 2008; Auffret et al., 2010; Chao et al., 2006; Chaudhury et al., 2003; Chen et al., 2000; Das and Golde, 2006; Mattson and Chan, 2003; Puglielli, 2008; Solerte et al., 1998; Town et al., 2008; Xia and Hyman, 1999). The genes from the enriched KEGG pathways (listed in Table 4.4) were then consolidated into a list of 37 common genes.

Table 4.4. Enriched KEGG pathways in Alzheimer's disease resulted from DAVID analysis.

KEGG Pathways	Gene count	FE [a]	Benjamini	Genes
hsa05014:Amyotrophic lateral sclerosis (ALS)	7	12.21	6.67E-04	CASP3, NOS1, GRIA2, GRIN2A, TP53, NEFL, NEFM
hsa04062:Chemokine signaling pathway	10	4.94	0.002	MAPK1, IL8, GRB2, ARRB1, PTK2B, JAK2, ADRBK1, STAT1, VAV1, STAT3
hsa04722:Neurotrophin signaling pathway	7	5.22	0.014	MAPK1, PSEN1, CAMK4, GRB2, JUN, GAB1, TP53
hsa05010:Alzheimer's disease	8	4.54	0.015	MAPK1, APP, CDK5R1, CASP3, NOS1, PSEN1, BACE1, GRIN2A
hsa04650:Natural killer cell mediated cytotoxicity	7	4.87	0.017	MAPK1, CASP3, GRB2, PTK2B, ICAM2, LCK, VAV1
hsa04720:Long-term potentiation	5	6.80	0.029	MAPK1, EP300, CAMK4, GRIA2, GRIN2A
hsa04660:T cell receptor signaling pathway	6	5.14	0.030	MAPK1, GRB2, JUN, LCK, CD4, VAV1
hsa04662:B cell receptor signaling pathway	5	6.16	0.034	MAPK1, CR2, GRB2, JUN, VAV1
hsa04520:Adherens junction	5	6.00	0.036	MAPK1, EP300, SMAD3, CDH1, SRC
hsa04020:Calcium signaling pathway	7	3.68	0.040	NOS1, CAMK4, SPHK2, CHRM3, PTK2B, GRIN2A, NOS3
hsa04012:ErbB signaling pathway	5	5.31	0.049	MAPK1, GRB2, JUN, GAB1, SRC
hsa04350:TGF-beta signaling pathway	5	5.31	0.049	MAPK1, EP300, SP1, SMAD3, DCN

[a]Fold Enrichment

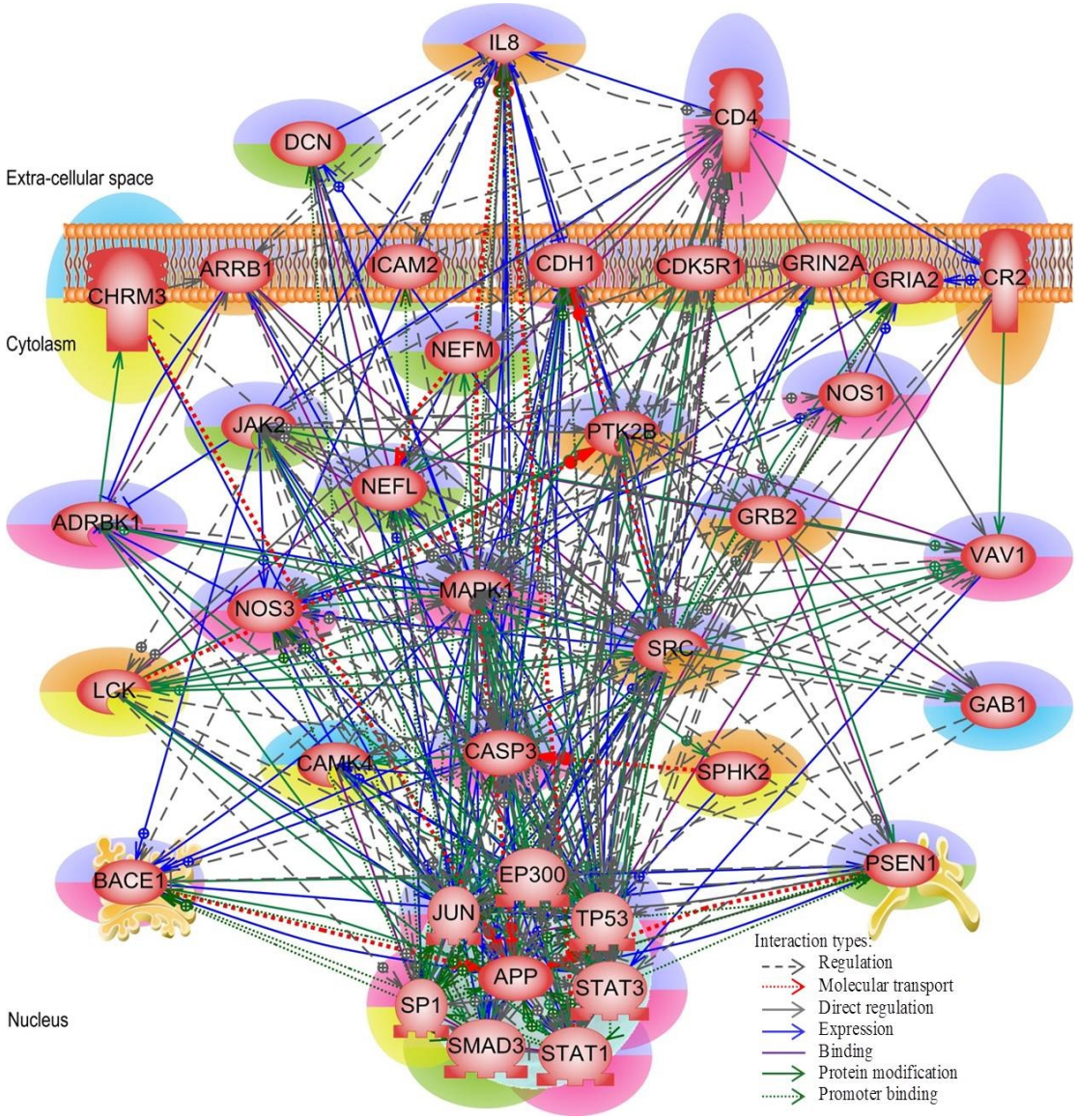


Figure 4.3. ***Integrated*** **Alzheimer's disease mechanism.** The 37 genes/proteins are found in common in all 12 enriched KEGG pathways. Genes/proteins implicated in AD pathology are highlighted in green/red and the genes/proteins of potential interest are highlighted in blue/orange (See Table 4.1. for details).

4.4.1 KEGG analysis

To analyze the interactions among these 37 common genes, we constructed an integrated AD mechanism network (see Figure 4.3). We found several routes of disease initiation via three extra-cellular proteins namely, CD4, DCN and IL8. Below we summarize the biological functions of the three proteins and their role in Alzheimer's disease pathogenesis. Increased accumulation of amyloid-beta plaques in AD brain has been shown to attract a surge of inflammatory mediators including chemokines, cytokines and interleukins leading to chronic inflammation that triggers neuronal death (Akiyama et al., 2000). IL8 (interleukin 8) is a member of the CXC chemokine family and one of the major mediators of the inflammatory response. Studies have shown overproduction of IL8 along with other interleukins found in specific regions affected in Alzheimer's disease. Under normal conditions, IL8 promotes cell-cell communications in central nervous system, but increased amyloid-beta production leads to chronic stimulation of IL8 response which is suspected as cause for abnormal physiology in AD brains (Gitter et al., 1995). Similarly, overproduction of CD4 has been associated with indirect neuronal damage in infectious and immune-mediated diseases of the central nervous system. Persistent amyloid-beta stimulus in AD brains has been implicated in chronic CD4 activation which results in premature immunosenescence (gradual deterioration of the immune system brought on by natural age advancement) manifestation (Larbi et al., 2009). DCN (decorin) is a small cellular or pericellular matrix proteoglycan. Earlier immunohistological study of AD brain tissues has found that decorin was primarily localized to the periphery of the amyloid plaque and to the edges of amyloid fibril bundles within the plaque periphery. It has been suggested that decorin molecules may contribute to regulating the size of amyloid accumulation and/or in maintaining their spherical shape (Snow

et al., 1992). In order to achieve continued homeostasis, a delicate balance has to be maintained between APP gene expression and various inflammatory mediators.

Thus, the three extra-cellular molecules of the integrated AD mechanism network have already been implicated in Alzheimer's disease pathogenesis in one way or another. The integrated network uncovered *direct* interactions between these three extra-cellular molecules and amyloid beta (A4) precursor protein, APP inside the nucleus. In addition, the network revealed mutually-connected interactions between APP, BACE1 and PSEN1 genes/proteins. Deregulation of APP, BACE1 (beta-site APP-cleaving enzyme 1) and PSEN1 are well-known causes of Alzheimer's disease where APP and PSEN1 mutations are associated with familial form of the disease and BACE1 gene codes for one of the proteases that cleaves APP whose accumulation is the hallmark of Alzheimer's disease. Apart from *direct* interactions with APP, BACE1 and PSEN1, the three extra-cellular molecules also trigger the downstream targets of MAP kinases, and caspases directing cell death.

4.4.2 Classification of disease causing and therapeutic mediators

Further, each of the 37 common genes of the enriched KEGG pathways was subjected to Google and NCBI's PubMed search for their biological/molecular functions, as well as for their potential role in AD pathogenesis. Based on this search, we categorized these 37 genes as either disease causing (leading to neuronal death) or possible therapeutics (helping with reducing the harmful effects and/or preventing neuronal loss) for Alzheimer's disease. This categorization is depicted in Figure 4.3 where the genes are highlighted either in *purple* or *yellow*, respectively.

As mentioned earlier, by modulating SRC tyrosine kinase gene expression we

could employ some inhibition of amyloid-beta accumulation in AD brains. But in general, the *integrated* disease mechanism network resulted into a well-connected destructive pathway with almost 80% of the 37 common genes found to play harmful roles in AD pathogenesis. We could try exerting gene/protein expression regulation through some of the transcription factor in the integrated network. However, the transcription factors found in this network were very generic and they are required for normal functioning of the cell. Thus, it was difficult for us to find/propose a node in the integrated Alzheimer's disease mechanism network through which we could modulate the expression of disease causing/aggravating genes. It was mentioned before that the amyloid-beta production has to be delicately balanced in order to maintain homeostasis within the cell. The integrated network revealed none to very little room to modulate APP gene/protein expression.

4.4.3 Proposed therapeutic measures

Apart from transcription factor's regulations, microRNAs provide additional post-transcriptional modulation of target gene expression. Even though many of microRNA's biological functions are still unclear, they offer many attractive features from a drug development standpoint such as small size, specific regulation, conserved nucleotide sequence etc., which makes them potential therapeutic entities. Currently, microRNAs as medical treatment means in various disease conditions are pursued either through the use of synthetic active microRNAs called mimics or through the use of anti-microRNAs. For example, miR-208 is used in cardiac diseases, miR-155 in chronic inflammatory diseases and miR-122 in Hepatitis C Virus (van Rooij et al., 2012). Many microRNA clinical trials are also ongoing.

We then sought after microRNA regulation in order to suppress the malicious

effects of Alzheimer's disease genes. For this, we used Pathway Studio's ResNet 9.0 database to identify the microRNAs that targets APP, BACE1, PSEN1, as well as the three extra-cellular molecules CD4, DCN and IL8. We found 22 such microRNAs including five of the already AD-known microRNAs namely, miR-153, miR-17, miR-20a, miR-29b-1 and miR-9. For example, BACE1 (beta-site APP-cleaving enzyme 1) is the target gene for miR-29a/b-1. BACE1 enzyme is essential for the generation of amyloid-beta, the hallmark neuropathological lesions in Alzheimer's disease brain. Earlier study shows that loss of miR-103, miR-29a/b-1 and miR-9 can contribute to increased BACE1 and Aβ levels in sporadic AD (Hébert et al., 2008). In addition to APP and BACE1 gene expression regulation, miR-9 was suspected to regulate PSEN1, a familial early-onset AD gene. Loss of PSEN1 impacts memory, synaptic plasticity and induces neurodegenerative changes (see above for more details). In a PSEN1 null mice model study, miR-9 down-regulation was correlated to severe brain defects (Barbato et al., 2009). Using transgenic mouse model of AD, it was found that miR-153 downregulated the expression of APP and APLP2 (a homologue of APP) proteins (Liang et al., 2012). Another study strongly suggests that miR-17-5p along with miR-20a and miR-106b may contribute, at least in part, to the developmental regulation of APP expression in brain and in differentiating neurons (Hébert et al., 2009).

With possible therapeutic role of microRNAs in mind, we incorporated the 22 microRNAs into our integrated Alzheimer's disease mechanism network (see Figure 4.4). Following this we examined the microRNAs and their targets. BACE1 had the highest microRNA regulation being targeted by nine miRNAs, next were APP and CD4, each of which was targeted by five microRNAs, then IL8 with three microRNA regulation, and one each for PSEN1 and CD4. In addition, miR-320a was found to target both APP and IL8, similarly miR-181a1 targeted both CD4 and DCN. In

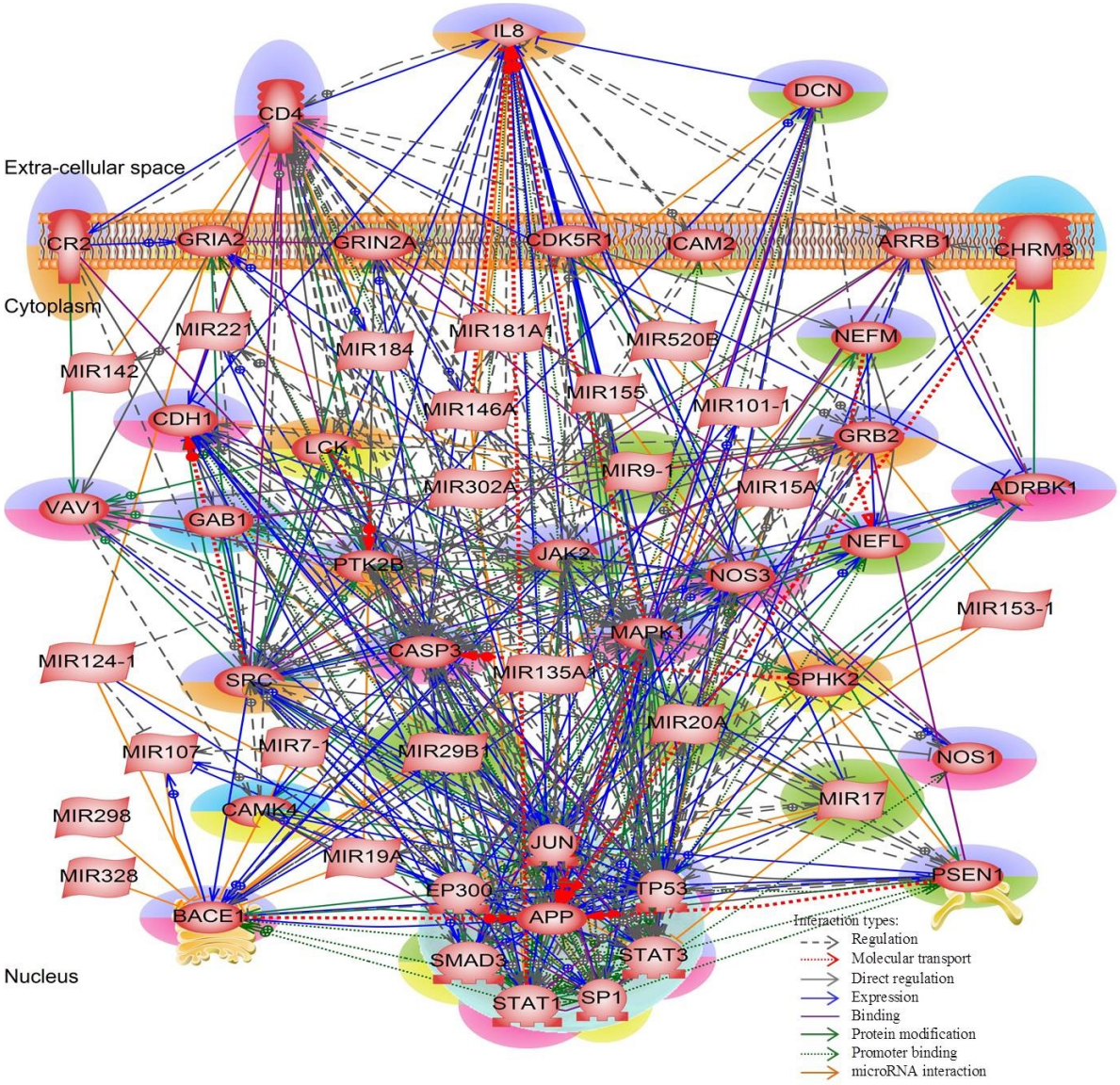


Figure 4.4. MicroRNA incorporated *integrated* Alzheimer's disease mechanism. The 37 genes/proteins found in common in all 12 enriched KEGG pathways are regulated by 22 microRNAs. Genes/proteins implicated in AD pathology are highlighted in green/red and the genes/proteins of potential interest are highlighted in blue/orange (See Table 4.1.) for details.

conclusion, by incorporating microRNAs into the integrated network, we were able to suggest some critical nodes to regulate Alzheimer's disease causing and/or aggravating genes. Using network techniques we could offer both narrow and specific range of entities through which the network could be modulated. Thus, we propose 17 microRNAs (miR-101-1, miR-107, miR-124-1, miR-135a1, miR-142, miR-146a, miR-155, miR-15a, miR-181a1, miR-184, miR-19a, miR-221, miR-298, miR-302a, miR-328, miR-520B and miR-7-1) in addition to the known AD-related five as potential therapeutic targets in Alzheimer's disease. Most of these novel microRNAs are not yet experimentally verified but one may expect that after such verification a sizable part of them will be found of therapeutic interest.

4.4.4 Alzheimer's disease *microRNA* regulatory network

Besides finding the microRNAs that could regulate the critical nodes such as APP, BACE1, CD4, DCN, IL8 and PSEN1, we searched to uncover additional regulatory mechanisms of Alzheimer's disease genes. For this we used all the 214 significantly differentially expressed AD genes and constructed a *shortest-path* network with only microRNA interactions found in Pathway Studio's ResNet 9.0 database. This SPNW revealed that 67 miRNAs regulate many of the 214 AD SDEGs. Utilizing these 67 miRNAs and 214 SDEGs, we then constructed a *direct* miRNA-mRNA targets interaction network shown in Figure 4.5.

In addition to miRNAs regulation, the network also incorporated regulation by ten transcription factors like BCL6, LMO4, HEY2 and SMAD3 which were not like the very generic TFs reported earlier. Thus, like in other complex diseases, cooperative two-level gene expression regulatory mechanisms occur in Alzheimer's disease pathogenesis via miRNAs and transcriptional factors.

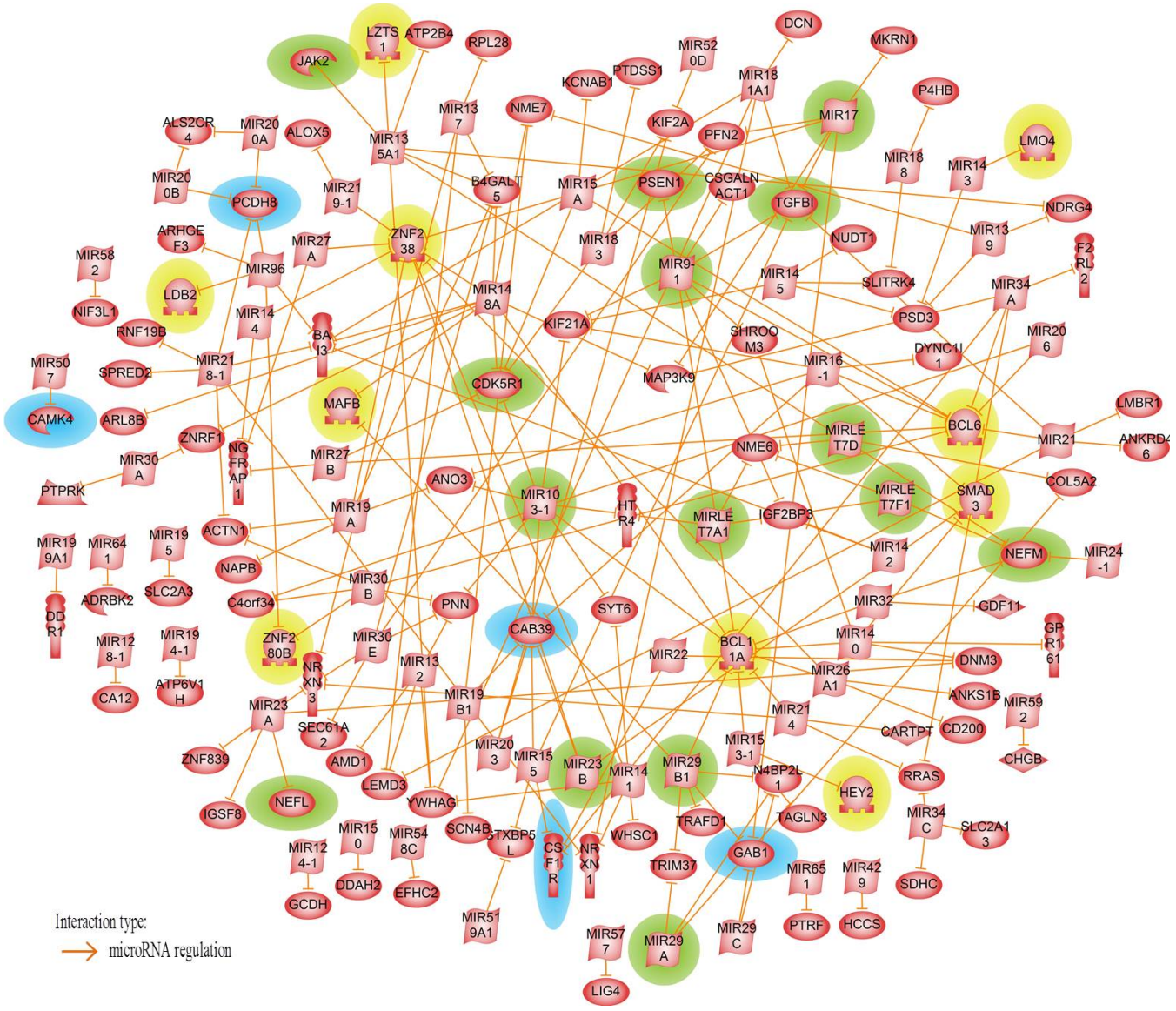


Figure 4.5. Alzheimer's disease *miRNA regulatory* network.The genes/proteins and miRNAs implicated in AD pathology are highlighted in green and the genes/proteins of potential interest are highlighted in blue. Genes/proteins that code for transcription factors (TFs) are highlighted in yellow.

On further examination of microRNA regulatory network (Figure 4.5), miR-148a was the node with the highest degree (hub), closeness and betweenness centrality scores. It regulates gene expression of ten genes including CDK5R1 whose expression is needed for proper CDK5 activity, as a part of the AD molecular mechanism discussed above. MicroRNA quantitative biochemical study has reported altered expression of miR-148a along with other microRNAs in hippocampus and medial frontal gyrus regions of human AD brain samples (Cogswell et al., 2008). In addition, our network analysis revealed that expression of CDK5R1 gene could also be regulated by five microRNAs (miR-103, miR-148a, miR-15a, miR-183, and miR-27b), of which regulation via miR-103 along with miR-107 has already been experimentally established (Moncini et al., 2011).

Additional microRNA regulation of some of the genes of interest in Alzheimer's disease was uncovered in the microRNA regulatory network. It includes modulation of CAB39 gene expression by six microRNAs (miR-103-1, miR-155, miR-16-1, miR-19b1, miR-203 and miR-23b), PCDH8 by four microRNAs (miR-200a/b, miR-218-1 and miR-96), CSF1R by three microRNAs (miR-155, miR-22 and miR-34a), GAB1 by three microRNAs (miR-17, and miR-29b1/c), CAMK4 by miR-507 and HEY2 by miR-153-1, which would require experimental validation. Finally, we found that miR-148a, miR-155 and miR-153 seem to target many of the critical genes, as well as genes of interest, in Alzheimer's disease.

4.5 Summary

Online Mendelian Inheritance in Man (OMIM) database search for Alzheimer's disease revealed many genes such as A2M, ACE, APOE, APP, NOS3, PSEN1, PSEN2 and SORL1 implicated in the disease mechanism. However, none of these genes except

PSEN1 were included in our "seed genes" since these genes were not strongly differentially expressed, and also their log-fold change was minimal (ranged from -0.3 to +0.1). We consider the lowered gene expression intensity in the post-mortem brain samples compared to a functioning brain as attributing factor to this gene set low variability. But it is valuable to note that APP and NOS3 along with other well-known AD genes like AR, BACE1, STAT1 etc., although not significantly changed were later included as one of the many connecting genes/proteins in our shortest-path network and thus became critical nodes in the closest network environment of the known significantly changed Alzheimer's genes. These findings reflect upon our ideas that it is beneficial to utilize network techniques to study complex systems like neurodegenerative diseases where even in the absence of an appropriate study medium (a fully functional brain) we could recreate some of missing links/nodes that are valuable to understand the underlying system mechanism.

It should be mentioned that our detailed network analysis was unable to confirm the AD role of APOJ, TIMP3 and IRAK1 genes reported by Dunckley et al. 2006. These genes were not statistically significant after the normalization we performed on their dataset, and they did not emerge in any way in the networks we constructed on the basis of well-established AD genes. Identification of PSEN1, a well-known familial Alzheimer's disease gene as an important node in both *direct*-interaction and shortest-path network is critical in number of ways. First, PSEN1 was found to be significantly differentially expressed gene (p-value < 0.01) in the microarray datasets. Secondly, it directly interacts with many already known AD genes such as APP, CDH1, CDK5R1, CREB1, JUN, SP1 and STAT3. It is also among the top 33 nodes with degree ≥ 8, and top 26 nodes with higher accessibility (closeness centrality index) to all other nodes in the *compact* shortest-path network, emerging thus with a major role in the molecular mechanism of Alzheimer disease.

Our network analysis provided sufficient arguments in favor of CAMK4 and GAB1 as novel candidate genes for Alzheimer's disease based on their *direct* network interactions with already known AD genes. With somewhat less confidence we propose also CSF1R gene, which despite being a second-level neighbor, has demonstrated important role in neuronal survival after injury and degeneration. All three genes are significantly differentially expressed in the post-mortem AD samples we used and they could have critical role as neuroprotective agents. From shortest-path network assessment, being first-level interacting partners with three or more already known AD-related genes we conclude that *connecting* genes/proteins added by the Pathway Studio software such as ESRRA, MDM2, NRF1, PTK2B, SRC and SRF could also be of potential interest in Alzheimer's disease domain. Among the above listed genes, MDM2 and SRC were directly connected with at least *ten* or more previously known AD genes (!). Additionally, one may consider including to the list of Alzheimer's genes five more from the connecting genes in the shortest-path network (CHRM3, CR2, HEY2, IL3 and NOX4) via their guilt-by-associations to one or two already known AD genes. The ARRB1 and SPHK2 connecting genes, which are second-level interacting partners with previously implicated AD genes, may also be viewed of interest via their role as modulators in amyloid-β production (Takasugi et al., 2011; Wolfe, 2013).

Apart from directly interacting with a large number of known AD-related genes/proteins as shown in the guilt-by-association list above, MDM2 and SRC are among the top 25 hub nodes (with degree > 10), among the top 25 nodes with higher accessibility to all other nodes (as measured by closeness centrality network descriptor) and among the top 25 traffic influential nodes in the network (as measured by the betweenness centrality index). All this may be considered as evidence for a potential important systems biology role of these genes/proteins in the integral mechanism

of Alzheimer's disease.

In addition to identifying novel individual genes, proceeding from our *integrated* Alzheimer's disease mechanism network, we propose three disease initiating routes via extra-cellular ligands like CD4, DCN and IL8. These genes directly interact with many known AD genes including APP, BACE1 and PSEN1 and influence AD pathogenesis mechanism. Since this network was well-connected with almost 80% of genes promoting neurodegeneration process, we were unable to propose any therapeutic targets or mechanisms using this network. However, by incorporating microRNA regulations to this network, we identified 17 microRNAs (miR-101-1, miR-107, miR-124-1, miR-135a1, miR-142, miR-146a, miR-155, miR-15a, miR-181a1, miR-184, miR-19a, miR-221, miR-298, miR-302a, miR-328, miR-520B and miR-7-1) in addition to the known five AD-related miRNAs as potential therapeutic targets in Alzheimer's disease. More generally, we propose a scheme of complex multi-level regulations taking place between the critical players of AD, such as APP, BACE1, and PSEN1, and other disease causing/alleviating gene/protein entities. In order to maintain continued homeostasis, a delicate balance should be sustained between the genes, and microRNA regulation could play an important role in this scenario. Perhaps, a well-designed cellular automata simulation would help to shed some light on this delicate but life-determining balance.

The genes discussed in the summary were also shown by gene ontology analysis to be key players in numerous specific biological mechanisms like neuroprotection against amyloid-beta accumulation, formation of neurofibrillary tangles, synapses, cognitive deficits, mental retardation, memory loss and attenuating APP accumulation, many of which are known to be modulated in the underlying Alzheimer's pathophysiology and potential compensatory responses.

Chapter 5

Huntington's Disease Network Analysis

5.1 Introduction

Unlike the Parkinson's and Alzheimer's disease network analysis, we found and used only one post-mortem human brain tissue sample dataset namely, GSE3790 for Huntington's disease (HD). However, this dataset had a larger number of samples (44 HD and 36 controls) than the used PD and AD samples. Hodges et al. published this dataset in 2006. One of the major conclusions of this study was that the differential gene expression in HD brains showed distinct regional pattern similar to an already known pattern of neuronal loss. The greatest number and magnitude of differentially expressed genes/proteins were detected in the caudate nucleus, followed by motor cortex, then in cerebellum tissue types. Our statistical analysis with renormalized data revealed a similar gene expression difference in these tissue samples.

The original authors had performed gene set enrichment analysis using DAVID tools to identify biological processes and pathways significantly affected in HD. In this study, we expanded upon their research work by subjecting the differentially

expressed genes to various network techniques to explore the underlying cellular mechanisms and molecular players of Huntington's disease. We initiated our HD network analysis using 531 "seed genes" (see Chapter 2 for details).

5.2 Huntington's disease *direct* interaction (DI) network

Compared to Parkinson's and Alzheimer's number of seed genes, for Huntington's disease we had a larger number of seed genes to deal with. Working with a larger gene set is beneficial in a couple of ways. On one side one is having a broader view of the network neighborhood of any gene of interest along with all its complex interactions. On the other side, at times one could also shrink the network size in order to have a closer look on the proximity of those nodes that are important. Figure 5.1 shows the primary *direct* interaction (DI) network of the 531 Huntington's disease "seed genes". 224 of these genes directly interact with each other and the network included such types of interactions like regulation, physical binding, co-expression, promoter binding, protein modification, molecular transport and direct regulation.

After exhausted literature review, we found in the *direct* interaction network 23 genes that have already been associated to Huntington's disease (more details on which are given in Chapter 1). Another eleven genes (CNTNAP1, CX3CL1, DPYSL5, FDFT1, FGFR1, FKBP5, PPARA, PRDX2, PRDX6, RCAN2 and ZNF148) were identified as being of potential interest in Huntington's disease pathogenesis due to their specific molecular functions as detailed below. More details on how we conducted our literature search and gene classification are given in Chapter 2. We constructed

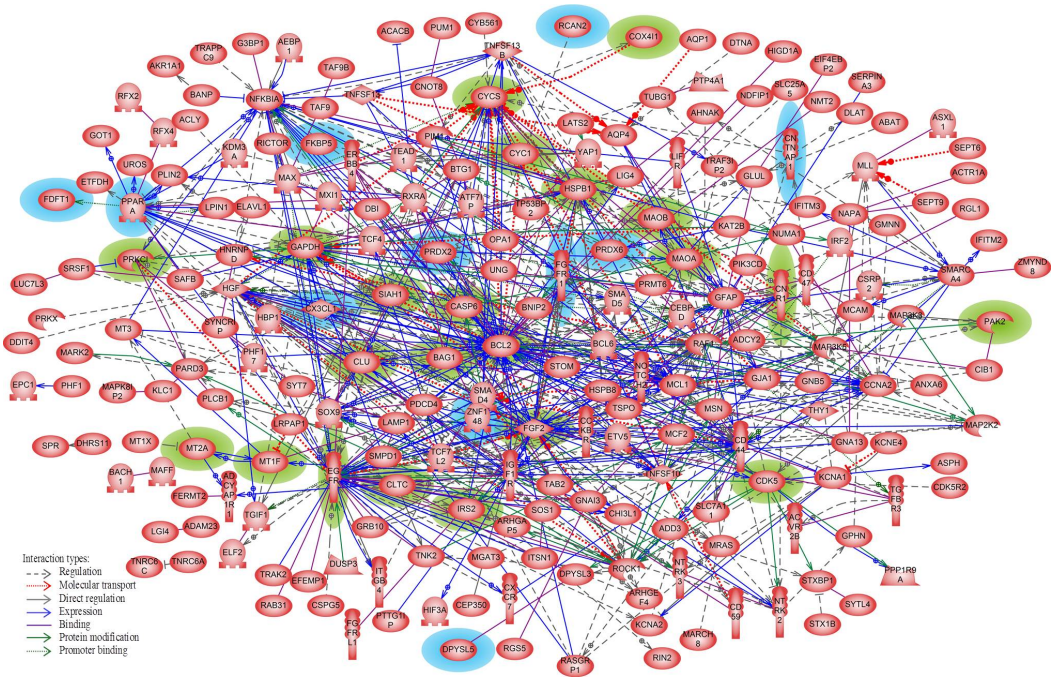


Figure 5.1. Huntington's disease *direct interaction* network. The 23 genes/proteins implicated in HD pathology are highlighted in green and the eleven genes/proteins of potential interest for that disease are highlighted in blue.

a *compact* direct interaction network using these 34 genes/proteins to understand whether and how they interact with each other, and to validate the eleven genes of interest by their guilt-by-association. The network shown in Figure 5.2 clearly reveals that except FKBP5 and DPYSL5, all other genes/proteins of potential interest are connected to the HD-known ones.

5.2.1 Molecular functions of some of the proposed candidate genes

The intrinsic biological functions of the proposed candidate genes seem promising for a neuroprotective mechanism in Huntington's disease realm. For instance, fractalkine (CX3CL1) is a known *Parkinson's* disease gene where it exhibited neuro-

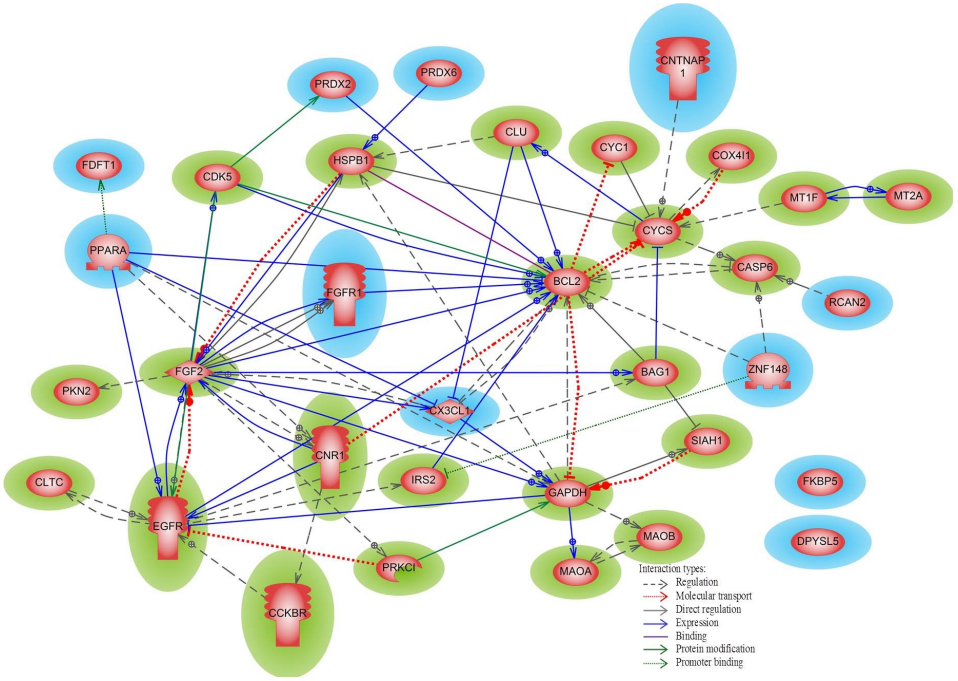


Figure 5.2. Huntington's disease *compact* direct interaction network. The 23 genes/proteins implicated in HD pathology are highlighted in green and the eleven genes/proteins of potential interest for that disease are highlighted in blue.

protective role against microglia activation as well as reduced motor coordination impairment (see Chapter 1 for details). Being a known neuroprotective agent for a similar neurodegeneration disease with movement disorder, it may have a potential therapeutic role in Huntington's disease domain too. As mentioned earlier, FGFs were proposed to improve the motor performance and to extend the lifespan in HD mouse model study. Being a receptor for fibroblast growth factors, FGFR1 could up-regulate FGF's beneficial activities in the cell. Studies have found that PPAR-gamma together with PGC-1α (a transcriptional co-activator) are required for the regulation of mitochondrial biogenesis. PPAR-γ agonists are thought to be neuroprotective in ALS and HD (Xiang et al., 2011). Similarly, PPAR-alpha could be another means for a neuroprotective pathway in HD.

Currently, there is no cure available for HD patients. In a mouse model of HD

study, zinc finger proteins were designed in such a way that these proteins were able to recognize and bind with CAG repeats in DNA, specifically. The study reported that there was considerable reduction of the mutant huntingtin gene expression at both protein and mRNA levels (95% and 78% reduction, respectively) (Garriga-Canut et al., 2012). Many zinc finger proteins including ZBTB10, ZFP36L1 and ZNF148 were found significantly differentially expressed in the microarray dataset used in this study. These zinc finger proteins could emerge as a promising new gene therapy tool for Huntington's disease which could be extended and tested in human HD patients.

5.2.2 Network attributes of some of the proposed candidate genes

The genes such as CX3CL1, FGFR1, PPARA and zinc fingers including seven others are recommended as novel candidate genes for HD due to their biological/cellular functions. Additionally, from network perspective (see Figure 5.2), many of these candidate genes are *direct* interacting partners with already HD-associated genes. For example, CX3CL1 and PPARA interacts with four already known-HD genes namely, BCL2, CLU, FGF2 and GAPDH and BCL2, CNR1, EGFR and GAPDH, respectively. In addition, CX3CL1 and PPARA directly interact with each other. The modulation of these previously known-HD genes have been shown beneficial in reducing the disease pathogenesis. Similarly, ZNF148 directly interacts with three known-HD genes BCL2, CASP6 and IRS2.

Besides being first-level interacting partners with already implicated-HD genes, the novel candidate gene PPARA is one of the top 10 nodes with highest connectivity (node degree ≥ 20) as well as betweenness (traffic-influential) centrality scores. In addition, it is one of the top 20 nodes with highest closeness (accessible to other

nodes) centrality measures in the network. Their innate physiological roles along with their vital network attributes, increases the chance of the eleven candidate genes to be involved in the HD pathology, which would require further investigation, particularly for FKBP5 and DPYSL5, which remained unconnected in the *direct* interaction network.

5.3 Huntington's disease *shortest-path* network (SPNW)

Different from the previous *direct* interaction network type, *shortest-path* network help us to identify indirect protein-protein interactions that take place through intermediary nodes in the absence of direct relationship. With this in mind, we attempted to build the *shortest-path* network using only those seed genes which had at least 25 neighbors in Pathway Studio ResNet 9.0 database (refer Chapter 2). 258 out of 531 seed genes met this cut-off criteria and the resulted *shortest-path* network included 208 Pathway Studio software-added connecting genes. Following our study methodology, we categorized these connecting genes into two groups. The genes that were already implicated in Huntington's disease belong to the known-HD genes group and those genes which could be of potential interest in HD due to their cellular functions were grouped separately. Table 5.1 shows the different categories and the number of genes in each.

We constructed a *compact* shortest-path network (CSPNW), using the 85 genes from Table 5.1 along with few additional connecting genes that were needed to have a unified well-connected network. The average node degree of this *compact* shortest-path network was 7.10. In this CSPNW, many of the known-HD genes such as AKT1, AR, BCL2, INSR and SP1 were among the top 25 nodes with node degree ≥ 10, as well as the top 25 with highest closeness (network monitors) and betweeness (traffic

Table 5.1. Summary of the genes of interest and genes already known in Huntington's disease.

Different categories	Number of genes	Node color in figure
Genes of interest from SDEGs	14	blue
Known HD genes from SDEGs	24	green
Genes of interest in SPNW connecting nodes	24	orange
Known HD genes in SPNW connecting nodes	23	red

influential) centrality measures. Figure 5.3 illustrates the interactions between the known and the genes of interest in the HD *compact* SPNW.

Even though PRNP gene/protein was not statistically differentially expressed in our microarray dataset, it was valuable to note that it emerged as connecting gene/protein in the *shortest-path* network. PRNP (prion protein) is a glycoprotein that tends to aggregate into rod-like structures causing neuronal cell death. Prion proteins have been associated with many neurodegenerative disorders including Huntington's, Creutzfeldt-Jakob diseases in human, and "mad cow" disease in cattle (Imran and Mahmood, 2011; Moore et al., 2001). We found that PRNP was of importance for network topology as one of the top 15 nodes with highest visibility (closeness) and most influence (betweeness) in the *compact* shortest-path network.

5.3.1 Guilt-by-association analysis

In the next few paragraphs, we summarize the innate molecular characteristics of various genes that could be of potential interest in HD pathogenesis, in addition to their "guilt-by-association" relationship to some of the already implicated HD genes. Table 5.2 lists our proposed Huntington's disease candidate genes along with the number of known-HD genes to which they directly interact with (see Figure 5.3).

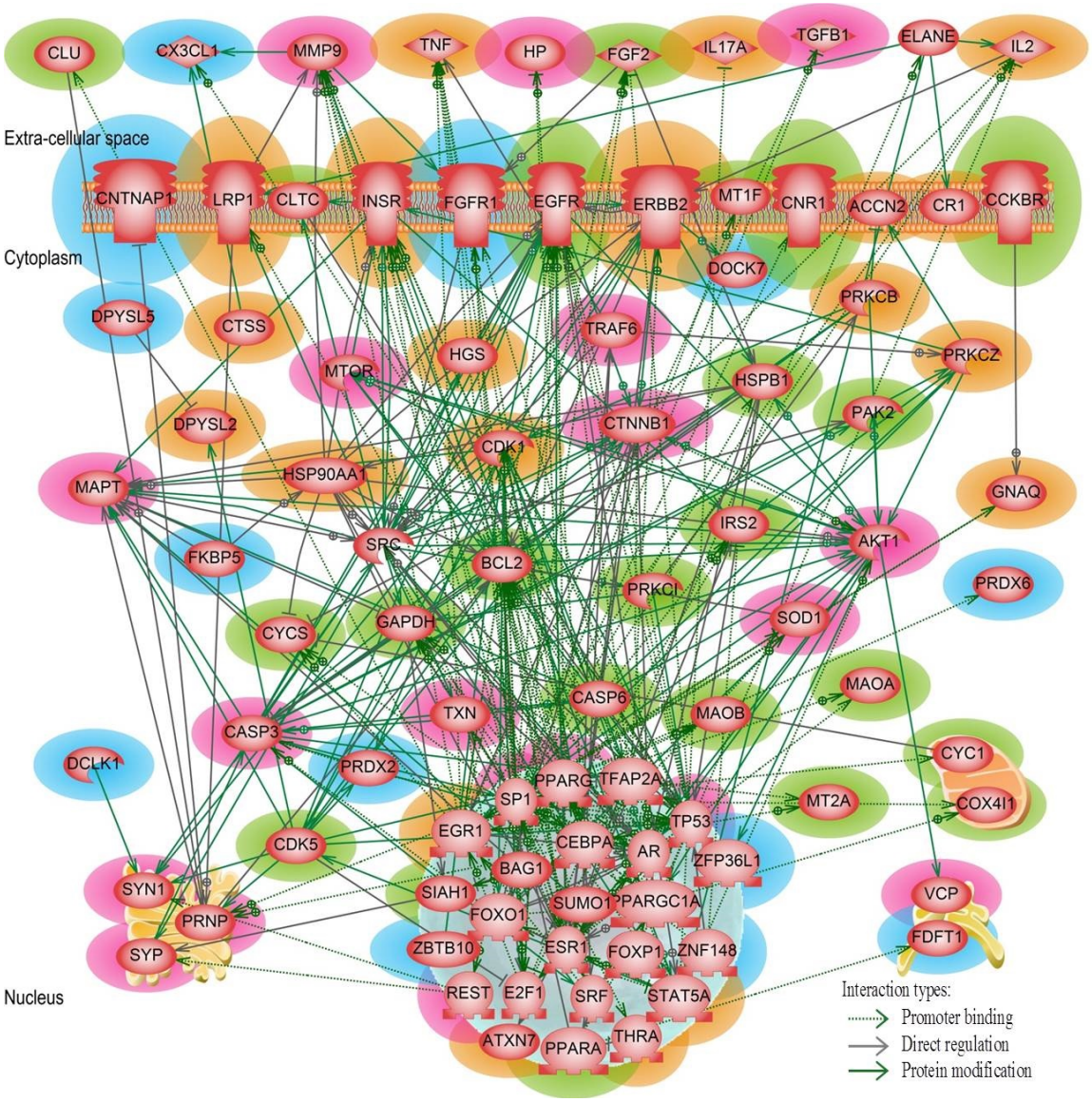


Figure 5.3. Huntington's disease *compact* shortest path network. The genes/proteins implicated in HD pathology are highlighted in green and red. The genes/proteins of potential interest are highlighted in blue and orange. (see Table 5.1. for gene highlighting details).

Table 5.2. Genes of interest for Huntington's Disease identified by guilt-by-association with the known HD-related genes.

Genes of Interest	Interacts with no. of known-HD genes
EGR1	**13**
CEBPA	9
CDK1	8
FOXO1, HSP90AA1	7
PRKCZ	6
E2F1, STAT5A	5
SRF	4
ERBB2, FGFR1, IL2, INSR, LRP1, PRKCB, TNF, ZNF148	3
CX3CL1, GNAQ, PPARA	2
ATXN7, CNTNAP1, DCLK1, DPYSL2, PRDX2, PRDX6, THRA, ZFP36L1	1

From the *compact* shortest-path network, we found that EGR1 (early growth response 1) was first-level interacting partner to 13 (!) previously known-HD genes (AKT1, AR, BCL2, CLU, CTNNB1, CYCS, FGF2, MAOB, MMP9, SOD1, SP1, TGFB1 and TP53). BIM (BCL2-like 11 apoptosis facilitator) plays an important role in neuronal apoptosis, a hallmark feature of many neurological diseases including Alzheimer's and Parkinson's. Previous research study had demonstrated that EGR1 directly transactivate BIM gene expression to promote neuronal apoptosis. EGR1/BIM pathway has been suggested as a pro-apoptotic mechanism in neurological diseases. Mithramycin A, a U.S. Food and Drug Administration clinically approved drug has been studied to improve motor symptoms and extend life span in a mouse model of Huntington's disease. Mithramycin A was suspected to exploit the EGR1/BIM pathway to promote neuroprotective mechanism in HD models and thus could be a promising drug for the treatment for the same (Ferrante et al., 2004; Xie et al., 2011).

Our next HD candidate gene is CEBPA (CCAAT/enhancer binding protein, alpha). It has been shown to bind to the promoter and modulate the expression of the gene encoding for leptin, a protein that plays an important role in body weight homeostasis. Leptin receptors are found in various brain regions such as the hippocampus and cerebral cortex, and have known roles in neural development and neuroendocrine functions. Studies have indicated that leptin could be neuroprotective and thus enhance neuronal survival (Tang, 2008). CEBPA, the promoter of leptin gene could also play a critical role in this neuroprotective mechanism. In the *compact* shortest-path network, CEBPA was the first-level interacting partners with nine known-HD genes.

Another recommendation for HD candidate gene is CDK1 (cyclin-dependent kinase 1).The abnormal activation of CDK1 is likely to be involved in the neuronal cell loss in neurodegenerative diseases including Alzheimer's disease and HIV (Castedo et al., 2002). Earlier, CDK5 was suspected to contribute to the deleterious protein accumulation in Alzheimer's disease (Iijima et al., 2000; Crews and Masliah, 2010; Mateo et al., 2009). In the *compact* SP network, CDK1 directly interacts with *eight* known-HD genes. CDK1 is a part of the kinase family that is actively contributing to the neurodegeneration process in similar disease conditions; it could have potential role in HD neurodegeneration mechanism too.

Thus, through their intrinsic molecular functions and their pivotal neighboring positions with many already implicated HD genes in the *compact* shortest-path network, we propose those genes listed in Table 5.2 as novel candidate genes for Huntington's disease. Undeniably, these candidate genes should be further investigated for their molecular role in HD. However, we expect many of these novel genes to surpass the experimental verification due to their "guilt-by-association" with previously established Huntington's disease genes.

5.3.2 Molecular functions and network attributes of some of the critical connecting genes

Apart from novel connecting genes, it is valuable to note that genes like EGFR, ESR1, HSBP1 and MAPT were also included in the *compact* SP network as connecting genes. They are previously known contributors as well as therapeutic agents in neurodegenerative disorders. Hyperphosphorylated tau (MAPT) is the major component of the neurofibrillary tangles, one of the hallmarks of neurodegenerative diseases (Jellinger, 2009; Pittman et al., 2006; Lei et al., 2010; Crews and Masliah, 2010). Similarly, huntingtin gene, the mutant in HD was suggested to be indirectly associated with EGFR, thus deregulating the downstream actions of EGFR leading to cell death (Liu et al., 1997). Considering the treatment measures, HSBP1 and ESR1 are suggested to offer such mechanisms. In general, heat shock proteins (HSBP1) are evaluated as therapeutic targets in mitigating or preventing protein aggregate formations (Mymrikov et al., 2011). One of the major conclusions of an animal HD model study was that the female sex hormone, estrogen (ESR1) could be a target for neuroprotective therapy aiming at postponing the onset and reducing the severity of HD. A similar pattern of late onset was also shown in a human HD study (Bode et al., 2008; Roos et al., 1991). Moreover, these four connecting genes were among the top 25 nodes with highest connectivity (degree > 10) as well as one of the top 25 nodes with highest visibility as measured by the closeness centrality scores. Except HSBP1, the other three genes were also among the top 25 nodes with highest accessibility to other nodes in the network as determined by the betweenness centrality in the *compact* SP network.

Table 5.3. Gene set DAVID enrichment analysis of Huntington's disease *compact* shortest-path network.

Category	Term	Gene count	Fold Enrichment	Benjamini
GOTERM_BP_FAT	GO:0048666 neuron development	14	6.50	5.99E-06
GOTERM_BP_FAT	GO:0030182 neuron differentiation	15	5.39	1.44E-05
GOTERM_BP_FAT	GO:0046324 regulation of glucose import	6	28.60	4.33E-05
GOTERM_CC_FAT	GO:0043005 neuron projection	13	6.07	1.37E-04
GOTERM_BP_FAT	GO:0050994 regulation of lipid catabolic process	5	30.25	3.09E-04
GOTERM_BP_FAT	GO:0010001 glial cell differentiation	6	17.81	3.09E-04
GOTERM_BP_FAT	GO:0043523 regulation of neuron apoptosis	7	12.23	3.22E-04
GOTERM_BP_FAT	GO:0031175 neuron projection development	10	6.14	4.40E-04
GOTERM_BP_FAT	GO:0043627 response to estrogen stimulus	7	10.49	6.42E-04
GOTERM_BP_FAT	GO:0042063 gliogenesis	6	14.52	6.76E-04
GOTERM_BP_FAT	GO:0048812 neuron projection morphogenesis	9	6.65	6.99E-04
GOTERM_BP_FAT	GO:0006979 response to oxidative stress	8	7.67	8.94E-04
GOTERM_BP_FAT	GO:0001836 release of cytochrome c from mitochondria	4	29.96	0.003
GOTERM_BP_FAT	GO:0048667 cell morphogenesis involved in neuron differentiation	8	6.02	0.003
GOTERM_CC_FAT	GO:0030425 dendrite	8	7.84	0.004
GOTERM_CC_FAT	GO:0030424 axon	8	8.04	0.004
GOTERM_BP_FAT	GO:0007568 aging	6	8.58	0.006
GOTERM_MF_FAT	GO:0008289 lipid binding	11	3.82	0.009
GOTERM_BP_FAT	GO:0000302 response to reactive oxygen species	5	10.49	0.010
GOTERM_BP_FAT	GO:0050727 regulation of inflammatory response	5	10.35	0.011
GOTERM_BP_FAT	GO:0007409 axonogenesis	7	5.71	0.011
GOTERM_BP_FAT	GO:0008286 insulin receptor signaling pathway	4	17.01	0.013
GOTERM_BP_FAT	GO:0050804 regulation of synaptic transmission	6	6.94	0.013
GOTERM_BP_FAT	GO:0031644 regulation of neurological system process	6	6.17	0.020
GOTERM_BP_FAT	GO:0016192 vesicle-mediated transport	11	3.00	0.023
GOTERM_BP_FAT	GO:0050767 regulation of neurogenesis	6	5.69	0.027
GOTERM_BP_FAT	GO:0030518 steroid hormone receptor signaling pathway	4	10.85	0.037
GOTERM_BP_FAT	GO:0006874 cellular calcium ion homeostasis	6	5.16	0.038
GOTERM_CC_FAT	GO:0030136 clathrin-coated vesicle	6	7.26	0.039
GOTERM_BP_FAT	GO:0055114 oxidation reduction	11	2.71	0.042
GOTERM_CC_FAT	GO:0045121 membrane raft	6	6.70	0.045

5.3.3 DAVID enrichment analysis

Continuing with our network analysis, we subjected the genes in the *compact* shortest-path network to DAVID analysis to identify various enriched biological processes and pathways in Huntington's disease. Table 5.3 lists some of the Gene Ontology categories/subcategories related to nervous system and functions that were statistically significantly enriched in HD (with Benjamini-Hochberg multiple correction). DAVID analysis uncovered biological processes involving in oxidative stress, reactive oxygen species, deregulation in inflammatory response, steroid hormone receptor signaling, lipid binding and insulin receptor signaling pathways to be significantly affected in Huntington's disease, some of which were mentioned above. Other neurodegenerative signaling pathway including Alzheimer's and ALS were also considerably affected in Huntington's disease which reinforce our view for similar underlying molecular pattern in all these diseases. In the next section, we will provide detailed information about our proposed model for Huntington's disease mechanism based on the various biological pathways that were affected in this disease.

5.4 *Integrated* Huntington's disease mechanism

In addition to various biological processes, DAVID analysis also recommended several KEGG pathways to be significantly (p-values < 0.05 with Benjamini-Hochberg multiple correction) affected in Huntington's disease of which we selected ten pathways for further evaluation. These pathways were selected (listed in Table 5.4) on the basis of previous implications in Huntington's disease research work. Either the entire pathway or many important players of the pathways were found deregulated in HD pathogenesis (Bae et al., 2005; Godin et al., 2010; Reis et al., 2011;

Andreassen et al., 2002; Phan et al., 2009; Lalić et al., 2008; Moreira Sousa et al., 2013; Apostol et al., 2006).

5.4.1 KEGG analysis

Enriched KEGG pathways belong to endocrine system, cell communication, cell growth and death, signal transduction, neurodegenerative diseases, and endocrine and metabolic diseases classification. We used the ten pathways (see Table 5.4) to search for any underlying molecular mechanism that could either cause or mitigate Huntington's disease pathology. To accomplish this task, we constructed an *integrated* HD mechanism network using the 41 genes found in common in all the ten enriched KEGG pathways (see Figure 5.4).

We then performed a Huntington's disease literature search to classify these 41 genes into two groups namely, genes that aid in the neuronal survival or cause loss. 23 out of 41 were implicated in neuronal loss and the remaining 18 genes were related to neuronal survival. This classification is depicted in Figure 5.4 where genes are highlighted in *purple* and *yellow*, respectively. Similar to our previous integrated neurodegenerative disease mechanism networks shown in Chapters 3 and 4, we found a pattern of three extra-cellular ligands (FGF2, TNF, and TGFB1) initiating various downstream signaling cascades in the *integrated* Huntington's disease mechanism network as well. As explained earlier, FGF2s are pursed as promising drug targets for its neuroprotective and neuroproliferative roles. TNF (tumor necrosis factor) and TGFB1 (transforming growth factor, beta 1) belong to inflammatory cytokine family which is involved in the regulation of a wide variety of biological processes including cell proliferation, differentiation, adhesion, apoptosis, lipid metabolism, and coagulation. Neuroinflammation has been implicated in a many

Table 5.4. Enriched KEGG pathways in Huntington's disease resulted from DAVID analysis.

KEGG Pathways	Gene count	FE [a]	Benjamini	Genes
hsa05016:Huntington's disease	12	5.14	4.40E-04	CASP3, GNAQ, SP1, CYCS, PPARG, CYC1, TP53, COX4I1, REST, CLTC, SOD1, PPARGC1A
hsa04010:MAPK signaling pathway	14	4.04	5.86E-04	EGFR, FGFR1, TNF, TP53, SRF, TGFB1, PRKCB, AKT1, CASP3, PAK2, MAPT, HSPB1, TRAF6, FGF2
hsa04012:ErbB signaling pathway	8	7.08	0.002	EGFR, AKT1, PAK2, ERBB2, STAT5A, MTOR, SRC, PRKCB
hsa05010:Alzheimer's disease	10	4.73	0.003	CASP3, TNF, LRP1, GNAQ, MAPT, CYCS, CYC1, COX4I1, GAPDH, CDK5
hsa05014:Amyotrophic lateral sclerosis (ALS)	6	8.72	0.006	CASP3, TNF, BCL2, CYCS, TP53, SOD1
hsa04910:Insulin signaling pathway	8	4.57	0.011	AKT1, PRKCZ, IRS2, PRKCI, FOXO1, MTOR, INSR, PPARGC1A
hsa04920:Adipocytokine signaling pathway	6	6.90	0.012	AKT1, PPARA, IRS2, TNF, MTOR, PPARGC1A
hsa04930:Type II diabetes mellitus	5	8.20	0.015	PRKCZ, IRS2, TNF, MTOR, INSR
hsa04520:Adherens junction	6	6.00	0.016	EGFR, FGFR1, ERBB2, INSR, SRC, CTNNB1
hsa04115:p53 signaling pathway	5	5.67	0.049	CDK1, CASP3, CYCS, TP53, SIAH1

[a]Fold Enrichment

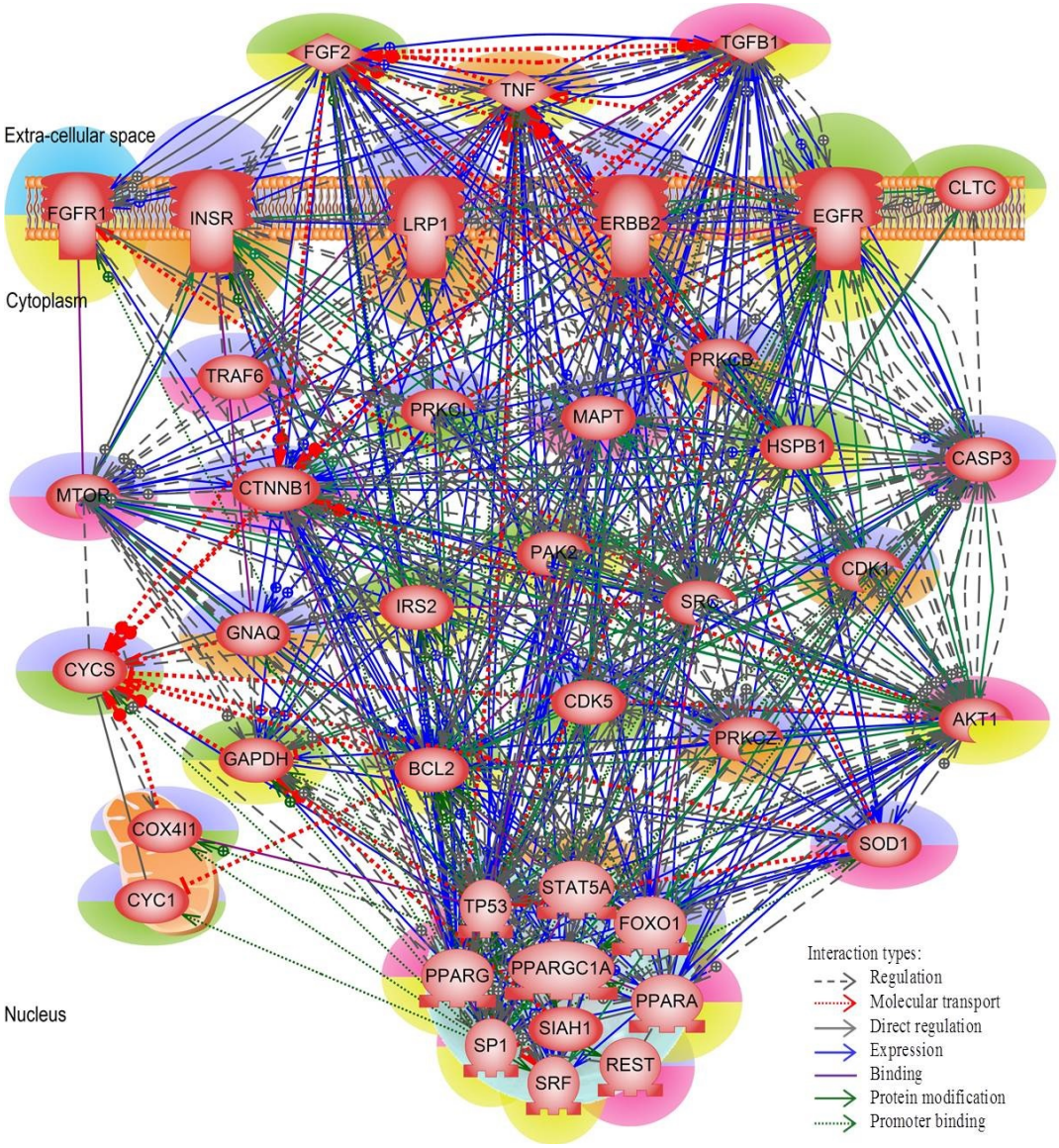


Figure 5.4. ***Integrated*** **Huntington's disease mechanism.** The 41 genes/proteins are found in common in all 10 enriched KEGG pathways. Genes/proteins implicated in HD pathology are highlighted in green/red and the genes/proteins of potential interest are highlighted in blue/orange (See Table 5.1. for details).

neurological disorders including Huntington's disease. In general, cytokines are required for normal functioning of cells. However, the formation of protein aggregates inside the cell triggers inflammatory mechanism which leads to increased cytokine activities thereby causing chronic cell stress (Battaglia et al., 2010; Möller, 2010). A delicate balance has to be maintained in order to sustain homeostasis within cell. On examining the *integrated* network, we propose that the hemostasis in Huntington's disease environment could be restored by regulating the three extra-cellular ligands FGF2, TGFB1 and TNF thereby controlling their downstream signaling cascades of various target genes expression. On the other hand, even though the *integrated* HD mechanism network size was relatively small, due to the high interconnectedness of all the nodes it was difficult to suggest a pathway or two to implement our proposed neuronal restoration in Huntington's disease regulated via the three ligands.

5.4.2 Existing treatment options

Once again we went back to HD research literature looking for some molecular mechanisms and/or therapeutic pathways that are currently being utilized in this field. A recent review article by Zuccato et al. in 2010 described the past achievements, the current status along with suspected disease mechanisms and therapeutic measures available in Huntington's disease realm. Several research works suggest few important players in HD whose regulation could promote neurogenesis. Under normal physiological conditions, HTT interacts with many genes/proteins including BDNF, MTOR and REST to promote the survival of striatal neurons, the ones that are subjected to cell death in Huntington's disease. Interaction between BDNF (brain-derived neurotrophic factor) and HTT is important for the survival of striatal neurons as well as promoting synapse formation. In addition, HTT bind and

sequester mechanistic target of rapamycin (MTOR) inside the cytoplasm inhibiting MTOR's downstream regulation. In general, MTORs are negative regulators of autophagy. Autophagy is an essential, homeostatic process by which cells break down their own components. They are the debris clearance machineries in the cell that is required to protect against infections, autoimmune and inflammatory diseases (Levine et al., 2011). Likewise, REST (RE1-silencing transcription factor) and HTT interaction is also important in HD pathogenesis. HTT binds with REST to maintain low levels of REST gene expression inside the cytoplasm thereby not affecting the transcription of BDNF gene.

In Huntington's disease, mutation in HTT causes protein aggregation formations which were not properly cleared from the cell thus disrupting the normal functioning of the stratial neurons. Due to transcription suppression by REST, BDNF level was found reduced in neurodegenerative diseases including Alzheimer's, Parkinson's and Huntington's disease (Zuccato and Cattaneo, 2007; Zuccato et al., 2001). Mutant HTTs were found inducing neuronal death via distinct but complementary pathways including deregulation of apoptosis and/or autophagy, altered transcription, metabolism and cellular stress responses. Currently one of the therapeutic measures suggested in Huntington's disease domain is by clearing the HTT protein aggregates from the cell through the induction of autophagy by the MTOR inhibitor rapamycin. Another indicated treatment is through increasing the beneficial BDNF gene expression (Zuccato et al., 2010). Animal HD model studies have shown that the use of rapamycin (MTOR inhibitor) improved striatal neuron survival and motor performance. However, due to deleterious side effects of rapamycin, it was not recommended for use as an exclusive drug in HD treatment. A combinatorial strategy with rapamycin or other drugs promoting autophagy has been suggested as relevant treatment for HD and other related diseases (Ravikumar et al., 2004).

Following these HD literature suggested treatment ideas, we modified our integrated Huntington's disease mechanism network to include only those nodes (13 genes: AKT1, BCL2, GAPDH, EGFR, FGF2, FGFR1, INSR, MTOR, PPARGC1A, REST, SP1, TGFB1 and TNF) that might play a critical role in both inhibiting MTOR and improving BDNF gene expression. BDNF was added to the reduced network, as was done with HTT. The resulted *enhanced* integrated HD mechanism network is shown in Figure 5.5.

5.4.3 Proposed treatment options

From this *enhanced* network we propose two pathways through which homeostasis in HD could be restored by initiating the downstream signaling cascade of various target genes expression via primarily through TGFB1, one of the three extra-cellular ligands. Our first proposal includes a two-step process of MTOR inhibition. Step 1: TGFB1 activates PPARGC1A gene expression in the nucleus, which in turn increases GAPDH gene/protein activity. Step 2: Up-regulation of GAPDH inhibits MTOR gene expression activity. Once MTOR is inhibited, autophagy mechanism could be boosted up in the cell. As soon as autophagy process is reestablished, HTT protein aggregates will be effectively cleared from the cell thus leading to neuronal survival.

Our second restoration pathway recommendation is via both FGF2 and TGFB1 ligand activation of EGFR receptors thereby initiating several downstream target genes expression. Among those upregulated genes, AKT1 (v-akt murine thymoma viral oncogene homolog 1) is a vital downstream target for EGFR and has been shown to be a critical mediator of neuronal survival. AKT1 has been suggested to be a promising therapeutic target to promote cell survival (Dudek, 1997). Apart from

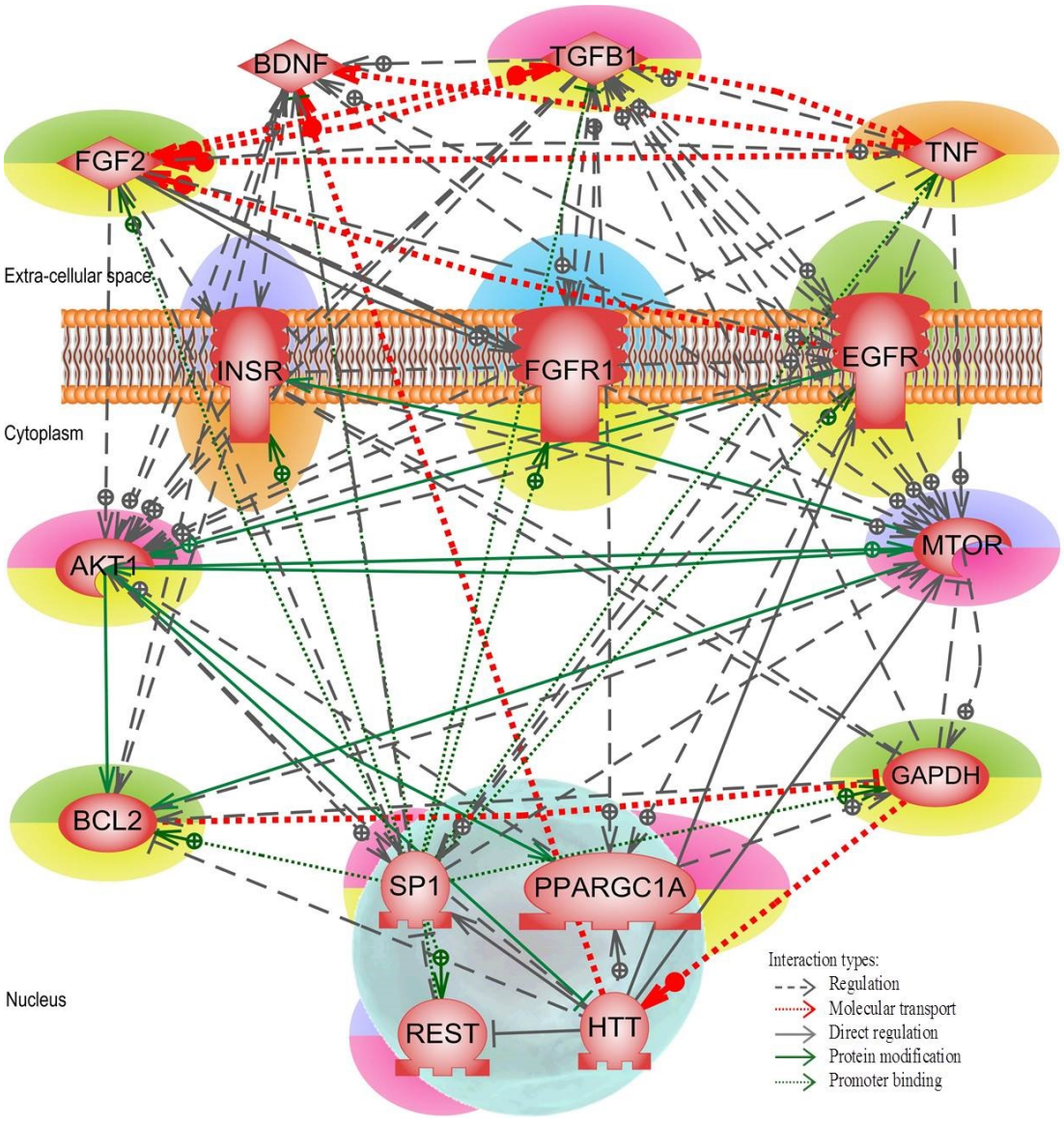


Figure 5.5. ***Enhanced*** **integrated Huntington's disease mechanism.**The 15 genes/proteins that were suggested to play major role in HD treatment. Genes/proteins implicated in HD pathology are highlighted in green/red and the genes/proteins of potential interest are highlighted in blue/orange (See Table 5.1. for details).

AKT1, EGFR interacts with SP1 (Sp1 transcription factor) which is involved in many cellular processes, including cell differentiation, cell growth, apoptosis, immune responses, response to DNA damage, and chromatin remodeling. SP1 fine-tunes the transcription of many genes including BCL2 and REST. BDNF transcription could be increased via maintaining tight regulation between REST and SP1. In addition, SP1 could also up-regulate BCL2 gene expression, promoting anti-apoptosis. Thus, eventually striatal neurons could be protected by promoting BDNF activity, as well as by reducing the apoptotic process in the cell. Additionally, EGFR also promote GAPDH gene expression eventually aiding in neuronal survival as detailed in our earlier pathway proposal.

From network analysis stand-point, the two proposed homeostasis restoration pathways show promising measures towards treatment plans in Huntington's disease. As a first step towards translating our proposed therapeutic networks into real world applications, such complex multi-player interconnected pathways could be evaluated using advanced dynamic modeling tools such as cellular automata (Ermentrout and Edelstein-Keshet, 1993).

5.5 Huntington's disease *microRNA* regulatory network

MicroRNAs perform important role in delivering post-transcriptional regulation of gene expression. In order to identify the microRNAs and their potential targets in Huntington's disease domain, we constructed a microRNA regulatory network (MRN) using the 514 "seed genes".

Before to proceed with the MRN construction, we first identified the microRNAs that could target our seed genes. This was accomplished using the *shortest-path* network option in Pathway Studio software where we subjected all the 514 HD "seed

genes" to only microRNA interactions type. We found 132 microRNAs to target our HD genes. In order to obtain the microRNA-target gene interactions, we constructed a *direct* interaction network using the 132 microRNAs and the 514 "seed genes". [Note: MicroRNA regulatory network figure is not included in this text due to its bigger size.]

5.5.1 Critical microRNA regulation in HD

The average node degree of the microRNA regulatory network was 4.3. Being the node with highest degree in the network, miR-9 was observed to target 35 genes (!). In addition, miR-9 was the node with highest closeness and betweeness centrality scores. Finding it in our microRNA regulatory network was exciting for a couple of reasons. First, miR-9 was previously known in Huntington's disease mechanism, as well as found to target REST (RE1-silencing transcription factor), one of the important players of HD pathogenesis. Secondly, miR-9 regulation has already been identified and found reduced in both Alzheimer's and Huntington's disease brains (Lukiw, 2007; Packer et al., 2008; Zuccato and Cattaneo, 2007; Zuccato et al., 2001), providing thus another evidence for the conjectured unified underlined mechanism of the neurodegenerative diseases. Apart from miR-9 regulation, the network included miR-132 and miR-29a/b1 miRNAs, both already associated with HD pathogenesis (Junn and Mouradian, 2012; Lau and de Strooper, 2010).

The next top five microRNAs found in the regulatory network were miR-124, miR-135a, miR-141, miR-182 and miR-19a. All these microRNAs were among the top 25 nodes with highest degree (node degree ≥ 12), and top 25 nodes with highest closeness and betweeness centrality scores. miR-124 is one of the most abundantly expressed miRNAs in the nervous system, being widely expressed in neurons

in the brain, retina, and spinal cord. It has been implicated in the modulation of neurite outgrowth, as well as cytoskeleton formation (Yu et al., 2008). There have been no indications so far for involvement in neurodegenerative processes of miR-135a, known to target genes involved in blood pressure regulation (Sõber et al., 2010). Similarly, miR-141 and miR-182 were known to be involved only in DNA methylation and cancer metastasis, respectively (Segura et al., 2009; Vrba et al., 2010). miR-19 and miR-21 have been found to target PTEN, a gene/protein found localized in the neurofibrillary tangles (NFTs) and senile plaques in Alzheimer's disease brains (Pezzolesi et al., 2008; Sonoda et al., 2010). MicroRNA regulatory network also uncovered that many of these top regulating microRNAs modulate several known-HD genes such as CNR1, FOXP1, GAPDH, and IRS2. Reiterating the ResNet 9.0 database's current status of microRNA-target interaction identifications, further experimental verification is recommended.

5.5.2 Highly targeted genes of interest in HD

Table 5.5. Genes of interest determined from Huntington's disease microRNA regulatory network

Genes of Interest	Targeted by no. of miRNAs
DOCK7, ZBTB10	7
DPYSL5, ZNF148	6
DCLK1, OSBPL11, RCAN2, ZFP36L1	5
CNTNAP1	4
FGFR1, FKBP5	2
CX3CL1, FDFT1	1

Table 5.5 shows the genes of interest in the HD microRNA regulatory network and how many microRNAs are targeting each gene. In this network, DOCK7 (dedicator of cytokinesis 7) was the gene with the highest number of microRNA regulations,

being regulated by seven members of miR-181 and miR-30 families. DOCK7 gene encodes for guanine nucleotide exchange factor (GEF) protein that plays a major role in axon formation and neuronal polarization. In general, GEFs are critical mediators of Rho GTPase activation by stimulating the exchange of GDP for GTP. Under normal physiological conditions, Rho GTPases act as molecular switches in intracellular signaling pathways and have many downstream targets. Mutations in GEFs and deregulated Rho GTPase signaling have been implicated in ALS, a debilitating motor neuron disease caused by neuronal degeneration. Based on its molecular function and its association with similar neurodegeneration disease, DOCK7 could be of potential interest in Huntington's disease mechanism as well.

Moving on with other genes of interest in the microRNA regulatory network, zinc finger proteins (ZBTB10, ZNF148, and ZFP36L1) were highly targeted by multiple microRNAs including miR-20a and miR-29b1, known microRNAs in Alzheimer's and Huntington's disease, respectively (Sonntag, 2010). As reported in the previous section, zinc finger proteins are demonstrated to be a promising new gene therapy tool for Huntington's disease. Such a therapy could be enhanced by these microRNA regulations.

Additional to microRNA regulation, the network also included 43 genes that code for transcription factors. These significantly differentially expressed TFs indicate a possible integrated gene expression regulation mechanism in Huntington's disease. Like we noticed for the other two neurodegenerative disorders, there is a likelihood of dual-level gene expression regulation also occurring in HD paradigm. Thus, Huntington's disease is becoming another complex disease system that involves highly interconnected molecular players and multi-level regulation.

5.6 Summary

As in our analysis on Alzheimer's disease some of the well-known Huntington disease genes like HTT, HDL3, JPH3 and PRNP have not been significantly expressed in the post-mortem microarrays, and were not included in our "seed gene" list. While this may be an unavoidable consequence of the overall decrease in the level of expression in the post-mortem probes, one of the major advantages of our network analysis is to offer additional ways for identifying such important genes as critical nodes in network topology. Thus, the PRNP gene emerged in our shortest path network as one of the many *connecting* genes/proteins, and in the *compact* version of this network as one of the top 15 critical nodes with highest visibility (closeness centrality) and most influence on the interaction traffic (betweenness centrality).

From our network analysis, we identified seven novel genes of importance for the Huntington disease: CNTNAP1, CX3CL1, FGFR1, PPARA, PRDX2, RCAN2 and ZNF148 based on their *direct* interaction with many already known-HD genes, as well as on their intrinsic molecular functions mostly in neuroprotective roles. All these seven genes were among the 514 "seed genes" that were significantly differentially expressed in HD postmortem samples. Next, via their "guilt-by associations" to many known-HD genes from *shortest-path* network, we concluded that 15 *connecting* genes/proteins added by the Pathway Studio software, such as CDK1, CEBPA, E2F1, EGR1, ERBB2, FOXO1, HSP90AA1, IL2, INSR, LRP1, PRKCB, PRKCZ, SRF, STAT5A and TNF, could also be of potential interest in Huntington's disease domain. These genes were first-level partners with *three* or more previously implicated HD genes. Especially, EGR1 was found interacting directly with thirteen (!) known-HD genes in the *compact* shortest-path network. One may also consider including to the list of potential Huntington's genes also those, which interact in that network

with two or one known HD genes, such as ATXN7, DCLK1, DPYSL2, GNAQ, PRDX2, PRDX6, THRA and ZFP36L1.

Besides their "guilt-by-associations" with many previously implicated Huntington's genes, the *connecting* genes/proteins CDK1, CEBPA, E2F1, EGR1 and INSR were among the top 25 nodes with highest degree ($\geq$ 12), highest visibility and highest traffic influence in the *compact* shortest-path network, as assessed by their closeness and betweenness centrality scores. Identifying these critical network measures may be considered as an evidence for a potentially important systems biological role of these five genes/proteins in the integral mechanism of Huntington's disease.

DAVID gene set enrichment analysis revealed that many of these candidate genes were involved in neuron and glial cell development, projection and differentiation process as well as regulation of synaptic transmission and response to oxidative stress mechanism. Deregulation of these biological processes was attributed to the dysfunction of many of the known-HD genes in addition to HTT gene mutation. Due to their critical role in the biological pathways that were significantly affected in Huntington's disease, as well as being directly associated with many known-HD genes, increases the probability that these proposed candidate genes could play a major part in the HD pathogenesis.

Through our microRNA regulatory network, we suggest that microRNAs could be potential regulators and drug targets in HD. The following twelve microRNAs namely, miR-101-1, miR-124-1, miR-128-1, miR-135A1, miR-141, miR-153-1, miR-15A, miR-16-1, miR-182, miR-19A, miR-27A and miR-96 could be of potential interest in HD. These microRNAs were regulating many known and genes of interest in HD. Like Parkinson and Alzheimer diseases, there is a possibility of dual-level gene regulation by both microRNAs and transcription factors in Huntington's dis-

ease mechanism as well. Like Parkinson and Alzheimer diseases, there is a possibility of dual-level gene regulation by both microRNAs and transcription factors in Huntington's disease mechanism as well.

From the *integrated* network assessment, we propose a couple of potential treatment plans for Huntington's disease initiated via the extra-cellular ligands TGFB, FGF2 and TNF. Restoring the homeostasis in Huntington's disease seems possible; one such plan is to up-regulate the innate autophagy process by inhibiting MTOR activity within the cell. Another, treatment plan is to promote striatal neuron survival via increasing BDNF gene expression.

Chapter 6

Towards unified underlying mechanisms of NDDs

6.1 Introduction

There is growing evidence that deregulation of several biological processes such as oxidative stress, inflammation, mitochondrial dysfunction, free radical formation, ubiquitin-proteosomal system and others contribute to Neurodegenerative disorders (NDDs) including Parkinson's (PD), Alzheimer's (AD), Huntington's (HD), Amyotrophic lateral sclerosis (ALS), Progressive supranuclear palsy (PSP) etc. Genetic and environmental factors, protein misfoldings and abnormal protein aggregations were some of the common hallmark characteristics of such diseases. Over many years, affected individuals display cognitive, motor and emotional disturbances with increasing disability ultimately leading to death (Schon and Manfredi, 2003; Jellinger, 2009; Akiyama et al., 2000; Chiti and Dobson, 2006; Möller, 2010; Di Monte et al., 2002; Gil and Rego, 2008; Mattson, 2000; Zuccato et al., 2010; Emerit et al., 2004; Liu et al., 2011; Jellinger, 2010).

With this background, using network biology tools we wanted to identify the molecular players and underlying mechanisms in three universal and debilitating neurodegenerative disorders namely, Parkinson's, Alzheimer's and Huntington's diseases. In earlier chapters, we identified and explained such critical molecular players, mechanisms and integrated approaches for possible disease initiators as well as potential therapeutic techniques for each of these diseases. Following our modest success, we further conducted similar network-based analysis to uncover crucial molecular partakers and biological processes that are common in all three disease conditions. Based on this we propose a unified underlying mechanism which includes several possible disease initiating as well as potential treatment scenarios common for Parkinson's, Alzheimer's and Huntington's neurodegenerative disorders.

By overlapping the significantly differentially expressed genes from the PD, AD and HD microarray datasets, 22 "seed genes" (ADAM23, AHNAK, ATP2B2, ATP6V0E1, BCL6, CALD1, CAPRIN1, DCLK1, GLT1D1, ITSN1, MCL1, MSI2, NTRK3, PLCB1, PREPL, RASGRP1, REPS2, SCAMP1, SSBP3, STXBP1, SYT1 and ZMAT3) were found in common in all disease conditions. Ten (ATP2B2, ITSN1, MCL1, NTRK3, PLCB1, RASGRP1, REPS2, SCAMP1, STXBP1 and SYT1) out of 22 overlapping genes were previously known to be related in one or the other neurodegnerative disorders under study. Among the remaining twelve overlapping genes, four genes AHNAK, BCL6, CALD1 and DCLK1 could be of potential interest in NDDs. More information about these novel candidate genes are given in the following sections. Utilizing the 22 seed genes, various types of biological networks were constructed using Pathway Studio software.

6.2 *Shortest-path* network of neurodegeneration mechanism based on seed genes found in common in all three NDDs

In general, *shortest-path* (SP) network are used to uncover indirect relationships between any pair of genes. By doing so, many intermediary genes (connecting genes) were revealed and their intrinsic molecular functions as well as their possible associations with neurodegeneration processes were identified. In this manner, the *shortest-path* network of 22 "seed genes" grew bigger by incorporating 30 more software-added genes. To construct the SP network, protein modification, promoter binding and direct regulation types of interactions were used. A *direct* interaction network with the 22 seed genes was constructed but it only revealed interactions between two pairs of seed genes. Figure 6.1 illustrates the *shortest-path* network of neurodegeneration mechanism.

Both the seed and the connecting genes were examined in OMIM, NCBI's PubMed and databases and Google website to confirm whether these genes were already or not implicated in any of the NDDs pathogenesis. MalaCards, the human maladies and their annotations database (available at `http://www.malacards.org/`) was also cross-referenced for such previous disease associations (Rappaport et al., 2013). If not, the intrinsic molecular functions of the genes were investigated to categorize them as genes of potential interest in NDDs mechanisms. The genes are highlighted in *red* or *green* if they were previously implicated in the neurodegenerative disease mechanism. If the genes are of potential interest in NDDs then they are highlighted in *orange* or *blue* in the network figures shown in this chapter.

The 22 "seed genes" used to identify and explore the common unified underlying

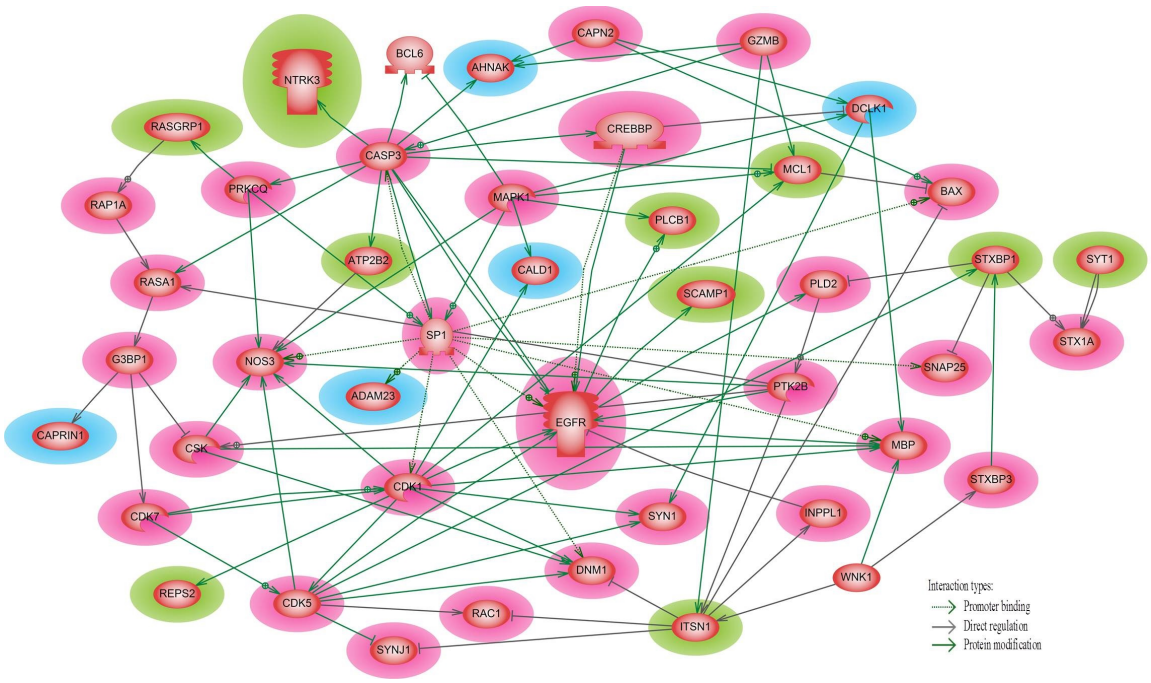


Figure 6.1. ***Shortest path*** **network of neurodegeneration mechanism.** The genes/proteins implicated in NDDs are highlighted in green and red. The genes/proteins of potential interest are highlighted in blue.

mechanisms in NDDs were chosen purely based on the overlapping and yet significantly differentially expressed genes in all three disease conditions. This set of "seed genes" included few already known NDDs genes like ATP2B2, ITSN1, MCL1, NTRK3 and SYT1 while some well-known NDDs genes like APP, HTT, PSEN1 and SNCA etc., were among the non-completely overlapping ones. However, the SP and other networks (explained below) incorporated many noteworthy genes like BNDF, CREBBP, CDK5, EGFR, MBP, NGF, PTEN and SP1 etc. as *connecting* genes/proteins. Emergence of such important nodes as *connectors* in a network primarily based on the inter-relationships is excellent example for network-based systems biology research work. Network-based approaches could be vital in studying a system as a whole even if it is missing some critical components. It can be used to recreate some of the missing pieces from the components at hand. For example, from the 22 "seed

genes" with limited associations to neurodegeneration process emerged a *shortest-path* network with 45 genes of which 28 genes were previously implicated in NDDs (see Figure 6.1).

6.2.1 Molecular functions and network attributes of some of the critical conecting genes

In the next few lines, we explain the molecular functions of some of the known NDDs genes, including those that reemerged in our shortest-path network, and their suspected role in neurodegeneration process. Abnormal *CDK5* activity has been suspected to contribute to the aberrant accumulation of hyperphosphorylated tau proteins in AD brains (Mateo et al., 2009). In addition to functioning as 'housekeeping' transcription factor, when deregulated *SP1* has been shown to promote neurodegeneration process in multiple ways. It is a positive regulator of *APP* (amyloid-beta precursor protein) gene expression as well as of genes like *tau*, *BACE1* and *CASP3*. All these factors initiate the characteristic features of any NNDs such as the protein misfoldings, abnormal accumulations and eventual neuronal loss (Citron et al., 2008). Another hallmark trait in NDDs is the presence of neurofibrillary tangles (NFTs). *PTEN* is one of the main components of such tangles in the brain (Sonoda et al., 2010).

The shortest-path and other networks constructed for this study not only included many disease related genes, it also featured some of the genes that aid in neuronal survival mechanisms. For instance, epidermal growth factor receptor (*EGFR*) is a potential target to treat amyloid-beta accumulation (Wang et al., 2012). Nerve growth factors (*NGF*) promote neuroprotection as well as nerve repair process (Sofroniew et al., 2001). Another growth factor namely, brain-derived neurotrophic

factor (*BDNF*) was explored as promising gene therapy tool for Alzheimer's and Huntington's disease (Zuccato et al., 2001; Tapia-Arancibia et al., 2008). Along with having significant molecular functions in neurodegeneration and protection processes, many of these genes occupy critical positions in the *shortest-path* network. Genes like CASP3, CDK1, CDK5, EGFR and SP1 were the top five nodes with highest connectivity (degree ≥ 9), visibility and traffic-influential as measured by the closeness and betweenness centrality scores. These attributes increase the significance of the genes in a network. Higher connectivity means the node is more central and critical for the proper functioning of the network. With higher visibility, a node could be a better monitor of the information flow, and with higher betweenness centrality score it could influence the outcome of network processes. Thus, based on these vital network parameters and relevant molecular functions some of the genes in the *shortest-path* network could be potential important players common to the three NDDs under study.

6.3 *Combined* common regulator and common target network of neurodegeneration mechanism based on seed genes found in common in all three NDDs

Implied by the name, using the common regulator (CR) and common target (CT) network one can identify the regulators and targets found in common for a given set of genes. Finding such nodes in a network is important since they can be promising drug targets. Next to shortest-path network, the CR and CT network constructions are popular options offered by the Pathway Studio software. We constructed the CR and CT networks individually using the 22 "seed genes" found in common in

all three disease conditions. Protein modification, promoter binding and microRNA regulatory interaction types were used. Next we merged the two networks to build one *combined* common regulators/targets network (see Figure 6.2).

In this *combined* network, genes like BCL6, ITSN1, MCL1, NTRK3 and RASGRP1 were the top five common regulators with out-degree (outgoing connections from the node) ≥ 10 as well as genes like BCL6, CALD1, MCL1, PLCB1 and SYT1 were the top five common targets with in-degree (incoming connections to the node) ≥ 10. In addition to their high network connectivity, BCL6 and MCL1 were the top two nodes with highest closeness and betweenness centrality scores. These high quantitative scores make BCL6 and MCL1 critical molecular players in all three studied neurodegenerative disorders - Parkinson's, Alzheimer's and Huntington's disease.

MCL1 (myeloid cell leukemia sequence 1) is an anti-apoptotic protein of the BCL-2 family. In general, it enhances the survival of the cell by inhibiting apoptosis. Promoting BCL-2 function has been proposed to treat neurodegenerative as well as neurological disorders like stroke. The neuroprotective functions of BCL-2 are mediated by activation of ERK/PI3K and CREB pathways. Valproate, the anticonvulsant and mood stabilizer medication has been explored as a potential treatment technique for NDDs where it was suspected to up-regulate BCL-2 (B-cell CLL/lymphoma 6) and promote BCL-2-related neuronal protective and neurotrophic effects (Creson et al., 2009; Shacka and Roth, 2005). *BCL6* is a transcription repressor which is frequently translocated and hypermutated in diffuse large-cell lymphoma (DLCL), and may contribute to the DLCL pathogenesis (Hans et al., 2004). However, the BCL6 association with neurodegenerative disorders is unclear. Based on its multiple significant network attributes it would be noteworthy to explore in detail the BCL6 role in NDDs.

The *shortest-path* and the *combined* CR and CT networks revealed novel genes

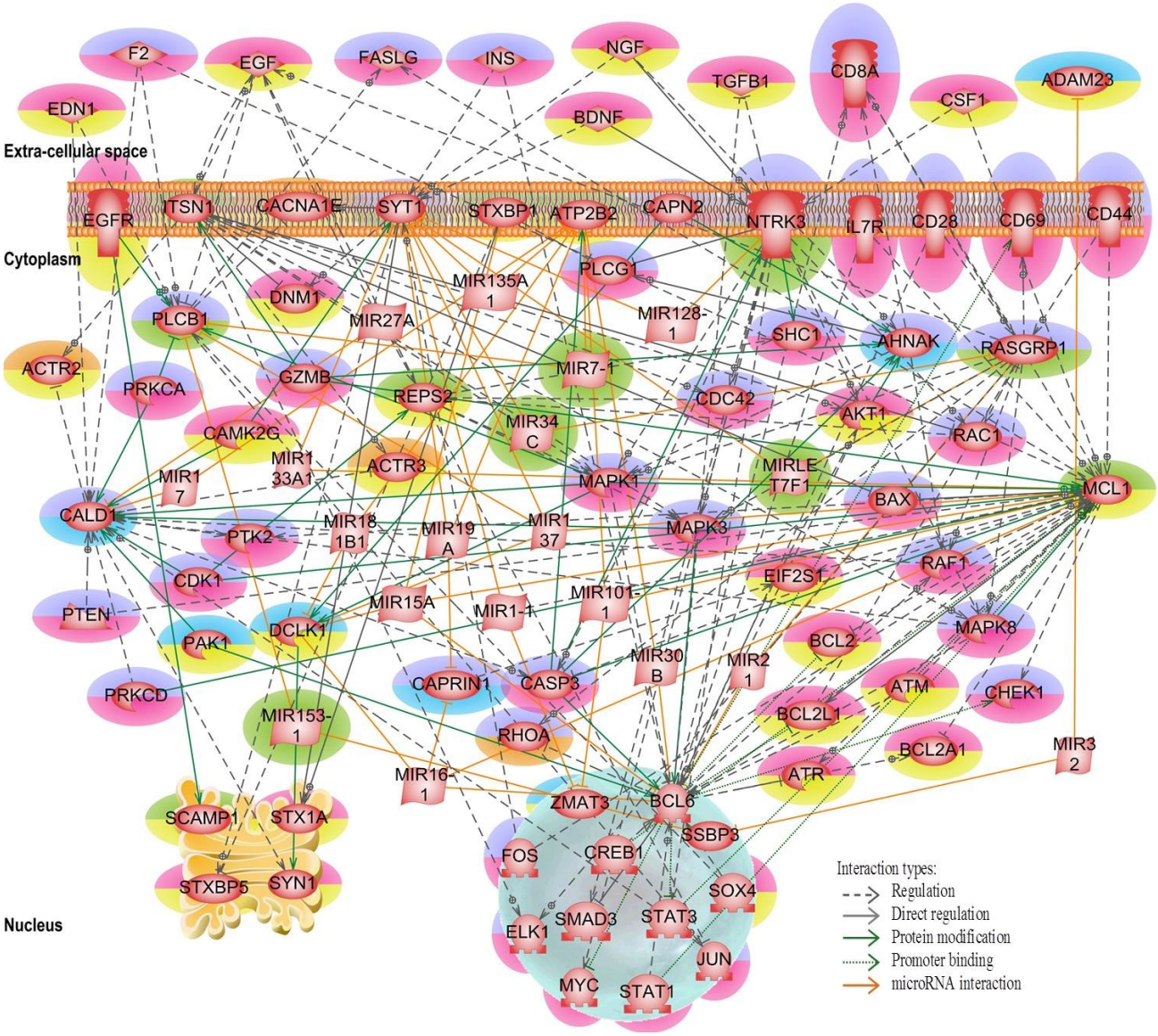


Figure 6.2. ***Combined*** **common regulator and common target network of neurodegeneration mechanism.** The genes/proteins implicated in NDDs are highlighted in green and red. The genes/proteins of potential interest are highlighted in blue and orange. Genes/proteins contributing to the disease pathogenesis are highlighted in purple and those in neuronal survival are in yellow.

of potential interest in neurodegenerative disorders based on their innate molecular functions and their importance for network topology. A more general analysis of the role of these candidate genes will be discussed in the next section within the framework of our modest attempt to explain the unified underlying molecular mechanisms in NDDs

6.4 *Integrated* neurodegeneration mechanism network based on the seed genes found in common in all NDDs.

The *integrated* neurodegeneration mechanism network was designed as an amalgamation of shortest-path, common regulators and common targets networks of 22 "seed genes" found in common in all Parkinson's, Alzheimer's and Huntington's disease. By combining the connecting genes/proteins from the *shortest-path* network, the 22 "seed genes" grew bigger into a 46 node network (Figure 6.1) which further increased in size as a 114 node *integrated* network (Figure 6.3) by incorporating the *common regulators* and *common targets* (Figure 6.2).

In the *integrated* network, MCL1 and BCL6 continue to be significant unifying players in neurodegeneration mechanism as one of the top five nodes with highest connectivity, closeness and betweenness centrality scores. Including these two genes there were 13 nodes higher connectivity node with degree > 10. In biological networks, such hubs are important since these nodes tend to be critical and conserved. In the constructed 114-gene network more than 60% of the genes like ATM, ATR, BDNF, CREB1, CREBP, EGFR, MAPK1, NGF, PTEN, SNAP25, SYN1 and TGFB1 etc. were previously implicated in NDDs either as disease contributing or as neuroprotec-

tive agents. The intrinsic molecular functions of all 114 genes were cross-referenced in OMIM, MalaCards, PubMed databases aided with Google search for their possible associations in neurodegenerative disorders. Based on this, the genes were classified either as disease aggravators (highlighted in *purple*) or as helpers in neuronal survival process (highlighted in *yellow*) (see Figure 6.3).

6.4.1 Guilt-by-association analysis

As mentioned earlier, the *integrated* network included some potential genes of interest in NDDs. The criteria used to select such genes are as follows. As determined by being neighbors of the known NDDs genes, *CALD1* has the highest number of 15 (!) nearest neighbors in the *integrated* neurodegeneration mechanism network (see Figure 6.3) all of which known to be involved in one or another NDD under study (CAMK2G, CASP3, CDK1, EDN1, F2, MAPK1, MAPK3, MAPK8, PRKCA, PRKCD, PTEN, PTK2, RAF1, STAT3 and SYN1). The guilt-by-association property makes CALD1 prime NDD candidate gene. CALD1 (caldesmon 1) gene encodes calmodulin- and actin-binding protein. Calcium together with calmodulin increases the activity of calmodulin binding proteins (CaBPs). One such CaBP is BACE1, the enzyme which is essential for the generation of beta-amyloid (Cole and Vassar, 2007). Many research studies point out that the increase in calcium and calmodulin leads to the abnormal protein accumulations in the brain, the key feature of NDDs. Calmodulin antagonists are explored as potential drug targets in Alzheimer's disease (O'Day and Myre, 2004). All this reinforces the prognosis made for CALD1 as an important NDD-related gene.

Our second choice for candidate gene is *AHNAK* (nucleoprotein). It neighbors with nine already known-NDDs genes (AKT1, CAPN2, CASP3, CD8A, GZMB, MAPK1,

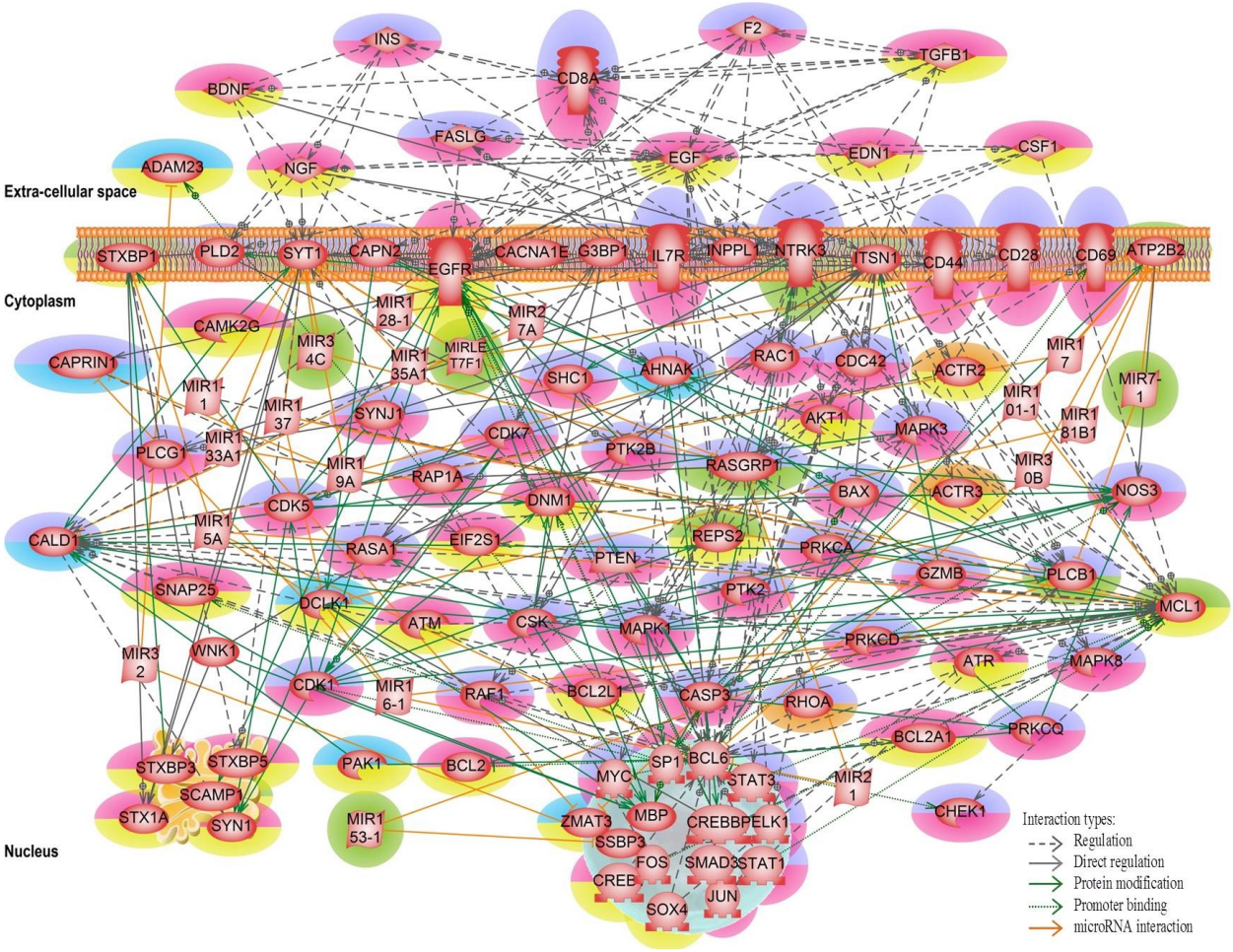


Figure 6.3. ***Integrated*** **network of neurodegeneration mechanism.** The genes/proteins implicated in NDDs are highlighted in green and red. The genes/proteins of potential interest are highlighted in blue and orange. Genes/proteins contributing to the disease pathogenesis are highlighted in purple and those in neuronal survival are in yellow.

NGF, PLCG1 and RAF1). S100B (S100 calcium binding protein B) is suspected to play a role in neurodegeneration by recruiting pro-inflammatory molecules which ultimately cause neuronal apoptosis (Rothermundt et al., 2003). AHNAK is a major and most specific calcium-dependent S100B target protein. The S100B/AHNAK interaction is supposed to participate in the S100B-mediated regulation of cellular calcium homeostasis (Salim et al., 2009). As mentioned earlier, disruption of calcium homeostasis is one of the crucial steps in neurodegeneration process. Thus, AHNAK could likely be involved in such deregulation activity promoting neuronal loss.

Another recommendation for candidate gene is *DCLK1* which directly interacts with five known NDD-genes (CAPN2, CREBBP, MAPK1, MBP and SYN1). DCLK1 (doublecortin-like kinase 1) is highly expressed in the hippocampus and in the cortices. Le Hellard et al. in 2009 have shown that up-regulation of DCLK1 by BDNF (brain-derived neurotrophic factor) could be associated with increased synaptic formation and general cognitive abilities. DCLK1 could be pursued as a potential drug target to enhance the memory formation and cognition.

Intrinsic molecular functions relevant to neurodegeneration process, as well as their associations with many previously known NDDs-genes, increases the chances of ANHAK, CALD1 and DCLK1 to be potential genes of interest in Parkinson's, Alzheimer's and Huntington's disease.

6.4.2 MicroRNA regulation in NDDs

The *integrated* network also incorporated regulatory interactions of 19 microRNAs. It included four miRNAs namely, miR-153, miR-34c, miR-7 and let-7, previously implicated in NDDs. They target APP and SNCA, the two well-known promoters of

neurodegeneration in Alzheimer's and Parkinson's disease (Miñones Moyano et al., 2011; Junn and Mouradian, 2012). Apart from these four known miRNAs ten others namely, miR-128-1, miR-135A1, miR-137, miR-15A, miR-16-1, miR-17, miR-19A, miR-27A, miR-30B and miR-32, were repeatedly identified in *this integrated* network as well as in the microRNA regulatory networks that we constructed for each disease conditions individually (see sections 3.5, 4.4 and 5.5). Based on this recurring pattern, these ten microRNAs could have potential role in NDDs, which requires further validation. Besides microRNAs, regulation by transcription factors (TFs) was also noticed in the *integrated* network. Twelve TFs regulates the expression of many genes in this network. Except BCL6 all other TFs CREB1, CREBBP, ELK1, FOS, JUN, MYC, SMAD3, STAT1, STAT3, SP1 and SOX4 were previously known in NDDs. The transcription factor regulation is an integral part of NDDs mechanisms which we also observed in the individual PD, AD and HD regulatory network-based analysis. Again, the possibility of dual-level TF/miRNA regulation in each of these neurodegenerative disorders is reinforced here.

6.4.3 DAVID enrichment analysis

Apart from identifying individual molecular players in NDDs, DAVID analysis was utilized to evaluate the biological processes and pathways enriched in the *integrated* network 114-genes set. Table 6.1 shows the identified statistically significantly enriched (Benjamini-Hochberg multiple correction) Gene Ontology (GO) categories and pathways related to brain and nervous system.

Biological processes like aging, memory or learning, neuron development and differentiation, synaptic vesicle trafficking and endocytosis and vasculature development etc. were some of the highly affected common mechanisms in the neu-

Table 6.1. DAVID enrichment analysis of the 114-genes network of the *integrated* neurodegeneration mechanism.

Category	Term	Gene count	Fold Enrichment	Benjamini
GOTERM_BP_FAT	GO:0030182 neuron differentiation	20	6.71	4.32E-09
GOTERM_BP_FAT	GO:0031175 neuron projection development	15	8.62	5.84E-08
GOTERM_BP_FAT	GO:0048666 neuron development	16	6.94	2.24E-07
GOTERM_BP_FAT	GO:0051960 regulation of nervous system development	12	9.19	1.60E-06
GOTERM_BP_FAT	GO:0060627 regulation of vesicle-mediated transport	9	13.79	5.52E-06
GOTERM_BP_FAT	GO:0001666 response to hypoxia	10	10.97	6.02E-06
GOTERM_BP_FAT	GO:0070482 response to oxygen levels	10	10.43	8.90E-06
GOTERM_BP_FAT	GO:0007568 aging	9	12.03	1.42E-05
GOTERM_BP_FAT	GO:0007611 learning or memory	9	11.92	1.51E-05
GOTERM_BP_FAT	GO:0046649 lymphocyte activation	11	8.13	1.66E-05
GOTERM_BP_FAT	GO:0007264 small GTPase mediated signal transduction	13	6.27	1.69E-05
GOTERM_BP_FAT	GO:0048667 cell morphogenesis involved in neuron differentiation	11	7.74	2.37E-05
GOTERM_BP_FAT	GO:0048812 neuron projection morphogenesis	11	7.59	2.72E-05
GOTERM_BP_FAT	GO:0042110 T cell activation	9	10.50	3.39E-05
GOTERM_BP_FAT	GO:0048489 synaptic vesicle transport	6	26.74	4.17E-05
GOTERM_BP_FAT	GO:0017157 regulation of exocytosis	6	25.95	4.79E-05
GOTERM_BP_FAT	GO:0007005 mitochondrion organization	9	9.59	6.34E-05
GOTERM_BP_FAT	GO:0007409 axonogenesis	10	7.62	8.75E-05
GOTERM_BP_FAT	GO:0001944 vasculature development	11	6.44	1.01E-04
GOTERM_BP_FAT	GO:0050767 regulation of neurogenesis	9	7.97	2.03E-04
GOTERM_BP_FAT	GO:0050804 regulation of synaptic transmission	8	8.65	4.15E-04
GOTERM_BP_FAT	GO:0051924 regulation of calcium ion transport	6	12.79	9.83E-04
GOTERM_BP_FAT	GO:0006979 response to oxidative stress	8	7.17	0.001
GOTERM_BP_FAT	GO:0019226 transmission of nerve impulse	11	4.62	0.001
GOTERM_BP_FAT	GO:0006887 exocytosis	7	8.95	0.001
GOTERM_BP_FAT	GO:0007268 synaptic transmission	10	4.93	0.002
GOTERM_BP_FAT	GO:0055074 calcium ion homeostasis	8	6.26	0.002
GOTERM_BP_FAT	GO:0050877 neurological system process	20	2.43	0.003
GOTERM_BP_FAT	GO:0016192 vesicle-mediated transport	13	3.32	0.004
GOTERM_BP_FAT	GO:0006897 endocytosis	8	5.35	0.005

rodegenerative disorders. The deregulation of such biological processes was also revealed in the each of the individual disease enrichment analysis. In addition, enriched *KEGG* pathways also resulted from DAVID analysis. Table 6.2 presents all the 31 significantly enriched (with Benjamini-Hochberg multiple correction) KEGG pathways in NDDs.

The enriched KEGG pathways belonged to signal transduction, cell communication, nervous system, immune system, endocrine system, cell motility, development, endocrine and metabolic diseases, neurodegenerative diseases and carbohydrate metabolism categories. It is remarkable to notice that other progressive and chronic neurodegenerative disorders like amyotrophic lateral sclerosis (ALS) and Prion diseases, although not directly included in our three NDDs project, were also significantly enriched. Prion disease is characterized by the abnormal prion protein misfoldings in the brain and ALS is caused by the degeneration and death of motor neurons. Similarly, it was valuable to detect Type I and Type II diabetes mellitus to be enriched in the NDDs gene set. Impaired glucose metabolism has been associated with the abnormal protein aggregate formation and vice versa. In addition, cognitive dysfunction and the incidence of dementia were related with impaired glucose metabolism (Umegaki, 2012; Ristow, 2004).

As mentioned in the earlier chapters, many of the enriched KEGG pathways like long-term potentiation, axon guidance, calcium signaling pathway, gap and tight junction and neurotrophin signaling pathway etc. were also significantly affected in Parkinson's, Alzheimer's and Huntington's diseases when analyzed separately. Unlike the individual disease-based DAVID analysis, we observed in our PD/AD/HD *integrated* network analysis several more KEGG pathways like inositol phosphate metabolism, phosphatidylinositol signaling system, JAK-STAT signaling pathway, leukocyte transendothelial migration, long-term depression, GnRH signal-

Table 6.2. DAVID enrichment analysis of KEGG pathways in the 114-genes *integrated* neurodegeneration mechanism.

KEGG Pathways	Gene count	FE [a]	Benjamini	Genes
hsa04722:Neurotrophin signaling pathway	21	12.13	1.01E-14	CAMK2G, FASLG, RAF1, PRKCD, AKT1, NTRK3, CDC42, MAPK1, BDNF, PLCG1, JUN, BAX, BCL2, MAPK3, RAC1, RHOA, RAP1A, MAPK8, SHC1, CSK, NGF
hsa04012:ErbB signaling pathway	16	13.17	7.00E-12	PRKCA, EGFR, CAMK2G, RAF1, ELK1, AKT1, MAPK1, PTK2, PLCG1, JUN, MAPK3, SHC1, MAPK8, PAK1, EGF, MYC
hsa04010:MAPK signaling pathway	24	6.44	8.52E-12	PRKCA, EGFR, ELK1, FASLG, RAF1, TGFB1, AKT1, CDC42, MAPK1, FOS, CASP3, BDNF, JUN, RASGRP1, MAPK3, RAC1, RAP1A, MAPK8, CACNA1E, PAK1, EGF, MYC, RASA1, NGF
hsa04510:Focal adhesion	20	7.13	1.32E-10	PRKCA, EGFR, RAF1, ELK1, CAPN2, PTEN, AKT1, CDC42, MAPK1, PTK2, JUN, BCL2, MAPK3, RAC1, RHOA, RAP1A, MAPK8, SHC1, PAK1, EGF
hsa04912:GnRH signaling pathway	14	10.23	4.93E-09	PRKCA, EGFR, PLD2, CAMK2G, RAF1, ELK1, PRKCD, MAPK1, CDC42, PTK2B, JUN, MAPK3, MAPK8, PLCB1
hsa04660:T cell receptor signaling pathway	14	9.28	1.51E-08	CD8A, RAF1, AKT1, PRKCQ, FOS, MAPK1, CDC42, PLCG1, RASGRP1, JUN, MAPK3, RHOA, PAK1, CD28
hsa04062:Chemokine signaling pathway	17	6.51	2.33E-08	RAF1, STAT1, PRKCD, STAT3, AKT1, CDC42, MAPK1, PTK2, PTK2B, MAPK3, RAC1, RHOA, RAP1A, SHC1, PAK1, PLCB1, CSK
hsa04370:VEGF signaling pathway	10	9.55	3.64E-06	PRKCA, AKT1, CDC42, MAPK1, PTK2, PLCG1, MAPK3, RAC1, RAF1, NOS3
hsa04650:Natural killer cell mediated cytotoxicity	12	6.46	8.05E-06	PRKCA, MAPK1, CASP3, PLCG1, PTK2B, MAPK3, RAC1, RAF1, FASLG, GZMB, SHC1, PAK1
hsa04810:Regulation of actin cytoskeleton	13	4.33	1.15E-04	EGFR, RAF1, MAPK1, CDC42, PTK2, INS, MAPK3, RAC1, F2, RHOA, PAK1, EGF, CSK
hsa04720:Long-term potentiation	8	8.43	1.30E-04	PRKCA, MAPK1, CAMK2G, MAPK3, CREBBP, RAF1, RAP1A, PLCB1
hsa04115:p53 signaling pathway	8	8.43	1.30E-04	CDK1, CASP3, ZMAT3, BAX, CHEK1, ATR, PTEN, ATM
hsa04520:Adherens junction	8	7.44	2.71E-04	EGFR, CDC42, MAPK1, MAPK3, CREBBP, RAC1, RHOA, SMAD3
hsa04350:TGF-beta signaling pathway	8	6.59	5.47E-04	MAPK1, SP1, MAPK3, CREBBP, RHOA, SMAD3, MYC, TGFB1
hsa04310:Wnt signaling pathway	10	4.74	5.93E-04	PRKCA, CAMK2G, JUN, CREBBP, RAC1, RHOA, SMAD3, MAPK8, PLCB1, MYC
hsa04540:Gap junction	8	6.44	6.10E-04	PRKCA, EGFR, CDK1, MAPK1, MAPK3, RAF1, EGF, PLCB1
hsa04360:Axon guidance	9	5.00	9.58E-04	CDC42, MAPK1, PTK2, MAPK3, RAC1, RHOA, PAK1, CDK5, RASA1
hsa04620:Toll-like receptor signaling pathway	8	5.67	0.001	AKT1, MAPK1, FOS, JUN, MAPK3, RAC1, MAPK8, STAT1
hsa04930:Type II diabetes mellitus	6	9.14	0.001	MAPK1, INS, MAPK3, MAPK8, CACNA1E, PRKCD
hsa04662:B cell receptor signaling pathway	7	6.68	0.001	AKT1, MAPK1, FOS, JUN, MAPK3, RAC1, RAF1
hsa04670:Leukocyte transendothelial migration	8	4.86	0.003	PRKCA, CDC42, PTK2, PLCG1, PTK2B, RAC1, RHOA, RAP1A
hsa04910:Insulin signaling pathway	8	4.24	0.005	AKT1, MAPK1, INS, MAPK3, RAF1, ELK1, MAPK8, SHC1
hsa04020:Calcium signaling pathway	9	3.66	0.006	PRKCA, EGFR, ATP2B2, PLCG1, PTK2B, CAMK2G, NOS3, CACNA1E, PLCB1
hsa04070:Phosphatidylinositol signaling system	6	5.81	0.007	PRKCA, PLCG1, INPPL1, SYNJ1, PLCB1, PTEN
hsa05014:Amyotrophic lateral sclerosis (ALS)	5	6.76	0.012	CASP3, BCL2, BAX, RAC1, BCL2L1
hsa00562:Inositol phosphate metabolism	5	6.63	0.012	PLCG1, INPPL1, SYNJ1, PLCB1, PTEN
hsa04530:Tight junction	7	3.74	0.019	PRKCA, AKT1, CDC42, PRKCQ, RHOA, PRKCD, PTEN
hsa05020:Prion diseases	4	8.19	0.023	MAPK1, BAX, MAPK3, ELK1
hsa04730:Long-term depression	5	5.19	0.027	PRKCA, MAPK1, MAPK3, RAF1, PLCB1
hsa04630:Jak-STAT signaling pathway	7	3.23	0.034	AKT1, CREBBP, BCL2L1, IL7R, STAT1, MYC, STAT3
hsa04940:Type I diabetes mellitus	4	6.82	0.034	INS, FASLG, GZMB, CD28

[a]Fold Enrichment

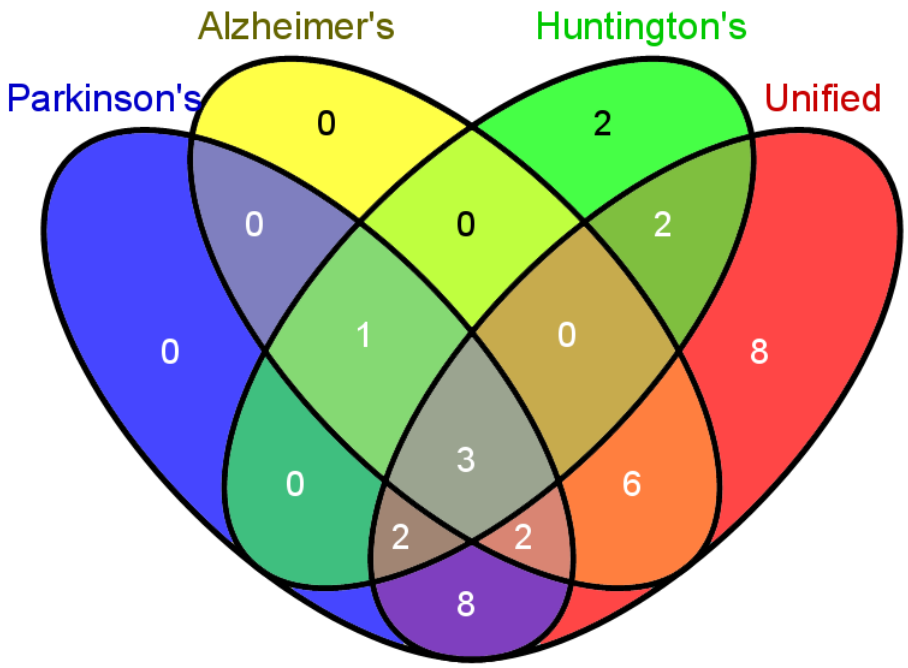


Figure 6.4. Four-set Venn diagram illustration of the overlap of enriched KEGG pathways found in Parkinson's, Alzheimer's, Huntington's and unified NDDs resulted from DAVID analysis. Courtesy: VENNY

ing pathway, type I diabetes mellitus and Prion diseases to be significantly affected. Deregulation of these pathways in NDDs has been claimed in the literature previously. Figure 6.4 and Table 6.3 illustrates the overlap of 34 enriched KEGG pathways that resulted from the DAVID analysis of the individual disease conditions as well as the unified NDDs.

Observing these kinds of *common* molecular players and biological processes is noteworthy suggesting that there is a unified underlying molecular mechanism

Table 6.3. Enriched KEGG pathways found in Parkinson's, Alzheimer's, Huntington's and unified NDDs resulted from DAVID analysis

KEGG pathways	Enriched in disease conditions
hsa04010:MAPK signaling pathway	PD, HD, unified NDDs
hsa04012:ErbB signaling pathway	PD, AD, HD, unified NDDs
hsa04020:Calcium signaling pathway	AD, unified NDDs
hsa04062:Chemokine signaling pathway	PD, AD, unified NDDs
hsa04115:p53 signaling pathway	PD, HD, unified NDDs
hsa04310:Wnt signaling pathway	PD, unified NDDs
hsa04350:TGF-beta signaling pathway	AD, unified NDDs
hsa04360:Axon guidance	PD, unified NDDs
hsa04370:VEGF signaling pathway	PD, unified NDDs
hsa04510:Focal adhesion	PD, unified NDDs
hsa04520:Adherens junction	PD, AD, HD, unified NDDs
hsa04530:Tight junction	PD, unified NDDs
hsa04540:Gap junction	PD, unified NDDs
hsa04620:Toll-like receptor signaling pathway	PD, unified NDDs
hsa04650:Natural killer cell mediated cytotoxicity	AD, unified NDDs
hsa04660:T cell receptor signaling pathway	AD, unified NDDs
hsa04662:B cell receptor signaling pathway	AD, unified NDDs
hsa04720:Long-term potentiation	AD, unified NDDs
hsa04722:Neurotrophin signaling pathway	PD, AD, unified NDDs
hsa04810:Regulation of actin cytoskeleton	PD, unified NDDs
hsa04910:Insulin signaling pathway	HD, unified NDDs
hsa04930:Type II diabetes mellitus	HD, unified NDDs
hsa05014:Amyotrophic lateral sclerosis (ALS)	PD, AD, HD, unified NDDs
hsa00562:Inositol phosphate metabolism	unified NDDs
hsa04070:Phosphatidylinositol signaling system	unified NDDs
hsa04630:Jak-STAT signaling pathway	unified NDDs
hsa04670:Leukocyte transendothelial migration	unified NDDs
hsa04730:Long-term depression	unified NDDs
hsa04912:GnRH signaling pathway	unified NDDs
hsa04920:Adipocytokine signaling pathway	HD
hsa04940:Type I diabetes mellitus	unified NDDs
hsa05010:Alzheimer's disease	PD, AD, HD
hsa05016:Huntington's disease	HD
hsa05020:Prion diseases	unified NDDs

across several NDDs. Based on this common phenomena and further analysis of the *integrated* mechanism network, we propose several possible ways of neurodegeneration initiation, as well as neuroprotection mechanisms common for Parkinson's, Alzheimer's and Huntington's disease. First, we will describe some possible disease induction routes. The *integrated* network shown in Figure 6.3, includes 11 extracellular ligands and 15 membrane bound proteins and receptors. Some of these entities such as BDNF, EGFR, INS and TGFB1 were also detected in the individual *integrated* Parkinson's, Alzheimer's and Huntington's disease networks (see section 3.4, 4.4 and 5.4).

6.4.4 Proposed neurodegeneration initiation routes

Reiterating the insulin deregulation in neurodegeneration hypothesis, we suggest that INS could be a major disease initiation origin. The extra-cellular ligand *INS* (insulin) interacts directly with many genes present not only in the membrane but also in the cytoplasm and nucleus. Chronic hyperinsulinaemia and insulin resistance, or reduced insulin effectiveness, is suggested to exert a negative influence on memory and cognitive functions in AD patients (Umegaki, 2012; Ristow, 2004). This could be realized by the INS interactions with *phospholipase* (PLD2) and *synaptotagmins* (SYT1) as shown in our network. In general, phospholipases are a class of enzymes that hydrolyzes phospholipids into fatty acids and other lipid substances. Prolonged stimulation of phospholipase along with the accumulation of reactive oxygen species (ROS) could contribute to both acute (brain injuries) and chronic neurodegeneration diseases like Alzheimer's. Recently in an AD mouse study, it has been suggested that reduction of phospholipase could mitigate cognitive deficits (Servitja et al., 2002; Farooqui et al., 1997; Sanchez-Mejia et al., 2008). Synapto-

tagmins function as calcium sensors in the regulation of neurotransmitter release and other hormonal secretion. They are one of the key players in the rapid exocytosis and endocytosis of synaptic vesicles (SV). During synapse formation, SYT1 acts as calcium sensor and triggers the SV release (Chapman, 2008). Disruption of calcium homeostasis could be detrimental to the calcium-dependent synaptotagmin activities which could have adverse effects in SV trafficking and fusion, thus contributing to neurodegeneration (Glavan et al., 2009). Another study has explained the role of synaptotagmin in Ca^{2+} mediated insulin exocytosis (Gauthier and Wollheim, 2008). In addition, the *integrated* network also revealed direct inhibitory interaction between INS and BCL6. As mentioned earlier, the role of BCL6 is unclear and further experimental validation is required to explore whether any neurodegeneration occurs as a result of such interactions between INS and BCL6. In general, many research work supports that deregulation of insulin potentially predisposes for the development of neurodegeneration (Schubert et al., 2004; Craft and Watson, 2004).

Another proposed route for disease induction is via a different extra-cellular ligand namely, *F2* (coagulation factor II (thrombin)). Animal studies have demonstrated that thrombin is neurotoxic to the dopaminergic neurons and induces microglial activity, which contributes to neuronal loss. At physiological concentrations, thrombin enhances the survival of neurons and astrocytes that were exposed to cellular injury. However, at higher concentrations thrombin promotes neuron degeneration. Caspase-3 mediates the neurotoxic apoptotic action of thrombin in motor neurons (Choi et al., 2003; Lee et al., 2006). In the *integrated* network, F2 directly interacts with many genes including phospholipase D2 which was mentioned in the foregoing as a calmodulin- and actin-binding protein potentially involved in the neurodegeneration pathogenesis. The CASP3 and PTEN are the major down-

stream targets of CALD1. PTEN is already implicated in Alzheimer's disease where it is associated with the hyperphosphorylated tau protein formations, a hallmark feature of NDDs (Kerr et al., 2006).

Third possible route for neurodegeneration process is through the *FAS* ligand namely, FASLG. Interaction of FAS and its receptor is critical in triggering apoptosis. In general, FAS is required to maintain the immune suppressed status in brain. But when up-regulated, it is suspected to promote neuronal cell death and inflammation in a variety of neurological disorders including Multiple sclerosis (MS), Prion diseases, PSPs etc. Studies have suggested that FAS-mediated signaling might contribute to immune-mediated oligodendrocyte injury and demyelination in MS (D'Souza et al., 1996; Choi and Benveniste, 2004). The *integrated* network reveals FAS interactions with phospholipase B1 and other downstream targets which increases the possibility of FAS-mediated neuronal loss in NDDs.

In addition to the above discussion, it should be noted that many T-cell molecules like CD8a, CD28, CD44 and CD69 were identified in the extra-cellular space of the *integrated* network, as well as in the cell membrane. Increasing biochemical research evidence suggests that receptors of the innate immune response could act as sensors that induce or amplify inflammation in NDDs like ALS, MS etc. (Glass et al., 2010). Neuroinflammation is one of the key features that accelerate neurodegeneration process and these T-cells could mediate such mechanism in NDDs.

6.4.5 Proposed therapeutic measures

Based on the *integrated* network (see Figure 6.3), until now we proposed many neurodegeneration initiation routes. Below we focus on similar kinds of analysis and proposals for neuroprotective mechanisms. Several growth factors, including

BDNF, EGF, NGF as well as colony stimulating factor, endothelin and TGFB1, were observed in the *integrated* network. Many of the above listed extra-cellular ligands have been pursued as neuroprotective measures in NDDs. Up-regulation of brain-derived neurotrophic factor (BDNF) has been widely explored as neuroprotective mechanism in NDDs including Alzheimer's and Huntington's disease (Ferrer et al., 2000; Zuccato et al., 2001; Tapia-Arancibia et al., 2008). Nerve growth factor (NGF) mediates memory formation, learning and other higher cognitive abilities in addition to assisting with neuronal development and repair process. NGF cell therapy has shown cognitive improvement in patients with Alzheimer's disease (Sofroniew et al., 2001; Chao et al., 2006). Another growth factor namely, epidermal growth factor (EGF) has neuromodulatory role in the central nervous system. EGF also stimulates neuronal growth, increases dopamine uptake and enhances the survival of dopaminergic neurons. Recently, it has been shown that EGF promotes neurogenesis in age-related neurological disorders. Its survival activities are realized through the activation of its receptor EGFR which is also identified in our *integrated* network as one of the many connecting genes/proteins (Jin et al., 2003; Wong and Guillaud, 2004). Reduced TGF-β signaling in neurons resulted in age-mediated neurodegeneration. Up-regulation of TGFB1 signaling has been proposed as a therapeutic measure in mitigating the amyloid-formation in Alzheimer's disease. In the *integrated* network such neuroprotective mechanism of TFGB1 could be enhanced by its interaction with BDNF, EGFR and other downstream targets (Krieglstein et al., 2002; Tesseur et al., 2006).

Besides the genes, microRNA regulation could also be exercised to promote neurogenesis as well as suppress the harmful gene expressions. Among the ten microRNAs proposed earlier, the following microRNAs miR-128-1, miR-135A1, miR-17, miR-19A, miR-27A, miR-30B and miR-32 could have a potential role in suppressing

the activity of some of the neurodegeneration promoting genes like CALD1, PLCB1 and SYT1. Similarly, epigenetic regulation of miR-137, miR-15A and miR-16-1 expressions (e.g., by hypermethylation) could be beneficial for the normal functioning of pro-survival genes like MCL1 and DCLK1.

Analyzing the *integrated* neurodegeneration mechanism network, it was evident that there are many complex interactions occurring between diseases promoting as well as protecting agents. Many of the genes in the network are essential for the general functioning of the cell including metabolism, growth, cytoskeletal organization, cell cycle control and transcriptional regulation etc. A delicate balance between these entities has to be established and maintained for the cell to survive. Identifying disease initiation and protective routes is vital in restoring cellular homeostasis and helps in identifying potential drug targets. Proceeding from similar underlying pathophysiological conditions and researching their *integrated* network, seems promising approach to reveal such disease initiating and neuroprotective mechanisms that are common for Parkinson's, Alzheimer's and Huntington's disease.

6.5 Summary

Motivation to undertake this research work is the growing evidence of deregulation of many biological processes that were found in common in several neurodegenerative disorders. Using network-based systems biology approach we set out to investigate the critical molecular players and underlying mechanisms in Parkinson's, Alzheimer's and Huntington's disease. After initial success with such network-based approaches in individual disease realm, we extended our inquiry for key players and mechanisms that could be common in all three NDDs. The unified network-based analysis was initiated using the overlapping 22 significantly differentially expressed

genes found in all three NDDs microarray datasets. Using the 22 "seed genes", several networks including shortest path, common targets and common regulators were constructed and merged to form one big *integrated* network which is used to examine the critical molecular partakers and biological processes that are common in all three NDDs. Four of the 22 seed genes remained unconnected in our networks. From the rest 18 ten (ATP2B2, ITSN1, MCL1, NTRK3, PLCB1, RASGRP1, REPS2, SCAMP1, STXBP1 and SYT1) were previously known to be modulated in NDDs, while the other eight (ADAM23, AHNAK, BCL6, CALD1, CAPRIN1, DCLK1, SSBP3 and ZMAT3) are reported for the first time in this function.

The *shortest path* network incorporated many previously known-NDDs genes like BNDF, CREBBP, CDK5, EGFR, MBP, NGF, PTEN and SP1 etc. as *connecting* genes/proteins. Many of these known genes were highly connected nodes which suggest that these genes could be critical in the neurodegeneration mechanism. MCL1 and BCL6 were the most highly connected common regulators as well as common targets in the gene set used for the study. Finding such critical nodes is significant since they can be promising drug targets in NDDs. On examining the *integrated* network it was evident that there is complex inter-connection existing between disease promoting and protective genes. It was difficult to have a line of separation between these two types of agents. Via their guilt by associations with several well-known NDDs genes, we propose AHNAK, CALD1 and DCLK1 as novel candidate genes common for Parkinson's, Alzheimer's and Huntington's disease. Apart from being neighbors with previously implicated disease genes, these three candidate genes have innate molecular functions relevant to neurodegeneration and protection mechanism. Disruption of CALD1 and AHNAK could increase neurodegeneration and the subsequent neuronal loss whereas DCLK1 could be pursued as potential drug target to enhance the memory formation and cognition.

Besides genes, many microRNAs-target interactions were also detected in the *integrated* network. Other than few known microRNAs, we identified ten novel microRNAs namely, miR-128-1, miR-135A1, miR-137, miR-15A, miR-16-1, miR-17, miR-19A, miR-27A, miR-30B and miR-32 which occurred repeatedly in individual and *integrated* networks of NDDs. Further experimental validation of these microRNAs could shed more light on the regulatory mechanisms in NDDs. Detection of transcription factors in the *integrated* network reveals the possibility of dual-level regulation in NDDs thereby adding more complexity to the disease mechanism. DAVID analysis uncovered many biological processes like aging, memory or learning, neuron development and differentiation, synaptic vesicle trafficking and endocytosis and vasculature development etc. including KEGG pathways like long-term potentiation, axon guidance, calcium signaling pathway, gap and tight junction, neurotrophin signaling pathway, ALS, AD and Prion diseases to be significantly affected in all three NDDs under study.

In addition, we propose several disease initiating and neuroprotecting routes that are common for Parkinson's, Alzheimer's and Huntington's disease. In the *integrated* network, extra-cellular ligands like INS, F2 and FASLG could have a major role in initiating the vicious cycle of inducing abnormal protein accumulation which in turn promotes neurodegeneration and eventual neuronal loss. This mechanism is mediated by the ligands and membrane receptors and is followed by disruption of the normal activity of their downstream targets. Similarly, we propose several more routes for neuroprotective mechanisms that could have a universal role in the three neurodegenerative disorders. Up-regulation of BDNF, EGF, NGF, TGFB1 along with EGFR could promote neurogenesis and restore homeostasis. In conclusion, our systems biology study reveals important details of neurodegenerative disorders as polygenic and highly complex in nature. Finding and maintaining the delicate

balance between NDD promoters and protectors is essential for restoring the homeostasis.

In conclusion, using network-based system biology techniques we have identified many critical molecular players as well as deregulation of various biological processes that were common in Parkinson's, Alzheimer's and Huntington's neurodegenerative diseases. Based on these findings, we have proposed a unified *integrated* neurodegeneration mechanism network with several possible routes for disease initiation and neuroprotection processes verifying thus the basic hypothesis of this research project.

Chapter 7

Conclusions and future directions

7.1 Major findings

Major findings of our research work are listed below. Table 7.1 summarizes the main results from Chapters 3-5 and lists the novel candidate genes, novel microRNAs as well as the proposed routes for disease initiation and/or neuroprotective modulation in Parkinson's, Alzheimer's and Huntington's disease. Table 7.2 lists the major findings of the unified Neurodegenerative disorders (NDDs) network-based analysis. The reported seed genes are common for the three studied NDDs, while the new genes, miRNAs, biological processes, pathways and routes for disease initiation or potential disease suppression are obtained from the network analysis of the unified mechanism of neurodegenerative disorders proposed in Chapter 6.

7.2 Future directions

Network-based approaches are powerful and beneficial tools to study complex systems like neurodegenerative disorders. By constructing and analyzing relevant net-

Table 7.1. Major findings in Parkinson's, Alzheimer's and Huntington's disease network-based analysis.

Categories	Parkinson's disease (PD)	Alzheimer's disease (AD)	Huntington's disease (HD)	Counts
Novel candidate genes	ACACB, ADAM17, BSN, CACNA1G, CDKN1B, CEBPA, CTNNB1, CTNND1, DCLK1, EGFR, KCNQ2, KLF1, NCAM1, NEDD4L, PAK1, PCHD8, ROCK1, SEMA6D, STXBP1, SYN1, SYNJ1, TIAM1, TUBB3, UBE2N, UNC13A, VAMP2	ARRB1, CAMK4, CHRM3, CR2, CSF1R, ESRRA, GAB1, HEY2, IL3, MDM2, NOX4, NRF1, PTK2B, SPHK2, SRC, SRF	ATXN7, CDK1, CEBPA, CNTNAP1, CX3CL1, DCLK1, DPYSL2, E2F1, EGR1, ERBB2, FGFR1, FOXO1, GNAQ, HSP90AA1, IL2, INSR, LRP1, PPARA, PRDX2, PRDX2, PRDX6, PRKCB, PRKCZ, RCAN2, SRF, STAT5A, THRA, TNF, ZFP36L1, ZNF148	26 in PD, 16 in AD, 30 in HD
Novel microRNAs	miR-132, miR-133a1, miR-181-1, miR-182, miR-218-1, miR-29a, miR-330	miR-101-1, miR-107, miR-124-1, miR-135a1, miR-142, miR-146a, miR-155, miR-15a, miR-181a1, miR-184, miR-19a, miR-221, miR-298, miR-302a, miR-328, miR-520B, miR-7-1	miR-101-1, miR-124-1, miR-128-1, miR-135A1, miR-141, miR-153-1, miR-15A, miR-16-1, miR-182, miR-19A, miR-27A, miR-96	7 in PD, 17 in AD, 12 in HD
ELDDIN routes[a]	CX3CL1, IL12B, SEMA6D	CD4, DCN, IL8	FGF2, TGFB, TNF	3 for each

For full details see Figures 3.4, 4.4, 5.4.

[a]Extra-cellular Ligand Downstream Disease Initiating/Neuroprotecting

Table 7.2. Major findings of unified network-based analysis of Neurodegenerative disorder's (NDDs).

Categories	Counts	Findings
Common "seed genes"	22	ADAM23, AHNAK, ATP2B2, ATP6V0E1, BCL6, CALD1, CAPRIN1, DCLK1, GLT1D1, ITSN1, MCL1, MSI2, NTRK3, PLCB1, PREPL, RASGRP1, REPS2, SCAMP1, SSBP3, STXBP1, SYT1, ZMAT3
Previously known NNDs genes	10	ATP2B2, ITSN1, MCL1, NTRK3, PLCB1, RASGRP1, REPS2, SCAMP1, STXBP1, SYT1
Novel candidate genes	5	ADAM23, AHNAK, BCL6, CALD1, DCLK1
Novel microRNAs	10	miR-128-1, miR-135A1, miR-137, miR-15A, miR-16-1, miR-17, miR-19A, miR-27A, miR-30B, miR-32
Novel therapeutic targets	2	BCL6, MCL1
Major common biological processes	7	Aging, memory or learning, neuron development and differentiation, synaptic vesicle trafficking and endocytosis, vasculature development
NDD common KEGG pathways	34	Adipocytokine, B-cell receptor, calcium, chemokine, ErbB, GnRH, Insulin, Jak-STAT, MAPK, neurotrophin, p53, T-cell receptor, Toll-like receptor, TGF-beta, VEGF, Wnt signaling pathways, axon guidance, focal adhesion, adherens, tight and gap juntions, natural killer cell mediated cytotoxicity, regulation of actin cytoskeleton, inositol phosphate metabolism, phosphatidylinositol signaling system, leukocyte transendothelial migration, long-term depression/potentiation, type I/II diabetes mellitus, ALS, AD, HD, Prion diseases [a]
ELDDIN routes[b]	11	ADAM23, BDNF, CD84, CSF1, EDN1, EGF, F2, FASLG, INS, NGF, TGFB1
Major disease initators	3	F2, FASLG, INS
Major neuroprotective mediators	5	BDNF, EGF, EGFR, NGF, TGFβ1

For full details see Figure 6.3.

[a] ALS - Amyotrophic lateral sclerosis; AD - Alzheimer's disease; HD - Huntington's disease

[b] Extra-cellular Ligand Downstream Disease Initiating/Neuroprotecting

works one could arrive at a broader view of the whole system. The structure of the network is critical for its functioning and reveals the key players and molecular pathways involved, along with the underlying mechanisms of the system.

In the past, guilt-by-association techniques and shortest-path network-based analysis have played a vital role in predicting protein-protein interactions in yeast as well as identifying novel genes that modulate longevity in Saccharomyces cerevisiae (Schwikowski et al., 2000; Managbanag et al., 2008). Especially such network-based candidate genes predictions have been successfully validated by experimentalists there by contributing important advancement in aging-related research realm.

Based on such network approach, we have only exposed the tip of the iceberg of the complex neurodegenerative mechanism. There are too many critical players with intertwined as well as delicate interactions among them. This opens new prospects for integrative studies of neurodegenerative diseases. Using network-based analysis, we have narrowed down the search for critical molecular players for each of the three neurodegenerative disorders under study, as well as novel common key players in this disease realm. Moving forward, both for the individual disease conditions and the unified approach, we could further refine our proposed candidate genes selections based on additional network attributes like eigenvector centrality scores, cliques etc. (Koschützki and Schreiber, 2008; Ozgür et al., 2008). Future research along these lines may include a broad collaboration with experimentalists to test some of the candidate genes and NDDs mechanisms proposed in our project with good chances for funded studies.

On computational side, the work done in this project may be considerably expanded to include more advanced and accurate biochemical assays like exon arrays, RNA-seq etc., as well as *integrated* network analysis of larger amount and broader types of experimental datasets. We believe such development could expand and

fine-tune our current findings thus contributing for the better understanding and treatment of these dangerous diseases. Apart from this, integrating diverse data types like proteomic data with gene expression data, one could have better understanding of the underlying complex molecular characteristics of the neurodegeneration mechanisms. This integrated approach will permit us to identify genes at the transcription level along with post-transcriptional regulations that are crucial during the neurodegeneration process which would otherwise be not apparent by examining either mRNA or protein expression alone (Ideker et al., 2001; Grant and Blackstock, 2001; Tian et al., 2004).

In addition, some of our proposed candidate genes and their critical functioning in neurodegeneration process could be tested in appropriate animal models. Often aging increases the risk for neurodegeneration in human which may not be directly studied using animal models. However, many genetically engineered animals have been instrumental in studying such complex diseases like NDDs (Phillips et al., 2009; Jucker, 2010; Langui et al., 2007). Utilizing these animal models, we believe at least a handful of our candidate genes could be further evaluated for potential drug targets for neurodegenerative disorders. Age factor in human neurodegeneration process could be studied in time-series experiments by including animals at different chronological age.

In recent years, cerebrospinal fluids and blood samples have been investigated for early disease detection, as well as for biomarker validation. Doecke et al. (2012) have identified a panel of 18 blood-based biomarkers for Alzheimer's disease where the presence of genes/proteins related to inflammation and especially insulin-like and epidermal growth factors had increased the sensitivity of the assay. From our network-based gene expression analysis, we also found genes/proteins related to inflammation (like interleukins) and EGFR, INS as well as genes/proteins related

to wound healing and aging to be highly enriched in Parkinson's, Alzheimer's and Huntington's disease conditions. Similarly, Molochnikov et al. (2012) identified seven blood-based protein biomarkers for Parkinson's disease, such as p19 S-phase kinase-associated protein 1A (SKP1), huntingtin interacting protein-2 (RNF2) and heat shock 70-kDa protein 8 (HSPA8). In future we anticipate that such network-based analysis will be vital in discovering better biomarkers from blood-based gene expression datasets. The blood-samples analysis will also be extended to study the protein-protein interaction networks that evolve over a time period.

Many prospective multi-disciplinary research opportunities could thus be opened in near future by our network-based analysis of neurodegenerative disorders for further refining our novel candidate genes collections and validating some of proposed disease initiating and neuroprotective mechanisms.

Bibliography

Agadjanyan, M. G., A. Ghochikyan, I. Petrushina, V. Vasilevko, N. Movsesyan, M. Mkrtichyan, T. Saing, and D. H. Cribbs, 2005: Prototype Alzheimer's disease vaccine using the immunodominant B cell epitope from beta-amyloid and promiscuous T cell epitope pan HLA DR-binding peptide. *Journal of immunology (Baltimore, Md. : 1950)*, **174**, 1580–6. 87

Akiyama, H., S. Barger, S. Barnum, B. Bradt, J. Bauer, G. M. Cole, N. R. Cooper, P. Eikelenboom, M. Emmerling, B. L. Fiebich, C. E. Finch, S. Frautschy, W. S. Griffin, H. Hampel, M. Hull, G. Landreth, L. Lue, R. Mrak, I. R. Mackenzie, P. L. McGeer, M. K. O'Banion, J. Pachter, G. Pasinetti, C. Plata-Salaman, J. Rogers, R. Rydel, Y. Shen, W. Streit, R. Strohmeyer, I. Tooyoma, F. L. Van Muiswinkel, R. Veerhuis, D. Walker, S. Webster, B. Wegrzyniak, G. Wenk, and T. Wyss-Coray, 2000: Inflammation and Alzheimer's disease. *Neurobiology of aging*, **21**, 383–421. 90, 128

Albin, R. L., A. Reiner, K. D. Anderson, J. B. Penney, and A. B. Young, 1990: Striatal and nigral neuron subpopulations in rigid Huntington's disease: implications for the functional anatomy of chorea and rigidity-akinesia. *Annals of neurology*, **27**, 357–65. 13

Anderson, L. R., R. Betarbet, M. Gearing, J. Gulcher, A. A. Hicks, K. Stefánsson, J. J. Lah, and A. I. Levey, 2007: PARK10 candidate RNF11 is expressed by vulnerable neurons and localizes to Lewy bodies in Parkinson disease brain. *Journal of neuropathology and experimental neurology*, **66**, 955–64. 7

Andreassen, O. A., A. Dedeoglu, V. Stanojevic, D. B. Hughes, S. E. Browne, C. A. Leech, R. J. Ferrante, J. F. Habener, M. F. Beal, and M. K. Thomas, 2002: Huntington's disease of the endocrine pancreas: insulin deficiency and diabetes mellitus due to impaired insulin gene expression. *Neurobiology of disease*, **11**, 410–24. 114

Apostol, B. L., K. Illes, J. Pallos, L. Bodai, J. Wu, A. Strand, E. S. Schweitzer, J. M. Olson, A. Kazantsev, J. L. Marsh, and L. M. Thompson, 2006: Mutant huntingtin alters MAPK signaling pathways in PC12 and striatal cells: ERK1/2 protects against mutant huntingtin-associated toxicity. *Human molecular genetics*, **15**, 273–85. 114

Arikkath, J. and L. F. Reichardt, 2008: Cadherins and catenins at synapses: roles in synaptogenesis and synaptic plasticity. *Trends in neurosciences*, **31**, 487–94. 87

Auffret, A., V. Gautheron, M. P. Mattson, J. Mariani, and C. Rovira, 2010: Progressive age-related impairment of the late long-term potentiation in Alzheimer's disease presenilin-1 mutant knock-in mice. *Journal of Alzheimer's disease : JAD*, **19**, 1021–33. 87

Ayuso-Sacido, A., J. A. Moliterno, S. Kratovac, G. S. Kapoor, D. M. O'Rourke, E. C. Holland, J. M. García-Verdugo, N. S. Roy, and J. A. Boockvar, 2010: Activated EGFR signaling increases proliferation, survival, and migration and blocks neuronal differentiation in post-natal neural stem cells. *Journal of neuro-oncology*, **97**, 323–37. 59

Bae, B.-I., H. Xu, S. Igarashi, M. Fujimuro, N. Agrawal, Y. Taya, S. D. Hayward, T. H. Moran, C. Montell, C. A. Ross, S. H. Snyder, and A. Sawa, 2005: p53 mediates cellular dysfunction and behavioral abnormalities in Huntington's disease. *Neuron*, **47**, 29–41. 113

Bak, M., A. Silahtaroglu, M. Mø ller, M. Christensen, M. F. Rath, B. Skryabin, N. Tommerup, and S. Kauppinen, 2008: MicroRNA expression in the adult mouse central nervous system. *RNA (New York, N.Y.)*, **14**, 432–44. 68

Barabási, A.-L., N. Gulbahce, and J. Loscalzo, 2011: Network medicine: a network-based approach to human disease. *Nature reviews. Genetics*, **12**, 56–68. 19

Barabási, A.-L. and Z. N. Oltvai, 2004: Network biology: understanding the cell's functional organization. *Nature reviews. Genetics*, **5**, 101–13. 17, 19

Barbato, C., F. Ruberti, and C. Cogoni, 2009: Searching for MIND: microRNAs in neurodegenerative diseases. *Journal of biomedicine & biotechnology*, **2009**, 871313. 93

Batagelj, V. and A. Mrvar, 1998: Pajek - Program for Large Network Analysis. *Connections*, **21**, 1–11. 42

Bates, G., 2003: Huntingtin aggregation and toxicity in Huntington's disease. *Lancet*, **361**, 1642–4. 13

Battaglia, G., M. Cannella, B. Riozzi, S. Orobello, M. Maat-Schieman, E. Aronica, C. L. Busceti, A. Ciarmiello, S. Alberti, and J. Sassone, 2010: Platform presentationâĂŤEarly Defect of Transforming Growth Factor β1 Formation in Huntington's Disease. *Neurotherapeutics*, **7**, 135–136. 117

Bekris, L. M., C.-E. Yu, T. D. Bird, and D. W. Tsuang, 2010: Genetics of Alzheimer disease. *Journal of geriatric psychiatry and neurology*, **23**, 213–27. 9, 10

Benarroch, E. E., 2012: ADAM proteins, their ligands, and clinical implications. *Neurology*, **78**, 914–20. 66

Benjamini, Y. and Y. Hochberg, 1995: Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, **57**, pp. 289–300. 32

Bergmann, S., J. Ihmels, and N. Barkai, 2004: Similarities and differences in genome-wide expression data of six organisms. *PLoS biology*, **2**, E9. 19

Betarbet, R., T. B. Sherer, and J. T. Greenamyre, 2005: Ubiquitin-proteasome system and Parkinson's diseases. *Experimental neurology*, **191 Suppl**, S17–27. 51

Blalock, E. M., H. M. Buechel, J. Popovic, J. W. Geddes, and P. W. Landfield, 2011: Microarray analyses of laser-captured hippocampus reveal distinct gray and white matter signatures associated with incipient Alzheimer's disease. *Journal of chemical neuroanatomy*, **42**, 118–26. 76

Bleiler, W. T. and Laura, 2012: 2012 Alzheimer's disease facts and figures. *Alzheimer's & dementia : the journal of the Alzheimer's Association*, **8**, 131–68. 1, 8

Bode, F. J., M. Stephan, H. Suhling, R. Pabst, R. H. Straub, K. A. Raber, M. Bonin, H. P. Nguyen, O. Riess, A. Bauer, C. Sjoberg, A. Petersén, and S. von Hörsten, 2008: Sex differences in a transgenic rat model of Huntington's disease: decreased 17beta-estradiol levels correlate with reduced numbers of DARPP32+ neurons in males. *Human molecular genetics*, **17**, 2595–609. 111

Burke, J. R., J. J. Enghild, M. E. Martin, Y. S. Jou, R. M. Myers, A. D. Roses, J. M. Vance, and W. J. Strittmatter, 1996: Huntingtin and DRPLA proteins selectively interact with the enzyme GAPDH. *Nature medicine*, **2**, 347–50. 15

Calderwood, M. A., K. Venkatesan, L. Xing, M. R. Chase, A. Vazquez, A. M. Holthaus, A. E. Ewence, N. Li, T. Hirozane-Kishikawa, D. E. Hill, M. Vidal, E. Kieff, and E. Johannsen, 2007: Epstein-Barr virus and virus human protein interaction maps. *Proceedings of the National Academy of Sciences of the United States of America*, **104**, 7606–11. 17

Camicioli, R., M. M. Moore, A. Kinney, E. Corbridge, K. Glassberg, and J. A. Kaye, 2003: Parkinson's disease is associated with hippocampal atrophy. *Movement disorders : official journal of the Movement Disorder Society*, **18**, 784–90. 68

Campbell, I. M., S. A. Yatsenko, P. Hixson, T. Reimschisel, M. Thomas, W. Wilson, U. Dayal, J. W. Wheless, A. Crunk, C. Curry, N. Parkinson, L. Fishman, J. J. Riviello, M. J. M. Nowaczyk, S. Zeesman, J. A. Rosenfeld, B. A. Bejjani, L. G. Shaffer, S. W.

Cheung, J. R. Lupski, P. Stankiewicz, and F. Scaglia, 2012: Novel 9q34.11 gene deletions encompassing combinations of four Mendelian disease genes: STXBP1, SPTAN1, ENG, and TOR1A. *Genetics in medicine : official journal of the American College of Medical Genetics*, **14**, 868–76. 50

Carlsson, T., F. R. Schindler, M. Höllerhage, C. Depboylu, O. Arias-Carrión, S. Schnurrbusch, T. W. Rösler, W. Wozny, G. P. Schwall, K. Groebe, W. H. Oertel, P. Brundin, A. Schrattenholz, and G. U. Höglinger, 2011: Systemic administration of neuregulin-1β1 protects dopaminergic neurons in a mouse model of Parkinson's disease. *Journal of neurochemistry*, **117**, 1066–74. 58

Castedo, M., J.-L. Perfettini, T. Roumier, and G. Kroemer, 2002: Cyclin-dependent kinase-1: linking apoptosis to cell cycle and mitotic catastrophe. *Cell death and differentiation*, **9**, 1287–93. 110

Cerami, E., E. Demir, N. Schultz, B. S. Taylor, and C. Sander, 2010: Automated network analysis identifies core pathways in glioblastoma. *PloS one*, **5**, e8918. 17

Cesca, F., P. Baldelli, F. Valtorta, and F. Benfenati, 2010: The synapsins: key actors of synapse function and plasticity. *Progress in neurobiology*, **91**, 313–48. 50

Chandrasekaran, S. and D. G. Bonchev, 2012: A network view on Schizophrenia related genes. *Network Biology*, **2**, 16–25. 21

Chao, M. V., R. Rajagopal, and F. S. Lee, 2006: Neurotrophin signalling in health and disease. *Clinical science (London, England : 1979)*, **110**, 167–73. 87, 149

Chapman, E. R., 2008: How does synaptotagmin trigger neurotransmitter release? *Annual review of biochemistry*, **77**, 615–41. 147

Chaudhury, A. R., K. M. Gerecke, J. M. Wyss, D. G. Morgan, M. N. Gordon, and S. L. Carroll, 2003: Neuregulin-1 and erbB4 immunoreactivity is associated with neuritic plaques in Alzheimer disease brain and in a transgenic model of Alzheimer disease. *Journal of neuropathology and experimental neurology*, **62**, 42–54. 87

Chen, H., B. Liu, and A. H. Neufeld, 2007: Epidermal growth factor receptor in adult retinal neurons of rat, mouse, and human. *The Journal of comparative neurology*, **500**, 299–310. 59

Chen, Q. S., B. L. Kagan, Y. Hirakura, and C. W. Xie, 2000: Impairment of hippocampal long-term potentiation by Alzheimer amyloid beta-peptides. *Journal of neuroscience research*, **60**, 65–72. 87

Chiba, T., M. Yamada, and S. Aiso, 2009: Targeting the JAK2/STAT3 axis in Alzheimer's disease. *Expert opinion on therapeutic targets*, **13**, 1155–67. 9

Chiti, F. and C. M. Dobson, 2006: Protein misfolding, functional amyloid, and human disease. *Annual review of biochemistry*, **75**, 333–66. 2, 128

Choi, C. and E. N. Benveniste, 2004: Fas ligand/Fas system in the brain: regulator of immune and apoptotic responses. *Brain Research Reviews*, **44**, 65–81. 148

Choi, S.-H., D. Y. Lee, J. K. Ryu, J. Kim, E. H. Joe, and B. K. Jin, 2003: Thrombin induces nigral dopaminergic neurodegeneration in vivo by altering expression of death-related proteins. *Neurobiology of disease*, **14**, 181–93. 147

Christen, Y., 2000: Oxidative stress and Alzheimer disease. *The American journal of clinical nutrition*, **71**, 621S–629S. 79

Chung, S. J., S. M. Armasu, J. M. Biernacka, T. G. Lesnick, D. N. Rider, S. J. Lincoln, A. I. Ortolaza, M. J. Farrer, J. M. Cunningham, W. A. Rocca, and D. M. Maraganore, 2011: Common variants in PARK loci and related genes and Parkinson's disease. *Movement disorders : official journal of the Movement Disorder Society*, **26**, 280–8. 6

Churcher, I., 2006: Tau therapeutic strategies for the treatment of Alzheimer's disease. *Current topics in medicinal chemistry*, **6**, 579–95. 61

Citron, B. A., J. S. Dennis, R. S. Zeitlin, and V. Echeverria, 2008: Transcription factor Sp1 dysregulation in Alzheimer's disease. *Journal of neuroscience research*, **86**, 2499–504. 132

Cogswell, J. P., J. Ward, I. A. Taylor, M. Waters, Y. Shi, B. Cannon, K. Kelnar, J. Kemppainen, D. Brown, C. Chen, R. K. Prinjha, J. C. Richardson, A. M. Saunders, A. D. Roses, and C. A. Richards, 2008: Identification of miRNA changes in Alzheimer's disease brain and CSF yields putative biomarkers and insights into disease pathways. *Journal of Alzheimer's disease : JAD*, **14**, 27–41. 97

Cole, S. L. and R. Vassar, 2007: The Alzheimer's disease beta-secretase enzyme, BACE1. *Molecular neurodegeneration*, **2**, 22. 84, 137

Craft, S. and G. S. Watson, 2004: Insulin and neurodegenerative disease: shared and specific mechanisms. *Lancet neurology*, **3**, 169–78. 147

Creson, T. K., P. Yuan, H. K. Manji, and G. Chen, 2009: Evidence for involvement of ERK, PI3K, and RSK in induction of Bcl-2 by valproate. *Journal of molecular neuroscience : MN*, **37**, 123–34. 134

Crews, L. and E. Masliah, 2010: Molecular mechanisms of neurodegeneration in Alzheimer's disease. *Human molecular genetics*, **19**, R12–20. 12, 110, 111

Cui, Y.-F., G. Hargus, J.-C. Xu, J. S. Schmid, Y.-Q. Shen, M. Glatzel, M. Schachner, and C. Bernreuther, 2010: Embryonic stem cell-derived L1 overexpressing neural aggregates enhance recovery in Parkinsonian mice. *Brain : a journal of neurology*, **133**, 189–204. 7, 66

Das, P. and T. Golde, 2006: Dysfunction of TGF-beta signaling in Alzheimer's disease. *The Journal of clinical investigation*, **116**, 2855–7. 87

Davies, J. E., S. Sarkar, and D. C. Rubinsztein, 2007: The ubiquitin proteasome system in Huntington's disease and the spinocerebellar ataxias. *BMC biochemistry*, **8 Suppl 1**, S2. 13

Dhar, S. S., S. Ongwijitwat, and M. T. T. Wong-Riley, 2008: Nuclear respiratory factor 1 regulates all ten nuclear-encoded subunits of cytochrome c oxidase in neurons. *The Journal of biological chemistry*, **283**, 3120–9. 84

Dhawan, G. and C. K. Combs, 2012: Inhibition of Src kinase activity attenuates amyloid associated microgliosis in a murine model of Alzheimer's disease. *Journal of neuroinflammation*, **9**, 117. 83

Di Monte, D. A., M. Lavasani, and A. B. Manning-Bog, 2002: Environmental factors in Parkinson's disease. *Neurotoxicology*, **23**, 487–502. 3, 128

Di Paolo, G. and T.-W. Kim, 2011: Linking lipids to Alzheimer's disease: cholesterol and beyond. *Nature reviews. Neuroscience*, **12**, 284–96. 83

Diekstra, F. P., P. W. J. van Vught, W. van Rheenen, M. Koppers, R. J. Pasterkamp, M. A. van Es, H. J. Schelhaas, M. de Visser, W. Robberecht, P. Van Damme, P. M. Andersen, L. H. van den Berg, and J. H. Veldink, 2012: UNC13A is a modifier of survival in amyotrophic lateral sclerosis. *Neurobiology of aging*, **33**, 630.e3–8. 51

Doecke, J. D., S. M. Laws, N. G. Faux, W. Wilson, S. C. Burnham, C.-P. Lam, A. Mondal, J. Bedo, A. I. Bush, B. Brown, K. De Ruyck, K. A. Ellis, C. Fowler, V. B. Gupta, R. Head, S. L. Macaulay, K. Pertile, C. C. Rowe, A. Rembach, M. Rodrigues, R. Rumble, C. Szoeke, K. Taddei, T. Taddei, B. Trounson, D. Ames, C. L. Masters, and R. N. Martins, 2012: Blood-based protein biomarkers for diagnosis of Alzheimer disease. *Archives of neurology*, **69**, 1318–25. 158

Dowie, M. J., E. L. Scotter, E. Molinari, and M. Glass, 2010: The therapeutic potential of G-protein coupled receptors in Huntington's disease. *Pharmacology & therapeutics*, **128**, 305–23. 15

Doyle, K. M., D. Kennedy, A. M. Gorman, S. Gupta, S. J. M. Healy, and A. Samali, 2011: Unfolded proteins and endoplasmic reticulum stress in neurodegenerative disorders. *Journal of cellular and molecular medicine*, **15**, 2025–39. 12

D'Souza, S. D., B. Bonetti, V. Balasingam, N. R. Cashman, P. A. Barker, A. B. Troutt, C. S. Raine, and J. P. Antel, 1996: Multiple sclerosis: Fas signaling in oligodendrocyte cell death. *The Journal of experimental medicine*, **184**, 2361–70. 148

Dudek, H., 1997: Regulation of Neuronal Survival by the Serine-Threonine Protein Kinase Akt. *Science*, **275**, 661–665. 119

Dumstrei, K., C. Nassif, G. Abboud, A. Aryai, and V. Hartenstein, 1998: EGFR signaling is required for the differentiation and maintenance of neural progenitors along the dorsal midline of the Drosophila embryonic head. *Development (Cambridge, England)*, **125**, 3417–26. 59

Dunckley, T., T. G. Beach, K. E. Ramsey, A. Grover, D. Mastroeni, D. G. Walker, B. J. LaFleur, K. D. Coon, K. M. Brown, R. Caselli, W. Kukull, R. Higdon, D. McKeel, J. C. Morris, C. Hulette, D. Schmechel, E. M. Reiman, J. Rogers, and D. A. Stephan, 2006: Gene expression correlates of neurofibrillary tangles in Alzheimer's disease. *Neurobiology of aging*, **27**, 1359–71. 76, 98

Emerit, J., M. Edeas, and F. Bricaire, 2004: Neurodegenerative diseases and oxidative stress. *Biomedicine & Pharmacotherapy*, **58**, 39–46. 2, 79, 84, 128

Ermentrout, G. B. and L. Edelstein-Keshet, 1993: Cellular automata approaches to biological modeling. *Journal of theoretical biology*, **160**, 97–133. 121

Esposito, G., F. Ana Clara, and P. Verstreken, 2012: Synaptic vesicle trafficking and Parkinson's disease. *Developmental neurobiology*, **72**, 134–44. 51

Estrada, E., 2011: *The Structure of Complex Networks*. Oxford University Press, hardback edition, 480 pp. 20, 42

Farooqui, A. A., H. C. Yang, and L. Horrocks, 1997: Involvement of phospholipase A2 in neurodegeneration. *Neurochemistry international*, **30**, 517–22. 146

Farrer, M. J., 2006: Genetics of Parkinson disease: paradigm shifts and future prospects. *Nature reviews. Genetics*, **7**, 306–18. 5

Ferrante, R. J., H. Ryu, J. K. Kubilus, S. D'Mello, K. L. Sugars, J. Lee, P. Lu, K. Smith, S. Browne, M. F. Beal, B. S. Kristal, I. G. Stavrovskaya, S. Hewett, D. C. Rubinsztein, B. Langley, and R. R. Ratan, 2004: Chemotherapy for the brain: the antitumor antibiotic mithramycin prolongs survival in a mouse model of Huntington's disease. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, **24**, 10335–42. 109

Ferrer, I., E. Goutan, C. Marín, M. J. Rey, and T. Ribalta, 2000: Brain-derived neurotrophic factor in Huntington disease. *Brain research*, **866**, 257–61. 16, 149

Flirski, M. and T. Sobow, 2005: Biochemical markers and risk factors of Alzheimer's disease. *Current Alzheimer research*, **2**, 47–64. 12

Galpern, W. R. and A. E. Lang, 2006: Interface between tauopathies and synucleinopathies: a tale of two proteins. *Annals of neurology*, **59**, 449–58. 61

Galvin, J. E., V. M. Lee, and J. Q. Trojanowski, 2001: Synucleinopathies: clinical and pathological implications. *Archives of neurology*, **58**, 186–90. 60

Galvin, J. E., K. Uryu, V. M. Lee, and J. Q. Trojanowski, 1999: Axon pathology in Parkinson's disease and Lewy body dementia hippocampus contains alpha-, beta-, and gamma-synuclein. *Proceedings of the National Academy of Sciences of the United States of America*, **96**, 13450–5. 6

Garriga-Canut, M., C. Agustín-Pavón, F. Herrmann, A. Sánchez, M. Dierssen, C. Fillat, and M. Isalan, 2012: Synthetic zinc finger repressors reduce mutant huntingtin expression in the brain of R6/2 mice. *Proceedings of the National Academy of Sciences of the United States of America*, **109**, E3136–45. 105

Gasser, T., 2009: Molecular pathogenesis of Parkinson disease: insights from genetic studies. *Expert reviews in molecular medicine*, **11**, e22. 5

Gauthier, B. R. and C. B. Wollheim, 2008: Synaptotagmins bind calcium to release insulin. *American journal of physiology. Endocrinology and metabolism*, **295**, E1279–86. 147

Gentleman, R. C., V. J. Carey, D. M. Bates, B. Bolstad, M. Dettling, S. Dudoit, B. Ellis, L. Gautier, Y. Ge, J. Gentry, K. Hornik, T. Hothorn, W. Huber, S. Iacus, R. Irizarry, F. Leisch, C. Li, M. Maechler, A. J. Rossini, G. Sawitzki, C. Smith, G. Smyth, L. Tierney, J. Y. H. Yang, and J. Zhang, 2004: Bioconductor: open software development for computational biology and bioinformatics. *Genome biology*, **5**, R80. 30

George, J. M., 2002: The synucleins. *Genome biology*, **3**, REVIEWS3002. 4

Gil, J. M. and A. C. Rego, 2008: Mechanisms of neurodegeneration in Huntington's disease. *The European journal of neuroscience*, **27**, 2803–20. 13, 79, 128

Gitler, A. D., B. J. Bevis, J. Shorter, K. E. Strathearn, S. Hamamichi, L. J. Su, K. A. Caldwell, G. A. Caldwell, J.-C. Rochet, J. M. McCaffery, C. Barlowe, and S. Lindquist, 2008: The Parkinson's disease protein alpha-synuclein disrupts cellular Rab homeostasis. *Proceedings of the National Academy of Sciences of the United States of America*, **105**, 145–50. 7

Gitter, B. D., L. M. Cox, R. E. Rydel, and P. C. May, 1995: Amyloid beta peptide potentiates cytokine secretion by interleukin-1 beta-activated human astrocytoma cells. *Proceedings of the National Academy of Sciences of the United States of America*, **92**, 10738–41. 90

Glass, C. K., K. Saijo, B. Winner, M. C. Marchetto, and F. H. Gage, 2010: Mechanisms underlying inflammation in neurodegeneration. *Cell*, **140**, 918–34. 148

Glavan, G., R. Schliebs, and M. Zivin, 2009: Synaptotagmins in neurodegeneration. *Anatomical record (Hoboken, N.J. : 2007)*, **292**, 1849–62. 147

Glynn Dennis, J., G. Dennis, B. T. Sherman, D. A. Hosack, J. Yang, W. Gao, H. C. Lane, and R. A. Lempicki, 2003: DAVID: Database for Annotation, Visualization, and Integrated Discovery. *Genome Biology*, **4**, P3. 45

Godin, J. D., G. Poizat, M. A. Hickey, F. Maschat, and S. Humbert, 2010: Mutant huntingtin-impaired degradation of beta-catenin causes neurotoxicity in Huntington's disease. *The EMBO journal*, **29**, 2433–45. 113

Goehler, H., M. Lalowski, U. Stelzl, S. Waelter, M. Stroedicke, U. Worm, A. Droege, K. S. Lindenberg, M. Knoblich, C. Haenig, M. Herbst, J. Suopanki, E. Scherzinger, C. Abraham, B. Bauer, R. Hasenbank, A. Fritzsche, A. H. Ludewig, K. Büssow, K. Buessow, S. H. Coleman, C.-A. Gutekunst, B. G. Landwehrmeyer, H. Lehrach, and E. E. Wanker, 2004: A protein interaction network links GIT1, an enhancer of huntingtin aggregation, to Huntington's disease. *Molecular cell*, **15**, 853–65. 20

Goh, K.-I., M. E. Cusick, D. Valle, B. Childs, M. Vidal, and A.-L. Barabási, 2007: The human disease network. *Proceedings of the National Academy of Sciences of the United States of America*, **104**, 8685–90. 19

Grant, S. G. and W. P. Blackstock, 2001: Proteomics in neuroscience: from protein to network. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, **21**, 8315–8. 158

Hans, C. P., D. D. Weisenburger, T. C. Greiner, R. D. Gascoyne, J. Delabie, G. Ott, H. K. Müller-Hermelink, E. Campo, R. M. Braziel, E. S. Jaffe, Z. Pan, P. Farinha, L. M. Smith, B. Falini, A. H. Banham, A. Rosenwald, L. M. Staudt, J. M. Connors, J. O. Armitage, and W. C. Chan, 2004: Confirmation of the molecular classification of diffuse large B-cell lymphoma by immunohistochemistry using a tissue microarray. *Blood*, **103**, 275–82. 134

Harraz, M. M., T. M. Dawson, and V. L. Dawson, 2011: MicroRNAs in Parkinson's disease. *Journal of chemical neuroanatomy*, **42**, 127–30. 70, 74

Harris, T. W., E. Hartwieg, H. R. Horvitz, and E. M. Jorgensen, 2000: Mutations in synaptojanin disrupt synaptic vesicle recycling. *The Journal of cell biology*, **150**, 589–600. 71

Hays, S. E., F. K. Goodwin, and S. M. Paul, 1981: Cholecystokinin receptors are decreased in basal ganglia and cerebral cortex of Huntington's disease. *Brain research*, **225**, 452–6. 15

Hébert, S. S., K. Horré, L. Nicolaï, B. Bergmans, A. S. Papadopoulou, A. Delacourte, and B. De Strooper, 2009: MicroRNA regulation of Alzheimer's Amyloid precursor protein expression. *Neurobiology of disease*, **33**, 422–8. 93

Hébert, S. S., K. Horré, L. Nicolaï, A. S. Papadopoulou, W. Mandemakers, A. N. Silahtaroglu, S. Kauppinen, A. Delacourte, and B. De Strooper, 2008: Loss of microRNA cluster miR-29a/b-1 in sporadic Alzheimer's disease correlates with increased BACE1/beta-secretase expression. *Proceedings of the National Academy of Sciences of the United States of America*, **105**, 6415–20. 93

Hertel, M., S. Braun, S. Durka, C. Alzheimer, and S. Werner, 2002: Upregulation and activation of the Nrf-1 transcription factor in the lesioned hippocampus. *The European journal of neuroscience*, **15**, 1707–11. 84

Hicks, A. A., H. Pétursson, T. Jónsson, H. Stefánsson, H. S. Jóhannsdóttir, J. Sainz, M. L. Frigge, A. Kong, J. R. Gulcher, K. Stefánsson, and S. Sveinbjörnsdóttir, 2002: A susceptibility gene for late-onset idiopathic Parkinson's disease. *Annals of neurology*, **52**, 549–55. 6

Hodges, A., A. D. Strand, A. K. Aragaki, A. Kuhn, T. Sengstag, G. Hughes, L. A. Elliston, C. Hartog, D. R. Goldstein, D. Thu, Z. R. Hollingsworth, F. Collin, B. Synek, P. A. Holmans, A. B. Young, N. S. Wexler, M. Delorenzi, C. Kooperberg, S. J. Augood, R. L. M. Faull, J. M. Olson, L. Jones, and R. Luthi-Carter, 2006: Regional and cellular gene expression changes in human Huntington's disease brain. *Human molecular genetics*, **15**, 965–77. 101

Holgado-Madruga, M. and A. J. Wong, 2003: Gab1 is an integrator of cell death versus cell survival signals in oxidative stress. *Molecular and cellular biology*, **23**, 4471–84. 79

Howard, R., R. McShane, J. Lindesay, C. Ritchie, A. Baldwin, R. Barber, A. Burns, T. Dening, D. Findlay, C. Holmes, A. Hughes, R. Jacoby, R. Jones, R. Jones, I. McKeith, A. Macharouthu, J. O'Brien, P. Passmore, B. Sheehan, E. Juszczak, C. Katona, R. Hills, M. Knapp, C. Ballard, R. Brown, S. Banerjee, C. Onions, M. Griffin, J. Adams, R. Gray, T. Johnson, P. Bentham, and P. Phillips, 2012: Donepezil and memantine for moderate-to-severe Alzheimer's disease. *The New England journal of medicine*, **366**, 893–903. 12

Hsu, L. J., Y. Sagara, A. Arroyo, E. Rockenstein, A. Sisk, M. Mallory, J. Wong, T. Takenouchi, M. Hashimoto, and E. Masliah, 2000: alpha-synuclein promotes mitochondrial deficit and oxidative stress. *The American journal of pathology*, **157**, 401–10. 5

Huang, D. W., B. T. Sherman, and R. A. Lempicki, 2009a: Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic acids research*, **37**, 1–13. 45

—, 2009b: Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature protocols*, **4**, 44–57. 45

Ideker, T. and R. Sharan, 2008: Protein networks in disease. *Genome research*, **18**, 644–52. 21

Ideker, T., V. Thorsson, J. A. Ranish, R. Christmas, J. Buhler, J. K. Eng, R. Bumgarner, D. R. Goodlett, R. Aebersold, and L. Hood, 2001: Integrated genomic and proteomic analyses of a systematically perturbed metabolic network. *Science (New York, N.Y.)*, **292**, 929–34. 26, 158

Iijima, K., K. Ando, S. Takeda, Y. Satoh, T. Seki, S. Itohara, P. Greengard, Y. Kirino, A. C. Nairn, and T. Suzuki, 2000: Neuron-specific phosphorylation of Alzheimer's beta-amyloid precursor protein by cyclin-dependent kinase 5. *Journal of neurochemistry*, **75**, 1085–91. 85, 110

Imran, M. and S. Mahmood, 2011: An overview of animal prion diseases. *Virology journal*, **8**, 493. 107

Inestrosa, N. C. and E. M. Toledo, 2008: The role of Wnt signaling in neuronal dysfunction in Alzheimer's Disease. *Molecular neurodegeneration*, **3**, 9. 79

Irizarry, R. A., B. Hobbs, F. Collin, Y. D. Beazer-Barclay, K. J. Antonellis, U. Scherf, and T. P. Speed, 2003: Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics (Oxford, England)*, **4**, 249–64. 31

Jeanmougin, M., A. de Reynies, L. Marisa, C. Paccard, G. Nuel, and M. Guedj, 2010: Should we abandon the t-test in the analysis of gene expression microarray data: a comparison of variance modeling strategies. *PloS one*, **5**, e12336. 32

Jeffery, I. B., D. G. Higgins, and A. C. Culhane, 2006: Comparison and evaluation of methods for generating differentially expressed gene lists from microarray data. *BMC bioinformatics*, **7**, 359. 32

Jellinger, K. A., 2009: Recent advances in our understanding of neurodegeneration. *Journal of neural transmission (Vienna, Austria : 1996)*, **116**, 1111–62. 3, 5, 79, 111, 128

—, 2010: Basic mechanisms of neurodegeneration: a critical update. *Journal of cellular and molecular medicine*, **14**, 457–87. 2, 5, 12, 13, 79, 128

Jin, K., M. LaFevre-Bernt, Y. Sun, S. Chen, J. Gafni, D. Crippen, A. Logvinova, C. A. Ross, D. A. Greenberg, and L. M. Ellerby, 2005: FGF-2 promotes neurogenesis and neuroprotection and prolongs survival in a transgenic mouse model of Huntington's disease. *Proceedings of the National Academy of Sciences of the United States of America*, **102**, 18189–94. 15

Jin, K., Y. Sun, L. Xie, S. Batteur, X. O. Mao, C. Smelick, A. Logvinova, and D. A. Greenberg, 2003: Neurogenesis and aging: FGF-2 and HB-EGF restore neurogenesis in hippocampus and subventricular zone of aged mice. *Aging Cell*, **2**, 175–183. 149

Johnson, K. A., P. J. Conn, and C. M. Niswender, 2009: Glutamate receptors as therapeutic targets for Parkinson's disease. *CNS & neurological disorders drug targets*, **8**, 475–91. 66

Jomova, K., D. Vondrakova, M. Lawson, and M. Valko, 2010: Metals, oxidative stress and neurodegenerative disorders. *Molecular and cellular biochemistry*, **345**, 91–104. 12, 84

Jordan-Sciutto, K. L., L. M. Malaiyandi, and R. Bowser, 2002: Altered distribution of cell cycle transcriptional regulators during Alzheimer disease. *Journal of neuropathology and experimental neurology*, **61**, 358–67. 85

Jucker, M., 2010: The benefits and limitations of animal models for translational research in neurodegenerative diseases. *Nature medicine*, **16**, 1210–4. 158

Junn, E. and M. M. Mouradian, 2012: MicroRNAs in neurodegenerative diseases and their therapeutic potential. *Pharmacology & therapeutics*, **133**, 142–50. 122, 140

Kalinowska, A. and J. Losy, 2006: PECAM-1, a key player in neuroinflammation. *European journal of neurology : the official journal of the European Federation of Neurological Societies*, **13**, 1284–90. 9

Kanehisa, M. and S. Goto, 2000: KEGG: kyoto encyclopedia of genes and genomes. *Nucleic acids research*, **28**, 27–30. 45, 62

Kang, H., L. D. Sun, C. M. Atkins, T. R. Soderling, M. A. Wilson, and S. Tonegawa, 2001: An important role of neural activity-dependent CaMKIV signaling in the consolidation of long-term memory. *Cell*, **106**, 771–83. 79

Kauffmann, A., R. Gentleman, and W. Huber, 2009: arrayQualityMetrics–a bioconductor package for quality assessment of microarray data. *Bioinformatics (Oxford, England)*, **25**, 415–6. 31

Kerr, F., A. Rickle, N. Nayeem, S. Brandner, R. F. Cowburn, and S. Lovestone, 2006: PTEN, a negative regulator of PI3 kinase signalling, alters tau phosphorylation in cells by mechanisms independent of GSK-3. *FEBS letters*, **580**, 3121–8. 61, 148

Kiechle, T., A. Dedeoglu, J. Kubilus, N. W. Kowall, M. F. Beal, R. M. Friedlander, S. M. Hersch, and R. J. Ferrante, 2002: Cytochrome C and caspase-9 expression in Huntington's disease. *Neuromolecular medicine*, **1**, 183–95. 15

Kim, E. K. and E.-J. Choi, 2010: Pathological roles of MAPK signaling pathways in human diseases. *Biochimica et biophysica acta*, **1802**, 396–405. 7

Klein, C. and M. G. Schlossmacher, 2006: The genetics of Parkinson disease: Implications for neurological care. *Nature clinical practice. Neurology*, **2**, 136–46. 5, 6, 79

Kooperberg, C., A. Aragaki, A. D. Strand, and J. M. Olson, 2005: Significance testing for small microarray experiments. *Statistics in medicine*, **24**, 2281–98. 32

Koschützki, D. and F. Schreiber, 2008: Centrality analysis methods for biological networks and their application to gene regulatory networks. *Gene regulation and systems biology*, **2**, 193–201. 157

Krieglstein, K., J. Strelau, A. Schober, A. Sullivan, and K. Unsicker, 2002: TGF-beta and the regulation of neuron survival and death. *Journal of physiology, Paris*, **96**, 25–30. 149

La Spada, A. R., 2005: Huntington's disease and neurogenesis: FGF-2 to the rescue? *Proceedings of the National Academy of Sciences of the United States of America*, **102**, 17889–90. 15

Lai, S. K., Y. C. Tse, M. S. Yang, C. K. C. Wong, Y. S. Chan, and K. K. L. Yung, 2003: Gene expression of glutamate receptors GluR1 and NR1 is differentially modulated in striatal neurons in rats after 6-hydroxydopamine lesion. *Neurochemistry international*, **43**, 639–53. 52

Lalić, N. M., J. Marić, M. Svetel, A. Jotić, E. Stefanova, K. Lalić, N. Dragasević, T. Milić, L. Lukić, and V. S. Kostić, 2008: Glucose homeostasis in Huntington disease: abnormalities in insulin sensitivity and early-phase insulin secretion. *Archives of neurology*, **65**, 476–80. 114

Langui, D., F. Lachapelle, and C. Duyckaerts, 2007: Animal models of neurodegenerative diseases. *Médecine sciences : M/S*, **23**, 180–6. 158

Larbi, A., G. Pawelec, J. M. Witkowski, H. M. Schipper, E. Derhovanessian, D. Goldeck, and T. Fulop, 2009: Dramatic shifts in circulating CD4 but not CD8 T cell subsets in mild Alzheimer's disease. *Journal of Alzheimer's disease : JAD*, **17**, 91–103. 90

Lau, P. and B. de Strooper, 2010: Dysregulated microRNAs in neurodegenerative disorders. *Seminars in cell & developmental biology*, **21**, 768–73. 122

Le Hellard, S., B. Hå vik, T. Espeseth, H. Breilid, R. Lø vlie, M. Luciano, A. J. Gow, S. E. Harris, J. M. Starr, K. Wibrand, A. J. Lundervold, D. J. Porteous, C. R. Bramham, I. J. Deary, I. Reinvang, and V. M. Steen, 2009: Variants in doublecortin-

and calmodulin kinase like 1, a gene up-regulated by BDNF, are associated with memory and general cognitive abilities. *PloS one*, **4**, e7534. 139

Lee, D. Y., K. W. Park, and B. K. Jin, 2006: Thrombin induces neurodegeneration and microglial activation in the cortex in vivo and in vitro: proteolytic and non-proteolytic actions. *Biochemical and biophysical research communications*, **346**, 727–38. 147

Lei, P., S. Ayton, D. I. Finkelstein, P. A. Adlard, C. L. Masters, and A. I. Bush, 2010: Tau protein: relevance to Parkinson's disease. *The international journal of biochemistry & cell biology*, **42**, 1775–8. 61, 111

L'episcopo, F., M. F. Serapide, C. Tirolo, N. Testa, S. Caniglia, M. C. Morale, S. Pluchino, and B. Marchetti, 2011: A Wnt1 regulated Frizzled-1/β-Catenin signaling pathway as a candidate regulatory circuit controlling mesencephalic dopaminergic neuron-astrocyte crosstalk: Therapeutical relevance for neuron survival and neuroprotection. *Molecular neurodegeneration*, **6**, 49. 59

Lesage, S. and A. Brice, 2009: Parkinson's disease: from monogenic forms to genetic susceptibility factors. *Human molecular genetics*, **18**, R48–59. 3, 4

Leuba, G., A. Vernay, D. Vu, C. Walzer, B. Belloir, R. Kraftsik, C. Bouras, and A. Savioz, 2004: Differential expression of LMO4 protein in Alzheimer's disease. *Neuropathology and applied neurobiology*, **30**, 57–69. 10

Levine, B., N. Mizushima, and H. W. Virgin, 2011: Autophagy in immunity and inflammation. *Nature*, **469**, 323–35. 118

Liang, C., H. Zhu, Y. Xu, L. Huang, C. Ma, W. Deng, Y. Liu, and C. Qin, 2012: MicroRNA-153 negatively regulates the expression of amyloid precursor protein and amyloid precursor-like protein 2. *Brain research*, **1455**, 103–13. 93

Lillien, L. and H. Raphael, 2000: BMP and FGF regulate the development of EGF-responsive neural progenitor cells. *Development (Cambridge, England)*, **127**, 4993–5005. 59

Liu, Q., F. Xie, A. Alvarado-Diaz, M. A. Smith, P. I. Moreira, X. Zhu, and G. Perry, 2011: Neurofilamentopathy in neurodegenerative diseases. *The open neurology journal*, **5**, 58–62. 128

Liu, Y. F., R. C. Deth, and D. Devys, 1997: SH3 domain-dependent association of huntingtin with epidermal growth factor receptor signaling complexes. *The Journal of biological chemistry*, **272**, 8121–4. 111

Loscalzo, J., I. Kohane, and A.-L. Barabasi, 2007: Human disease classification in the postgenomic era: a complex systems approach to human pathobiology. *Molecular systems biology*, **3**, 124. 19

Lukiw, W. J., 2007: Micro-RNA speciation in fetal, adult and Alzheimer's disease hippocampus. *Neuroreport*, **18**, 297–300. 71, 122

Luo, J., F. Elwood, M. Britschgi, S. Villeda, H. Zhang, Z. Ding, L. Zhu, H. Alabsi, R. Getachew, R. Narasimhan, R. Wabl, N. Fainberg, M. L. James, G. Wong, J. Relton, S. S. Gambhir, J. W. Pollard, and T. Wyss-Coray, 2013: Colony-stimulating factor 1 receptor (CSF1R) signaling in injured neurons facilitates protection and survival. *The Journal of experimental medicine*, **210**, 157–72. 79

Majed, H. H., S. Chandran, S. P. Niclou, R. S. Nicholas, A. Wilkins, M. G. Wing, K. E. Rhodes, M. G. Spillantini, and A. Compston, 2006: A novel role for Sema3A in neuroprotection from injury mediated by activated microglia. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, **26**, 1730–8. 67

Managbanag, J. R., T. M. Witten, D. Bonchev, L. A. Fox, M. Tsuchiya, B. K. Kennedy, and M. Kaeberlein, 2008: Shortest-path network analysis is a useful approach toward identifying genetic determinants of longevity. *PloS one*, **3**, e3802. 40, 157

Manczak, M., P. Mao, M. J. Calkins, A. Cornea, A. P. Reddy, M. P. Murphy, H. H. Szeto, B. Park, and P. H. Reddy, 2010: Mitochondria-targeted antioxidants protect against amyloid-beta toxicity in Alzheimer's disease neurons. *Journal of Alzheimer's disease : JAD*, **20 Suppl 2**, S609–31. 12

Mantamadiotis, T., T. Lemberger, S. C. Bleckmann, H. Kern, O. Kretz, A. Martin Villalba, F. Tronche, C. Kellendonk, D. Gau, J. Kapfhammer, C. Otto, W. Schmid, and G. Schütz, 2002: Disruption of CREB function in brain leads to neurodegeneration. *Nature genetics*, **31**, 47–54. 61, 85

Martinez-Vicente, M., Z. Talloczy, E. Wong, G. Tang, H. Koga, S. Kaushik, R. de Vries, E. Arias, S. Harris, D. Sulzer, and A. M. Cuervo, 2010: Cargo recognition failure is responsible for inefficient autophagy in Huntington's disease. *Nature neuroscience*, **13**, 567–76. 13

Martins, M., A. Rosa, L. C. Guedes, B. V. Fonseca, K. Gotovac, S. Violante, T. Mestre, M. Coelho, M. M. Rosa, E. R. Martin, J. M. Vance, T. F. Outeiro, L. Wang, F. Borovecki, J. J. Ferreira, and S. A. Oliveira, 2011: Convergence of miRNA expression profiling, α-synuclein interacton and GWAS in Parkinson's disease. *PloS one*, **6**, e25443. 6

Mateo, I., J. L. Vázquez-Higuera, P. Sánchez-Juan, E. Rodríguez-Rodríguez, J. Infante, I. García-Gorostiaga, J. Berciano, and O. Combarros, 2009: Epistasis between tau phosphorylation regulating genes (CDK5R1 and GSK-3beta) and Alzheimer's disease risk. *Acta neurologica Scandinavica*, **120**, 130–3. 10, 110, 132

Matigian, N., G. Abrahamsen, R. Sutharsan, A. L. Cook, A. M. Vitale, A. Nouwens, B. Bellette, J. An, M. Anderson, A. G. Beckhouse, M. Bennebroek, R. Cecil, A. M. Chalk, J. Cochrane, Y. Fan, F. Féron, R. McCurdy, J. J. McGrath, W. Murrell, C. Perry, J. Raju, S. Ravishankar, P. A. Silburn, G. T. Sutherland, S. Mahler, G. D. Mellick, S. A. Wood, C. M. Sue, C. A. Wells, and A. Mackay-Sim, 2010: Disease-specific, neurosphere-derived cells as models for brain disorders. *Disease models & mechanisms*, **3**, 785–98. 55

Mattson, M. P., 2000: Apoptosis in neurodegenerative disorders. *Nature reviews. Molecular cell biology*, **1**, 120–9. 15, 128

Mattson, M. P. and S. L. Chan, 2003: Neuronal and glial calcium signaling in Alzheimer's disease. *Cell calcium*, **34**, 385–97. 87

Mazzola, J. L. and M. A. Sirover, 2001: Reduction of glyceraldehyde-3-phosphate dehydrogenase activity in Alzheimer's disease and in Huntington's disease fibroblasts. *Journal of neurochemistry*, **76**, 442–9. 15

Mehta, J. P. and S. Rani, 2011: Software and tools for microarray data analysis. *Methods in molecular biology (Clifton, N.J.)*, **784**, 41–53. 30

Meriin, A. B. and M. Y. Sherman, 2005: Role of molecular chaperones in neurodegenerative disorders. *International journal of hyperthermia : the official journal of European Society for Hyperthermic Oncology, North American Hyperthermia Group*, **21**, 403–19. 61

Miñones Moyano, E., S. Porta, G. Escaramís, R. Rabionet, S. Iraola, B. Kagerbauer, Y. Espinosa-Parrilla, I. Ferrer, X. Estivill, and E. Martí, 2011: MicroRNA profiling of Parkinson's disease brains identifies early downregulation of miR-34b/c which modulate mitochondrial function. *Human molecular genetics*, **20**, 3067–78. 70, 74, 140

Michael, G. J., S. Esmailzadeh, L. B. Moran, L. Christian, R. K. B. Pearce, and M. B. Graeber, 2011: Up-regulation of metallothionein gene expression in parkinsonian astrocytes. *Neurogenetics*, **12**, 295–305. 7

Micieli, G., P. Tosi, S. Marcheselli, and A. Cavallini, 2003: Autonomic dysfunction in Parkinson's disease. *Neurological sciences : official journal of the Italian Neurological Society and of the Italian Society of Clinical Neurophysiology*, **24 Suppl 1**, S32–4. 3

Mirnics, Z. K., C. Yan, C. Portugal, T.-W. Kim, H. U. Saragovi, S. S. Sisodia, K. Mirnics, and N. F. Schor, 2005: P75 neurotrophin receptor regulates expression of neural cell adhesion molecule 1. *Neurobiology of disease*, **20**, 969–85. 50

Möller, T., 2010: Neuroinflammation in Huntington's disease. *Journal of neural transmission (Vienna, Austria : 1996)*, **117**, 1001–8. 117, 128

Molochnikov, L., J. M. Rabey, E. Dobronevsky, U. Bonucelli, R. Ceravolo, D. Frosini, E. Grünblatt, P. Riederer, C. Jacob, J. Aharon-Peretz, Y. Bashenko, M. B. H. Youdim, and S. A. Mandel, 2012: A molecular signature in blood identifies early Parkinson's disease. *Molecular neurodegeneration*, **7**, 26. 159

Moncini, S., A. Salvi, P. Zuccotti, G. Viero, A. Quattrone, S. Barlati, G. De Petro, M. Venturin, and P. Riva, 2011: The role of miR-103 and miR-107 in regulation of CDK5R1 expression and in cellular migration. *PloS one*, **6**, e20038. 97

Moore, R. C., F. Xiang, J. Monaghan, D. Han, Z. Zhang, L. Edström, M. Anvret, and S. B. Prusiner, 2001: Huntington disease phenocopy is a familial prion disease. *American journal of human genetics*, **69**, 1385–8. 107

Moran, L. B., D. C. Duke, M. Deprez, D. T. Dexter, R. K. B. Pearce, and M. B. Graeber, 2006: Whole genome expression profiling of the medial and lateral substantia nigra in Parkinson's disease. *Neurogenetics*, **7**, 1–11. 47, 72

Moran, L. B. and M. B. Graeber, 2008: Towards a pathway definition of Parkinson's disease: a complex disorder with links to cancer, diabetes and inflammation. *Neurogenetics*, **9**, 1–13. 20, 47

Moreira Sousa, C. a., J. R. McGuire, M. S. Thion, D. Gentien, P. de la Grange, S. Tezenas du Montcel, A. Vincent-Salomon, A. Durr, and S. Humbert, 2013: The Huntington disease protein accelerates breast tumour development and metastasis through ErbB2/HER2 signalling. *EMBO molecular medicine*, **5**, 309–25. 114

Morganti, J. M., K. R. Nash, B. A. Grimmig, S. Ranjit, B. Small, P. C. Bickford, and C. Gemma, 2012: The soluble isoform of CX3CL1 is necessary for neuroprotection in a mouse model of Parkinson's disease. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, **32**, 14592–601. 4

Mouradian, M. M., 2012: MicroRNAs in Parkinson's disease. *Neurobiology of disease*, **46**, 279–84. 70, 74

Müller, T., D. Blum-Degen, H. Przuntek, and W. Kuhn, 1998: Interleukin-6 levels in cerebrospinal fluid inversely correlate to severity of Parkinson's disease. *Acta neurologica Scandinavica*, **98**, 142–4. 67

Murie, C., O. Woody, A. Y. Lee, and R. Nadon, 2009: Comparison of small n statistical tests of differential expression applied to microarrays. *BMC bioinformatics*, **10**, 45. 32

Mymrikov, E. V., A. S. Seit-Nebi, and N. B. Gusev, 2011: Large potentials of small heat shock proteins. *Physiological reviews*, **91**, 1123–59. 111

Nagahara, A. H., D. A. Merrill, G. Coppola, S. Tsukada, B. E. Schroeder, G. M. Shaked, L. Wang, A. Blesch, A. Kim, J. M. Conner, E. Rockenstein, M. V. Chao, E. H. Koo, D. Geschwind, E. Masliah, A. A. Chiba, and M. H. Tuszynski, 2009: Neuroprotective effects of brain-derived neurotrophic factor in rodent and primate models of Alzheimer's disease. *Nature medicine*, **15**, 331–7. 12

Nakamura, T., H. Yamashita, T. Takahashi, and S. Nakamura, 2001: Activated Fyn phosphorylates alpha-synuclein at tyrosine residue 125. *Biochemical and biophysical research communications*, **280**, 1085–92. 64

Niikura, T., T. Chiba, S. Aiso, M. Matsuoka, and I. Nishimoto, 2004: Humanin: after the discovery. *Molecular neurobiology*, **30**, 327–40. 9

Nikitin, A., S. Egorov, N. Daraselia, and I. Mazo, 2003: Pathway studio–the analysis and navigation of molecular networks. *Bioinformatics (Oxford, England)*, **19**, 2155–7. 36

Nikolić, M., 2008: The Pak1 kinase: an important regulator of neuronal morphology and function in the developing forebrain. *Molecular neurobiology*, **37**, 187–202. 59

Nunez-Iglesias, J., C.-C. Liu, T. E. Morgan, C. E. Finch, and X. J. Zhou, 2010: Joint genome-wide profiling of miRNA and mRNA expression in Alzheimer's disease cortex reveals altered miRNA regulation. *PloS one*, **5**, e8898. 10

O'Brien, R. J. and P. C. Wong, 2011: Amyloid precursor protein processing and Alzheimer's disease. *Annual review of neuroscience*, **34**, 185–204. 60

O'Day, D. H. and M. A. Myre, 2004: Calmodulin-binding domains in Alzheimer's disease proteins: extending the calcium hypothesis. *Biochemical and biophysical research communications*, **320**, 1051–4. 137

Okun, M. S., 2012: Deep-Brain Stimulation for Parkinson's Disease. *New England Journal of Medicine*, **367**, 1529–1538. 8

Ozgür, A., T. Vu, G. Erkan, and D. R. Radev, 2008: Identifying gene-disease associations using centrality on a literature mined gene-interaction network. *Bioinformatics (Oxford, England)*, **24**, i277–85. 157

Pabon, M. M., A. D. Bachstetter, C. E. Hudson, C. Gemma, and P. C. Bickford, 2011: CX3CL1 reduces neurotoxicity and microglial activation in a rat model of Parkinson's disease. *Journal of neuroinflammation*, **8**, 9. 7, 66

Packer, A. N., Y. Xing, S. Q. Harper, L. Jones, and B. L. Davidson, 2008: The bifunctional microRNA miR-9/miR-9* regulates REST and CoREST and is downregulated in Huntington's disease. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, **28**, 14341–6. 122

Parkinson, J., 2002: An essay on the shaking palsy. 1817. *The Journal of neuropsychiatry and clinical neurosciences*, **14**, 223–36; discussion 222. 3

Pezzolesi, M. G., P. Platzer, K. A. Waite, and C. Eng, 2008: Differential expression of PTEN-targeting microRNAs miR-19a and miR-21 in Cowden syndrome. *American journal of human genetics*, **82**, 1141–9. 123

Phan, J., M. A. Hickey, P. Zhang, M.-F. Chesselet, and K. Reue, 2009: Adipose tissue dysfunction tracks disease progression in two Huntington's disease mouse models. *Human molecular genetics*, **18**, 1006–16. 114

Phillips, W., A. Michell, H. Pruess, and R. A. Barker, 2009: Animal models of neurodegenerative diseases. *Methods in molecular biology (Clifton, N.J.)*, **549**, 137–55. 158

Pittman, A. M., H.-C. Fung, and R. de Silva, 2006: Untangling the tau gene association with neurodegenerative disorders. *Human molecular genetics*, **15 Spec No**, R188–95. 61, 111

Pollak, D., N. Cairns, and G. Lubec, 2003: Cytoskeleton derangement in brain of patients with Down syndrome, Alzheimer's disease and Pick's disease. *Journal of neural transmission. Supplementum*, 149–58. 79

Polymeropoulos, M. H., C. Lavedan, E. Leroy, S. E. Ide, A. Dehejia, A. Dutra, B. Pike, H. Root, J. Rubenstein, R. Boyer, E. S. Stenroos, S. Chandrasekharappa, A. Athanassiadou, T. Papapetropoulos, W. G. Johnson, A. M. Lazzarini, R. C. Duvoisin, G. Di Iorio, L. I. Golbe, and R. L. Nussbaum, 1997: Mutation in the alpha-synuclein gene identified in families with Parkinson's disease. *Science (New York, N.Y.)*, **276**, 2045–7. 5

Puglielli, L., 2008: Aging of the brain, neurotrophin signaling, and Alzheimer's disease: is IGF1-R the common culprit? *Neurobiology of aging*, **29**, 795–811. 87

Qin, Q., G. Liao, M. Baudry, and X. Bi, 2010: Cholesterol Perturbation in Mice Results in p53 Degradation and Axonal Pathology through p38 MAPK and Mdm2 Activation. *PloS one*, **5**, e9999. 83

Qin, S., X.-Y. Hu, H. Xu, and J.-N. Zhou, 2004: Regional alteration of synapsin I in the hippocampal formation of Alzheimer's disease patients. *Acta neuropathologica*, **107**, 209–15. 79

Rappaport, N., N. Nativ, G. Stelzer, M. Twik, Y. Guan-Golan, T. Iny Stein, I. Bahir, F. Belinky, C. P. Morrey, M. Safran, and D. Lancet, 2013: MalaCards: an integrated compendium for diseases and their annotation. *Database : the journal of biological databases and curation*, **2013**, bat018. 130

Rascol, O., A. Lozano, M. Stern, and W. Poewe, 2011: Milestones in Parkinson's disease therapeutics. *Movement disorders : official journal of the Movement Disorder Society*, **26**, 1072–82. 6

Raux, G., L. Guyant-Maréchal, C. Martin, J. Bou, C. Penet, A. Brice, D. Hannequin, T. Frebourg, and D. Campion, 2005: Molecular diagnosis of autosomal dominant early onset Alzheimer's disease: an update. *Journal of medical genetics*, **42**, 793–5. 9

Ravikumar, B., C. Vacher, Z. Berger, J. E. Davies, S. Luo, L. G. Oroz, F. Scaravilli, D. F. Easton, R. Duden, C. J. O'Kane, and D. C. Rubinsztein, 2004: Inhibition of mTOR induces autophagy and reduces toxicity of polyglutamine expansions in fly and mouse models of Huntington disease. *Nature genetics*, **36**, 585–95. 118

Recchia, A., P. Debetto, A. Negro, D. Guidolin, S. D. Skaper, and P. Giusti, 2004: Alpha-synuclein and Parkinson's disease. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology*, **18**, 617–26. 4

Reiner, A., I. Dragatsis, and P. Dietrich, 2011: Genetics and neuropathology of Huntington's disease. *International review of neurobiology*, **98**, 325–72. 13

Reis, S. A., M. N. Thompson, J.-M. Lee, E. Fossale, H.-H. Kim, J. K. Liao, M. A. Moskowitz, S. Y. Shaw, L. Dong, S. J. Haggarty, M. E. MacDonald, and I. S. Seong, 2011: Striatal neurons expressing full-length mutant huntingtin exhibit decreased N-cadherin and altered neuritogenesis. *Human molecular genetics*, **20**, 2344–55. 113

Rentzos, M., C. Nikolaou, E. Andreadou, G. P. Paraskevas, A. Rombos, M. Zoga, A. Tsoutsou, F. Boufidou, E. Kapaki, and D. Vassilopoulos, 2009: Circulating interleukin-10 and interleukin-12 in Parkinson's disease. *Acta neurologica Scandinavica*, **119**, 332–7. 67

Ristow, M., 2004: Neurodegenerative disorders associated with diabetes mellitus. *Journal of molecular medicine (Berlin, Germany)*, **82**, 510–29. 142, 146

Roos, R. A., M. Vegter-van der Vlis, J. Hermans, H. M. Elshove, A. C. Moll, J. J. van de Kamp, and G. W. Bruyn, 1991: Age at onset in Huntington's disease: effect of line of inheritance and patient's sex. *Journal of Medical Genetics*, **28**, 515–519. 111

Ross, C. A. and S. J. Tabrizi, 2011: Huntington's disease: from molecular pathogenesis to clinical treatment. *Lancet neurology*, **10**, 83–98. 16

Rothermundt, M., M. Peters, J. H. M. Prehn, and V. Arolt, 2003: S100B in brain damage and neurodegeneration. *Microscopy research and technique*, **60**, 614–32. 139

Sõber, S., M. Laan, and T. Annilo, 2010: MicroRNAs miR-124 and miR-135a are potential regulators of the mineralocorticoid receptor gene (NR3C2) expression. *Biochemical and biophysical research communications*, **391**, 727–32. 123

Sadagurski, M., Z. Cheng, A. Rozzo, I. Palazzolo, G. R. Kelley, X. Dong, D. Krainc, and M. F. White, 2011: IRS2 increases mitochondrial dysfunction and oxidative stress in a mouse model of Huntington disease. *The Journal of clinical investigation*, **121**, 4070–81. 15

Salim, C., Y. V. Boxberg, J. Alterio, S. Féréol, and F. Nothias, 2009: The giant protein AHNAK involved in morphogenesis and laminin substrate adhesion of myelinating Schwann cells. *Glia*, **57**, 535–49. 139

Samantaray, S., S. K. Ray, and N. L. Banik, 2008: Calpain as a potential therapeutic target in Parkinson's disease. *CNS & neurological disorders drug targets*, **7**, 305–12. 67

Sanchez-Mejia, R. O., J. W. Newman, S. Toh, G.-Q. Yu, Y. Zhou, B. Halabisky, M. Cissé, K. Scearce-Levie, I. H. Cheng, L. Gan, J. J. Palop, J. V. Bonventre, and L. Mucke, 2008: Phospholipase A2 reduction ameliorates cognitive deficits in a mouse model of Alzheimer's disease. *Nature neuroscience*, **11**, 1311–8. 146

Schon, E. A. and G. Manfredi, 2003: Neuronal degeneration and mitochondrial dysfunction. *The Journal of clinical investigation*, **111**, 303–12. 2, 128

Schubert, M., D. Gautam, D. Surjo, K. Ueki, S. Baudler, D. Schubert, T. Kondo, J. Alber, N. Galldiks, E. Küstermann, S. Arndt, A. H. Jacobs, W. Krone, C. R. Kahn, and J. C. Brüning, 2004: Role for neuronal insulin resistance in neurodegenerative diseases. *Proceedings of the National Academy of Sciences of the United States of America*, **101**, 3100–5. 147

Schwikowski, B., P. Uetz, and S. Fields, 2000: A network of protein-protein interactions in yeast. *Nature biotechnology*, **18**, 1257–61. 40, 157

Sedelis, M., R. K. Schwarting, and J. P. Huston, 2001: Behavioral phenotyping of the MPTP mouse model of Parkinson's disease. *Behavioural brain research*, **125**, 109–25. 3

Segura, M. F., D. Hanniford, S. Menendez, L. Reavie, X. Zou, S. Alvarez-Diaz, J. Zakrzewski, E. Blochin, A. Rose, D. Bogunovic, D. Polsky, J. Wei, P. Lee, I. Belitskaya-Levy, N. Bhardwaj, I. Osman, and E. Hernando, 2009: Aberrant miR-182 expres-

sion promotes melanoma metastasis by repressing FOXO3 and microphthalmia-associated transcription factor. *Proceedings of the National Academy of Sciences of the United States of America*, **106**, 1814–9. 123

Selkoe, D. J., 2001: Alzheimer's Disease: Genes, Proteins, and Therapy. *Physiol Rev*, **81**, 741–766. 9, 10, 12

Serretti, A., P. Artioli, R. Quartesan, and D. De Ronchi, 2005: Genes involved in Alzheimer's disease, a survey of possible candidates. *Journal of Alzheimer's disease : JAD*, **7**, 331–53. 10

Servitja, J.-M., R. Masgrau, R. Pardo, E. Sarri, and F. Picatoste, 2002: Effects of Oxidative Stress on Phospholipid Signaling in Rat Cultured Astrocytes and Brain Slices. *Journal of Neurochemistry*, **75**, 788–794. 146

Shacka, J. J. and K. A. Roth, 2005: Regulation of neuronal cell death and neurodegeneration by members of the Bcl-2 family: therapeutic implications. *Current drug targets. CNS and neurological disorders*, **4**, 25–39. 134

Sharif, A. and V. Prevot, 2010: ErbB receptor signaling in astrocytes: a mediator of neuron-glia communication in the mature central nervous system. *Neurochemistry international*, **57**, 344–58. 55

Shulman, J. M., P. L. De Jager, and M. B. Feany, 2011: Parkinson's disease: genetics and pathogenesis. *Annual review of pathology*, **6**, 193–222. 5

Simunovic, F., M. Yi, Y. Wang, L. Macey, L. T. Brown, A. M. Krichevsky, S. L. Andersen, R. M. Stephens, F. M. Benes, and K. C. Sonntag, 2009: Gene expression profiling of substantia nigra dopamine neurons: further insights into Parkinson's disease pathology. *Brain : a journal of neurology*, **132**, 1795–809. 70

Smith, M. Z., Z. Nagy, and M. M. Esiri, 1999: Cell cycle-related protein expression in vascular dementia and Alzheimer's disease. *Neuroscience letters*, **271**, 45–8. 85

Smyth, G. K., 2004: Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Statistical applications in genetics and molecular biology*, **3**, Article3. 32

Snow, A. D., H. Mar, D. Nochlin, H. Kresse, and T. N. Wight, 1992: Peripheral distribution of dermatan sulfate proteoglycans (decorin) in amyloid-containing plaques and their presence in neurofibrillary tangles of Alzheimer's disease. *The journal of histochemistry and cytochemistry : official journal of the Histochemistry Society*, **40**, 105–13. 90

Sofroniew, M. V., C. L. Howe, and W. C. Mobley, 2001: Nerve growth factor signaling, neuroprotection, and neural repair. *Annual review of neuroscience*, **24**, 1217–81. 132, 149

Solerte, S. B., M. Fioravanti, A. Pascale, E. Ferrari, S. Govoni, and F. Battaini, 1998: Increased natural killer cell cytotoxicity in Alzheimer's disease may involve protein kinase C dysregulation. *Neurobiology of aging*, **19**, 191–9. 87

Sonntag, K.-C., 2010: MicroRNAs and deregulated gene expression networks in neurodegeneration. *Brain research*, **1338**, 48–57. 124

Sonoda, Y., H. Mukai, K. Matsuo, M. Takahashi, Y. Ono, K. Maeda, H. Akiyama, and T. Kawamata, 2010: Accumulation of tumor-suppressor PTEN in Alzheimer neurofibrillary tangles. *Neuroscience letters*, **471**, 20–4. 10, 61, 123, 132

Spillantini, M. G. and M. Goedert, 2000: The alpha-synucleinopathies: Parkinson's disease, dementia with Lewy bodies, and multiple system atrophy. *Annals of the New York Academy of Sciences*, **920**, 16–27. 60

—, 2001: Tau and Parkinson Disease. *JAMA: The Journal of the American Medical Association*, **286**, 2324–2326. 61

Stefanis, L., 2012: α-Synuclein in Parkinson's disease. *Cold Spring Harbor perspectives in medicine*, **2**, a009399. 60

Steinacker, P., A. Aitken, and M. Otto, 2011: 14-3-3 proteins in neurodegeneration. *Seminars in cell & developmental biology*, **22**, 696–704. 58

Stelzl, U., U. Worm, M. Lalowski, C. Haenig, F. H. Brembeck, H. Goehler, M. Stroedicke, M. Zenkner, A. Schoenherr, S. Koeppen, J. Timm, S. Mintzlaff, C. Abraham, N. Bock, S. Kietzmann, A. Goedde, E. Toksöz, A. Droege, S. Krobitsch, B. Korn, W. Birchmeier, H. Lehrach, and E. E. Wanker, 2005: A human protein-protein interaction network: a resource for annotating the proteome. *Cell*, **122**, 957–68. 18

Takasugi, N., T. Sasaki, K. Suzuki, S. Osawa, H. Isshiki, Y. Hori, N. Shimada, T. Higo, S. Yokoshima, T. Fukuyama, V. M.-Y. Lee, J. Q. Trojanowski, T. Tomita, and T. Iwatsubo, 2011: BACE1 activity is modulated by cell-associated sphingosine-1-phosphate. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, **31**, 6850–7. 99

Tamariz, E., N. E. Díaz-Martínez, N. F. Díaz, C. M. García-Peña, I. Velasco, and A. Varela-Echavarría, 2010: Axon responses of embryonic stem cell-derived dopaminergic neurons to semaphorins 3A and 3C. *Journal of neuroscience research*, **88**, 971–80. 67

Tang, B. L., 2008: Leptin as a neuroprotective agent. *Biochemical and biophysical research communications*, **368**, 181–5. 60, 110

Tapia-Arancibia, L., E. Aliaga, M. Silhol, and S. Arancibia, 2008: New insights into brain BDNF function in normal aging and Alzheimer disease. *Brain research reviews*, **59**, 201–20. 16, 133, 149

Taylor, I. W., R. Linding, D. Warde-Farley, Y. Liu, C. Pesquita, D. Faria, S. Bull, T. Pawson, Q. Morris, and J. L. Wrana, 2009: Dynamic modularity in protein interaction networks predicts breast cancer outcome. *Nature biotechnology*, **27**, 199–204. 17

Tesseur, I., K. Zou, L. Esposito, F. Bard, E. Berber, J. V. Can, A. H. Lin, L. Crews, P. Tremblay, P. Mathews, L. Mucke, E. Masliah, and T. Wyss-Coray, 2006: Deficiency in neuronal TGF-beta signaling promotes neurodegeneration and Alzheimer's pathology. *The Journal of clinical investigation*, **116**, 3060–9. 149

Thanvi, B., N. Lo, and T. Robinson, 2007: Levodopa-induced dyskinesia in Parkinson's disease: clinical features, pathogenesis, prevention and treatment. *Postgraduate medical journal*, **83**, 384–8. 8

Tian, Q., S. B. Stepaniants, M. Mao, L. Weng, M. C. Feetham, M. J. Doyle, E. C. Yi, H. Dai, V. Thorsson, J. Eng, D. Goodlett, J. P. Berger, B. Gunter, P. S. Linseley, R. B. Stoughton, R. Aebersold, S. J. Collins, W. A. Hanlon, and L. E. Hood, 2004: Integrated genomic and proteomic analyses of gene expression in Mammalian cells. *Molecular & cellular proteomics : MCP*, **3**, 960–9. 26, 158

Tichauer, J. E. and R. von Bernhardi, 2012: Transforming growth factor-β stimulates β amyloid uptake by microglia through Smad3-dependent mechanisms. *Journal of neuroscience research*, **90**, 1970–80. 9

Town, T., Y. Laouar, C. Pittenger, T. Mori, C. A. Szekely, J. Tan, R. S. Duman, and R. A. Flavell, 2008: Blocking TGF-beta-Smad2/3 innate immune signaling mitigates Alzheimer-like pathology. *Nature medicine*, **14**, 681–7. 87

Tsang, S. W. Y., H. V. Vinters, J. L. Cummings, P. T.-H. Wong, C. P. L.-H. Chen, and M. K. P. Lai, 2008: Alterations in NMDA receptor subunit densities and ligand binding to glycine recognition sites are associated with chronic anxiety in Alzheimer's disease. *Neurobiology of aging*, **29**, 1524–32. 79

Tuppo, E. E. and H. R. Arias, 2005: The role of inflammation in Alzheimer's disease. *The international journal of biochemistry & cell biology*, **37**, 289–305. 12

Ueberham, U., E. Ueberham, H. Gruschka, and T. Arendt, 2006: Altered subcellular location of phosphorylated Smads in Alzheimer's disease. *The European journal of neuroscience*, **24**, 2327–34. 9

Uetz, P., Y.-A. Dong, C. Zeretzke, C. Atzler, A. Baiker, B. Berger, S. V. Rajagopala, M. Roupelieva, D. Rose, E. Fossum, and J. Haas, 2006: Herpesviral protein networks and their interaction with the human proteome. *Science (New York, N.Y.)*, **311**, 239–42. 17

Umegaki, H., 2012: Neurodegeneration in diabetes mellitus. *Advances in experimental medicine and biology*, **724**, 258–65. 142, 146

van Rooij, E., A. L. Purcell, and A. A. Levin, 2012: Developing microRNA therapeutics. *Circulation research*, **110**, 496–507. 6, 92

Vladimir Kuznetsov, S. T. and D. Bonchev, 2008: Data-driven Networking Reveals 5-Genes Signature for Early Detection of Lung Cancer. *BioMedical Engineering and Informatics, International Conference on*, **1**, 413–417. 21

von Bohlen und Halbach, O., A. Schober, and K. Krieglstein, 2004: Genes, proteins, and neurotoxins involved in Parkinson's disease. *Progress in neurobiology*, **73**, 151–77. 5

Vrba, L., T. J. Jensen, J. C. Garbe, R. L. Heimark, A. E. Cress, S. Dickinson, M. R. Stampfer, and B. W. Futscher, 2010: Role for DNA methylation in the regulation of miR-200c and miR-141 expression in normal and cancer cells. *PloS one*, **5**, e8697. 123

Walker, D. G., K. Terai, A. Matsuo, T. G. Beach, E. G. McGeer, and P. L. McGeer, 1998: Immunohistochemical analyses of fibroblast growth factor receptor-1 in the human substantia nigra. Comparison between normal and Parkinson's disease cases. *Brain research*, **794**, 181–7. 7

Walker, F. O., 2007: Huntington's disease. *Lancet*, **369**, 218–28. 16

Wang, L., H.-C. Chiang, W. Wu, B. Liang, Z. Xie, X. Yao, W. Ma, S. Du, and Y. Zhong, 2012: From the Cover: Epidermal growth factor receptor is a preferred target for treating Amyloid- -induced memory loss. *Proceedings of the National Academy of Sciences*, **109**, 16743–16748. 59, 132

Wang, X., S. Zhu, Z. Pei, M. Drozda, I. G. Stavrovskaya, S. J. Del Signore, K. Cormier, E. M. Shimony, H. Wang, R. J. Ferrante, B. S. Kristal, and R. M. Friedlander, 2008: Inhibitors of cytochrome c release with therapeutic potential for Huntington's disease. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, **28**, 9473–85. 15

Witten, T. M. and D. Bonchev, 2007: Predicting aging/longevity-related genes in the nematode Caenorhabditis elegans. *Chemistry & biodiversity*, **4**, 2639–55. 21, 40

Wolfe, M. S., 2013: Alzheimer's γ-secretase under arrestin. *Nature medicine*, **19**, 22–4. 99

Wolfe, M. S., W. Xia, B. L. Ostaszewski, T. S. Diehl, W. T. Kimberly, and D. J. Selkoe, 1999: Two transmembrane aspartates in presenilin-1 required for presenilin endoproteolysis and gamma-secretase activity. *Nature*, **398**, 513–7. 9

Wong, R. W. C. and L. Guillaud, 2004: The role of epidermal growth factor and its receptors in mammalian CNS. *Cytokine & growth factor reviews*, **15**, 147–56. 149

Wu, Z., R. A. Irizarry, R. Gentleman, F. Martinez-Murillo, and F. Spencer, 2004: A Model-Based Background Adjustment for Oligonucleotide Expression Arrays. *Journal of the American Statistical Association*, **99**, pp. 909–917. 31

Xia, M. Q. and B. T. Hyman, 1999: Chemokines/chemokine receptors in the central nervous system and Alzheimer's disease. *Journal of neurovirology*, **5**, 32–41. 87

Xiang, Z., M. Valenza, L. Cui, V. Leoni, H.-K. Jeong, E. Brilli, J. Zhang, Q. Peng, W. Duan, S. A. Reeves, E. Cattaneo, and D. Krainc, 2011: Peroxisome-proliferator-activated receptor gamma coactivator 1 α contributes to dysmyelination in experimental models of Huntington's disease. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, **31**, 9544–53. 104

Xie, B., C. Wang, Z. Zheng, B. Song, C. Ma, G. Thiel, and M. Li, 2011: Egr-1 transactivates Bim gene expression to promote neuronal apoptosis. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, **31**, 5032–44. 109

Yasuhara, T., T. Shingo, K. Kobayashi, A. Takeuchi, A. Yano, K. Muraoka, T. Matsui, Y. Miyoshi, H. Hamada, and I. Date, 2004: Neuroprotective effects of vascular endothelial growth factor (VEGF) upon dopaminergic neurons in a rat model of Parkinson's disease. *The European journal of neuroscience*, **19**, 1494–504. 64

Yu, J.-Y., K.-H. Chung, M. Deo, R. C. Thompson, and D. L. Turner, 2008: MicroRNA miR-124 regulates neurite outgrowth during neuronal differentiation. *Experimental cell research*, **314**, 2618–33. 123

Zambrano, N., D. Gianni, P. Bruni, F. Passaro, F. Telese, and T. Russo, 2004: Fe65 is not involved in the platelet-derived growth factor-induced processing of Alzheimer's amyloid precursor protein, which activates its caspase-directed cleavage. *The Journal of biological chemistry*, **279**, 16161–9. 83

Zhang, X., Y.-w. Zhang, S. Liu, A. Bulloj, G. Tong, Z. Zhang, F.-F. Liao, and H. Xu, 2006: Tumor suppressor PTEN affects tau phosphorylation: deficiency in the phosphatase activity of PTEN increases aggregation of an FTDP-17 mutant Tau. *Molecular Neurodegeneration*, **1**, 7. 61

Zhang, Y., M. James, F. A. Middleton, and R. L. Davis, 2005: Transcriptional analysis of multiple brain regions in Parkinson's disease supports the involvement of specific protein processing, energy metabolism, and signaling pathways, and suggests novel disease mechanisms. *American journal of medical genetics. Part B, Neuropsychiatric genetics : the official publication of the International Society of Psychiatric Genetics*, **137B**, 5–16. 48, 72

Zhang, Y., V. O. Ona, M. Li, M. Drozda, M. Dubois-Dauphin, S. Przedborski, R. J. Ferrante, and R. M. Friedlander, 2003: Sequential activation of individual caspases, and of alterations in Bcl-2 proapoptotic signals in a mouse model of Huntington's disease. *Journal of neurochemistry*, **87**, 1184–92. 15

Zuccato, C. and E. Cattaneo, 2007: Role of brain-derived neurotrophic factor in Huntington's disease. *Progress in neurobiology*, **81**, 294–330. 118, 122

Zuccato, C., A. Ciammola, D. Rigamonti, B. R. Leavitt, D. Goffredo, L. Conti, M. E. MacDonald, R. M. Friedlander, V. Silani, M. R. Hayden, T. Timmusk, S. Sipione, and E. Cattaneo, 2001: Loss of huntingtin-mediated BDNF gene transcription in Huntington's disease. *Science (New York, N.Y.)*, **293**, 493–8. 118, 122, 133, 149

Zuccato, C., M. Valenza, and E. Cattaneo, 2010: Molecular mechanisms and potential therapeutical targets in Huntington's disease. *Physiological reviews*, **90**, 905–81. 16, 117, 118, 128

Zujovic, V., J. Benavides, X. Vigé, C. Carter, and V. Taupin, 2000: Fractalkine modulates TNF-alpha secretion and neurotoxicity induced by microglial activation. *Glia*, **29**, 305–15. 66

Appendix A

Statistical analysis of microarray dataset

A.1 R Program code

Below is the R program source code for statistical analysis of microarray dataset.

```
#R Code for Statistical Analysis of microarray dataset. This includes
#1. Pre-processing --- Density Histogram and boxplot, NUSE and RLE
#2. Normalization --- RMA normalization
#3. Post-processing --- ArrayQualityMetrix, Linear Model with eBayes.
   MA plot and Clusters from the eBayes data.Valcano plot and Heat Map
   from the toptable results

source("http://bioconductor.org/biocLite.R")
biocLite("hgu133a.db")
biocLite("hgu133acdf")
biocLite("gplots")

library(affy); library(limma); library(gcrma);
library(arrayQualityMetrics);library(hgu133acdf);
library(gplots);library(affyPLM);library(genefilter);
library(hgu133a.db);library(annotate);


#STEP 1: PRE-PROCESSING
#Read the CEL data
pd <- read.AnnotatedDataFrame("NBD_PD_TargetFile.txt")
data<-ReadAffy(filenames=rownames(pData(pd)),phenoData=pd)
```

```
#Quality assessment plots
hist(data)
boxplot(data)
Pset <- fitPLM(data)
NUSE(Pset)
RLE(Pset)
RNAdeg <- AffyRNAdeg(data)
plotAffyRNAdeg(RNAdeg,transform="shift.scale",log.it=F)

#STEP 2: PROCESSING/NORMALIZING
#Normalize the data
eset_rma <- rma(data)

#To run array Quality Metrics after normalization
arrayQualityMetrics(expressionset = eset_rma, outdir="QA_NBD_PD_Norm",
   force = FALSE, intgroup = fac)

#Analysis - Do a linear model of the expression dataset described by
   diagnosis (Control and case)
design <- model.matrix(~diagnosis,pData(eset_rma))
fit <- lmFit(eset_rma, design)
efit <- eBayes(fit)

#STEP 3: POST-PROCESSING
#MA Plot
M <- efit$coefficients[,2]
A <- efit$Amean
ma.plot(A,M,plot.method="normal")

#Clustering of different samples
eset_cluster <- exprs(eset_rma)
distance <- dist(t(eset_cluster),method="maximum")
clusters <- hclust(distance)
plot(clusters,main="Cluster dendrogram of normalized dataset")

#Volcano plot
plot(results$logFC,-log10(results$P.Value),pch=16,cex=0.1,
     xlab="Log Fold-Changes",
     ylab="-log10 P-Values",
     main='Volcano Plot of normalized NBD_PD dataset')
abline(h = 2, v=c(0.5,-0.5))
t <- is.element(results$ID,results$ID[results$adj.P.Val<0.05])
points(results$logFC[t],-log10(results$P.Value[t]),pch=16,col='red',cex=0.1)
t <- is.element(results$ID,results$ID[results$adj.P.Val<0.05][1:25])
```

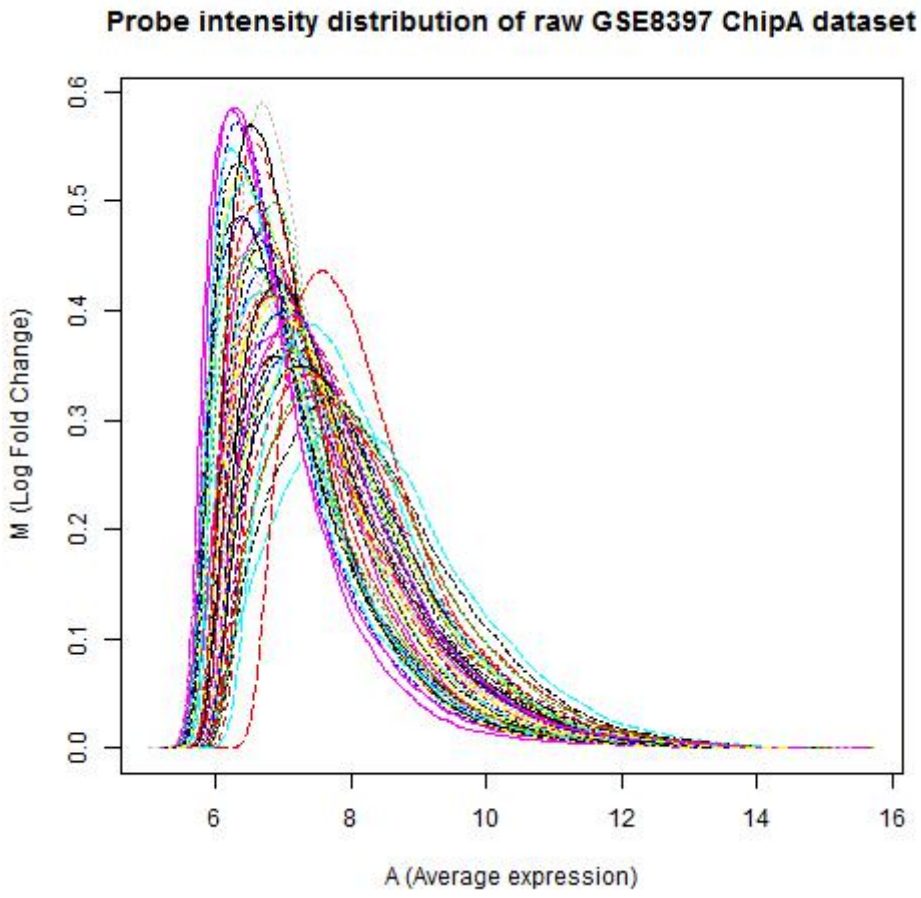
```
points(results$logFC[t],-log10(results$P.Value[t]),pch=7,col='black')
legend("topright",
c('SDEGs (BH p-value<0.05)','Top 25 SDEGs (BH p-value<0.05)','Other'),
pch=c(16,7,16),col=c('red','black','black'),cex=1)

#Heatmap of the arrays using the significant genes
topgenes = results[results[, "adj.P.Val"] < 0.00001, ]
m = exprs(eset_rma[topgenes[, "ID"], ])
par(mfrow=c(1,1))
colnames(m) = pData(pd)$Diagnosis
rownames(m) = topgenes$Symbol
colours = c("Red", "Blue")
col = colours[diagnosis]
heatmap.2(m, ColSideColors = col, margin = c(5,
   8),scale="none",cexRow=0.5,col=redgreen(75),
         key=FALSE,symkey=FALSE,density.info="none",trace="none",
            dendrogram="both",
         main = "Top SDGEs (p value <0.00001)",
         xlab="<-- Samples -->", ylab="<-- Genes -->")
```

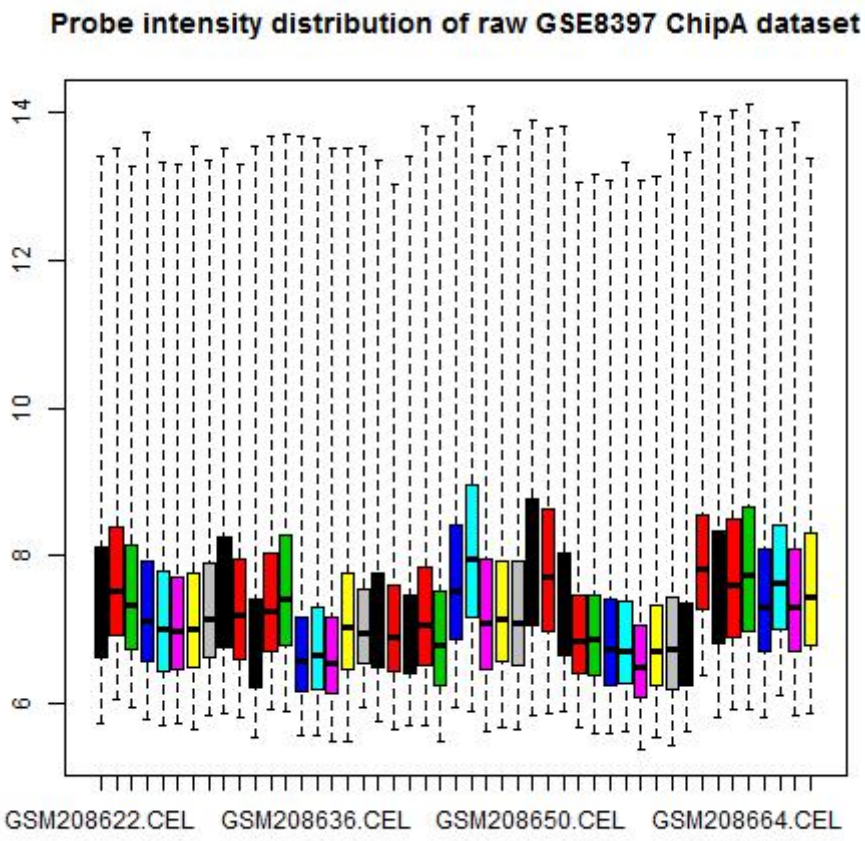
A.2 Microarray data processing plots/graphs

Following are the various plots/graphs drawn during the statistical analysis of a microarray dataset.

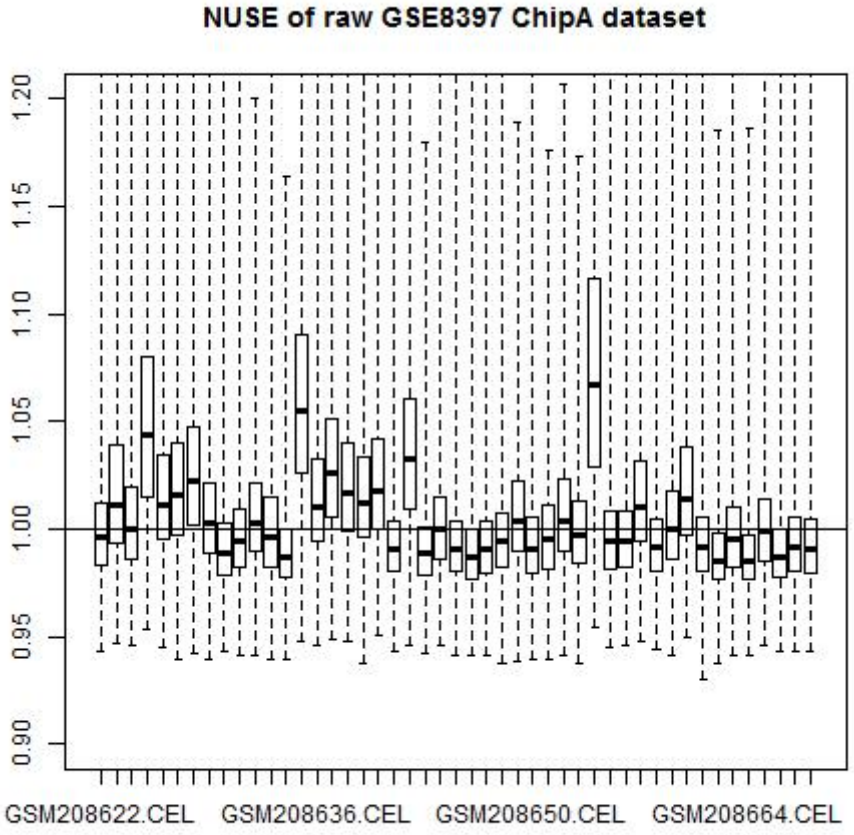
Pre-normalization: Histogram



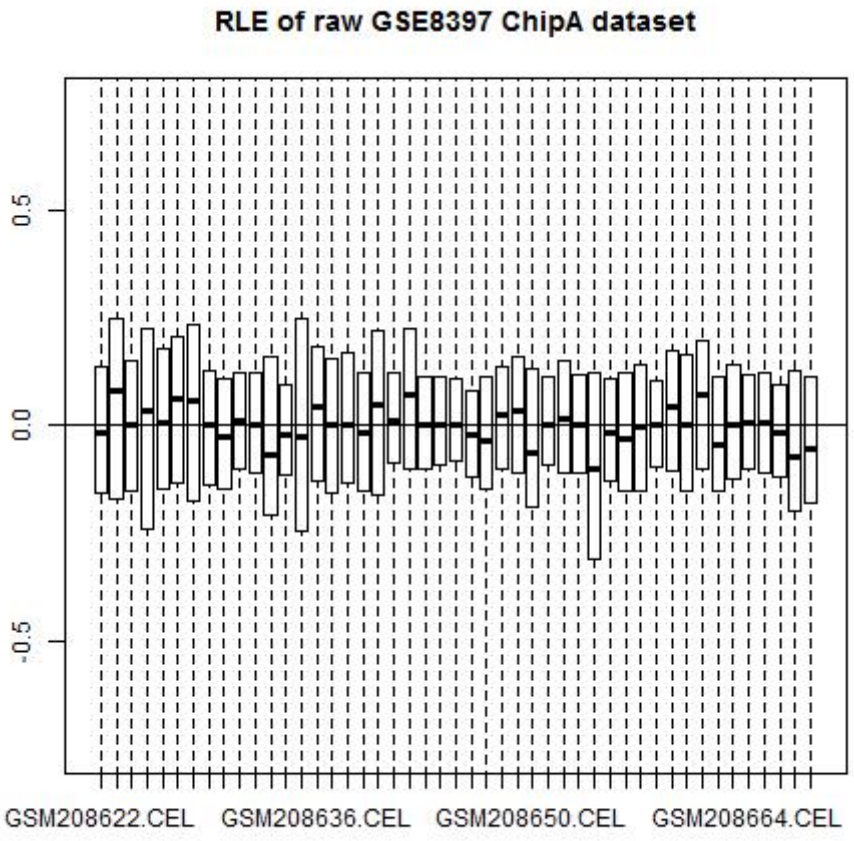
Pre-normalization: Boxplot



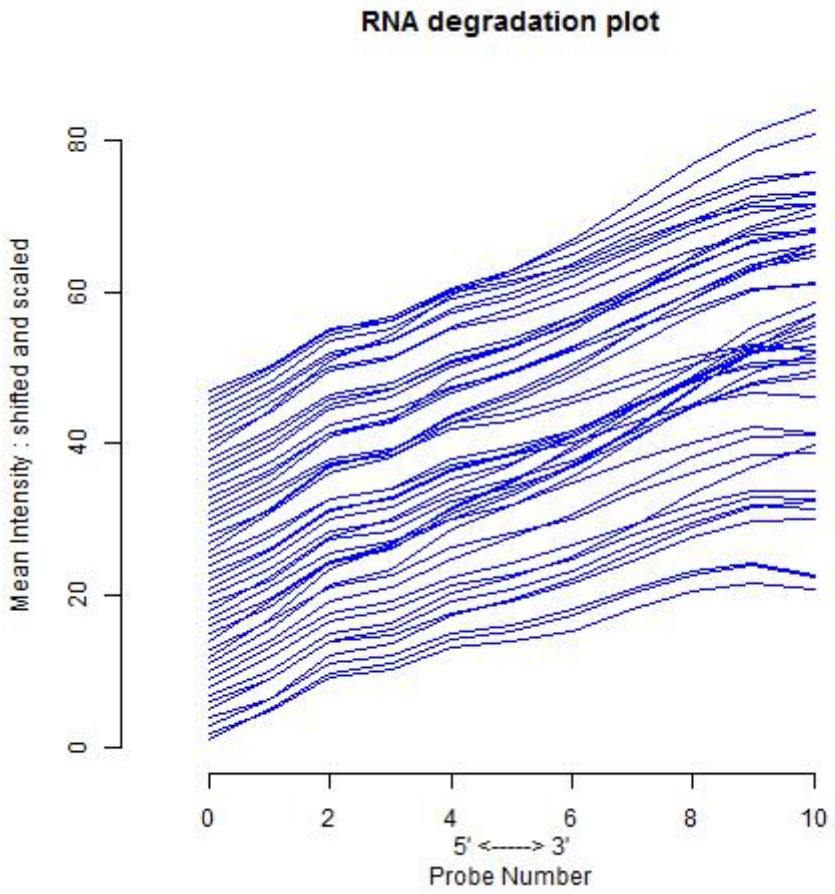
Pre-normalization: Normalized Unscaled Standard Error (NUSE) plot



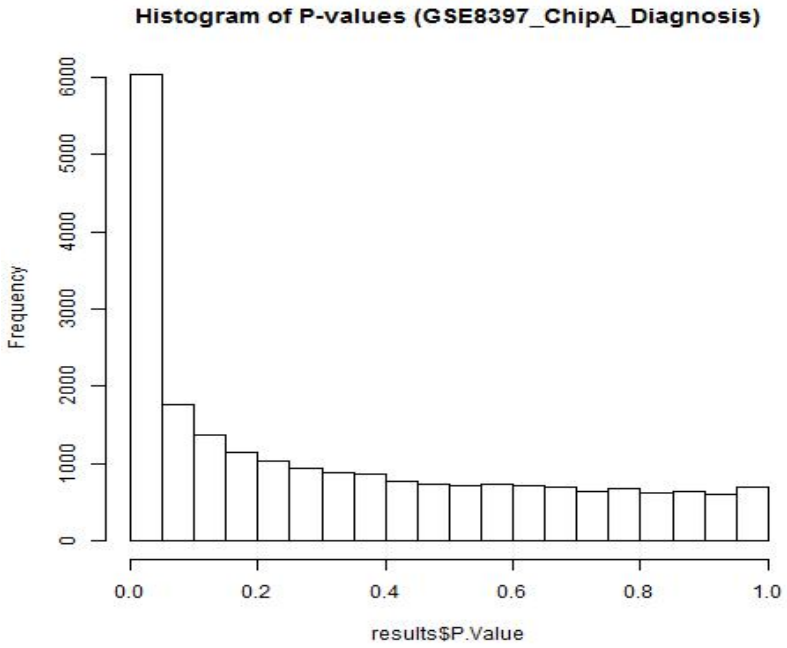
Pre-normalization: Relative Log Expression (RLE) plot

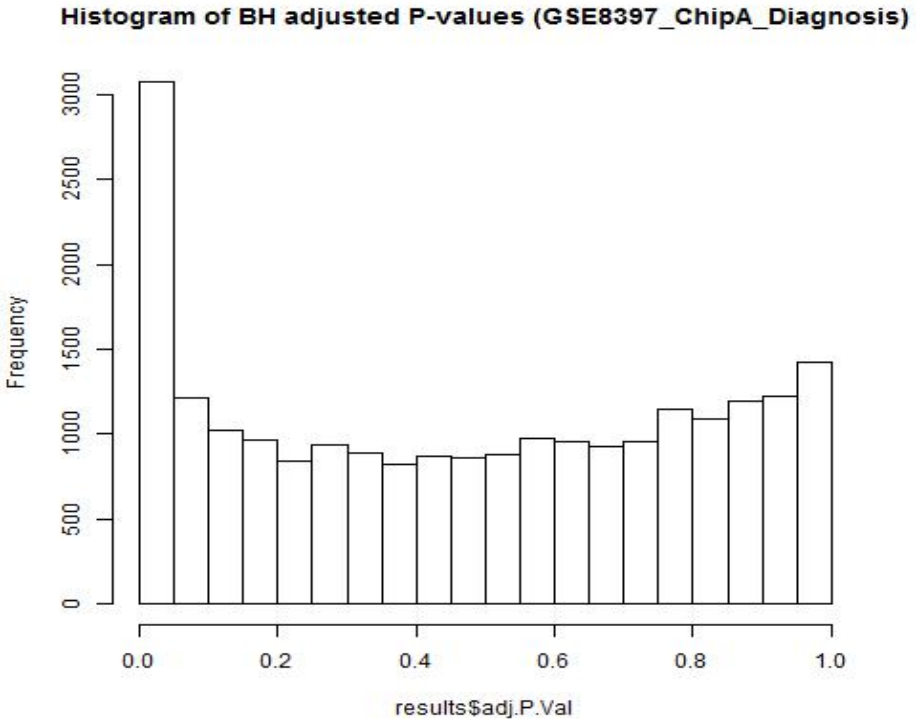


Pre-processing: RNA degradation plot

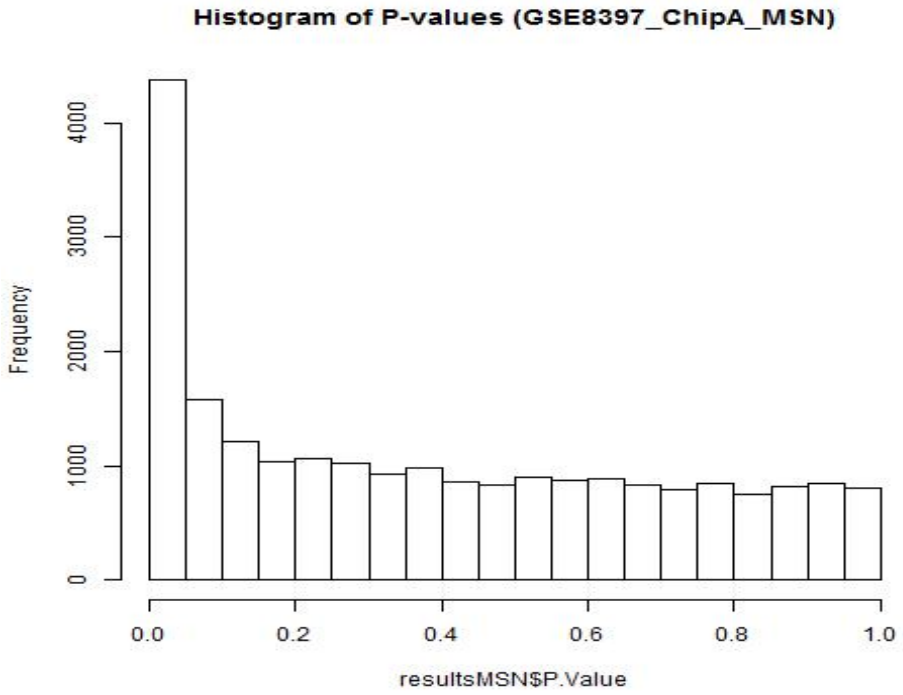


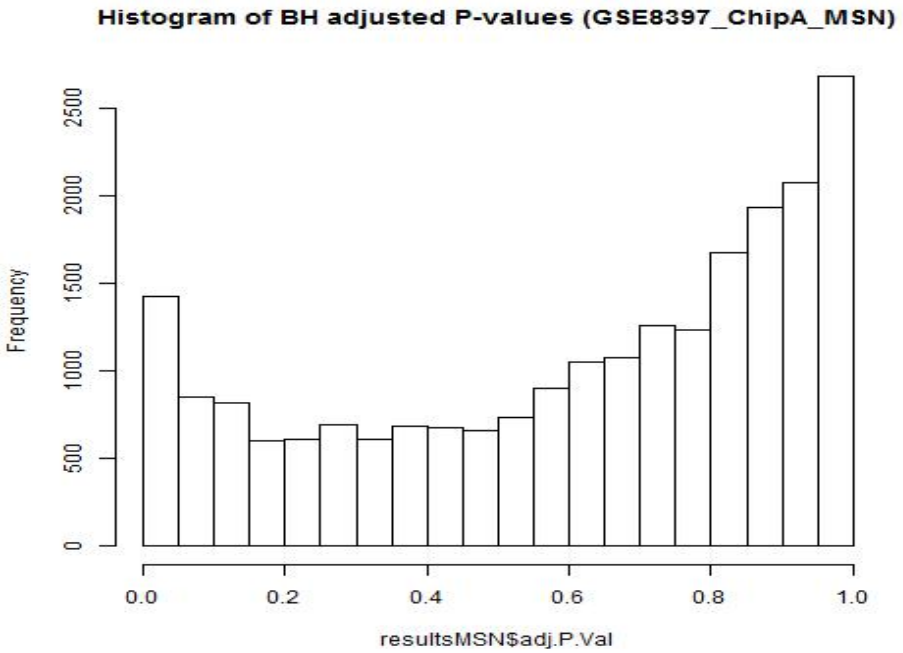
Post-processing: Histogram of p-and q-values



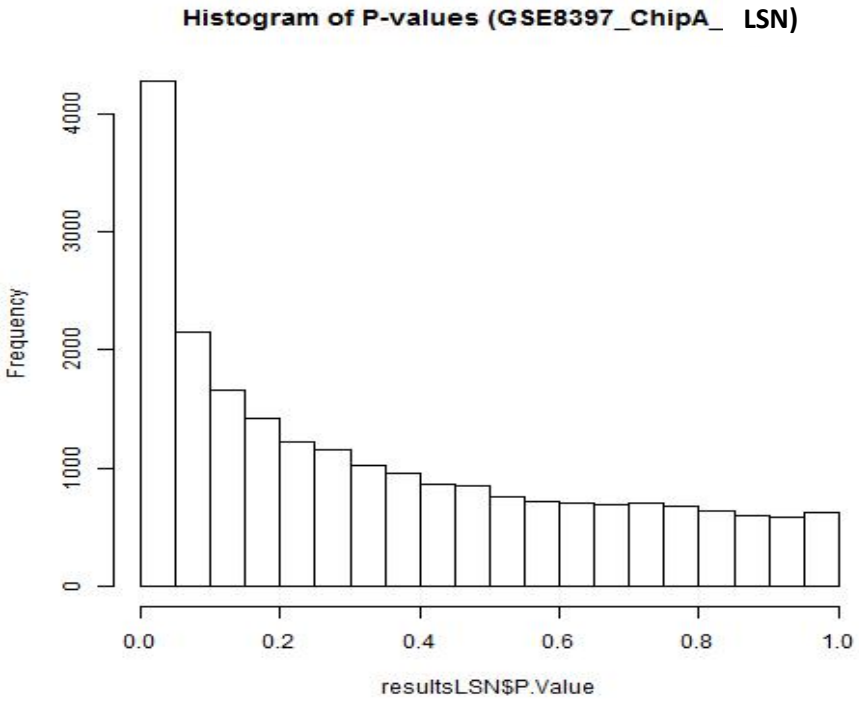


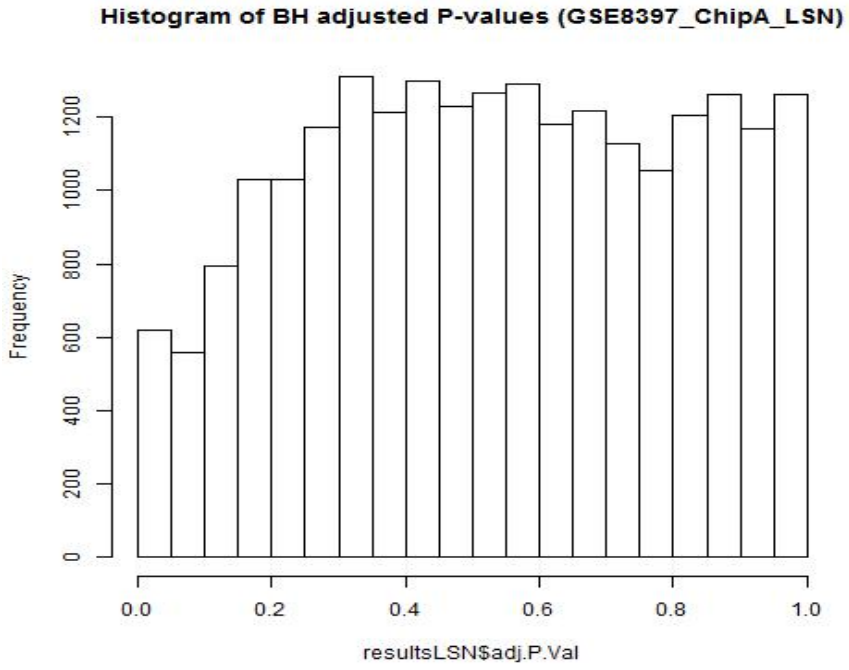
Post-processing: Histogram of p-and q-values



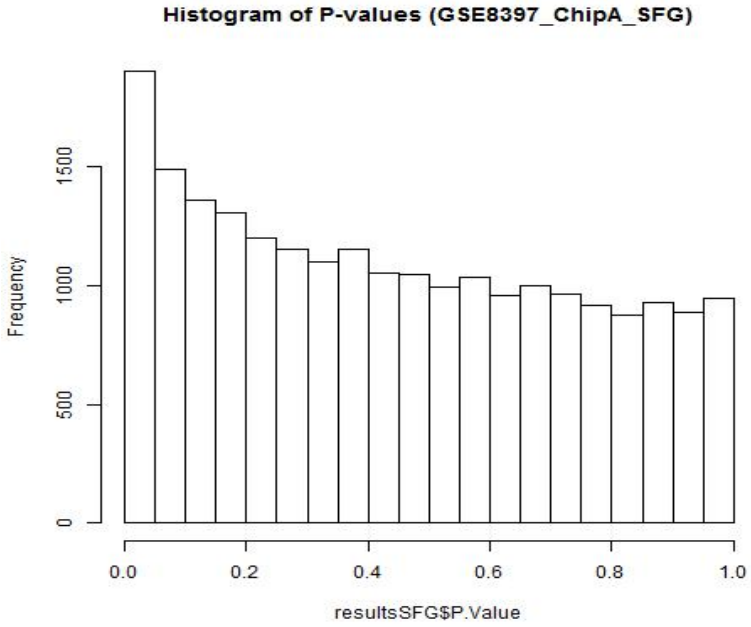


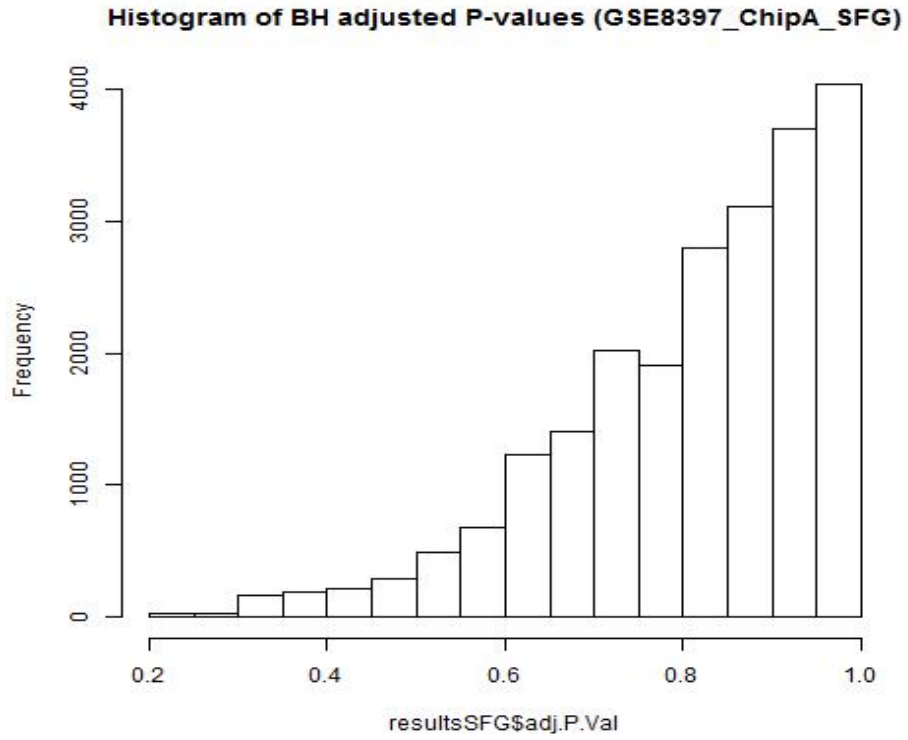
Post-processing: Histogram of p-and q-values



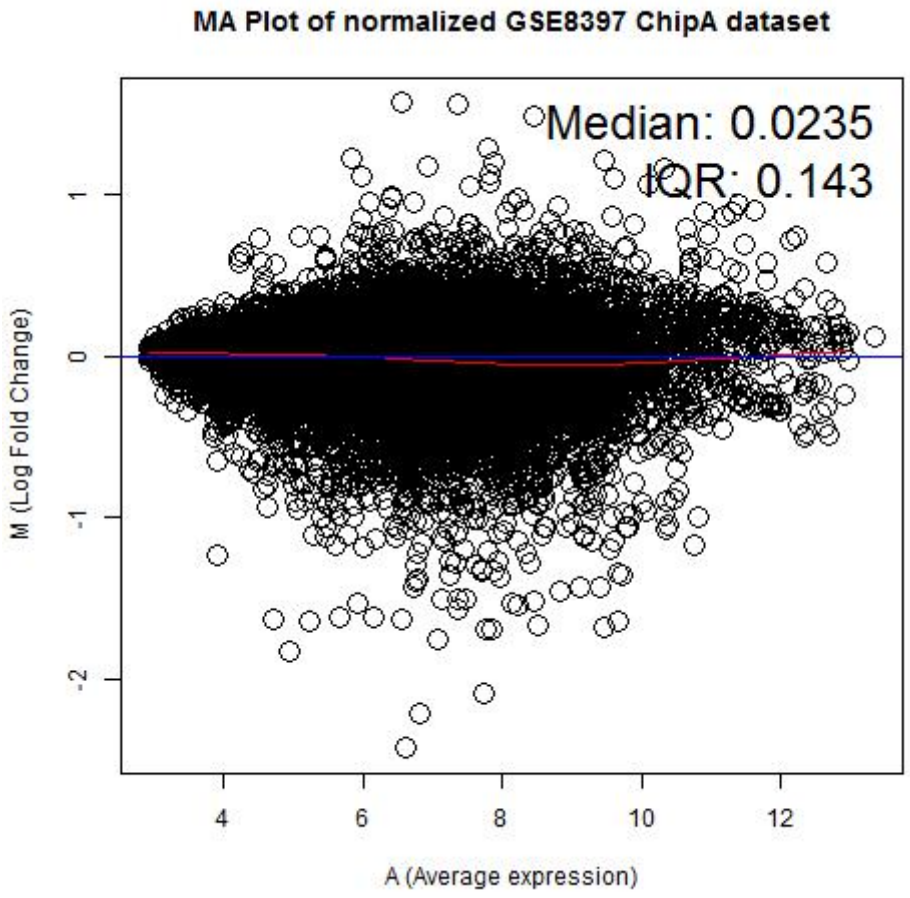


Post-processing: Histogram of p-and q-values

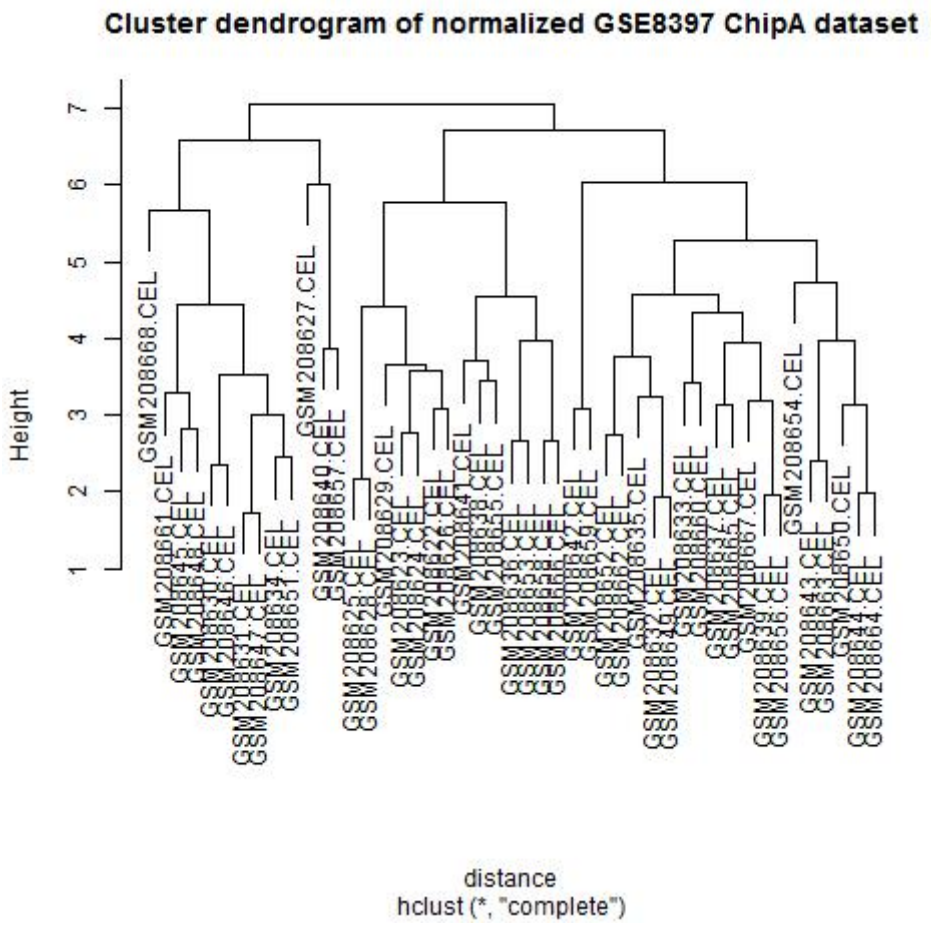




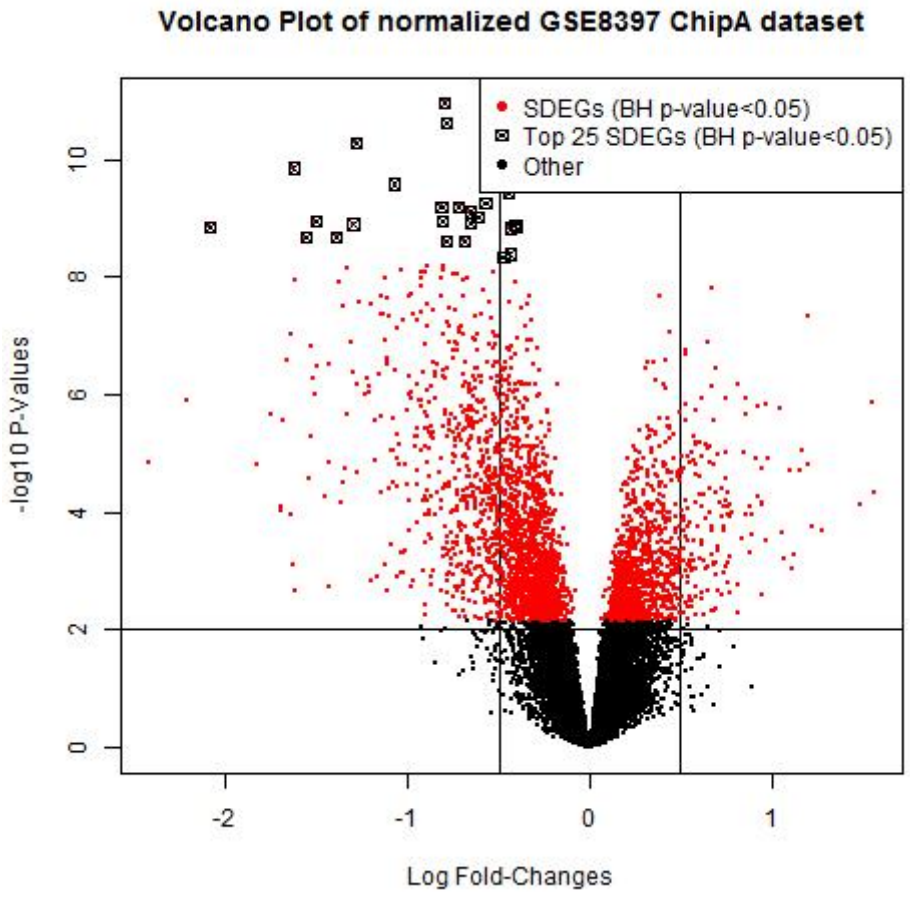
Post-processing: MA plot



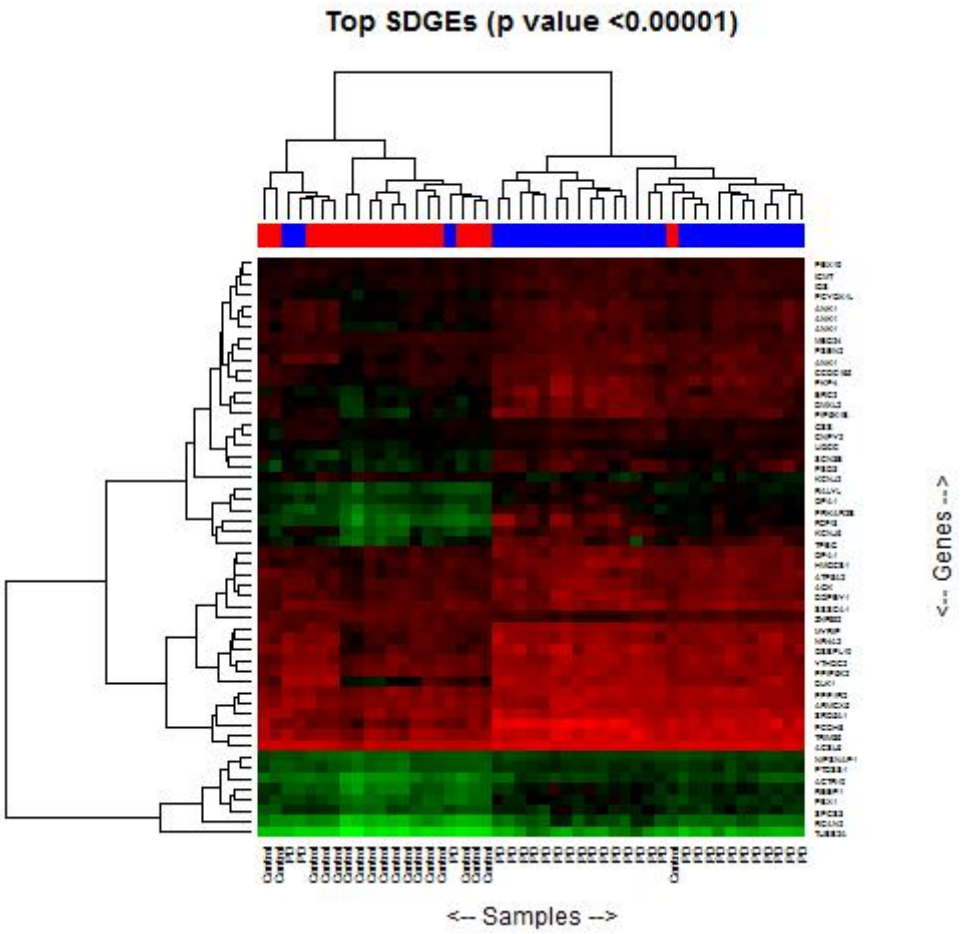
Post-processing: Clustering



Post-processing: Volcano plot



Post-processing: Heatmap



A.3 Sample *arrayQualityMetrics* report

Next few pages contains the *arrayQualityMetrics* report for a microarray dataset used in this study.

GSE8397QA Norm Report

Browser compatibility

This report uses recent features of HTML 5 which have not yet been implemented by all browsers. Thus, unfortunately, browser compatibility currently needs to be considered:

- Firefox 4 - tested, works well,
- Chrome 10 - tested, works well,
- Safari 5 - the interactive (SVG) plots will be missing, since this browser does not support the embedding of the <svg> tag in HTML.

+ Array metadata and outlier detection overview

Section 1: Between array comparison

- Figure 1: Distances between arrays.

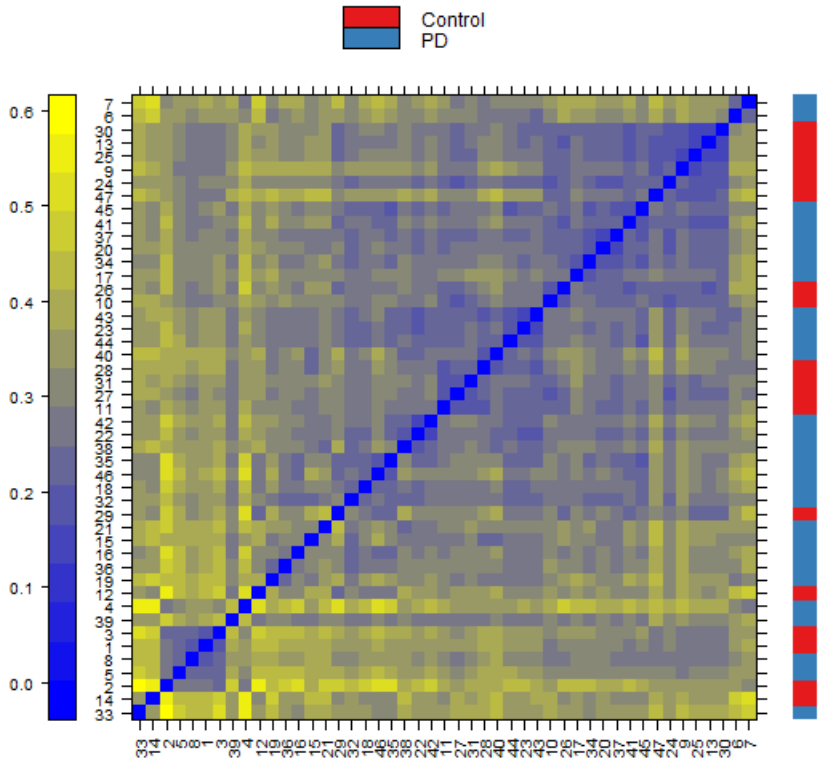


Figure 1 (PDF file) shows a false color heatmap of the distances between arrays. The color scale is chosen to cover the range of distances encountered in the dataset. Patterns in this plot can indicate clustering of the arrays either because of intended biological or unintended experimental factors (batch effects). The distance d_{ab} between two arrays a and b is computed as the mean absolute difference (L_1-distance) between the data of the arrays (using the data from all probes without filtering). In formula, d_{ab} = mean | M_{ai} - M_{bi} |, where M_{ai} is the value of the i-th probe on the a-th array. Outlier detection was performed by looking for arrays for which the sum of the distances to all other arrays, $S_a = \Sigma_b\ d_{ab}$ was exceptionally large. No such arrays were detected.

- Figure 2: Outlier detection for Distances between arrays.

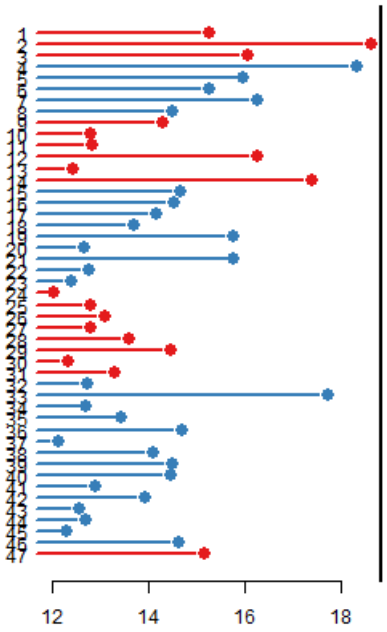


Figure 2 (PDF file) shows a bar chart of the sum of distances to other arrays S_a, the outlier detection criterion from the previous figure. The bars are shown in the original order of the arrays. Based on the distribution of the values across all arrays, a threshold of 18.8 was determined, which is indicated by the vertical line. None of the arrays exceeded the threshold and was considered an outlier.

- Figure 3: Principal Component Analysis.

array
sampleNames
Target
Name
TissueType
Age
ScanDate

Figure 3 (PDF file) shows a scatterplot of the arrays along the first two principal components. You can use this plot to explore if the arrays cluster, and whether this is according to an intended experimental factor, or according to unintended causes such as batch effects. Move the mouse over the points to see the sample names.
Principal component analysis is a dimension reduction and visualisation technique that is here used to project the multivariate data vector of each array into a two-dimensional plot, such that the spatial arrangement of the points in the plot reflects the overall data (dis)similarity between the arrays.

Section 2: Array intensity distributions

- Figure 4: Boxplots.

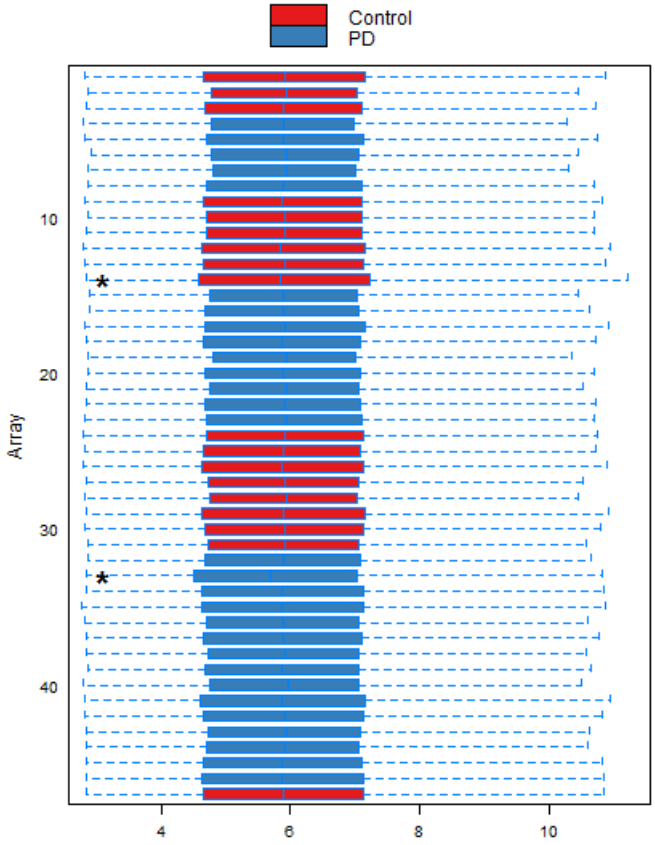


Figure 4 (PDF file) shows boxplots representing summaries of the signal intensity distributions of the arrays. Each box corresponds to one array. Typically, one expects the boxes to have similar positions and widths. If the distribution of an array is very different from the others, this may indicate an experimental problem. Outlier detection was performed by computing the Kolmogorov-Smirnov statistic K_a between each array's distribution and the distribution of the pooled data.

- Figure 5: Outlier detection for Boxplots.

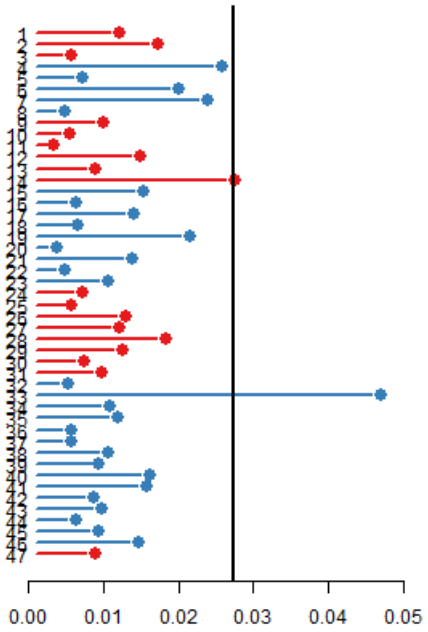


Figure 5 (PDF file) shows a bar chart of the Kolmogorov-Smirnov statistic K_a, the outlier detection criterion from the previous figure. The bars are shown in the original order of the arrays. Based on the distribution of the values across all arrays, a threshold of 0.0274 was determined, which is indicated by the vertical line. 2 arrays exceeded the threshold and were considered outliers.

- Figure 6: Density plots.

array
sampleNames
Target
Name
TissueType
Age
ScanDate

Figure 6 (PDF file) shows density estimates (smoothed histograms) of the data. Typically, the distributions of the arrays should have similar shapes and ranges. Arrays whose distributions are very different from the others should be considered for possible problems. Various features of the distributions can be indicative of quality related phenomena. For instance, high levels of background will shift an array's distribution to the right. Lack of signal diminishes its right right tail. A bulge at the upper end of the intensity range often indicates signal saturation.

Section 3: Variance mean dependence

- Figure 7: Standard deviation versus rank of the mean.

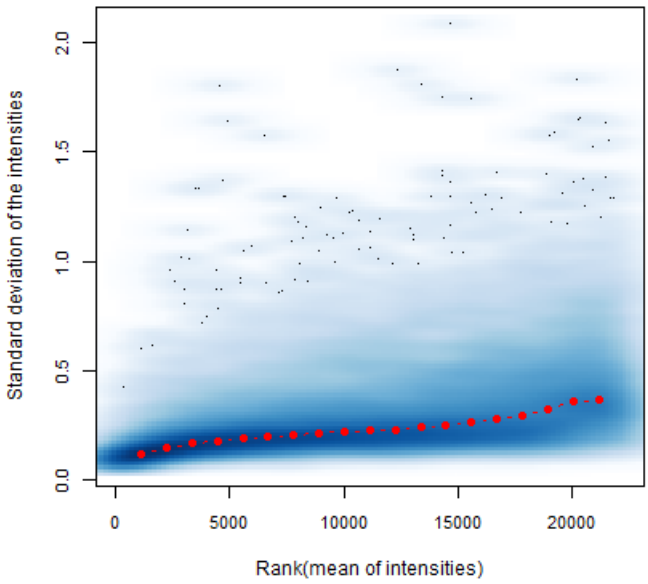


Figure 7 (PDF file) shows a density plot of the standard deviation of the intensities across arrays on the *y*-axis versus the rank of their mean on the *x*-axis. The red dots, connected by lines, show the running median of the standard deviation. After normalisation and transformation to a logarithm(-like) scale, one typically expects the red line to be approximately horizontal, that is, show no substantial trend. In some cases, a hump on the right hand of the x-axis can be observed and is symptomatic of a saturation of the intensities.

Section 4: Individual array quality

- Figure 8: MA plots.

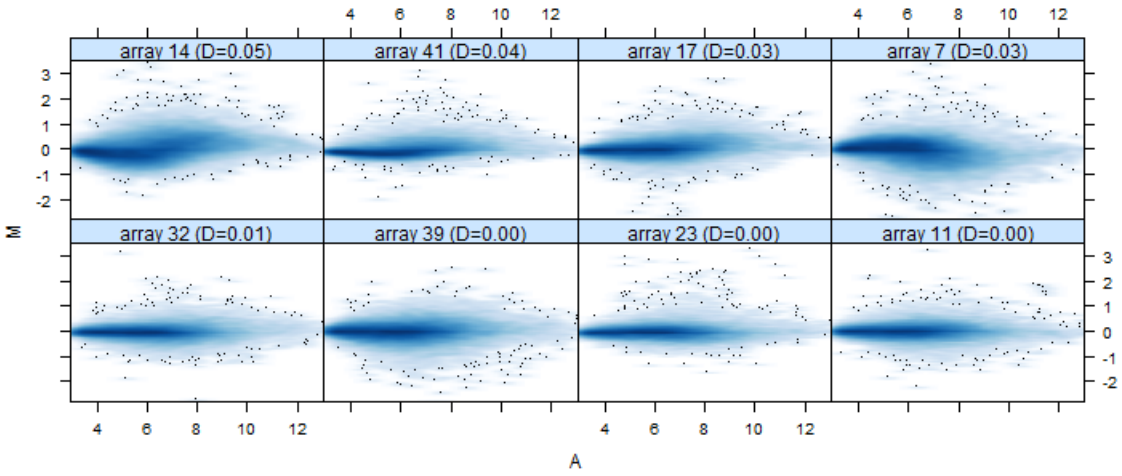


Figure 8 (PDF file) shows MA plots. M and A are defined as:
$M = \log_2(I_1) - \log_2(I_2)$
$A = 1/2\ (\log_2(I_1)+\log_2(I_2))$,
where I_1 is the intensity of the array studied, and I_2 is the intensity of a "pseudo"-array that consists of the median across arrays. Typically, we expect the mass of the distribution in an MA plot to be concentrated along the M = 0 axis, and there should be no trend in M as a function of A. If there is a trend in the lower range of A, this often indicates that the arrays have different background intensities; this may be addressed by background correction. A trend in the upper range of A can indicate saturation of the measurements; in mild cases, this may be addressed by non-linear normalisation (e.g. quantile normalisation).
Outlier detection was performed by computing Hoeffding's statistic D_a on the joint distribution of A and M for each array. Shown are the 4 arrays with the highest value of D_a (top row), and the 4 arrays with the lowest value (bottom row). The value of D_a is shown in the panel headings. 0 arrays had D_a>0.15 and were marked as outliers. For more information on Hoeffing's *D*-statistic, please see the manual page of the function `hoeffd` in the `Hmisc` package.

- Figure 9: Outlier detection for MA plots.

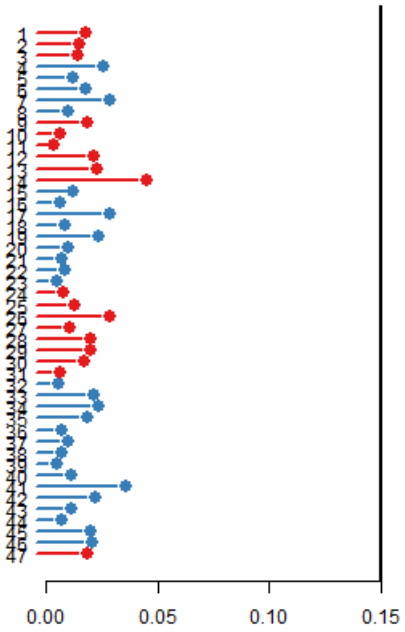


Figure 9 (PDF file) shows a bar chart of the Hoeffding's statistic D_a, the outlier detection criterion from the previous figure. The bars are shown in the original order of the arrays. A threshold of 0.15 was used, which is indicated by the vertical line. None of the arrays exceeded the threshold and was considered an outlier.

This report has been created with arrayQualityMetrics 3.10.0 under R version 2.14.0 (2011-10-31).

(Page generated on Sat Nov 17 16:43:33 2012 by hwriter)

Appendix B

Find “seed genes”

Four-set Venn diagram illustrations of the overlap of significantly differentially expressed genes in various microarray datasets used in the study. The key to the Venn diagram is also included at the end. Courtesy: Oliveros, J.C. (2007) VENNY. An interactive tool for comparing lists with Venn Diagrams. http://bioinfogp.cnb.csic.es/tools/venny/index.html.

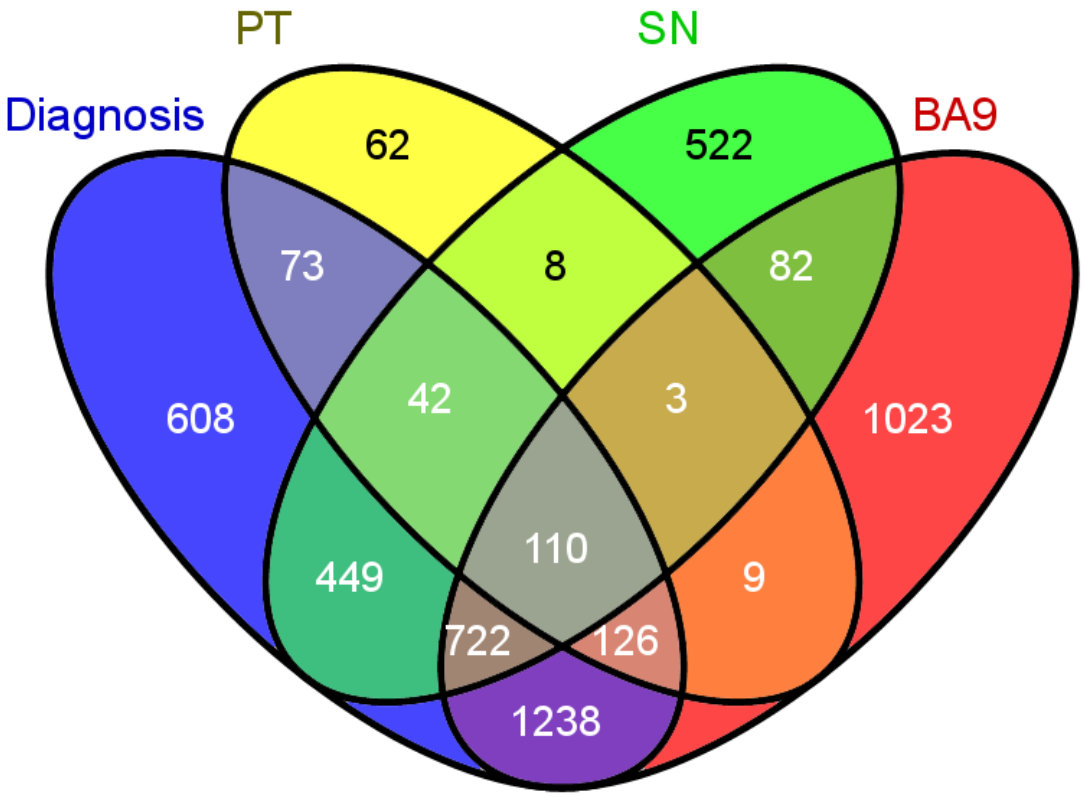


Figure B.1. Four-set Venn diagram illustration of the overlap of SDEGs in GSE20295 HG-U133A gene expression dataset for Parkinson's disease. Courtesy: VENNY

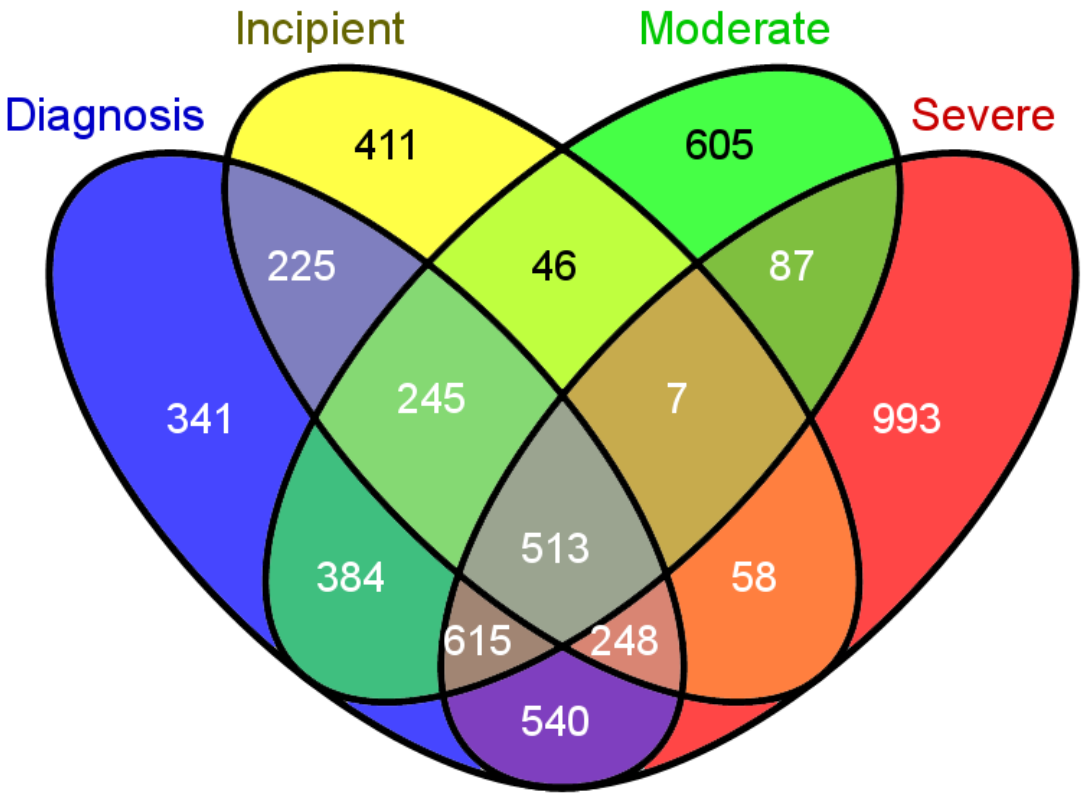


Figure B.2. Four-set Venn diagram illustration of the overlap of SDEGs in GSE28146 HG-U133 Plus 2.0 gene expression dataset for Alzheimer's disease. Courtesy: VENNY

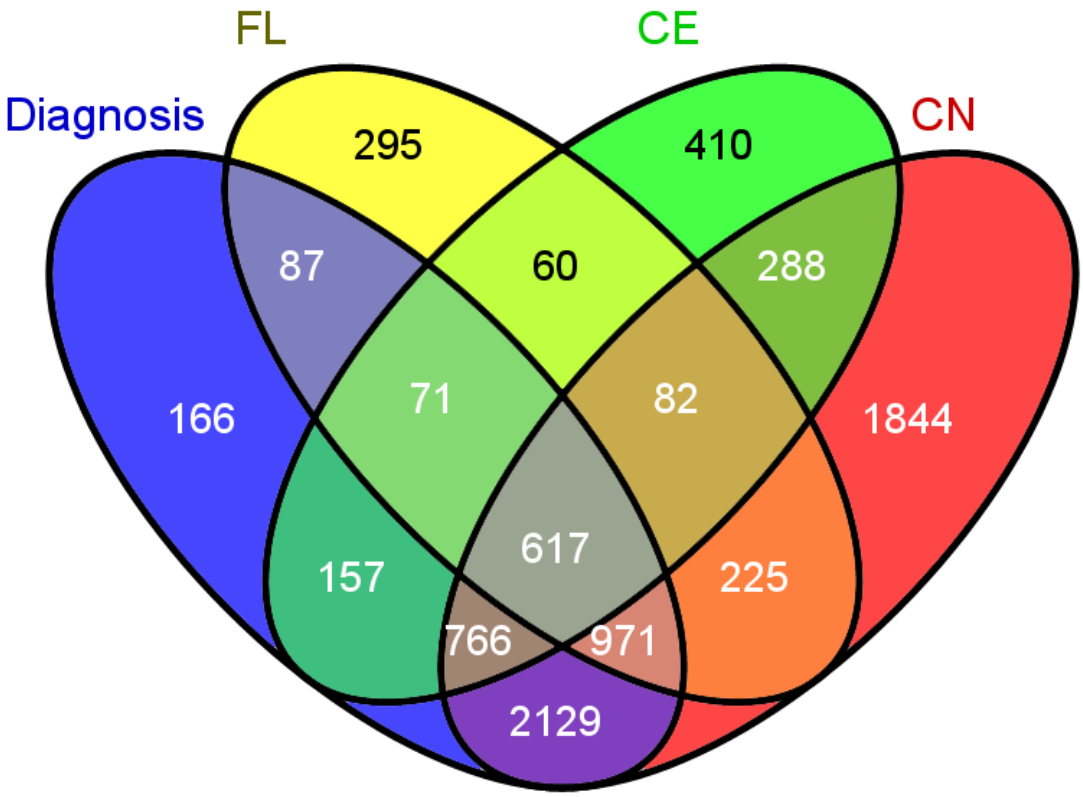


Figure B.3. Four-set Venn diagram illustration of the overlap of SDEGs in GSE3970A HG-U133A gene expression dataset for Huntington's disease. Courtesy: VENNY

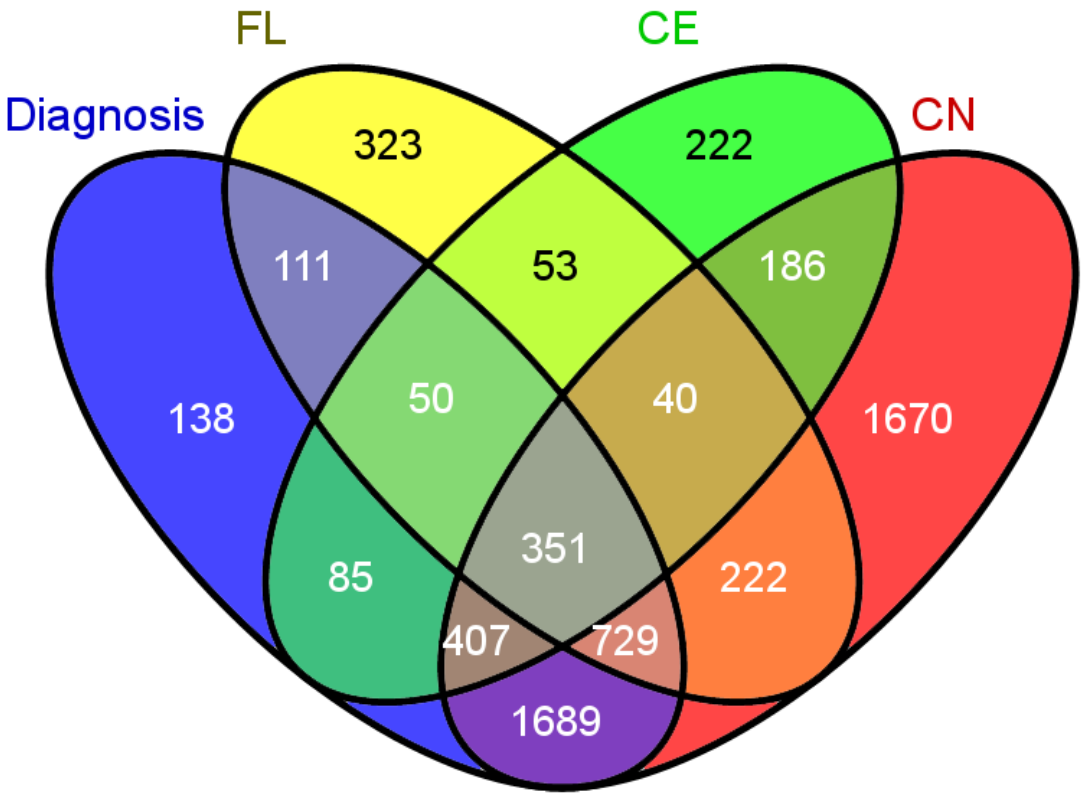


Figure B.4. Four-set Venn diagram illustration of the overlap of SDEGs in GSE3970B HG-U133B gene expression dataset for Huntington's disease. Courtesy: VENNY