IDENTIFICATION OF MIRNA REGULATORY PATHWAYS IN COMPLEX DISEASES

by

ILKNUR MELIS DURASI KUMCU

Submitted to the Graduate School of Engineering and Natural Sciences

in partial fulfilment of the requirements for the degree of

Doctor of Philosophy

Sabancı University

July 2018

IDENTIFICATION OF MIRNA REGULATORY PATHWAYS IN COMPLEX DISEASES

APPROVED BY:

Prof. Dr. Devrim Gözüaçık
(Thesis Supervisor) .....

Prof. Dr. Osman Uğur Sezerman .....

Prof. Dr. İsmail Çakmak .....

Prof. Dr. Yücel Saygın .....

Assoc. Prof. Dr. Emel Timuçin .....

DATE OF APPROVAL: 31/07/2018

To my beloved brother…

"You're only given one little spark of madness. You mustn't lose it."

IDENTIFICATION OF MIRNA REGULATORY PATHWAYS IN COMPLEX DISEASES

Ilknur Melis DURASI KUMCU

BIO, Doctor of Philosophy, 2018

Thesis Supervisor: Prof.Dr.Devrim Gözüaçık

Keywords: DEGs, miRNAs, complex diseases, directed signaling networks, regulatory pathways

ABSTRACT

MicroRNAs, small endogenous non-coding RNAs are one of the most important components in the cell and they play a critical role in many cellular processes and have been linked to the control of signal transduction pathways. Identifying disease related miRNAs and using that knowledge to understand the disease pathogenesis at the molecular level, new molecular tools can be designed for reducing the time and cost of diagnosis, treatment and prevention. Computational models have become very useful and practical in terms of discovering new miRNA disease associations to be used in experimental validations.

Omics studies demonstrated that changes in miRNA profiles of various tissues correlate with many complex diseases, such as Alzheimer's, Parkinson's or Huntington's and various cancers.

The aim of our study was to identify the potential active TF-miRNA-gene regulatory pathways involved in complex diseases Huntington's and Parkinson's, via integrating miRNA and gene expression profiles with known experimentally verified miRNAs/genes and directed signaling network.

We downloaded the miRNA and gene expression profiles from gene expression omnibus (GEO) database. We derived the differentially expressed genes (DEGs) and differentially expressed miRNAs (DEmiRs). SIGNOR database of causal relationships between signaling entities is used

as a signed directed network and TF-miRNA-gene bidirectional regulatory network is constructed. Then, DEGs and DEmiRs are mapped to the TF-miRNA-gene regulatory network. We connected the mapped DEGs and DEmiR nodes with their third-degree neighbors, hence, the potential regulatory TF-miRNA-gene subnetwork was built. By using BFS algorithm, the potential disease related TF-miRNA-gene regulatory pathways were identified.

In this study, we analyzed Huntington's and Parkinson's related mRNA and miRNA expression profiles with transcription factors (TF) and miRNAs known to be related to diseases. miRNA-TF-gene regulatory mechanisms and disease specific TF and miRNA regulatory pathways were aimed to be identified systematically.

This study provides bioinformatic support for further research on the molecular mechanism of complex diseases.

KOMPLEKS HASTALIKLARDA MİRNA DÜZENLEYİCİ YOLAKLARIN BELİRLENMESİ

Ilknur Melis DURASI KUMCU

BIO, Doktora Tezi, 2018

Tez Danışmanı: Prof.Dr.Devrim Gözüaçık

Anahtar kelimeler: Gen ifadesi, miRNA ifadesi, kompleks hastalıklar, yönlü sinyal ağları, düzenleyici yolaklar

ÖZET

mikroRNA'lar, küçük, endojen, kodlamayan RNA molekülleridir ve pek çok hücresel süreçte kritik rol oynarlar ve sinyal iletimi yolaklarının kontrolüyle bağdaştırılmışlardır. Hücrenin en önemli bileşenlerinden biri olarak, farklı biyolojik süreçlerle ilgili önemli role sahiptirler. Hastalık ilişkili miRNAların tanımlanması ve bu bilginin moleküler düzeyde hastalıkların patogenezinin anlaşılabilmesi için teşhis, tedavi ve koruma için harcanan zamanı ve maliyeti düşüren yeni moleküler araçlar geliştirilebilir. Bilgisayımsal modeller hastalık ilişkili yeni miRNA'ların keşfedilmesi ve deneysel validasyonlarda kullanılabilmesi için oldukça kullanışlı ve pratik hale gelmiştir.

Omik çalışmalar, çeşitli dokulardaki miRNA profillerindeki değişimlerin Alzheimer, Parkinson, Huntington ve kanser çeşitleri gibi kompleks hastalıklar ile korele olduğunu göstermiştir.

Çalışmamızdaki amacımız, miRNA ve gen ifade profillerini, ilgili hastalıkla iligisi olduğu bilinen ve deneysel olarak doğrulanmış miRNA/gen ve yönlü sinyal ağlarını birleştirerek, Huntington ve Parkinson kompleks hastalıklarında yer alan potansiyel aktif Transkripsiyon Faktör(TF)–miRNA–gen düzenleyici yolaklarını tanımlayabilmekti.

Omics studies demonstrated that changes in miRNA profiles of various tissues correlate with many complex diseases, such as Alzheimer's, Parkinson's or Huntington and various cancers.

miRNA ve gen ifade profillerini Gene Expression Omnibus (GEO) veri bankasından indirdik. Kademeli ifade edilen genleri ve miRNA'ları belirledik. Sinyalleşen birimler arası nedensel ilişkiler bilgisini barındıran SIGNOR veri bankası, yönlü sinyal ağın oluşturulması için kullanıldı, TF-miRNA-gen çift yönlü düzenleyici ağ yapılandırıldı. İfade edilen genler ve miRNA'lar organize edilmiş TF-miRNA-gen düzenleyici ağ üzerine aktif düğümler olarak işaretlendi. Aktif düğümler, birinci derece komşuluğuklarıyla birleştirilerek potansiyel düzenleyici ilgili hastalığa özgü TF-miRNA-gen alt ağı elde edildi. BFS algoritması kullanılarak, potansiyel aktif TF-miRNA-gen düzenleyici yolakları tanımlandı.

Bu çalışmada, sistemik olarak Huntington ve Parkinson ile ilişkili mRNA ve miRNA ifade profillerini, organize edilmiş TF ve miRNA düzenleyici mekanizmalarını, aktif TF ve miRNA düzenleyici yolaklarını tanımlamak için analiz ettik.

Bu çalışma gelecekte yapılacak kompleks hastalıkların mekanizması üzerine yapılacak araştırmalar için biyoenformatiksel destek sağlayacaktır.

ACKNOWLEDGEMENTS

I would like to express my deepest appreciation and gratitude to my supervisor, Prof. Dr. Osman Uğur Sezerman for his support and constructive critique, for his understanding and guidance not only for this project but also for my career. He has been there as my supervisor for more than 10 years and it has been a pleasure to be his student. Without his guidance and help this dissertation would not have been possible. Thank you so much for pushing me to look at and work on my research in different ways and thank you for opening my mind. I would like to thank my dissertation supervisor Prof. Dr. Devrim Gözüaçık for providing indispensable advice, information and cooperation and all jury members Prof. Dr. İsmail Çakmak, Prof. Dr. Yücel Saygın and Assoc. Prof. Dr. Emel Timuçin for their constructive comments.

There are number of people without whom this thesis might not have been written, and to whom I am greatly indebted.

To my father and mother, who have been a source of encouragement and inspiration to me, a very special thank you for supporting me in my determination to find and realize my potential in life. To my dear brother, thank you for not letting me feel hopeless for the future, and thank you for always finding a way to make me smile and change my mood. I know that you all will always be there for me in every step of my life.

To my beloved husband, a very special thank you for your practical and emotional support and believing in me. It was such a pleasure and motivation for me to see the proud in your eyes when talking about my research.

I have been supported by many friends and colleagues. Without their motivation, this journey would have been tougher and longer. I would also like to thank Senem Avaz Seven and Utku Seven, Gökşin Liu, Kadriye Kahraman, Aslı Yenenler for being there whenever I need them. I have also been a part of a precious team "Sezerman Lab". My lab mates, Begüm Özemek, Nogayhan Seymen, Ceren Saygı, Ege Ülgen and Rüçhan Ekren, "thank you so much". I will never forget their support during the end of my thesis writing. I also would like to thank to my FENS G022 family for being a part of my journey: Ahmet Sinan Yavuz, Beyza Vuruşaner, Cem

Meydan, Deniz Adalı, Hazal Yılmaz, Zoya Khalid, Serkan Sırlı. It was fun to share the office with you.

Last but not least, I would like to express my gratitude to all my teachers who put their faith in me and urged me to do better. Thank you for molding me into someone I can be proud of.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS AND ABBREVIATIONS

AGO family proteins	Argonaute family proteins
BFS	Breath First Search
CNS	Central Nervous System
CR	Coverage Rate
DEGs	Differentially Expressed Genes
DEmiRs	Differentially Expressed miRNAs
DGCR8	Drosha-DiGeorge syndrome critical region 8
dsRNA	double stranded RNA
EXP5	Exportin 5
FDR	False Discovery Rate
HD	Huntington's Disease
LCFAs	Long Chain Fatty Acids
miRNA	microRNA
nt	nucleotide
PD	Parkinson's Disease
piRNAs	PIWI-interacting RNAs
pri-miRNA	primary miRNA
Pol II	RNA polymerase II
PPI	protein-protein interaction

PPIN	Protein-protein interaction network
RISC	RNA-induced silencing complex
SIGNOR	Signaling Network Open Resource
TF	Transcription Factor
UTR	untranslated region

BACKGROUND

1.1 Understanding the Mechanism of Complex Diseases

Complex diseases are caused by a combination of genetic perturbations and environmental factors. Scientists know that a single genetic mutation in other words Mendelian patterns of inheritance cannot explain the pattern of a complex disease.

Understanding the molecular mechanisms through which factors affects a phenotype is complicated. Moreover, it is more difficult to understand the complex relationships of genetic and environmental factors in affected individuals as the complete view of complex diseases might be changeable among them. In recent years, systems biology approaches and network-based approaches were discovered and catch researchers' attention. Their powerful potential for studying complex diseases were expected to be a new era for the development of precision medicine. Network-based approaches generally use the physical and functional interactions between molecules to represent the interaction data as a network. An interaction network contains both the binary relationships between individual nodes and hidden higher level organization of cellular communication. That is why, it is crucial to combine multi-omics data into an integrated network to constitute enough knowledge for the interpretation of the disease molecular mechanism[1].

Many diseases fall in the category of complex disease including cancer, autism, diabetes, obesity, Huntington's disease, Parkinson's disease, and coronary artery disease. Recently, there is a huge amount of data such as genomic, transcriptomic, proteomic and metabolomic data related to these diseases. They are available to scientists to be used to do significantly facilitated research into complex diseases. However, extracting useful information from biological databases is a complex task. Recently, there are many studies just using individual type of biological layer which do not declare any interconnection between them. The task of revealing the molecular perturbations of diseases becomes even more complicated when it comes to gene regulation, TFs a transcriptional regulators and miRNAs as post-transcriptional regulators[2].

1.2 microRNAs (miRNAs)

Multiple types of small RNAs exist in eukaryotes and these RNAs regulate gene expression not only in the cytoplasm but also in the nucleus. Small RNAs suppress unwanted genetic materials and transcripts by different regulatory mechanisms: a) post-transcriptional gene silencing, b) chromatin-dependent gene silencing or c) RNA activation. That is why, their roles in health and disease development is important and need to be understood [3].

Small RNAs are defined as non-coding RNA molecules and their length is about 18–30 nucleotides. Three classes of small RNAs have been defined: microRNAs (miRNAs), siRNAs and Piwi-interacting RNAs (piRNAs) [4].

In eukaryotes, miRNAs are ~22 nucleotides in length. They are produced by Drosha and Dicer which are RNase III proteins and they dominate other classes of small RNAs. The domain at the 5′ end from nucleotide position 2 to 7 which is responsible for target recognition is called 'miRNA seed' and miRNA binding regions are generally located in the 3′ untranslated region (UTR) of mRNA sequences[5,6]. It was thought that, perfect seed matching was the only mechanism for miRNA silencing process but recent studies showed that downstream nucleotides of miRNAs specifically nucleotide 8 and nucleotides 13–16 which are outside the seed, reported to promote binding to mRNA nucleotides [7]. It is also known that, more than 60% of human protein-coding genes are in tendency to construct a pairing with miRNAs. Hence, it becomes more apparent why many miRNA binding sites have conserved sites, in addition to non-conserved sites. It can be concluded that, most protein-coding genes may be under the control of miRNAs [5]. Moreover, not only the expression of genes is regulated by miRNAs but also the expression of miRNAs themselves are regulated by regulatory mechanisms[8], and their dysregulation is revealed to be related to human diseases, including cancer, neurodevelopmental disorders, cardiovascular disease, diabetes, kidney and liver disease and infectious diseases [9].

1.2.1 miRNA transcription

miRNA genes are transcribed by RNA polymerase II (Pol II) and primary transcripts (pri-miRNAs) are generated. One transcript with a local hairpin structure is longer than the other

one. pri-miRNAs are processed by the Drosha-DiGeorge syndrome critical region gene 8 (DGCR8) complex, in other words Microprocessor complex and ~70 nucleotide (nt) long pre-miRNAs are generated. Nuclear export factor exportin 5 binds to nuclear pre-miRNAs from the 3' overhang. They are transferred from nucleus to cytoplasm and the cytoplasmic RNase III Dicer catalyses the production of miRNA duplexes. RNA-induced silencing complex (RISC) removes one strand of the miRNA duplex. The single stranded miRNAs are resulted to be partially complementary to target mRNA from its 'seed' sequence from the 5' end to the 3' UTR of mRNA targets (Figure 2).

miRNA genes can be observed in animals, plants, protists and viruses and they are one of the largest gene family [10]. miRBase a miRNA database has been constructed for collecting existing or discovered miRNAs. The latest release of the miRNA database (miRBase) has catalogued 2,588 miRNAs in humans, and not all miRNAs' functional importance has been understood, most of the miRNA annotations are still need to be determined [11,12].

miRNA sequences are hidden in different genomic regions. In humans, although there exist some miRNAs which are encoded by intergenic (exonic) regions, most of the accepted miRNAs are generated by introns of transcriptional units. Some miRNA genes have the same promoter with their host gene. In this case miRNA genes have been detected to be in the introns of protein-coding genes. The miRNAs in the same transcription unit are called clusters and are generally co-transcribed. Generally, several miRNA loci constitutes a polycistronic transcription unit [13]. Transcription regulation is not the only regulation mechanism for miRNAs. Individual miRNAs can also be regulated at the post-transcriptional level. In addition to this, it has been revealed that miRNA genes generally have more than one transcription start sites and that the promoters of intronic miRNAs can be sometimes different from the promoters of their host genes [14,15].

Transcription of miRNAs is mainly controlled by RNA Pol II, and transcription factors associated with RNA Pol II protein [16,17]. Transcription factors are known to regulate the expression of miRNAs [18,19] and there may be even more interesting cases in regulation of miRNAs by TFs. For example, there is a feedback loop between PTEN and has-miR-21 in which PTEN directly regulates the hsa-miR-21 and hsa-mir-21 regulates the expression of PTEN.

Apart from TFs, also epigenetic regulators, such as DNA methylation in miRNAs' respective promoter regions and histone modifications in transcription sites also have regulatory affect in miRNA expression (Figure 1) [20].

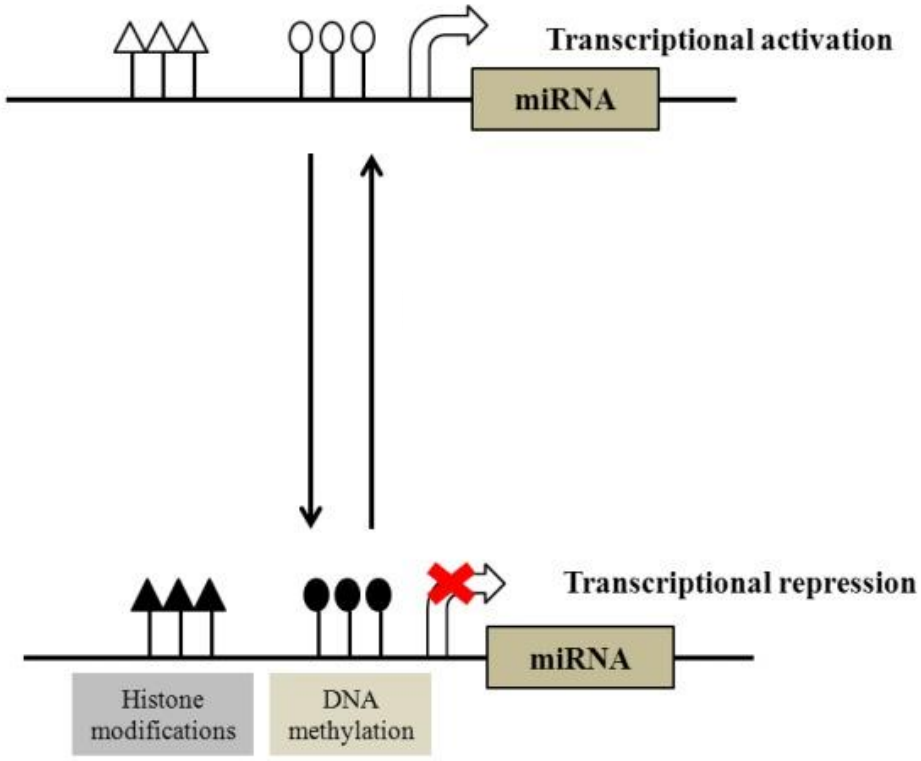


Figure 1: DNA Methylation and Histone Modifications play critical role in miRNA transcription. Republished from the original publication [21].

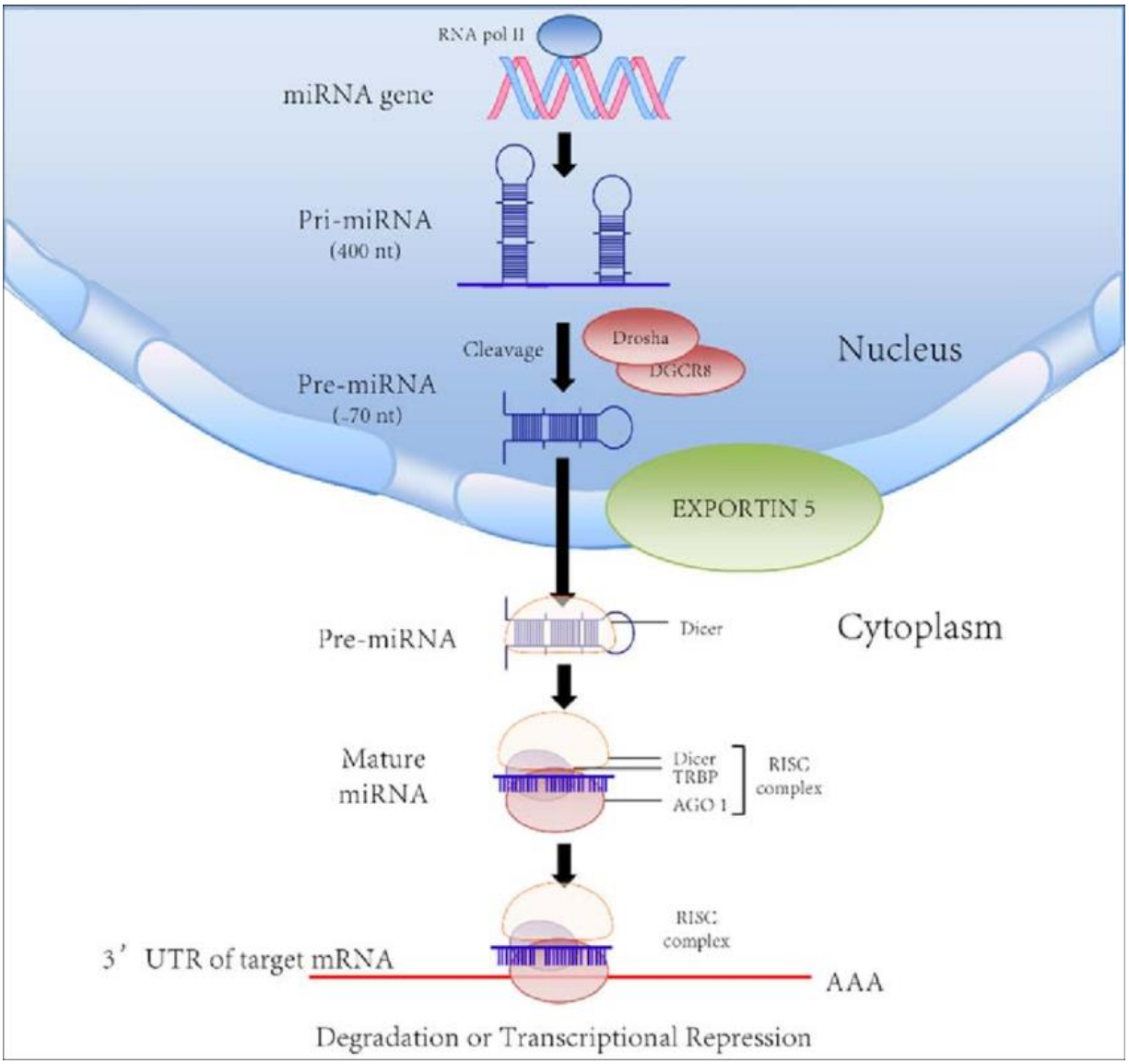


Figure 2: Schematic model of microRNA (miRNA) biogenesis.
Republished from the original publication [22].

1.2.2 miRNA Nuclear Processing

Following transcription in the nucleus and formation of pri-miRNA transcripts, they need to be converted to the mature forms. pri-miRNA is over 1 kb and contains a stem–loop structure and harbors the mature miRNA sequences in it. Pri-miRNA stem length is 33–35 bp, and it has a terminal loop and single-stranded RNA sites at the 3′ and 5′ regions. The Drosha crops the stem-loop and a small hairpin-shaped RNA of ~65 nucleotides in length (pre-miRNA) is released [23]. Drosha with its cofactor DGCR8, forms a protein complex called, the

Microprocessor complex. Drosha is a nuclear protein and is effective on double-stranded RNA (dsRNA). It belongs to the family of RNase III-type endonucleases. Drosha and DGCR8 are conserved in mammals and together they fractionates at 650 kDa [24,25].

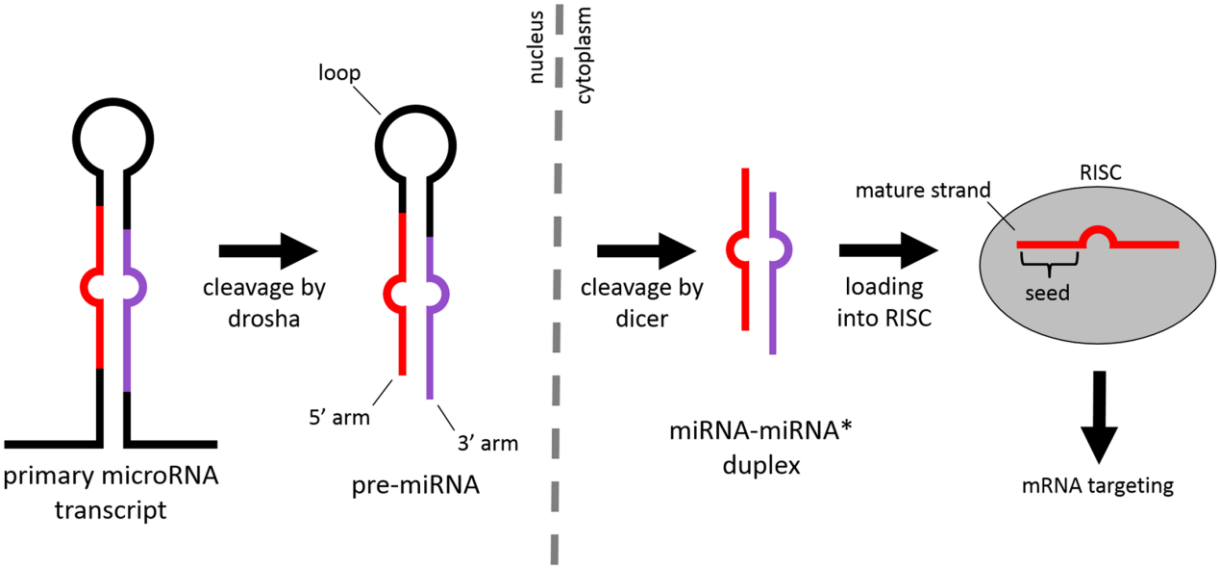


Figure 3: Translocation of microRNA from nucleus to cytoplasm

Drosha cleaves pri-miRNA to the hairpin structured pre-miRNA (Figure 3) [26]. Pri-miRNA processing is an important stage in defining the miRNA abundance. There are more than one regulatory mechanisms controlling the expression level, activity and specificity of Drosha and DGCR8. Post-translational modifications can affect the protein stability [27,28], nuclear localization [29] and processing activity of Microprocessor [30]. But, it is still ambiguous how Drosha and DGCR8 participate in the maturation process of pri-miRNA.

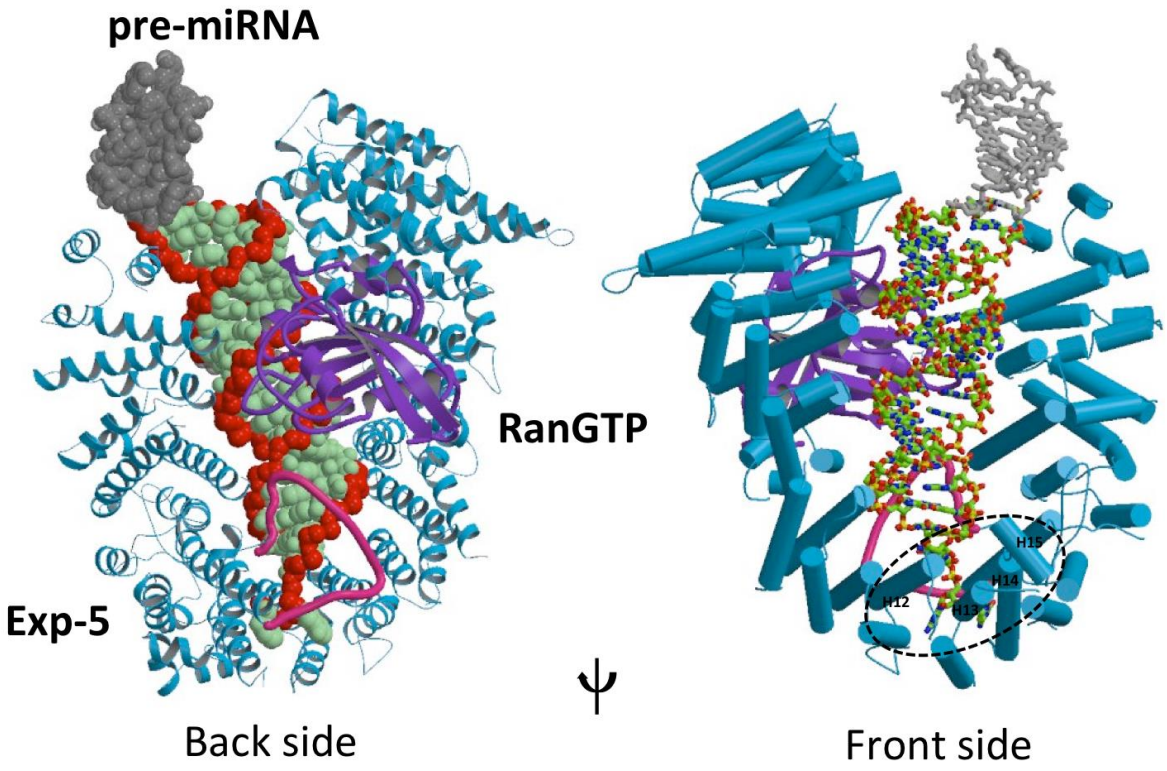


Figure 4: pre-miRNA export by EXP5- RAN•GTP transport complex

1.2.3 pre-miRNA Nuclear Export

Upon Drosha processing, pre-miRNA is translocated from the nucleus to the cytoplasm by exportin 5 (EXP5). EXP5, with GTP-binding nuclear protein forms RAN•GTP and together with a pre-miRNA forms a protein complex responsible from transportation of pre-miRNA (Figure 3) [31]. After the transport to cytoplasm, pre-miRNA is released, GTP is hydrolyzed and the transport complex is disassembled.

1.2.4 pre-miRNA Processing in Cytoplasm

Following the transport of pre-miRNA to the cytoplasm, Dicer cleaves pre-miRNA near the terminal loop to a small RNA duplex (Figure 2) [24].

1.2.5 RNA-induced silencing complex (RISC) formation

Following the formation of the small RNA duplex by Dicer, AGO protein binds to miRNA-miRNA* duplex and after passenger strand ejection, together they form the effector complex

named as RNA-induced silencing complex (RISC) (miRNA* stands for the passenger strand). RISC assignment has two sequential steps: 1) the loading of the RNA duplex and 2) unwinding of the miRNA duplex. miRNA duplexes are loaded onto AGO proteins and AGO protein selects only one of the strands as a guide which will also be its stablemate until the end of its life. After loading, the pre-RISC (in which AGO proteins associate with RNA duplexes) removes the passenger strand to generate a mature RISC. Another mechanism which is used more frequently is the unwinding of miRNA duplex without passenger strand cleavage because most of the miRNAs cannot match and bind completely to AGO protein because of the central mismatches. That's why human AGO1, AGO3 and AGO4 do not have slicer activity [32,33,34]. But, it also indicates that AGO protein family is capable to be coordinated with different types of RNAs [35]. Thus, miRNA passenger strand cleavage although seems to be the general process, there are many cases showing miRNA duplex unwinding without cleavage is preferred in miRNA processing. In miRNA duplex unwinding mechanism without cleavage, there exists mismatches in the guide strand at nucleotide positions 2–8 and 12–15 which trigger unwinding of miRNA duplexes [36]. miRNAs have important roles in diverse regulatory pathways so that it is explicable why they are strongly connected to signaling pathways. TFs and miRNA processing molecules are under the control of cell signaling. That is why it is important to uncover the relationship between signaling molecules and upstream and downstream of miRNAs to understand the miRNA biogenesis.

Previous studies showed that miRNAs are often involved in mechanisms like feedback loops, which support their crucial role in regulation. There are several good examples explaining their regulatory role like LIN28 proteins and let-7 in mammals. It is observed that, let-7 maturation is blocked by LIN28 proteins and let-7 downregulates LIN28 proteins by binding to their 3′ UTR [37]. Furthermore, MYC is one of the targets of let-7 and it is known that MYC activates the transcription of LIN28 proteins in mammals [38]. It can be concluded that, there is a regulatory loop mechanism among LIN28, MYC proteins and let-7. Hence, it will be interesting to identify additional miRNA regulatory mechanisms as their wide coverage of protein coding genes make them interesting to be used in defining disease regulatory mechanism.

1.3 Regulatory Networks

Genes, proteins, signaling molecules in a cell are generally in a system of interacting network modules like biological pathways. By working systematically with each other, the biological system can actualize its biological functions. Proteins by binding to each other can form a stable protein complex to regulate gene expression or instead they can interact with each other to generate biological signals. Similarly, regulation of number of genes involved in the same biological process may be in homeostasis with each other so that they can respond effectively to different biological conditions. They are some good examples explaining the modularity of interactions. Revealing the transcription process of co-regulated genes and the regulatory mechanism of expression of genes encoding proteins in a biological system would be a significant approach to study biological mechanisms underlying various cell activities. High throughput microarray and RNA-sequencing techniques have been developed for genome-wide profiling of transcriptomes under different biological conditions. The analysis of these profiles can provide information about gene expression reflecting gene regulation activities. These techniques give important data to develop and test new computational models or tools that can reveal transcriptional mechanisms of different molecular processes [39]. There are number of computational methods developed for this purpose and constructing gene regulatory networks using gene expression data is one of the important approaches that is used by different computational models [40,41]. By these methods [40] it becomes possible to combine multiple omics data such as transcriptomics, metabolomics, proteomics etc. to reveal the description of the complex systems with its regulators and the elements. But, it was not enough to integrate the data in transcriptional level only to understand the function and structure of regulation mechanism. It is understood that both physical and genetic interaction of molecules are important when speaking of complex biological systems. In recent years, molecular network construction, such as transcription regulatory networks and protein-protein interaction networks (PPINs) have driven interest but further development of networks is essential. There exist many concepts focusing on detecting topological, structural and architectural properties when analyzing the network. However, although the PPINs and transcription regulatory networks

have been constructed for identification of pathways and modules, they are not sufficient enough to integrate important post-transcriptional regulations.

1.4 Role of miRNAs in Human Organism

Transcription factors (TFs) contribute to biological processes at the transcription level of the genes and TFs are not the only regulatory factors of gene expression. Compared to transcriptional regulators, miRNAs act as posttranscriptional regulators, being active in the cytoplasmic compartment. They disturb/cancel out the effect of upstream processes of transcription in the nucleus. They are capable of regulating transcripts in different special tissues. They can also be in high concentrations around 10.000s of molecules in a cell, providing stableness [42].

In recent years, studies suggest that miRNAs play critical roles in a variety of essential biological processes that is why disruptions in the expression of miRNAs would effect cell functions such as cell cycle regulation, differentiation, development, metabolism, neuronal patterning, aging etc. [6]. It is determined that miRNA-gene, TF-miRNA relations and regulations are complicated and also evolutionarily conserved [43,44]. Although miRNAs represent only about ~1% of the genome, their authority in regulating gene expression is undeniable. Different from the mechanism of complete base pairing between miRNAs and the mRNA, multiple miRNAs can synergistically regulate one or more pathways [45,46]. It has been also shown that, a single miRNA can bind to more than one mRNA, in other words a target gene can be targeted by multiple miRNAs [47]. Different tissues or a specific tissue under different conditions would have different miRNA expression profiles as well. Therefore, with increasing evidences it is revealed that, deregulations of miRNAs are responsible and effective in the development of various human diseases like cancer and neurological disorders. The different expression levels of miRNAs affect the initiation, progression and metastasis of different cancer types such as breast cancer [48], lung cancer [49], prostate cancer [50], colon cancer [51], ovarian cancer[52], brain cancer[53]. New disease related-miRNAs are emerging with the new results coming up from the experimental literature.

Thereby, miRNAs have become an important potential biomarker for understanding the molecular mechanisms of complex diseases leading to obtain new potential biomarkers for the diagnosis, treatment, prognosis and potential drug targets in drug discovery and clinical treatment.

1.5 Approaches for Detecting miRNA-Disease Regulatory Relations

In the past few years, based on the assumption that miRNAs which have similar functions are generally related to similar disease and vice versa, studies have been focused on developing computational methods to infer potential miRNA-disease associations. [54] developed a model which uses hypergeometric distribution on the integrated data which includes miRNA functional interactions network, disease phenotype similarity network and the known phenome-microRNAome network and the prediction accuracy is not that high. [55], again makes predictions about miRNA-disease associations by integrating the functional link information between miRNA targets and disease related genes in protein-protein interaction network. But, these methods both strongly rely on the predicted miRNA-target interactions, that is why they have high number of false positive and false negative results.

Apart from these methods, RWRMDA [56] and HDMP [57] have given good results for miRNA-disease association prediction, the only obstacle about them is, they cannot be applied to the diseases without related miRNAs. RWRMDA uses the implementation of random walk on the miRNA functional similarity network and it does not rely on predicted miRNA-target interactions. HDMP predicts potential miRNAs associated with human disease based on weighted k most similar neighbors.

In addition to miRNA-disease regulatory networks, miRNA-regulated networks are such as miRNA co-regulated networks, miRNA-mRNA networks and miRNA-TF networks are studied. On the other hand, research on miRNA-regulated protein-protein interaction networks have barriers because of both the complex working mechanism of miRNAs and complexity of protein-protein interactions.

1.6 miRNAs and Protein-Protein Interaction Networks

For the continuation of biological functions like DNA replication, transcription, translation, signal transduction, protein-protein interaction (PPI) is inevitable for a living cell [58]. PPI can be represented as an undirected graph structure with topological properties like edges, nodes and clusters and mathematical and computational analysis can be applied to understand the organization of the cell [59].

In 1989, the yeast two-hybrid system was introduced to construct PPI networks[60]. In 2000, first PPI network of yeast was published [61] and in 2005 first human PPI network was released [62]. Recently, PPI network studies generally focus on PPI network detection and prediction [63], signal transduction pathways[64,65,66], protein function prediction based on PPI networks and protein complex prediction in PPI networks [67,68].

Studies about miRNA-regulated PPI networks are developed mainly in two areas: a) revealing the correlation between miRNAs and protein-protein interaction networks, using bioinformatics approaches and statistical means. This method tries to find new miRNA-regulated gene expressions beside seed matching. The unfavorable things about these studies are, they suffer from poor coverage rates, false positives and false negatives; b) identification of the impact of miRNA regulation on PPI networks in diseases is the second way of developing miRNA-regulated PPI networks. Signal transduction pathways are one of the important components of PPIs and they are the primary factors of miRNA targeting modulators in animal cells [69].

miRNAs can serve as mediators of crosstalk between signaling pathways [69] and it can be understood that miRNAs act as an indirect regulator in PPI networks. Additionally, as signaling pathways are the most important sub-graphs of the PPI network, understanding the miRNA-regulated signaling pathways relationship mechanism becomes very important.

Causal interactions between proteins are not that easy to capture in a structured format but it is obvious that it would be more informative for representing the direction and sign of information flow in signal transduction. Recently, it is difficult to construct activity flow diagrams with sufficient high coverage rates and to support each interaction with experiments. To handle these

considerations a new tool called SIGNOR has been developed, capturing causal interactions between proteins [70]. It offers a comprehensive network of experimentally validated functional relationships between signaling proteins. During writing this thesis work, SIGNOR has about 16,000 manually curated interactions connecting about 4,000 biological molecules like chemicals, metabolites, proteins or protein complexes which have significant role in signal transduction pathways[71]. SIGNOR is a source of signaling information and uses the functional relevance information of two interactors according to the probability of their citation in the same paper. It stores the causal relationships as lists of interactions between two molecules. One of the molecule would be the regulator and the other would be the regulated molecule. Most of the molecules in the network are proteins but other chemicals, phenotypes, stimuli, complexes and protein families are included as well. That is why it provides comprehensive directional interaction information for data analysis, computational modeling and prediction.

In this thesis, only the protein entities are used for constructing the protein-protein interaction network directionally.

INTRODUCTION

Neurodegenerative diseases, are today's one of the most important groups of diseases [72] that have a high impact on society because of their high incidence, mortality and decrease in the quality of living.

Huntington's and Parkinson's Diseases (HD and PD) are neurodegenerative disorders. In one hand, they all share a similar ability to cause damage when they capture brain cells, on the other hand the specific proteins and types of neurons are affected differently.

Transcriptional dysregulation has been observed in HD and PD [73]. Transcription, neuroinflammation and developmental processes are dysregulated in the brains with HD and inflammation and mitochondrial dysfunction were obtained in the brains of patients with PD[73].

Understanding the molecular mechanisms underlying complex diseases (in this case HD, PD neurodegenerative disorders) is necessary for the diagnosis and treatment of the disorders. It is therefore important to detect the most important genes and miRNAs and studying their interactions for recognition of disease mechanisms. It seems that miRNAs are involved in deregulation of neurodegenerative diseases[74]. Many studies demonstrated the expression of specific miRNA in the central nervous system (CNS) with different roles. Therefore, a comprehensive study in miRNAs involved in neurodegenerative diseases could be conveniently used in innovative therapies.

The aim of this study is by focusing on miRNAs involved in HD, PD and their target genes, to determine the most important miRNAs, TFs, genes and their pathways in the diseases. In this way, a systematic analysis of the mechanism of HD and PD is done to understand biological processes common to all of them and differences if there is any. In this model, disease specific (HD and PD) transcriptional and post-transcriptional regulatory pathways, using disease related miRNA and mRNA databases and mRNA and miRNA expression profiles were identified (Figure 5).

For this purpose, to obtain stable signatures we identified disease related differentially expressed genes (DEGs) only in the prefrontal cortex of the brains of HD and PD human subjects compared to neuropathologically normal control brain tissues using mRNA-Seq. In addition to this, we also identified differentially expressed miRNAs (DEmiRs) in the prefrontal cortex of the brains of HD and PD and in the parietal lobe cortex of the brains of AD as it is the only miRNA expression analysis done in AD.

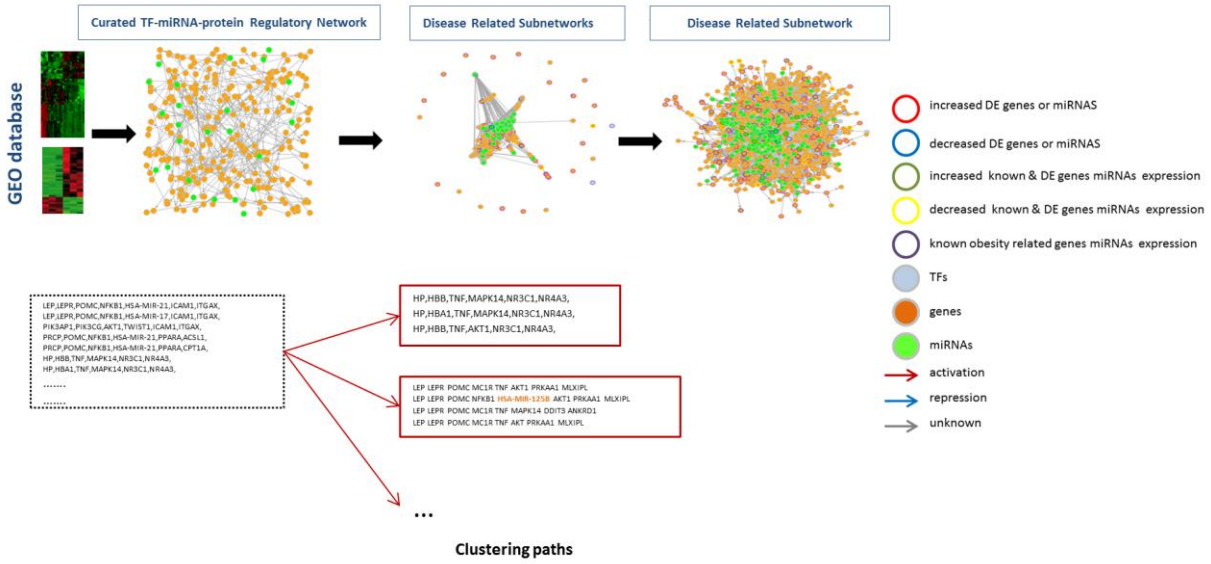


Figure 5: Overview of the proposed approach.

In addition to this, breadth-first-search (BFS) algorithm was used to find the disease related pathways of a complex regulatory network which is constructed by using directed protein-protein interaction network, TF-miRNA, miRNA-mRNA, TF-gene relations. Consequently, these pathways may contain non-DE genes and miRNAs as well. To attain the significance scores of the potential pathways hypergeometric test was used. Resulted significant pathways were clustered according to their resemblance and KEGG pathway analysis was done to reveal the functional enrichment of the genes and miRNAs in the final disease related network.

MATERIALS & METHODS

3.1 Studying RNA-seq data

HD and PD are complex diseases, and different brain regions of these diseases have diverse gene expression patterns[75]. That is why, to get accurate results and to compare truly, we searched the GEO database for RNA-seq data with the same brain tissue for each disease. Expression profiles of GSE64810 for HD and GSE68719 for PD have been used. For HD, analysis was done by next-generation sequencing in human (BA9) in 20 HD and 49 neuropathologically normal individuals using Illumina high-throughput sequencing[76]. For PD, brain tissue from the prefrontal cortex Brodmann Area 9 of 29 PD and 44 control samples were used and any AD-type pathology beyond normal signs of aging were excluded[77].

Differentially expressed genes with adjusted p-value less than 0.0002 were selected.

3.2 Differentially Expressed miRNAs in HD, PD

High-throughput techniques to investigate miRNA expression in HD and PD have rarely been used. For HD, GSE64977 with 26 HD patients, 49 neurologically normal control prefrontal cortex samples are used. For PD, GSE72962 with 29 PD patients, 33 control prefrontal cortex samples are used.

	Huntington's Disease (HD)	**Parkinson's Disease (PD)**
RNA-seq Gene Expression Data	GSE64810	GSE68719
miRNA Expression Data	GSE64977	GSE72962

Table 1: Directed Protein-Protein Interaction Data

To attain directed PPI disease data, we used SIGNOR (SIGnaling Network Open Resource) database[71]. The output of SIGNOR database provided us to construct the directed graph

between signaling entities. We created directed PPI network of 12315 interactions from 4627 nodes.

3.3 Identification of Transcription Factors (TFs)

For the identification of TF in the directed PPI network, union of TRANSFAC (version 11.4) and TRED databases are used[78,79].

Within the curated disease specific regulatory network, all the self-loops were removed from the graph and if there were more than one interaction with same directionality between two nodes, the interactions were represented with a single edge.

3.4 Identification of HD, PD related miRNAs and genes

Disease related experimentally verified genes are obtained from the database DISGENET[80]. The disease genes presented in DISGENET which offers one of the most comprehensive collections of human gene-disease associations. For each disease, genes with DISGENET PMID score >=2 are selected.

Disease related experimentally verified miRNAs are derived from HMDD[81] and the miR2Disease database[82]. Both HMDD and miR2Disease databases collect the miRNA-disease associations manually from experimentally verified published data.

3.5 TF-miRNA-mRNA Regulatory Network Construction

The construction of curated TF-miRNA-mRNA regulatory network was done by combining various databases. Four data sources were used: a) TransmiR database (version 1.2) represented the curated TF-miRNA relations [83] and; b) miRTarBase database (version 4.5), c) miRecords (version 3); d) TarBase (version 5.0) represented the curated miRNA-mRNA regulations [84,85,86].

3.6 Construction of Regulatory Subnetwork of HD, PD

RNA-seq method provides important capabilities like high resolution and broad dynamic range and it enriches to the progress of transcriptomics research. Important amount of data was

detected as a result of this sequencing method. It is known that RNA-seq data is complex and it is not easy to get meaningful results from a huge data [87]. To hold the information about disease related genes and miRNAs with DEGs and DEmiRs, we mapped them with their third-degree neighbors to construct the TF-miRNA-mRNA regulatory subnetwork. The nodes represent TFs and miRNAs which were in the databases and the edges represent the regulating relationships between miRNAs, TFs and genes. To get a global view of this subnetwork, we used R, igraph package.

3.7 Pathway Analysis of Disease Regulatory Networks

The subnetworks of each disease have complex structures, although they are simplified from the background TF-miRNA-gene network. To get meaningful information from this complex network structure, regulatory pathways which include multiple TFs, miRNAs and target genes were considered first. Identification of regulatory pathways in HD and PD by uncovering transcriptional and post-transcriptional regulations, revealed the molecular regulatory mechanisms.

The regulatory cascades are detected by using the shortest path algorithm in the package igraph [88]. shortest.path() function uses Breadth-First Search Algorithm (BFS).

breadth first search:

```
choose some starting vertex x
mark x
list L = x
tree T = x
while L nonempty
choose some vertex v from front of list
visit v
for each unmarked neighbor w
        mark w
        add it to end of list
        add edge vw to T
```

BFS is one of the most important and fundamental algorithm used to traverse graph structures. The breadth first search tree holds a list of nodes to be added to the tree. It starts traversing from the selected source node. Algorithm traverses the graph layer by layer by visiting the neighbor nodes directly connected to the starting node (Figure 6).

The directed regulatory subnetwork was scanned and all the paths between every two 0-indegree and 0-outdegree differentially expressed genes (DEGs) and differentially expressed miRNAs (DEmiRs) with more than two nodes were identified (Figure 7).

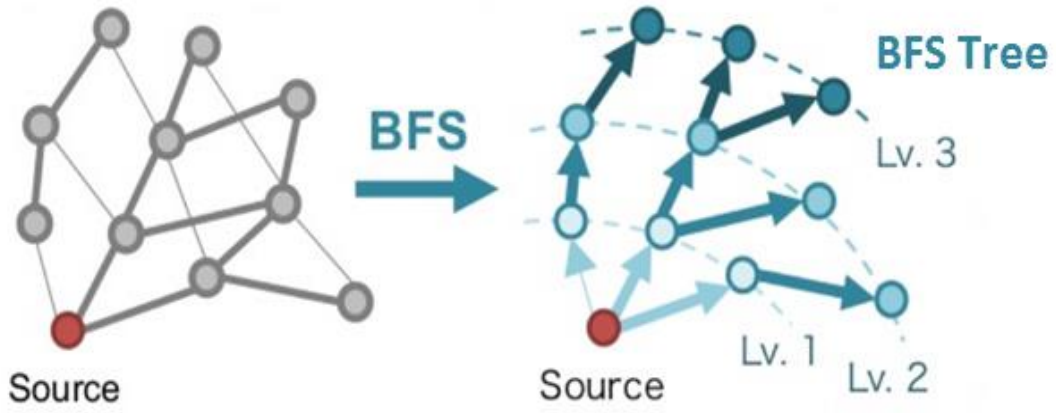


Figure 6: Breath-First Search Algorithm

3.8 Evaluation of Disease Related Cascades/Pathways

For each cascade between DEGs and DEmiRs, we evaluated a coverage rate (CR). CR value is calculated to determine the relationship strength of the pathways identified and the disease of interest. CR value is calculated as,

$$CR = \frac{ND}{NT} \tag{3.1}$$

Equation 1: *ND* represents number of disease related nodes, *NT* represents the length of the cascade

To evaluate the statistical significance of the CR value, hypergeometric test is used. Hence, rate of observing if CR value is likely to occur by chance or not is evaluated.

$$f(k) = \frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}} \quad (3.2)$$

Equation 2: Hypergeometric test calculates the probability of k successes in n selections with replacement. N represents the population size, k represents the number of successes and n represents the sample size. $\binom{N}{n}$ represents the number of ways a sample of size n can be selected from a population of size N. $\binom{N-M}{n-k}$ represents the number of ways $n - k$ failures can be selected from a total of $N - M$ failures in the population. $\binom{M}{k}$ represents the number of ways x successes can be selected from a total of r successes in the population

Finally, multiple testing correction via false discovery rate (FDR) was performed using Benjamini-Hochberg procedure and assigned to pathways. Cascades which have FDR value smaller than 0.2 are selected as functional disease related pathways.

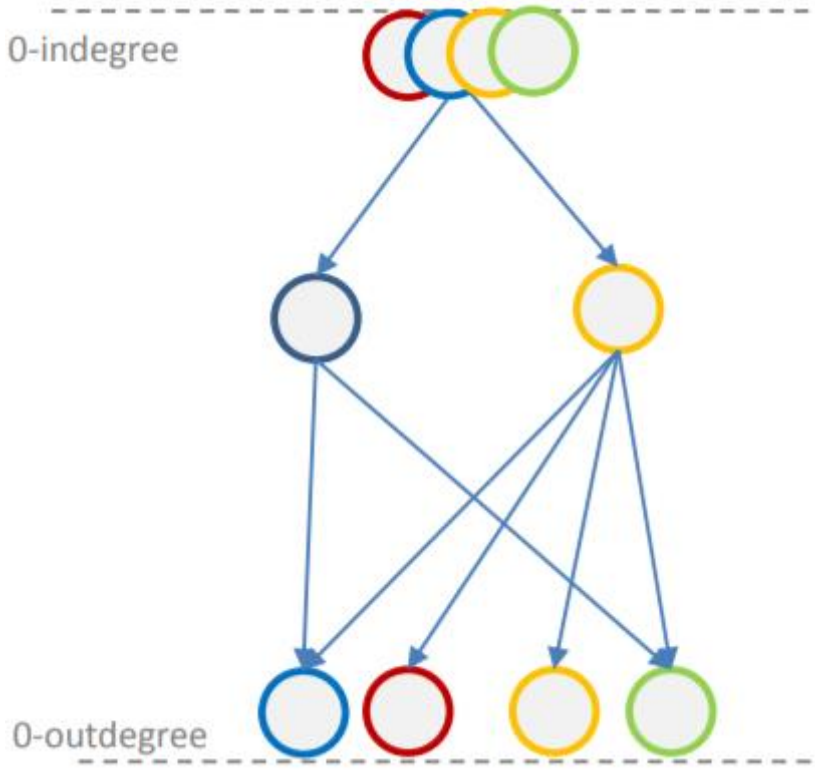


Figure 7: Pathways between 0-indegree and 0-outdegree nodes are determined

3.9 KEGG Pathway Analysis of Disease Related Cascades

Our method groups the potential pathways according to their resemblance. If the sequences of cascades are 50% the same, then those are put to the same group. To determine the functional relation of the groups with the related disease, PathFindR pathway analysis was done to the genes involved in each subgroup [89].

Databases	**Main Feature**
TransmiR	the experimentally validated microRNA-target interactions database
miRTarBase	the experimentally validated microRNA-target interactions database
miRecords	manually curated database of experimentally validated miRNA-target interactions
TarBase	manually curated database of experimentally validated miRNA targets
TRANSFAC	the database of eukaryotic transcription factors
TRED	a transcriptional regulatory element database
HMDD (the Human microRNA Disease Database)	a database of curated experiment-supported evidence for human microRNA (miRNA)
miR2Disease	a manually curated database, aims at providing a comprehensive resource of miRNA deregulation in various human diseases
DISGENET (v5.0)	collections of genes and variants associated to human diseases

Table 2: Database list used for Disease Related Network Construction

RESULTS

4.1 Disease Related Regulatory Network Construction

The Signaling Network Open Resource (SIGNOR), warehouses the signaling information in a structured format. It stores only the interactions that were validated in the scientific literature. The captured information is stored as cause and effect relationship between the source molecules and the target molecules. By this means, this structured format can be represented as a directional network. The information can be downloaded from (https://signor.uniroma2.it/). The network is constructed by using R, igraph package. There were 4731 number of unique nodes and 12447 number of unique interactions.

MiRNA-gene, TF-miRNA experimentally validated relations were downloaded from TransmiR, miRTarBase, miRecords and TarBase databases. There were 2829 number of relations integrated to directed PPI network. With the addition of new relations, network was extended. miRNA regulatory network had 5241 number of nodes with 15276 number of unique relations. There were 468 number of TFs, 4231 number of genes and 392 number of miRNAs in the extended regulatory network. TFs were detected by using TRANSFAC and TRED databases. Figure 8 shows the TF-miRNA-gene directed regulatory network.

From GEO database, GSE64977 miRNA expression profile and GSE64810 gene expression profile were used for Huntington's Disease. 20 HD patients and 49 neuropathologically normal controls were analyzed for genome-wide analysis of mRNA expression in human prefrontal cortex using next generation high-throughput sequencing. For Parkinson's Disease (PD), GSE72962 for miRNA and GSE68719 for gene expression profile were used. 29 PD and 33 neuropathologically normal controls were included for genome-wide analysis of mRNA expression in human prefrontal cortex using next generation high-throughput sequencing. Genes and miRNAs with FDR values smaller than 0.0001 were selected. Differentially expressed miRNAs and genes were identified and separated according to their increased and decreased expressions (Table 3). DE genes and miRNAs mapped to the network. The HD network had 191 number of increased DE genes/miRNAs and 33 number of decreased DE genes/miRNAs. The PD network had 42 number of increased and 34 number of decreased DE genes/miRNAs.

	Huntington's Disease		Parkinson's Disease	
	DE miRNAs	DE genes	DE miRNAs	DE genes
increased expression	26	165	31	11
decreased expression	16	17	33	1

Table 3: Differentially expressed miRNAs/genes for Huntinton's Disease and Parkinson's Disease with their increased, decreased information

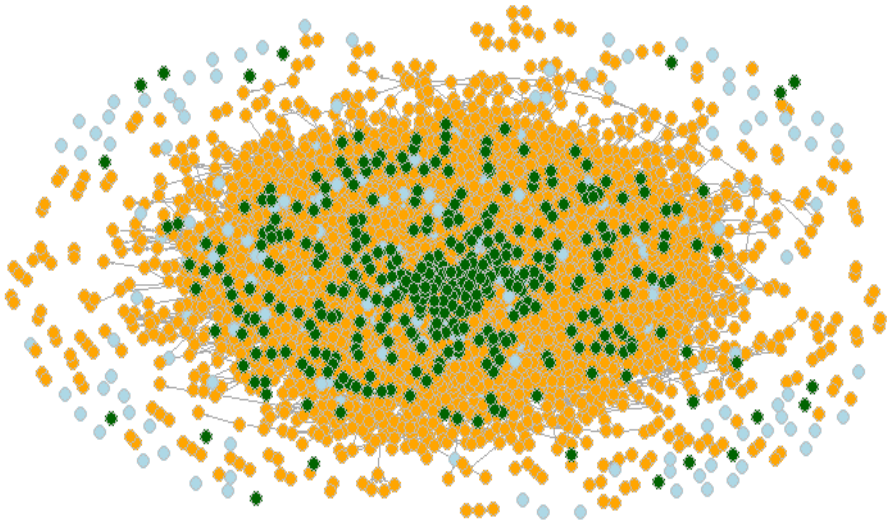


Figure 8: TF-miRNA-gene Directed Regulatory Network

The potential disease specific TF-miRNA-mRNA regulatory subnetwork was constructed by connecting all disease related nodes which were comprised from DE genes/miRNAs and their 3rd degree neighbours. The subgraph for HD had 4724 number of nodes and 14922 number of relations. The subgraph for PD had 4474 number of nodes and 14605 number of relations. Our tool has the options to select among 1st, 2nd and 3rd degree neighbor nodes. In this analysis we chose 3rd degree to include most of the disease related known nodes in the subnetwork. There were 634 known HD related genes [80] and 14 HD related miRNAs were detected. 352 number of them were mapped to the regulatory network and 332 number of them were included in the HD related active subnetwork (Figure 9a). There were 443 known PD related genes and 38 PD related known miRNAs, 259 number of them were mapped to the regulatory network and 235 number of them were included in the PD related active subnetwork (Figure 9b).

4.2 Identifying Disease Related Potential Regulatory Pathways

In this study, all directed acyclic paths were found by using BFS algorithm between 0-indegree and 0-outdegree DE nodes. For HD, we got 9167 and for PD we got 614 number of directed acyclic paths. The length of all the potential cascades were longer than 2 and these cascades were accepted as potential active disease related pathways.

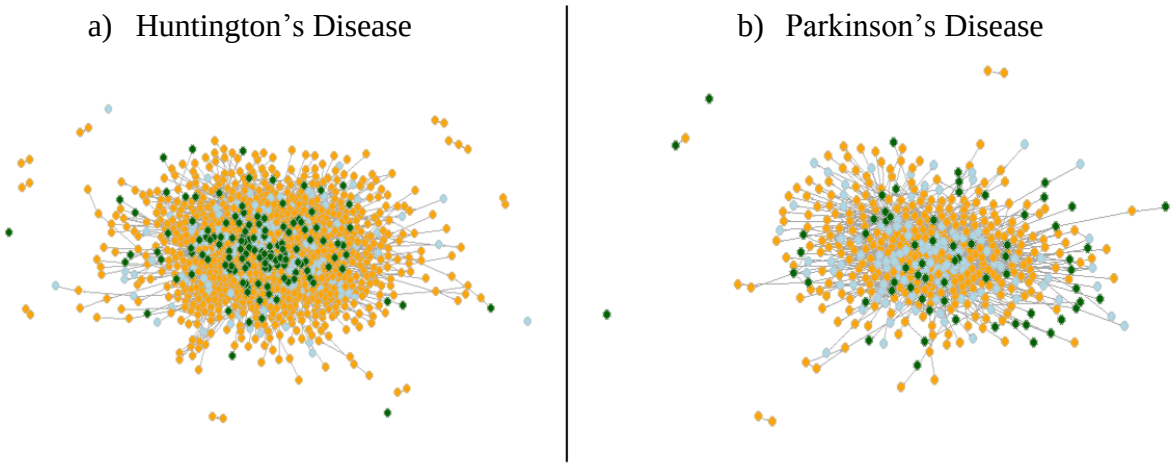


Figure 9: The orange nodes represent genes, green nodes represent miRNAs, blue nodes represent TFs. Red and Blue borders indicate increased and decreased expressions of miRNAs/genes a) Huntington's Disease (HD) Regulatory Network b) Parkinson's Disease (PD) Regulatory Network.

For each pathway CR values were calculated to measure the relevance of the pathways and the disease of interest. By applying hypergeometric test, significant pathways were selected. Multiple testing using FDR values were done and for HD we got 42, for PD we got 27 number of pathways with FDR-value < 0.2.

Significant disease related active pathways were grouped according to their similar cascades. If they had equal or larger than 50% similar cascades, they were put into the same subgroup. For HD we got 8 and for PD we got 17 number of subgroups (Figure 10,11,12) (Appendix A).

In Figure 13 and Figure 14, significant HD and PD related pathways can be observed on the HD and PD related networks. Some edges are larger than the other ones. The edge thickness was adjusted according to the frequency of the edges in the significant pathways designated.

4.3 Comparison of Cascades in miRNA Regulatory Pathways in HD and PD

Significant regulatory pathways for each disease were analyzed according to their frequent cascades. For Huntington's Disease, there were 86 and for Parkinson's Disease there were 117 number of unique relations. The common relations between HD and PD were detected and is shown in Table 4.

Common Cascades in HD and PD	45 HD significant pathways	61 PD significant pathways
BCL2L1 → CASP9	13	6
CASP9 → CASP3	13	6
CASP3 → AKT1	12	3
AKT1 → GSK3B	1	4
TP53 → FGF2	1	4
AKT1 → PRKACA	6	4

CASP3 → AKT	1	2
CDX2 → INS	2	2

Table 4: It shows the common cascades between Huntington's Disease and Parkinson's Disease

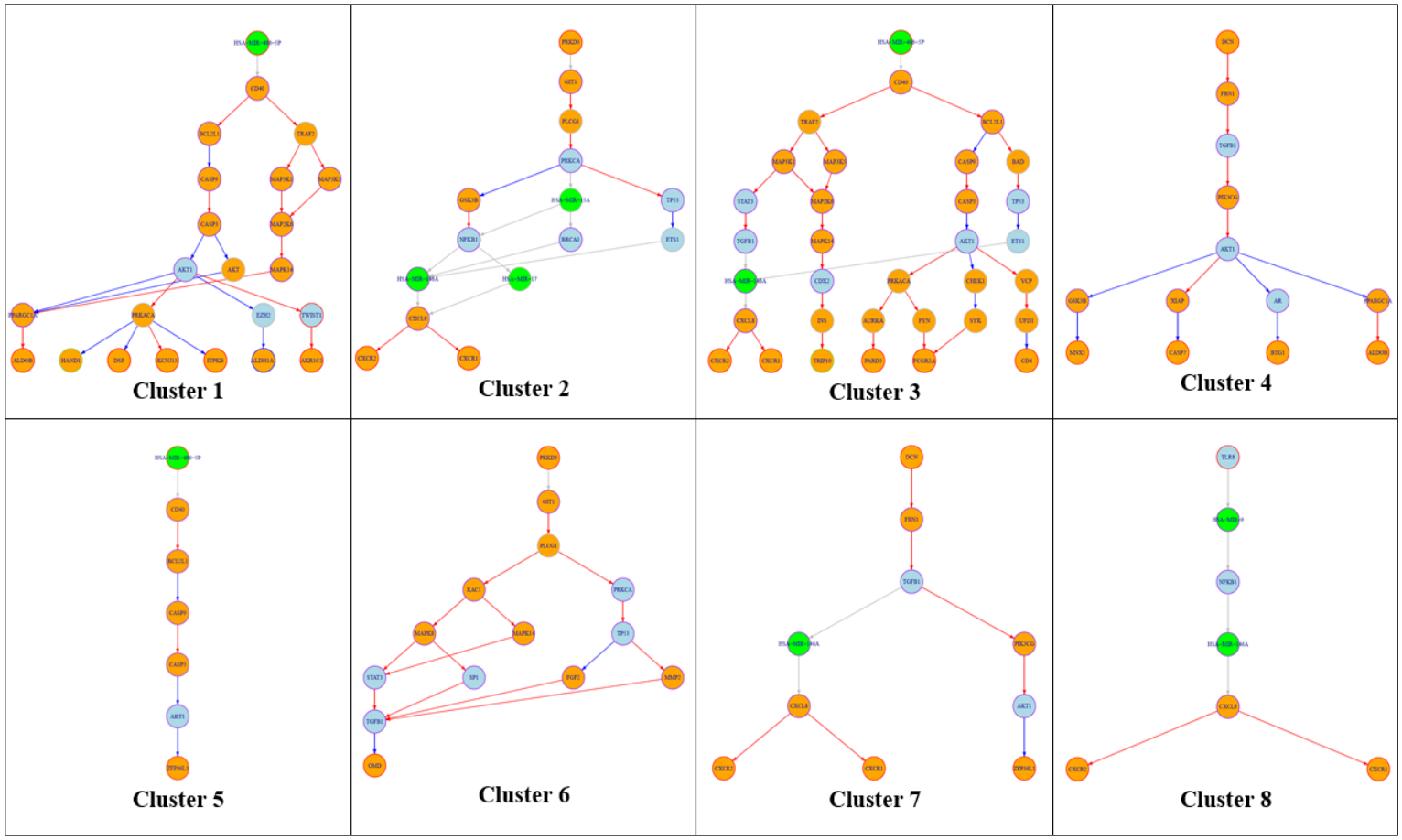


Figure 10: Huntington's Disease active pathways were grouped according to their ratio of similar relations. There were 8 number of pathway groups. Color of arrows indicates the relationship type: 1) →: activation,→: repression, →: not-known. Orange node represents the genes, green nodes represent miRNAs, blue nodes represent TFs. Red border colors indicates the increased DE, blue border color indicates repression in DE genes/miRNAs. Yellow and Green border colors show decreased and increased DE genes/miRNAs are also known to be related to disease of interest. Purple border color represents the known disease related miRNAs/genes.

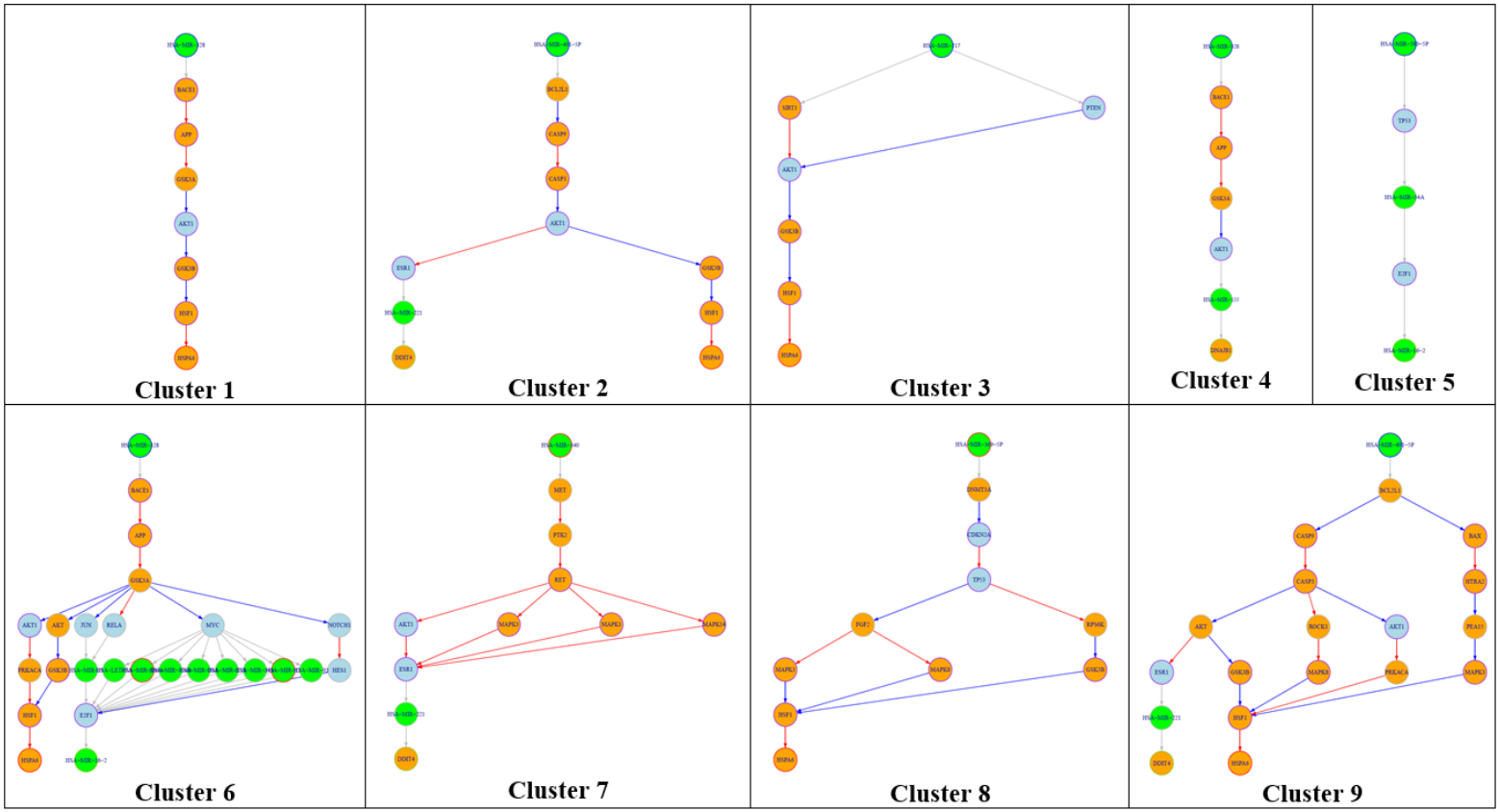


Figure 11: Parkinson's Disease related active pathways were grouped according to their ratio of similar relations. There were 17 pathway groups. 1-9 groups are shown. Color of arrows indicates the relationship type: 1) →: activation,→: repression, →: not-known. Orange node represents the genes, green nodes represent miRNAs, blue nodes represent TFs. Red border colors indicates the increased DE, blue border color indicates repression in DE genes/miRNAs. Yellow and Green border colors show decreased and increased DE genes/miRNAs are also known to be related to disease of interest. Purple border color represents the known disease related miRNAs/genes.

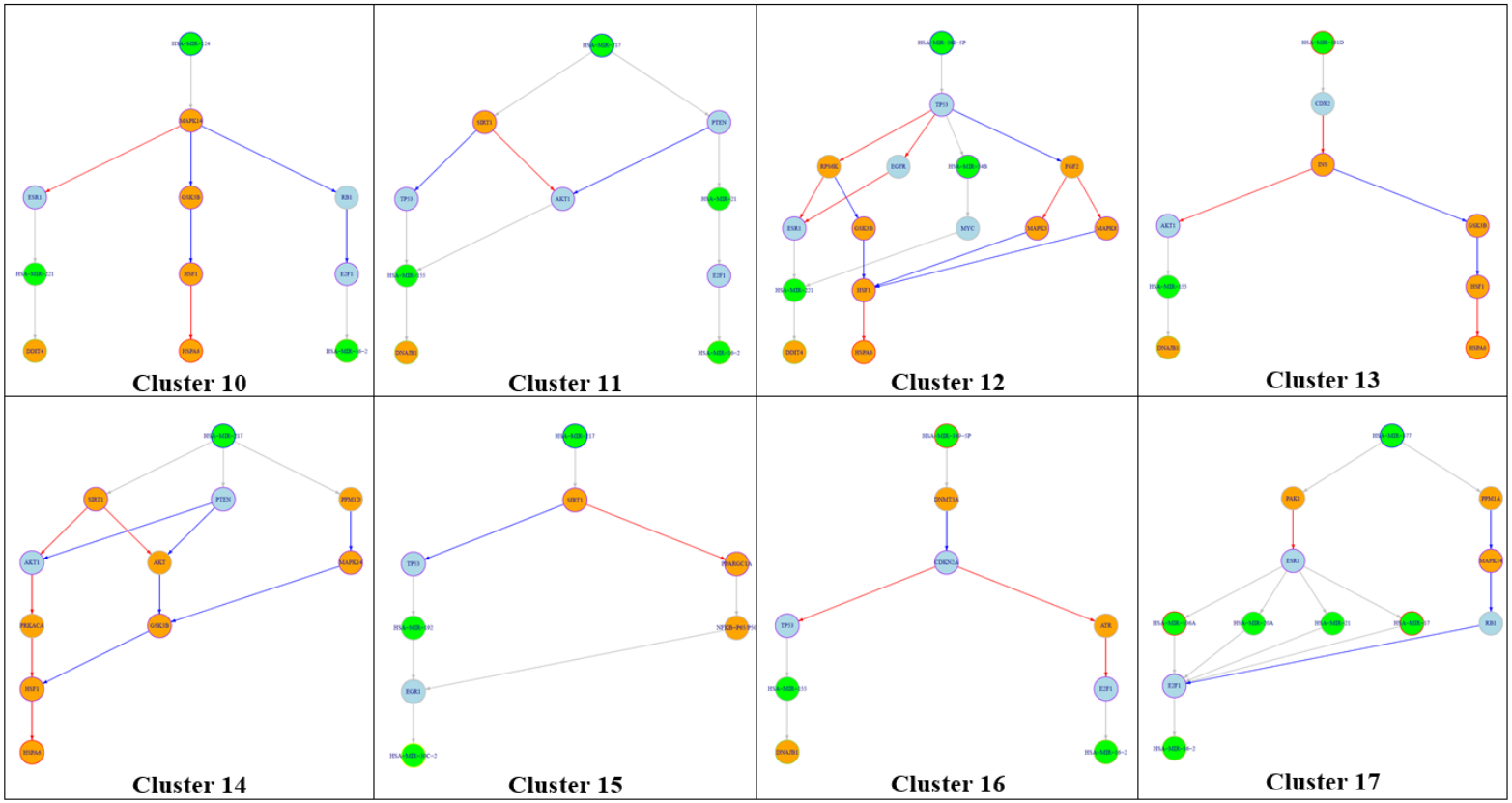


Figure 12: Parkinson's Disease related active pathways were grouped according to their ratio of similar relations. There were 17 number of pathway groups. 10-17 subgroups are shown. Color of arrows indicates the relationship type: 1) →: activation, →: repression, →: not-known. Orange node represents the genes, green nodes represent miRNAs, blue nodes represent TFs. Red border colors indicates the increased DE, blue border color indicates repression in DE genes/miRNAs. Yellow and Green border colors show decreased and increased DE genes/miRNAs are also known to be related to disease of interest. Purple border color represents the known disease related miRNAs/genes.

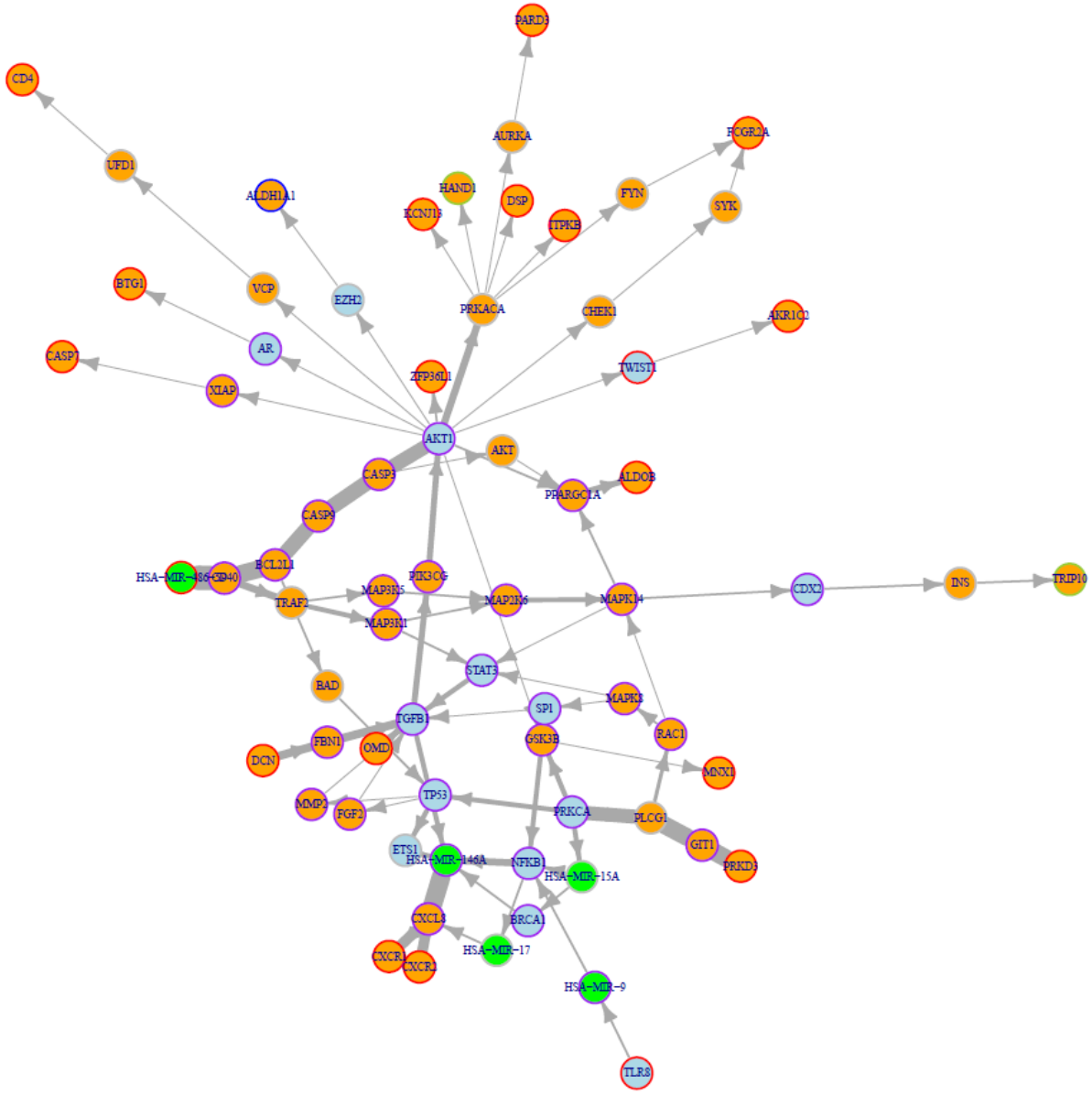


Figure 13: Huntington Disease, Significant Pathways are represented as graph. The edge width represents the frequency of relations among active pathways. Orange nodes represent the genes, green nodes represent miRNAs, blue nodes represent TFs. Red border color indicates the increased DE, blue border color indicates repression in DE genes/miRNAs. Yellow/Green border colors show decreased/increased DE genes/miRNAs are also known to be related to disease of interest. Purple border color represents the known disease related miRNAs/genes.

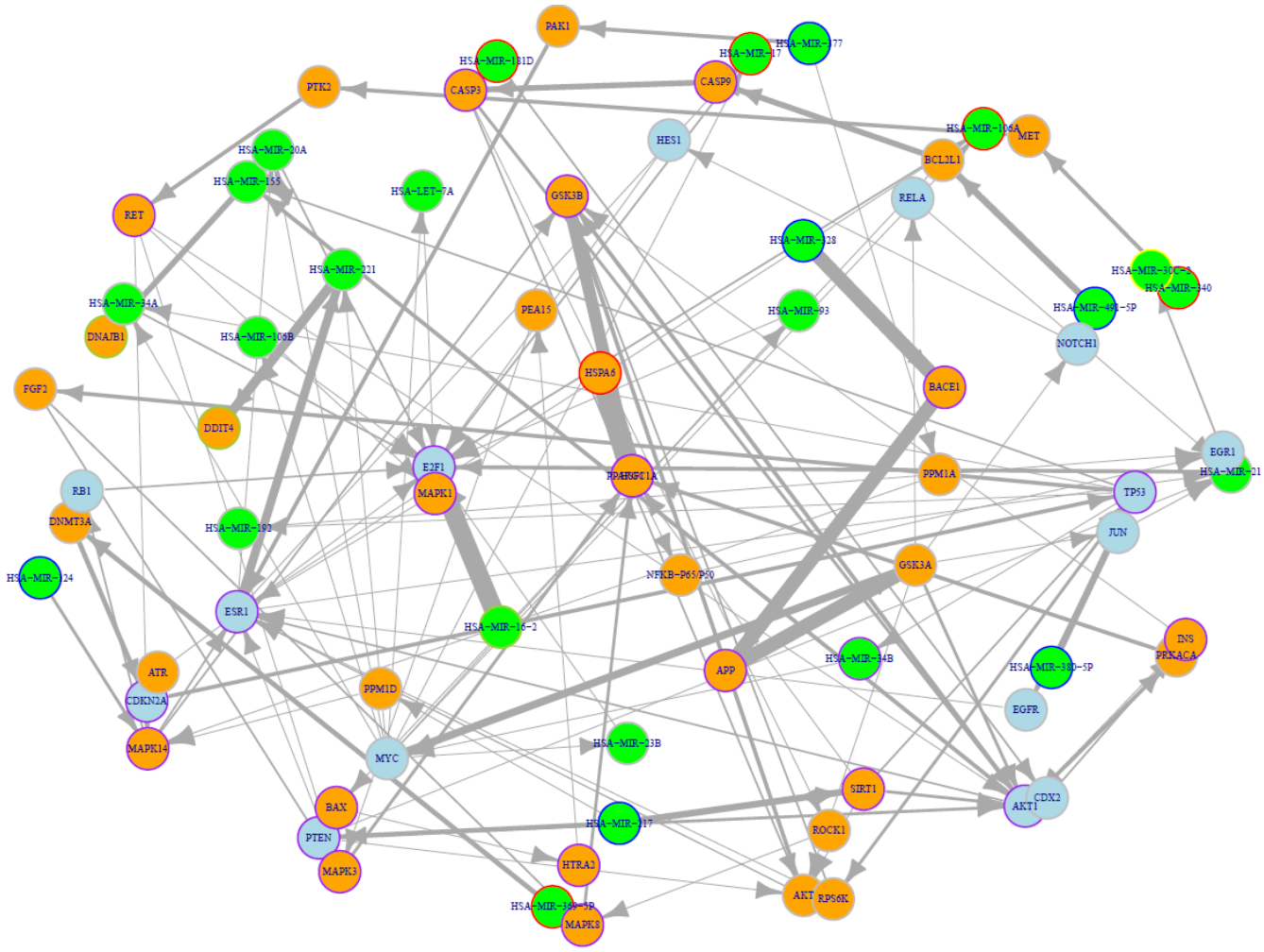


Figure 14: Parkinson's Disease, Significant Pathways are represented as graph. The edge width represents the frequency of relations among active pathways. Orange node represents the genes, green nodes represent miRNAs, blue nodes represent TFs. Red border colors indicates the increased DE, blue border color indicates repression in DE genes/miRNAs. Yellow and Green border colors show decreased and increased DE genes/miRNAs are also known to be related to disease of interest. Purple border color represents the known disease related miRNAs/genes.

4.4 KEGG Pathway Analysis of miRNAs/genes in miRNA Regulatory Pathways

The functions of miRNAs which were included in the significant regulatory pathways were predicted my using the miRpath v.3 software. miRpath, assigns pathways to the miRNA targets using KEGG database (Table 5,7) [90]. Also, KEGG pathway analysis for the genes which were in the cascades of important regulatory pathways was done by using PathFindR Tool in R (Table 6,8).

miRNAs in HD Regulatory Pathways	**KEGG Pathway Analysis**
HSA-MIR-146A **HSA-MIR-9**	Hippo signaling pathway Glycosphingolipid biosynthesis - lacto and neolacto series Protein processing in endoplasmic reticulum Glycosaminoglycan biosynthesis - keratan sulfate ErbB signaling pathway Chronic myeloid leukemia Lysine degradation Allograft rejection Measles
HSA-MIR-486-5P	Arrhythmogenic right ventricular cardiomyopathy (ARVC)
HSA-MIR-15A **HSA-MIR-17**	Proteoglycans in cancer

Table 5: Huntington Disease KEGG Pathway analysis results of the miRNAs in the significant miRNA regulatory pathways (miRpath v.3 was used). Purple colored miRNA names indicate known HD related miRNAs. Red colored miRNA name indicates DE miRNA with increase expression. Grey colored miRNA names indicate unknown miRNAs

ID	KEGG Pathway	Genes
hsa05205	Proteoglycans in cancer	**AKT1**, RAC1, STAT3, **TP53**, PRKCA, DCN, TGFB1, MMP2, **FGF2**, PLCG1, **PRKACA**, MAPK14, TWIST1, **CASP3**
hsa04010	MAPK signaling pathway	NFKB1, PRKCA, **PRKACA**, **TP53**, MAPK14, MAP2K6, **AKT1**, MAP3K5, TRAF2, **CASP3**, TGFB1, RAC1, MAPK8, MAP3K1, **FGF2**, **INS**
hsa04071	Sphingolipid signaling pathway	**AKT1**, PRKCA, MAP3K5, MAPK8, MAPK14, **TP53**, NFKB1, RAC1, FYN, TRAF2
hsa04210	Apoptosis	**BCL2L1**, **TP53**, XIAP, **CASP9**, CASP7, **CASP3**, BAD, NFKB1, **AKT1**, TRAF2, MAP3K5, MAPK8
hsa05014	Amyotrophic lateral sclerosis (ALS)	**CASP3**, BAD, **BCL2L1**, **CASP9**, MAP3K5, MAP2K6, MAPK14, RAC1, **TP53**
hsa04064	NF-kappa B signaling pathway	**BCL2L1**, TRAF2, NFKB1, CD40, PLCG1, CXCL8, SYK, XIAP
hsa05162	Measles	NFKB1, FYN, **AKT1**, STAT3, **TP53**, **GSK3B**
hsa04012	ErbB signaling pathway	**GSK3B**, BAD, MAPK8, PRKCA, PLCG1, **AKT1**
hsa04912	GnRH signaling pathway	PRKCA, MAP2K6, MAPK14, MAP3K1, MMP2, **PRKACA**, MAPK8
hsa05418	Fluid shear stress and atherosclerosis	MAPK14, **AKT1**, MAPK8, MAP2K6, RAC1, MMP2, **TP53**, MAP3K5, NFKB1
hsa04151	PI3K-Akt signaling pathway	**AKT1**, **GSK3B**, BAD, **BCL2L1**, **TP53**, NFKB1, **FGF2**, **INS**, **CASP9**, PIK3CG, SYK, RAC1, BRCA1, PRKCA
hsa04014	Ras signaling pathway	**AKT1**, **FGF2**, **INS**, RAC1, PRKCA, BAD, **BCL2L1**, NFKB1, MAPK8, ETS1, PLCG1, **PRKACA**
hsa04115	p53 signaling pathway	CHEK1, **CASP9**, **TP53**, **CASP3**, **BCL2L1**
hsa04068	FoxO signaling pathway	MAPK8, **INS**, **AKT1**, STAT3, MAPK14, TGFB1
hsa04621	NOD-like receptor signaling pathway	MAPK8, MAPK14, NFKB1, XIAP, TRAF2, **BCL2L1**, CXCL8
hsa04072	Phospholipase D signaling pathway	**AKT1**, CXCR1, CXCR2, PLCG1, PRKCA, PIK3CG, **INS**, CXCL8, FYN, SYK
hsa05131	Shigellosis	RAC1, NFKB1, MAPK8, MAPK14, CXCL8
hsa04928	Parathyroid hormone synthesis, secretion and action	**PRKACA**, PRKCA, SP1
hsa05016	Huntington’s disease	**TP53**, **CASP3**, **CASP9**, PPARGC1A, SP1
hsa05146	Amoebiasis	NFKB1, CXCL8, **CASP3**, PRKCA, **PRKACA**, TGFB1

hsa05010	Alzheimer's disease	**CASP9**, **CASP3**, BAD, CASP7, **GSK3B**
hsa04140	Autophagy - animal	**INS**, **AKT1**, MAPK8, **PRKACA**, **BCL2L1**, BAD
hsa04150	mTOR signaling pathway	**AKT1**, PRKCA, **INS**, **GSK3B**

Table 6: Huntington Disease KEGG Pathway Analysis of the genes in the miRNA regulatory pathways. Bold gene names indicate the genes included in the common cascades of Huntington and Parkinson Diseases

miRNAs in PD Regulatory Pathways	**KEGG Pathways**
HSA-MIR-16-2 **HSA-MIR-30C-2** **HSA-MIR-34B**	Fatty acid biosynthesis Prion diseases Fatty acid metabolism Glycosaminoglycan degradation Proteoglycans in cancer Central carbon metabolism in cancer
HSA-MIR-328 **HSA-MIR-217** **HSA-MIR-380-5P** **HSA-MIR-491-5P** **HSA-MIR-377** **HSA-MIR-124**	ECM-receptor interaction Adherens junction Fatty acid elongation Transcriptional misregulation in cancer Proteoglycans in cancer Fatty acid degradation Lysine degradation Amoebiasis Long-term depression
HSA-MIR-369-5P **HSA-MIR-106A** **HSA-MIR-17** **HSA-MIR-340** **HSA-MIR-181D**	Prion diseases Proteoglycans in cancer Fatty acid biosynthesis TGF-beta signaling pathway Hippo signaling pathway FoxO signaling pathway Adherens junction
HSA-MIR-221 **HSA-MIR-155** **HSA-MIR-21** **HSA-MIR-20A** **HSA-MIR-34A**	Prion diseases Fatty acid biosynthesis Fatty acid metabolism Cell cycle ECM-receptor interaction Lysine degradation Hepatitis B

HSA-LET-7A	Proteoglycans in cancer
HSA-MIR-106B	Hippo signaling pathway
HSA-MIR-192	Adherens junction
HSA-MIR-23B	Protein processing in endoplasmic reticulum
HSA-MIR-93	Thyroid hormone signaling pathway
	p53 signaling pathway
	Steroid biosynthesis
	FoxO signaling pathway

Table 7: Parkinson Disease KEGG Pathway analysis results of the miRNAs in the significant miRNA regulatory pathways (miRpath v.3 was used). Purple colored miRNA names indicate known PD related miRNAs. Red colored miRNA name indicates DE miRNA with increase expression. Grey colored miRNA names indicate unknown miRNAs

ID	**KEGG Pathway**	**Genes**
hsa04012	ErbB signaling pathway	EGFR, GSK3B, PAK1, MAPK8, PTK2, JUN, AKT1, MYC, MAPK1, MAPK3
hsa04210	Apoptosis	**BCL2L1**, **TP53**, **CASP9**, **CASP3**, RELA, **AKT1**, HTRA2, BAX, MAPK8, JUN, MAPK1, MAPK3
hsa04010	MAPK signaling pathway	RELA, MAPK1, MAPK3, **PRKACA**, EGFR, MET, **TP53**, MAPK14, PPM1A, **AKT1**, **CASP3**, PAK1, MAPK8, HSPA6, JUN, MYC, **FGF2**, **INS**
hsa05205	Proteoglycans in cancer	ROCK1, **AKT1**, PAK1, MAPK1, ESR1, MAPK3, **TP53**, PTK2, MYC, MET, **FGF2**, **PRKACA**, MAPK14, EGFR, **CASP3**
hsa04014	Ras signaling pathway	MAPK1, MAPK3, **AKT1**, **FGF2**, **INS**, EGFR, MET, **BCL2L1**, RELA, MAPK8, PAK1, **PRKACA**
hsa04926	Relaxin signaling pathway	**AKT1**, RELA, MAPK1, MAPK3, **PRKACA**, JUN, MAPK14, MAPK8, EGFR
hsa04115	p53 signaling pathway	ATR, PPM1D, PTEN, CDKN2A, BAX, **CASP9**, **TP53**, **CASP3**, **BCL2L1**
hsa04722	Neurotrophin signaling pathway	**AKT1**, MAPK1, MAPK3, **GSK3B**, RELA, MAPK8, **TP53**, JUN, MAPK14, BAX
hsa04510	Focal adhesion	ROCK1, **AKT1**, PTEN, PTK2, MAPK1, MAPK3, EGFR, MET, MAPK8, JUN, PAK1, **GSK3B**

hsa04071	Sphingolipid signaling pathway	MAPK1, MAPK3, **AKT1**, ROCK1, PTEN, MAPK8, MAPK14, BAX, **TP53**, RELA
hsa05014	Amyotrophic lateral sclerosis (ALS)	**CASP3**, BAX, **BCL2L1**, **CASP9**, MAPK14, **TP53**
hsa04140	Autophagy - animal	**INS**, PTEN, **AKT1**, MAPK1, MAPK3, DDIT4, MAPK8, **PRKACA**, **BCL2L1**
hsa04151	PI3K-Akt signaling pathway	**AKT1**, PTEN, EGFR, MET, **GSK3B**, MYC, **BCL2L1**, **TP53**, RELA, **FGF2**, **INS**, DDIT4, **CASP9**, MAPK1, MAPK3, PTK2
hsa04728	Dopaminergic synapse	**PRKACA**, **AKT1**, GSK3A, **GSK3B**, MAPK14, MAPK8
hsa04068	FoxO signaling pathway	MAPK8, MAPK1, MAPK3, **INS**, **AKT1**, EGFR, PTEN, SIRT1, MAPK14
hsa04657	IL-17 signaling pathway	RELA, JUN, MAPK14, MAPK1, MAPK3, MAPK8, **CASP3**, **GSK3B**
hsa04933	AGE-RAGE signaling pathway in diabetic complications	RELA, MAPK8, EGR1, MAPK14, MAPK1, MAPK3, JUN, BAX, **CASP3**, **AKT1**
hsa04664	Fc epsilon RI signaling pathway	**AKT1**, MAPK14, MAPK1, MAPK3, MAPK8
hsa04620	Toll-like receptor signaling pathway	MAPK1, MAPK3, MAPK14, MAPK8, **AKT1**, RELA, JUN
hsa04024	cAMP signaling pathway	**PRKACA**, **AKT1**, MAPK1, MAPK3, MAPK8, ROCK1, RELA, JUN, PAK1
hsa04662	B cell receptor signaling pathway	RELA, **GSK3B**, **AKT1**, JUN, MAPK1, MAPK3
hsa05131	Shigellosis	RELA, MAPK8, MAPK1, MAPK3, MAPK14, ROCK1
hsa04932	Non-alcoholic fatty liver disease (NAFLD)	**INS**, **AKT1**, RELA, GSK3A, **GSK3B**, **CASP3**, BAX, MAPK8, JUN
hsa04550	Signaling pathways regulating pluripotency of stem cells	**GSK3B**, **AKT1**, MAPK1, MAPK3, **FGF2**, MAPK14, MYC
hsa04072	Phospholipase D signaling pathway	MAPK1, MAPK3, **AKT1**, EGFR, **INS**
hsa04630	Jak-STAT signaling pathway	MYC, **AKT1**, EGFR, **BCL2L1**
hsa05031	Amphetamine addiction	**PRKACA**, SIRT1, JUN
hsa05010	Alzheimer's disease	BACE1, APP, **CASP9**, **CASP3**, MAPK1, MAPK3, **GSK3B**
hsa05162	Measles	RELA, **AKT1**, HSPA6, **TP53**, **GSK3B**
hsa04723	Retrograde endocannabinoid signaling	**PRKACA**, MAPK14, MAPK1, MAPK3, MAPK8
hsa05016	Huntington's disease	**TP53**, **CASP3**, **CASP9**, PPARGC1A, BAX

hsa05134	Legionellosis	RELA, **CASP9**, **CASP3**, HSF1, HSPA6
hsa04621	NOD-like receptor signaling pathway	MAPK8, MAPK1, MAPK3, MAPK14, RELA, JUN, **BCL2L1**
hsa05012	Parkinson's disease	HTRA2, **CASP9**, **CASP3**, **PRKACA**

Table 8: Parkinson Disease KEGG Pathway analysis results of the genes in the significant miRNA regulatory pathways. Bold gene names indicate the genes included in the common cascades of Huntington and Parkinson Diseases.

4.5 Comparison of Cascades in miRNA Regulatory Pathways in HD and PD

Significant disease specific regulatory pathways of HD and PD were compared (Table 9).

Common Cascades in HD and PD	**45 HD significant pathways**	**61 PD significant pathways**
BCL2L1 → CASP9	13	6
CASP9 → CASP3	13	6
CASP3 → AKT1	12	3
AKT1 → GSK3B	1	4
TP53 → FGF2	1	4
AKT1 → PRKACA	6	4
CASP3 → AKT	1	2
CDX2 → INS	2	2

Table 9: Common cascades in HD and PD. There are 45 HD related significant pathways and 61 PD related significant pathways. Table shows how many times each common relation is included among significant pathways.

Common interactions between HD and PD is shown in Table 9. There were 45 HD specific and 61 PD specific significant pathways observed. Table 9 shows the amount of occurrences of each cascade in these significant pathways. In Table 10, the function of each gene is shown individually. The information is detected from GeneCards database [91].

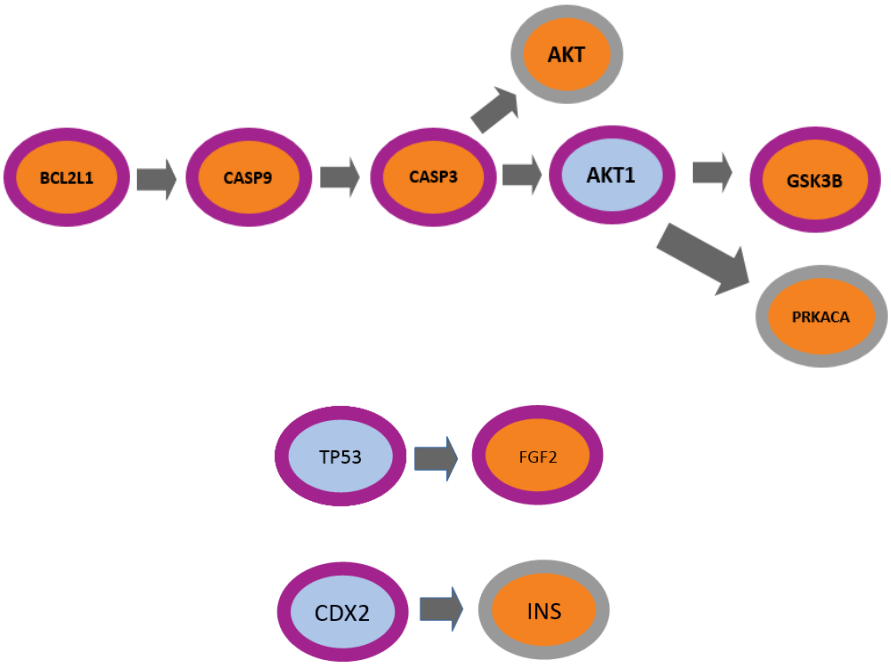


Figure 15: Huntington Disease (Common Cascades between HD and PD)

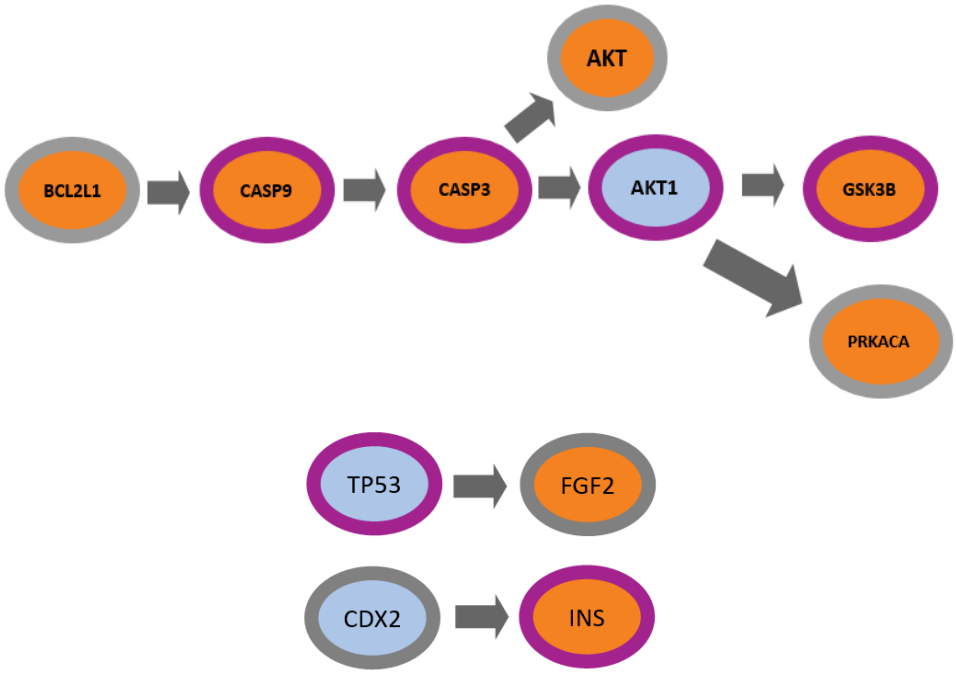


Figure 16: Parkinson Disease (Common Cascades between HD and PD)

When common interactions were gathered, the differences between the common pathways were observed (Figure 15-16).

For each participant of the cascade groups in HD and PD, functional information were presented in Table 10.

Gene Name	Gene Summary from GeneCards
BCL2L1	The protein encoded by this gene belongs to the BCL-2 protein family. BCL-2 family members form hetero- or homodimers and act as **anti- or pro-apoptotic regulators**
CASP9	Caspase 9, Apoptosis-Related Cysteine Peptidase. Sequential activation of caspases plays a central role in the execution-phase of cell apoptosis
CASP3	Caspase 3, Apoptosis-Related Cysteine Peptidase. The protein encoded by this gene is a cysteine-aspartic acid protease that plays a central role in the execution-phase of cell apoptosis.
AKT1	In the developing nervous system AKT is a critical mediator of growth factor-induced neuronal survival. Survival factors can suppress apoptosis in a transcription-independent manner by activating the serine/threonine kinase AKT1, which then phosphorylates and inactivates components of the apoptotic machinery.
GSK3B	The protein encoded by this gene is a serine-threonine kinase belonging to the glycogen synthase kinase subfamily. It is a negative regulator of glucose homeostasis and is involved in energy metabolism, inflammation, ER-stress, mitochondrial dysfunction, and apoptotic pathways
PRKACA	Protein Kinase CAMP-Activated Catalytic Subunit Alpha

TP53	This gene encodes a tumor suppressor protein containing transcriptional activation, DNA binding, and oligomerization domains. The encoded protein responds to diverse cellular stresses to regulate expression of target genes, thereby inducing cell cycle arrest, apoptosis, senescence, DNA repair, or changes in metabolism.
FGF2	FGF family members bind heparin and possess broad mitogenic and angiogenic activities. This protein has been implicated in diverse biological processes, such as limb and nervous system development, wound healing, and tumor growth.
CDX2	The encoded protein is a major regulator of intestine-specific genes involved in cell growth and differentiation. This protein also plays a role in early embryonic development of the intestinal tract.
INS	Binding of insulin to the insulin receptor (INSR) stimulates glucose uptake.

Table 10: Summary of Genes included in the common cascades from GeneCards database[91].

DISCUSSION

In this dissertation our aim was to understand the TF-miRNA-gene regulatory mechanisms of Huntington and Parkinson Diseases, and reveal the significant regulatory signaling pathways by using the differential expression analysis results of miRNAs and genes, known disease related miRNAs/genes by mapping the information on directed PPI network. Thus, it would be possible to uncover the unknown but possibly important cascades.

5.1 Disease Related Regulatory Network

We used Signor Database to include the type of the regulations between two entities. We extended the directed PPI network by adding miRNA-gene and TF-miRNA regulatory information. The resulting TF-miRNA-gene directed regulator network had 5241 number of entities and 15276 number of unique relations. This network had 468 number of TFs, and 392 number of miRNAs.

miRNA regulatory network had 5241 number of nodes with 15276 number of unique relations. There were 468 number of TFs, 4231 number of genes and 392 number of miRNAs in the extended regulatory network. Thus, we integrated the transcriptional and post-transcriptional regulation information to molecular interaction network. All the relations between the entities were selected and mapped on the network if they were experimentally validated, to provide and increase the reliability of the results.

Disease related known miRNA/genes were selected from the databases. All the disease related known miRNA/genes were again experimentally validated. They were mapped on the regulatory network by changing the border colors of the nodes to purple.

To include new informative differential expression analysis, we also integrated the miRNAs/genes that were detected to have significant expression changes. They were mapped to the network with red/blue border colors. If one miRNA/gene existed in both lists, then the border color was changed to yellow/green (explained in detail in section Materials &Methods). By this way, we got the disease related directed regulatory network both for Huntington's and Parkinson's Diseases.

5.2 Disease Related Regulatory Subnetwork

Directed regulatory network was too complex to analyze and time consuming to get the information we need that is why, interaction networks are useful models to understand the functional interpretations of molecules.

To analyze how connected parts compose the whole network system, to better understand the relative importance of system components and make quantitative predictions for understanding of Huntington's and Parkinson's diseases systematically, we detected the DE miRNAs/genes on the network and for each DE miRNA/gene we determined the 3rd degree neighbors [92]. The disease related genes do not show difference in their expressions significantly in some situations[93,94]. That is why, there may be some disease-related genes in the disease related subnetwork among non-DE genes. The tool we developed have the options to select among 1st, 2nd and 3rd neighbors as well. We connected those DE miRNAs/genes with their 3rd degree neighbors. Thus, potential active disease related regulatory subnetwork was produced. Huntington's subnetwork contained 94% and Parkinson's subnetwork contained 90.4% of known-disease related miRNA's/genes. For Huntington's Disease the subnetwork included 4724 number of nodes with 14922 relations and Parkinson's Disease subnetwork included 4474 number of nodes with 14605 relations. As we expected, subnetworks of diseases were different from each other.

To reveal the regulatory pathways that changed the expression values of DE miRNAs/genes, 0-indegree and 0-outdegree of DE miRNAs/genes were selected. We had a directed regulatory subnetwork, so BFS algorithm was a good option to traverse all the network to get the pathways between 0-indegree and 0-outdegree DE miRNAs/genes. For Huntington's Disease we got 9167, for Parkinson's we got 614 number of directed acyclic paths. Finally, based on the known disease-related genes and miRNAs, hypergeometric test was used to evaluate the significance of the association of the pathways and the disease of interest. Thus, we revealed 42 TF-miRNA-protein regulatory pathways significantly related to HD and 27 TF-miRNA-protein regulatory pathways significantly related to PD.

When significant regulatory pathways of HD and PD were compared, common cascades were found. When those cascades were integrated to each other, a unique pathway was formed (Figure 15-16). The interesting thing was, although HD and PD shared the relations, not every node was identical to its form in other disease. BCL2L1, FGF2, CDX2 genes were known to be related to HD, but they were not known to be related to PD. Furthermore, INS gene was not known to be related to HD, but it was known to be related to PD.

5.3 KEGG Pathway Analysis of miRNAs/genes in miRNA Regulatory Pathways

From disease related significant regulatory pathways, miRNAs and genes were identified. Separately the KEGG analysis of miRNAs and genes were done.

In our analysis, *hsa-mir-146a* and *hsa-mir-9* were known to be related to HD. Their KEGG analysis enriched *hippo signaling*. A recent study investigated the possible role of this pathway and observed alterations in human Huntington's Disease brain [95].

Glycosaminoglycan biosynthesis was the other pathway that was detected. Quantitative and qualitative fractioning of GAGs have been widely used in screening for MPS disorders for many years. The accumulation of undegraded GAGs in lysosomes affects functions of most cell types, tissues and organs, including viscera, connective tissue and central nervous system (CNS)[96].The other informative and interesting pathway which was enriched in all groups of miRNAs was *proteoglycans in cancer* pathway. Proteoglycans consist of a core protein and one or more covalently attached glycosaminoglycan chains [97]. Proteoglycans and glycosaminoglycans play important roles in neurological disorder development. Thus, we can conclude that *hsa-mir-15a* and *hsa-mir-17* are not known to be related to HD, and not DE but may be important for HD development. Similarly, for Parkinson Disease which is again a neurological disorder, it was revealed that known disease related miRNAs enriched the pathway *proteoglycans in cancer* and *glycosaminoglycan degradation* pathways. Interestingly, miRNAs with increased and decreased differential expressions and miRNAs with unknown information enriched the *proteoglycan in cancer* pathway.

Glycosphingolipid biosynthesis pathway was enriched by the known disease related miRNAs *hsa-mir-146a* and *hsa-mir-9*. A study done in 2007 on HD mice/humans brains showed disrupted patterns of glycolipids and ganglioside levels and hypothesized that glycolipid metabolism changes may be used as an alternative therapeutic target for HD [98]. Sphingolipids are responsible from organizing the functions of many neuronal ion channels and receptors. Role and mechanism of membrane microdomains in development of neurological diseases is not certain and known yet[99] but alterations in sphingolipid metabolism found to be related to many neurological disorders not only Huntington's Disease but also Parkinson's as well[98]. In addition to this, our PD results showed that, all groups of miRNAs enriched fatty acid related KEGG pathways such as *fatty acid biosynthesis*, *fatty acid metabolism*, *fatty acid elongation* and *fatty acid degradation.* Fatty acids are components of most cellular lipids, such as sphingolipids and cholesterol esters and major fatty acid species include long-chain fatty acids (LCFAs). Most of the LCFAs in mammals are components of sphingolipids and when the polar head group of sphingolipids are sugar, it is called glycosphingolipids [100]. It seems that, our results promote the significance of fatty acid mechanism in the development of HD and PD and possible regulatory roles of miRNAs in this system.

Furthermore, we discovered that *has-mir-486-5p* was one miRNA which was included in the significant pathways and it was differentially expressed. It does not take part in HMDD or miR2disease databases when searched for HD related miRNAs. A previous study done in 2015 about plasma-derived miRNAs mimicking behavior in Huntington's Disease brain tissue showed that *hsa-mir-486-5p* had increased levels of expression in HD patients' plasma, compared to control subjects. In addition to this, results were consistent with the expression changes in the brain [101].

Apart from this, *hsa-mir-15a* and *hsa-mir-17* did not belong neither to known-disease related miRNA group, nor to DE miRNA group. They were found to be on the path of significant regulatory cascades. When a literature search was done about their relation to HD, we found the study of Marti et.al. showing the miRNA variability in human brain and both *has-miR-15a* and *hsa-mir-17* were found to be significantly upregulated [102]. Their KEGG analysis enriched fatty acid biosynthesis and fatty acid metabolism pathways. In previous studies, it was shown

that the lipid dysregulation had effects on sterol regulatory element binding proteins in Huntington's Disease[103]. Prion diseases pathway was enriched and it was already known that prion diseases are included in a group of neurodegenerative disorders and both Huntington and Parkinson Diseases belong to this group[104].

Lysosomal storage disorders and neurodegenerative diseases (Alzheimer's, Parkinson's and Huntington diseases) share many features related to the mechanisms that characterize the disease pathogenesis.

5.4 Analysis of Disease Related Directional Pathway Subgroups in HD

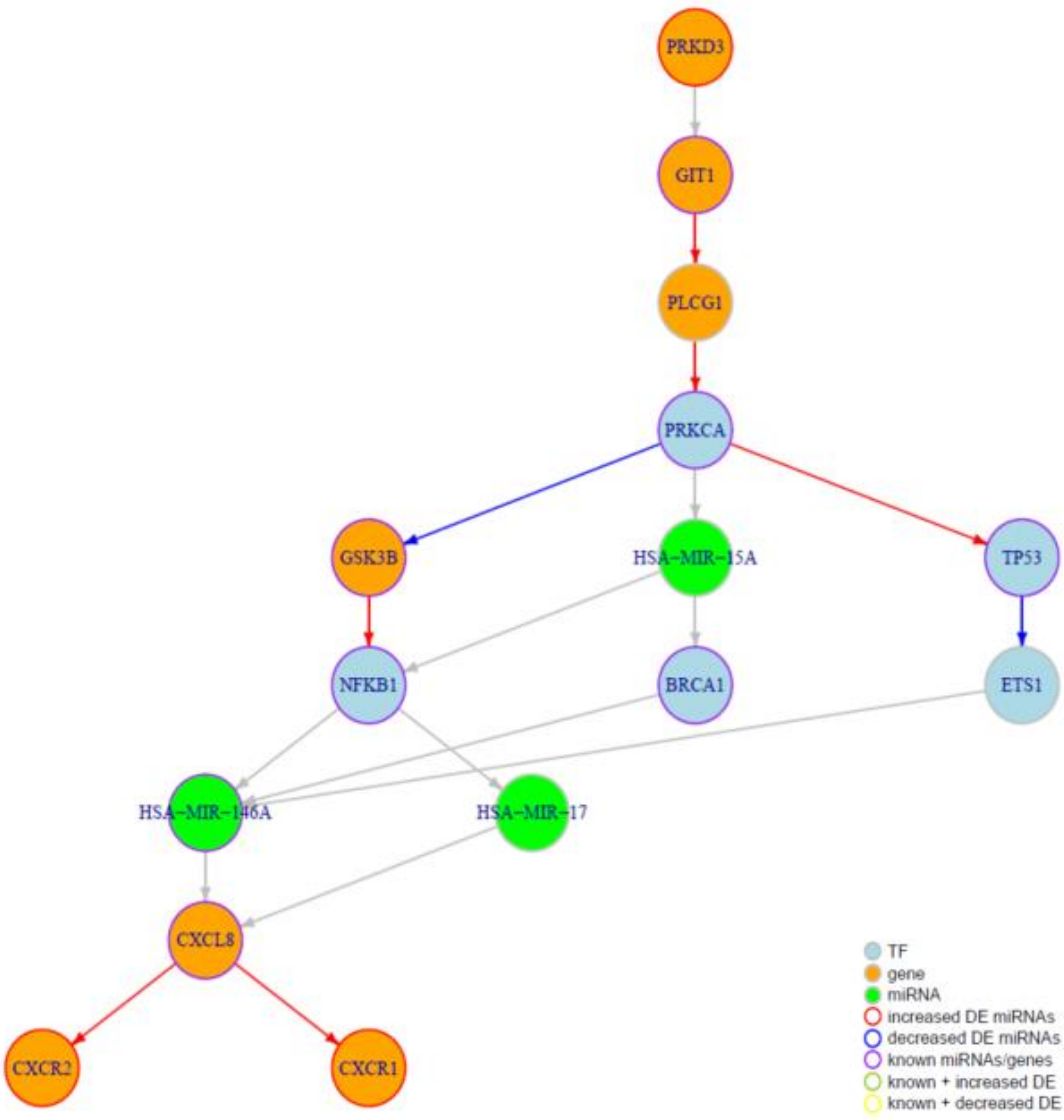


Figure 17: Huntington Disease related active pathways were grouped according to their ratio of similar relations. 2nd pathway subgroup of HD. Color of arrows indicates the relationship type: 1) →: activation, →: repression, →: not-known.

In our study to understand the relationship of miRNAs with their targets and TFs with their target miRNAs, the significant regulatory pathways were grouped according to their sequence of cascades. If they had same length with each other and if they had the same beginning node,

they were put in to the same group. Tree view of the cascades provides user to analyze and follow the regulatory sequences apparently.

To understand how to analyze the tree view of the cascade groups, Figure 17 which is the 2nd cascade group of HD will be used.

The tree starts with the root node PRKD3 gene and ends with leaf nodes CXCR2 and CXCR1 genes which have red border colour, representing the DE genes in HD. The pathways in this subgroup is listed in Table 9. Each subgroup is shown in detail in Appendix A.

According to HD subgroupss, we could observe that, hsa-mir-146a →CXCL8 → CXCR1/2 were included in most of the cascades. CXCR1/2 were activated by CXCL8 and it was shown by the directed red arrow in Figure 17. CXCR1 and CXCR2, known as Interleukin-8 receptor A (IL-8RA) and Interleukin-8 receptor B (IL-8RB) form a gene cluster. C-X-C motif ligand 8 (CXCL8) is a chemokine that acts as an important multifunctional cytokine and CXCR1/2 are the receptors of CXCL8. CXCL8-CXCR1/2 pathway was found to be related to various cancer types such as, breast cancer, prostate cancer, lung cancer etc. [105]. The chemokine receptors CXCR1/2 and their ligand CXCL8 are essential for the activation and trafficking of inflammatory mediators as well [106]. Reported sources of CXCL1, CXCL2, and CXCL8 include activated microglia [107]. Microglia when stimulated seem to express CXCR2 receptor after demyelination [108]. Demyelination in HD compared to other neurodegenerative diseases remain relatively unexplored and a recent study investigated the myelin breakdown in HD suggested that myelin breakdown contributed to white-matter impairment in human HD [109]. Thus, our pathway results supported the hypothesis about the possible relation HD to demyelination process. In addition to this, a study and data showed the function of CXCL8-CXCR2 network involvement in neuronal electrical activity, neurotransmitter release and synaptic plasticity in the central nervous system (CNS) [110].

Apart from HD and PD, multiple sclerosis (MS) also is known to have demyelination in CNS [111]. Studies emphasized the crucial role of Zn in the pathogenesis of MS [112]and it was also shown that differences in Zn levels stimulated different amount of cytokine release according

to different type of cells [113]. Thus, like in the case of MS, alterations in Zn levels may be important to control the demyelination process of neurological diseases like HD and PD.

electrical activity was lost in HD brain [114]. Another study showed that, CXCR2 expression increase on Purkinje neurons enhancing glutaminergic activity. It is already known that, HD is caused by a mutation in HD gene and huntingtin protein get built incorrectly by getting extra glutamine proteins. So, CXCL8-CXCR2 pathway mechanism may be informative for understanding the development process of HD.

CXCL8-CXCR1/2 cascades were controlled by hsa-mir-146a which was known to be related to HD. Transcription factor NFkB is known to regulate the expression of hsa-mir-146a by binding to the upstream sequences[115]. A study investigating the NFkB, TP53 and miRNA involvement in the regulation of cell death mechanism in HD cells, showed that, there were expression changes in NFkB and TP53 and decrease in hsa-mir-146a. The study also showed that the NFkB activation was decreased while TP53 was increased [116]. Our regulatory figure explains the decrease in the expression of hsa-mir-146a in pursuit of the decrease in NFkB. [116] confirmed earlier studies done, showing increase in the level of TP53 in other cell and animal models of HD[117,118,119]. However, none of the studies could explain the mechanism of increased level of TP53. Our results showed that, TP53 did not have direct regulatory effect on hsa-mir-146a. There exists a negative relationship between TP53 and ETS1 and this regulation was shown by blue coloured arrow in Figure 17. It was found that ETS1 physically had association with TP53 and its C-terminal part was found to be essential for inhibition of ETS1 transcriptional activity [120]. That is why according to our results, ETS1 was involved in the regulation of hsa-mir-146a, and as it has grey coloured border meaning neither it is known disease related nor differentially expressed, both its function in regulating the hsa-mir-146a, hence CXCL8-CXCR1/2 cascades with TP53 and concordantly its importance for HD mechanism may be clarified.

PRKD3 →GIT1 →PLCG1 →PRKCA→GSK3B→NFKB1→HSA-MIR-146A→CXCL8→CXCR2
PRKD3→GIT1→PLCG1→PRKCA→GSK3B→NFKB1→HSA-MIR-146A→CXCL8→CXCR1
PRKD3→GIT1→PLCG1→PRKCA→GSK3B→NFKB1→HSA-MIR-17→CXCL8→CXCR2
PRKD3→GIT1→PLCG1→PRKCA→HSA-MIR-15A→BRCA1→HSA-MIR-146A→CXCL8→CXCR2
PRKD3→GIT1→PLCG1→PRKCA→HSA-MIR-15A→NFKB1→HSA-MIR-146A→CXCL8→CXCR2
PRKD3→GIT1→PLCG1→PRKCA→TP53→ETS1→HSA-MIR-146A→CXCL8→CXCR2
PRKD3→GIT1→PLCG1→PRKCA→GSK3B→NFKB1→HSA-MIR-17→CXCL8→CXCR1
PRKD3→GIT1→PLCG1→PRKCA→HSA-MIR-15A→BRCA1→HSA-MIR-146A→CXCL8→CXCR1
PRKD3→GIT1→PLCG1→PRKCA→HSA-MIR-15A→NFKB1→HSA-MIR-146A→CXCL8→CXCR1
PRKD3→GIT1→PLCG1→PRKCA→TP53→ETS1→HSA-MIR-146A→CXCL8→CXCR1

Table 9: Pathway list of directed regulatory network in Fig. 14.

GSK3B is a known HD related gene and [121] demonstrated levels of total GSK3 were decreased in the HD-affected frontal cortex. Our HD related important pathways that we found with their relation information added to the network explains the flow of the cascade. If GSK3B expression decreases in HD, then NFkB1 activity would decrease as expected because of the direct proportion between them. A study on Rett Syndrome (RTT) which is the second leading cause of mental impairment in girls, also showed that inhibition of GSK3B reduces NFkB1 signaling [122].

PLCG1 and PRKCA gene which has a regulatory effect on both NFkB and TP53, have a regulatory interaction between each other. PLCG1 has grey border color in Figure 17, meaning its relation to Huntington Disease is unknown and it was not detected to have significant expression change according to the gene expression analysis results we used but it appeared in the path of significant HD related directed pathways and may be an important factor effecting the HD mechanism. PLCG1 is an important signaling regulator involved in cellular processes effecting the development of brain and synaptic transmission and its abnormal expression and activation was observed in various brain disorders like Alzheimer's and Huntington's disease. PLC enzyme (protein of PLCG1 gene) by triggering calcium mobilization activates the protein kinase C (PKC) (protein of PRKCA gene). PLC-alpha enzyme activity was observed to be decreased in HD models [123]. PLCG1 is between two entities GIT1 and PRKCA transcription factor. They are both known to be related to Huntington Disease and there may exist a new regulatory mechanism that have not been discovered yet between GIT1 and PRKCA. As PLCG1 was found to be at the beginning of the signaling cascade, it may be important for the progression of the HD mechanism. It should be investigated.

The root of the tree, the PRKD3 gene as we mentioned before, was differentially expressed and showed an increase in expression. It is known that PRKD3 is responsible from different kind of processes like cell proliferation, cell survival, immune cell signaling and neuronal development selectively phosphorylates GIT1 to enhance cell spreading and motility [124,125] . Thus, there may be an important regulatory cascade mechanism between PRKD3 and PRKCA and may reveal a novel target for therapeutic applications for HD.

Another miRNA hsa-mir-15a, which exists in the HD related pathway [PRKCA → hsa-mir-15a → BRCA1 → has-mir-146a] has an unknown relation to HD. It has a role in the regulation of NFkB and BRCA1 genes. BRCA1 gene actually is a major tumor suppressor gene and it is targeted by hsa-mir-15 post-transcriptionally [126]. The study was done in pancreatic epithelium, colorectal and pancreatic adenocarcinoma cell lines and protein abundances were different in each cell line, meaning that the regulation mechanism of hsa-mir-15a and BRCA1 is complex and cell-specific. Thus, both previous studies and our results propose that hsa-mir-15a, BRCA1 interaction may be an important regulatory mechanism and hsa-mir-15a may be newly discovered miRNA that plays an important role in HD development.

Hsa-mir-17 again may be a newly discovered miRNA, as it has a regulatory role on the expression of CXCL8. With the decrease of hsa-mir-146a in HD, CXCL8 expression was expected to increase [127]. According to this disease related subnetwork mechanism, hsa-mir-17 function may be an important regulatory factor with hsa-mir-146a in triggering the inflammation mediators IL-8. Activation of inflammation related cascades in HD also opens a new point of view to the relatively unexplored area of HD.

5.5 Analysis of Disease Related Directional Pathway Subgroups in HD

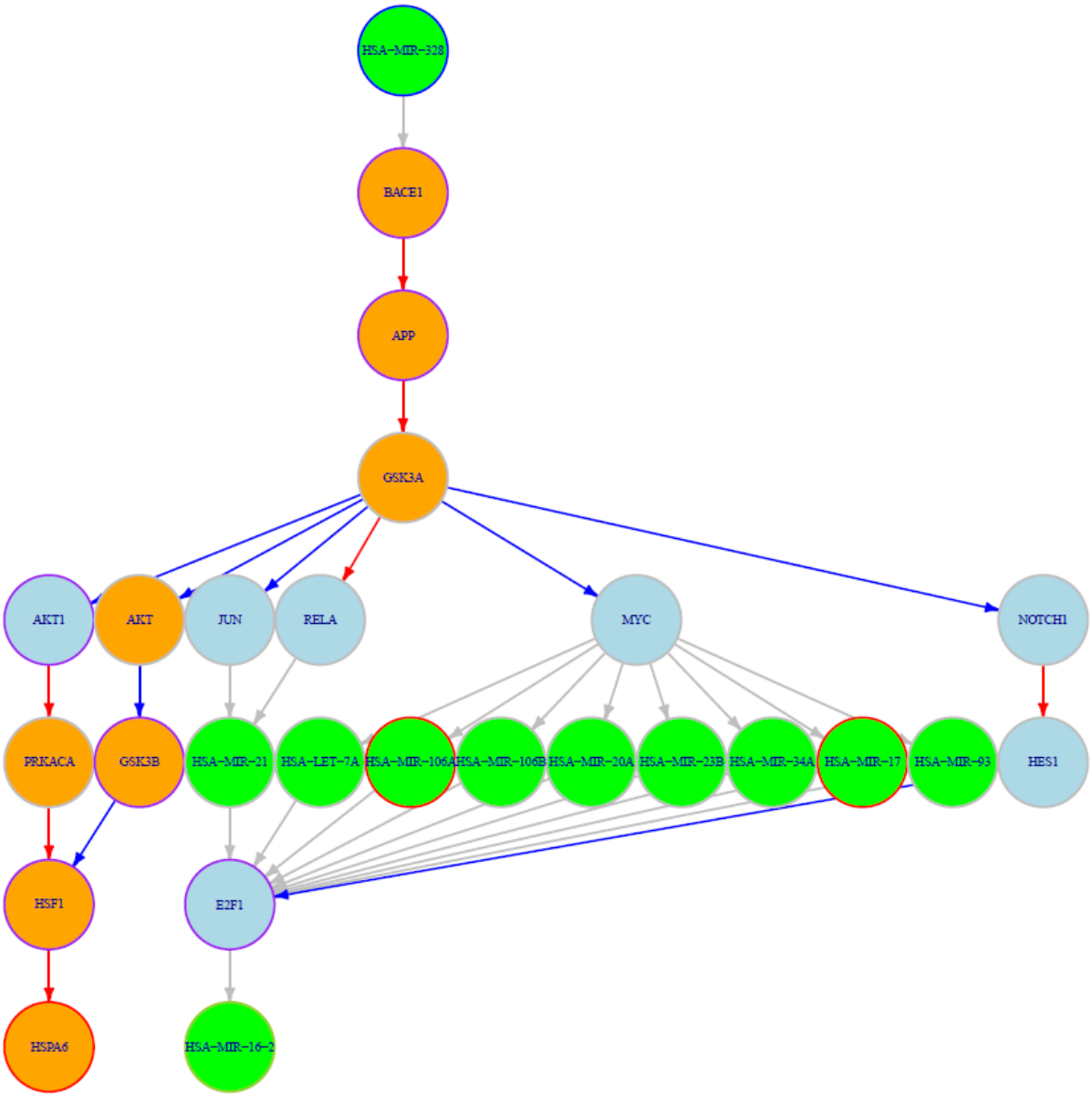


Figure 18: Parkinson Disease related active pathways were grouped according to their ratio of similar relations. 5th pathway subgroup of PD. Colour of arrows indicates the relationship type: 1) →: activation, 2)→: repression, 3) →: not-known.

HSA-MIR-328 → BACE1 → APP→GSK3A→AKT1→PRKACA→HSF1→HSPA6
HSA-MIR-328→BACE1→APP→GSK3A→AKT→GSK3B→HSF1→HSPA6
HSA-MIR-328→BACE1→APP→GSK3A→JUN→HSA-MIR-21→E2F1→HSA-MIR-16-2
HSA-MIR-328→BACE1→APP→GSK3A→RELA→HSA-MIR-21→E2F1→HSA-MIR-16-2
HSA-MIR-328→BACE1→APP→GSK3A→MYC→HSA-LET-7A→E2F1→HSA-MIR-16-2
HSA-MIR-328→BACE1→APP→GSK3A→MYC→HSA-MIR-106A→E2F1→HSA-MIR-16-2
HSA-MIR-328→BACE1→APP→GSK3A→MYC→HSA-MIR-106B→E2F1→HSA-MIR-16-2
HSA-MIR-328→BACE1→APP→GSK3A,→MYC→HSA-MIR-20A→E2F1→HSA-MIR-16-2
HSA-MIR-328→BACE1→APP→GSK3A→MYC→HSA-MIR-23B→E2F1→HSA-MIR-16-2
HSA-MIR-328→BACE1→APP→GSK3A→MYC→HSA-MIR-34A→E2F1→HSA-MIR-16-2
HSA-MIR-328→BACE1→APP→GSK3A→MYC→HSA-MIR-17→E2F1→HSA-MIR-16-2
HSA-MIR-328→BACE1→APP→GSK3A→MYC→HSA-MIR-93→E2F1→HSA-MIR-16-2
HSA-MIR-328→BACE1→APP→GSK3A→NOTCH1→HES1→E2F1→HSA-MIR-16-2

Table 10: Pathway list of directed regulatory network in Fig. 15.

When we analyze the regulatory cascades in Figure 18 for PD, cascades start with hsa-mir-328. It has a blue border color which indicates, its expression was decreased in PD patients. Like in the case of HD, GSK3B, PRKACA, AKT interactions seemed to be important. At the same level of regulatory pathways between hsa-mir-328 and both known PD related and also upregulated hsa-mir-16-2 there exists different type of miRNAs. It seemed like E2F1 gene regulation is highly balanced by miRNAs. According to this analysis, hsa-mir-106a and hsa-mir-17 were upregulated and may be active role in triggering the disease related regulatory mechanisms. All the miRNAs regulatory transcription factor is MYC but its function in PD is unknown and again it may be an important revealing for understanding the disease mechanism.

Hsa-mir-17 had a role in HD as we mentioned in section 4.5. It was in the regulatory pathways of HD, but its function was unknown and it was not DE. But in PD, its expression was upregulated and again it exists in the regulatory pathway of PD. Hsa-mir-17 may be an important miRNA in neurological disorders. One recent study done about differentially expressed miRNAs in PD revealed hsa-mir-17 upregulation [128]. The results of [128] showed SH-SY5Y cells in PD stress conditions affected the expression levels of miRNAs. They found that, hsa-mir-155, hsa-mir-16-5p, hsa-mir-17, hsa-mir-181d were upregulated and hsa-mir-23b, hsa-mir-7a were down regulated. All these miRNAs were also detected by our pathway results. According to our study hsa-mir-155, hsa-mir-23b and hsa-mir-7a were not known to be related to PD but our results showed that they were included and took part in the significant regulatory pathways. Their role in the regulatory pathways can be observed in depth from supplementary figures (Appendix A).

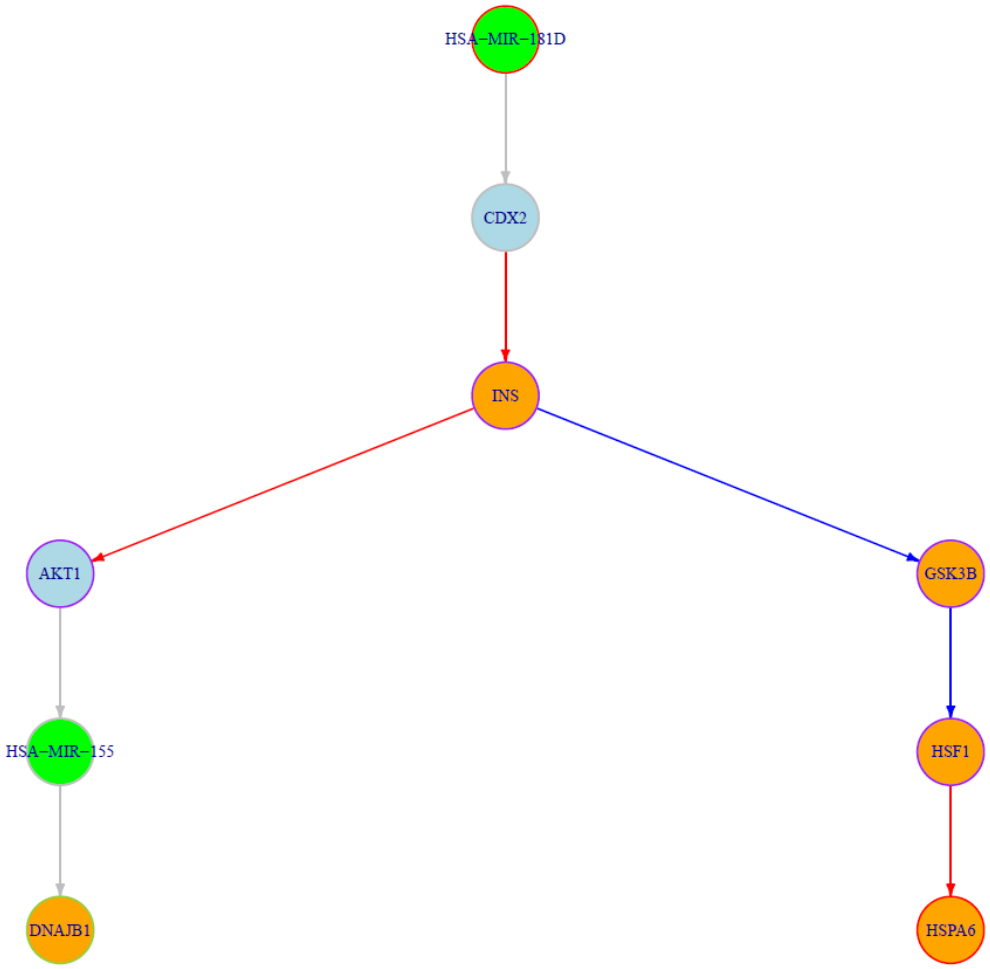


Figure 19: Parkinson Disease related active pathways were grouped according to their ratio of similar relations. 13th pathway subgroup of PD. Colour of arrows indicates the relationship type: 1) →: activation, 2)→: repression, 3)→: not-known.

In Figure 19, it is represented that, has-mir-155 and CDX2 may be PD related miRNA and gene. According to the graphical representation, they have grey border colour which indicates their relation to PD is unknown. A recently published paper in 2018 suggested that, anti-Tumor Necrosis Factor (anti-TNF) which have been used in inflammatory bowel disease (IBD) treatment may reduce the risk of Parkinson's disease. The results indicated that, the IBD patients taking anti-TNF therapy compared to patients receiving no anti-TNF were 78% less likely to have PD[129]. In addition to this, when CDX2 function was analysed in GeneCards, this gene was found to be the important regulator of intestine-specific genes which are involved in cell

growth and differentiation. Its role in early embryonic development of the intestinal tract was also interesting. These findings indicated that there may be a possible relation between the condition of intestinal tract and PD development.

According to the results, we can say that, our method is successful in revealing the significant regulatory pathways of HD and PD. It gives reliable results, by combining the transcriptional and post-transcriptional regulatory factors in one disease related network. Each disease related network will be unique and will have different patterns.

CONCLUSION AND FUTURE WORK

In this study, the core hypothesis was developing a network modeling method to reveal the active TFs, miRNAs, genes and their regulatory mechanism by using directed interaction and signaling information and understand the systemic active transcriptional and post-transcriptional molecular mechanisms of the complex diseases.

For model complex diseases, Huntington's and Parkinson's diseases were used. miRNA functions and their interactions were important to understand the complex disease mechanism.

Hippo signaling, *glycosaminoglycan* and *glycosphingolipid biosynthesis*, *fatty acid elongation* and *biosynthesis* KEGG pathways were enriched by the miRNAs and genes included in significant pathways. These pathways were interesting to focus on as they were revealed to be related to neurological disorders and their possible relation to each other were discussed in terms of HD and PD. HD and PD although only shared just one miRNA which was hsa-mir-17, the enriched pathways were connected to each other and indicated the function of the method we developed.

When the regulatory cascades of both diseases were compared, shared cascades were found and with further study on these cascades, different but informative regulations can be revealed. BCL2L1 gene in shared cascades was interesting. It was known to be HD related, but its relation to PD was not known. The protein encoded by this gene belongs to the BCL-2 protein family, which acts as anti- or pro- apoptotic regulators.

Apoptotic regulations of HD and PD can be analyzed in deep for further studies as apoptosis KEGG pathway was enriched in both diseases as well.

This tool also can be used to discover new possible disease related miRNAs. For HD, *hsa-mir-15a* and *hsa-mir-17*; for PD, *hsa-mir-155*, *hsa-mir-23b* and *hsa-mir-7a* may be possible disease related miRNAs. Further work targeting these miRNAs may be productive in explaining the regulatory mechanism of disease development.

In conclusion, out method is successful in discovering new potential miRNAs related to disease of interest. Significant disease specific pathways can be revealed for further examination. Our

approach is flexible to analyze the integrated networks and it is the first time a directed signaling network is used to explain the regulatory cascades. Integrating miRNA-mediated TF and mRNA-protein interactions is helpful for constructing more comprehensive regulatory networks.

With the increase in available miRNA expression profiles for complex diseases and the accumulation of validated miRNA regulations, the results of the tool will become more precise.

This approach is general and could be applied to other complex diseases and for different model organisms.

BIBLIOGRAPHY

[1] D.-Y. Cho, Y.-A. Kim, and T. M. Przytycka, "Chapter 5: Network Biology Approach to Complex Diseases," *PLoS Comput. Biol.*, vol. 8, no. 12, p. e1002820, 2012.

[2] M. Baglioni, F. Russo, F. Geraci, M. Rizzo, G. Rainaldi, and M. Pellegrini, "A new method for discovering disease-specific MiRNA-target regulatory networks," *PLoS One*, vol. 10, no. 4, pp. 1–15, 2015.

[3] M. Ha and V. N. Kim, "Regulation of microRNA biogenesis," *Nat. Rev. Mol. Cell Biol.*, vol. 15, no. 8, pp. 509–524, 2014.

[4] C. Zhang, "Novel functions for small RNA molecules.," *Curr. Opin. Mol. Ther.*, vol. 11, no. 6, pp. 641–651, 2009.

[5] E. Huntzinger and E. Izaurralde, "Gene silencing by microRNAs: Contributions of translational repression and mRNA decay," *Nat. Rev. Genet.*, vol. 12, no. 2, pp. 99–110, 2011.

[6] D. P. Bartel, "MicroRNA Target Recognition and Regulatory Functions," *Cell*, vol. 136, no. 2, pp. 215–233, 2009.

[7] G. Brancati and H. Großhans, "An interplay of miRNA abundance and target site architecture determines miRNA activity and specificity," *Nucleic Acids Res.*, vol. 46, no. 7, pp. 3259–3269, 2018.

[8] L. F. Gulyaeva and N. E. Kushlinskiy, "Regulatory mechanisms of microRNA expression," *J. Transl. Med.*, vol. 14, no. 1, pp. 1–10, 2016.

[9] H. Im and P. Kenny, "MicroRNAs in neuronal function and dysfunction," *Trends Neurosci.*, vol. 35, no. 5, pp. 325–334, 2012.

[10] S. Griffiths-Jones, H. K. Saini, S. Van Dongen, and A. J. Enright, "miRBase: Tools for microRNA genomics," *Nucleic Acids Res.*, vol. 36, no. SUPPL. 1, pp. 154–158, 2008.

[11] A. Kozomara and S. Griffiths-Jones, "MiRBase: Annotating high confidence microRNAs

using deep sequencing data," *Nucleic Acids Res.*, vol. 42, no. D1, pp. 68–73, 2014.

[12] H. R. Chiang *et al.*, "Mammalian microRNAs: Experimental evaluation of novel and previously annotated genes," *Genes Dev.*, vol. 24, no. 10, pp. 992–1009, 2010.

[13] Y. Lee, K. Jeon, J. Lee, S. Kim, and V. N. Kim, "4663.Full," vol. 21, no. 17, pp. 4663–4670, 2002.

[14] F. Ozsolak, L. Poling, Z. Wang, H. Liu, and XS, "Chromatin structure analyses identify miRNA promoters," *Genes Dev.*, vol. 23, pp. 3172–3183, 2008.

[15] A. M. Monteys *et al.*, "Structure and activity of putative intronic miRNAs promoters," *RNA (New York, NY)*, pp. 495–505, 2010.

[16] X. Cai, C. H. Hagedorn, and B. R. Cullen, "Human microRNAs are processed from capped, polyadenylated transcripts that can also function as mRNAs," *Rna*, vol. 10, no. 12, pp. 1957–1966, 2004.

[17] Y. Lee *et al.*, "MicroRNA genes are transcribed by RNA polymerase II," *EMBO J.*, vol. 23, no. 20, pp. 4051–4060, 2004.

[18] P. Graves and Y. Zeng, "Biogenesis of Mammalian MicroRNAs: A Global View," *Genomics, Proteomics Bioinforma.*, vol. 10, no. 5, pp. 239–245, 2012.

[19] J. Krol, I. Loedige, and W. Filipowicz, "The widespread regulation of microRNA biogenesis, function and decay," *Nat. Rev. Genet.*, vol. 11, no. 9, pp. 597–610, 2010.

[20] B. N. Davis-Dusenbery and A. Hata, "Mechanisms of control of microRNA biogenesis," *J. Biochem.*, vol. 148, no. 4, pp. 381–392, 2010.

[21] P. K. Singh and M. J. Campbell, "The interactions of microRNA and epigenetic modifications in prostate cancer," *Cancers (Basel).*, vol. 5, no. 3, pp. 998–1019, 2013.

[22] M. Pisarello, L. Loarca, T. Ivanics, L. Morton, and N. LaRusso, "MicroRNAs in the Cholangiopathies: Pathogenesis, Diagnosis, and Treatment," *J. Clin. Med.*, vol. 4, no. 9, pp. 1688–1712, 2015.

[23] Y. Lee *et al.*, "The nuclear RNase III Drosha initiates microRNA processing," *Nature*, vol. 425, no. 6956, pp. 415–419, 2003.

[24] J. Han, Y. Lee, K. Yeom, Y. Kim, H. Jin, and V. N. Kim, "The Drosha – DGCR8 complex in primary microRNA processing," *Genes Dev.*, pp. 3016–3027, 2004.

[25] V. Filippov, V. Solovyev, M. Filippova, and S. S. Gill, "A novel type of RNase III family proteins in eukaryotes," *Gene*, vol. 245, no. 1, pp. 213–221, 2000.

[26] J. Han *et al.*, "Molecular Basis for the Recognition of Primary microRNAs by the Drosha-DGCR8 Complex," *Cell*, vol. 125, no. 5, pp. 887–901, 2006.

[27] K. M. Herbert, G. Pimienta, S. J. DeGregorio, A. Alexandrov, and J. A. Steitz, "Phosphorylation of DGCR8 Increases Its Intracellular Stability and Induces a Progrowth miRNA Profile," *Cell Rep.*, vol. 5, no. 4, pp. 1070–1081, 2013.

[28] X. Tang *et al.*, "Acetylation of Drosha on the N-Terminus Inhibits Its Degradation by Ubiquitination," *PLoS One*, vol. 8, no. 8, 2013.

[29] A. Shiohama, T. Sasaki, S. Noda, S. Minoshima, and N. Shimizu, "Nucleolar localization of DGCR8 and identification of eleven DGCR8-associated proteins," *Exp. Cell Res.*, vol. 313, no. 20, pp. 4196–4207, 2007.

[30] T. L. Cheng *et al.*, "MeCP2 Suppresses Nuclear MicroRNA Processing and Dendritic Growth by Regulating the DGCR8/Drosha Complex," *Dev. Cell*, vol. 28, no. 5, pp. 547–560, 2014.

[31] M. T. Bohnsack, K. Czaplinski, and D. Görlich, "Exportin 5 is a RanGTP-dependent dsRNA-binding protein that mediates nuclear export of pre-miRNAs," *Rna*, vol. 10, no. 2, pp. 185–191, 2004.

[32] R. Lakes and S. Saha, "American Association for the Advancement of Science," *Nature*, vol. 204, no. May, pp. 501–503, 1979.

[33] K. Förstemann, M. D. Horwich, L. Wee, Y. Tomari, and P. D. Zamore, "Drosophila

microRNAs Are Sorted into Functionally Distinct Argonaute Complexes after Production by Dicer-1," *Cell*, vol. 130, no. 2, pp. 287–297, 2007.

[34] G. Meister, M. Landthaler, A. Patkaniowska, Y. Dorsett, G. Teng, and T. Tuschl, "Human Argonaute2 mediates RNA cleavage targeted by miRNAs and siRNAs," *Mol. Cell*, vol. 15, no. 2, pp. 185–197, 2004.

[35] K. Nakanishi, "Anatomy of RISC: how do small RNAs and chaperones activate Argonaute proteins?," *Wiley Interdiscip. Rev. RNA*, vol. 7, no. 5, pp. 637–660, 2016.

[36] T. Kawamata, H. Seitz, and Y. Tomari, "Structural determinants of miRNAs for RISC loading and slicer-independent unwinding," *Nat. Struct. Mol. Biol.*, vol. 16, no. 9, pp. 953–960, 2009.

[37] J. E. Thornton and R. I. Gregory, "NIH Public Access," vol. 22, no. 9, pp. 474–482, 2013.

[38] T.-C. Chang *et al.*, "Lin-28B transactivation is necessary for Myc-mediated let-7 repression and proliferation.," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 106, no. 9, pp. 3384–9, 2009.

[39] M. Zhu, X. Deng, T. Joshi, D. Xu, G. Stacey, and J. Cheng, "Reconstructing differentially co-expressed gene modules and regulatory networks of soybean cells," *BMC Genomics*, vol. 13, no. 1, 2012.

[40] E. Segal *et al.*, "Module networks: Identifying regulatory modules and their condition-specific regulators from gene expression data," *Nat. Genet.*, vol. 34, no. 2, pp. 166–176, 2003.

[41] A. Zinovyev, J. Lages, and D. L. Shepelyansky, "Inferring hidden causal relations between pathway members using reduced Google matrix of directed biological networks Inferring hidden causal relations between pathway members using reduced Google matrix of directed biological networks," *bioRxiv*, no. December, pp. 1–18, 2016.

[42] M. S. Ebert and P. A. Sharp, "Roles for MicroRNAs in conferring robustness to biological processes," *Cell*, vol. 149, no. 3, pp. 505–524, 2012.

[43] K. Chen and N. Rajewsky, “Deep Conservation of MicroRNA-target Relationships and 3 ’ UTR Motifs in Vertebrates , Flies , and Nematodes Deep Conservation of MicroRNA-target Relationships and 3 ′ UTR Motifs in Vertebrates , Flies , and Nematodes,” *Cold Spring Harb. Symp. Quant. Biol.*, vol. 71, pp. 149–156, 2006.

[44] X. Xie *et al.*, “Systematic discovery of regulatory motifs in human promoters and 3′ UTRs by comparison of several mammals,” *Nature*, vol. 434, no. 7031, pp. 338–345, 2005.

[45] G. Boross, K. Orosz, and I. J. Farkas, “Human microRNAs co-silence in well-separated groups and have different predicted essentialities,” *Bioinformatics*, vol. 25, no. 8, pp. 1063–1069, 2009.

[46] J. Xu *et al.*, “MiRNA-miRNA synergistic network: Construction via co-regulating functional modules and disease miRNA topological features,” *Nucleic Acids Res.*, vol. 39, no. 3, pp. 825–836, 2011.

[47] X. Chen *et al.*, “RBMMMDA: Predicting multiple types of disease-microRNA associations,” *Sci. Rep.*, vol. 5, pp. 1–13, 2015.

[48] M. V. Iorio *et al.*, “MicroRNA gene expression deregulation in human breast cancer,” *Cancer Res.*, vol. 65, no. 16, pp. 7065–7070, 2005.

[49] N. Yanaihara *et al.*, “Unique microRNA molecular profiles in lung cancer diagnosis and prognosis,” *Cancer Cell*, vol. 9, no. 3, pp. 189–198, 2006.

[50] K. P. Porkka, M. J. Pfeiffer, K. K. Waltering, R. L. Vessella, T. L. J. Tammela, and T. Visakorpi, “MicroRNA expression profiling in prostate cancer,” *Cancer Res.*, vol. 67, no. 13, pp. 6130–6135, 2007.

[51] Y. Akao, Y. Nakagawa, and T. Naoe, “MicroRNA-143 and -145 in Colon Cancer,” *DNA Cell Biol.*, vol. 26, no. 5, pp. 311–320, 2007.

[52] H. Yang *et al.*, “MicroRNA Expression Profiling in Human Ovarian Cancer: miR-214 Induces Cell Survival and Cisplatin Resistance by Targeting PTEN,” *Cancer Res.*, vol. 68, no. 2, pp. 425–433, 2008.

[53] L. F. Sempere, S. Freemantle, I. Pitha-Rowe, E. Moss, E. Dmitrovsky, and V. Ambros, "Expression profiling of mammalian microRNAs uncovers a subset of brain-expressed microRNAs with possible roles in murine and human neuronal differentiation," *Genome Biol.*, vol. 5, no. 3, p. R13, 2004.

[54] Q. Jiang *et al.*, "Prioritization of disease microRNAs through a human phenome-microRNAome network," *BMC Syst. Biol.*, vol. 4, no. SUPPL. 1, pp. 1–9, 2010.

[55] H. Shi *et al.*, "Walking the interactome to identify human miRNA-disease associations through the functional link between miRNA targets and disease genes," *BMC Syst. Biol.*, vol. 7, 2013.

[56] X. Chen, M. X. Liu, and G. Y. Yan, "RWRMDA: Predicting novel human microRNA-disease associations," *Mol. Biosyst.*, vol. 8, no. 10, pp. 2792–2798, 2012.

[57] P. Xuan *et al.*, "Prediction of microRNAs Associated with Human Diseases Based on Weighted k Most Similar Neighbors," *PLoS One*, vol. 8, no. 8, 2013.

[58] W. Zhu and Y. P. P. Chen, "Computational developments in microRNA-regulated protein-protein interactions," *BMC Syst. Biol.*, vol. 8, no. 1, pp. 1–11, 2014.

[59] L. H. Hartwell, J. J. Hopfield, S. Leibler, and A. W. Murray, "From molecular to modular cell biology," *Nature*, vol. 402, no. 6761supp, pp. C47–C52, 1999.

[60] S. Fields and O. Song, "A novel genetic system to detect protein-protein interactions. [Yeast two hybrid]," *Nature*, vol. 340, no. 6230, pp. 245–246, 1989.

[61] B. Schwikowski, P. Uetz, and S. Fields, "A network of protein-protein interactions in yeast.," *Nat. Biotechnol.*, vol. 18, no. 12, pp. 1257–1261, 2000.

[62] U. Stelzl *et al.*, "A human protein-protein interaction network: A resource for annotating the proteome," *Cell*, vol. 122, no. 6, pp. 957–968, 2005.

[63] T. Ito, T. Chiba, R. Ozawa, M. Yoshida, M. Hattori, and Y. Sakaki, "A comprehensive two-hybrid analysis to explore the yeast protein interactome," *Proc. Natl. Acad. Sci.*, vol.

98, no. 8, pp. 4569–4574, 2001.

[64] C. E. Rogler *et al.*, "MicroRNA-23b cluster microRNAs regulate transforming growth factor-beta/bone morphogenetic protein signaling and liver stem cell differentiation by targeting Smads," *Hepatology*, vol. 50, no. 2, pp. 575–584, 2009.

[65] J. Mei, R. Bachoo, and C.-L. Zhang, "MicroRNA-146a Inhibits Glioma Development by Targeting Notch1," *Mol. Cell. Biol.*, vol. 31, no. 17, pp. 3584–3592, 2011.

[66] T. Papagiannakopoulos *et al.*, "Pro-neural miR-128 is a glioma tumor suppressor that targets mitogenic kinases," *Oncogene*, vol. 31, no. 15, pp. 1884–1895, 2012.

[67] A. J. Enright, S. Van Dongen, and C. A. Ouzounis, "An efficient algorithm for large-scale detection of protein families.," *Nucleic Acids Res.*, vol. 30, no. 7, pp. 1575–1584, 2002.

[68] G. Liu, L. Wong, and H. N. Chua, "Complex discovery from weighted PPI networks," *Bioinformatics*, vol. 25, no. 15, pp. 1891–1897, 2009.

[69] M. Inui, G. Martello, and S. Piccolo, "MicroRNA control of signal transduction," *Nat. Rev. Mol. Cell Biol.*, vol. 11, no. 4, pp. 252–263, 2010.

[70] L. Perfetto *et al.*, "SIGNOR: A database of causal relationships between biological entities," *Nucleic Acids Res.*, vol. 44, no. D1, pp. D548–D554, 2016.

[71] P. Lo Surdo, A. Calderone, G. Cesareni, and L. Perfetto, "SIGNOR: A database of causal relationships between biological entities-a short guide to searching and browsing," *Curr. Protoc. Bioinforma.*, vol. 2017, no. June, p. 8.23.1-8.23.16, 2017.

[72] P. de Haan, H. C. Klein, and B. A. 't Hart, "Autoimmune aspects of neurodegenerative and psychiatric diseases: A template for innovative therapy," *Front. Psychiatry*, vol. 8, no. APR, pp. 1–11, 2017.

[73] A. Labadorf, S. H. Choi, and R. H. Myers, "Evidence for a Pan-Neurodegenerative Disease Response in Huntington's and Parkinson's Disease Expression Profiles," *Front.*

Mol. Neurosci., vol. 10, no. January, pp. 1–12, 2018.

[74] S. Moradifard, M. Hoseinbeyki, S. M. Ganji, and Z. Minuchehr, "Analysis of microRNA and Gene Expression Profiles in Alzheimer's Disease: A Meta-Analysis Approach," *Sci. Rep.*, vol. 8, no. 1, pp. 1–17, 2018.

[75] X. Li *et al.*, "Identification of active transcription factor and miRNA regulatory pathways in Alzheimer's disease," *Bioinformatics*, vol. 29, no. 20, pp. 2596–2602, 2013.

[76] A. Labadorf *et al.*, "RNA sequence analysis of human huntington disease brain reveals an extensive increase in inflammatory and developmental gene expression," *PLoS One*, vol. 10, no. 12, pp. 1–21, 2015.

[77] A. Dumitriu *et al.*, "Integrative analyses of proteomics and RNA transcriptomics implicate mitochondrial processes, protein folding pathways and GWAS loci in Parkinson disease," *BMC Med. Genomics*, vol. 9, no. 1, pp. 1–17, 2016.

[78] E. Wingender, P. Dietze, H. Karas, and R. Knüppel, "TRANSFAC: A database on transcription factors and their DNA binding sites," *Nucleic Acids Res.*, vol. 24, no. 1, pp. 238–241, 1996.

[79] C. Jiang, Z. Xuan, F. Zhao, and M. Q. Zhang, "TRED: A transcriptional regulatory element database, new entries and other development," *Nucleic Acids Res.*, vol. 35, no. SUPPL. 1, pp. 140–143, 2007.

[80] J. Piñero *et al.*, "DisGeNET: A discovery platform for the dynamical exploration of human diseases and their genes," *Database*, vol. 2015, pp. 1–17, 2015.

[81] Y. Li *et al.*, "HMDD v2.0: A database for experimentally supported human microRNA and disease associations," *Nucleic Acids Res.*, vol. 42, no. D1, pp. 1070–1074, 2014.

[82] Q. Jiang *et al.*, "miR2Disease: A manually curated database for microRNA deregulation in human disease," *Nucleic Acids Res.*, vol. 37, no. SUPPL. 1, pp. 98–104, 2009.

[83] J. Wang, M. Lu, C. Qiu, and Q. Cui, "TransmiR: A transcription factor microRNA

regulation database," *Nucleic Acids Res.*, vol. 38, no. SUPPL.1, pp. 119–122, 2009.

[84] S. Da Hsu *et al.*, "MiRTarBase: A database curates experimentally validated microRNA-target interactions," *Nucleic Acids Res.*, vol. 39, no. SUPPL. 1, pp. 163–169, 2011.

[85] F. Xiao, Z. Zuo, G. Cai, S. Kang, X. Gao, and T. Li, "miRecords: An integrated resource for microRNA-target interactions," *Nucleic Acids Res.*, vol. 37, no. SUPPL. 1, pp. 105–110, 2009.

[86] G. L. Papadopoulos, M. Reczko, V. A. Simossis, P. Sethupathy, and A. G. Hatzigeorgiou, "The database of experimentally supported targets: A functional update of TarBase," *Nucleic Acids Res.*, vol. 37, no. SUPPL. 1, pp. 155–158, 2009.

[87] F. Finotello and B. Di Camillo, "Measuring differential gene expression with RNA-seq: Challenges and strategies for data analysis," *Brief. Funct. Genomics*, vol. 14, no. 2, pp. 130–142, 2015.

[88] G. Csárdi and T. Nepusz, "The igraph software package for complex network research," *InterJournal Complex Syst.*, vol. 1695, pp. 1–9, 2006.

[89] E. Ulgen, O. Ozisik, and O. U. Sezerman, "pathfindR: An R Package for Pathway Enrichment Analysis Utilizing Active Subnetworks," *bioRxiv*, p. 272450, 2018.

[90] I. S. Vlachos *et al.*, "DIANA-miRPath v3.0: Deciphering microRNA function with experimental support," *Nucleic Acids Res.*, vol. 43, no. W1, pp. W460–W466, 2015.

[91] M. Safran *et al.*, "GeneCards Version 3: the human gene integrator.," *Database (Oxford).*, vol. 2010, pp. 1–16, 2010.

[92] F. Jordan, T.-P. Nguyen, and W. -c. Liu, "Studying protein-protein interaction networks: a systems view on diseases," *Brief. Funct. Genomics*, vol. 11, no. 6, pp. 497–504, 2012.

[93] D. Nitsch *et al.*, "Network analysis of differential expression for the identification of disease-causing genes," *PLoS One*, vol. 4, no. 5, 2009.

[94] J. Zhao, T.-H. Yang, Y. Huang, and P. Holme, "Ranking Candidate Disease Genes from

Gene Expression and Protein Interaction: A Katz-Centrality Based Approach," *PLoS One*, vol. 6, no. 9, p. e24306, 2011.

[95] K. A. Mueller *et al.*, "Hippo Signaling Pathway Dysregulation in Human Huntington's Disease Brain and Neuronal Stem Cells," *Sci. Rep.*, vol. 8, no. 1, p. 11355, 2018.

[96] M. Salvalaio *et al.*, "Brain RNA-seq profiling of the mucopolysaccharidosis type II mouse model," *Int. J. Mol. Sci.*, vol. 18, no. 5, 2017.

[97] J. D. Esko, K. Kimata, and U. Lindahl, "Chapter 16Proteoglycans and Sulfated Glycosaminoglycans," no. Chapter 16, 1960.

[98] P. A. Desplats *et al.*, "Glycolipid and ganglioside metabolism imbalances in Huntington's disease," *Neurobiol. Dis.*, vol. 27, no. 3, pp. 265–277, 2007.

[99] A. S. B. Olsen and N. J. Færgeman, "Sphingolipids: membrane microdomains in brain development, function and neurological diseases.," *Open Biol.*, vol. 7, no. 5, p. 170069, 2017.

[100] A. Kihara, "Very long-chain fatty acids: Elongation, physiology and related disorders," *J. Biochem.*, vol. 152, no. 5, pp. 387–395, 2012.

[101] and D. S. H. Xiwei Zheng, Cong Bi, Marissa Brooks, "HHS Public Access," *Anal Chem.*, vol. 25, no. 4, pp. 368–379, 2015.

[102] E. Martí *et al.*, "A myriad of miRNA variants in control and Huntington's disease brain regions detected by massively parallel sequencing," *Nucleic Acids Res.*, vol. 38, no. 20, pp. 7219–7235, 2010.

[103] R. C. Block, E. R. Dorsey, C. A. Beck, J. Thomas, and I. Shoulson, "NIH Public Access," vol. 4, no. 1, pp. 17–23, 2011.

[104] R. Goold, C. McKinnon, and S. J. Tabrizi, "Prion degradation pathways: Potential for therapeutic intervention," *Mol. Cell. Neurosci.*, vol. 66, no. PA, pp. 12–20, 2015.

[105] Q. Liu *et al.*, "The CXCL8-CXCR1/2 pathways in cancer," *Cytokine Growth Factor Rev.*,

vol. 31, pp. 61–71, 2016.

[106] H. Ha, B. Debnath, and N. Neamati, “Role of the CXCL8-CXCR1/2 axis in cancer and inflammatory diseases,” *Theranostics*, vol. 7, no. 6, pp. 1543–1588, 2017.

[107] A. Vallès, L. Grijpink-Ongering, F. M. de Bree, T. Tuinstra, and E. Ronken, “Differential regulation of the CXCR2 chemokine network in rat brain trauma: Implications for neuroimmune interactions and neuronal survival,” *Neurobiol. Dis.*, vol. 22, no. 2, pp. 312–322, 2006.

[108] M. Lindner, C. Trebst, S. Heine, T. Skripuletz, P. N. Koutsoudaki, and M. Stangel, “The chemokine receptor CXCR2 is differentially regulated on glial cells in vivo but is not required for successful remyelination after cuprizone-induced demyelination,” *Glia*, vol. 56, no. 10, pp. 1104–1113, 2008.

[109] J. Bourbon-Teles, S. Bells, D. K. Jones, E. Coulthard, A. Rosser, and C. Metzler-Baddeley, “Myelin breakdown in human Huntington’s disease: Multi-modal evidence from diffusion MRI and quantitative magnetization transfer,” *Neuroscience*, 2017.

[110] S. Mélik-Parsadaniantz and W. Rostène, “Chemokines and neuromodulation,” *J. Neuroimmunol.*, vol. 198, no. 1–2, pp. 62–68, 2008.

[111] B. F. G. Popescu, I. Pirko, and C. F. Lucchinetti, “Pathology of multiple sclerosis: where do we stand?,” *Continuum (Minneap. Minn).*, vol. 19, no. 4 Multiple Sclerosis, pp. 901–21, 2013.

[112] M. Bredholt and J. L. Frederiksen, “Zinc in multiple sclerosis: A systematic review and meta-analysis,” *ASN Neuro*, vol. 8, no. 3, pp. 1–9, 2016.

[113] M. Foster and S. Samman, “Zinc and Regulation of inflammatory cytokines: Implications for cardiometabolic disease,” *Nutrients*, vol. 4, no. 7, pp. 676–694, 2012.

[114] J. F. Atherton, E. L. McIver, M. R. M. Mullen, D. L. Wokosin, D. J. Surmeier, and M. D. Bevan, “Early dysfunction and progressive degeneration of the subthalamic nucleus in mouse models of huntington’s disease,” *Elife*, vol. 5, no. DECEMBER2016, pp. 1–26,

2016.

[115] K. D. Taganov, M. P. Boldin, K.-J. Chang, and D. Baltimore, "NF- B-dependent induction of microRNA miR-146, an inhibitor targeted to signaling proteins of innate immune responses," *Proc. Natl. Acad. Sci.*, vol. 103, no. 33, pp. 12481–12486, 2006.

[116] J. Ghose, M. Sinha, E. Das, N. R. Jana, and N. P. Bhattacharyya, "Regulation of miR-146a by RelA/NFkB and p53 in STHdh q111/Hdh q111 cells, a cell model of Huntington's disease," *PLoS One*, vol. 6, no. 8, 2011.

[117] F. Trettel, "Dominant phenotypes produced by the HD mutation in STHdhQ111 striatal cells," *Hum. Mol. Genet.*, vol. 9, no. 19, pp. 2799–2809, 2000.

[118] Z. Feng *et al.*, "P53 Tumor Suppressor Protein Regulates the Levels of Huntingtin Gene Expression," *Oncogene*, vol. 25, no. 1, pp. 1–7, 2006.

[119] D. Son *et al.*, "NIH Public Access," vol. 86, no. 3, pp. 573–579, 2007.

[120] E. Kim *et al.*, "Tumor suppressor p53 inhibits transcriptional activation of invasion gene thromboxane synthase mediated by the proto-oncogenic factor ets-1," *Oncogene*, vol. 22, no. 49, pp. 7716–7727, 2003.

[121] N. K. H. Lim *et al.*, "Localized changes to glycogen synthase kinase-3 and collapsin response mediator protein-2 in the Huntington's disease affected brain," *Hum. Mol. Genet.*, vol. 23, no. 15, pp. 4051–4063, 2014.

[122] O. C. Jorge-Torres *et al.*, "Inhibition of Gsk3b Reduces Nfkb1 Signaling and Rescues Synaptic Activity to Improve the Rett Syndrome Phenotype in Mecp2-Knockout Mice," *Cell Rep.*, vol. 23, no. 6, pp. 1665–1677, 2018.

[123] A. Giralt *et al.*, "Brain-derived neurotrophic factor modulates the severity of cognitive alterations induced by mutant huntingtin: Involvement of phospholipaseCγ activity and glutamate receptor expression," *Neuroscience*, vol. 158, no. 4, pp. 1234–1250, 2009.

[124] R. J. Hoefen, "The multifunctional GIT family of proteins," *J. Cell Sci.*, vol. 119, no. 8,

pp. 1469–1475, 2006.

[125] B. Huck, R. Kemkemer, M. Franz-Wachtel, B. Macek, A. Hausser, and M. A. Olayioye, "GIT1 Phosphorylation on Serine 46 by PKD3 regulates paxillin trafficking and cellular protrusive activity," *J. Biol. Chem.*, vol. 287, no. 41, pp. 34604–34613, 2012.

[126] K. Quann, Y. Jing, and I. Rigoutsos, "Post-transcriptional regulation of BRCA1 through its coding sequence by the miR-15/107 group of miRNAs," *Front. Genet.*, vol. 6, no. JUL, pp. 1–10, 2015.

[127] D. Bhaumik *et al.*, "MicroRNAs miR-146a/b negatively modulate the senescence-associated inflammatory mediators IL-6 and IL-8.," *Aging (Albany. NY).*, vol. 1, no. 4, pp. 402–411, 2009.

[128] L. Leggio *et al.*, "MicroRNAs in parkinson's disease: From pathogenesis to novel diagnostic and therapeutic approaches," *Int. J. Mol. Sci.*, vol. 18, no. 12, 2017.

[129] I. Peter *et al.*, "Anti–Tumor Necrosis Factor Therapy and Incidence of Parkinson Disease Among Patients With Inflammatory Bowel Disease," *JAMA Neurol.*, vol. 10029, pp. 1–8, 2018.

APPENDIX A SUPPLEMENTARY TABLES AND FIGURES

8.1 Huntington Disease miRNA Regulatory Pathways Subgroups

Figure A.4.1.1 Huntington Disease Subgroup 1

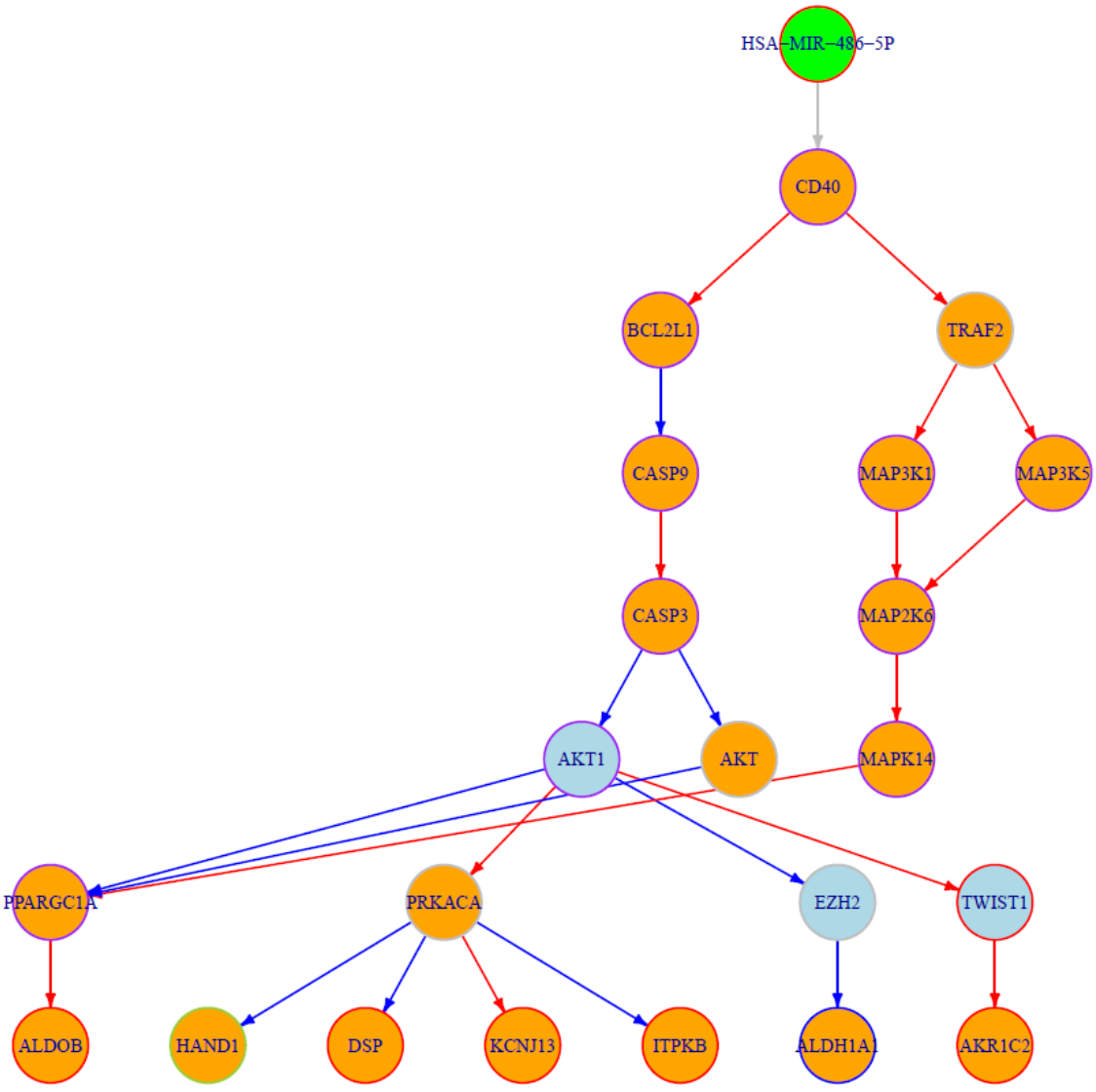


Figure A.4.1.2. Huntington Disease Subgroup 2

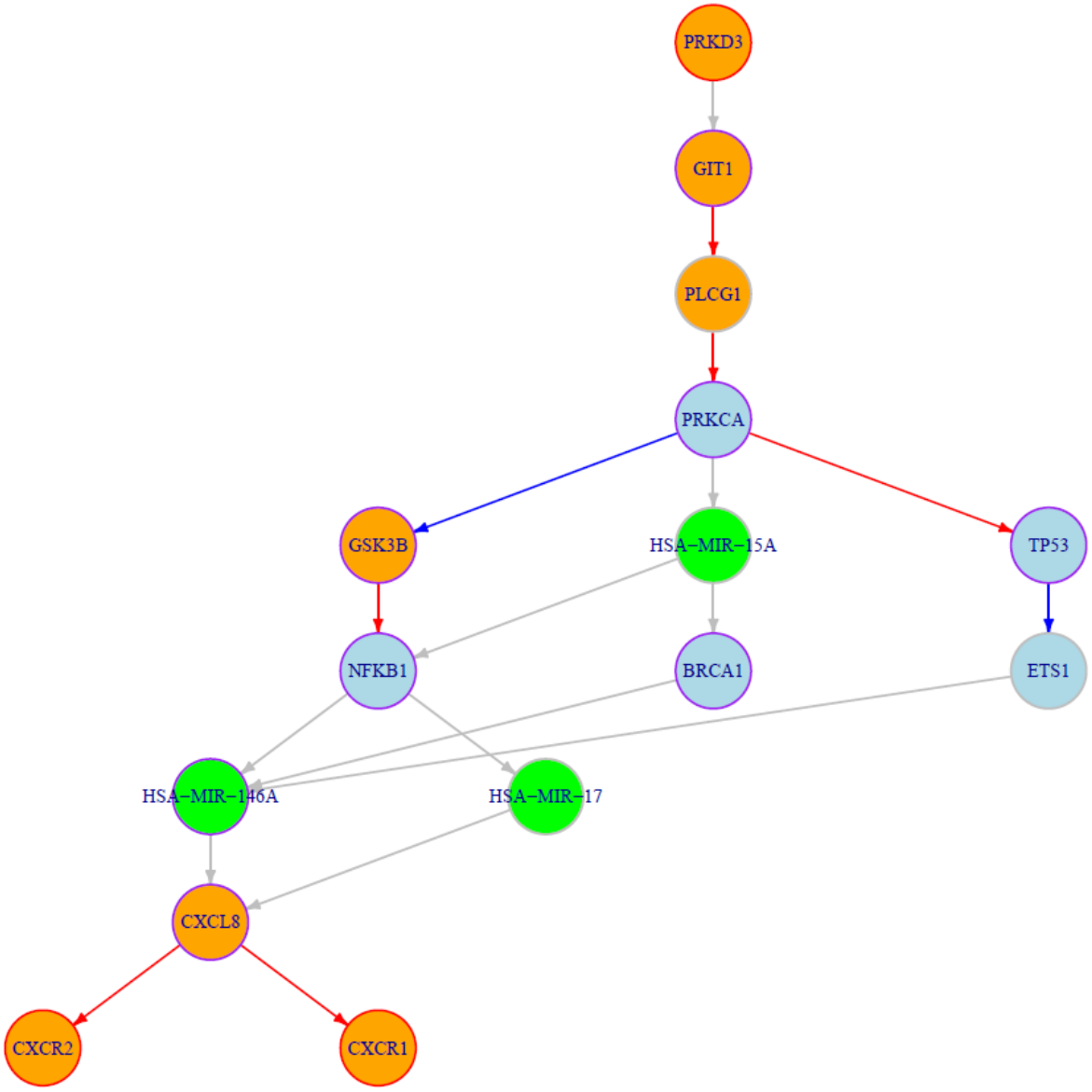


Figure A.4.1.3. Huntington Disease Subgroup 3

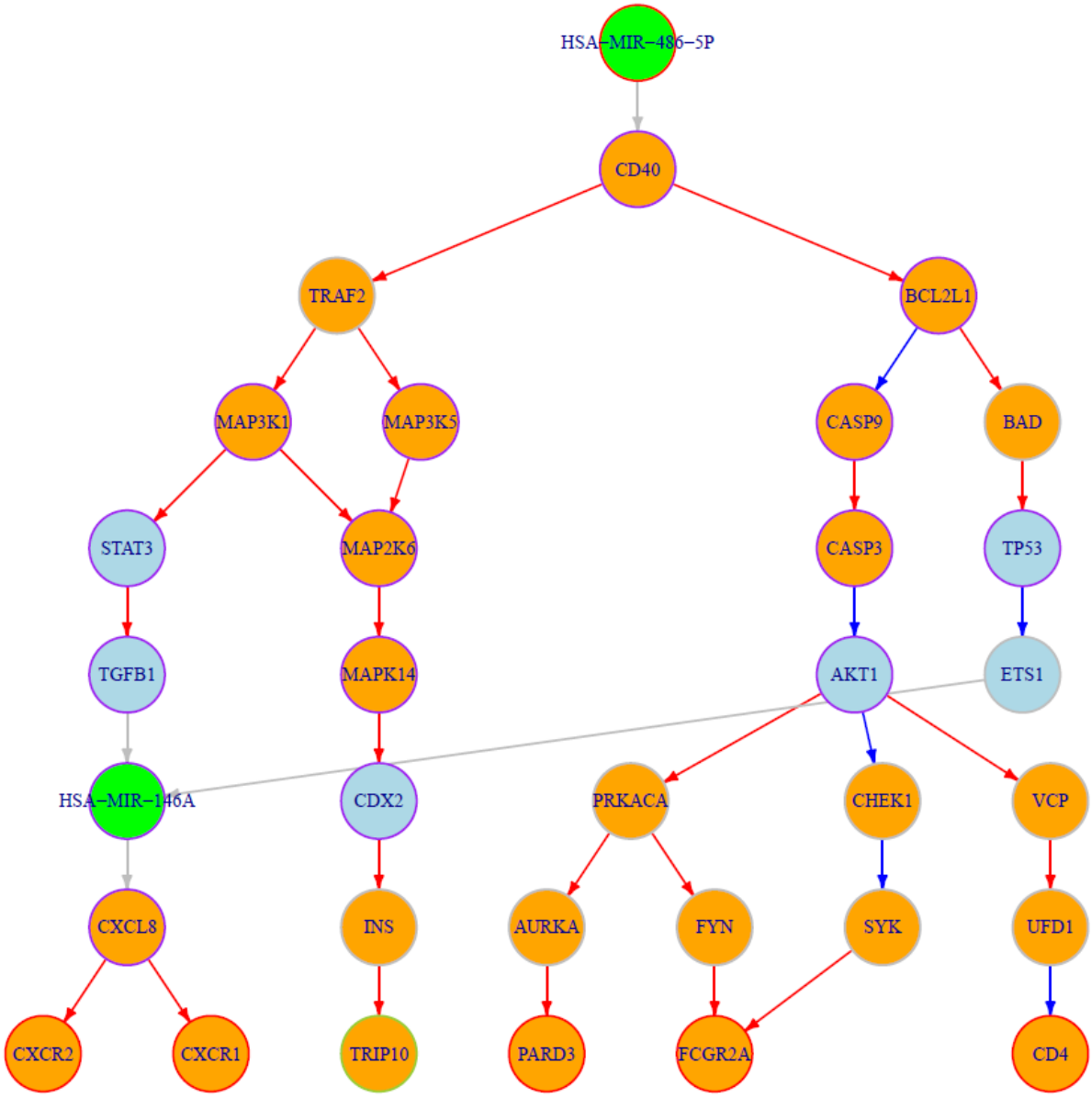


Figure A.4.1.4. Huntington Disease Subgroup 4

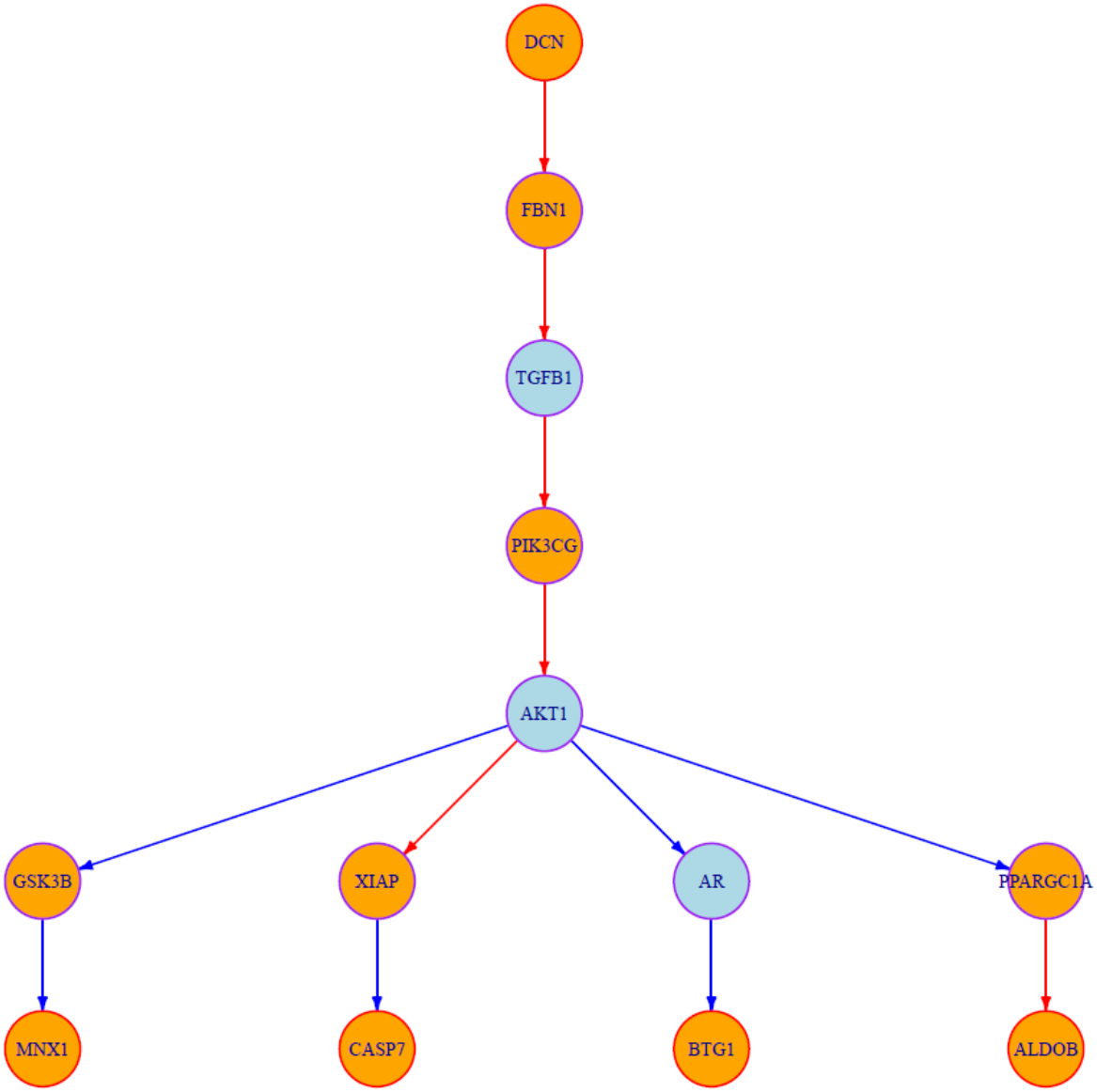


Figure A.4.1.5. Huntington Disease Subgroup 5

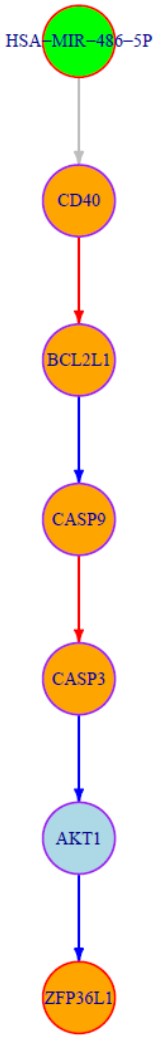


Figure A.4.1.6. Huntington Disease Subgroup 6

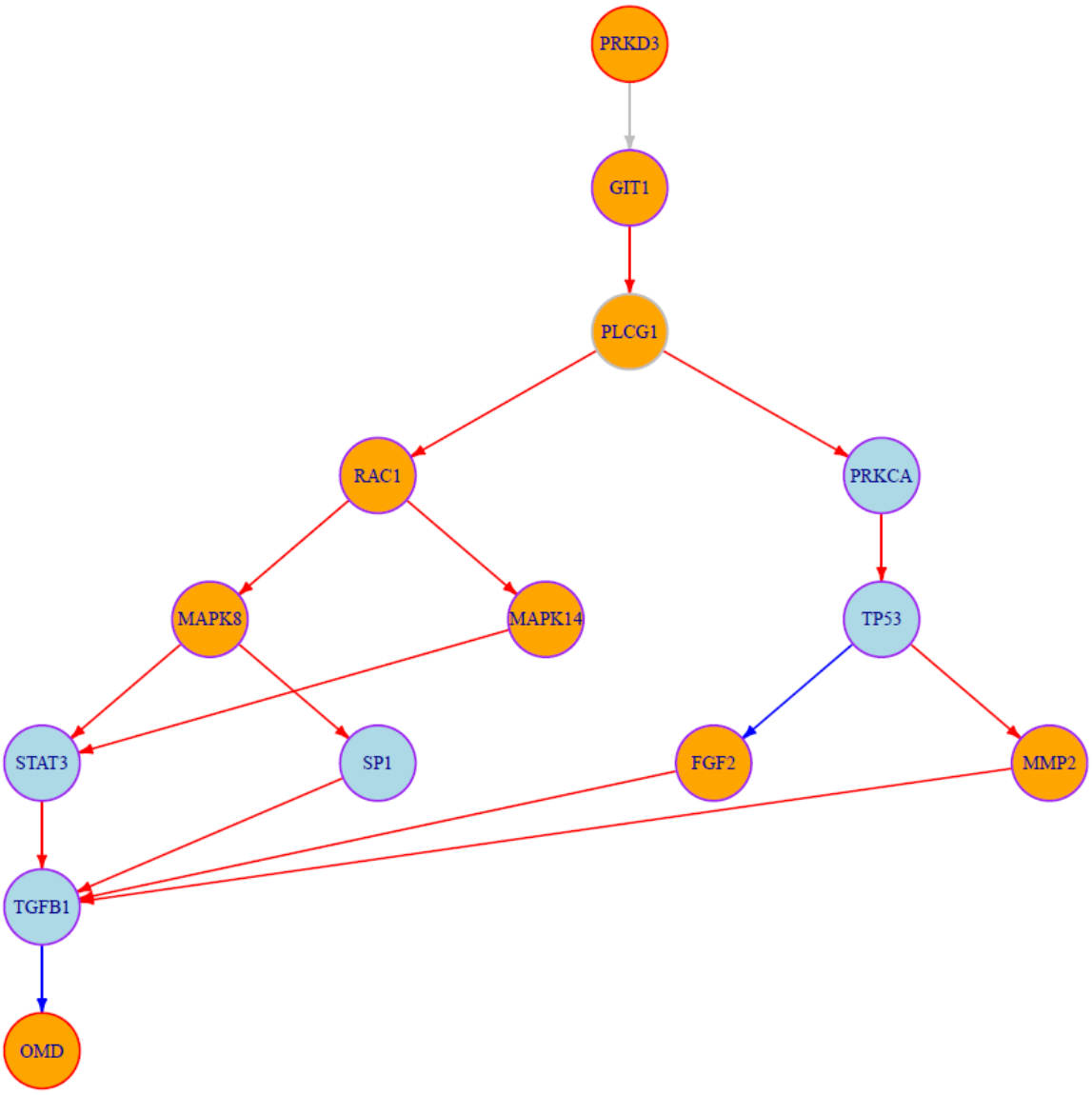


Figure A.4.1.7. Huntington Disease Subgroup 7

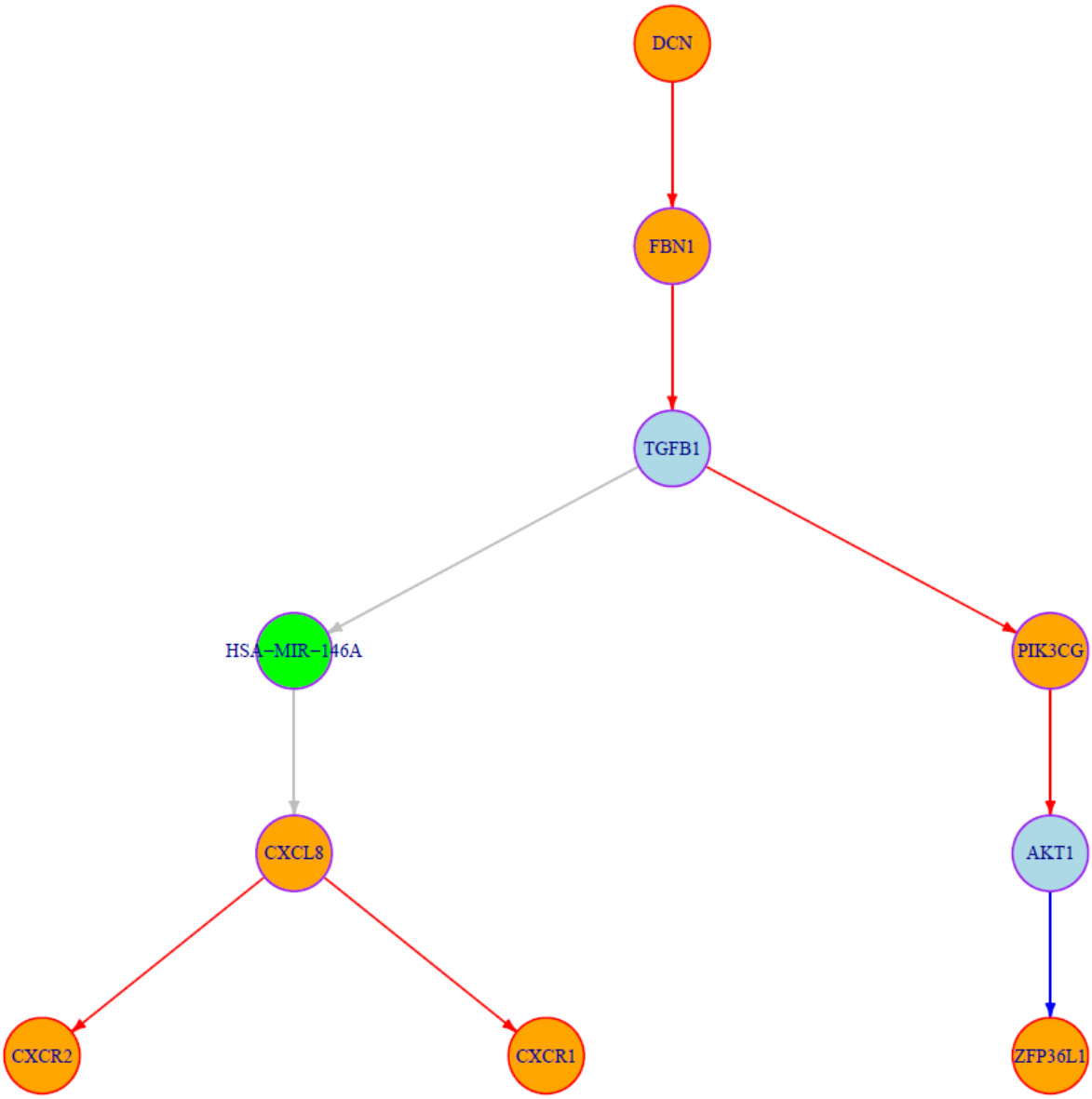
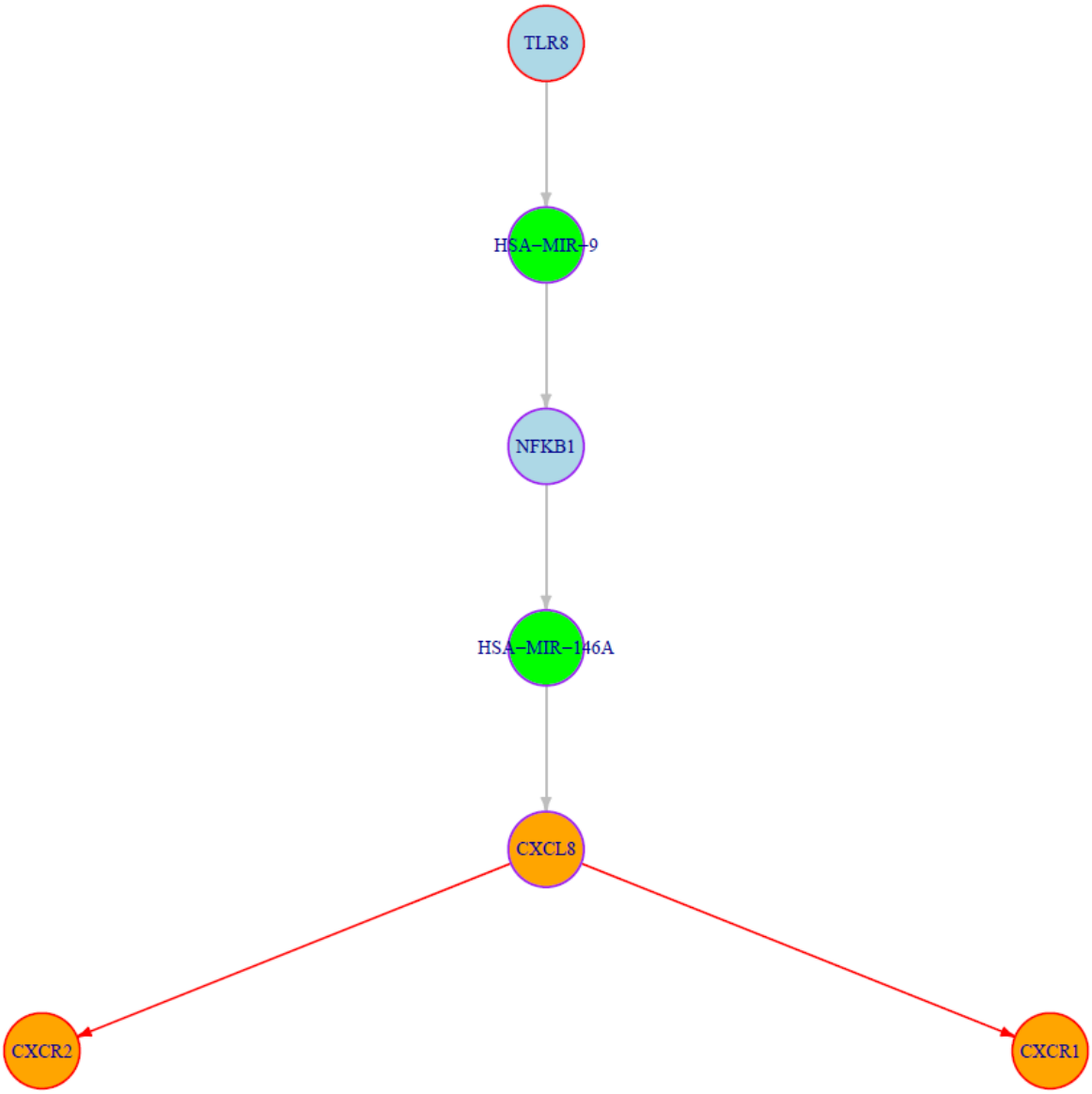


Figure A.4.1.8. Huntington Disease Subgroup 8



8.2 Parkinson Disease miRNA Regulatory Pathways Subgroups

Figure A.4.2.1. Parkinson Disease Subgroup 1 - 2

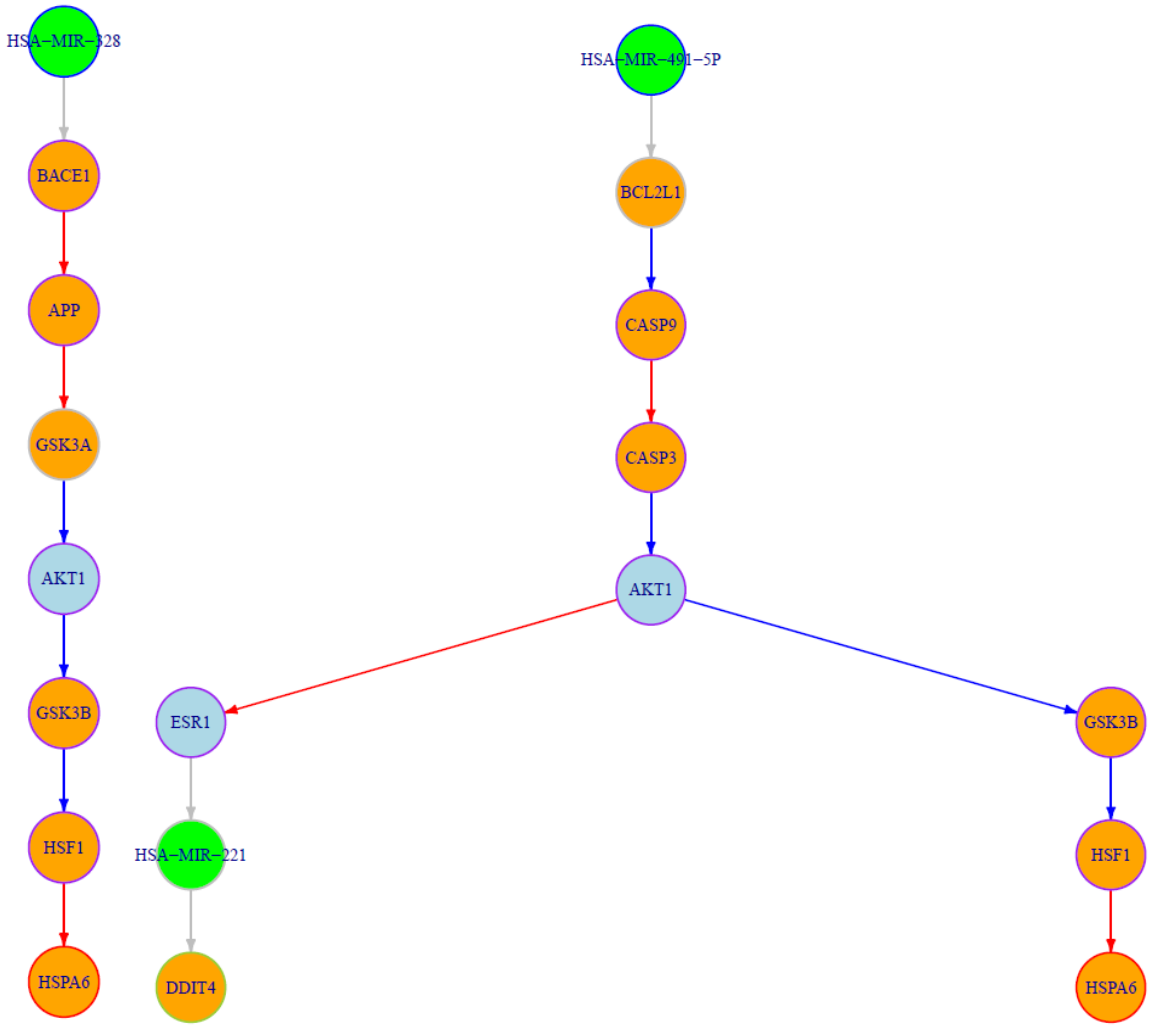


Figure A.4.2.2. Parkinson Disease Subgroup 3 - 4

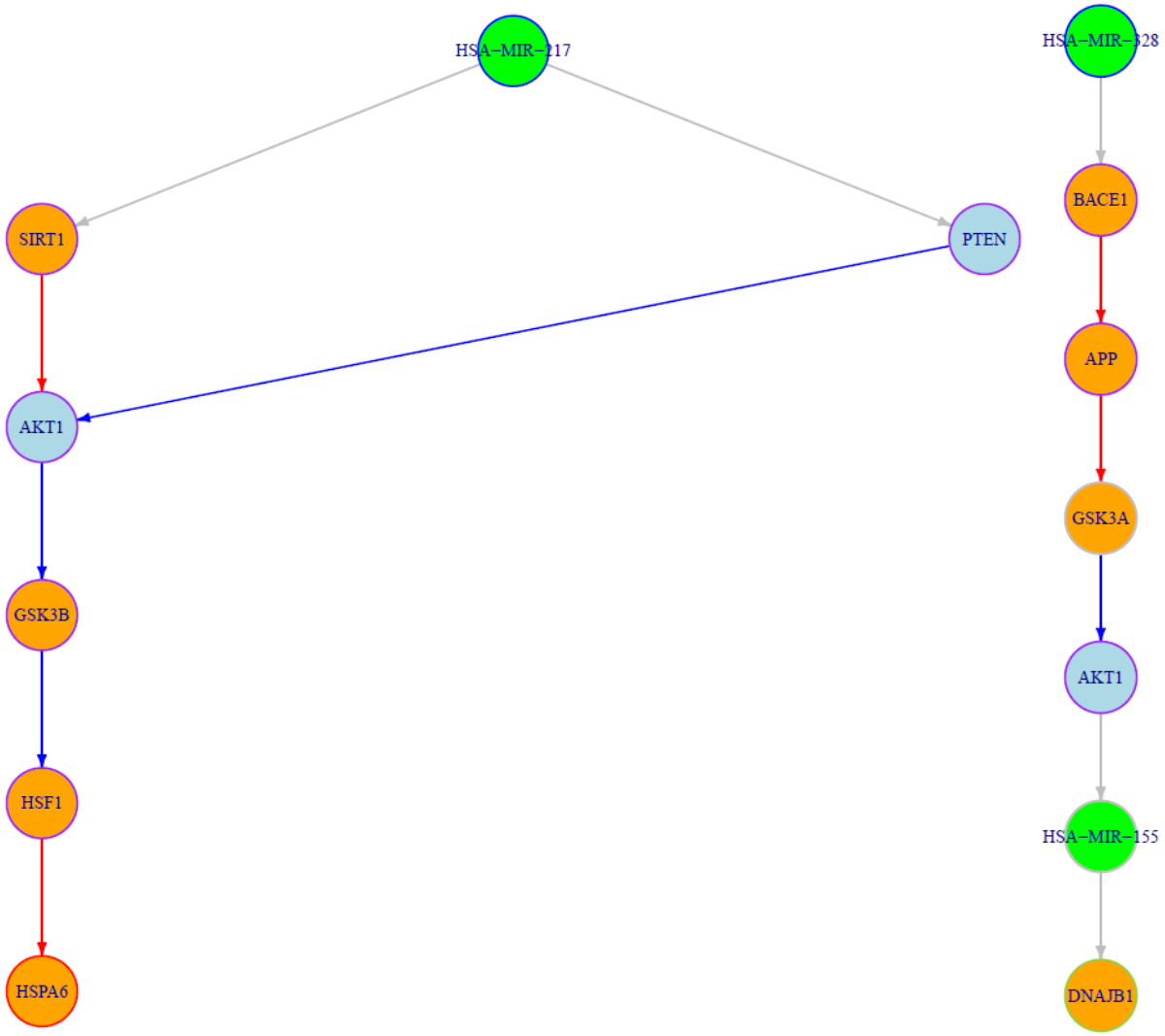


Figure A.4.2.3. Parkinson Disease Subgroup 5 - 6

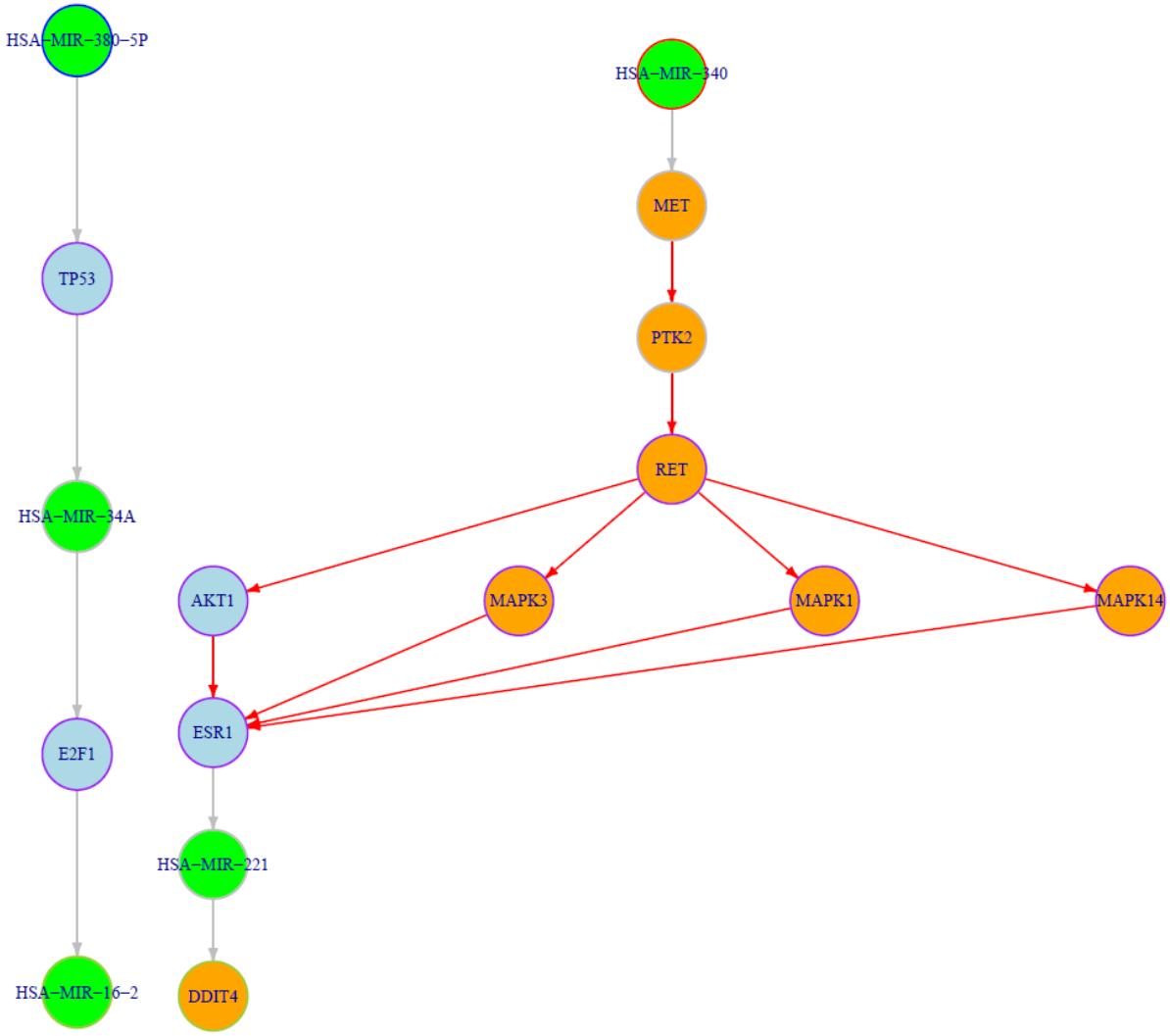


Figure A.4.2.4. Parkinson Disease Subgroup 7

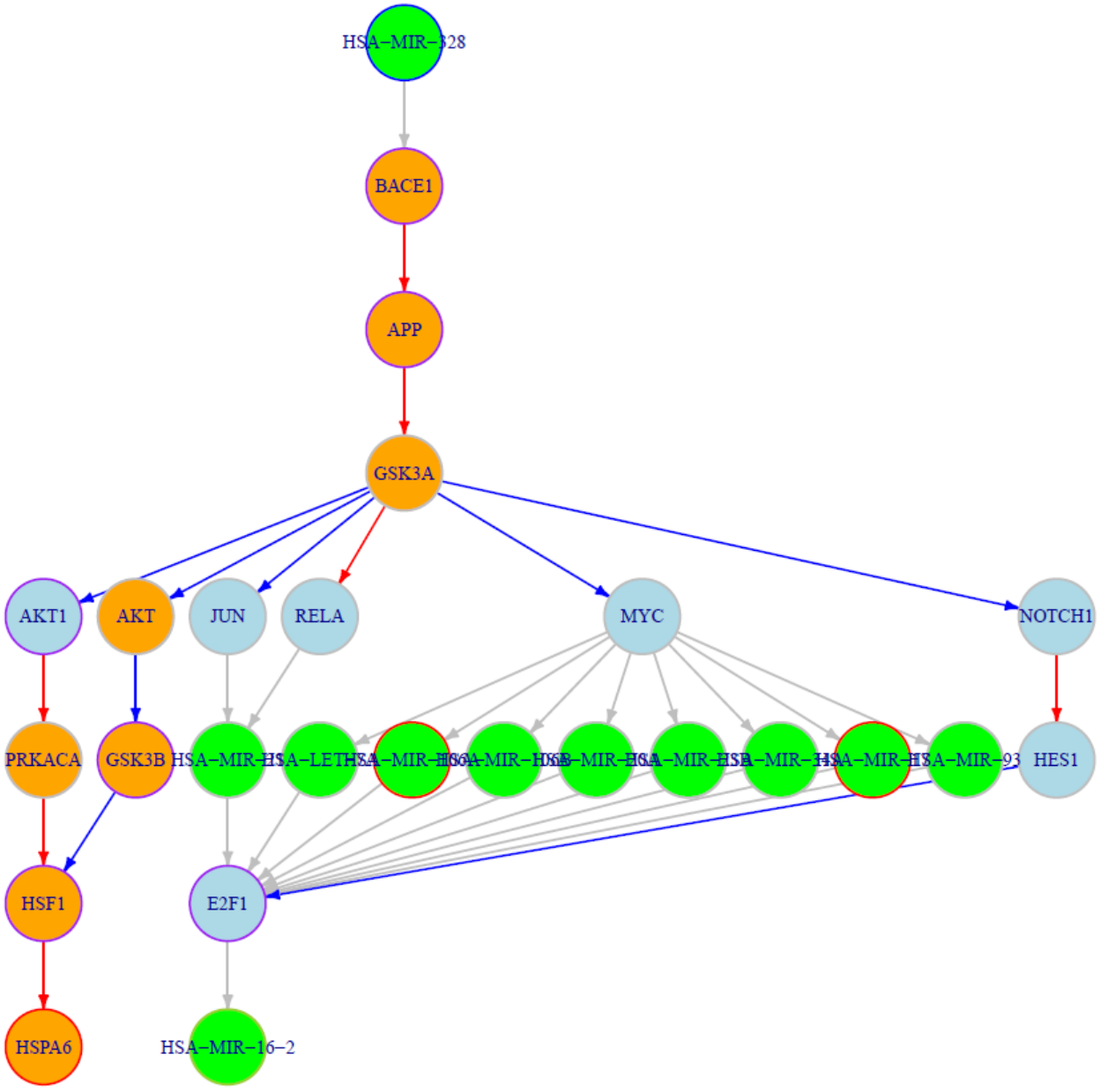


Figure A.4.2.5. Parkinson Disease Subgroup 8

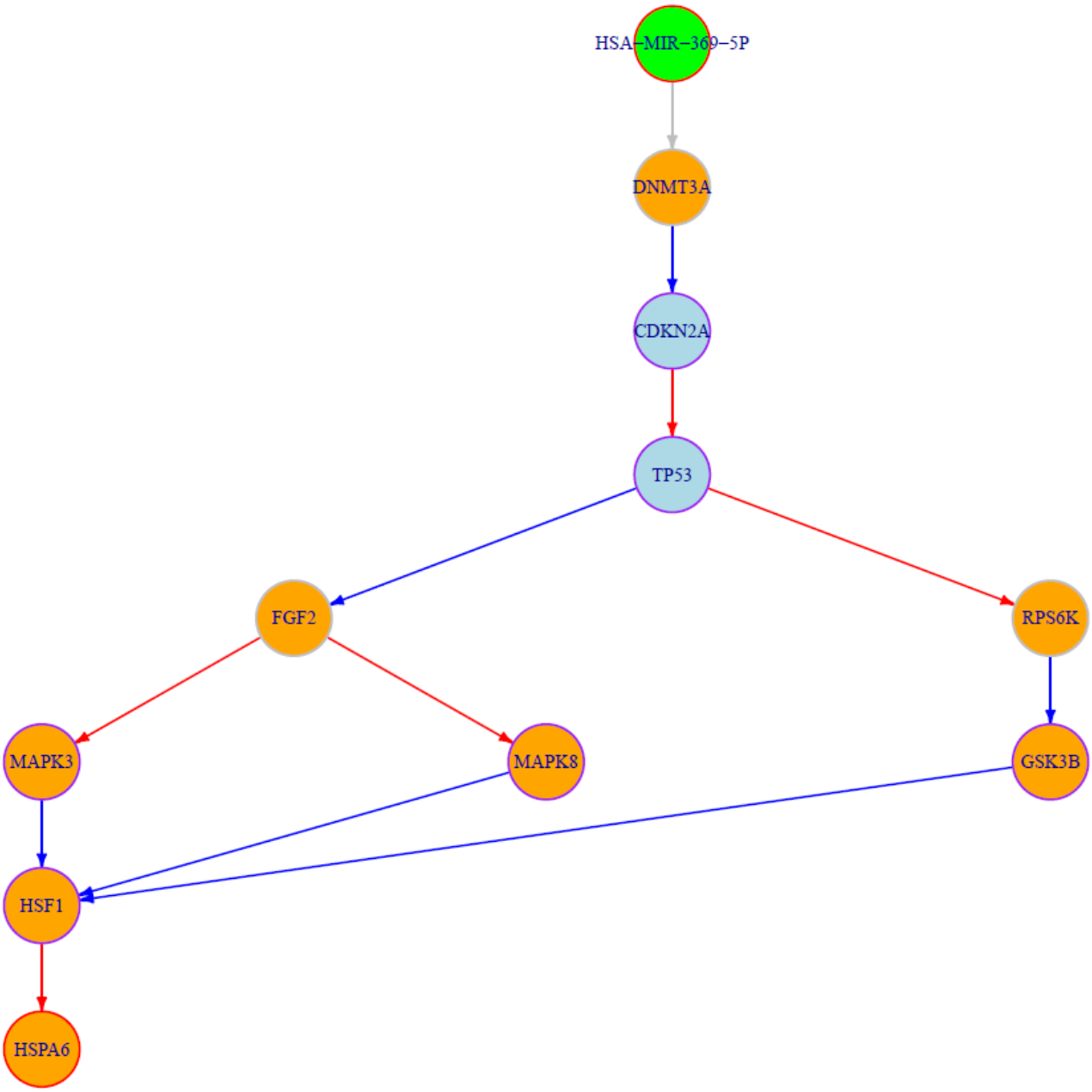


Figure A.4.2.6. Parkinson Disease Subgroup 9

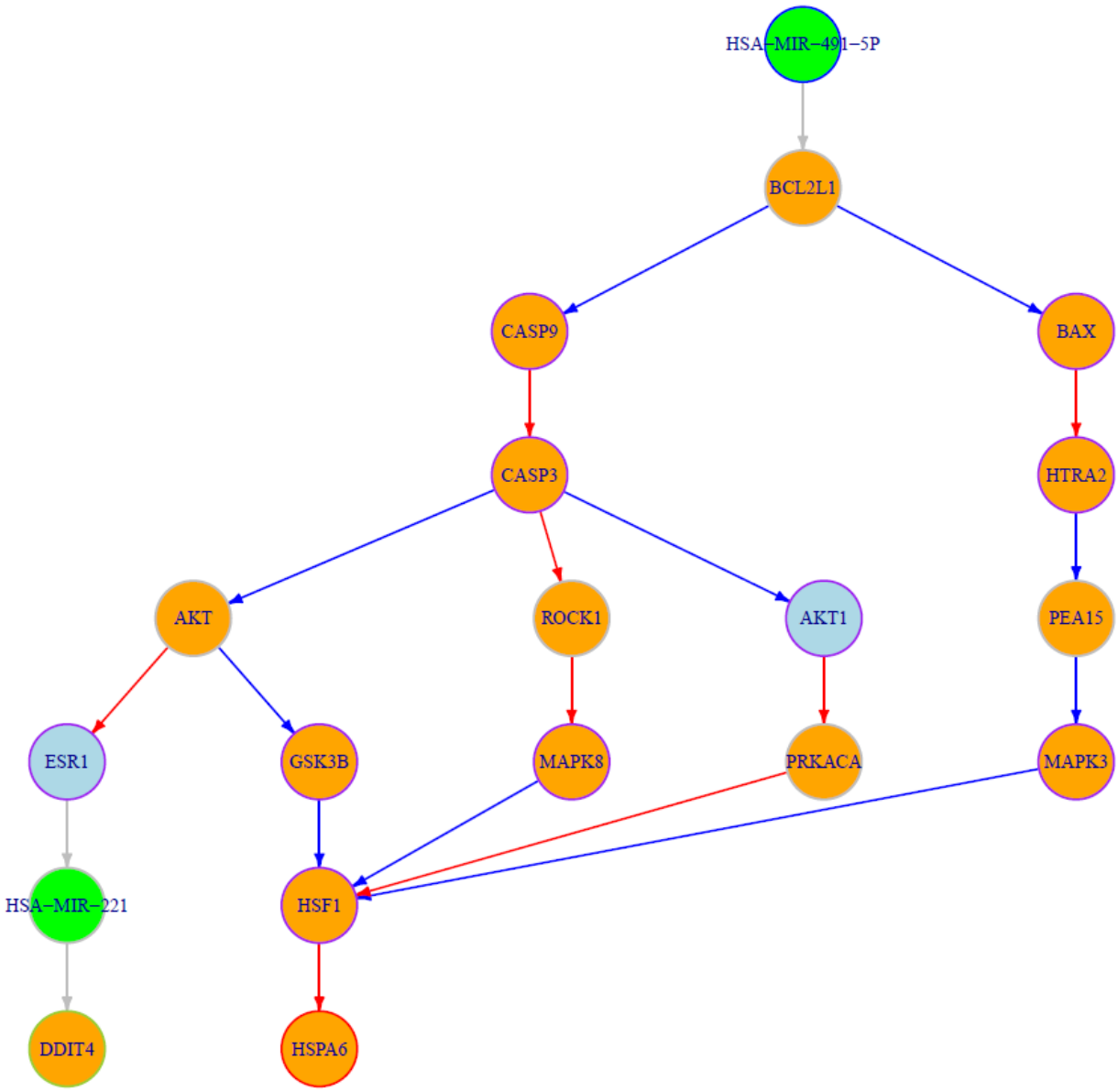


Figure A.4.2.6. Parkinson Disease Subgroup 10

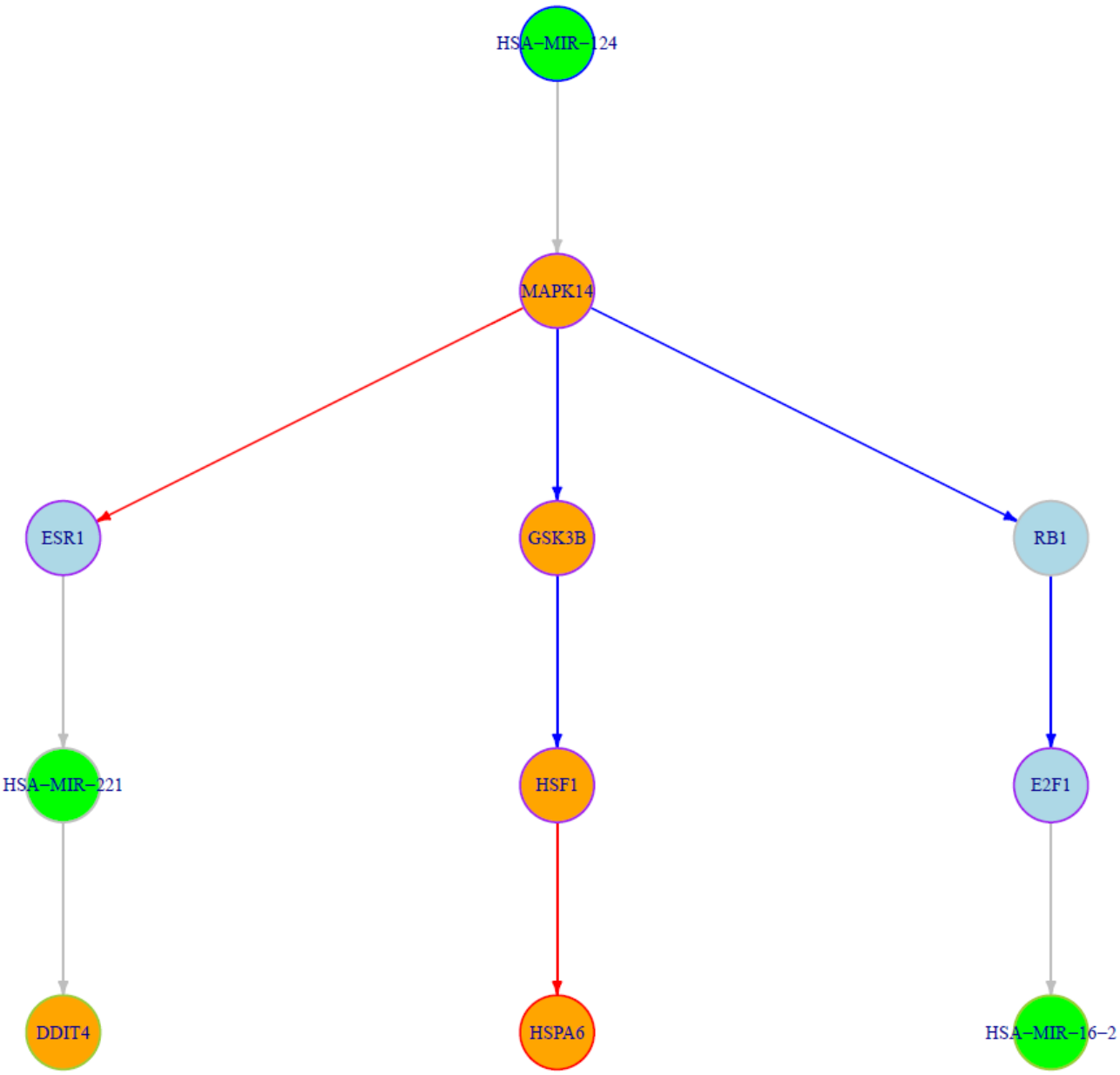


Figure A.4.2.7. Parkinson Disease Subgroup 11

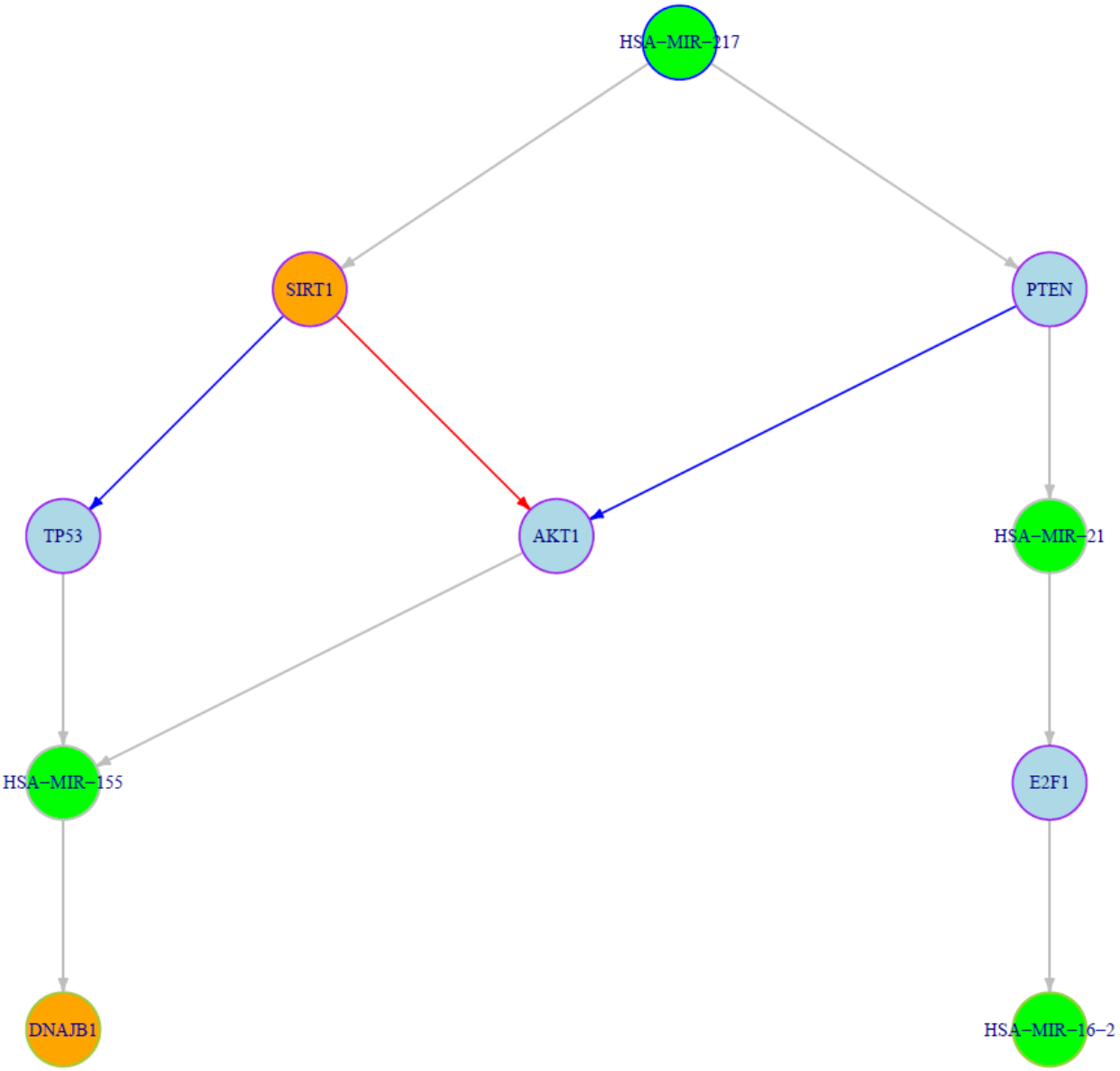


Figure A.4.2.8. Parkinson Disease Subgroup 12

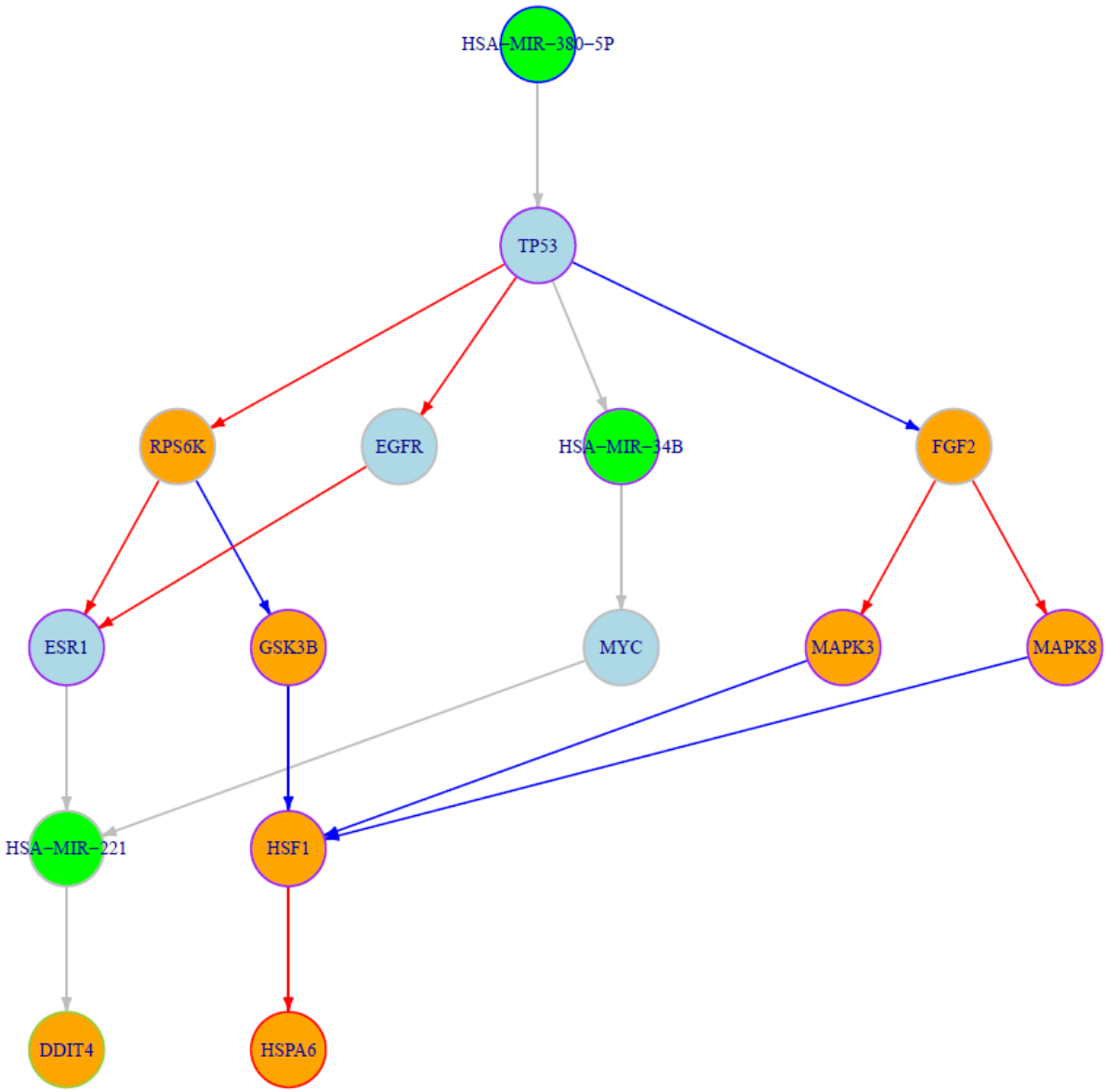


Figure A.4.2.9. Parkinson Disease Subgroup 13

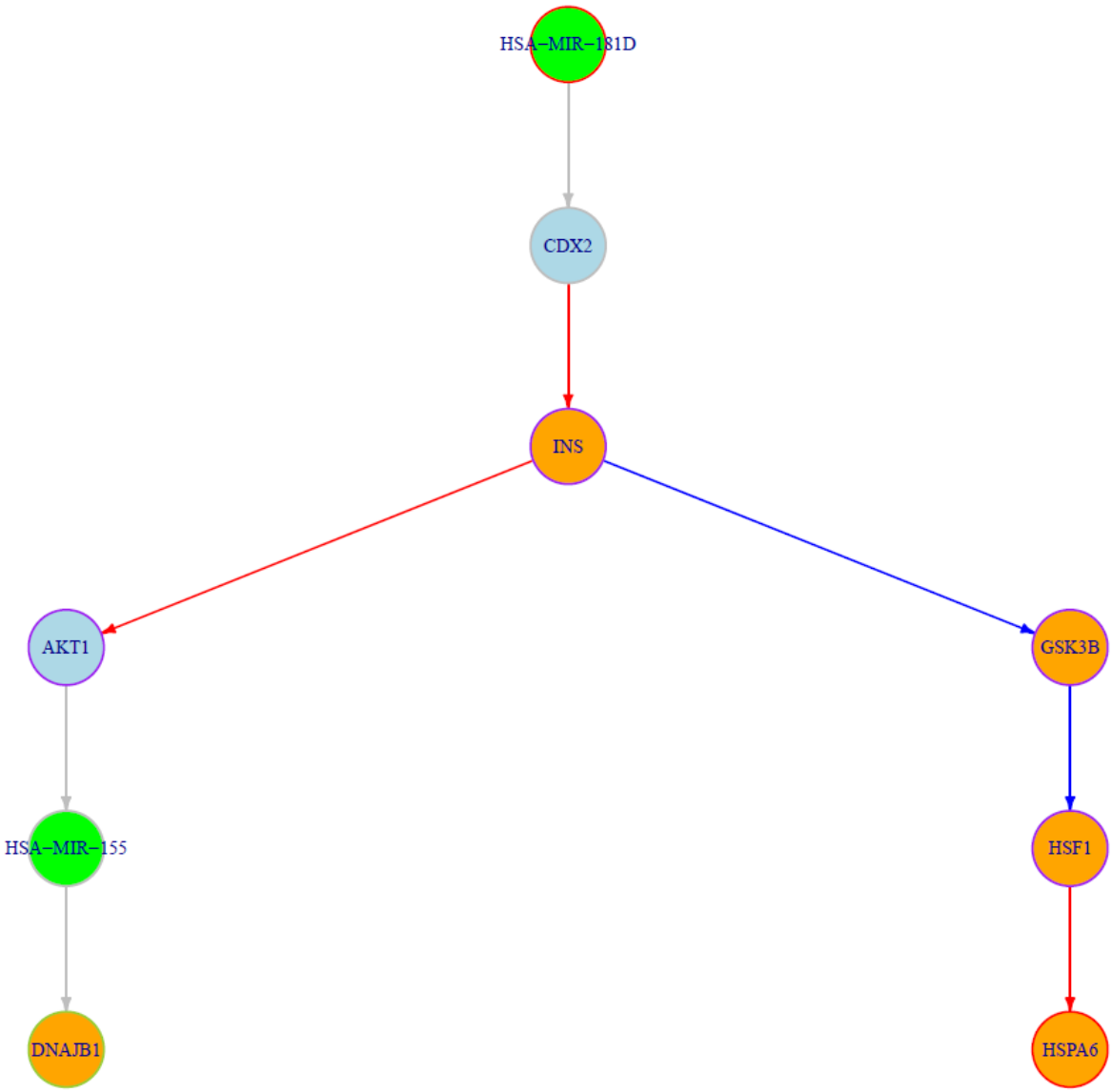


Figure A.4.2.10. Parkinson Disease Subgroup 14

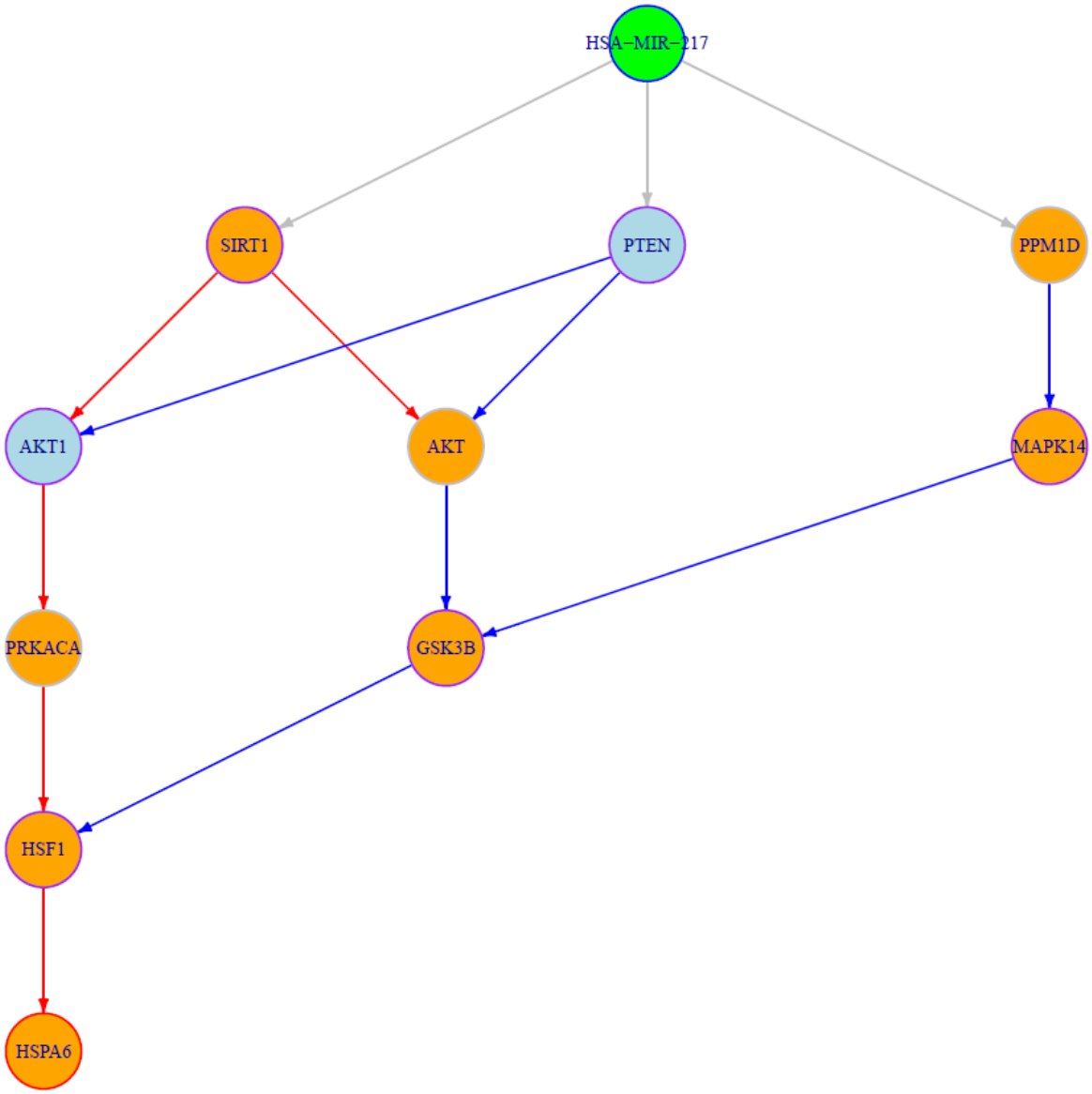


Figure A.4.2.11. Parkinson Disease Subgroup 15

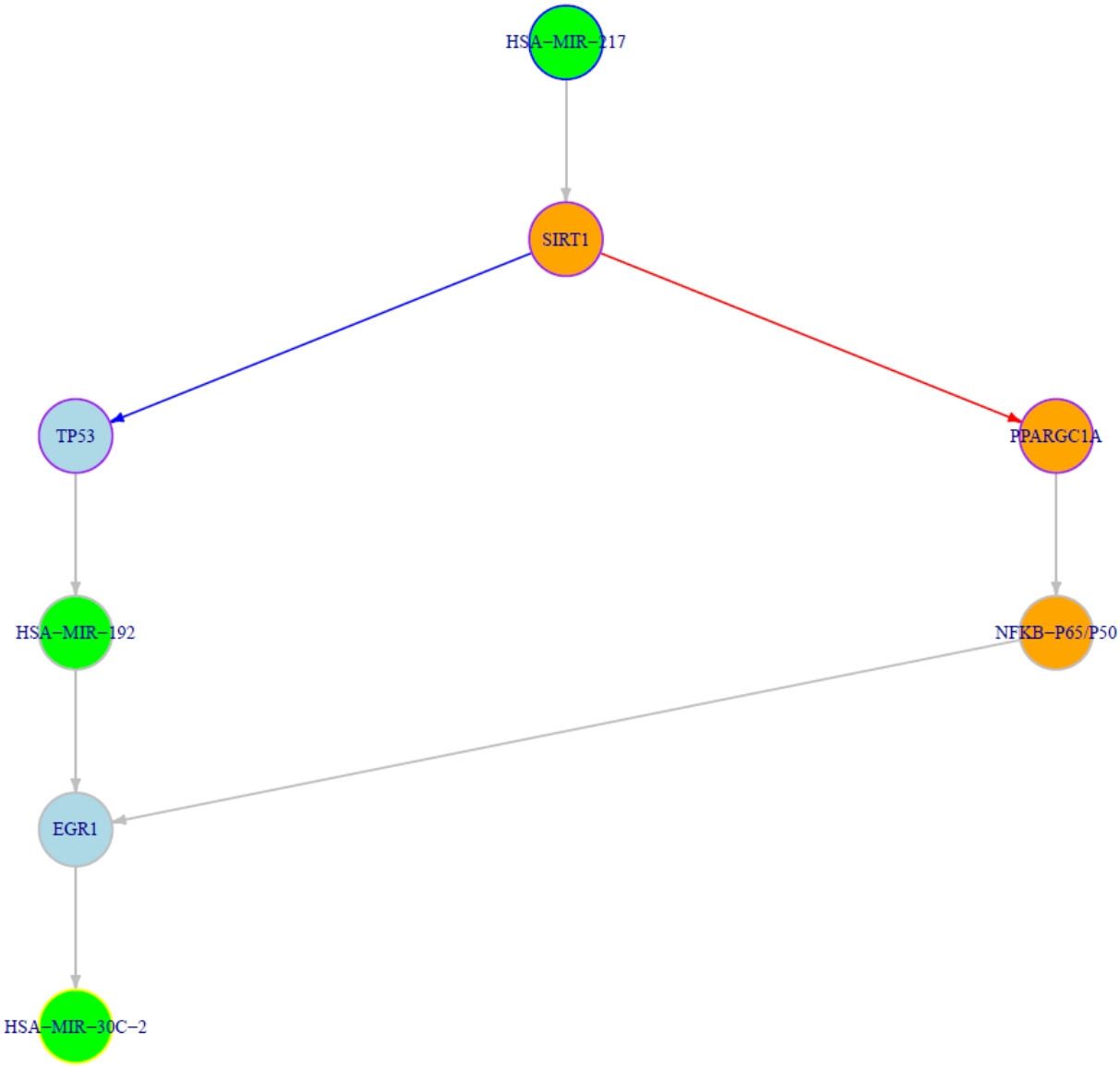


Figure A.4.2.12. Parkinson Disease Subgroup 16

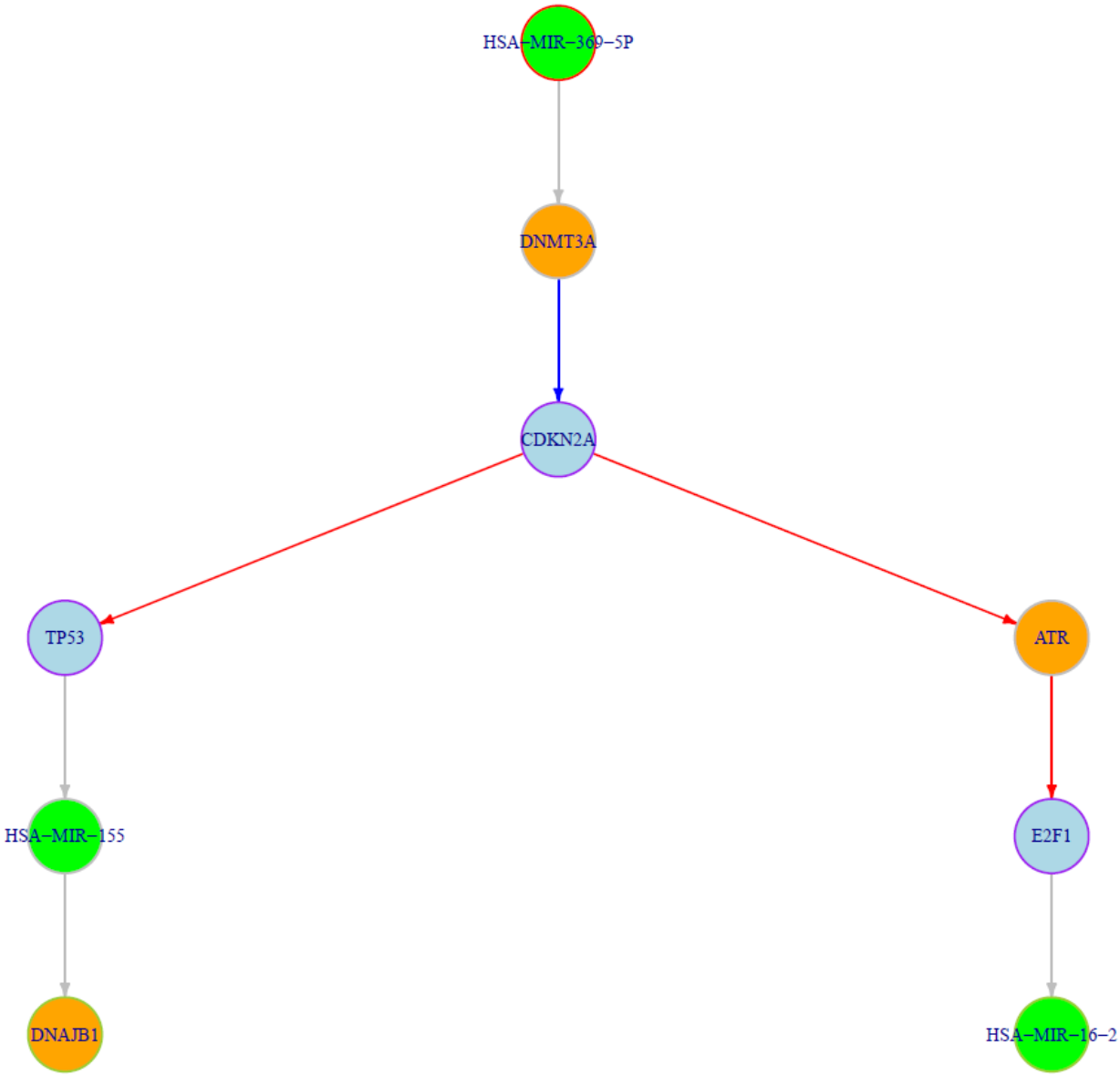
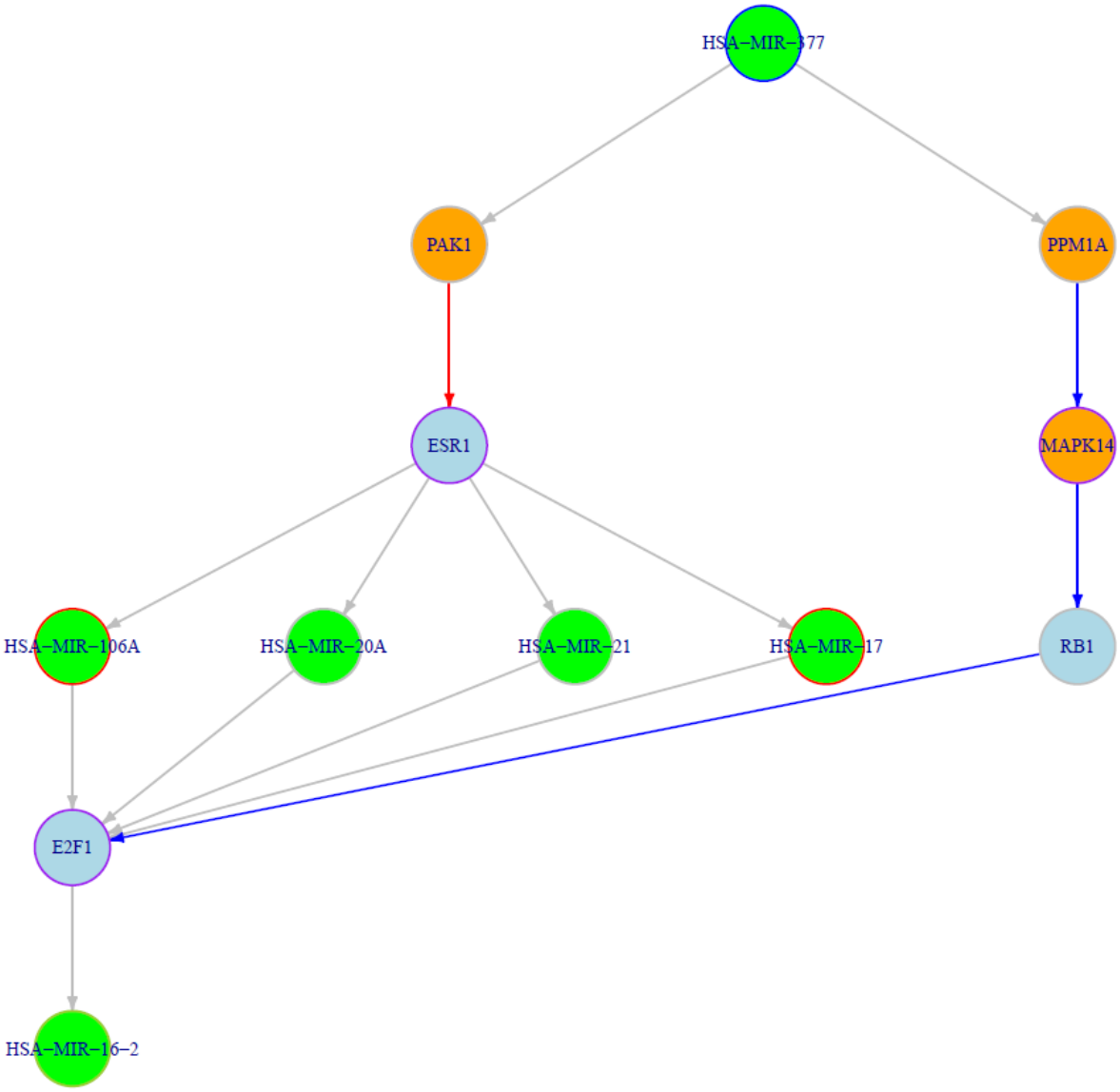


Figure A.4.2.13. Parkinson Disease Subgroup 17



CURRICULUM VITAE

1. General

Name Surname:	İ. Melis Durası Kumcu
Address:	Sabanci University Faculty of Engineering and Natural Sciences, 34956, Istanbul, TR
Date of Birth:	10.12.1989
Phone:	+905301143129
E-mail:	melisdurasi@sabanciuniv.edu

2. Education

Year	Degree	School	Department
2018 (*July*)	Ph.D.	Sabancı University	Molecular Biology, Genetics and Bioengineering
2017	B.Sc.	Acıbadem University	Faculty of Health Sciences / Nutrition and Dietetics
2011	B.Sc.	Sabancı University	Molecular Biology, Genetics and Bioengineering
2007	High School	Burak Bora Anatolian High School (1st in Class of 2007)	-

3. Publications & Conferences

- 3rd European Summer School on Nutrigenomics, Selected Presentation, "*Analysis of Expression Data Using Directed Signalling Network Identi-es Regulatory Pathways in Human Obesity"*, June 25-29, 2018, Italy.
- 11th Congress of the International Society of Nutrigenetics/Nutrigenomics, September 16-19, 2017, Los Angeles, CA, USA (participant).
- 6. Ulusal Sağlıklı Yaşam Sempozyumu ve 1. Yaşam İçin Beslenme ve Spor Kongresi, *"Besin Seçim Anketinin Türkiye'de 12-21 Yaş Grubunda Uygulanması"*, Accepted Oral Presentation, Best Project & Best Oral Presentation Reward, 2017.
- Plos One (Research Article), "*In silico* analyses and global transcriptional profiling reveal novel putative targets for Pea3 transcription factor related to its function in neurons", 2017.
- Sci Reports, "*IDH-mutant Glioma specific association of rs55705857 located at 8q24.21 involves Myc deregulation*", 2016.
- Neuroscience Letters, "*Potential of GRID2 receptor gene for preventing TNF-induced neurodegeneration in autism*", May 4, 2016.
- Toxicology and Applied Pharmacology, "*Assessment of global and gene-specific DNA methylation in rat liver and kidney in response to non-genotoxic carcinogen exposure*". December 1,2015.

- Plos One (Research Article), "*CSF Proteomics Identifies Specific and Shared Pathways for Multiple Sclerosis Clinical Subtypes*", May 5,2015.
- 67th American Academy of Neurology (AAN) Annual Meeting in Washington, DC, "*Proteomic Analysis in CSF Identified Subtype Specific and Shared Molecular Pathways for Multiple Sclerosis Clinical Phenotypes*", April 18-25, 2015.
- International Conference on Applied Informatics for Health and Life Sciences in conjunction with Turkish-German Workshop on Bioinformatics: Recent Developments from Health to Nanotechnology, Tutorial Given as a Speaker:*"Introduction to DAVID and PANOGA Bioinformatics Resources*", 2014.
- International Conference on Applied Informatics for Health and Life Sciences Turkish-German Workshop on Bioinformatics: Recent Developments from Health to Nanotechnology, "*Identification of Pathways from Proteomic Analysis with DAVID and PANOGA*", Kuşadası-TÜRKİYE, 19-22 October 2014.
- 62nd American Society Human Genetics Annual Meeting in San Francisco, "*Disease Specific Pathway Analysis in Multiple Sclerosis",* 2012
- International Symposium on Health Informatics and Bioinformatics, "*Identification of Transcription Factor Binding Sites in Promoter Databases",* 2011

4. Theses

Ph.D.	July, **2018**

5. Academic Experiences

Year	Title	University	Department/Course
2012- present	Research Assistant	Sabancı University	Molecular Biology, Genetics and Bioengineering
2012-2014	Teaching Assistant	Sabancı University	Academic Support Program
2013-2014	Teaching Assistant	Sabancı University	Natural Sciences Course
2011-2013	Teaching Assistant	Sabanci University	First Year Math Course

6. Skills

Computer Skills: R, Perl, C++, Microsoft Office Programs.

7. Personal Activities

- 2016-… Writer at "**Mutfak Magazin**"
- 2016-… Member of "**Adım Adım**", TEV
- 2014 Ayşe ARMAN, "**Hürriyet Gazetesi**", interview "Genetik Doktoralı Tescilli Bir Güzel"
- 2013 Member of "**Bilim Kahramanları**" and "**First Lego Leauge Tournament**"
- 2012 **Miss Turkey** 2nd Runner-up
- 2012 **Miss Earth 2012** (representer of Turkey) Miss Compassion Award
- 2011-2012 Sabanci University - Academic Support Program Teaching Asistant *Work & Study Scholarship*

- 2011 Representer of Turkey in "**SanFrancisco Women's Marathon**" as an ambassador of "**NIKE**"
- 2011 SABANCI RUN- "**NIKE**" Run Freeks, 6 weeks
- 2007-2011 Sabanci University Voleybol Team Member and License Holder
- 2007-2011 Sabanci University Dance Tango Team
- 2007 Sabanci University Civic Involvement Projects, 9 months
  Worked with "**Bir Dilek Tut Derneği**"(Make a Wish Foundation), to give hope, strength and happiness to children with life-threatening medical conditions
- 2005 The Elizabeth Johnson Organization English Languauge Teaching Centre certificate
- 2004 1st ranking award for the essay on the topic "Traffic", in Istanbul
- 2004 The Ardmore Group English Language Teaching Centre certificate of attendance