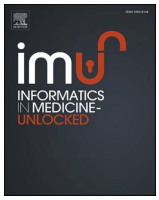
Informatics in Medicine Unlocked 37 (2023) 101174

Contents lists available at ScienceDirect

Informatics in Medicine Unlocked

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



Identification of drug and protein-protein interaction network among stress and depression: A bioinformatics approach

Md. Abul Basar [a], Md. Faruk Hosen [a], Bikash Kumar Paul [a,b,*], Md. Rakibul Hasan [b], S.M. Shamim [a], Touhid Bhuyian [c]

[a] Department of Information and Communication Technology (ICT), Mawlana Bhashani Science and Technology University (MBSTU), Santosh, Tangail, 1902, Bangladesh
[b] Department of Software Engineering (SWE), Daffodil International University (DIU), Ashulia, Dhaka, Bangladesh
[c] Department of Computer Science and Engineering (CSE), Daffodil International University (DIU), Ashulia, Dhaka, Bangladesh

ARTICLE INFO

Keywords:
Stress
Protein-protein interaction
Protein-drug interaction
Gene-miRNA interaction
TF-gene interaction
Depression

ABSTRACT

The fields of data mining, computational biology, and statistics have been combined to form the massive research area of bioinformatics. In the areas of genetics, education, and healthcare, bioinformatics integrates the tools available in different fields such as computing, inventorying, performing statistical analyses, and collecting and processing genomic data. Stress and depression are two of the most severe mental disorders that affect people of all ages, including children and adults. The goal of this study was to look into the relationship between genetic alterations and the two diseases mentioned above as well as to develop a PPI network or related channel. The first step is to determine whether or not there is a biological relationship between them. This would assist us in connecting both of them as well as building therapeutic drugs that are effective against stress and depression disorders. Using R programs, the genes that are responsible for different diseases are acquired, pre-processed, analyzed, and mined in order to better understand them. During the study, a novel pathway was discovered. Based on common genes between the two diseases studied, the PPI network, gene-miRNA interaction, TF-gene interaction, and PDI network were established. This data can help us better understand how the PPI network binds to its ligands. We anticipate that our study will contribute to the development of new drugs for stress and depression.

1. Introduction

Nowadays, people suffer from different kinds of psychological disorders, such as anxiety disorder, bipolar disorder, panic disorder, stress, depression, schizophrenia, etc. It is also referred to as a "mental" or "psychiatric" disorder [1]. Mental disorders are affecting an increasing number of people worldwide. That's why mental disorders are considered the most fatal problem globally [2]. Mental health concerns are common at all ages and have many potential triggers, including stress at home and work and financial difficulties (Mental Health Foundation, 2016) [3]. Data, uncertainty, fear, and worry seem to heighten public concern, leading to mental health difficulties such as stress and depression [4]. Lazard et al. (2015) noted that previous findings and observations with significant outbreaks of deadly diseases have revealed "panic potential that is frequently much greater than disease risk" [5]. According to the World Health Organization, behavioral and mental disorders account for approximately 12% of the worldwide infection rates. According to the estimates provided by the Mental Health Statistical Analysis in 2017, 792 million people were living with a mental health issue. This equates to slightly more than one in every ten people worldwide (10.7%) [6].

The non-specific manner in which the body responds to any kind of demand placed on it is referred to as stress disorder (SD). It is possible for it to be an issue of both positive and negative experiences [7]. When social pressure becomes severe, SD can emerge, and in certain cases, it can induce anxiety disorders, including depression. 74% of participants felt stressed while unable to complete necessary tasks, according to mental health data (https://www.mentalh ealth.org.uk). Stress causes changes in the body's physiology that lead to heart disease, high blood pressure, genetic susceptibility, growth insults, and psychosocial

* Corresponding author. Department of Information and Communication Technology (ICT), Mawlana Bhashani Science and Technology University (MBSTU), Santosh, Tangail, 1902, Bangladesh.
E-mail addresses: bikash.k.paul@ieee.org, bikash@mbstu.ac.bd (B. Kumar Paul).

https://doi.org/10.1016/j.imu.2023.101174
Received 25 September 2022; Received in revised form 25 December 2022; Accepted 14 January 2023
Available online 18 January 2023

problems, all of which are known to increase the risk of depression [8]. Depression risk can be raised by stress disorder, even though it is considered a separate factor. There is a clear link between stress and depression, as shown in research; nevertheless, many processes may be at work [9]. The American Stress Institute says that worldwide, around 33% of people report experiencing significant stress. Stress has a negative impact on the physical health of 77% of people and the mental health of 73% of people [10]. Depressed feelings were expressed by 51% of stressed individuals, while anxiety was recorded by 61%. 16% of those who indicated they had experienced stress at some point in their lives had self-harmed, and 32% had suicidal thoughts and feelings [11].

Depression is among the most prevalent and debilitating neuropsychiatric disorders, and it has emerged as a global health concern [12]. Studies indicate that approximately one-third of depressed people do not respond to traditional antidepressant treatment [13]. It is of the utmost importance that novel biomarkers for the treatment of depression be developed as soon as possible. A deeper investigation of the molecular pathways involved in the pathological process of depression may uncover novel approaches to tackle this complex disorder [14]. Depression has been linked to shrinking hippocampuses and changes in the function or interconnectivity of brain networks [15]. At its worst, depression may lead to suicide. Suicide is the leading cause of death among people aged 15 to 19, according to the World Health Organization. Approximately 4.4% of the world's population suffers from depression, which affects 350 million people globally [16]. Depression is more common in those who live in poverty than in those who are well-off, and it affects people of all ages and walks of life. Unemployment, physical disease, and drug use are all factors that contribute to depression [17].

Protein-protein interactions (PPIs) are essential in all organisms' cellular activities and biological functions. Protein interactions can contribute to a better knowledge of infection processes, as well as the creation of numerous pharmaceutical medicines and therapy optimization. PPIs are very important in regulating selectivity along signaling pathways. Thus, assessing the competitive interaction of PPIs in a cell is critical for understanding basic cellular regulation as well as designing treatments that target those diseases [18]. The networks of PPI interactions are significant tools for studying cellular processes, mechanisms for diseases, and the creation of medications. Interpreting a PPI is challenging because of the network's complexity [19].

Bioinformatics is used in a variety of applications, including data collection, data organization, data sequencing, huge dataset management, data storage, data transmission, DNA summarization, PPI networking, and treatment scheduling [20]. Bioinformatics is a hybrid science that stores biological information by means of a combination of biological and technical data. Bioinformatics plays a massive role in gaining extremely significant information on disorders that is useful in treatment development. Modern bioinformatics technologies, such as PPI network construction and drug design, provide a significant contribution to gene analysis [21].

Finding a genetic relationship and protein-drug interaction between stress and depression was one of the major objectives of this work. The new aspect of the research was the discovery of shared genes between the two diseases and the subsequent analysis of these genes [22].

Using the R programming language, we looked at a specific gene for two diseases: stress and depression disorders. created protein-protein interactions, topological characteristics, gene-miRNA interactions, protein-drug interactions, and TF-gene interactions after uncovering certain common genes.

2. Proposed methodology

The proposed technique is depicted step by step in Fig. 1.

2.1. Gene collection

The genes related to a disease must be gathered in order to study it. The NCBI (National Center for Biotechnology Information) is a valuable resource for online gene databases that are free to browse and download, as well as a large collection of bioinformatics services and tools. The genes are downloaded in ascending order based on their weight. Based on the behavior of large datasets, they retained data in many databases for analysis. Several databases, including GenBank, PubMed, and OMIM, can be used for various reasons. The NCBI gene database was used to compile a list of genes for stress and depression.

2.2. Data preprocessing

The gene database is selected in NCBI, and a list of genes associated with a certain disease is searched for. To search for genes associated with a certain disease, go to the National Center for Biotechnology

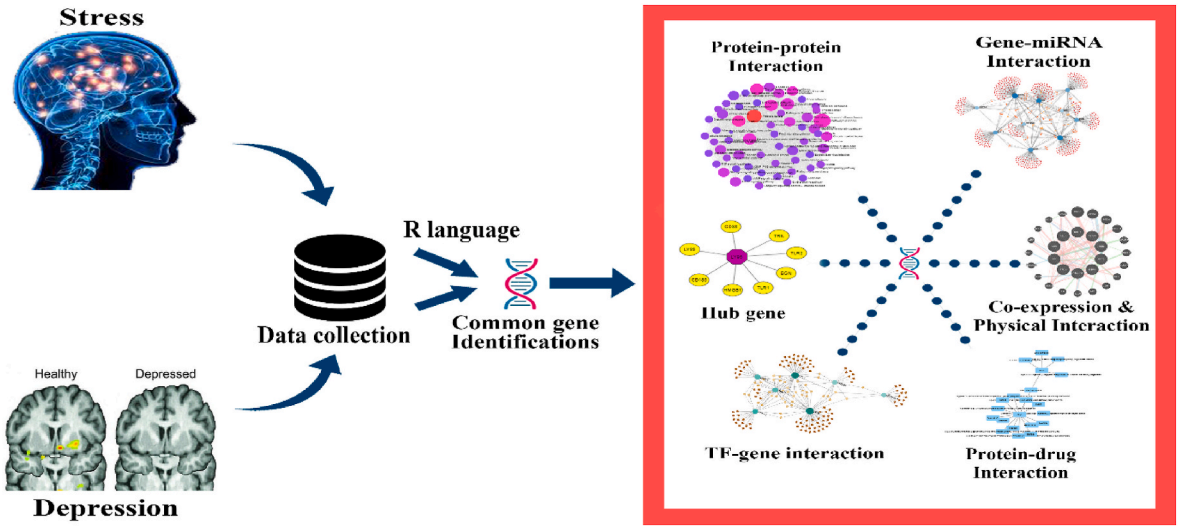


Fig. 1. Procedures are to be followed in an orderly fashion in the ongoing research work. From NCBI, collected two datasets on stress and depression. The genes related to a disease must be gathered in order to study it. The genes are downloaded in ascending order based on their weight. The programming language R was implemented in order to search both datasets for genes that were common to both. Based on the common genes, a protein-protein interaction network is created, along with its topological properties. From the PPI network, the top 10 hub genes are identified. The top ten hub genes were used to create gene-miRNA interactions, TF-gene interactions, protein-drug interactions, co-expression and physical interactions.

Information (NCBI). In increasing order, the genes are downloaded based on their weight. The database was last accessed on March 15, 2021. However, just preprocess and gather genes that are exclusive to *Homo sapiens* in this initial phase. However, in this initial phase, only *Homo sapiens* genes are preprocessed and collected. That is, all obtained genes are filtered, and only those associated with human diseases are preserved for further investigation.

2.3. Gene mining

The main objective of disease gene mining is to detect an essential gene as opposed to optimizing classification accuracy on a test set. The concordant genes are identified between those two diseases using R programming [23]. Data mining's major purpose is to make the data useable for research and analysis. Everything relevant to the genes is collected and stored in a text file, while a separate text file includes non-research-related material or genes. Gene mining is a critical component of our study, and we shall achieve our goal via gene extraction. This study is crucial since a single omission might lead a gene to be backtracked, resulting in a false result. The genes are identified and categorized according to their relationship to one another. Data mining was used to find potential genes associated with stress and depression.

2.4. Protein-protein interaction

Protein functions have a variety of characteristics, the most important of which is protein-protein interaction [24]. PPIs define the connection between two or more proteins, and these interactions are facilitated by biochemical, hydrophobic, and electrostatic factors. PPI networks provide a wealth of new information on protein function [25]. Proteins are the primary determinants of how biological processes work. Moreover, proteins regulate molecular and cellular functions, allowing us to assess the healthy situation or diseased states of organisms. PPI refers to several processes occurring inside the cell. A PPI network holds information about the protein-protein interactome of a specific organism. The Protein-Protein Interaction Network (PPI) represents interconnections between genes and proteins between genes that are related. The PPI network serves an essential role in bioinformatics research. PPI networks are created using Cytoscape, a well-known and trustworthy bioinformatics research tool [26]. For this study, a PPI network is built using NetworkAnalyst, an online bioinformatics application.

2.5. Protein-drug interaction

The core tasks of drug design are thought to be increasing drug efficiency and decreasing toxicity [27]. Finding new inhibitors is a major issue during drug development. Structure-based design is a fundamental technique for this endeavor, and it is increasingly becoming a part of drug development. A substantial number of therapeutic targets have been shown to benefit from the protein's precise three-dimensional architecture [28]. Protein-drug interactions are essential for properly understanding the structural properties required for ligand affinity. These interactions relate to the regulation of drug distribution, disposal rate, and toxicity. Protein-drug interactions are developed for all genes that are related, causative, and common in the targeted disorders. The enrichment analysis for this study was calculated utilizing the online web-based program NetworkAnalyst, a prominent bioinformatics tool.

2.6. Gene-miRNA interaction

MicroRNAs play a vital role in transcriptomic regulation. MicroRNAs are essential post transcriptional mediators of gene expression in many biological processes in animals and plants. Identifying the genes that are controlled by a miRNA is essential for a complete understanding of its biological role. MicroRNAs are naturally occurring, single-stranded, tiny RNA molecules that regulate gene expression by binding and inhibiting translation and destruction by targeting mRNAs. The paucity of experimental data identifying their associated mRNA targets contrasts with the identification and confirmation of numerous miRNA genes. The definition of the principles of miRNA target identification is the most fundamental problem in miRNA biology [29]. The accuracy and durability of the experimental approaches used to find novel miRNA targets and confirm anticipated interactions might vary substantially [30]. Functional miRNA interaction sites are defined by factors such as miRNA "seed" complementarity, structural accessibility, and sequence and positional biases. These components facilitate modulative interactions with RNA-binding proteins [31].

2.7. TF-gene interactions

TF (transferrin) is a gene that codes for proteins. Transcriptional factors (TFs) are thought to be important in the control of gene transcription as well as the determination of cellular identity and activities. A variety of target genes are regulated by individual transcription factors (TFs) through direct or indirect interactions with other transcription factors. The TF-gene interaction explores the effect of TF on functional pathways and gene expression levels. In order to understand the functions of pleiotropic global regulators, it is essential to identify the significant TF-gene interactions. So, several transcription factors work together to control life activities [32].

3. Results and discussion

3.1. Gene collection

All relevant genes for stress and depression were gathered for this study from the NCBI's online gene database. collected *Homo sapiens* genes from the NCBI. There are 3483 and 545 genes for stress and depression, respectively. Table 1 shows the number of NCBI-retrieved genes associated with various diseases as well as the number of *Homo sapiens* genes.

3.2. Gene mining, intersection and common gene finding

In this study, utilized the R programming language to discover genes that are shared by stress and depression. When it comes to data mining, the sequential pattern approach and the FP tree method are used to process the necessary information. Data mining techniques (sequential pattern approach and FP tree method) were used to find the common genes leading to stress and depression because the method of data mining can assist in the discovery or identification of common patterns or trends in transaction data for a specific time period. A sequential pattern is a type of data mining that is used to examine patterns in data that come in a particular order. Despite other algorithms such as Apriori, FP-tree is proposed for faster performance when there are a large number of patterns to select for mining [33]. It identified connections between stress and depression. Fig. 2 depicts the number of common genes. After linking the two diseases, 304 common responsible genes were found [34]. To avoid confusing the data, the top ten weighted genes were retained. The top ten weighted genes were APP, ESR1, TP53,

Table 1
The representation of the gene collection of targeted diseases is from the NCBI database. There are 3483 for stress and 545 for depression after processing and sorting the associated genes for *Homo sapiens*. The genes are sorted by their weight in ascending order. The numerical values of the identified liable genes are shown in Table 1.

Diseases	Total number of gene	Total number of *Homo sapiens* gene
Stress Dataset	105020	**3483**
Depression Dataset	**1165**	**545**

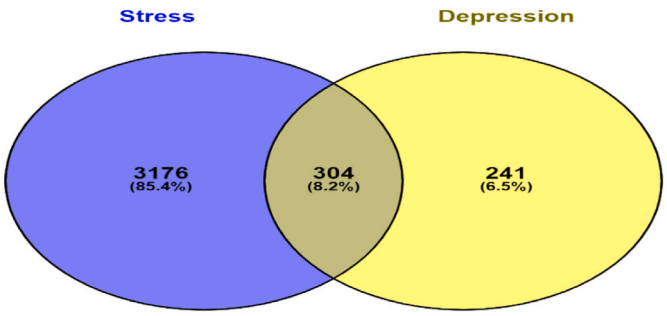


Fig. 2. A Venn diagram shows the relationship between the common genes. Out of all 4028 genes, 304 were identified as common, which is 8.2% of both stress and depression.

CTNNB1, SIRT1, AR, ABL1, HSPA4, AKT1, and CREB1. We called the top ten genes weighted. We found the top ten genes according to their degree values: 1958, 798, 659, 330, 314, 276, 275, 251, 248, 245. We found that those ten genes are mostly connected to each other. The protein-protein interaction of the top ten weighted nodes in the network has 3927 nodes and 5326 edges. The top ten genes are mostly responsible for those diseases. If we treat those genes accurately, we can prevent these diseases [35]. Fig. 2 shows a venn diagram showing the number of genes and their common gene ratio. Genes were obtained from a reliable database to begin this study. A data mining technique was then used to mine the data. Furthermore, there is a clear understanding of the relationship between two diseases. Stress and depression genes in the human genome totaled 3483 and 545, respectively. Fig. 2 depicts the outcome of the verification inquiry.

3.3. Protein-protein interaction

Protein functions have a variety of characteristics, the most important of which is protein-protein interaction. PPIs define the connection between two or more proteins, and these interactions are facilitated by biochemical, hydrophobic, and electrostatic factors. PPI networks provide a wealth of new information on protein function. The Protein-Protein Interaction Network (PPI) represents interconnections between genes and proteins between genes that are related. The PPI network serves an essential role in bioinformatics research. Researchers may perform frequent and sophisticated conceptual analyses of gene expression profiles using NetworkAnalyst 3.0, a powerful Internet tool with a natural online interface that enables them to do so easily [36]. The PPI network connects genes and hub proteins that are connected directly as well as those that are linked indirectly. The NetworkAnalyst web-based tool and the STRING interactome database are used to generate SIF files for the network diagram. The goal of the STRING database (http://string-db.org) is to provide a critical evaluation and integration of direct (physical) and indirect (functional) interactions between proteins [37]. PPI networks were created and evaluated using Cytoscape in order to identify the most important genes among the common genes. The top 10 common genes of APP, ESR1, TP53, CTNNB1, SIRT1, AR, ABL1, HSPA4, AKT1, and CREB1 were then investigated using the PPI network. Every organism for whom data on the relationship is available is included in the database [38]. Fig. 3 depicts the protein-protein interaction network for the top ten genes examined.

3.4. Topological properties

Topological properties are important to identify the key nodes in a network. The main objective of analyzing topological properties is to identify drug-target proteins and understand biological networks and mechanisms of drug activity. With the help of the Network Analyzer program, the PPI network in the Cytoscape tool is utilized to identify

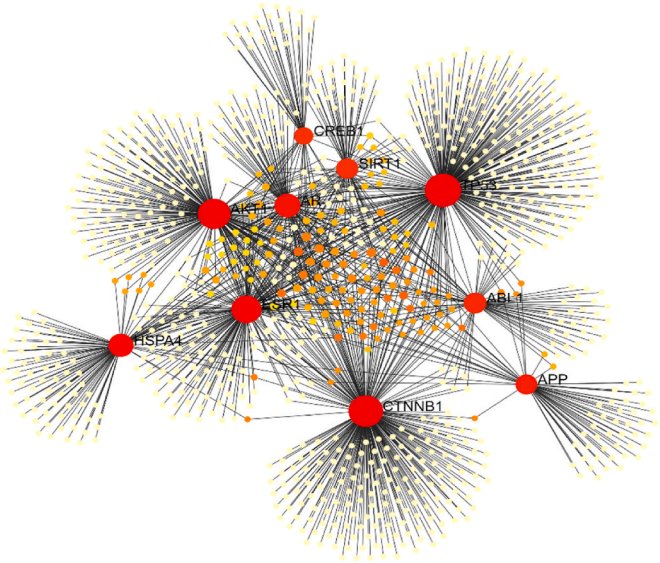


Fig. 3. This network represents protein-protein interaction for top 10 hub genes. The network has 988 nodes, 1390 edges, and 10 seeds (APP, ESR1, TP53, CTNNB1, SIRT1, AR, ABL1, HSPA4, AKT1, CREB1). Nodes in red represents hub genes, edges in yellow represents the connections that exist between the proteins. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

topological features of the network. Network Analyzer using the Cytoscape tool's Simple Interaction Format (SIF) file. Table 2 lists the top 10 weighted genes. Figs. 4–7 depicts the PPI network's various topological features. Closeness centrality, as seen in Fig. 4, is a method of detecting objects that may effectively transfer data across a network. The average distance between nodes is determined by a node's closeness centrality. With a high closeness score, nodes are closer to each other than any other nodes. The average clustering coefficient is determined by the degree of average clustering of nodes in a network, as shown in Fig. 5. Fig. 6 depicts, Betweenness centrality measures a vertex's effect on data flow between two vertices, assuming data flows largely down the shortest pathways. Fig. 7 shows, the topological coefficient is a numerical measure of how well a node links its neighbors to other nodes. Nodes with one or no neighbors are given a topological coefficient of 0.

3.5. Gene regulatory network

In order to represent the relationship between genes, the term "gene regulatory network" (GRN) is employed. Using web-based network analyst tools, we defined the gene regulatory network. Gene regulatory networks may be divided into three categories: gene-miRNA interactions, TF-gene interactions, and the TF-miRNA co-regulatory network [39].

RegNetwork TFs-miRNA coregulatory interactions to identify transcriptional and post-transcriptional regulators of genes The TF-miRNA coregulatory network was visualized using NetworkAnalyst. Gene-miRNA Small non-coding RNAs are being found to have more and more types and functions. This suggests that there are regulatory mechanisms that are much more complicated than the ones that are presently used to analyze and design gene regulatory networks. The evaluation of inter-pathway regulatory reasons can be useful during the process of finding new therapeutic targets. The activity of noncoding miRNAs is essential in this regulatory context because miRNAs are responsible for activating pathways [40]. Through the use of the miRBase database, miRNAs can be retrieved. miRBase is a database that contains published miRNA sequences as well as annotations that can be searched. Each entry in the miRBase database for a miRNA is correlated with the relevant information on the genetic location, which can then be

Table 2
Topological properties of the PPI network were investigated in order to understand the biological process, and the correlation between closeness centrality, cluster coefficients, betweenness centrality, topological coefficient, degree, number of neighbors, and other variables was discovered. Using the Cytoscape tool, we examined the topological properties of the top ten responsible genes in the PPI network.

Name of Protein	Degree	Betweenness Centrality	Closeness Centrality	Clustering Coefficient	Topological Coefficient
APP	1958	0.4382706130	0.4837275878	4.3582582208	0.0013860241
ESR1	798	0.1196689212	0.4085905186	0.0034905331	0.0039145678
TP53	659	0.1024568117	0.4173003185	0.0050735433	0.0040595482
CTNNB1	330	0.0521111492	0.4412755137	0.0126922723	0.0058149058
SIRT1	314	0.0418426250	0.3907619279	0.0085468346	0.0067302307
AR	276	0.0305328534	0.3825667302	0.0101712779	0.0082048563
ABL1	275	0.0476278674	0.4161104270	0.0069276709	0.0058116696
HSPA4	251	0.0361502317	0.4351155384	0.0167011952	0.0072242347
AKT1	248	0.0322482364	0.4332366237	0.0177615254	0.0073262968
CREB1	245	0.0453843127	0.4299296953	0.0117430579	0.0068597337

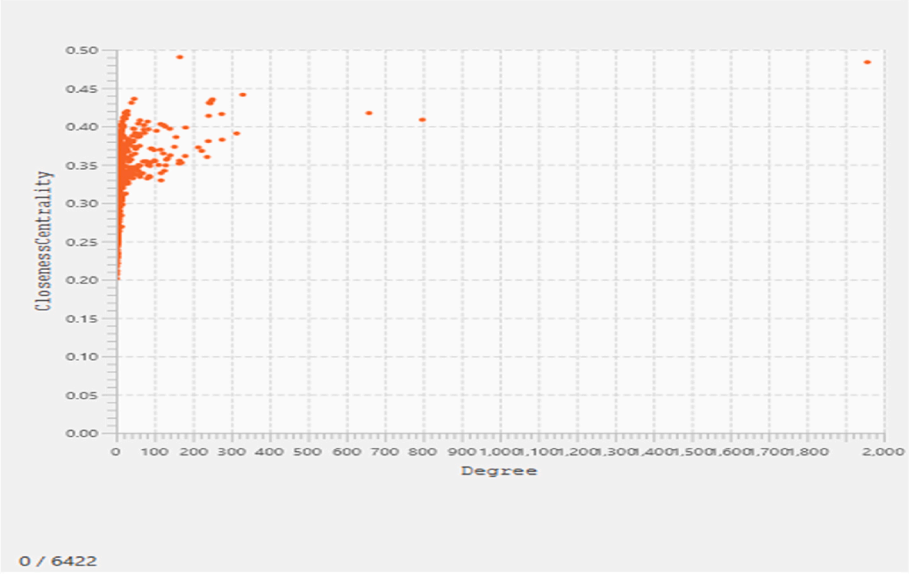


Fig. 4. Figure depicts the relationship between the degree and closeness centrality of the cytoscape-generated protein-protein interaction network.

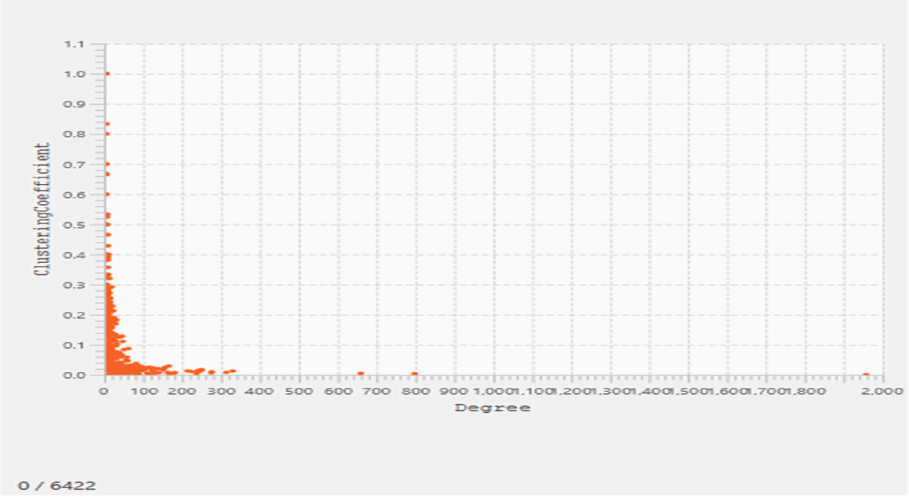


Fig. 5. Figure depicts the relationship between the degree and clustering coefficient of the cytoscape-generated protein-protein interaction network.

used to identify the host genes. NetworkAnalyst is web-based software that is used to make the gene-miRNA interaction network. NetworkAnalyst was made so that biologists, not just bioinformatics experts, could utilize it. The top ten hub genes' gene-miRNA interactions are depicted in Fig. 8.

Through TF-gene interaction analysis, the common genes are used to assess the effect of TF on the functional pathways and expression levels of the genes. Transferrin glycoproteins form a strong yet reversible bond

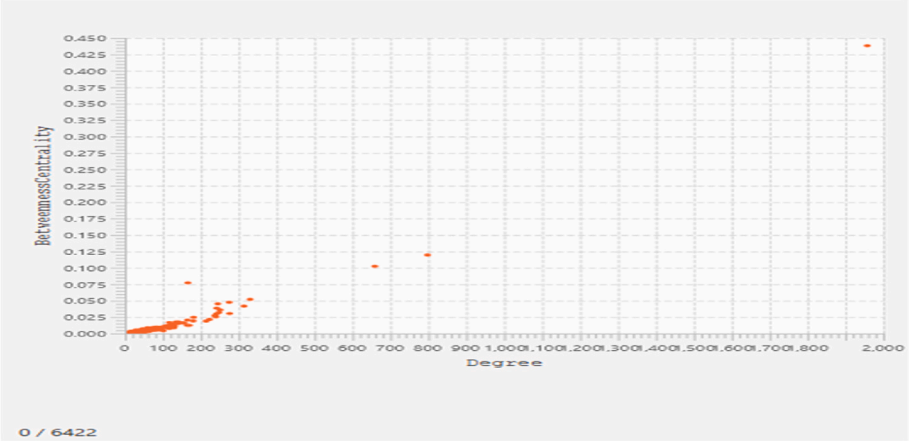


Fig. 6. Figure depicts the relationship between the degree and betweenness centrality of the cytoscape-generated protein-protein interaction network.

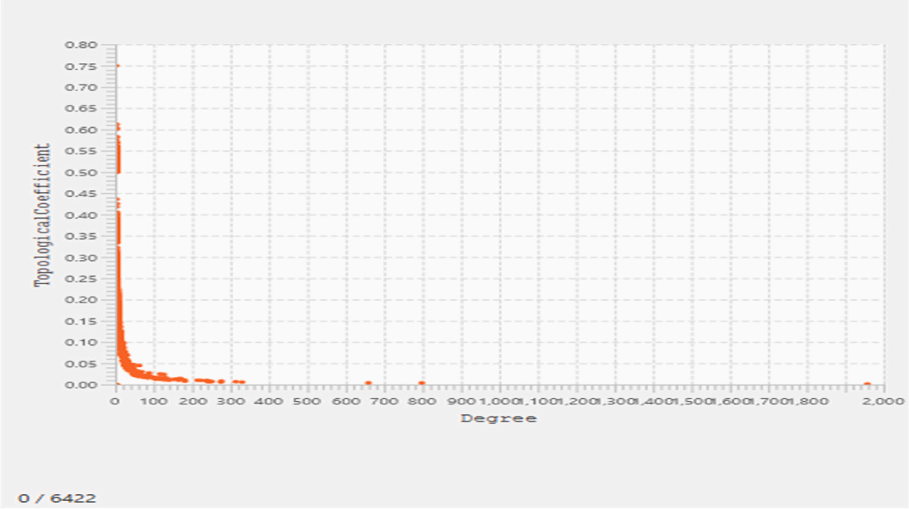


Fig. 7. Figure depicts the relationship between the degree and topological coefficient of the cytoscape-generated protein-protein interaction network.

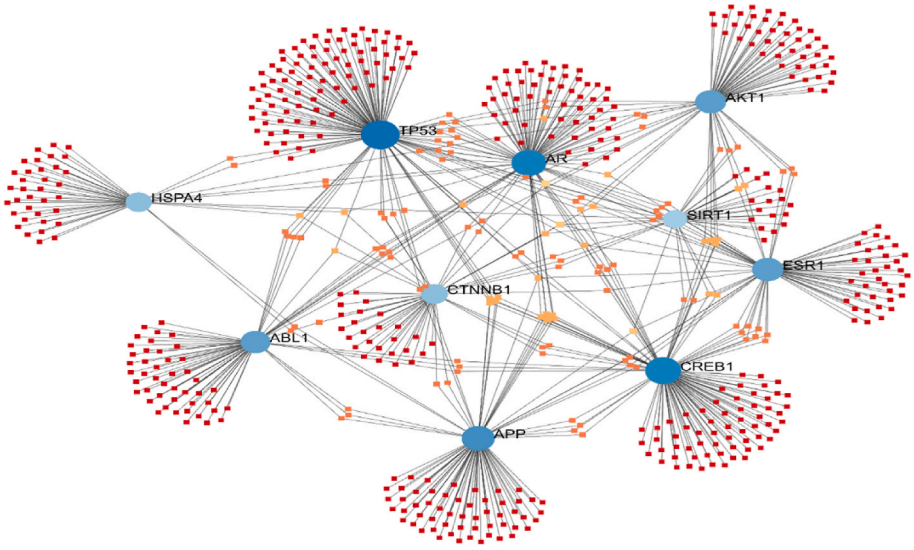


Fig. 8. Gene-miRNA interaction network was built according to the top ten hub genes. Nodes in sky blue indicated hub genes, and nodes in red indicated gene-miRNA. This network consists of 603 nodes and 754 edges. There are 10 seed nodes (APP, ESR1, TP53, CTNNB1, SIRT1, AR, ABL1, HSPA4, AKT1, CREB1). (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

with iron molecules [41]. The TF gene makes the 76-kDa glycoprotein transferrin, which is made in the liver. Vertebrate serum is composed of a variety of proteins that bind iron. Transferrin is assumed to be the most important of all these. The NetworkAnalyst platform is used to identify TF-gene interactions with commonly known genes. Analysis of biological modules can be performed with that, which also offers tasks that are typical of network topologies [42]. The ENCODE database, which is part of the NetworkAnalyst platform, is used to make the network for the TF-gene interaction network. TF-gene interactions are depicted in Fig. 9.

3.6. Protein Drug Interaction

In the cells and tissues of a live organism, activities of intermolecular recognition and/or mediation, including membrane transport events, are facilitated by protein-drug interactions [43]. During drug development, the protein target must be studied structurally and mechanistically, and this should be accompanied by a thorough understanding of how ligand binding impacts protein conformation and biological activity on several levels [44]. The DrugBank database was used to create the PDI network, which was produced using the NetworkAnalyst program [45]. The PDI network has been analyzed in this research work very carefully which has been demonstrated in Fig. 10(a) and (b) respectively.

3.7. Co-expression and physical interaction

Typically, a co-expression network is an undirected graph; this network demonstrated a strong connection between two genes. If two genes are discovered to interact in protein-protein interaction research, they are linked. These ligand-based protein networks, which anticipate the capacity of adjacent proteins to bind linked molecules indirectly, may be used to supplement genetically oriented gene networks, which predict the importance of an operation or disease. This paper examines in detail the extent to which such interactions between ligand-based proteins may allow functional genomic correlations, taking into consideration genetic overlap, physical protein-protein interactions, co-expression, and disease gene annotations [46]. Two or more proteins may interact physically, resulting in binary interactions and complex proteins [47]. Physical interaction between ligands and proteins establishes protein-protein interactions, and these ligands are often developed within protein families [48]. Certain protein interactions occur in signaling or metabolic pathways with other ligands such as nucleic acids, lipids, and certain tiny compounds. Genetic interaction refers to the functional relationship between genes. Epistasis is a non-allelic gene interaction in which one gene's function is hidden by another, resulting in repression or a new characteristic [49]. Standard pathway analysis includes software or web services that evaluate transcriptomics, proteomics with protein–protein interactions, and metabolomics data. The three most common categories of high-throughput data have been visualized and analyzed [50]. Using public data, we test the premise that interacting proteins should be co-localized. We show that a considerable number of interacting proteins are co-located using a fully filtered PPI dataset [51]. For this study, used GeneMANIA to construct networks of physical interaction and co-expression for 10 weighted genes. Fig. 11

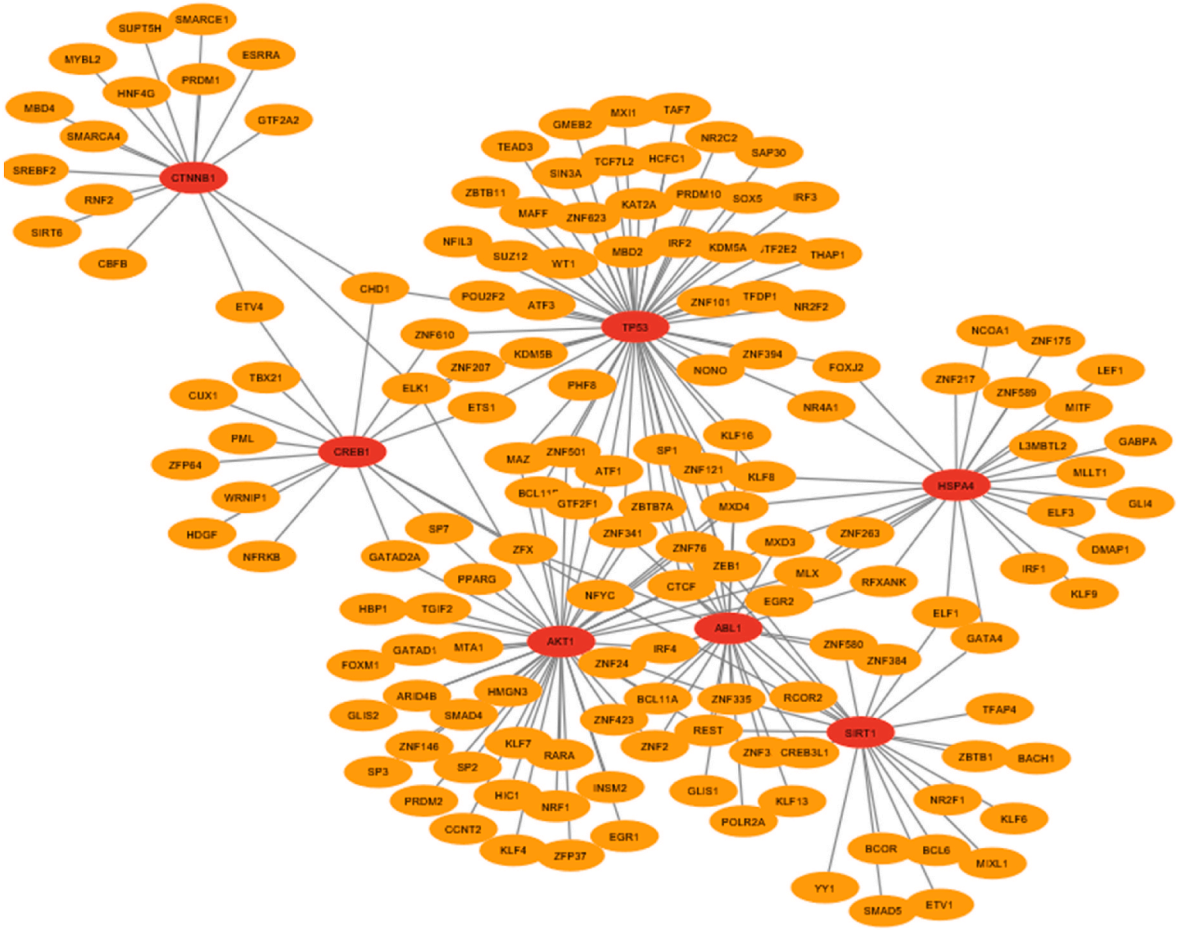
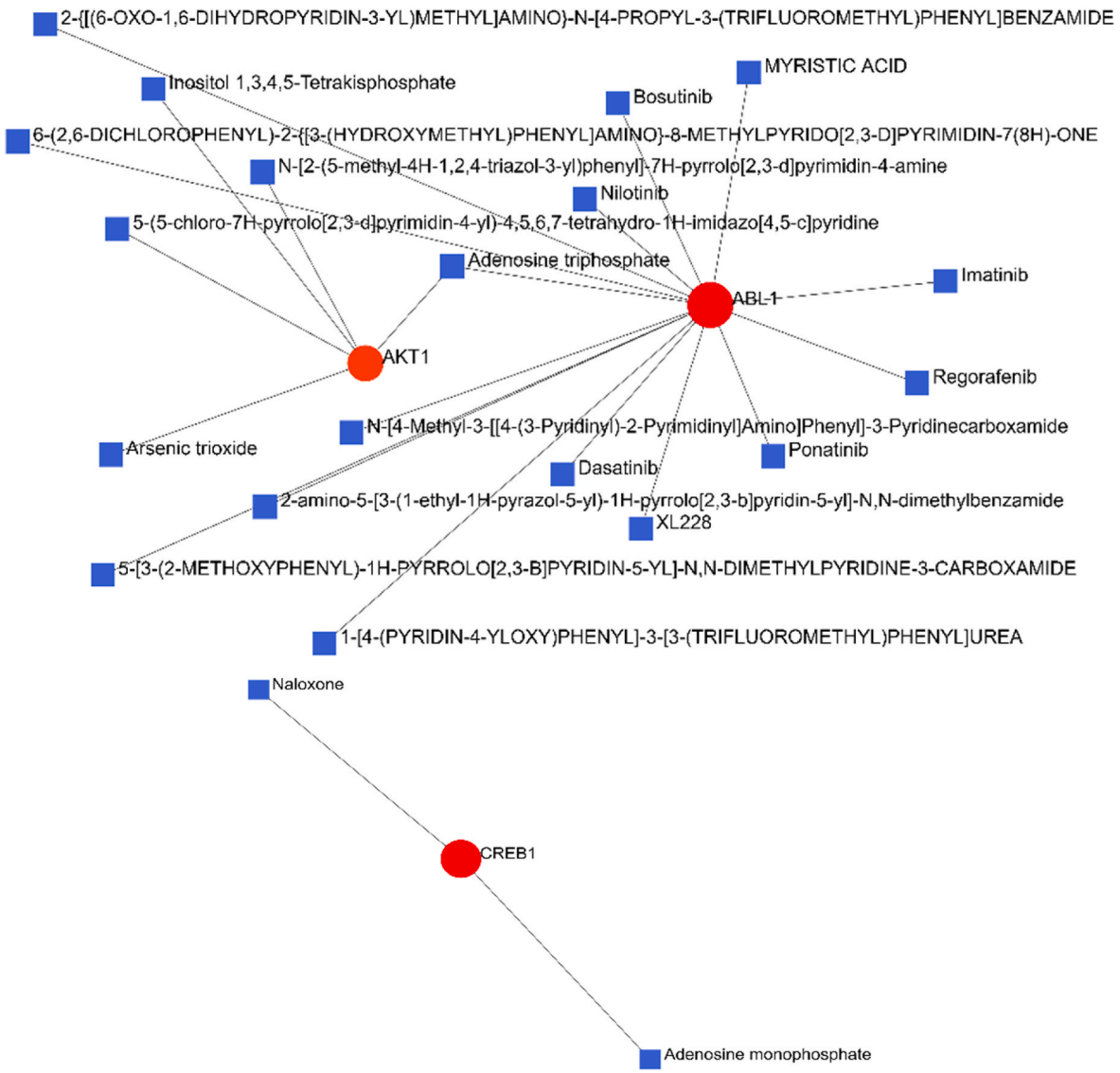


Fig. 9. TF-gene interaction network was built according to the top ten hub genes. Nodes in orange indicated hub genes, and nodes in yellow indicated TF genes. This network consists of 157 nodes and 206 edges. There are 7 seed nodes (TP53, CTNNB1, SIRT1, ABL1, HSPA4, AKT1, CREB1). (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)



(b)

Fig. 10. (a). The graphs show subnetwork 1 of the protein-drug interactions of the 10 most significant hub genes for stress and depression. Each node represents a gene, and each edge denotes a relationship between two genes. The nodes that are placed in the middle are seed nodes. There are 21 nodes in the network, 20 edges, and two seeds (ABL1 and AKT1). Based ondegree value, ABL1 is the most significant hub protein in the network. ABL1 comprises 80.5% of the overall network's edges. (b). The graphs show subnetwork 2 of the protein-drug Interactions of the 10 most significant hub genes for stress and depression Each node represents a gene, and each edge denotes a relationship between two genes. The nodes that are placed in the middle are seed nodes. The network has 3 nodes, 2 edges, and 1 (CREB1) seed node. Based on the degree of interactions it has, CREB1 is the most essential hub protein in the network.

displays the physical contact and co-expression between the two individuals' diseases. The figure indicates the physical interaction is 61%, co-localization is 11.29%, pathway is 10.49%, genetic interaction is 8.23%, co-expression is 6.58%, predicted is 1.30%, and shared protein domains are 1.12% between those two diseases. Physical interaction: 61% of us truly recognize that, according to our diseases, they have physical interaction between them.

4. Discussion

Finding a genetic relationship and protein-drug interaction between stress and depression was one of the major objectives of this research work. The new aspect of the research was the discovery of shared genes between the two diseases and the subsequent analysis of these genes. Genes were acquired from NCBI, and common genes were retrieved from these diseases as a result of this research. The quantity of genes gathered from the NCBI gene database was huge. As a result, the pre-processing, filtering, and gene mining phases had to be followed. PPI networks were created and evaluated using Cytoscape in order to identify the most important genes among the common genes. The top 10 common genes of APP, ESR1, TP53, CTNNB1, SIRT1, AR, ABL1, HSPA4, AKT1, and CREB1 were then investigated using the PPI network. By creating the top ten hub genes, different types of interactions between those two diseases were created. The topological properties of the PPI network were investigated in order to understand the biological process, and the correlation between cluster coefficients, betweenness, degree, number of neighbors, and other variables was discoverd. Topological properties are important to identify the key nodes in a network. The main objective of analyzing topological properties is to identify drug–target proteins and understand biological networks and mechanisms of drug activity. The findings of the topological properties, which focused on an understanding of co-expression and pathway analysis in a less complicated

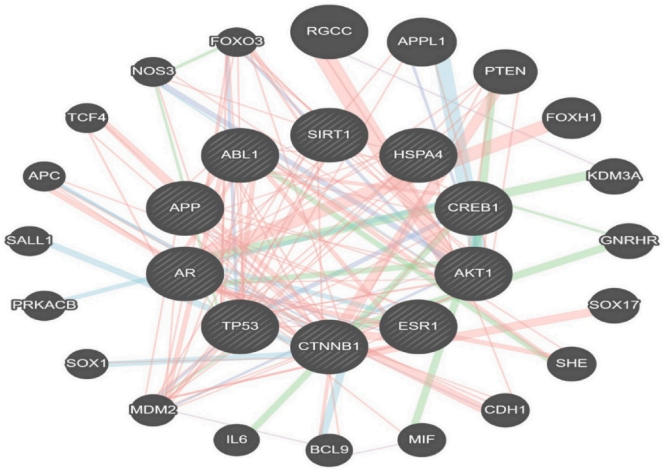


Fig. 11. Physical interaction and co-expression were investigated using the GeneMANIA software. Physical interaction is shown by the red line, co-expression is represented by the purple line, genetic interaction is represented by the green line, and pathway and co-location are represented by the blue line. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

manner, were presented. Three kinds of gene regulatory networks were employed in this study to identify the functional analysis of genomic programs. Gene-miRNA and TF-gene networks demonstrated that therapeutic molecules interacted with the genes they were targeting. The coregulatory interactions between RegNetwork TFs and miRNAs to identify transcriptional and post-transcriptional regulators of genes. The TF-miRNA coregulatory network was visualized using NetworkAnalyst. Through the use of nanosystems, a therapeutic approach that involves inserting exogenous nucleic acids into target cells in order to prevent or treat genetically based disorders can be combined with other treatment methods and imaging agents in order to produce much more precise and effective diagnostic and therapeutic approaches for stress and depression. In order to better understand how a cell's many biological and medicinal qualities interact with one another, a number of these genes were studied. Based on the study of different analyses such as PPI, topological properties, gene regulatory networks, physical interaction, and co-expression, a protein-drug interaction is recommended for stress and depression. The recommended drugs should be efficient in the therapeutic field of stress and depression.

5. Conclusion

We have given an overview of two diseases in this study: depression and stress. For drug design and development, this bioinformatics study classified collective genes from two different related diseases, which will be useful for classifying possible new drugs as a result. Four other critical characteristics have also been thoroughly studied in this work: PPI, gene-miRNA interaction, TF-gene interaction, and PDI. Modeling includes choosing the appropriate datasets, methods, variables, and formatting techniques for mining data. There must be an understanding of the genes that are affected by a disease before a drug can be recommended for therapeutic purposes. The current study used inter-gene mapping to reveal the biological relationship between the detected disorders. The protein-protein interaction network (PPI) represents interconnections between genes and proteins between genes that are related. The PPI network serves an essential role in bioinformatics research. The top 10 common genes of APP, ESR1, TP53, CTNNB1, SIRT1, AR, ABL1, HSPA4, AKT1, and CREB1 were then investigated using the PPI network. Using the top ten hub genes, different types of interactions between those two diseases were created. The topological properties of the PPI network were investigated in order to understand the biological process, and we discovered the correlation between cluster coefficients, betweenness, degree, number of neighbors, and other variables. PPI network, topological properties, PDI network, gene-miRNA interaction, TF-gene interaction, and co-expressions all help in drug discovery for the two diseases selected. Topological properties are created by the PPI network analysis program cytoscape to identify the disease-related genes that are connected to each other. Using cytoscape, find the top 10 hub genes from the PPI network that are mostly connected to each other. These top ten genes are truly responsible for these two diseases. The co-expression and physical interaction of genes were established using GeneMANIA. The PPI network, topological properties, gene-regulatory network, protein-drug interaction, and physical interaction amongst the genes that were responsible for the scheme ensured a similar therapeutic design. This study may contribute to our knowledge of the PDI, but further system biology and bioinformatics research will be required. The goal of the study in the future is to develop a generic drug for the two diseases that have been identified. It is possible that this study may offer useful information for drug development.

Funding

This work is not financially supported.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

This paper has not been previously published and is not currently being considered for publication anywhere. The authors are really thankful to those who have taken part in this research work.

Abbreviation

SD	Stroke Disorder
PPI	Protein-Protein Interaction
GRN	Gene Regulatory Network
PDI	Protein Drug Interaction
TF	Transferrin
NCBI	National Center of Biotechnology Information

References

[1] National Collaborating Centre for Mental Health (Great Britain). National Institute for health, clinical excellence (great britain). British Psychological Society and Royal College of Psychiatrists; 2011 [Common mental health disorders: identification and pathways to care].
[2] World Health Organization. Global action plan on physical activity 2018-2030: more active people for a healthier world. World Health Organization; 2019.
[3] Iseselo MK, Kajula L, Yahya-Malima KI. The psychosocial problems of families caring for relatives with mental illnesses and their coping strategies: a qualitative urban based study in Dar es Salaam, Tanzania. BMC Psychiatr 2016;16(1):1–12.
[4] Nishiura H, Kobayashi T, Miyama T, Suzuki A, Jung SM, Hayashi K, Kinoshita R, Yang Y, Yuan B, Akhmetzhanov AR, Linton NM. Estimation of the asymptomatic ratio of novel coronavirus infections (COVID-19). Int J Infect Dis 2020;94:154–5.
[5] Nor NFM, Jeffree NB, Nor HAM. Health is wealth: a corpus-driven analysis of the portrayal of mental health in Malaysian English online newspapers. GEMA Online Journal of Language Studies 2021;21(2):46–71.
[6] Bidisha M, Gouranga M. Epigenomics: the interplay of genetic and epigenetic factors behind psychological disorders. Research Journal of Biotechnology 2021; 16:2.
[7] Selye H. Stress without distress. In: Psychopathology of human adaptation. Boston, MA: Springer; 1976. p. 137–46.
[8] Dobson KS, Dozois DJ, editors. Risk factors in depression. Elsevier; 2011.
[9] Hammen C. Stress and depression. 2005 Annu Rev Clin Psychol 2005;1(1): 293–319.
[10] Arian A. Israeli public opinion on national security 1999. Jaffee center for strategic studies. Tel Aviv University; 1999.

[11] Gatt JM, Nemeroff CB, Dobson-Stone C, Paul RH, Bryant RA, Schofield PR, Gordon E, Kemp AH, Williams LM. Interactions between BDNF Val66Met polymorphism and early life stress predict brain and arousal pathways to syndromal depression and anxiety. Mol Psychiatr 2009;14(7):681–95.
[12] Etkin A, Büchel C, Gross JJ. The neural bases of emotion regulation. Nat Rev Neurosci 2015;16(11):693–700.
[13] Rush AJ, Trivedi MH, Wisniewski SR, Nierenberg AA, Stewart JW, Warden D, Niederehe G, Thase ME, Lavori PW, Lebowitz BD, McGrath PJ. Acute and longer-term outcomes in depressed outpatients requiring one or several treatment steps: a STAR* D report. Am J Psychiatr 2006;163(11):1905–17.
[14] Yang J, Chen C, Jin X, Liu L, Lin J, Kang X, Zhu S. Wfs1 and related molecules as key candidate genes in the hippocampus of depression. Front Genet 2021;11:589370.
[15] Ma L, Yue L, Zhang Y, Wang Y, Han B, Cui S, Liu FY, Wan Y, Yi M. Spontaneous pain disrupts ventral hippocampal CA1-infralimbic cortex connectivity and modulates pain progression in rats with peripheral inflammation. Cell Rep 2019;29(6):1579–93.
[16] Busch Y, Menke A. Blood-based biomarkers predicting response to antidepressants. J Neural Transm 2019;126(1):47–63.
[17] Shadhin KA, Hasan MR, Paul BK, Ahmed K, Moni MA, Kahrizi D, Bhuyian T, Ibrahim SM. Analysis of topological properties and drug discovery for bipolar disorder and associated diseases: a bioinformatics approach. Cell Mol Biol 2020;66(7):152–60.
[18] De Las Rivas J, Fontanillo C. Protein–protein interactions essentials: key concepts to building and analyzing interactome networks. PLoS Comput Biol 2010;6(6):e1000807.
[19] Vella D, Marini S, Vitali F, Di Silvestre D, Mauri G, Bellazzi R. MTGO: PPI network analysis via topological and functional module identification. Sci Rep 2018;8(1):1–13.
[20] Tchankouo-Nguetcheu S, Khun H, Pincet L, Roux P, Bahut M, Huerre M, Guette C, Choumet V. Differential protein modulation in midguts of Aedes aegypti infected with chikungunya and dengue 2 viruses. PLoS One 2010;5(10):e13149.
[21] Islam MR, Ahmed ML, Paul BK, Asaduzzaman S, Ahmed K. Common gene regulatory network for anxiety disorder using cytoscape: detection and analysis. In: International work-conference *on Bioinformatics and biomedical engineering*. Cham: Springer; 2019, May. p. 209–18.
[22] Khoobi A, Shahdost-fard F, Arbabi M, Akbari M, Mirzaei H, Nejati M, Lotfinia M, Sobhani-Nasab A, Banafshe HR. Sonochemical synthesis of ErVO4/MnWO4 heterostructures: application as a novel nanostructured surface for electrochemical determination of tyrosine in biological samples. Polyhedron 2020;177:114302.
[23] Ghoreishi SM, Behpour M, Khoobi A, Masoum S. Application of experimental design for quantification and voltammetric studies of sulfapyridine based on a nanostructure electrochemical sensor. Arab J Chem 2017;10:S3156–66.
[24] Wang X, Ren X, Li B, Yue J, Liang L. Applying modularity analysis of PPI networks to sequenced organisms. Virulence 2012;3(5):459–63.
[25] Ahmed ML, Islam MR, Paul BK, Ahmed K, Bhuyian T. Computational modeling and analysis of gene regulatory interaction network for metabolic disorder: a bioinformatics approach. Biointerface Res Appl Chem 2020;10(4):5910–7.
[26] Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res 2003;13(11):2498–504.
[27] Khatri P, Sirota M, Butte AJ. Ten years of pathway analysis: current approaches and outstanding challenges. PLoS Comput Biol 2012;8(2):e1002375.
[28] Canduri F, de Azevedo J, Walter F. Protein crystallography in drug discovery. Curr Drug Targets 2008;9(12):1048–53.
[29] Kuhn DE, Martin MM, Feldman DS, Terry Jr AV, Nuovo GJ, Elton TS. Experimental validation of miRNA targets. Methods 2008;44(1):47–54.
[30] Karagkouni D, Paraskevopoulou MD, Chatzopoulos S, Vlachos IS, Tastsoglou S, Kanellos I, Papadimitriou D, Kavakiotis I, Maniou S, Skoufos G, Vergoulis T. DIANA-TarBase v8: a decade-long collection of experimentally supported miRNA–gene interactions. Nucleic Acids Res 2018;46(D1):D239–45.
[31] Hausser J, Zavolan M. Identification and consequences of miRNA–target interactions—beyond repression of gene expression. Nat Rev Genet 2014;15(9):599–612.
[32] Ye Y, Gao L, Zhang S. Integrative analysis of transcription factor combinatorial interactions using a bayesian tensor factorization approach. Front Genet 2017;8:140.
[33] Han J, Pei J, Yin Y, Mao R. Mining frequent patterns without candidate generation: a frequent-pattern tree approach. Data Min Knowl Discov 2004;8(1):53–87.
[34] Habib N, Paul BK, Islam MS, Ahmed K. Design regulatory interaction network for anxiety disorders using R: a bioinformatics approach. Beni-Suef Univ. J. Basic. Appl. Sci. 2018;7(3):326–35.
[35] Nazari F, Ghoreishi SM, Khoobi A. Bio-based Fe3O4/chitosan nanocomposite sensor for response surface methodology and sensitive determination of gallic acid. Int J Biol Macromol 2020;160:456–69.
[36] Xia J, Gill EE, Hancock RE. NetworkAnalyst for statistical, visual and network-based meta-analysis of gene expression data. Nat Protoc 2015;10(6):823–44.
[37] Szklarczyk D, Franceschini A, Wyder S, Forslund K, Heller D, Huerta-Cepas J, Simonovic M, Roth A, Santos A, Tsafou KP, Kuhn M. STRING v10: protein–protein interaction networks, integrated over the tree of life. Nucleic Acids Res 2015;43(D1):D447–52.
[38] Orchard S, Kerrien S, Abbani S, Aranda B, Bhate J, Bidwell S, Bridge A, Briganti L, Brinkman FS, Cesareni G, Chatr-Aryamontri A. Protein interaction data curation: the International Molecular Exchange (IMEx) consortium. Nat Methods 2012;9(4):345–50.
[39] Szklarczyk D, Santos A, Von Mering C, Jensen LJ, Bork P, Kuhn M. STITCH 5: augmenting protein–chemical interaction networks with tissue and affinity data. Nucleic Acids Res 2016;44(D1):D380–4.
[40] Khoobi A, Salavati-Niasari M. High performance of electrocatalytic oxidation in direct glucose fuel cell using molybdate nanostructures synthesized by microwave-assisted method. Energy 2019;178:50–6.
[41] Ghoreishi SM, Behpour M, Mortazavi M, Khoobi A. Fabrication of a graphene oxide nano-sheet modified electrode for determination of dopamine in the presence of tyrosine: a multivariate optimization strategy. J Mol Liq 2016;215:31–8.
[42] Behpour M, Attaran AM, Sadiany MM, Khoobi A. Adsorption effect of a cationic surfactant at carbon paste electrode as a sensitive sensor for study and detection of folic acid. Measurement 2016;77:257–64.
[43] Hasan MR, Paul BK, Ahmed K, Mahmud S, Dutta M, Hosen MS, Hassan MM, Bhuyian T. Computational analysis of network model-based relationship of mental disorder with depression. Biointerface Res Appl Chem 2020;10(5):6293–305.
[44] Eyers CE, Vonderach M, Ferries S, Jeacock K, Eyers PA. Understanding protein–drug interactions using ion mobility–mass spectrometry. Curr Opin Chem Biol 2018;42:167–76.
[45] Wishart DS, Feunang YD, Guo AC, Lo EJ, Marcu A, Grant JR, Sajed T, Johnson D, Li C, Sayeeda Z, Assempour N. DrugBank 5.0: a major update to the DrugBank database for 2018. Nucleic Acids Res 2018;46(D1):D1074–82.
[46] O'Meara MJ, Ballouz S, Shoichet BK, Gillis J. Ligand similarity complements sequence, physical interaction, and co-expression for gene function prediction. PLoS One 2016;11(7):e0160098.
[47] Frishman D, Valencia A. Modern genome annotation. The Bio-Sapiens Network; 2009. p. 213–38.
[48] Finn RD, Mistry J, Tate J, Coggill P, Heger A, Pollington JE, Gavin OL, Gunasekaran P, Ceric G, Forslund K, Holm L. The Pfam protein families database. Nucleic Acids Res 2010;38(suppl_1):D211–22.
[49] Boucher B, Jenna S. Genetic interaction networks: better understand to better predict. Front Genet 2013;4:290.
[50] Cirillo E, Parnell LD, Evelo CT. A review of pathway-based analysis tools that visualize genetic variants. Front Genet 2017;8:174.
[51] Shin CJ, Wong S, Davis MJ, Ragan MA. Protein-protein interaction as a predictor of subcellular location. BMC Syst Biol 2009;3(1):1–20.