A System Biology and Bioinformatics approach to determine the molecular signature, core ontologies, functional pathways, drug compounds in between Stress and Type 2 Diabetes

Chapter · July 2023

DOI: 10.1007/978-3-031-34953-9

CITATIONS
2

READS
107

5 authors, including:

Md. Abul Basar
Uttara University
6 PUBLICATIONS **15** CITATIONS
SEE PROFILE

Md. Rakibul Hasan
Daffodil International University
7 PUBLICATIONS **47** CITATIONS
SEE PROFILE

Bikash Kumar
Indira Gandhi National Open University (IGNOU)
10 PUBLICATIONS **16** CITATIONS
SEE PROFILE

Khairul alam Shadhin
Daffodil International University
4 PUBLICATIONS **6** CITATIONS
SEE PROFILE

A System Biology and Bioinformatics approach to determine the molecular signature, core ontologies, functional pathways, drug compounds in between Stress and Type 2 Diabetes

Md. Abul Basar[1], Md. Rakibul Hasan[2,*], Bikash Kumar Paul[1], Khairul Alam Shadhin[2], and Md. Sarwar Mollah[3]

[1] Department of Information and Communication Technology (ICT), Mawlana Bhashani Science and Technology University (MBSTU), Santosh, Tangail, 1902, Bangladesh

[2] Department of Software Engineering (SWE), Daffodil International University (DIU), Ashulia, Dhaka, Bangladesh

[3] Department of Computing and Information System (CIS), Daffodil International University (DIU), Ashulia, Dhaka, Bangladesh

hasan35-148@diu.edu.bd

Abstract. Bioinformatics is the application of computer science and information technology to the field of biology and medicine. It involves the analysis of large amounts of biological data, such as DNA sequences, protein structures, and gene expression patterns. Bioinformatics is used to develop new methods for understanding and analyzing biological data, as well as to develop new tools and technologies for biological research. Bioinformatics is used in a variety of fields, including genomics, proteomics, and drug discovery. In this study, focus on two severe diseases which affect millions of people globally such as stress and type 2 diabetes. Stress can have a significant impact on people with type 2 diabetes. Stress can cause blood sugar levels to rise, making it difficult to manage diabetes. The purpose of this research is to use various bioinformatics methods to discover potential therapeutic drugs and functional pathways between stress and type 2 diabetes. The microarray datasets GSE183648 and GSE20966 are used for the analysis of stress and type 2 diabetes samples respectively. After the datasets have been preprocessed and filtered through the use of the R programming language, identified the common DEGs. The depiction of common DEGs is shown by venn diagram. Next, the most active genes are identified through topological properties, and PPIs are built from the similar differential expressed genes (DEGs). These five genes NTRK2, SOCS3, NEDD9, MAP3K8, and SIRPA are the most important hub genes with in the interaction network of protein-protein. According to the common DEGs, GO terms molecular function (MF), KEGG and WikiPathways are shown in this study. Gene-miRNA interaction, TF-gene regulatory network, module analysis, GO terms (Biological Process, Cellular Component), Pathways (Reactome, BioCarta, BioPlanet) are all things that could be done with this research work in the future. In last, a therapeutic drug compounds are recommended on the basis of common DEGs.

Keywords: Stress, · Microarray, · System Biology, · Transcriptomic, · Differentially expressed genes, · Protein-protein interactions, · Gene Ontology, · TF-gene interaction, · Hub gene, · drug compounds, · Type 2 Diabetes.

1 Introduction

In today's world, stress has become one of the most common as well as dangerous diseases [1]. Unchecked and uncontrolled stress has been linked to a wide range of negative health outcomes that put individuals' mental, physical, and social health at risk. The body adapts to any kind of demand is to engage its natural defense system, which is known as stress. It's really essential that both negative and positive situations to contribute to the issue [2]. When things get much more challenging socially, it can be tough for some people, and this can cause stress. According to the World Health Organization, an estimated 3.4 billion people worldwide are affected by stress [3]. In the United States, approximately 8 in 10 adults report feeling stressed at least occasionally. Additionally, 1 in 5 adults report feeling extreme levels of stress [4]. Stress is a known risk factor for type 2 diabetes [5]. When you experience stress, your body releases hormones such as adrenaline and cortisol, which can cause physical reactions such as increased heart rate, sweating, and muscle tension. These hormones can also affect your blood sugar levels, making it more difficult for your body to regulate them. This can lead to an increased risk of developing type 2 diabetes. Additionally, stress can lead to unhealthy behaviors such as overeating, which can also increase your risk of developing type 2 diabetes [6]. Type 2 diabetes is a chronic condition in which the body does not produce enough insulin or does not use insulin effectively. This can lead to high blood sugar levels, which can cause a variety of health problems [7]. Common symptoms of type 2 diabetes include increased thirst, frequent urination, fatigue, blurred vision, and slow healing of cuts and bruises. Treatment for type 2 diabetes typically includes lifestyle changes such as eating a healthy diet and exercising regularly, as well as medications to help control blood sugar levels [8]. According to the World Health Organization, approximately 422 million people worldwide were living with diabetes in 2014. Of those, 90-95% had type 2 diabetes. In the United States, approximately 30.3 million people, or 9.4% of the population, had diabetes in 2015. Of those, approximately 23.1 million had type 2 diabetes [9]. High throughput methods have been drastically enhanced by the study of microarray data along with the data provided by that expression dataset [10]. Following the screening of genes taken from datasets GSE183648 and GSE20966, a total of 47 differentially expressed genes (DEGs) that were shared by both datasets were identified. For this study, the protein-protein interaction (PPI) is the most significant thing, so we will proceed by analyzing that. Through a degree-topological analysis, the PPIs network found hub genes and ranked the 10 most important ones. Various bioinformatics investigations have been carried out to merge DEGs with shared features and to use DEGs to pinpoint specific drug compounds. Analysis of frequently

occurring DEGs may also help identify Gene Ontology (GO) terms and other functional pathways. Therapeutic studies including the recommended drugs for stress and type 2 diabetes should yield positive results. Microarray data contain biological information that may be gleaned via computer investigation, which is useful for the work of biological researchers. The overarching goal of this study is to use gene-based research to uncover the biological connection between stress and type 2 diabetes, and to then use that information to identify promising biomarkers. Differentially expressed genes must be discovered before the genes underlying stress and type 2 diabetes may be identified. The relationship between the DEGs is then visualized using a PPIs network. The next step is to analyze the biological system's pathways using the KEGG database and then draw conclusions about the system's overall functionality. We propose therapeutic compounds targeting the common DEGs in stress and type 2 diabetes once hub genes have been identified as shared among both conditions. The procedure that was followed for this research can be seen in figure 1.

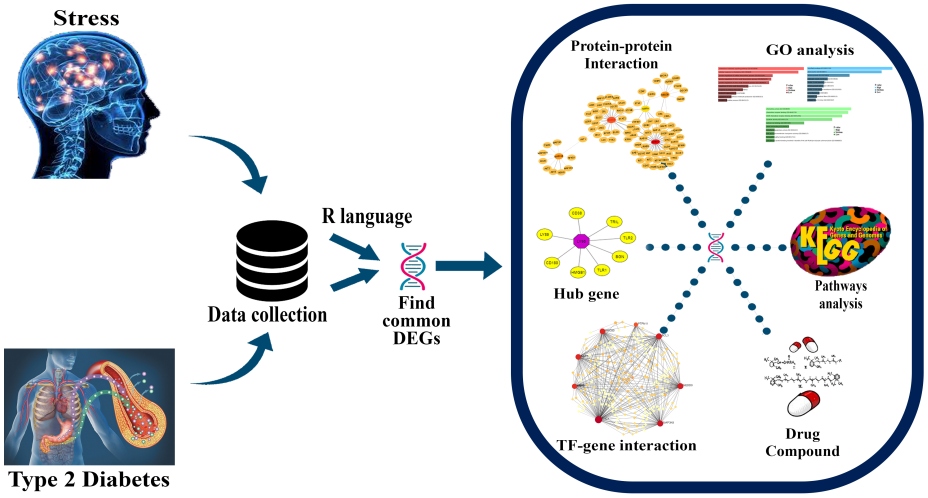

Fig. 1. Analysis of data and research procedure shows the flow chart. We combed through the samples (normal and infected) in the GSE183648 and GSE20966 databases to find what we required. Loss of p53 prevents replicating strain Damage to dna in ATRX-deficient neuroblastoma, as shown by the microarray dataset GSE183648. Conversely, GSE20966 is a microarray dataset that was obtained by collecting beta-cell enriched tissue from persons with type 2 diabetes and then analyzing their gene expression levels by Laser Capture Microdissection. We found the most prevalent DEGs by using R to analyze the two data sets. PPIs network TF-gene interaction, Hub genes, GO keywords, KEGG and WikiPathways, TF-gene interactions, and therapeutic drug compounds can all be found by using the common DEGs.

2 Proposed Mehtodology

2.1 Data Collection

In this study, we selected two diseases one is stress and other type 2 diabetes. We selected two datasets form NCBI are GSE183648 and GSE20966. Gene Expression Omnibus (GEO) database was used to gather for the GSE183648 and GSE20966 datasets, respectively [11]. Microarray and high-throughput sequencing datasets are both included in the GEO database, which is hosted on the platform that the NCBI oversees [12]. Microarray dataset GSE183648 demonstrates that p53 deficiency prevents DNA damage inflicted by replication stress in ATRX-deficient neuroblastoma cancer. The dataset is analyzed using GPL21185, an agilent-072363 SurePrint G3 Human GE v3 8x60K Microarray 039494 platform. GSE183648 has a total of 12 samples. Four of them are normal cells, while the remaining eight are stress cells. GSE183648 is comprised of a total of twelve samples, four of which are representative of normal cells and the remaining eight representing stress cells. Dataset GSE183648 was imparted by J. Akter et al [13]. Besides, GSE20966 is a microarray dataset that was collected beta-cell enriched tissue from people who had type 2 diabete and their gene expression profiles were analyzed using Laser Capture Microdissection. Affymetrix Human X3P Array GPL1352 [U133 X3P] platform was utilized to evaluate this dataset. Additionally, there were 10 samples taken from people who did not have diabetes and 10 samples taken from diabetics that were used in the microarray dataset. The GSE20966 dataset was contributed by Marselli et al., who gathered pancreatic beta cells via LCM [14].

2.2 Data filtering and finding of DEGs, and identification of concordant DEGs between Stress and Type 2 Diabetes

Transcriptomic datasets GSE183648 for Stress and GSE20966 for Type 2 Diabetes is employed for this study. This study begins by retrieving DEGs from the datasets GSE183648 and GSE20966. With the assistance of a R programming language, identify the differential expressed genes (DEGs). Preprocessing procedures were carried out on both sets of data with an adjusted p-value< 0.05 and a $|\log FC|>1$. The Benjamini-Hochberg approach was applied to the GSE183648 and GSE20966 datasets with the purpose of reducing the false discovery rate [15]. Identify the DEGs from both datasets and also find the common genes between them that scenario is depicted in a Venn diagram.

2.3 Design of protein-protein interaction network and analysis

The study of protein interactions, which is considered as the first step in drug discovery and systems biology, yields significant information about the functions of proteins [16]. PPIs networks [17] are being thoroughly investigated to find out how numerous complex biological processes there are [18]. Stress and Type 2 Diabetes DEGs that were found to be common were put into the

String [19] database with the help of the NetworkAnalyst tool [20]. Cytoscape (https://cytoscape.org/) was used to perform additional modifications to the network. Cytoscape software is a good way to combine how proteins interact with each other and how genes interact with each other.

2.4 Analysis of hub genes

The protein-protein interaction network shows the relationship between nodes and edges, it also illustrates the connections between the nodes. A hub gene is a node in a network that has the most connections to other nodes. Identifying the functional genes for this work is accomplished by the use of the degree topological method. Cytoscape [21] has been used in recent studies to do an analysis of the networks of PPIs. It is feasible to discover the genes that function as hubs for the related protein protein interaction network by applying the Cytoscape bioinformatics software plugin called as cytoHubba [22]. That plugin may be found at (http://apps.cytoscape.org/apps/cytohubba).

2.5 Enrichment analysis: Gene Ontology and Pathways

Enrichment analysis of gene is a computational and statistical technique for determining if a group of genes is significantly enriched in one biological context compared to another [23]. For the annotation of gene products, it is referred to as the Gene Ontology (GO) word, which is categorized into three subheadings: biological process, molecular function, and cellular component [24]. When it comes to learning about metabolic pathways, the KEGG pathway is quite popular due to the many benefits it offers in comparison to more conventional gene annotation approaches. Alongside the use of the KEGG pathway, the WikiPathways are also utilized. A web-based tool called Enrichr (https://amp.pharm.mssm.edu/Enrichr/) supplied all the pathways and GO keywords for the shared genes discovered in the initial step. Enrichr is a web service that performs enrichment analysis on gene sets for genes that have been subjected to genome-wide analysis [25].

2.6 TF-gene interactions

TF-gene interactions with similar DEGs analyze TF's impact on functional pathways and gene expression [26]. After determining which genes are similar throughout organisms, the NetworkAnalyst (https://www. networkanalyst.ca/) platform is used to find the TFs that interact with those genes. NetworkAnalyst is a powerful web tool that permits users to do gene expression for a wide variety of species, as well as meta-analysis. The NetworkAnalyst software utilizes the ENCODE [27] data to build the TF-gene interaction network.

2.7 Finding of candidate drug

Finding new drug molecules is a crucial aspect of ongoing study. A number of therapeutic molecules can be estimated using the DSigDB database and the Enrichr platform, both of which are based on common differential expressed genes.

There are a total of 22527 gene sets, 19531 genes, and 17389 distinct substances in DSigDB database [28]. DSigDB is mainly responsible for the prediction of drugs through the use of gene expression-based datasets, and each group of genes is taken into consideration whenever a drug is investigated.

3 Results

3.1 Findings of differential expressed genes and common DEGs among stress and type 2 diabetes

A comprehensive analysis of the GSE183648 dataset revealed 1547 genes to be differentially expressed. Total of 1547 genes, 654 of them are up-regulated, while 893 are down-regulated. Another dataset GSE20966 was analyzed, and it was found that 538 genes exhibit differential expression. Of these genes, 180 are up-regulated, and 358 are down-regulated. In the analysis of those two datasets, we found two common up-regulated genes and 23 common down-regulated genes. There were 2038 DEGs found with opposite behaviors among two diseases. Among the DEGs 47 were identified as common differential expressed genes between two datasets, and this filtered data is being used to perform a more comprehensive analysis of the ongoing research. The common DEGs among two datasets depicts by venn diagram in figure 2.

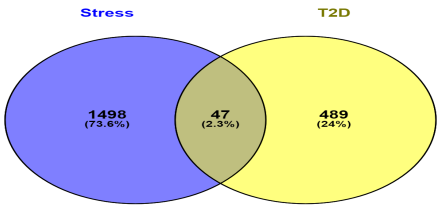

Fig. 2. *Common DEGs from the stress and T2D dataset depicted by that figure. There are a total of 1547 DEGs in the stress dataset, while there are 538 DEGs in the type 2 diabetes dataset. Of these, 47 were found to be highly significant. Only 2.3% of the 2085 DEGs were between the two data sets.*

3.2 Analysis of protein-protein interaction network

As an input, the common differential expressed genes (DEGs) were provided to NetworkAnalyst. This study focused mostly on the protein-protein interaction

network, analyzing hub genes and specific modules in relation to a network. NetworkAnalyst, a web-based tool, is used in connection with the STRING database to produce SIF files for the network diagram. The protein-protein interaction network and top 5 five hub gene connection are shown in figure 3.

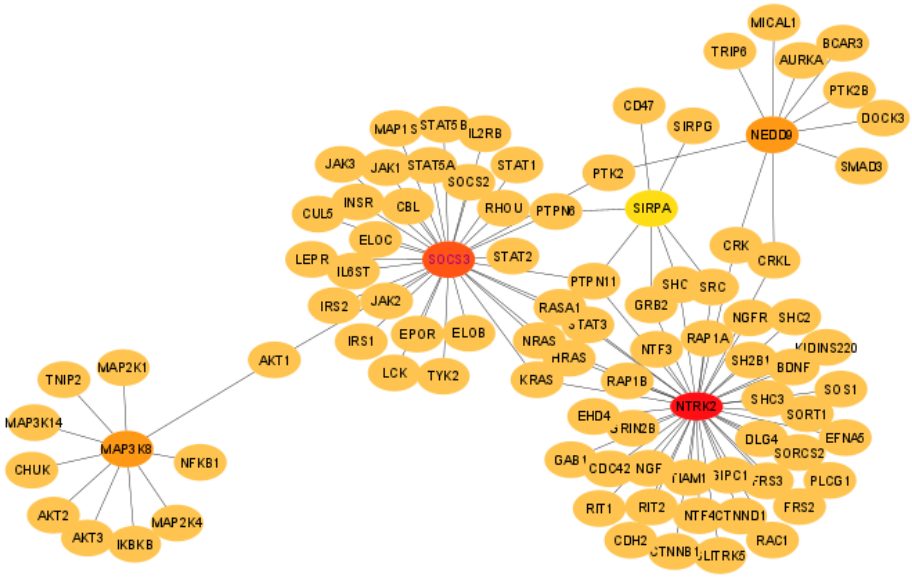

Fig. 3. Protein-protein interactions (PPIs) network created by the 47 common DEGs. Orange/red color nodes indicates hub genes, and yellow nodes represents the connections between the hub genes and other. The PPIs network has a total of 97 nodes, 106 edges, and 6 seeds.

3.3 Hub gene identification

The nodes in a network that are the most connected to other nodes are referred to as hub nodes. After doing an analysis of the PPIs network with cytohubba, the top ten most active genes have been determined for the purposes including further study. These ten genes are the most important ones: NTRK2, SOCS3, NEDD9, MAP3K8, SIRPA, BCL3, PTPN11, AKT1, HRAS, and SHC1. The degree value of NTRK2 is greater than that of any other DEG in the network. Cytoscape's PPI network is examined with that program's Network Analyzer in order to determine the topological characteristics of the network. Table I contains an overview of the topological analysis performed on the top five genes.

Table 1. The topological properties of the five most important hub genes, as examined through cytoscape.

Hub gene	*Degree*	*Stress*	*Closeness Centrality*	*Betweenness Centrality*
NTRK2	43	23438.0	58.95476	5683.97922
SOCS3	32	22570.0	53.38333	5898.04545
NEDD9	10	2652.0	36.4381	1333.92078
MAP3K8	10	8076.0	32.61667	2448.0
SIRPA	7	2142.0	33.77143	480.05455

3.4 According to the common DEGs analyze GO and Pathways

Following the identification of shared DEGs between stress and type 2 diabetes, several databases (KEGG, WikiPathways) were used to discover GO keywords and cell pathways. To further understand the relationships between these three GO terms, we take a look at their molecular components in figure 4 (A). According to the results presented in the molecular function section, a transition metal ion binding factor is magnificently implicated in the widespread DEGs. KEGG and WikiPathways are also analyzed by the common DEGs between two diseases and illustrates in figure 4(B).

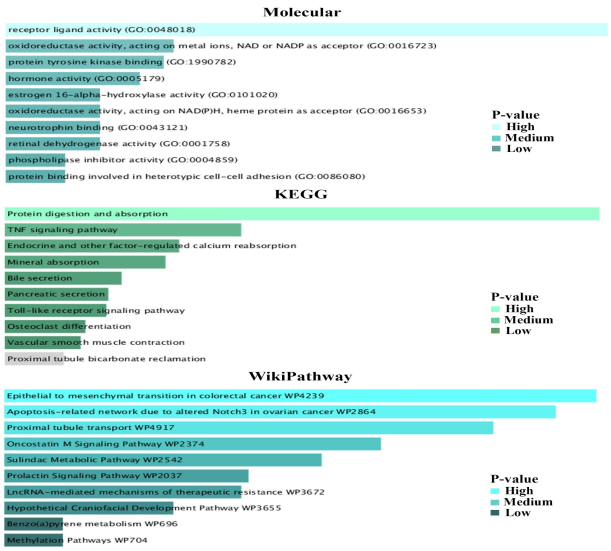

Fig. 4. (A) Identification of Molecular Component-Related GO Terms Based on a Weighted Average of Their Individual Scores. If an ontology has a high enrichment score, then a large proportion of relevant genes are represented in that ontology. (B) Locating KEGG and WikiPathways pathway analysis outcomes. The cumulative score was used to determine which pathway terms had yielded significant results.

3.5 TF-gene interactions

TF-gene interactions mainly show the connections between TF-gene and common genes between those two diseases. The TF-gene is determined by the common DEGs using NetworkAnalyst online tool. In Figure 5, the similar DEGs and the TF regulators that interact with them are displayed. There are 104 nodes, 321 edges, and 42 seeds in this network by using JASPAR database. The network of interactions between TFs and common genes is shown in Figure 5.

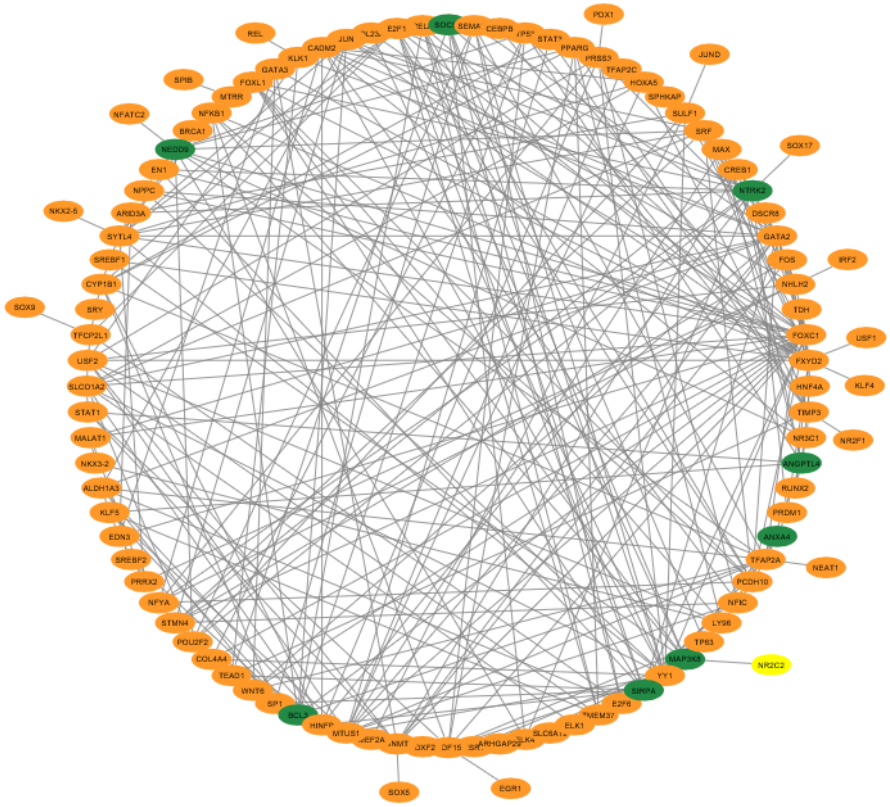

Fig. 5. Commonly differentially expressed genes connected to a network of TF-gene interactions. Genes that are shared by multiple organisms are represented by the green node, whereas the other yellow nodes represent different TF genes. There are 104 nodes in the network, 321 edges between them, and 42 seeds.

3.6 Identification of drug compounds

The Enrichr online program is used to extract the drug compounds from the DSigDB database to provide the desired results. The following drugs were sug-

gested based on a combination of the p-value and the modified p-value. The table below describes potential therapeutic compounds for stress and type 2 diabetes that combine common DEGs. Table 2 is utilized to evaluate which drug compounds have been the most successful by employing the DEGs that are most commonly encountered.

Table 2. Drug compounds estimates based on the degree's most heavily linked to stress and type 2 diabetes.

Drugs	*p-value*	*Adjusted p-value*	*Genes*
STOCK1N-35696 MCF7 UP	5.69E-06	0.003823864	ALDH1A3, GDF15, CYP1B1, NEAT1; ANGPTL4
Caspan CTD 00000180	1.46E-05	0.003823864	NHLH2,CADM2,ANXA4, MTUS1, NEDD9, CYP1B1, ARHGAP29, TFCP2L1, SULF1, NEAT1
trichostatin A CTD 00000660	2.01E-05	0.003823864	FOXC1, PCDH10, GDF15, ANXA4, COL23A1
puromycin MCF7 UP	2.09E-05	0.003823864	GDF15,BCL3,NEDD9,MAP3K8
GW-8510 MCF7 UP	2.32E-05	0.003823864	NHLH2, ALDH1A3, GDF15, CYP1B1

4 Discussion

Stress can have a significant impact on people with type 2 diabetes. Stress can cause blood sugar levels to rise, making it difficult to manage diabetes. Stress can also lead to unhealthy behaviors, such as overeating, which can further increase blood sugar levels. Additionally, stress can lead to depression and anxiety, which can make it difficult to manage diabetes. To help manage stress, people with type 2 diabetes should practice relaxation techniques, such as deep breathing, yoga, and meditation. In a component of the ongoing study, bioinformatics analysis was performed to combine shared DEGs between stress and type 2 diabetes, and protein-protein interaction network analysis was conducted once the common genes were identified. The drug compounds are proposed for the therapeutic of stress and type 2 diabetes based on the differential expressed genes (DEGs). Analysis of differential expressed genes (DEGs) between GSE183648 and GSE20966 datasets, identified 47 common DEGs among these two diseases. Owing of its necessity to the ongoing study, the PPI network is the next target of this evaluation.Hub genes as well as the 10 most prominent hub genes were found using the PPIs network. Using a degree topological strategy, the PPIs network singled out 10 major hub genes. Therapeutic drug compounds included in the top ten include those for STOCK1N-35696 MCF7 UP, Caspan

CTD 00000180, trichostatin A CTD 00000660, puromycin MCF7 UP, and GW-8510 MCF7 UP. Several bioinformatics analyses have been carried out to merge similar DEGs and to use them to pinpoint specific drug compounds. Stress and type 2 diabetes therapeutic are promising areas for the suggested drugs.

5 Conclusion

There has never been a study of stress and type 2 diabetes in the field of transcriptomic analysis. In genetic innovation, the bioinformatics field has made substantial developments. Using the technique of system biology and the functional annotation method, we were able to identify the primary pathways and biomolecules that are responsible for stress and type 2 diabetes. A diverse range of bioinformatics analysis methods are utilized in the process of gene filtering, which is an indispensable step in the field of systems biology. After that, the genes are examined with one another to find a regular healing drug for stress and type 2 diabetes. According to similar differentially expressed genes (DEGs), a protein-protein interaction network was designed and also found the top ten responsible genes (NTRK2, SOCS3, NEDD9, MAP3K8, SIRPA, BCL3, PTPN11, AKT1, HRAS, and SHC1) depending on their degree. The components of mass-produced successful remedial drugs can be identified when two diseases contain the same DEGs because the hub DEGs are highly harmful. The most essential five biomarkers were prognostic, which helped make a good therapeutic drug molecule for stress and type 2 diabetes. Through our research work, we identified novel molecular biomarkers and provided guidelines for the effective diagnosis and treatment of stress and type 2 diabetes at an early stage.

Acknowledgements. This paper is neither published nor being considered for publication anywhere else at this time. Each and every person who helped with this study is greatly appreciated by the authors.

Funding. This work is not finantially supported.

References

1. Rahal, A., Kumar, A., Singh, V., Yadav, B., Tiwari, R., Chakraborty, S. and Dhama, K., 2014. Oxidative stress, prooxidants, and antioxidants: the interplay. BioMed research international, 2014.
2. Selye, H., 1976. Stress without distress (pp. 137-146). Springer US.
3. Basar, M.A., Hosen, M.F., Paul, B.K., Hasan, M.R., Shamim, S.M. and Bhuyian, T., 2023. Identification of drug and protein-protein interaction network among stress and depression: A bioinformatics approach. Informatics in Medicine Unlocked, p.101174.
4. Theodore, W.H., Spencer, S.S., Wiebe, S., Langfitt, J.T., Ali, A., Shafer, P.O., Berg, A.T. and Vickrey, B.G., 2006. Epilepsy in North America: a report prepared under

the auspices of the global campaign against epilepsy, the International Bureau for Epilepsy, the International League Against Epilepsy, and the World Health Organization. Epilepsia, 47(10), pp.1700-1722.
5. Cosgrove, M.P., Sargeant, L.A., Caleyachetty, R. and Griffin, S.J., 2012. Work-related stress and Type 2 diabetes: systematic review and meta-analysis. Occupational Medicine, 62(3), pp.167-173.
6. Hosen, M.F., Basar, M.A., Paul, B.K., Hasan, M.R. and Uddin, M.S., 2022, December. A bioinformatics approach to identify candidate biomarkers and common pathways between bipolar disorder and stroke. In 2022 12th International Conference on Electrical and Computer Engineering (ICECE) (pp. 429-432). IEEE.
7. Joint National Committee on Prevention, Evaluation, A., Treatment of High Blood Pressure and National High Blood Pressure Education Program, 1997. Report of the Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure (Vol. 6). Public Health Service, National Institutes of Health, National Heart, Lung, and Blood Institute.
8. Goyal, S., Morita, P., Lewis, G.F., Yu, C., Seto, E. and Cafazzo, J.A., 2016. The systematic design of a behavioural mobile health application for the self-management of type 2 diabetes. Canadian journal of diabetes, 40(1), pp.95-104.
9. Gesinde, B., 2019. An Avatar video intervention on type 2 diabetes for women of color using brief motivational interviewing: Predictors of self-efficacy post-video for performing the American Association of Diabetes Educator's Seven Self-care Behaviors (Doctoral dissertation, Teachers College, Columbia University).
10. Hasan, M.R., Paul, B.K., Ahmed, K. and Bhuyian, T., 2020. Design protein-protein interaction network and protein-drug interaction network for common cancer diseases: A bioinformatics approach. Informatics in Medicine Unlocked, 18, p.100311.
11. Clough, E. and Barrett, T., The gene expression omnibus database. InStatistical genomics 2016 (pp. 93-110). Humana Press, New York, NY. DOI: DOI, 10, pp.978-1
12. Edgar, R., Domrachev, M. and Lash, A.E., 2002. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. Nucleic acids research, 30(1), pp.207-210.
13. Akter, J., Katai, Y., Sultana, P., Takenobu, H., Haruta, M., Sugino, R.P., Mukae, K., Satoh, S., Wada, T., Ohira, M. and Ando, K., 2021. Loss of p53 suppresses replication stress-induced DNA damage in ATRX-deficient neuroblastoma. Oncogenesis, 10(11), p.73.
14. Marselli, L., Thorne, J., Dahiya, S., Sgroi, D.C., Sharma, A., Bonner-Weir, S., Marchetti, P. and Weir, G.C., 2010. Gene expression profiles of Beta-cell enriched tissue obtained by laser capture microdissection from subjects with type 2 diabetes. PloS one, 5(7), p.e11499.
15. Benjamini, Y. and Hochberg, Y., 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal statistical society: series B (Methodological), 57(1), pp.289-300.
16. Shadhin, K.A., Hasan, M.R., Paul, B.K., Ahmed, K., Moni, M.A., Kahrizi, D., Bhuyian, T. and Ibrahim, S.M., 2020. Analysis of topological properties and drug discovery for bipolar disorder and associated diseases: A bioinformatics approach. Cellular and Molecular Biology, 66(7), pp.152-160.
17. Šikić, M., Tomić, S. and Vlahoviček, K., 2009. Prediction of protein–protein interaction sites in sequences and 3D structures by random forests. PLoS computational biology, 5(1), p.e1000278.
18. Pagel, P., Kovac, S., Oesterheld, M., Brauner, B., Dunger-Kaltenbach, I., Frishman, G., Montrone, C., Mark, P., Stümpflen, V., Mewes, H.W. and Ruepp, A., 2005.

The MIPS mammalian protein–protein interaction database. Bioinformatics, 21(6), pp.832-834.
19. Mering, C.V., Huynen, M., Jaeggi, D., Schmidt, S., Bork, P. and Snel, B., 2003. STRING: a database of predicted functional associations between proteins. Nucleic acids research, 31(1), pp.258-261.
20. Xia, J., Gill, E.E. and Hancock, R.E., 2015. NetworkAnalyst for statistical, visual and network-based meta-analysis of gene expression data. Nature protocols, 10(6), pp.823-844.
21. Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B. and Ideker, T., 2003. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome research, 13(11), pp.2498-2504.
22. Chin, C.H., Chen, S.H., Wu, H.H., Ho, C.W., Ko, M.T. and Lin,C.Y., 2014. cytoHubba: identifying hub objects and sub-networks from complex interactome. BMC systems biology, 8(4), pp.1-7.
23. Subramanian A, Kuehn H, Gould J, et al. GSEA-P: a desktop application for Gene Set Enrichment Analysis. Bioinformatics 2007;23(23):3251–3.
24. Delfs, R., Doms, A., Kozlenkov, A. and Schroeder, M., 2004. GoPubMed: ontology-based literature search applied to Gene Ontology and PubMed. In German Conference on Bioinformatics 2004, GCB 2004.Gesellschaft fur Informatik eV.
25. Kuleshov, M.V., Jones, M.R., Rouillard, A.D., Fernandez, N.F., Duan, Q., Wang, Z., Koplev, S., Jenkins, S.L., Jagodnik, K.M., Lachmann, A. and McDermott, M.G., 2016. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. Nucleic acids research, 44(W1), pp.W90-W97.
26. Ye, Z., Wang, F., Yan, F., Wang, L., Li, B., Liu, T., Hu, F., Jiang, M., Li,W. and Fu, Z., 2019. Bioinformatic identification of candidate biomarkers and related transcription factors in nasopharyngeal carcinoma. World journal of surgical oncology, 17(1), pp.1-10.
27. Davis, C.A., Hitz, B.C., Sloan, C.A., Chan, E.T., Davidson, J.M., Gabdank, I., Hilton, J.A., Jain, K., Baymuradov, U.K., Narayanan, A.K. and Onate, K.C., 2018. The Encyclopedia of DNA elements (ENCODE): data portal update. Nucleic acids research, 46(D1), pp.D794-D801.
28. Yoo M, Shin J, Kim J, et al. DSigDB: drug signatures database for gene set analysis. Bioinformatics 2015;31(18):3069–71..