Drug design and analysis for bipolar disorder and associated diseases: A bioinformatics approach

Nahida Habib¹, Kawsar Ahmed²³, Iffat Jabin¹, Mohammad Motiur Rahman¹
¹Department of Computer Science and Engineering (CSE), Mawlana Bhashani Science and Technology University (MBSTU), Santosh, Tangail-1902, Bangladesh
²Department of Information and Communication Technology (ICT), Mawlana Bhashani Science and Technology University (MBSTU), Santosh, Tangail-1902, Bangladesh
³Group of Bio-photomaterial, Santosh, Tangail-1902, Bangladesh
E-mail: nahidahabib164@yahoo.com, kawsar.ict@mbstu.ac.bd, k.ahmed.bd@ieee.org, kawsarit0805@gmail.com, iffat.jabin02@gmail.com, mm73rahman@gmail.com

Received 27 April 2017; Accepted 19 May 2017; Published 1 June 2017

Abstract
Bioinformatics deals with biological data and analyzes or processes the data using computer science techniques. With the appearance of modern bioinformatics tools, it is now possible to design a drug using these high technologies and open a new area of drug design and development. This research predicts to design a common drug for four associated mental disorders that include bipolar disorder, schizophrenia, coronary heart diseases and stroke. The key to drug design is a biomolecule or protein. To show the protein interactions and evolutions, a protein–protein interaction network is created among the common genes of the four diseases. The genes corresponding to each disease are collected from NCBI gene database. These genes are preprocessed, mined and verified to find the common genes among the diseases. After getting common genes (7 genes), PPI network is created with them. Then a common drug is designed that will work on four investigated diseases. This structure based drug design research will open a new era to discover and develop new drug compounds using different bioinformatics tools.

Keywords bioinformatics; mental disorders; bipolar disorder; schizophrenia; coronary heart diseases; stroke; PPI Network; drug design.

1 Introduction
Bioinformatics plays an important role in data processing, sequencing, handling large dataset, storing, transformation, visualizing PPI network and last but not least drug design. Modern bioinformatics tool made it easier to design a common drug for bipolar disorder and associated diseases. Drug targets are typically key
molecules involved in a specific metabolic or cell signaling pathway that is known, or believed, to be related to a particular disease state and are most often proteins and enzymes in these pathways (Zhang, 2016b; TechTarget, 2017; Zhang and Feng, 2017). A common pathway shared (Nahida, 2016), by bipolar disorder, schizophrenia, coronary heart diseases and stroke can be designed to show the association and interaction among the proteins of these diseases which leads to the way to drug design. Moreover, designing a common drug for four associated diseases will decrease the amount of drug one should absorb for the diseases separately.

Mental disorder, mental illness, major depressive disorder or psychiatric disorder include a wide range of problems, such as depression, major depression, bipolar disorder, schizophrenia, coronary heart diseases, stroke, anxiety disorders etc. There may be multiple causes of mental disorders. Genes, biological factors, environment and lifestyle, family history, a stressful job or home life, traumatic brain injury, mother's exposure to viruses or toxic chemicals while pregnant may play a part (NIH). One in 5 adults experiences a mental health condition every year and one in 17 lives with a serious mental illness such as schizophrenia or bipolar disorder (Nami National Alliance on Mental Illness). Childhood disorders can lead to adult disorders or personality disorders.

Bipolar disorder and schizophrenia are among the most severe and complex mental disorders. Bipolar disorder-I (BD-I) is a highly prevalent and often chronic mood disorder with a lifetime prevalence of 1% to 5% (Akiskal et al., 2000), which is frequently characterized by episodic recurrent mania or hypomania and major depression (Belmaker, 2004). Bipolar disorder (BD) that is known as manic-depressive illness, affects approximately 1% of the general population (Merikangas et al., 2011) with high heritability (Edvardsen et al., 2008). Schizophrenia is a devastating disorder affecting ~1% of the population (Ayalew, 2012). Overall, the incidence of schizophrenia was found to be higher in males than females (McGrath et al., 2004). The ratio of incidence rates between men and women was 1.4 (Aleman et al., 2003).

CHD and stroke are also complex disorders. Coronary heart disease (CHD) is a leading cause of death in the United States (Keenan, 2011) and the world (Lozano et al., 2012). Within the studies focusing on medical illnesses among patients with bipolar disorder, the most common medical problems cited are obesity, diabetes mellitus and subsequent cardiovascular disease; all of these medical conditions, as well as depressive symptoms, are recognized as risk factors for stroke (Everson et al., 1998; Krishnan, 2005). Nevertheless, there is scant information on the risk of developing stroke among patients with bipolar disorder, despite cerebrovascular diseases having been reported as one of the major causes of death among this particular patient population (Hoyer et al., 2000; Joukamaa et al., 2001; Tsai et al., 2005). CVD is the leading cause of death in BD, with a standardized mortality ratio of 1.5 to 2.5 (Osby, 2001; Weeke, 1987). Approximately 38% of all individuals who suffer a heart attack will die from it (Rosamond et al., 2008).

This research investigates with BD, SZ, CHD and stroke whose are directly or indirectly associated. For the first time we tried to design a common drug for these associated diseases in this research. This paper is organized in 5 sections. Section 2 discusses about the background or previous works related to the research, section 3 describes the proposed methodology and working principle, section 4 discusses and analyzes the result and last but not least section 5 includes conclusion and future work.

2 Background
Prior research shows the interaction between the investigated four diseases (BD, SZ, CHD and stroke). Also different bioinformatics tools were used for this purpose. Information about the bio-informatics tools that can be used to show interaction, finding common genes and represent the PPI network among genes as well as proteins are described in the article (Klingstrom and Plwczynski, 2010). It has recently been shown that
mental diseases like schizophrenia, major depression and bipolar disorder may be quite closely related genetically, meaning that there is a considerable genetic overlap between the three different diagnoses (Rasic, 2013). Based on this recent meta-analysis it was found that offspring of parents with one of the mentioned diseases not only had an increased risk of developing the same illness as their ill parent, but also an increased risk of developing the other two disorders (Thorup et al., 2015). The detail information about associations of schizophrenia and psychotic bipolar disorder for default mode network were provided on the article by Medaa (2014).

There is substantial evidence for partial overlap of genetic influences on schizophrenia and bipolar disorder, with family, twin, and adoption studies showing a genetic correlation between the disorders of around 0.6 (Cardno and Owen, 2014). The article (Osby et al., 2001), said that patients with BD are up to twice as likely to die from cardiovascular causes as their counterparts in the general population. The association between major depression, bipolar disorder & cardiovascular diseases are shown in the paper (Baune et al., 2006). The article (Sowden and Huffman, 2009) interpreted that depression, anxiety disorders, schizophrenia and bipolar disorder (BD) have all been identified as risk factors for the onset and progression of cardiovascular disease (CVD). Individuals with chronic heart diseases or stroke have a significantly increased incidence and prevalence of affective disorders (Baune et al., 2006). The paper (De Hert et al., 2009), mentioned that compared with healthy controls, people with SMI who were not prescribed any antipsychotics were at increased risk of CHD and stroke than controls, whereas those prescribed such agents were at even greater risk. The paper also revealed that those receiving the higher doses were at greatest risk of death from both CHD and stroke. In the United States alone, more than 16.3 million adults have CHD and an estimated 935,000 heart attacks occur each year (Roger et al., 2012). With regard to vascular outcomes, lifetime rates of stroke, myocardial infarction, and other forms of CVD are elevated among people who report having been maltreated as children (Goldstein et al., 2015).

![Collected Genes for Specific Diseases](image-url)

Fig. 1 Collected genes for specific diseases.
Jesmin et al. (2016) provided a common disease regulatory network for metabolic disorders by investigating on associated diseases. The UNIHI tool that is used to predict PPI network, common metabolic pathway and drug design is referred by Kalathur et al. (2014). The role of gene duplication for creating gene network evolution has been investigated in the paper of Teichmann and Babu (2014). In addition, Li and Zhang (2013) identified crucial metabolites/reactions in tumor signaling networks. Zhang (2016a) developed a mathematical model to describe dynamics of occurrence probability of missing links in predicted missing link list. Zhang and Feng (2017) used network analysis to analyze metabolic pathway of non-alcoholic fatty liver disease.

Various previous research have been held on bioinformatics that includes analyzing genes, finding PPI network and common pathway for associated diseases. The presented research is a descendant of previous research which aimed to design a common drug for BD, SZ, CHD and stroke.

![Cross Linkage Genes between Diseases](image)

**Fig. 2** Cross linkage genes between selected diseases.

3 Proposed Methodology

Drug design is a step by step process. Several steps are performed in this section to reach the desire goal. Each of the steps is described below in the following subsections through 3.1 to 3.4 respectively.

3.1 Gene filtering

The NCBI (National Center for Biotechnology Information) is an important resource for bioinformatics tools...
and services. It maintains a huge database of all the DNA and protein sequence data. For this research project, genes associated with bipolar disorder, schizophrenia, coronary heart disease and stroke diseases are collected from NCBI gene database (Nahida, 2016). Collected genes are preprocessed and filtered and the genes only responsible for Homo sapiens are stored for further processing.

3.2 Gene mining

Gene mining is one of the most important parts of this research. Corresponding genes are downloaded in the sorted increasing order by their weight. The collected genes for each disease are merged to find gene linkage. The interrelated genes between 2 selected diseases (like BD & CHD; BD & Stroke etc.) and among 3 selected diseases (like BD, CHD & stroke; BD, CHD & schizophrenia, etc.) are identified and collected. From the sorted linkage gene files only the genes are mined (Nahida, 2016). Then the common genes from top 100 and top 50 interrelated genes are searched and collected.

3.3 PPI network

Protein-Protein Interaction Network or PPI network is used to show the protein interaction and common pathway among the interrelated genes. UniHi, a very popular reliable bioinformatics tools is used for this purpose. In this step, from the interrelated common genes, PPI networks and common pathways are created using UniHi tool.

![Cross Linkage Genes among Diseases](image)

**Fig. 3** Cross linkage genes among diseases.
Fig. 4 Protein-protein interaction (PPI) network with 17 common genes.

3.4 Design drug
This is the most vital step of the current research project. A drug is any substance (other than food that provides nutritional support) that, when inhaled, injected, smoked, consumed, absorbed via a patch on the skin, or dissolved under the tongue, causes a physiological change in the body (Wikipedia). In other word, a drug can be defined as a substance used to treat, cure and prevent an illness, relieve from a pain, or modify some specific process in the body for some specific cure. Therapeutic response for a disease is the key root to the invention of a drug. It is the devising process of finding new drugs based on the investigated protein or molecules. The drug should be designed and developed in such a way that it does not disturb the normal chemical process of the body and only affect the target protein. In order to dispose a disease, specific drug needs to be exhibited along with target identification to affect the target genes or proteins. Bioinformatics tools made it easier for researcher to research on specific diseases, leading to the design and development of drug for those diseases. Rapid and revolutionary developments in genome sciences, combinatorial chemistry, informatics and robotics are having major impacts on drug discovery (Blundell, 2002). Protein structure can influence drug discovery at every stage in the design process and can also be used in target identification and
selection (Blundell et al., 2006). So, before designing a drug, analyzing or creating PPI Network is important. From the PPI network and common pathway a drug target can be designed using UniHi tool.

4 Results and Discussion
A disordered or any abnormal condition caused by some interrelated genes is called a disease. The interrelated genes are downloaded from trustable gene databases using bioinformatics tools whose are then used to drug design.

4.1 Gene filtering
The collected responsible genes without preprocessing & filtering are estimated as 640 for BD, 2555 for schizophrenia, 294 for CHD and 1028 for Stroke and after preprocessing as well as filtering the corresponding genes for *Homo sapiens* are 619 for BD, 1612 for Schizophrenia, 288 for CHD and 638 for
stroke (Nahida, 2016). Before and after filtering the number of genes for each specific disease is displayed using a bar plot as shown in Fig. 1.

4.2 Gene mining
After performing cross linkage between and among the investigated corresponding genes, the resultant cross linkage genes between BD & SZ, BD & CHD, BD & stroke, SZ & CHD, CHD & stroke, SZ & stroke are 363, 39, 66, 96, 153, 171 respectively and among BD & SZ & CHD, BD & SZ & stroke, BD & CHD & stroke, SZ & CHD & stroke are 34, 54, 33, 67 respectively (Nahida, 2016). Finally, cross linkage genes among the investigated four diseases (BD, SZ, CHD, stroke) are 29. After mining these 29 common genes and taking the common genes from top 50 weighted genes 17 common genes results. They are TNF, MTHFR, IL6, ACE, ESR1, SLC6A4, ABCB1, CRP, NOS3, IL1B, CXCL8, GSTM1, IGF1, IFNG, GSTT1, PON1 and CCL2. The bar plot in Fig. 2 and Fig. 3 shows the cross linkage genes between and among associated diseases.

4.3 PPI network and regulatory interaction
Using UniHI tool, a PPI network and a regulatory interaction network is created. These networks are used to represent the directly or indirectly connected gene and protein interaction. Fig. 4 and Fig. 5 shows the PPI and Regulatory Interaction Network respectively. From Fig. 4 and Fig. 5, it is clear that only 7 genes are

![Fig. 6 Protein-protein interaction (PPI) network with 7 common genes.](image)
responsible for direct interconnection with each other in the networks. They are TNF, IL6, ESR1, IL1B, CXCL8, IFNG, CCL2. So, these common 7 genes are now used to create a PPI Network as displayed in Fig. 6. Regulatory interaction network displays the directly interacted proteins with the target diseases. The regulatory interaction network of Fig. 7 represents the only interacted proteins of associated genes of target diseases.

![Fig. 7 Regulatory interaction network with 7 common genes.](image)

Cytoscape is an open source software project for integrating biomolecular interaction networks with high-throughput expression data and other molecular states into a unified conceptual framework (Shannon et al., 2003). The network in Fig. 8 is generated in Cytoscape with 7 common genes. STRING (Search Tool for the Retrieval of Interacting Genes/Proteins) is a biological database that contains various information and shows predicted protein–protein interactions. In Fig. 9, the network with 7 common genes is generated using STRING.
Fig. 8 Network with 7 common genes using Cytoscape.

Fig. 9 Network with 7 common genes using String.
GeneMANIA is a bioinformatics tool that shows functional association data, genetic interactions, pathways and co-expression for a set of input data. From the common 7 genes GeneMANIA creates a network as shown in Fig. 10.

![Network with 7 common genes using GeneMania.](image)

**Fig. 10** Network with 7 common genes using GeneMania.

### 4.4 Design drug

The ultimate goal of this research is designing a common drug for the investigated diseases. Target identification is insufficient for achieving a successful treatment of a disease. Real drugs need to be developed. The target proteins must be influenced by drug in such a way that it does not interfere with normal metabolism. To achieve this activity of protein, various bioinformatics tools are developed. UniHi tool is one of popular bioinformatics tool. Using UniHi tool, a common drug target is designed for selected diseases. Fig. 11 and Fig. 12 are exhibited the structure based drug design for investigated 7 common genes from different side of views. These figures also reveal that there is a strong correlation among 7 investigated genes via some proteins. Figure 11 and 12 also demonstrate the affected and unaffected proteins in the structure. Red color proteins has directed compound with targeted genes and yellow color proteins are vice versa.

It is well-known that a drug must bind to a particular spot on a particular protein or nucleotide. We need to identify and study the lead compounds that have direct activity against a disease. To find out the proteins whose are directly interacted with target diseases, filtering technique is used here. Fig. 13 and 14 illustrate the only interacted proteins of associated genes of target diseases from different point of view. We often apply in several techniques and test a large numbers of compounds from a database that have available structures.
Fig. 11 Drug target network with 7 common genes from side view-1.

Fig. 12: Drug target with 7 common genes from side view-2.
Fig. 13 Drug target network with 7 common genes after filter from side view-1.

Fig. 14 Drug target network with 7 common genes after filter from side view-2.
5 Conclusions
In the genomic revolution, the contribution of Bioinformatics is incredible. The developments of Bioinformatics tools have disclosed new research area and made uncompromising task easier. Researcher can now analyze diseases, identify and process disease genes to cure the diseases. Computer aided drug design aimed to design a drug to redress certain diseases. Target identification and drug design leads to the development of a drug. According to Anderson (2003), many years of research may be necessary to convert a drug lead into a drug that will be both structure-based drug design and includes, primarily, effective and tolerated by the human body.

Major depression is the key causes of BD, schizophrenia, CHD and stroke diseases. To create a common drug for these diseases the corresponding genes of each disease are studied and particular operations are performed on them. Finally, in this research project a common drug is designed that will remedy these diseases. This research is mainly helpful to understand the PPI, Regulatory Interaction Network and to target drug design. The future work of the research is to work on several other interrelated diseases to design a common drug for those diseases.

Abbreviations
BD= Bipolar Disorder; SZ=Schizophrenia; CHD= Coronary Heart Disease; CVD= Cardiovascular disease.

Acknowledgment
The authors are grateful to the participants who contributed to this research.

References
Anderson AC. 2003. The process of structure-based drug design. Chemistry & Biology, 10: 787-797
De Hert et al. 2009. Cardiovascular disease and diabetes in people with severe mental illness position
statement from the European Psychiatric Association (EPA), supported by the European Association for the Study of Diabetes (EASD) and the European Society of Cardiology (ESC). European Psychiatry, 24: 412-424


Everson SA., Roberts RE., Goldberg DE, Kaplan GA. 1998. Depressive symptoms and increased risk of stroke mortality over a 29-year period. Archives of Internal Medicine, 158: 1133-1138


Sweden. Arch Gen Psychiatry, 58: 844–850
Thorup et al. 2015. The Danish high risk and resilience study. BMC Psychiatry, 15: 233