Article

Application of R to investigate common gene regulatory network pathway among bipolar disorder and associate diseases

Nahida Habib[1], **Kawsar Ahmed**[2], **Iffat Jabin**[1], **Mohammad Motiur Rahman**[1]

[1]Department of Computer Science and Engineering (CSE), Mawlana Bhashani Science and Technology University (MBSTU), Santosh, Tangail-1902, Bangladesh

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

E-mail:nahidahabib164@yahoo.com,kawsar.ict@mbstu.ac.bd,k.ahmed.bd@ieee.org,kawsarit08050@gmail.com,iffat.jabin02@gmail.com, mm73rahman@gmail.com

Received 5 July 2016; *Accepted 10 August 2016*; *Published 1 December 2016*

(cc) BY

Abstract

Depression, Major Depression or mental disorder creates severe diseases. Mental illness such as Unipolar Major Depression, Bipolar Disorder, Dysthymia, Schizophrenia, Cardiovascular Diseases (Hypertension, Coronary Heart Disease, Stroke) etc., are known as Major Depression. Several studies have revealed the possibilities about the association among Bipolar Disorder, Schizophrenia, Coronary Heart Diseases and Stroke with each other. The current study aimed to investigate the relationships between genetic variants in the above four diseases and to create a common pathway or PPI network. The associated genes of each disease are collected from different gene database with verification using R. After performing some preprocessing, mining and operations using R on collected genes, seven (7) common associated genes are discovered on selected four diseases (SZ, BD, CHD and Stroke). In each of the iteration, the numbers of collected genes are reduced up to 51%, 36%, 10%, 2% and finally less than 1% respectively. Moreover, common pathway on selected diseases has been investigated in this research.

Keywords bipolar disorder; major depressive disorder; Schizophrenia; gene mining; PPI network and R toolkit.

Network Biology
ISSN 2220-8879
URL: http://www.iaees.org/publications/journals/nb/online-version.asp
RSS: http://www.iaees.org/publications/journals/nb/rss.xml
E-mail: networkbiology@iaees.org
Editor-in-Chief: WenJun Zhang
Publisher: International Academy of Ecology and Environmental Sciences

1 Introduction

Clinical depression, major depression or major depressive disorder is the more severe form of depression. Symptoms of depression can include agitation, restlessness, irritability, anger, lack of concentration or interest, self-hate, hopelessness, helplessness, changes in movement, sleep and appetite, thoughts of death or suicide. Today, depression is estimated to affect 350 million people (WHO). At its worst, depression can lead to

suicide and almost 1 million lives are lost yearly due to suicide, which translates to 3000 suicide deaths every day (Marcus et al., 2016).

Bipolar disorder, also called manic-depressive illness, is not as common as major depression or persistent depressive disorder is characterized by cycling mood changes-from extreme highs (e.g., mania) to extreme lows (e.g., depression) (NIMH, 2016). About 5.7 million American adults, or about 2.6 percent of the population, age 18 and older in any given year, have bipolar disorder (IBF, 2014). BD 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). Approximately 20% of the patients die of suicide (Kilbane et al., 2009). Co-morbid disorders are frequent and include adult attention-deficit/hyperactivity disorder (aADHD), anxiety disorders, and substance abuse (Kessler et al., 2006; Merikangas et al., 2006). In addition, BD-I is associated with higher rates of general medical comorbidities including premature mortality from cardiovascular, respiratory, and endocrine causes (Roshanaei-Moghaddam and Katon, 2009; Carney and Jones, 2006; Weber et al., 2011).

Schizophrenia and bipolar disorder are among the most severe and disabling mental disorders, influencing both patients, relatives and society (Thorup et al., 2015, Laursen et al., 2009). Schizophrenia is complex biochemical brain disorder that affects a person's thinking and behaviors. Patients with schizophrenia may have the difficulty to distinguish between the reality and the imaginary. In 2012 according to Ayalew M et al., Schizophrenia is a devastating disorder affecting ~1% of the population. 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). Another study has recently 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 (Thorup et al., 2015).

Coronary artery disease and its sequelae (angina, myocardial infarction, cardiac revascularization or transplantation) affect more than 16 million Americans (Rosamond W et al., 2008) and untold numbers of people world-wide (Dorn and Cresci, 2009). It is a condition for the buildup of plaque in the heart's arteries called atherosclerosis that leads to blockages or heart attack. It's responsible for more than 73,000 deaths in the UK each year. About 1 in 6 men and 1 in 10 women die from CHD (National Health Service (NHS, 2016).In the United States alone, more than 16.3 million adults have CHD and an estimated 935,000 heart attacks occur each year (Joehanes et al., 2013). The article (Goldstein et al., 2015) has shown the association between Major Depressive Disorder (MDD), BD & CVD. People with severe mental illness (SMI) have an increased risk of mortality associated with physical illness, with the commonest cause of death being Cardiovascular Disease (CVD) (Brown et al., 2009; Brown et al., 2005; Brown et al., 1997). There is emerging evidence that modifiable cardiovascular risk factors are also increased in patients with bipolar disorders and in those with a history of depression or taking drugs to treat depression (Brown et al., 2009; Brown et al., 2005; Angst et al., 2002).

Stroke is the 5th leading cause of death in the US, with one person dying every 4 minutes as a result and for black people. Stroke is also the 3rd leading cause of death (MNT, 2016). A stroke is a serious, life-threatening medical condition that occurs when the blood supply to part of the brain is cut off (NHS, 2016). Approximately 800,000 people have a stroke each year; about one every 40 seconds (MNT, 2016). Stroke occurred among 2.97% of patients with bipolar disorder and 1.50% of patients undergoing appendectomy between 1998 and 2003 (Lin et al., 2007). As stroke is one of the CVD it is also associated with SZ, BD and CHD.

The above discussion on schizophrenia, bipolar disorder, coronary heart disease & stroke, can conclude

that they may be directly or indirectly associated with each other. They are also genetically associated with each other. So they have common genes by which they are inter-related. The common genes will be maintained a gene regulatory network path. This research investigates the common genes as well as common gene regulatory path among bipolar disorder and associate diseases using R.

2 Background

This section is design to discuss prior studies related to this research. Different kinds of bioinformatics tools were used to show the interaction or to find the common genes. The article of Klingstrom and Plwczynski (2010) provided information about the bio-informatics tool that can be used to show interaction, finding common genes and represent the PPI network among genes as well as proteins. About half of the individuals diagnosed with bipolar disorder suffer from distorted experiences of reality, known as hallucinations and delusions (Kerner B, 2014) that are also the symptoms of schizophrenia. Baune et al. (2006) showed the association between major depression, bipolar disorder & cardiovascular diseases. In 2016, the detail information about associations of schizophrenia and psychotic bipolar disorder for default mode network were provided on the article by Medaa SA. The connectivity between bipolar disorder and schizophrenia has also shown on the article of Thorup et al. (2015). This cohort study describes the Danish High Risk and Resilience Study on 7-year-old children, who were born to parents with either schizophrenia, bipolar disorder or neither of the two diagnoses. Ayalew et al. (2012) explained an overlap between top candidate genes for schizophrenia and candidate genes for anxiety and bipolar disorder.

Depression, anxiety disorders, schizophrenia and bipolar disorder (BD) have all been identified as risk factors for the onset and progression of cardiovascular disease (CVD) (Sowden and Huffman, 2009). Environmental exposures, including infections, nutritional deficits, and neurotoxins, are known causes of neuropsychiatric disorders, and are potent disruptors of brain development, which has been proposed to play a major role in the etiology of schizophrenia (Brown et al., 2005a). Bipolar I disorder, which occurs in approximately 1% of the general population is significantly more prevalent in patients with cardiac disease (Baune et al. 2006). 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). MRI studies of first episode cases of schizophrenia have revealed morphologic brain abnormalities, including ventricular enlargement (Lawrie and Abukmeil, 1998; Vita et al., 2006), decreased hippocampal volume (Bogerts et al., 1990; Nelson et al., 1998), and an increased prevalence of cavum septum pellucidum (CSP) (Degreef et al., 1992; DeLisi et al., 1993; Nopoulos et al., 1997). Patients with BD are up to twice as likely to die from cardiovascular causes as their counterparts in the general population (Osby et al. 2001).

It is said that the heart and mind are intimately linked.The paper (Cardno and Owen, 2014) has shown strong evidence for partial overlap of genetic influences on schizophrenia and bipolar disorder, with a genetic correlation of around 0.6. Now-a-days a leading cause of death is coronary heart disease (CHD). The article (Joehanes et al., 2013) mentioned that the estimated direct cost of CHD in 2010 was $272.5 billion and it is projected to reach $818 billion by 2030. The article (Goldstein BI et al., 2015) described compelling evidence regarding excessive and premature cardiovascular disease (CVD) among adults with major depressive disorder (MDD) and bipolar disorder (BD). The paper also disclosed the recent epidemiological studies in the United States, where the prevalence of CVD among adults with MDD was nearly 3-fold greater than among adults without mood disorders, and adults with CVD and MDD were ≈7.5 years younger than adults with CVD who did not have mood disorders. Approximately 38% of all individuals who suffer a heart attack will die from it (Dorn and Cresci, 2009).

In 2011, Fiedorowicza JG et al. predicted in their article that increased obesity in mood disorders,

particularly bipolar disorder, could also explain the links with cardiovascular diseases and mortality. Epidemiological studies have consistently shown excess CVD mortality in patients with schizophrenia, bipolar disorder and depression (De Hert et al., 2009). Baune (2006) insisted that individuals with chronic heart diseases or stroke have a significantly increased incidence and prevalence of affective disorders.

From the above discussion, it is nicely visualized that schizophrenia and bipolar disorder have a genetically relationship. Besides this, there are several diseases like coronary heart disease and stroke whose have also association with bipolar disorder. But is any common pathway shared by bipolar disorder and associated diseases? It is important to assess the interactions among gene-gene network. PPI network is helpful for this purpose. The research work employed a system in bioinformatics approach for developing the common genes network and predicts to create a PPI network for SZ, BD, CHD and Stroke. It will be also a nice achievement if any programming language can help to investigate the desired goal. This research paper first time investigates the common gene lists and gene regulatory network for bipolar disorder and associated diseases using programming language R.

3 Proposed Methodology

Several steps are performed in this research to reach the desire goal. Fig. 1 shows the step by step graphical representation of the research methodology. The procedure with code using R is briefly described to find out the common genes, to create a PPI network, and also a random network. And each of the steps is also reported below in the following subsections through 3.1 to 3.9 respectively.

3.1 Gene collection

The NCBI (National Center for Biotechnology Information) is freely accessible and downloadable on-line gene database. It is also an important resource for bioinformatics tools and services. Based on the nature of diverse data, they are stored in different databases. For example-Gene is searchable database storage for different kinds of genes, focusing on genomes. PubMed, GENE Bank & OMIM databases are also used for some perspectives. For this research project genes associated with Bipolar Disorder, Schizophrenia, Coronary Heart Disease and Stroke diseases are collected from NCBI Gene database using R. Data collecting code is shown in Fig. 2 respectively.

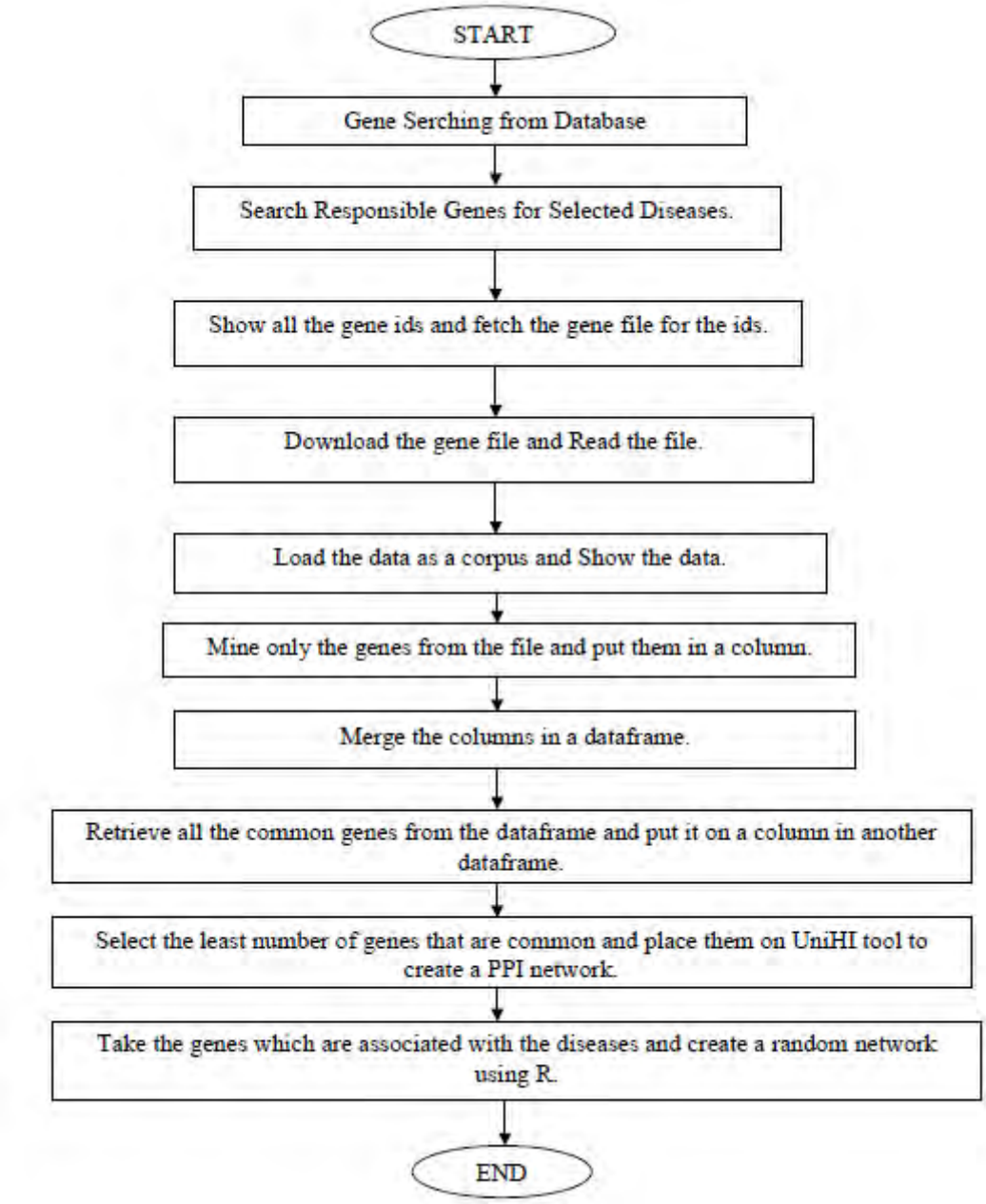

Fig. 1 Flowchart of proposed methodology.

```
1. library(rentrez)
2. entrez_dbs()
3. entrez_db_searchable("gene")
4. r_search<- entrez_search(db="gene", term="(disease_name[ALL])",retmax=700)
5. r_search
6. r_search$ids
7. r_seqs<-entrez_fetch(db="gene", id=r_search$ids, rettype="txt",retmode="text")
8. r_seqs
```

Fig. 2 Gene collection procedure using R.

3.2 Preprocessing and filtering

In the previous step all the genes associated with BD, Schizophrenia, CHD, Stroke are collected. But current step only select or collect those genes that are only for Homo sapiens. So collected data is needed to modify that is called here preprocessing. That is, all the collected genes are filtered and only the genes which are responsible for human diseases are kept. Use the same R query as Fig. 2 but modify r_search by adding filter Homo sapiens that is shown in Fig. 3.

```
1. r_search<-entrez_search(db="gene",term="(disease_name[ALL]) AND homo
   sapiens[ORGN]", retmax=700)
2. r_search
```

Fig. 3 Gene filtering procedure using R.

3.3 Gene sorting using R

The genes are downloaded in the increasing order by their weight using R. For each of the four diseases, genes are downloaded separately and stored in a text file. Fig. 4 displays the sorting technique of genes using R.

```
1. r_seqs1<-write(r_seqs, file="file_name.txt")
```

Fig. 4 Gene sorting procedure using R.

3.4 Gene linkage using R

This step is to identify the interrelated genes among diseases. 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. To get the linkage among diseases write the diseases name on the term in the search query as shown in Fig. 5.

```
1. r_search<-entrez_search(db="gene",term="(bipolar disorder[ALL] AND
   schizophrenia[ALL])", retmax=700)
2. r_search
3. r_search$ids
4. r_seqs<-entrez_fetch(db="gene",id=r_search$ids,rettype="txt",retmode="text")
```

Fig. 5 Gene linkage procedure using R.

3.5 Gene mining using R

Data mining technique is used mainly for making appropriate data. Gene mining is one of the most important parts of this research because any kind of mistake can discard an important gene that results wrong output. On the other hand larger number of gene can make the result complex. From the sorted linkage gene files only the genes are mined. After that interrelated genes are identified and sorted. Take the interrelated top 100 genes and saved them in a text file using EXPASY database and read them as a dataframe. Fig. 6 represents the gene mining, identifying and sorting procedure in R.

```
1. library(tm)
2. text <- readLines(filePath)
3. docs <- Corpus(VectorSource(text))
4. inspect(docs)
5. docs <-docs[-(1:3)]
6. docs <- tm_map(docs, removeWords, c("dd", "dt", "class", "dl","desig","details"))
7. docs <- tm_map(docs, removePunctuation)
8. df<-head(docs,100)
9. inspect(df)
10. //copy df into a text file
11. df<read.table("textfile_name.txt",fill=TRUE,header = FALSE)
12. df
13. df <-df[,-(3:227)] //delete column 3 to rest
```

Fig. 6 Gene mining procedure using R.

3.6 Display the mined genes on a data frame using R

Mining of associated genes is the critical part in this research. Fig. 7 depicts the mined genes on a dataframe called gene_DB1.

3.7 Find and display the common genes

At first the common genes among all are collected. Then the common genes from top 100 and top 50 interrelated genes are searched and collected using R and compared with EXPASY database. These common genes are displayed as 3 columns in another dataframe called comparison_DB as shown in Fig.8.

3.8 Create PPI network

Unified Human Interactome or UniHI is a tool for visualization of PPI or human molecular interaction network. It is now a very popular reliable bioinformatics tool, which is user friendly and very easy to use. In Bioinformatics research protein-protein interaction or PPI network plays an important role. UniHI tools are used here to create PPI network.

3.9 Create a random network using R

From the PPI network 7 genes are found responsible for our four diseases. A random network is created with these genes Using R and shown in Fig. 9.

```
1. bipolar_disorder<-c("TNF","EGFR","MTHFR","IL6","ACE","ESR1","SLC6A4","ABCB1","CRP"
   ,"NOS3","IL1B","AKT1","BDNF","COMT","TLR4","CXCL8","CTNNB1","GSTM1","IGF1","B
   CL2","IFNG","BRCA2","GSTT1","MTOR","PON1","CCL2","CTLA4","CYP2D6","DRD2","KIT
   ","NR3C1","IL18","NOD2","VWF","PRNP","MBL2","GSK3B","SLC6A3","IL1RN","APOB","F
   MR1","DRD4","HTR2A","FTO","FGF2","TCF7L2","HTT","MAOA","TLR9","BCL2L1","TNFR
   SF1A","FGFR2","ALDH2","CREBBP","PRKCD","DMD","NRG1","HSPA5","CREB1","GNB3",
   "FCGR2A","ADM","TNFRSF1B","FYN","S100B","FOXO3","P2RX7","HTR1A","FCGR3A","A
   NXA2","NOS1","C9orf72","DRD3","CBS","UCP2","NCAM1","IL6R","MYD88","NGF","TPH2"
   ,"CHRNA7","CACNA1C","PPARD","HTR2C","NTRK2","ANXA1","HLAE","MTRR","WWOX
   ","CRHR1","CDH2","DISC1","DRD1","GRIN2B","GFAP","KAT2B","TPH1","GCH1","SLC6A2
   ","DPYD")
2. schizophrenia<-c("TP53","TNF","APOE","EGFR","VEGFA","MTHFR","IL6","TGFB1","ACE","
   ESR1","SLC6A4","IL10","HLADRB1","ADIPOQ","NFKB1","ABCB1","MMP9","PTGS2","AR"
   ,"CRP","NOS3","IL1B","AKT1","BDNF","COMT","PPARG","HLAB","CXCL8","CTNNB1","G
   STM1","SNCA","IGF1","BCL2","LEP","MAPT","IFNG","CCR5","GSTT1","MTOR","PON1","
   HLAA","GSTP1","CCL2","HLADQB1","HFE","SOD1","CTLA4","CYP2D6","FAS","DRD2","T
   LR2","ATM","XRCC1","CXCL12","ICAM1","NR3C1","ESR2","MAPK14","HMOX1","IL18","I
   L17A","GHRL","IL4","RELA","APOA1","ITGB3","VWF","MET","PRNP","CASP3","HLAG","
   SOD2","LRRK2","GSK3B","MAPK3","SLC6A3","IL1RN","SIRT1","AGER","JUN","APC","SP
   1","HLAC","CYP3A4","FMR1","DRD4","HTR2A","LPL","IL1A","EP300","UGT1A1","ABCA1
   ","KCNH2","FTO","TTR","IL2","PRKCA","TCF7L2","MECP2","HSP90AA1")
3. coronary_heart_disease<-c("TNF","APOE","VEGFA","MTHFR","IL6","TGFB1","ACE","ESR1",
   "SLC6A4","IL10","HIF1A","ADIPOQ","ABCB1","MMP9","AR","CRP","NOS3","CDKN2A","I
   L1B","PPARG","CXCL8","GSTM1","TERT","IGF1","LEP","IFNG","MYC","CCR5","MMP2","
   GSTT1","SERPINE1","PON1","CCL2","HFE","NPPB","CXCL12","ICAM1","ESR2","CYP2C19
   ","AGT","HMOX1","IL18","GHRL","IL4","APOA1","ITGB3","VWF","CYP1A1","INS","EDN1"
   ,"AGTR1","AGER","APOB","CFH","LDLR","ABCG2","HBA1","CYP3A4","HTR2A","CD14","
   KDR","LPL","IL1A","ABCA1","FTO","FGF2","CETP","MIF","TCF7L2","IGFBP3","MMP3","L
   EPR","OPRM1","ALB","ALDH2","RETN","CD40LG","C3","NQO1","CXCL10","F7","PPARA",
   "CNR1","CLU","LTA","MMP14","VCAM1","APOA5","FCGR2A","HNF1A","PROC","CD36","
   PCSK9","IL15","SELE","HP","TNNT2","CCR2","VKORC1","SELP")
4. stroke<-c("TNF","EGFR","MTHFR","IL6","ACE","ESR1","SLC6A4","ABCB1","CRP","NOS3",
   "IL1B","AKT1","BDNF","COMT","TLR4","CXCL8","CTNNB1","GSTM1","IGF1","BCL2","IF
   NG","BRCA2","GSTT1","MTOR","PON1","CCL2","CTLA4","CYP2D6","DRD2","KIT","NR3C
   1","IL18","NOD2","VWF","PRNP","MBL2","GSK3B","SLC6A3","IL1RN","APOB","FMR1","D
   RD4","HTR2A","FTO","FGF2","TCF7L2","HTT","MAOA","TLR9","BCL2L1","TNFRSF1A","F
   GFR2","ALDH2","CREBBP","PRKCD","DMD","NRG1","HSPA5","CREB1","GNB3","FCGR2A
   ","ADM","TNFRSF1B","FYN","S100B","FOXO3","P2RX7","HTR1A","FCGR3A","ANXA2","N
   OS1","C9orf72","DRD3","CBS","UCP2","NCAM1","IL6R","MYD88","NGF","TPH2","CHRNA
   7","CACNA1C","PPARD","HTR2C","NTRK2","ANXA1","HLAE","MTRR","WWOX","CRHR1
   ","CDH2","DISC1","DRD1","GRIN2B","GFAP","KAT2B","TPH1","GCH1","SLC6A2","DPYD")
5. geneDB1<-data.frame(bipolar_disorder,schizophrenia,coronary_heart_disease,stroke)
6. geneDB1
```

Fig. 7 Representation of mined genes on a data frame using R.

```
1. common_from_allDB<-c("TNF","MTHFR","IL6","ACE","ESR1","SLC6A4","ABCB1","CRP
   ","NOS3","IL1B","CXCL8","GSTM1","IGF1","IFNG","GSTT1","PON1","CCL2","IL18","V
   WF","HTR2A","FTO","TCF7L2","ALDH2","FCGR2A","NOS1","IL6R","HTR2C","MTRR",
   "FADS2")
2. commom_from100<-Reduce(intersect,list(geneDB1$bipolar_disorder,geneDB1$schizophrenia
   ,geneDB1$c oronary_heart_disease,geneDB1$stroke))
3. geneDB2<-head(geneDB1,50)
4. geneDB2
5. commom_from50<-Reduce(intersect,list(geneDB2$bipolar_disorder,geneDB2$schizophrenia,
   geneDB2$coronary_heart_disease,geneDB2$stroke))
6. comparison_DB<-data.frame(common_from_allDB)
7. comparison_DB$commom_from100<c(commom_from100,rep("",nrow(comparison_DB)-leng
   th(commom_from100)))
8. comparison_DB$commom_from50<-c(commom_from50,rep("",nrow(comparison_DB)-length
   (commom_from50)))
9. comparison_DB
```

Fig. 8 Common genes identification procedure using R

```
1. library(GGally)
2. library(ggnet)
3. library(network)
4. library(sna)
5. library(ggplot2)
6. library(grid)
7. net = rgraph(7, mode = "graph", tprob = 0.5)
8. net = network(net, directed = FALSE)
9. ggnet2(net,mode="circle",size=17,label=c("TNF","IL6","ESR1","IL1B","CXCL8","IFNG","C
   CL2"),color=rep(c("red","darksalmon","darkolivegreen","palevioletred","tomato","goldenrod",
   "black")),label.color="white",edge.color="blue") + theme(panel.background = element_rect(fill
   = "skyblue"))
```

Fig. 9 Random network procedure using R.

4 Results and Discussion

A disease is a particular abnormal condition caused by a single gene or more than one gene that is directly or indirectly related to each other. The trustable gene database and Bioinformatics tools make it possible to identify each of the genes associated with diseases and their correlation. The collected mined genes are used to create database that shows the resulting common genes. The architectural view of the database that is created using R is given below. The three level architectural designs in Fig. 10 shows three form of schemas External, Logical and Internal which describes how data is stored, processed and presented.

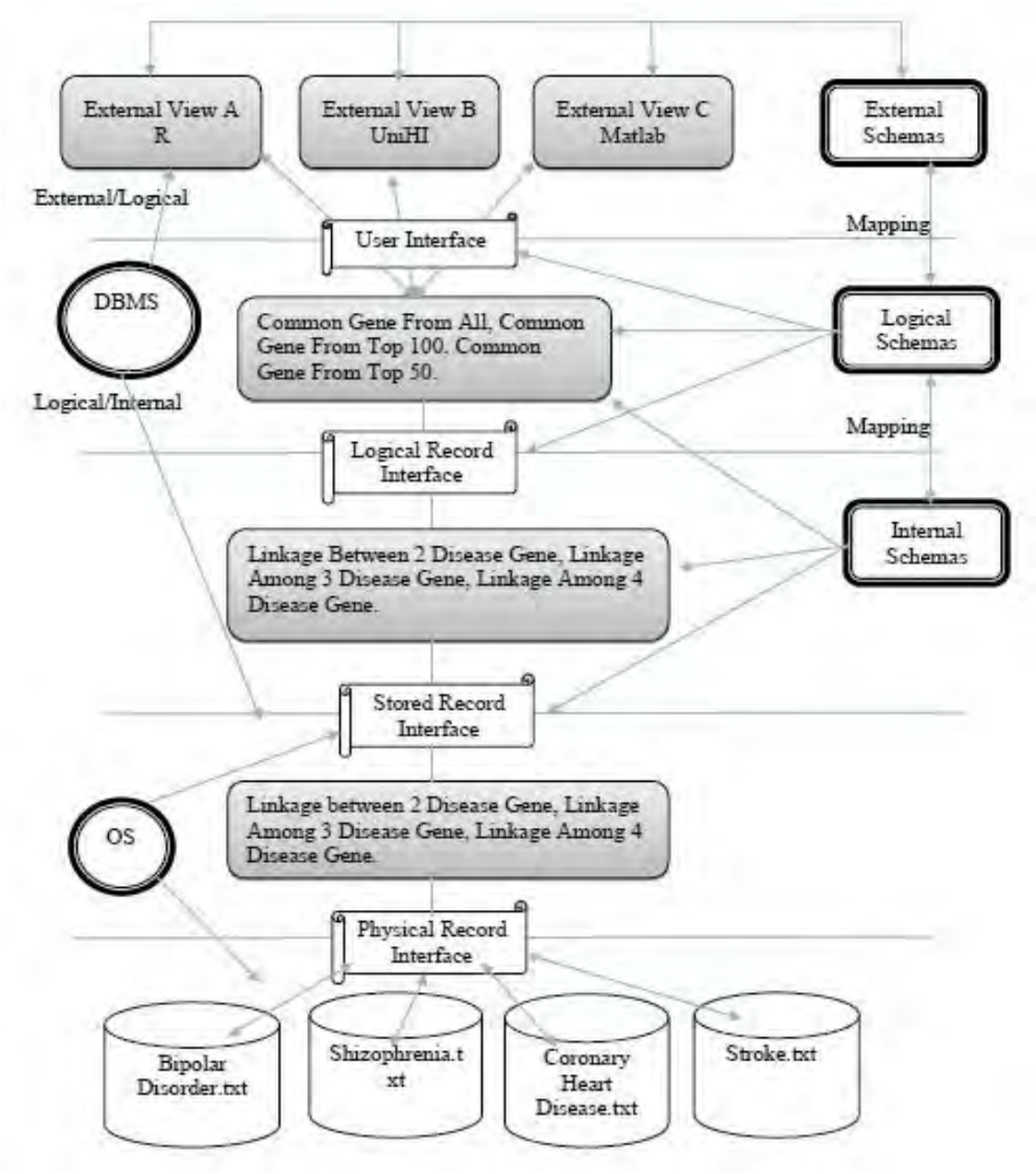

Fig. 10 Architectural view of database**.**

4.1 Gene collection, preprocessing & filtering, gene sorting

The collected responsible genes without preprocessing & filtering are estimated as 640 for BD, 2555 for Schizophrenia, 294 for CHD and 1028 for Stroke. 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. The identified genes are then sorted increasing order by their weight. Gene optimization at each step can be shown by using a pie chart as given in Fig. 11. The pie chart shows the percentage of genes used at each step. Table 1 shows resultant responsible genes at each step.

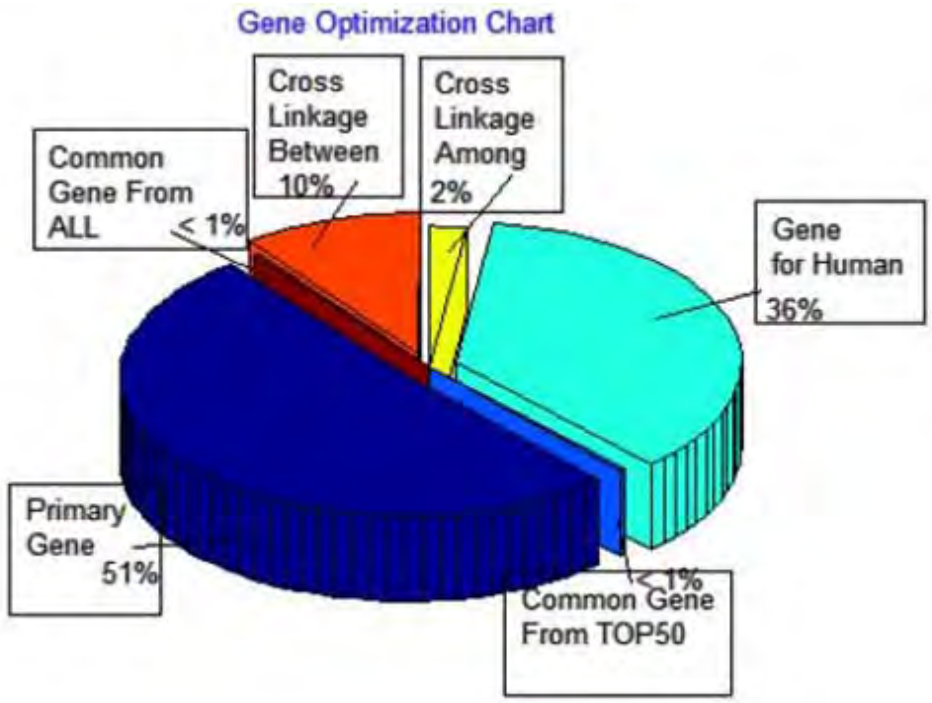

Fig. 11 Scenario of gene optimization by pie chart.

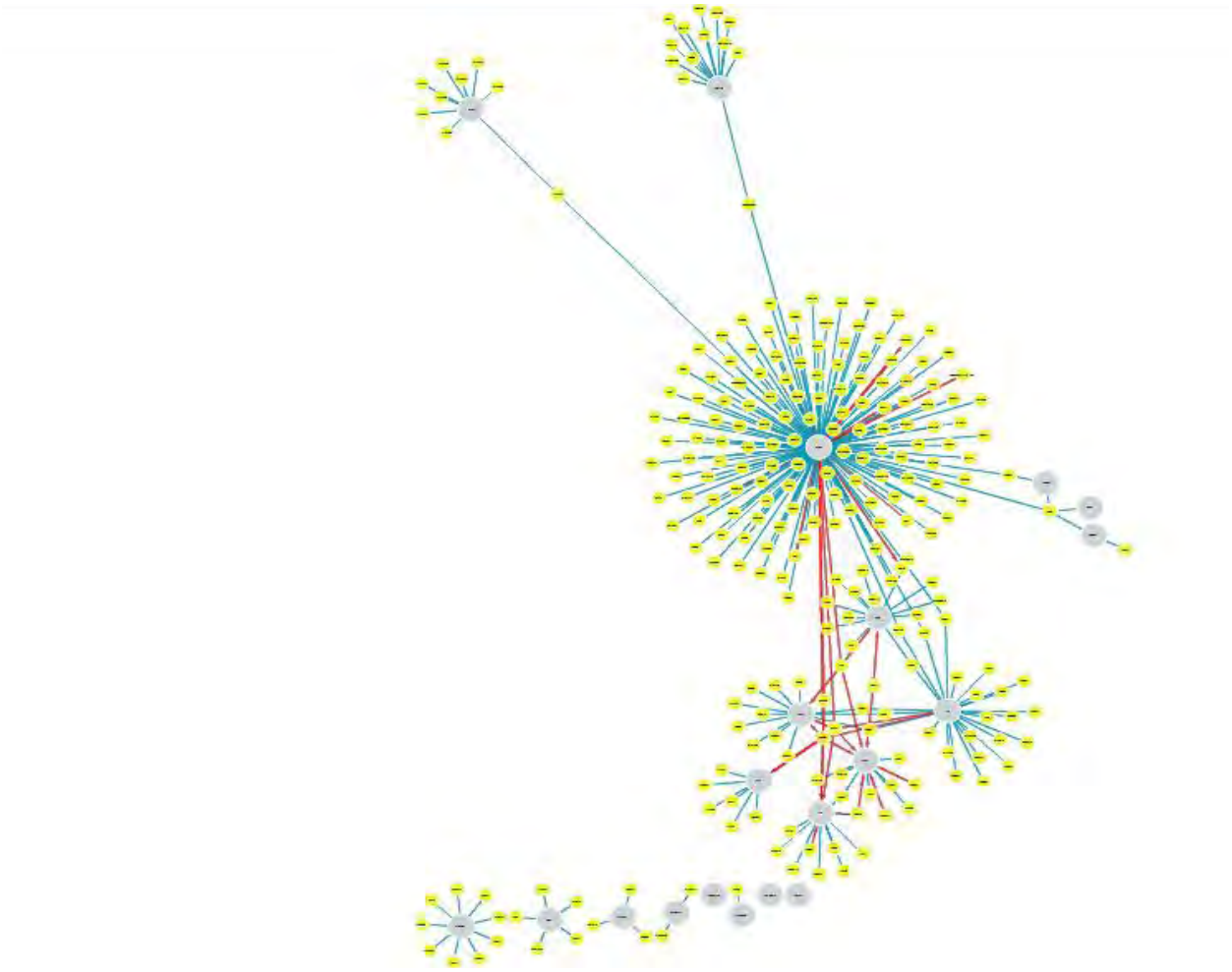

Fig. 12 PPI network with 17 common gene of selected diseases.

Table 1 Collected genes for specific diseases.

Name of Disease	Primary Number of Gene Collection	Number of Corresponding Gene for Human
Bipolar Disorder(BD)	640	619
Schizophrenia	2555	1612
Coronary Heart Disease(CHD)	294	288
Stroke	1028	638

4.2 Gene linkage, gene mining & common gene finding

The linkage between the investigated genes is identified. The resultant cross linkage genes are shown in Table 2. After gene linkage there are 29 common genes. These 29 common genes are mined using data mining technique. Now compare with EXPASY Database genes are collected which results 17 common genes. The resultant 17 common genes are TNF, MTHFR, IL6, ACE, ESR1, SLC6A4, ABCB1, CRP, NOS3, IL1B, CXCL8, GSTM1, IGF1, IFNG, GSTT1, PON1 and CCL2.

4.3 Common regulatory pathway or PPI network

The 17 common genes are then used to create a PPI network using UniHI tool. The PPI network represents the interaction among the genes and hub protein, some of which are connected directly and some are connected indirectly. The resultant network is shown in Fig. 12. From the Fig. 12, it is clearly visualized that only 7 genes are directly interconnected with each other in PPI network. Now directly connected 7 genes PPI network is noticed in Fig. 13.

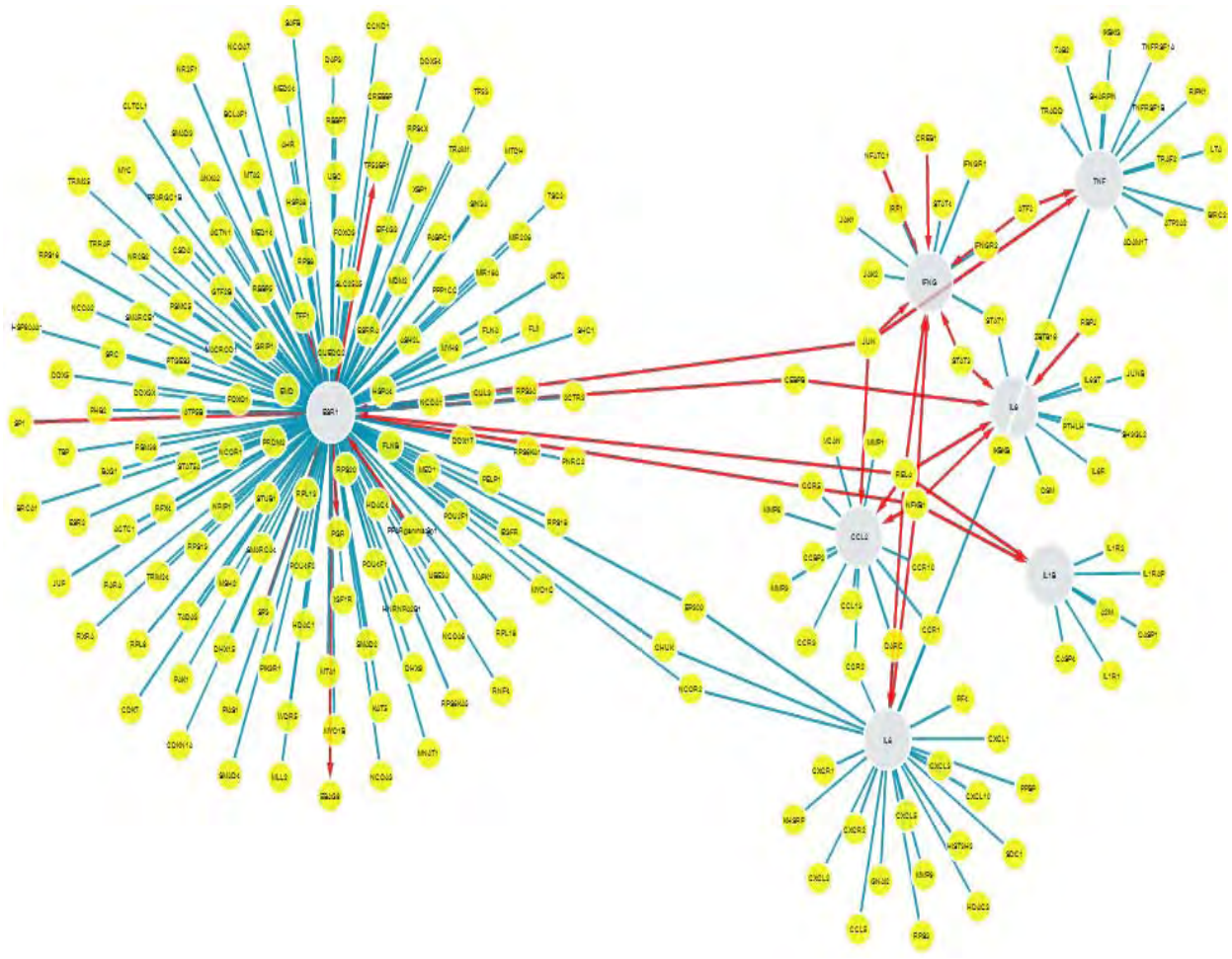

Fig. 13 PPI Network with finally selected 7 common genes.

Table 2 Cross linkage gene chart.

Cross Linkage	No. of Gene	Cross Linkage	No. of Gene
BD & Schizophrenia	364	BD, Schizophrenia & CHD	34
BD & CHD	39	BD, Schizophrenia & Stroke	54
BD & Stroke	66	BD, CHD & Stroke	33
Schizophrenia & CHD	96	Schizophrenia, CHD & Stroke	67
CHD & Stroke	153	Schizophrenia, CHD, BD & Stroke	29
Schizophrenia & Stroke	171		

4.4 Creating random network using R

Finally, using investigated 7 genes random network is created through R. Fig. 14 depicts the scenario of random network of finally selected 7 common associated genes.

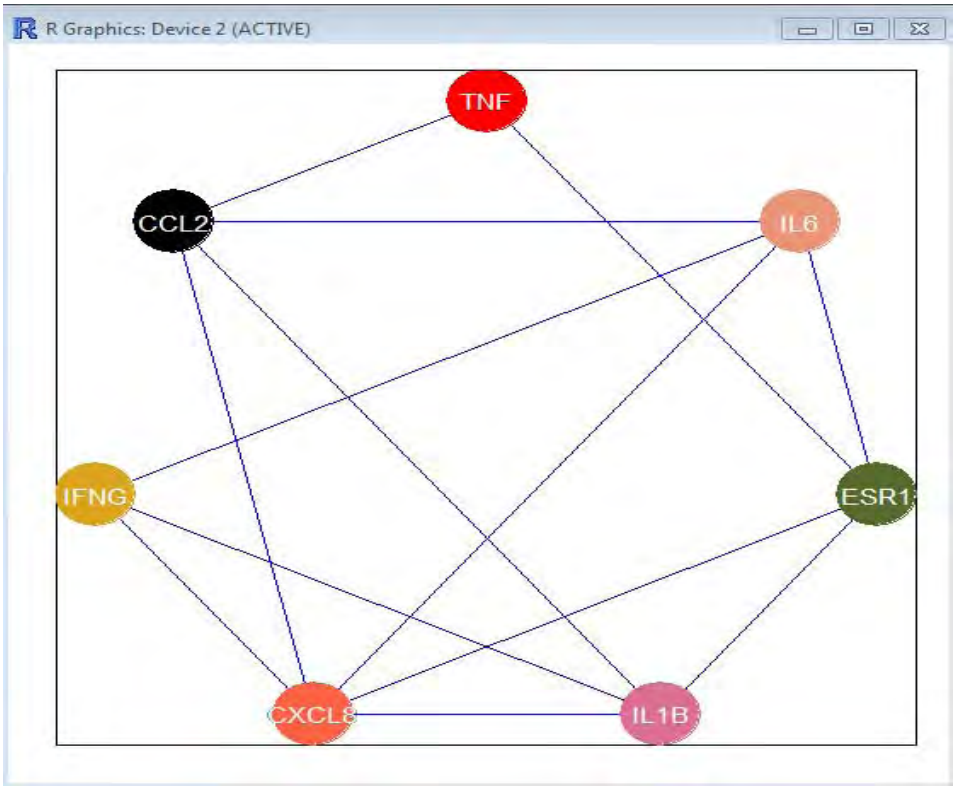

Fig. 14 Random network using R.

5 Conclusions

To design a drug for a disease, it is necessary to know the affected genes associated with the disease. And to study with more than one disease, it is important to know the linkage of the genes among the associated diseases (Zhang, 2016a, 2016b; Zhang and Li, 2015). Functional cross-links studies between genes are still in their early stages. BD, Schizophrenia, CHD and stroke diseases are caused because of major depression. Findings of the interconnected genes, linkage genes and common genes among associated diseases will be helpful to analysis the diseases as well as to design drug accurately. In this research constructed a cross talking sub pathway via mapping inter-genes to PPI to show the association among the selected diseases through the genetic level. Higher-level relationship network among gene and disease has been gained by cross-talking sub pathways network. The network-based analysis depicts a rather than promising insight of a common gene

regulatory path between gene and disease. This research creates a new dimension of using R in gene regulatory network which is the main potentiality of the research work. This research is mainly helpful to understand the bipolar network among genes and to target drug design.

Abbreviation

BD=Bipolar Disorder; CHD=Coronary Heart Disease; MDD=Major Depression Disorder.

Acknowledgment

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

References

Akiskal HS, Bourgeois Ml, Angst J, et al. 2000. Re-evaluating the prevalence of and diagnostic composition within the broad clinical spectrumof bipolar disorders. Journal of Affective Disorders, 59: 5-30

Aleman A, Kahn RS, Selten JP. 2003. Sex differences in the risk of schizophrenia: evidence from metaanalysis. Archives of General Psychiatry, 60: 565-571

Angst F, Stassen HH, Clayton PJ, et al. 2002. Mortality of patients with mood disorders: follow-up over 34-38 years. Journal of Affective Disorders, 68(2-3): 167-181

Ayalew M, Le-Niculescu H, Levey DF, et al. 2012. Immediate communication: Convergent functional genomics of schizophrenia: from comprehensive understanding to genetic risk prediction. Molecular Psychiatry, 17: 887-905

Baune BT, Adrian I, Arolt V, et al. 2006. Associations between major depression, bipolar disorders, dysthymia and cardiovascular diseases in the general adult population. Psychotherapy and Psychosomatics, 75: 319-326

Belmaker RH. 2004. Bipolar disorder. The New England Journal of Medicine, 351: 476-486

Bogerts B, Ashtari M, Degreef G, et al. 1990. Reduced temporal limbic structure volumes on magnetic resonance images in first episode schizophrenia. Psychiatry Research, 35: 1-13

Brown AD, Barton DA, Lambert GW. 2009. Cardiovascular abnormalities in patients with major depressive disorder: autonomic mechanisms and implications for treatment. CNS Drugs, 23(7): 583-602

Brown AS, Bresnahan M, Susser ES. 2005a. Schizophrenia: Environmental Epidemiology. Comprehensive Textbook of Psychiatry. 1371-1380, Lippincott, Williams, Wilkins; Baltimore, MD, USA

Brown LC, Majumdar SR, Newman SC, et al. 2005. History of depression increases risk of type 2 diabetes in younger adults. Diabetes Care, 28: 1063-1067

Brown S. 1997. Excess mortality of schizophrenia. A meta-analysis. British Journal of Psychiatry, 171: 502-508

Cardno AG and Owen MJ. 2014. Genetic Relationships Between Schizophrenia, Bipolar Disorder, and SchizoaffectiveDisorder. Schizophrenia Bulletin, 40(3): 504-515

Carney CP and Jones LE. 2006. Medical comorbidity in women and men with bipolar disorders: a population-based controlledstudy. Psychosomatic Medicine, 68: 684-691

Degreef G, Bogerts B, Falkai P, et al. 1992. Increased prevalence of the cavum septum pellucidum in magnetic resonance scans and post-mortem brains of schizophrenic patients. Psychiatry Research, 45: 1-13

DeLisi LE, Tew W, Xie SH, et al. 1995. A prospective follow-up study of brain morphology and cognition in first-episode schizophrenic patients: preliminary findings. Biological psychiatry, 38(6): 349-60

Dorn GW and Cresci S. 2009. Genome-wide association studies of coronary artery diseaseand heart failure: where are we going?. Pharmacogenomics, 10(2): 213-223

Fiedorowicz JG, He J and Merikangas KR. 2011. The association between mood and anxiety disorders with vascular diseases and risk factors in a nationally-representative sample. Journal of Psychosom Research, 70(2): 145-154

Goldstein BI, Carnethon MR, Matthews KA, et al. 2015. Major Depressive Disorder and Bipolar Disorder Predispose Youth to Accelerated Atherosclerosis and Early Cardiovascular Disease: A Scientific Statement From the American Heart Association. Circulation, 132: 965-986

International Bipolar Foundation. 2014. The Application of Clinical Genetics: 7. IBF

Joehanes R, Ying S, Huan T, et al. 2013. Gene Expression Signatures of Coronary Heart Disease. Arteriosclerosis, Thrombosis, and Vascular Biology, 33(6): 1418-1426

Kerner B. 2014. Genetics of bipolar disorder. The Application of Clinical Genetics, 7: 33-42

Kessler RC, Akiskal HS, Ames M, et al. 2006. Prevalence and effects of mood disorders on work performance in a nationally representativesample of U.S. workers. American Journal of Psychiatry, 163: 1561-1568

Kilbane EJ, Gokbayrak NS, Galynker I, et al. 2009. A review of panic and suicide in bipolar disorder: does comorbidityincrease risk? Journal of Affective Disorders, 115: 1-10

Klingstrom T and Plwczynski D. 2010. Protein-protein interaction and pathway databases, a graphical review. Briefings in Bioinformatics, 12(6): 702-713

Laursen TM, Munk-Olsen T, Agerbo E, et al. 2009. Somatic hospital contacts, invasive cardiac procedures, and mortality from heart disease in patients with severe mental disorder. Archives of General Psychiatry, 66(7): 713-720

Lawrie SM, Abukmeil SS. 1998. Brain abnormality in schizophrenia: A systematic and quantitative review of volumetric magnetic resonance imaging studies. British Journal of Psychiatry, 172: 110-120

Lin HC, Tsai SY, Lee HC. 2007. Increased risk of developing stroke among patients with bipolar disorder after an acute mood episode: A six-year follow-up study. Journal of Affective Disorders, 100: 49-54

Marcus M, Yasamy MT, Ommeren MV, et al. 2016. Depression a global public health concern developed. Department of mental health and substance abuse. WHO, Switzerlands

McGrath J, Saha S, Welham J, et al. 2004. A systematic review of the incidence of schizophrenia: the distribution of rates and the influence of sex, urbanicity, migrant status and methodology. BMC Medicine, 2: 13

Medaa SA, Ruanob G, Windemuthb A, et al. 2014. Multivariate analysis reveals genetic associations of the resting default mode network in psychotic bipolar disorder and schizophrenia. PNAS, E2066-E2075

Medical News Today. 2016. MNT

Merikangas KR, Akiskal HS, Angst J, et al. 2006. Lifetime and 12- month prevalence of bipolar spectrum disorder in the NationalComorbidity Survey replication. Archives of General Psychiatry, 64: 543-552

National Health Service. 2016. NHS

National Institute of Mental Health. 2016. NIMH

Zhang WJ. 2016a. A node degree dependent random perturbation method for prediction of missing links in the network. Network Biology, 6(1): 1-11

Zhang WJ. 2016b. Network pharmacology: A further description. Network Pharmacology, 1(1): 1-14

Zhang WJ, Li X. 2015. Linear correlation analysis in finding interactions: Half of predicted interactions are undeterministic and one-third of candidate direct interactions are missed. Selforganizology, 2(3): 39-45