Journal of Entomology and Zoology Studies

Available online at www.entomoljournal.com

E-ISSN: 2320-7078
P-ISSN: 2349-6800
JEZS 2015; 3(4): 462-466

Received: 22-04-2015
Accepted: 25-05-2015

Tasnuba Jesmin
Department of Information and Communication Technology Mawlana Bhashani Science and Technology University, Santosh, Tangail-1902, Bangladesh.

Dr. Sajjad Waheed
Department of Information and Communication Technology Mawlana Bhashani Science and Technology University, Santosh, Tangail-1902, Bangladesh.

Abdullah-Al-Emran
Department of Biotechnology and Genetic Engineering, Mawlana Bhashani Science and Technology University, Santosh, Tangail-1902, Bangladesh

Correspondence:
Tasnuba Jesmin
Department of Information and Communication Technology Mawlana Bhashani Science and Technology University, Santosh, Tangail-1902, Bangladesh.

Investigation of common genes for metabolic disorders: A bioinformatics approach

Tasnuba Jesmin, Sajjad Waheed, Abdullah-Al-Emran

Abstract
The failure of metabolism process is resulted by metabolic disorder and growing interest worldwide. The purpose of this research is to investigate common genes that are shared by four metabolic diseases like OB, T2D, HT and CVD. Associated genes with specific metabolic disease have been accumulated from different world recognize databases. Initially 2794, 1520, 2531 and 4713 genes are collected respectively. Collected genes are verified and mined for the establishment of gene interaction models. It reduced 91.05%, 89.74%, 92.70% and 96.22% genes respectively. Sorting technique has been applied for arranging the collected genes separately. After completion of analysis 125, 108, 121 and 110 genes have been achieved respectively. 62 common candidate genes are determined. Finally, cross linkage process discovered 10 common genes and reduced 83.87% genes for four diseases. This research will be helpful not only to construct simple common metabolic pathway but also to target drug for metabolic diseases.

Keywords: Common Genes, Data Mining, Metabolic Disorders, Metabolic Diseases, Gene Regulatory Network.

1. Introduction
Overweight and Obesity are among the five leading cause of death worldwide. The occurrence of overweight and obesity have increased rapidly and become the most common cause of death in most countries. According to the WHO, among living populations in the world more than 65% death are occurring due to overweight and obesity [1]. Obesity is the cause for overweight and therefore they are proportionally interrelated. On the other hand obesity is not a disease but leads to death for various metabolic diseases like Type 2 diabetes, Hypertension, Cardiovascular diseases etc. The death level is increasing at a dreadful level mainly among adults and children due to Obesity [2]. The articles [3, 4] show that obesity shortens the life expectancy 6-7 year's average. Another article [5] proposed that 2-4 years life expectancy is reduced for BMI level 30-35 kg/m^2, on the other hand around 10 years life expectancy is reduced for more than 40 kg/m^2 BMI level. So, it is easily seen that obesity is responsible for shortening of life expectancy and also for causing Type-2 Diabetes. Now why is type-2 diabetes awful?
Type-2 diabetes is more dangerous because its increasing level may remain in undetected for many years. It is more affective among adult persons. Last year it is founded 9% of adults 18 years and older had diabetes. Last year it was found in a study that 9% of adults (18 years and above) has diabetes. It is high, around 80% in low and middle income countries. To initiate Type-2 diabetes both lifestyle and genetic factors play a vital role [6-9]. So there is a strong correlation between Type-2 diabetes and metabolic disorder. According to Chobanian *et al*. (2003), hypertension is a chronic metabolic disorder and is common for both the developed and developing nations. Hypertension is general both for developed and developing countries. The number of hypertension patients are 333 million and 639 million for both developed and developing countries respectively [11]. Type-2 diabetes is 2.5 times higher as compare to blood pressure in developing countries in subjects with hypertension [12]. Like as Type-2 diabetes, obesity has a strong proportional correlation with hypertension [13]. From the literature review [14], it is clearly visualized that obesity and hypertension are interrelated. Different studies reported that obesity is one of the major risk factors for occurring hypertension [15, 16]. They also described vastly about hypertension with different stages. In 2012, a common metabolic pathway between hypertension and diabetes was established [17]. Through this discussion it is seen that obesity, type-2 diabetes and hypertension have association with each other. But what is about cardiovascular disease?

Morrish *et al*. 2001[18], proposed that heart disease and stroke are highly responsible for deaths in diabetic patients. In their multinational study they viewed that cardiovascular disease causes around 50% diabetes patients death. Metabolic syndrome plays an important role to develop cardiovascular disease [19, 20]. The article [21] also focused that type-2 diabetes is a main fronted leader for causing cardiovascular disease. Microvascular and macrovascular complications of diabetes are major cause of occurring of coronary heart diseases. On the other hand to initiate metabolic syndrome, obesity and type-2 diabetes are highly responsible [22]. Now cardiovascular diseases have a connection with obesity and Type-2 diabetes [23]. From the above discussion it clear that obesity, hypertension, type-2 diabetes and cardiovascular disease have association with each other. And they also biologically interconnected. Now question arises, how the analysis can be done?

For last few years, the researchers are more concerned in the biological interaction networks at the genetic levels. Major goals of those studies are to identify basic structural relationships among genes, proteins and corresponding diseases. Gene regulatory interaction network models established for analysis. There are also different types of bioinformatics tools that are helpful to represent interaction with predicted pathway among genes, proteins and corresponding diseases [14]. KEGG database is one of them. Its use and analysis are briefly described in article [24]. The UniHi is another better tool to analysis common metabolic pathway. This tool is referred by Ravi Kiran Reddy Kalathur *et al*. 2014 [25]. In 2014, the authors proposed a community based approach gene regulatory network for complete and incomplete topology using genes [26]. During the development of cells, gene regulatory networks regulate critical event [27]. New links in the gene regulatory network are identified which are responsible for developing this critical cell population [28]. Unprecedented characterization of the migratory CNC transcriptome has been also provided by them. Now it is clear that gene regulatory network is one approach to analysis biological interaction. Is there another approach to analysis?

An Nrf2 regulatory network was designed by John D. Hayes and Albena T. Dinkova-Kostova [29] in 2014. Through is network an interface between intermediary metabolism and redox can be easily gained. Another network called Parkinson's disease gene regulatory network was designed by Dusonchet J *et al*. [21] in 2014. This network has capability for finding LRRK2 gene regulatory pathway. In 2014, a gene regulatory network with metabolic pathways for integrative genomics of coronary artery disease was proposed by Makinen VP, 2014[30]. At the same year, Teichmann SA and Babu MM [31] investigated the role of gene duplication for gene network evolution. A genetic association's mode network for psychotic bipolar disorder and schizophrenia was proposed by Medaa SA *et al*. [32] at 2014. Xi-Mei Zhang *et al*. [33] designed a miRNA-TF-gene regulatory pathway in obesity in 2015. The above background study clarify that gene regulatory network is best approach to analysis biological interaction. For analysis another question comes to front how to make analyze easy?

The literature review focuses that for development of any gene regulatory network firstly we needs the list of genes that are associated with selected diseases. Generally gene interaction network is complex. The larger gene list is mainly responsible for this complexity. Removing the unnecessary genes without those biological functions make the analysis easy. Biological functions are complex interaction network among the cell's numerous constituents such as RNA, DNA, protein and other small molecules which are responsible to create it. So it is very important to access interactions among gene-protein, gene-gene and metabolic levels. This research paper investigates the common genes associated with four metabolic diseases like obesity, hypertension, type-2 diabetes and cardiovascular disease applying bioinformatics approach. If we able to find out the abnormal genes for a particular disease then it will be most helpful for easy analysis and push drug for particular disease. With the best knowledge of authors there are no research articles which are investigate the common genes among the proposed four specific metabolic diseases.

2. Materials and Methods

To find out common genes among metabolic diseases six steps were followed. Detailed description is shown in Figure-1 and full history in below subsections, step by step. The study was approved by Department of Biotechnology and Genetic Engineering Research and Ethics Committee, Mawlana Bhashani Science and Technology University.

2.1 Data Collection

Associated genes of this research are collected from PubMed due to its reliability and authentic storage. It is also freely accessible and downloadable gene database. Data also collected from OMIM database and Gene Bank data warehouse to enrich research data collection.

2.2 Data Processing

According to specific disease genes were enlisted and merged that were collected from different databases. To avoid duplication and unnecessary genomic data the collected genes were individually processed. With the help of KEGG database the collected genes data were verified.

2.3 Data Mining

To making data appropriate for analysis or application data mining technique has been applied. After mining the collected data genes on metabolic diseases were stored in UniGene data warehouse. Because it refers a cluster of genes what perform a particular function.

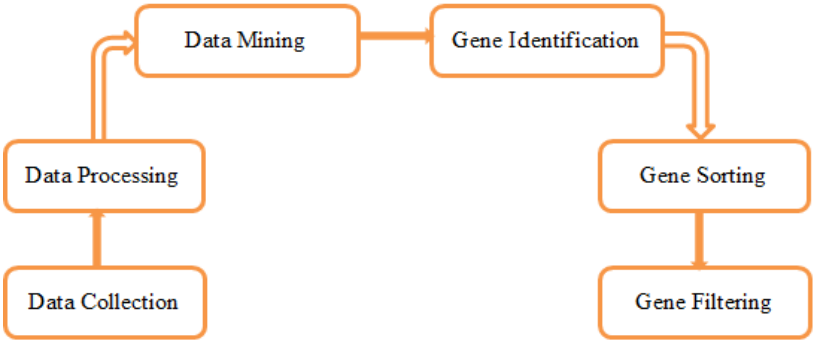

Fig 1: Working Step by Step Procedure

To making data appropriate for analysis or application data mining technique has been applied. After mining the collected data genes on metabolic diseases were stored in UniGene data warehouse. Because it refers a cluster of genes what perform a particular function

2.4 Gene Identification
EXPASY database was used to identify the interrelated genes among metabolic diseases. It performs to find out genes which are related to diseases. It also finds out the action of genes for the cause of crashing metabolism process of specific diseases.

2.5 Gene Sorting
The displacement of any important gene may give a wrong result. For setting the actual place of each gene according to their responsibility sorting technique was used in this research. The identified genes were sorted here using sorting algorithm of Taxonomy database which are internally correlated among T2D, HT, OBS and CVD diseases.

2.6 Gene Filtering
This is the crucial part of this research. To find out the common genes among selected metabolic diseases the UniHI (Unified Human Interactome) linkage network filtering technique has been applied. This technique is helpful to find out the genes those have minimum binary and also complex interaction among themselves.

3 Results and Discussions
Sequence abnormality in a single gene causes a disease. In a disease, there are the effect of more than one gene directly or indirectly even through proteomic level. Various bioinformatics tools are established to analysis the disease process newly by analyzing genes and proteins interaction among themselves with their structure also.

3.1 Gene Collection, Integration, Mining and Sorting
Responsible genes for target diseases (T2D, HT, OBS and CVD) were collected from trustable databases named PubMed, OMIM and Gene bank etc. After the collection, the lists obtained from different databases were merged and preprocessed. The results showed that primarily 2794, 1520, 2531 and 4713 responsible genes were collected for T2D, HT, OB, CVD respectively.

UniGene database has been used to perform the mining operation at preprocess level into the list of responsible genes. After the mining operation the results were visualized that the genes were reduced compared to primary collection. The mining procedure reduced 91.05%, 89.74%, 92.70% and 96.22% genes for T2D, HT, OB and CVD respectively. Therefore, the responsible genes for T2D, HT, OB and CVD were 250, 156, 185 and 178, respectively.

The genes which were responsible for anyone of the above mentioned diseases were mention here as candidate gene and selected among mining result. The result has also been justified experimentally through KEGG pathways. Resulting gene reduction in this section were 50.00%, 22.44%, 41.62% and 38.20% for T2D, HT, OB and CVD respectively. After analyzing the result, the candidate genes for T2D, HT, OB and CVD were 125,121, 108 and110 respectively. 62 common genes were identified for T2D, HT, OB and CVD diseases for applying sorting procedure.

3.2 Gene Linkage Filtering among T2D, OB, HT and CVD
The investigation procedure of molecular cross-talk among four metabolic interrelated diseases (like T2D, HT, OBS and CVD) mechanisms has drawn using cross linkage technique. The individual disease networks were also looked thoroughly at their 'hubs'. By analyzing the gene patterns and their relationship within different diseases, the genes were collected for four diseases. Gene list is created by investigating their connecting procedure and cross talk which contains all types of connections among four diseases. During this process 83.87% common genes were reduced for all four diseases. Finally ten common genes were identified i.e., APOA1, APOB, CCL2, IL6, LPL, NFKB1, NR3C1, PPARGC1A, STAT3, TNF for the common genes among four investigated metabolic diseases. Figure-2 represents the full gene reduction history step by step.

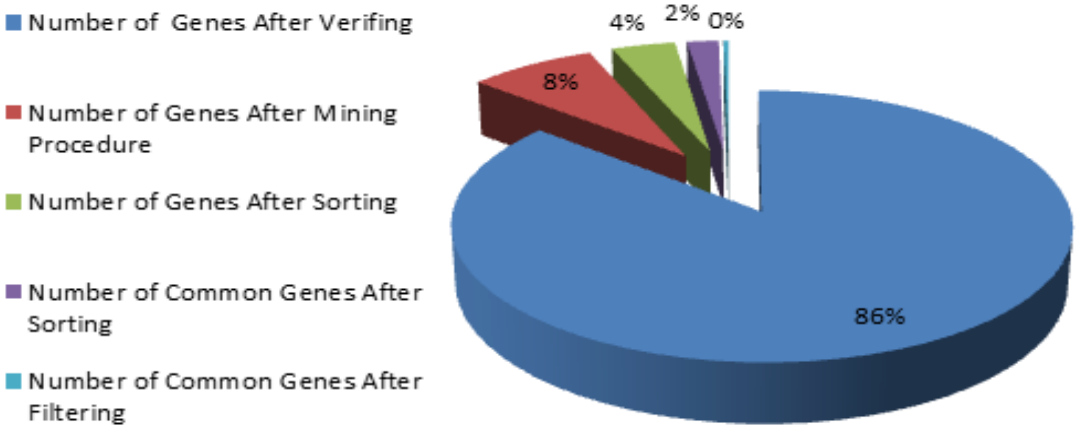

Fig 2: Full Gene Reduction History Step by Step Result

4 Conclusion
The connectivity knowledge of genes and diseases are mostly helpful to understand the genetic mechanisms of diseases. To gather connectivity knowledge firstly there is need to know about genes which are associated with specific disease. Larger numbers of associated genes make the analysis complex. So smaller number of genes investigation is the key point not only to build the simplest disease network but also to make easiest analysis. Building of a metabolic diseases network the investigation of candidate genes whose are associated with T2D, OB, HT and CVD diseases is topologically important. The genes which have perfect biological relation to the specific disease have been collected by verifying the interrelationship among those metabolic diseases. Finally the

completion of the step by step procedure was done and ten common genes are identified in the present study. This research will be helpful to construct a common metabolic pathway or metabolic diseases network and also to target drug for metabolic diseases.

Acknowledgement
Financial Support has been given by Ministry of Information and Communication Technology, Bangladesh. I, Tasnuba Jesmin, thank to all ICT ministry personnel, staffs for their supporting and special thanks to my supervisors for their valuable guidance and insight, encouragement.

Abbreviations
T2D=Type-2 Diabetes
OB=Obesity
HT= Hypertension
CVD=Cardiovascular Disease
BMI= Body Mass Index

References

1. World Health Organization. World Diabetes. Fact sheet N312, 2008.
2. Barness LA, Opitz JM, Gilbert BE. Obesity: genetic, molecular, and environmental aspects. American Journal of Medical Genetics. 2007; 143:3016-34.
3. Haslam DW, James WP. Obesity Lancet 2005; 366:1197-209.
4. Peeters A, Barendregt JJ, Willekens F, Mackenbach JP, Al Mamun A, Bonneux L. Obesity in adulthood and its consequences for life expectancy: a life-table analysis. Annals of Internal Medicine 2003; 138(1):24-32.
5. Prospective Studies Collaboration. Body-mass index and cause-specific mortality in 900 000 adults: collaborative analyses of 57 prospective studies. The Lancet 2009; 373(9669):1083-1096.
6. Ripsin CM, Kang H, Urban RJ. Management of blood glucose in type 2 diabetes mellitus. Am Fam Physician 2009; 79:29-36.
7. Ahmed K, Jesmin T, Fatima U, Moniruzzaman M, Emran AA, Rahman MZ. Intelligent and Effective Diabetes Risk Prediction System Using Data Mining. Oriental Journal of Computer Science & Technology. 2012; 5(2):215-221.
8. Riserus U, Willett WC, Hu FB. Dietary fats and prevention of type 2 diabetes. Progress in Lipid Research 2009; 48:44-51.
9. Eckel N, Muhlenbruch K, Meidtner K, Boeing H, Stefan N, Schulze MB. Characterization of metabolically unhealthy normal-weight individuals: Risk factors and their associations with type 2 diabetes. Metabolism Clinical and Experimental 2015; 64(1):862:871.
10. Chobanian AV, Bakris GL, Black HR, Cushman WC, Green LA, Izzo JL *et al*. Seventh report of the Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure. Hypertension 2003; 42(6):1206-52.
11. Kearney PM, Whelton M, Reynolds K, Muntner P, Whelton PK, He J. Global burden of hypertension: analysis of worldwide data. The Lancet 2005, 365(9455): 217-223.
12. Gress TW, Nieto FJ, Shahar E, Wofford MR, Brancati FL. Hypertension and antihypertensive therapy as risk factors for type 2 diabetes mellitus. New England Journal of Medicine. 2000; 342(13):905-912.
13. Cheung BM. This is a brief review of the overlap between hypertension and type-2 diabetes that proposes there is a spectrum ranging from hypertension without dysglycemia to type-2 diabetes without elevated blood pressure. The hypertension-diabetescontinuum. J Cardiovasc Pharmacol. 2010; 55:333-9.
14. Tomas K, Dariusz P. Protein-protein interaction and pathway databases, a graphical review. Briefings in Bioinformatics 2010; 12(6):702-713.
15. Aneja A, El-Atat F, McFarlane SI, Sowers JR. Hypertension and Obesity. The Endocrine Society 2004; 59:169-206.
16. Carretero OA, Oparil S. Essential Hypertension: Part I: Definition and Etiology. Circulation 2000; 101(1):329-335.
17. Cheung MBY, Li C. Diabetes and Hypertension: Is There a Common Metabolic Pathway? Curr Atheroscler Rep 2012; 14:160-166.
18. Morrish NJ, Wang SL, Stevens LK, Fuller JH, Keen H. Mortality and causes of death in the WHO Multinational Study of Vascular Disease in Diabetes. Diabetologia 2001; 44(2):S14-S21.
19. Norhammar A, Malmberg K, Diderhol E. Diabetes mellitus: the major risk factor in unstable coronary artery disease even after consideration of the extent of coronary artery disease and benefits of revascularization. J Am Coll Cardiol. 2004; 43:585-591.
20. Kvan E, Pettersen KI, Sandvik L, Reikvam A, Inpharm. Study Investigators. High mortality in diabetic patients with acute myocardial infarction: cardiovascular co-morbidities contribute most to the high risk. International journal of cardiology. 2007; 121(2)184-188.
21. Dusonchet J, Li H, Guillily M, Liu M, Stafa K, Troletti CD *et al*. A Parkinson's disease gene regulatory network identifies the signaling protein RGS2 as a modulator of LRRK2 activity and neuronal toxicity. Human Molecular Genetics 2014; 23(18):4887-4905.
22. Isomaa B, Henricsson M, Almgren P, Tuomi T, Taskinen MR, Groop L. The metabolic syndrome influences the risk of chronic complications in patients with type II diabetes. Diabetologia 2001; 44(9):1148-1154.
23. Highlander P, Shaw GP. Current pharmacotherapeutic concepts for the treatment of cardiovascular disease in diabetics. Ther Adv Cardiovasc Dissease 2010; 4 (1): 43–54.
24. Kanehisa M, Araki M, Goto S, Hattori M, Hirakawa M, Itoh M *et al*. KEGG for linking genomes to life and the environment.Nucleic acids research 2008, 36(suppl 1): D480-D484.
25. Kalathur RKR, Pinto JP, Hernandez-Prieto MA, Machado RS, Almeida D, Chaurasia G *et al*. UniHI 7: an enhanced database for retrieval and interactive analysis of human molecular interaction networks. Nucleic Acids Research 2014, 42: D408–D414.
26. Meyer P, Cokelaer T, Chandran D, Kim KH, Loh PR, Tucker G *et al*. Network topology and parameter estimation: from experimental design methods to gene regulatory network kinetics using a community based approach. BMC Systems Biology 2014; 8 (1): 1-13.
27. Wang S, Sengel C, Emerson MM, Cepko CL. A Gene Regulatory Network Controls the Binary Fate Decision of Rod and Bipolar Cells in the Vertebrate Retina. Developmental Cell 2014; 30 (1): 513–527.
28. Costa MS, Cabugao JT, Antoshechkin I, Spengler TS, Bronner ME. Transcriptome analysis reveals novel players in the cranial neural crest gene regulatory network. Genome Research 2014; 24 (1): 281-290.
29. Hayes JD, Kostova ATD. The Nrf2 regulatory network

provides an interface between redox and intermediary metabolism. Trends in Biochemical Sciences 2014; 39 (4): 199-218.
30. Makinen VP, Cevelek M, Meng Q, Zhang B, Zhu J, Levian C *et al*. Integrative Genomics Reveals Novel Molecular Pathways and Gene Networks for Coronary Artery Disease. PLOS Genetics 2014; 10(7):1-14.
31. Teichmann SA, Babu MM. Gene regulatory network growth by duplication. NATURE GENETICS 2014; 36 (5):492-496.
32. Meda SA, Ruano G, Windemuth A, Neil KO, Berwise C, Dunn SM *et al*. Multivariate analysis reveals genetic associations of the resting default mode network in psychotic bipolar disorder and schizophrenia. PNAS 2014; E2066-E2075.
33. Zhang XM, Guo L, Chi MH, Sun HM, Chen XW. Identification of active miRNA and transcription factor regulatory pathways in human obesity-related inflammation. BMC Bioinformatics 2015; 16(76):1-7