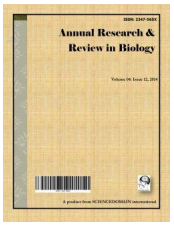


Annual Research & Review in Biology

29(5): 1-9, 2018; Article no.ARRB.45510
ISSN: 2347-565X, NLM ID: 101632869

Comparative Analysis of Some Proteins Encoded by Genes Significantly Related to Diabetes

Rajika Roy Chowdhury[1] and Malavika Bhattacharya[1*]

[1]Department of Biotechnology, Techno India University, West Bengal, EM-4, Sector-V, Salt Lake, Kolkata-700091, West Bengal, India.

Authors' contributions

This work was carried out in collaboration between both authors. Both authors read and approved the final manuscript.

Article Information

DOI: 10.9734/ARRB/2018/45510
Editor(s):
(1) Dr. Farouk El-Sabban, Department of Food Science and Nutrition, College of Life Sciences, Kuwait University, Kuwait.
(2) Dr. George Perry, Dean and Professor of Biology, University of Texas at San Antonio, USA.
Reviewers:
(1) AS Prasad, Gitam Dental Collage, India.
(2) Meral Mert, Kayseri Training and Research Hospital, Turkey.
Complete Peer review History: http://www.sciencedomain.org/review-history/27769

Original Research Article

Received 23 September 2018
Accepted 03 December 2018
Published 16 December 2018

ABSTRACT

Aims: The study was performed with the aim of understanding the role of protein structures encoded by a few of those genes which show the most significant alterations in their expression under normal versus diabetic conditions.
Study Design: The study involved identifying a few relevant genes and analysis of various components of their protein structures.
Methodology: Nine genes were shortlisted based on the extensive search of available secondary data. The structures of proteins encoded by them were generated using standard online tools. Comparative models of each of them were also generated in reference to the gene PPARγ due to its high significance in both diabetes as well as obesity, one of its predominant contributing factors.
Results: Our studies indicate that the protein structures have domains which can interact with each other as well as other signaling molecules and thereby contribute towards the transfer of information across the cells. Moreover, some of these proteins show significant overlap with the protein encoded by the gene PPARγ, indicating probable interactions between them.

**Corresponding author: E-mail: malavikab@gmail.com;*

Conclusion: These preliminary observations are indicative of probable protein-protein interactions which may contribute towards disease pathology. Further studies on interactions between these domains of various proteins may throw light on this aspect. Since diabetes incidences are increasing exponentially across the world, further detailed analysis of the individual components of the protein structures may help in obtaining a better understanding of the molecular mechanisms that are involved in this disease. This study substantiates those findings which have reported the importance of genetics in diabetes.

Keywords: Diabetes; protein-protein interactions; protein domains; non-communicable diseases.

1. INTRODUCTION

Non-communicable diseases (NCDs) are increasing their grip worldwide. They include cancers, diabetes, cardiovascular diseases and respiratory diseases (mainly asthma and chronic obstructive pulmonary disease/COPD). In the past several years, they have replaced communicable diseases as the main contributors towards morbidity and mortality [1,2]. According to a WHO report, NCDs were responsible for around 53% of disease-related deaths in India in the year 2011 [3]. This figure increased to around 62% by the year 2016 [4] indicating the pace at which this dreaded disease in engulfing people of this country. In addition, the combined problem of an exponentially growing population and inadequate healthcare facilities can pose a huge economic burden due to NCDs here. It has already been projected that incidences of NCDs will account for a loss of around $4.58 trillion for India between 2013 and 2030 [5]. Both Type 1 and Type 2 Diabetes have indicated the role of inheritance in their occurrence. While the effect of genetic inheritance is more pronounced in case of Type 1 Diabetes [6], it has also been observed that those who have a family history of Type 2 Diabetes (T2D) are more likely to suffer from it compared to those who do not have any such family history [7,8,9]. In fact, several studies have stressed on the importance of being able to identify candidate genes which can aid in predicting the susceptibility of T2D. These genes have been found to be involved in various pathways which either directly or indirectly increase risk of T2D, such as glucose metabolism, action of insulin and function of pancreatic β cell [10,11]. The importance of genetic composition of an individual in T2D is also highlighted from the findings that in case of people suffering from Maturity-Onset Diabetes of the Young (MODY), a rare type of T2D, the inheritance pattern is autosomal dominant and the inheritance spans upto three generations [12]. Moreover, mutations in distinct genes involved in pancreatic β cell function have been found to be responsible for causing at least six distinct forms of MODY [13]. All these studies indicate that genetics has a vital role in various types of diabetes. Since the roles of only a few genes have been studied in this context, further research to identify other target genes and to understand their role in diabetes pathology is essential.Thus, focused research is needed to get a better understanding of these diseases and for developing accurate and rapid prognosis/diagnosis and treatment.

The specialised area of research involving network biology has gained importance in recent past [14]. These studies not only help in deciphering the roles of various gene/protein networks, but also help in identifying novel members of various signaling pathways. Such studies, in addition, help to understand the detailed molecular mechanisms that get altered when the physiology changes from normal to diseased state [15,16]. Several reports have highlighted the significance of such studies at various levels, ranging from genes to proteins to small compounds [17,18,19].

2. METHODOLOGY

In our previous study, we have searched through available secondary data with the help of online tools such as Pubmed and Google scholar to shortlist 10 genes which have been reported to be most involved or affected by diabetes (primarily T2D). These genes include EGFR, IL1β, BDNF, ITGβ3, Pparγ, MIR21, ICAM1, FOXO3, PARP1 and Itg3 (R Roy Chowdhury and M Bhattacharya, Techno India University, West Bengal, India, personal communication). Out of these 10 genes, 1 gene encodes a microRNA (MIR210). The remaining 9 genes encode proteins. In this study we have concentrated on the structural aspects of the proteins that are encoded by these 9 genes [20]. The work was done with the help of various freely available online tools (For ExampleExPASy).

3. RESULTS

3.1 Determination of Protein Structures

As a first step, the amino acid sequences of the proteins encoded by these 9 shortlisted genes were obtained with the help of NCBI website. Table 1 summarises these findings.

In the next step, these amino acid sequences were used to generate structures of the corresponding proteins through online available resources, such as SWISSPROT (ExPASy). Table 2 shows images obtained for the various proteins generated with the help of this online tool.

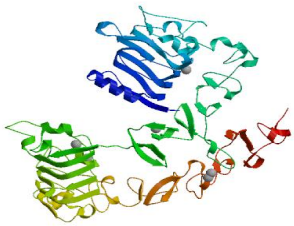
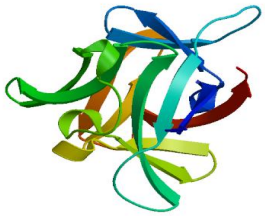
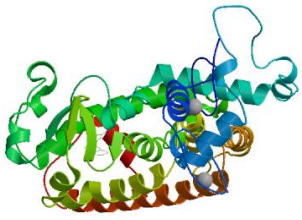
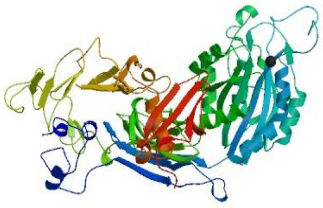
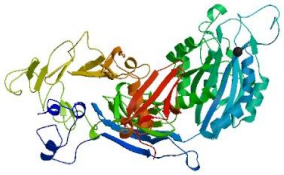
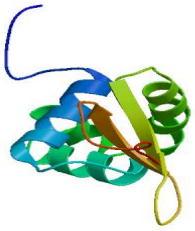
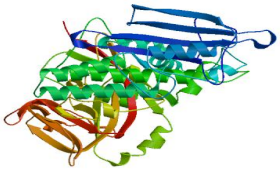
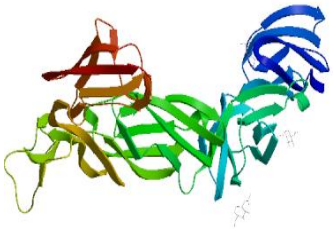
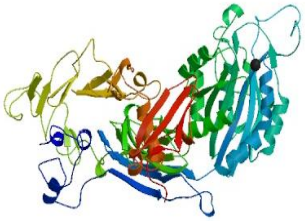
Table 1. Amino acid sequences of the proteins encoded by the shortlisted genes

Name of Genes	Amino acid sequence
- EGFR	MRPSGTAGAALLALLAALCPASRALEEKKVCQGTSNKLTQLGTFEDHFLSLQRM FNNCEVVLGNLEITYVQRNYDLSFLKTIQEVAGYVLIALNTVERIPLENLQIIRGNM YYENSYALAVLSNYDANKTGLKELPMRNLQEILHGAVRFSNNPALCNVESIQWR DIVSSDFLSNMSMDFQNHLGSCQKCDPSCPNGSCWGAGEENCQKLTKIICAQQ CSGRCRGKSPSDCCHNQCAAGCTGPRESDCLVCRKFRDEATCKDTCPPLMLY NPTTYQMDVNPEGKYSFGATCVKKCPRNYVVTDHGSCVRACGADSYEMEEDG VRKCKKCEGPCRKVCNGIGIGEFKDSLSINATNIKHFKNCTSISGDLHILPVAFRG DSFTHTPPLDPQELDILKTVKEITGFLLIQAWPENRTDLHAFENLEIIRGRTKQHG QFSLAVVSLNITSLGLRSLKEISDGDVIISGNKNLCYANTINWKKLFGTSGQKTKII SNRGENSCKATGQVCHALCSPEGCWGPEPRDCVSCRNVSRGRECVDKCNLL EGEPREFVENSECIQCHPECLPQAMNITCTGRGPDNCIQCAHYIDGPHCVKTCP AGVMGENNTLVWKYADAGHVCHLCHPNCTYGCTGPGLEGCPTNGPKIPSIATG MVGALLLLLVVALGIGLFMRRRHIVRKRTLRRLLQERELVEPLTPSGEAPNQALL RILKETEFKKIKVLGSGAFGTVYKGLWIPEGEKVKIPVAIKELREATSPKANKEILD EAYVMASVDNPHVCRLLGICLTSTVQLITQLMPFGCLLDYVREHKDNIGSQYLLN WCVQIAKGMNYLEDRRLVHRDLAARNVLVKTPQHVKITDFGLAKLLGAEEKEYH AEGGKVPIKWMALESILHRIYTHQSDVWSYGVTVWELMTFGSKPYDGIPASEIS SILEKGERLPQPPICTIDVYMIMVKCWMIDADSRPKFRELIIEFSKMARDPQRYLVI QGDERMHLPSPTDSNFYRALMDEEDMDDVVDADEYLIPQQGFFSSPSTSRTPL LSSLSATSNNSTVACIDRNGLQSCPIKEDSFLQRYSSDPTGALTEDSIDDTFLPV PEYINQSVPKRPAGSVQNPVYHNQPLNPAPSRDPHYQDPHSTAVGNPEYLNTV QPTCVNSTFDSPAHWAQKGSHQISLDNPDYQQDFFPKEAKPNGIFKGSTAENA EYLRVAPQSSEFIGA
- IL1β	MAEVPKLASEMMAYYSGNEDDLFFEADGPKQMKCSFQDLDLCPLDGGIQLRIS DHHYSKGFRQAASVVVAMDKLRKMLVPCPQTFQENDLSTFFPFIFEEEPIFFDT WDNEAYVHDAPVRSLNCTLRDSQQKSLVMSGPYELKALHLQGQDMEQQVVFS MSFVQGEESNDKIPVALGLKEKNLYLSCVLKDDKPTLQLESVDPKNYPKKKMEK RFVFNKIEINNKLEFESAQFPNWYISTSQAENMPVFLGGTKGGQDITDFTMQFV SS
- BDNF	MRARPRPRPLWATVLALGALAGVGVGGPNICTTRGVSSCQQCLAVSPMCAWC SDEALPLGSPRCDLKENLLKDNCAPESIEFPVSEARVLEDRPLSDKGSGDSSQV TQVSPQRIALRLRPDDSKNFSIQVRQVEDYPVDIYYLMDLSYSMKDDLWSIQNL GTKLATQMRKLTSNLRIGFGAFVDKPVSPYMYISPPEALENPCYDMKTTCLPMF GYKHVLTLTDQVTRFNEEVKKQSVSRNRDAPEGGFDAIMQATVCDEKIGWRND ASHLLVFTTDAKTHIALDGRLAGIVQPNDGQCH VGSDNHYSASTTMDYPSLGLMTEKLSQKNINLIFAVTENVVNLYQNYSELIPGTT VGVLSMDSSNVLQLIVDAYGKIRSKVELEVRDLPEELSLSFNATCLNNEVIPGLK SCMGLKIGDTVSFSIEAKVRGCPQEKEKSFTIKPVGFKDSLIVQVTFDCDCACQA QAEPNSHRCNNGNGTFECGVCRCGPGWLGSQCECSEEDYRPSQQDECSPRE GQPVCSQRGECLCGQCVCHSSDFGKITGKYCECDDFSCVRYKGEMCSGHGQ CSCGDCLCDSDWTGYYCNCTTRTDTCMSSNGLL CSGRGKCECGSCVCIQPGSYGDTCEKCPTCPDACTFKKECVECKKFDRGALH DENTCNRYCRDEIESVKELKDTGKDAVNCTYKNEDDCVVRFQYYEDSSGKSILY

Name of Genes	Amino acid sequence
	VVEEPECPKGPDILVVLLSVMGAILLIGLAALLIWKLLITIHDRKEFAKFEEERARA KWDTANNPLYKEATSTFTNITYRGT
- ITGβ3	MRARPRPRPLWATVLALGALAGVGVGGPNICTTRGVSSCQQCLAVSPMCAWC SDEALPLGSPRCDLKENLLKDNCAPESIEFPVSEARVLEDRPLSDKGSGDSSQV TQVSPQRIALRLRPDDSKNFSIQVRQVEDYPVDIYYLMDLSYSMKDDLWSIQNL GTKLATQMRKLTSNLRIGFGAFVDKPVSPYMYISPPEALENPCYDMKTTCLPMF GYKHVLTLTDQVTRFNEEVKKQSVSRNRDAPEGGFDAIMQATVCDEKIGWRND ASHLLVFTTDAKTHIALDGRLAGIVQPNDGQCHVGSDNHYSASTTMDYPSLGLM TEKLSQKNINLIFAVTENVVNLYQNYSELIPGTTVGVLSMDSSNVLQLIVDAYGKI RSKVELEVRDLPEELSLSFNATCLNNEVIPGLKSCMGLKIGDTVSFSIEAKVRGC PQEKEKSFTIKPVGFKDSLIVQVTFDCDCACQAQAEPNSHRCNNGNGTFECGV CRCGPGWLGSQCECSEEDYRPSQQDECSPREGQPVCSQRGECLCGQCVCH SSDFGKITGKYCECDDFSCVRYKGEMCSGHGQCSCGDCLCDSDWTGYYCNC TTRTDTCMSSNGLLCSGRGKCECGSCVCIQPGSYGDTCEKCPTCPDACTFKKE CVECKKFDRGALHDENTCNRYCRDEIESVKELKDTGKDAVNCTYKNEDDCVVR FQYYEDS
- Pparγ	MTMVDTEMPFWPTNFGISSVDLSVMEDHSHSFDIKPFTTVDFSSISTPHYEDIPF TRTDPVVADYKYDLKLQEYQSAIKVEPASPPYYSEKTQLYNKPHEEPSNSLMAIE CRVCGDKASGFHYGVHACEGCKGFFRRTIRLKLIYDRCDLNCRIHKKSRNKCQY CRFQKCLAVGMSHNAIRFGRMPQAEKEKLLAEISSDIDQLNPESADLRALAKHL YDSYIKSFPLTKAKARAILTGKTTDKSPFVIYDMNSLMMGEDKIKFKHITPLQEQS KEVAIRIFQGCQFRSVEAVQEITEYAKSIPGFVNLDLNDQVTLLKYGVHEIIYTMLA SLMNKDGVLISEGQGFMTREFLKSLRKPFGDFMEPKFEFAVKFNALELDDSDLA IFIAVIILSGNRPGLLNVKPIEDIQDNLLQALELQLKLNHPESSQLFAKLLQKMTDL RQIVTEHVQLLQVIKKTETDMSLHPLLQEIYKDLY
- ICAM1	MAPSSPRPALPALLVLLGALFPGPGNAQTSVSPSKVILPRGGSVLVTCSTSCDQ PKLLGIETPLPKKELLLPGNNRKVYELSNVQEDSQPMCYSNCPDGQSTAKTFLT VYWTPERVELAPLPSWQPVGKNLTLRCQVEGGAPRANLTVVLLRGEKELKREP AVGEPAEVTTTVLVRRDHHGANFSCRTELDLRPQGLELFENTSAPYQLQTFVLP ATPPQLVSPRVLEVDTQGTVVCSLDGLFPVSEAQVHLALGDQRLNPTVTYGND SFSAKASVSVTAEDEGTQRLTCAVILGNQSQETLQTVTIYSFPAPNVILTKPEVSE GTEVTVKCEAHPRAKVTLNGVPAQPLGPRAQLLLKATPEDNGRSFSCSATLEVA GQLIHKNQTRELRVLYGPRLDERDCPGNWTWPENSQQTPMCQAWGNPLPELK CLKDGTFPLPIGESVTVTRDLEGTYLCRARSTQGEVTRKVTVNVLSPRYEIVIITV VAAAVIMGTAGLSTYLYNRQRKIKKYRLQQAQKGTPMKPNTQATPP
- FOXO3	MAEAPASPAPLSPLEVELDPEFEPQSRPRSCTWPLQRPELQASPAKPSGETAA DSMIPEEEDDEDDEDGGGRAGSAMAIGGGGGSGTLGSGLLLEDSARVLAPGG QDPGSGPATAAGGLSGGTQALLQPQQPLPPPQPGAAGGSGQPRKCSSRRNA WGNLSYADLITRAIESSPDKRLTLSQIYEWMVRCVPYFKDKGDSNSSAGWKNSI RHNLSLHSRFMRVQNEGTGKSSWWIINPDGGKSGKAPRRRAVSMDNSNKYTK SRGRAAKKKAALQTAPESADDSPSQLSKWPGSPTSRSSDELDAWTDFRSRTN SNASTVSGRLSPIMASTELDEVQDDDAPLSPMLYSSSASLSPSVSKPCTVELPR LTDMAGTMNLNDGLTENLMDDLLDNITLPPSQPSPTGGLMQRSSSFPYTTKGS GLGSPTSSFSTVFGPSSLNSLRQSPMQTIQENKPATFSSMSHYGNQTLQDLLTS DSLSHSDVMMTQSDPLMSQASTAVSAQNSRRNVMLRNDPMMSFAAQPNQGS LVNQNLLHHQHQTQGALGGSRALSNSVSNMGLSESSSLGSAKHQQQSPVSQS MQTLSDSLSGSSLYSTSANLPVMGHEKFPSDLDLDMFNGSLECDMESIIRSELM DADGLDFNFDSLISTQNVVGLNVGNFTGAKQASSQSWVPG
- Itg3	MRARPRPRPLWATVLALGALAGVGVGGPNICTTRGVSSCQQCLAVSPMCAWC SDEALPLGSPRCDLKENLLKDNCAPESIEFPVSEARVLEDRPLSDKGSGDSSQV TQVSPQRIALRLRPDDSKNFSIQVRQVEDYPVDIYYLMDLSYSMKDDLWSIQNL GTKLATQMRKLTSNLRIGFGAFVDKPVSPYMYISPPEALENPCYDMKTTCLPMF GYKHVLTLTDQVTRFNEEVKKQSVSRNRDAPEGGFDAIMQATVCDEKIGWRND ASHLLVFTTDAKTHIALDGRLAGIVQPNDGQCH VGSDNHYSASTTMDYPSLGLMTEKLSQKNINLIFAVTENVVNLYQNYSELIPGTT

Name of Genes	Amino acid sequence
	VGVLSMDSSNVLQLIVDAYGKIRSKVELEVRDLPEELSLSFNATCLNNEVIPGLK SCMGLKIGDTVSFSIEAKVRGCPQEKEKSFTIKPVGFKDSLIVQVTFDCDCACQA QAEPNSHRCNNGNGTFECGVCRCGPGWLGSQCECSEEDYRPSQQDECSPRE GQPVCSQRGECLCGQCVCHSSDFGKITGKYCECDDFSCVRYKGEMCSGHGQ CSCGDCLDSDWTGYYCNCTTRTDTCMSSNGLL CSGRGKCECGSCVCIQPGSYGDTCEKCPTCPDACTFKKECVECKKFDRGALH DENTCNRYCRDEIESVKELKDTGKDAVNCTYKNEDDCVVRFQYYEDSSGKSILY VVEEPECPKGPDILVVLLSVMGAILLIGLAALLIWKLLITIHDRKEFAKFEEERARA KWDTANNPLYKEATSTFTNITYRGT
- PARP1	MAEAPASPAPLSPLEVELDPEFEPQSRPRSCTWPLQRPELQASPAKPSGETAA DSMIPEEEDDEDDEDGGGRAGSAMAIGGGGGSGTLGSGLLLEDSARVLAPGG QDPGSGPATAAGGLSGGTQALLQPQQPLPPPQPGAAGGSGQPRKCSSRRNA WGNLSYADLITRAIESSPDKRLTLSQIYEWMVRCVPYFKDKGDSNSSAGWKNSI RHNLSLHSRFMRVQNEGTGKSSWWIINPDGGKSGKAPRRRAVSMDNSNKYTK SRGRAAKKKAALQTAPESADDSPSQLSKWPGSPTSRSSDELDAWTDFRSRTN SNASTVSGRLSPIMASTELDEVQDDDAPLSPMLYSSSASLSPSVSKPCTVELPR LTDMAGTMNLNDGLTENLMDDLLDNITLPPSQPSPTGGLMQRSSSFPYTTKGS GLGSPTSSFNSTVFGPSSLNSLRQSPMQTIQENKPATFSSMSHYGNQTLQDLLT SDSLSHSDVMMTQSDPLMSQASTAVSAQNSRRNVMLRNDPMMSFAAQPNQG SLVNQNLLHHQHQTQGALGGSRALSNSVSNMGLSESSSLGSAKHQQQSPVSQ SMQTLSDSLSGSSLYSTSANLPVMGHEKFPSDLDLDMFNGSLECDMESIIRSEL MDADGLDFNFDSLISTQNVVGLNVGNFTGAKQASSQSWVPG

Table 2. Predicted structures of the proteins

		
EGFR	IL1β	Pparγ
		
ITGβ3	Itg3	FOXO3
		
PARP1	ICAM1	BDNF

3.2 Identification of Domains Present in These Protein Structures

Since distinct domains of the protein structure are involved with performing specific functions, the proteins sequences were analysed for presence of already characterised domains. Table 3 summarises the findings related to presence of various domains in these proteins.

3.3 Determination of Oligo State and Ligands for These Proteins

In the subsequent step, information was collected on the oligo state of these proteins. Since a lot of proteins become functional only after achieving their oligomeric conformation and also because of the fact that several proteins depend upon inter-meric interactions for their optimum functioning, such information about oligomerisation may help in understanding their mechanism of action. In addition, information was also collected about their ligands. Table 4 summarises the findings.

3.4 Determination of Overlapping of Structures of Proteins with Pparγ Encoded Protein

The peroxisome proliferator-activated receptors (PPARs) (also known as the glitazone receptor), a group of nuclear receptor proteins, are known to function as transcription factors [21,22]. Out of the three known variants, γ1 is expressed in all tissues, whereas γ2 is expressed mainly in adipose tissue and γ3 in macrophages, large intestine and white adipose tissue. Functions of PPARγ include regulation of fatty acid storage and glucose metabolism. In addition, genes activated by PPAR γ stimulate the uptake of lipid and adipogenesis by the fat cell [23]. Since PPARγ has been reported for its role in the process of adipogenesis and is thus a key player in obesity, and also because of the fact this gene is reported to be significantly affected in diabetes, it may be one of the most vital genes linking diabetes with obesity. In addition, it is increasingly being observed that obesity is one of the main causative factors which can predispose a person towards diabetes. Keeping in view these trends, the protein structures were further analysed for their possible overlap with the protein encoded by this gene. Table 5 summarises the structures that have been found to show such overlap.

4. DISCUSSION

In the present study, we have tried to understand the structural characteristics of proteins encoded by nine genes which have been reported to be most prominently related to diabetes. On analysis of the domains present in these proteins, we found that they include those involved in cellular communication (Protein kinase, nuclear receptor binding) to movement (Integrin) to gene regulation (Forkhead DNA binding). The oligomeric status of these proteins also revealed wide variance. In addition, when the structure of each of these proteins was superimposed with that encoded by PPARΓ, proteins encoded by two of them (BDNF and ITG3) could not be matched and superimposed. Superposition of the other proteins hints towards probable regions which may be rendering similar function or acting as hetero-oligomer. Further studies are needed to validate these ideas. Clinically, these findings may be able to identify compensations at genetic levels which may be responsible for protection/susceptibility of an individual towards this disease.

Table 3. List of domains present in the proteins

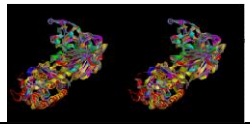
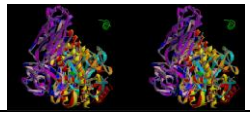
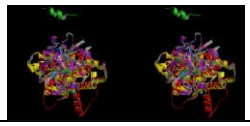
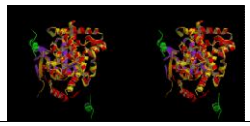
Name of genes	Name of proteins	Domain avalable
- EGFR	EPIDERMAL GROWTH FACTOR RECEPTOR	PROTEIN _KINASE_ATP PROTEIN_KINASE_TYR
- IL1β	INTERLUKIN 1, BETA	INTERLUKIN_1
- BDNF	BRAIN DERIVED NEURO FACTOR	EG-1, EG-2,
- ITGβ3	INTERGIN SUBUNIT BETA	INTEGRIN BETA
- Pparγ	PEROXISOME PROLIFTER ACTIVATED RECEPTOR GAMMA	NUCLEAR_REC_DBD_2 NR_LBD NUCLEAR_REC_DBD-1
- ICAM1	INTERCELLULAR ADHESION MOLECULES 1	IMMUNOGLOBULIN (Ig)

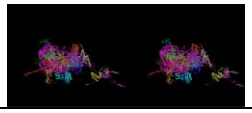
Name of genes	Name of proteins	Domain avalable
- FOXO3	FORHEAD BOX 03	FORK HEAD DNA-BINDING
- Itg3	INTERGRIN BETA 3	INTEGRIN BETA
- PARP1	POLY (ADP- RIBOSOMES POLYMERASE)	PARP_ZN_FINGER_2 BRCT PARP_ALPHA_HD PARP_CATALYTIC PARP_ZN_FINGER_1

Table 4. Summary of oligomeric state and ligands of these proteins

Name of genes	Oligo State	Ligand
- EGFR	Homodimer	1 x CA: CALCIUM ION;
- IL1β	Monomer	None
- BDNF	Monomer	1 x CA: CALCIUM ION;
- ITGβ3	Heterodimer	1 x CA: CALCIUM ION;
- Pparγ	Monomer	1 x GW9: 2-chloro-5-nitro-N-phenylbenzamide; 2 x ZN: ZINC ION;
- ICAM1	Homodimer	3 x NAG: SUGAR (N-ACETYL-D-GLUCOSAMINE);
- FOXO3	Monomer	None
- Itg3	Heterodimer	1 x CA: CALCIUM ION;
- PARP1	Monomer	None

Table 5. Summary of superposition of these protein structures with that of the protein encoded by PPARγ

Name of gene	Superposition possible / Not possible	Comparative models
- EGFR	Possible	
- IL1B	Possible	
- BDNF	Not possible	
- ITGβ	Possible	
- ICAM1	Possible	
- FOXO3	Possible	
- Itg3	Not possible	

Name of gene	Superposition possible / Not possible	Comparative models
- PARP1	Possible	

5. CONCLUSION

In this study, we focused on few structural aspects of the proteins encoded by nine genes which have been most extensively reported to be affected in diabetes or play a siginificant role in pathology of this disease. Since most of the cellular communications take place through proteins, understanding their individual components and their interactions with other proteins/ligands can provide insight into their mode of cellular communication. This information can also be helpful in designing strategies for combating diabetes through future research.

COMPETING INTERESTS

Authors have declared that no competing interests exist.

REFERENCES

1. Terzic A, Waldman S. Chronic diseases: The emerging pandemic. Clin Transl Sci. 2011;4(3):225–226.
2. Geneau R, Stuckler D, Stachenko S, McKee M, Ebrahim S, Basu S, et al. Raising the priority of preventing chronic diseases: A political process. Lancet. 2010;376(9753):1689–1698.
3. World health organization. NCD country profiles. [Online]. Available:http://www.who.int/nmh/countries/ind_en.pdf?ua=1. 2011.
4. Indian Council of Medical Research, Public Health Foundation of India, and Institute for Health Metrics and Evaluation. India: Health of the Nation's States – The India State-level Disease Burden Initiative. New Delhi, India: ICMR, PHFI, and IHME; 2017.
5. Economics of Non-Communicable Diseases in India: Executive Summary (World Economic Forum, 2014).
6. Dorman JS, Bunker CH. HLA-DQ locus of the human leukocyte antigen complex and type 1 diabetes mellitus: A HuGE review. Epidemiol Rev. 2000;22:218-227.
7. Flores JC, Hirschhorn J, Altshuler D. The inherited basis of diabetes mellitus: Implications for the genetic analysis of complex traits. Ann Rev Genomics Hum Genet. 2003;4:257-291.
8. Hansen L. Candidate genes and late-onset type 2 diabetes mellitus. Susceptibility genes or common polymorphisms? Dan Med Bull. 2003;50:320-346.
9. Gloyn AL. The search for type 2 diabetes genes. Ageing Res Rev. 2003;2:111-127.
10. Barroso I, Luan J, Middelberg RPS, et al. Candidate gene association study in type 2 diabetes indicates a role for genes involved in B-Cell function as well as insulin action. PLoS Biol. 2003;1:41-55.
11. Stumvoll M. Control of glycaemia: From molecules to men. Minkowski Lecture 2003. Diabetologia. 2004;47:770-781.
12. Stride A, Hattersley AT. Different genes, different diabetes: Lessons from maturity-onset diabetes of the young. Ann Med. 2002;34:207-216.
13. Winter WE. Newly defined genetic diabetes syndromes: Maturity onset diabetes of the young. Rev Endocr Metab Disord. 2003;4:43-51.
14. Hood L, Heath JR, Phelps ME, Lin B. Systems biology and new technologies enable predictive and preventative medicine. Science. 2004;306:640-643.
15. Bernardo DD, Thompson MJ, Gardner TS, Chobot SE, Eastwood EL, et al. Chemogenomic profiling on a genome-wide scale using reverse-engineered gene networks. National Biotechnology. 2005; 23:377-383.
16. Hwang D, Lee IY, Yoo H, Gehlenborg N, Cho JH, et al. A systems approach to prion disease. Molecular Systems Biology. 2009;5:252.
17. Li BQ, Huang T, Liu L, Cai YD, Chou KC. Identification of colorectal cancer related genes with mRMR and shortest path in protein-protein interaction network. PLoS One. 2012;7:e33393.
18. Hu LL, Chen C, Huang T, Cai YD, Chou KC. Predicting biological functions of compounds based on chemical-chemical interactions. PLoS One. 2011;6:e29491.
19. Segal E, Friedman N, Kaminski N, Regev A, Koller D. From signatures to models:

Understanding cancer using microarrays. Nature Genetics. 2005;37:38-45.
20. UniProt: A hub for protein information. Nucleic acids research. 2014;43,Database issue D204-12.
21. Greene ME, Blumberg B, McBride OW, Yi HF, Kronquist K, Kwan K, et al. Isolation of the human peroxisome proliferator activated receptor gamma cDNA: expression in hematopoietic cells and chromosomal mapping. Gene Expression. 1995;4(4-5):281–299.
22. Michalik L, Auwerx J, Berger JP, Chatterjee VK, Glass CK, Gonzalez FJ, et al. International Union of Pharmacology. LXI. Peroxisome proliferator-activated receptors. Pharmacol. Rev. 2006;58(4): 726–741.
23. Jones JR, Barrick C, Kim KA, Lindner J, Blondeau B, Fujimoto Y et al. Deletion of PPARgamma in adipose tissues of mice protects against high fat diet-induced obesity and insulin resistance. Proceedings of the National Academy of Sciences (USA). 2005;102(17):6207-6212.

Peer-review history:
The peer review history for this paper can be accessed here:
http://www.sciencedomain.org/review-history/27769