Current Genomics, 2017, *18*, 298-304

RESEARCH ARTICLE

System Level Meta-analysis of Microarray Datasets for Elucidation of Diabetes Mellitus Pathobiology

Aditya Saxena[1,2,*], Kumar Sachin[3] and Ashok Kumar Bhatia[1]

[1]*Department of Biotechnology, Institute of Applied Sciences & Humanities, GLA University, Mathura (U.P.) India;* [2]*Uttarakhand Technical University, Dehradun (U.K.) India;* [3]*Department of Biochemistry and Biotechnology, S.B.S. (PG) Institute of Biomedical Sciences & Research, Dehradun (U.K.) India*

ARTICLE HISTORY

Received: May 18, 2016
Revised: June 21, 2016
Accepted: November 16, 2016

DOI:
10.2174/1389202918666170105093339

Abstract: ***Background:*** Type 2 diabetes (T2D) is a common multi-factorial disease that is primarily accounted to ineffective insulin action in lowering blood glucose level and later escalates to impaired insulin secretion by pancreatic β cells. Deregulation in insulin signaling to its target organs is attributed to this disease phenotype. Various genome-wide microarray studies from multiple insulin responsive tissues have been conducted in past but due to inherent noise in microarray data and heterogeneity in disease etiology; reproduction of prioritized pathways/genes is very low across various studies.

Objective: In this study, we aim to identify consensus signaling and metabolic pathways through system level meta-analysis of multiple expression-sets to elucidate T2D pathobiology.

Method: We used 'R', an open source statistical environment, which is routinely used for Microarray data analysis particularly using special sets of packages available at Bioconductor. We primarily focused on gene-set analysis methods to elucidate various aspects of T2D.

Result: Literature-based evidences have shown the success of our approach in exploring various known aspects of diabetes pathophysiology.

Conclusion: Our study stressed the need to develop novel bioinformatics workflows to advance our understanding further in insulin signaling.

Keywords: Type 2 Diabetes, Insulin-signaling, Microarray, Meta-analysis, Bioconductor, Gene-set analysis.

1. INTRODUCTION

Diabetes has been considered a serious health problem; World Health Organization (WHO) has placed it among one of the four priority non-communicable diseases (NCD) and called for action on World Health Day 2016 [1]. There are an estimated 422 million adults with diabetes [1]. This complex metabolic condition depends on various factors such as high-calorie intake, sedentary lifestyle, obesity, and stress, as well as aging. Though global estimates of diabetes prevalence for type 1 and type 2 do not exist, the majority of people with diabetes are affected by type 2 diabetes which is generally attributed to insulin resistance in target organs and tissues; liver, skeletal, adipose and peripheral blood cells. T2D progresses through different disease phenotypes *viz*. Normal Glucose Tolerance (NGT) to Impaired Glucose Tolerance (IGT), Hyperinsulinemia (HI) and Insulin Resistance (IR) and eventually leads to overt diabetes. Different cultural settings clearly implicate the prevalence of T2D. There is definitely a pressing need to elucidate the mechanism underlying diabetes pathogenesis through the identification of disease-implicated signaling, and metabolic pathways at genome-wide level in order to design effective treatment.

Various genome-wide microarray studies pertaining to different disease phenotypes (IGT, IR, and T2D) have been conducted in past by various research groups and data is available at public repositories like Gene Expression Omnibus (GEO) [2] and Array Express [3] *etc*. However, due to inherent noise in microarray data as well as genetic differences in study population, disease heterogeneity, and non-uniform physiologic condition at the time of sampling and method of sampling, replication of most of the genes in different studies was very poor [4].

This situation clearly demands a need for robust bioinformatics meta-analysis of these microarray datasets, independent of geographical or other variables. In this study, we described a novel approach of the meta-analysis of microarray datasets to elucidate common signaling and metabolic pathways in an attempt to shed light on diabetes pathobiology.

2. MATERIALS AND METHODS

2.1. Microarray Data & Analysis Tools

We have downloaded twelve microarray datasets pertaining to various case-control diabetic studies from GEO. Those

*Address correspondence to this author at the Department of Biotechnology, Institute of Applied Sciences & Humanities, GLA University, Mathura P.O. Box: 281 406, Mathura, India; Tel: +91-8171630011;
E-mails: aditya.saxena@gla.ac.in; aditya.235@gmail.com

1875-5488/17 $58.00+.00

datasets containing samples from more than one disease phenotypes were further sub-divided and resulting twenty subsets were grouped tissue-wise: skeletal (11 sets), subcutaneous adipose (4 sets), peripheral (3 sets) and liver (2 sets). Selection of these datasets and their sub-division was based on a previous study of candidate gene identification in type 2 diabetes mellitus [5]. Some studies also included human subjects monitored for other diabetes-related clinical parameters *i.e.* disease family history, diet, and physical training regimen. Hence, these datasets can be considered an excellent collection of T2D gene-expression profiles due to the inclusion of disparate environmental conditions that seem vital for the elucidation of diabetes pathophysiology. Metadata of microarray datasets used in this study are shown in Table (**1**). We used 'R' an open source Statistical software environment, used for Microarray data analyses particularly using a special set of packages available at bioconductor.

2.2. R Analysis Workflow for Meta-analysis of Microarray Datasets

After acquisition of microarray datasets from GEO, we first downloaded Custom Chip Description Format (CDF) files from Brainarray [6] (www.brainarray.mbni.med.umich.edu) as due to rapid updates in genome and transcriptome databases, microarray probesets generally suffer from outdated/incorrect annotations that might affect downstream functional analysis for biological interpretation. Separate AffyBatch objects for each dataset were created. These were subsequently background corrected, normalized and summarized into ExpressionSet class objects using GeneChip Robust Multiarray Analysis (GC-RMA) method from package *gcrma* [7]. This method is capable of achieving precise gene expression values due to the use of sequence specific probe affinities.

Various algorithms for microarray statistical meta-analysis have been described in recent years [8]. In this study, the 'direct merging of expression values' approach was adapted as all studies used Affymetrix gene chip platform. Normalized values obtained in different studies were expected to have statistical coherence and appropriate for creating tissue-specific consolidated Expression Set objects.

In order to improve the analysis of differentially expressed genes (DEGs), a nonspecific filtering step was performed using package *genefilter* [9] that first estimates the overall variability across the array of each probe set and then uses it to estimate the shortest interval with half of the data. This value was chosen as variance cutoff to create the reduced ExpressionSet objects. As our Meta-analysis employed different Affymatrix chip types, to create tissue-specific consolidated Expression set, the probes must be converted to their official gene symbols and/or Entrezgene ids as later has been reported to be more effective in the computational analysis because gene symbols tend to inadvertently mutate into 'date format' during excel sheet operations [10]. We used package *org.Hs.eg.db* [11] to obtain annotations of probes based on Entrez Gene identifiers. Finally, five tissue-specific consolidated sets were obtained: Adipose, Liver, Peripheral, Skeletal-IGT, and Skeletal-NGT. Last two skeletal sets differ with each other based on T2D phenotype contrasted with either Impaired Glucose Tolerant or Normal Glucose Tolerant phenotypes. For Differential Gene Expression (DGE) analysis, an empirical Bayes moderation approach based on modified t-statistics was used that was implemented in package *limma* [12].

Downstream analysis of the obtained gene list for biological interpretation is a more challenging task than engineering a data analysis pipeline, particularly for a complex disease like type 2 diabetes. Genes in the gene list, obtained through DGE analysis can be grouped either on the basis of their numerical expression values using statistical clustering or classification. However these approaches have been criticized, due to the fact that weakly expressed genes may also contribute to disease phenotype but generally failed to qualify statistical thresholds [13]. Alternatively genes can also be grouped according to shared biological features *i.e.* involvement in metabolic/signaling pathway or Gene Ontology (GO) Consortium [14]. Therefore, gene-set based methods are biologically more meaningful than single gene methods due to the fact that even small expression changes in the members of a pathway dramatically alter the flux through the pathway. Various methods have been in use in microarray community for functional categorization of genes in the form of gene-sets that are statistically associated with expression data. We used two different types of gene set enrichment methods; over-representation analysis based on Hypergeometric or Fisher test; these methods create gene-sets using hard cut-off in the form of p-value or expression fold-change or a combination of both using package *GOstats* [15]. Another type of method for gene-set analysis is based on regression test to find gene-sets statistically associated with studied phenotypes; package *globaltest* [16] provided necessary software infrastructure to carry out this analysis. In addition to canonical Kyoto Encyclopedia of Genes and Genomes (KEGG) [17] and Gene Ontology (GO) [18] categories, *globaltest* also facilitates gene set analysis against Molecular Signature Database (MSigDB v5.1) maintained by Broad Institute [19]. MSigDB is an inventory of 13311 annotated gene sets divided into eight major collections. Though Broad Institute also offered a method for gene-set analysis termed, "Gene Set Enrichment Analysis (GSEA)", [19] that is based on Kolmogorov-Smirnov-like statistics but regression model based Global test implementation has been reported to have higher power than GSEA [20]. We downloaded complete Molecular Signatures Database (MSigDB v5.1) in XML format and package *GSEABase* [21] was used to obtain complete gene set collection in R environment. This collection was sub-setted into the Hallmark gene set (H), Curetted gene set (C2) and Motif gene set (C3). Further splitting of C2 and C3 yielded reactome gene set collection and microRNA target collection, respectively.

3. RESULTS

Insulin Receptor (IR) being an Receptor Tyrosine Kinase (RTK) may bind to various downstream signaling molecules with SH2 domain (Src Homology 2) *viz*. members of Insulin Receptor Substrate family (IRS-1, IRS-2), Growth factor receptor-bound protein 2 (Grb2), Phospholipase C γ (PLC-γ), and Phosphatidylinositol-4, 5-bisphosphate 3-kinase (PI3-K).

Our microarray meta-analysis of various insulin-responsive tissues has enriched disparate disease-related

Table 1. Microarray samples used in meta-analysis.

S. No.	GEO Series	Tissue	Affymetrix chip Type	Place of Study	Number of Samples		
					Normal Glucose Tolerant (NGT)/Healthy	Impaired Glucose Tolerant (IGT)/ Prediabetics/ Insulin Resistant (IR)	T2DM
1	GSE18732 (G1)	Skeletal	Affymetrix Gene Chip Human Genome U133 Plus 2.0 Array	Tissue Injury & Repair, Edinburgh University, Edinburgh, UK	20		20
2	GSE18732 (G2)					20	20
3	GSE19420(G1)		Affymetrix Human Genome U133 Plus 2.0 Array	Genetics and Cell biology, Maastricht University, Maastricht, Netherlands		12	10
4	GSE19420(G2)					12	8 (after 52 weeks Training)
5	GSE19420(G3)				12		10
6	GSE19420(G4)				12		8 (after 52 weeks Training)
7	GSE25462(G1)		Affymetrix Human Genome U133 Plus 2.0 Array	Research Division, Joslin Diabetes Center, Boston, MA, USA	15 (Family History-)		10
8	GSE25462(G2)					25 (Family History +)	10
9	GSE12643		Affymetrix Human Genome U95 Version 2 Array	Odense University Hospital, Odense, Denmark		10	10
10	GSE22309(G1)		Affymetrix Human Genome U95A Array	Biostatistics, University of Alabama at Birmingham, AL, USA	20		15
11	GSE22309(G2)					20	15
12	GSE26637(G1)	Adipose tissue	Affymetrix Human Genome U133 Plus 2.0 Array	Institute for Molecular Medicine Finland (FIMM), University of Helsinki, Helsinki, Finland	5 (Fasting)	5 (Fasting)	
13	GSE26637(G2)				5 (Hyperinsulinemia)	5 (Hyperinsulinemia)	
14	GSE15773 (G1)		Affymetrix Human Genome U133 Plus 2.0 Array	Department of Molecular Medicine, University of Massachusetts, MA, USA	5	4	
15	GSE15773 (G2)				5	5	
16	GSE13760	Peripheral Blood	Affymetrix Human Genome U133A 2.0 Array	Department of Hematology, Roskilde Hospital, Roskilde, Denmark	11		
17	GSE9006		Affymetrix Human Genome U133A Array	UTSW Medical Center, Dallas, TX, USA	24		12
18	GSE15932		Affymetrix Human Genome U133 Plus 2.0 Array	Department of surgey, Second Affiliated Hospital, School of Medicine, Zhejiang University, Hangzhou, China	8		
19	GSE15653	Liver	Affymetrix Human Genome U133A Array	Department of Research Division, Joslin Diabetes Center, Boston, MA, USA	5		5
20	GSE23343		Affymetrix Human Genome U133 Plus 2.0 Array	Kanazawa University, Kanazawa, Japan	7		10

signaling and metabolic pathways and gene–sets. Tables **2**, **3** and **4** enlisted top 10 pathways/gene-sets for each functional analysis. Furthermore, literature based exploration pointed their connection with T2D through SH2 domain-containing downstream molecules. Fig. (**1**) further illustrates the signaling cascades downstream of Insulin Receptor initiated by PLC-γ, Grb2, and PI3-K.

Table 2. Top 10 common KEGG pathways enriched in GOstat analysis.

S. No.	KEGG Pathway id	Pathway Name
1	5200	Pathways in cancer
2	4510	Focal adhesion
3	4012	ErbB signaling pathway
4	5220	Chronic myeloid leukemia
5	4722	Neurotrophin signaling
6	5100	Bacterial invasion of epithelial cells
7	5212	Pancreatic cancer
8	4520	Adherens junction
9	5211	Renal cell carcinoma
10	5223	Non-small cell lung cancer

Table 3. Top 10 common KEGG pathways enriched in global-test analysis.

S. No.	KEGG id	Pathway Name
1	4940	Type I diabetes mellitus
2	760	Nicotinate and nicotinamide metabolism
3	5330	Allograft rejection
4	5332	Graft-*versus*-host disease
5	5340	Primary immunodeficiency
6	5310	Asthma
7	4672	Intestinal immune network for IgA production
8	400	Phenylalanine, tyrosine and tryptophan biosynthesis
9	270	Cysteine and methionine metabolism
10	471	D-Glutamine & D-glutamate metabolism

3.1. Results from *GOstats* Pathways Analysis

*GOstat*s based KEGG analysis ($P < 0.05$) enriched 'Focal adhesions' among top-ranking pathways in all the tissue sets. The high glucose level in extra cellular fluid exerts a detrimental effect on Focal Adhesion Kinase (FAK) mediated wound healing process [22] which has been reported to associate with lower limb amputation (LLA) [23]; the prevalence of later was found to be eight times higher in diabetic individuals than in nondiabetic individuals [24]. Interestingly FAK can also bind with Grb2-SoS complex, PI 3-kinase, and phospholipase C-γ through its phosphotyrosine residues like that of insulin receptor and this might be a reason for clinical co-occurence of T2D and LLA. Another top enriched pathway was 'Pathways in cancer' and the link between metabolic syndromes and cancer has previously been reported [25]. Insulin-signaling induces two major signaling pathways: Mitogenic MAPK pathway and Anti-apoptotic PI3-K pathway. PI3K activates Protein Kinase B (PKB) or Akt resulting in the activation of mTOR/Raptor complex which mediate its effects on mitogenesis and cell growth. Impaired insulin- signaling may, therefore, leads to deregulation in mTOR-signaling which has been linked to numerous human cancers [26]. The analysis results in this study can be found in Online Supporting Information S1.

3.2. Results from Globaltest Pathways Analysis

Global test based KEGG analysis primarily enrich various immune system related pathways in all tissue types *i.e.*, 'Graft- *versus*- Host disease', 'Allograft reaction', 'Asthma', 'Intestinal Immune network for IgA', 'primary Immunodeficiency' and 'Type I Diabetes'. The analysis in this study, clearly established the role of immune deregulation in Type 2 diabetes as reported earliar by Jaffe & group in 2004 [27]. In addition, various metabolic pathways were also enriched including 'Cysteine and methionine', 'D-Glutamine and D-glutamate', 'Nicotinate and nicotinamide', 'Phenylalanine', 'tyrosine and tryptophan' and 'Ubiquinone and other terpenoid- quinone metabolism'pathway. The implication of these metabolic pathways in T2D pathophysiology seems more challenging than the exposition of the role of signaling pathways in disease etiology and requires further literature support. The results of the analysis in this study can be found in Online Supporting Information S2.

3.3. Results from MSigDB's Gene-sets Analysis

The MSigDB gene-sets analysis also contributed significantly in exposing various aspects of disease etiology. Reactome analysis revealed 'T Cell receptor (TCR) signaling' among top five pathways in all the studies on tissue. TCR-signaling involves the formation of a multi-molecular structure 'Signalosomes' that includes Grb2 and PLC- γ [28]; it is interesting to note that Grb2 is a positive regulator of insulin-signaling pathway. It binds to phosphorylated Tyr of IRS-1 and propagates signal downstream of insulin receptor. While another component of signalosome, PLC- γ is a negative regulator of insulin- signaling, PI3K phosphorylates Phosphatidylinositol 4, 5-bisphosphate (PIP_2) into Phosphatidylinositol (3,4,5)-trisphosphate (PIP_3), PIP_3 intern recruits Akt to the plasma membrane and Akt facilitates the movement of glucose transporter GLUT4 from the internal membrane vesicle to the plasma membrane; hence, stimulating the uptake of glucose [29, 30]. PLC- γ hydrolyzes PIP_2 into diacylglycerol (DAG) and Inositol triphosphate (IP3) and may probably lead to insulin resistance. Presumably, T cell Receptor signaling and impaired insulin signaling cross-talk with each other through these shared components and give rise to the diabetic or IR phenotype. Deregulated T cell receptor signaling may also lead to inappropriate Antibody Dependent Cell Cytotoxicity (ADCC) [31] which may result

Table 4. Top 10 MSigDB gene-sets enriched in global-test analysis.

S. No	Gene-set	Pathway/ Mi RNA
1	Reactome	Phosphorylation of CD3 and TCR ZETA Chains
2		Signaling by TGF beta Receptor Complex
3		Regulation of Pyruvate DehydrogenasePDG complex
4		Hormone Sensitive Lipase HSL Mediated Triacylglycerol Hydrolysis
5	Hallmark	IL6-JAK STAT3 Signaling
6		TGF BETA Signaling
7	MiRNA	TCCGTCC,MIR-184
8		CGGTGTG,MIR-220
9		ATACTGT,MIR-144
10		GGCAGAC,MIR-346

in various immune complications like those enriched previously in the KEGG analysis. Top enriched pathway reported in the liver study was 'Glutamate synaptic neurotransmission'; this pathway includes vesicle-associated membrane protein (VAMP) which is also involved in insulin secretion from pancreatic β cells as well as insulin signaling to responsive tissues by SNARE-Mediated Vesicle Exocytosis [32]. Top enriched pathway in skeletal muscles was 'Pyruvate dehydrogenase (PDH) complex-mediated glucose metabolism' and 'Hormone-sensitive lipase (HSL)-mediated triacylglycerol hydrolysis'; HSL has also been reported to be associated with abdominal obesity in patients with type 2 diabetes [33].

Top ranking Hallmark pathway included 'IL6 signaling *via* JAK/STAT pathway' in all the studies on tissue. IL6 is an inflammatory cytokine which is released from the expanded visceral adipose tissue and adversely affects the insulin signaling cascade. The key element in the JAK/STAT pathway is the Signal Transducer and Activator of Transcription 3 (STAT3) protein that acts as a transcription factor and regulates the Suppressor of Cytokine Signaling (SOCS3) mRNA synthesis. Interestingly SOCS3 is a negative regulator of both cytokine signaling as well as insulin signaling. It binds to phosphorylated tyrosine residues on IR/IRS2 and enhances their proteolytic degradation through ubiquitin-proteasomal pathway. STAT3 signaling has also been reported to contribute to skeletal muscle insulin resistance in T2D [34]. Another pathway that was observed as one of the top pathways was 'TGF-β signaling pathway'. TGF-β signaling is involved in the regulation of insulin gene transcription, pancreatic β cell function and glucose tolerance and energy homeostasis [35-38]. *Invitro* studies have confirmed its role in the inhibition of insulin-stimulated tyrosine phosphorylation of IRS-1 [39]. Moreover, it also inhibits insulin/insulin-like growth factor-1 (IGF-1)-dependent adipose cell differentiation (adipogenesis) process. TGF-β exerts this effect through the induction of extracellular matrix (ECM) synthesis, which in turn inhibits adipogenesis and leads towards metabolic syndrome.

MicroRNA target analysis revealed miR-184 in all four studies on tissue as the top 10 microRNA targets. miR-184 expression has been reported to be negatively correlated with insulin secretion in the pancreas; [40] it also inhibits Protein Kinase C- Beta (PRKCB) expression, a positive regulator of insulin signaling. PRKCB functions downstream of IRS1 in the muscle cells and mediates insulin-dependent DNA synthesis through the RAF1-MAPK/ERK signaling cascade. miR-184 was also reported to participate in the regulation of glucose transport in adipocytes by negatively modulating the insulin-stimulated translocation of the glucose transporter SLC2A4/GLUT4 [40]. The aforementioned facts implicate its role in the pathogenesis of diabetes. IGT and NGT studies on skeletal tissue shared miR-346 among the top 10 targets and an miRNA microarray study on Diabetic nephropathy mice confirmed its role in negative regulation of SMAD 3/4 mediated signaling which has been reported in the pathogenesis of obesity and T2D [41]. Together, these findings proved miR-346 a novel therapeutic target against diabetes. The results of analysis can be found in Online Supporting Information S3.

CONCLUSION

Insulin-signaling and its deregulation in the form of diabetes phenotype is quite an intractable problem and our present understanding of insulin signaling is based on recent advances in biochemistry and medicine [42]. Though insulin is a peptide hormone, insulin-signaling also shares features with growth factor signaling, it is also mediated by Receptor tyrosine kinases (RTKs) like those of EGF, TGFα, PDGF, FGF, VEGFR and CSF-1 signaling. It is speculated that the main objective of insulin is to manage under-utilization of glucose in order to support anabolic reactions like DNA replication, protein synthesis, mitogenesis, cell growth and differentiation [42].

Experimental evidence suggests that insulin signaling actually cross-talks with the signaling pathways of growth factors. The present functional analysis revealed that this cross-talk might be mediated through SH-2 domain containing signaling molecules participating in different signaling

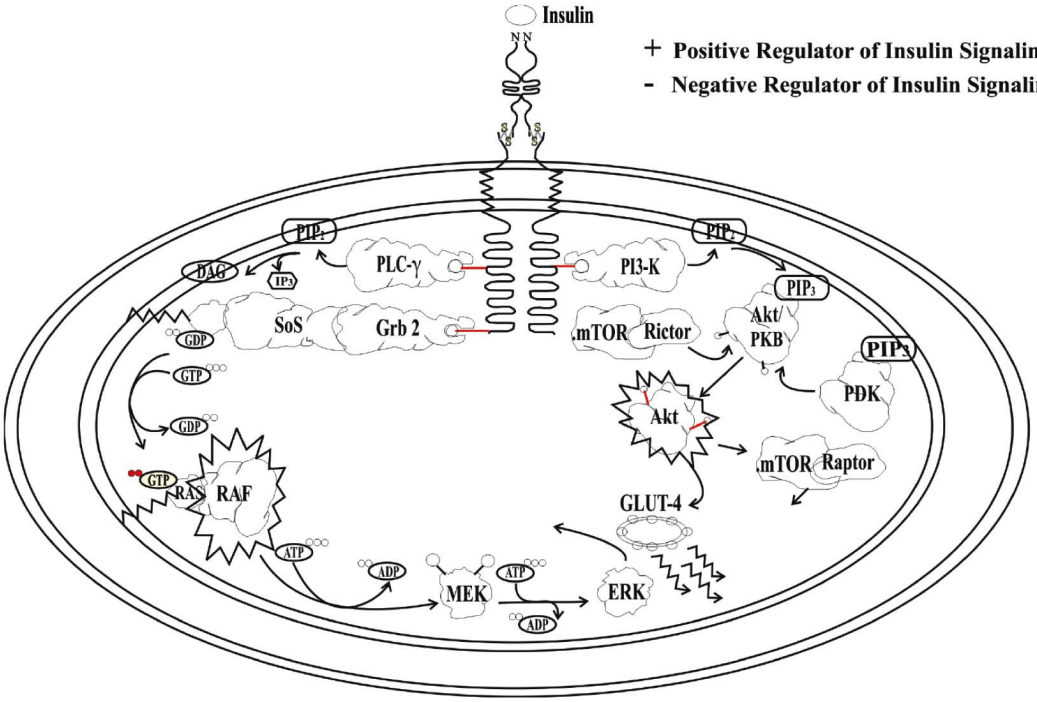

Fig. (1). Signal Transduction pathways downstream of Insulin Receptor.

pathways. It might also be a probable reason of clinical interlink of T2D with cancer, immune-system, and focal adhesion related pathways. In addition, the present study further highlights the need to develop bioinformatics tools and workflows to advance our understanding in insulin signaling in an effort to combat diabetes.

CONFLICT OF INTEREST

The authors confirm that this article content has no conflict of interest.

ACKNOWLEDGEMENTS

AS is grateful to Vice-Chancellor, Uttarakhand Technical University, Dehradun for carrying out the study as part of his doctoral thesis and Director, IAH, GLA University, Mathura for providing necessary facility to conduct the study.

SUPPLEMENTRY MATERIAL

Supplementary material is available on the publisher's web site along with the published article.

REFERENCES

[1] *Global report on diabetes*, World Health Organization **2016.** ISBN 978 92 4 156525 7.

[2] Barrett, T.; Troup, D.B.; Wilhite, S.E.; Ledoux, P.; Evangelista, C.; Kim, I.F.; Tomashevsky, M.; Marshall, K.A.; Phillippy, K. H.; Sherman, P.M.; Muertter, R. N.; Holko, M.; Ayanbule, O.; Yefanov, A.; Soboleva, A. NCBI GEO: archive for functional genomics data sets--10 years on. *Nucleic Acids Res.*, **2010**, *39*(Database).

[3] Parkinson, H. ArrayExpress--a public repository for microarray gene expression data at the EBI. *Nucleic Acids Res.,* **2004**, 33(Database issue).

[4] Mathur, S. K.; Jain, P.; Mathur, P. Microarray evidences the role of pathologic adipose tissue in insulin resistance and their clinical implications. *J. Obes.,* **2011**, 1-16.

[5] Chen, J.; Meng, Y.; Zhou, J.; Zhuo, M.; Ling, F.; Zhang, Y.; Du, H.; Wang, X. Identifying candidate genes for type 2 diabetes mellitus and obesity through gene expression profiling in multiple tissues or cells. *J. Diabetes Res.,* **2013**, 1-9.

[6] Sandberg, R.; Larsson, O. Improved precision and accuracy for microarrays using updated probe set definitions. *BMC Bioinform.,* **2007**, 8(1),48-55.

[7] Gharaibeh, R.Z.; Fodor, A.A.; Gibas, C.J. Background correction using dinucleotide affinities improves the performance of GCRMA. *BMC Bioinform.,* **2008**, *9*(1), 452.

[8] Tseng, G.C.; Ghosh, D.; Feingold, E. Comprehensive literature review and statistical considerations for microarray meta-analysis. *Nucleic Acids Res.,* **2012**, *40*(9), 3785-3799.

[9] Gentleman, R.; Carey, V.; Huber, W.; Hahne, F. genefilter: Methods for filtering genes from high-throughput experiments. *R package version 1.54.2,* 2016.

[10] Zeeberg, B.R.; Riss, J.; David, W.K.; Bussey, K.J. ; Uchio, E. ; Linehan, W.M.; Barrett, J.C.; Weinstein, J.N. Mistaken Identifiers: Gene name errors can be introduced inadvertently when using Excel in bioinformatics. *BMC Bioinform.,* **2004**, *5*, 80.

[11] Carlson, M. org.Hs.eg.db: Genome wide annotation for Human. *R package version 3.3.0*., **2016**.

[12] Ritchie, M.E.; Phipson, B.; Wu, D.; Hu, Y.; Law, C.W.; Shi, W.; Smyth, G.K. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.*, **2015**, *43*(7), e47.

[13] Natarajan, J.; Berrar, D.; Dubitzky, W.; Hack, C.; Zhang, Y.; DeSesa, C.; Brocklyn, J.R.V.; Bremer, E.G. Text mining of full-text journal articles combined with gene expression analysis reveals a relationship between sphingosine-1-phosphate and invasiveness of a glioblastoma cell line. *BMC Bioinform.,* **2006**, *7*, 373.

[14] The Gene Ontology Consortium,; Ashburner, M.; Ball, C. A.; Blake, J.A.; Botstein, D.; Butler, H.; Sherlock, G. Gene Ontology: Tool for the unification of biology. *Nat. Genet.,* **2000**, *25*(1), 25-29. http://doi.org/10.1038/75556.

[15] Falcon, S.; Gentleman, R. Using GOstats to test gene lists for GO term association. Bioinformatics. *Bioinformatics,* **2006**, *23*(2), 257-258.

[16] Goeman, J.J.; Geer, S.A.V.D.; Kort, F.D.; Houwelingen, H. C. V. A global test for groups of genes: Testing association with a clinical outcome. *Bioinformatics,* **2003**, *20*(1), 93-99.

[17] Kanehisa, M.; Sato, Y.; Kawashima, M.; Furumichi, M.; Tanabe, M. KEGG as a reference resource for gene and protein annotation. *Nucleic Acids Res.,* **2015**, *44*(D1), D457-62.

[18] Thomas, P.D. Methods in Molecular Biology, The Gene Ontology

Handbook, **2016**, 15-24.

[19] Subramanian, A.; Tamayo, P.; Mootha, V.K.; Mukherjee, S.; Ebert, B.L.; Gillette, M.A.; Paulovich, A.; Pomeroy, S.L.; Golub, T.R.; Lander, E.S.; Mesirov, J.P. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci.*, **2005**, *102*(43), 15545-15550.

[20] Lee, S.; Kim, J.; Lee, S. comparative study on gene-set analysis methods for assessing differential expression associated with the survival phenotype. *BMC Bioinform.,* **2011**, *12*(1), 377.

[21] Morgan, M.; Falcon, S.; Gentlemanm R. GSEABase: Gene set enrichment data structures and methods. *R package version 1.36.0*, 2016.

[22] Tamura, M.; Osajima, A.; Nakayamada, S.; Anai, H.; Kabashima, N.; Kanegae, K.; Ota, T.; Tanaka, Y.; Nakashima, Y. High glucose levels inhibit focal adhesion kinase-mediated wound healing of rat peritoneal mesothelial cells. *Kidney Int.,* **2003**, *63*(2), 722-731.

[23] Jensen, J.S.; Poulsen, T.M.; Krasnik, M. Wound healing complications following major amputations of the lower limb. *Prosthet. Orthot. Int.,* **1982**, *6,* 105-107.

[24] Johannesson, A.; Larsson, G.-U.; Ramstrand, N.; Turkiewicz, A.; Wirehn, A.-B.; Atroshi, I. Incidence of lower-limb amputation in the diabetic and nondiabetic general population, A 10-year population-based cohort study of initial unilateral and contralateral amputations and reamputations. *Diabetes Care*, **2008**, *32*(2), 275-280.

[25] Braun, S. The link between the metabolic syndrome and cancer. *Int. J. Biol. Sci.,* **2011**, *7*(7), 1003-1015.

[26] Ellisen, L.W. Growth control under stress: mTOR regulation through the REDD1-TSC pathway. *Cell Cycle,* **2005**, *4*(11), 1500-1502.

[27] Jaffe, R.; Mani, J.; Trocki, T.; Mehl, M.L. First line comprehensive care. The diabetes continuum of insulin, glucose, energy dysregulation: Better clinical management of causes improves outcomes, reduces risks and vascular complications. *Original Internist.,* **2004**, *11,* 11-27.

[28] Gorentla, B.K.; Zhong, X.P. T cell Receptor Signal Transduction in T lymphocytes. *J. Clin. Cell Immunol.,* **2012**(*Suppl* 12), 005.

[29] Saltiel, A.R.; Kahn, C.R. Insulin signaling and the regulation of glucose and lipid metabolism. *Nature,* **2001**, 414(6865), 799-806.

[30] Mejia, C.F. Molecular basis of type-2 Diabetes. *Mole. Endocrinol.,* **2006**, 87-108 ISBN: 81-308-0100-0.

[31] Hesslein, D.G.T.; Takaki, R.; Hermiston, M.L.; Weiss, A.; Lanier, L. L. Dysregulation of signaling pathways in CD45-deficient NK cells leads to differentially regulated cytotoxicity and cytokine production. *Proc. Natl. Acad. Sci.,* **2006**, *103*(18), 7012-7017.

[32] Saltiel, A.R.; Pessin, J.E. Mechanisms of Insulin Action. *Springer Science + Business Media, LLC* **2007**, ISBN 978-0-387-72203-0.

[33] Kraemer, F. B. Hormone-sensitive lipase: Control of intracellular tri-(di-) acylglycerol and cholesteryl ester hydrolysis. *J. Lipid Res.*, **2002**, *43*(10), 1585-1594.

[34] Mashili, F.; Chibalin, A.V.; Krook, A.; Zierath, J. R. Constitutive STAT3 phosphorylation contributes to skeletal muscle insulin resistance in type 2 diabetes. *Diabetes*, **2012**, *62*(2), 457-465.

[35] Tan, C.K.; Leuenberger, N.; Tan, MJ.; Yan, Y.W.; Chen, Y.; Kambadur, R.; Wahli, W.; Tan, N.S. Smad3 deficiency in mice protects against insulin resistance and obesity induced by a high-fat diet. *Diabetes*, **2011**, *60*(2), 464-476.

[36] Yadav, H.; Quijano, C.; Kamaraju, A.K.; Gavrilova, O.; Malek, R.; Chen, W.; Zerfas, P.; Zhigang, D.; Wright, E.C.; Stuelten, C.; Sun, P.; Lonning, S.; Skarulis, M.; Sumner, A.E.; Finkel, T.; Rane, S.G. Protection from obesity and diabetes by blockade of TGF-b/Smad3 signaling. *Cell Metab.*, **2011**, *14*(1), 67-79.

[37] Brown, M.L.; Schneyer, A.L. Emerging roles for the TGFbeta family in pancreatic beta-cell homeostasis. *Trends Endocrinol. Metab.,* **2010**, *21*(7), 441-448.

[38] Lin, H.-M.; Lee, J.-H.; Yadav, H.; Kamaraju, A.K.; Liu, E.; Zhigang, D.; Vieira, A.; Kim, S.-J.; Collins, H.; Matschinsky, F.; Harlan, D. M.; Roberts, A. B.; Rane, S. G. Transforming growth factor-beta/Smad3 signaling regulates insulin gene transcription and pancreatic islet beta-cell function. *J. Biol. Chem.*, **2009**, *284*(18), 12246-12257.

[39] Gagnon, A.M.; Chabot, J.; Pardasani, D.; Sorisky, A. Extracellular matrix induced by TGFβ impairs insulin signal transduction in 3T3-L1 preadipose cells. *J. Cell. Physiol.*, **1998**, *175*(3), 370-378.

[40] Morita, S.; Horii, T.; Kimura, M.; Hatada, I. MiR-184 regulates insulin secretion through repression of Slc25a22. *Peer J*., **2013**, *1*, e162.

[41] Zhang, Y.; Xiao, H.Q.; Wang, Y.; Yang, Z.S.; Dai, L.J.; Xu, Y. C. Differential expression and therapeutic efficacy of microRNA-346 in diabetic nephropathy mice. *Exp. Ther. Med.*, **2015,** *10*(1), 106-112.

[42] Ernest, J.M.; Helmreich. *The Biochemistry of Cell Signaling*, Oxford University Press. ISBN 0-19-567384-0.