SVMRFE based approach for prediction of most discriminatory gene target for type II diabetes

Atul Kumar a,⁎, D. Jeya Sundara Sharmila b, Sachidanand Singh a

a Departament of Biotechnology and Health Sciences, Karunya University, Coimbatore, Tamil Nadu, India
b Department of Nanosciences and Technology, Tamil Nadu Agriculture University, Coimbatore, Tamil Nadu, India

1. Introduction

Support Vector Machine (SVM), a machine learning technique implied in the area of time series prediction and classification [31,36] has widely been applied in the life science fields, especially in Bioinformatics. It can handle nonlinear classification tasks efficiently by mapping the samples into a higher dimensional feature space by using a nonlinear kernel function. Since the SVM approach is data-driven and model-free, it has important discriminating power for classification. This characteristic of SVM is obvious in cases where the sample sizes are negligible and numerous variables are involved (high-dimensional space).

Expression profile come under such a category, which contain a large number of attributes (genes). This type of expression data is used to predict the type and occurrence of the disease in a patient [39]. An important aspect while analyzing such type of expression data is the feature selection or dimensionality reduction. Most algorithms lose their potency when genes are large in number with different time series data or dimensionality [7].

To accomplish the task of dimensionality reduction a modified version of SVM known as SVMRFE (Support Vector Machine Recursive Feature Elimination) has been used in this work. SVMRFE was used to identify the most discriminatory target gene in four different microarray

data samples of type II diabetes. These samples have been taken from the Gene Expression Omnibus database (GEO) [13] and Diabetes Genome Anatomy Project (DGAP) (http://www.diabetesgenome.org/). The idea was to build a model wherein the least important features (genes) can be eliminated at each iterative step based on the weight assigned to each gene through SVM. The genes identified through this approach were then classified as essential and non-essential through this. The protein-protein interaction of these non-essential genes revealed vital information regarding interacting proteins. Functional enrichment about these proteins shed a light on their regulatory pathways associated with type II diabetes which can be further explored and confirmed using experimental approach.

2. Materials and methods

2.1. Collection of data sample

71 samples from Pancreatic Islet and Skeletal muscle of Homo sapiens were collected from the GEO and DGAP. Out of these 37 samples are of normal human beings and 34 are of diabetic humans. Table 1 shows the detail description of each of the data sets which were undertaken for studies.

Fisher linear discriminant was applied to all the above-mentioned data sets to rank them based on the Fischer score [21] which was continued with a redundancy reduction step to reduce the redundant data in the microarray dataset [22]. The gene number present in each data set

⁎ Corresponding author.
E-mail address: atulkumar@karunya.edu (A. Kumar).
A $t$-test [3] with a significance level of 0.05 was applied to the datasets to filter out the genes which are not involved in causing type II diabetes. After this reduction step SVMRFE approach (with linear kernel function and 6 subsets of the training data) [24] was applied to train the data samples for 5 iterations. As a result, discriminatory genes based on the weighted ranking were obtained. The identified genes were identified as being essential and non-essential using the database of essential genes. A gene interaction and pathway analysis of the potential non-essential genes was performed to identify the novel targets for type II diabetes (Fig. 1).

### 3. Result and discussion

#### 3.1. $t$-test analysis

For each of the T2D datasets, a $t$-test analysis was performed with a significance level of 0.05. As a result, there was a high dimensionality was still high. A $t$-test [3] with a significance level of 0.05 was applied to the datasets to filter out the genes which are not involved in causing type II diabetes. After this reduction step SVMRFE approach (with linear kernel function and 6 subsets of the training data) [24] was applied to train the data samples for 5 iterations. As a result, discriminatory genes based on the weighted ranking were obtained. The identified genes were identified as being essential and non-essential using the database of essential genes. A gene interaction and pathway analysis of the potential non-essential genes was performed to identify the novel targets for type II diabetes (Fig. 1).

### Table 1

Microarray dataset undertaken for studies.

| Source | Data                                                                 | No. of samples | No. of genes | Country          |
|--------|----------------------------------------------------------------------|----------------|--------------|------------------|
| GEO    | Effect of insulin infusion on human skeletal muscle [33]             | 6              | 22,215       | Sweden           |
| DGAP   | Human pancreatic islets from normal and Type 2 diabetic subjects (A) [18] | 7              | 22,191       | Caucasian and Asian |
| DGAP   | Human pancreatic islets from normal and Type 2 diabetic subjects (B) [18] | 7              | 22,550       |                   |
| DGAP   | Human skeletal muscle - type 2 diabetes [29]                        | 17             | 22,177       | Sweden           |

### Table 2

Number of input and output genes from each dataset for $t$-test analysis.

| Name of dataset | No of inputted genes | No of genes rejecting the null hypothesis |
|-----------------|----------------------|------------------------------------------|
| Effect of insulin infusion on human skeletal muscle | 1223                | 24                                       |
| Human pancreatic islets from normal and type II diabetic subjects (A) | 1210                | 17                                       |
| Human pancreatic islets from normal and type II diabetic subjects (B) | 803                  | 21                                       |
| Human skeletal muscle-type II diabetes               | 1238                | 28                                       |

### Table 3

$p$-value of genes following the alternative hypothesis for the dataset “GSE7146”.

| Probe id | Gene                                                                 | p-Value   |
|----------|---------------------------------------------------------------------|-----------|
| 213524_s_at | G0/G1switch 2                                                      | 0.00001   |
| 216599_s_at | Solute carrier family 22 (organic anion transporter), member 6     | 0.00005   |
| 207295_at   | Sodium channel, non-voltage-gated 1, gamma                         | 0.0001    |
| 218409_s_at | Dnaj (Hsp40) homolog, subfamily C, member 1                         | 0.0003    |
| 203221_at   | Transducin-like enhancer of split 1 (E (sp1) homolog, (Drosophila)| 0.0004    |
| 210452_x_at | Cytochrome P450, family 4, subfamily F, polypeptide 2               | 0.001     |
| 201630_s_at | Acid phosphatase 1, soluble                                         | 0.001     |
| 207955_at   | Chemokine (C-C motif) ligand 2                                      | 0.002     |
| 208507_at   | Olfactory receptor, family 7, subfamily C, member 2                 | 0.002     |
| 210889_s_at | Fc fragment of IgG, low affinity IIb, receptor (CD32)               | 0.002     |
| 207732_s_at | Discs, large homolog 3 (neuroendocrine-dlg, Drosophila)            | 0.002     |
| 220636_at   | Dynemin, axonemal, intermediate polypeptide 2                       | 0.002     |
| 205863_at   | S100 calcium binding protein A12                                    | 0.002     |
| 205603_s_at | Diaphanous homolog 2 (Drosophila)                                   | 0.003     |
| 220979_s_at | ST6 (alpha-N-acetyl-neuraminyl 1-2, 3-beta-galactosyl 1-4, 2)       | 0.003     |
| 2063110_at  | Serine peptidase inhibitor, K25II Type B (acrosin-trypsin inhibitor) | 0.004     |
| 210442_at   | Interleukin 1 receptor-like 1                                        | 0.004     |
| 201214_s_at | Protein phosphatase 1, regulatory subunit 7                         | 0.004     |
| 220385_at   | Junctophilin 2                                                      | 0.004     |
| 205490_x_at | Gap junction protein, beta 3, 31 kDa (connexin 31)                  | 0.004     |
| 213772_s_at | Golgi-associated, gamma adaptin ear containing, ARF binding protein 2 | 0.004     |
| 213950_s_at | Protein phosphatase 3 (formerly 2B), catalytic subunit, gamma isoform (calcinexin A gamma) | 0.004     |
| 201681_s_at | Discs, large homolog 5 (Drosophila)                                | 0.004     |
| 220782_x_at | Kalikrein-related peptidase 12                                     | 0.004     |

![Fig. 1. Flow chart of the analysis.](image-url)
The subsets of genes based on the p-value were given as an input to the support vector machine. Recursive Feature Elimination (RFE) is an iterative procedure for SVM classifier. The recursive feature elimination algorithm of the support vector machine assigns a weight to each gene. The weight was calculated based on the expression value of genes in the disease and normal sample for all the dataset. The algorithm classified the genes (with a classification accuracy of 83.9%) based on the descending order of the weight. Then it generated the list of genes which were found to be the most discriminatory in the normal and disease samples (Tables 7–10). The outline for SVMRFE in the linear kernel is presented below:

**Inputs:**
- Training samples: \( X_0 = [x_1, x_2, \ldots, x_n]^T \)
- Class labels (1 for normal or 0 for diseased): \( y = [y_1, y_2, \ldots, y_n]^T \)

**Initialize:**
- Surviving genes: \( s = [1, 2, \ldots, n] \)
- Gene-ranking list: \( r = [] \)
- Limit training samples to good genes: \( X = \chi_0 s, \chi \)
- Train the classifier: \( \alpha = SVM-train(X, y) \)

\[ w = \sum_k \alpha_k y_k x_k \] where \( k \) indicates the \( k \text{th} \) training pattern

Compute the ranking criterion for the \( i \text{th} \) gene
\[ R(i) = (w_j)^2 \]
Mark the gene with the lowest ranking
\( g = \arg \min R \)
Renew the gene-ranking list
\[ r = [s(g), r] \]
3.3. Identification of degree of essentiality and non-essentiality of genes

To identify significant and reliable targets, the work was concentrated on non-essential genes. Essential genes were ruled out based on the hits obtained from the Database of Essential Genes (DEG 10.9) (http://tubic.tju.edu.cn/deg/) [46]. Essential genes sustain an organism. Therefore, having them as a potential gene target may induce side effects of the drugs. Hence, it is important to identify only the non-essential genes which may be used as a potential drug target. Tables 11–14 show the non-essential genes from the microarray dataset which is under study.

3.4. Gene interaction studies

After obtaining the non-essential genes from the top ranked coding genes for each of the datasets, gene regulatory network was constructed using STRING (Search Tool for the Retrieval of Interacting Genes/Proteins) database [40]. The study was mainly done to observe the interaction between non-essential protein-coding genes with other proteins which are a result of biochemical events and/or electrostatic forces [23]. The function and activity of a protein are often modulated by other proteins with which it interacts.

3.4.1. Gene regulatory network of dataset “GSE7146”

In this dataset, out of the ten best coding genes obtained through the SVMRFE approach, only 5 genes (ACP1, FCGR2B, SCNN1G, CCL27, and DLG3) showed interaction with other protein coding genes (Fig. 6). The ACP1 showed a direct interaction with EPHA2, which is reported to increase the chance of myocardial infarction and reduce the survival
rate of hyperglycemic mice [12]. LYN showed indirect interaction with ACP1 via EPHA2 and direct interaction with FCGR2B. Its kinase activation modulation has been reported to be a novel insulin receptor-potentiating agent. This potentiating agent produces a rapid-onset and a durable blood glucose-lowering activity in diabetic animals [32]. FCGR2B also showed direct interaction with PTPN6 which is been reported to negatively regulate insulin action on glucose homeostasis in the liver and muscle [44]. An analysis of DLG3 has shown its direct interaction with GRIN2A and GRIN2B. Both these genes have been reported to play a potential role in diabetes [11,37,42]. UBC has been reported to play a major role in the diabetes pathway [8,16,26] and its direct interaction with SCNN1G shows that SCNN1G may also play a role in diabetes [30].

3.4.2. Gene regulatory network of dataset “human pancreatic islets from normal and type II diabetic subjects (A)”

Except for ABCC4 and FMO5, all the other four proteins showed a significant and strong interaction with other neighboring proteins (Fig. 7). Purine Nucleoside Phosphorylase (PNP) and Nucleoside Phosphate Kinase (NPK) have reportedly played a major role in diabetes either by positive or negative metabolic regulation [9]. These two molecules also showed interaction with the HPRT1 and the CDA. Caveolin has already been reported to mediate insulin signaling thereby affecting the glucose uptake [6]. In the other subgroup network FUT3 has three direct neighbors: FUT1, FUT2, and B4GALT1 of which the B4GALT1 expression level has been shown to be affected by hyperglycemia [25].

3.4.3. Gene regulatory network of the dataset “human pancreatic islets from normal and type II diabetic subjects (B)”

Both the protein coding genes in this dataset (RNASEK and APCDD1) have shown a significant interaction with the neighboring proteins (Fig. 8). The involvement of RNASEK in diabetes is still an unanswered question, but APCDD1 interaction with its neighbors shows that it may be involved in the pathophysiology of diabetes. LPAR6 (Lysophosphatidic Acid Receptor 6) interacting directly with APCCD1 has shown its activity with PPARγ which is a potential target for diabetes [38]. Aranda et al., in 2012 also showed that the DM/HG (Diabetes mellitus/High Glucose)

Fig. 4. p-Value corresponding to all the genes in the training set for dataset “human pancreatic islets from normal and type II diabetic subjects (B)”.

Fig. 5. p-Value corresponding to all the genes in the training set for dataset “human skeletal muscle-type II diabetes”.
diabetes interaction network of USP6NL have shown its significant till now, for diabetes. The three genes (SOS1, EGFR, and EGF) in the genes (Fig. 9). This selective network of ProSAPiP1 has not been reported with diabetes. SOS1 has shown its association with reference to the insulin action [4], in differential expression of EGFR which is a responsible for diabetes susceptibility. Therefore it may be a potential reprograms signaling pathways in RECs (Retinal Endothelial Cells) to induce a state of LPA (Lysoosphaphatic Acid) resistance. In the year 2000, Figueroa et al. [14] showed that alterations in LRPS5 expression may be responsible for diabetes susceptibility. Therefore it may be a potential target for therapeutic intervention. It has been reported that Wnt/LRPS5 (lipoprotein receptor-related protein 5) signaling contributes to the glucose-induced insulin secretion in the islets [15].

3.4.4. Gene regulatory network of dataset “human skeletal muscle-type II diabetes”

The two prominent protein coding genes (USP6NL and ProSAPiP1) as per SVMRF analysis showed interaction with a different set of genes (Fig. 9). This selective network of ProSAPiP1 has not been reported till now, for diabetes. The three genes (SOS1, EGF, and EGF) in the interaction network of USP6NL have shown its significance in connection with diabetes. SOS1 has shown its association with reference to the insulin action [4], in differential expression of EGFR which is a major impact on diabetes and associated diseases [1,5,27,28,41,45]. Kasayama et al. [19] long back in 1989 reported that EGF deficiency occurs in diabetes mellitus hence insulin may be important in maintaining the normal level of EGF in the submandibular gland and plasma.

3.5. Functional enrichment of significant genes implying pathway analysis

To further validate the involvement of the identified genes in type II diabetes, pathway enrichment was considered. This was solely meant for all the interacting proteins with the identified significant protein(s). The study was carried out using Biointerpreter, a web-based biological interpretation tool for Microarray data analysis (Genotypic Technology Pvt. Ltd., Bangalore, India). The pathway analysis showed that some of the interacting proteins were involved in pathways which were directly or indirectly associated with type II diabetes.

| Table 7 | Best ranked genes for dataset “GSE7146”. |
|--------|----------------------------------------|
| Gene name |  |
| G0/G1switch 2 | Transducin-like enhancer of split 1 (E(sp1) homolog, Drosophila) |
| Acid phosphatase 1, soluble | DnaJ (Hsp40) homolog, subfamily C, member 1 |
| Golgi-associated, gamma adaptin ear containing, ARF binding protein 2 | Protein phosphatase 1, regulatory subunit 7 |
| Interleukin 1 receptor-like 1 | Discs, large homolog 5 (Drosophila) |
| Cytotoxic protein P450, family 4, subfamily F, polypeptide 2 | Protein phosphatase 3 (formerly 28), catalytic subunit, gamma isoform (calcineurin A gamma) |
| Gap junction protein, beta 3, 31 kDa (connexin 31) | Diaphanous homolog 2 (Drosophila) |
| Olfactory receptor, family 7, subfamily C, member 2 | Solute carrier family 22, organic anion transporter, member 6 |
| Serine peptidase inhibitor, Kazal Type II (acrosin-trypsin inhibitor) | Chemokine (C-C motif) ligand 27 |
| Dynemin, axonemal, intermediate chain 2 | Junctophilin 2 |
| Kallikrein-related peptidase 12 | Discs, large homolog 3 (neuroendocrine-dlg, Drosophila) |
| S100 calcium binding protein A12 | Sodium channel, non-voltage-gated 1, gamma subunit |
| ST6 (alpha-N-acetyl-neuraminyl-2,3-beta-galactosyl-1,3) | STG (alpha-N-acetylneuraminyl-2,3-beta-galactosyl-1,3) |
| -N-acetylgalactosaminidase alpha-2, 6-sialyltransferase 5 | Fe fragment of IgC, low affinity Iib, receptor (CD32) |

| Table 8 | Best ranked genes for dataset “human pancreatic islets from normal and type II diabetic subjects (A)”. |
|--------|----------------------------------------|
| Gene name |  |
| Glucuronidase, beta | Enoyl-CoA delta isomerase 1 |
| C-terminal binding protein 1 | Inhibitor of DNA binding 2, dominant negative helix-loop-helix protein |
| Hypoxanthine phosphoribosyltransferase 1 | Solute carrier family 26, member 6 |
| Solute carrier family 26, member 6 | TGF-beta activated kinase 1/MAP3K7 binding protein 2 |
| Caveolin 2 | Nephrosis 1, congenital, Finnish type (nephrin) |
| Fucosyltransferase 9 (alpha (1,3) fucosyltransferase) | Cytidine deaminase |
| Cytochrome P450, family 7, subfamily A, polypeptide 1 |  |

| Table 9 | Best ranked genes for dataset “human pancreatic islets from normal and type II diabetic subjects (B)”. |
|--------|----------------------------------------|
| Gene name |  |
| Adenomatosis polyposis coli down-regulated 1 | Ribonuclease, RNase K |

| Table 10 | Best ranked genes for dataset “human skeletal muscle-type II diabetes”. |
|--------|----------------------------------------|
| Gene name |  |
| Protocadherin beta 3 | Leucine zipper, putative tumor suppressor family member 3 |
| USP6 N-terminal like | Ubiquitously transcribed tetratricopeptide repeat containing, Y-linked |

| Table 11 | Non-essential genes for dataset “GSE7146”. |
|--------|----------------------------------------|
| Gene symbol | Gene name |
| GOS2 | G0/G1switch 2 |
| ACPI | Acid phosphatase 1, soluble |
| CCL27 | Chemokine (C-C motif) ligand 27 |
| JPH2 | Junctophilin 2 |
| KLK12 | Kallikrein-related peptidase 12 |
| S100A12 | S100 calcium binding protein A12 |
| DLG3 | Discs, large homolog 3 (neuroendocrine-dlg, Drosophila) |
| SCNN1G | Sodium channel, non-voltage-gated 1, gamma subunit |
| ST6GALNAC3 | ST6 (alpha-N-acetylneuraminyl-2,3-beta-galactosyl-1,3) |
| ST6GALNAC5 | -N-acetylgalactosaminidase alpha-2, 6-sialyltransferase 5 |
| FCGR2B | Fe fragment of IgC, low-affinity Iib, receptor (CD32) |

| Table 12 | Non-essential genes for dataset “human pancreatic islets from normal and type II diabetic subjects (A)”. |
|--------|----------------------------------------|
| Gene symbol | Gene name |
| HRPT1 | Hypoxanthine phosphoribosyltransferase 1 |
| ABC4 | ATP-binding cassette, sub-family C (CFTR/MRP), member 4 |
| FM05 | Flavin-containing monoxygenase 5 |
| CAV2 | Caveolin 2 |
| FLT3 | Fucosyltransferase 9 (alpha (1,3) fucosyltransferase) |
| CDA | Cytidine deaminase |
proteins interacting mainly with the identified protein DLG3 have been shown to be involved in 3 different pathways viz. Neuroactive ligand-receptor interaction, circadian entrainment and Long-term potentiation (Fig. 10). The proteins present in the Neuroactive ligand-receptor interaction have shown a significant role in the pathobiology of obesity and type II diabetes [10]. The second pathway, circadian entrainment is the biological process that displays an endogenous oscillation of about 24 h. Studies show that exposure to light at night lowers glucose-stimulated insulin secretion due to a decrease in insulin secretory pulse mass. Potential mechanisms have been identified by which disturbances in the circadian rhythms due to modern lifestyle can lead to islet failure in the type II diabetes [35]. It has also been reported that the impaired energy utilization from insulin deficiency impairs a long-term potentiation in diabetes [47].

Table 13
Non-essential genes for dataset “human pancreatic islets from normal and type II diabetic subjects [B]”.

| Gene symbol | Gene name                                      |
|-------------|-----------------------------------------------|
| APCDD1      | Adenomatosis polyposis coli down-regulated    |
| RNASEK      | Ribonuclease, RNase K                         |

Table 14
Non-essential genes for dataset “human skeletal muscle-type II diabetes”.

| Gene symbol | Gene name                                      |
|-------------|-----------------------------------------------|
| USP6NL      | Leucine zipper, putative tumor suppressor family member 3 |
| PROSAPIP1   | USP6 N-terminal like                          |

Fig. 6. Gene regulatory network of dataset “GSE7146”.

Fig. 7. Gene regulatory network of dataset “human pancreatic islets from normal and type II diabetic subjects (A)”.
3.5.2. Pathway enrichment for the interacting proteins of the dataset “human pancreatic islets from normal and type II diabetic subjects (A)”

The protein B4GALT1, interacting with the identified protein FUT3 is involved in several metabolic pathways, connected to type II diabetes (Fig. 11). The protein B4GALT1 participates both in glycoconjugate and lactose biosynthesis. It has shown to be a biomarker in hepatocellular carcinoma, mainly caused due to the insulin resistance syndrome. Finally, the ailment manifests as obesity and later as diabetes [17].

3.5.3. Pathway enrichment for the interacting proteins of the dataset “human pancreatic islets from normal and type II diabetic subjects (B)”

The protein PNPT1 interacting with the RNASEK is reported to be involved in pyrimidine and purine metabolism and the RNA degradation (Fig. 12). Effects of the insulin regulation of purine and pyrimidine metabolism had shown to cause some late complications of the diabetic disease [34]. In 2009, Kocic et al. [20] reported that an impaired dsRNA metabolism may lead to increased levels of different sized RNAs in type II diabetic patients and may have an influence on further ineffective response against the different pathogens.

3.5.4. Pathway enrichment for the interacting proteins of dataset “human skeletal muscle-type II diabetes”

EGFR protein interacting with the identified protein USP6NL has already been reported by many researchers to be involved in diabetes [1, 5, 27, 28, 41, 45]. With the pathway studies, it was identified that the main pathways in which EGFR is involved, is also leading directly to or indirectly to diabetes (Fig. 13). Hypoxia-inducible factor 1 alpha (HIF-1α) is regulated precisely by hypoxia and hyperglycemia. It had also been

Fig. 8. Gene regulatory network of dataset “human pancreatic islets from normal and type II diabetic subjects (B)”.

Fig. 9. Gene regulatory network of dataset “human skeletal muscle-type II diabetes”.

A. Kumar et al. / Genomics Data 12 (2017) 28–37
shown that the HIF-1α and glucose can sometimes influence each other [43]. It has been reported that the components of the MAPK/ERK pathway act as modifiers of the cellular insulin responsiveness. The insulin resistance was due to downregulation of the insulin-like receptor gene expression following persistent MAPK/ERK inhibition. The mechanism permits physiological adjustment of insulin sensitivity and the subsequent maintenance of the circulating glucose at appropriate levels [48]. MAPK and GnRh-Glp-1 pathways in the ileum have also been reported to be involved in the improvement of the blood glucose level [45].

4. Conclusion

Analysis of type II diabetes expression data from two different tissue samples i.e. skeletal muscle and pancreatic islet has given a deep insight into genes which may be possibly involved in the pathophysiology of the disease. The most discriminatory genes obtained in each dataset after complete analysis, have been found to be associated with diabetes either directly or indirectly. However, the majority of the genes have not been previously reported in association with diabetes. The genes identified in the current study viz. FCGR2B, DLG3, SCNN1G, FUT3, HPRT1, APCDD1, USP6NL, ProSAPI and RNASEK may act as a potential drug target. The significant pathways identified through the overall approach were Neuroactive ligand-receptor interaction, circadian entrainment, Long-term potentiation, pyrimidine and purine metabolism, dsRNA metabolism, MAPK/ERK pathway, and GnRh-Glp-1. This study gave the insight to focus on these associated pathways with the above-reported proteins to study in pathway models or mouse model to elucidate them as drug targets or markers for type II diabetes.

Conflict of interest

The authors declare that there is no conflict of interest in the present work.

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.gdata.2017.02.008.

References

[1] A. Advani, K.J. Wiggins, A.J. Cox, Y. Zhang, R.E. Gilbert, D.J. Kelly, Inhibition of the epidermal growth factor receptor preserves podocytes and attenuates albuminuria in experimental diabetic nephropathy. Nephrology 16 (6) (2011 Aug 1) 573 – 581.

[2] J. Aranda, R. Motiejunaite, E. Im, A. Kazlauskas, Diabetes disrupts the response of retinal endothelial cells to the angiomodulator lysophosphatidic acid. Diabetes 61 (5) (2012 May 1) 1225 – 1233.

[3] P. Baldi, A.D. Long, A Bayesian framework for the analysis of microarray expression data: regularized t-test and statistical inferences of gene changes. Bioinformatics 17 (6) (2001 Jun 1) 509 – 519.
A.M. Davalli, A.R. Ochman, C.A. Lipinski, J.A. Handler, A.G. Reaume, M.S. Saporito, The potential role of glutamate in the current diabetes mellitus. Diabetes Metab. 26 (1) (2010 Jan 1) 12–16.

P.J. Miettinnen, J. Uustin, P. Ormio, R. Gao, J. Palgi, E. Hakonen, L. Juntti-Berggren, P.O. Berggren, T. Otonkoski, Downregulation of EGF receptor signaling in pancreatic islets causes diabetes due to impaired postnatal β-cell growth. Diabetes 55 (12) (2006 Dec 1) 3290–3298.

V.K. Mootha, C.M. Lindgren, K.F. Eriksson, A. Subramanian, S. Sihag, J. Lehar, P. Puigserver, E. Carlsson, M. Ridderstråle, E. Laurila, N. Houstis, PGC-1α-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. Nat. Genet. 34 (3) (2003 Jul 1) 267–273.

P.R. Nagib, J. Gameiro, L.G. Stivanin-Silva, M.S. de Arruda, D.M. Villa-Verde, W. Savino, L. Veriaud, Thymic microenvironmental alterations in experimentally induced diabetes. Immunobiology 215 (12) (2010 Dec 31) 971–979.

K.L. Ng, S.K. Mishra, De novo SVM classification of precursor microRNAs from genomic pseudo hairpins using global and intrinsic folding measures. Bioinformatics 23 (11) (2007 Nov 1) 1321–1330.

A.R. Ochman, C.A. Lipinski, J.A. Handler, A.G. Reaume, M.S. Saporito, The Lyn kinase activator MLR-1023 is a novel insulin receptor potentiator that elicits a rapid-onset and durable improvement in glucose homeostasis in animal models of type 2 diabetes. J. Pharmacol. Exp. Ther. 342 (1) (2012 Jul 1) 23–32.

H. Parikh, E. Carlsson, W.A. Ch rutlow, L.E. Johansson, H. Storgaard, P. Poulsen, R. Saxena, C. Ladd, P.C. Schulze, M.J. Mazzini, C.R. Jensen, TNXP regulates peripheral glucose metabolism in humans. PLoS Med. 4 (5) (2007 May 1), e158.

K. Pillwein, M.A. Reardon, H.N. Jayaram, Y. Natsuma, W.M. Elliott, M.A. Faderan, N. Pajda, W. Sperli, G. Weber, Insulin regulatory effects on purine-and pyrimidine metabolism in alloxan diabetic rat liver. Pädiatr. 23 (2) (1987 Dec) 135–144.

J. Qian, G.D. Block, C.S. Colwell, A.V. Matveyenko, Consequences of exposure to light at night on the pancreatic islet circadian clock and function in rats. Diabetes 62 (10) (2013 Oct 1) 3469–3478.

S.B. Rice, G. Nenadic, B.J. Staple, Mining protein function from text using term-based support vector machines. BMC Bioinformatics 6 (1) (2005 May 24) 1.

A.R. Santiago, J.M. Gaspar, F.J. Baptista, A. Cristovão, F.P. Santos, W. Kamphuis, A. Ambrosio, Diabetes changes the levels of iostotopic glucose receptors in the rat retina. Mol. Vis. 15 (2009) 1620–1630.

C.M. Stapleton, D.G. Masek, S. Wang, C.A. Nagle, C.W. Gline, P. Thullier, L.M. Leesnitzer, L.O. Li, J.B. Stimmel, G.I. Shulman, R.A. Coleman, Lysophosphatidic acid activates peroxisome proliferator activated receptor-γ in CHO cells that over-express glucol 3-acyltransferase-1. PLoS One 6 (4) (2011 Apr 20), e18932.

S. Van Dieren, J.W. Beulens, A.P. Kengne, L.M. Peelen, G.E. Rutten, M. Woodward, Y.T. Van der Schouw, K.G. Moons, Prediction models for the risk of cardiovascular disease in patients with type 2 diabetes: a systematic review. Heart 98 (5) (2012 Mar 1) 360–369.

C. Von Mering, M. Huynen, D. Jaeggi, S. Schmidt, P. Bork, B. Snel, STRING: a database of predicted functional associations between proteins. Nucleic Acids Res. 31 (1) (2003 Jan 1) 258–261.

L. Wassef, D.J. Kelly, R.E. Gilbert, Epidermal growth factor receptor inhibition attenuates early kidney enlargement in experimental diabetes. Kidney Int. 66 (5) (2004 Nov 1) 1805–1814.

L. Wilson, Diabetes: pathogenesis of diabetes mellitus: does glutamate have a role? Nat. Rev. Endocrinol. 7 (5) (2011 May 1) 248.

H. Xiao, Z. Gu, G. Wang, T. Zhao, The possible mechanisms underlying the impairment of HIF-1α pathway signaling in hyperglycemia and the beneficial effects of certain therapies. Int. J. Med. Sci. 10 (10) (2013 Jan 1) 1412–1421.

E. Xu, A. Charbonneau, Y. Bolland, K. Bellmann, L. Pao, K.A. Simonovitch, B.G. Neel, N.M. Beauchemin, A. Marrete, Hepatocyte-specific Pparg deletion protects from obesity-linked hepatic insulin resistance. Diabetes 61 (8) (2012 Aug 1) 1940–1958.

M.Z. Zhang, Y. Wang, P. Paeselkson, R.C. Harris, Epidermal growth factor receptor inhibition slows progression of diabetic nephropathy in association with a decrease in endoplasmic reticulum stress and an increase in autophagy. Diabetes 63 (4) (2014 Jun 1) 1803–2072.

R. Zhang, H.Y. Ou, C.T. Zhang, DEG: a database of essential genes. Nucleic Acids Res. 32 (Suppl. 1) (2004 Jan 1) D271–D272.

Y. Izumi, K.A. Yamada, M. Matsukawa, C.F. Zorumski, Effects of insulin on long-term potentiation in hippocampal slices from diabetic rats. Diabetologia 46 (7) (2003 July) 1421–1427.

W. Zhang, B.J. Thompson, V. Hietakangas, S.M. Cohen, MAPK/ERK signaling regulates insulin sensitivity to control glucose metabolism in Drosophila. PLoS Genetics 7 (12) (2011), e1002429.