1H NMR Based Serum Metabolic Profiling Reveals Differentiating Biomarkers in Patients with Diabetes and Diabetes Comorbidity

Atul Rawat[1,2], Gunjan Misra[2,5$], Madhukar Saxena[2$], Sukanya Tripathi[3], Durgesh Dubey[1,2], Sulekha Saxena[4], Avinash Aggarwal[4], Varsha Gupta[3], M Y Khan[2] and Anand Prakash[2,6#]

[1]Centre of Biomedical Research, Lucknow; [2]Department of Biotechnology, Babasaheb Bhimrao Ambedkar University, Lucknow; [3]Department of Zoology, Lucknow University, Lucknow; [4]Department of Critical Care, King George Medical University, Lucknow; [5]Department of Biotechnology, CSJMU, Kanpur; [6]Department of Biotechnology, Mahatma Gandhi Central University, Bihar India

Author's Contributed Equally -: $

Corresponding Author -: #

 Dr.Anand Prakash,
Dean, School of Life Sciences,
Professor & Head, Department of Biotechnology,
Mahatma Gandhi Central University, Bihar, INDIA
E-mail: anandprakash@mgcub.ac.in
Tel (Off): +91-9628282357

Abstract

Background: Diabetes is among the most prevalent diseases worldwide, of all the affected individuals a significant proportion of the population remains undiagnosed because of a lack of specific symptoms early in this disorder and inadequate diagnostics. Diabetes and its associated sequela, i.e., comorbidity are associated with microvascular and macrovascular complications. As diabetes is characterized by an altered metabolism of key metabolites and regulatory pathways. Metabolic phenotyping can provide us with a better understanding of the unique set of regulatory perturbations that predispose to diabetes and its associated comorbidities.

Methodology: The present study utilizes the analytical platform NMR spectroscopy coupled with Random Forest statistical analysis to identify the discriminatory metabolites of diabetes (DB) and diabetes-related comorbidity (DC) along with the healthy control (HC) subjects. A combined and pairwise analysis was performed, between the serum samples of HC (n=50), and DB (n=38), and DC (n=35) individuals to identify the discriminatory metabolites responsible for class separation. The perturbed metabolites were further rigorously validated using t-test, AUROC analysis to examine the statistical significance of the identified metabolites.

Results: The DB and DC patients were well discriminated from HC. However, 15 metabolites were found to be significantly perturbed in DC patients compared to DB, the identified panel of metabolites are TCA cycle (succinate, citrate), methylamine metabolism (trimethylamine, methylamine, betaine), -intermediates; energy metabolites (glucose, lactate, pyruvate); and amino acids (valine, arginine, glutamate, methionine, proline and threonine). The metabolites were further used to identify the perturbed metabolic pathway and correlation of metabolites in DC patients.

Conclusion: The ^{1}H NMR metabolomics may prove a promising technique to differentiate and predict diabetes and its comorbidities on their onset or progression by determining the altered levels of the metabolites in serum.

1. Introduction

Diabetes is an enigmatic, genetically inherited, primarily a metabolic disorder characterized by multifaceted perturbation in metabolism mainly glucose and lipid[1] in both -type 1 (insulin deficiency due to autoimmune destruction of the pancreatic b-cells)[2] and -type 2 (insulin resistance or reduced insulin secretion due to islet cell dysfunction)[3] diabetes. From being considered as the disease of the elderly, diabetes has now become the major cause of morbidity and mortality among youth and middle age population as well[4-5]. Owing to perpetually increasing industrialization, contracting urban-rural division, augmented economic growth, varying dietary norms, lesser or no physical activity and alleviated stress levels among all strata of society are the risk factors behind the diabetes[6]. Study performed by the Medical Expenditure Panel Survey showed that most adults with diabetes have at least one comorbid chronic disease and as many as 40% have at least three[7]. Diabetes in its course develops many co-morbidities associated with microvascular (neuropathy, nephropathy, and retinopathy) and macrovascular (atherosclerotic-related vascular disease, coronary heart disease, cerebrovascular disease, and peripheral vascular disease) complications[8-9]. However, patients with diabetes not only have diabetes-related comorbidity, they as well have nondiabetes-related dysfunctions, such as musculoskeletal diseases and depression which take its toll on overall quality of life[8, 10]. The condition of multimorbidity in diabetic patients further worsen the patient's situation and the treatment strategy as a consequence affecting the success of the treatment. These co-morbidities are often challenging for physicians to manage because they can be present for years before becoming clinically apparent. Diagnosis and the treatment aspect of diabetes have been extensively studied, but the identification of early biomarkers/novel pathways suggestive of metabolic aberrations related to co-morbidity development is still unclear. Till now, most studies involving metabolomics of diabetes have focussed on its prediction rather than its co-morbidities. Improved prognosis may encourage early treatment interventions and reduce diabetes burden. Metabolomics is the study of small molecules and biochemical intermediates (metabolites), which are highly relevant to other regulatory mechanisms (e.g., genomics, transcriptome, and proteome) and sensitive to environmental stimuli, forming detailed representations of organismal phenotypes[11]. Over the past decade, the application of metabolomics has been used to gain new insights into the pathology of numerous diseases including diabetes[12-13]. Therefore, the application of metabolomics to study (diabetes and diabetes-related co-morbidities) pathophysiology represents a promising avenue of research to identify candidate biomarkers related to disease development and progression[14]. Metabolomics offers a new approach for the identification of novel risk markers of diseases with the advent of high-throughput analytical platforms in which simultaneous measurements of hundreds of analytes is possible[15]. Among the analytical techniques that can be used for metabolomics applications, nuclear magnetic resonance (NMR) spectroscopy and mass spectrometry (MS) are the most commonly used. NMR technique enjoys a vast range of simultaneous detectable metabolites not to mention a simple sample preparation and excellent reproducibility and non-destructive nature of the analysis[16]. Thus, combining metabolomics data with clinical and molecular data may open new vistas in the prediction of diabetes as well as its co-morbidities.

2. Methodology

2.1 Recruitment of Subjects and Sample Collection:

The diabetic patients were carefully chosen from those attending the OPD (outpatient department) of government hospital in Lucknow and Kanpur. To compare the findings an equal number of age-matched healthy individuals were studied. The study was spearheaded after due approval by the institutional research and ethics committee, (G.S.V.M. Medical College, Kanpur, Uttar Pradesh, ethical code No.14/Steno, on 13 January 2011) and acquiring written informed consent from all the study subjects. The serum samples used in this study were obtained from patients, diagnosed with diabetes and diabetes-related complication. A rigorous inclusion and exclusion criteria was followed for sample collection. Inclusion Criteria: (a) The healthy control (HC) group (n=50) were age and sex-matched healthy subjects taken from the general population, free from any ailment and were not on any medication which could affect the parameters under study. (b) Diabetes (DB) patients with good glycaemic control (n=38), disease duration less than 7 years and glycated haemoglobin (HbA1C) level less than 8%. They were on lifestyle modifications and oral hypoglycemic drugs. (c) diabetes-related complication (DC) patients with inadequate glycaemic control (n=35), disease duration more than 7 years and Glycated haemoglobin (HbA1C) level more than 8%. They were on lifestyle modifications, oral hypoglycemic drugs, insulin or combination of all three. Exclusion criteria: The exclusion criteria for patients extended to those diagnosed with type 1 Diabetes Mellitus, acute complications for instance severe infections or major surgery, trauma, severe cardiovascular/respiratory diseases, pregnant and breastfeeding females. Recently diagnosed cases and those suffering from chronic diabetic co-morbidities were also excluded from the study. Based on the above inclusion and exclusion criteria, a total of 131 patients were recruited in this study. In each case, the 3.0 ml of blood sample was drawn and processed to extract the serum. The collected serum was transferred into a sterile centrifuge tubes and stored at $-80°C$ immediately after processing until the NMR experiments were performed. The serum samples of patients those diagnosed with diabetes (DB) and diabetes-related complication (DC) along with age and sex-matched healthy control (HC) without any age-dependent problems were finally used in this study to profile the metabolic differences.

2.2 NMR- Sample preparation, Data Acquisition, and Processing:

The stored serum samples were thawed and centrifuged at 10,000 rpm for 5 minutes to remove precipitates or cellular debris if any. The serum samples (250 µl) was mixed with 250 µl of 0.9% saline sodium phosphate buffer of strength 20 mM and pH 7.4 prepared in D_2O. A 500 µl aliquot was transferred into 5 mm NMR tubes with a coaxial insert containing a known concentration of TSP (0.1mM prepared in D_2O) to provide a lock for NMR experiments, and as a standard external reference to aid chemical shift referencing for metabolite quantification and assignment. D_2O (Deuterium oxide-as a co-solvent and to provide a deuterium field/frequency lock) and TSP (sodium salt of trimethylsilylpropionic acid-d4) used for NMR experiments were acquired from Sigma-Aldrich (Rhode Island, USA). For each sample, the NMR spectra were recorded at 300 Kelvin (K) on a Bruker Biospin Avance-III 800 MHz NMR spectrometer. Using standard parameters from the Bruker's pulse program library sequence (cpmgpr1d), a transverse relaxation-edited CPMG (Carr–Purcell–Meiboom–Gill) 1D ^{1}H NMR

spectra were acquired, mainly to get information about micromolecular metabolites and remove broad signals from proteins, cholesterols, and phospholipids. The obtained spectra were processed using Bruker NMR data Processing Software Topspin-v3.5. The ^{1}H NMR spectra of all the serum samples were manually phased, and baseline corrected followed by calibrating each spectrum to the lactate resonance at δ(1.31) ppm. Further, all the spectra were subjected to visual inspection for acceptability and finally used for binning in the region δ(0.5 to 8.5) ppm. The CPMG data, spectral regions of δ(0.50 to 4.57) ppm and δ(5.1 to 8.5) ppm were subdivided into integrated regions of δ(0.01) ppm using AMIX 3.8.7, (Bruker, BioSpin), to reduce the complexity of the NMR data and aid data visualization using pattern recognition methods. The region distorted due to water suppression δ(5.1-4.7) ppm, were excluded. The obtained data were exported to Excel and subjected to multivariate pattern recognition and univariate analysis in the statistical analysis module of MetaboAnalyst, an open access web-based tool for metabolomics studies. The ^{1}H-NMR signals of all common metabolites including amino acids, organic acids, and carbohydrates were assigned in reference to previous publications.

2.3 Statistical Analysis:

In the present article, we chose Random forest (RF) to select differentiating biomarkers between the different cohorts, i.e. diabetes (DB), diabetes-related co-morbidity (DC) and healthy control (HC). Prior to multivariate statistical analysis, the binned dataset was Pareto scaled, which is the recommended method for untargeted metabolomics studies. Subsequently, the normalized dataset was subjected to supervised classification methods— RF. It is a classification and regression technique which involves constructing a multitude of trees in the training phase. More precisely saying, it is an ensemble method of trees developed from a training dataset and validated internally to achieve an accurate prediction of the target variable from the predictors for the purpose of future observations. RF will create multiple classification and regression trees (CART) based on the bootstrap sample from the original training data. It also randomly searches the feature to determine the splitting point for growing a tree. In addition, the RF does not overfit as the number of trees increases but it will produce a restrictive value on the generalization error. Random Forest is best suited for the analysis of complex data structures embedded in small to moderate data sets. The strengths of RF lies in identifying outliers and anomalies in the data, displaying proximity clusters, predicting future outcomes, identifying relevant predictors, discovering data patterns, replacing missing values with imputations, and providing insightful graphics. The important variable/features were ranked by their contributions to classification accuracy, i.e., Mean Decrease Accuracy (MDA). Next, unpaired t-test was also applied for pairwise analysis to assess the significance of the change in the metabolic profile and a p-value of < 0.05 was used as the criterion for statistical significance. Receiver operating characteristic (ROC) analysis was also carried out to verify the robustness of discriminatory metabolites, which are generated by plotting the false positive rate against true positive rate and area under the curve (AUC) values were calculated as an indication of the prediction accuracy. ANOVA was performed for boxplot representation to visualize the comparative variation in the levels of significantly altered metabolites in different cohorts (DB, DC, and HC) identified in the multivariate analysis. The identified significant metabolites were further used for metabolic pathway analysis, to identify the altered metabolic pathways associated with the disease.

Table 1: Clinical and demographic details.

Parameter	HC	DB	DC
Number	50	38	35
Sex (M/F)	29/21	21/17	20/15
Age	47.83±4.97	46.64±3.40	50.76±4.39
BMI (Kg/m^2)	24.64±0.8	27.27±0.91*	26.66±0.89^{ns}
HbA1C (%)	5.4±0.34	7.79±0.27*	8.3±0.3^{ns}
Disease duration (yrs)	NA	5.45±1.13*	10.16±1.54^s

*Values expressed in mean±SEM, ns=difference in values not significant, s= difference in values significant ($p < 0.05$).

3. Results

3.1 Metabolite Assignment:

The standard one-dimension (1D) spectrum gave an overview of all metabolites. A typical 1D ^{1}H CPMG NMR spectra of serum samples obtained from the three groups are shown in **Figure 1**. Chemical shifts were annotated as far as possible, by comparing them with the chemical shifts available with the software Chenomx[17] (NMR Suite, v8.1, Chenomx Inc., Edmonton, Canada). The assigned resonances of the metabolite peaks were validated using: (a) previously reported NMR assignments of metabolites, data obtained from BMRB database (Biological Magnetic Resonance Data Bank)[18] and HMDB (The Human Metabolome Database)[19]. The NMR spectra presented signals mainly from lipids/lipoproteins [e.g., low-density lipoprotein (LDL), very low-density lipoprotein (VLDL), polyunsaturated fatty acids (PUFAs)], membrane metabolites [e.g., choline, phosphocholine (PC), and Glycerophosphocholine (GPC)], N-acetyl glycoproteins (NAG), and amino acids (e.g. leucine, isoleucine, valine, alanine, arginine, lysine, proline, glutamine, glutamate, glycine, threonine, histidine, tyrosine, and phenylalanine etc.). Other identified metabolites were glucose, lactate, acetate, citrate, pyruvate, betaine, creatine, creatinine, etc.
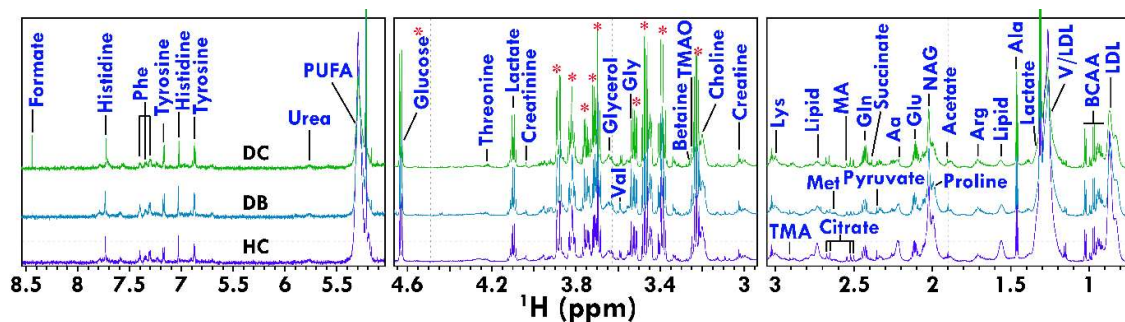

Figure 1. Stack plot of representative 1D ^{1}H NMR spectra of serum obtained from different groups.

The peaks annotated in the figure show the assignments of serum metabolites. The abbreviations used are LDL/VLDL: Low/very-low-density lipoproteins; PUFA: polyunsaturated fatty acids; BCAA (Branched-chain amino acids): Isoleucine; Leucine; Valine; Aa: Acetoacetate; Lys: Lysine; Phe: Phenylalanine; Gly: Glycine; Val: Valine; Gln: Glutamine; Glu: Glutamate; Arg: Arginine; Ala: Alanine; Met: Methionine; MA: Methylamine; TMA: Trimethylamine.

3.2 Classification & Feature Selection:

Random Forest was used for multivariate analysis of the 1D ^{1}H NMR CPMG data sets. RF is a powerful supervised classification method available in the statistical analysis module of MetaboAnalyst[20]. RF uses an ensemble of classification trees, each of which is grown by random feature selection from a bootstrap sample at each branch. The class prediction is based on the majority vote of the ensemble. During tree construction, around one-third of the instances are left out of the bootstrap sample. This remaining data is then used as test sample to obtain an unbiased estimate of the classification and regression (OOB = out of the bag) errors as trees are added to the forest[21]. RF analysis was performed to get a discriminatory overview of the three cohorts and further pairwise analysis was also performed between the groups to identify the differentiating metabolites, i.e. DB vs. HC, DC vs. HC and DC vs. DB as shown in **Figure 2**. In RF analysis 500 trees were grown and 7 features were randomly selected at each node. The generalization error was estimated on the OOB samples.
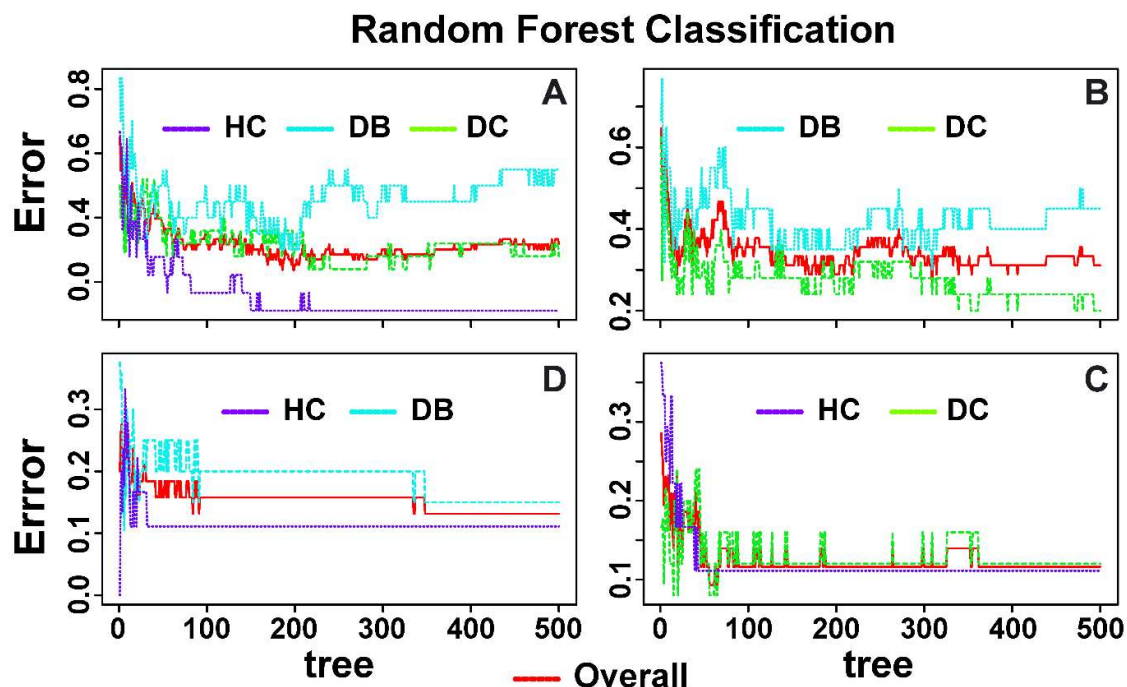

Figure 2. Random Forest classification.

RF analysis of 1D 1H CPMG NMR spectra comprising of all the groups (A) and pairwise analysis (B, C, D). The overall error rate is shown by the red line; the respective colors show the error rates for each class.

Next, we determined the important biomarkers in metabolomics dataset of pairwise analysis. Variable importance (VI) is based on MDA, which is the result of the permutation of the average over all trees and is used to measure the importance of the variables in RF. After manual data mining and processing, we identified metabolites significantly perturbed in sera of different cohorts (MDA ≥0.01) as enlisted in **Table 2**. The quantitative difference in the metabolite concentration was assessed using a t-test, (p-Value < 0.05). Further receiver's operating characteristic (ROC) curves analysis was performed for the significant metabolite

markers to evaluate their predictive power or diagnostic accuracy. The area under the ROC curve (AUC) is generally considered as the method of choice for evaluating the performance of potential biomarkers. The higher the AUC, the better the prediction of the model and discriminatory ability ($0.5 \leq$ no discrimination; $1 =$ perfect discrimination). The highest and lowest AUC values in our study range from 0.98 to 0.52 (**Table 2**) which indicated that these metabolites could be potential biomarkers for diagnosis, surveillance, and early detection of metabolic perturbations in such patients. A cumulative ROC curves of serum metabolites significantly altered in DB vs. HC, DC vs. HC, and DC vs. DB are shown in **Figure S1**. Henceforth these metabolites can be aptly used for clinical diagnosis and surveillance in diabetic patients. The discriminatory metabolites were mainly related to lipid, amino acid, glucose, and energy metabolism.
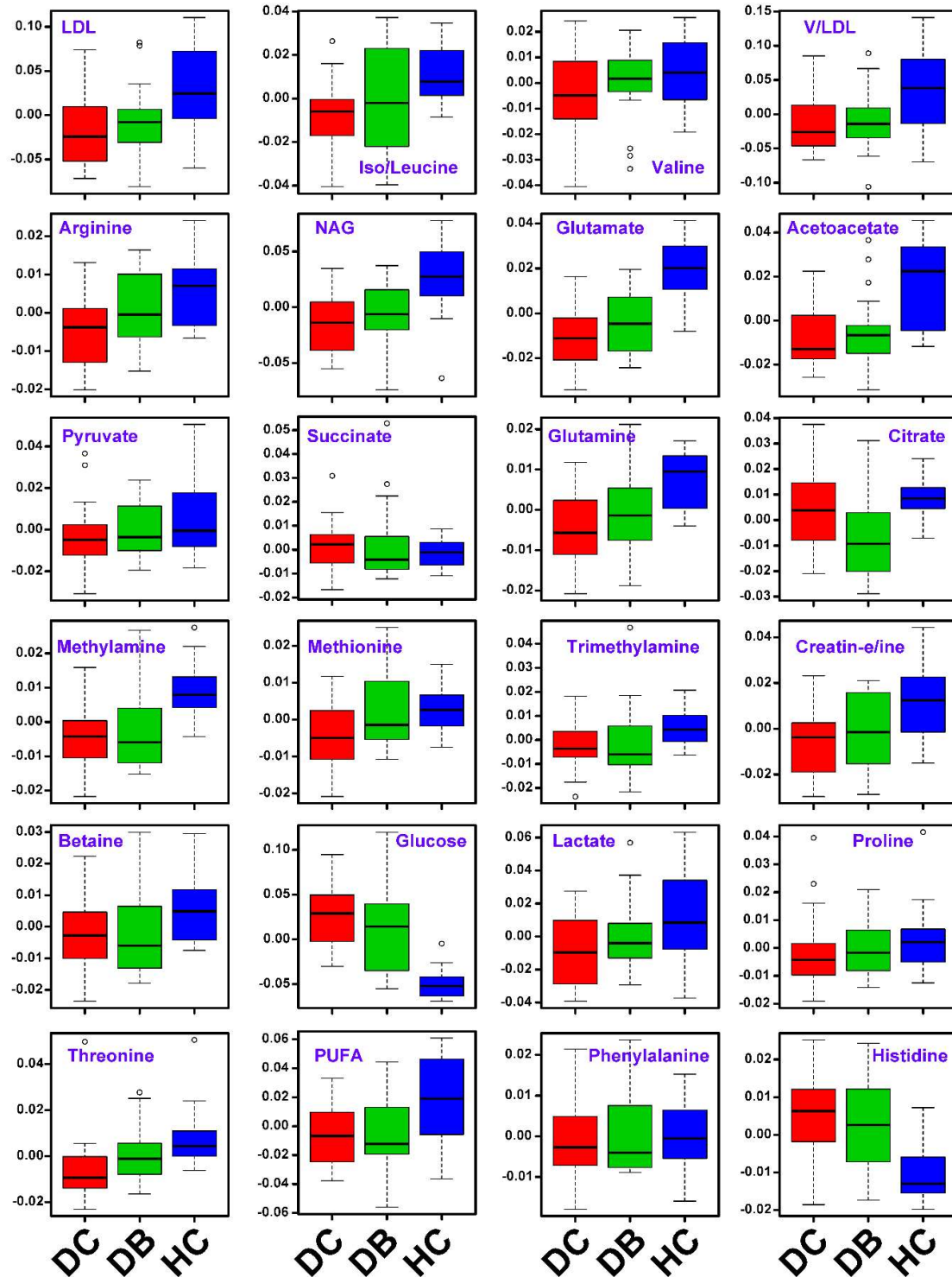
Figure 3. Box Plot's.

Box-whisker plots of metabolites that were significantly perturbed across the groups derived from 1D 1H CPMG NMR spectra using RF analysis. In the box plots, the boxes denote interquartile ranges, horizontal line inside the box denote the median, and bottom and top boundaries of boxes are 25th and 75th percentiles, respectively. Lower and upper whiskers are 5th and 95th percentiles, respectively.

The increased and decreased levels of metabolites in the body fluids can differentiate between healthy and diseased states. Compared with those from healthy controls, serum samples of patients with DB and DC showed (a) increased levels of glucose, methionine, histidine (b) by the decreased levels of LDL/VLDL, lipids, polyunsaturated lipids, N-acetyl glycoproteins(NAG), lipids, lactate, pyruvate, citrate, succinate, acetoacetate, creatinine/creatine, and most of the amino acids (like alanine, valine, isoleucine, proline, glutamate, glutamine, arginine, histidine, and phenylalanine, etc). Most of the metabolites in DB and DC followed a similar pattern in a matter of their increased or decreased levels of serum with respect to the control group. However, when compared to the DB, the DC patients have an increased levels of succinate, citrate, glucose, threonine, PUFA and the decreased levels of LDL/VLDL, lipids, N-acetyl glycoproteins (NAG), pyruvate, creatine, and amino acids (like valine, arginine, glutamate, methionine, and proline. The results are summarized in **Table 2** which clearly shows that protein biosynthesis, amino-acid metabolism, glucose-energy metabolism and glycerolipid metabolism are disturbed in patients with DC. Hence metabolic perturbation in patients with DC has been discussed in detail in the discussion section.

Table 2: Metabolic differences between the DB and DC with respect to HC; and DC with respect to DB. The up (↑) and down (↓) arrows represent, increased and decreased levels of the metabolites respectively.

#	^{1}H ppm	Metabolite	DB vs. HC	DC vs. HC	DC vs. DB
			AUC↓↑	AUC↓↑	AUC↓↑
1	0.805-0.885	HDL/LDL	0.75 (↓)	0.71 (↓)	0.59 (↓)*
2	0.935	Iso/Leucine	0.68 (↓)*	0.84 (↓)	x (↓)
3	1.015	Valine	x (↓)	x (↓)	0.58 (↓)*
4	1.215-1.245	LDL/VLDL	0.71 (↓)	0.74 (↓)	0.57 (↓)*
5	1.665-1.705	Arginine	x (↓)	0.73 (↓)	0.63 (↓)*
6	2.015	NAG	0.78 (↓)	0.82 (↓)	0.60 (↓)*
7	2.115-2.135	Glutamate	0.86 (↓)	0.92 (↓)	0.63 (↓)*
8	2.205	Acetoacetate	x (↓)	0.86 (↓)	x (↓)
9	2.355	Pyruvate	x (↓)	x (↓)	0.57 (↓)*
10	2.385	Succinate	x(↓)	x (↑)	0.52 (↑)*
11	2.425-2.465	Glutamine	0.70 (↓)	0.84 (↓)	x (↓)
12	2.515,2.655	Citrate	0.82 (↓)	0.97 (↓)	0.70 (↑)
13	2.545	Methylamine	0.78 (↓)	0.89 (↓)	x (↑)
14	2.615	Methionine	0.59 (↓)*	x (↓)	0.68 (↓)
15	2.905	Trimethylamine	x (↓)	0.76 (↓)	x (↑)
16	3.015	Creatin-e/ine	0.71 (↓)	0.86 (↓)	0.58 (↓)*
17	3.265	Betaine	0.67 (↓)*	0.87 (↓)	x (↑)
18	3.475,3.895,4.625	Glucose	0.92 (↑)	0.98 (↑)	0.67 (↑)*
19	4.085	Lactate	x (↓)	0.70 (↓)	x (↓)
20	4.125	Proline	x (↓)	x (↓)	0.58 (↓)*
21	4.225-4.245	Threonine	0.76 (↓)	0.94 (↓)	0.53 (↑)*
22	5.245-5.325	PUFA	0.68 (↓)	0.79 (↓)	0.64 (↑)
23	7.295	Phenylalanine	0.53 (↓)*	0.85 (↓)	x (↑)
24	7.025, 7.725	Histidine	0.80 (↑)	0.87 (↑)	x (↑)

Note: * = p-Value > 0.05, x = MDA value < 0.001 (ie. Metabolite not significant in the respective group).

3.3 Metabolic Pathway Analysis:

To gain insight into the metabolic mechanism of DC, the metabolic pathways of the significantly altered metabolites (identified through RF analysis) were analyzed using the "pathway analysis" module within the MetaboAnalyst software. This web-based Metabolic Pathway Analysis (MetPA) function implemented in MetaboAnalyst, It (MetPA) enables identification of altered metabolic pathways, metabolomic data interpretation, and visulization[22]. The metabolites as enlisted in **Table S1** were only used for the detailed analysis of the perturbed metabolites in DC compared to HC and DB, metabolites such as lipids (V/LDL, PUFA), NAG, are not recognized by the programme, hence were excluded from the list. The final list of altered metabolites were uploaded and analyzed in MetPA. The built-in human (Homo sapiens) pathway library and pathway analysis algorithm -hypergeometric test and -Relative-betweenness Centrality was employed for Over-Representation Analysis and Pathway Topology Analysis respectively. All matched pathways are shown according to their p-values from the pathway enrichment analysis (vertical axis or y-axis, the intensity of color) and pathway impact values (horizontal axis or x-axis, the size of circle) from pathway topology analysis, with the most impacted pathways colored in red, results as shown graphically in **Figure 4**. The metabolic pathways with an impact value greater than 0.1 were considered to be target pathways with high significance to disease morbidity. These pathways can be considered for further exploration through biochemical analysis. Taken together, DC patients have distinctive metabolic pathways compared to HC and DB patients.
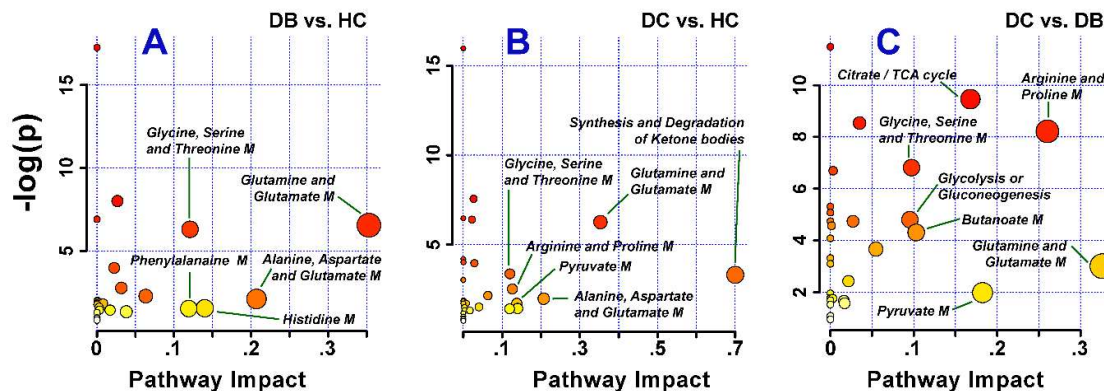

Figure 4. Pathway Analysis.

Identification of the perturbed metabolic pathways by overrepresentation analysis (ORA) using the significantly altered metabolites identified by RF VIP score. The analysis was done by using a pathway library restricted to Humans, and p-values for ORA stand for hypergeometric test. Test p-value (vertical axis, intensity of color) and impact factor (horizontal axis, size of circle).

3.4 Correlation analysis:

Metabolites are fundamental to the biological system, to evaluate the complex interaction of perturbed metabolites we also performed correlation analysis (CA). CA is important to find the asymptomatic post diabetic associated complications among diabetic individual on the basis of metabolic profiling. The identified significantly changed serum metabolites in DC patients were

subjected to CA. The Colour Heat Map was designed using Pearson's correlation coefficients analysis in statistical analysis Module of Metaboanalyst of 15 serum metabolites in DC patients to find the interrelation as depicted in **Figure 5**. To analyze the result of Color Heat Map we have taken only those metabolites which have a correlation score of $r > 0.61$ (**Table S2).** From this correlation color Heat Map we found that the amino acid intermediates of TCA cycle, glucogenic amino acids and methylamine metabolism were predominantly involved viz. succinate, citrate, as well as glutamate, arginine, and valine and methylamine are the most significantly positive correlated metabolites. Moreover, to this the correlated metabolites and the intermediated of TCA cycle will undoubtedly mark a print on asymptomatic post diabetic complication and probably these metabolites will be used as a prognostic marker for the prior onset of disease. This correlation map and metabolites will be a boon for future therapeutic strategies and also patient care beforehand.
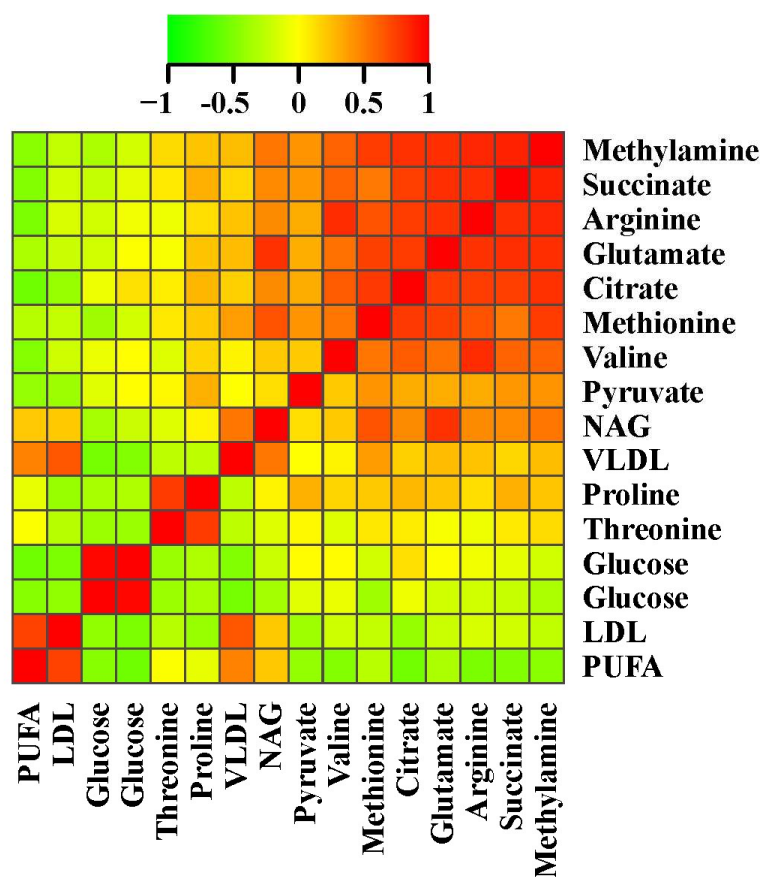

Figure 5. Correlation Analysis.

A color heat map of the Pearson's correlation coefficients computed for the 15 metabolites observed in the DC patients. The colors refer to the correlation coefficient ranging from 1 (red) to -1 (green).

4. Discussion

Diabetes is a lifestyle-related disorder, and thus it is crucial to focus efforts on the identification of biomarkers, which can help in the diagnosis and early intervention of the disease. In recent times metabolomics has emerged as an imperative tool in diabetes research[13, 23]. The present

study used the NMR based metabolomics coupled with Randon Forest to identify the discriminatory metabolites. Our results show significant variations in metabolic profiles of patients with DC compared to DB, suggesting that NMR-based metabolomics has a definite say in monitoring the metabolic signatures associated with DC. As evident from **Table 2**, 26 NMR-based metabolite entities were found to be perturbed in sera of DB and DC patients compared to HC, whereas 15 metabolites were found to be perturbed in DC as compared to DB patients. The major group of altered metabolites in the sera of DC patients contained (a) intermediates of the tricarboxylic acid cycle (TCA cycle), products of glycolysis and energy metabolism (such as glucose, lactate, pyruvate, succinate, citrate, creatine and creatinine), (b) amino acids (glutamate, valine, proline, arginine) and (c) molecules related to lipid and membrane metabolism like lipoproteins (V/LDL), PUFA, and NAG. The implications of these metabolic differences and pathways in the pathophysiology of the DC have been discussed further.

4.1 Energy Metabolism

An elevated glucose is the metabolic hallmark of diabetes, and is due to the enhanced glucose production from noncarbohydrate sources (gluconeogenesis) and the breakdown of glycogen (glycogenolysis). As evident from our study, the DC patients has relatively higher levels of glucose compared to DB and HC. The increased level of glucose followed by decreased levels of lactate suggests disturbed glucose metabolism with dampened aerobic glycolytic activity in DC patients[24]. The elevated glucose (hyperglycemia) as evident could be one of the possible reasons inducing oxidative stress in such patients, a feature common to all cell types that are damaged by hyperglycemia owing to the increased production of Reactive oxygen species (ROS) and Reactive nitrogen species (RNS)[25]. Oxidative stress (imbalance between pro-oxidants and antioxidants) causes oxidation of various biomolecules including protein, lipids, carbohydrates, DNA, and mitochondrial DNA (mtDNA) within the body[26-27]. The damage to mitochondria further increases the production of ROS, rendering it susceptible to oxidative damage and mutations[28]. Mitochondrial dysfunction aggravates the energy crisis since that ATP is produced mainly from aerobic glucose oxidation in mitochondria. Creatine and creatinine plays a pivotal role in energy homeostasis, their depleted levels further support the notion of disorder of energy metabolism, moreover they also reflect the deteriorating condition of the DC patients, as evident in various reports creatine has protective effects against hyperglycemia and is an antioxidant[29-30], whereas creatinine is a marker of low skeletal muscle mass and insulin resistance[31]. With respect to HC, DB and DC patients have decreased levels of energy metabolites, pyruvate, succinate, and citrate. Whereas, the decreased levels of pyruvate to the increased levels of succinate and citrate in DC compared to the DB patients, suggest that TCA cycle is altered to cause the accumulation of intermediate products of the TCA cycle and impaired aerobic glycolysis and, thus dampened oxidative phosphorylation and ATP production in such patients. As evident TCA cycle and its intermediates play an eminent role in the metabolic and physiological state of the co-morbid patients.

4.2 Lipid metabolism

Lipids are one of the primary sources of energy for metabolism. Lipid abnormalities are common in diabetic patients and have been reported to be the major predictors of metabolic disturbances and critical medical conditions, such as dyslipidemia, hypertension, diabetes and

cardiovascular diseases[32]. The serum samples of DC patients have significantly decreased levels of HDL, V/LDL and increased levels of PUFA compared to DB and HC. The increased levels of PUFA are thought to be widely beneficial to human health, because of their association with mitigation of the inflammatory response in disease conditions such as autoimmune, heart, arthritis and diabetes[33]. The possible reason for decreased serum levels of LDL could be due to peroxidation of LDL to Ox-LDL which is aberrantly involved in inflammatory processes through the formation of higher molecular weight complexes with distinct inflammatory mediators[34]. It is one of the significant risk factors for cardiovascular diseases and is also regarded as comorbidity affecting the development of kidney complications in diabetes[35-36] and is also influential in the progression of diabetic neuropathy[37]. Further acetoacetate was also found to be down-regulated and its decrease reflects a disturbed lipid metabolism, for its reliance on ketone bodies to meets its energy requirements in diabetic patients[38]. Taken together, the lipid metabolic profiles in sera of DC patients suggested dysregulation of lipid metabolism and metabolic alterations in energy production. In addition to changes in lipoprotein levels, residual signals of N-acetyl glycoproteins (NAG) were found to be decreased in DC patients compared to DB and HC. N-acetyl glycoproteins (N-acetylglucosamine and N-acetylgalactosamine glycan moieties) are mainly acute phase proteins with anti-inflammatory properties and are expressed more during inflammation and immune responses, these proteins also help to protect the body from oxidative stress owing to their antioxidant properties[39]. Furthermore, methylamine, trimethylamine, and betaine are related to methylamine metabolism and are processed by gut microbiota[40]. The decrease in the serum levels of these metabolites suggests a reduced number and/or altered gut microbiota metabolism. Their decrease in the serum suggests liver dysfunction consistent in several human and animal studies[41]. Consequently, the observed reduction in major structural lipids may be a consequence of increased fatty acid β-oxidation and reduced bioavailability of fatty acids for phospholipid/structural lipid synthesis, followed by decreased levels of anti-inflammatory proteins owing to oxidative damage and hyperglycemia.

4.3 Amino acid Metabolism

Amino acids are small organic molecules that play a key metabolic and physiological roles in all living organisms. They serve as substrates for protein synthesis, glyconeogenesis, ureagenesis and precursors for syntheses of hormones and low-molecular-weight nitrogenous substances and various other catabolic processes[42]. Metabolic alterations in amino acids levels (a) essential, non-essential, conditionally essential, (b) glucogenic and ketogenic, (c) branched-chain amino acids (BCAA) and aromatic amino acids (AAA) in the serum are common in conditions of disease (e.g., cancer, diabetes)[43-44]. Amino acids serve as a key source of energy, especially during conditions in which glucose availability is limited. Their significance is not limited to this only but are involved in protein biosynthesis as well as biosynthesis of several biogenic amines essential for survival in conditions of acute stress[45]. Muscles along with liver release a high quantity of amino acids present in the body to maintain the cellular homeostasis in conditions of energy deprivation[46]. The glucogenic amino acids are broken down into pyruvate, alpha-ketoglutarate, succinyl-CoA, fumarate, and oxaloacetate that can be predominantly converted into glucose or glycogen via TCA cycle or gluconeogenesis. As evident from our study the up and down-regulated metabolites thus suggest perturbed glycolysis and beta-oxidation pathway. Consistent with this report, our results also indicate decreased levels of several amino acids in

the sera of DC patients (such as glutamate, glutamine, methionine, arginine, proline, BCAA - isoleucine, leucine, valine) and increased levels of threonine and AAA-histidine, phenylalanine, suggesting the aberrant amino acid catabolism and protein biosynthesis. The BCAA and AAA are catabolized to replenish the depleted levels of TCA cycle intermediates or as acetyl derivatives to generate energy during stress suggesting dampened glycolysis alongside muscle and protein break down[47]. Serum levels of amino acids, particularly BCAA and AAA, have long been recognized as a marker of obesity-associated insulin resistance[48-49]. A noteworthy finding from our study was that the levels of proline and arginine were significantly lower in DC patients compared to DB and HC. Arginine being the most versatile amino acids in cells serves as a precursor for the synthesis of proteins, nitric oxide, urea, polyamines, proline, glutamate, and creatine[50]. Proline is a key precursor in the synthesis of skeletal muscle protein collagen which is a primary response to repair an injury[51]. Arginine and proline as found through studies play a pivotal role in wound healing when supplemented in diabetic patients. Thus their decreased level shows that the healing capacity is more perturbed in DC patients as compared to DB and HC[52]. The decreased levels of glutamate and glutamine suggest that they are being actively metabolized to meet the energy requirements of the cell via gluconeogenesis. Both glutamate and glutamine along with methionine are also important components of the pathway, which leads to the synthesis of glutathione, an important intracellular antioxidant rendering cells more susceptible to oxidative stress.

5. Conclusions

In the present study, ^{1}H NMR-based serum metabolomics combined with random forest analysis was used to identify the discriminatory metabolites. About 24 metabolites were found be responsible for differentiating DC and DB independently from HC with high reliability. However, 15 metabolites were identified as the discriminatory biomarkers of DC from DB. Our findings suggest that progression to DC is characterized by increased oxidative stress, perturbed states of inflammation, altered lipid metabolism and gut microbiota. More importantly, these findings serve as a basis for the identification of metabolic states in conjunction with metabolic pathways that can be used as diagnostic and prognostic indicators and also identify novel therapeutic targets owing to the perturbed metabolic pathways in pathophysiology of DC.

Acknowledgment

Declaration of Conflicting Interests

The authors declare no conflicts of interest.

Funding

None.

References

1.   Parhofer, K. G., Interaction between glucose and lipid metabolism: more than diabetic dyslipidemia. *Diabetes & metabolism journal* **2015,** *39* (5), 353-362.

2.   Yoon, J.-W.; Jun, H.-S., Autoimmune destruction of pancreatic β cells. *American journal of therapeutics* **2005,** *12* (6), 580-591.

3.   Prentki, M.; Nolan, C. J., Islet β cell failure in type 2 diabetes. *The Journal of clinical investigation* **2006,** *116* (7), 1802-1812.

4.   Tabish, S. A., Is diabetes becoming the biggest epidemic of the twenty-first century? *International Journal of health sciences* **2007,** *1* (2), V.

5.   Saxena, M.; Srivastava, N.; Banerjee, M., Cytokine gene variants as predictors of type 2 diabetes mellitus. *Current diabetes reviews* **2017**.

6.   Ramachandran, A.; Snehalatha, C.; Shetty, A. S.; Nanditha, A., Trends in prevalence of diabetes in Asian countries. *World journal of diabetes* **2012,** *3* (6), 110.

7.   Piette, J. D.; Kerr, E. A., The impact of comorbid chronic conditions on diabetes care. *Diabetes care* **2006,** *29* (3), 725-731.

8.   Guh, D. P.; Zhang, W.; Bansback, N.; Amarsi, Z.; Birmingham, C. L.; Anis, A. H., The incidence of co-morbidities related to obesity and overweight: a systematic review and meta-analysis. *BMC public health* **2009,** *9* (1), 88.

9.   Salek, R. M.; Maguire, M. L.; Bentley, E.; Rubtsov, D. V.; Hough, T.; Cheeseman, M.; Nunez, D.; Sweatman, B. C.; Haselden, J. N.; Cox, R., A metabolomic comparison of urinary changes in type 2 diabetes in mouse, rat, and human. *Physiological genomics* **2007,** *29* (2), 99-108.

10.   Trikkalinou, A.; Papazafiropoulou, A. K.; Melidonis, A., Type 2 diabetes and quality of life. *World journal of diabetes* **2017,** *8* (4), 120.

11.   Dunn, W. B.; Broadhurst, D. I.; Atherton, H. J.; Goodacre, R.; Griffin, J. L., Systems level studies of mammalian metabolomes: the roles of mass spectrometry and nuclear magnetic resonance spectroscopy. *Chemical Society Reviews* **2011,** *40* (1), 387-426.

12.   Guasch-Ferré, M.; Hruby, A.; Toledo, E.; Clish, C. B.; Martínez-González, M. A.; Salas-Salvadó, J.; Hu, F. B., Metabolomics in prediabetes and diabetes: a systematic review and meta-analysis. *Diabetes care* **2016,** *39* (5), 833-846.

13.   Floegel, A.; Stefan, N.; Yu, Z.; Mühlenbruch, K.; Drogan, D.; Joost, H.-G.; Fritsche, A.; Häring, H.-U.; de Angelis, M. H.; Peters, A., Identification of serum metabolites associated with risk of type 2 diabetes using a targeted metabolomic approach. *Diabetes* **2013,** *62* (2), 639-648.

14.   Suhre, K., Metabolic profiling in diabetes. *Journal of endocrinology* **2014,** *221* (3), R75-R85.

15.   Gowda, G. N.; Zhang, S.; Gu, H.; Asiago, V.; Shanaiah, N.; Raftery, D., Metabolomics-based methods for early disease diagnostics. *Expert review of molecular diagnostics* **2008,** *8* (5), 617-633.

16.   Emwas, A.-H. M., The strengths and weaknesses of NMR spectroscopy and mass spectrometry with particular focus on metabolomics research. *Metabonomics: Methods and Protocols* **2015**, 161-193.

17.   Chenomx, N., Suite 7.0.(2010). *Chenomx Inc*.

18.   Ulrich, E. L.; Akutsu, H.; Doreleijers, J. F.; Harano, Y.; Ioannidis, Y. E.; Lin, J.; Livny, M.; Mading, S.; Maziuk, D.; Miller, Z., BioMagResBank. *Nucleic acids research* **2008,** *36* (suppl 1), D402-D408.

19.   Wishart, D. S.; Tzur, D.; Knox, C.; Eisner, R.; Guo, A. C.; Young, N.; Cheng, D.; Jewell, K.; Arndt, D.; Sawhney, S., HMDB: the human metabolome database. *Nucleic acids research* **2007,** *35* (suppl 1), D521-D526.

20.   Xia, J.; Psychogios, N.; Young, N.; Wishart, D. S., MetaboAnalyst: a web server for metabolomic data analysis and interpretation. *Nucleic acids research* **2009,** *37* (suppl 2), W652-W660.

21.   Xia, J.; Psychogios, N.; Young, N.; Wishart, D. S., MetaboAnalyst: a web server for metabolomic data analysis and interpretation. *Nucleic acids research* **2009,** *37* (suppl_2), W652-W660.

22.   Xia, J.; Wishart, D. S., MetPA: a web-based metabolomics tool for pathway analysis and visualization. *Bioinformatics* **2010,** *26* (18), 2342-2344.

23.   Padilha, K.; Venturini, G.; de Farias Pires, T.; Horimoto, A. R.; Malagrino, P. A.; Gois, T. C.; Kiers, B.; Oliveira, C. M.; de Oliveira Alvim, R.; Blatt, C., Serum metabolomics profile of type 2 diabetes mellitus in a Brazilian rural population. *Metabolomics* **2016,** *12* (10), 156.

24.   Cox, K.; Cocchi, M. N.; Salciccioli, J. D.; Carney, E.; Howell, M.; Donnino, M. W., Prevalence and significance of lactic acidosis in diabetic ketoacidosis. *Journal of critical care* **2012,** *27* (2), 132-137.

25.   Asmat, U.; Abad, K.; Ismail, K., Diabetes mellitus and oxidative stress—a concise review. *Saudi Pharmaceutical Journal* **2016,** *24* (5), 547-553.

26.   Rahal, A.; Kumar, A.; Singh, V.; Yadav, B.; Tiwari, R.; Chakraborty, S.; Dhama, K., Oxidative stress, prooxidants, and antioxidants: the interplay. *BioMed research international* **2014,** *2014*.

27.   Liemburg-Apers, D. C.; Willems, P. H.; Koopman, W. J.; Grefte, S., Interactions between mitochondrial reactive oxygen species and cellular glucose metabolism. *Archives of toxicology* **2015,** *89* (8), 1209-1226.

28.   Kwak, S. H.; Park, K. S.; Lee, K. U.; Lee, H. K., Mitochondrial metabolism and diabetes. *Journal of diabetes investigation* **2010,** *1* (5), 161-169.

29.   Rocic, B.; Bajuk, N. B.; Rocic, P.; Weber, D.; Boras, J.; Lovrencic, M. V., Comparison of antihyperglycemic effects of creatine and metformin in type II diabetic patients. *Clinical & Investigative Medicine* **2009,** *32* (6), 322-326.

30.   Mahmood, R.; Qasim, N. M., Protective effect of creatine against hyperglycemia induced oxidative damage in human erythrocytes. *Free Radical Biology and Medicine* **2017,** *112*, 32.

31.   Hjelmesæth, J.; Røislien, J.; Nordstrand, N.; Hofsø, D.; Hager, H.; Hartmann, A., Low serum creatinine is associated with type 2 diabetes in morbidly obese women and men: a cross-sectional study. *BMC endocrine disorders* **2010,** *10* (1), 6.

32.   Park, S.; Sadanala, K. C.; Kim, E.-K., A metabolomic approach to understanding the metabolic link between obesity and diabetes. *Molecules and cells* **2015,** *38* (7), 587.

33.   Wang, Y.; Huang, F., N-3 polyunsaturated fatty acids and inflammation in obesity: local effect and systemic benefit. *BioMed research international* **2015,** *2015*.

34.   Rawat, A.; Srivastava, R. K.; Dubey, D.; Guleria, A.; Singh, S.; Prakash, A.; Modi, D. R.; Khetrapal, C. L.; Tiwari, S.; Kumar, D., Serum Metabolic Disturbances Hailing in Initial Hours of Acute Myocardial Infarction Elucidated by NMR based Metabolomics. *Current Metabolomics 5* (1), 55-67.

35.   Gao, S.; Liu, J., Association between circulating oxidized low-density lipoprotein and atherosclerotic cardiovascular disease. *Chronic Diseases and Translational Medicine* **2017**.

36.   Florens, N.; Calzada, C.; Lyasko, E.; Juillard, L.; Soulage, C. O., Modified lipids and lipoproteins in chronic kidney disease: A new class of uremic toxins. *Toxins* **2016,** *8* (12), 376.

37.   Vincent, A. M.; Hayes, J. M.; McLean, L. L.; Vivekanandan-Giri, A.; Pennathur, S.; Feldman, E. L., Dyslipidemia-induced neuropathy in mice. *Diabetes* **2009,** *58* (10), 2376-2385.

38.   Mahendran, Y.; Vangipurapu, J.; Cederberg, H.; Stančáková, A.; Pihlajamäki, J.; Soininen, P.; Kangas, A. J.; Paananen, J.; Civelek, M.; Saleem, N. K., Association of ketone body levels with hyperglycemia and type 2 diabetes in 9,398 Finnish men. *Diabetes* **2013**, DB_121363.

39.   Dijk, W.; Mackiewicz, A., Interleukin‐6‐Type Cytokine‐induced Changes in Acute Phase Protein Glycosylationa. *Annals of the New York Academy of Sciences* **1995,** *762* (1), 319-330.

40.   Ruan, Z.; Yang, Y.; Wen, Y.; Zhou, Y.; Fu, X.; Ding, S.; Liu, G.; Yao, K.; Wu, X.; Deng, Z., Metabolomic analysis of amino acid and fat metabolism in rats with l-tryptophan supplementation. *Amino Acids* **2014,** *46* (12), 2681-2691.

41.   Rawat, A.; Dubey, D.; Guleria, A.; Kumar, U.; Keshari, A. K.; Chaturvedi, S.; Prakash, A.; Saha, S.; Kumar, D., 1H NMR-based serum metabolomics reveals erythromycin-induced liver toxicity in albino Wistar rats. *Journal of pharmacy & bioallied sciences 8* (4), 327.

42.   Wu, G., Amino acids: metabolism, functions, and nutrition. *Amino acids* **2009,** *37* (1), 1-17.

43.   Siomkajło, M.; Rybka, J.; Mierzchała-Pasierb, M.; Gamian, A.; Stankiewicz-Olczyk, J.; Bolanowski, M.; Daroszewski, J., Specific plasma amino acid disturbances associated with metabolic syndrome. *Endocrine* **2017,** *58* (3), 553-562.

44.   Bi, X.; Henry, C., Plasma-free amino acid profiles are predictors of cancer and diabetes development. *Nutrition & diabetes* **2017,** *7* (3), e249.

45.   Medina, M. Á.; Urdiales, J. L.; Rodríguez-Caso, C.; Ramírez, F. J.; Sánchez-Jiménez, F., Biogenic amines and polyamines: similar biochemistry for different physiological missions and biomedical applications. *Critical reviews in biochemistry and molecular biology* **2003,** *38* (1), 23-59.

46.   Evans, R. D.; Heather, L. C., Metabolic pathways and abnormalities. *Surgery-Oxford International Edition* **2016,** *34* (6), 266-272.

47.   Monirujjaman, M.; Ferdouse, A., Metabolic and physiological roles of branched-chain amino acids. *Advances in Molecular Biology* **2014,** *2014*.

48.   Würtz, P.; Tiainen, M.; Mäkinen, V.-P.; Kangas, A. J.; Soininen, P.; Saltevo, J.; Keinänen-Kiukaanniemi, S.; Mäntyselkä, P.; Lehtimäki, T.; Laakso, M., Circulating metabolite predictors of glycemia in middle-aged men and women. *Diabetes care* **2012,** *35* (8), 1749-1756.

49.   Würtz, P.; Soininen, P.; Kangas, A. J.; Rönnemaa, T.; Lehtimäki, T.; Kähönen, M.; Viikari, J. S.; Raitakari, O. T.; Ala-Korpela, M., Branched-chain and aromatic amino acids are predictors of insulin resistance in young adults. *Diabetes care* **2013,** *36* (3), 648-655.

50.   Lind, D. S., Arginine and cancer. *The Journal of nutrition* **2004,** *134* (10), 2837S-2841S.

51.   Ponrasu, T.; Jamuna, S.; Mathew, A.; Madhukumar, K. N.; Ganeshkumar, M.; Iyappan, K.; Suguna, L., Efficacy of L-proline administration on the early responses during cutaneous wound healing in rats. *Amino acids* **2013,** *45* (1), 179-189.

52.   Raynaud-Simon, A.; Belabed, L.; Le Naour, G.; Marc, J.; Capron, F.; Cynober, L.; Darquy, S., Arginine plus proline supplementation elicits metabolic adaptation that favors wound healing in diabetic rats. *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology* **2012,** *303* (10), R1053-R1061.

.