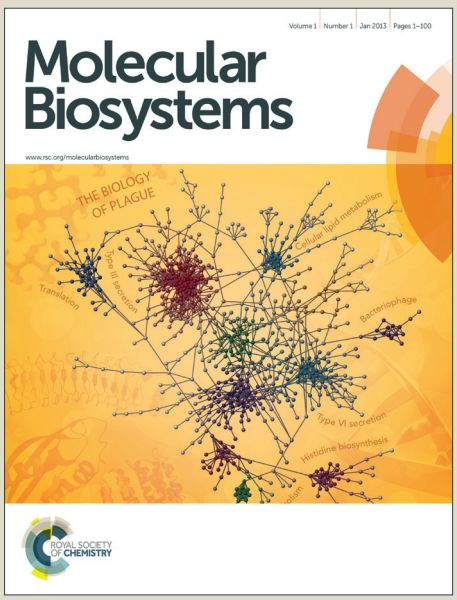
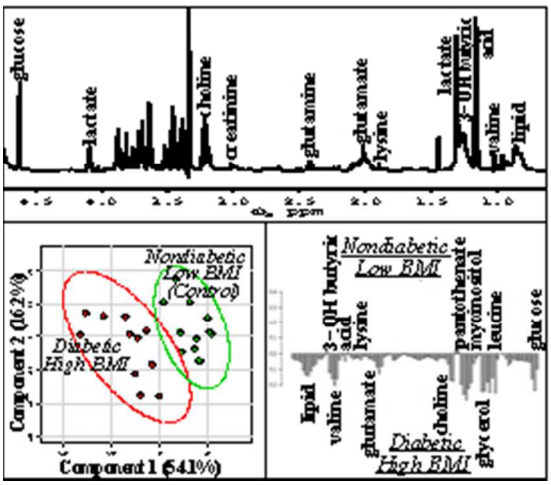
Molecular BioSystems

Accepted Manuscript



ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/molecularbiosystems



46x39mm (150 x 151 DPI)

ARTICLE TYPE www. rsc.org/xxxxxx | XXXXXXXX

Investigating correlations in the altered metabolic profiles of obese and diabetic subjects in a South Indian Asian population using an NMR-based metabolomic approach †

Navdeep Gogna[a], Murahari Krishna[b], Anup Mammen Oommen[b], and Kavita Dorai*[a]

Received Xth XXXXXXXXXX 20XX, Accepted Xth XXXXXXXXX 20XX
First published on the web Xth XXXXXXXXXX 200X
DOI: 10.1039/b000000x

It is well known that obesity/high body mass index (BMI) plays a key role in the evolution of insulin resistance and Type-2 diabetes mellitus (T2DM). However, the exact mechanism underlying its contribution is still not fully understood. This work focuses on an NMR-based metabolomic investigation of the serum profiles of diabetic, obese South Indian Asian subjects. ^{1}H 1D and 2D NMR experiments were performed to profile the altered metabolic patterns of obese diabetic subjects and multivariate statistical methods were used to identify metabolites that contributed significantly to the differences in the samples of four different subject groups: diabetic and non-diabetic with low and high BMIs. Our analysis revealed that the T2DM-high BMI group has higher concentrations of saturated fatty acids, certain amino acids (leucine, isoleucine, lysine, proline, threonine, valine, glutamine, phenylalanine, histidine), lactic acid, 3-hydroxybutyric acid, choline, 3,7-dimethyluric acid, pantothenic acid, myoinositol, sorbitol, glycerol, and glucose, as compared to the non-diabetic-low BMI (control) group. Of these 19 identified significant metabolites, the levels of saturated fatty acids, lactate, valine, isoleucine, and phenylalanine are also higher in obese non-diabetic subjects as compared to control subjects, implying that this set of metabolites could be identified as potential biomarkers for the onset of diabetes in subjects with a high BMI. Our work validates the utility of NMR-based metabolomics in conjunction with multivariate statistical analysis to provide insights into the underlying metabolic pathways that are perturbed in diabetic subjects with a high BMI.

1 Introduction

Diabetes has become an epidemic of the modern world and caused 4.8 million deaths in 2012 alone. Current trends envisage a proportional increase of diabetic patients in developing countries and the past decade has witnessed a dramatic increase in the diabetic population in these countries.[1] While it is well-known that the predominant triggers for T2DM are related to lifestyle factors such as diet, lack of physical activity, obesity and stress, a detailed mechanistic picture of the pathogenesis of T2DM is yet to emerge. Metabolomics is an emerging science that can characterize the change in metabolite patterns to identify multiple tissue pathophysiology underlying diabetes. NMR spectroscopy based metabolomic investigations have led to a fuller understanding of the dysregulation in metabolism and the environmental factors that contribute to complex diseases such as T2DM.[2–6] The utility of metabolomic profiling in the diagnosis and treatment of obesity and diabetes has been comprehensively reviewed recently.[7,8] Several NMR metabolomics studies have been published that profiled multiple metabolites correlated with insulin resistance,[9] obesity [10] and T2DM.[11–14] Recent NMR studies of serum metabolites in diabetic rats have correlated the time-varying metabolite levels with glycolysis, the TCA cycle and branched-chain amino acid metabolism in the evolution of diabetes.[15,16]

There have been several investigations that correlate phenotype and metabolism and helped in identifying serum or plasma metabolic markers that are involved separately in obesity[17,18] and in diabetes.[19,20] It is well known that obesity (high BMI) plays a key role in insulin resistance and T2DM; however the exact mechanism underlying this role is still not fully understood. Several unifying hypotheses have identified obesity as a risk factor for the development of T2DM: for instance, "metainflammation" i.e. dysregulated adipokine profile with chronic low grade inflammation due to excess adiposity[21] as well as "metabolic derangement" due to excess lipid flux and ectopic fat storage exemplify such risk factors.[22] Despite the increasing prevalence of diabetes in the South Asian population, there have been only a few attempts at systematic metabolic profiling in this population .[23,24] The prevalence of diabetes has been shown to be positively correlated with in-

[a] *Indian Institute of Science Education & Research (IISER) Mohali, Knowledge City Sector 81, Mohali PO Manauli 140306 Punjab India. Tel: +91-172-2240266; E-mail: kavita@iisermohali.ac.in*
[b] *Connexios Life Sciences Pvt Ltd, Shilpa Vidya, JP Nagar, Bangalore 560078 Karnataka India.*

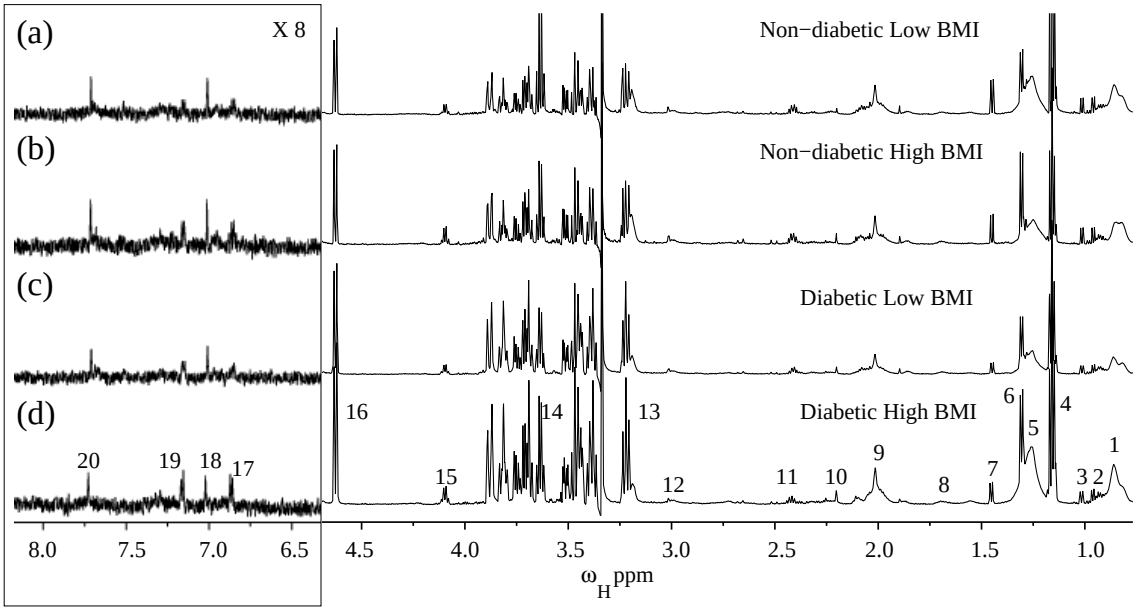


Fig. 1 Serum ^{1}H 600 MHz 1D solution state NMR spectrum of (a) T2DM-High BMI, (b) T2DM-Low BMI, (c) Non-Diabetic High BMI and (d) Non-Diabetic Low BMI subjects, with labeling of specific resonances of metabolites. Peaks numbering: 1,Lipid terminal CH3 (0.91ppm); 2,Isoleucine (0.94ppm); 3,Valine (0.98ppm); 4,3-hydroxybutyric acid (1.19ppm); 5,Alloisoleucine (1.29ppm); 6,Lactate (1.33ppm); 7,Chenodeoxycholic acid (1.49ppm); 8,Lysine (1.68ppm); 9,Glutamate (2.12ppm); 10,Ketoleucine (2.13ppm); 11,Glutamine (2.44ppm); 12,Creatinine (3.03ppm); 13,Choline (3.21ppm); 14,Ribitol (3.65ppm); 15,Lactate (4.12ppm); 16,-glucose (4.63ppm); 17,Tyrosine (6.87ppm); 18,Histidine (7.05ppm); 19,Tyrosine (7.18ppm); 20,Histidine (7.73ppm).

creasing BMI across multiple ethnic groups,[25,26] while other studies have divulged the positive correlation of visceral fat in Asian Indians to dyslipidemia and insulin resistance even in the non-obese BMI category.[27]

2 Results

2.1 Clinical characteristics

The subjects were grouped as non-diabetic, pre-diabetic (IFG : FPG of $\geq 5.6- \leq 6.9$ mM/l or $\geq 100- \leq 125$ mg/dl and IGT : 2hr PPPG of $\geq 7.8- \leq 11.0$ mM/l or ≥ 140 and ≤ 200 mg/dl) and T2DM (FPG > 7.0 mmol/l and PPPG > 11.1 mmol/l) based on clinical history and routine biochemical tests that included FPG and PPPG (WHO definition). The population consisted of 128 non-diabetics, 17 pre-diabetics and 165 diabetics with each group comprising equal numbers of both sexes. All subjects were aged between 35 to 45 years. The diabetics had a documented history of 5 years or more of the disease. The subjects were further classified with respect to their BMI into low BMI (18-23 kg/m2), moderate BMI (23-27 kg/m2) and high BMI (27-30 kg/m2) groups (Table 1). Both T2DM and non-diabetic groups were divided into insulin-sensitive and insulin-resistant groups, based on HOMA-IR. The cut-off range for HbA1c and FBS in the T2DM group was kept at $> 8.0\%$ and > 160 mg/dl respectively. HbA1C levels between the T2DM insulin-resistant low BMI and high BMI groups was statistically significantly different ($p = 0.010$). Samples from a total of 55 subjects (both male and female) were screened for NMR analysis, of which 26 were non-diabetics and 29 were T2DM. A further classification of these two groups was performed based on BMI, before proceeding for the NMR experiments.

Table 1 Clinical characteristics of the study populations

Subjects	Age (Yrs)	BMI	FBS	HbA1c%
Non Diabetic Low BMI-IS	40.75 ± 1.96	21.23 ± 1.38	95.00 ± 0.00	6.37 ± 1.09
Non Diabetic High BMI-IS	40.46 ± 2.63	28.20 ± 0.79	93.31 ± 6.24	5.65 ± 0.31
T2DM Low BMI-IR	40.20 ± 2.86	20.86 ± 1.49	241.47 ± 70.50	9.40 ± 0.24
T2DM High BMI-IR	39.57 ± 2.59	28.46 ± 1.24	223.64 ± 48.81	10.83± 1.87

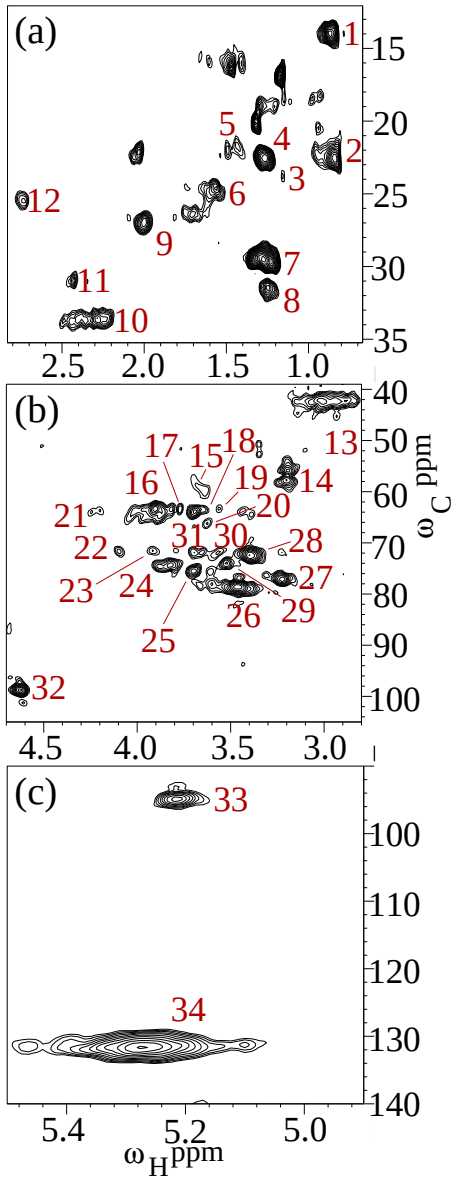


Fig. 2 Serum 2D HMQC NMR spectrum of Non-Diabetic-Low BMI subjects, showing metabolites with peaks labeled (as detailed in text) in the regions 0-2.5 ppm, 3.0-4.5 ppm and 5.0-5.5 ppm respectively.

2.2 NMR Analysis

Serum NMR spectra revealed the presence of a wide variety of metabolites. Representative 600 MHz ^{1}H NMR spectra of serum samples from non-diabetic and T2DM subjects with low and high BMI, are shown in Figure 1. The single pulse proton experiment for serum has dominance of broad resonances from lipoproteins and plasma proteins, masking and making the identification and quantification of small molecular weight metabolites difficult. This problem was overcome by using the CPMG pulse sequence to attenuate the broad resonances associated with macromolecules and other molecules bound to them, that have shorter T_2 relaxation times. CPMG NMR spectra are useful for providing relative quantitative information for low molecular weight metabolites, when samples are recorded under identical experimental conditions.[28] Metabolites were identified based on comparison with standard reference databases (HMDB, MMCD and BMRB) as well as with 2D ^{1}H-^{1}H COSY and ^{1}H-^{13}C HSQC and HMQC NMR experiments. The detailed validation of metabolite fingerprinting using cross-peaks in 2D spectra is illustrated by the representative serum 2D ^{1}H-^{13}C HMQC spectrum of Non-Diabetic-Low BMI subjects shown in Figure 2: metabolites are labeled with peak numbering in the region 0-2.5 ppm according to (a): 1, Lipid terminal CH3; 2, Leucine; 3, Valine; 4, Lactate; 5, Leucine; 6, Lipid; 7, Isoleucine; 8, Lipid; 9, Lipid; 10, Valine; 11, Glutamine/Glutamate; 12, Lipid; in the region 3.0-4.5 ppm (b): 13, Asparagine; 14, Choline; 15, Leucine; 16, beta-glucose; 17, Glycerol; 18, Glycogen; 19, Histidine; 20, Isoleucine; 21, Proline; 22, Threonine; 23, Phosphocholine; 24, Glycogen; 25, alpha-glucose; 26, beta-glucose; 27, Myoinositol; 28, alpha-glucose; 29, Choline; 30, Phosphocholine; 31, Ethanolanine; 32, alpha-glucose; and in the region 5.0-5.5 ppm (c): 33, alpha-glucose; 34, Lipid. The 2D HSQC and COSY NMR spectra are given in Figures S1-S2 of the Electronic Supplementary Information (ESI). The results of the metabolite identification performed using 1D and 2D NMR experiments is summarized in Table ST1 (Electronic Supplementary Information), which contains metabolites broadly grouped according to the biomolecular groups to which they belong. NMR metabolite fingerprinting was able to identify a large number of amino acids and carbohydrates along with lipids, keto acids, sugar alcohols etc. The presence of glycerol in the serum samples was confirmed from 1D and 2D NMR spectra and 100% matching peak scores in Metabohunter output. Glycerol peaks (3.78 ppm) overlap with some of the peaks of glucose and we observed from the loading plot and VIP scores that the bin containing these overlapping peaks was significant in contributing to separation between the groups. In order to identify the individual contributions of glucose and glycerol to group separation, we determined the concentrations of both the metabolites. We followed the principle that each metabolite could have more than one peak in the NMR spectrum and the quantitative contribution of each peak is determined by its stoichiometric properties and hence any resolved, non-overlapping peak can be integrated for metabolite quantification. If one metabolite in an overlapped region has a well-resolved peak in some other spectral region, then a simple arithmetic subtraction of its peak area (based on stoichiometric ratio) from the total integral of overlapping region

can give the integral and concentration of the other metabolite in the overlapped region. The concentration of glucose was first determined by using its isolated peak at 4.63 ppm and then arithmetic subtraction of the peak area for glucose at 3.78 ppm from the total peak area was done to calculate the concentration of glycerol.[28,29]

2.3 Multivariate Analysis

Multivariate analysis was performed to correlate diabetes and obesity-induced changes in metabolism and to identify potential metabolic pathways associated with onset of the disease. In order to identify metabolic markers associated with both diabetes and obesity, the serum samples from T2DM-High BMI subjects, were compared with the Non-Diabetic-Low BMI group (used as control in our analysis). The binned data was first analyzed by PCA, which enabled detection of any outliers (located outside the 95% confidence region of the model). Outliers were then removed from further analysis. This was followed by PLS-DA analysis. The PLS-DA score plot (Figure 3(a)) with component one explaining 54.1% of the variation and component two accounting for 16.2% of the variation, shows a clear separation between these two groups. The loadings plot given in Figure 3(b) displays the variables (in bins) responsible for the clear separation in the score plot. The metabolites in the bins were identified from 1D and 2D NMR analysis. The negative regions in the loading plot (below the baseline) correspond to metabolite levels that are increased in the serum of the T2DM-High-BMI group, whereas the positive regions (above the baseline) correspond to metabolite levels that are decreased in the serum of the T2DM-High-BMI group, as compared to the serum of the control group. The VIP parameter was also used to identify variables that explain most of the variance between metabolomic profiles of the T2DM-High BMI and Non-Diabetic-Low BMI groups (Figure 3(d)). The T2DM-High BMI group has higher concentrations of saturated fatty acids, certain amino acids (leucine, isoleucine, lysine, proline, threonine, valine, glutamine, asparagine, phenylalanine, histidine), lactic acid, 3-hydroxybutyric acid, choline, 3,7-dimethyluric acid, pantothenic acid, myoinositol, sorbitol, glycerol, and glucose, as compared to the Non-Diabetic-Low BMI (control) group. Validation of the PLS-DA model, performed using permutation test showed that the PLS-DA model is robust and credible (Figure 3(c)). The PLS-DA analysis showed distinct separation ($R^2Y = 0.86$) and good predictability (Q^2=0.73). In order to confirm that the variation in concentration of the identified metabolites in the loadings plot was statistically significant, a t-test was done to identify the important features with a threshold of $p < 0.05$. Of all the metabolites identified from the loadings plot and VIP scores, the concentration variation of 19 metabolites was found to be statistically significant between the groups and hence contributed significantly to group separation (Table 2).

Table 2 Relative concentrations (normalized integrals (%)) of serum metabolites present in T2DM-High BMI (DH), T2DM-Low BMI (DL), Non-diabetic-High BMI (NDH) and Non Diabetic-Low BMI (NDL) groups. Data are represented as mean ± SD and a t-test and ANOVA analysis were performed to confirm statistical significance (labels $a-g$ explained in text)

Metabolite	ppm	DL	DH	NDL	NDH
Saturated fatty acids	0.91	3.20^{d} ± 0.86	3.32^{b} ± 0.81	2.79 ± 0.69	$2.97^{f,g}$ ± 0.64
Valine	0.98	1.69^{d} ± 0.32	1.79^{b} ± 0.44	1.37 ± 0.55	$1.43^{f,g}$ ± 0.41
3-hydroxy-butyric acid	1.20	2.59^{c} ± 0.44	2.72^{a} ± 0.51	2.03 ± 0.54	2.37^{e} ± 0.56
Isoleucine	1.24	1.98^{d} ± 0.42	2.29^{b} ± 0.46	1.59 ± 0.41	$1.77^{f,g}$ ± 0.50
Lactate	1.33	2.16^{d} ± 0.61	2.27^{b} ± 0.84	1.68 ± 0.49	$1.83^{f,g}$ ± 0.58
Lysine	1.46	1.25^{c} ± 0.33	1.34^{a} ± 0.49	0.84 ± 0.29	0.98^{e} ± 0.35
Glutamine	2.14	1.48^{c} ± 0.62	1.85^{a} ± 0.69	0.98 ± 0.47	1.27^{e} ± 0.34
Choline	3.19	1.06^{c} ± 0.41	1.45^{a} ± 0.47	0.98 ± 0.38	0.85^{e} ± 0.38
Phenylalanine	3.30	1.95^{d} ± 0.27	2.12^{b} ± 0.47	1.54 ± 0.36	$1.71^{f,g}$ ± 0.34
Proline	3.34	1.16^{c} ± 0.31	1.36^{a} ± 0.35	0.92 ± 0.30	0.97^{e} ± 0.34
3,7-dimethy-luric acid	3.39	1.61^{c} ± 0.56	1.83^{a} ± 0.54	1.34 ± 0.48	1.45^{e} ± 0.49
Pantothenic acid	3.43	1.24^{c} ± 0.43	1.42^{a} ± 0.55	0.91 ± 0.32	1.02^{e} ± 0.46
Myoinositol	3.52	1.58^{c} ± 0.35	1.69^{a} ± 0.38	1.25 ± 0.18	1.36^{e} ± 0.36
Threonine	3.58	1.40^{c} ± 0.21	1.67^{a} ± 0.28	0.91 ± 0.17	1.13^{e} ± 0.31
Leucine	3.73	1.47^{c} ± 0.41	1.65^{a} ± 0.69	1.22 ± 0.46	1.31^{e} ± 0.44
Sorbitol	3.75	1.77^{c} ± 0.66	2.26^{a} ± 0.79	1.07 ± 0.67	1.17^{e} ± 0.70
Glycerol	3.78	2.22^{c} ± 0.68	2.37^{a} ± 0.71	1.26 ± 0.61	1.44^{e} ± 0.69
Glucose	4.63	5.38^{c} ± 0.78	5.58^{a} ± 0.99	4.55 ± 0.87	4.83^{e} ± 0.91
Histidine	7.73	1.80^{c} ± 0.38	1.91^{a} ± 0.52	1.53 ± 0.36	1.67^{e} ± 0.25

The list of 19 statistically significant metabolites (with rela-

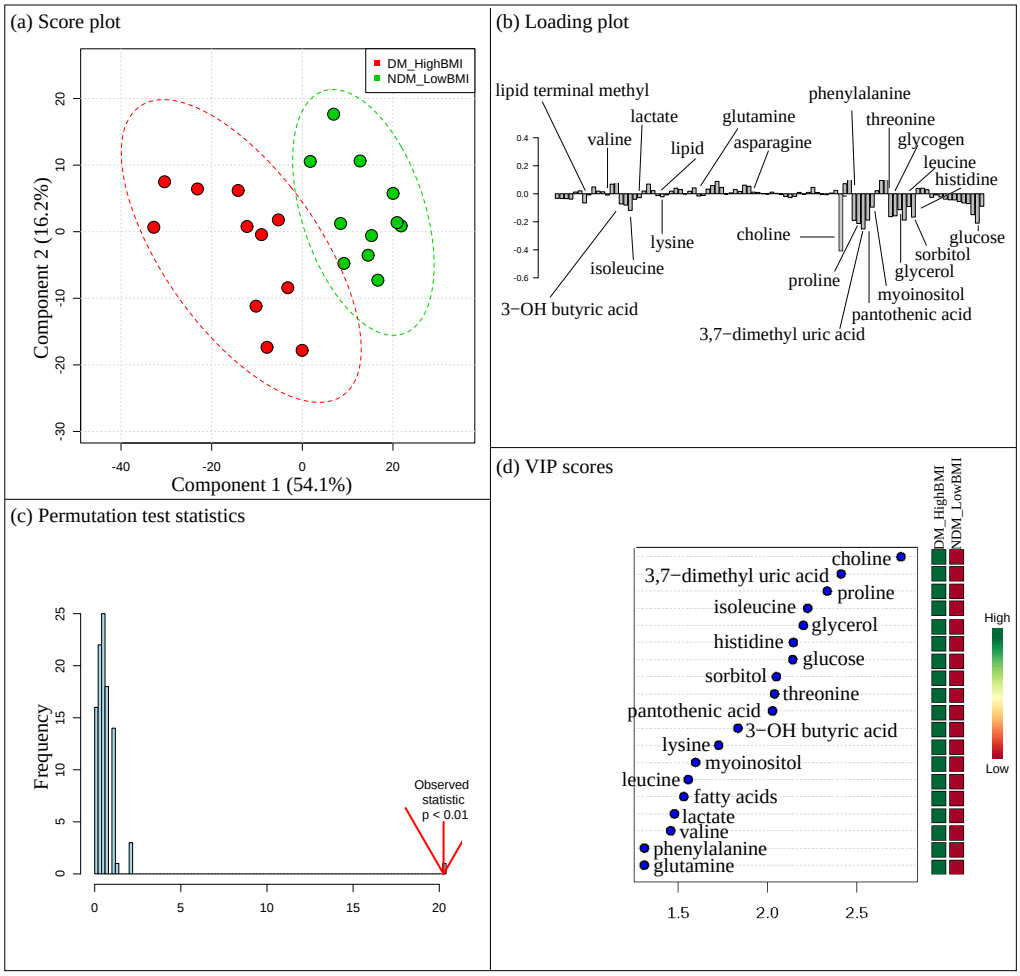


Fig. 3 (a) PLS-DA score plot of serum NMR spectra from T2DM-High BMI and Non-Diabetic-Low BMI subjects (R^2Y=0.86, Q^2=0.73). (b) Loadings plot with metabolites marked below the baseline being present in greater amounts and those marked above the baseline being present in lesser amounts in T2DM-High BMI subjects as compared to the control. (c) PLS-DA model validation by permutation tests based on separation distance ($p < 0.01$). (d) (VIP) scores with the corresponding expression heat map. Green and red indicate increased and decreased metabolite levels, respectively. The marked metabolites have t-test scores of $p < 0.05$.

tive integral concentrations) that contribute to differences between all four groups is tabulated in Table 2. The t test and ANOVA analysis revealed the following differences (labeled in Table 2): a, b label metabolites with significant differences between T2DM-High BMI and Non diabetic-Low BMI groups ($p < 0.01$ and $p < 0.05$ respectively) using the t-test; c, d label metabolites with significant differences between T2DM-Low BMI and Non diabetic-Low BMI groups ($p < 0.01$ and $p < 0.05$ respectively) using ANOVA; e, f label metabolites with significant differences between T2DM-High BMI and Non diabetic-High BMI groups ($p < 0.01$ and $p < 0.05$ respectively) using ANOVA; and g labels metabolites with significant differences between Non Diabetic-High BMI and Non diabetic-Low BMI groups ($p < 0.05$) using ANOVA.

The 19 significant metabolites identified above, are associated with both diabetes and with high BMI/obesity. In order to identify metabolites that can be specifically correlated with obesity (regardless of whether the subject is T2DM or not), we looked at the clustering of the T2DM-High BMI and Non-Diabetic-High BMI groups, with the Non-Diabetic-Low-BMI group acting as control (Figure 4). The PLS-DA score plot (Figure 4(a)) with component one explaining 50.4% of the variation and component two accounting for 16.8% of the variation, shows a clustering of both the Non-Diabetic groupings with some separation from the T2DM High-BMI group. However, there is some overlap of the Non-diabetic-High BMI group with the T2DM-High BMI group as well, indicating that BMI indeed contributes to group separation

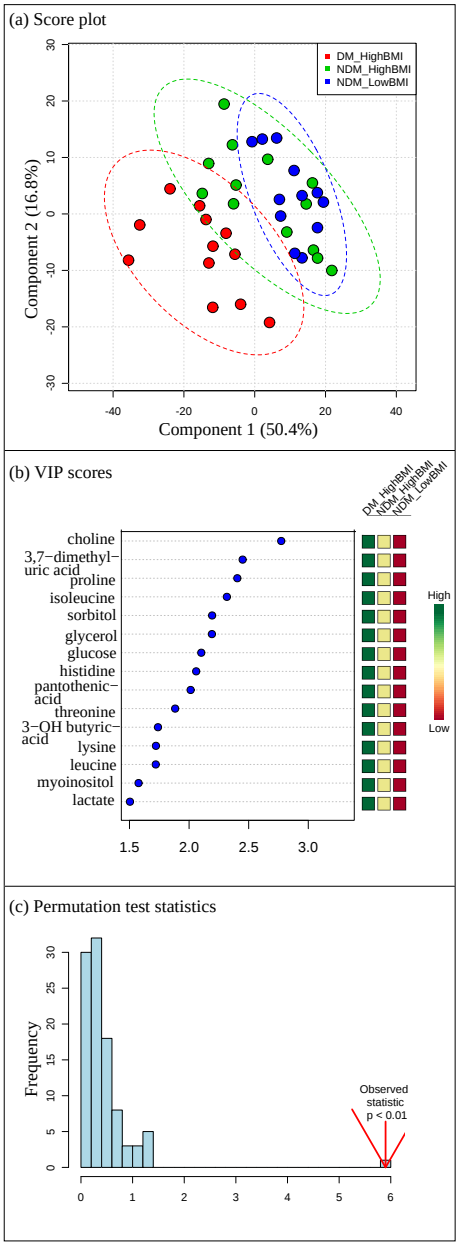


Fig. 4 (a) PLS-DA score plot of serum NMR spectra from T2DM-High BMI, Non-Diabetic-High BMI and Non Diabetic-Low BMI (control) subjects (R^2Y=0.81, Q^2=0.68), (b) VIP scores from PLS-DA models with an expression heat map; the colors red and green indicate decreased and increased metabolite levels respectively ($p < 0.01$ using ANOVA), and (c) PLSDA model validation by permutation tests based on separation distance between the three groups. The p value based on permutation is $p < 0.01$.

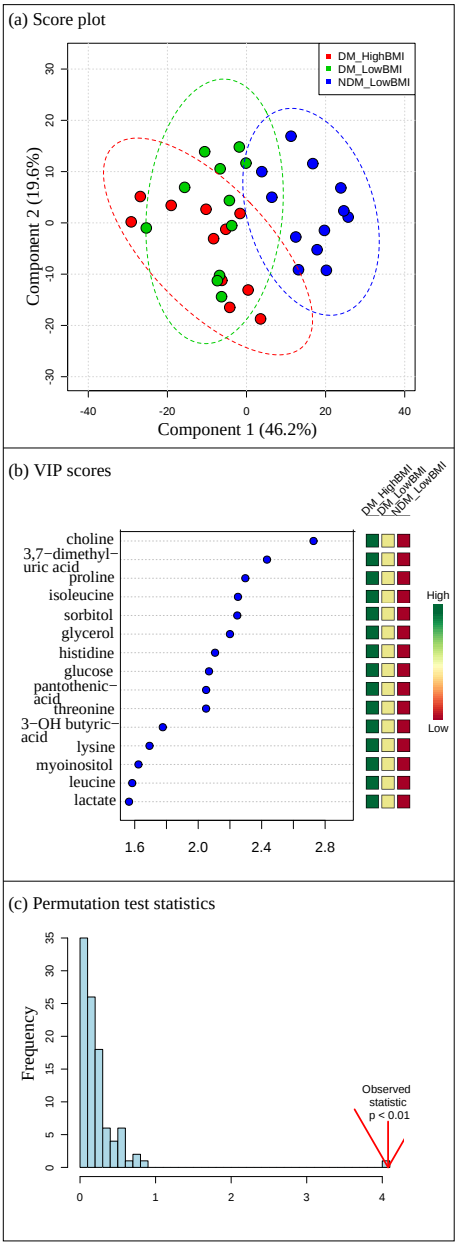


Fig. 5 (a) PLS-DA score plot of serum NMR spectra from T2DM-High BMI, T2DM-Low BMI and Non Diabetic-Low BMI (control) subjects (R^2Y=0.81, Q^2=0.63), (b) VIP scores from PLS-DA models with an expression heat map; the colors red and green indicate decreased and increased metabolite levels respectively ($p < 0.01$ using ANOVA), and (c) PLSDA model validation by permutation tests based on separation distance between the three groups. The p value based on permutation is $p < 0.01$.

and that the metabolic dysregulation occurring due to high BMI/obesity can be correlated with dysregulation due to diabetes. The PLSDA model was validated using permutation test and showed good separation (R^2Y= 0.81) and predictability (Q^2 = 0.68). Since the number of groups were more than two in this case, one-way ANOVA was performed to see if the overall comparison is significant or not, followed by post-hoc analysis to identify which two groups are different (Table ST2, ESI). The post-hoc analysis showed that for all the 19 previously identified metabolites, the concentration differences between (i) T2DM-High BMI and Non diabetic-Low BMI and (ii) T2DM-High BMI and Non diabetic-High BMI groups were responsible for overall group separation. However for some metabolites, the concentration differences between (i) T2DM-High BMI and Non diabetic-Low BMI, (ii) T2DM-High BMI and Non diabetic-High BMI and (iii) Non diabetic-High BMI and Non diabetic-Low BMI groups were responsible for overall group separations. The concentration of five metabolites: saturated fatty acids, lactate, valine, isoleucine, and phenylalanine, varied significantly in all three groups, with the highest concentrations in the T2DM-High BMI group, followed by the Non diabetic-High BMI group and lowest concentrations in the Non diabetic-Low BMI group. These five metabolites also show up in the serum of obese but non-diabetic subjects, and can hence provide insights into the dysregulation of metabolic pathways during the early onset of diabetes.

We also looked at the clustering of the T2DM-High BMI and T2DM-Low BMI groups, with the Non-Diabetic-Low-BMI group acting as the control (Figure 5). The PLS-DA score plot (Figure 5(a)) with component one explaining 46.2% of the variation and component two accounting for 19.6% of the variation, shows a clustering together of both the T2DM (High BMI and Low BMI) groups, with a clear separation from the Non-Diabetic (control) group. The PLSDA model was validated using permutation test and showed good separation (R^2Y= 0.81) and predictability (Q^2 = 0.63). The high overlap between T2DM-High BMI and Low BMI groups shows that after the onset of the disease, the impact of BMI is less significant and disruption in metabolic pathways is more due to the disease than due to BMI. The metabolites contributing significantly to the clustering of the T2DM groups and separation from Non-Diabetic-Low BMI group can thus be correlated with dysregulation in the diabetic metabolism, with BMI having a perceptible but concomitantly less significant effect. ANOVA with a threshold of 0.05 was performed as before and used to identify metabolites that were statistically significant in contributing towards group separations and the same 19 metabolites were found to be responsible for group separation (Table 2). Post-hoc analysis (Fisher's LSD) for Figure 5 (Table ST3 ESI) showed that differences between (i) T2DM-High BMI and Non diabetic-Low BMI groups and (ii) T2DM-Low BMI and Non diabetic-Low BMI groups were responsible for separations with not much differentiation between both the T2DM groups, thus confirming that group separations were impacted more by diabetes. The findings detailed in Figures 3-5 indicate that alterations in metabolism are dominated by the diabetic state (the T2DM-High-BMI group is clearly separated from the control Non-Diabetic-Low-BMI group in all the score plots) and that BMI-related changes in the profiles, while contributing to group differences, are less obvious in this study.

3 Discussion

The changes in the metabolic profiles of T2DM-High-BMI, T2DM-Low-BMI and Non-Diabetic-High-BMI groups as compared to the control group (Non-Diabetic-Low-BMI), were used to make inferences about the alterations in metabolic pathways associated with obesity and T2DM in South Indian Asian subjects. Our analysis revealed a panel of statistically significant 19 metabolites and their associated metabolic pathways (Table 3), that clearly correlated with the T2DM phenotype irrespective of the BMI background. Interestingly, among these metabolites, a panel of 5 metabolites which includes saturated fatty acids, lactate, valine, isoleucine and phenyl alanine (Figures 4(a), 4(b)) shows significant correlation with the BMI phenotype both in T2DM as well as in non-diabetic subjects. Based on this observation we hypothesize that these 5 metabolites could represent dysregulation in metabolic pathways during early onset of diabetes in obese subjects. Corroborating our hypothesis there are several small scale clinical studies that reports similar alteration in plasma levels of these metabolites in obese subjects.[30–32] Figure 6 shows a coarse-grained picture of the metabolic pathways to which these metabolites belong and their interrelationships identified using the Connexios Network Knowledge Base. Figure 7 shows a schematic diagram of the possible correlation between the 19 significantly altered metabolites detected by ^{1}H NMR serum analysis, and their associated pathway, with obesity related chronic diseases such as diabetes and cardiovascular disease. The clinical significance of these 19 metabolites, based on previously published clinical and pre-clinical observation, is discussed below.

3.1 Glucose and Polyol pathway

One obvious metabolite that was significant in the serum of obese T2DM subjects in our list was D-glucose (Figures 3(b), 3(d), 4(b), 5(b)). Increased glucose is shown to increase the flux of metabolite into polyol and other glycosylation pathways contributing to several aspects of diabetes complications. Interestingly, we were also able to identify increased levels of sorbitol in T2DM phenotype (Figures 3(b), 3(d), 4(b), 5(b)),

indicating increased flux of glucose through the polyol pathway. Plasma levels of sorbitol has been shown previously to be high in T2DM subjects with an established contribution to diabetic complications.[33] The incomplete oxidative disposal of glucose, a characteristic feature of T2DM phenotype due to progressive development of insulin resistance, can also be tracked using the plasma levels of lactate.[34] Studies by Chen et. al have shown that plasma lactate levels are high in non-diabetic glucose intolerant obese subjects as well.[32] In congruence with these observations, we also report a positive correlation between the elevated levels of plasma lactate with the high BMI phenotype (Table 2, Fig 4(a)) which correlates with impaired glucose metabolism due to insulin resistance.

3.2 Ketone Body synthesis

Enhanced ketone body production due to dysregulated fat metabolism is a hall mark of advanced T2DM phenotype. Increased level of 3-hydroxy butyric acid, a major plasma marker of ketogenesis, is shown to be increased in insulin dependent diabetics as well as obese subjects .[35] Moreover norepinephrine infusion in obese T2DM subjects has shown increased levels of ketone bodies which correlated well with the insulin resistant phenotype .[36] Consistent with these studies, we also observed a positive correlation between 3-hydroxy butyric acid with the T2DM phenotype (Figures 3(b),3(d),4(b),5(b)), which could indicate dysregulated whole body fat metabolism as a result of insulin resistance.

3.3 Amino acid metabolism

Recent cross sectional studies involving participants of STRRIDE study, a large cross-sectional study of Asian Indian and Chinese men and a detailed study by Wang et. al,[11] have identified branched amino acids (BCAA) and related metabolites to be significantly correlated with insulin resistance and diabetes development.[37,38] In our study we were also able to identify BCAA metabolism related metabolites (such as isoleucine and leucine) as significantly contributing to the T2DM phenotype (Figures 3(b), 3(d), 4(b), 5(b)). Other correlations previously reported between the altered aminoacid metabolism and T2DM phenotype, such as impaired vasodilatory function due to high plasma arginase activity,[39,40] loss of lean mass,[41–43] and associated increase in amino acid metabolism,[44] add credence to our analysis in identifying a positive correlation between the increased levels of proline, histidine, lysine and threonine respectively in the serum of T2DM subjects with the diabetic phenotype (Figures 3(b), 3(d), 4(b), 5(b)). Hyperaminoacidemia, a state of increased plasma levels of amino acids, is also a characteristic feature of the obese phenotype.[30] Recent metabolic profiling studies[11,45] have established a causal relationship between the plasma elevation of branched chain amino acids and aromatic amino acids with progressive development of insulin resistance and T2DM in obese subjects. In consonance with these observations our study also highlights a significant correlation between three amino acids namely, valine, isoleucine and phenylalanine with the high BMI phenotype (Table 2, Figure 4(a)).

Table 3 List of metabolites with significantly altered concentrations associated with T2DM and high BMI (Obesity) and their related metabolic pathway

Metabolic Pathway	Key metabolites
Ketone body synthesis	3-hydroxy butyric acid
Purine degradation	3,7-dimethyl uric acid
Amino acid metabolism	Threonine Lysine Histidine Isoleucine Leucine Proline Valine Phenylalanine Glutamine
Glucose and Polyol pathway	D-glucose Sorbitol
CoA metabolism	Pantothenic acid
Choline metabolism	Choline
Inositol metabolism	Myo-inositol
Adipose lipolysis	Glycerol
Glycolysis	Lactate
Fatty acid metabolism	Saturated fatty acids

3.4 Purine Degradation

Increased levels of 3,7 dimethyl uric acid observed in the serum of T2DM subjects (Figures 3(b), 3(d), 4(b), 5(b)), might indicate an augmented purine metabolism in this phenotype. Previously, semi-ischemic forearm test in T2DM subjects has revealed augmented purine metabolism, especially hypoxanthine degradation and associated increase in serum uric acid levels.[46] This might be an adaptive mechanism to support cellular energy in the absence of adequate glucose oxidation.

3.5 CoA metabolism

One of the metabolites in our list that has not been previously reported in T2DM serum samples is Pantothenic acid (Figures 3(b), 3(d), 4(b), 5(b)). Studies in rodent models of T2DM have shown an impaired uptake of pantothenic acid into skeletal muscle as well as a parallel increase in uptake and incorporation into Coenzyme A in liver tissue.[47,48] This has been

articulated as an adaptive mechanism in response to the dysregulation in fuel oxidation, switching from muscle to liver. A similar mechanism might be cooperating in the T2DM subjects we studied, which can be further substantiated by the increase in serum glucose as well as ketone bodies observed in this phenotype.

3.6 Inositol metabolism

Myo-inositol is one of the key intracellular signaling molecules that can modulate key aspects of cell and tissue function.[49] Alteration in uptake of myo-inositol by cells can lead to its serum elevation and at the same time tissue inositol reduction can affect signaling cascades as observed in T2DM subjects.[50] In our study we observed an increase in serum myo-inositol levels in T2DM subjects, indicating a possible impairment in tissue inositol signaling pathways.

3.7 Adipose Lipolysis

Large population based studies have shown that increased levels of plasma glycerol correlate well with development of hyperglycemia and T2DM.[51] One of the primary sources for this elevated level of plasma glycerol is adipose tissue. Insulin resistance that develops in adipose tissue during the diabetic condition, results in altered ability of insulin signaling cascade to store triglyceride in adipose tissue and leads to uncontrolled release of free fatty acids (NEFA) and glycerol.[52] An elevated level of glycerol, observed in our study (Figures 3(b), 3(d), 4(b), 5(b)), correlates well with established T2DM phenotype and indicates peripheral insulin resistance in these subjects. Interestingly, in our study saturated fatty acids also show a significant correlation with the High BMI category in addition to the T2DM phenotype (Table 2, Fig 4(a)). Previous studies have clearly correlated the rise in plasma levels of saturated fatty acids or NEFA, to the insulin resistant state in adipocytes under obese and type 2 diabetic conditions,[10] and it is one metabolite that has an established causal relationship with obesity to the T2DM phenotype.[53]

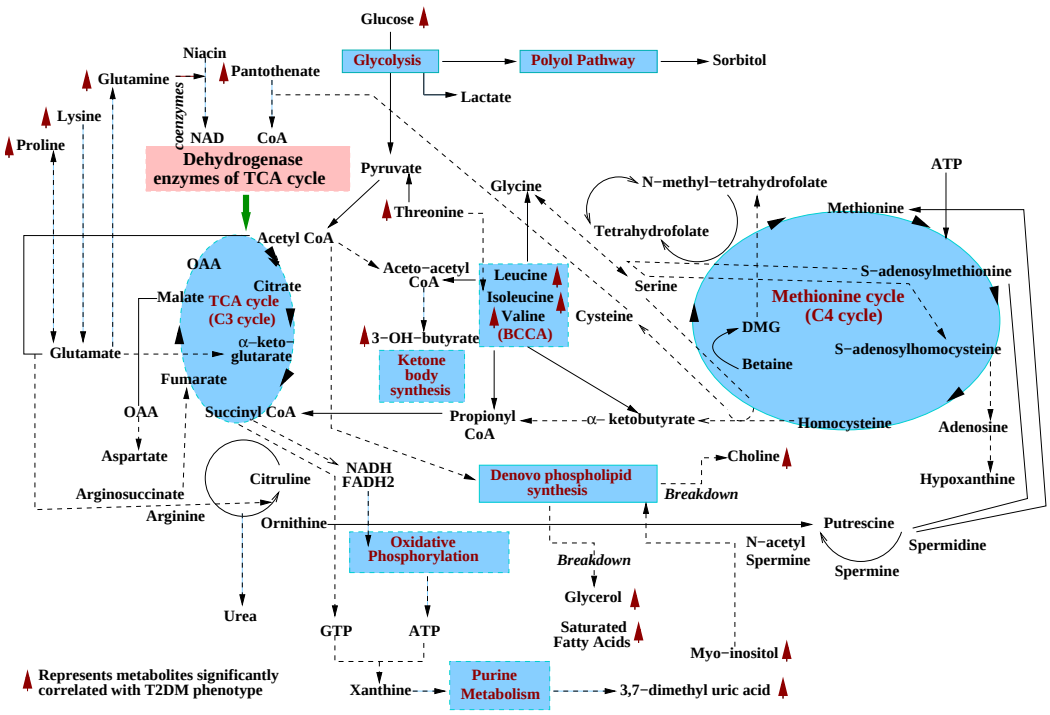


Fig. 6 Coarse-grained pictorial depiction of the metabolic pathways and their inter-relationships, useful in locating key metabolites significantly expressed in the serum of T2DM subjects.

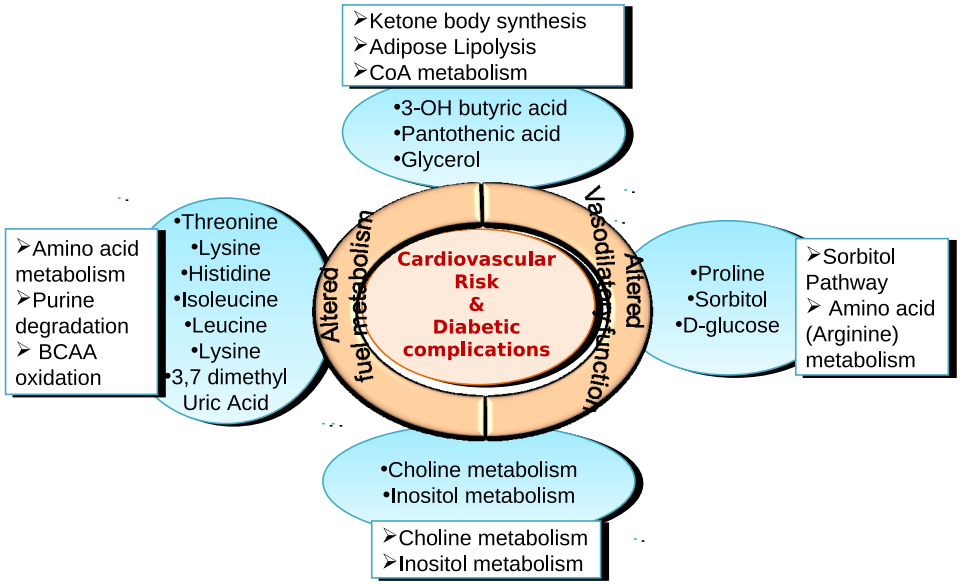


Fig. 7 Schematic diagram of the dysregulated metabolic pathways associated with the T2DM high-BMI phenotype and significant metabolites (identified from serum ^{1}H NMR).

3.8 Choline Metabolism

Choline is one of the major precursors for synthesis of phospholipids as well as for betaine (a metabolite involved in methylation). Elevation in choline to betaine ratio has been positively correlated with components of metabolic syndrome.[54] In our study an elevated level of choline was observed in T2DM subjects, which might indicate an alteration in phospholipid metabolites as well as an increased cardiovascular risk in T2DM subjects[55] due to defective choline metabolism.

4 Experimental

4.1 Ethics Statement

The study protocol was approved by a local Ethics Committee, Bangalore (ChairPerson, Dr Manorama Thomas) and all subjects provided their informed consent before entering the study.

4.2 Study Setting

The study was conducted at an out-patient stand-alone clinic set-up specifically for screening, examining and documenting study subjects and performing phlebotomy.

4.3 Study Population

A total of 316 subjects (South Indian Asians) were enrolled in the study. Subject recruitment was done by a combination of advertisement at diagnostic centers, and through the clinics of general practitioners. The subjects were invited to come to the facility of the sponsor. The first step was the consent process. After explaining the details of the study, a comprehensive case history was recorded on a semi-structured, close-ended pro forma. All subjects with Type-1 diabetes and Type-2 diabetics on insulin therapy were excluded from the study. Basic data on age, sex, education, occupation, smoking and tobacco-chewing status, alcohol consumption, diet, and physical activity were collected from all the subjects. Blood and urine samples were collected for routine laboratory assessment. The clinical manager, a physician, scrutinized the results of the clinical examination as well as the biochemical tests to determine whether they matched the inclusion and exclusion criteria in order to formally enroll the subjects into the study. For the diabetic subjects, no lifestyle or drug interventions were introduced prior to obtaining the blood samples.

4.4 Anthropometric Parameters and Biochemical Tests

A detailed clinical history including significant past medical and surgical history, personal history including dietary and exercise habits, a detailed family history with particular emphasis on indicators of the metabolic syndrome, cardiovascular disease, and cerebrovascular disease in first degree relatives, and list of current medications were recorded. This was followed by a complete physical examination by a physician. Baseline ECG and Chest-X ray were done within a week of the physical examination at a designated diagnostic centre. All subjects had an extensive biochemical examination of their blood. These tests were performed in a designated diagnostic laboratory. The blood sample was initially screened serologically for HIV, hepatitis B, hepatitis C. If the sample tests were negative for all the three viral antibodies, then subsequent biochemical and hematological testing were performed. These included liver function tests, TSH, fasting blood glucose, post-prandial blood glucose, fasting insulin, fasting lipid profile (triglycerides, HDL, LDL and total cholesterol), BUN, creatinine, C-reactive protein, lipoprotein A, homocysteine, HbA1c, serum insulin, uric acid, PT and PTT. Insulin resistance (IR) was calculated by the homeostasis model assessment (HOMA) using the following equation: IR = [fasting insulin (μIU/mL) $\times$ fasting glucose (mmol/L)]/22.5. A complete hemogram was obtained and routine urine examination and estimation of micro albumin/creatinine ratio was performed. Female subjects had their FSH and LH levels estimated.

4.5 NMR Sample Preparation

NMR studies were conducted by collecting blood from subjects in serum separator tubes and keeping the blood for 30 min to allow for clot formation. Serum was then separated from the blood clot and collected in a fresh clean tube, followed by further centrifugation at 5000 rpm for 5 min. Aliquots of serum samples were then stored under identical conditions at -80 deg C until the NMR experiments were performed. Prior to NMR analysis, serum samples were thawed and 400 micro liters aliquots were mixed with 200 micro liters of 0.9% saline solution. The saline solution of 0.9% NaCl (wt/vol) was prepared by weighing 0.9g of NaCl into a 250ml volumetric flask. To this, 10 ml of D_2O (as a field lock) was added and the volume made up to 100 ml with H_2O and shaken till the salt had completely dissolved. Each sample was then transferred into a 5 mm high quality NMR tube. NMR experiments were carried out over four consecutive days with approximately 13-14 samples being recorded per day. Sample coding (with no other information) was performed to allow samples to be randomly picked for NMR experiments. The information about the samples was collated at the end of all the NMR experiments and samples were grouped on the basis of diabetes and/or obesity to allow for fair random sampling.

4.6 NMR spectroscopy

NMR spectra were recorded on a Bruker Biospin 600 Avance-III spectrometer operating at a 1H frequency of 600.13 MHz at 300 K using a 5 mm QXI probe. Gradient shimming was performed prior to signal acquisition. 1D ^{1}H NMR spectra were acquired using the water suppressed Car-Purcell-Meiboom-Gill (CPMG) spin-echo pulse sequence optimized with a spin-echo delay τ of 300μs and $n = 400$ and a total spin-spin relaxation delay ($2n\tau$) time of 240 ms to achieve attenuation of fast-relaxing broad signals from larger molecules. The proton spectra were collected with a 90 degree pulse width of 9.15 μs, a relaxation delay of 2 s, 16 scans, 16K data points and a spectral width of 7211.54 Hz. Data were zero-filled by a factor of 2 and the FIDs were multiplied by an exponential weighting function equivalent to a line broadening of 1 Hz prior to Fourier transformation. The spectra were phase and baseline corrected and referenced to an internal methyl peak of lactate at 1.33 ppm. For resonance assignment and metabolite identification, two-dimensional NMR spectra were recorded, including 1H-1H correlation spectroscopy (COSY) and 1H-13C heteronuclear and homonuclear single quantum coherence spectroscopy (HSQC, HMQC). 2D 1H-13C HMQC and HSQC spectra were obtained with a spectral width of 12 ppm and 280 ppm in the proton and carbon dimensions respectively, 1K data points, 32 scans, 256 t_1 increments and a recycle delay of 1.5 s. The COSY spectra were acquired with a spectral width of 12 ppm in both dimensions, 2K data points, 32 scans and 128 t_1 increments.

4.7 Pattern recognition and statistical analysis

The NMR variables used for statistical analysis were processed using the software package Mnova from the MestReC Lab (`http://www.mestrec.com`). Spectra were segmented into 0.04 ppm chemical shift "bins" between 0.6 ppm and 8 ppm. Spectral regions between 4.66 and 4.8 ppm were excluded from the analysis to mitigate errors due to any residual peak from the suppressed water signal. Data were normalized to a total integral of 100 to compensate for possible differences in signal-to-noise ratios between spectra.

Univariate and multivariate statistical analysis was done using the Metaboanalyst web portal (`www.metaboanalyst.ca`) and MetaboAnalyst 2.0, a web-based suite for high-throughput metabolomic data analysis.[56] The binned NMR spectra were divided into the classes of interest for identification of significant bins. Metaboanalyst was used to get the required Discriminant Analysis between the four groups: T2DM-High BMI, T2DM-Low BMI, Non-Diabetic-High BMI and and Non-Diabetic-Low BMI. Principal component analysis (PCA) and partial least-squares discriminant analysis (PLS-DA) were performed on the binned data. Outliers in the data defined as observations located outside the 95% confidence region of the Hotelling's T2 ellipses in the PCA score plots were excluded from further analysis. PCA was followed by supervised PLS-DA; PCA detects intrinsic clusters and outliers within the data set whereas PLS-DA maximizes the class discrimination. The PLS-DA model was validated using the leave one out cross-validation (LOOCV) method and the quality of the model was assessed based on R^2 and Q^2 scores, representing the predictive capability and the explained variance respectively. Permutation analysis was further performed on the best model using 1000 permutation tests with a threshold p value of < 0.01 indicating that none of the results are better than the original one. Significant metabolites were ranked according to their variable influence on the projection (VIP) score, which is a weighted sum of squares of the PLS weights, accounting for the Y variance in each dimension. VIP analysis displays the metabolites ordered according to their influence on group separation, with metabolites arranged according to their VIP values, and the y-axis denoting their relative intensities. After the significant metabolites were identified from the PLS-DA loadings plots, the data for two-group analysis were analyzed using unpaired t-tests (for comparison of the means of two samples with equal variances) to determine the relative intensities of the significant metabolites that contribute to group separation. The data were presented as mean $\pm$ SD and a p-value < 0.05 was considered to be statistically significant. In order to determine diabetes and BMI correlations, Analysis of Variance (ANOVA) was performed. ANOVA can determine whether the comparison between multi-group analysis is significant or not. Input variables for ANOVA was based on VIP scores, where VIP > 1 indicated contribution to the model. This was followed by post-hoc analysis in order to perform all pairwise comparisons between group means by least significant difference (LSD) t-test.

4.8 Metabolite identification

The significant bins were identified and the corresponding peak lists and intensities were obtained from TopSpin2.0 (Bruker Biospin). The peak and intensities list were input to MetaboHunter,[57] for metabolite identification. MetaboHunter is a web server application that can be used for automatic assignment of ^{1}H-NMR spectra and identification of metabolites based on three different search methods and with possibility of peak drift in a user defined spectral range. The results of the search were matched with the compound libraries (HMDB) for samples in the pH range of 6-8. For every identified metabolite, its corresponding 2D NMR experimental (HSQC, HMQC and COSY) plot and peak/intensity value was cross-validated by visual inspection with the reference 2D lists from HMDB.

4.9 Metabolic Pathway Mapping

The identified metabolites were mapped to their respective pathways using the Connexios Network Biology Platform.[58] The Connexios Network Biology Platform is an integrated repository of molecular events spanning numerous cellular pathways across core metabolic tissues.

5 Conclusions

Using our Connexios Network Knowledge Base and literature evidence, we linked altered metabolites in the serum of T2DM obese subjects (identified using ^{1}H NMR), to key cellular processes which are dysregulated in diabetic conditions. We hypothesize a correlation between these 19 significant metabolites and the associated metabolic pathways with cardiovascular risk and complications underlying the diabetic condition. Increased levels of valine, isoleucine, lactate and saturated fatty acids were observed in the Non-Diabetic-High BMI category, which represent alterations in fuel oxidation that correlate well with vascular dysfunction in diabetic phenotype and could indicate an early dysregulation in amino acid, glucose and lipid metabolic pathways. Our results match well with previous studies[2] which have also noted increased levels of branched chain amino acids and pantothenic acid in diabetic subjects. Since our observations correlate well with previously reported metabolic alterations in T2DM subjects, we conclude that our study provides a robust metabolic fingerprinting of diabetic phenotype independent of ethnicity. NMR-based metabolomics of serum in conjunction with multivariate statistics is hence a viable approach to understand the underlying dysregulation of metabolic pathways associated with both diabetes and obesity. Further molecular studies are required to provide a detailed mechanistic basis for these observed dysregulations.

Acknowledgements:
The experiments were performed on a Bruker Avance-III 600 MHz FT-NMR spectrometer at the NMR Research Facility at IISER Mohali. AMO acknowledges Mr. M. Y. Reddy for help with the statistical analysis and Mrs. Rashi Vikram Patrikar and the clinical team for helping in setting up the clinical trials and Vinita Radhakrishnan, Manager of Intellectual Property at Connexios Life Sciences Pvt. Ltd., for critical reading and discussion regarding this manuscript.

References

1 L. D. Roberts, A. Koulman and J. L. Griffin, *Lancet Diab. Endocr.*, 2014, **2**, 66.
2 S. C. Connor, M. K. Hansen, R. F. Smith and T. E. Ryan, *Mol. Biosyst.*, 2010, **6**, 909.
3 I. R. Lanza, S. Zhang, L. E. Ward, H. Karakelides, D. Raftery and K. S. Nair, *PLoS One*, 2010, **5**, 10538.
4 S. Zhang, G. N. Gowda, V. Asiago, N. Shanaiah, C. Barbas and D. Raftery, *Anal. Biochem.*, 2008, **383**, 76.
5 R. Salek, K. K. Cheng and J. L. Griffin, *Meth. Enzym.*, 2011, **500**, 337.
6 Y. Yang, L. Wang, S. Wang, R. Huang, L. Zheng, S. Liang, L. Zhang and J. Xu, *Mol. Biosyst.*, 2014, **10**, 1803.
7 O. Y. Kim, J. H. Lee and G. Sweeney, *Expert Rev. Cardiovasc. Ther.*, 2013, **11**, 61.
8 J. Lu, G. Xie, W. Jia and W. Jia, *Front. Med.*, 2013, **7**, 4.
9 C. B. Newgard, J. An, J. R. Bain, M. J. Muehlbauer, R. D. Stevens, L. F. Lien, A. M. Hagg, S. H. Shah, M. Arlotto, C. A. Slentz, J. Rochon, D. Gallup, O. Ilkayeva, B. R. Wenner, W. S. Y. Jr, H. Eisenson, G. Musante, R. S. Surwit, D. S. Millington, M. D. Butler and L. P. Svetkey, *Cell Metab.*, 2009, **9**, 311–326.
10 J. Y. Kim, J. Y. Park, O. Y. Kim, B. M. Ham, H.-J. Kim, D. Y. Kwon, Y. Jang and J. H. Lee, *J. Proteome Res.*, 2010, **9**, 4368.
11 T. J. Wang, M. G. Larson, R. S. Vasan, S. Cheng, E. P. Rhee, E. McCabe, G. D. Lewis, C. S. Fox, P. F. Jacques, C. Fernandez, C. J. O'Donnell, S. A. Carr, V. K. Mootha, J. C. Florez, A. Souza, O. Melander, C. B. Clish and R. E. Gerszten, *Nature Med.*, 2011, **17**, 448–453.
12 R. M. Salek, M. L. Maguire, E. Bentley, D. V. Rubtsov, T. Hough, M. Cheeseman, D. Nunez, B. C. Sweatman, J. N. Haselden, R. D. Cox, S. C. Connor and J. L. Griffin, *Physiological Genomics*, 2007, **29**, 99–108.
13 J. L. Griffin, H. J. Atherton, C. Steinbeck and R. M. Salek, *BMC Research Notes*, 2011, **4**, 272.
14 P. M. Nissen, C. Nebel, N. Oksbjerg and H. C. Bertram, *J. Biomed. Biotech.*, 2011, **2011**, 378268.
15 M. Guan, L. Xie, C. Diao, N. Wang, W. Hu, Y. Zheng, L. Jin, Z. Yan and H. Gao, *PLoS One*, 2013, **8**, 60409.
16 C. Diao, L. Zhao, M. Guan, Y. Zheng, M. Chen, Y. Yang, L. Lin, W. Chen and H. Gao, *Mol. Biosyst.*, 2014, **10**, 686.
17 R. Williams, E. M. Lenz, A. J. Wilson, J. Granger, I. D. Wilson, H. Major, C. Stumpf and R. A. Plumb, *Mol. Biosyst.*, 2006, **2**, 174.
18 Q. He, P. Ren, X. Kong, Y. Wu, G. Wu, P. Li, F. Hao, H. Tang, F. Blachier and Y. Yin, *J. Nutri. Biochem.*, 2012, **23**, 133.
19 J. Chen, X. Zhao, J. Fritsche, P. Yin, P. Schmitt-Kopplin, W. Wang, X. Lu, H. U. Haring, E. D. Schleicher, R. Lehmann and G. Xu, *Anal. Chem.*, 2008, **80**, 1280–1289.
20 Y. Bao, T. Zhao, X. Wang, Y. Qiu, M. Su and W. Jia, *J. Proteome Res.*, 2009, **8**, 1623.
21 C. N. L. A. R. Saltiel, *J. Clin. Invest.*, 2011, **121**, 2111–2117.
22 D. M. Muoio, M. Deborah and C. B. Newgard, *Annu. Rev. Biochem.*, 2006, **75**, 367–401.
23 Y. Wang, H. J. Chen, S. Shaikh and P. Mathur, *Obes. Rev.*, 2009, **10**, 456–474.
24 D. K. Guptar, P. Shah, A. Misra, S. Bharadwaj, S. Gulati, N. Gupta, R. Sharma, R. M. Pandey and K. Goel, *PLoS One*, 2011, **6**, e17221.
25 G. Maskarinec, A. Grandinetti, G. Matsuura, S. Sharma, M. Mau, B. E. Henderson and L. N. Kolonel, *Ethnicity and Disease*, 2009, **19**, 49–55.
26 J. C. Chan, V. Malik, W. Jia, T. Kadowaki, C. S. Yajnik, K. H. Yoon and F. Hu, *J. Am. Med. Assoc.*, 2009, **301**, 2129–2140.

† Electronic Supplementary Information (ESI) available: Figures S1 and S2 display 2D HSQC and 2D COSY NMR spectra respectively for all four groups. Table ST1 shows the results of NMR metabolite fingerprinting of serum samples. Tables ST2 and ST3 display the results of post-hoc analysis for groupings of T2DM-High BMI, T2DM-Low BMI and control and T2DM-High BMI, Non-diabetic-High BMI and control subjects, respectively.

27 M. A. Banerji, N. Faridi, R. Atluri, R. L. Chaiken and H. E. Lebovitz, *J. Clin. Endocrinol. Metab.*, 1999, **84**, 137–144.

28 P. Tripathi, L. Bala, R. Saxena, S.K.Yachha, R. Roy and C. L. Khetrapal, *J. Gastrointestin. Liver Dis.*, 2009, **18**, 329.

29 J.A.Williams, R.Lalonde, J.Koup and D.D.Christ, *Predictive Approaches in Drug Discovery and Development-Biomarkers and In Vitro/In Vivo Correlations*, John Wiley and Sons, New Jersey, USA, 2012.

30 S. J. Mihalik, B. H. Goodpaster, D. E. Kelley, D. H. Chace, J. Vockley, F. G. Toledo and J. P. DeLany, *Obesity*, 2010, **18**, 1695.

31 B. Caballero, N. Finer and R. J. Wurtman, *Metabolism*, 1988, **37**, 672.

32 Y. D. Chen, B. B. Varasteh and G. M. Reaven, *Diabetes Metab.*, 1993, **19**, 348–354.

33 G. M. Preston and R. A. Calle, *Biomark. Insights*, 2010, **5**, 33–38.

34 S. O. Crawford, R. C. Hoogeveen, F. L. Brancati, B. C. Astor, C. M. Ballantyne, M. I. Schmidt and J. H. Young, *Int. J. Epidemiol.*, 2010, **39**, 1647.

35 S. E. Hall, M. E. Wastney, T. M. Bolton, J. T. Braaten and M. Berman, *J. Lipid Res.*, 1984, **25**, 1184–94.

36 B. Willms, M. Boettcher, V. Wolters, N. Sakamoto and H.-D. Soeling, *Diabetologia*, 1969, **5**, 88–96.

37 C. B. Newgard, *Cell Metab.*, 2012, **15**, 606.

38 E. S. Tai, M. L. S. Tan, R. D. Stevens, Y. L. Low, M. J. Muehlbauer, D. L. M. Goh, O. R. Ilkayeva, B. R. Wenner, J. R. Bain, J. J. M. Lee, S. C. Lim, C. M. Khoo, S. H. Shah and C. B. Newgard, *Diabetologia*, 2010, **53**, 757–767.

39 S. R. Kashyap, A. Lara, R. Zhang, Y. M. Park and R. A. DeFronzo, *Diab. Care*, 2007, **31**, 134.

40 S. M. Morris, *J. Nutr.*, 2007, **137**, 1602S.

41 T. Chen, X. Zhang, Y. Long, H. Yu, X. Ran, Y. Gao, H. Lu, X. Xie, X. Chen, Y. Ren, J. Shi and H. Tian, *J. Endocrinol. Invest.*, 2012, **35**, 772.

42 D. L. Chinkes, *Curr. Opin. Clin. Nutr. Metab. Care*, 2005, **8**, 534.

43 R. Gougeon, *Can. J. Diab.*, 2013, **37**, 115.

44 H. Norrelund, H. Wiggers, M. Halbirk, J. Frystyk, A. Flyvbjerg, H. E. Botker, O. Schmitz, J. O. Jorgensen, J. S. Christiansen and N. Moller, *J. Intern. Med.*, 2006, **260**, 11.

45 S. E. McCormack, O. Shaham, M. A. McCarthy, A. A. Deik, T. J. Wang, R. E. Gerszten, C. B. Clish, V. K. Mootha, S. K. Grinspoon and A. Fleischman, *Pediatr. Obes.*, 2013, **8**, 52.

46 Y. Tanaka, I. Hisatome, T. Kinugawa, H. Tanaka, Y. Tomikura, F. Ando, T. Matsuura, G. Igawa, K. Matsubara, Y. Yamamoto, M. Yamawaki, R. Yamamoto, K. Ogino, O. Igawa, T. Ikeda, C. Shigemasa and A. Takeda, *J. Intern. Med.*, 2003, **42**, 788–792.

47 D. K. Reibel, B. Wyse, D. A. Berkich, W. M. Palko and J. R. Neely, *Am. J. Physiol.*, 1981, **240**, E597.

48 G. Veerababu, J. Tang, R. T. Hoffman, M. C. Daniels, L. F. Hebert, E. D. Crook, R. C. Cooksey and D. A. McClain, *Diabetes*, 2000, **49**, 2070–2078.

49 B. J. Holub, *Annu. Rev. Nutr.*, 1986, **6**, 563.

50 C. Servo, *Acta Med. Scand.*, 1977, **201**, 59.

51 Y. Mahendran, H. Cederberg, J. Vangipurapu, A. J. Kangas, P. Soininen, J. Kuusisto, M. Uusitupa, M. Ala-Korpela and M. Laakso, *Diab. Care*, 2013, **36**, 3732.

52 M. T. van der Merwe, G. P. Schlaphoff, N. J. Crowther, I. H. Boyd, I. P. Gray, B. I. Joffe and P. N. Lonnroth, *J. Clin. Endocrinol. Metab.*, 2001, **86**, 3296.

53 S. E. Kahn, R. L. Hull and K. M. Utzschneider, *Nature*, 2006, **444**, 840.

54 S. V. Konstantinova, G. S. Tell, S. E. Vollset, O. Nygard, O. Bleie and P. M. Ueland, *J. Nutr.*, 2008, **138**, 914.

55 W. H. Tang, Z. Wang, B. S. Levison, R. A. Koeth, E. B. Britt, X. Fu, Y. Wu and S. L. Hazen, *New Eng. J. Med.*, 2013, **368**, 1575.

56 J. Xia, R. Mandal, I. V. Sinelnikov, D. Broadhurst and D. S. Wishart, *Nucl. Acids Res.*, 2012, **40**, W127–W133.

57 D. Tulpan, S. Leger, L. Belliveau, A. Culf and M. Cuperlovic-Culf, *BMC Bioinform.*, 2011, **12**, 400.

58 A. M. Oommen, U. Narayanan and M. R. Jagannath, *ISRN Cell Biol.*, 2012, **2012**, ID278636.