Plasma Metabonomic Profiling of Diabetic Retinopathy

Running Title: Plasma Markers of Diabetic Retinopathy

Keywords: Retinopathy, Metabonomics, Metabolomics, Metabolic Profiling, Metabolite, GC-MS

Liyan Chen, PhD, 1,2 Ching-Yu Cheng, MD, PhD, 1,3,4 Hyungwon Choi, PhD, 5,6 Mohammad Kamran Ikram, MD, PhD, 1,4 Charumathi Sabanayagam, MD, PhD, 1,3,4 Gavin SW Tan, MD, 1 Dechao Tian, MSC, 1,8 Liang Zhang, BSC, 1,8 Gopalakrishnan Venkatesan PhD, 2 E Shyong Tai, MRCP,PhD, 5,7,9, Jie Jin Wang, MD, PhD, 10 Paul Mitchell, MD, PhD, 10 Chiu Ming Gemmy Cheung, MD, 1, Roger Wilmer Beuerman,PhD 1,3,11,12 Lei Zhou, PhD, 1,3,11,* Eric Chun Yong Chan, Ph.D, 2,* Tien Yin Wong, MD, Ph.D 1,3,4,5,*

Affiliations:
1 Singapore Eye Research Institute, Singapore National Eye Centre, Singapore
2 Department of Pharmacy, Faculty of Science, National University of Singapore, Singapore
3 Department of Ophthalmology, Yong Loo Lin School of Medicine, National University of Singapore, Singapore
4 Ophthalmology and Visual Sciences Academic Clinical Research Program, Duke-NUS Graduate Medical School, Singapore
5 Saw Swee Hock School of Public Health, National University of Singapore, Singapore
6 Institute of Molecular and Cell Biology, Agency for Science, Technology and Research, Singapore
7 Department of Medicine, Yong Loo Lin School of Medicine, National University of Singapore, Singapore
8 Department of Statistics and Applied Probability, Faculty of Science, National University of Singapore
9 Cardiovascular & Metabolic Disorders Program, Duke-NUS Graduate Medical School, Singapore

10 Department of Ophthalmology, Centre for Vision Research, Westmead Millennium Institute, University of Sydney, Sydney, New South Wales, Australia

11 Signature Research Program in Neuroscience and Behavioral Disorders, Duke-NUS Graduate Medical School, Singapore

12 Department of Biochemistry, Yong Loo Lin School of Medicine, National University of Singapore, Singapore

*Corresponding authors: zhou.lei@seri.com.sg, phaccye@nus.edu.sg, ophwty@nus.edu.sg
**Abstract**

Diabetic retinopathy (DR) is the most common microvascular complication of diabetes and the leading cause of visual impairment in working-age adults. Diabetic patients often develop DR despite appropriate control of systemic risk factors, suggesting the involvement of other pathogenic factors. We hypothesize a plasma metabolic signature of DR that is distinct and resolvable from that of diabetes alone. A nested population-based case-control metabonomic study was first performed on 40 DR cases and 40 diabetic controls using gas chromatography-mass spectrometry. Eleven metabolites were found to be correlated with DR, and the majority were robust when adjusted for metabolic risk factors and confounding kidney disease. The metabolite markers 2-deoxyribononic acid; 3,4-dihydroxybutyric acid; erythritol; gluconic acid and ribose were validated in an independent sample set with 40 DR cases, 40 diabetic controls and 40 non-diabetic individuals. DR cases and diabetic controls were matched by HbA1c in the validation set. Activation of the pentose phosphate pathway was identified from the list of DR metabolite markers. The identification of novel metabolite markers for DR provides insights into potential new pathogenic pathways for this microvascular complication and holds translational value in DR risk stratification and the development of new therapeutic measures.

(195/200 words)
Introduction

Diabetic retinopathy (DR) is the most common microvascular complication of diabetes and the leading cause of visual impairment in working-age adults worldwide (1; 2). The global prevalence of diabetes is rising and the number of diabetic persons is projected to increase by 54% in 2030, compared with 2010 (3). The public health burden of diabetes and DR would thus increase correspondingly. The major risk factors of DR are poor glycemic control and hypertension, as well as the duration of diabetes (4; 5), but their relative importance varies between studies (2; 6). While the risks of DR progression and vision loss are reduced with intensive control of risk factors (7; 8), many patients with diabetes continue to develop complications despite tight glycemic and blood pressure control. There is increasing evidence to suggest that “metabolic memory” is responsible for this observation. The term metabolic memory refers to the persistent epigenetic modifications caused by early exposure to hyperglycemia that in turn, predispose individuals to development of diabetic complications even after good glycemic control is achieved (9). More recently, Zavrelova et al. found that the progression of DR in a treated cohort occurs in distinct patterns, ranging from a stable condition that is free from DR to rapid progression to the proliferative DR within six years (10). The mechanisms responsible for these patterns in disease progression are unclear (11). The discovery of biomarkers that characterize the different developmental phenotypes of DR thus become important as these biomarkers may provide insight on pathogenic pathways that are currently unknown, and may allow clinical stratification of patients for DR monitoring and treatment (5).

The variation in the metabolome represents the interplay of genetic and environmental factors, and provides information that is complementary to genomic, transcriptomic or proteomic data. The study of metabolic phenotypes (metabotypes) in association with disease states may reveal new knowledge in disease mechanism and pathophysiology (12). Recent metabonomic studies have uncovered plasma and sera metabolic signatures associated with, or predictive of impaired glucose tolerance and diabetes (13-22). Furthermore, DR is a complex disease where findings from genome-wide-association studies have not been conclusive (23). We postulate that a distinct metabolic signature of DR exists, and can be resolved from that of diabetes alone. Barba et al. identified metabolite markers
of DR in the vitreous humour (24). However, the invasive nature of vitreous sampling limits study replication and the translational potential of any biomarkers identified from vitreous fluid. In contrast, plasma or sera have remained the biofluid of choice in metabonomic studies. With these biological matrices, several research groups have reported metabolite markers of global (25; 26) and targeted (27-29) metabonomic analysis. These studies however, did not account for confounders and co-morbidities, such as medication use and kidney disease. This study’s objective is to investigate if the plasma metabotype of DR is resolvable from the metabolic perturbations associated with underlying diabetes, and to identify novel metabolite markers of DR.

Research Design and Methods

Sample Selection

We first conducted a nested population-based case-control study on 40 diabetes patients with DR and 40 diabetes patients without DR (discovery set) selected from banked plasma collected as part of the Singapore Indian Eye Study (30). Metabolite markers of DR identified from the discovery set were separately quantified in a separate sample set with 40 diabetes patients with DR, 40 diabetes patients without DR and 40 participants with no diabetes (validation set).

For the discovery set, cases were 40 participants with type 2 diabetes with moderate non-proliferative DR (level 43 on the ETDRS scale) or worse in at least one eye, documented from retinal photographs (31). Equal numbers of samples were selected from participants who had been diagnosed with diabetes for ten years or less, and for over ten years. Controls were 40 participants with diabetes who had no DR, matched to cases by diabetes duration (in 5 year bands).

For the validation set, participants with chronic kidney disease were excluded from selection. Cases were 40 participants with type 2 diabetes with moderate non-proliferative DR (level 43 on the ETDRS scale) or worse in at least one eye. Controls were 40 participants with diabetes who had no DR, matched to cases by HbA1c levels (in 0.5% bands), age (in 5 year bands).
bands) and gender. 40 participants with no diabetes were also selected for the validation set, matched to DR cases by age and gender.

**Singapore Indian Eye Study Protocol**

The Singapore Indian Eye Study was a population-based study of Indian adults aged 40 years and above living in Singapore. The study was approved by the SingHealth Institutional Review Board, conducted in accordance to the Declaration of Helsinki, and written informed consent was obtained from all participants. Detailed population selection and methodology have been published elsewhere (30). Briefly, from 2007 to 2009, the study was conducted in the southwestern part of Singapore using an age-stratified random sampling strategy. Of these, 4497 individuals were deemed eligible to participate and 3400 participants took part in the study, giving a 75.6% participation rate (31).

Participants had a comprehensive examination with standardized questionnaire, systemic and ocular examination, and had plasma and serum samples collected and stored. Diabetes was defined as self-reported of a previous diagnosis of the disease by a doctor, use of diabetic medication, or hemoglobin A1c (HbA1c) of 6.5% (48 mmol/mol) or greater (32). DR was graded from retinal photographs using a scale modified from the Airlie House classification system. Clinical and biochemical variables like body mass index (BMI), total cholesterol, low density lipoprotein (LDL) cholesterol and high density lipoprotein (HDL) cholesterol, glycosylated hemoglobin (HbA1c), and blood pressure (BP) were measured from standardized protocols. Chronic kidney disease was defined as having estimated glomerular filtration rate (eGFR) of less than 60 mL/min/1.73 m².

**Plasma Collection Protocol**

K$_2$EDTA tubes were used to collect blood samples. Tubes were kept on ice and transported to the laboratory within the same working day. Tubes were centrifuged at 1,000 g for 10 min (4°C) to separate plasma from whole blood. Plasma aliquots were stored at −80 °C.

**GC-MS Metabonomic Profiling**

Plasma samples (100 µL aliquots in microcentrifuge tubes) were thawed to room temperature (25 °C). D-27 myristic acid (20 µL of 200 µg/mL solution in methanol) and 300
µL of methanol were added to each sample. Samples were mixed for 15 min at room
temperature and centrifuged for 10 min at 16,000 g (4°C). The top 300 µL of each
supernatant was transferred to a glass tube and dried under nitrogen gas. The dried samples
were re-suspended in 100 µL of toluene, vortexed vigorously for 10 s and dried again under
nitrogen. A two-step derivatization method was employed for chemical derivatisation of
metabolites. Samples were first incubated with 50 µL of 2 % methoxyamine chloride in
pyridine (Pierce Biotechnology) for 1.5 h at 60 °C. Next, 50 µL of N-methyl-N-
trifluoroacetamide (Thermo Scientific) was added and the samples were incubated again for
1 h at 60 °C. 80 µL of derivatised samples were transferred to silanised glass vials for gas
chromatography-mass spectrometry (GC-MS) analysis. Ten aliquots of pooled K₂EDTA
human plasma (Innovative Research Inc., batch 052511), three aliquots of distilled water
and 10 mM glucose in distilled water were also similarly prepared. Two pooled plasma
samples were injected before the first study sample and one pooled plasma sample was
injected after every ten study samples to act as quality controls (QC). The blanks and glucose
standards were injected at the beginning, at the middle and at the end of the sample batch.

GC-MS analysis was performed using an Agilent 6890N gas chromatograph (Agilent
Technologies) coupled to a LECO Pegasus III (4D) GC × GC/MS time-of-flight mass analyzer
operating in GC/MS mode (Leco Corporation). Sample injection was performed by a CTC PAL
autosampler (CTC Analytics AG). The injection temperature was 250 °C and injection volume
was set to 1 µL, with a split ratio of 1:2. A capillary column with DB-1 stationary phase of
21.5 m × 0.25 mm, 0.25 µm film thickness (Agilent Technologies) was used with a constant
helium carrier gas flow of 1.5 mL/min. The temperature gradient was held at 70 °C for 2 min,
increased at a rate of 10 °C/min to 310 °C, and held at 310 °C for 6 min. The transfer line and
ion source were set at 280 and 250 °C respectively. Mass spectra were acquired after a
delay of 240 s post-injection, from m/z 45 to 600 at 20 Hz using ionization energy of 70 eV
and a detector voltage of 1650 V.

Data Processing
Deconvolution and peak finding were performed with ChromaTOF (version 4.44, Leco
Corporation) using (i) a signal-to-noise (S/N) threshold of 150, (ii) a minimum peak width of
2.5 s and (iii) a minimum of two apexing masses. Instrument performance was assessed using the peak area of D-27 myristic acid, the internal standard.

Peak alignment was performed with the “calibration” function in ChromaTOF (33). A reference table was first built using data obtained from plasma samples of participants who had declared no medication usage. There were three such participants among the selected samples. Samples from an additional six diabetic participants who had declared no medication usage were analysed for the purpose of building the reference table. Each peak was only included in the reference table if (i) it was detected in at least seven out of nine medication-free samples (78%) and (ii) its average integrated peak area was at least five times greater than that of the glucose standards and blanks. This ensured that the peaks in the reference table included endogenous metabolites as represented in the medication-free samples and excluded xenobiotics such as drugs and their associated metabolites. Peaks derived from glucose were saturated in most samples and were thus excluded from the reference table. Furthermore, plasma glucose levels in the DR group were not significantly different from diabetic controls in this study.

Peak data from DR patients and diabetic controls were then aligned to the reference table using a retention time tolerance of 3 s. Retention times were converted to Kovats retention indices using C8–C40 alkane standards (Sigma Aldrich). Peaks with missing values in more than 20% of the samples were discarded. For peaks with missing values in less than 20% of the samples, the missing values were filled with half of the lowest detected peak area. The raw peak areas was normalised to the sum of peak areas within the 90th percentile of peak area distribution in each sample, using the median sample as the denominator in the calculation of normalisation factors. Normalised data was then log transformed (base 2). Peaks with coefficient of variation (CV) greater than 30% in the QC samples were not considered for further analysis.

**Statistical Analysis**

Principal component analysis (PCA) with unit variance scaling was performed to evaluate the presence of dominating trends in the peak data. Mann-Whitney U tests were first performed to compare the DR group with diabetic controls. To correct for multiple testing,
false discovery rates (FDR) was calculated using q-values (34). Metabolites were considered significant if their q-values were less than or equal to 0.2, and if their fold-changes were greater than 1.2 or less than 0.8. The area under the curve (AUC) in receiver operating characteristic (ROC) analysis was calculated to evaluate the discriminating power of the metabolite markers (35). Logistic regression models were fit to evaluate the association of metabolite peak areas with the presence of DR. Odds ratios with 95% confidence intervals (CI) were calculated based on one standard deviation change in metabolite peak areas. Clinical variables representing known risk factors of DR (4; 5) such as HbA1c, were added as covariates to logistic regression models to calculate adjusted odds ratios. Adjusted odds ratios were also calculated for dominating trends that were identified from PCA, and for major classes of diabetes medications. PCA was performed using SIMCA (Umetrics AB), all other statistical tests were performed in the R environment.

**Metabolite Identification**

Experimentally-obtained spectra were searched against entries in the National Institutes of Science and Technology (NIST) library using ChromaTOF. Metabolite identities were confirmed if both the forward and reverse similarity scores were above 800 and the Kovats retention index was within 2% from database entries. Peaks that could not be identified in the NIST library were exported to text spectra and searched against the Golm metabolite database (36). Ambiguous identifications were resolved by injection of derivatised analytical standards. For metabolites forming multiple derivatisation peaks, the peaks with lower CV in QC samples were reported. Peaks with unresolvable identifications were reported as mixtures of compounds.

**Pathway Analysis**

Pathway analysis was performed with MetaboAnalyst (37). The list of identified metabolites with $p<0.05$ and a fold change greater than 1.2 or less than 0.8 were used as input. Canonical pathways were considered to be significantly enriched in the dataset if their FDR were less than 0.2.

**Metabolite Quantitation in Validation Set**
Standards for 1,5-gluconolactone, erythritol, gluconic acid, lactose, maltose, ribose and trehalose were purchased from Sigma-Aldrich. The compounds 2-deoxyribonic acid and 3,4-dihydroxybutyric acid were synthesized in-house and characterized by proton nuclear magnetic resonance. Synthesis methods are described in the supplementary. Calibration curves were constructed with standard solutions from 0.050 to 2.00 µg/mL. Plasma samples were similarly derivatised and analysed by GC-MS. Besides the column length (28.5 m) and detector voltage (1800 V), all other settings were the same as those used in for discovery metabonomic profiling. Kruskal-Wallis tests (non-parametric ANOVA) were performed to identify differences in plasma metabolite concentrations between groups.

Results

Sample Characteristics

The clinical characteristics of participants selected for discovery metabonomic profiling are shown in Table 1. Sample groups were comparable for most metabolic and clinical characteristics, such as BMI, total cholesterol, LDL cholesterol and HDL cholesterol. Participants with DR had higher HbA1c than the diabetic controls, mirroring typical epidemiological findings (4; 31). While hypertension identified as a risk factor for DR in the Singapore Indian Eye Study cohort (4; 5; 31), the increased systolic BP between DR cases and diabetic control groups in our study was not statistically significant.

Data Table and Chemometric Analysis

There were 263 peaks in the reference table constructed from medication-free samples and 258 were present in at least 80% of the samples. The CV of D-27 myristic acid, the internal standard was 7.9% and 193 peaks had a CV of less than 30% in the QC samples.

The PCA model constructed from aligned peak data was optimized at 5 PCs, with $R^2$ and $Q^2$ values of 0.48 and 0.27 respectively. The first and second PCs explain 15% and 6% of model variation respectively. Most samples from participants with chronic kidney disease are located towards the positive scores along the first PC (Figure S1A). When labelled by HbA1c values, the majority of samples with greater HbA1c levels are found on the positive axis of the second PC (Figure S1B). These clustering trends show that chronic kidney disease and
glycemic control are significant contributors to the overall plasma metabotype of diabetic persons, and could be confounders in metabonomic analysis.

**Metabolite Markers of DR**

There were fourteen candidate metabolites identified from discovery metabonomic profiling ($p < 0.05$, Mann-Whitney U test), and eleven met a $q$-value cut-off of 0.2 after correction for multiple testing (Table 2). DR samples had decreased levels of 1,5-anhydroglucitol and increased levels of 1,5-gluconolactone; 2-deoxyribonic acid; 3,4-dihydroxybutyric acid; erythritol; gluconic acid; lactose/cellobiose; maltose/trehalose; mannose; ribose and urea. A heatmap showing the relative peak area distribution of these metabolites in DR cases and diabetic controls are shown in (Figure 1). Analytical information of all identified metabolite markers are provided in Table S1.

The odds ratios of the metabolite markers in the basic logistic regression models and those adjusted for systolic BP, HbA1c levels and diabetes duration are shown in Table 3. While HbA1c was a significant covariate for all metabolite markers ($p < 0.05$ in likelihood-ratio test), adjusted odds ratios for 1,5-anhydroglucitol; 1,5-gluconolactone; 2-deoxyribonic acid; gluconic acid; lactose/cellobiose and urea remained significant (Table 3). Increased systolic BP and diabetes duration are known risk factors of DR (4; 5) but these factors were not significant covariates for any metabolite markers in our study (Table 3). The metabolite markers 1,5-gluconolactone; 2-deoxyribonic acid; gluconic acid, lactose/cellobiose and urea retained significant odds ratios when adjusted for systolic BP; HbA1c levels and diabetes duration together (Table 3). Adjusted odds ratios for age, gender, BMI, diastolic BP, serum creatinine, glucose, triglycerides, total cholesterol, LDL cholesterol, HDL cholesterol, urine albumin-to-creatinine ratio (ACR) and eGFR are shown in Table S2, Table S3 and Table S4. Age, gender, triglycerides, total cholesterol, LDL cholesterol and HDL cholesterol were not significant covariates for any metabolite marker (Table S2, Table S3, Table S4).

The number of participants with prescriptions for each class of diabetes medications is shown in Table S5. Biguanides, sulphonylureas, thiazolidinediones and alpha glucosidase inhibitors were not significant covariates for any metabolite marker, and insulin reduced the significance of only erythritol (Table S6).
Validation of Metabolite Markers

The reproducibility of metabolite markers identified from discovery metabonomic profiling was investigated in another set of samples selected from Singapore Indian Eye Study cohort. To minimize the contributions of potentially confounding kidney disease and glycemic control, participants with eGFR below 60 mL/min/1.73 m² were excluded, and diabetic controls were matched to DR cases for HbA1c. Samples from non-diabetic individuals were also analysed to establish reference ranges of these metabolites. The clinical characteristics of participants selected for the validation set are summarized in Table S7.

In the validation set, DR cases showed increased plasma concentrations of 2-deoxyribonic acid; 3,4-dihydroxybutyric acid, erythritol, gluconic acid and ribose when compared with diabetic controls (Table 4). In particular, diabetic controls also showed increased plasma concentrations of gluconic acid when compared with non-diabetic individuals (Figure 2). The previously unresolved metabolite pairs, lactose and cellobiose, maltose and trehalose were separated with the longer analytical column. Cellobiose was not detected in plasma samples. Trehalose was detected only in 26 DR cases, 22 diabetic controls and 4 non-diabetic individuals; and was excluded from further statistical analysis. Mannose was not quantified as it could not be resolved from glucose in this set of analyses.

Discussion

Metabolite Markers Complementary to Known Risk Factors

It is relevant to identify metabolite markers that are complementary to known risk factors such as glycemic control (4; 5) to improve existing risk stratification in DR-free diabetic patients and those with early stages of DR, (Figure 1). A panel of eleven metabolite markers was identified from discovery metabonomic profiling (Table 2). These metabolite markers are not only independent of HbA1c, but also remained statistically significant when adjusted for known metabolic risk factors for DR and potentially confounding kidney disease. We found that 1,5-gluconolactone, 2-deoxyribonic acid, gluconic acid and urea retain significant odds ratios when adjusted for key clinical variables such systolic blood pressure, HbA1c,
duration of diabetes, urine albumin-to-creatinine ratio, and glomerular filtration rate (Table S4).

Metabolite Markers of DR are Distinct from Diabetes

Increased levels of branched chain (18-20) and aromatic (16-19) amino acids in plasma and sera have been identified as markers and predictors of type 2 diabetes, while decreased levels of glycine were found to be correlated with diabetes (19; 21). Our results from discovery metabonomic profiling showed no significant differences in the levels of these amino acids between DR cases and diabetic controls. (Table S8). This contrast is likely due to the definition and choice of the control group. In the aforementioned studies, control samples were derived from healthy persons, while the controls in the discovery set were sampled from diabetic persons who have potentially had pre-existing increased catabolism of branched chain and aromatic amino acids. In addition, the majority of participants in our study had been prescribed with diabetes-controlling medications (Table 1). These medications might have helped lower the magnitude of perturbations in amino acid metabolism than untreated diabetes.

Non-Replicable Metabolite Markers of DR

In Li et al.’s investigation on plasma metabolite markers of DR, lowered levels of linoleic and arachidonic acids were found in DR cases (26). While those findings are in agreement with the view that arachidonic acid generally produces pro-inflammatory eicosanoids (38), no significant differences linoleic acid and arachidonic acid levels were found between DR cases and diabetic controls in the discovery set (Table S8). The levels of arachidonic acid are likely correlated to the use of anti-hypertensive medications than DR-status (39). This is supported by the similar numbers of participants on anti-hypertensive medications among the DR cases and diabetic control groups (27 and 24 respectively).

Omega-3 polyunsaturated fatty acids (PUFAs) are thought to aid in diabetes prevention but meta-analyses on omega-3 fatty acids intake show heterogeneous results, stratified by study locale (40). The omega-3 PUFAs, docosahexaenoic and eicosapentaenoic acids were detected on our platform, yet we did not find significant differences in these PUFAs between the DR group and diabetic controls in the discovery set (Table S8). Similarly,
Mäkinen et al. found that the baseline serum docosahexaenoic levels did not predict progressive kidney disease among patients with type 1 diabetes (41). Although omega-3 PUFAs have been shown to protect against DR in rodent models (42; 43), their efficacy in preventing vascular complications of diabetes can only be proven through randomized controlled trials.

**Unique Signatures of Hydroxy Fatty Acids in DR**

Earlier metabonomic studies on diabetes have demonstrated the association of several hydroxylated fatty acids with the development of diabetes. 2-Hydroxybutyric acid, a derivative of α-ketobutyrate, was previously identified as a marker (14; 21) and predictor (44) for diabetes. The ketone body, 3-hydroxybutyric acid was identified as a marker for impaired glucose tolerance (13) and type 2 diabetes (20; 21), and was also identified as a marker for DR in Li et al.’s study (26). No significant differences in the levels of 2-hydroxybutyric acid and 3-hydroxybutyric acid were found between DR cases and diabetic controls in both the discovery and validation sets (Table S8). DR cases and diabetic controls showed increased levels of 2-hydroxybutyric acid than non-diabetic individuals, consistent with a diagnosis of diabetes (Figure S2). These findings, together with aforementioned trends in amino acid levels suggest that plasma metabotype of DR is unique and not mere extension of the plasma metabotype of diabetes.

There is accumulating evidence to support the contribution of altered gut microbiota to obesity and diabetes (45) and gut microbiota-derived short-chain fatty acids (acetate, propionate, butyrate) from gut microbiota have been identified as signaling molecules in glucose and lipid metabolism (46). In our study, increased levels of 3,4-dihydroxybutyric acid and 2-deoxyribonic acid in DR cases were identified using discovery metabonomic profiling (Figure 1, Table 2). Similar increases in the concentrations of 3,4-dihydroxybutyric acid and 2-deoxyribonic acid were also found in the validation set (Table 4), though no significant differences in the levels of these hydroxy fatty acids were found between diabetic controls and non-diabetic individuals. These are no previous reports on the association of these hydroxy fatty acids with diabetes or DR. At present, it is known that 3,4-dihydroxybutyric acid is a urinary marker for succinate semialdehyde dehydrogenase deficiency (47) and the presence of 2-deoxyribonic acid within DNA is an indication of oxidative damage to the
sugar moiety (48). These hydroxy fatty acids are not part of major metabolic pathways and may be products of gut microbial metabolism on butyrate or other short-chain fatty acids. Metabolic flux analysis of these 2-deoxyribonic acid and 3,4-dihydroxybutyric acid could be performed to investigate their metabolic origins and their role in DR pathogenesis.

**Metabonomic Profiling Identifies DR-associated pathways**

Pathway mapping on MetaboAnalyst (37) showed that the pentose phosphate and galactose metabolism pathways are significantly enriched (FDR<0.2) among the metabolite markers identified from discovery metabonomic profiling. The pentose phosphate pathway was identified with increased levels of 1,5-gluconolactone, gluconic acid and ribose; while galactose metabolism was identified with lactose, mannose and myo-inositol (Table S9). Myo-inositol levels were increased among DR cases in the discovery set (1.25 fold) but did not meet significance after correction for multiple testing. These results suggest that the implications of selected metabolite classes (organic acids, polyols and sugars, oxidative stress markers) in DR pathogenesis.

Erythritol is recommended as a diabetic-safe sweetener as it is a metabolically inert antioxidant (49). Earlier metabonomic studies have identified erythritol as a marker of diabetes (20) and impaired fasting glucose (15). In our study, DR cases in both the discovery and validation sets showed increased levels of erythritol when compared with diabetic controls (Figure 1, Table 2, Table 4). In the absence of dietary survey data, it is unclear if this identified trend in erythritol levels was due to increased consumption of sweeteners or DR-related metabolic dysregulation. Mannose, a metabolite in galactose metabolism was previously identified as a marker for impaired fasting glucose (14; 15) and diabetes (14; 20). Increased plasma levels of mannose were identified in DR cases from discovery metabonomic profiling (Figure 1, Table 2). Therefore, future studies targeting the quantitation of polyols and sugars become pertinent since metabolites in the pentose phosphate and galactose pathways are potential clinical metabolite markers of DR. Dietary information should also be incorporated in such studies to detect distinct identification of metabolite markers associated with DR resolved from potential confounding by related nutritional factors.
Increased polyol pathway flux is one of the main driving hypotheses responsible for diabetic vascular complications (50). Increased aldose reductase activity under hyperglycemic conditions consumes NADPH, affecting regeneration of reduced glutathione, thereby causing or worsening underlying oxidative stress. While Barba et al. established indirect evidence of increased polyol pathway flux in the analysis of vitreous humour of DR patients (24), increases in plasma levels sorbitol and fructose, products of the polyol pathway, were not observed among participants with DR in this study. This is likely attributable to localisation of aldose reductase activity on glucose in susceptible cell types—in this case, retina capillary endothelium (50), and as a result, this perturbation remains undetected in plasma samples. The pentose phosphate pathway is the main source of cellular NADPH. Liu et al. had identified activation of the pentose phosphate pathway in skeletal muscle mitochondria of adiponectin knock-out mice on a high-fat diet, and these increases in pentose phosphate metabolites was associated with disturbances in glucose homeostasis and lipid metabolism (51). More recently Gaprov et al. also identified increased plasma concentrations of gluconic acid in non-obese diabetic mice (52). The identification of the pentose phosphate pathway from discovery metabonomic profiling (Figure 3), and the confirmation of increased gluconic acid (Figure 2) and ribose levels in DR cases (Table 4) in the validation set demonstrate that oxidative stress is implicated in DR pathogenesis.

Limitations

Our study focused on the Indian (South Asian) ethnic group as the age-adjusted prevalence of diabetes in Indians is the highest among the major ethnic groups of Singapore (53). As cross-sectional sampling only captures a snapshot of plasma metabotypes, some of the identified markers may represent short-term metabolic perturbations instead of chronic risk factors associated with development or DR. For example, discovery metabonomic profiling identified 1,5-anhydroglucitol, a marker for short-term glycemic control (lowered values indicate hyperglycemia) was significantly lowered in DR cases with higher HbA1c values than diabetic controls. Adjustment for covariates in the logistic regression models might not be sufficiently robust, given the limited sample size. Although several metabolite markers were validated in an independent set where DR cases and diabetic controls were matched for HbA1c, the generalizability of these findings to other populations is unclear.
GC-MS is generally less sensitive than LC-MS-based metabonomic profiling and requires sample derivatization and elevated temperatures for analysis. Therefore, the list of identified metabolite markers in our study could be limited to compounds that form stable silylated derivatives.

Conclusions

Using global metabonomic profiling, this study has identified plasma metabolite signatures that distinguish diabetic patients with retinopathy from those without retinopathy. These metabolite markers remained significant and robust while controlling for established risk factors. The metabolite markers 2-deoxyribonic acid; 3,4-dihydroxybutyric acid; erythritol; gluconic acid and ribose were validated in an independent, HbA1c-matched sample set. 2-Deoxyribonic acid and 3,4-dihydroxybutyric acid are novel metabolite markers with no prior reports of association with diabetes or DR. The pentose phosphate pathway was identified as a key metabolic dysregulation associated with DR, demonstrating the involvement of oxidative stress in disease pathogenesis. These findings provide the foundation for longitudinal metabonomic studies to establish the correlation and predictive value of metabolite profiles with the rate of DR progression in diabetic patients.
### Table 1. Clinical characteristics of samples selected for discovery metabonomic profiling

<table>
<thead>
<tr>
<th>Clinical Characteristics*</th>
<th>DR</th>
<th>Diabetic Controls</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>40</td>
<td>40</td>
<td>-</td>
</tr>
<tr>
<td>Gender (male)</td>
<td>20</td>
<td>23</td>
<td>0.501†</td>
</tr>
<tr>
<td>Age (years)</td>
<td>59 (53 - 66 )</td>
<td>62 (52 - 69 )</td>
<td>0.846‡</td>
</tr>
<tr>
<td>Diastolic BP (mm Hg)</td>
<td>74 (68 - 81 )</td>
<td>77 (68 - 83 )</td>
<td>0.516‡</td>
</tr>
<tr>
<td>Systolic BP (mm Hg)</td>
<td>146 (132 - 155 )</td>
<td>133 (122 - 146 )</td>
<td>0.086‡</td>
</tr>
<tr>
<td>BMI (kg/m$^2$)</td>
<td>24.5 (23.0 - 28.1 )</td>
<td>28.6 (23.4 - 30.3 )</td>
<td>0.070‡</td>
</tr>
<tr>
<td>Serum creatinine (µM)</td>
<td>84 (71 - 134 )</td>
<td>80 (58 - 117 )</td>
<td>0.092‡</td>
</tr>
<tr>
<td>Blood glucose, random (mM)</td>
<td>10.4 (6.8 - 14.6 )</td>
<td>8.3 (6.9 - 11.2 )</td>
<td>0.109‡</td>
</tr>
<tr>
<td>Total Cholesterol (mM)</td>
<td>4.50 (3.99 - 5.39 )</td>
<td>4.38 (3.66 - 5.25 )</td>
<td>0.320‡</td>
</tr>
<tr>
<td>LDL (mM)</td>
<td>2.84 (2.17 - 3.49 )</td>
<td>2.59 (2.08 - 3.32 )</td>
<td>0.332‡</td>
</tr>
<tr>
<td>HDL (mM)</td>
<td>0.98 (0.78 - 1.23 )</td>
<td>1.00 (0.84 - 1.23 )</td>
<td>0.486‡</td>
</tr>
<tr>
<td>Triglycerides (mM)</td>
<td>1.79 (1.39 - 2.59 )</td>
<td>1.35 (0.94 - 2.19 )</td>
<td>0.076‡</td>
</tr>
<tr>
<td>Blood HbA1c (%)</td>
<td>8.2 (7.4 - 10.2 )</td>
<td>7.4 (6.5 - 8.0 )</td>
<td>0.001‡</td>
</tr>
<tr>
<td>Duration of diabetes (years)</td>
<td>12 (7 - 21 )</td>
<td>11 (5 - 19 )</td>
<td>0.374‡</td>
</tr>
<tr>
<td>Urine ACR (µg/mg creatinine)</td>
<td>46.0 (16.8 - 161.6 )</td>
<td>25.7 (13.2 - 56.4 )</td>
<td>0.053‡</td>
</tr>
<tr>
<td>eGFR ( mL/min/1.73 m$^2$)</td>
<td>68.7 (44.6 - 95.9 )</td>
<td>87.5 (50.7 - 100.9 )</td>
<td>0.141‡</td>
</tr>
<tr>
<td>Microalbuminuria</td>
<td>21</td>
<td>14</td>
<td>0.115†</td>
</tr>
<tr>
<td>Macroalbuminuria</td>
<td>5</td>
<td>2</td>
<td>0.235†</td>
</tr>
<tr>
<td>Chronic Kidney Disease§</td>
<td>16</td>
<td>11</td>
<td>0.237†</td>
</tr>
<tr>
<td>On Diabetes Medication</td>
<td>34</td>
<td>32</td>
<td>0.556†</td>
</tr>
</tbody>
</table>

*Values indicate median with interquartile range, where applicable

† Pearson chi-square test
‡ Mann-Whitney U test
§ Defined as eGFR below 60 mL/min/1.73 m$^2$
Table 2. Metabolite markers identified from discovery metabonomic profiling

<table>
<thead>
<tr>
<th>Metabolite</th>
<th>Fold-change*</th>
<th>P-value</th>
<th>Q-value</th>
<th>AUC</th>
</tr>
</thead>
<tbody>
<tr>
<td>1,5-Anhydroglucitol†</td>
<td>0.50</td>
<td>&lt;0.001</td>
<td>0.03</td>
<td>0.74 (0.63-0.84)</td>
</tr>
<tr>
<td>1,5-Gluconolactone</td>
<td>1.45</td>
<td>0.001</td>
<td>0.05</td>
<td>0.71 (0.60-0.83)</td>
</tr>
<tr>
<td>2-Deoxyribonic acid</td>
<td>1.60</td>
<td>0.007</td>
<td>0.12</td>
<td>0.68 (0.56-0.79)</td>
</tr>
<tr>
<td>3,4-Dihydroxybutyric acid</td>
<td>1.26</td>
<td>0.019</td>
<td>0.20</td>
<td>0.65 (0.53-0.78)</td>
</tr>
<tr>
<td>Erythritol</td>
<td>1.25</td>
<td>0.019</td>
<td>0.20</td>
<td>0.63 (0.51-0.76)</td>
</tr>
<tr>
<td>Gluconic acid</td>
<td>1.57</td>
<td>0.001</td>
<td>0.03</td>
<td>0.72 (0.61-0.84)</td>
</tr>
<tr>
<td>Lactose/Cellobiose</td>
<td>1.75</td>
<td>0.010</td>
<td>0.13</td>
<td>0.67 (0.55-0.79)</td>
</tr>
<tr>
<td>Maltose/Trehalose</td>
<td>1.74</td>
<td>0.003</td>
<td>0.08</td>
<td>0.70 (0.58-0.81)</td>
</tr>
<tr>
<td>Mannose</td>
<td>1.21</td>
<td>0.009</td>
<td>0.13</td>
<td>0.67 (0.55-0.79)</td>
</tr>
<tr>
<td>Ribose</td>
<td>1.42</td>
<td>0.016</td>
<td>0.19</td>
<td>0.66 (0.54-0.78)</td>
</tr>
<tr>
<td>Urea</td>
<td>1.35</td>
<td>0.004</td>
<td>0.08</td>
<td>0.69 (0.57-0.80)</td>
</tr>
</tbody>
</table>

*Values above 1 indicate higher levels in DR group
† Classification in ROC analysis reversed for metabolites with fold-change below 1
### Table 3. Odds ratios of metabolite markers adjusted for systolic BP, HbA1c and diabetes duration

<table>
<thead>
<tr>
<th>Metabolite</th>
<th>Odds Ratios (95% CI) in Basic Model</th>
<th>Odds Ratios (95% CI) in Adjusted Models with Covariates</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>+ Systolic BP</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1,5-Anhydroglucitol</td>
<td>0.37 (0.19-0.65)*</td>
<td>0.36 (0.19-0.64)*</td>
</tr>
<tr>
<td>1,5-Gluconolactone</td>
<td>2.40 (1.38-4.65)*</td>
<td>2.58 (1.46-5.05)*</td>
</tr>
<tr>
<td>2-Deoxyribonic acid</td>
<td>2.08 (1.25-3.76)*</td>
<td>2.22 (1.32-4.11)*</td>
</tr>
<tr>
<td>3,4-Dihydroxybutyric acid</td>
<td>1.61 (1.01-2.70)†</td>
<td>1.67 (1.05-2.80)*</td>
</tr>
<tr>
<td>Erythritol</td>
<td>1.68 (1.04-2.85)*</td>
<td>1.75 (1.08-2.97)*</td>
</tr>
<tr>
<td>Gluconic acid</td>
<td>2.65 (1.50-5.33)*</td>
<td>2.76 (1.55-5.57)*</td>
</tr>
<tr>
<td>Lactose/Cellobiose</td>
<td>1.78 (1.10-3.01)*</td>
<td>1.93 (1.19-3.29)*</td>
</tr>
<tr>
<td>Maltose/Trehalose</td>
<td>2.14 (1.29-3.84)*</td>
<td>2.01 (1.21-3.57)*</td>
</tr>
<tr>
<td>Mannose</td>
<td>1.81 (1.09-3.27)*</td>
<td>1.65 (1.00-2.93)**‡</td>
</tr>
<tr>
<td>Ribose</td>
<td>1.65 (1.02-2.80)*</td>
<td>1.68 (1.04-2.85)**‡</td>
</tr>
<tr>
<td>Urea</td>
<td>1.95 (1.18-3.48)*</td>
<td>2.17 (1.30-3.95)**‡</td>
</tr>
</tbody>
</table>

*P(model)< 0.05
†0.05≤P(model)<0.1
‡P(likelihood-ratio)<0.05
§0.05≤P(likelihood-ratio)<0.1
Table 4. Performance of metabolite markers in the validation set

<table>
<thead>
<tr>
<th>Metabolite</th>
<th>P-value (Kruskal-Wallis test)</th>
<th>Fold-change</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>DR cases vs</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Diabetic Controls</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Diabetic Controls vs Non-Diabetic Individuals</td>
</tr>
<tr>
<td>1,5-Gluconolactone</td>
<td>0.893</td>
<td>1.06</td>
</tr>
<tr>
<td>2-Deoxyribose acid</td>
<td>0.014</td>
<td>1.27*</td>
</tr>
<tr>
<td>3,4-Dihydroxybutyric acid</td>
<td>0.029</td>
<td>1.39*</td>
</tr>
<tr>
<td>Erythritol</td>
<td>0.008</td>
<td>1.21*</td>
</tr>
<tr>
<td>Gluconic acid</td>
<td>&lt;0.001</td>
<td>1.28*</td>
</tr>
<tr>
<td>Lactose</td>
<td>0.543</td>
<td>1.05</td>
</tr>
<tr>
<td>Maltose</td>
<td>0.010</td>
<td>0.93</td>
</tr>
<tr>
<td>Ribose</td>
<td>0.001</td>
<td>1.38*</td>
</tr>
</tbody>
</table>

*P<0.05, Mann-Whitney U test
### Figure 1.

<table>
<thead>
<tr>
<th></th>
<th>DR</th>
<th>Diabetic Controls</th>
</tr>
</thead>
<tbody>
<tr>
<td>1,5-Arihydruoglucitol</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1,5-Gluconolactone</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2-Deoxyribo nucleic acid</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3,4-Dihydroxybutyric acid</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Erythritol</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gluconic acid</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lactose/Cellobiose</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maltose/Trehalose</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mannose</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ribose</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Urea</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Color Key**
-3 0 2
Fold-change compared to global mean, log 2

Value
Figure 2

Gluonic acid (µg/ml)

Non-Diabetic Individuals  Diabetic Controls  DR Cases
Figure 3.
Figure Legends

Figure 1. Heatmap showing relative peak areas of metabolites marker identified from discovery metabonomic profiling.

Figure 2. Plasma concentrations of gluconic acid from 40 non-diabetic individuals, 40 diabetic controls and 40 DR cases in the validation set.

Figure 3. Activation of the pentose phosphate pathway identified from elevated levels of gluconic acid and ribose (red arrows) among DR cases in both the discovery and validation sets. Reactions marked with an asterisk produce NADPH. Green arrows denote the accumulation of glycolytic metabolites leading activation of the polyol, hexosamine, protein kinase C and AGE pathways in the unifying mechanism responsible for diabetic complications, as proposed by M. Brownlee.
Article Information

Funding
This study was funded by the Singapore National Medical Research Council’s Centre Grant CG 2013 to the Singapore Eye Research Institute and the Singapore Ministry of Education’s Academic Research Fund (AcRF) Tier 1 Grant R-148-000-135-112 to Eric C.Y. Chan. Liyan Chen is supported by the Singapore National Eye Centre’s Health Research Endowment Fund Learning Award.

Author Contributions
L.Chen. performed the experiments, analyzed the data, interpreted the findings and wrote the manuscript. C.Y.Cheng, M.K.Ikram., C.Sabanayagam., G.Tan, E.S.Tai, C.M.G.Cheung and T.Y.Wong conducted the Singapore Indian Eye Study, conceptualized the manuscript’s research questions and edited the manuscript. P.Mitchell. and J.Wang. contributed to the methodology of the Singapore Indian Eye Study, discussion of the manuscript’s research questions and edited the manuscript. C.Y.Cheng, H.W.Choi. and R.W. Beuerman. provided guidance for statistical analyses and data presentation and edited the manuscript. D.Tian and L.Zhang performed the statistical analyses. G. Venkatesan synthesized and characterized 3,4-dihydroxybutyric acid and 2-deoxyribonic acid. L.Zhou. conceptualized the manuscript’s research questions, interpreted the findings and edited the manuscript. E.C.Y.Chan provided guidance for the experiments, interpreted the findings and edited the manuscript. L.Zhou, E.C.Y.Chan and T.Y.Wong are guarantors of this work.

Duality of Interest
No conflicts of interest relevant to this article are reported.
References

10. Zavrelova H, Hoekstra T, Alssema M, Welschen LMC, Nijpels G, Moll AC, de Vet HCW, Polak BCP, Dekker JM: Progression and Regression: Distinct Developmental Patterns of Diabetic Retinopathy in Patients With Type 2 Diabetes Treated in the Diabetes Care System West-Friesland, the Netherlands. Diabetes Care 2011;34:867-872
Fernandez C, O'Donnell CJ, Carr SA, Mootha VK, Florez JC, Souza A, Melander O, Clish CB, Gerszten 
Pischon T: Identification of Serum Metabolites Associated With Risk of Type 2 Diabetes Using a 
20. Suhre K, Meisinger C, Döring A, Altmaier E, Belcredi P, Gieger C, Chang D, Milburn MV, Gall WE, 
Weinberger KM, Mewes H-W, Hrabé de Angelis M, Wichmann HE, Kronenberg F, Adamski J, Illig T: 
PLoS ONE 2010;5:e13953
Reflective of Glucose Homeostasis in Non-Diabetic and Type 2 Diabetic Obese African-American 
Women. PloS ONE 2010;5:e15234
chromatography/time-of-flight mass spectrometry for metabonomics: Biomarker discovery for 
diabetes mellitus. Analytica Chimica Acta 2009;633:257-262
23. Kuo JZ, Wong TY, Rotter JI: CHallenges in elucidating the genetics of diabetic retinopathy. JAMA 
Ophthalmology 2014;132:96-107
Metabolic Fingerprints of Proliferative Diabetic Retinopathy: An 1H-NMR–Based Metabolomic 
Approach Using Vitreous Humor. Investigative Ophthalmology and Visual Science 2010;51:4416-
4421
25. Lehmann M, Yanes O, Krohne TU, Dorsey AL, Aguilar E, Marchetti V, Moreno SK, Trombley J, 
Siuzdak G, Friedlander M: Metabolomic Analysis of Serum from Diabetic Patients With and Without 
Retinopathy. In Association for Research in Vision and Ophthalmology (ARVO) Meeting 2011 Fort 
Lauderdale, FL, 2011, p. A386
chromatography-mass spectrometry: a comparison of stages and subtypes diagnosed by Western 
S-D: Homocysteine and other biochemical parameters in Type 2 diabetes mellitus with different 
diabetic duration or diabetic retinopathy. Clinica Chimica Acta 2006;366:293-298
dimethylarginine levels in type 2 diabetic patients with diabetic retinopathy. Diabetes Research and 
Clinical Practice 2009;84:219-223
metabolites and diabetic retinopathy in Chinese type 2 diabetic patients. Clinica Chimica Acta 
2011;412:940-945
Young TL, Cajucom-Uy H, Foster PJ, Aung T, Saw SM, Wong TY: Methodology of the Singapore Indian 
Chinese Cohort (SiCC) Eye Study: Quantifying ethnic variations in the epidemiology of eye diseases in 
S-M, Wong TY: Prevalence and Risk Factors of Diabetic Retinopathy in Migrant Indians in an 
Urbanized Society in Asia: The Singapore Indian Eye Study. Ophthalmology 2012;119:2119-2124
Impact of Migration and Acculturation on Prevalence of Type 2 Diabetes and Related Eye 