Research Article

For reprint orders, please contact: reprints@futuremedicine.com

Biomarkers in Medicine

Utilization of N-glycosylation profiles as risk stratification biomarkers for suboptimal health status and metabolic syndrome in a Ghanaian population

Eric Adua[†,1], Elham Memarian[†,2], Alyce Russell[1], Irena Trbojević-Akmačić[2], Ivan Gudelj[2], Julija Jurić[2], Peter Roberts[1], Gordan Lauc[‡,2,3] & Wei Wang*[,‡,1,4,5]

[1]School of Medical & Health Sciences, Edith Cowan University, WA 6027, Australia

[2]Genos Glycoscience Research Laboratory, Zagreb 10000, Croatia

[3]University of Zagreb, Faculty of Pharmacy & Biochemistry, Zagreb 10000, Croatia

[4]School of Public Health, Taishan Medical University, Shandong, Taian 271000, PR China

[5]Beijing Key Laboratory of Clinical Epidemiology, School of Public Health, Capital Medical University, Beijing 100069, PR China

*Author for correspondence: Tel.: +61 8 6304 3717; wei.wang@ecu.edu.au

† Authors contributed equally

‡ Authors contributed equally

Aim: The study sought to apply N-glycosylation profiles to understand the interplay between suboptimal health status (SHS) and metabolic syndrome (MetS). **Materials & methods:** In this study, 262 Ghanaians were recruited from May to July 2016. After completing a health survey, plasma samples were collected for clinical assessments while ultra performance liquid chromatography was used to measure plasma N-glycans. **Results:** Four glycan peaks were found to predict case status (MetS and SHS) using a step-wise Akaike's information criterion logistic regression model selection. This model yielded an area under the curve of MetS: 83.1% (95% CI: 78.0–88.1%) and SHS: 67.1% (60.6–73.7%). **Conclusion:** Our results show that SHS is a significant, albeit modest, risk factor for MetS and N-glycan complexity was associated with MetS.

First draft submitted: 3 January 2019; Accepted for publication: 22 July 2019; Published online: 26 September 2019

Keywords: biomarker • glycan peaks • metabolic syndrome • N-glycans • population genetics • prediction • risk factors • suboptimal health status • Type 2 diabetes mellitus • ultra performance liquid chromatography

Metabolic diseases, including metabolic syndrome (MetS) and Type II diabetes mellitus (T2DM), remain major health problems worldwide with their burden linked to increased healthcare costs and premature deaths [1–4]. To date, only symptomatic treatments are available for cardiometabolic diseases and this, in the context of preventative medicine, is a delayed intervention [5–7]. Early detection would be beneficial by providing patients with tailored treatments that can postpone the onset of these diseases [6,8].

On the one hand, a possible approach is to screen for people who may be experiencing poor health without being diagnosed with a clinical condition, otherwise known as suboptimal health status (SHS) [9–11]. The SHS is a physical state between health and disease, characterized by the perception of general weakness, health complaints, chronic fatigue and decreased vitality [9,10,12]. The SHS can be determined with a noninvasive and validated instrument: SHS questionnaire-25 (SHSQ-25) and, hitherto, has been used to establish the relationship between SHS and cardiometabolic diseases in three major ethnic groups: Caucasians [10], Chinese [9,11] and West Africans [13].

Early detection and diagnosis could also be possible by targeting associated metabolic alterations in glycosylation [14–17]. Glycosylation is a common co- and post-translational modification process where complex oligosaccharides (glycans) covalently bind to protein backbones [18–20]. When pinned to proteins, they affect their function, structure, stability, folding or trafficking, and thus they are crucial for intra/extracellular interaction or signaling, cell adhesion, translocation, cell–cell or molecule–cell interaction, inflammation and immune function [18,21–25]. While recognizing other glycan types, the focus of this study will be on N-glycans. N-glycans are those that attach

Future Medicine

N-acetylglucosamine (GlcNAc) to the nitrogen of asparagine (Asn) residues of proteins in the consensus sequence Asn-X-threonine (Thr) or Asn-X-serine (Ser), where X is any amino acid apart from proline [26,27]. N-glycan is the most ubiquitous glycan subtype and can modify up to 90% of plasma proteins, contributing to the plasma protein's complexity and heterogeneity [14,28,29]. In fact, over 100 different N-glycans may be bound to a single protein site forming different glycoforms of the same protein [25,30]. Thus, they are useful targets for biological investigations by many glycobiologists.

In recent years, N-glycan profiles have been investigated for biomarker discovery since the total human glycome reflects the biological state of the organism. That is, under normal physiological conditions, N-glycans are stable, but change when the physiological conditions become abnormal or when there are unfavorable environmental conditions [31]. Environmental factors that affect N-glycan composition include smoking, alcohol consumption, dietary and socioeconomic status [23,32], while several studies suggest N-glycans are altered in chronic disease such as with cancers [33–35], rheumatoid arthritis [36,37], Parkinson's disease [38,39], hypertension [40], dyslipidemia [41] and also as part of the ageing process [40,42].

Despite its implication in multiple chronic diseases, studies on aberrant N-glycosylation profiles and risk of MetS or T2DM are still scanty [28,43,44]. Recently, one study established that ten N-glycan traits were either positively or negatively associated with plasma glucose, blood pressure and BMI in Chinese Han and Croatian populations [28]. Another study by McLachlan and colleagues identified N-glycan traits that were either associated with increased or decreased risk for developing MetS in Orcadian populations [44]. While both of these studies have reported significant findings such as revealing the protective role of core-fucosylated N-glycans, they were mainly confined to a few MetS components, abandoning other potential risk factors and markers of liver function. Because MetS is derived from a complex pathway, examining a broad spectrum of cardiometabolic risk factors will provide a better understanding of the disease's pathophysiology.

In addition, different populations are exposed to different epigenetic and genetic factors, both of which may influence N-glycosylation patterns [23,45]. Thus far, nearly all N-glycome profiling studies related to metabolic health have been conducted among Asian and European populations, while none to date has focused on geographically, climatically and genetically isolated populations such as those in West Africa. Therefore, this study aims to profile plasma N-glycome in healthy individuals, as well as those with SHS and MetS in a Ghanaian population. In addition, the study will compare the findings with the already analyzed and reported findings from Chinese Han, Croatian and Orcadian populations.

Materials & methods

Study design

In this cross-sectional study we recruited 262 participants from May to July 2016. The method of recruitment was based on convenience sampling where individuals residing within three suburbs (Ash Town, Pankrono and Abrepo) in Kumasi, Ghana volunteered to participate.

Inclusion criteria

In order to screen for individuals with undiagnosed risk factors, we excluded all participants who had been previously diagnosed with diabetes and/or hypertension and those who were taking antihypertensive, lipid or glucose lowering medicines at the time of blood sampling. In addition, we excluded individuals who were suffering from other chronic diseases related to the genitourinary, digestive, respiratory and hematological systems. The age range was 18–80 years.

Data collection

Participants were interviewed by trained research assistants to obtain information on demographic characteristics, medical and family history, environmental exposures, medication use and lifestyle factors. We then used the SHSQ-25 to determine SHS [12]. The SHSQ-25 comprises 25 items, which are subgrouped into five domains. Fatigue domain comprise nine items, the immune system comprises three items, digestive system three items, cardiovascular system three items and mental health seven items. Participants answered and selected specific items based on their health status over the past 3 months and chosen items were rated as follows: never or almost never, occasionally, often, very often and always. We then recorded the raw scores from 1–5 to 0–4 [12].

Anthropometric examination

Participants were asked to remove their shoes and their height (cm) and weight (kg) were measured with a standard stadiometer (SECA, Hamburg, Germany). Their measured height and weight were used to calculate their body BMI using the formula: BMI = weight (kg)/height $(m)^2$. We then used a measuring tape to measure the hip and the waist circumference and from this, the waist-to-height ratio was calculated (WHtR = waist [cm]/height [cm]). After allowing participants to rest for 5 min, diastolic and systolic blood pressures were measured with a standard sphygmomanometer (Omron HEM711DLX, Milton Keynes, UK).

Clinical data

Fasting blood samples from the veins of each participant was collected by a qualified phlebotomist. The collected blood samples were dispensed into EDTA anticoagulated, gel separator and fluoride oxalate tubes. Samples in these tubes were span in a centrifuge at 3000 g at 4°C for 10 min (centrifuge Eppendorf 5702R; Eppendorf AG, Hamburg, Germany). An automated chemistry analyzer (Roche Diagnostics, COBAS INTEGRA 400 Plus, IN, USA) was used to measure fasting plasma glucose (FPG) from samples in fluoride oxalate tubes. Similarly, serum albumin, globulin, ALP, GGT and direct bilirubin (DBIL), triglycerides (TG), HDL-c, LDL-c and total cholesterol (TC) were measured on the same analyzer. Aliquots of processed plasma samples were stored at -80°C until N-glycan analysis.

Diagnostic criteria for SHS & metabolic syndrome

The criteria for establishing MetS was based on the National Cholesterol Education Program – Adult treatment Panel (ATP) III guidelines [46]. A person is considered to have MetS if they have any three of the following: SBP ≥130 mmHg and/or DBP ≥85 mmHg, FPG ≥100 mg/dl (5.55 mmol/l), waist measurement of ≥102 cm in men and ≥88 cm in women; TG ≥150 mg/dl (1.7 mmol/l) or <50 mg/dl (1.3 mmol/l) in women and HDL-c of <40 mg/dl (1.0 mmol/l) in men. We did not use conventional obesity index (BMI >30) in this instance since BMI is not part of the criteria established by National Cholesterol Education Program-ATP. Thus, we used abdominal/central obesity. In Table 1, we provide the demographic characteristics of the population including BMI for males and females. These values indicate the proportion of the participants who were obese. The total SHS score was obtained from the sum of the 25 items. A median score greater than 22 and less than 22 illustrates poor health and good health, respectively [12].

Ethical approval

The study was conducted in accordance to the principles of the Declaration of Helsinki. Ethical approval was obtained from the Human Research Ethics Committee, Edith Cowan University, Australia and the Committee on Human Research, Publication and Ethics, Kwame Nkrumah University of Science and Technology, Ghana. Participants signed informed consent prior to the study. No animals were included in the study.

N-glycan release & labeling

To limit experimental errors and avoid bias, samples were first randomized on multiple plates. Plasma samples (10 µl) were aliquoted in 96-well plates and denatured with 20 µl 2 % (w/v) sodium dodecyl sulphate (Invitrogen, MA, USA), incubated at 65°C for 10 min and cooled to room temperature for 30 min [15,17,37]. Subsequently, 10 µl of 4 % (v/v) Igepal CA-630 (Sigma-Aldrich, MO, USA) was added and mixed. N-glycans were removed from glycoproteins after adding 1.2 U of Peptide *N*-glycosidase F (PNGase F; Promega, CA, USA) in 10 µl 5× PBS and incubated for 18 h at 37°C. A 2-aminobenzamide (2-AB; Sigma-Aldrich) solution was then used to label the released N-glycans. Subsequently, a solution of 2-picoline borane (2-PB, 44.8 mg/ml; Sigma-Aldrich) in DMSO (Sigma-Aldrich), 2-AB (19.2 mg/ml) and glacial acetic acid (Merch, Germany; 70:30 v/v) was prepared. To the glycans in the plate, 25 µl of the labeling mixture was added, covered and allowed to incubate at 65°C for 2 h. Afterward, excess label and reducing agent in samples were cleaned by hydrophilic interaction liquid chromatography solid phase extraction (HILIC) on hydrophilic 0.2 µm AcroPrep GHP filter plate (Pall Corporation, NY, USA) [15,17,37]. Samples were cooled for 30 min after which 700 µl of cold (4°C) acetonitrile (ACN) were added to each sample. Excess solvent was then removed using a vacuum manifold (Pall Corporation). Before loading samples in the GHP filter plate wells, the wells were washed with 200 µl of 70 % (v/v) ethanol, 200 µl of ultra pure water and equilibrated using 200 µl of cold (4°C) 96 % (v/v) ACN, followed by a brief sample incubation and washing with 5× 200 µl of cold (4°C) 96 % ACN. N-glycans were eluted with 2× 90 µl of ultra pure water by centrifugation at

Table 1. Characteristics of study participants.

Variable	Males	Females	Statistic	p-value	q-value
Age, years (n = 262)	51.95 ± 11.99	50.97 ± 12.45	7742.5†	0.7027	0.7027
BMI:			43.149‡	0.0001	0.0003*
– Underweight	7 (7.4)	5 (3.0)			
– Normal weight	58 (61.1)	50 (29.9)			
– Overweight	28 (29.5)	58 (34.7)			
– Obese	2 (2.1)	54 (32.3)			
Education:					
– Tertiary	25 (26.3)	10 (6)	24.47‡	0.0001	0.0003*
– Senior high school	25 (26.3)	57 (34.3)			
– Junior high school	33 (34.7)	58 (34.9)			
– Lower primary	6 (6.3)	25 (15.1)			
– No formal education	6 (6.3)	16 (9.6)			
Occupation:					
– Employed	73 (76.8)	110 (66.3)	19.53‡	0.002	0.005*
– Retired	11 (11.6)	10 (6.0)			
– Keeping house	1 (1.1)	16 (9.6)			
– Unemployed	0 (0)	15 (9)			
– Informal	10 (10.5)	15 (9)			
T2DM	39 (41.1)	78 (47.3)			
Clinical data:					
– WHtR (n = 262)	0.51 ± 0.06	0.58 ± 0.08	3833†	0.0001	0.0003*
– BMI (kg/m2) (n = 262)	23.15 ± 3.51	27.30 ± 5.24	4179.5†	0.0001	0.0003*
– SBP (mmHg) (n = 262)	146.99 ± 26.96	141.58 ± 22.18	7248†	0.223	0.2624
– DBP(mmHg) (n = 262)	81.94 ± 15.71	85.67 ± 13.46	6670.5†	0.0281	0.0401*
– FPG (mmol/l) (n = 260)	5.73 ± 0.75	5.87 ± 0.99	7306.5†	0.3329	0.3504
– TC (mmol/l) (n = 242)	4.24 ± 1.02	4.80 ± 1.35	5067†	0.0008	0.0021*
– TG (mmol/l) (n = 242)	1.16 ± 0.77	1.39 ± 0.98	5324.5†	0.004	0.0073*
– HDL-C (mmol/l) (n = 242)	1.17 ± 0.29	1.27 ± 0.33	5335.5†	0.004	0.0073*
– LDL-C (mmol/l) (n = 242)	2.55 ± 0.91	2.95 ± 1.13	5409†	0.0066	0.0101*
– VLDL-C (mmol/l) (n = 242)	5.21 ± 1.29	5.46 ± 1.60	5300†	0.0034	0.0073*
– Albumin	45.85 ± 2.62	42.73 ± 7.32	4101.5†	0.0001	0.0003*
– ALP	187.69 ± 47.55	202.11 ± 65.96	6158†	0.1467	0.1834
– Globulins	34.17 ± 5.59	35.15 ± 6.56	6133.5†	0.1343	0.1791
– GGT	33.16 ± 22.49	26.11 ± 18.03	5433†	0.0049	0.0082*
– DBIL	7.42 ± 3.68	5.52 ± 2.80	4468†	0.0001	0.0003*

Data presented as mean ± standard deviation and n (%).
†Mann–Whitney U tests. Tests of significance were two tailed.
‡$\chi 2$ test of independence.
* q-values obtained following a correction of multiple testing.
DBIL: Direct bilirubin; DBP: Diastolic blood pressure; FDR: False discovery rate; FPG: Fasting plasma glucose; SBP: Systolic blood pressure; T2DM; Type 2 diabetes mellitus; TC:Total cholesterol; TG: Triglyceride.

164 g for 5 min [15,17,37] (centrifuge 5804, rotor A-2-DWP, Eppendorf) in each step and stored at -20°C awaiting further analyses.

Ultra performance liquid chromatography

Labeled N-glycans were separated using HILIC on an Acquity UPLC instrument (Waters, MA, USA) with fluorescent detector in the presence of 100 mM ammonium formate as solvent A and ACN as solvent B. The following conditions were applied: linear gradient of 30–47 % solvent A, flow rate = 0.56 ml/ml, separation temperature = 25°C, sample temperature = 10°C, excitation wavelength = 250 nm, emission wavelength = 428 nm.

External standards of 2-AB labeled glucose oligomers were used to calibrate the system in the lead up to data processing and integration. The integration was followed by separation of total plasma N-glycome into 39 peaks (Supplementary Figure 1). The abundance of individual glycan peaks (GPs) was determined by dividing the area of each glycan by the total integrated area and expressed as a percentage. Derived traits consisting of low branching (LB), high branching (HB), neutral or no sialylation (S0), monosialylated (S1), disialylated (S2), trisialylated (S3), tetrasialylated (S4), agalactosylated (G0), monogalactosylated (G1), digalactosylated (G2), trigalactosylated (G3), tetragalactosylated (G4), antennary fucosylated (FUC_A), core fucosylated (FUC_C), biantennary (BA), biantennary agalactosylated (A2), biantennary galactosylated (A2G), monosialylated biantennary (BAMS), disialylated biantennary (BADS), triantennary (TRIA) and tetraantennary (TA) traits (details are shown in Supplementary Table 1). A flow chart of the study is shown in Supplementary Figure 2.

Statistical analysis

Batch correction and normalization of the UPLC data were performed to control nonbiological variability and experimental errors. Normality distribution was determined with Kolmogov–Smirnoff test and via visual distribution checks. Categorical variables were expressed as frequencies (percentages) while continuous data were reported as mean ± standard deviation. However, since the N-glycan data were skewed, the interquartile range was used to describe the distribution. Between groups comparison for continuous variables were performed with Mann–Whitney U-tests while intergroup comparisons of categorical variables were performed with χ^2 tests. Correlations between biochemical parameters and N-glycans were determined using Spearman correlation rank correlation method. A test for the reliability of the SHSQ-25 was performed using Cronbach's α coefficient. Benjamini–Hochberg method was used to control the false discovery rate (FDR). All statistical analyses were performed using the R statistical package software version 3.4.3 (R Core team, 2017) and the Statistical Package for Social Sciences version 23. Here, q-value obtained after multiple testing was calculated, and a q-value of <0.05 was considered significant.

Results

This study recruited 262 individuals from a Ghanaian population comprising 36.88% males and 63.11% females with mean ages 51.95 ± 11.99 and 50.97 ± 12.45 years, respectively. Females were generally obese when BMI (27.30 ± 5.24, q = 0.0003) and WHtR (0.58 ± 0.08, q = 0.0003) were used as obesity indices. Serum lipids (TC, TG, HDL-c and LDL-c) were generally higher in females ($q < 0.05$). However, blood pressure and FPG did not differ significantly between males and females ($q > 0.05$). Most participants had a formal education and were employed (Table 1).

Based on a median SHS score of 22, participants were grouped into low SHS (SHS ≤22) and high SHS (SHS >22). In Table 2, high SHS individuals were older (U = 6483.5, q = 0.0128), had higher SBP, DBP, FPG, TG, VLDL-c ($q < 0.05$) and a primarily sedentary lifestyle. In addition, higher levels of education and being employed were associated with low SHS ($q < 0.05$).

There were differential levels of N-glycan, albeit moderate, between low and high SHS. GPs: GP 34 (A4G4S [3,3,6,]3), GP 36 (A4G4S [3,3,3,3]4) and GP38 (A4G4S [3,3,3,6]4) were higher in high SHS ($p < 0.05$) while GP31 (FA3G3S[3,3,6]3) was lower in high SHS ($p < 0.05$). However, after controlling for FDR, only GP 34 (A4G4S [3,3,6]3) was significantly increased in the high SHS (Supplementary Table 2). Also, S4 was higher among high SHS ($p < 0.05$) whereas G2 and A2G were lower in high SHS compared with the low SHS group ($p < 0.05$). There were no significant differences in the derived N-glycan traits after correction for multiple testing ($q > 0.05$; Supplementary Table 3). Additionally, participants with MetS had higher total SHS scores compared with those without it (Figure 1). Further, LB, A2G, BAMS and BA were higher in the normal group compared with MetS whereas HB, G3, FUC_A and TRIA were significantly higher in the MetS compared with normal (Figure 2).

LB, S1, G2, A2G, BAMS, BA were statistically significantly lower in people with raised BP compared with normal BP. S4 was lower in raised FPG compared with S4 in normal FPG; however, this association was not statistically significant after FDR correction. In Table 3, there were high levels of S3 and S4 N-glycans in obese individuals compared to controls. That is S3 (normal [11.08 ± 1.80] obese [12.17 ± 1.75]; q = 0.0060) and S4 (normal [2.37 ± 0.48] obese [2.64 ± 0.66]; q = 0.0412). Likewise, G3 (11.84 ± 2.37) structures were higher in obese individuals compared with controls (13.44 ± 1.92, p = 0.0032). In addition, higher BMI was statistically significantly associated with higher HB, S3, S4, G3, FUC_A and TRIA whereas LB, S1, A2G, BAMS and BA were statistically significantly lower in individuals with high BMI. G2 level was also lower in individuals with high BMI, but this association was not statistically significant after FDR correction (Table 3).

Table 2. Distribution of participants in low and high suboptimal health status.

Variable	Low SHS (n = 131)	High SHS (n = 126)	Statistic	p-value	q-value
Age (years)	49.24 ± 11.00	54.22 ± 12.40	6483.5	0.0039	0.0168*
Female	76 (58.0)	87 (69.0)	3.37†	0.44	0.543529
BMI (kg/m^2):					
– Underweight	4 (3.1)	8 (6.4)		0.876	0.877
– Normal weight	54 (41.2)	50 (40.0)			
– Overweight	46 (35.1)	38 (30.4)			
– Obese	27 (20.6)	29 (23.2)			
Education:			19.28†	0.001	0.0105*
– Tertiary	23 (17.6)	12 (9.7)			
– Senior high school	52 (39.7)	28 (22.6)			
– Junior high school	40 (30.5)	49 (39.5)			
– Lower primary	11 (8.4)	19 (15.3)			
– No formal education	5 (3.8)	16 (12.9)			
Occupation:			31.67†	0.0001	0.00021*
– Employed	110 (84.0)	68 (54.8)			
– Retired	5 (3.8)	16 (12.9)			
– Keeping house	8 (6.1)	8 (6.5)			
– Unemployed	5 (3.8)	11 (8.9)			
– Informal	3 (2.3)	21 (16.9)			
Clinical data:					
– BMI (kg/m^2)	0.55 ± 0.08	0.56 ± 0.08	8160.5‡	0.877	0.877
– WHtR	25.94 ± 5.02	25.81 ± 5.19	7243.5‡	0.09	0.189
– SBP (mmHg)	139.60 ± 23.01	148.16 ± 24.71	6442.5‡	0.0032	0.0168*
– DBP (mmHg)	82.13 ± 13.39	86.66 ± 15.23	6648.5‡	0.0093	0.0217*
– FPG (mmol/l)	5.70 ± 0.83	5.96 ± 0.98	6506‡	0.0079	0.021*
– TC (mmol/l)	4.48 ± 1.19	4.71 ± 1.34	6571‡	0.2423	0.36345
– TG (mmol/l)	1.20 ± 0.79	1.41 ± 1.01	5683‡	0.0048	0.0168*
– HDL-C (mmol/l)	1.21 ± 0.30	1.26 ± 0.33	6491‡	0.1849	0.323575
– LDL-C (mmol/l)	2.74 ± 1.03	2.86 ± 1.12	6831‡	0.4932	0.5754
– VLDL-C (mmol/l)	0.55 ± 0.36	0.62 ± 0.33	5675.5‡	0.0046	0.0168*
– Albumin	44.07 ± 5.92	43.99 ± 5.52	7156.5‡	0.682	0.753789
– Globulin	34.14 ± 5.11	35.17 ± 6.54	6658.5‡	0.224	0.361846
– ALP	193.16 ± 59.39	200.44 ± 61.3	6734.5‡	0.282	0.387188
– GGT	27.21 ± 18.95	30.34 ± 21.16	6750.5‡	0.295	0.387188
– DBIL	6.48 ± 3.37	5.95 ± 3.19	6513.5‡	0.138	0.263455
Physical activity:					
– Primary sedentary	27 (22.1)	54 (40.6)	11.93	0.008	0.021*
– Moderate activity	93 (76.3)	79 (59.4)			
– Very active	2 (1.6)	0 (0)			

Data presented as mean ± standard deviation and n (%).
†$\chi 2$ test of independence.
‡Mann-Whitney U tests. Tests of significance were two tailed, *q<0.05 significant after correction for FDR and the latter is bold.
DBIL: Direct bilirubin; FDR: False discovery rate; SHS: Suboptimal health status; TG: Triglyceride.

The level of correlations between BMI and derived plasma N-glycan traits in different populations (Ghanaian, Scottish, Chinese and Croatian) are displayed in Supplementary Table 4A. It was obvious that G3, TRIA and S3 were positively correlated and S1 negatively correlated with BMI in all four populations. However, correlations between BMI and other derived traits were restricted to particular populations. For example, G0 correlated with BMI among Ghanaians, Scottish and Croatians; BA correlated with BMI among Ghanaians, Scottish and Chinese; FUC_C correlated with BMI among Scottish, Chinese and Croatians, BAMS correlated with BMI among Ghanaians, Scottish and Croatians; G1, G2, BADS correlated with BMI among Scottish and Croatians, FUC_A correlated

Table 3. Distribution of derived plasma N-glycan traits in blood pressure, fasting plasma glucose and BMI.

Peak	Blood pressure				Fasting plasma glucose				BMI			
	Normal	Raised	p-value	q-value	Normal	raised	p-value	q-value	Normal	Obese	p-value	q-value
Branching												
LB	83.06 ± 2.68	82.24 ± 2.61	0.0154	0.0404*	82.69 ± 2.75	82.76 ± 2.62	0.9642	0.9669	83.04 ± 2.59	81.45 ± 2.62	0.0002	0.0049*
HB	16.54 ± 2.67	17.29 ± 2.59	0.0388	0.0815	16.87 ± 2.71	16.83 ± 2.62	0.9245	0.9669	16.50 ± 2.57	18.24 ± 2.52	0.0000	0.0046*
Level of sialylation												
S0	25.23 ± 4.52	24.91 ± 4.16	0.7317	0.8087	25.16 ± 4.27	25.04 ± 4.49	0.7699	0.9669	25.24 ± 4.53	24.66 ± 3.72	0.5582	0.8889
S1	21.04 ± 1.51	20.48 ± 1.42	0.0026	0.0137*	20.87 ± 1.61	20.73 ± 1.39	0.4526	0.7479	20.93 ± 1.54	20.28 ± 1.21	0.0038	0.0412*
S2	39.91 ± 3.39	39.96 ± 3.29	0.8398	0.8398	39.71 ± 3.27	40.16 ± 3.43	0.2407	0.7479	39.92 ± 3.47	39.94 ± 2.84	0.9753	0.9753
S3	11.08 ± 1.86	11.62 ± 1.79	0.0286	0.0667	11.32 ± 1.90	11.29 ± 1.81	0.8264	0.9669	11.08 ± 1.80	12.17 ± 1.75	0.0002	0.0060*
S4	2.33 ± 0.48	2.55 ± 0.58	0.0007	0.0095*	2.49 ± 0.55	2.36 ± 0.52	0.0344	0.7224	2.37 ± 0.48	2.64 ± 0.66	0.0052	0.0412*
Level of galactosylation												
G0	9.50 ± 2.71	9.96 ± 2.49	0.0794	0.1229	9.48 ± 2.61	9.91 ± 2.64	0.2322	0.7479	9.81 ± 2.72	9.34 ± 2.26	0.2404	0.6710
G1	10.72 ± 1.99	10.52 ± 1.88	0.6018	0.7021	10.73 ± 2.00	10.55 ± 1.91	0.5116	0.7674	10.65 ± 2.01	10.63 ± 1.69	0.9195	0.9753
G2	62.84 ± 3.76	61.76 ± 3.69	0.007	0.0291*	62.49 ± 3.91	62.30 ± 3.66	0.9669	0.9669	62.58 ± 3.97	61.49 ± 2.8	0.0490	0.2175
G3	12.04 ± 2.41	12.35 ± 2.33	0.3275	0.4298	12.03 ± 2.37	12.27 ± 2.38	0.4280	0.7479	11.84 ± 2.37	13.44 ± 1.92	0.0000	0.0032*
G4	4.50 ± 1.15	4.94 ± 1.53	0.0549	0.0961	4.83 ± 1.46	4.55 ± 1.22	0.2370	0.7479	4.66 ± 1.26	4.80 ± 1.60	0.6118	0.8889
A2	8.51 ± 2.67	8.98 ± 2.45	0.0819	0.1229	8.49 ± 2.57	8.92 ± 2.60	0.2440	0.7479	8.83 ± 2.68	8.34 ± 2.23	0.2301	0.6710
A2G	72.46 ± 3.11	71.14 ± 2.98	0.0009	0.0095*	72.12 ± 3.30	71.72 ± 2.97	0.4630	0.7479	72.13 ± 3.25	70.94 ± 2.37	0.0049	0.0412
Position of fucose												
FUC_A	2.09 ± 0.48	2.28 ± 0.58	0.0083	0.0291*	2.16 ± 0.58	2.18 ± 0.50	0.3050	0.7479	2.09 ± 0.50	2.46 ± 0.56	0.0001	0.0046*
FUC_C	38.82 ± 5.23	38.40 ± 4.97	0.5676	0.7012	38.82 ± 5.23	38.52 ± 5.04	0.8380	0.9669	38.78 ± 5.24	38.22 ± 4.64	0.7374	0.9753
Level of sialylation of biantennary glycans												
BAMS	21.04 ± 1.51	20.48 ± 1.42	0.0026	0.0137*	20.87 ± 1.61	20.73 ± 1.39	0.4530	0.7479	20.93 ± 1.54	20.28 ± 1.21	0.0038	0.0412*
BADS	35.68 ± 3.20	35.72 ± 3.16	0.8166	0.8398	35.56 ± 3.10	35.85 ± 3.25	0.3850	0.7479	35.78 ± 3.27	35.34 ± 2.76	0.3144	0.7206
Degree of branching												
BA	80.97 ± 2.70	80.12 ± 2.63	0.0132	0.0396*	80.61 ± 2.75	80.63 ± 2.66	0.9350	0.9669	80.96 ± 2.61	79.29 ± 2.6	0.0001	0.0046*
TRIA	12.90 ± 2.47	13.21 ± 2.43	0.3114	0.4298	12.86 ± 2.44	13.15 ± 2.46	0.3200	0.7479	12.68 ± 2.45	14.34 ± 1.96	0.0000	0.0032*
TA	4.74 ± 1.24	5.21 ± 1.64	0.0549	0.0961	5.09 ± 1.57	4.81 ± 1.32	0.2760	0.7479	4.91 ± 1.37	5.06 ± 1.70	0.5978	0.8889

Data presented as mean ± standard deviation tests of significance were two tailed.
*$q < 0.05$ significant after correction for false discovery rate.
A2: Biantennary agalactosylated; A2G: Biantennary galactosylated; BA: Biantennary; BADS: Disialylated biantennary; BAMS: Monosialylated biantennary; BMI: Body mass index; FUC_A: Antennary fucosylated; FUC_C: Core fucosylated; G0: Agalactosylated; G1: Monogalactosylated; G2: Digalactosylated; G3: Trigalactosylated; G4: Tetragalactosylated; HB: High branching; LB: Low branching; S0: Neutral or no sialylation; S1: Monosialylated; S2: Disialylated; S3: Trisialylated; S4: Tetrasialylated; TA: Tetraantennary; TRIA: Triantennary.

10.2217/bmm-2019-0005

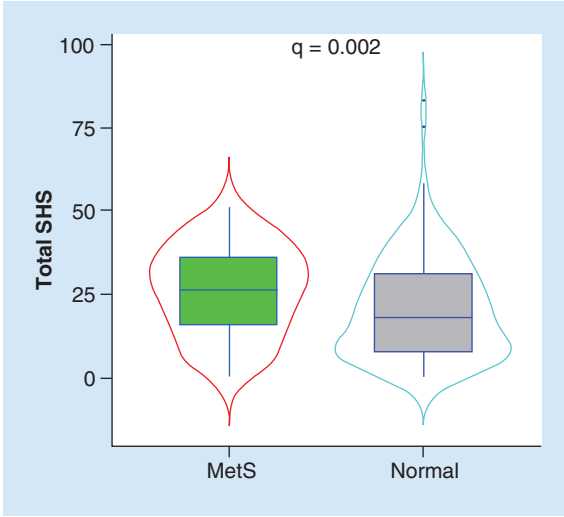

Figure 1. Distribution of total suboptimal health status score in metabolic syndrome and normal individuals.
Overall, increased in total SHS score was associated with MetS.
MetS: Metabolic syndrome; SHS: Suboptimal health status.

with BMI among Ghanaians and Scottish, S2 correlated with BMI among Scottish and Chinese; whereas A2 and S4 were only correlated with BMI in Ghanaians.

In Supplementary Table 4B, G0 and A2 were positively correlated with SBP in all populations, whereas FUC_C was negatively correlated with SBP in all four populations. G2 was negatively correlated with SBP among Scottish, Chinese and Croatians, whereas S3 was positively correlated with SBP among Scottish, Chinese and Croatians. BA, TRIA and S1 correlated with SBP among Scottish and Chinese, TA correlated with SBP among Ghanaians and Scottish, and BADS correlated with SBP among Scottish and Croatians. G3 and S2 were correlated with SBP only in Chinese, whereas G4 and BAMS correlated with SBP among Ghanaians and Scottish, respectively.

G0 and A2 positively correlated with DBP among Scottish and Croatians, G2 correlated with DBP among Ghanaians, Scottish and Croatians, TRIA and FUC_C correlated with DBP among Scottish and Chinese, BAMS correlated with Ghanaians and Scottish, S1 and S3 correlated with DBP among Ghanaians and Chinese, and S4 correlated with DBP among Ghanaians and Croatians. However, G3 and S2 correlated with DBP among Chinese, whereas FUC_A and BADS correlated with DBP Ghanaians and Croatian, respectively (Supplementary Table 4C).

Discussion

To our knowledge, this was the first attempt to apply complementary biomarkers for investigating SHS and MetS in a Ghanaian population. We hypothesized that SHS was associated with MetS and that there was N-glycosylation pattern similarity between SHS and MetS at the level of total plasma glycoproteins. In addition, increasing plasma N-glycan complexity was associated with MetS components and these differences vary from one population to the other. To a large extent, our results agreed with these hypotheses and details are provided below.

Association between SHS & metabolic syndrome

The SHS has been recognized as a precursor for chronic disease, but the extent of contribution of SHS to MetS has not been reported. Therefore, we employed the noninvasive SHSQ-25 to measure SHS within the population. Utilizing a median cutoff of 22, participants were categorized into low SHS and high SHS after which the clinical data of each group was assessed. We confirmed the findings of previous investigations by showing a positive association of SHS with age, education and occupation [9,11,13]. Additionally, high SHS was associated with cardiometabolic risk factors including higher SBP, DBP, FPG and TG [13,47,48]. Further, as displayed in Figure 1, having a higher total SHS score was associated with MetS as a whole. However, since the association between SHS and MetS cannot be unraveled with the subjective SHSQ-25 alone, we employed the cutting-edge technique of HILIC-UPLC to examine the glycosylation patterns (objective markers) for SHS and MetS.

N-glycosylation profile of total plasma glycoproteins as a biomarker for suboptimal health status & metabolic syndrome

As shown in Supplementary Tables 2 and 3, S4 ($p = 0.0371$; GP31 [FA3G3S(3,3,6)3]; $p = 0.0437$; GP34 A4G4S(3,3,6)3]; $p = 0.0110$; and GP38 [A4G4S(3,3,3,6)4]; $p = 0.0493$) were increased in high SHS whereas

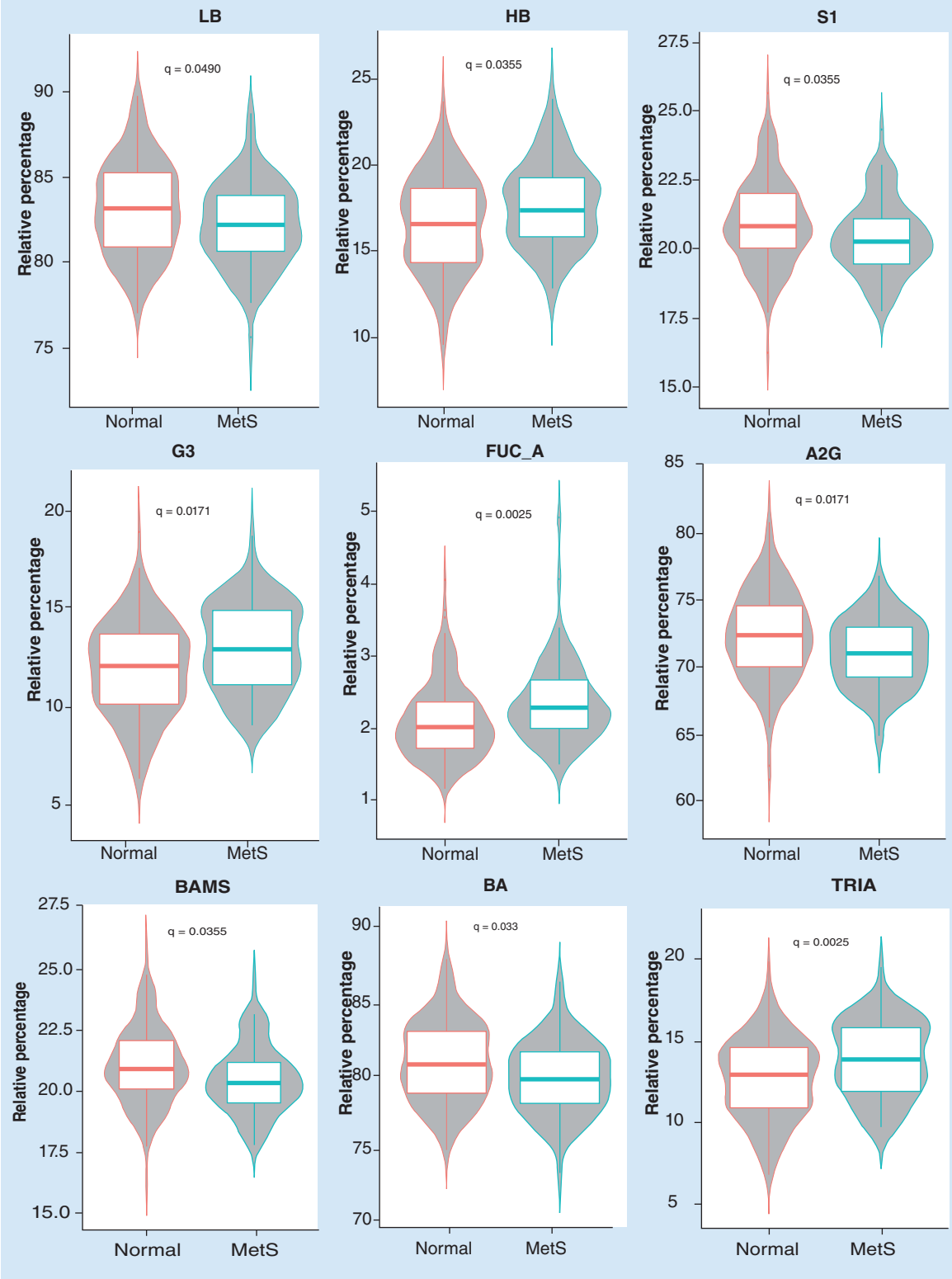

Figure 2. N-glycosylation differences between normal and metabolic syndrome. The levels of plasma N-glycan derived traits in normal and MetS are displayed in the form of violin/boxplots. In general, LB, A2G, BAMS and BA are higher in the normal group compared with MetS. In contrast, HB, G3, FUC_A and TRIA were significantly higher in the MetS compared with normal.
A2G: Biantennary galactosylated; BA: Biantennary; BAMS: Monosialylated biantennary; FUC_A: Antennary fucosylated; G3: Trigalactosylated; HB: High branching; LB: Low branching; MetS: Metabolic syndrome; S1: Monosialylated; TRIA: Triantennary.

10.2217/bmm-2019-0005

A2G (p = 0.0214) and G2 (p = 0.0297) were lowered in high SHS compared with low SHS. Although these associations were not significant after correction for multiple testing, there was an indication of clinical significance in the context of MetS progression. For example, even after adjusting for multiple testing, GP31, GP34 and GP38 were significantly increased in MetS with q = 0.0000, q = 0.0000 and q = 0.0274, respectively, whereas A2G was decreased in MetS (q = 0.0171). This indicates that there was a possible similarity in the plasma N-glycosylation patterns in SHS and MetS. It is also worth noting that other altered N-glycosylation levels existed in MetS that were not present in SHS. The most striking was the presence of increased HB and decreased LB (Figure 2). These findings have been previously reported by Keser *et al.* [37], when they were investigating N-glycosylation complexities that were associated with increased risk of developing T2DM. Higher branching indicates the abundance of *N*-acetylglucosamine residues which are due to the overexpression of *N*-acetylglucosaminyltransferase, the enzyme that catalyzes β 1, 6 GIcNAc branching [21]. In turn, this drives an inflammatory state that characterizes MetS [37]. Further, increases in antennarity and fucosylation of triantennary structures were evident in MetS. In particular, GP26 (A3G3S [3,3]2), GP30 (A3G3S [3,3,6]3), FUC_A and TRIA were higher in MetS compared with normal individuals and may signal inflammation. To emphasize these associations and the possible link with inflammation, Reiding *et al.* [49] showed that an increase in tri-and tetra-antennary structures corresponded with an increase in levels of plasma hsCRP and IL-6, both of which are widely recognized inflammation markers [50–52]. It could also mean that during MetS, there is increased plasma concentrations of a bi-and tri-antennary N-glycan carrier protein called α-1-antitrypsin [49,53], that may be in response to acute inflammation.

In regard to liver function markers, significant correlations were pronounced for globulin and ALP and, to a lesser extent, albumin. Surprisingly, the direction and magnitude of correlations between ALP and derived N-glycan traits were similar to that of BMI, WHtR and to a lesser extent LDL-c. At the same time, the correlations between globulin and derived traits were opposite to that of body fat markers. This indicates a possible role of high body fat in the development of liver disease (Supplementary Table 5). These results confirm those of Ko *et al.* [54] who showed a positive correlation between BMI, waist circumference, visceral fat and body fat with nonalcoholic fatty liver disease.

Interpopulation comparison of selected metabolic syndrome components & derived plasma N-glycan traits

It was evident that N-glycosylation was gender specific, as previously reported [15,22,23]. For example, females had higher S3, S4, FUC_A and TRIA and decreased levels of G2, A2G and BA compared with males. While it is apparent that these differences are due to hormonal differences, our results appear slightly different from those reported for other populations (Supplementary Figure 3). Lu *et al.* [28] showed that the levels of G1, FUC_C and S1 were significantly higher in Han Chinese females than in males, whereas G3 and S2 were higher in Han Chinese males (Lu *et al.* [28]). Among Croatian populations, S4, TRIA, TA, BAMS, G3 and G4 were higher in females whereas FUC_A, S2, BA and BADS were significantly higher in males. This study has showed that N-glycans correlate significantly with multiple clinical measures including BMI, SBP, DBP, FPG, LDL-c, TC, TG (Figure 3; Supplementary Table 4) as well as ageing (Supplementary Figure 3).

Many of these correlations confirm previous reports in the literature [22,41,42,55,56]. From Table 3 evident that increasing sialylation and galactosylation of N-glycan structures was associated with increasing BMI or obesity. These associations are consistent with earlier findings from other populations. For example, G3, S3 and TRIA were positively associated with BMI in all four populations (i.e., Ghanaians, Scottish, Chinese and Croatians) (Supplementary Table 4A) [22,28,55,56]. Among the derived traits, G0 and A2 were shown to be positively correlated and FUC_C negatively correlated with SBP in all four populations (Supplementary Table 4B). However, compared with normal individuals, those with raised BP had increased S4 and FUC_A. Although there were no common correlations between derived traits and DBP in all four populations, some findings were replicated in specific populations. For example, S4 was positively and G2 was negatively correlated with DBP in both Ghanaian and Croatian populations but not Chinese. In addition, FUC_C was negatively correlated with DBP in Scottish and Chinese but not in Ghanaians (Supplementary Table 4C). Moreover, it has been suggested that FUC_C may have a direct impact on special molecules such as EGF receptors, TGFβ receptor as well the phosphate inositol 3 kinase (PI3K) insulin signaling pathways. As such, abnormal expression of FUC_C may indicate metabolic abnormality [28,44,57]. However, the present study could not confirm this suggestion since there was no significant difference in the level of expression of FUC_C between normal and MetS.

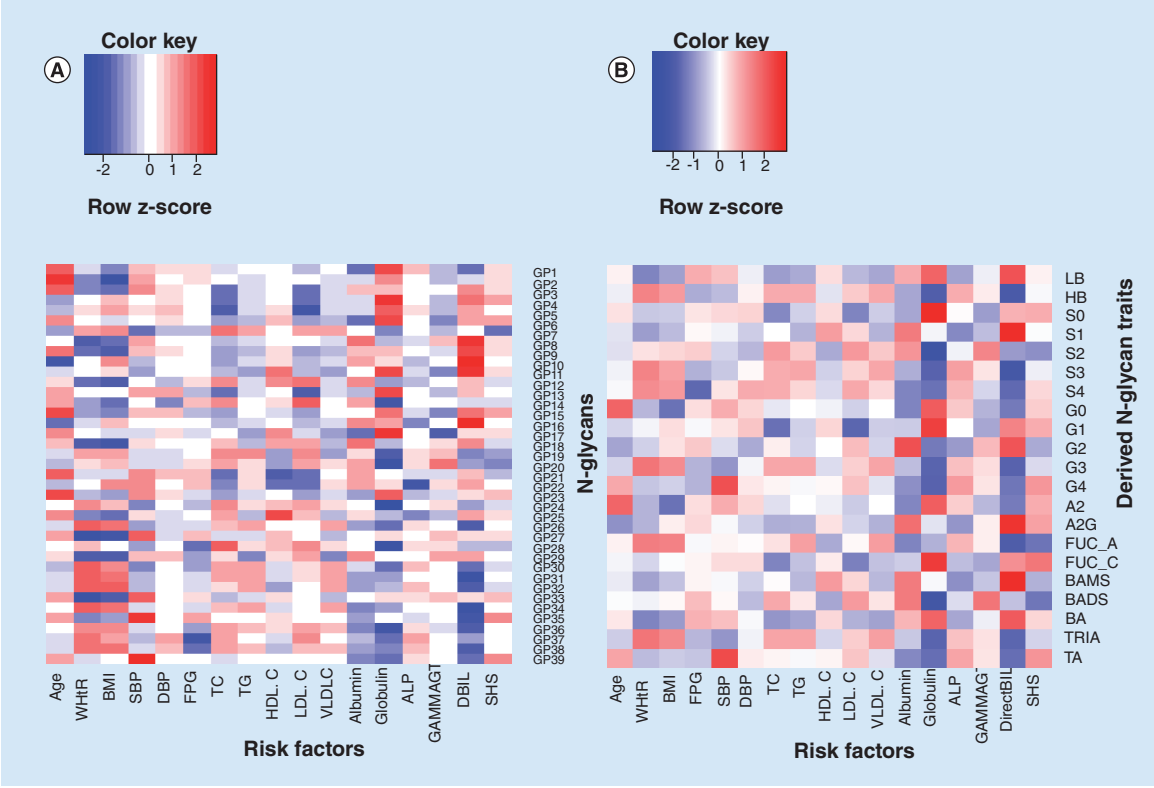

Figure 3. Heatmaps displaying the correlations between plasma N-glycans and risk factors for cardiometabolic diseases. The color key for the heat map and the histogram shows the overall distribution of Spearman correlation coefficients. **(A)** Correlations between plasma N-glycan peaks (GP1-GP39) and cardiometabolic risk factors. **(B)** Correlations between derived plasma N-glycan traits and cardiometabolic risk factors.
A2: Biantennary agalactosylated; A2G: Biantennary galactosylated; BA: Biantennary; BADS: Disialylated biantennary; BAMS: Monosialylated biantennary; BIL: Bilirubin; BMI: Body mass index; FUC_A: Antennary fucosylated; FUC_C: Core fucosylated; G0: Agalactosylated; G1: Monogalactosylated; G2: Digalactosylated; G3: Trigalactosylated; G4: Tetragalactosylated; HB: High branching; LB: Low branching; S0: Neutral or no sialylation; S1: Monosialylated; S2: Disialylated; S3: Trisialylated; S4: Tetrasialylated; SHS: Suboptimal health status; TA: Tetraantennary; TRIA: Triantennary; WHtR: Waist-to-height ratio.

While this study has identified common N-glycosylation patterns of total plasma glycoproteins in different populations, it should also be noted that there were some discrepancies in the extent of correlation or levels of N-glycans in these populations (Supplementary Tables 4A–C). These discrepancies may be driven by ethnic, environmental or genetic differences. For example, while utilizing two population genetics parameters: population differentiation and positive selection, Dall'Olio *et al.* [58] demonstrated that specific genes are involved in N-glycan biosynthesis and that there is interpopulation differences in the distribution of N-glycan structures. While certain genes are involved in the constitutive pathway, others are linked to a variable trait or the environment. Genes involved in the upstream N-glycosylation pathway (for proper protein folding) include *ST8SIA6*, *MAN2A2* and *MGAT3* in central south Asia; *MGAT4A* in European and East Asia; *ST8SIA3* in Europeans; *ST3GAL4* in sub-Saharan Africa (SSA) and *B4GALT2* in middle east–north Africa [58]. Therefore, given the significant roles of genetic and environmental factors in the integrity of N-glycans, and that the expression of these factors may not be the same across populations, the findings were somewhat expected.

Overall, the present study confirms the fact that N-glycosylation profiles are promising biomarkers for subclinical disease as four GPs could predict disease status (MetS and SHS) using a step-wise Bayesian information criterion and Akaike's information criterion logistic regression model selection (Table 4). From this, the area under curve were determined. MetS: 83.1% (95% CI: 78.0–88.1%) and SHS: 67.1% (60.6–73.7%; Figure 4A & B).

Before concluding, some limitations are worth mentioning. The sample size was small and, therefore, did not provide enough power to make our assumptions absolute and generalizable. Second, we did not control for

Table 4. Adjusted logistic regression model with metabolic syndrome as binary outcome.

N-glycans	Model fitting criteria						
	β	SE	AIC	BIC	-2Log likelihood	χ^2	q-value
GP8	0.365	0.184	265.918	283.741	255.918	4.035	0.045
GP18	0.452	0.176	268.662	286.485	258.662	6.779	0.009
GP21	-0.585	0.2	271.069	288.892	261.069	9.186	0.002
GP34	-0.724	0.212	275.16	292.982	265.16	13.276	0.000

Tests of statistical significance are two tailed and q < 0.05 is bold.
AIC: Akaike information criterion; BIC: Bayesian information criterion; SE: Standard error.

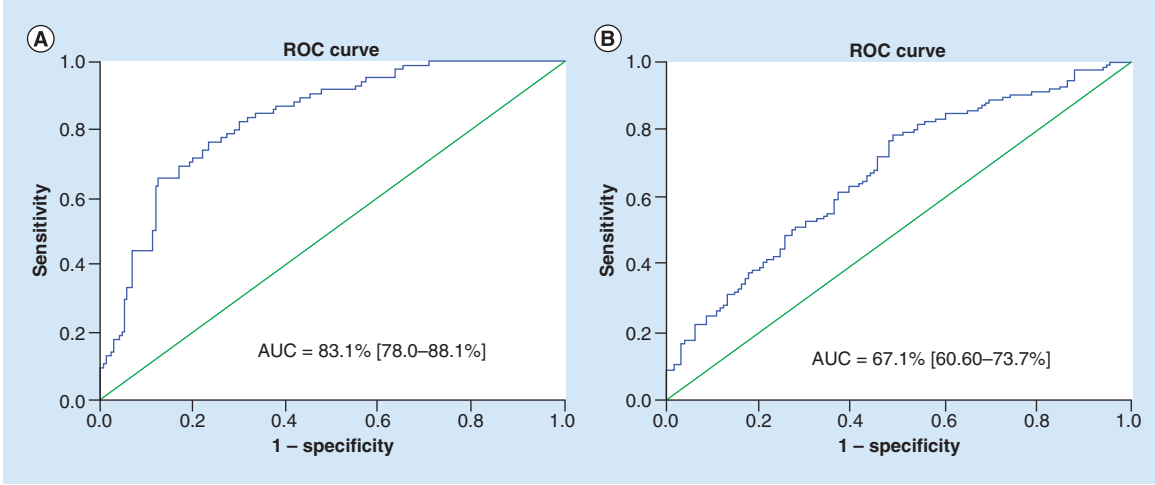

Figure 4. Receiver operating characteristics curves for metabolic syndrome and suboptimal health status. Receiver operative characteristic curve illustrating the performance of the age and sex adjusted logistic regression model in predicting the status of patients with **(A)** metabolic syndrome as binary outcome **(B)** suboptimal health status as binary outcome.
AUC: Area under curve; ROC: Receiver operating characteristic.

heritability relationships and hence our results may be under- or overestimated. The third relates to the cross-sectional design which does not explain whether aberrant N-glycan profiles were the cause or consequence of metabolic abnormality. Future longitudinal studies exploring hazard ratios from Cox regression models will give better insights into this association.

In summary, there is an association between SHS and MetS, and MetS is characterized by increased levels of more complex N-glycan structures on plasma glycoproteins or increase in plasma concentration of glycoproteins predominately carrying these glycans. Many of the findings in this study agree with earlier studies on other populations; the differences among populations are largely due to genetic and environmental factors.

Supplementary data

To view the supplementary data that accompany this paper please visit the journal website at: www.futuremedicine.com/doi/suppl/10.2217/bmm-2019-0005

Acknowledgments

We wish to thank Xia Li, University of La Trobe and Ebenezer Afrifa-Yamoah of Edith Cowan University, Australia for their advice on statistical analyses. Also, the authors appreciate the support from the staff at the Diabetic Clinic, KATH, Kumasi.

Financial & competing interests disclosure

This study was supported partially by the Joint Project of the Australian National Health & Medical Research Council and the National Natural Science Foundation of China (NHMRC APP1112767-NSFC 81561128020), National Natural Science Foundation of China (grant numbers 8177120753, 81673247 and 81573215), Edith Cowan University Collaboration Enhancement Scheme 2017 (round 1), the National Key Technology Support Program of China (2012BAI37B03), as well as by funding from the European

Structural and Investments funds for 'Croatian National Centre of Research Excellence in Personalized Healthcare' (contract no. KK.01.1.1.01.0010) and funding from the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska–Curie grant for project GlySign (contract no. 722095). The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

No writing assistance was utilized in the production of this manuscript.

Ethical conduct of research

All the patient investigations conformed to the principles outlined in the Declaration of Helsinki, and have been performed with the permission from the Human Research Ethics Committee, Edith Cowan University, Australia, and the Committee on Human Research, Publication and Ethics, Kwame Nkrumah University of Science and Technology, Kumasi. All the patients were informed about the purposes of the study and signed 'consent of the patient'. This article does not contain any studies with animals performed by any of the authors.

Summary points

- N-glycosylation profiles are intermediate phenotype between suboptimal health status (SHS) and metabolic syndrome (MetS).
- Total 262 Ghanaians were recruited from May to July 2016.
- Ultra performance liquid chromatography was used to analyze plasma N-glycans after which statistical methods were applied for data analyses.
- The SHS score was associated with age, education, occupation, physical inactivity, systolic blood pressure, diastolic blood pressure, fasting plasma glucose, triglycerides and MetS as a whole.
- MetS was associated with increased high branching, trigalactosylated, antennary fucosylated, triantennary and decreased low branching N-glycans.
- Four glycan peaks were found to predict case status (MetS and SHS) using a step-wise Akaike's information criterion and Bayesian information criterion logistic regression model selection.
- The model yielded an area under curve of MetS: 83.1% (95% CI: 78.0–88.1%) and SHS: 67.1% (60.6–73.7%).
- Our results show that SHS is a significant, albeit modest, risk factor for MetS and N-glycan complexity was associated with MetS.

References

Papers of special note have been highlighted as: • of interest; •• of considerable interest

1. Wilson PW, D'Agostino RB, Parise H, Sullivan L, Meigs JB. Metabolic syndrome as a precursor of cardiovascular disease and Type 2 diabetes mellitus. *Circulation* 112(20), 3066–3072 (2005).
2. Shah AD, Langenberg C, Rapsomaniki E *et al.* Type 2 diabetes and incidence of cardiovascular diseases: a cohort study in 1.9 million people. *Lancet Diabetes Endocrinol.* 3(2), 105–113 (2015).
3. Tao Z, Shi A, Zhao J. Epidemiological perspectives of diabetes. *Cell Biochem. Biophys.* 73(1), 181–185 (2015).
4. Zhong Y, Lin PJ, Cohen JT, Winn AN, Neumann PJ. Cost-utility analyses in diabetes: a systematic review and implications from real-world evidence. *Value Health* 18(2), 308–314 (2015).
5. Golubnitschaja O. Time for new guidelines in advanced diabetes care: paradigm change from delayed interventional approach to predictive, preventive & personalized medicine. *EPMA J.* 1(1), 3–12 (2010).

•• **Comprehensively describes approaches for chronic disease prediction and prevention.**

6. Golubnitschaja O, Kinkorova J, Costigliola V. Predictive, preventive and personalised medicine as the hardcore of 'Horizon 2020': EPMA position paper. *EPMA J.* 5(1), 1–29 (2014).
7. Lemke HU, Golubnitschaja O. Towards personal health care with model-guided medicine: long-term PPPM-related strategies and realisation opportunities within 'Horizon 2020'. *EPMA J.* 5(1), 1–9 (2014).
8. Golubnitschaja O, Baban B, Boniolo G *et al.* Medicine in the early twenty-first century: paradigm and anticipation-EPMA position paper 2016. *EPMA J.* 7(1), 1–13 (2016).
9. Yan YX, Dong J, Liu YQ *et al.* Association of suboptimal health status and cardiovascular risk factors in urban Chinese workers. *J. Urban Health* 89(2), 329–338 (2012).
10. Kupaev V, Borisov O, Marutina E, Yan YX, Wang W. Integration of suboptimal health status and endothelial dysfunction as a new aspect for risk evaluation of cardiovascular disease. *EPMA J.* 7(19), 1–7 (2016).
11. Wang Y, Ge S, Yan Y *et al.* China suboptimal health cohort study: rationale, design and baseline characteristics. *J. Transl. Med.* 14(1), 1–12 (2016).

- **Applies suboptimal health questionnaire (SHS)-25 as a novel tool for predicting the onset of cardiometabolic diseases.**

12. Yan YX, Liu YQ, Li M *et al.* Development and evaluation of a questionnaire for measuring suboptimal health status in urban Chinese. *J. Epidemiol.* 19(6), 333–341 (2009).
13. Adua E, Roberts P, Wang W. Incorporation of suboptimal health status as a potential risk assessment for Type 2 diabetes mellitus: a case-control study in a Ghanaian population. *EPMA J.* 8(4), 345–355 (2017).
14. Lauc G, Pezer M, Rudan I, Campbell H. Mechanisms of disease: the human N-glycome. *Biochim. Biophys. Acta* 1860(8), 1574–1582 (2016).
15. Adua E, Memarian E, Russell A *et al.* High throughput profiling of whole plasma N-glycans in Type 2 diabetes mellitus patients and healthy individuals: a perspective from a Ghanaian population. *Arch. Biochem. Biophys.* 661(1), 10–21 (2018).

- **This first N-glycan study involving a West African population and sheds light on how N-glycan profiles differ between people with Type 2 diabetes and those without it.**

16. Russell A, Adua E, Ugrina I, Laws S, Wang W. Unravelling immunoglobulin G Fc N-glycosylation: a dynamic marker potentiating predictive, preventive and personalised medicine. *Int. J. Mol. Sci.* 19(2), 1–13 (2018).
17. Russell A, Kepka A, Akmacic IT *et al.* Higher levels of abdominal body fat are associated with an increase in pro-inflammatory immunoglobulin G N-glycans: results from the Busselton Healthy Ageing Study. Paper presented at the 12th Jenner Glycobiology and Medicine Symposium Translational Glycobiology: from bench to bedside, May 6–9, 2017, Dubrovnik, Croatia. (2017).
18. Kukuruzinska M, Lennon K. Protein N-glycosylation: molecular genetics and functional significance. *Crit. Rev. Oral Biol. Med.* 9(4), 415–448 (1998).
19. Taylor ME, Drickamer K. *Introduction to Glycobiology.* Oxford University Press, NY, USA (2011).
20. Gu J, Isaji T, Xu Q *et al.* Potential roles of N-glycosylation in cell adhesion. *Glyconj. J.* 29(8–9), 599–607 (2012).
21. Dube DH, Bertozzi CR. Glycans in cancer and inflammation – potential for therapeutics and diagnostics. *Nat. Rev. Drug Discov.* 4(6), 477–488 (2005).
22. Knežević A, Gornik O, Polašek O *et al.* Effects of aging, body mass index, plasma lipid profiles, and smoking on human plasma N-glycans. *Glycobiology* 20(8), 959–969 (2010).
23. Knezevic A, Polasek O, Gornik O *et al.* Variability, heritability and environmental determinants of human plasma N-glycome. *J. Proteome Res.* 8(2), 694–701 (2009).

- **Indicates how environmental factors influence the human N-glycome and it is particularly relevant for the present study.**

24. Bieberich E. Synthesis, processing, and function of N-glycans in N-glycoprotein. In: *Glycobiology of the Nervous System. Advances in Neurobiology*. Yu R, Schengrund CL (Eds). Springer, NY, USA 9, 45–70 (2014).
25. Lauc G, Zoldoš V. Protein glycosylation – an evolutionary crossroad between genes and environment. *Mol. Biosyst.* 6(12), 2373–2379 (2010).
26. Varki A, Cummings RD, Esko JD *et al. Essentials of Glycobiology*. Cold Spring Habour, NY, USA (2009).
27. Stanley P, Cummings RD. Structures common to different glycans. In: *Essentials of Glycobiology 3rd edition 2009*. Cold Spring Harbor Laboratory Press, NY, USA 6, 760–784 (2009).
28. Lu JP, Knezevic A, Wang YX *et al.* Screening novel biomarkers for metabolic syndrome by profiling human plasma N-glycans in Chinese Han and Croatian populations. *J. Proteome Res.* 10(11), 4959–4969 (2011).

- **This is one of our previous studies that showed the association between N-glycans and components of metabolic syndrome. Using the Ghanaians as the test population, our results were compared with results of this published study.**

29. Zoldoš V, Horvat T, Lauc G. Glycomics meets genomics, epigenomics and other high throughput omics for system biology studies. *Curr. Opin. Chem. Biol.* 17(1), 34–40 (2013).
30. Adua E, Russell A, Roberts P, Wang Y, Song M, Wang W. Innovation analysis on postgenomic biomarkers: glycomics for chronic diseases. *OMICS* 21(4), 183–196 (2017).

- **This is a review paper that comprehensively describes N-glycosylation from biosynthesis to quantitation in the laboratory.**

31. Gornik O, Wagner J, Pučić M, Knežević A, Redžić I, Lauc G. Stability of N-glycan profiles in human plasma. *Glycobiology* 19(12), 1547–1553 (2009).
32. Vasseur JA, Goetz JA, Alley WR, Novotny MV. Smoking and lung cancer-induced changes in N-glycosylation of blood serum proteins. *Glycobiology* 22(12), 1684–1708 (2012).
33. Liu L, Yan B, Huang J *et al.* The identification and characterization of novel N-glycan-based biomarkers in gastric cancer. *PLoS ONE* 8(10), e77821 (2013).
34. Taniguchi N, Kizuka Y. Chapter two – glycans and cancer: role of N-glycans in cancer biomarker, progression and metastasis, and therapeutics. *Adv. Cancer Res.* 126, 11–51 (2015).
35. Theodoratou E, Thaci K, Agakov F *et al.* Glycosylation of plasma IgG in colorectal cancer prognosis. *Sci. Rep.* 6, 28098 (2016).

36. Sebastian A, Alzain MA, Asweto CO *et al.* Glycan biomarkers for rheumatoid arthritis and its remission status in Han Chinese patients. *OMICS* 20(6), 343–351 (2016).

37. Keser T, Gornik I, Vučković F *et al.* Increased plasma N-glycome complexity is associated with higher risk of Type 2 diabetes. *Diabetologia* 60(12), 2352–2360 (2017).

38. Akasaka-Manya K, Manya H, Sakurai Y *et al.* Protective effect of N-glycan bisecting GlcNAc residues on β-amyloid production in Alzheimer's disease. *Glycobiology* 20(1), 99–106 (2009).

39. Russell AC, Šimurina M, Garcia MT *et al.* The N-glycosylation of immunoglobulin G as a novel biomarker of Parkinson's disease. *Glycobiology* 27(5), 501–510 (2017).

40. Wang Y, Klaric L, Yu X *et al.* The association between glycosylation of immunoglobulin g and hypertension: a multiple ethnic cross-sectional study. *Medicine* 95(17), e3379 (2016).

41. Liu D, Chu X, Wang H *et al.* The changes of immunoglobulin GN-glycosylation in blood lipids and dyslipidaemia. *J. Transl. Med.* 16(1), 235 (2018).

42. Yamada E, Tsukamoto Y, Sasaki R, Yagyu K, Takahashi N. Structural changes of immunoglobulin G oligosaccharides with age in healthy human serum. *Glycoconj. J.* 14(3), 401–405 (1997).

43. Itoh N, Sakaue S, Nakagawa H *et al.* Analysis of N-glycan in serum glycoproteins from db/db mice and humans with Type 2 diabetes. *Am. J. Physiol. Endocrinol. Metab.* 293(4), E1069–E1077 (2007).

44. McLachlan F, Timofeeva M, Bermingham M *et al.* A case–control study in an orcadian population investigating the relationship between human plasma N-glycans and metabolic syndrome. *J. Glycomics Lipidomics* 6(139), 2153–0637 (2016).

45. Grarup N, Sandholt CH, Hansen T, Pedersen O. Genetic susceptibility to Type 2 diabetes and obesity: from genome-wide association studies to rare variants and beyond. *Diabetologia* 57(8), 1528–1541 (2014).

46. Lorenzo C, Williams K, Hunt KJ, Haffner SM. The National Cholesterol Education Program – Adult Treatment Panel III, International Diabetes Federation, and World Health Organization definitions of the metabolic syndrome as predictors of incident cardiovascular disease and diabetes. *Diabetes Care* 30(1), 8–13 (2007).

47. Wang W, Yan Y. Suboptimal health: a new health dimension for translational medicine. *Clin. Transl. Med.* 1(1), 1–6 (2012).

48. Yan YX, Dong J, Liu YQ *et al.* Association of suboptimal health status with psychosocial stress, plasma cortisol and mRNA expression of glucocorticoid receptor α/β in lymphocyte. *Stress* 18(1), 29–34 (2014).

49. Reiding KR, Ruhaak LR, Uh HW *et al.* Human plasma N-glycosylation as analyzed by matrix-assisted laser desorption/ionization-fourier transform ion cyclotron resonance-MS associates with markers of inflammation and metabolic health. *Mol. Cell. Proteomics* 16(2), 228–242 (2017).

50. Visser M, Bouter LM, McQuillan GM, Wener MH, Harris TB. Elevated C-reactive protein levels in overweight and obese adults. *JAMA* 282(22), 2131–2135 (1999).

51. Tamakoshi K, Yatsuya H, Kondo T *et al.* The metabolic syndrome is associated with elevated circulating C-reactive protein in healthy reference range, a systemic low-grade inflammatory state. *Int. J. Obes.* 27(4), 443 (2003).

52. Castelblanco E, Hernández M, Castelblanco A *et al.* Low-grade inflammatory marker profile may help to differentiate patients with LADA, classic adult-onset Type 1 diabetes, and Type 2 diabetes. *Diabetes Care* 41(4), 862–868 (2018).

53. Ruhaak LR, Koeleman CA, Uh HW *et al.* Targeted biomarker discovery by high throughput glycosylation profiling of human plasma alpha1-antitrypsin and immunoglobulin A. *PLoS ONE* 8(9), e73082 (2013).

54. Ko YH, Wong TC, Hsu YY, Kuo KL, Yang SH. The correlation between body fat, visceral fat, and nonalcoholic fatty liver disease. *Metab. Syndr. Relat. Disord.* 15(6), 304–311 (2017).

55. Perkovic MN, Bakovic MP, Kristic J *et al.* The association between galactosylation of immunoglobulin G and body mass index. *Prog. Neuropsychopharmacol Biol. Psych.* 48, 20–25 (2014).

56. Lemmers RF, Vilaj M, Urda D *et al.* IgG glycan patterns are associated with Type 2 diabetes in independent European populations. *Biochim. Biophys. Acta Gen. Subj.* 1861(9), 2240–2249 (2017).

•• **Sheds light on aberrant N-glycosylation and Type 2 diabetes in European populations.**

57. Matrougui K. Diabetes and microvascular pathophysiology: role of epidermal growth factor receptor tyrosine kinase. *Diabetes Metab. Res. Rev.* 26(1), 13–16 (2010).

58. Dall'Olio GM, Laayouni H, Luisi P, Sikora M, Montanucci L, Bertranpetit J. Distribution of events of positive selection and population differentiation in a metabolic pathway: the case of asparagine N-glycosylation. *BMC Evol. Biol.* 12(1), 1–13 (2012).

•• **This is the first study to explain how positive selection and population differentiation are distributed in the metabolic pathway.**