

RESEARCH ARTICLE

# Healthy human serum *N*-glycan profiling reveals the influence of ethnic variation on the identified cancer-relevant glycan biomarkers

Abrha G. Gebrehiwot<sup>1</sup>, Daniel Seifu Melka<sup>2</sup>, Yimenashu Mamo Kassaye<sup>2</sup>, Ibrahim F. Rehan<sup>1,3</sup>, Shobith Rangappa<sup>1</sup>, Hiroshi Hinou<sup>1</sup>, Toshiya Kamiyama<sup>4</sup>, Shin-ichiro Nishimura<sup>1\*</sup>

**1** Division of Drug Discovery Research, Faculty of Advanced Life Science and Graduate School of Life Science, Hokkaido University, Kita-ku, Sapporo, Japan, **2** Department of Biochemistry, School of Medicine, Addis Ababa University, Addis Ababa, Ethiopia, **3** Department of Animal Behaviour and Husbandry, Faculty of Veterinary Medicine, South Valley University, Qena, Egypt, **4** Department of Gastroenterological Surgery I, Graduate School of Medicine, Hokkaido University, N15, W7, Kita-ku, Sapporo, Japan

\* [shin@sci.hokudai.ac.jp](mailto:shin@sci.hokudai.ac.jp)



**OPEN ACCESS**

**Citation:** Gebrehiwot AG, Melka DS, Kassaye YM, Rehan IF, Rangappa S, Hinou H, et al. (2018) Healthy human serum *N*-glycan profiling reveals the influence of ethnic variation on the identified cancer-relevant glycan biomarkers. PLoS ONE 13 (12): e0209515. <https://doi.org/10.1371/journal.pone.0209515>

**Editor:** Wei Wang, Edith Cowan University, AUSTRALIA

**Received:** September 25, 2018

**Accepted:** December 6, 2018

**Published:** December 28, 2018

**Copyright:** © 2018 Gebrehiwot et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All relevant data are within the manuscript.

**Funding:** This work was partly supported by the Ministry of Education, Culture, Sports, Science, and Technology Japan (MEXT) and JSPS KAKENHI Grant Number 25220206, JSPS Core-to-Core Program B. Asia-Africa Science Platforms, and by Japan Science and Technology Agency (JST) through a grant for "Development of Systems and Technology for Advanced Measurement and

## Abstract

### Background

Most glycomics studies have focused on understanding disease mechanisms and proposing serum markers for various diseases, yet the influence of ethnic variation on the identified glyco-biomarker remains poorly addressed. This study aimed to investigate the inter-ethnic serum *N*-glycan variation among US origin control, Japanese, Indian, and Ethiopian healthy volunteers.

### Methods

Human serum from 54 healthy subjects of various ethnicity and 11 Japanese hepatocellular carcinoma (HCC) patients were included in the study. We employed a comprehensive glyco-blotting-assisted MALDI-TOF/MS-based quantitative analysis of serum *N*-glycome and fluorescence HPLC-based quantification of sialic acid species. Data representing serum *N*-glycan or sialic acid levels were compared among the ethnic groups using SPSS software.

### Results

Total of 51 *N*-glycans released from whole serum glycoproteins could be reproducibly quantified within which 33 glycoforms were detected in all ethnicities. The remaining *N*-glycans were detected weakly but exclusively either in the Ethiopians (13 glycans) or in all the other ethnic groups (5 glycans). Highest abundance ( $p < 0.001$ ) of high mannose, core-fucosylated, hyperbranched/hypersialylated *N*-glycans was demonstrated in Ethiopians. In contrast, only one glycan ( $m/z$  2118) significantly differed among all ethnicities being highest in Indians and lowest in Ethiopians. Glycan abundance trend in Ethiopians was generally close to that of Japanese HCC patients. Glycotyping analysis further revealed ethnic-based disparities mainly in the branched and sialylated structures. Surprisingly, some of the

Analysis (SENTAN)" and "The Matching Program for innovations in Future Drug Discovery and Medical Care". The funders had no role in study design, data collection and data analysis, decision to publish, or preparation of the manuscript.

**Competing interests:** The authors have declared that no competing interests exist.

glycoforms greatly elevated in the Ethiopian subjects have been identified as serum biomarkers of various cancers. Sialic acid level was significantly increased primarily in Ethiopians, compared to the other ethnicities.

## Conclusion

The study revealed ethnic-specific differences in healthy human serum *N*-glycome with highest abundance of most glycoforms in the Ethiopian ethnicity. The results strongly emphasized the need to consider ethnicity matching for accurate glyco-biomarker identification. Further large-scale study employing various ethnic compositions is needed to verify the current result.

## Introduction

Glycosylation, the process in which sugars are attached to proteins or lipids, is the most abundant and complex post-translational process, causing immense structural and functional variabilities in majority of eukaryotic cell proteins [1, 2]. Glycan parts of glycoconjugates are known to facilitate essential roles in almost all physio-pathological processes including fertilization, cell differentiation, cell adhesion, cell recognition, molecular trafficking, signal transduction, protein folding, immunological regulation, aging, and even malignant alterations [1–3]. The indispensable role of glycosylation in complex organisms like humans is apparent from the fact that many eukaryotic cells can function and survive without nucleus; however, none of them can function normally without glycans at least on their surface [4–6], and hence total absence of glycans is embryologically lethal [7]. Many of the human serum proteins including alpha-1-acid glycoprotein, alpha-1-antitrypsin, alpha-2-macroglobulin, antithrombin-III, apolipoproteins, ceruloplasmin, fibrinogen, immunoglobulins, haptoglobin, hemopectin, and serotransferrin, are heavily glycosylated, making them targets for glyco-biomarker discovery and therapeutic opportunities [8].

In contrast to nucleic acids and proteins, biosynthesis of glycans is not template-driven but, rather, is a result of a complex network of metabolic and enzymatic reactions. Because of this and subsequent methodological difficulties, the field of glycomics has been lagging behind genomics and proteomics [9, 10]. Tremendous advancements in analytical techniques and bioinformatics platforms have recently revolutionized the area, enabling comprehensive profiling of glycans and glycoproteins to be released from various biological samples [11–13] and suggested as biomarkers. However, bringing these glycan or glycoprotein markers to clinical practice has been hindered as their potential to distinguish cases from controls or disease stages seems to be inadequate and varies from country to country population wise, complicating their validity and clinical utility [14, 15]. For example, the FDA approved cancer biomarkers based on serum level of *O*-glycosylated mucin glycoproteins of carbohydrate antigens (CA125 for ovarian, CA27.29 and CA15-3 for breast, CA19-9 for pancreatic cancers) and *N*-glycosylated glycoproteins ( $\alpha$ -fetoprotein for hepatocellular carcinoma, prostate-specific antigen for prostate cancer) lack the specificity and sensitivity to be used for early detection of cancer [16, 17].

Previous population-based studies have reported the association of plasma *N*-glycan structure alterations (mostly increased plasma *N*-glycan complexity) with metabolic syndrome and higher risk of type 2 diabetes [18, 19]. In several glycomics studies profiling immunoglobulin G (IgG) focused *N*-glycome among various ethnic populations, it was emphasized that changes

in IgG-linked glycan composition and abundance were correlated with hypertension [20–22], cardiovascular disease [23], blood lipids and dyslipidaemia [24]. Apart from several pathological studies that have profiled glycosylation pattern, comparative glycosylation studies among healthy subjects with various characteristics are often overlooked. Nevertheless, such studies on healthy people would be of clinical benefit not only by providing anticipatory insights but also by elucidating confounding factors that could lead to controversies on the identified biomarker [25]. In this context, a comprehensive study on human plasma *N*-glycan profile provided an evidence for the variability of some glycan levels with aging, life style and environmental factors [26]. Similarly, Ding N. and his group demonstrated that a healthy human serum *N*-glycan profile had shown considerable variations in age and sex dependent manner [27]. Furthermore, IgG focused *N*-glycomic study in a Han Chinese population revealed that changes in IgG *N*-glycan features significantly correlate with age [28]. To our knowledge, little is known about the association between ethnicity and healthy human serum *N*-glycome profile. Thus, this study was aimed to address the inter-ethnic differences in serum *N*-glycome among US origin control, South Indian, Japanese, and Ethiopian ethnic populations using a rapid glycoblotting-assisted MALDI-TOF/MS-based quantitative analysis. The present study clearly demonstrated that various high mannose, core-fucosylated, multiply branched and sialylated glycoforms illustrated an ethnic-specific expression pattern and marked alterations in their serum abundance.

## Materials and methods

### 2.1 Human serum samples

The study was performed in accordance with the ethical guidelines and protocols of the Declaration of Helsinki upon approval by the ethical review boards of Hokkaido University, Faculty of Advanced Life Sciences, Japan and Addis Ababa University, School of Medicine, Ethiopia. Informed consent was obtained from all volunteer participants. As summarized in [S1 Table](#), total of 54 healthy subjects having a relatively narrow age range and various ethnic composition were involved in the study. Serum from male Japanese ( $n = 10$ , age = 20s-30s y/o), and male South Indians ( $n = 10$ , age =  $32.5 \pm 5$  y/o) were collected as part of Asia-Africa Science Platforms project. Similarly, serum samples from 24 female Ethiopians ( $31.54 \pm 7$  y/o) were collected in Black Lion Specialized Teaching Hospital of Addis Ababa University and carefully transported to Japan after freeze dried. US origin healthy control samples were from serum pool of different male donors whose age information was not provided as these samples were purchased from Sigma-Aldrich company (product # H4522). Serum samples from male Japanese hepatocellular carcinoma (HCC) patients ( $n = 11$ , age = 50s-60s y/o) were also included in the study. As inclusion criteria for the healthy ethnic groups, fully healthy (mentally and physically) adults ( $> 18$  y/o), non-smokers, and non-obese individuals were included in the present study. Whereas, individuals who were pregnant, were receiving any treatment, or were with a medical history of disease such as cancer, diabetes, neurodegenerative, liver, or cardiovascular diseases which can affect the glycan profile, were excluded from the study. All serum samples were kept at  $-80^{\circ}\text{C}$  until used in the subsequent experiment.

### 2.2 Reagents and Equipment

Ammonium bicarbonate 99% (ABC), 1-propanesulfonic acid, 2-hydroxyl-3-myristamido (PHM), 3-methyl-1-*p*-tolyltriazen (MTT), *O*-benzylhydroxylamine hydrochloride (BOA), and disialyloctasaccharide were purchased from Tokyo Chemical Industry Co., Ltd. (Tokyo, Japan). Peptide *N*-glycosidase F (PNGase F) was purchased from New England Biolabs Inc. (Ipswich, Massachusetts, USA), whereas BlotGlyco H beads were purchased from Sumitomo

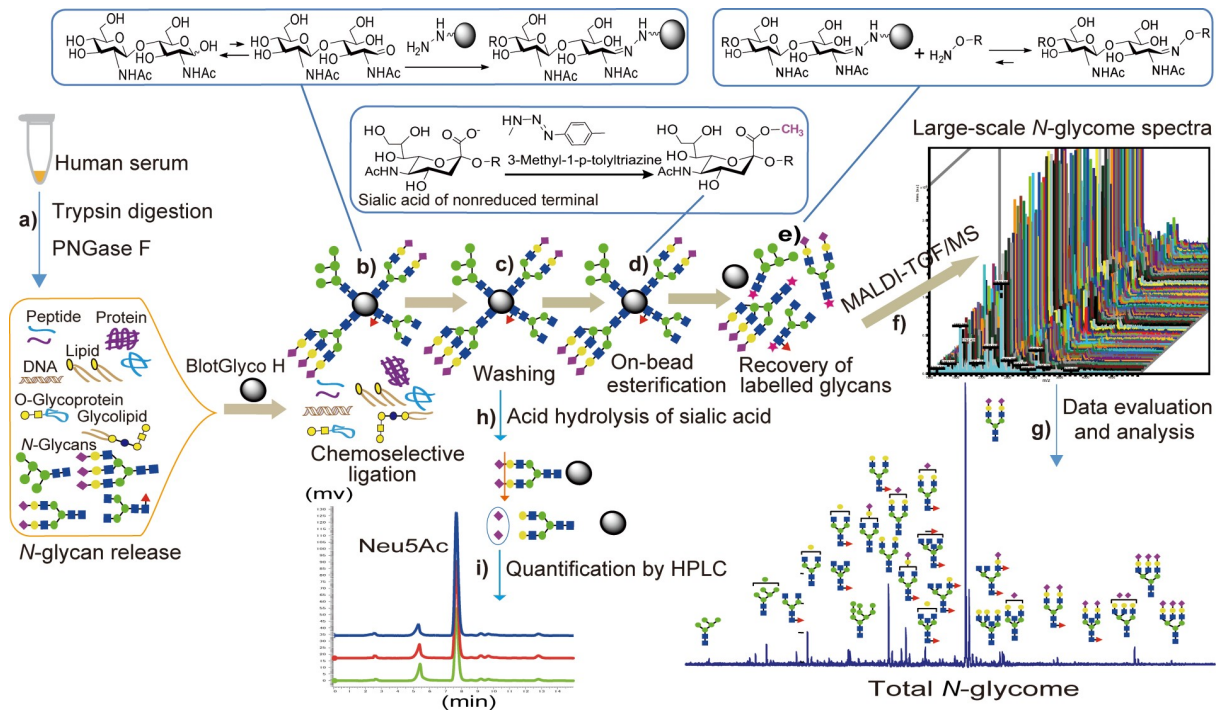
Bakelite, Co. Ltd. (Tokyo, Japan). Dithiothreitol (DTT), iodoacetamide (IAA), trypsin,  $\alpha$ -cyano-4-hydroxycinnamic acid diethylamine salt, 2-mercaptoethanol, and 5-*N*-Glycolylneuraminic acid (Neu5Gc) were from Sigma-Aldrich, Inc. (St. Louis, MO, USA). 5-*N*-acetylneuraminic acid (Neu5Ac) was purchased from Japan Food & Liquor Alliance Inc., Food & Bio Research Center, Inc. (Kyoto, Japan). 1,2-Diamino-4,5-methylenedioxybenzene (DMB) was from Dojindo Laboratories (Kumamoto, Japan). Other reagents and solvents were obtained from Wako Pure Chemicals Co., (Tokyo, Japan), unless otherwise stated. SweetBlot<sup>TM</sup> (automated glycan processing and incubating machine) was from System Instruments Co., Ltd., (Hachioji, Japan). MultiScreen Solvinert<sup>R</sup> filter plates were purchased from Millipore Co., Inc. (Tokyo, Japan). All mass quantifications were done using MALDI-TOF/MS (Ultraflex III, Bruker Daltonics, Germany). HPLC analysis was performed on Hitachi D-7000, Hitachi High Technologies Co., Ltd. (Tokyo, Japan).

### 2.3 Release of total *N*-glycans from human serum glycoproteins

A method based on high-throughput glyco-technology previously developed and optimized in our laboratory [11–13, 29–32] was used for glycan release, purification, labeling and spotting. Initially, 10  $\mu$ L of each serum sample was transferred to a 96 well polymerase chain reaction (PCR) plate and dissolved with 30  $\mu$ L of freshly prepared 0.33M ammonium bicarbonate (ABC) containing 0.1% of PHM in 10mM ABC and the solution was incubated at 37°C for 10 min., serially diluted human serum standards were prepared by mixing about 50  $\mu$ L serum aliquot of each healthy participant to give pooled sera, from which the concentration of standard serial dilutions (0.5 $\times$ , 0.75 $\times$ , 1.0 $\times$ , 1.25 $\times$ , 1.5 $\times$ , 1.75 $\times$ , 2.0 $\times$ , and 2.25 $\times$ ) were adjusted using Milli-Q water and then included in the experiment as the resulting calibration curve helps to evaluate the linearity and reproducibility of the detect peaks. As an internal standard, 12  $\mu$ L of 60  $\mu$ M disialyloctasaccharide was also added and mixed in each well to aid eventual quantification of detection *N*-glycans. Solubilized proteins were reduced by 10  $\mu$ L of 120 mM 1, 4-dithiothreitol (DTT) at 60°C for 30 min followed by alkylation with 20  $\mu$ L of 123 mM iodoacetamide by incubation in dark for 1 hr. The mixture was then treated with 10  $\mu$ L of 40 U/ $\mu$ L trypsin in 1 mM HCl at 37°C for 2 hrs. After heat-inactivation of the enzyme at 90°C for 10 minutes and then cooling to room temperature, *N*-glycans were released from trypsin-digested samples by incubation with 2 U of Peptide *N*-glycosidase F (PNGase F) at 37°C for 6 hrs.

### 2.4 Selective *N*-glycan enrichment by glycoblotting method

Once *N*-glycans were enzymatically released, oligosaccharides carrying reducing terminal were chemically ligated with hydrazide-functionalized BlotGlyco H bead, allowing their selective capturing from the complex mixtures of serum originated biomolecules as diagrammed in Fig 1. In this glycoblotting-based quantitative *N*-glycomics strategy, 250  $\mu$ L of BlotGlyco H bead was placed into each well of a MultiScreen Solvinert filter plate (Millipore) with vacuuming. 20  $\mu$ L of PNGase F digested mixture containing released *N*-glycans was then mixed with the bead in each well, followed by the addition of 180  $\mu$ L of 2% acetic acid (AcOH) in acetonitrile (ACN). To capture the *N*-glycans specifically onto beads via reversible hydrazone bonds, the plate was incubated at 80°C for 45 minutes. Next, the plate was washed twice with each 200  $\mu$ L of 2 M guanidine-HCl in ABC, water, and 1% triethyl amine in methanol (TEA in MeOH). To cap unreacted hydrazide functional groups on beads, 10% acetic anhydride in MeOH was added with incubation for 30 min at room temperature and then removing the solution by vacuum. The beads were then washed twice with each 200  $\mu$ L of 10 mM HCl, methanol, and dioxane. To prevent sialic acid dissociation under mild acidic condition or when directly ionized by MALDI-TOF/MS, on-bead esterification of its carboxyl groups was



**Fig 1. Schematic workflow of glycoblotting-assisted MALDI-TOF/MS and HPLC analysis.** (a) Enzymatic treatment for serum protein digest and deglycosylation using trypsin and PNGase F, respectively, (b) Chemoselective capturing of reducing sugars onto a hydrazide-functionalized BlotGlyco H beads from complex mixture of biomolecules, (c) Washing to remove all impurities, (d) On-bead methyl esterification of sialic acid residues, (e) Recovery of BSA-labeled glycans, (f) MALDI-TOF/MS analysis, (g) Data processing and evaluation, (h) Cleavage of free sialic acid residues from captured *N*-glycans, (i) Free sialic acid labeling and quantification by fluorescence-HPLC.

<https://doi.org/10.1371/journal.pone.0209515.g001>

carried out by incubation with fresh 100 mM 3-methyl-1-*p*-tolyltriazene (MTT) in dioxane at 60°C for 90 minutes. This approach allows simultaneous analysis of neutral and acidic (sialylated) glycans in positive-ion detection mode. Subsequently, each well was washed twice using 200 µL of dioxane, water, methanol, and water. The glycans blotted on the beads were labeled by trans-iminization reaction with 20 µL of 50 mM *O*-benzyloxyamine hydrochloride (BOA) and 180 µL of 2% AcOH in ACN with incubation at 80°C for 45 minutes. BOA-tagged *N*-glycans were finally eluted with 100 µL of Milli-Q water.

### 2.5 Serum *N*-glycan analysis by MALDI-TOF/MS

BOA-labeled *N*-glycans were directly dissolved with an equivalent volume of matrix solution (100 mM  $\alpha$ -cyano-4-hydroxycinnamic acid diethylamine salt), after which 2.5 µL of each sample-matrix mixture was auto-spotted in quadruplicate on MTP 384 target plate (polished steel TF, Bruker Daltonics). Ultraflex III mass spectrometry that works based on matrix-assisted laser desorption/ionization-time of flight/mass spectrometry (MALDI-TOF/MS) was used during which mass spectra were acquired in an automated manner using AutoXecute flexControl software (Bruker Daltonics, Germany) in reflector, positive ion mode, typically summing 1000 shots for each spot. The obtained mass spectra were further analyzed using FlexAnalysis v. 3 Software (Bruker Daltonics, Germany). The intensities from monoisotopic peaks of each quadruplicated spectra were normalized using known concentration of an internal standard and then averaged. This data was used for further statistical and quantitative comparison. Detected *N*-glycans were selected based on their quantitative reproducibility after evaluated using calibration curve of serially diluted human serum standards. *N*-glycan structural

compositions were assigned by GlycoMod (ExPASy proteomics server, Swiss Institute of Bioinformatics: <http://br.expasy.org/tools/glycomod/>) using experimental masses, and by CFG database (<http://www.functionalglycomics.org/glycomics/publicdata/home.jsp>).

## 2.6 Quantification of free sialic acids cleaved from *N*-glycans of human serum

To demonstrate the versatility of our comprehensive glycoblotting method for HPLC based sialic acid quantification, previously reported methods [33, 34] were slightly modified and integrated to our automated glycan enrichment approach. Briefly, after enzymatic release of serum *N*-glycans under the digestion conditions described above, 20  $\mu$ L from each sample containing released *N*-glycans was subjected to glycoblotting with BlotGlyco H beads in a similar procedure as described above except sialic acid esterification with MTT and subsequent procedures were excluded in this case. Next to acetyl capping and subsequent washing, 100  $\mu$ L of 25 mM HCl was added to each well, with sealing and incubating of the plate at 80°C for 1 hr to release terminal sialic acids via selective cleavage of the  $\alpha$ -glycosidic bond from the adjacent galactose residues (Fig 1H). The hydrolysate containing trimmed sialic acid was filtered and collected into PCR tubes. The filtrate was then reacted with 100  $\mu$ L of 7 mM 1,2-diamino-4,5-methylenedioxybenzene (DMB) reagent (prepared by dissolving DMB-2HCl powder in equal volumes of 1 M 2-mercaptoethanol and 18 mM Na<sub>2</sub>S<sub>2</sub>O<sub>4</sub>) with heating at 60°C for 2.5 hrs in dark, allowing labeling of sialic acids. This DMB reagent effectively derivatizes all type of sialic acids without any side reaction as it specifically reacts with their  $\alpha$ -keto acid moiety (see Fig 2). After stopping the reaction by cooling in ice water, 50  $\mu$ L of the mixture solution was transferred into vial tube from which 10  $\mu$ L was auto-injected into fluorescence HPLC for analysis on a reversed-phase column. The column was eluted at a flow rate of 1 mL/min using MeOH/ACN/H<sub>2</sub>O solution mixture (3:1:10, v/v). The fluorescence was detected at 448 nm using excitation at 373 nm in a D-7000 HPLC system equipped with an L-7485 fluorescence detector (Hitachi High-Technologies Co., Tokyo, Japan). Standard mixtures of *N*-acetylneuraminic acid (Neu5Ac) and *N*-glycolylneuraminic acid (Neu5Gc, a hydroxylated form of the common sialic acid, Neu5Ac) were run parallel with the samples to normalize the retention time of each peak on the column, allowing quantitative analysis.

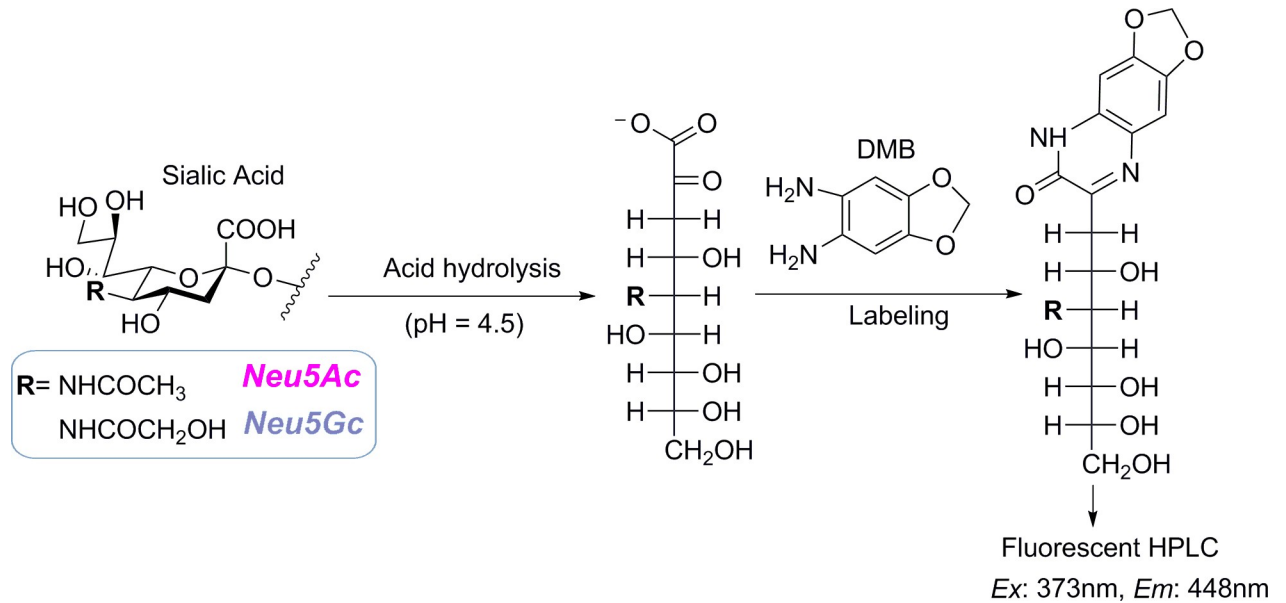
## 2.7 Statistical analysis

*N*-glycan peaks detected in MALDI-TOF/MS spectra were annotated using FlexAnalysis 3.0 software (Bruker Daltonics, Germany). Normalized data for expression levels of *N*-glycans were analyzed using SPSS software. Multiple comparisons among the ethnic groups were done using one-way analysis of variance (ANOVA). P-values were adjusted for multiple testing using Bonferroni method and mean value differences were considered significant at 95% confidence interval ( $p \leq 0.05$ ). We also used Graph Pad prism 5 to show the serum *N*-glycan and free sialic acid level of individual samples in a scatter dot plot.

## Results

### 3.1 Quantitative reproducibility test using standard human serum

To select the reliably detected *N*-glycans, each peak was evaluated for its quantitative reproducibility using serially diluted standard human serum samples (0.5 $\times$ , 0.75 $\times$ , 1.0 $\times$ , 1.25 $\times$ , 1.5 $\times$ , 1.75 $\times$ , 2.0 $\times$ , and 2.25 $\times$ ) that were simultaneously experimented in the same plate beside the study samples. The peak intensity of each glycan was first normalized using known concentration of an internal standard. Standard calibration curve for each *N*-glycan across the dilution



**Fig 2. Scheme for hydrolytic cleavage of terminal sialic acid and its labeling for fluorescence detection.** DMB: 1,2-diamino-4,5-methylenedioxybenzene, Neu5Ac: N-acetylneuraminic acid, Neu5Gc: N-glycolylneuraminic acid.

<https://doi.org/10.1371/journal.pone.0209515.g002>

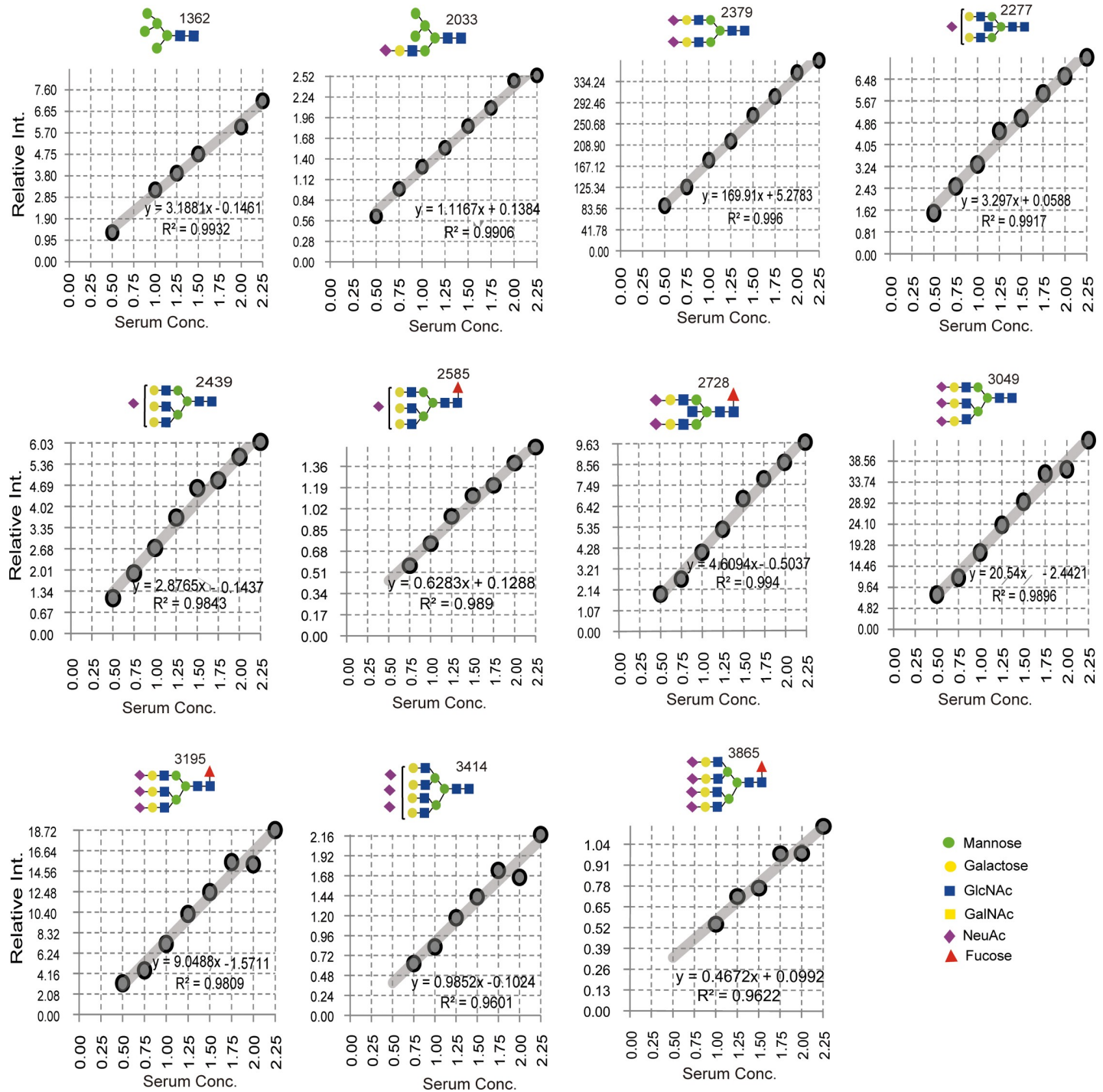
series was plotted using the normalized data as shown in Fig 3. N-glycans peaks that met the criteria of at least  $p < 0.05$  detection reliability, accessibility in at least six of the total eight human serum standards, and minimum outlier scores were selected and considered for quantitative comparison in the result of main study samples.

### 3.2 Inter-ethnic variation in the total serum N-glycan profile

The comprehensive N-glycomics study mainly emphasized investigating the association between serum glycan expression pattern and ethnic variation among 54 healthy individuals comprising Japanese, South Indians, Ethiopians, and US origin controls. Mass spectral data acquired by MALDI-TOF/MS analysis demonstrated noticeable variation in the peak intensities of several glycans with the highest signals appeared in the Ethiopian ethnic group (Fig 4). Taking the entire study samples, detection of 51 N-glycans (Table 1) was quantitatively verified of which 33 N-glycans were common to all ethnic groups. Detection level of the remaining glycans was  $< 5 \mu\text{M}$  and exclusive either to the Ethiopian ethnic group (13 glycans, Table 1 highlighted in gold color shade) or to all the other ethnic groups (5 glycans, Table 1 highlighted in light green color shade). Here, it should be clear that glycans whose detection profile was not quantitatively reliable or limited to some samples were not considered to avoid biased conclusions. Overall, 40 (78.43%) of the detected N-glycans belong to complex type whereas high-mannose and hybrid types comprise 5 (9.8%) and 6 (11.76%), respectively.

The  $m/z$  values indicate experimental masses of N-glycans tagged with BOA for enhancing their ionization potentials. Glycans whose peaks # is highlighted by color shade had generally weaker detection profile in which those highlighted in gold color shade were specific to the Ethiopian group while those highlighted in light green shade were common to the remaining groups. Peak number 26 detected at  $m/z$  2175.78 represents an internal standard (I.S).

We quantitatively compared the expression levels of commonly detected N-glycans among the ethnic groups by ANOVA test. Ethnicity-associated variation over a wide range of serum N-glycans with a consistent highest abundance (except one glycan drastically declined;  $m/z$  of

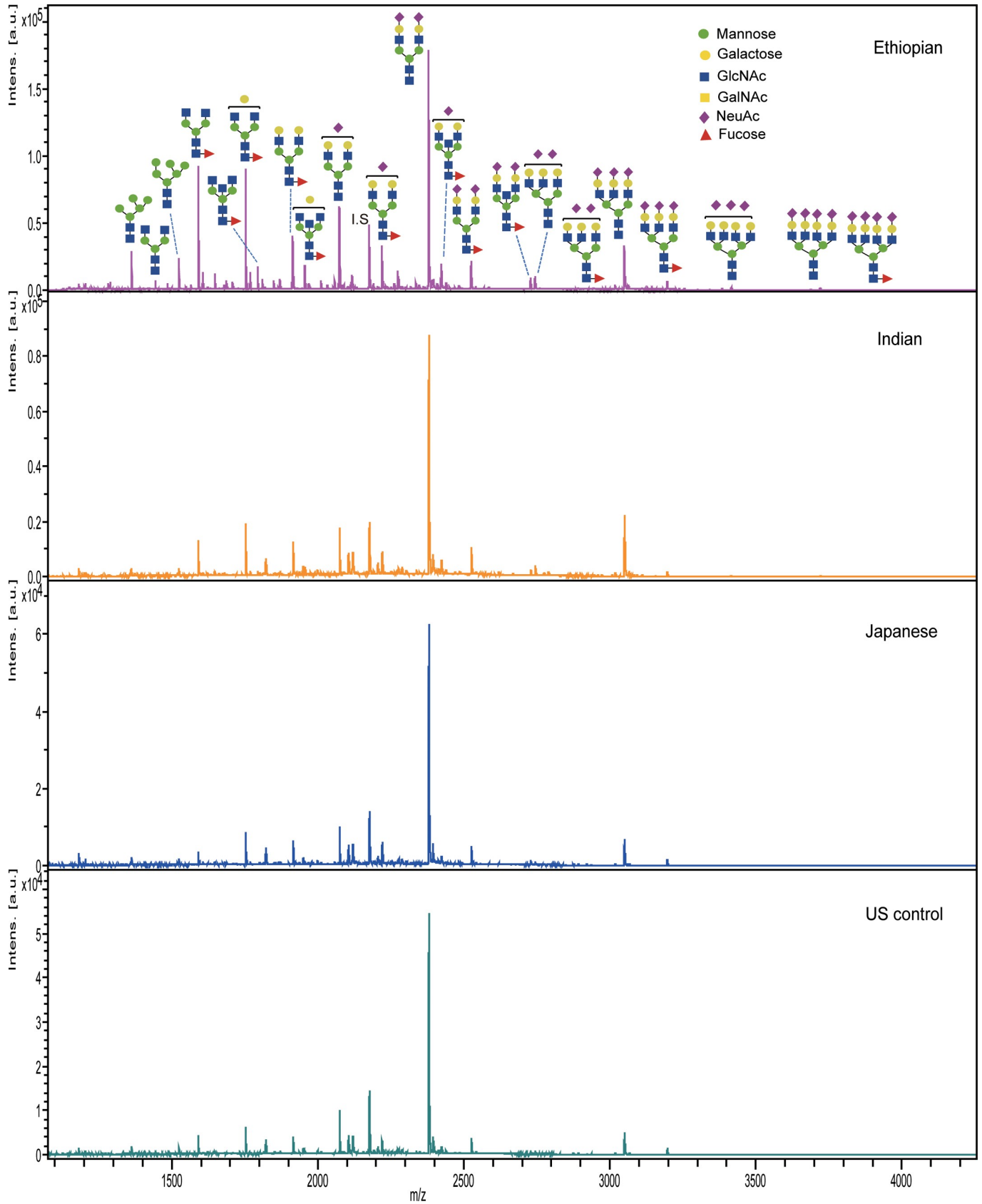


**Fig 3. Human serum standard calibration curve of some common N-glycans.** Glycans that have shown quantitatively reproducible detection across the various standard serum concentrations were considered and selected in the case of the main study samples. For identification purpose, individual glycans are coded or labeled by their molecular weight (*m/z* value) shown near to their structure.

<https://doi.org/10.1371/journal.pone.0209515.g003>

2118) in the Ethiopian subjects ( $p < 0.001$ ) was observed [Fig 5A–5C]. Structurally, these significantly altered glycoforms are predominantly composed of high mannose, core-fucosylated, hyperbranched or hypersialylated types. Moreover, the trend for glycan abundance in





**Fig 4. Representative MALDI-TOF/MS spectra of human serum *N*-glycans acquired from subjects of various ethnicity.** Structures for some *N*-glycans that demonstrated differential peak intensities are indicated over the respective glycan peak.

<https://doi.org/10.1371/journal.pone.0209515.g004>

Ethiopians tends to be isolated from the other healthy ethnic groups, rather seems closer to that of Japanese hepatocellular carcinoma (HCC) patients [29] whose serum was concurrently re-examined for *N*-glycan profiling. Compared to the US controls, the Japanese and the Indian healthy human serum glycome showed a unidirectional increase in expression level whereas a statistically non-significant expression difference was noticed between the two Asian ethnic groups. As an exception, the result pointed out that a bisect-type *N*-glycan having *m/z* of 2118 showed a significant ( $p < 0.001$ ) concentration alteration among all ethnic groups with the highest in the Indian and the lowest in the Ethiopian groups [see *m/z* 2118 in Fig 5B].

Since Ethiopian subjects were females, for more clarity on whether differences in ethnicity, gender, or age seem to have more marked effects, we further included one Indian female (age = 43) and one Japanese female (age = 40s), whose serum *N*-glycome pattern in comparison with one Ethiopian female (age = 40) is diagramed in S1 Fig. For further emphasis on the gender effect, spectral *N*-glycome profile of Indian and Japanese subjects of both sexes is provided in S2 Fig. within which peak intensity of many *N*-glycans varied mainly in an ethnicity dependent manner, but irrespective of the gender difference. Altogether, these results show clear differences in the glycan abundance that were more strongly associated with ethnic differences than gender or age variation. The effect of age on serum glycan level within each ethnic group of this study was statistically non-significant.

To address the inter-individual variation of each *N*-glycan level from one ethnic data series to another, we determined coefficient of variation (CoV) for each glycan as the ratio of the standard deviation to the mean of glycan expression. The degree of variation was found to be lowest across the entire glycans in the US control sample. On the other hand, there were no consistent inter-subject differences among Japanese, Indian and Ethiopian ethnicities up to *m/z* of 2500, after which the hyperbranched and hypersialylated glycans showed widely dispersed expression pattern in the Indian and the Japanese groups (Fig 6).

### 3.3 Glycotyping analysis

After *N*-glycans were stratified into glyco-subclasses that share certain structural features of core-fucosylation, bisecting, sialylation, and branching (more details provided in S2 Table), their overall expression pattern among the study groups is shown in Fig 7. The result clearly revealed that profound abundance of these glycotypes (except the bisecting) was found to be associated with the Ethiopian population. Ethnic-based greater disparities in the serum level of the glycoforms was mainly pronounced towards the higher *m/z* tri-/tetra-sialylated and tri-/tetra-antennary structures. Next to Ethiopians, consistent declining trend in the expression level was noticed in Indians, Japanese, and finally US controls, respectively. Minimum inter-ethnic differences in the level of bisecting glycoforms was observed across all ethnicities.

### 3.4 Sialic acid quantification in human serum

We also performed a glycoblotting-assisted HPLC-based quantification of free sialic acid species cleaved from sialylated *N*-glycans captured by glycoblotting of human serum glycoproteins. 1,2-diamino-4,5-methylenedioxybenzene (DMB) facilitated selective labeling of sialic acid moieties was performed prior to HPLC-fluorescence detection. Standard solutions of *N*-acetylneuraminic acid (Neu5Ac) and *N*-glycolylneuraminic acid (Neu5Gc) in 50–750  $\mu$ M concentration ranges were analyzed in parallel with the serum samples. Samples from four ethnic

**Table 1. Estimated composition of 51 N-linked glycans identified from human serum glycoproteins.**

Peak #	m/z	Glycan composition	Type
1	1362.48109	(Hex) <sub>2</sub> + (Man) <sub>3</sub> (GlcNAc) <sub>2</sub>	High-Mannose
2	1444.53419	(HexNAc) <sub>2</sub> + (Man) <sub>3</sub> (GlcNAc) <sub>2</sub>	Complex
3	1524.53392	(Hex) <sub>3</sub> + (Man) <sub>3</sub> (GlcNAc) <sub>2</sub>	High-Mannose
4	1565.56047	(Hex) <sub>2</sub> (HexNAc) <sub>1</sub> + (Man) <sub>3</sub> (GlcNAc) <sub>2</sub>	Hybrid
5	1590.59210	(HexNAc) <sub>2</sub> (Deoxyhexose) <sub>1</sub> + (Man) <sub>3</sub> (GlcNAc) <sub>2</sub>	Complex
6	1606.58702	(Hex) <sub>1</sub> (HexNAc) <sub>2</sub> + (Man) <sub>3</sub> (GlcNAc) <sub>2</sub>	Complex
7	1686.58675	(Hex) <sub>4</sub> + (Man) <sub>3</sub> (GlcNAc) <sub>2</sub>	High-Mannose
8	1708.61871	(Hex) <sub>1</sub> (HexNAc) <sub>1</sub> (NeuAc) <sub>1</sub> + (Man) <sub>3</sub> (GlcNAc) <sub>2</sub>	Complex
9	1752.64493	(Hex) <sub>1</sub> (HexNAc) <sub>2</sub> (Deoxyhexose) <sub>1</sub> + (Man) <sub>3</sub> (GlcNAc) <sub>2</sub>	Complex
10	1768.63985	(Hex) <sub>2</sub> (HexNAc) <sub>2</sub> + (Man) <sub>3</sub> (GlcNAc) <sub>2</sub>	Complex
11	1793.67148	(HexNAc) <sub>3</sub> (Deoxyhexose) <sub>1</sub> + (Man) <sub>3</sub> (GlcNAc) <sub>2</sub>	Complex
12	1848.63958	(Hex) <sub>5</sub> + (Man) <sub>3</sub> (GlcNAc) <sub>2</sub>	High-Mannose
13	1854.67662	(Hex) <sub>1</sub> (HexNAc) <sub>1</sub> (Deoxyhexose) <sub>1</sub> (NeuAc) <sub>1</sub> + (Man) <sub>3</sub> (GlcNAc) <sub>2</sub>	Complex
14	1870.67154	(Hex) <sub>2</sub> (HexNAc) <sub>1</sub> (NeuAc) <sub>1</sub> + (Man) <sub>3</sub> (GlcNAc) <sub>2</sub>	Hybrid
15	1873.874	(Hex) <sub>3</sub> (HexNAc) <sub>1</sub> (Deoxyhexose) <sub>1</sub> + (Man) <sub>3</sub> (GlcNAc) <sub>2</sub>	Hybrid
16	1911.69809	(Hex) <sub>1</sub> (HexNAc) <sub>2</sub> (NeuAc) <sub>1</sub> + (Man) <sub>3</sub> (GlcNAc) <sub>2</sub>	Complex
17	1914.69776	(Hex) <sub>2</sub> (HexNAc) <sub>2</sub> (Deoxyhexose) <sub>1</sub> + (Man) <sub>3</sub> (GlcNAc) <sub>2</sub>	Complex
18	1955.72431	(Hex) <sub>1</sub> (HexNAc) <sub>3</sub> (Deoxyhexose) <sub>1</sub> + (Man) <sub>3</sub> (GlcNAc) <sub>2</sub>	Complex
19	1996.967	(HexNAc) <sub>4</sub> (Deoxyhexose) <sub>1</sub> + (Man) <sub>3</sub> (GlcNAc) <sub>2</sub>	Complex
20	2010.69241	(Hex) <sub>6</sub> + (Man) <sub>3</sub> (GlcNAc) <sub>2</sub>	High-Mannose
21	2032.72437	(Hex) <sub>3</sub> (HexNAc) <sub>1</sub> (NeuAc) <sub>1</sub> + (Man) <sub>3</sub> (GlcNAc) <sub>2</sub>	Hybrid
22	2057.75600	(Hex) <sub>1</sub> (HexNAc) <sub>2</sub> (Deoxyhexose) <sub>1</sub> (NeuAc) <sub>1</sub> + (Man) <sub>3</sub> (GlcNAc) <sub>2</sub>	Complex
23	2073.75092	(Hex) <sub>2</sub> (HexNAc) <sub>2</sub> (NeuAc) <sub>1</sub> + (Man) <sub>3</sub> (GlcNAc) <sub>2</sub>	Complex
24	2117.77714	(Hex) <sub>2</sub> (HexNAc) <sub>3</sub> (Deoxyhexose) <sub>1</sub> + (Man) <sub>3</sub> (GlcNAc) <sub>2</sub>	Complex
25	2166.284	(Hex) <sub>3</sub> (HexNAc) <sub>1</sub> (Deoxyhexose) <sub>3</sub> + (Man) <sub>3</sub> (GlcNAc) <sub>2</sub>	Hybrid
26	2175.78261	(Hex) <sub>2</sub> (HexNAc) <sub>2</sub> (NeuAc) <sub>2</sub> + (Man) <sub>3</sub> (GlcNAc) <sub>1</sub>	I.S
27	2219.80883	(Hex) <sub>2</sub> (HexNAc) <sub>2</sub> (Deoxyhexose) <sub>1</sub> (NeuAc) <sub>1</sub> + (Man) <sub>3</sub> (GlcNAc) <sub>2</sub>	Complex
28	2260.83538	(Hex) <sub>1</sub> (HexNAc) <sub>3</sub> (Deoxyhexose) <sub>1</sub> (NeuAc) <sub>1</sub> + (Man) <sub>3</sub> (GlcNAc) <sub>2</sub>	Complex
29	2263.83505	(Hex) <sub>2</sub> (HexNAc) <sub>3</sub> (Deoxyhexose) <sub>2</sub> + (Man) <sub>3</sub> (GlcNAc) <sub>2</sub>	Complex
30	2276.83030	(Hex) <sub>2</sub> (HexNAc) <sub>3</sub> (NeuAc) <sub>1</sub> + (Man) <sub>3</sub> (GlcNAc) <sub>2</sub>	Complex
31	2301.444	(HexNAc) <sub>4</sub> (Deoxyhexose) <sub>1</sub> (NeuAc) <sub>1</sub> + (Man) <sub>3</sub> (GlcNAc) <sub>2</sub>	Complex
32	2336.85144	(Hex) <sub>3</sub> (HexNAc) <sub>4</sub> + (Man) <sub>3</sub> (GlcNAc) <sub>2</sub>	Complex
33	2378.86199	(Hex) <sub>2</sub> (HexNAc) <sub>2</sub> (NeuAc) <sub>2</sub> + (Man) <sub>3</sub> (GlcNAc) <sub>2</sub>	Complex
34	2422.88821	(Hex) <sub>2</sub> (HexNAc) <sub>3</sub> (Deoxyhexose) <sub>1</sub> (NeuAc) <sub>1</sub> + (Man) <sub>3</sub> (GlcNAc) <sub>2</sub>	Complex
35	2438.88313	(Hex) <sub>3</sub> (HexNAc) <sub>3</sub> (NeuAc) <sub>1</sub> + (Man) <sub>3</sub> (GlcNAc) <sub>2</sub>	Complex
36	2483.89335	(Hex) <sub>3</sub> (HexNAc) <sub>1</sub> (Deoxyhexose) <sub>1</sub> (NeuAc) <sub>2</sub> + (Man) <sub>3</sub> (GlcNAc) <sub>2</sub>	Hybrid
37	2520.93623	(Hex) <sub>1</sub> (HexNAc) <sub>5</sub> (NeuAc) <sub>1</sub> + (Man) <sub>3</sub> (GlcNAc) <sub>2</sub>	Complex
38	2524.91990	(Hex) <sub>2</sub> (HexNAc) <sub>2</sub> (Deoxyhexose) <sub>1</sub> (NeuAc) <sub>2</sub> + (Man) <sub>3</sub> (GlcNAc) <sub>2</sub>	Complex
39	2581.94137	(Hex) <sub>2</sub> (HexNAc) <sub>3</sub> (NeuAc) <sub>2</sub> + (Man) <sub>3</sub> (GlcNAc) <sub>2</sub>	Complex
40	2584.94104	(Hex) <sub>3</sub> (HexNAc) <sub>3</sub> (Deoxyhexose) <sub>1</sub> (NeuAc) <sub>1</sub> + (Man) <sub>3</sub> (GlcNAc) <sub>2</sub>	Complex
41	2727.99928	(Hex) <sub>2</sub> (HexNAc) <sub>3</sub> (Deoxyhexose) <sub>1</sub> (NeuAc) <sub>2</sub> + (Man) <sub>3</sub> (GlcNAc) <sub>2</sub>	Complex
42	2743.99420	(Hex) <sub>3</sub> (HexNAc) <sub>3</sub> (NeuAc) <sub>2</sub> + (Man) <sub>3</sub> (GlcNAc) <sub>2</sub>	Complex
43	2787.89	(Hex) <sub>3</sub> (HexNAc) <sub>4</sub> (Deoxyhexose) <sub>1</sub> (NeuAc) <sub>1</sub> + (Man) <sub>3</sub> (GlcNAc) <sub>2</sub>	Complex
44	2890.05211	(Hex) <sub>3</sub> (HexNAc) <sub>3</sub> (Deoxyhexose) <sub>1</sub> (NeuAc) <sub>2</sub> + (Man) <sub>3</sub> (GlcNAc) <sub>2</sub>	Complex
45	3007.09472	(Hex) <sub>4</sub> (HexNAc) <sub>5</sub> (NeuAc) <sub>1</sub> + (Man) <sub>3</sub> (GlcNAc) <sub>2</sub>	Complex
46	3049.10527	(Hex) <sub>3</sub> (HexNAc) <sub>3</sub> (NeuAc) <sub>3</sub> + (Man) <sub>3</sub> (GlcNAc) <sub>2</sub>	Complex
47	3109.12641	(Hex) <sub>4</sub> (HexNAc) <sub>4</sub> (NeuAc) <sub>2</sub> + (Man) <sub>3</sub> (GlcNAc) <sub>2</sub>	Complex

(Continued)

Table 1. (Continued)

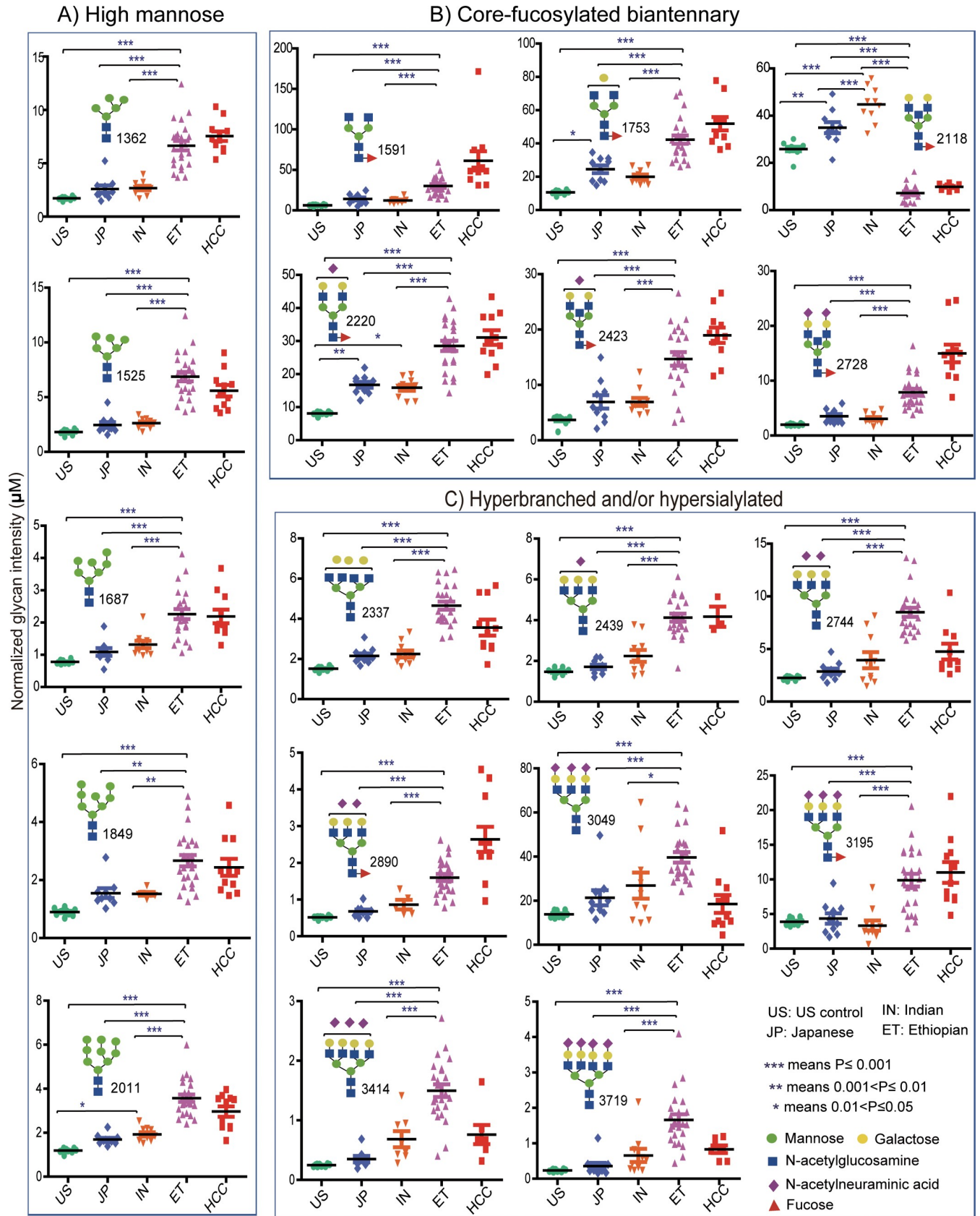
Peak #	m/z	Glycan composition	Type
48	3195.16318	(Hex)3 (HexNAc)3 (Deoxyhexose)1 (NeuAc)3 + (Man)3 (GlcNAc)2	Complex
49	3414.23748	(Hex)4 (HexNAc)4 (NeuAc)3 + (Man)3 (GlcNAc)2	Complex
50	3560.29539	(Hex)4 (HexNAc)4 (Deoxyhexose)1 (NeuAc)3 + (Man)3 (GlcNAc)2	Complex
51	3719.34855	(Hex)4 (HexNAc)4 (NeuAc)4 + (Man)3 (GlcNAc)2	Complex
52	3865.40646	(Hex)4 (HexNAc)4 (Deoxyhexose)1 (NeuAc)4 + (Man)3 (GlcNAc)2	Complex

<https://doi.org/10.1371/journal.pone.0209515.t001>

groups (US controls, Japanese, Indians and Ethiopians) were used in this experiment. The result demonstrated a consistent signal intensity pattern for the standard Neu5Gc and Neu5Ac in a concentration dependent manner as shown by HPLC profiles in Fig 8A), suggesting the reliability of the quantification method. Focusing on the study samples, highest peak intensity for Neu5Ac was primarily found in the Ethiopian group (Fig 8B) whose absolute concentration (average  $\mu\text{M} \pm \text{SE} = 234.23 \pm 24.15$ ), as determined based on the standard Neu5Ac calibration curve (Fig 9A), was significantly higher ( $p < 0.001$ ,  $p = 0.002$ ,  $p = 0.01$ ) comparing with that of US controls ( $105.51 \pm 9.39$ ), Japanese ( $139.84 \pm 20.86$ ), and Indians ( $155.69 \pm 14.27$ ), respectively (Fig 9B). Apart from inter-ethnic variation, Neu5Ac level was observed to be very low in only a few of the Ethiopian subjects, implicating the need to consider inter-individual disparities in the glycosylation process. However, Neu5GC was identified in none of the study subjects which is not surprising as this type of sialic acid normally does not exist in healthy human serum glycoproteins [35]. Furthermore, HPLC chromatogram produced from serum Neu5Ac level of three age-matched females varying in ethnicity (Ethiopian 40 y/o, Indian, 43 y/o, Japanese 40s y/o) is provided in S3 Fig which shows marked differences in their Neu5Ac peak intensity, being most abundant in the Ethiopian subject. This observation was consistent with the result of total serum N-glycome spectra of the same individuals [S1 Fig], all of which indicating that healthy human serum N-glycosylation pattern seems to be affected more strongly by ethnic difference than gender or age variation of the participants.

## Discussion

Aiming at investigating the inter-ethnic physiological variations in serum N-glycome profile, we performed a glycoblotting-assisted MALDI-TOF/MS-based quantitative analysis focusing on 54 healthy subjects of various ethnicity. Total of 51 N-glycoforms could be identified and reproducibly quantified with an evident differential expression pattern among the ethnic groups. Among all, Ethiopian ethnic group exhibited the most isolated detection trend with greatly increased expression levels of particularly high mannose, Core-fucosylated, multi-antennary, and multi-sialylated glycans. Notably, some less intense glycans had exclusive occurrence in which 13 glycan structures were specific to Ethiopians whereas other 5 glycoforms were identified to be associated with the remaining ethnicities. Healthy Indian and Japanese serum glycoproteins seem to have mostly similar N-glycan composition and abundance with slight differences in the expression level of few N-glycoforms. It is interesting to note that one glycoform carrying a core-fucose and a bisecting GlcNAc structures (m/z 2118) showed a clear-cut expression difference among the four study groups. It is the only glycoform whose serum level was markedly decreased in Ethiopians comparing with the other ethnic groups. Declining of this glycan in Ethiopian subjects is most probably due to its involvement as a substrate in the subsequent biosynthetic steps, the product of which (m/z 2423 and m/z 2728 in Fig 5B) could be up-regulated comparing with the other ethnic populations. We have found

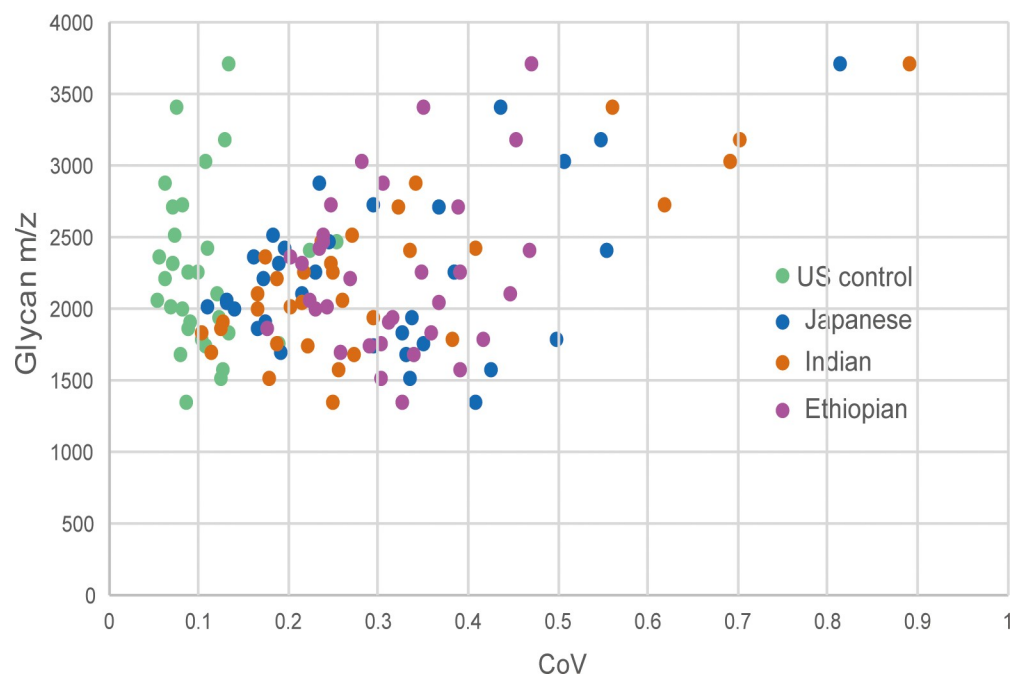


**Fig 5. Dot plot illustrating comparative expression abundance of selected serum *N*-glycans.** Proposed glycan structure and statistically significant differences among the groups is indicated for each glycoform. **A)** High mannose type, **B)** Core-fucosylated and biantennary structures, **C)** Hyperbranched and/or hypersialylated structures. **HCC:** Hepatocellular carcinoma patients.

<https://doi.org/10.1371/journal.pone.0209515.g005>

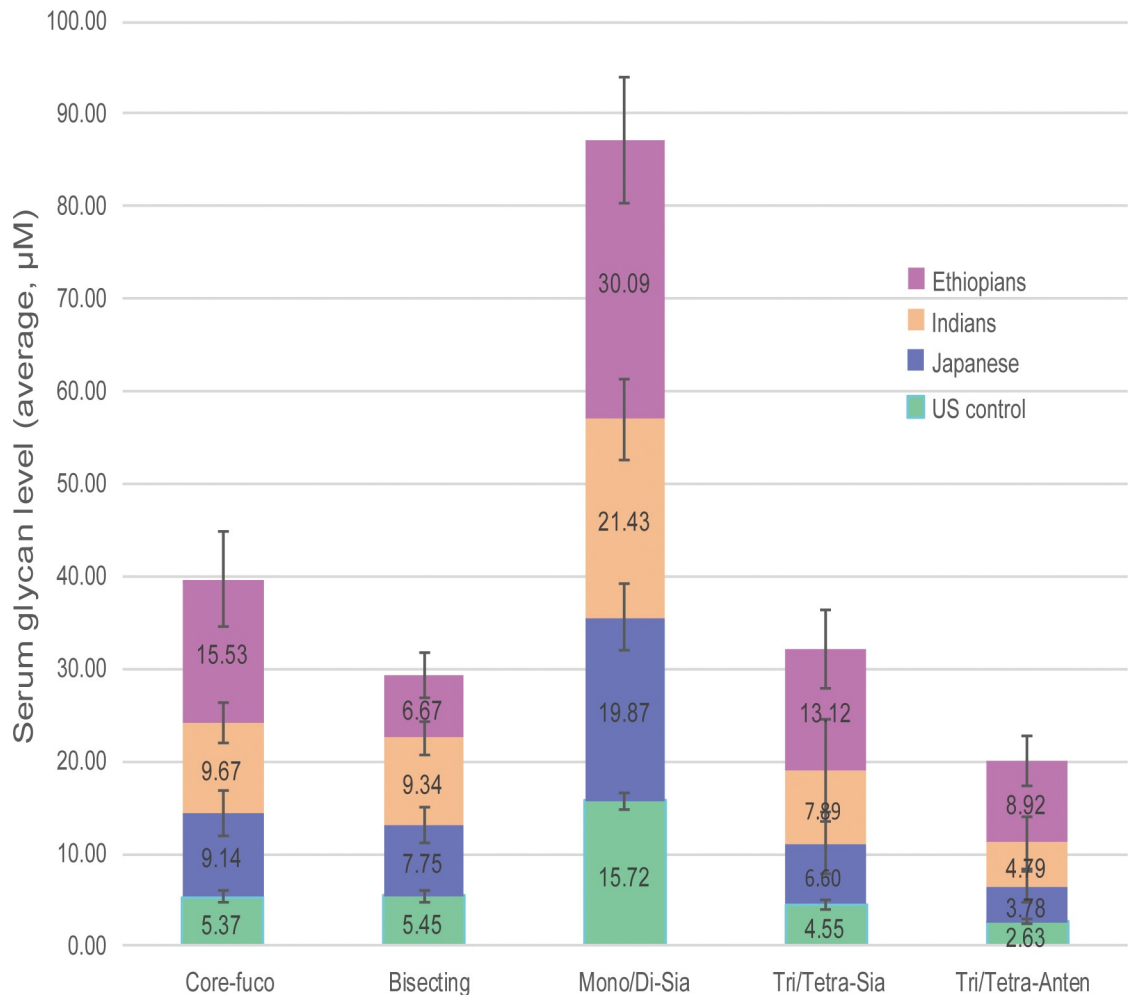
high inter-individual variations in the expression level of multiply branched and sialylated glycan structures in the Indian and Japanese subjects (Fig 6). Combining both observations of lower inter-individual variation and higher glycan concentration (compared with Indians and Japanese) found in Ethiopians, it can be suggested that total serum glycan concentration was elevated across the entire Ethiopian subjects. Marked differences in glycotyping analysis, in which up to 2–3 times greater abundance particularly in the multi-branched or multi-sialylated glycan features, were associated with Ethiopian ethnic group (Fig 7). This observation was consistent with the above-mentioned differential expression pattern of the individual *N*-glycans.

Increasing evidences show that glycan abundance is regulated by multiple molecular mechanisms that rely on metabolic interplay among genes, sugars, proteins, and lipids [1, 9]. In glycoproteins, the amino acid sequences of the core protein mostly remain stable while the glycan moiety undergoes faster alterations in response to physiological, pathological, and environmental stimuli [36–38]. Despite previous studies that have reported the correlation between changes in plasma/IgG glycosylation patterns and aging process [26–28], there was no age wise significant variation of healthy serum *N*-glycome profile in the current study, probably due to the relatively younger age range of the healthy participants. With quite different background of our study subjects, the observed changes in the present study may be explained by genetic, environmental, (nutritional habits, physical activity, exposure to pathogens, stress level), and sociocultural variations, while the degree to which each factor influences serum



**Fig 6. CoV illustrating the inter-subject extent of variation within each ethnic group.** Y-axis shows *m/z* values of detected *N*-glycans, X-axis shows the coefficient of variation (CoV) value representing how far the glycan concentration was scattered within subjects of each ethnic group.

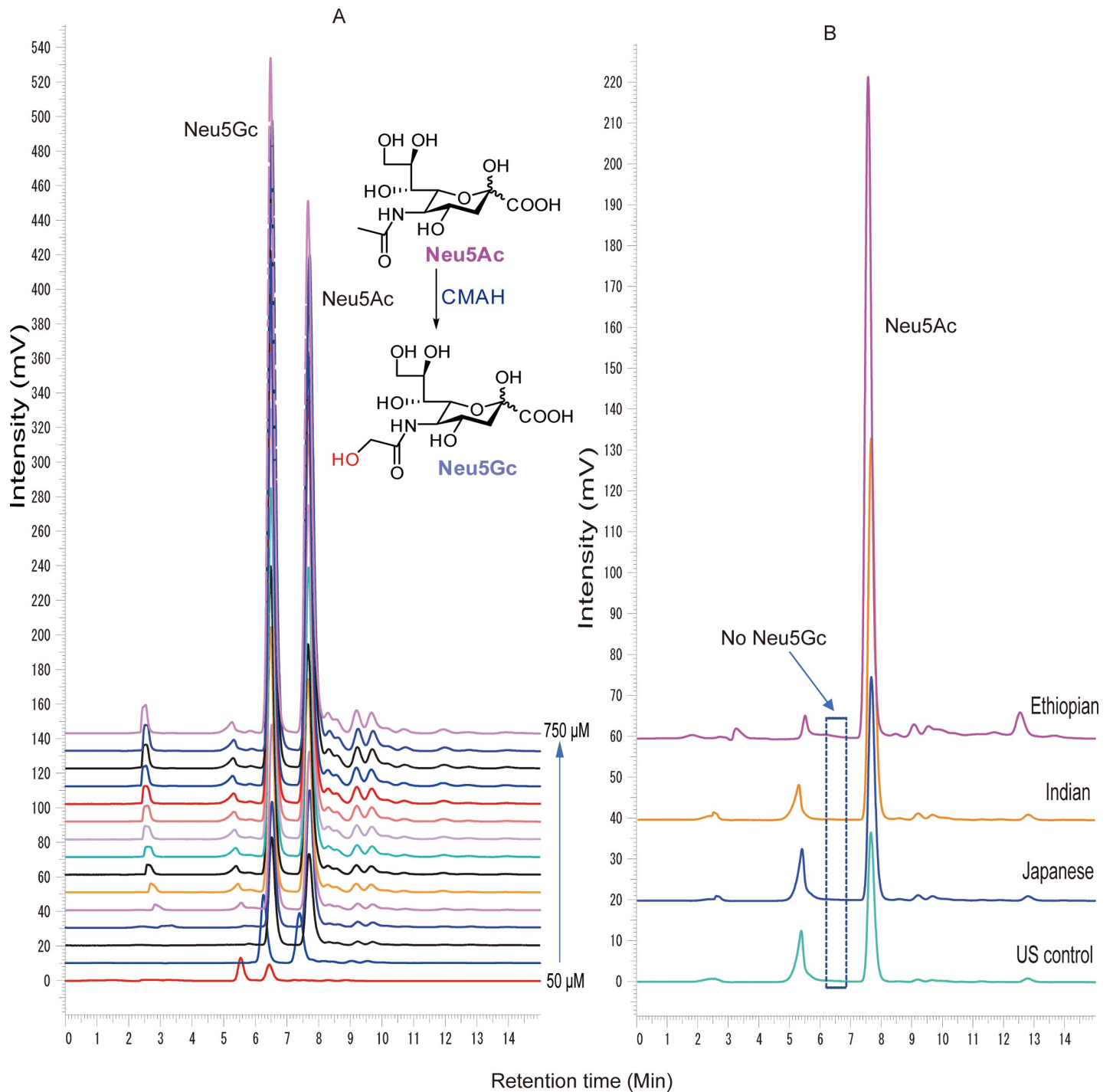
<https://doi.org/10.1371/journal.pone.0209515.g006>



**Fig 7. Ethnic-based serum level of some common glyco-subclasses.** *N*-glycans were grouped in terms of core-fucosylation, bisecting, mono- or di-sialylation, tri- or tetra-sialylation, tri- or tetra-antennary features. Error bars indicate standard errors of the means.

<https://doi.org/10.1371/journal.pone.0209515.g007>

glycan level may depend on the specific ethnic population. These variations have been emphasized by previous reports addressing the variability, heritability, genetic and environmental determinants of human plasma and IgG *N*-glycome using chromatographic techniques [26, 39, 40]. Given the complexity of glycosylation at cellular and molecular levels, these diverse factors can ultimately affect the activity of glycosyltransferases and glycosidases enzymes that orchestrate the glycan biosynthetic and degradative pathways [41]. In this context, despite the non-template-based synthesis of glycans, their structures and expression abundance have been reported to be regulated by genetic and epigenetic factors [42, 43], partly accounting for the inter-individual glycan signature variations observed in health and disease conditions. It was also pointed out that variations in the composition of plasma protein *N*-glycans, particularly increased branching, galactosylation and sialylation features, have been associated with metabolic syndrome related risk factors and higher risk of developing type 2 diabetes [18, 19]. While literatures on the association of healthy human serum glycosylation signatures with ethnicity are scarcely available, numerous studies have witnessed ethnic-specific differences in a number of biochemical markers [44, 45]. Although appreciating population diversity during

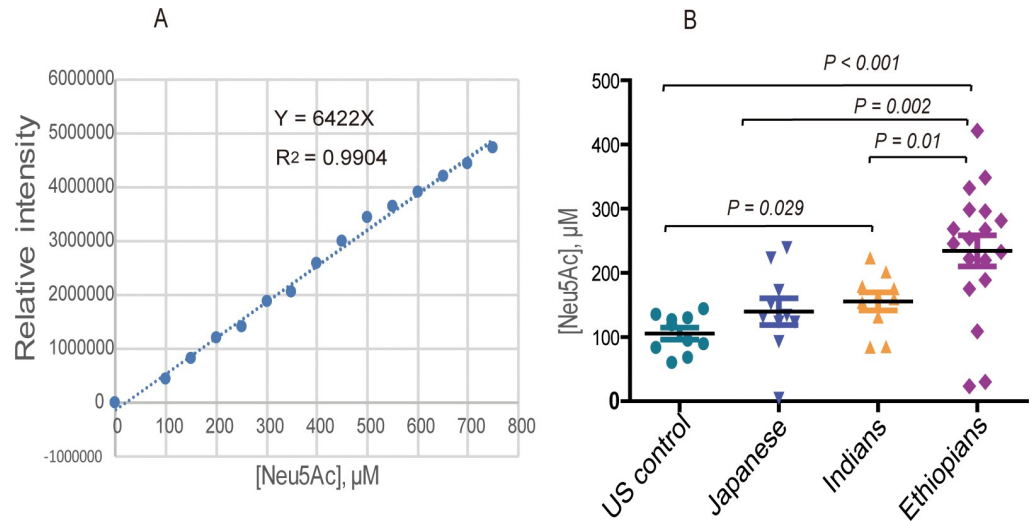


**Fig 8. HPLC chromatogram showing quantitation of free sialic acids.** A) standard Neu5Gc and standard Neu5Ac, B) Neu5Ac derived from captured serum *N*-glycans. Neu5Ac: *N*-acetylneuraminic acid, Neu5Gc: *N*-glycolylneuraminic acid. CMAH: Cytidine monophosphate *N*-acetylneuraminic acid hydroxylase (enzyme that catalyzes conversion of Neu5Ac to Neu5Gc).

<https://doi.org/10.1371/journal.pone.0209515.g008>

comprehensive experimental studies is mostly uncommon, we have considered black Ethiopians/Africans in the present study which can be taken as an advantage as pattern of glycosylation in black population, from physiological or pathological perspective, has not been addressed elsewhere.





**Fig 9. HPLC-based Neu5Ac quantification.** A) standard Neu5Ac calibration curve, B) Dot plot showing Neu5Ac expression levels derived and quantified from human serum *N*-glycans of four ethnic groups.

<https://doi.org/10.1371/journal.pone.0209515.g009>

Surprisingly, some of the specific *N*-glycoforms observed to be exclusive or elevated in the Ethiopian ethnic group including the hyper-branched structures with sialic acid residues (m/z 1362, 1591, 3195, 3560, 3865) have previously been identified as sensitive serum biomarkers as a significant recurrence factor of HCC using large-scale Japanese samples [29]. In the present study, these glycans could demonstrate nearly similar serum expression pattern between the Ethiopian group and the HCC group (11 Japanese patients) up on simultaneous experimentation (Fig 5). Similarly, among the hyperbranched glycans strongly increased in Ethiopians ( $p < 0.001$ ), it was demonstrated that the glycans with m/z 2337, 2439, and 2890 could become promising prognostic biomarkers in renal cell carcinoma [46] while m/z 3049 and 3414 have been associated significantly with metastatic castration-resistant status in prostate cancer [47]. Hence, alteration in serum *N*-glycan profile seems not exclusive for pathological conditions as the present results also clearly demonstrated among healthy subjects in an ethnicity dependent manner. Altogether, these observations emphasize the substantial impact of ethnic differences in human serum *N*-glycome variation, the ignorance of which may provide unclear and imprecise conclusion of the diagnosis by using glycan-related disease biomarkers.

Free sialic acid quantitation result among the four ethnic groups revealed nonnegligible ethnic differences in the serum Neu5Ac level in which the highest abundance has been shown in the Ethiopians, compared to the remaining groups [Figs 8B and 9B]. These informative ethnic-associated variations in the free sialic acid residue further strengthen our MALDI-TOF/MS-based quantification results that demonstrated a consistent declining trend in the expression levels of sialylated glyco-subclasses among Indians, Japanese, and US controls, respectively. The non-detection of Neu5Gc in the current result agrees with the fact that humans do not naturally produce it because of the species-specific embryonic inactivating mutation of the gene encoding for CMP-Neu5Ac hydroxylase enzyme that converts Neu5Ac to Neu5Gc [35]. In addition, it is clear that there is little influence of exogenously incorporated Neu5Gc to the *N*-glycan biosynthesis of major serum glycoproteins. Neu5Ac, being a chief contributor of the anion layer of cellular surfaces in human, greatly modulates cell to cell repulsion, ligand-receptor interaction, immunogenicity, half-life of circulatory proteins, glomerular filtration, neural plasticity and cognitive development [48–50].

Given that more than half of human proteins are glycosylated [51], considerable interest still exists in identifying the specific carrier proteins to which those glycans attach. Bi-antennary serum *N*-glycans are reported to be carried mainly by IgG, a major serum glycoprotein and an essential part of the immune system, whose structural stability, binding and effector functions are greatly influenced by the type of *N*-glycan attached [52]. Recent inter- and intra-population studies on IgG *N*-glycome profile have increasingly evidenced the association between alteration in IgG *N*-glycome profile and hypertension [20–22]. Particularly, ethnic-based differences in IgG *N*-glycome have been observed with significantly reduced galactosylation and sialylation features in European hypertensive subjects (but non-significantly in Chinese cases), comparing to their healthy counterparts [20]. In another IgG subclass-specific *N*-glycomics study, Liu JN et al consistently found a marked decrease in galactosylation of IgG1, IgG2, and IgG4, as well as sialylation of IgG1 and IgG2 among northwestern Chinese hypertensive individuals of four different ethnic categories [22]. Further association of IgG glycosylation alterations (loss of galactose and sialic acid, along with addition of bisecting GlcNAc) with blood lipid profile was proposed to cause dyslipidaemia [24], whereas more core fucose and bisecting GlcNAc structures were found to be strongly associated with atherosclerotic plaque [23]. Apart from the role of IgG glycovariants in switching on and off the pro- and anti-inflammatory functions of IgG (and hence its contribution in disease pathogenesis), these reports are suggestive for the possibility that individual variation in IgG *N*-glycan profile may influence the extent of susceptibility to the conditions of diseases. The hyperbranched and hypersialylated glycans most of which have shown ethnic-based differences in the present study are possibly originated from Alpha-1-acid glycoprotein (AGP). It is one of the heavily *N*-glycosylated serum proteins carrying mainly high molecular weight glycans [53] among of which our group has recently succeeded in developing a focused glycoproteomics strategy to directly quantify serum level of AGP carrying a tri-antennary glycoform in multiple cancer types [54].

The Versatility of our comprehensive glycoblotting method is evident because many glycoforms have reproducibly been profiled from diverse biological samples including serum, cell lines, tissues, and cerebrospinal fluid [11–13, 29–32, 46, 47, 55, 56]. Importantly, unlike several prior reports that had measured relative abundances, our systematic strategy could concurrently quantify the absolute concentrations of whole serum glycome and their free sialic acid terminals only from 10  $\mu$ L of serum aliquot.

In conclusion, our inter-ethnic group glycomics result strongly revealed noticeable variations among the ethnic populations with high mannose, core-fucosylated, multi-antennary and multi-sialylated glycans, as well as the predominant sialic acid (Neu5Ac) demonstrated highest abundance in Ethiopians. The result further indicated some of the glycans that have shown profound expression alteration may not be useful candidates to be biomarkers of various diseases due to their large inter-ethnic and inter-individual variation. Despite the general scope of the present study, we were able to obtain interesting and informative results on the associations between ethnic difference and distinct changes in protein glycosylation which may become helpful for further in-depth investigations in the area. Due to limited samples to comprehensively address the gender and age effects, the current results are preliminary, and thus cannot be generalized to the target populations. In a large-scale study employing these and other ethnic compositions, we further need to investigate the correlation between human glycoforms and various confounding factors including gender and age. Establishing database for healthy human glycome variations among multi-ethnic populations is important as it further improves and accelerates the clinical utility of glycomics and glycoproteomics fields.

## Supporting information

**S1 Fig. Comparative serum *N*-glycome spectra of age-matched females varying in ethnicity.** In attempt to provide more clarity on whether difference in ethnicity or gender had marked effect on the *N*-glycan profile, we have included one Indian female (age = 43) and one Japanese female (age = 40s), whose serum *N*-glycome spectra is comparatively presented with that of one Ethiopian female (age = 40). Some of the *N*-glycan peaks showing marked variations among the three female subjects are highlighted by the dotted line shape. Considering this result from gender and age matched samples, most serum glycoforms showed abundant peak intensity in the Ethiopian subject, while one triantennary trisialylated glycan ( $m/z$  3049) demonstrated highest intensity in the Indian sample. These results from few female subjects intensify the variations observed in serum *N*-glycan profile result when Indian, Japanese, and US male subjects were considered as well, evidencing the profound influence of ethnicity on the *N*-glycosylation signature of the study groups.  
(TIF)

**S2 Fig. Serum *N*-glycome spectra showing Indian and Japanese of both sexes.** Age of subjects: Indian female = 43, Indian male = 39, Japanese female = 40s, Japanese male = 30s. Variation in *N*-glycan peak intensity is clearly shown in an ethnicity dependent manner in which glycans such as the core-fucosylated ( $m/z$  1591 and 1753) illustrated abundant expression in the Japanese subjects, whereas hyperbranched and hypersialylated glycans ( $m/z$  3049 and 3195) demonstrated up-regulated expression in the Indian subjects, irrespective of the gender difference.  
(TIF)

**S3 Fig. HPLC chromatogram of free Neu5Ac captured from serum of three age-matched females varying in ethnicity.** The female subjects were each of Ethiopian (age = 40), Indian (age = 43), and Japanese (age = 40s). This result from age- and gender-matched subjects provides an evidence for the strong influence of ethnic deference on the sialylation pattern of human serum glycoproteins and strengthens our total sialylated *N*-glycan result that demonstrated ethnic-associated variation in detection profile.  
(TIF)

**S1 Table. Demographic characteristics of study participants.** While recruiting the Japanese subjects, age of each subject was not recorded in its exact value, rather as 20s (within 20–29.9 y/o), 30s (within 30–39.9 y/o), 50s (within 50–59.9 y/o), 60s (within 60–69.9 y/o). The US origin control serum is a pool collected from several male donors whose age information was not provided as it was purchased sample from Sigma-Aldrich company, product # H4522.  
(DOCX)

**S2 Table. List of *N*-glycans considered for glyco-subclass analysis.** Only glycans that were detected in all the ethnic groups have been considered for the glyco-subclass analysis. There is a chance that a glycan can be counted in more than one group when it contains more than one structural features as per the grouping mechanism.  $m/z$  values are given as label with each glycan structure.  
(DOCX)

## Acknowledgments

We would like to acknowledge Mr. Masakazu Tanaka and Mr. Tetsu Ohashi for their kind help during data processing and evaluation. We also thank the study participants for the generous donation of their time.

## Author Contributions

**Conceptualization:** Shin-Ichiro Nishimura.

**Data curation:** Daniel Seifu Melka, Yimenashu Mamo Kassaye, Ibrahim F. Rehan, Shobith Rangappa, Hiroshi Hinou, Toshiya Kamiyama.

**Formal analysis:** Abrha G. Gebrehiwot, Ibrahim F. Rehan.

**Funding acquisition:** Shin-Ichiro Nishimura.

**Investigation:** Abrha G. Gebrehiwot, Daniel Seifu Melka, Yimenashu Mamo Kassaye, Shobith Rangappa, Hiroshi Hinou, Toshiya Kamiyama.

**Methodology:** Shin-Ichiro Nishimura.

**Project administration:** Shin-Ichiro Nishimura.

**Resources:** Daniel Seifu Melka, Yimenashu Mamo Kassaye, Toshiya Kamiyama, Shin-Ichiro Nishimura.

**Supervision:** Shin-Ichiro Nishimura.

**Validation:** Toshiya Kamiyama.

**Writing – original draft:** Abrha G. Gebrehiwot.

**Writing – review & editing:** Shin-Ichiro Nishimura.

## References

1. Ohtsubo K, Marth JD. Glycosylation in cellular mechanisms of health and disease. *Cell* 2006 Sep 8; 126(5):855–67. <https://doi.org/10.1016/j.cell.2006.08.019> PMID: 16959566.
2. Pilobello KT, Mahal LK. Deciphering the glycode: the complexity and analytical challenge of glycomics. *Curr Opin Chem Biol.* 2007 Jun; 11(3):300–5. <https://doi.org/10.1016/j.cbpa.2007.05.002> PMID: 17500024.
3. Contessa JN, Bhojani MS, Freeze HH, Rehemtulla A, Lawrence TS. Inhibition of *N*-linked glycosylation disrupts receptor tyrosine kinase signaling in tumor cells. *Cancer Res.* 2008 May 15; 68(10):3803–9. <https://doi.org/10.1158/0008-5472.CAN-07-6389> PMID: 18483264.
4. Ajit V. Evolutionary Forces Shaping the Golgi Glycosylation Machinery: Why Cell Surface Glycans Are Universal to Living Cells. *Cold Spring Harb Perspect Biol.* 2011 Jun; 3(6):a005462. <https://doi.org/10.1101/cshperspect.a005462> PMID: 21525513.
5. Ajit V, John BL. Biological roles of glycans: Essentials of Glycobiology. 2<sup>nd</sup> ed. Cold Spring Harbor Laboratory Press; 2009.
6. Lauc G, Krištić J, Zoldoš V. Glycans—the third revolution in evolution. *Front Genet.* 2014 May 23; 5:145. <https://doi.org/10.3389/fgene.2014.00145> PMID: 24904645.
7. Marek KW, Vijay IK, Marth JD. A recessive deletion in the GlcNAc-1-phosphotransferase gene results in peri-implantation embryonic lethality. *Glycobiology* 1999 Nov; 9(11):1263–71. <https://doi.org/10.1093/glycob/9.11.1263> PMID: 10536042.
8. Clerc F, Reiding KR, Jansen BC, Kammeijer GS, Bondt A, Wuhler M. Human plasma protein N-glycosylation. *Glycoconj J.* 2016 Jun; 33(3):309–43. <https://doi.org/10.1007/s10719-015-9626-2> PMID: 26555091.
9. Hart GW, Copeland RJ. Glycomics Hits the Big Time. *Cell* 2010 Nov 24; 143(5):672–6. <https://doi.org/10.1016/j.cell.2010.11.008> PMID: 21111227.
10. Cummings RD, Pierce JM. The Challenge and Promise of Glycomics. *Chem Biol.* 2014 Jan 16; 21(1):1–15. <https://doi.org/10.1016/j.chembiol.2013.12.010> PMID: 24439204.
11. Nishimura S-I., Niikura K, Kuroguchi M, Matsushita T, Fumoto M, Hinou H, et al. High-throughput protein glycomics: combined use of chemoselective glycoblotting and MALDI-TOF/TOF mass spectrometry. *Angew Chem Int Ed Engl.* 2004 Dec 17; 44(1):91–6. <https://doi.org/10.1002/anie.200461685> PMID: 15599899.
12. Nishimura S-I. Toward automated glycan analysis. *Adv Carbohydr Chem Biochem.* 2011; 65:219–71. <https://doi.org/10.1016/B978-0-12-385520-6.00005-4> PMID: 21763513.

13. Gizaw ST, Koda T, Amano M, Kamimura K, Ohashi T, Hinou H, Nishimura S-I. A comprehensive glycome profiling of Huntington's disease transgenic mice. *Biochim Biophys Acta* 2015 Sep; 1850(9):1704–18. <https://doi.org/10.1016/j.bbagen.2015.04.006> PMID: 25907331.
14. Phuong LM, Nicolas W, Mark HG. Challenges Related to Developing Serum-based Biomarkers for Early Ovarian Cancer Detection. *Cancer Prev Res (Phila)*. 2011 Mar; 4(3): 303–306. <https://doi.org/10.1158/1940-6207.CAPR-11-0053> PMCID: PMC3077065. PMID: 21372029
15. Eleftherios PD. The failure of protein cancer biomarkers to reach the clinic: why, and what can be done to address the problem? *BMC Med*. 2012; 10:87. <https://doi.org/10.1186/1741-7015-10-87>.
16. Duffy MJ. Serum tumor markers in breast cancer: are they of clinical value? *Clin Chem*. 2006 Mar; 52(3):345–51. <https://doi.org/10.1373/clinchem.2005.059832> PMID: 16410341.
17. Kirwan A, Utratna M, O'Dwyer ME, Joshi L, Kilcoyne M. Glycosylation-Based Serum Biomarkers for Cancer Diagnostics and Prognostics. *Biomed Res Int*. 2015; 2015:490531. <http://dx.doi.org/10.1155/2015/490531> PMID: 26509158.
18. Lu JP, Knežević A, Wang YX, Rudan I, Campbell H, Zou ZK, et al. Screening novel biomarkers for metabolic syndrome by profiling human plasma N-glycans in Chinese Han and Croatian populations. *J Proteome Res*. 2011 Nov 4; 10(11):4959–69. <https://doi.org/10.1021/pr2004067> PMID: 21939225.
19. Keser T, Gornik I, Vučković F, Selak N, Pavić T, Lukić E, Gudelj I, et al. Increased plasma N-glycome complexity is associated with higher risk of type 2 diabetes. *Diabetologia* 2017 Dec; 60(12):2352–2360. <https://doi.org/10.1007/s00125-017-4426-9> PMID: 28905229.
20. Wang Y, Klarić L, Yu X, Thaqi K, Dong J, Novokmet M, et al. The Association Between Glycosylation of Immunoglobulin G and Hypertension: A Multiple Ethnic Cross-Sectional Study. *Medicine (Baltimore)*. 2016 Apr; 95(17):e3379. <https://doi.org/10.1097/MD.0000000000003379> PMCID: PMC4998686. PMID: 27124023
21. Gao Q, Dolikun M, Štambuk J, Wang H, Zhao F, Yiliham N, et al. Immunoglobulin G N-Glycans as Potential Postgenomic Biomarkers for Hypertension in the Kazakh Population. *OMICS*. 2017 Jul; 21(7):380–389. <https://doi.org/10.1089/omi.2017.0044> PMID: 28692416.
22. Liu JN, Dolikun M, Štambuk J, Trbojević-Akmačić I, Zhang J, Wang H, et al. The association between subclass-specific IgG Fc N-glycosylation profiles and hypertension in the Uygur, Kazak, Kirgiz, and Tajik populations. *J Hum Hypertens*. 2018 Sep; 32(8–9):555–563. <https://doi.org/10.1038/s41371-018-0071-0> PMID: 29867134.
23. Menni C, Gudelj I, Macdonald-Dunlop E, Mangino M, Zierer J, Bešić E, et al. Glycosylation Profile of Immunoglobulin G Is Cross-Sectionally Associated With Cardiovascular Disease Risk Score and Subclinical Atherosclerosis in Two Independent Cohorts. *Circ Res*. 2018 May 25; 122(11):1555–1564. <https://doi.org/10.1161/CIRCRESAHA.117.312174> PMCID: PMC5970566. PMID: 29535164
24. Liu D, Chu X, Wang H, Dong J, Ge SQ, Zhao ZY, et al. The changes of immunoglobulin G N-glycosylation in blood lipids and dyslipidaemia. *J Transl Med*. 2018 Aug 29; 16(1):235. <https://doi.org/10.1186/s12967-018-1616-2> PMCID: PMC6114873. PMID: 30157878
25. Eleftherios PD. Cancer Biomarkers: Can We Turn Recent Failures into Success? *J Natl Cancer Inst*. 2010 Oct 6; 102(19):1462–1467. <https://doi.org/10.1093/jnci/djq306> PMID: 20705936.
26. Knezevic A, Gornik O, Polasek O, Pucic M, Redzic I, Novokmet M, et al. Effects of aging, body mass index, plasma lipid profiles, and smoking on human plasma N-glycans. *Glycobiology* 2010 Aug; 20(8):959–69. <https://doi.org/10.1093/glycob/cwq051> PMID: 20356825.
27. Ding N, Nie H, Sun X, Sun W, Qu Y, Liu X, et al. Human serum N-glycan profiles are age and sex dependent. *Age Ageing*. 2011 Sep; 40(5):568–75. <https://doi.org/10.1093/ageing/afr084> PMID: 21807702.
28. Yu X, Wang Y, Kristic J, Dong J, Chu X, Ge S, et al. Profiling IgG N-glycans as potential biomarker of chronological and biological ages: A community-based study in a Han Chinese population. *Medicine (Baltimore)* 2016 Jul; 95(28):e4112. <https://doi.org/10.1097/MD.0000000000004112> PMCID: PMID: PMC4956791.
29. Kamiyama T, Yokoo H, Furukawa J, Kuroguchi M, Togashi T, Miura N, et al. Identification of novel serum biomarkers of hepatocellular carcinoma using glycomic analysis. *Hepatology* 2013 Jun; 57(6):2314–25. <https://doi.org/10.1002/hep.26262> PMID: 23322672.
30. Inafuku S, Noda K, Amano M, Ohashi T, Yoshizawa C, Saito W, et al. Alteration of N-Glycan Profiles in Diabetic Retinopathy. *Invest Ophthalmol Vis Sci*. 2015 Aug; 56(9):5316–22. <https://doi.org/10.1167/iov.15-16747> PMID: 26258617.
31. Kita Y, Miura Y, Furukawa J, Nakano M, Shinohara Y, Ohno M, et al. Quantitative glycomics of human whole serum glycoprotein based on the standardized protocol for liberating N-glycans. *Mol Cell Proteomics* 2007 Aug; 6(8):1437–45. <https://doi.org/10.1074/mcp.T600063-MCP200> PMID: 17522412.
32. Furukawa J, Shinohara Y, Kuramoto H, Miura Y, Shimaoka H, Kuroguchi M, et al. Comprehensive approach to structural and functional glycomics based on chemoselective glycoblotting and sequential

- tag conversion. *Anal Chem.* 2008 Feb 15; 80(4):1094–101. <https://doi.org/10.1021/ac702124d> PMID: 18205388.
33. Hirose K, Amano M, Hashimoto R, Lee YC, Nishimura S-I. Insight into glycan diversity and evolutionary lineage based on comparative avio-N-glycomics and sialic acid analysis of 88 egg whites of Gallonaseriae. *Biochemistry* 2011 May 31; 50(21):4757–74. <https://doi.org/10.1021/bi101940x> PMID: 21495680.
  34. Rehan IF, Ueda K, Mitani T, Amano M, Hinou H, Ohashi, et al. Large-Scale Glycomics of Livestock: Discovery of Highly Sensitive Serum Biomarkers Indicating an Environmental Stress Affecting Immune Responses and Productivity of Holstein Dairy Cows. *J. Agric. Food Chem.* 2015; 63(48):10578–10590. <https://doi.org/10.1021/acs.jafc.5b04304> PMID: 26595672.
  35. Chou HH, Takematsu H, Diaz S, Iber J, Nickerson E, Wright KL, et al. A mutation in human CMP-sialic acid hydroxylase occurred after the *Homo-Pan* divergence. *Proc Natl Acad Sci U S A.* 1998 Sep 29; 95(20):11751–6. PMID: 9751737.
  36. Pinho SS, and Reis CA. Glycosylation in cancer: mechanisms and clinical implications. *Nat Rev Cancer* 2015 Sep; 15(9):540–55. <https://doi.org/10.1038/nrc3982> PMID: 26289314.
  37. Clerc F, Reiding KR, Jansen BC, Kammeijer GS, Bondt A, Wuhrer M. Human plasma protein N-glycosylation. *Glycoconj J.* 2016 Jun; 33(3):309–43. <https://doi.org/10.1007/s10719-015-9626-2> PMID: 26555091.
  38. Anderson NL, Anderson NG. The human plasma proteome: history, character, and diagnostic prospects. *Mol Cell Proteomics.* *Mol Cell Proteomics* 2002 Nov; 1(11):845–67. <https://doi.org/10.1074/mcp.R200007-MCP200> PMID: 12488461.
  39. Knezević A1, Polasek O, Gornik O, Rudan I, Campbell H, Hayward C, et al. Variability, heritability and environmental determinants of human plasma N-glycome. *J Proteome Res.* 2009 Feb; 8(2):694–701. <https://doi.org/10.1021/pr800737u> PMID: 19035662.
  40. Pucić M, Knezević A, Vidic J, Adamczyk B, Novokmet M, Polasek O, et al. High throughput isolation and glycosylation analysis of IgG-variability and heritability of the IgG glycome in three isolated human populations. *Mol Cell Proteomics* 2011 Oct; 10(10):M111.010090. <https://doi.org/10.1074/mcp.M111.010090> PMID: 21653738.
  41. Bieberich E. Synthesis, processing, and function of N-glycans in N-glycoproteins. *Adv Neurobiol.* 2014; 9:47–70. [https://doi.org/10.1007/978-1-4939-1154-7\\_3](https://doi.org/10.1007/978-1-4939-1154-7_3) PMID: 25151374.
  42. Menni C, Keser T, Mangino M, Bell JT, Erte I, Akmačić I, et al. Glycosylation of Immunoglobulin G: Role of Genetic and Epigenetic Influences. *PLoS One.* 2013 Dec 6; 8(12):e82558. <https://doi.org/10.1371/journal.pone.0082558> PMCID: PMC3855797. PMID: 24324808
  43. Nairn AV, York WS, Harris K, Hall EM, Pierce JM, Moremen KW. Regulation of glycan structures in animal tissues: transcript profiling of glycan-related genes. *J Biol Chem.* 2008 Jun 20; 283(25):17298–313. <https://doi.org/10.1074/jbc.M801964200> PMID: 18411279.
  44. Tahmasebi H, Trajcevski K, Higgins V, Adeli K. Influence of ethnicity on population reference values for biochemical markers. *Crit Rev Clin Lab Sci.* 2018; 55(5):359–375. <https://doi.org/10.1080/10408363.2018.1476455> PMID: 29874957
  45. Ellman N, Keswell D, Collins M, Tootla M, Goedecke JH. Ethnic differences in the association between lipid metabolism genes and lipid levels in black and white South African women. *Atherosclerosis* 2015 Jun; 240(2): 311–7. <https://doi.org/10.1016/j.atherosclerosis.2015.03.027> PMID: 25864161.
  46. Hatakeyama S, Amano M, Tobisawa Y, Yoneyama T, Tsuchiya N, Habuchi T, et al. Serum N-glycan alteration associated with renal cell carcinoma detected by high-throughput glycan analysis. *J Urol.* 2014 Mar; 191(3):805–13. <https://doi.org/10.1016/j.juro.2013.10.052> PMID: 24140550.
  47. Ishibashi Y, Tobisawa Y, Hatakeyama S, Ohashi T, Tanaka M, Narita S, et al. Serum Tri- and tetra-antennary N-glycan is a potential predictive biomarker for castration-resistant prostate cancer. *Prostate* 2014 Nov; 74(15):1521–9. <https://doi.org/10.1002/pros.22869> PMID: 25154914.
  48. Wang B. Molecular Mechanism Underlying Sialic Acid as an Essential Nutrient for Brain Development and Cognition. *Adv Nutr.* 2012 May 1; 3(3):465S–72S. <https://doi.org/10.3945/an.112.001875> PMID: 22585926
  49. Varki A. Sialic acids in human health and disease. *Trends Mol Med.* 2008 Aug; 14(8):351–60. <https://doi.org/10.1016/j.molmed.2008.06.002> PMID: 18606570.
  50. Varki A, Gagneux P. Multifarious roles of sialic acids in immunity. *Ann N Y Acad Sci.* 2012 Apr; 1253:16–36. <https://doi.org/10.1111/j.1749-6632.2012.06517.x> PMID: 22524423.
  51. Apweiler R, Hermjakob H, Sharon N. On the frequency of protein glycosylation, as deduced from analysis of the SWISS-PROT database. *Biochim Biophys Acta* 1999 Dec 6; 1473(1):4–8. [https://doi.org/10.1016/S0304-4165\(99\)00165-8](https://doi.org/10.1016/S0304-4165(99)00165-8) PMID: 10580125.

52. Arnold JN, Wormald MR, Sim RB, Rudd PM, Dwek RA. The Impact of Glycosylation on the Biological Function and Structure of Human Immunoglobulins. *Annu Rev Immunol.* 2007; 25:21–50. <https://doi.org/10.1146/annurev.immunol.25.022106.141702> PMID: 17029568.
53. Kevin DS, Jennifer B, Gerardine MS, Anthony MM. Alpha-1-Acid Glycoprotein (AGP) as a Potential Biomarker for Breast Cancer. InTechOpen 2012. <http://dx.doi.org/10.5772/48177>.
54. Yogesh KV, Kamiyama T, Ohyama C, Yoneyama T, Nouse K Kimura S. Synthetic glycopeptides as a designated standard in focused glycoproteomics to discover serum cancer biomarkers. *Med. Chem. Commun.* 2018; 9:1351–1358. <https://doi.org/10.1039/C8MD00162F> PMID: 30151090
55. Terashima M, Amano M, Onodera T, Shin-Ichiro Nishimura S-I, Iwasaki N. Quantitative glycomics monitoring of induced pluripotent- and embryonic stem cells during neuronal differentiation. *Stem Cell Res.* 2014 Nov; 13(3):454–64. <https://doi.org/10.1016/j.scr.2014.10.006> PMID: 23500645.
56. Gizaw ST, Ohashi T, Tanaka M, Hinou H, Nishimura S. Glycoblotting method allows for rapid and efficient glycome profiling of human Alzheimer's disease brain, serum and cerebrospinal fluid towards potential biomarker discovery. *Biochim Biophys Acta* 2016 Aug; 1860(8):1716–27. <https://doi.org/10.1016/j.bbagen.2016.03.009> PMID: 26968461.