Glycobiology vol. 20 no. 8 pp. 970–975, 2010
doi:10.1093/glycob/cwq052
Advance Access publication on April 8, 2010

Common aberrations from the normal human plasma *N*-glycan profile

Maja Pučić[2], Sofia Pinto[3], Mislav Novokmet[2], Ana Knežević[4], Olga Gornik[4], Ozren Polašek[5], Kristian Vlahoviček[3], Wei Wang[6], Pauline M Rudd[7], Alan F Wright[8], Harry Campbell[9], Igor Rudan[9,10,11], and Gordan Lauc[1,2,4]

[2]Genos Ltd., Glycobiology Division, Planinska 1, 10000 Zagreb, Croatia, [3]University of Zagreb, Faculty of Science, Division of Biology, Bioinformatics Group, Horvatovac 102, Zagreb, Croatia, [4]University of Zagreb, Faculty of Pharmacy and Biochemistry, Ante Kovačića 1, 10000 Zagreb, Croatia, [5]Andrija Štampar School of Public Health, University of Zagreb Medical School, Zagreb, Croatia, [6]School of Public Health and Family Medicine, Capital Medical University, Beijing, China, [7]National Institute for Bioprocessing Research and Training, Dublin-Oxford Glycobiology Lab., Conway Institute, University College Dublin, Belfield, Dublin 4, Ireland, [8]MRC Human Genetics Unit, Institute of Genetics and Molecular Medicine, Western General Hospital, Crewe Road, Edinburgh, UK, [9]Department of Public Health Sciences, The University of Edinburgh Medical School, Teviot Place, Edinburgh, UK, [10]Croatian Centre for Global Health, University of Split Medical School, Split, Croatia, and [11]Gen-info Ltd, Zagreb, Croatia

Received on February 6, 2010; revised on March 11, 2010; accepted on March 24, 2010

After performing hydrophilic interaction and weak anion exchange high-performance liquid chromatography to analyze *N*-glycans in the plasma of 1991 people, we identified several individuals that differed significantly from the "normal" profile of *N*-glycans. By performing consensus scoring of pairwise distances between vectors containing measured glycan values, we formed six groups of individuals with specific glyco-phenotypes. Some aberrations from the normal plasma protein patterns were found to be associated with clinical conditions (such as renal problems in people with increased monosialylated biantennary glycans, A2G2S1), while other substantial changes in *N*-glycan structure, such as the near complete absence of neutral glycans or antennary fucosylated tri- and tetraantennary glycans, were not associated with any observed adverse health outcomes. These results demonstrate the existence of specific altered glyco-phenotypes in some individuals and indicate that in some cases they might represent risk factors for the development of specific diseases.

Keywords: glycan analysis/human plasma glycome/*N*-glycans/protein glycosylation

Introduction

In contrast to RNA and proteins, glycans are being synthesized without a genetic template that pre-determines their final structure. Instead of being molded by a single gene, glycans are a product of complex interactions between hundreds of different genes. Therefore, the exact structure of glycans on a glycoprotein is determined by both genetic polymorphisms in the genome and the past environmental influences on the cell producing the given glycoprotein (Lauc et al. 2010). Consequently, the variability of glycans both between species and within a species is larger than the variability of proteins or RNA molecules (Royle et al. 2008; Knežević et al. 2009). N-glycosylation is essential for multicellular life, and its complete absence is embryonically lethal (Marek et al. 1999). In contrast to core *N*-glycans which are essential for functions of many glycoproteins, variability in monosaccharides at the end of glycan antennae is common (e.g., ABO blood groups). This structural diversity of glycans contributes to glycoproteome heterogeneity in a population that can be advantageous for evading pathogens and adapting to changing environment (Varki 1993; Lauc and Zoldos 2009).

In general, the human plasma glycome is very stable and changes very little over time (Gornik et al. 2009). In a majority of people, this glycan profile is very similar and can be referred to as "the normal profile" (Knežević et al. 2009). However, sometimes a combination of genetic polymorphisms or environmental conditions will occur that will result in a significantly different profile that might be a predisposition for a specific disease or could also be a consequence of the same disease. After analyzing *N*-glycans in plasma of 1991 individuals from Croatian islands Vis and Korčula, here we describe several groups of outliers with similar glycan profiles that differ from the general population in both glycan profiles and (patho)physiological characteristics.

Results

By combining hydrophilic interaction and weak anion exchange high-performance liquid chromatography (HPLC) analysis of 2AB-labeled released *N*-glycans, we determined basic characteristics of the plasma *N*-glycome in 1991 individuals. Glycan profiles in the majority of the analyzed individuals were rather similar, but in some individuals significantly different glycan profiles were observed (Figure 1). When several individuals shared the same "outlying" profile, they were treated as a group. If only one individual was signif-

[1]To whom correspondence should be addressed: Fax: +385 1 639 4400; e-mail: glauc@pharma.hr

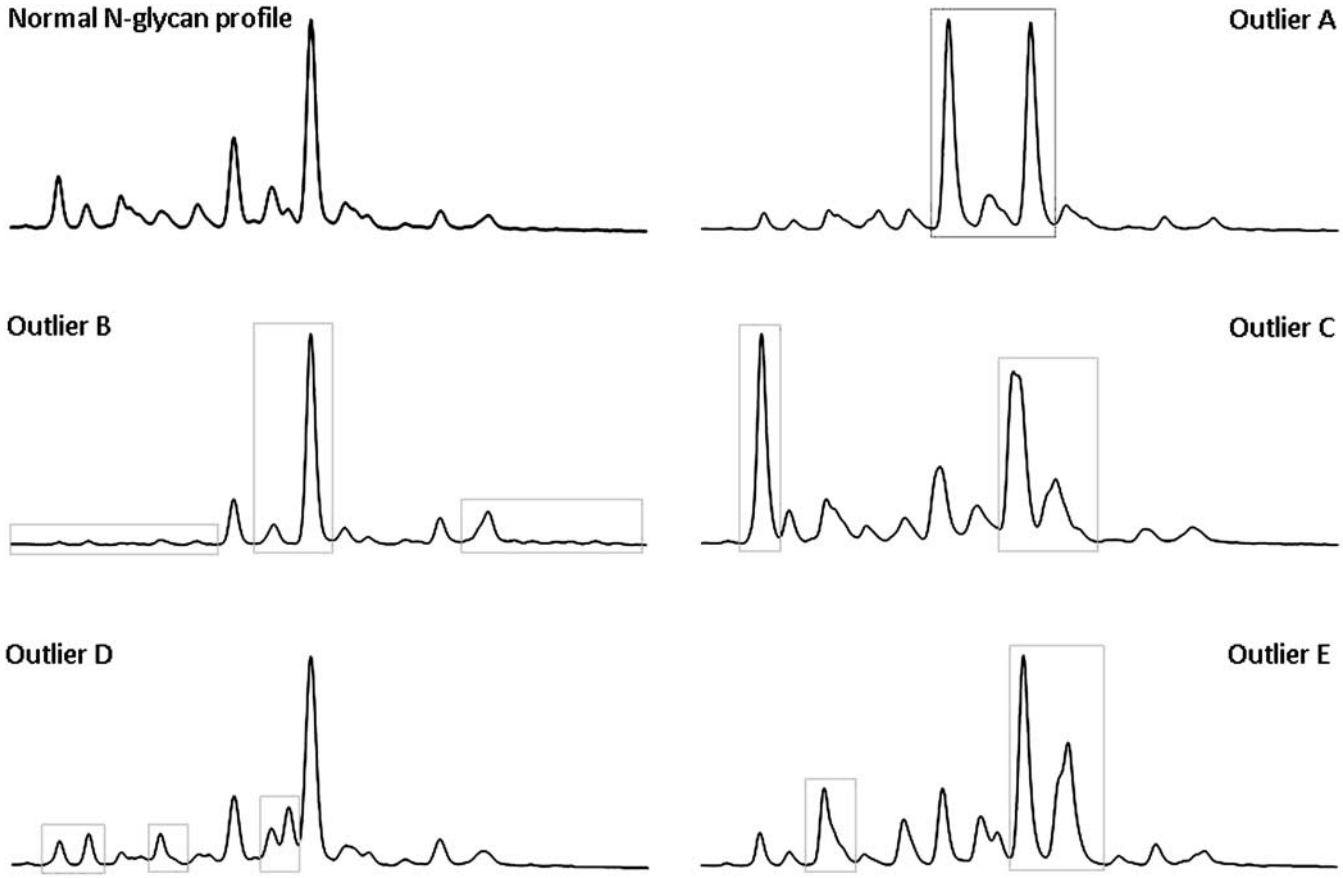

Fig. 1. Outlier A, individual with elevated A2G2S1 glycans; outlier B, individual with glycan changes that mirror premature aging; outlier C, individual with elevated biantennary nongalactosylated glycans; outlier D, individual with elevated biantennary monosialylated glycans; outlier E, individual with increased core fucosylated glycans.

icantly different, then its "first neighbors" were identified using consensus scoring of pairwise distances between vectors containing measured glycan values as described in the Materials and methods section. Comprehensive statistical analysis of over 100 biochemical traits and other medical data was performed to identify shared phenotypic characteristics within each outlier group.

Individuals lacking antennary fucose on triantennary and tetraantennary glycans

Four individuals who nearly completely lacked glycans in DG9 and DG12 glycan groups were observed (Figure 2). From the complete listing of structures in each glycan group presented in Supplementary Table 1, it is clearly visible that these individuals nearly completely lacked antennary fucosylated triantennary and tetraantennary glycans. At the same time, they all had increased levels of both triantennary and tetraantennary structures as well as some antennary fucosylated biantennary structures, indicating that all relevant enzymes were functional in these individuals. However, for some currently unknown reason, fucose was not being added to tri- and tetraantennary glycans. Interestingly, the same glycosylation phenotype was observed in three out of 300 analyzed individuals from China. These individuals were not characterized biochemically and are not included in this study, but the presence of the same aberrant glycosylation phenotype in both Croatia and China indicates that the mutation causing this phenotype is probably evolutionarily very old. The four individuals from Croatia were all females (51, 65, 67 and 74 years of age). Two of them were sisters, while the others were unrelated. Their biochemical status was characterized by increased levels of cholesterol and low-density lipoprotein (LDL) or increased level of insulin, while all of them had decreased levels of creatinine.

Individuals with elevated A2G2S1 glycans

A single individual with extremely high A2G2S1 glycan (the main component of peak 7, GP7 in hydrophilic interaction liquid chromatography (HILIC) analysis of plasma glycans) was observed in a population (Figure 1, outlier A). Five first neighbors with the most similar glycome profiles were identified (Figure 3, Supplementary Table 2). All these neighbors (30-, 49- and 58-year-old males and 58- and 76-year-old females) were characterized by an increase in the ratio of monosialylated to bisialylated biantennary glycans and a decrease in triantennary structures (complete listing of all glycan values for these and all other individuals is presented in Supplementary Table 3). As for their biochemical status, three had increased levels of cholesterol and LDL, and four had increased fibrinogen level. The most interesting piece of data

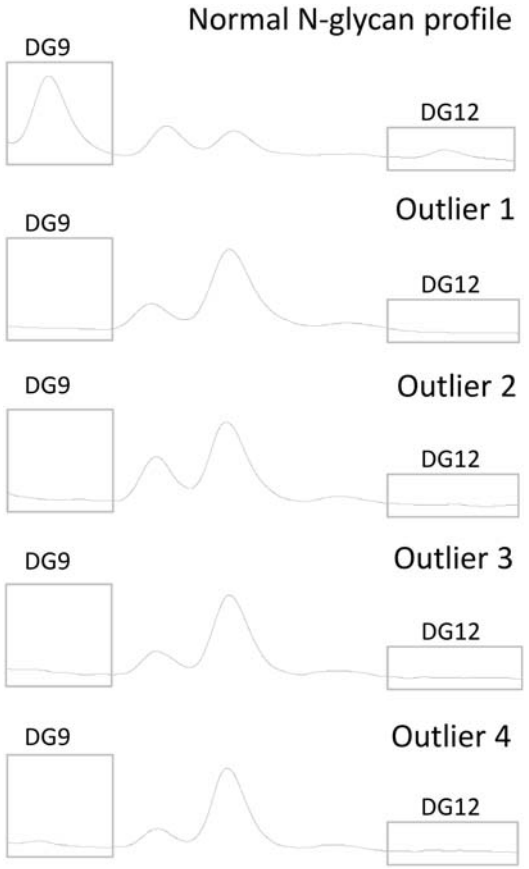

Fig. 2. Individuals lacking antennary fucose on triantennary and tetraantennary glycans. Section of HPLC chromatograms showing affected glycan peaks are showed.

in their medical records was the fact that 50% of these individuals had some type of renal disease, which represents a nearly 6-fold increase compared to a prevalence of 8.4% in the whole population. The association of A2G2S1 with renal problems was confirmed by the analysis of GP7 levels in the whole studied population. Only 4.3% of healthy individuals had GP7 values above the 95th percentile, 8.9% above the 90th percentile and 11% above the 75th percentile. In people with renal diseases, these values were approximately double; 9% of individuals had GP7 values above the 95th percentile, 17.6% above the 90th percentile and 28% above the 75th percentile.

Individuals with glycan changes that mirror premature aging

An individual that nearly completely lacked neutral glycans was identified (Figure 1, outlier B, 44-year-old female). The four first neighbors (67-year-old male and 57-, 58- and 62-year-old females) with the most similar glycome profiles were identified (Supplementary Table 2). These five individuals shared a pattern of decreased or increased levels of nine glycans that change with age: monosialylated, monosialylated biantennary, disialylated biantennary, trisialylated, biantennary, triantennary, core fucosylated, trigalactosylated and tetragalactosylated glycans (Knežević et al. 2010). In all five individuals, these glycans were either below the 5th percentile or above the 95th percentile of the general population. All five individuals had increased fibrinogen level, and three of them had decreased creatinine and albumin levels (Supplementary Table 4), but they did not share any peculiar medical status.

Individuals with elevated biantennary nongalactosylated glycans

A notable individual (Figure 1, outlier C, 75-year-old female) with extremely high levels of biantennary nongalactosylated glycans (GP2 and DG2 glycan groups) was identified. Four first neighbors (38-, 47- and 50-year-old males and 72-year-old female) with the most similar glycan profiles were identified (Figure 4, Supplementary Table 2). In addition to high biantennary nongalactosylated and monogalactosylated glycans, the common feature of this group was a decreased level of triantennary and tetraantennary glycans. Three of these five individuals had increased levels of fibrinogen and insulin. The original outlier (75-year-old female) and her most similar neighbor 1 (50-year-old male) had increased levels of glucose, insulin, HbA1c and fibrinogen. Both of them also suffered from hypertension (Supplementary Table 4). In the whole population, 7.3% of individuals had diabetes, among which 16.5% had levels of triantennary glycans below the 10th percentile.

Individuals with altered levels of different IgG glycans

Another individual identified as an outlier displayed complex changes in levels of glycans that mostly originate from IgG (Figure 1, outlier D, 81-year-old female). From the performed experiments, it is not possible to unequivocally determine exact structural changes in this individual, but it appears that the main feature was an increase in core-fucosylated biantennary glycans with bisecting GlcNAc. Four closest neighbors were identified (73-year-old male and 27-, 47- and 72-year-old females). A shared biochemical feature of this group was a high level of fibrinogen in all individuals.

Individuals with increased core fucosylated glycans

A female individual, 57 years old, who was an outlier with an extremely high level of core fucosylated glycans was identified (Figure 1, outlier E). Most of this core fucosylation was restricted to biantennary monogalactosylated or digalactosylated structures which were also showing distinct increase. Distance matrix calculation identified four additional individuals (57- and 65-year-old males and 45- and 78-year-old females), all of which had similar plasma glycan profiles. Three of these four individuals had increased levels of triglycerides and LDL, while their high-density lipoprotein and creatinine were decreased (Supplementary Table 4). There were no other shared characteristics, and they appeared to be healthy despite highly increased core-fucosylated glycans.

Discussion

After analyzing plasma *N*-glycans in 1991 individuals, we identified six groups of people who were significantly different from the rest of the studied population. One of the groups was identified manually, while the others were formed around outlying individuals by advanced computing algorithms. In the manually identified group, two individuals were siblings; all the other individuals were not related. Computational identifi-

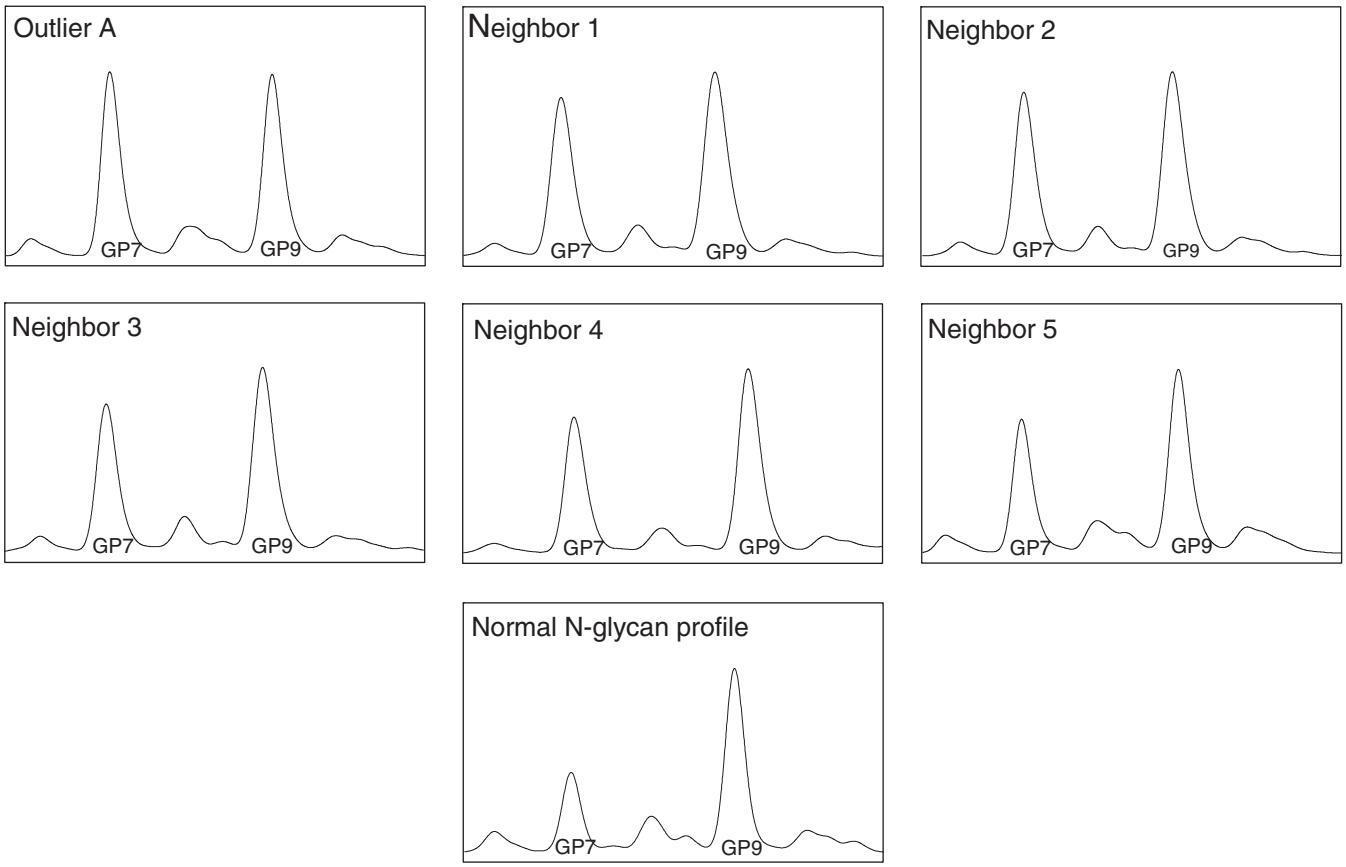

Fig. 3. An outlying individual with elevated A2G2S1 glycans was identified in a population. Further five individuals with similar glyco-phenotype (neighbors 1–5) were identified by consensus scoring of pairwise distances between vectors containing measured glycan values. Sections of HPLC chromatograms showing affected glycan peaks are showed.

cation of people with similar glycome characteristics using consensus scoring of pairwise distances between vectors containing measured glycan values was subsequently verified by manual scanning of the chromatograms. Glycan vectors used to identify first neighbors were adjusted for age and sex before computation, and the original chromatograms were not, but the successfulness of this approach is still clearly visible in Figures 3 and 4 which show similarities among the identified individuals.

The observed deviations from the normal plasma glycan profiles in these six groups of individuals were much more pronounced than changes that were reported to occur in common diseases (Alavi and Axford 2008; Gornik and Lauc 2008), and the incidence of these deviations in the studied population was much lower than the incidence of any common disease. These all indicate that the observed deviations were not a result of altered physiological conditions but were instead a consequence of some relatively rare mutations and/or rare combinations of common polymorphisms. Hundreds of different genes are involved in protein glycosylation, and numerous common polymorphisms in these genes are among the principal reasons for high variability of glycans. Contrary to proteins which are being made according to a specific genetic template, glycans are products of complex interactions between hundreds of proteins. Specific combination of several slightly altered protein functions could result in specific changes in the efficacy of the glycosylation machinery, which would be reflected in altered patterns of plasma glycans (glyco-phenotypes). Altered glycosylation of numerous proteins would result in a specific phenotype and could be associated with predispositions to some diseases.

Our study population was characterized in great detail; thus, we were able to identify some of the shared characteristics for different glyco-phenotypes. Some of the aberrations from the normal plasma proteins were associated with rather serious conditions like renal problems in people with increased A2G2S1 (Figure 1, outlier A). Some other changes, although apparently rather dramatic like the absence of neutral glycans (Figure 1, outlier B) or increased core-fucosylation (Figure 1, outlier E), were apparently healthy. Determining the exact physiological and/or medical consequences of specific aberrations from the normal glycan profiles would require more detailed examination of many additional individuals with the same aberration; thus, the results reported here should be considered as only an indication and not the definite proof of association between specific glyco-phenotypes and medical conditions. However, the results presented here represent a solid proof that specific common glyco-phenotypes do exist in the general population and that they might represent risk factors for the development of specific diseases. However, in some

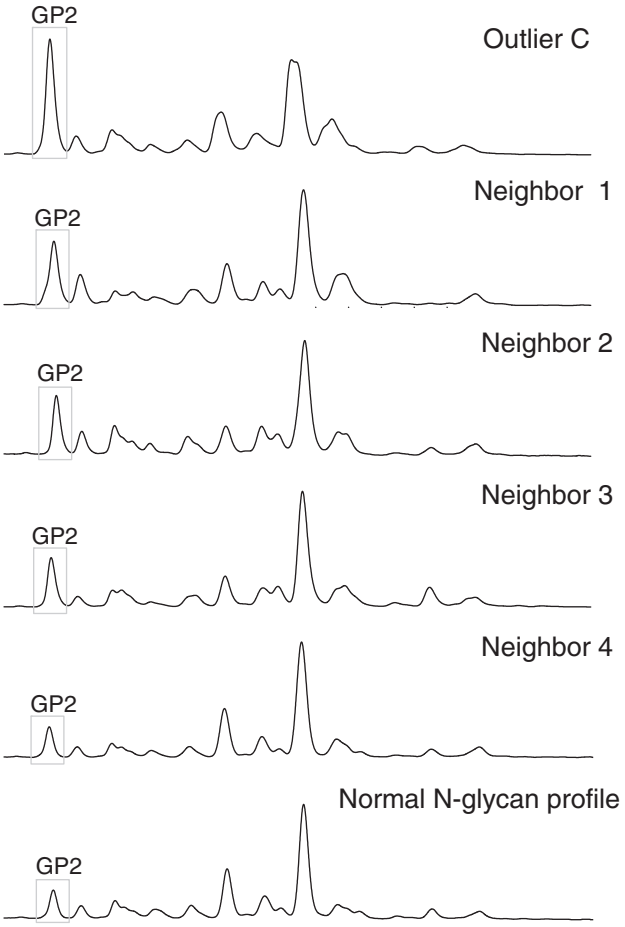

Fig. 4. An outlying individual with extremely high biantennary nongalactosylated glycans was observed. Further four individuals with similar glyco-phenotype (neighbors 1–4) were identified by consensus scoring of pairwise distances between vectors containing measured glycan values.

cases, even dramatic changes in the glyco-phenotype, like the near complete absence of neutral glycans, or antennary fucosylated tri- and tetraantennary glycans were apparently not detrimental to health.

Materials and methods

Plasma samples and glycan analysis

Blood samples were collected from unselected examinees from two Croatian Adriatic islands, Vis and Korčula, as a part of a larger genetic epidemiology program, and glycans were analyzed as described in detail in the accompanying article published in the same issue of *Glycobiology* (Knežević et al. 2010). Levels of glycans sharing the same structural features were approximated by adding the structures having the same characteristic, from either HILIC, weak anion exchange or after sialidase treatment-integrated glycan profiles (individual glycan structures present in each glycan group are listed in Supplementary Table 1).

Identification of outliers and their nearest neighbors

Individuals containing the most similar glycan profiles to identified outliers were determined using a consensus scoring of pairwise distances between vectors containing measured glycan values. Glycan profiles from both cohorts were normalized for age and gender differences and scaled to the mean residuals of linear regression. Five nearest neighbors (the individuals with the smallest respective profile distances) were calculated for each candidate using five distance calculation methods: maximum value (maximum difference in any coordinate dimension); Manhattan (city block); Euclidean (square root of the sum of squared vector coordinates); Canberra (sum of differences between the vector coordinates); and Minkowski generalized distance of order 4 (fourth root of the sum of vector coordinates raised to the fourth power). Neighbors occurring in a group of five nearest neighbors using at least two different methods were selected as true neighbors and are listed in Supplementary Table 2. Calculations were performed using the R package for statistical computing (http://www.r-project.org).

Supplementary data

Supplementary data for this article is available online at http://glycob.oxfordjournals.org/.

Funding

This work was supported by the Croatian Ministry of Science, Education and Sport grants [#309-0061194-2023] to GL [#216-1080315-0302], to IR and [#119-0982913-1211] to KV; by European Commission [EuroPharm, INTEGERS and EuroGlycoArrays] grants; by EMBO Young Investigator Program [Installation grant 1431/2006] and ICGEB collaborative research program grant CRP/CRO07-03. The Vis and Korčula studies in the Croatian islands of Vis and Korčula were supported through the grants from the Medical Research Council (UK) to IR, HC and AW.

Acknowledgements

The authors collectively thank a large number of individuals for their individual help in obtaining funding support, organizing, planning and carrying out the field work related to the project and for assistance with data management, analysis and interpretation: the teams from the Medical Research Council Human Genetics Unit in Edinburgh and The University of Edinburgh, UK, led by Professors Nicholas D. Hastie, Alan Wright and Harry Campbell, including Caroline Hayward and Veronique Vitart (for funding support, intellectual input in study design and their role in data management, analysis and interpretation); Professor Pavao Rudan and the staff of the Institute for Anthropological Research in Zagreb, Croatia (organization of the field work, anthropometric and physiological measurements and DNA extraction in Vis); Professor Ariana Vorko-Jović and the staff and medical students of the Andrija Štampar School of Public Health of the Medical School, University of Zagreb, Croatia (questionnaires, genealogical reconstruction and data entry); and the employees of several other Croatian institutions who participated in the field work, including but not limited to the University of Rijeka; Croatian Institute of Public Health; Institutes of Public Health in Split and Dubrovnik. KV and

SP acknowledge the help of Vedran Franke with statistical procedures in R.

Abbreviations

HILIC, hydrophilic interaction liquid chromatography; HPLC, high-performance liquid chromatography; LDL, low-density lipoprotein.

References

Alavi A, Axford JS. 2008. Sweet and sour: the impact of sugars on disease. *Rheumatology.* Oxford, 47:760–770.

Gornik O, Lauc G. 2008. Glycosylation of serum proteins in inflammatory diseases. *Dis Markers.* 25:267–278.

Gornik O, Wagner J, Pučić M, Knežević A, Redžić I, Lauc G. 2009. Stability of *N*-glycan profiles in human plasma. *Glycobiology.* 19:1547–1553.

Knežević A, Gornik O, Polašek O, Pučić M, Novokmet M, Redžić I, Rudd PM, Wright AF, Campbell H, Rudan I, Lauc G. 2010. Effects of aging, smoking and lifestyle on human plasma *N*-glycans. *Glycobiology.* in press.

Knežević A, Polašek O, Gornik O, Rudan I, Campbell H, Hayward C, Wright A, Kolčić I, O'Donoghue N, Bones J, et al. 2009. Variability, heritability and environmental determinants of human plasma *N*-glycome. *J Proteome Res.* 8:694–701.

Lauc G, Rudan I, Campbell H, Rudd PM. 2010. Complex genetic regulation of protein glycosylation. *Mol Biosyst.* 6:329–335.

Lauc G, Zoldos V. 2009. Epigenetic regulation of glycosylation could be a mechanism used by complex organisms to compete with microbes on an evolutionary scale. *Med Hypotheses.* 73:510–512.

Marek KW, Vijay IK, Marth JD. 1999. A recessive deletion in the GlcNAc-1-phosphotransferase gene results in peri-implantation embryonic lethality. *Glycobiology.* 9:1263–1271.

Royle L, Matthews E., Corfield A, Berry M, Rudd PM, Dwek RA, Carrington SD. 2008. Glycan structures of ocular surface mucins in man, rabbit and dog display species differences. *Glycoconjugate J.* 25:763–773.

Varki A. 1993. Biological roles of oligosaccharides: all of the theories are correct. *Glycobiology.* 3:97–130.