University of Zagreb

FACULTY OF SCIENCE
DIVISION OF BIOLOGY

Maja Pučić Baković

VARIABILITY AND HERITABILITY OF IMMUNOGLOBULIN G GLYCOSYLATION

DOCTORAL THESIS

Zagreb, 2013.

Sveučilište u Zagrebu

PRIRODOSLOVNO-MATEMATIČKI FAKULTET
BIOLOŠKI ODSJEK

Maja Pučić Baković

VARIJABILNOST I HERITABILNOST GLIKOZILACIJE IMUNOGLOBULINA G

DOKTORSKI RAD

Zagreb, 2013.

The work presented in this doctoral thesis was performed at Genos Ltd., Zagreb, Croatia under supervision of Prof. Dr. Gordan Lauc, Ph.D., as a part of postgraduate doctoral programme at Division of Biology, Faculty of Science, University of Zagreb. In collaboration with Prof. Dr. Djuro Josić, Ph.D., and Dr. Manfred Wuhrer, Ph.D., parts of this work were performed at Proteomics Core, COBRE Center for Cancer Research and Development, Rhode Island Hospital, Providence, USA and at Biomolecular Mass Spectrometry Unit, Department of Parasitology, Leiden University Medical Center, Leiden, Netherlands, respectively.

Zahvale/Acknowledgements

Hvala mom mentoru prof. dr. sc. Gordanu Laucu koji me uveo u glikobiologiju, omogućio mi rad i dao priliku da ostavim malen trag u svijetu znanosti, podržavao moju želju za upoznavanjem novih ljudi i drugih laboratorija, strpljivo trpio moj perfekcionizam i sklonost detaljima te me puštao da sama otkrivam i isprobavam nove stvari.

Hvala prof. dr. sc. Djuri Josiću što mi je dopustio da na tri mjeseca „okupiram" čitav Proteomics Core laboratorij, potrošim ogromne količine laboratorijske plastike i kemikalija te iskoristim i najmanji slobodan prostor u zamrzivačima s tisućama pripremljenih uzoraka. Hvala Vam na iskazanom povjerenju, prenesenom znanju i životnom iskustvu i ponajviše hvala na volji i želji da i dalje surađujemo. My American Dream would not be complete without kindness, friendship and support of prof. Hixson, Elizabeth, Kate, Martina, Marijana, Lulu, Feilei and many more dear people who shared this experience with me and never made me feel like I don't belong there.

Many thanks to dr. Manfred Wuhrer for letting me scratch the surface of MS analysis and for rarely just giving me the answer but instead leading me to the answer. Thank you for your tremendous patience, kindness and willingness to share your knowledge. Months spent in Leiden gave me a new perspective on glycan analysis and more importantly a chance to meet many wonderful people: Manu, Gerhild, Carolien, Maurice, Marcus, Niels, Ruben, Wesley, Crina, Rico, Angela, Nicole, Ron, Marco, Cees, Stephanie, Rosina, Agnes, Karli, Albert, Yoann, Dennis, Bas, Leonardo, Irina and those whose names I can't remember but were there to at least share a smile or goedemorgen.

Velika hvala Lidiji, Urški, Jani in celotnom Bia Separations ker so mi dali priložnost da bodem del zelo, zelo kul ekipe v Ajdovščini. Upam da se bomo še kdaj družili.

A debt of gratitude to all our other collaborators for making all of this possible and giving us a chance to explore and discover.

Beskrajno hvala Ani, Irmi, Olgi i Mislavu što su uvijek bili neiscrpan izvor podrške, pomoći, razumijevanja i smijeha. AMMIGO tim u svakom smislu.

Hvala Genos ekipi na zanimljivoj svakodnevici: najdražoj Ivani, šefici Vedrani, mami Žaklini, sestri po grickalicama AniM, Juliji, Petru i juniorima Tomi, Ivoni, Jasminki, Ireni, Luciji, Frani i Jerku.

Hvala mojim roditeljima na bezuvjetnoj podršci i potpunoj slobodi da sama biram svoj put i svoje bitke. Hvala vam na svim odricanjima kojih često nisam ni bila svjesna. Mom bratu, vrlo kratko, ali potpuno iskreno: hvala ti što postojiš.

Mom Ivanu, hvala ti što mi vjeruješ i koračaš pored mene dok lutam i tražim svoje/naše mjesto pod suncem. „All we gotta do is be brave and be kind...".

Veliko hvala mojoj maloj velikoj prijateljici Katarini što je svojim crtežom ukrasila naslovnu stranicu ovog doktorata.

This bumpy road trip brought many special people into my life from whom I taught about work and about life. There were many hard moments but happy ones are those which I remember better. Despite that, I wouldn't like to take the same road again. Mainly because I strongly believe, certainly hope so, that better and more exciting challenges are jet to come.

University of Zagreb
Faculty of Science
Division of Biology

Doctoral thesis

VARIABILITY AND HERITABILITY OF IMMUNOGLOBULIN G GLYCOSYLATION

MAJA PUČIĆ BAKOVIĆ
Genos Ltd.
Hondlova 2/11, Zagreb, Croatia

Immunoglobulin G (IgG) is the most abundant glycoprotein in the human plasma and a major effector molecule of the humoral immune response. Glycans are essential structural components of the IgG antibody and even small changes in glycan composition can have a profound influence on IgG effector function by modulating binding to the Fc receptors. A development of the high-throughput IgG purification from plasma and optimization of chromatographic method for IgG glycan analysis of high resolution and sensitivity enabled a first large-scale study of IgG *N*-glycome in a population. Results from over 2000 individuals revealed a very high variability while heritability of IgG glycans was generally between 30 and 50%. Subclass-specific IgG Fc *N*-glycosylation analysis showed a significant age and sex-dependance. The most prominent changes in glycosylation in females were observed during menopausal age. Age-dependant changes in children differed from changes in adult population in both, direction and intensity. Genome-wide association study (GWAS) of the IgG *N*-glycome identified nine genetic loci that control IgG glycosylation.

Keywords: IgG, *N*-glycans, HILIC, variability, heritability, aging, gender, IgG subclass specificity, GWAS

Supervisor:	Prof. Dr. Gordan Lauc, Ph.D., Faculty of Pharmacy and Biochemistry, University of Zagreb
Reviewers:	Assoc. Prof. Dr. Olga Gornik, Ph.D., Faculty of Pharmacy and Biochemistry, University of Zagreb
	Prof. Dr. Igor Rudan, M.D., Ph.D., University of Edinburgh
	Prof. Dr. Nada Oršolić, Ph.D., Faculty of Science, University of Zagreb

Sveučilište u Zagrebu
Prirodoslovno-matematički fakultet
Biološki odsjek

Doktorska disertacija

VARIJABILNOST I HERITABILNOST GLIKOZILACIJE IMUNOGLOBULINA G

MAJA PUČIĆ BAKOVIĆ
Genos d.o.o.
Hondlova 2/11, Zagreb, Hrvatska

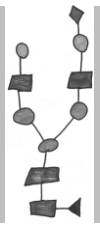
Imunoglobulin G (IgG) je najzastupljeniji glikoprotein u ljudskoj plazmi i glavna efektorska molekula u humoralnom imunološkom odgovoru. Glikani su esencijalne strukturne komponente IgG antitijela pa čak i male promjene u sastavu glikana mogu promjeniti vezanje na Fc receptore i time značajno utjecati na efektorske funkcije IgG-a. Razvoj visoko protočnog pročišćavanja IgG-a iz plazme i optimizacija kromatografske metode visoke razlučivosti i osjetljivosti za analizu IgG glikana omogućili su prvu veliku studiju IgG *N*-glikoma u populaciji. Rezultati istraživanja na preko 2000 pojedinca otkrili su vrlo veliku varijabilnost dok je heritabilnost IgG glikana bila između 30 i 50%. Fc *N*-glikozilacija specifična za pojedine potklase IgG-a pokazala je značajnu spolnu i dobnu ovisnost. Najveće promjene u glikozilaciji opažene su kod žena koje prolaze kroz menopauzu. Smjer i intenzitet dobne ovisnosti promjena u glikozilaciji razlikovali su se između djece i odraslih. Cjelogenomskom asocijacijskom studijom IgG *N*-glikoma identificirano je devet genskih lokusa koji kontroliraju glikozilaciju IgG-a.

Ključne riječi: IgG, *N*-glikani, HILIC, varijabilnost, heritabilnost, starenje, spol, specifičnost IgG potklasa, cjelogenomska asocijacijska studija

Mentor: Prof. dr. sc. Gordan Lauc, Farmaceutsko-biokemijski fakultet Sveučilišta u Zagrebu
Ocjenjivači: Doc. dr. sc. Olga Gornik, Farmaceutsko-biokemijski fakultet Sveučilišta u Zagrebu
Prof. dr. sc. Igor Rudan, dr. med., Sveučilište u Edinburgu
Prof. dr. sc. Nada Oršolić, Prirodoslovno-matematički fakultet Sveučilišta u Zagrebu

Table of Contents

1 INTRODUCTION



1.1 Immunoglobulin G as glycoprotein

Immunoglobulins (Igs) play an important role in the adaptive immune system by providing defence against many different antigens. Just like the majority of eukaryotic proteins,[1, 2] all five distinct classes (IgG, IgM, IgA, IgD and IgE) of human Igs are glycoproteins. The Igs show an extensive diversity in the position and number of the conserved *N*-linked glycosylation sites present both on the Fc (crystallizable fragment) and Fab (antigen-binding fragment).[3] IgG is the most abundant antibody class in the human blood (approx. 10 mg/ml) and a major effector molecule of the humoral immune response. IgG antibodies play an important role in defending the body by mediating activation of a wide range of effector functions which result in destruction and removal of the pathogen. Four subclasses of human IgG (IgG1, IgG2, IgG3 and IgG4) are glycoproteins composed of two heavy and two light chains linked together by interchain disulphide bonds (Figure 1). The two light chains together with the parts of the heavy chains (V_H and C_H1 domains) form two Fab moieties which are linked by a flexible hinge region to one Fc moiety formed by the remainders of the two heavy chains (C_H2 and C_H3 domains).[3] The length of the hinge region and the number of interchain disulphide bonds differ significantly between the IgG subclasses and influence mobility and conformation of the Fab and Fc moieties, with respect to each other.[4, 5]

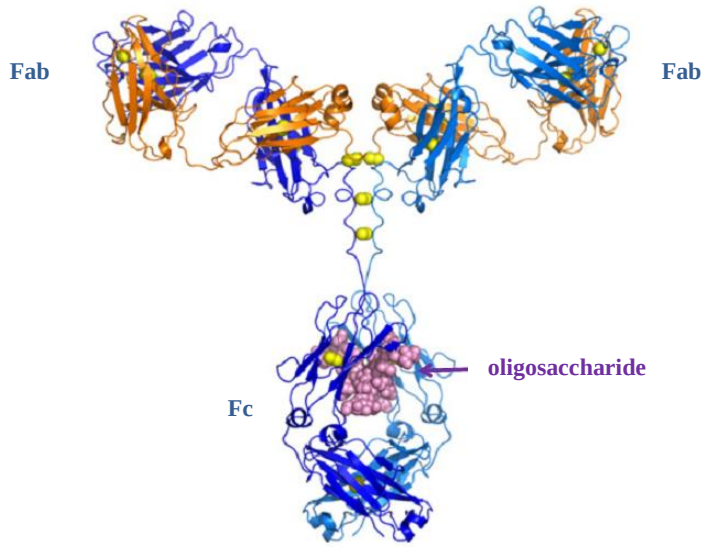


Figure 1. Structure of an IgG1 molecule. Heavy chains (blue), light chains (orange) and disulphide bonds (yellow).[6]

Each heavy chain in the Fc region carries a single covalently attached biantennary *N*-glycan at the highly conserved asparagine 297 in the C_H2 domain.[7] The *N*-glycans present in the Fc of normal polyclonal IgG are biantennary complex-type structures which are mostly core-fucosylated and may contain a bisecting *N*-acetylglucosamine (GlcNAc) and a small portion of sialic acid (Figure 2).[3, 7, 8] The biantennary glycans containing zero, one, or two galactose residues in their outer arms are commonly known as G0, G1, and G2, respectively. Each Asn-297 glycan makes multiple noncovalent interactions with the inner protein surface of the C_H2 domain which stabilize the Fc and help to maintain the structure.[9-12] The majority of IgG *N*-glycans are attached to the heavy chains of the Fc region, but 20% of polyclonal human IgG molecules also contain *N*-glycans within the Fab regions of the light chain, the heavy chain or both.[13] Analysis of Fab regions revealed the presence of highly bisected, substantially galactosylated and sialylated glycans, in contrast to glycans released from the Fc.[13-17]

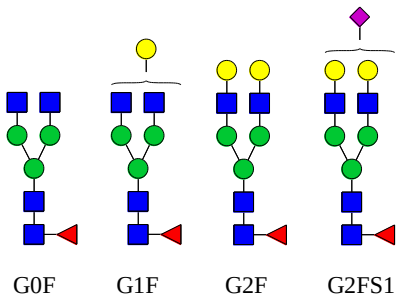


Figure 2. Typical IgG *N*-glycans. Glycan species are given in terms of number of galactoses (G0, G1, G2), fucose (F) and *N*-acetylneuraminic acid (S). Structural schemes are given in terms of blue square (*N*-acetylglucosamine), red triangle (fucose), green circle (mannose), yellow circle (galactose), and purple diamond (*N*-acetylneuraminic acid).[18]

Fc glycans are essential structural components of the IgG molecule and even minor changes in glycan composition can have a profound influence on IgG effector functions by modulating binding to Fc receptors.[12, 13] The influence of individual components of the Fc glycans to antibody activity and efficacy has been a focus of interest over the last years, making controlled and specific glycosylation of therapeutic antibodies an important challenge in biopharmaceutical industry.[13] In general, Fc *N*-glycans are crucial for the antibody's effector functions, whereas Fab *N*-glycans have modulating influence on antigen binding properties.[8] Glycosylation of polyclonal human IgG varies significantly in health and disease, showing dependence on various physiological parameters (age, sex, pregnancy) and pathological conditions (particularly inflammatory and autoimmune diseases).[3, 7]

1.2 Impact of glycosylation on IgG function

The IgG subclasses exhibit considerable differences in their ability to mediate effector functions.[19] Next to the variation in the primary amino acid sequence between different IgG subclasses, an additional diversity, even within the same subclass, is introduced by the glycan moiety.[3] A key effector functions for IgGs are antibody-dependent cellular cytotoxicity (ADCC) and complement-dependent cytotoxicity (CDC). These functions are dependent on the Fc glycan which is crucial for the interaction with cellular Fcγ receptors (FcγRs) and the complement component C1q.[5] Presence of Fc glycan enables binding to FcγRs by maintaining an open conformation of the two heavy chains while removal of the glycan moiety generates a closed conformation unfavorable for the interaction with the Fcγ receptors.[10, 20] Binding of IgG Fc to the FcγRs expressed on effector cells or activating other immune mediators, such as C1q, initiates inflammatory cascades that eliminate pathogen. Truncation or elimination of the Fc glycan leads to reduced or ablated binding of aglycosylated IgG Fc to FcγRs and complement component C1q.[10-12, 21] In contrast, binding to neonatal Fc receptor (FcRn), which determines antibody half life, is not dependent on the glycan moiety.[4, 22]

Influence of the attached glycan on efficacy of effector functions can vary between different antibody glycoforms.[8] For example, the lack of core fucose residue on the Fc *N*-glycan dramatically enhances the IgG1 binding to FcγRIIIa.[7, 22-24] By binding to this activating Fc receptor expressed primarily on natural killer (NK) cells, antibodies initiate ADCC which leads to destruction of target cells.[7, 22, 24] Besides, presence of complex glycans attached to Asn45 and Asn162 of FcγRIIIa has a crucial influence on IgG Fc binding.[22] Subsequently, improved ADCC by increased interaction with Fc receptors was shown in CHO cells transfected with the human β1,4-*N*-acetylglucosaminyltransferase III (GnT-III) gene which adds bisecting GlcNAc.[25, 26] The addition of bisecting GlcNAc, a relatively early event in glycoprotein processing, inhibits α(1,6)-fucosyltransferase (FUT8) and the addition of core fucose.[27] However, it seems that the lack of core fucose, not the presence of bisecting GlcNAc, has the most critical role in enhanced ADCC.[28] Due to drastic enhancement of ADCC, afucosylated monoclonal antibodies exhibit strong therapeutic potential in anti-cancer therapy.[29] It was shown that improved FcγRIIIa binding allows the low doses of afucosylated IgG1 to overcome the competition with high concentrations of heavily fucosylated serum IgG.[24, 30]

The presence of complex glycan structures with galactoses on IgG Fc is required for binding to complement component C1q and triggering classical pathway of complement activation. Removal of terminal galactose residues from monoclonal antibodies was shown to reduce complement activation while FcγR-mediated functions stayed intact.[31, 32] Moreover, IgG Fc glycans with terminal *N*-acetylglucosamine residues were shown to bind mannose-binding lectin (MBL) *in vitro*.[33] The MBL is a serum protein involved in clearance of immune complexes and in the lectin pathway of complement activation.[34] Being a structural homologue of the C1q molecule, the MBL forms a complex with

structural homologues of C1 complement components and when activated triggers CDC.[33, 35] The MBL has also been implicated in certain diseases such as rheumatoid arthritis (RA), an autoimmune disease with a significant increase of agalactosylated (G0) IgG glycoforms. Based on the increased binding of MBL to agalactosylated glycans *in vitro*, it has been suggested that the MBL may contribute to an additional inflammation by activating complement.[34] However, more recent *in vivo* studies suggest that the enhanced MBL binding is not a major factor for antibody activity *in vivo* and that the activity of IgG-G0 glycoforms is fully dependent on activating FcγRs.[36] Very recently Karsten et al.[37] reported anti-inflammatory properties to be mediated by Fc galactosylation via the formation of immune complexes. High *N*-glycan galactosylation of IgG1 in immune complexes was shown to promote the association between the inhibitory IgG receptor FcγRIIB and C-type lectin-like receptor dectin-1, resulting in a blockage of pro-inflammatory effector functions.

Altered affinities of differentially glycosylated IgG antibodies to distinct FcγRs may modulate inflammatory responses. Intravenous immunoglobulin (IVIG), a therapeutic preparation of highly purified IgG from pooled human plasma, has been widely used to treat a number of autoimmune diseases, including immune thrombocytopenic purpura, rheumatoid arthritis, Guillain–Barré syndrome, Kawasaki Disease, and systemic lupus erythematosus.[38-40] When given at high doses (1 to 2 g/kg), IVIG has anti-inflammatory properties. The anti-inflammatory activity of IVIG is a property of the Fc, as Fc fragments were found to be sufficient to suppress inflammation.[41-43] Recent studies by Ravetch and colleagues suggested that immunosupressive function of IVIG is limited to the IgG Fcs with fully processed *N*-glycan terminating in α2,6 sialic acid.[44, 45] The authors used enrichment of sialylated IVIG species by Sambucus nigra agglutinin (SNA) lectin affinity chromatography and fully recombinant, sialylated IgG1 Fc fragments to prevent pathology in a murine model of rheumatoid arthritis. Subsequently, the authors defined the mechanism by which α2,6-sialylated Fc mediate an anti-inflammatory response.[46] It was demonstrated that the FcγRs and FcRn are not involved in this pathway and that a C-type lectin receptor, SIGN-R1, on murine splenic macrophages specifically recognizes sialic acid on the Fcs attenuating autoantibody-initiated inflammation. Likewise, DC-SIGN (Dendritic Cell-Specific Intercellular adhesion molecule-3-Grabbing Non-integrin), the human orthologue of SIGN-R1, was found to bind sialylated IgG Fcs.[46] It was also reported that administration of α2,6-sialylated Fc moieties to a humanized DC-SIGN mouse model suppressed inflammation by promoting IL-33 production through a novel T_H2 pathway.[47] Because SIGN-R1 and DC-SIGN differ significantly in their cellular and tissue distribution, Bayry and colleagues explored the interaction of DC-SIGN and α2,6-sialylated Fc.[48] They found that this interaction is not important for the anti-inflammatory activity of IVIG on human dendritic cells (DC). Later, Stadlmann *et al.*[49] revealed that binding of IVIG to SNA lectin is primarily mediated by Fab glycosylation and that binding of the Fc moieties to the lectin requires at least two sialic acids. More recently, preferential binding of Fab-sialylated IgG to SNA lectin was reported by Guhr *et al.*[50] and

Käsermann *et al.*[51] Moreover, use of IVIG enriched for Fab-sialylated IgG resulted in a decrease rather than an increase of the IVIG efficiency in the murine model of passive immune thrombocytopenia, suggesting that SNA lectin fractionation is not a suitable method to enrich Fc sialylated IgG.[50] Käsermann *et al.*[51] stressed the importance of SNA affinity chromatography process since different elution fraction showed different anti-inflammatory potential. Considering the results of these studies, further reseach is required to clarify the role of sialylated IgG within IVIG in different autoimmune diseases and caution should be taken when extrapolating from mouse models to humans.

1.3 IgG glycosylation in health and disease

For more than three decades, specific patterns of IgG glycosylation have been described in both, healthy and diseased individuals. A number of studies have reported age- and sex-dependence of certain IgG glycosylation features in healthy population. Parekh *et al.* were the first to describe decreasing levels of galactosylation with aging.[52] They have focused solely on galactosylation levels of total IgG and obtained data from 151 individuals of both sexes varying in age from 1 to 70 years. Levels of agalactosylated *N*-glycans were shown to change continuously with age, with a decrease from 30% on average at birth to 20% at the age of 25 and then a steady increase up to 40% at 70 years of age. Interestingly, the levels of monogalactosylated *N*-glycans stayed constant over the whole age range, while digalactosylated *N*-glycans showed the opposite trend to that of agalactosylated glycans. Differences between the sexes were not found.

Sex-related differences in galactosylation were found by Yamada *et al.*[53] Enzimatically released *N*-glycans were labeled with 2-aminopyridine, desialylated and separated by reverse phase high performance liquid chromatography (RP-HPLC). The study included 176 females and 227 males of 0 to 85 years of age. A difference in the level of galactosylation was found between females and males in their twenties with males showing higher level of agalactosylated glycans than females. In addition, the incidence of bisecting GlcNAc was found to increase with age and seemed to reach a plateau at the age of 50. However, no sex differences were found for the bisecting GlcNAc.

A small cohort consisting of 43 female and 37 male individuals ranging in age from 18 to 73 years was used to study IgG glycosylation changes in aging by Shikata *et al.*[54] They have released *N*-glycans by hydrazinolysis, labeled them with *p*-aminobenzoic acid ethyl ester (ABEE) and analyzed by HPLC. Age-dependent changes in galactosylation were only observed in females, while the incidence of bisecting GlcNAc was found to increase with age in both sexes. Moreover, in females, the incidence of monosialylated *N*-glycans also decreased with age which may be explained by a lower level of galactosylated, acceptor glycan structures, at older age.

More recent studies confirmed the earlier findings regarding galactosylation and bisecting GlcNAc.[55, 56] In a large-scale study, Ruhaak et al.[55] analyzed tryptic Fc *N*-

glycopeptides by MALDI-TOF-MS and revealed a potential of the levels of agalactosylated glycoforms with bisecting GlcNAc as an early marker of familial longevity.

Altered IgG glycosylation has been reported during pregnancy, with an increase in galactosylation and sialylation and a decrease in bisecting GlcNAc.[57-59] These changes were reported for both, healthy pregnant females and female patients with RA during pregnancy when remission of the disease is often observed. Moreover, Williams et al.[60] compared total glycosylation between fetal and maternal IgG and described a higher level of galactosylation of fetal IgG than maternal IgG, indicating a preferential transport of highly galactosylated IgG to the fetus. Alternatively, these pronounced differences between fetal and maternal IgG galactosylation levels may partly be explained by the difference in subclass ratios since the portion of IgG1 which tends to be more galactosylated is elevated in cord blood.[60, 61] In contrast, when subclass-specific IgG Fc N-glycosylation was analyzed, remarkably similar Fc glycosylation of all the subclasses between fetal and maternal IgG was observed indicating that placental IgG transport is not Fc glycosylation selective.[61]

It has long been known that IgG glycosylation patterns are skewed toward specific glycoforms in various diseases. As well as in healthy individuals, galactosylation levels are a major source of IgG heterogeneity in diseased individuals. More than 25 years ago, low IgG galactosylation was associated with rheumatoid arthritis and osteoarthritis[62] and since then many reports have followed describing hypogalactosylation in a number of different autoimmune diseases (rheumatoid arthritis[59, 62-69], juvenile chronic arthritis[64, 65, 70], osteoarthritis[62], Chron's disease[64, 71], ulcerative colitis[71], systemic lupus erythematosus with Sjörgen's syndrome[72], myositis[73]), infectious diseases (hepatitis C infection[74], HIV infection[75]) and cancer (ovarian cancer[76]).

In RA, levels of IgG-G0 glycans have been shown to correlate with clinical parameters, disease progression and disease activity.[57, 66] Rademacher *et al.*[77] demonstrated association of IgG-G0 glycoforms with pathogenicity in a murine collagen-induced arthritis model. Interestingly, the disease goes into remission in RA patients during pregnancy when serum levels of IgG-G0 reduce to normal serum levels.[57, 59] Furthermore, association has been reported for these glycoforms and lowered galactosyltransferase activity.[78] IgG Fc glycosylation of anti-citrullinated protein antibodies (ACPA), autoantibodies exibiting unique specificity for RA, showed considerable difference from total serum IgG1.[79, 80] However, the pathogenic role of IgG-G0 glycoforms is not yet fully understood and more research is required to established factors implicated in this inflammatory process.

More recent studies described IgG Fc *N*-glycosylation changes in well-defined antibody-mediated autoimmune diseases, i.e., Lambert-Eaton myasthenic syndrome (LEMS) and myasthenia gravis (MG).[81] Although the described changes of the overall IgG Fc *N*-glycosylation may not reflect the glycosylation of antigen specific IgG, they

could provide insight into potential association of a certain glycosylation feature and antibody pathogenicity.[81]

The regulation of IgG glycosylation is still largely not understood. However, a recent *in vitro* study using a primary human B cell culture has shown that various systemic and microenvironmental factors, such as CpG oligodeoxynucleotide, interleukin 21 and all-*trans* retinoic acid, are able to modulate IgG glycosylation profiles with respect to galactosylation, sialylation and bisecting GlcNAc.[82]

Changes in IgG glycosylation upon immunological challenge have been shown in animal studies. In a murine nephrotoxic serum nephritis model, Kaneko *et al.*[45] found a decrease in total IgG sialylation during an immune response. Lastra *et al*.[83] immunised specific pathogen free mice with bovine serum albumin (BSA) and showed decrease in galactosylation of anti-BSA IgG when antibody titers rose, and later, when titers fell, they became more galactosylated. IgG1 Fc *N*-glycosylation changes induced by influenza and tetanus vaccination in humans were reported by Selman *et al.*[84] Upon active immunization, anti-vaccine IgG showed increased levels of galactosylation and sialylation. Future studies are neccessary to elucidate variation and regulation of IgG glycosylation in an immune response.

1.4 High-throughput glycomics

Due to the glycan complexity and technological limitations, until only a few years ago glycan analysis was extremely laborious and complex, hampering large-scale studies of the glycome. Several factors including template-less biosynthesis, microheterogeneity, lack of a natural chromophore and existence of structural isomers (both position and linkage) contribute to complexity of glycan structural analysis.[85] However, major progress has been made in the last few years resulting in several high-throughput analytical techniques for glycan analysis.[56, 86-90]

In principle, three major strategies are used for glycan analysis: liquid chromatography (e.g. reverese-phase, hydrophilic interaction), capillary electrophoresis and mass spectrometry. When analyzing a single protein, such as IgG, a first step in glycosylation analysis is IgG purification from plasma or serum. Protein purification from a large number of samples is one of the major bottlenecks in a large-scale proteomics and glycoproteomics studies. The most widely used IgG purification technique is affinity-purification using immobilized Protein A or G.[14, 91] Both of these bacterial proteins may be applied for the purification of IgG from human plasma, however, in contrast to Protein G which binds all four IgG subclasses, Protein A does not bind IgG3 subclass.[92] High-throughput purification in a 96 well format of IgG has been previously performed by affinity chromatography with protein A (or G)-Sepharose beads,[92] but this protocol includes incubation which prolonges the whole procedure. Fast purification of large volumes of complex biological compounds, such as plasma, can be achieved by the use of monoliths as stationary phase. Monoliths, a continuous stationary phase cast in a single piece, show many advantageous properties when compared with conventional

particulate supports.[93-96] Due to very large and highly interconnected pores, their dynamic binding capacity is practically independent of the flow rate which significantly reduces purification time.[97, 98]

Release of the *N*-glycans from glycoproteins is most commonly done by use of the enzyme endoglycosidase peptide-*N*-glycosidase F (PNGase F).[86, 91] This enzyme releases all asparagine-linked glycans unless they are α1-3 core fucosylated, a modification observed in plants and invertebrates. The released glycans can be fluorescently tagged (e.g. 2-aminobenzamide, 2-AB; 2-aminobenzoic acid, 2-AA; 1-aminopyrene-3,6,8-trisulfonic acid, APTS) at their reducing end by reductive amination[99-101] and separated by liquid chromatography (LC) or capillary electrophoresis (CE) followed by fluorescence detection or mass spectrometry. Liquid chromatography is a widely applied technique for separation of both neutral and charged oligosaccharides.[76, 86-88, 91] Glycans are predominantly separated using hydrophilic interaction liquid chromatography (HILIC) mode, however, weak anion exchange (WAX) and reverse phase (RP) chromatography are also employed. In HILIC, glycans are resolved based on differences in the hydrophilicity and hydrophobicity with glycan composition, linkage and arm specificity all contributing to retention times.[88, 102] HILIC is often reffered to as „size separation“ since larger glycans tend to elute later. A capillary gel electrophoresis with laser induced fluorescence (CGE-LIF) detection allows rapid separation of glycans in terms of their charge to size ratio.[89, 90, 103, 104] In general, both LC and CE enable reliable glycan quantification. Although the use of smaller HILIC particles (sub-2-µm) and introduction of ultra performance liquid chromatography (UPLC) has led to a marked reduction of analysis times,[87, 105] CE has a much higher throughput capability.[89, 104] However, LC has greater resolving power and is more amenable to coupling to MS enabling in-depth analysis of complex oligosaccharide mixtures.[85]

The released glycans can be analyzed by MS in their native reducing form, after labeling of the reducing end or after permethylation.[106] The two main MS techniques used for glycan analysis are matrix-assisted laser desorption/ionization (MALDI) and electrospray ionization (ESI).[105-109] MALDI is harsher ionization techique and often results in massive desialylation due to in-source and metastable decay.[110] However, this phenomenon can be efficiently avoided by stabilizing sialic acid by methylation of carboxylic acid residues or by permethylation.[111, 112] In addition, the degree of desialylation is strongly influenced by the chosen MALDI matrix.[56, 113, 114] Alternatively, high-performance anion-exchange chromatography with pulsed amperometric detection (HPAEC-PAD) may be used to analyze native reducing-end glycans.[115]

Although in-depth analysis of released glycans may provide a detailed picture of the glycan structure, no information on the original glycan attachment sites and the identity of the carrier protein is provided. Such information can be obtained by the direct analysis of glycopeptides. IgG glycopeptide analysis allows discrimination between different IgG subclasses on the basis of the peptide moieties and masses, and

additionally provides *N*-glycan profiles that are Fc specific.[92] Analysis of IgG glycopeptides can be achieved by MALDI-MS[56, 116] or LC-MS.[15, 49, 92] MALDI-MS analysis of human polyclonal IgG has been shown to allow subclass-specific Fc *N*-glycosylation profiling in a high-throughput manner.[56] However, due to the identical peptide moieties of their tryptic Fc glycopeptides, IgG2 and IgG3 cannot be distinguished.[81, 92] In addition, existence of isomeric glycopeptides causes overlap of certain signals and prevents proper determination of IgG2 fucosylation and IgG4 glycosylation.[56] Therefore, a separation of subclasses prior to MS detection is essential for a detailed characterization of Fc *N*-glycosylation of all IgG subclesses.[92] In general, LC-MS is less prone to in-source and metastable decay of highly sialylated glycopeptides than MALDI-MS which on the other hand offers shorter analysis times.

None of the currently available analytical techiques is capable of performing a detailed structural analysis of protein glycosylation in a single step. The existence of many thousands of different glycans attached to human proteins,[117] multiple glycosylation sites, large variety of the attached glycans in a single glycosylation site and structural complexity of glycans, require implementation of several strategies for a detailed charecterization of protein glycosylation in a complex biological samples.

1.5 Genome wide association studies of the human glycome

Contrary to proteins which are defined by the sequence of nucleotides in the corresponding genes, glycans are synthesized without the direct genetic template. Instead, it has been estimated that over 700 proteins, including various glycosyltransferases, glycosidases, enzymes for sugar nucleotide biosynthesis, transporters, etc., are involved in the complex pathway of glycan biosynthesis.[117, 118] In addition, changes in the abundance and/or localization of any of the enzymes, glycoprotein substrates and activated sugar donors involved in glycan biosynthesis, will affect the final structure of the glycan.[119] Therefore, glycome is shaped by dynamic interactions of both genetic background and environmental influences.[120]

Due to experimental limitations in quantifying glycans in complex biological samples, understanding of the genetic regulation of glycosylation is still very limited.[121] Recent technological advances in glycan analytics allowed reliable, high-throughput quantification of glycans and glycopeptides and enabled large population studies of the human plasma and IgG glycome.[55, 86, 122] Knežević *et al.*[86] performed the first large-scale analysis of human plasma glycome, revealing a high variability in glycome composition between individuals with the median difference between the minimal to maximal values of glycans being over six-fold. This variability appears to be mostly genetically predetermined since environmental factors were found to have a limited influence on the majority of analyzed glycans.[122] In addition, individual plasma glycome appears to change very little even after a prolonged period of time.[123]

The majority of human variability originates from single nucleotide polymorphisms (SNPs) which individually do not have visible phenotypes, but if present in specific

combinations within the same individual can have significant phenotypic effects.[124-126] Due to interaction of hundreds of genes, glycosylation is particularly prone to this type of variability. Some combinations of individual SNPs can be manifested as specific glyco-phenotypes, which might represent potential evolutionary advantages or disadvantages. The most prominent examples are various forms of congenital disorders of glycosylation which are usually caused by a combination of several individual mutations, which, if present individually, do not have visible phenotypes.[121]

Next to the development of high-throughput glycome analysis, in recent years we have also witnessed success of high-throughput genome analysis in gene identification.[127, 128] Genome-wide association studies (GWAS) represent a powerful tool in detecting associations between common SPNs and common disease, as well as other complex traits.[129] Three GWA studies of glycosylation-related traits have been published until now.[130-132] The first comprehensive analysis of common polymorphisms affecting protein glycosylation combined genome-wide association and high-throughput glycomics analysis of 2705 individuals from three European populations.[132] Desialylated total plasma *N*-glycans were separated by HPLC into 13 structurally related groups of glycans from which two more traits were calculated, namely the percentage of glycan structures containing core or antennary fucose, yielding a total of 15 glycan traits. Significant associations with particular SNPs were found for five original peaks, as well as for antennary fucose.[132] The identified SNPs were located within three genes (*FUT6*, *FUT8* and *HNF1α*), all of which are involved in fucosylation. *FUT6* and *FUT8* are known glycosyltransferase genes and they were found to strongly associate with glycan structures that are known substrates or products of these two fucosyltransferases, thus molecular mechanisms behind these associations were clear. However, a third identified gene encoding the transcription factor HNF1α had no previous biological links to glycosylation. Subsequent functional studies revealed a new role for HNF1α as a master transcriptional regulator of antennary fucosylation of plasma proteins. This transcription factor was shown to promote both *de novo* and salvage pathways of GDP-fucose synthesis, expression of fucosyltransferases, *FUT3*, *FUT5* and *FUT6*, and suppressing the expression of *FUT8*.[132]

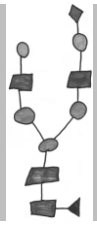
The second published GWAS of the plasma glycome was an extension of the first study, with more individuals included (3533) and a more detailed glycome analysis.[130] All of the previous findings were confirmed and a new association of *HNF1α* with glycan branching was revealed. Moreover, three new genes which associate with plasma glycome have been identified, namely *MGAT5*, *B3GAT1* and *SLC9A9*. In accordance with its biological function, *MGAT5* was found to associate with highly branched glycans. *B3GAT1* is a member of glucuronyltransferase gene family whose product adds glucuronic acid on a subset of human lymphocytes[133], but was not previously reported to exist on plasma proteins. To explain the association of *B3GAT1* with the plasma glycome, a detailed structural analysis was performed confirming the existence of glucuronic acid on a subset of *N*-glycans released from plasma glycoproteins. The third identified gene, *SLC9A9*, was not previously related to glycosylation. However, its

association with tetrasialylated glycans is biologically plausible since *SLC9A9* codes for a proton pump which regulates pH in endosomes[134] and changes in Golgi pH are known to impair protein sialylation.

Plasma glycans originate from different glycoproteins produced in different cell types where they undergo cell type-specific glycosylation. These cell type-specific effects are blurred when analyzing pooled glycans from different glycoproteins whose concentrations can vary in many physiological processes. Therefore, to identify genes that regulate cell type-specific glycosylation it is essential to perform GWAS of glycomes of individual proteins.

High throughput isolation and glycosylation analysis of IgG -variability and heritability of the IgG glycome in three isolated human populations

2



High Throughput Isolation and Glycosylation Analysis of IgG–Variability and Heritability of the IgG Glycome in Three Isolated Human Populations*[S]

Maja Pučić‡, Ana Knežević‡, Jana Vidič§, Barbara Adamczyk¶, Mislav Novokmet‡, Ozren Polašek‖, Olga Gornik, Sandra Šupraha-Goreta**, Mark R. Wormald‡‡, Irma Redžić**, Harry Campbell§§, Alan Wright¶¶, Nicholas D. Hastie¶¶, James F. Wilson§§, Igor Rudan‖§§, Manfred Wuhrer‖‖, Pauline M. Rudd¶, Djuro Josić[ab], and Gordan Lauc‡**[c]**

All immunoglobulin G molecules carry *N*-glycans, which modulate their biological activity. Changes in *N*-glycosylation of IgG associate with various diseases and affect the activity of therapeutic antibodies and intravenous immunoglobulins. We have developed a novel 96-well protein G monolithic plate and used it to rapidly isolate IgG from plasma of 2298 individuals from three isolated human populations. *N*-glycans were released by PNGase F, labeled with 2-aminobenzamide and analyzed by hydrophilic interaction chromatography with fluorescence detection. The majority of the structural features of the IgG glycome were consistent with previous studies, but sialylation was somewhat higher than reported previously. Sialylation was particularly prominent in core fucosylated glycans containing two galactose residues and bisecting GlcNAc where median sialylation level was nearly 80%. Very high variability between individuals was observed, approximately three times higher than in the total plasma glycome. For example, neutral IgG glycans without core fucose varied between 1.3 and 19%, a difference that significantly affects the effector functions of natural antibodies, predisposing or protecting individuals from particular diseases. Heritability of IgG glycans was generally between 30 and 50%. The individual's age was associated with a significant decrease in galactose and increase of bisecting GlcNAc, whereas other functional elements of IgG glycosylation did not change much with age. Gender was not an important predictor for any IgG glycan. An important observation is that competition between glycosyltransferases, which occurs *in vitro*, did not appear to be relevant *in vivo*, indicating that the final glycan structures are not a simple result of competing enzymatic activities, but a carefully regulated outcome designed to meet the prevailing physiological needs. ***Molecular & Cellular Proteomics 10: 10.1074/mcp.M111.010090, 1–15, 2011.***

From the ‡Genos Ltd., Glycobiology Division, Planinska 1, 10000 Zagreb, Croatia; §BIA Separations Ltd., Teslova 30, 1000 Ljubljana, Slovenia; ¶National Institute for Bioprocessing Research and Training, Dublin-Oxford Glycobiology Laboratory, Conway Institute, University College Dublin, Belfield, Dublin 4, Ireland; ‖University of Split School of Medicine, Šoltanska 2, Split, Croatia; **University of Zagreb, Faculty of Pharmacy and Biochemistry, Ante Kovačića 1, 10000 Zagreb, Croatia; ‡‡Oxford Glycobiology Institute, Department of Biochemistry, University of Oxford, Oxford, UK; §§Centre for Population Health Sciences, The University of Edinburgh Medical School, Edinburgh, UK; ¶¶MRC Human Genetics Unit; Institute of Genetics and Molecular Medicine, Western General Hospital, Edinburgh, UK; ‖‖Biomolecular Mass Spectrometry Unit, Department of Parasitology, Leiden University Medical Center, Leiden, The Netherlands; [a]COBRE Center for Cancer Research Development, Rhode Island Hospital and Brown University, Providence, RI 02903; [b]Department of Biotechnology, University of Rijeka, Trg Braće Mažuranića 10, 51 000 Rijeka, Croatia

[S] This article contains supplemental Tables S1 and S2.

Received April 6, 2011, and in revised form, June 6, 2011

⌘ *Author's Choice*—Final version full access.

Published, MCP Papers in Press, June 8, 2011, DOI 10.1074/mcp.M111.010090

Glycosylation is a widespread post-translational modification capable of producing significant structural changes to proteins. Contrary to the core *N*-glycan structure, which is essential for multicellular life (1), mutations in genes involved in modifications of glycan antennae are common and cause a large part of individual phenotypic variations that exist in humans and other higher organisms. Glycosylation of membrane receptors modulates adaptive properties of the cell membrane and affects communication between the cell and its environment (2). Deregulation of glycosylation is associated with a wide range of diseases, including cancer, diabetes, cardiovascular, congenital, immunological and infectious disorders (3–5). Variations in glycosylation are of great physiological significance because it has been demonstrated that changes in glycans significantly modulate the structure and function of polypeptide parts of glycoproteins (6), and a prom-

inent example for this type of regulation is the immunoglobulin G (IgG).

Each heavy chain of IgG carries a single covalently attached bi-antennary *N*-glycan at the highly conserved asparagine 297 residue in each of the C_H2 domains of the Fc region of the molecule. The attached oligosaccharides are structurally important for the stability of the antibody and its effector functions (7). In addition, 15–20% of normal IgG molecules also bear complex bi-antennary oligosaccharides attached to the variable regions of the light chain, heavy chain or both (8, 9). Decreased galactosylation of IgG glycans in rheumatoid arthritis was reported over 25 years ago (10) and numerous subsequent studies of IgG glycosylation revealed a number of important functional consequences of structural alterations in IgG glycans. For example, the addition of sialic acids dramatically changes the physiological role of IgGs by converting them from pro-inflammatory into anti-inflammatory agents (11, 12). Another structural change to IgG glycans, the addition of fucose to the glycan core, interferes with binding of IgG to FcγRIIIa and dampens its ability to destroy target cells through antibody dependent cell-mediated cytotoxicity (ADCC) (13, 14). Lack of core fucose enhances the clinical efficacy of monoclonal antibodies, which exert their therapeutic effect by ADCC mediated killing (15–17). However, despite the undisputed importance of glycosylation for the function of IgGs, a large scale study that identifies the variability and heritability of IgG glycosylation in human populations has not been attempted.

One of the major bottlenecks in large scale proteomics and glycomics studies is protein purification from a large number of samples. Affinity chromatography and liquid chromatography have been widely used, as they are versatile techniques for this purpose. A combination of affinity chromatography and monolithic supports exhibits many advantageous properties when compared with conventional particulate supports (18–22). Monoliths are continuous stationary phases cast in a single piece with very large and highly interconnected pores (23). In comparison to particulate supports where molecules are transferred by diffusion, the high porosity of monoliths allows convective mass transport. This makes resolution and dynamic binding capacity practically independent of the flow rate (24–27). High dynamic binding capacity for large molecules and high flow rates at a very low pressure drop enable rapid processing of large volumes of complex biological mixtures (28). Polymethacrylate monoliths, specifically poly(glycidyl methacrylate-*co*-ethylene dimethacrylate), possess all of the above mentioned characteristics of monolithic supports. In addition, they are also known for their good mechanical strength, pH resistance, high surface area, high porosity, and simple attachment of ligands to the epoxy groups (29). One of the most commonly used bioaffinity ligands for the isolation of IgG is protein G (30–32). All four subclasses of human IgG strongly bind to protein G through their Fc fragments. Here we present the development and application of a 96-well Protein G monolithic plate for high throughput isolation of IgG and its application for the first large scale population study of the IgG glycome.

EXPERIMENTAL PROCEDURES

Chemicals—Glycidyl methacrylate, ethylene dimethacrylate, cyclohexanol, and 1-dodecanol were purchased from Sigma-Aldrich (St. Louis, MO). Photoinitiator was purchased from CIBA (Basel, Switzerland) and Protein G from GE Healthcare (Uppsala, Sweden). Sodium acetate, sulfuric acid, and hydrochloric acid (37%) were obtained from Merck (Darmstadt, Germany). All the buffers were filtered through a 0.45 μm pore size filter composed of Sartolon polyamide (Sartorius, Goettingen, Germany). The 96-well plates with frits, mean pore size 36 microns, were purchased from Chromacol (Welwyn Garden City, United Kingdom).

Chemicals for buffer preparations (phosphate buffered saline (PBS), Tris, HCl, NaOH, formic acid, ammonium bicarbonate, propan-2-ol) were purchased from Fisher Scientific (Pittsburgh, PA) and Sigma-Aldrich. Chemicals for running the SDS-PAGE were purchased from Invitrogen (Carlsbad, CA). Sodium bicarbonate, DL-dithiothreitol, iodoacetamide, ammonium persulfate, 2-aminobenzamide, sodium cyanoborohydride, acetic acid, and dimethyl sulfoxide were from Sigma-Aldrich and ultra pure water (Purite Fusion 40 water purification system, Purite Ltd., Thame, UK) were used throughout.

Human Samples—This study was based on samples from respondents who were residents of the Croatian Adriatic islands Vis and Korčula or the Northern Scottish Orkney Islands and who were recruited within a larger genetic epidemiology program that sought to investigate genetic variability and map genes influencing common complex diseases and disease traits in genetically isolated populations (33, 34). The genetic-epidemiology program on the islands began in 2002, and continues today. The sampling framework was based on the voting register in Croatia, which was used to send postal invitations to all adult inhabitants (over 18 years of age); in Orkney subjects were volunteers from the Orkney Complex Disease Study, again aged over 18 years.

The sample for this study consisted of 906 subjects from the Vis island (39.4%), 915 (39.8%) from Korčula island and 477 from the Orkney islands (20.8%) totaling to 2298 individuals. The age range for the entire sample was 18–100 years (median age 56, interquartile range 22 years). There were 894 men (39.2%) and 1384 women in the sample (60.8%), for 20 people gender data were missing. Heritability analysis was performed for the Vis Island sample only, because of a more extensive number of familial links. The genealogical information was reconstructed based on the Church Parish records and information provided by the subjects, and then checked against genetic data on allele sharing between relatives as a quality control measure to exclude data errors. The sample contained a total of 809 genealogical relationships (including 205 parent-child, 123 sibling, and 481 other relationships). The Korčula sample contained a much lower number of familial links and because of large standard errors arising from rather shallow genealogical records, we did not calculate heritability estimates for the Korčula island sample.

All of the members of the three sample groups were interviewed by one of the trained surveyors, based on an extensive questionnaire (35). The questionnaire collected data on personal characteristics (name, date, and place of birth, gender, marital status, education level and occupation), selected health-related lifestyle variables (such as diet and smoking status), health complaints, drug intake and hospitalization records. Blood was taken in epruvetes containing anticoagulant and immediately processed; plasma was separated by centrifugation and stored at −70 °C. This study conformed to the ethical guidelines of the 1975 Declaration of Helsinki. All respondents signed

an informed consent form before participating in the study and the study was approved by the appropriate Ethics Board of the University of Zagreb Medical School and by Research Ethics Committees in Orkney and Aberdeen.

Preparation of Protein G Monolithic Plates—The 96-well plates consisting of a polymethacrylate (poly(glycidyl methacrylate-*co*-ethylene dimethacrylate)) monolithic stationary phase with protein G coupled to the epoxy groups and casted inside each well was custom designed and prepared by BIA Separations (Ljubljana, Slovenia). The basic monolith was synthesized by a free-radical polymerization of GMA and a cross-linking agent, EDMA, in the presence of porogenic solvents, cyclohexanol and dodecanol (60 vol.% of the reaction mixture) as described by Tennikova *et al.* (36), but instead of thermally initiated polymerization, UV polymerization was used. The preparation of the monolithic stationary phase is a simple process and the polymerization mixture, which consists of monomers and porogens, is polymerized by applying heat and UV light. In both types of polymerization, an important property of a monolithic macroporous material is the pore size distribution. The photoinduced copolymerization of 150 μl of the mixtures of monomers, cross-linking agent, photoinitiator, and porogenic solvents was performed at room temperature directly in each well of 96 plates. The mixture was irradiated with a constant intensity from a 5 $\times$ 8 W mercury lamp using a wavelength of 312 nm (UVItec Ltd, Cambridge, UK) with an exposure time of up to 180 min. Although the instrument does not enable active cooling, the temperature did not exceed 30 °C thus effectively excluding thermal initiation. After the polymerization was completed, each well of the 96-well plate was extensively washed with ethanol to wash out the porogenic solvents and other soluble compounds. The average pore size was determined by intrusive mercury porosimetry (PASCAL 440 porosimeter, Thermoquest Italia, Rodano, Italy). The pore size distribution of the monoliths were around 700 nm, which is comparable to thermally polymerized monoliths (37). The immobilization of protein G on the monoliths in the 96-well plate was performed by flushing the monoliths with protein G solution prepared in a buffer solution of sodium acetate. Afterward the monoliths were flushed with deionized water and the deactivation of the remaining epoxy groups was performed with 0.5 M solution of sulfuric acid.

Isolation of IgG—Before use, the monolithic plate was washed with 10 column volumes (CV) of ultra pure water and then equilibrated with 10 CV of binding buffer (1X PBS, pH 7.4). Plasma samples (50 μl) were diluted 10 $\times$ with the binding buffer and applied to the Protein G plate. The filtration of the samples was completed in ~5 min. The plate was then washed five times with 5 CV of binding buffer to remove unbound proteins. IgG was released from the protein G monoliths using 5 CV of elution solvent (0.1 M formic acid, pH 2.5). Eluates were collected in a 96-deep-well plate and immediately neutralized to pH 7.0 with neutralization buffer (1 M ammonium bicarbonate) to maintain the IgG stability. After each sample application, the monoliths were regenerated with the following buffers: 10 CV of 10 $\times$ PBS, followed by 10 CV of 0.1 M formic acid and afterward 10 CV of 1 $\times$ PBS to re-equilibrate the monoliths. Each step of the chromatographic procedure was done under vacuum (cca. 60 mmHg pressure reduction while applying the samples, 500 mmHg during elution and washing steps) using a manual set-up consisting of a multichannel pipet, a vacuum manifold (Beckman Coulter, Brea, CA) and a vacuum pump (Pall Life Sciences, Ann Arbor, MI). If the plate was not used for a short period, it was stored in 20% ethanol (v/v) at 4 °C.

After repeated use of the plate contaminants present in the sample sometimes did not completely elute from the monolithic stationary phase. A specific cleaning protocol was developed that included washing with 0.1 M NaOH to remove precipitated proteins and with 30% propan-2-ol to remove strongly bound hydrophobic proteins or lipids. This procedure effectively removed all precipitates and did not significantly diminish IgG binding capacity of the immobilized protein G.

The purity of the isolated IgG was verified by SDS-PAGE with NuPAGE Novex 4–12% Bis-Tris gels in an Xcell SureLock Mini-Cell (Invitrogen) according to the manufacturer. Precision Plus Protein All Blue Standards (BioRad, Hercules, CA) was used as the molecular weight marker. The gels were run at 180 V for 45 min, stained with GelCode Blue (Pierce) and visualized by a VersaDoc Imaging System (BioRad).

Glycan Release and Labeling—Glycan release and labeling was performed as reported previously (38). Plasma proteins were immobilized in a block of SDS-polyacrylamide gel and *N*-glycans were released by digestion with recombinant N-glycosidase F (ProZyme, CA). This was done in a 96-well microtiter plate to achieve the best throughput of sample preparation. After extraction, glycans were fluorescently labeled with 2-aminobenzamide.

Exoglycosidase Digestions of 2-AB Labeled IgG N-Glycans—The following enzymes, all purchased from ProZyme (San Leandro, CA), were used for digestions: Sialidase A™/NANase III (recombinant gene from *Arthrobacter ureafaciens*, expressed in *Escherichia coli*), 5 mU; α(1–2,3,4,6)fucosidase (bovine kidney), 1.25 mU; α(1–3,4)-fucosidase (almond meal), 1.6 mU; β(1–3,4)-galactosidase (bovine testis), 5 mU; β(1–4)-galactosidase (*Streptococcus pneumoniae*), 2 mU; β-N-acetylhexosaminidase/HEXase I (recombinant gene from *Streptococcus pneumoniae*, expressed in *E. coli*), 40 mU; α(1–2,3,6)-mannosidase (jack bean), 150 mU. Aliquots of the 2-AB labeled glycan pool were dried down and digested in a mixture of enzymes, corresponding 1X concentrated manufacturers buffer and water in total volume of 5 μl. After overnight incubation at 37 °C, enzymes were removed by filtration through the AcroPrep 96 Filter Plates, 10K (Pall Corporation, MI, USA). Digested glycans were then separated by HILIC-UPLC for comparison against an undigested equivalent.

Hydrophilic Interaction Chromatography—Fluorescently labeled *N*-glycans were separated by ultra performance liquid chromatography on a Waters Acquity UPLC instrument consisting of a quaternary solvent manager, sample manager and a FLR fluorescence detector set with excitation and emission wavelengths of 330 and 420 nm, respectively. The instrument was under the control of Empower 2 software, build 2145 (Waters, Milford, MA). Labeled *N*-glycans were separated on a Waters BEH Glycan chromatography column, 100 $\times$ 2.1 mm i.d., 1.7 μm BEH particles, with 100 mM ammonium formate, pH 4.4, as solvent A and acetonitrile as solvent B. Recently reported methods for UPLC profiling of glycans (39, 40) were used as a starting point for the development of the separation method that used linear gradient of 75–62% acetonitrile at flow rate of 0.4 ml/min in a 20 min analytical run. Samples were maintained at 5 °C before injection, and the separation temperature was 60 °C. The system was calibrated using an external standard of hydrolyzed and 2-AB labeled glucose oligomers from which the retention times for the individual glycans were converted to glucose units. Data processing was performed using an automatic processing method with a traditional integration algorithm after which each chromatogram was manually corrected to maintain the same intervals of integration for all the samples. The chromatograms obtained were all separated in the same manner into 24 peaks and the amount of glycans in each peak was expressed as % of total integrated area.

MS Analysis of Glycans—Before MS analysis of each glycan peak, the 2-AB labeled IgG *N*-glycan pool was fractionated by hydrophilic interaction high performance liquid chromatography (HILIC) on a 100 $\times$ 2.1 mm i.d., 1.7 μm BEH particles column using a linear gradient of 75–62% acetonitrile with 100 mM ammonium formate, pH 4.4, as solvent A and acetonitrile as solvent B. UltiMate Dual Gradient LC system (Dionex, Sunnyvale, CA) controlled by Chromeleon soft-

ware and connected to *FP-2020* Plus fluorescence detector (Jasco, Easton, MD) was used. To obtain the same separation as with UPLC system, flow was adjusted to 0.3 ml/min and analytical run time was prolonged to 60 min. Collected fractions were dried by vacuum centrifugation and resuspended in water.

Nano-LC-ESI-MS/MS. MS analysis of the collected glycan fractions was performed using an Ultimate 3000 nano-LC system (Dionex/LC Packings, Amsterdam, The Netherlands) equipped with a reverse phase trap column (C_{18} PepMap 100Å, 5 μm, 300 μm $\times$ 5 mm; Dionex/LC Packings) and a nano column (C_{18} PepMap 100Å, 3 μm, 75 μm $\times$ 150 mm; Dionex/LC Packings).

The column was equilibrated at room temperature with eluent A (0.1% formic acid in water) at a flow rate of 300 nL/min. For fractions with disialylated glycans, extra 0.04% of trifluoroacetic acid was added to the eluent A. After injection of the samples, a gradient was applied to 25% eluent B (95% acetonitrile) in 15 min and to 70% eluent B at 25 min followed by an isocratic elution with 70% eluent B for 5 min. The eluate was monitored by UV absorption at 214 nm. The LC system was coupled *via* an online nanospray source to an Esquire HCTultra ESI-IT-MS (Bruker Daltonics, Bremen, Germany) operated in the positive ion mode. For electrospray (1100–1250 V), stainless steel capillaries with an inner diameter of 30 μm (Proxeon, Odense, Denmark) were used. The solvent was evaporated at 170 °C employing a nitrogen stream of 7 L/min. Ions from *m/z* 500 to 1800 were registered. Automatic fragment ion analysis was enabled, resulting in MS/MS spectra of the most abundant ions in the MS spectra. Glycan structures were assigned using GlycoWorkbench (41).

MALDI-TOF-MS. 2-AB labeled glycan fractions were spotted onto an AnchorChip target plate (Bruker Daltonics, Bremen, Germany). Subsequently 1 μl of 5 mg/ml 2,5-dihydroxybenzoic acid in 50% acetonitrile was applied on top of each sample and allowed to dry at room temperature. MALDI-TOF-MS was performed on an UltrafleX II mass spectrometer (Bruker Daltonics). Calibration was performed on a peptide calibration standard. Spectra were acquired in reflector positive mode over the *m/z* range from 700 to 3500 Da for a total of 2000 shots. Glycan structures were assigned using GlycoWorkbench (41).

Calculation of Derived Glycosylation Traits—Derived glycosylation traits were approximated from the ratios of glycan peaks (GP1-GP24) each of which combined the glycans with the same structural characteristics (see Table I). The minor glycan peak GP3 was excluded from all the calculations because in some samples it co-eluted with a contaminant that significantly affected its value. Derived traits were defined as: the percentage of sialylation of fucosylated galactosylated structures without bisecting GlcNAc in total IgG glycans-FGS/(FG + FGS) = SUM(GP16 + GP18 + GP23)/SUM(GP16 + GP18 + GP23 + GP8 + GP9 + GP14)* 100; the percentage of sialylation of fucosylated galactosylated structures with bisecting GlcNAc in total IgG glycans-FBGS/(FBG + FBGS) = SUM(GP19 + GP24)/SUM(GP19 + GP24 + GP10 + GP11 + GP15)* 100; the percentage of sialylation of all fucosylated structures without bisecting GlcNAc in total IgG glycans-FGS/(F + FG + FGS) = SUM(GP16 + GP18 + GP23)/SUM(GP16 + GP18 + GP23 + GP4 + GP8 + GP9 + GP14)* 100; the percentage of sialylation of all fucosylated structures with bisecting GlcNAc in total IgG glycans-FBGS/(FB + FBG + FBGS) = SUM(GP19 + GP24)/SUM(GP19 + GP24 + GP6 + GP10 + GP11 + GP15)* 100; the percentage of monosialylation of fucosylated monogalactosylated structures in total IgG glycans-FG1S1/(FG1 + FG1S1) = GP16/SUM(GP16 + GP8 + GP9)* 100; the percentage of monosialylation of fucosylated digalactosylated structures in total IgG glycans-FG2S1/(FG2+FG2S1+FG2S2) = GP18/SUM(GP18 + GP14 + GP23)* 100; the percentage of disialylation of fucosylated digalactosylated structures in total IgG glycans-FG2S2/(FG2 + FG2S1 + FG2S2) = GP23/SUM(GP23 + GP14 + GP18)* 100; the percentage of monosialylation of fucosylated digalactosylated structures with bisecting GlcNAc in total IgG glycans-FBG2S1/(FBG2 + FBG2S1 + FBG2S2) = GP19/SUM(GP19 + GP15 + GP24)* 100; the percentage of disialylation of fucosylated digalactosylated structures with bisecting GlcNAc in total IgG glycans-FBG2S2/(FBG2 + FBG2S1 + FBG2S2) = GP24/SUM(GP24 + GP15 + GP19)* 100; ratio of all fucosylated (± bisecting GlcNAc) monosialylated and disialylated structures in total IgG glycans-$F^{total}S1/F^{total}S2$ = SUM(GP16 + GP18 + GP19)/SUM(GP23 + GP24); ratio of fucosylated (without bisecting GlcNAc) monosialylated and disialylated structures in total IgG glycans-FS1/FS2 = SUM(GP16 + GP18)/GP23; ratio of fucosylated (with bisecting GlcNAc) monosialylated and disialylated structures in total IgG glycans - FBS1/FBS2 = GP19/GP24; ratio of all fucosylated sialylated structures with and without bisecting GlcNAc-FBS^{total}/FS^{total} = SUM(GP19 + GP24)/SUM(GP16 + GP18 + GP23); ratio of fucosylated monosialylated structures with and without bisecting GlcNAc-FBS1/FS1 = GP19/SUM(GP16 + GP18); the incidence of bisecting GlcNAc in all fucosylated monosialylated structures in total IgG glycans-FBS1/(FS1 + FBS1) = GP19/SUM(GP16 + GP18 + GP19); ratio of fucosylated disialylated structures with and without bisecting GlcNAc - FBS2/FS2 = GP24/GP23; the incidence of bisecting GlcNAc in all fucosylated disialylated structures in total IgG glycans - FBS2/(FS2 + FBS2) = GP24/SUM(GP23 + GP24). The following derived traits were approximated only from the ratios of glycan peaks containing neutral glycan as a major structure. First, the percentage of each neutral glycan peak ($GP1^n$ - $GP15^n$) was calculated from the total neutral glycan fraction (SUM(GP1:GP15)) and then traits were defined as: the percentage of agalactosylated structures in total neutral glycan fraction-$G0^n$ = SUM($GP1^n$ + $GP2^n$ + $GP4^n$ + $GP6^n$); the percentage of monogalactosylated structures in total neutral glycan fraction - $G1^n$ = SUM($GP7^n$ + $GP8^n$ + $GP9^n$ + $GP10^n$ + $GP11^n$); the percentage of digalactosylated structures in total neutral glycan fraction - $G2^n$ = SUM($GP12^n$ + $GP13^n$ + $GP14^n$ + $GP15^n$); the percentage of all fucosylated (±bisecting GlcNAc) structures in total neutral glycan fraction - $F^{n\ total}$ = SUM($GP1^n$ + $GP4^n$ + $GP6^n$ + $GP8^n$ + $GP9^n$ + $GP10^n$ + $GP11^n$ + $GP14^n$ + $GP15^n$); the percentage of fucosylation of agalactosylated structures-$FG0^{n\ total}/G0^n$ = SUM($GP1^n$ + $GP4^n$ + $GP6^n$)/$G0^n$ * 100; the percentage of fucosylation of monogalactosylated structures-$FG1^{n\ total}/G1^n$ = SUM($GP8^n$ + $GP9^n$ + $GP10^n$ + $GP11^n$)/$G1^n$ * 100; the percentage of fucosylation of digalactosylated structures-$FG2^{n\ total}/G2^n$ = SUM($GP14^n$ + $GP15^n$)/$G2^n$ * 100; the percentage of fucosylated (without bisecting GlcNAc) structures in total neutral glycan fraction-F^n = SUM($GP1^n$ + $GP4^n$ + $GP8^n$ + $GP9^n$ + $GP14^n$); the percentage of fucosylation (without bisecting GlcNAc) of agalactosylated structures - $FG0^n/G0^n$ = SUM($GP1^n$ + $GP4^n$)/$G0^n$ * 100; the percentage of fucosylation (without bisecting GlcNAc) of monogalactosylated structures -$FG1^n/G1^n$ = SUM($GP8^n$ + $GP9^n$)/$G1^n$ * 100; the percentage of fucosylation (without bisecting GlcNAc) of digalactosylated structures-$FG2^n/G2^n$ = $GP14^n/G2^n$ * 100; the percentage of fucosylated (with bisecting GlcNAc) structures in total neutral glycan fraction-FB^n = SUM($GP6^n$ + $GP10^n$ + $GP11^n$ + $GP15^n$); the percentage of fucosylation (with bisecting GlcNAc) of agalactosylated structures-$FBG0^n/G0^n$ = $GP6^n/G0^n$ * 100; the percentage of fucosylation (with bisecting GlcNAc) of monogalactosylated structures-$FBG1^n/G1^n$ = SUM($GP10^n$ + $GP11^n$)/$G1^n$ * 100; the percentage of fucosylation (with bisecting GlcNAc) of digalactosylated structures-$FBG2^n/GP2^n$ = $GP15^n/G2^n$ * 100; ratio of fucosylated structures with and without bisecting GlcNAc-FB^n/F^n = FB^n/F^n; the incidence of bisecting GlcNAc in all fucosylated structures in total neutral glycan fraction-$FB^n/F^{n\ total}$ = $FB^n/F^{n\ total}$ * 100; ratio of fucosylated non-bisecting GlcNAc structures and all structures with bisecting GlcNAc-$F^n/(B^n + FB^n)$ = $F^n/(GP13^n + FB^n)$; ratio of structures with bisecting GlcNAc and all fucosylated structures (± bisecting GlcNAc)-$B^n/(F^n + FB^n)$ (‰) = $GP13^n/(F^n$+ % $FB^n)$ * 1000; ratio of

fucosylated digalactosylated structures with and without bisecting GlcNAc-FBG2^n/FG2^n = GP15^n/GP14^n; the incidence of bisecting GlcNAc in all fucosylated digalactosylated structures in total neutral glycan fraction-FBG2^n/(FG2^n + FBG2^n) = GP15^n/(GP14^n + GP15^n) × 100; ratio of fucosylated digalactosylated nonbisecting GlcNAc structures and all digalactosylated structures with bisecting GlcNAc-FG2^n/(BG2^n + FBG2^n) = GP14^n/(GP13^n + GP15^n); ratio of digalactosylated structures with bisecting GlcNAc and all fucosylated digalactosylated structures (±bisecting GlcNAc)-BG2^n/(FG2^n + FBG2^n) (‰) = GP13^n/(GP14^n + GP15^n) * 1000.

Overview of derived traits and glycans structures present in each chromatographic peak is available as supplemental Table S1.

Molecular Modeling—Molecular modeling was performed on a Silicon Graphics Fuel work station using InsightII and Discover software (MSI Inc.). The crystal structure of IgG Fc (42) was used as the basis for modeling (pdb code 1H3V; obtained from the Protein Data Bank (43). The IgG Fc used for crystallization contains A2G2F glycans at Asn 297, but the majority of the 3-arm is disordered in the crystal. Preferred conformations for the bisecting GlcNAc were obtained from the database of glycosidic linkage conformations (44, 45).

Statistical Analysis—The descriptive part of this study was based on non-parametric methods because deviations from normal distribution were observed. Correlations were performed with Spearman's rank test and gender differences were tested with the Mann-Whitney test. The basic analysis was performed in genealogy unlinked individuals, in order not to bias the results with sample relatedness. The sample size for the subset of these analyses was thus reduced to 612 samples from Vis Island (67.5% of the full sample size), 520 samples (56.8%) from Korčula Island and 477 from Orkney islands, or 1609 samples in total. SPSS version 13 was used in the analysis (SPSS Inc, Chicago, IL). In the last step of analysis we used pedigree information and entire sample size to establish relationships between respondents to calculate heritability estimates. Heritability analysis was conducted using polygenic models in Sequential Oligogenic Linkage Analysis Routines (46). Age and sex were used as predictor variables in these models. Significance was set at $p < 0.05$.

RESULTS AND DISCUSSION

Development of a New Affinity Material and Purification of IgG from 2298 Human Plasma Samples—The newly developed protein G monolithic plate with the bed volume of a single protein G column of 150 μl was used for IgG purification. Plasma samples (V = 50 μl) were diluted ten times with PBS, pH 7.4, and loaded onto the columns. The dynamic binding capacity for IgG was not exceeded. The purity of eluted fractions was examined by SDS-PAGE revealing two clearly visible bands corresponding to the molecular masses of heavy (~50 kDa) and light chains (~25 kDa) of IgG (supplemental Fig. S1).

The newly developed 96-well protein G plates were used to purify IgG from 2298 plasma samples. The entire chromatographic procedure for 96 samples, including the binding, washing and elution steps, was performed in less than 30 min. The concentration of IgG in human plasma varies between 6.6 and 14.5 mg/ml (47). The average amount of IgG isolated from 50 μl of plasma with the use of 96-well protein G monolithic plates was 640 μg, indicating that the majority of IgG in the sample was successfully captured and released.

Because the elution of IgG from protein G requires very low pH, there is a certain risk of loss of sialic acids because of acid hydrolysis. Isolation with monoliths minimized this risk as elution occurs within seconds and therefore the pH was quickly restored to neutrality preserving the integrity and activity of the IgG molecules. The use of a vacuum for liquid transfer enabled easy and efficient handling of large sample sets.

Analysis of IgG Glycans—*N*-glycans attached to IgG were released using PNGase F and labeled with 2-aminobenzamide (2-AB). Labeled glycans were separated by hydrophilic interaction chromatography on a recently introduced Waters BEH Glycan chromatography column. Because this was the first application of this column for the analysis of IgG glycans in our laboratory, each chromatographic peak was collected and analyzed by exoglycosidase digestion (data not shown) and mass spectrometry to determine the glycan structures that elute in each peak (supplemental Table S2). A total of 23 major and 10 minor glycan structures were successfully resolved (Fig. 1, Table I). An additional 40 glycosylation traits (galactosylation, core fucosylation, sialylation, etc.) were derived from ratios of these glycan peaks as described in the *Experimental Procedures* section.

Approximately 96% of all neutral IgG glycans contained core fucose (Table II: F$^{n\ \mathrm{total}}$). In contrast with immunoglobulins, which are mainly produced by B-cells, other major plasma glycoproteins generally originate from hepatocytes, which express only very low levels of the FUT8 fucosyltransferase and thus contain a low percentage of core fucosylated glycans (48). They are also generally more highly sialylated and consequently over 35% of all glycans in human plasma are A2G2S2 structures (38). The median level of A2G2S2 glycans in our IgG preparations was slightly over 3% (Table II: GP21), which is very similar to previously reported values (49). This indicated that the level of contaminating plasma proteins in our IgG preparations was very low, but because disialylated structures without core fucose on IgG are of rather low frequency, even low level of contaminating plasma proteins would cause significant errors in calculation of the sialylation level. To minimize this problem, all calculations of sialylation-related traits were performed using only core fucosylated structures, which are predominant on IgG and less abundant on other plasma proteins.

Galactosylation—Galactosylation of IgG is one of the most studied glycosylation feature of any glycoprotein. Because the discovery of decreased galactosylation of IgG in rheumatoid arthritis more than 25 years ago (10), over 50 different studies have analyzed the role of IgG galactosylation in different inflammatory diseases (50). In our three populations, neutral glycans without galactoses (G0^n) were slightly below 40%, neutral glycans with one terminal galactose (G1^n) slightly above 40% and neutral glycans with two terminal galactoses (G2^n) were ~20% of the neutral IgG glycome (Table II). Because only 11.6% of G1 was sialylated (Table II:

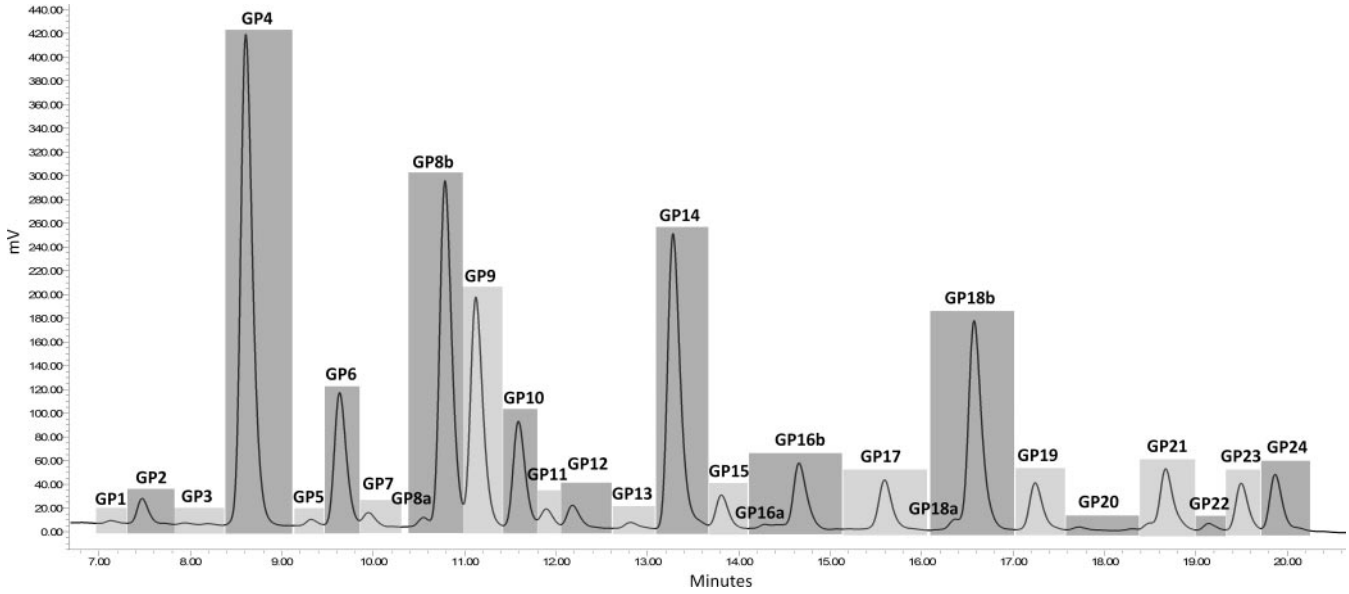


FIG. 1. **UPLC analysis of the IgG glycome.** IgG glycome was separated into 24 chromatographic peaks by hydrophilic interaction chromatography. Compositions and structural schemes of glycans in each chromatographic peak and the average percentage of individual structures are shown in Table I.

FG1S1/(FG1+FG1S1)), whereas over 50% of G2 was sialylated (Table II: FG2S2/(FG2+FG2S1+FG2S2) and FBG2S2/(FBG2+FBG2S1+FBG2S2)), in the total pool of IgG glycans, median levels of a galactosylated, mono galactosylated, and digalactosylated structures-the latter two with or without sialic acid-were approximately the same. However, the ratio of G0^n to G2^n increased significantly with age, (see section "*Effects of Gender and Age*"), thus the observed equilibrium between glycans with different number of galactoses could be a peculiarity of our relatively old study cohorts.

A clear preference for the addition of the first galactose to the antennae, which extend from the α1–6 linked mannose residues of the trimannosyl core (6-arm), was observed because over 65% of G1^n structures contained galactose on the 6-arm (Table II: GP8/(GP8+GP9)). This difference was even larger in structures containing bisecting GlcNAc where over 85% of all G1^n structures contained galactose on the 6-arm (Table II: GP10/(GP10+GP11)). This is in accordance with previous studies, which demonstrated that on native IgG it is the 6-arm, which is preferentially galactosylated (51), despite the fact that galactosyltransferase preferentially galactosylates the 3-arm of free biantennary glycans *in vitro* (52). This apparent paradox results from the fact that glycans attached to Asn_{297} of IgG are located in a cleft between the two heavy chains (53), which affects their accessibility to glycosyltransferases. Oligosaccharides and the polypeptide chains of the C_H2 domain form multiple noncovalent bonds (54). The majority of these interactions occur between the elongated antennae that extend from the α1–6 linked mannose residues of the trimannosyl cores whereas the 3-arm extends into the interstitial space between the C_H2 domains and is therefore less accessible to glycosyltransferases (42). Moreover, the addition of a galactose to the 3-arm does not directly affect the accessibility of the 6-arm allowing the G1 glycan to be converted to G2, whereas the addition of galactose to the 6-arm reduces accessibility of the 3-arm and prevents further processing (55).

Sialylation—Sialylation of IgG recently attracted much attention after it was shown that it is responsible for the anti-inflammatory activity of intravenously administered immunoglobulins (11, 56). IgG antibodies have long been recognized as proinflammatory mediators of the humoral immune response. Appropriate glycosylation of Asn_{297} is essential for the proinflammatory activity of IgG antibodies by maintaining the heavy chains in conformation, which favors binding to Fcγ receptors (57). Enzymatic removal of this glycan significantly reduces FcγR binding and, consequently, the proinflammatory activity *in vivo* (58). However, if the glycan is sialylated, the proinflammatory effect of IgG is reversed and it now exerts anti-inflammatory effects (56), most probably through interactions with the lectin receptor SIGN-R1 or DC-SIGN (12). This has significant implications for both the normal function of IgG in the immune response and the application of intravenous immunoglobulins for the treatment of a number of autoimmune diseases (59).

Approximately 10% of IgG glycans were found to terminate in sialic acid in a number of early studies (57). We found that in our population the percentage of sialylated glycans was significantly higher with median values being ~20% in core fucosylated glycans without bisecting GlcNAc (Table II: FGS/(F+FG+FGS)) and 29.5% in core fucosylated glycans with bisecting GlcNAc (Table II: FBGS/(FB+FBG+FBGS)). The most probable cause for this apparent difference is the improvement in methodology, which results in decreased hy-

TABLE I

Composition of the IgG glycome. IgG glycome was separated into 24 chromatographic peaks by hydrophilic interaction chromatography. Structures of glycans in each chromatographic peak and the average percentage of individual structures (%) were determined by mass spectrometry. Structure abbreviations: all N-glycans have core sugar sequence consisting of two N-acetylglucosamines (GlcNAc) and three mannose residues; F indicates a core fucose α1–6 linked to the inner GlcNAc; Mx, number (×) of mannose on core GlcNAcs; Ax, number of antenna (GlcNAc) on trimannosyl core; A2, biantennary glycan with both GlcNAcs as β1–2 linked; B, bisecting GlcNAc linked β1–4 to β1–3 mannose; Gx, number of β1–4 linked galactose (G) on antenna; [3]G1 and [6]G1 indicates that the galactose is on the antenna of the α1–3 or α1–6 mannose; Sx, number (×) of sialic acids linked to galactose. Structural schemes are given in terms of N-acetylglucosamine (square), mannose (circle), fucose (rhomb with a dot), galactose (rhomb) and sialic acid (star)

Glycan peak	Peak composition	Structure	%
GP1	F(6)A1		100
GP2	A2		100
GP3	A2B		100
GP4	F(6)A2		100
GP5	M5		63
	F(6)A2		37
GP6	F(6)A2B		97
	A2[6]G1		3
GP7	A2[3]G1		75
	F(6)A2B		25
GP8a	A2BG1		93
	F(6)A2[6]G1		7
GP8b	F(6)A2[6]G1		100
GP9	F(6)A2[3]G1		100
GP10	F(6)A2[6]BG1		100
GP11	F(6)A2[3]BG1		100
GP12	A2G2		91
	F(6)A2[3]BG1		9
GP13	A2BG2		87
	F(6)A2G2		13
GP14	F(6)A2G2		100

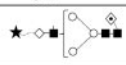
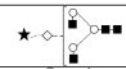
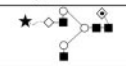
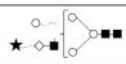
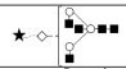
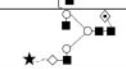
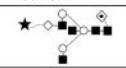
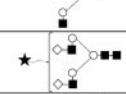
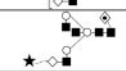
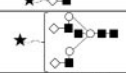
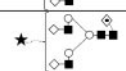
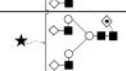
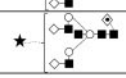
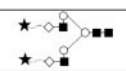
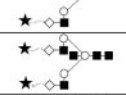
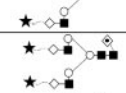
Glycan peak	Peak composition	Structure	%
GP15	F(6)A2BG2		83
	F(6)A1G1S1		8
	A2G1S1		5
	F(6)A2G2		4
GP16a	F(6)A2[6]G1S1		63
	M4A1G1S1		25
	A2BG1S1		13
GP16b	F(6)A2[3]G1S1		91
	F(6)A2[6]BG1S1		9
GP17	A2G2S1		89
	F(6)A2[3]BG1S1		11
GP18a	A2BG2S1		91
	F(6)A2G2S1		9
GP18b	F(6)A2G2S1		100
GP19	F(6)A2BG2S1		100
GP20	n.d.		/
GP21	A2G2S2		100
GP22	A2BG2S2		100
GP23	F(6)A2G2S2		100
GP24	F(6)A2BG2S2		100

TABLE II

Descriptives of the IgG glycome in three populations. IgG glycans were analyzed for individuals from the Croatian Adriatic islands Vis (n = 915), Korčula (n = 906) and Orkney (n = 477). Median values, interquartile ranges (IQR), and minimal and maximal values are shown. Calculations and descriptions of derived glycosylation traits are explained in the Experimental Procedures section. An extended list of derived glycan features is available in supplementary Table S3

IgG glycan	Population: Vis		Population: Korčula		Population: Orkney	
	Median (IQR)	Min-Max	Median (IQR)	Min-Max	Median (IQR)	Min-Max
GP1	0.16 (0.19)	0.02–2.38	0.15 (0.12)	0.03–1.02	0.16 (0.1)	0.04–1.24
GP2	0.74 (0.55)	0.11–5.2	0.72 (0.57)	0.14–9.39	0.71 (0.53)	0.17–4.59
GP4	20.39 (8.61)	6.56–42.37	20.05 (8.54)	6.07–41.34	20.14 (8.17)	8.08–49.47
GP5	0.31 (0.13)	0.09–0.97	0.29 (0.14)	0.12–0.97	0.27 (0.09)	0.13–0.84
GP6	5.24 (2.15)	1.93–12.86	5.5 (2.2)	1.79–11.09	4.82 (2.11)	1.95–10.65
GP7	0.67 (0.52)	0.14–4.01	0.68 (0.46)	0.15–4.06	0.48 (0.31)	0.12–1.72
GP8	16.42 (2.61)	9.84–23.7	16.2 (2.53)	8.93–23.56	18.08 (2.21)	11.89–25.4
GP9	7.9 (1.57)	4.27–11.8	7.95 (1.69)	4.6–12.19	8.98 (1.59)	5.01–12.5
GP10	4.59 (1.3)	2.48–10.2	4.6 (1.22)	2.59–8.19	4.48 (1.11)	2.64–13.37
GP11	0.75 (0.2)	0.33–2.92	0.76 (0.2)	0.4–1.82	0.77 (0.19)	0.44–1.41
GP12	0.96 (0.64)	0.26–6.14	0.98 (0.69)	0.23–3.91	0.8 (0.5)	0.21–3.87
GP13	0.25 (0.16)	0.08–1.59	0.22 (0.07)	0.1–1.01	0.24 (0.07)	0.13–0.58
GP14	11.02 (5.06)	3.39–22.97	10.77 (4.83)	3.5–24.81	12.26 (5.07)	3.66–23.91
GP15	1.41 (0.48)	0.71–2.89	1.51 (0.46)	0.68–2.81	1.64 (0.45)	0.77–2.83
GP16	3.13 (0.68)	1.99–5.11	3.21 (0.61)	1.82–5.22	3.21 (0.61)	1.69–5.25
GP17	2.61 (1.85)	1–13.61	2.43 (1.67)	0.94–10.45	1.58 (0.53)	0.92–7.73
GP18	8.58 (3.65)	3.96–18.02	8.83 (3.52)	3.58–26.02	9.33 (3.67)	3.45–18.92
GP19	2.45 (0.7)	1.16–9.03	2.41 (0.63)	1.15–5.1	2.38 (0.48)	1.3–4.54
GP20	0.43 (0.43)	0.07–3.25	0.62 (0.49)	0.11–3.46	0.49 (0.22)	0.25–2.56
GP21	3.41 (2.38)	1.03–23.82	3.2 (2.83)	0.98–24.6	1.56 (0.68)	0.67–6.01
GP22	0.36 (0.21)	0.06–1.33	0.31 (0.13)	0.08–1.46	0.29 (0.12)	0.06–0.88
GP23	1.98 (0.79)	0.71–13.73	2.27 (0.83)	0.9–4.67	2.3 (0.89)	0.88–4.62
GP24	2.7 (0.79)	1.02–7.66	2.72 (0.78)	0.67–5.84	2.66 (0.71)	0.97–8.86
FGS/(FG+FGS)	28.15 (4.52)	17.59–47.45	29.5 (4.54)	19.02–42.8	27.32 (4.09)	19.63–37.27
FBGS/(FBG+FBGS)	43.14 (8.86)	14.07–74.67	43.05 (7.84)	21.95–60.72	41.96 (7.16)	12.06–66.91
FGS/(F+FG+FGS)	19.71 (6.03)	9.67–42.34	20.82 (5.82)	9.63–39.62	19.99 (5.39)	9.14–32.28
FBGS/(FB+FBG+FBGS)	29.91 (7.99)	10.16–65.45	29.38 (7.64)	14.17–48.83	29.70 (7.36)	10.37–51.19
FG1S1/(FG1+FG1S1)	11.49 (2.59)	7.38–22.13	11.79 (2.33)	7.21–22.05	10.67 (2.02)	5.99–15.84
FG2S1/(FG2+FG2S1+FG2S2)	39.99 (3.28)	32.05–49.45	40.72 (3.22)	33.24–50.07	39.16 (3.26)	25.21–48.48
FG2S2/(FG2+FG2S1+FG2S2)	8.99 (3.3)	3.61–33.21	10.08 (3.44)	2.18–30.64	9.45 (2.79)	4.18–20.31
FBG2S1/(FBG2+FBG2S1+FBG2S2)	37.2 (4.73)	26.27–49.34	36.2 (4.66)	25.86–46.49	35.92 (4)	26.32–44.75
FBG2S2/(FBG2+FBG2S1+FBG2S2)	41.09 (5.92)	23.09–59.58	40.73 (5.77)	15.88–54.25	39.48 (4.59)	18.91–62.28
FBS1/(FS1+FBS1)	0.17 (0.05)	0.08–0.52	0.17 (0.05)	0.04–0.36	0.16 (0.05)	0.07–0.31
FBS2/(FS2+FBS2)	0.58 (0.08)	0.36–0.75	0.55 (0.08)	0.37–0.78	0.54 (0.08)	0.32–0.78
$G0^n$	37.09 (12.75)	14.69–69.57	37.22 (11.66)	13.84–62.82	34.62 (11.46)	16.17–69.42
$G1^n$	42.6 (4.54)	24.18–50.96	42.73 (4.38)	29.59–49.26	44.45 (3.78)	24.73–60.27
$G2^n$	19.17 (8.83)	5.93–43.76	18.88 (8.24)	7.05–47.09	20 (8.47)	5.64–40.27
$F^{n\ total}$	95.77 (2.32)	84.19–98.64	95.76 (2.37)	80.7–98.64	96.52 (1.8)	88.16–98.67
$FG0^{n\ total}/G0^n$	97.11 (1.95)	81.33–99.59	97.26 (2.09)	77.55–99.52	97.15 (1.89)	86.42–99.41
$FG1^{n\ total}/G1^n$	97.78 (1.67)	87.05–99.57	97.76 (1.57)	86.77–99.54	98.52 (0.91)	94.55–99.6
$FG2^{n\ total}/G2^n$	91.03 (5.05)	72.7–97.35	90.88 (4.32)	73.66–98.08	92.95 (3.37)	80.66–97.51
F^n	78.3 (4.62)	63.83–87.48	77.99 (4.72)	60.95–86.52	80.44 (4.05)	67.55–88.92
$FG0^n/G0^n$	77.02 (5.22)	49.91–86.55	76.18 (5.93)	53.13–86.54	78.3 (5.24)	59.25–88.33
$FG1^n/G1^n$	79.5 (4.5)	66.38–87.63	79.68 (4.46)	67.99–87.69	82.28 (4.1)	68.72–90.2
$FG2^n/G2^n$	80.42 (5.74)	57.71–90.2	79.48 (5.58)	60.13–89.57	81.59 (5.28)	63.28–90.49
FB^n	16.92 (3.45)	10.87–25.9	17.27 (3.61)	10.49–26.43	15.74 (3.37)	9–26.36
$FBG0^n/G0^n$	19.85 (4.57)	12.16–33.2	20.64 (4.7)	11.66–36.83	18.74 (4.14)	10.51–32.08
$FBG1^n/G1^n$	17.78 (3.91)	11.1–28.42	17.74 (4)	10.73–28.46	16.17 (4.07)	8.85–30.76
$FBG2^n/G2^n$	10.27 (2.39)	6.36–30.92	11.13 (3.06)	5.78–24.33	10.95 (3.06)	5.72–20.94
$FB^n/F^{n\ total}$	17.76 (3.72)	11.05–27.64	18.18 (3.94)	10.85–27.81	16.35 (3.51)	9.23–26.91

drolysis of sialic acids during the purification of IgG and glycan analysis. Both rapid elution and neutralization enabled by the use of newly developed protein G monolithic plates, and the replacement of hydrazide (used in early studies) with PNGase F contributed to the better preservation of sialic acid. However, it should be noted that some previous studies analyzed only heavy chains of IgG, thus the Fab light chain glycans (which are generally more sialylated) bound to light chains of IgG were excluded from the analysis. In addition, a recent interlaboratory comparison of glycan analysis methods revealed that mass spectrometry routinely underestimated sialic acid on IgG by nearly 3-fold, whereas HPLC analysis reported sialylation levels very similar to the sialylation levels observed in our populations (60). However, this interlaboratory comparison was performed on a small number of samples, and our study is the first large scale study to reported

TABLE III

Heritability of the IgG glycome. Heritability of individual glycans in the population of the Croatian Adriatic island Vis was estimated using polygenic models in Sequential Oligogenic Linkage Analysis Routines (SOLAR)

IgG glycan	Covariates significance					
	H2	P (H2)	S.E. (H2)	Gender	Age	R^{2a}
GP1	26.8%	0.010	0.117	0.771	2.00E-02	0.7%
GP2	26.1%	0.023	0.134	0.411	6.63E-24	11.1%
GP4	16.9%	0.061	0.114	0.128	2.52E-81	33.7%
GP5	0.0%	0.500	-	0.001	5.66E-01	1.3%
GP6	32.8%	0.001	0.113	0.218	4.68E-66	28.0%
GP7	13.0%	0.128	0.118	0.017	2.14E-01	0.7%
GP8	29.7%	0.005	0.116	0.271	2.60E-10	4.7%
GP9	29.1%	0.009	0.127	0.007	2.36E-01	1.0%
GP10	43.3%	0.000	0.118	0.037	2.18E-01	0.5%
GP11	17.0%	0.080	0.126	0.009	1.12E-14	7.2%
GP12	10.9%	0.170	0.118	0.049	3.65E-12	5.6%
GP13	1.5%	0.445	0.108	0.383	3.30E-02	0.6%
GP14	44.5%	0.000	0.112	0.040	4.85E-72	30.5%
GP15	36.2%	0.007	0.113	0.001	1.62E-30	14.9%
GP16	29.8%	0.008	0.127	0.074	7.30E-02	0.7%
GP17	27.9%	0.003	0.105	0.005	7.50E-02	1.2%
GP18	47.8%	0.000	0.112	0.329	7.86E-61	26.3%
GP19	40.3%	0.000	0.123	0.000	6.00E-02	2.2%
GP20	0.0%	0.500	-	0.001	8.46E-01	1.3%
GP21	19.3%	0.055	0.124	0.004	8.90E-03	1.8%
GP22	0.0%	0.500	-	0.309	6.12E-01	0.1%
GP23	17.9%	0.063	0.121	0.996	3.03E-11	4.8%
GP24	46.0%	0.000	0.113	0.008	1.22E-01	1.1%
FGS/(FG+FGS)	19.8%	0.052	0.126	0.099	7.74E-09	4.0%
FBGS/(FBG+FBGS)	39.7%	0.002	0.131	0.079	5.00E-03	1.4%
FGS/(F+FG+FGS)	21.7%	0.027	0.117	0.503	5.14E-58	25.4%
FBGS/(FB+FBG+FBGS)	11.5%	0.183	0.129	0.376	1.17E-07	3.2%
FG1S1/(FG1+FG1S1)	14.4%	0.118	0.123	0.740	1.75E-06	2.5%
FG2S1/(FG2+FG2S1+FG2S2)	40.2%	0.002	0.134	0.107	2.29E-02	0.8%
FG2S2/(FG2+FG2S1+FG2S2)	42.4%	0.000	0.121	0.443	2.03E-23	10.4%
FBG2S1/(FBG2+FBG2S1+FBG2S2)	47.9%	0.000	0.129	0.022	4.24E-19	9.3%
FBG2S2/(FBG2+FBG2S1+FBG2S2)	61.3%	0.000	0.110	0.810	2.30E-11	4.9%
FBS1/(FS1+FBS1)	56.4%	0.000	0.119	0.000	1.31E-47	22.1%
FBS2/(FS2+FBS2)	42.5%	0.000	0.122	0.000	5.64E-28	13.9%
$G0^n$	35.0%	0.001	0.118	0.779	2.96E-88	35.7%
$G1^n$	26.7%	0.014	0.124	0.948	7.74E-29	13.1%
$G2^n$	42.0%	0.000	0.109	0.009	1.02E-70	30.2%
$F^{n\ total}$	37.0%	0.001	0.118	0.595	2.00E-01	0.1%
$FG0^{n\ total}/G0^n$	47.2%	0.000	0.115	0.530	8.19E-01	0.1%
$FG1^{n\ total}/G1^n$	32.0%	0.004	0.123	0.726	4.39E-01	0.1%
$FG2^{n\ total}/G2^n$	29.6%	0.005	0.116	0.814	1.93E-07	3.2%
F^n	45.4%	0.000	0.115	0.129	9.00E-05	2.3%
$FG0^n/G0^n$	45.4%	0.001	0.115	0.162	2.15E-01	0.3%
$FG1^n/G1^n$	43.1%	0.000	0.124	0.006	3.00E-04	2.8%
$FG2^n/G2^n$	13.6%	0.122	0.120	0.268	2.76E-23	10.7%
FB^n	33.3%	0.002	0.118	0.001	6.44E-18	8.8%
$FBG0^n/G0^n$	29.8%	0.008	0.126	0.000	3.15E-01	1.6%
$FBG1^n/G1^n$	32.5%	0.003	0.124	0.000	1.25E-07	4.7%
$FBG2^n/G2^n$	47.3%	0.000	0.123	0.001	1.83E-01	1.3%
$FB^n/F^{n\ total}$	29.6%	0.006	0.120	0.000	1.84E-16	8.4%

[a] Percent of variance explained in the model.

such a high level of IgG sialylation. A more extensive international interlaboratory study of glycan quantification concluded that permethylation of glycans can enable reliable quantification of sialic acid on transferrin and IgG (61).

The variability of sialylation (of galactosylated glycans) in the population was also found to be rather high, ranging from 17.6% to 47.5% in IgG without bisecting GlcNAc (Table II: FGS/(FG+FGS)), and between 12.1% and 74.7% in IgG with bisecting GlcNAc (Table II: FBGS/(FBG+FBGS)). Because our study participants came from a cross section study, it was not possible to clearly identify whether high sialylation in some individuals was a genetic predisposition or a temporary physiological condition. However, the extent of sialylation was found to be the most heritable element of IgG glycosylation (Table III) indicating strong genetic regulation. This was further supported by relatively small changes in sialylation observed in several individuals that were sampled a second time, after a period of ~12 months (supplemental Fig. S2).

The majority of IgG glycans are attached to Asn_{297} in the constant region of the heavy chain, but 15–20% of IgG also contain glycans on the variable regions and some studies indicated that these glycans are more sialylated, and particularly more bisialylated than Fc glycans (49, 62, 63). The most comprehensive comparison of Fc and Fab IgG glycans was unfortunately performed on only two control individuals and the difference of Fc sialylation in these two patients was rather large (49). However, in one of them over 70% of bigalactosylated glycans with both core fucose and bisecting GlcNAc were sialylated, what is in line with our observations. Whether the observed variability of IgG sialylation originates from Fab or Fc glycans is not known, but because individual differences in Fc sialylation could have strong effects on the inflammatory response, this issue should be addressed in future studies.

Fucosylation—Core fucosylation of IgG has been intensively studied because of its role in ADCC. Natural killer cells have receptors for the Fc domain of IgG. They bind to the Fc portion of IgG antibodies on the surface of target cells, such as tumor cells, and release cytolytic components that kill the target cell. This mechanism of killing is considered to be the major mechanism of antibody-based therapeutics against tumors. Core fucose is very important in this process because IgG deficient in core fucose on the Fc glycan was found to have ADCC activity enhanced by up to 100-fold (64). Endogenous serum IgG inhibits therapeutic antibody-induced ADCC by competing for FcγRIIIa binding sites (17), but nonfucosylated therapeutic IgG was reported to be able to evade this through higher affinity FcγRIIIa binding (16). However, this might be highly dependent on the extent of core fucosylation of host IgGs. In our populations the fraction of neutral IgG glycans without the core fucose was found to vary between 1.3% and 19.3% (Table II; $F^{n\ total}$) and this large variability could affect the efficacy of therapeutic antibodies. Individuals with lower core fucose might need higher doses of the drug, thus the extent of host IgG core fucosylation may need to be one of the parameters in the calculation of the exact therapeutic dose.

Bisecting GlcNAc—The addition of bisecting GlcNAc to the IgG glycan by GlcNAc transferase III (GnTIII) significantly affects its accessibility to other glycosyltransferases. For example, the addition of bisecting GlcNAc prevents further branching because glycans with bisecting GlcNAc are not a substrate for GnTs IV, V, and VI (65). Some studies also indicate that the presence of bisecting GlcNAc diminishes galactosylation by GalT (66) and the addition of core fucose (67).

In our study, on average, ~18% of neutral glycans contained bisecting GlcNAc (Table II: $FB^n/F^{n\ total}$). However, when the percentages of bisecting GlcNAc in $G0^n$, $G1^n$, and $G2^n$ structures were compared, a significant decrease in the percentage of bisecting GlcNAc (nearly 50%) in $G2^n$ structures was observed (Table II: $FBG0^n/G0^n$, $FBG1^n/G1^n$, and $FBG2^n/G2^n$). At first sight, this seemed to confirm the results of transfection assays, which indicated that GnTIII and GalT compete for an agalactosyl nonbisected biantennary sugar chain as a common substrate (66). However, after considering this in the context of not only neutral, but also sialylated glycans, the presence of a bisecting GlcNAc did not seem to have a significant effect on IgG galactosylation. As presented in Table II and discussed in the "sialylation" section, whereas only slightly below 10% of FA2G2 structures contained two sialic acids, over 40% of FA2BG2 structures contained two sialic acids. Consequently, the decrease in bisecting GlcNAc in neutral bigalactosylated structures (from 18% in mono galactosylated to 11% in digalactosylated) was compensated by the increase in bisecting GlcNAc in bisialylated digalactosylated glycans (from 17% in mono sialylated to over 55% in disialylated digalactosylated glycans). Therefore, the percentage of digalactosylated structures with bisecting GlcNAc was approximately the same as for agalactosylated, and monogalactosylated structures. The presence of bisecting GlcNAc apparently associated with the increased addition of a second sialic acid, and consequently bisecting GlcNAc was underrepresented in neutral digalactosylated and overrepresented in sialylated digalactosylated glycan pools.

Interactions Between Different Functional Elements of IgG Glycosylation—Mutual interactions of the four above described functional elements of IgG glycosylation have been addressed in the past, but different experimental approaches (transfections, enzymatic assays, etc) often yielded conflicting results. Our study is the first large-scale detailed analysis of IgG glycosylation in well-characterized human populations, which has therefore enabled reliable conclusions about the interrelation of different functional elements. However, it is important to note that *in vivo* differences in the ratios of different glycosylation elements in IgG may not simply reflect the kinetic characteristics of relevant glycosyltransferases, but may result from regulatory elements, which govern IgG glycosylation and modulate their glycans, presumably according to the prevailing functional needs of the host organism. Only the most interesting associations are presented below, whereas the complete set of correlation coefficients and their p values is available as supplemental material (supplemental Table S4).

Recently we reported the existence of a negative correlation between galactosylation and sialylation in the human plasma glycome (68 + unpublished observations). A similar association was also found in the IgG glycome where the percentage of bigalactosylated structures in the neutral IgG glycome strongly negatively correlated with the percentage of disialylated glycans ($r = -0.39$, $p = 7.3E\text{-}43$; supplemental Fig. S3). One interpretation of this observation is that sialylation might be a rate-limiting step and that smaller amount of bigalactosylated structures can be more efficiently sialylated. The same observation could also be a reflection of the simple fact that bigalactosylated glycan is a substrate for sialyl-transferase and that efficient sialylation is decreasing

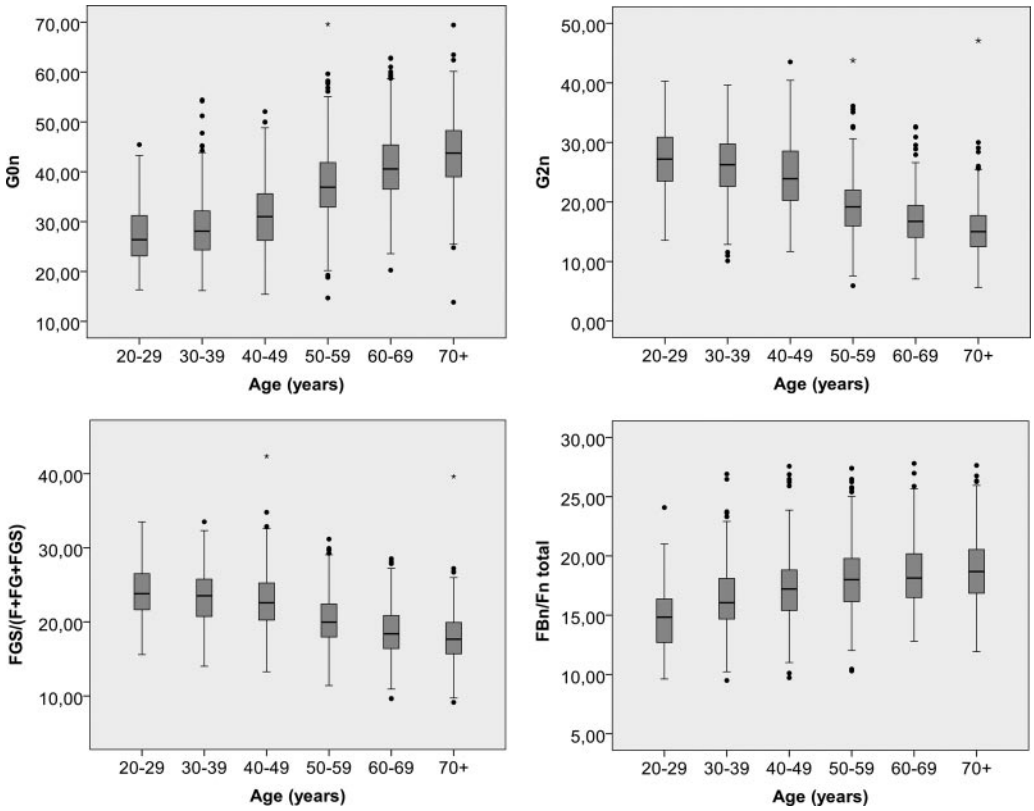


FIG. 2. **Association of IgG glycosylation with age.** Distribution of G0^n glycans, G2^n glycans, the percent of structures with sialic acid (FGS/(F+FG+FGS)) and bisecting GlcNAc (FB^n/F$^{n\ total}$) in fucosylated glycans between different age-groups are shown. Central box represents the values from the lower to upper quartile (25 to 75 percentile). The middle line represents the median. The horizontal line extends from the minimum to the maximum value, excluding "outside" and "far out" values that are displayed as separate points.

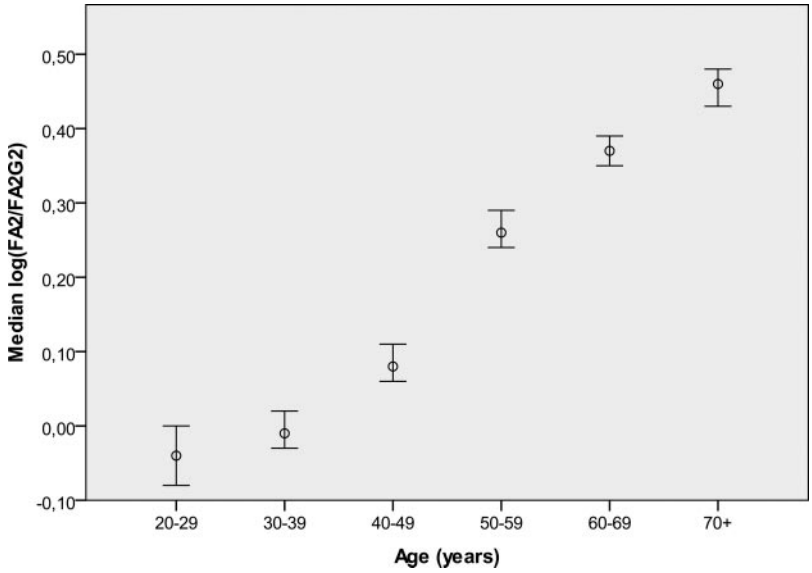


FIG. 3. **The Glyco-Age index.** Glyco-Age index calculated as the logarithm of the ratio of fucosylated G2 and G0 structures (FA2/FA2G2) was recently suggested to be a good indicator of individual's age (78). Median values of the Glyco-Age index (with 95% confidence intervals as error bars) in our study population are shown.

the level of bigalactosylated glycans in the neutral glycome by converting them into charged monosialylated and disialylated digalactosylated glycans. However, in that case monosialylated glycans should also negatively correlate with the percentage of digalactosylated glycans, what was not the case (supplemental Table S4).

Bisecting GlcNAc was reported to negatively affect both galactosylation (66) and fucosylation of IgG (67). A moderate negative ($r = -0.27$), but statistically highly significant ($p = 1.75E-20$) correlation was observed between bisecting GlcNAc (FB^n/F^n total) and core fucosylation of IgG (F^n total), which confirmed the results from *in vitro* studies. However, in our study we did not observe any negative correlation between the percentage of bisecting GlcNAc and galactosylation, indicating that *in vivo* there is no competition between GnTIII and GalT. GnTIII, which adds bisecting GlcNAc, and GalT, which adds galactoses clearly compete for the same substrate (66). The fact that this is not happening *in vivo* is a very important observation because it demonstrates that final glycan structures are not a simple result of competing enzymatic activities, but a predesigned outcome, presumably fitted to the needs of the producing cells by some still unknown regulatory mechanisms. Confirmation of this interpretation was provided by the comparison of Fab and Fc glycans in different myeloma lines, which revealed strong site-specific regulation of glycosylation (69).

The most surprising observation was the large increase in the proportion of bisecting GlcNAc in disialylated structures (FA2BG2S2, GP24), or alternatively the increase in sialylation of FA2BG2 glycans. Although ~18% of all IgG Glycans (Table II: $FB^n/F^{n\ total}$) contained bisecting GlcNAc, over 58% of all disialylated glycans contained bisecting GlcNAc (supplemental Table S3: FBS2/(FBS+FBS2). Because FA2BG2S2 structure (GP24) represents only 2.7% of the total glycome, an additional check was performed to confirm that this structure did not originate from contaminating plasma proteins. In the total serum glycome, the ratio of A2G2S2 and FA2BG2S2 structures is ~50:1 (70). Because the median value of A2G2S2 (GP21) in our population was 3.3, even if all of it originated from contaminating plasma proteins, the amount of FA2BG2S2 coming from the plasma proteins could be only up to 0.06%, which is far below the level that could significantly contribute to the observed increase in the proportion of bisecting GlcNAc in disialylated IgG glycans.

Molecular modeling of IgG Fc with a bisecting GlcNAc on the Asn 297 glycan shows that the bisect can be accommodated in a low energy linkage conformation with the rest of the glycan remaining in its crystallographic position (*i.e.* with the 6-arm bound to the surface of the protein). However, the N-acetyl group of the bisect is oriented close to the 6-arm GlcNAc and Phe 243 and may alter or disrupt the hydrophobic stacking between these two residues. This could result in the glycan 6-arm interacting less strongly with the protein surface, making the entire glycan more mobile and both 3- and 6-arms more likely to be available to glycosyltransferases.

Variability and Heritability of IgG Glycosylation—In our recent analysis of the total plasma glycome in 915 individuals from the Croatian island of Vis we observed a median ratio of minimal to maximal values of 6.17 and significant age- and gender-specific differences (38). This study (which also includes the Vis cohort) analyzed only the IgG glycome and revealed even higher variability in the population with median ratio of minimal to maximal values of 17.2 for the whole IgG glycome and 19.7 for the neutral IgG glycome. The variability of neutral glycans primarily originated from various percentage of bisecting GlcNAc and core fucose on different neutral structures that represent a relatively small part of the glycome, but even in some main glycosylation features the variability was rather high. For example the proportion of G0 ranged between 14 and 70% and the proportion of G2 between 6 and 47% (Table II). Bisecting GlcNAc and core fucose also varied significantly, but the most variable was the sialylation of different glycan structures. The disialylated form of FA2G2 varied between 2.2 and 33% (Table II: FG2S2/(FG2+FG2S1+FG2S2)), whereas the disialylated form of FA2BG2 varied between 16 and 60% (Table II: FBG2S2/(FBG2+FBG2S1+FBG2S2)).

Variations observed in a human phenotype are generally a combination of genetic differences and environmental factors. Heritability is one of the most basic and often one of the first analyses to be made in a genetic study, because it represents the proportion of the trait variance that can be attributed to genetic factors, and it is often used as a screening tool to determine whether a trait may be suitable for gene mapping (71). The fact that there are hundreds of genes involved in the complex glycan metabolic pathways argues in favor of a strong genetic influence, but environmental effects on glycan structures have also been reported (72–75). Recently we reported that there is a broad range of variation in heritability levels of plasma glycans, from insignificant or very low to over 50% for some glycans (38).

Using the large number of known genealogical relationships in our isolated populations we were able to reliably estimate heritability of IgG glycans (Table III). Sialylation appeared to be the most endogenously defined glycosylation feature, with up to 60% of variance explained by heritability. Age had very little effect on the extent of sialylation, whereas gender was nearly irrelevant. Contrary to the total plasma glycome, where gender played an important role in many glycans (68), gender was a significant predictor in the IgG glycome, although only for some structures containing bisecting GlcNAc. For the incidence of core fucose, bisecting GlcNAc and galactose, the heritability was generally between 25 and 45%, indicating that a significant part of the variability of IgG glycosylation can be explained by genetic polymorphisms.

Effects of Gender and Age—A number of strong associations between IgG glycans and age were observed and the most prominent ones are shown in Fig. 2. Individual's age was the most important predictor for the level of IgG galactosylation, with 35% of variance of G0 explained by age (Table III). This observation is in accordance with previous studies of galactosylation in aging (76, 77) and our recent study of the total plasma glycome (68). The proportion of structures with bisecting GlcNAc also increased with age, what is in accord-

ance with the results of the recent large study of IgG glycans in aging (77). However, contrary to the previous observations in plasma, the extent of core fucosylation of IgG was not found to change with age (data not shown), indicating that the observed decrease in core fucose with age in the total plasma glycome of women (68) could be the consequence of decreased IgG concentration, and not alterations in the glycosylation metabolism in plasma cells. Another difference between effects of age on glycosylation of IgG and total plasma proteins was observed in sialylation. Although in the total plasma glycome sialylation did not change with age (68), in IgG glycans there was a significant decrease of sialylation of IgG with age (Fig. 2: FGS/(F + FG + FGS)). Very recently, the logarithm of the ratio of fucosylated G0 to G2 structures (FA2/FA2G2), the so called GlycoAge index, was suggested as a good indicator of individual's age (78). In our population this index was also a reliable predictor of age with good separation of individuals in different decades of life (Fig. 3).

Acknowledgments—Plasma samples used for the isolation of IgG were collected with support of the FP6 EuroSpan project, the Royal Society, the Chief Scientist Office of the Scottish Government. We thank Carolien Koeleman for expert technical assistance, Lorraine Anderson for data collection, Dr. Eoin Cosgrave and Dr. Jonathan Bones for critical reading of the manuscript and the people of Orkney, Vis, and Korčula for volunteering.

* This work was supported by the Croatian Ministry of Science, Education, and Sport (grants #309-0061194-2023 and #216-1080315-0302), by the Croatian Science Foundation (grant # 04-47), by the European Commission (grants EuroGlycoArrays and GlycoBioM), and by the grant #P20RR017695 from the National Institute of Health, Bethesda, Maryland. In addition, M.P. gratefully acknowledges financial support of the Federation of European Biochemical Societies for her study visit to Biomolecular Mass Spectrometry Unit in Leiden and A.K. acknowledges the support of the "Unity Through Knowledge fund."

[c] To whom correspondence should be addressed: University of Zagreb, Faculty of Pharmacy and Biochemistry, A. Kovačića 1, 10000 Zagreb, Croatia. Fax: +385-1-639-4400; E-mail: glauc@pharma.hr.

REFERENCES

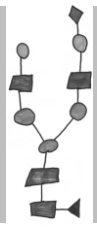
1. Marek, K. W., Vijay, I. K., and Marth, J. D. (1999) A recessive deletion in the GlcNAc-1-phosphotransferase gene results in peri-implantation embryonic lethality. *Glycobiology* **9,** 1263–1271
2. Dennis, J. W., Lau, K. S., Demetriou, M., and Nabi, I. R. (2009) Adaptive Regulation at the Cell Surface by N-Glycosylation. *Traffic* **10,** 1569–1578
3. Crocker, P. R., Paulson, J. C., and Varki, A. (2007) Siglecs and their roles in the immune system. *Nat. Rev. Immunol.* **7,** 255–266
4. Marth, J. D., and Grewal, P. K. (2008) Mammalian glycosylation in immunity. *Nat. Rev. Immunol.* **8,** 874–887
5. Ohtsubo, K., and Marth, J. D. (2006) Glycosylation in cellular mechanisms of health and disease. *Cell* **126,** 855–867
6. Skropeta, D. (2009) The effect of individual N-glycans on enzyme activity. *Bioorg. Med. Chem.* **17,** 2645–2653
7. Kobata, A. (2008) The N-linked sugar chains of human immunoglobulin G: their unique pattern, and their functional roles. *Biochim. Biophys. Acta* **1780,** 472–478
8. Jefferis, R. (2005) Glycosylation of recombinant antibody therapeutics. *Biotechnol. Prog.* **21,** 11–16
9. Zhu, D., Ottensmeier, C. H., Du, M. Q., McCarthy, H., and Stevenson, F. K. (2003) Incidence of potential glycosylation sites in immunoglobulin variable regions distinguishes between subsets of Burkitt's lymphoma and mucosa-associated lymphoid tissue lymphoma. *Br. J. Haematol.* **120,** 217–222
10. Parekh, R. B., Dwek, R. A., Sutton, B. J., Fernandes, D. L., Leung, A., Stanworth, D., Rademacher, T. W., Mizuochi, T., Taniguchi, T., and Matsuta, K. (1985) Association of rheumatoid arthritis and primary osteoarthritis with changes in the glycosylation pattern of total serum IgG. *Nature* **316,** 452–457
11. Kaneko, Y., Nimmerjahn, F., and Ravetch, J. V. (2006) Anti-inflammatory activity of immunoglobulin G resulting from Fc sialylation. *Science* **313,** 670–673
12. Anthony, R. M., and Ravetch, J. V. (2010) A novel role for the IgG Fc glycan: the anti-inflammatory activity of sialylated IgG Fcs. *J Clin Immunol 30 Suppl* **1,** S9–14
13. Nimmerjahn, F., and Ravetch, J. V. (2008) Fcgamma receptors as regulators of immune responses. *Nat Rev Immunol* **8,** 34–47
14. Ferrara, C., Stuart, F., Sondermann, P., Brünker, P., and Umana, P. (2006) The carbohydrate at FcgammaRIIIa Asn-162. An element required for high affinity binding to non-fucosylated IgG glycoforms. *J. Biol. Chem.* **281,** 5032–5036
15. Shinkawa, T., Nakamura, K., Yamane, N., Shoji-Hosaka, E., Kanda, Y., Sakurada, M., Uchida, K., Anazawa, H., Satoh, M., Yamasaki, M., Hanai, N., and Shitara, K. (2003) The absence of fucose but not the presence of galactose or bisecting N-acetylglucosamine of human IgG1 complex-type oligosaccharides shows the critical role of enhancing antibody-dependent cellular cytotoxicity. *J. Biol. Chem.* **278,** 3466–3473
16. Iida, S., Misaka, H., Inoue, M., Shibata, M., Nakano, R., Yamane-Ohnuki, N., Wakitani, M., Yano, K., Shitara, K., and Satoh, M. (2006) Nonfucosylated therapeutic IgG1 antibody can evade the inhibitory effect of serum immunoglobulin G on antibody-dependent cellular cytotoxicity through its high binding to FcgammaRIIIa. *Clin Cancer Res.* **12,** 2879–2887
17. Preithner, S., Elm, S., Lippold, S., Locher, M., Wolf, A., da Silva, A. J., Baeuerle, P. A., and Prang, N. S. (2006) High concentrations of therapeutic IgG1 antibodies are needed to compensate for inhibition of antibody-dependent cellular cytotoxicity by excess endogenous immunoglobulin G. *Mol. Immunol.* **43,** 1183–1193
18. Barut, M., Podgornik, A., Urbas, L., Gabor, B., Brne, P., Vidic, J., Plevcak, S., and Strancar, A. (2008) Methacrylate-based short monolithic columns: enabling tools for rapid and efficient analyses of biomolecules and nanoparticles. *J. Sep. Sci.* **31,** 1867–1880
19. Josić, D., and Buchacher, A. (2001) Application of monoliths as supports for affinity chromatography and fast enzymatic conversion. *J. Biochem. Biophys. Methods* **49,** 153–174
20. Jungbauer, A., and Hahn, R. (2008) Polymethacrylate monoliths for preparative and industrial separation of biomolecular assemblies. *J. Chromatogr. A* **1184,** 62–79
21. Roberts, M. W., Ongkudon, C. M., Forde, G. M., and Danquah, M. K. (2009) Versatility of polymethacrylate monoliths for chromatographic purification of biomolecules. *J. Sep. Sci.* **32,** 2485–2494
22. Podgornik, A., and Strancar, A. (2005) Convective Interaction Media (CIM)–short layer monolithic chromatographic stationary phases. *Biotechnol. Annu. Rev.* **11,** 281–333
23. Barut, M., Podgornik, A., Brne, P., and Strancar, A. (2005) Convective interaction media short monolithic columns: enabling chromatographic supports for the separation and purification of large biomolecules. *J. Sep. Sci.* **28,** 1876–1892
24. Brne, P., Podgornik, A., Bencina, K., Gabor, B., Strancar, A., and Peterka, M. (2007) Fast and efficient separation of immunoglobulin M from immunoglobulin G using short monolithic columns. *J. Chromatogr. A* **1144,** 120–125
25. Iberer, G., Hahn, R., and Jungbauer, A. (1999) Monoliths as stationary phases for separation of biopolymers: the fourth generation of chromatography sorbents. *LC-GC*, 998–1005
26. Urbas, L., Brne, P., Gabor, B., Barut, M., Strlic, M., Petric, T. C., and Strancar, A. (2009) Depletion of high-abundance proteins from human plasma using a combination of an affinity and pseudo-affinity column. *J. Chromatogr. A* **1216,** 2689–2694
27. Vlakh, E. G., and Tennikova, T. B. (2009) Applications of polymethacrylate-based monoliths in high-performance liquid chromatography. *J. Chromatogr. A* **1216,** 2637–2650
28. Tennikova, T. B., and Reusch, J. (2005) Short monolithic beds: history and introduction to the field. *J. Chromatogr. A* **1065,** 13–17
29. Tetala, K. K., and van Beek, T. A. (2010) Bioaffinity chromatography on

monolithic supports. *J. Sep. Sci.* **33,** 422–438
30. Berruex, L. G., Freitag, R., and Tennikova, T. B. (2000) Comparison of antibody binding to immobilized group specific affinity ligands in high performance monolith affinity chromatography. *J. Pharm. Biomed. Anal.* **24,** 95–104
31. Gupalova, T. V., Lojkina, O. V., Pàlàgnuk, V. G., Totolian, A. A., and Tennikova, T. B. (2002) Quantitative investigation of the affinity properties of different recombinant forms of protein G by means of high-performance monolithic chromatography. *J. Chromatogr. A* **949,** 185–193
32. Burnouf, T., Goubran, H., and Radosevich, M. (1998) Application of bioaffinity technology in therapeutic extracorporeal plasmapheresis and large-scale fractionation of human plasma. *J. Chromatogr. B Biomed. Sci. Appl.* **715,** 65–80
33. Rudan, I., Campbell, H., and Rudan, P. (1999) Genetic epidemiological studies of eastern Adriatic Island isolates, Croatia: objective and strategies. *Coll. Antropol.* **23,** 531–546
34. Rudan, I., Biloglav, Z., Vorko-Jović, A., Kujundzic-Tiljak, M., Stevanović, R., Ropac, D., Puntarić, D., Cucević, B., Salzer, B., and Campbell, H. (2006) Effects of inbreeding, endogamy, genetic admixture, and outbreeding on human health: a (1001 Dalmatians) study. *Croat. Med. J.* **47,** 601–610
35. Campbell, H., Carothers, A. D., Rudan, I., Hayward, C., Biloglav, Z., Barac, L., Pericic, M., Janicijevic, B., Smolej-Narancic, N., Polasek, O., Kolcic, I., Weber, J. L., Hastie, N. D., Rudan, P., and Wright, A. F. (2007) Effects of genome-wide heterozygosity on a range of biomedically relevant human quantitative traits. *Hum. Mol. Genet.* **16,** 233–241
36. Tennikova, T. B., Belenkii, B. G., and Švec, F. (1990) *J. Liq. Chromatogr.* **13,** 63–70
37. Barut, M., Podgornik, A., Merhar, M., Štrancar, A. F. Švec, T. B. Tennikova, and Z. Deyl, (2003) (Eds.), Monolithic Materials: Preparation, Properties and Applications, *Journal of Chromatography Library*, vol. 67 p. 51, Elsevier, Amsterdam.
38. Knežević, A., Polašek, O., Gornik, O., Rudan, I., Campbell, H., Hayward, C., Wright, A., Kolčić, I., O'Donoghue, N., Bones, J., Rudd, P. M., and Lauc, G. (2009) Variability, Heritability and Environmental Determinants of Human Plasma N-Glycome. *J. Proteome Res.* **8,** 694–701
39. Ahn, J., Bones, J., Yu, Y. Q., Rudd, P. M., and Gilar, M. (2010) Separation of 2-aminobenzamide labeled glycans using hydrophilic interaction chromatography columns packed with 1.7 microm sorbent. *J Chromatogr B Analyt Technol Biomed Life Sci* **878,** 403–408
40. Bones, J., Mittermayr, S., O'Donoghue, N., Guttman, A., and Rudd, P. M. (2010) Ultra performance liquid chromatographic profiling of serum N-glycans for fast and efficient identification of cancer associated alterations in glycosylation. *Anal Chem* **82,** 10208–10215
41. Ceroni, A., Maass, K., Geyer, H., Geyer, R., Dell, A., and Haslam, S. M. (2008) GlycoWorkbench: a tool for the computer-assisted annotation of mass spectra of glycans. *J. Proteome Res.* **7,** 1650–1659
42. Krapp, S., Mimura, Y., Jefferis, R., Huber, R., and Sondermann, P. (2003) Structural analysis of human IgG-Fc glycoforms reveals a correlation between glycosylation and structural integrity. *J. Mol. Biol.* **325,** 979–989
43. Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N., and Bourne, P. E. (2000) The Protein Data Bank. *Nucleic Acids Res.* **28,** 235–242
44. Wormald, M. R., Petrescu, A. J., Pao, Y. L., Glithero, A., Elliott, T., and Dwek, R. A. (2002) Conformational studies of oligosaccharides and glycopeptides: complementarity of NMR, X-ray crystallography, and molecular modelling. *Chem Rev* **102,** 371–386
45. Petrescu, A. J., Petrescu, S. M., Dwek, R. A., and Wormald, M. R. (1999) A statistical analysis of N- and O-glycan linkage conformations from crystallographic data. *Glycobiology* **9,** 343–352
46. Almasy, L., and Blangero, J. (1998) Multipoint quantitative-trait linkage analysis in general pedigrees. *Am. J. Hum. Genet.* **62,** 1198–1211
47. Beeck, H., and Hellstern, P. (1998) In vitro characterization of solvent/detergent-treated human plasma and of quarantine fresh frozen plasma. *Vox Sang 74* Suppl **1,** 219–223
48. Noda, K., Miyoshi, E., Uozumi, N., Yanagidani, S., Ikeda, Y., Gao, C., Suzuki, K., Yoshihara, H., Yoshikawa, K., Kawano, K., Hayashi, N., Hori, M., Taniguchi, N., and Yoshikawa, M. (1998) Gene expression of alpha1–6 fucosyltransferase in human hepatoma tissues: a possible implication for increased fucosylation of alpha-fetoprotein. *Hepatology* **28,** 944–952
49. Holland, M., Yagi, H., Takahashi, N., Kato, K., Savage, C. O., Goodall, D. M., and Jefferis, R. (2006) Differential glycosylation of polyclonal IgG, IgG-Fc and IgG-Fab isolated from the sera of patients with ANCA-associated systemic vasculitis. *Biochim. Biophys. Acta* **1760,** 669–677
50. Gornik, O., and Lauc, G. (2008) Glycosylation of serum proteins in inflammatory diseases. *Dis. Markers* **25,** 267–278
51. Fujii, S., Nishiura, T., Nishikawa, A., Miura, R., and Taniguchi, N. (1990) Structural heterogeneity of sugar chains in immunoglobulin G. Conformation of immunoglobulin G molecule and substrate specificities of glycosyltransferases. *J. Biol. Chem.* **265,** 6009–6018
52. Narasimhan, S., Freed, J. C., and Schachter, H. (1985) Control of glycoprotein synthesis. Bovine milk UDPgalactose:N-acetylglucosamine beta-4-galactosyltransferase catalyzes the preferential transfer of galactose to the GlcNAc beta 1,2Man alpha 1,3- branch of both bisected and nonbisected complex biantennary asparagine-linked oligosaccharides. *Biochemistry* **24,** 1694–1700
53. Sutton, B. J., and Phillips, D. C. (1983) The three-dimensional structure of the carbohydrate within the Fc fragment of immunoglobulin G. *Biochem. Soc Trans* **11,** 130–132
54. Lund, J., Takahashi, N., Pound, J. D., Goodall, M., and Jefferis, R. (1996) Multiple interactions of IgG with its core oligosaccharide can modulate recognition by complement and human Fc gamma receptor I and influence the synthesis of its oligosaccharide chains. *J. Immunol.* **157,** 4963–4969
55. Wormald, M. R., Rudd, P. M., Harvey, D. J., Chang, S. C., Scragg, I. G., and Dwek, R. A. (1997) Variations in oligosaccharide-protein interactions in immunoglobulin G determine the site-specific glycosylation profiles and modulate the dynamic motion of the Fc oligosaccharides. *Biochemistry* **36,** 1370–1380
56. Anthony, R. M., Nimmerjahn, F., Ashline, D. J., Reinhold, V. N., Paulson, J. C., and Ravetch, J. V. (2008) Recapitulation of IVIG anti-inflammatory activity with a recombinant IgG Fc. *Science* **320,** 373–376
57. Arnold, J. N., Wormald, M. R., Sim, R. B., Rudd, P. M., and Dwek, R. A. (2007) The impact of glycosylation on the biological function and structure of human immunoglobulins. *Annu. Rev. Immunol.* **25,** 21–50
58. Nimmerjahn, F., and Ravetch, J. V. (2006) Fcgamma receptors: old friends and new family members. *Immunity* **24,** 19–28
59. Nimmerjahn, F., and Ravetch, J. V. (2008) Anti-inflammatory actions of intravenous immunoglobulin. *Annu. Rev. Immunol.* **26,** 513–533
60. Thobhani, S., Yuen, C. T., Bailey, M. J., and Jones, C. (2009) Identification and quantification of N-linked oligosaccharides released from glycoproteins: an inter-laboratory study. *Glycobiology* **19,** 201–211
61. Wada, Y., Azadi, P., Costello, C. E., Dell, A., Dwek, R. A., Geyer, H., Geyer, R., Kakehi, K., Karlsson, N. G., Kato, K., Kawasaki, N., Khoo, K. H., Kim, S., Kondo, A., Lattova, E., Mechref, Y., Miyoshi, E., Nakamura, K., Narimatsu, H., Novotny, M. V., Packer, N. H., Perreault, H., Peter-Katalinic, J., Pohlentz, G., Reinhold, V. N., Rudd, P. M., Suzuki, A., and Taniguchi, N. (2007) Comparison of the methods for profiling glycoprotein glycans–HUPO Human Disease Glycomics/Proteome Initiative multi-institutional study. *Glycobiology* **17,** 411–422
62. Stadlmann, J., Pabst, M., Kolarich, D., Kunert, R., and Altmann, F. (2008) Analysis of immunoglobulin glycosylation by LC-ESI-MS of glycopeptides and oligosaccharides. *Proteomics* **8,** 2858–2871
63. Youings, A., Chang, S. C., Dwek, R. A., and Scragg, I. G. (1996) Site-specific glycosylation of human immunoglobulin G is altered in four rheumatoid arthritis patients. *Biochem. J.* **314** (Pt 2), 621–630
64. Shields, R. L., Lai, J., Keck, R., O'Connell, L. Y., Hong, K., Meng, Y. G., Weikert, S. H., and Presta, L. G. (2002) Lack of fucose on human IgG1 N-linked oligosaccharide improves binding to human Fcgamma RIII and antibody-dependent cellular toxicity. *J. Biol. Chem.* **277,** 26733–26740
65. Schachter, H. (1986) Biosynthetic controls that determine the branching and microheterogeneity of protein-bound oligosaccharides. *Adv. Exp. Med. Biol.* **205,** 53–85
66. Fukuta, K., Abe, R., Yokomatsu, T., Omae, F., Asanagi, M., and Makino, T. (2000) Control of bisecting GlcNAc addition to N-linked sugar chains. *J. Biol. Chem.* **275,** 23456–23461
67. Ferrara, C., Brunker, P., Suter, T., Moser, S., Puntener, U., and Umana, P. (2006) Modulation of therapeutic antibody effector functions by glycosylation engineering: influence of Golgi enzyme localization domain and co-expression of heterologous beta1, 4-N-acetylglucosaminyltransferase III and Golgi alpha-mannosidase II. *Biotechnol. Bioeng.* **93,** 851-861

68. Knežević, A., Gornik, O., Polašek, O., Pučić, M., Novokmet, M., Redžić, I., Rudd, P. M., Wright, A. F., Campbell, H., Rudan, I., and Lauc, G. (2010) Effects of aging, body mass index, plasma lipid profiles, and smoking on human plasma N-glycans. *Glycobiology* **20,** 959–969
69. Mimura, Y., Ashton, P. R., Takahashi, N., Harvey, D. J., and Jefferis, R. (2007) Contrasting glycosylation profiles between Fab and Fc of a human IgG protein studied by electrospray ionization mass spectrometry. *J. Immunol. Methods* **326,** 116–126
70. Kita, Y., Miura, Y., Furukawa, J., Nakano, M., Shinohara, Y., Ohno, M., Takimoto, A., and Nishimura, S. (2007) Quantitative glycomics of human whole serum glycoproteins based on the standardized protocol for liberating N-glycans. *Mol. Cell Proteomics* **6,** 1437–1445
71. Visscher, P. M., Hill, W. G., and Wray, N. R. (2008) Heritability in the genomics era–concepts and misconceptions. *Nat. Rev. Genet.* **9,** 255–266
72. Körner, C., Lehle, L., and von Figura, K. (1998) Carbohydrate-deficient glycoprotein syndrome type 1: correction of the glycosylation defect by deprivation of glucose or supplementation of mannose. *Glycoconjugate J.* **15,** 499–505
73. Ota, H., Nakayama, J., Momose, M., Hayama, M., Akamatsu, T., Katsuyama, T., Graham, D. Y., and Genta, R. M. (1998) Helicobacter pylori infection produces reversible glycosylation changes to gastric mucins. *Virchows Arch* **433,** 419–426
74. Tangvoranuntakul, P., Gagneux, P., Diaz, S., Bardor, M., Varki, N., Varki, A., and Muchmore, E. (2003) Human uptake and incorporation of an immunogenic nonhuman dietary sialic acid. *Proc. Natl. Acad. Sci. U.S.A.* **100,** 12045–12050
75. Barišić, K., Lauc, G., Dumić, J., Pavlović, M., and Flögel, M. (1996) Changes of glycoprotein patterns in sera of humans under stress. *Eur. J. Clin. Chem. Clin. Biochem.* **34,** 97–101
76. Parekh, R., Roitt, I., Isenberg, D., Dwek, R., and Rademacher, T. (1988) Age-related galactosylation of the N-linked oligosaccharides of human serum IgG. *J. Exp. Med.* **167,** 1731–1736
77. Ruhaak, L. R., Uh, H. W., Beekman, M., Koeleman, C. A., Hokke, C. H., Westendorp, R. G., Wuhrer, M., Houwing-Duistermaat, J. J., Slagboom, P. E., and Deelder, A. M. (2010) Decreased levels of bisecting GlcNAc glycoforms of IgG are associated with human longevity. *PLoS One* **5,** e12566
78. Vanhooren, V., Dewaele, S., Libert, C., Engelborghs, S., De Deyn, P. P., Toussaint, O., Debacq-Chainiaux, F., Poulain, M., Glupczynski, Y., Franceschi, C., Jaspers, K., van der Pluijm, I., Hoeijmakers, J., and Chen, C. C. (2010) Serum N-glycan profile shift during human ageing. *Exp. Gerontol.* **45,** 738–743

High-throughput IgG Fc *N*-glycosylation profiling by mass spectrometry of glycopeptides

3



Journal of proteome research

Article

pubs.acs.org/jpr

High-Throughput IgG Fc N-Glycosylation Profiling by Mass Spectrometry of Glycopeptides

Maja Pučić Baković,[†] Maurice H. J. Selman,[‡] Marcus Hoffmann,[‡,§] Igor Rudan,[∥] Harry Campbell,[∥] André M. Deelder,[‡] Gordan Lauc,[†,⊥] and Manfred Wuhrer*[,‡]

[†]Laboratory of Glycobiology, Genos Ltd., Planinska 1, 10000 Zagreb, Croatia

[‡]Biomolecular Mass Spectrometry Unit, Leiden University Medical Center, Leiden, The Netherlands

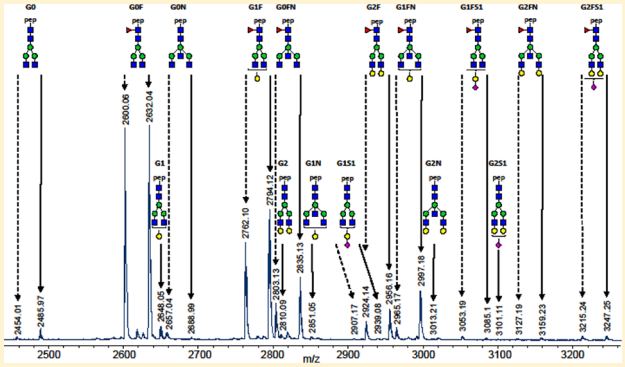
[§]Max Planck Institute for Dynamics of Complex Technical Systems, Sandtorstrasse 1, 39106 Magdeburg, Germany

[∥]Centre for Population Health Sciences, The University of Edinburgh Medical School, Edinburgh, U.K.

[⊥]Faculty of Pharmacy and Biochemistry, University of Zagreb, Ante Kovačića 1, 10000 Zagreb, Croatia

Supporting Information

ABSTRACT: Age and sex dependence of subclass specific immunoglobulin G (IgG) Fc *N*-glycosylation was evaluated for 1709 individuals from two isolated human populations. IgGs were obtained from plasma by affinity purification using 96-well protein G monolithic plates and digested with trypsin. Fc *N*-glycopeptides were purified and analyzed by negative-mode MALDI-TOF-MS with 4-chloro-α-cyanocinnamic acid (Cl-CCA) matrix. Age-associated glycosylation changes were more pronounced in younger individuals (<57 years) than in older individuals (>57 years) and in females than in males. Galactosylation and sialylation decreased with increasing age and showed significant sex dependence. Interestingly, the most prominent drop in the levels of galactosylated and sialylated glycoforms in females was observed around the age of 45 to 60 years when females usually enter menopause. The incidence of bisecting *N*-acetylglucosamine increased in younger individuals and reached a plateau at older age. Furthermore, we compared the results to the total IgG *N*-glycosylation of the same populations recently analyzed by hydrophilic interaction liquid chromatography (HILIC). Significant differences were observed in the levels of galactosylation, bisecting *N*-acetylglucosamine and particularly sialylation, which were shown to be higher in HILIC analysis. Age and sex association of glycosylation features was, to a large extent, comparable between MALDI-TOF-MS and HILIC IgG glycosylation profiling.



KEYWORDS: *immunoglobulin G, glycopeptides, N-glycosylation, glycomics, glycome, mass spectrometry, hydrophilic interaction liquid chromatography*

INTRODUCTION

Immunoglobulin G (IgG) is the most abundant glycoprotein in human serum and a major effector molecule of the humoral immune response. Human IgG occurs in four subclasses (IgG1−4), and each molecule consists of two heavy and two light chains. The two light chains together with the parts of the heavy chains (V_H and C_H1 domains) form two Fab (fragment antigen-binding) moieties, which are linked by a flexible hinge region to one Fc (fragment crystallizable) moiety formed by the remainders of the two heavy chains (C_H2 and C_H3 domains).[1] Each heavy chain in the Fc region carries a single covalently attached biantennary *N*-glycan at the highly conserved asparagine 297.[2] Fc glycans are essential structural components of the IgG molecule and minor changes in glycan composition can significantly alter the conformation of the Fc region changing the interaction with receptor proteins and thus modulating the effector functions of IgG.[3,4] The lack of core fucose enhances the IgG1 binding to activating Fc receptor FcγRIIIa leading to increased antibody-dependent cellular cytotoxicity (ADCC) and destruction of target cells.[2,5−7] Moreover, the presence of sialic acid on the Fc *N*-glycans confers anti-inflammatory properties to IgG.[8] In a mouse arthritis model, the anti-inflammatory effect of sialylated IgG was shown to be mediated through interaction with SIGN-R1, a lectin receptor on mouse splenic macrophages, the human orthologue of this receptor being DC-SIGN.[9,10] Very recently Karsten et al.[11] reported anti-inflammatory properties to be mediated by Fc galactosylation via the formation of immune complexes. High *N*-glycan galactosylation of IgG1 in immune complexes was shown to promote the association between the inhibitory IgG receptor FcγRIIB and C-type lectin-like receptor dectin-1, resulting in a blockage of pro-inflammatory effector functions.[11] Following numerous reports on the importance of

Received: September 20, 2012
Published: January 8, 2013

IgG Fc *N*-glycosylation for effector functions, engineering the Fc region of therapeutic antibodies has become a major challenge and goal for the biopharmaceutical industry.[3]

The majority of IgG *N*-glycans are attached to heavy chains of the Fc region, but 15−20% of polyclonal human IgG molecules also contain glycans within the Fab regions, which were found to be more highly bisected, galactosylated, and sialylated than Fc glycans.[3,12−15]

Microheterogeneity of human IgG glycans is known to be dependent on various physiological (age, sex, and pregnancy) and pathological parameters (tumors, infections, autoimmune diseases, etc.).[1] A number of earlier studies have reported that specific patterns of IgG glycosylation change with age and depend on gender. Parekh et al. were the first to describe decreasing levels of galactosylation with aging.[16] Subsequently, it has been shown that age-dependent galactosylation levels are sex-specific and that levels of bisected glycoforms change with age as well.[17−19] Moreover, in a recent large-scale study of IgG glycans, decreased levels of agalactosylated glycoforms with bisecting *N*-acetylglucosamine (GlcNAc) were found to be an early marker of familial longevity.[19]

In this study, we performed large-scale IgG1 and IgG2&3 Fc *N*-glycosylation profiling by negative-mode MALDI-TOF-MS with 4-chloro-α-cyanocinnamic acid (Cl-CCA) matrix in 1709 individuals and investigated age and sex specificity of the glycosylation features: sialylation, galactosylation, fucosylation, and the occurrence of bisecting GlcNAc. We compare the results to the total IgG glycosylation recently analyzed by hydrophilic interaction liquid chromatography (HILIC) of enzymatically-released and fluorescently-labeled glycans.[20]

EXPERIMENTAL SECTION

Study Population

The study was based on plasma samples from respondents who were residents of the Croatian Adriatic islands, Vis and Korčula, and who were recruited within a larger genetic epidemiology program previously described.[21,22] In total, 1709 individuals were included with 795 subjects from Vis (46.5%) and 914 (53.5%) from Korčula. The age range was 18−98 years (median age 57 and interquartile range 21 years). There were 61.7% females and 38.3% males. The study conforms to the ethical guidelines of the 1975 Declaration of Helsinki and was approved by the appropriate Ethics Board of the University of Zagreb Medical School. An informed consent was signed by all participants prior to participation.

IgG Purification

Immunoglobulin G was isolated from plasma by affinity chromatography using 96-well protein G monolithic plates as previously reported.[20] Briefly, 50 μL of plasma was diluted 10× with PBS, applied to the protein G plate and instantly washed. IgGs were eluted using 1000 μL of 100 mM formic acid and immediately neutralized to pH 7.0 with 1 M ammonium bicarbonate.

Trypsin Digestion of Human Polyclonal IgG

Aliquots (1/20; 50 μL) of the protein G eluates were brought to 96-well polypropylene V-bottom microtitration plates, and a standard IgG sample was added in 6-fold to each plate to allow evaluation of interbatch variation. TPCK trypsin (Sigma-Aldrich, St. Louis, MO) was first dissolved in ice-cold 20 mM acetic acid (Merck, Darmstadt, Germany) to a final concentration of 0.4 μg/μL after which it was further diluted to 0.02 μg/μL with ice-cold ultrapure water. For overnight digestion at 37 °C, 20 μL of diluted trypsin was added to each IgG sample.

Reverse-Phase Solid Phase Extraction (RP-SPE) of Glycopeptides

Glycopeptides were purified and desalted by reverse phase (RP) solid phase extraction (SPE) using Chromabond C_{18}ec beads (Marcherey-Nagel, Düren, Germany) as described previously.[23] Briefly, C-18 beads were activated with 80% ACN containing 0.1% trifluoroacetic acid (TFA; Fluka) and conditioned with 0.1% TFA. Tryptic IgG digests were diluted with 0.1% TFA, loaded onto C-18 beads, and washed with 0.1% TFA. IgG glycopeptides were eluted with low concentration of ACN (18%) containing 0.1% TFA to minimize coelution of interfering peptides. Eluates were dried by vacuum centrifugation and stored at −20 °C until mass spectrometric analysis. The heating and acid steps of the sample preparation method were evaluated and found not to lead to any noticeable degradation of the Fc glycan moieties with regard to sialylation and fucosylation.

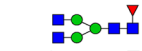
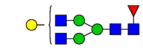
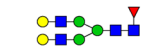
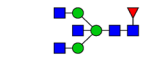
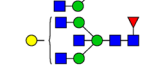
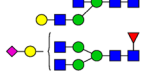
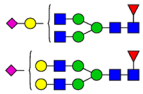
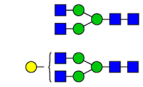
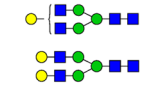
MALDI-TOF-MS of IgG Glycopeptides

Large-scale IgG glycosylation analysis was performed using MALDI-TOF-MS. Samples were dissolved in 20 μL of water, and 3 μL aliquots were spotted onto MTP 384 polished steel target plates (Bruker Daltonics, Bremen, Germany) and allowed to dry at room temperature. Subsequently, 1 μL of 5 mg/mL 4-chloro-α-cyanocinnamic acid (Cl-CCA; 95% purity; Bionet Research, Camelford, Cornwall, U.K.) in 50% acetonitrile was applied on top of each sample and allowed to dry.[24,25] Glycopeptides were analyzed on an UltrafleX II MALDI-TOF/TOF mass spectrometer (Bruker Daltonics), which was operated in the negative-ion reflectron mode. Ions between m/z 1000 and 3800 were recorded. To allow homogeneous spot sampling, a random walk laser movement with 50 laser shots per raster spot was applied, and each IgG glycopeptide sum mass spectrum was generated by accumulation of 2000 laser shots. Mass spectra were internally calibrated using a list of known glycopeptides. Data processing and evaluation were performed with FlexAnalysis Software (Bruker Daltonics) and Microsoft Excel, respectively. The data were baseline subtracted and the intensities of a defined set of 27 glycopeptides (16 glycoforms for IgG1 and 11 for IgG2&3) were automatically defined for each spectrum (Table 1). IgG2 and IgG3 have identical peptide moieties (E_{293}EQFNSTFR$_{301}$) of their tryptic Fc glycopeptides and are, therefore, not distinguished by the profiling method.[23]

Relative intensities of IgG Fc glycopeptides were obtained by integrating and summing four isotopic peaks followed by normalization to the total subclass specific glycopeptide intensities. An exception was made for the IgG1 G1S1 (monogalactosylated, monosialylated biantennary glycan without core fucose; see Table 1) glycoform for which only the first isotopic peak could be reliably determined as the second, third, and fourth isotopic peaks overlap with peaks of the IgG4 G2F glycoform and thus cannot be used to calculate the sum intensity. Instead, the signal intensity for the sum of four isotopic peaks of the IgG1 G1S1 glycoform was determined by multiplying the signal of the first isotopic peak of the IgG1 G1S1 glycoform with a correction factor (4.348, which is the average ratio of the sum of all four isotopic peaks versus the first isotopic peak determined for the related G2S1 glycoform of 126 standard samples). The level of galactosylation was

dx.doi.org/10.1021/pr300887z | J. Proteome Res. 2013, 12, 821−831

Table 1. Theoretical Masses of Tryptic Glycopeptides of Human IgG Subclasses 1, 2, and 3; Glycan Compositions Are Given in Terms of Hexose (H), *N*-Acetylhexoamine (N), Deoxyhexose (fucose; F), and *N*-Acetylneuraminic Acid (Sialic Acid; S); Structural Schemes Are As Given in Figure 1

Glycan species	Glycan structure	IgG1 E_{293}EQYNSTYR$_{301}$[a] [M-H]$^-$	IgG2&3 E_{293}EQFNSTFR$_{301}$[a] [M-H]$^-$
no glycan	-	1189.5120	1157.5222
G0F		2632.0460[b]	2600.0561
G1F		2794.0988[b]	2762.1089
G2F		2956.1516	2924.1617
G0FN		2835.1253[b]	2803.1354
G1FN		2997.1781[b]	2965.1882
G2FN		3159.2309	3127.2410
G1FS1		3085.1942[b]	3053.2043
G2FS1		3247.2470	3215.2571
G0		2485.9880	2453.9981
G1		2648.0408	n.d.[c]
G2		2810.0936	n.d.[c]
G0N		2689.0673	2657.0774
G1N		2851.1201	n.d.[c]
G2N		3013.1729	n.d.[c]
G1S1		2939.1362	2907.1463
G2S1		3101.1890	n.d.[c]

[a]Tryptic IgG glycopeptide sequence. [b]Isomeric glycopeptide species of IgG1 and IgG4. [c]Not determined due to the occurrence of isomeric glycopeptide species of IgG2&3 and IgG4.

calculated from the relative intensities of various Fc *N*-glycoforms (Table 1) according to the formula: (G1 + G1F + G1FN + G1N + G1S1 + G1FS1)0.5 + G2 + G2F + G2FN + G2N + G2S1 + G2FS1 for the IgG1 subclass and (G1F + G1FN + G1S1 + G1FS1)0.5 + G2F + G2FN + G2FS1 for the IgG2&3 subclasses. The prevalence of bisecting GlcNAc was determined by summing the relative intensities of all bisected glycoforms (G0N, G1N, G2N, G0FN, G1FN, and G2FN for IgG1 and G0N, G0FN, G1FN, and G2FN for IgG2&3). The level of sialylation was defined by summation of all sialylated glycoforms (G1S1, G1FS1, G2S1, and G2FS1 for IgG1 and G1S1, G1FS1, and G2FS1 for the IgG2&3 subclasses). The incidence of IgG1 fucosylation was evaluated by summing all fucosylated IgG1 Fc *N*-glycoforms (G0F + G0FN + G1F + G1FN + G1FS1 + G2F + G2FN + G2FS1). The incidence of IgG2&3 fucosylation was not evaluated, as a large portion of the afucosylated IgG2&3 glycoforms could not be determined due to mass spectrometric overlap with isomeric IgG4 glycoforms (Table 1).[23]

HILIC of 2-AB Labeled *N*-Glycans with Fluorescence Detection

Total IgG *N*-glycans were enzymatically released, fluorescently labeled with 2-aminobenzamide, and analyzed by hydrophilic interaction liquid chromatography (HILIC) on a Waters Acquity UPLC instrument (Waters, Milford, MA).[20] 2-AB labeled *N*-glycans were separated on a Waters BEH Glycan chromatography column, 100 × 2.1 mm i.d., 1.7 μm BEH particles, with 100 mM ammonium formate, pH 4.4, as solvent A and acetonitrile as solvent B. A linear gradient of 75−62% of acetonitrile at flow rate of 0.4 mL/min in 20 min was used to separate *N*-glycans into 24 peaks.

Comparison of MALDI-TOF-MS and Chromatography Analysis

In order to compare MALDI-TOF-MS IgG Fc *N*-glycopeptide profiles with the HILIC IgG *N*-glycan profiles, we recalculated glycosylation features of the MS measurement to be in accordance with published HILIC calculations. Hence, galactosylation and bisecting GlcNAc levels were determined only from neutral glycoforms. From MALDI-TOF-MS analysis, 12 out of 16 detected IgG1 glycoforms and 8 out of 11 detected IgG2&3 glycoforms were included in this comparison, and a normalization was performed (total signal intensity per subclass was set to 100%). As for HILIC glycan analysis, 12 neutral glycoforms from 14 chromatographic peaks (G1F and G1FN glycoforms were separated into two peaks each) were considered, and normalization was performed (total signal of the 14 chromatographic peaks = 100%). The level of galactosylation in neutral glycoforms was calculated as follows: (G1 + G1F + G1FN + G1N)0.5 + G2 + G2F + G2FN + G2N for the IgG1 subclass, (G1F + G1FN)0.5 + G2F + G2FN for the IgG2&3 subclasses, and (G1 + G1F + G1FN) 0.5 + G2 + G2N + G2F + G2FN for the total IgG. The incidence of bisecting GlcNAc in neutral glycoforms was determined by summing the relative intensities of all neutral bisected glycoforms: (G0N + G1N + G2N + G0FN + G1FN + G2FN) for the IgG1, (G0N + G0FN + G1FN + G2FN) for the IgG2&3, and (G0FN + G1FN + G2N + G2FN) for the total IgG. For comparison of sialylation levels between MALDI-TOF-MS and HILIC analysis, we used a degree of sialylation of fucosylated glycoforms without bisecting GlcNAc (FGS/(F + FG + FGS)) calculated from the total detected glycoforms: (G1FS1 + G2FS1)/(G0F + G1F + G2F + G1FS1 + G2FS1)100 for IgG1 and IgG2&3 and (G1FS1 + G2FS1 + G2FS2)/(G0F + G1F + G2F + G1FS1 + G2FS1 + G2FS2)100 for total IgG (HILIC data).

Additionally, we have compared MALDI-TOF-MS profiles of 2-AA-labeled *N*-glycan with HILIC profiles of 2-AB-labeled *N*-glycans for two standard samples, i.e., IgG and IgG Fc (Athens Research & Technology, Athens, GA). Sample preparation and analysis of 2-AB glycans was as described above. Details on sample preparation and analysis of 2-AA glycans can be found in Supplementary Figure 5. Comparison of *N*-glycopeptide patterns with HILIC profiles was performed for IgG1 and IgG Fc standard samples as described in Supplementary Figure 6.

Statistical Analysis

IgG glycosylation variables were tested for normality using the Kolmogorov−Smirnov test, and nonparametric statistical tests were further used. Correlations were determined with Spearman's rank test and gender differences were tested with the Mann−Whitney test. Statistical analysis was performed with SPSS 13 (SPSS Inc., Chicago, IL).

dx.doi.org/10.1021/pr300887z | J. Proteome Res. 2013, 12, 821−831

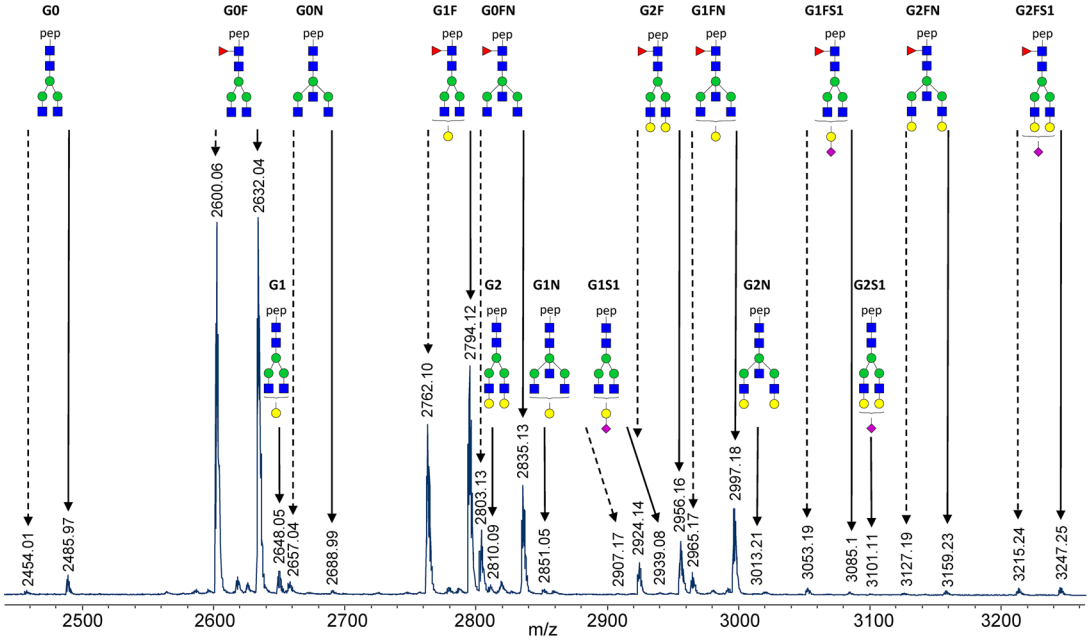


Figure 1. MALDI-TOF-MS spectra of tryptic IgG Fc *N*-glycopeptides. Dashed arrows represent IgG2&3 glycopeptides, while continued arrows represent IgG1 glycopeptides. Glycan species are given in terms of number of galactoses (G0, G1, and G2), fucose (F), bisecting *N*-acetylglucosamine (N), and *N*-acetylneuraminic acid (S). Structural schemes are given in terms of pep (peptide moiety), blue square (*N*-acetylglucosamine), red triangle (fucose), green circle (mannose), yellow circle (galactose), and purple diamond (*N*-acetylneuraminic acid).

Table 2. Correlation Coefficients of IgG Glycosylation Features and Age Stratified for Sex[a]

		all ages		ages ≤ 57		ages > 57	
		female	male	female	male	female	male
IgG subclass	glycosylation feature	*r*(*P*)	*r*(*P*)	*r*(*P*)	*r*(*P*)	*r*(*P*)	*r*(*P*)
IgG1 Fc	Gal	**−0.66** (<0.001)	**−0.44** (<0.001)	**−0.54** (<0.001)	**−0.39** (<0.001)	**−0.21** (<0.001)	**−0.23** (<0.001)
	Bis GlcNAc	**0.27** (<0.001)	**0.19** (<0.001)	**0.28** (<0.001)	**0.21** (<0.001)	0.04 (0.440)	0.07 (0.262)
	Core F	−0.01 (0.880)	−0.12 (0.009)	**−0.14** (0.003)	**−0.19** (0.003)	0.01 (0.767)	−0.03 (0.580)
	Sial	**−0.43** (<0.001)	**−0.17** (<0.001)	**−0.39** (<0.001)	−0.13 (0.044)	−0.01 (0.910)	−0.06 (0.372)
IgG2&3 Fc	Gal	**−0.69** (<0.001)	**−0.46** (<0.001)	**−0.55** (<0.001)	**−0.34** (<0.001)	**−0.27** (<0.001)	**−0.27** (<0.001)
	Bis GlcNAc	**0.17** (<0.001)	**0.14** (<0.001)	**0.18** (<0.001)	**0.16** (0.005)	−0.05 (0.267)	0.04 (0.451)
	Sial	**−0.60** (<0.001)	**−0.32** (<0.001)	**−0.48** (<0.001)	**−0.27** (<0.001)	**−0.22** (<0.001)	−0.15 (0.009)

[a]Positive correlation coefficients (*r*) for age indicate increased levels with increasing age, while negative correlation coefficients indicate decreased levels with increasing age. Correlations found to be significant after Bonferroni correction for gender and glycosylation features ($P \leq 0.006$ for IgG1 and $P \leq 0.008$ for IgG2&3) are in bold. Gal, level of galactosylation; Bis GlcNAc, level of bisecting *N*-acetylglucosamine; Core F, level of core fucosylation,; Sial, level of sialylation.

RESULTS

MALDI-TOF-MS Glycopeptide Profiling

IgG was purified from plasma samples of 1709 individuals (61.7% female and 38.3% male) ranging in age between 18 and 98 years. All four human IgG subclasses (IgG1, IgG2, IgG3, and IgG4) were obtained by high-throughput affinity purification using 96 well protein G monolithic plates. After tryptic digestion, IgG Fc *N*-glycopeptides were enriched and desalted by RP-SPE and analyzed by negative-ion reflectron mode MALDI-TOF-MS using 4-chloro-α-cyanocinnamic acid as the matrix substance.[24] IgG3 Fc *N*-glycopeptides have a peptide moiety, which is identical to the one of IgG2 (Table 1).[26] Therefore, these two subclasses were determined together. IgG4 Fc *N*-glycopeptides were not determined due to the low abundance of their signals and overlap with other signals (Table 1).[26] Mass spectra were processed automatically, which resulted in the determination of 16 IgG1 Fc glycoforms and 11 IgG2&3 Fc glycoforms (Table 1 and Figure 1). Glycoforms were assigned on the basis of their composition and literature data on human plasma IgG glycosylation.[13,20,27] Relative intensities of glycoforms were obtained by integration and summation of four isotopic peaks followed by normalization to the total subclass specific glycopeptide intensities.

To determine inter- and intrabatch variation of the analytical method, a standard IgG sample was added in six replicates to each sample plate. The intrabatch and interbatch variation were determined for the five major glycoforms of the standard IgG samples, and the relative standard deviation (RSD) was found to be ≤6% for each sample plate (intrabatch) and below 10% over the entire 20 plates (interbatch) for both IgG1 and IgG2&3. No batch correction was performed.

Association of IgG Glycosylation with Age and Sex

We evaluated the correlation of IgG glycosylation features and glycoforms with age (Table 2). The obtained data showed a pronounced decrease of IgG1 and IgG2&3 Fc galactosylation with age. The relative abundances of agalactosylated IgG

dx.doi.org/10.1021/pr300887z | J. Proteome Res. 2013, 12, 821−831

Table 3. Descriptives of Glycosylation Features in Females and Males with Statistical Significance (P) of Sex Differences and Differences between Age Groups[a]

		ages ≤ 57			ages > 57			differences between age groups (P)	
		female	male		female	male			
IgG subclass	glycosylation feature	median (IQR)	median (IQR)	sex differences (P)	median (IQR)	median (IQR)	sex differences (P)	female	male
IgG1 Fc	Gal	43.0 (12.8)	39.0 (8.9)	**<0.001**	31.3 (8.9)	33.9 (9.6)	**<0.001**	**<0.001**	**<0.001**
	Bis GlcNAc	11.8 (4.4)	11.9 (4.2)	0.785	13.6 (4.8)	12.9 (4.3)	0.008	**<0.001**	**0.001**
	Core F	92.0 (5.4)	92.6 (5.4)	0.122	92.4 (4.2)	91.9 (5.2)	0.147	0.159	0.117
	Sial	4.4 (2.2)	3.8 (2.1)	**<0.001**	3.0 (1.8)	3.2 (1.9)	**0.004**	**<0.001**	**0.001**
IgG2&3 Fc	Gal	33.0 (13.9)	29.7 (9.3)	**<0.001**	21.2 (7.6)	24.1 (7.9)	**<0.001**	**<0.001**	**<0.001**
	Bis GlcNAc	8.3 (3.3)	8.2 (3.1)	0.051	9.3 (3.6)	8.5 (3.1)	**<0.001**	**<0.001**	0.010
	Sial	3.8 (2.6)	3.4 (1.9)	**0.002**	2.0 (1.3)	2.6 (1.5)	**<0.001**	**<0.001**	**<0.001**

[a]Sex differences found to be significant after Bonferroni correction for age and glycosylation features ($P \leq 0.006$ for IgG1 and $P \leq 0.008$ for IgG2&3) and differences between age groups found to be significant after Bonferroni correction for sex and glycosylation features ($P \leq 0.006$ for IgG1 and $P \leq 0.008$ for IgG2&3) are in bold. IQR, interquartile range. Abbreviations of glycosylation features as in Table 2.

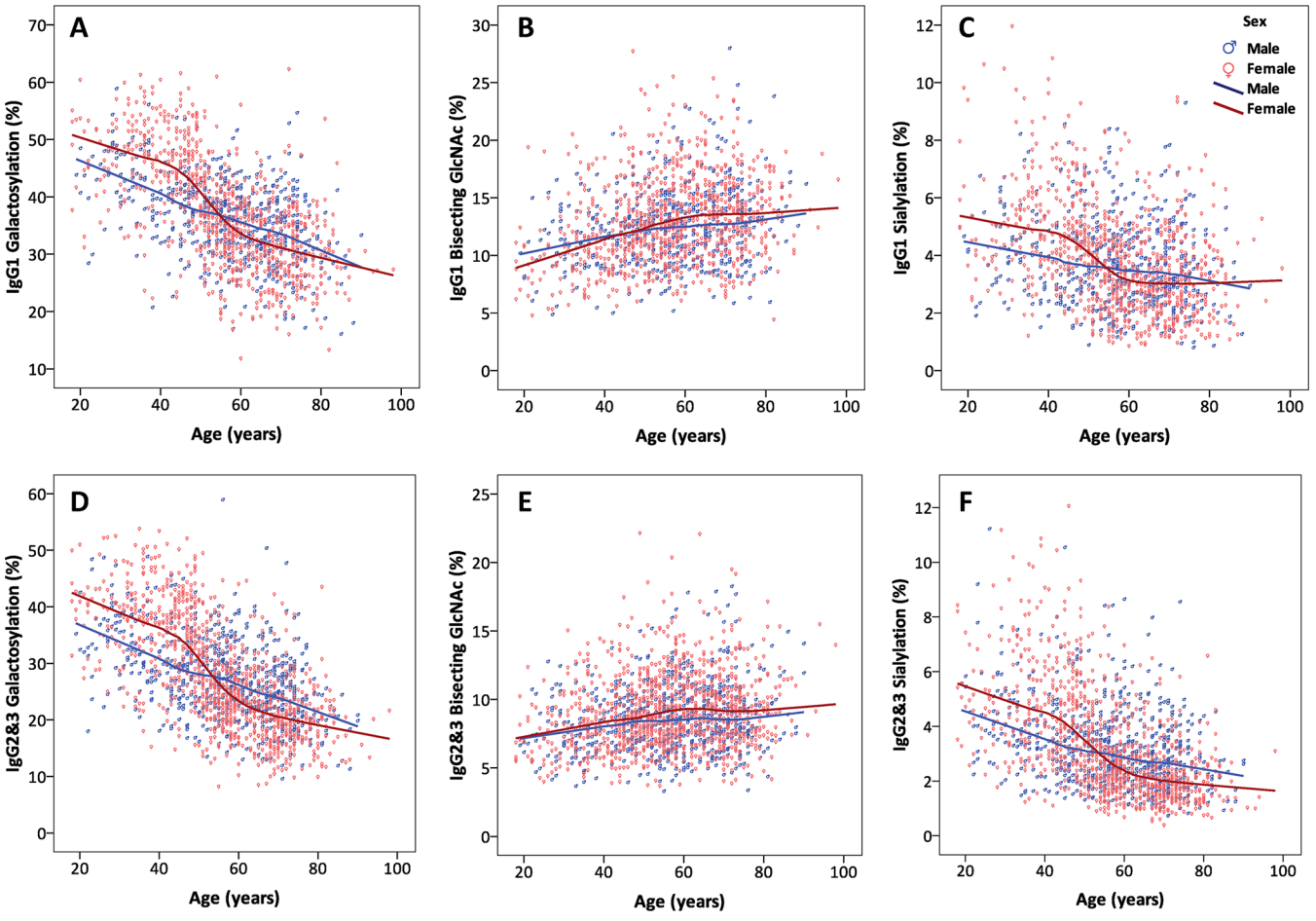


Figure 2. Age dependence of IgG1 and IgG2&3 glycosylation features. Females are plotted in red with a fitted line in dark red, while males are plotted in blue with a fitted line in dark blue. Both lines were fitted using the lowss (locally weighted scatterplot smoothing) method.

glycoforms (represented by G0F glycoform) increased at higher age (Supplementary Figure 1A,D). The opposite was observed for glycoforms with one or two galactoses (represented by G1F and G2F, respectively), whose abundances decreased with age (Supplementary Figure 1B,C,E,F). Previously, it has been shown that IgG glycosylation changes with age are sex-specific and that those associations are more evident and stronger in individuals of up to 50−60 years of age.[18,19,28] Therefore, to reveal more details, we divided our whole population into two age groups at median age (57 years of age). Younger individuals, both female and male, showed stronger negative correlation between age and galactosylation in all tested IgG subclasses. Stronger correlation coefficients in the younger group were also observed in all of the rest of glycosylation features, independent of the sex and subclass (Table 2). Statistically significant differences ($p \leq 0.001$) between the two age groups were revealed for all the glycosylation features that showed significant age dependence (Table 3).

dx.doi.org/10.1021/pr300887z | J. Proteome Res. 2013, 12, 821−831

Table 4. Correlation Coefficients (r) of Glycosylation Features[a]

		all ages			ages ≤ 57			ages > 57		
IgG subclass	glycosylation feature	Bis GlcNAc (r)	Core F (r)	Sial (r)	Bis GlcNAc (r)	Core F (r)	Sial (r)	Bis GlcNAc (r)	Core F (r)	Sial (r)
IgG1 Fc	Gal	**−0.09** (<0.001)	**−0.17** (<0.001)	**0.68** (<0.001)	**−0.12** (<0.001)	**−0.15** (<0.001)	**0.67** (<0.001)	**0.13** (<0.001)	**−0.24** (<0.001)	**0.60** (<0.001)
	Bis GlcNAc		**−0.30** (<0.001)	**0.13** (<0.001)		**−0.31** (<0.001)	**0.13** (0.001)		**−0.31** (<0.001)	**0.28** (<0.001)
	Core F			**−0.38** (<0.001)			**−0.35** (<0.001)			**−0.44** (<0.001)
IgG2&3 Fc	Gal	**−0.11** (<0.001)		**0.84** (<0.001)	−0.08 (0.026)		**0.83** (<0.001)	−0.01 (0.078)		**0.75** (<0.001)
	Bis GlcNAc			0.04 (0.081)			0.07 (0.049)			**0.16** (<0.001)

[a]Correlations found to be significant after Bonferroni correction ($P \leq 0.004$ for IgG1 and $P \leq 0.008$ for IgG2&3) are in bold. Abbreviations of glycosylation features as in Table 2.

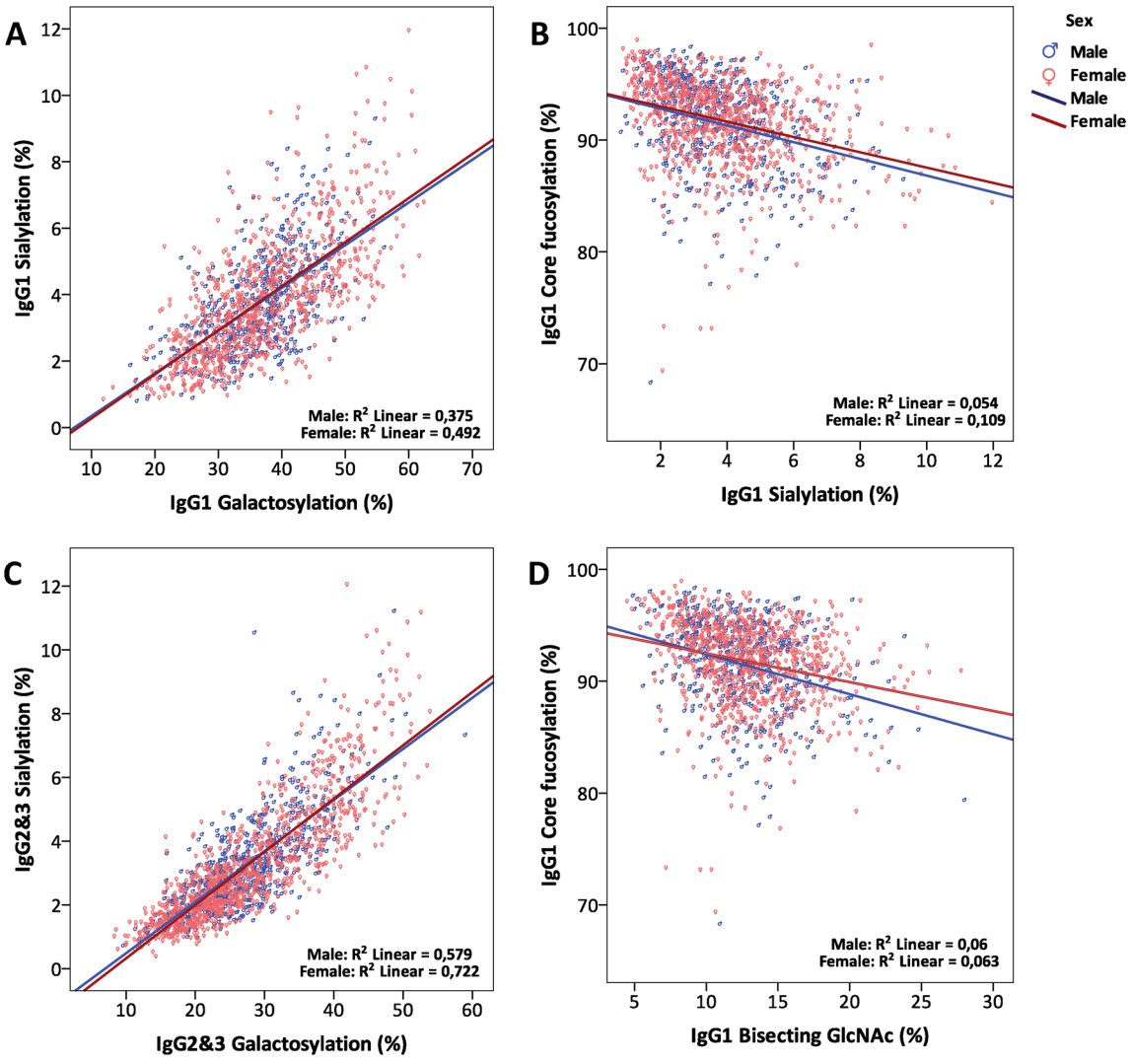


Figure 3. Correlations between IgG Fc glycosylation features stratified for sex. Females are plotted in red with a fitted line in dark red, while males are plotted in blue with a fitted line in dark blue. R^2, coefficient of determination.

When galactosylation was evaluated for the entire age range, no sex difference was revealed (data not shown). However, by analyzing the data stratified for ages below and above the median age, statistically significant differences in IgG galactosylation between females and males emerge. Younger females showed a stronger correlation between galactosylation and age than males for both IgG1 and IgG2&3 (Table 2). The level of galactosylation reached similar values for both sexes around median age after which sex differences in galactosylation changes with age were less obvious (Figure 2A,D). Interestingly, while the younger group of females (age ≤ 57) appeared to have higher galactosylation than males (median of 43.0%

dx.doi.org/10.1021/pr300887z | J. Proteome Res. 2013, 12, 821−831

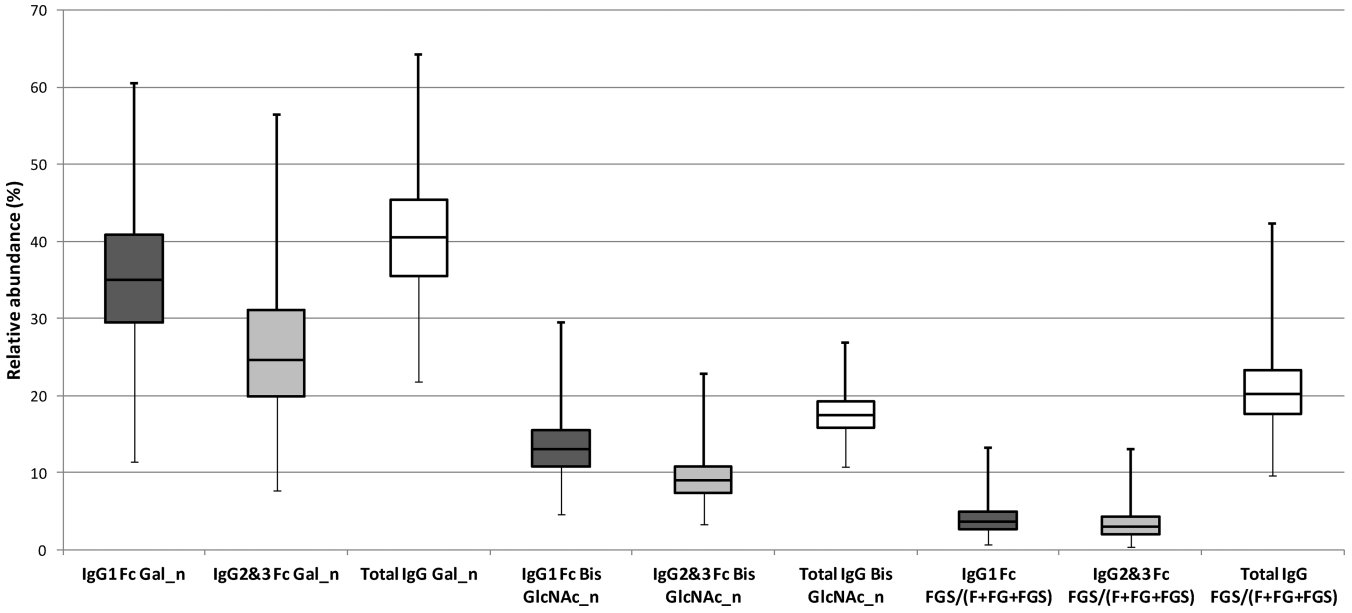


Figure 4. Boxplot representations of the levels of glycosylation features analyzed by MALDI-TOF-MS (IgG1 Fc and IgG2&3 Fc) and by HILIC (total IgG). Gal_n, level of galactosylation in neutral glycoforms; BisGlcNAc_n, incidence of bisecting GlcNAc in neutral glycoforms; FGS/(F + FG + FGS), degree of sialylation of fucosylated glycoforms without bisecting GlcNAc. The bottom of the box represents the lower quartile, while the top represents the upper quartile (25th and 75th percentile, respectively). The middle line represents the median. The whiskers extend to the minimum and to the maximum values.

versus 39.0%, respectively), at older age, this relationship seemed to be reversed with males showing higher level of galactosylation than females (median of 33.9% versus 31.3%, respectively; Table 3).

Association of IgG sialylation with age showed a trend similar to that of galactosylation. The level of sialylated glycoforms significantly decreased with increasing age for both IgG1 and IgG2&3 (Table 2). This negative correlation was more pronounced for younger individuals than for older ones. Correlation of age with sialylation was stronger in females than in males. Moreover, just like in the case of galactosylation, IgG sialylation for younger individuals was higher in females than in males (Figure 2C,F). Around the age of 60, sialylation reached similar levels for both sexes after which males exhibited higher level of sialylation.

The incidence of bisecting GlcNAc increased with age (Table 2; Figure 2B,E). However, significant positive correlations were observed only for the younger individuals, while in the older age group, the abundance of bisecting GlcNAc-containing glycoforms seemed to reach a plateau (Figure 2B,E). Within both age groups, no sex differences were found except for the IgG2&3 subclasses, which showed higher level of bisecting GlcNAc in older females compared to males of similar age.

In addition, age- and sex-dependent variations in IgG1 core fucosylation were evaluated. A weak but nevertheless significant negative age effect on the level of IgG1 core fucosylation was observed only in younger individuals (Table 2). Within our cohort, there was no clear difference in fucosylated IgG1 glycoforms between females and males (Table 3).

Next, we analyzed correlations between different IgG glycosylation features with Spearman's rank test (Table 4). The strongest relationship was found between galactosylation and the level of sialylation for IgG1 (r = 0.68) as well as IgG2&3 (r = 0.84) (Figure 3A,C). IgG1 galactosylation showed a negative correlation (r = −0.17) with the level of core fucosylation (Supplementary Figure 2A). An even stronger negative correlation was observed between IgG1 sialylation and core fucosylation, r = −0.38 (Figure 3B). Over the whole age range, the galactosylation of the analyzed IgG subclasses exhibited a weak negative (r = −0.09 for IgG1 and r = −0.11 for IgG2&3) but significant correlation with the incidence of bisecting GlcNAc (Supplementary Figure 2D,E,F). Interestingly, this effect showed a pronounced age-dependence for the IgG1 subclass: while younger individuals revealed a negative correlation (r = −0.12), at higher age, galactosylation appeared to positively correlate (r = 0.13) with bisecting GlcNAc. We additionally observed a negative correlation (r = −0.30) between the level of core fucosylated IgG1 *N*-glycans and the incidence of bisecting GlcNAc (Figure 3D). In addition, a weak positive correlation (r = 0.13) was observed between the bisecting GlcNAc and sialylated IgG1 *N*-glycans (Supplementary Figure 2B). A similar positive correlation (r = 0.16) was observed for bisecting GlcNAc and sialylation of IgG2&3 subclasses but only for older individuals (Supplementary Figure 2C).

Comparison of MALDI-TOF-MS and HILIC IgG Glycosylation Profiles

In order to compare the subclass-linked IgG Fc glycosylation analyzed by MALDI-TOF-MS with the total IgG (Fc and Fab of all subclasses) *N*-glycosylation profiles analyzed by HILIC, we chose to use calculations of glycosylation features as performed previously with HILIC data.[20] Hence, we compared IgG1 Fc glycosylation, IgG2&3 Fc glycosylation and total IgG glycosylation by evaluating the levels of galactosylation and bisecting GlcNAc from neutral glycoforms and the degree of sialylation of fucosylated glycoforms (without bisecting GlcNAc) from the total measured glycoforms. When levels of glycosylation features calculated from MALDI-TOF-MS profiles were correlated with the features calculated from HILIC analysis, highly significant and strong positive correlation coefficients of IgG1 Fc or IgG2&3 Fc with the

dx.doi.org/10.1021/pr300887z | J. Proteome Res. 2013, 12, 821−831

total IgG were obtained (Supplementary Figure 3). The strongest correlation was observed for the galactosylation of neutral glycoforms of total IgG and subclass-specific Fc glycopeptides ($r \geq 0.90$). The degree of sialylation of total fucosylated glycoforms as well as the levels of bisecting GlcNAc in neutral glycoforms analyzed by the two methods also revealed a positive correlation but with weaker coefficients ($r \geq 0.51$ and $r \geq 0.70$, respectively).

Next, MALDI-TOF-MS and HILIC data were compared with respect to age and sex dependence. Correlation coefficients of age and glycosylation features observed for IgG1 Fc, IgG2&3 Fc, and the total IgG (Supplementary Table 1) were very similar to those described above for the total Fc glycoforms. Correlation of age with the level of galactosylation was almost the same for all three data sets (IgG1 Fc, IgG2&3 Fc, and total IgG). As for the association of the degree of sialylation with age, IgG2&3 Fc and the total IgG were rather similar showing a strong negative correlation, evident even in the older group of individuals, while this correlation was much less pronounced for IgG1 Fc. Minor distinctions were noticed in the strength of the correlations of age with bisecting GlcNAc. The total IgG (Fc and Fab of all four subclasses) measured by HILIC showed stronger positive correlation of bisecting GlcNAc with age than IgG1 Fc and IgG2&3 Fc measured by MALDI-TOF-MS. Similarly to the total glycoforms, this age effect on bisecting GlcNAc was only significant for the younger individuals (≤57 years) as seen for both methods.

Very similar differences regarding sex and age groups were consistently observed for IgG1 Fc, IgG2&3 Fc, and the total IgG (Supplementary Table 2). Significant differences between younger and older females were observed for the levels of galactosylation, bisecting GlcNAc, and the degree of sialylation. In males, only the galactosylation levels measured by both methods notably differed between the two age groups (below and above age 57). As for the bisecting GlcNAc, the difference between younger and older males was revealed only for IgG1 Fc, while IgG2&3 Fc and the total IgG showed a difference in the degree of sialylation. Regardless of the analytical method, older individuals showed sex differences in all glycosylation features, with the exception of bisecting GlcNAc of IgG1. Similarly, younger females and males had significantly different levels of galactosylation and sialylation. Median levels of glycosylation features for the whole studied population were compared between the total IgG and subclass-specific Fc fragments (Supplementary Table 2). Levels of galactosylation and bisecting GlcNAc as well as the degree of sialylation were higher for the total IgG than for the IgG Fc glycosylation data sets (Figure 4). The difference was most prominent in the degree of sialylation of fucosylated glycoforms since the total IgG exhibited approximately 7 times higher level of sialylation than IgG1 Fc and IgG2&3 Fc. For all three data sets, the level of galactosylation and the level of sialylation showed the same trend of sex-associated differences in both age groups: while younger females had higher levels of galactosylation and sialylation than males, at older age, males showed higher levels of both features than females (Supplementary Figure 4). Another trait for which similar results were obtained in the different data sets is the incidence of the bisecting GlcNAc (Supplementary Table 2), which was increasing only in younger individuals and showed no difference between the sexes. At older age (>57 years), no significant change in the incidence of bisecting GlcNAc was observed anymore, but sex differences became apparent for IgG2&3 Fc and total IgG, with females showing a higher level of this feature.

In the next analysis step, we looked at the relations between glycosylation features for each of the three data sets and compared these relations between IgG1 Fc, IgG2&3 Fc, and total IgG (Supplementary Table 3). In all three cases, a strong positive correlation between the level of galactosylation and sialylation was noticed ($r \geq 0.65$ for the whole age range). Negative association ($r \geq -0.17$) of the galactosylation and the incidence of the bisecting GlcNAc for the total IgG was observed over the whole age range and in younger individuals following the stratification at median age. Similarly, IgG2&3 Fc showed a negative association, while in the case of IgG1 Fc, the only significant correlation was a positive one in older individuals. Correlations of the degree of sialylation and the level of bisecting GlcNAc showed the opposite direction for the Fc fragments and the total IgG. IgG1 Fc and IgG2&3 Fc had a significant positive correlation ($r \geq 0.14$) between those two features for the whole age range as well as for the stratified age groups. By contrast, total IgG showed a negative correlation significant for the whole age range ($r = -0.19$) as well as for the younger group of individuals ($r = -0.17$).

In addition, to evaluate the observed difference in sialylation between Fc glycopeptides (MALDI-TOF-MS) and total IgG glycans (HILIC), we analyzed two sialylated IgG standards by both methods. HILIC glycan profiles of IgG1 and IgG Fc standards were compared with their MALDI-TOF-MS Fc glycopeptide profile obtained in reflectron and linear negative-ion mode (Supplementary Figure 6). The most abundant sialylated glycan, G2FS1, of IgG1 standard showed 2× higher relative intensity in negative linear mode in comparison to negative reflectron mode MS. Relative abundance of this glycoform as determined by HILIC of 2-AB glycans (Supplementary Table 6) was even 3× higher than in linear negative mode MS analysis of glycopeptides.

Next, we analyzed the total released glycans of IgG and IgG Fc standards by negative linear mode MALDI-TOF-MS (after 2-AA-labeling; see Supplementary Table 4 for mass list) and HILIC (after 2-AB-labeling; Supplementary Figure 5). The coefficient of variation (CV) between MS and HILIC data for IgG standard of the major sialylated species, G1FS1 and G2FS1, was found to be less than 5%. Less abundant sialylated glycans (<1%) show much higher CV between two methods (data not shown).

DISCUSSION

We performed a large-scale IgG glycopeptide profiling by MALDI-TOF-MS with Cl-CCA matrix for 1709 individuals from two isolated human populations. In accordance with literature findings, our results showed a clear tendency of decreased galactosylation and sialylation with increasing age.[16−20,23,28] In both age groups (above and below the median age of 57 years), we observed sex-related differences with females showing higher levels at a young age, and males showing slightly higher levels at older age. The most prominent drops in the levels of galactosylation and sialylation in females were observed around the age of between 45 and 60 years, which might be linked to changes in the hormonal status due to the transition from pre- to postmenopausal stage. Altered IgG galactosylation and sialylation levels have been reported during pregnancy[29] suggesting that changes in hormone levels could in part explain observed changes in IgG glycosylation. While the regulation of IgG glycosylation is still largely not understood, a

recent in vitro study using a primary human B cell culture has established regulatory roles for various systemic or microenvironmental factors such as cytokines and all-*trans* retinoic acid.[30]

In line with the parallel decrease of galactosylation and sialylation, we observed, as expected, a strong positive and highly significant correlation between these two features for all the analyzed subclasses. It was reported that GalT-I (β-1,4-galactosyltransferase I) and ST6Gal-I (α-2,6-sialyltransferase I) form a complex that increases their enzymatic activity, suggesting that these two *N*-glycosyltransferases act cooperatively in *N*-glycan synthesis.[31] Additionally, we noticed that both galactosylation and sialylation of IgG1 tend to negatively correlate with the level of core fucosylation. IgG galactosylation and fucosylation changes going in different directions have been previously described for juvenile chronic arthritis and rheumatoid arthritis.[32,33]

For both sexes and both IgG1 Fc *N*-glycans and IgG2&3 Fc *N*-glycans, we confirmed an increasing incidence of bisecting GlcNAc with age.[17−20,23] In accordance with Yamada et al. who have reported the level of bisecting GlcNAc reaching a plateau at the age of 50 years,[18] we did not detect a correlation between age and bisecting GlcNAc in older individuals.

The addition of a bisecting GlcNAc by GnT-III (β-1,4-*N*-acetylglucosaminyltransferase III) has been shown to prohibit the subsequent addition of a core fucose.[34] In line with this, we found for the IgG1 subclass a significant negative correlation between bisection and fucosylation. Furthermore, GnT-III has also been reported to negatively affect the addition of galactose by GalT (β-1,4-galactosyltransferase).[35] In our study, we observed a weak negative correlation between galactosylation and bisection for both IgG1 Fc and IgG2&3 Fc over the whole age range. Interestingly, upon stratification to younger and older groups of individuals, correlation coefficients of almost the same strength but different direction were revealed. While the younger group followed the direction of the whole age range showing a negative correlation of IgG1 galactosylation and bisection, at older age, this correlation appeared to be positive, indicating the occurrence of pronounced changes in the regulation of IgG Fc glycosylation with increasing age.

IgG glycosylation of our cohort was analyzed by hydrophilic interaction chromatography and by MALDI-TOF mass spectrometry. In the first approach, a mixture of released Fab and Fc *N*-glycans of all four IgG subclasses was analyzed in detail.[20] In the second approach, presented in this article, we have thoroughly analyzed IgG Fc glycosylation in a subclass-specific manner. Very strong correlation between galactosylation levels determined by HILIC and those determined by MS of glycopeptides unambiguously demonstrates the high quality of both data sets. One should keep in mind that the very limited scattering of the data is only in part caused by measurement inaccuracies and may for a large part be caused by the inclusion of Fab glycans in the HILIC but not in the MS analyses.

By comparing glycosylation features measured by both methods, we observed many parallels, next to some striking differences: HILIC analysis of total IgG revealed higher levels of galactosylation and bisection in neutral glycoforms and sialylation, which was on average 7-fold higher for HILIC than for IgG1 Fc and IgG2&3 Fc. Increased levels of these glycosylation features observed in the HILIC analysis may largely be linked to the inclusion of Fab glycans, as Fab glycans are known to show an increased incidence of fully galactosylated, sialylated, and bisected structures compared to the Fc.[12,14,15,36] However, the analysis of sialylation levels by mass spectrometry is complicated by two phenomena: first, the charge introduced by the sialic acid will influence ionization. It may be assumed that negative-mode ionization of the sialylated species was more efficient than the ionization of glycopeptides with nonsialylated, neutral glycan chains. Second, sialic acids are known to be labile and will, therefore, be lost to some extent upon MALDI-ionization via in-source decay or metastable decay.[37] These phenomena may explain the lower relative signal heights of sialylated species observed in negative linear mode MALDI-TOF-MS as compared to their detection using HILIC with fluorescence analysis.

HILIC glycosylation profiling of the total IgG revealed a significant negative correlation of bisecting GlcNAc and sialylation, while in subclass-specific Fc glycosylation profiling, this association was significantly positive for both IgG1 and IgG2&3 regardless of the age. The presence of bisecting GlcNAc is known to negatively affect the addition of a galactose,[35] and since galactosylated glycoforms represent a substrate for sialylatransferase, the negative effect of bisecting GlcNAc on galactosylation could also inhibit sialylation by decreasing the level of galactosylated substrate.

The HILIC and Fc glycopeptide MALDI-TOF-MS IgG glycosylation profiling methods differ in various respects. First, sample preparation is different. After the common IgG affinity purification step, HILIC of fluorescently labeled glycans requires enzymatic glycan release, labeling, and sample cleanup, while the glycopeptides profiling with MALDI-TOF-MS stays more closely to a proteomics workflow with trypsin treatment and SPE. Alternatively, trypsin treatment may be directly followed by reverse phase LC−MS analysis of glycopeptides, thereby minimizing sample preparation steps.[38] Second, HILIC and MALDI-TOF-MS differ in sample throughput. While HILIC has been tremendously speeded up by UPLC technology allowing the analysis of a couple of samples per hour, the speed of MALDI-TOF-MS in mass spectrometric profiling is much faster. Third, the sample purity requirements of the two analysis methods are very different. HILIC of released glycans is quite sensitive to contaminants such as other glycoproteins, as the released glycans from glycoprotein contaminants will interfere with the IgG glycosylation profile. Hence, high sample purity is pivotal. In contrast, MALDI-TOF-MS of glycopeptides allows distinguishing between IgG Fc glycopeptides and glycopeptides of other glycoproteins on the basis of the mostly different masses of the peptide moieties, and the presence of low quantity amounts of contaminating glycoproteins will therefore, in most cases, not interfere with IgG Fc glycosylation profiling. Fourth, the assignment of glycans to the specific Fc glycosylation sites of IgG subclasses as achieved by the MALDI-TOF-MS method is pivotal for deducing the functional implications of the observed glycosylation features. IgG Fc glycans have very distinct functions as they modulate the interaction with Fc receptors[2,5−7,11] and other cell-surface receptors of immune cells.[9,10] Moreover, most of the modulating effects of IgG Fc glycans have been reported for IgG1[9,11,39] and may not apply to IgG2, stressing the relevance of subclass and site specific IgG glycosylation profiling as achieved by mass spectrometry of glycopeptides.

■ ASSOCIATED CONTENT

Supporting Information

Age dependence of IgG1 and IgG2&3 glycoforms; MALDI-TOF-MS spectra; correlation coefficients of IgG glycosylation features and age stratified for sex for the IgG1 Fc, IgG2&3 Fc, and the total IgG; descriptives of glycosylation features in females and males; calculated monoisotopic m/z values of 2-AA labeled glycans; comparison of relative abundance of *N*-glycan. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*Fax: +31-71-526-6907. E-mail: m.wuhrer@lumc.nl.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported by the EC FP7 program (HighGlycan project) under grant agreement 278535 and the Croatian Ministry of Science, Education, and Sport grant #309-0061194-2023. M.P.B. gratefully acknowledges financial support of the Federation of European Biochemical Societies for her study visit to the LUMC where mass spectrometric analysis was performed. We thank Agnes Hipgrave-Ederveen for technical support.

■ REFERENCES

(1) Arnold, J. N.; Wormald, M. R.; Sim, R. B.; Rudd, P. M.; Dwek, R. A. The impact of glycosylation on the biological function and structure of human immunoglobulins. *Annu. Rev. Immunol.* **2007**, *25*, 21−50.

(2) Huhn, C.; Selman, M. H.; Ruhaak, L. R.; Deelder, A. M.; Wuhrer, M. IgG glycosylation analysis. *Proteomics* **2009**, *9* (4), 882−913.

(3) Jefferis, R. Glycosylation of recombinant antibody therapeutics. *Biotechnol. Prog.* **2005**, *21* (1), 11−16.

(4) Mimura, Y.; Sondermann, P.; Ghirlando, R.; Lund, J.; Young, S. P.; Goodall, M.; Jefferis, R. Role of oligosaccharide residues of IgG1-Fc in Fc gamma RIIb binding. *J. Biol. Chem.* **2001**, *276* (49), 45539−45547.

(5) Iida, S.; Misaka, H.; Inoue, M.; Shibata, M.; Nakano, R.; Yamane-Ohnuki, N.; Wakitani, M.; Yano, K.; Shitara, K.; Satoh, M. Nonfucosylated therapeutic IgG1 antibody can evade the inhibitory effect of serum immunoglobulin G on antibody-dependent cellular cytotoxicity through its high binding to FcgammaRIIIa. *Clin. Cancer Res.* **2006**, *12* (9), 2879−2887.

(6) Shields, R. L.; Lai, J.; Keck, R.; O'Connell, L. Y.; Hong, K.; Meng, Y. G.; Weikert, S. H.; Presta, L. G. Lack of fucose on human IgG1 N-linked oligosaccharide improves binding to human Fcgamma RIII and antibody-dependent cellular toxicity. *J. Biol. Chem.* **2002**, *277* (30), 26733−26740.

(7) Ferrara, C.; Grau, S.; Jager, C.; Sondermann, P.; Brunker, P.; Waldhauer, I.; Hennig, M.; Ruf, A.; Rufer, A. C.; Stihle, M.; Umana, P.; Benz, J. Unique carbohydrate−carbohydrate interactions are required for high affinity binding between FcgammaRIII and antibodies lacking core fucose. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108* (31), 12669−12674.

(8) Kaneko, Y.; Nimmerjahn, F.; Ravetch, J. V. Anti-inflammatory activity of immunoglobulin G resulting from Fc sialylation. *Science* **2006**, *313* (5787), 670−673.

(9) Anthony, R. M.; Wermeling, F.; Karlsson, M. C.; Ravetch, J. V. Identification of a receptor required for the anti-inflammatory activity of IVIG. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105* (50), 19571−19578.

(10) Anthony, R. M.; Kobayashi, T.; Wermeling, F.; Ravetch, J. V. Intravenous gammaglobulin suppresses inflammation through a novel T(H)2 pathway. *Nature* **2011**, *475* (7354), 110−113.

(11) Karsten, C. M.; Pandey, M. K.; Figge, J.; Kilchenstein, R.; Taylor, P. R.; Rosas, M.; McDonald, J. U.; Orr, S. J.; Berger, M.; Petzold, D.; Blanchard, V.; Winkler, A.; Hess, C.; Reid, D. M.; Majoul, I. V.; Strait, R. T.; Harris, N. L.; Kohl, G.; Wex, E.; Ludwig, R.; Zillikens, D.; Nimmerjahn, F.; Finkelman, F. D.; Brown, G. D.; Ehlers, M.; Kohl, J. Anti-inflammatory activity of IgG1 mediated by Fc galactosylation and association of FcgammaRIIB and dectin-1. *Nat. Med.* **2012**, *18* (9), 1401−1406.

(12) Holland, M.; Yagi, H.; Takahashi, N.; Kato, K.; Savage, C. O.; Goodall, D. M.; Jefferis, R. Differential glycosylation of polyclonal IgG, IgG-Fc and IgG-Fab isolated from the sera of patients with ANCA-associated systemic vasculitis. *Biochim. Biophys. Acta* **2006**, *1760* (4), 669−677.

(13) Stadlmann, J.; Pabst, M.; Kolarich, D.; Kunert, R.; Altmann, F. Analysis of immunoglobulin glycosylation by LC−ESI−MS of glycopeptides and oligosaccharides. *Proteomics* **2008**, *8* (14), 2858−2871.

(14) Wormald, M. R.; Rudd, P. M.; Harvey, D. J.; Chang, S. C.; Scragg, I. G.; Dwek, R. A. Variations in oligosaccharide−protein interactions in immunoglobulin G determine the site-specific glycosylation profiles and modulate the dynamic motion of the Fc oligosaccharides. *Biochemistry* **1997**, *36* (6), 1370−1380.

(15) Youings, A.; Chang, S. C.; Dwek, R. A.; Scragg, I. G. Site-specific glycosylation of human immunoglobulin G is altered in four rheumatoid arthritis patients. *Biochem. J.* **1996**, *314* (Pt 2), 621−630.

(16) Parekh, R.; Roitt, I.; Isenberg, D.; Dwek, R.; Rademacher, T. Age-related galactosylation of the N-linked oligosaccharides of human serum IgG. *J. Exp. Med.* **1988**, *167* (5), 1731−1736.

(17) Shikata, K.; Yasuda, T.; Takeuchi, F.; Konishi, T.; Nakata, M.; Mizuochi, T. Structural changes in the oligosaccharide moiety of human IgG with aging. *Glycoconjugate J.* **1998**, *15* (7), 683−689.

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

(19) Ruhaak, L. R.; Uh, H. W.; Beekman, M.; Koeleman, C. A.; Hokke, C. H.; Westendorp, R. G.; Wuhrer, M.; Houwing-Duistermaat, J. J.; Slagboom, P. E.; Deelder, A. M. Decreased levels of bisecting GlcNAc glycoforms of IgG are associated with human longevity. *PLoS One* **2010**, *5* (9), e12566.

(20) Pučić, M.; Knežević, A.; Vidič, J.; Adamczyk, B.; Novokmet, M.; Polašek, O.; Gornik, O.; Šupraha-Goreta, S.; Wormald, M. R.; Redžić, I.; Campbell, H.; Wright, A.; Hastie, N. D.; Wilson, J. F.; Rudan, I.; Wuhrer, M.; Rudd, P. M.; Josić, D.; Lauc, G. High throughput isolation and glycosylation analysis of IgG-variability and heritability of the IgG glycome in three isolated human populations. *Mol. Cell. Proteomics* **2011**, *10* (10), M111 010090.

(21) Rudan, I.; Campbell, H.; Rudan, P. Genetic epidemiological studies of eastern Adriatic Island isolates, Croatia: objective and strategies. *Coll. Antropol.* **1999**, *23* (2), 531−546.

(22) Rudan, I.; Biloglav, Z.; Vorko-Jovic, A.; Kujundzic-Tiljak, M.; Stevanovic, R.; Ropac, D.; Puntaric, D.; Cucevic, B.; Salzer, B.; Campbell, H. Effects of inbreeding, endogamy, genetic admixture, and outbreeding on human health: a (1001 Dalmatians) study. *Croat. Med. J.* **2006**, *47* (4), 601−610.

(23) Selman, M. H.; McDonnell, L. A.; Palmblad, M.; Ruhaak, L. R.; Deelder, A. M.; Wuhrer, M. Immunoglobulin G glycopeptide profiling by matrix-assisted laser desorption ionization Fourier transform ion cyclotron resonance mass spectrometry. *Anal. Chem.* **2010**, *82* (3), 1073−1081.

(24) Selman, M. H.; Hoffmann, M.; Zauner, G.; McDonnell, L. A.; Balog, C. I.; Rapp, E.; Deelder, A. M.; Wuhrer, M. MALDI-TOF-MS analysis of sialylated glycans and glycopeptides using 4-chloro-alpha-cyanocinnamic acid matrix. *Proteomics* **2012**, *12* (9), 1337−1348.

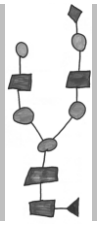
(25) Jaskolla, T. W.; Lehmann, W. D.; Karas, M. 4-Chloro-alpha-cyanocinnamic acid is an advanced, rationally designed MALDI matrix. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105* (34), 12200−12205.

(26) Selman, M. H.; Niks, E. H.; Titulaer, M. J.; Verschuuren, J. J.; Wuhrer, M.; Deelder, A. M. IgG fc N-glycosylation changes in

Lambert−Eaton myasthenic syndrome and myasthenia gravis. *J. Proteome Res.* **2011**, *10* (1), 143−152.
(27) Parekh, R. B.; Dwek, R. A.; Sutton, B. J.; Fernandes, D. L.; Leung, A.; Stanworth, D.; Rademacher, T. W.; Mizuochi, T.; Taniguchi, T.; Matsuta, K.; et al. Association of rheumatoid arthritis and primary osteoarthritis with changes in the glycosylation pattern of total serum IgG. *Nature* **1985**, *316* (6027), 452−457.
(28) Vanhooren, V.; Desmyter, L.; Liu, X. E.; Cardelli, M.; Franceschi, C.; Federico, A.; Libert, C.; Laroy, W.; Dewaele, S.; Contreras, R.; Chen, C. N-glycomic changes in serum proteins during human aging. *Rejuvenation Res.* **2007**, *10* (4), 521−531a.
(29) van de Geijn, F. E.; Wuhrer, M.; Selman, M. H.; Willemsen, S. P.; de Man, Y. A.; Deelder, A. M.; Hazes, J. M.; Dolhain, R. J. Immunoglobulin G galactosylation and sialylation are associated with pregnancy-induced improvement of rheumatoid arthritis and the postpartum flare: results from a large prospective cohort study. *Arthritis Res. Ther.* **2009**, *11* (6), R193.
(30) Wang, J.; Balog, C. I.; Stavenhagen, K.; Koeleman, C. A.; Scherer, H. U.; Selman, M. H.; Deelder, A. M.; Huizinga, T. W.; Toes, R. E.; Wuhrer, M. Fc-glycosylation of IgG1 is modulated by B-cell stimuli. *Mol. Cell. Proteomics* **2011**, *10* (5), M110 004655.
(31) Hassinen, A.; Pujol, F. M.; Kokkonen, N.; Pieters, C.; Kihlstrom, M.; Korhonen, K.; Kellokumpu, S. Functional organization of Golgi N- and O-glycosylation pathways involves pH-dependent complex formation that is impaired in cancer cells. *J. Biol. Chem.* **2011**, *286* (44), 38329−40.
(32) Flogel, M.; Lauc, G.; Gornik, I.; Macek, B. Fucosylation and galactosylation of IgG heavy chains differ between acute and remission phases of juvenile chronic arthritis. *Clin. Chem. Lab. Med.* **1998**, *36* (2), 99−102.
(33) Gornik, I.; Maravic, G.; Dumic, J.; Flogel, M.; Lauc, G. Fucosylation of IgG heavy chains is increased in rheumatoid arthritis. *Clin. Biochem.* **1999**, *32* (8), 605−608.
(34) Ferrara, C.; Brunker, P.; Suter, T.; Moser, S.; Puntener, U.; Umana, P. Modulation of therapeutic antibody effector functions by glycosylation engineering: influence of Golgi enzyme localization domain and co-expression of heterologous beta1, 4-*N*-acetylglucosaminyltransferase III and Golgi alpha-mannosidase II. *Biotechnol. Bioeng.* **2006**, *93* (5), 851−861.
(35) Fukuta, K.; Abe, R.; Yokomatsu, T.; Omae, F.; Asanagi, M.; Makino, T. Control of bisecting GlcNAc addition to N-linked sugar chains. *J. Biol. Chem.* **2000**, *275* (31), 23456−61.
(36) Mimura, Y.; Ashton, P. R.; Takahashi, N.; Harvey, D. J.; Jefferis, R. Contrasting glycosylation profiles between Fab and Fc of a human IgG protein studied by electrospray ionization mass spectrometry. *J. Immunol. Methods* **2007**, *326* (1−2), 116−126.
(37) Harvey, D. J.; Hunter, A. P.; Bateman, R. H.; Brown, J.; Critchley, G. Relationship between in-source and post-source fragment ions in the matrix-assisted laser desorption (ionization) mass spectra of carbohydrates recorded with reflectron time-of-flight mass spectrometers. *Int. J. Mass Spectrom.* **1999**, *188* (1−2), 131−146.
(38) Selman, M. H.; Derks, R. J.; Bondt, A.; Palmblad, M.; Schoenmaker, B.; Koeleman, C. A.; van de Geijn, F. E.; Dolhain, R. J.; Deelder, A. M.; Wuhrer, M. Fc specific IgG glycosylation profiling by robust nano-reverse phase HPLC-MS using a sheath-flow ESI sprayer interface. *J. Proteomics* **2012**, *75* (4), 1318−1329.
(39) Davies, J.; Jiang, L.; Pan, L. Z.; LaBarre, M. J.; Anderson, D.; Reff, M. Expression of GnTIII in a recombinant anti-CD20 CHO production cell line: Expression of antibodies with altered glycoforms leads to an increase in ADCC through higher affinity for FC gamma RIII. *Biotechnol. Bioeng.* **2001**, *74* (4), 288−294.

Loci associated with *N*-Glycosylation of human immunoglobulin G show pleiotropy with autoimmune diseases and haematological cancers

4



OPEN ACCESS Freely available online

PLOS | GENETICS

Loci Associated with *N*-Glycosylation of Human Immunoglobulin G Show Pleiotropy with Autoimmune Diseases and Haematological Cancers

Gordan Lauc[1,2☯], Jennifer E. Huffman[3☯], Maja Pučić[1☯], Lina Zgaga[4,5☯], Barbara Adamczyk[6☯], Ana Mužinić[1], Mislav Novokmet[1], Ozren Polašek[7], Olga Gornik[2], Jasminka Krištić[1], Toma Keser[2], Veronique Vitart[3], Blanca Scheijen[8], Hae-Won Uh[9,10], Mariam Molokhia[11], Alan Leslie Patrick[12], Paul McKeigue[4], Ivana Kolčić[7], Ivan Krešimir Lukić[7], Olivia Swann[4], Frank N. van Leeuwen[8], L. Renee Ruhaak[13], Jeanine J. Houwing-Duistermaat[9], P. Eline Slagboom[10,14], Marian Beekman[10,14], Anton J. M. de Craen[15], André M. Deelder[16], Qiang Zeng[17], Wei Wang[18,19,20], Nicholas D. Hastie[3], Ulf Gyllensten[21], James F. Wilson[4], Manfred Wuhrer[16], Alan F. Wright[3], Pauline M. Rudd[6¶], Caroline Hayward[3¶], Yurii Aulchenko[4,22¶], Harry Campbell[4¶], Igor Rudan[4¶*]

1 Glycobiology Laboratory, Genos, Zagreb, Croatia, **2** Faculty of Pharmacy and Biochemistry, University of Zagreb, Zagreb, Croatia, **3** Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh, United Kingdom, **4** Centre for Population Health Sciences, The University of Edinburgh Medical School, Edinburgh, United Kingdom, **5** Faculty of Medicine, University of Zagreb, Zagreb, Croatia, **6** National Institute for Bioprocessing Research and Training, Dublin-Oxford Glycobiology Laboratory, Dublin, Ireland, **7** Faculty of Medicine, University of Split, Split, Croatia, **8** Radboud University Nijmegen Medical Centre, Nijmegen, The Netherlands, **9** Department of Medical Statistics and Bioinformatics, Leiden University Medical Center, Leiden, The Netherlands, **10** Netherlands Consortium for Healthy Aging, Leiden, The Netherlands, **11** Department of Primary Care and Public Health Sciences, Kings College London, London, United Kingdom, **12** Kavanagh St. Medical Centre, Port of Spain, Trinidad and Tobago, **13** Department of Chemistry, University of California Davis, Davis, California, United States of America, **14** Department of Molecular Epidemiology, Leiden University Medical Center, Leiden, The Netherlands, **15** Department of Gerontology and Geriatrics, Leiden University Medical Center, Leiden, The Netherlands, **16** Biomolecular Mass Spectrometry Unit, Department of Parasitology, Leiden University Medical Center, Leiden, The Netherlands, **17** Chinese PLA General Hospital, Beijing, China, **18** School of Public Health, Capital Medical University, Beijing, China, **19** Graduate University, Chinese Academy of Sciences, Beijing, China, **20** School of Medical Sciences, Edith Cowan University, Perth, Australia, **21** Department of Immunology, Genetics, and Pathology, SciLifeLab Uppsala, Rudbeck Laboratory, Uppsala University, Uppsala, Sweden, **22** Institute of Cytology and Genetics SD RAS, Novosibirsk, Russia

Abstract

Glycosylation of immunoglobulin G (IgG) influences IgG effector function by modulating binding to Fc receptors. To identify genetic loci associated with IgG glycosylation, we quantitated *N*-linked IgG glycans using two approaches. After isolating IgG from human plasma, we performed 77 quantitative measurements of *N*-glycosylation using ultra-performance liquid chromatography (UPLC) in 2,247 individuals from four European discovery populations. In parallel, we measured IgG *N*-glycans using MALDI-TOF mass spectrometry (MS) in a replication cohort of 1,848 Europeans. Meta-analysis of genome-wide association study (GWAS) results identified 9 genome-wide significant loci ($P<2.27\times10^{-9}$) in the discovery analysis and two of the same loci (*B4GALT1* and *MGAT3*) in the replication cohort. Four loci contained genes encoding glycosyltransferases (*ST6GAL1*, *B4GALT1*, *FUT8*, and *MGAT3*), while the remaining 5 contained genes that have not been previously implicated in protein glycosylation (*IKZF1*, *IL6ST-ANKRD55*, *ABCF2-SMARCD3*, *SUV420H1*, and *SMARCB1-DERL3*). However, most of them have been strongly associated with autoimmune and inflammatory conditions (e.g., systemic lupus erythematosus, rheumatoid arthritis, ulcerative colitis, Crohn's disease, diabetes type 1, multiple sclerosis, Graves' disease, celiac disease, nodular sclerosis) and/or haematological cancers (acute lymphoblastic leukaemia, Hodgkin lymphoma, and multiple myeloma). Follow-up functional experiments in haplodeficient *Ikzf1* knock-out mice showed the same general pattern of changes in IgG glycosylation as identified in the meta-analysis. As *IKZF1* was associated with multiple IgG *N*-glycan traits, we explored biomarker potential of affected *N*-glycans in 101 cases with SLE and 183 matched controls and demonstrated substantial discriminative power in a ROC-curve analysis (area under the curve = 0.842). Our study shows that it is possible to identify new loci that control glycosylation of a single plasma protein using GWAS. The results may also provide an explanation for the reported pleiotropy and antagonistic effects of loci involved in autoimmune diseases and haematological cancer.

Citation: Lauc G, Huffman JE, Pučić M, Zgaga L, Adamczyk B, et al. (2013) Loci Associated with *N*-Glycosylation of Human Immunoglobulin G Show Pleiotropy with Autoimmune Diseases and Haematological Cancers. PLoS Genet 9(1): e1003225. doi:10.1371/journal.pgen.1003225

Editor: Greg Gibson, Georgia Institute of Technology, United States of America

Received August 30, 2012; **Accepted** November 21, 2012; **Published** January 31, 2013

Funding: The CROATIA-Vis study in the Croatian island of Vis was supported by grants from the Medical Research Council (UK); the Ministry of Science, Education, and Sport of the Republic of Croatia (grant number 108-1080315-0302); and the European Union framework program 6 European Special Populations Research Network project (contract LSHG-CT-2006-018947). ORCADES was supported by the Chief Scientist Office of the Scottish Government, the Royal Society, and the European Union Framework Programme 6 EUROSPAN project (contract LSHG-CT-2006-018947). Glycome analysis was supported by the Croatian Ministry of Science, Education, and Sport (grant numbers 309-0061194-2023 and 216-1080315-0302); the Croatian Science Foundation (grant number 04-47); the European Commission EuroGlycoArrays (contract #215536), GlycoBioM (contract #259869), and HighGlycan (contract #278535) grants; and National Natural Science Foundation and Ministry of Science and Technology, China (grant numbers NSFC31070727 and 2012BAI37B03). The Leiden Longevity Study was supported by a grant from the Innovation-Oriented Research Program on Genomics (SenterNovem IGE05007), the Centre for Medical Systems Biology, and the Netherlands Consortium for Healthy Ageing (grant 050-060-810), all in the framework of the Netherlands Genomics Initiative, Netherlands Organization for Scientific Research (NWO). The research leading to these results has received funding from the European Union's Seventh Framework Programme (FP7/2007-2011) under grant agreement 259679. Research for the Systemic Lupus Erythematosus (SLE) cases and con_rols series was supported with help from the ARUK (formerly ARC) grants M0600 and M0651 and by the National Institute for Health Research (NIHR) Biomedical Research Centre at Guy's and St. Thomas' NHS Foundation Trust and King's College London. The work of YA was supported by Helmholtz-RFBR JRG grant (12-04-91322-СИГ_a). BS acknowledges supported from a grant from the Dutch childhood oncology foundation KiKa. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

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

* E-mail: Igor.rudan@ed.ac.uk

☯ These authors contributed equally to this work.

¶ These authors also contributed equally to this work.

Introduction

Glycosylation is a ubiquitous post-translational protein modification that modulates the structure and function of polypeptide components of glycoproteins [1,2]. *N*-glycan structures are essential for multicellular life [3]. Mutations in genes involved in modification of glycan antennae are common and can lead to severe or fatal diseases [4]. Variation in protein glycosylation also has physiological significance, with immunoglobulin G (IgG) being a well-documented example. Each heavy chain of IgG carries a single covalently attached bi-antennary *N*-glycan at the highly conserved asparagine 297 residue in each of the CH2 domains of the Fc region of the molecule. The attached oligosaccharides are structurally important for the stability of the antibody and its effector functions [5]. In addition, some 15–20% of normal IgG molecules have complex bi-antennary oligosaccharides in the variable regions of light or heavy chains [6,7]. 36 different glycans (Figure 1) can be attached to the conserved Asn297 of the IgG heavy chain [8,9], leading to hundreds of different IgG isomers that can be generated from this single glycosylation site.

Glycosylation of IgG has important regulatory functions. The absence of galactose residues in association with rheumatoid arthritis was reported nearly 30 years ago [10]. The addition of sialic acid dramatically changes the physiological role of IgGs, converting them from pro-inflammatory to anti-inflammatory agents [11,12]. Addition of fucose to the glycan core interferes with the binding of IgG to FcγRIIIa and greatly diminishes its capacity for antibody dependent cell-mediated cytotoxicity (ADCC) [13,14]. Structural analysis of the IgG-Fc/FcγRIIIa complex has demonstrated that specific glycans on FcγRIIIa are also essential for this effect of core-fucose [15] and that removal of core fucose from IgG glycans increases clinical efficacy of monoclonal antibodies, enhancing their therapeutic effect through ADCC mediated killing [16–18].

New high-throughput technologies, such as high/ultra performance liquid chromatography (HPLC/UPLC), MALDI-TOF mass spectrometry (MS) and capillary electrophoresis (CE), allow us to quantitate *N*-linked glycans from individual human plasma proteins. Recently, we performed the first population-based study to demonstrate physiological variation in IgG glycosylation in three European founder populations [19]. Using UPLC, we showed exceptionally high individual variability in glycosylation of a single protein - human IgG - and substantial heritability of the observed measurements [19]. In parallel, we quantitated IgG *N*-glycans in another European population (Leiden Longevity Study – LLS) by mass spectrometry. In this study, we combined those high-throughput glycomics measurements with high-throughput genomics to perform the first genome wide association (GWA) study of the human IgG *N*-glycome.

Results

Genome-wide association study and meta-analysis

We separated a single protein (IgG) from human plasma and quantitated its *N*-linked glycans using two state-of-the-art technologies (UPLC and MALDI-TOF MS). Their comparative advantages in GWA studies were difficult to predict prior to the conducted analyses, so both were used - one in each available cohort. We performed 77 quantitative measurements of IgG *N*-glycosylation using ultra performance liquid chromatography (UPLC) in 2247 individuals from four European discovery populations (CROATIA-Vis, CROATIA-Korcula, ORCADES, NSPHS). In parallel, we measured IgG *N*-glycans using MALDI-TOF mass spectrometry (MS) in 1848 individuals from another European population (Leiden Longevity Study (LLS)). Descriptions of these population cohorts are found in Table S11. Aiming to identify genetic loci involved in IgG glycosylation, we performed a GWA study in both cohorts. Associations at 9 loci reached genome-wide significance ($P<2.27\times10^{-9}$) in the discovery meta-analysis and at two loci in the replication cohort. The two loci identified in the latter cohort were associated with the analogous glycan traits in the former cohort as detailed in the subsection "Replication of our findings". Both UPLC and MS methods for quantitation of *N*-glycans were found to be amenable to GWA studies. Since our UPLC study gave a considerably greater yield of significant findings in comparison to MS study, the majority of our results section focuses on the findings from the discovery population cohort, which was studied using the UPLC method.

Among the nine loci that passed the genome-wide significance threshold, four contained genes encoding glycosyltransferases (*ST6GAL1*, *B4GALT1*, *FUT8* and *MGAT3*), while the remaining five loci contained genes that have not been implicated in protein glycosylation previously (*IKZF1, IL6ST-ANKRD55, ABCF2-SMARCD3, SUV420H1-CHKA* and *SMARCB1-DERL3*). As a rule, the implicated genes were associated with several *N*-glycan traits. The explanation and notation of the 77 N-glycan measures is presented in Table S1. It comprises 23 directly measured quantitative IgG glycosylation traits (shown in Figure 1) and 54 derived traits. Descriptive statistics of these measures in the discovery cohorts are presented in Table S2. GWA analysis was performed in each of the populations separately and the results were combined in an inverse-variance weighted meta-analysis. Summary data for each gene region showing genome-wide association ($p<27.2\times10^{-9}$) or found to be strongly suggestive ($2.27\times10^{-9}<p<5\times10^{-8}$) are presented in Table 1. Summary data for all single-nucleotide polymorphisms (SNPs) and traits with suggestive associations ($p<1\times10^{-5}$) are presented in Table S3, with population-specific and pooled genomic control (GC) factors reported in Table S4.

The most statistically significant association was observed in a region on chromosome 3 containing the gene *ST6GAL1* (Table 1, Figure S1A). *ST6GAL1* codes for the enzyme sialyltransferase 6 which adds sialic acid to various glycoproteins including IgG glycans (Figure 2), and is therefore a highly biologically plausible candidate. In this region of about 70 kilobases (kb) we identified 37genome-wide significant SNPs associated with 14 different IgG glycosylation traits, generally reflecting sialylation of different glycan structures (Table 1). The strongest association was observed for the percentage of monosialylation of fucosylated digalactosylated structures in total IgG glycans (IGP29, see Figure 1 and Table S1 for notation), for which a SNP rs11710456 explained 17%, 16%, 18% and 3% of the trait variation for CROATIA-Vis,

Author Summary

After analysing glycans attached to human immunoglobulin G in 4,095 individuals, we performed the first genome-wide association study (GWAS) of the glycome of an individual protein. Nine genetic loci were found to associate with glycans with genome-wide significance. Of these, four were enzymes that directly participate in IgG glycosylation, thus the observed associations were biologically founded. The remaining five genetic loci were not previously implicated in protein glycosylation, but the most of them have been reported to be relevant for autoimmune and inflammatory conditions and/or haematological cancers. A particularly interesting gene, *IKZF1* was found to be associated with multiple IgG *N*-glycans. This gene has been implicated in numerous diseases, including systemic lupus erythematosus (SLE). We analysed *N*-glycans in 101 cases with SLE and 183 matched controls and demonstrated their substantial biomarker potential. Our study shows that it is possible to identify new loci that control glycosylation of a single plasma protein using GWAS. Our results may also provide an explanation for opposite effects of some genes in autoimmune diseases and haematological cancer.

CROATIA-Korcula, ORCADES and NSPHS respectively (meta-analysis $p = 6.12 \times 10^{-75}$). NSPHS had a very small sample size in this analysis (N = 179) and may not provide an accurate portrayal of the variance explained in this particular population (estimated as 3%). Although the allele frequency is similar between all populations, in the forest plot (Figure S1A) although NSPHS does overlap with the other populations, the 95% CI is much larger. It is also possible that there are population-specific genetic and/or environmental differences in NSPHS that are affecting the amount of variance explained by this SNP. After analysis conditioning on the top SNP (rs11710456) in this region, the SNP rs7652995 still reached genome-wide significance ($p = 4.15 \times 10^{-13}$). After adjusting for this additional SNP, the association peak was completely removed. This suggests that there are several genetic factors underlying this association. Conditional analysis of all other significant and suggestive regions resulted in the complete removal of the association peak.

We also identified 28 SNPs showing genome-wide significant associations with 11 IgG glycosylation traits ($2.70 \times 10^{-11} < p < 4.73 \times 10^{-8}$) at a locus on chromosome 9 spanning over 60 kb (Figure S1B). This region includes *B4GALT1*, which codes for the galactosyltransferase responsible for the addition of galactose to IgG glycans (Figure 2). The glycan traits showing genome-wide association included the percentage of FA2G2S1 in the total fraction (IGP17), the percentage of FA2G2 in the total and neutral fraction (IGP13, IGP53), the percentage of sialylation of fucosylated structures without bisecting GlcNAc (IGP24, IGP26), the percentage of digalactosylated structures in the total neutral fraction (IGP57) and, in the opposite direction, the percentage of bisecting GlcNAc in fucosylated sialylated structures (IGP36–IGP40).

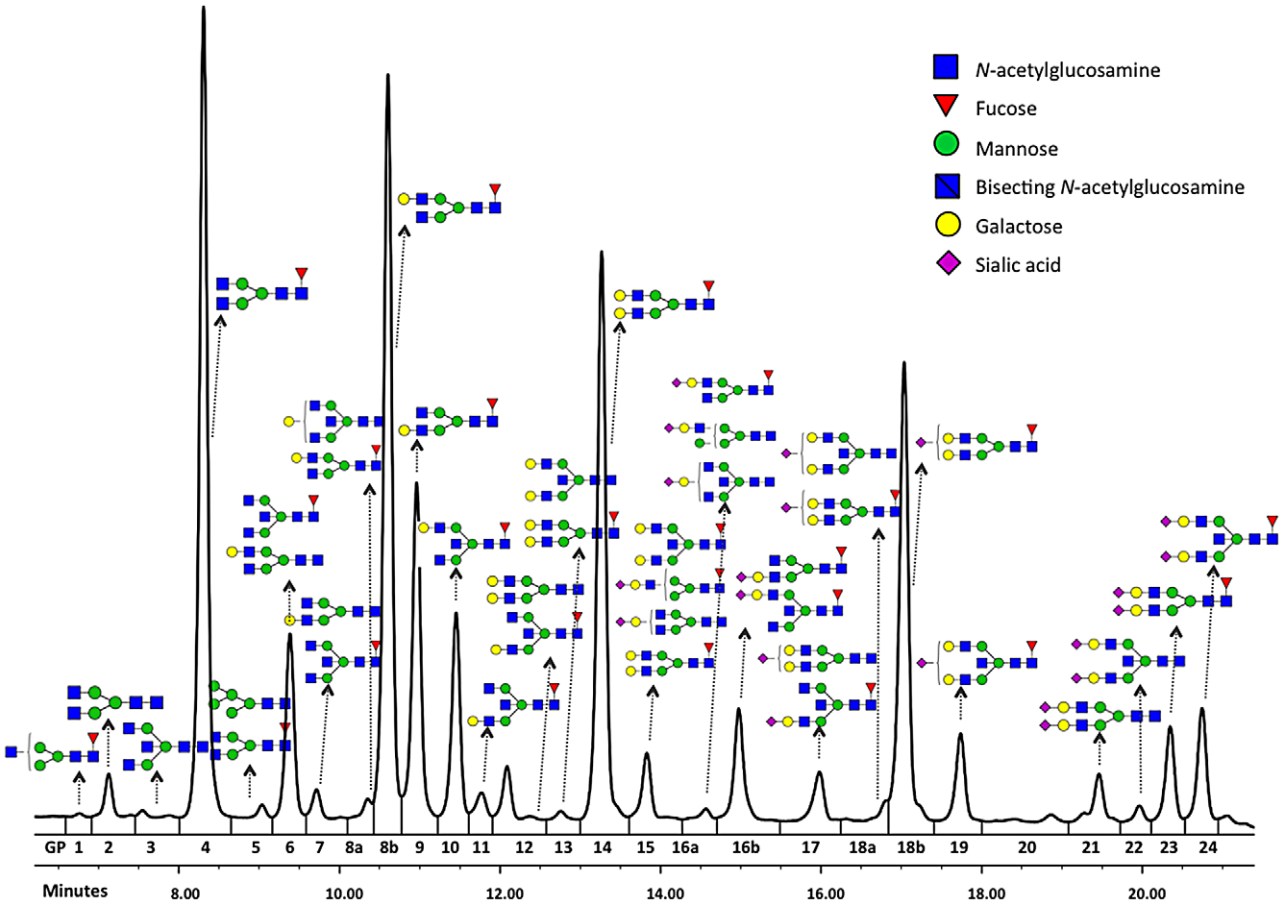


Figure 1. Structures of glycans separated by HILIC-UPLC analysis of the IgG glycome.
doi:10.1371/journal.pgen.1003225.g001

Table 1. A complete list of genetic markers that showed genome-wide significant (P<2.27E-9) or strongly suggestive (P≤5E-08) association with glycosylation of Immunoglobulin G analysed by UPLC in the discovery meta-analysis.

Chr.	SNP with lowest P-value	Lowest P-value	Effect size* (s.e.)	MAF	Interval size, kb	nHits	nTraits	Genes in the interval	Trait with lowest P-value+	Other Associated Traits+
Genome-wide Significant										
3	rs11710456	6.12E-75	0.64 (0.04)	0.30	14.2	20	14	*ST6GAL1*	IGP29	IGP14$, IGP15, IGP17, IGP23, IGP24, IGP26, IGP28, IGP30, IGP31$, IGP32, IGP35$, IGP37$, IGP38$
5	rs17348299	6.88E-11	0.29 (0.04)	0.16	16.1	4	6	*IL6ST-ANKRD55*	IGP53	IGP3, IGP13, IGP43, IGP55, IGP57
7	rs6421315	1.87E-13	0.23 (0.03)	0.37	21.4	11	13	*IKZF1*	IGP63	IGP2$, IGP6$, IGP42$, IGP46$, IGP58, IGP59, IGP60, IGP62, IGP67$, IGP70$, IGP71$, IGP72
7	rs1122979	2.10E-10	0.31 (0.05)	0.12	62.3	3	4	*ABCF2-SMARCD3*	IGP2	IGP5, IGP42, IGP45
9	rs12342831	2.70E-11	−0.24 (0.04)	0.26	60.1	28	11	*B4GALT1*	IGP17	IGP13, IGP24, IGP26, IGP36$, IGP37$, IGP38$, IGP39$, IGP40$, IGP53, IGP57
11	rs4930561	8.88E-10	0.19 (0.03)	0.49	58.7	5	2	*SUV420H1*	IGP41	IGP1
14	rs11847263	1.08E-22	−0.31 (0.03)	0.39	17.1	167	12	*FUT8*	IGP59	IGP2$, IGP6$, IGP11$, IGP42$, IGP46$, IGP51$, IGP58, IGP60, IGP61, IGP63, IGP65
22	rs2186369	8.63E-17	0.35 (0.04)	0.19	49.4	10	20	*SMARCB1-DERL3*	IGP72	IGP9$, IGP10$, IGP14$, IGP39$, IGP40$, IGP49$, IGP50$, IGP62, IGP63, IGP64, IGP66$, IGP67$, IGP68$, IGP69$, IGP70$, IGP71$, IGP74$, IGP75$, IGP76
22	rs909674	9.66E-25	0.34 (0.03)	0.30	27.9	60	17	*SYNGR1-TAB1-MGAT3-CACNA1I*	IGP40	IGP5, IGP9, IGP22$, IGP34, IGP39, IGP45, IGP49, IGP62$, IGP63$, IGP64$, IGP66, IGP67, IGP68, IGP70, IGP71, IGP72$
Strongly Suggestive										
6	rs9296009	3.79E-08	−0.21 (0.04)	0.20	–	1	1	*PRRT1*	IGP23	–
6	rs1049110	1.64E-08	0.19 (0.03)	0.35	32.3	1	2	*HLA-DQA2, HLA-DQB2*	IGP42	IGP2
6	rs404256	7.49E-09	−0.21 (0.04)	0.44	–	1	1	*BACH2*	IGP7	–
7	rs2072209	1.16E-08	−0.37 (0.07)	0.06	–	1	1	*LAMB1*	IGP69	–
9	rs4878639	3.51E-08	−0.20 (0.04)	0.26	14.4	1	1	*RECK*	IGP17	–
12	rs12828421	4.48E-08	−0.18 (0.03)	0.49	29.6	2	1	*PEX5*	IGP41	–
17	rs7224668	3.33E-08	0.17 (0.03)	0.48	45.9	2	1	*SLC38A10*	IGP31	–

Interval: size (kb) of the genomic interval containing SNPs with $R^2 >= 0.6$ with top associated SNP; nHits: number of SNPs with GW-significant association; nTraits: number of IgG glycosylation traits associated with the region at GW-significant level;

*effect size is in z-score units after adjustment for sex, age and first 3 principal components.

+Description of the traits provided in Table S1;

$the SNP effect in opposite direction to most significant trait.

doi:10.1371/journal.pgen.1003225.t001

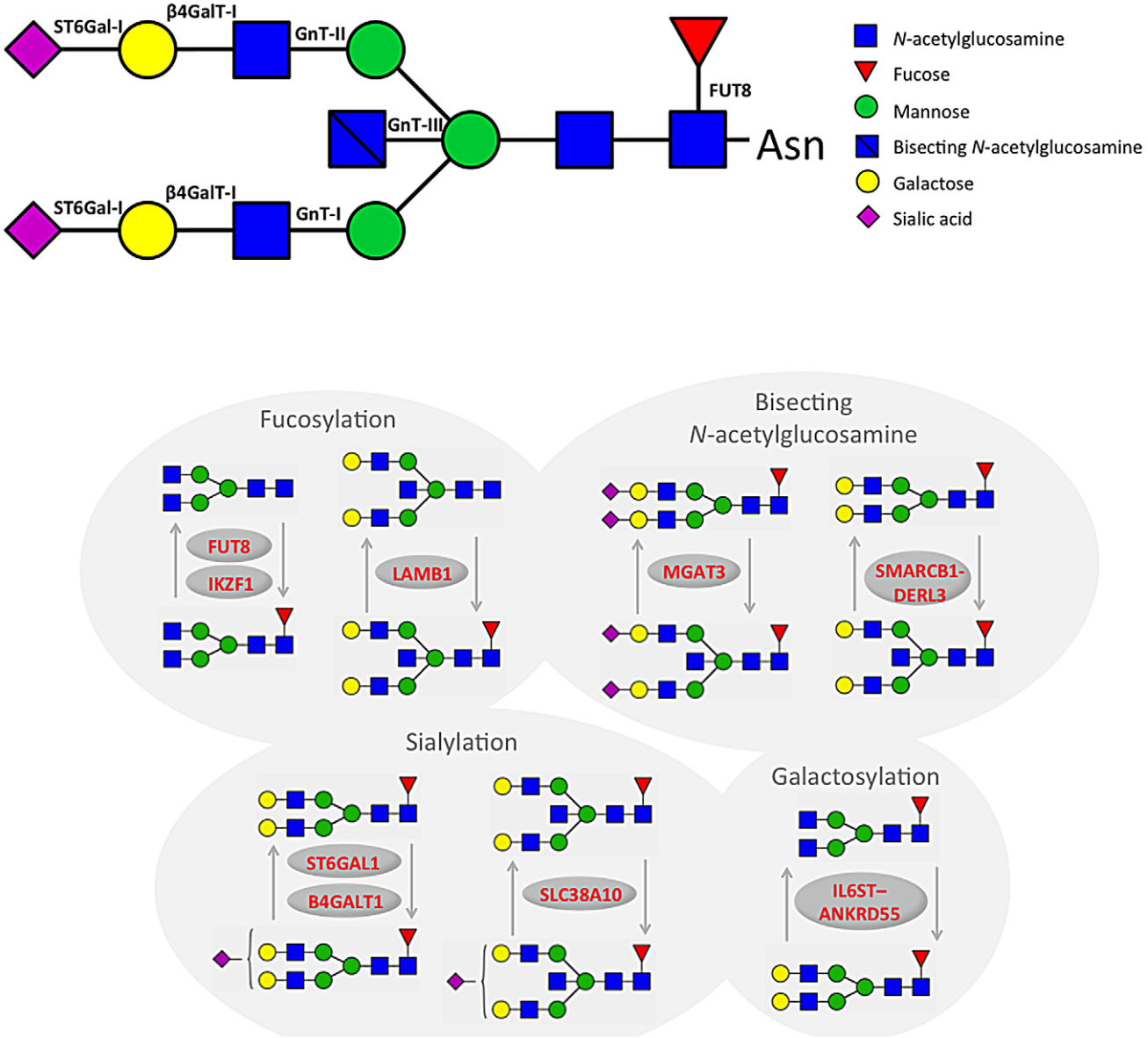


Figure 2. A summary of changes to IgG N-glycan structures that were associated with 16 loci identified through GWA study.
doi:10.1371/journal.pgen.1003225.g002

A large (541 kb) region on chromosome 14 harbouring the *FUT8* gene contained 167 SNPs showing significant associations with 12 IgG glycosylation traits reflecting fucosylation of IgG glycans (Figure S1C). *FUT8* codes for fucosyltransferase 8, an enzyme responsible for the addition of fucose to IgG glycans (Figure 2). The strongest association ($1.08\times10^{-22}<p<2.60\times10^{-17}$) was observed with the percentage of A2 glycans in total and neutral fractions (IGP2, IGP42) and for derived traits related to the proportion of fucosylation (IGP58, IGP59 and IGP61; all in the opposite direction). In summary, SNPs at the *FUT8* locus influence the proportion of fucosylated glycans, and, in the opposite direction, the percentages of A2, A2G1 and A2G2 glycans which are not fucosylated.

On chromosome 22, two loci were associated with IgG glycosylation. The first region, containing *SYNGR1-TAB1-MGAT3-CACNA1I* genes, spans over 233 kb. This region harboured 60 SNPs showing genome-wide significant association with 17 IgG glycosylation traits (Figure S1D). Association was strongest between SNP rs909674 and the incidence of bisecting GlcNAc in all fucosylated disialylated structures (IGP40, $p=9.66\times10^{-25}$) and the related ratio IGP39 ($p=8.87\times10^{-24}$). In summary, this locus contained variants influencing levels of fucosylated species and the ratio between fucosylated (especially disialylated) structures with and without bisecting GlcNAc (Figure 2). Since *MGAT3* codes for the enzyme *N*-acetylglucosaminyltransferase III (beta-1,4-mannosyl-glycoprotein-4-beta-*N*-acetylglucosaminyltransferase), which is responsible for the addition of bisecting GlcNAc to IgG glycans, this gene is the most biologically plausible candidate.

Bioinformatic analysis of known and predicted protein-protein interactions using String 9.0 software (http://string-db.org/) showed that interactions between the clusters of *FUT8-B4GALT1-MGAT3* genes and *ST6GAL1-B4GALT1-MGAT3* genes had high confidence score: *FUT8-B4GALT1* of 0.90; *FUT8-MGAT3* of 0.95; *ST6GAL1-B4GALT1* of 0.90; and *ST6GAL1-*

MGAT3 of 0.73. The glycosyltranferase genes at the four GWAS loci - *ST6GAL1*, *B4GALT1*, *FUT8*, and *MGAT3* – are responsible for adding sialic acid, galactose, fucose and bisecting GlcNAc to IgG glycans, thus demonstrating the proof of principle that a single protein glycosylation GWAS approach can identify biologically important glycan pathways and their networks. Interestingly, *ST6GAL1* has been previously associated with Type 2 diabetes [20], *MGAT3* with Crohn's disease [21], primary biliary cirrhosis [22] and cardiac arrest [23], and FUT8 with multiple sclerosis, blood glutamate levels [24] and conduct disorder [25] (Table 2). We have recently shown changes in plasma N-glycan profile between patients with attention-deficit hyperactivity disorder (ADHD), autism spectrum disorders and healthy controls, and identified loci influencing plasma *N*-glycome with pleiotropic effects on ADHD [26,27].

Novel candidate genes involved with N-glycosylation

In addition to four loci containing genes for enzymes known to be involved in IgG glycosylation, our study also found five unexpected associations showing genome-wide significance. In the second region on chromosome 22 we observed genome-wide significant associations of 10 SNPs with 20 IgG glycosylation traits. The region spans 49 kb and contains the genes *SMARCB1-DERL3* (Figure S1E). The strongest associations ($8.63\times10^{-17}<p<3.00\times10^{-13}$) were observed between SNP rs2186369 and the percentage of FA2[6]BG1 in total and neutral fractions (IGP9, IGP49) and levels of fucosylated structures with bisecting GlcNAc (IGP66, IGP68, IGP70, IGP71 in the same direction and IGP72 in the opposite direction). Thus, the *SMARCB1-DERL3* locus appears to specifically influence levels of fucosylated monogalactosylated structures with bisecting GlcNAc (Figure 2). *DERL3* is a promising functional candidate, because it encodes a functional component of endoplasmic reticulum (ER)-associated degradation for misfolded luminal glycoproteins [28]. However, *SMARCB1* is also known to be important in antiviral activity, inhibition of tumour formation, neurodevelopment, cell proliferation and differentiation [29]. The region has also been implicated in the regulation of γ-glutamyl-transferase (GGT) [30] (Table 2).

A locus on chromosome 7 spanning 26kb contained 11 SNPs showing genome-wide significant associations with 13 IgG glycosylation traits (Figure S1F). The strongest association ($p=1.87\times10^{-13}$) was observed between SNP *rs6421315* located in *IKZF1* and the percentage of fucosylation of agalactosylated structures without bisecting GlcNAc (IGP63). Thus, SNPs at this locus influence the percentage of non-fucosylated agalactosylated glycans, the fucosylation ratio in agalactosylated glycans (in opposite directions for glycan species with and without bisecting GlcNAc), and the ratio of fucosylated structures with and without bisecting GlcNAc (Figure 2). The *IKZF1* gene encodes the DNA-binding protein Ikaros, acting as a transcriptional regulator and associated with chromatin remodelling. It is considered to be the important regulator of lymphocyte differentiation and has been shown to influence effector pathways through control of class switch recombination [31], thus representing a promising functional candidate [32]. There is overwhelming evidence that *IKZF1* variants are associated with childhood acute lymphoblastic leukaemia [33,34] and several diseases with an autoimmune component: systemic lupus erythematosus (SLE) [35–37], type 1 diabetes [38,39], Crohn's disease [40], systemic sclerosis [41], malaria [42] and erythrocyte mean corpuscular volume [43] (Table 2).

SNPs at several other loci also showed genome-wide significant association with a number of different IgG glycosylation traits (Figure S1G–S1P). Chromosome 5 SNP rs17348299, located in *IL6ST-ANKRD55* was significantly associated ($6.88\times10^{-11}<p<2.39\times10^{-9}$) with six IgG glycosylation traits, including FA2 and FA2G2 in total and neutral fractions (IGP3, IGP13, IGP43, IGP53) and the percentage of agalactosylated and digalactosylated structures in total neutral IgG glycans (IGP55, IGP57) (Figure 2). The protein encoded by *IL6ST* is a signal transducer shared by many cytokines, including interleukin 6 (IL6), ciliary neurotrophic factor (CNTF), leukaemia inhibitory factor (LIF), and oncostatin M (OSM). Variants in *IL6ST* have been associated with rheumatoid arthritis and multiple myeloma, but also with components of metabolic syndrome [44–46].

The chromosome 7 SNP rs2072209 located in *LAMB1* was strongly suggestively associated with the percentage of fucosylation of digalactosylated (with bisecting GlcNAc) structures (IGP69; $p=1.16\times10^{-8}$) (Figure 2). *LAMB1* (laminin beta 1) is a member of a family of extracellular matrix glycoproteins that are the major non-collagenous constituent of basement membranes. It is thought to mediate the attachment, migration and organization of cells into tissues during embryonic development by interacting with other extracellular matrix components. It has been associated with ulcerative colitis in several large-scale studies in European and Japanese populations, suggesting that changes in the integrity of the intestinal epithelial barrier may contribute to the pathogenesis of the disease [47–51] (Table 2).

Another particularly interesting finding was the suggestive association between rs404256 in the *BACH2* gene on chromosome 6 and IGP7, defined through proportional contribution of FA2[6]G1 in all IgG glycans ($p=7.49\times10^{-9}$). *BACH2* is B-cell-specific transcription factor that can act as a suppressor or promoter; among many known functions, it has been shown to "orchestrate" transcriptional activation of B-cells, modify the cytotoxic effects of anticancer drugs and regulate IL-2 expression in umbilical cord blood $CD4^+$ T cells [52]. *BACH2* has been previously associated with a spectrum of diseases with autoimmune component: type 1 diabetes [53–56], Graves' disease [57], celiac disease [58], Crohn's disease [21] and multiple sclerosis [59] (Table 2).

The chromosome 11 SNP rs4930561 located in the *SUV420H1-CHKA* gene was associated with percentage of FA1 in neutral (IGP41; $p=8.88\times10^{-10}$) and total (IGP1; $p=1.30\times10^{-8}$) fractions of IgG glycans. *SUV420H1* codes for histone-lysine N-methyltransferase which specifically trimethylates lysine 20 of histone H4 and could therefore affect activity of many different genes; it is thought to be involved in proviral silencing in somatic and germ line cells through epigenetic mechanisms [60]. *CHKA* has a key role in phospholipid biosynthesis and may contribute to tumour cell growth. We recently reported a number of strong associations between lipidomics and glycomics traits in human plasma [61]. Thus, an enzyme involved in phospholipid synthesis is also a possible candidate because the lipid environment is known to affect glycosyltransferases activity [61].

Three further loci were identified as strongly suggestive through GWAS and deserve attention for their possible pleiotropic effects. SNP rs9296009 in *PRRT1* (proline-rich transmembrane protein 1) was associated with IGP23 ($p=3.79\times10^{-08}$) while variants in *PRRT1* previously showed associations with nodular sclerosis and Hodgkin lymphoma [62]. Moreover, rs1049110 in *HLA-DQA2-HLA-DQB2* was associated with IGP2 and IGP42 ($p=1.64\times10^{-08}$ and 4.44×10^{-08}, respectively). This SNP is in nearly complete linkage disequilibrium with two other SNPs in this region that have previously been associated with SLE and hepatitis B [63] (Table 2). Another SNP in this region has been linked with narcolepsy [64]. Finally, rs7224668 in SLC38A10, a putative sodium-dependent amino acid/proton antiporter, showed significant association with

Table 2. An analysis of pleiotropy between loci associated with IgG glycans and previously reported disease/trait susceptibility loci, with linkage disequilibrium computed between the most significantly associated SNPs.

Gene	IgG Glycan Top SNP	Disease	Top Disease SNP	Risk Allele	P-Value	Reference	Ancestry	HapMap 2		1000G Pilot 1	
								R^2	D′	R^2	D′
IKZF1	rs6421315	SLE	rs921916	C	2.00E-06	Gateva et al Nat Genet 2009	European	0.021	0.388	0.070	0.771
		SLE	rs2366293	G	2.33E-09	Cunninghame Graham et al PLoS Genet 2011	European	0.030	0.484	0.057	0.748
		SLE	rs4917014	A	3.00E-23	Han et al Nat Genet 2009	Han Chinese	0.001	0.040	0.053	0.277
		ALL	rs11978267	G	8.00E-11	Trevino et al Nat Genet 2009	European	0.002	0.047	0.012	0.130
		ALL	rs4132601	C	1.00E-19	Papeammanuil et al Nat Genet 2009	European	0.002	0.047	0.012	0.130
		Hippocampal atrophy (AD qt)	rs10276619	–	3.00E-06	Potkin et al PLoS One 2009	European	0	0.005	0	0.013
		Total ventricular volume (AD qt)	rs7805803	–	9.00E-06	Furney et al Mol Psychiatry 2010	European	0.071	0.280	0.087	0.332
		Crohn's disease	rs1456893	A	5.00E-09	Barrett et al Nat Genet 2008	European	0.011	0.117	0	0.007
		Mean corpuscular volume	rs12718597	A	5.00E-13	Ganesh et al Nat Genet 2009	European	0.018	0.181	0.019	0.161
		Malaria	rs1451375	–	6.00E-06	Jallow et al Nat Genet 2009	Gambian	0.014	0.204	0.002	0.097
		Systemic sclerosis	rs1240874	–	1.00E-06	Gorlova et al PLoSGenet 2011	European	–	–	–	–
		T1D	rs10272724	C	1.10E-11	Swafford et al Diabetes 2011	European	0.002	0.047	0.012	0.130
ST6GAL1	rs11710456	Drug-induced liver injury (flucloxacillin)	rs10937275	–	1.00E-08	Daly et al Nat Genet 2009	European	0	0.048	0.017	1
		T2D	rs16861329	G	3.00E-08	Kooner et al Nat Genet 2011	South Asian	0.011	0.221	0.000	0.005
IL6ST-ANKRD55	rs17348299	Rheumatoid arthritis	rs6859219	C	1.00E-11	Stahl et al Nat Genet 2010	European	0.012	0.487	0.044	1.000
LAMB1	rs2072209	Ulcerative colitis	rs2158836	A	7.00E-06	Silverberg et al Nat Genet 2009	European	0.027	0.716	0.055	1.000
		Ulcerative colitis	rs4598195	A	8.00E-08	McGovern et al Nat Genet 2010	European	0.071	0.807	0.115	1.000
		Ulcerative colitis	rs886774	G	3.00E-08	Barrett et al Nat Genet 2009	European	0.031	0.733	0.067	1.000
		Ulcerative colitis	rs4510766	A	2.00E-16	Anderson et al Nat Genet 2011	European	–	–	0.107	1.000
		Ulcerative colitis	rs4730276	–	9.00E-06	Silverberg et al Nat Genet 2009	European	–	–	0.038	0.534
		Ulcerative colitis	rs4730273	–	5.00E-06	Silverberg et al Nat Genet 2009	European	0.032	1.000	0.027	0.931
		Ulcerative colitis	rs2108225	A	1.00E-07	Asano et al Nat Genet 2009	Japanese	0.024	0.548	0.017	0.482
FUT8	rs11847263	N-Glycans (DG6)	rs10483776	G	1.00E-08	Lauc et al PLoS Genet 2010	European	0.369	0.864	0.416	0.758
		N-Glycans (DG1)	rs7159888	A	3.00E-18	Lauc et al PLoS Genet 2010	European	0.714	1	0.727	1
		Conduct disorder (symptom count)	rs1256531	–	4.00E-06	Dick et al Mol Psychiatry 2010	European, African, other	0.092	0.796	0.071	1
		Waist Circumference	rs7158173	–	4.00E-06	Polasek et al Croat Med J 2009	European	0.011	0.182	0.002	0.081
		Multiple Sclerosis - brain glutamate levels	rs8007846	–	9.00E-06	Baranzini et al Brain 2010	American	0.196	0.571	0.261	0.692

Table 2. Cont.

Gene	IgG Glycan Top SNP	Disease	Top Disease SNP	Risk Allele	P-Value	Reference	Ancestry	HapMap 2		1000G Pilot 1	
								R^2	D′	R^2	D′
SYNGR1-TAB1-MGAT3-CACNA1I	rs909674	Sudden cardiac arrest	rs54211	–	8.00E-07	Aouizerat et al BMC Car Diso 2011	European	0.053	0.362	0.043	0.360
		Primary biliary cirrhosis	rs968451	T	1.00E-09	Mells et al Nat Genet 2011	European	0.041	0.682	0.080	1
		Crohn's disease	rs2413583	C	1.00E-26	Franke et al Nat Genet 2010	European	0.043	0.313	0.053	0.292
SMARCB1-DERL3	rs2186369	GGT	rs2739330	T	2.00E-09	Chambers et al Nat Genet 2011	European	0.009	0.255	0	0.012
PRRT1	rs9296009	Nodular sclerosis Hodgkin lymphoma	rs204999	–	8.00E-18	Cozen et al Blood 2012	European	0.125	1	–	–
		Phospholipid levels	rs1061808	–	8.00E-10	Demirkan et al PLoS Genet 2012	European	0.137	0.626	–	–
HLA-DQA2 - HLA-DQB2	rs1049110	SLE	rs2301271	T	2.00E-12	Chung et al PLoS Genet 2011	European	0.967	1	–	–
		Hepatitis B	rs7453920	G	6.00E-28	Mbarek et al Hum Mol Genet 2011	Japanese	0.967	1	–	–
		Narcolepsy	rs2858884	A	3.00E-08	Hor et al Nat Genet 2010	European	0.193	1	–	–
BACH2	rs404256	Graves' disease	rs370409	T	2.00E-06	Chu et al Nat Genet 2011	Chinese	0.010	0.187	0.009	0.166
		Celiac disease	rs10806425	A	4.00E-10	Dubois et al Nat Genet 2010	European	0.005	0.103	0.006	0.09
		T1D	rs3757247	A	1.00E-06	Grant et al Diabetes 2009	European	0.039	0.196	0.042	0.204
		T1D	rs11755527	G	3.00E-08	Plagnol et al PLoSGenet 2011	European	0.031	0.186	0.031	0.179
		T1D	rs11755527	–	5.00E-08	Barrett et al Nat Genet 2009	European	0.031	0.186	0.031	0.179
		T1D	rs11755527	G	5.00E-12	Cooper et al Nat Genet 2008	European	0.031	0.186	0.031	0.179
		Crohn's disease	rs1847472	G	5.00E-09	Franke et al Nat Genet 2010	European	0	0.011	0.009	0.124
		Multiple Sclerosis	rs12212193	G	4.00E-08	Sawcer et al Nature 2011	European	0.001	0.036	0.027	0.166
SLC38A10	rs7224668	Longevity	rs10445407	–	1.00E-06	Yashin et al Aging 2010	European	0.714	1	0.692	1

Associations are those found in the GWAS Catalog track of USCS Genome browser (accessed 04/07/2012) and LD has been calculated using SNAP (http://www.broadinstitute.org/mpg/snap/;Johnson, A. D., Handsaker, R. E., Pulit, S., Nizzari, M. M., O'Donnell, C. J., de Bakker, P. I. W. SNAP: A web-based tool for identification and annotation of proxy SNPs using HapMap *Bioinformatics, 2008 24(24):2938–2939*).

doi:10.1371/journal.pgen.1003225.t002

IGP31 ($p = 3.33 \times 10^{-08}$). Although the function of this gene is not understood, it has been associated with autism and longevity [65,66].

The remaining three signals implicated *ABCF2-SMARCD3* region (rs1122979 was associated with IGP 2, 5, 42, 45, with p-value ranging between $2.10 \times 10^{-10} < p < 1.89 \times 10^{-9}$), *RECK* (rs4878639 was suggestively associated with IGP17; $p = 3.51 \times 10^{-8}$) and *PEX5* (rs12828421 suggestively associated with IGP41; $p = 4.48 \times 10^{-8}$). The function of *ABCF2* (ATP-binding cassette, sub-family F, member 2) is not well understood. *SMARCD3* stimulates nuclear receptor mediated transcription; it belongs to the neural progenitors-specific chromatin remodelling complex (npBAF complex) and the neuron-specific chromatin-remodelling complex (nBAF complex). *RECK* is known to be a strong suppressor of tumour invasion and metastasis, regulating metalloproteinases which are involved in cancer progression. *PEX 5* binds to the C-terminal PTS1-type tripeptide peroxisomal targeting signal and plays an essential role in peroxisomal protein import (www.genecards.org).

Results from an independent cohort using MS quantitation method

The parallel effort in the outbred Leiden Longevity Study (LLS) was based on a different *N*-glycan quantitation method (MS). While UPLC groups glycans according to structural similarities, MS groups them by mass. Furthermore, MS analysis focused on Fc glycans while UPLC measures both Fc and Fab glycans, thus traits measured by the two methods could not have been directly compared. Glycosylation patterns of IgG1 and IgG2 were investigated by analysis of tryptic glycopeptides, with six glycoforms per IgG subclass measured. The intensities of all glycoforms were related to the monogalactosylated, core-fucosylated biantennary species, providing five relative intensities registered per IgG subclass (Tables S5 and S6). The analysis identified two loci as genome-wide significant - implicating *MGAT3* ($p = 1.6 \times 10^{-10}$ for G1FN, analogous to UPLC IGP9; $p = 3.12 \times 10^{-8}$ for G0FN, analogous to UPLC IGP5), and *B4GALT1* ($p = 5.4 \times 10^{-8}$ for G2F, analogous to UPLC IGP13) confirming GWAS signals in the discovery meta-analysis.

Replication of our findings

We then sought a separate independent replication of the other 14 genome-wide significant and strongly suggestive signals identified in the discovery analysis, which was performed in the LLS cohort, appreciating that the quantitated *N*-glycan traits do not exactly match between the two cohorts. SNPs were chosen for replication based on initial meta-analysis results of genotype data prior to imputed analysis. All five traits measured in LLS cohort were tested for association with all the selected SNPs (Table S6). We were able to reproduce association to *ST6GAL1* ($p = 8.1 \times 10^{-7}$ for G2F, substrate for sialyltransferase) and *SMARCB1-DERL3* ($p = 1.6 \times 10^{-7}$ for G1N, analogous to UPLC IGP9). Weaker, though nominally significant associations were confirmed at *IKZF1* ($p = 2.3 \times 10^{-3}$ for G1N), *SLC38A10* ($p = 4.8 \times 10^{-3}$ for G2N), *IL6ST-ANKRD55* ($p = 1.3 \times 10^{-2}$ for G0N) and *ABCF2-SMARCD3* ($p = 2.7 \times 10^{-2}$ for G2N). The fact that we did not replicate associations at the other 8 loci was not unexpected, because those 8 loci showed association with UPLC-measured *N*-glycan traits that do not compare to any of the traits measured by MS (see Table S5 for comparison of MS and UPLC traits).

Functional experiment: Ikzf1 haplodeficiency results in altered N-glycosylation of IgG

IKZF1 is considered to be the important regulator governing differentiation of T cells into CD4+ and CD8+ T cells [67]. Since glycan traits associated with *IKZF1* were related to the presence and absence of core-fucose and bisecting GlcNAc, we analysed the promoter region of *MGAT3* (codes for enzyme that adds bisecting GlcNAc to IgG glycans) *in silico* and identified two binding sites for *IKZF1* that were conserved between humans and mice, while recognition sites for *IKZF1* were not found in the promoter region of *FUT8* (which codes for an enzyme that adds core-fucose to IgG glycans). Since the promoter regions of *MGAT3* were conserved between humans and mice, we used *Ikzf1* knockout mice [68] as a model to study the effects of *IKZF1* deficiency on IgG glycosylation. IgG was isolated from the plasma of 5 heterozygous knockout mice and 5 wild-type controls. The summary of the results of IgG glycosylation analysis is presented in Table 3, while complete results are presented in Table S7. We observed a number of alterations in glycome composition that were all consistent with the role of *IKZF1* in the down-regulation of fucosylation and up-regulation of the addition of bisecting GlcNAc to IgG glycans; 12 out of 77 IgG N-glycans measures showed statistically significant difference ($p < 0.05$) between wild type and heterozygous *Ikzf1* knock-outs, where 5 mice from each group were compared (Table 3). The empirical version of Hotelling's test demonstrated global significance ($p = 0.03$) of difference between distributions of IgG glycome between wild type and *Ikzf1* knock-out mice, where 5 mice from each group were compared. While the tests for differences between individual glycome measurements did not reach strict statistical significance after conservative Bonferroni correction ($p = 0.05/77 = 0.0006$), we observed that 12 out of 77 (15%) IgG N-glycans measures showed nominally significant difference ($p < 0.05$) between wild type and heterozygous *Ikzf1* knock-outs (Table 3). Significant results from the global difference test ensure that difference between the two groups does exist, and it is most likely due to the difference between (at least some of) the measurements which demonstrated nominal significance. Observed alterations in glycome composition were all consistent with the role of *IKZF1* in the down-regulation of fucosylation and up-regulation of the addition of bisecting GlcNAc to IgG glycans.

Investigating the biomarker potential of IgG N-glycans in Systemic Lupus Erythematosus (SLE)

Given that *IKZF1* has been convincingly associated with SLE in previous studies [35–37], and that functional studies in heterozygous knock-out mice in our study showed clear differences in profiles of several IgG *N*-glycan traits, we explored an intriguing hypothesis: whether the same IgG *N*-glycan traits that were significantly affected in *Ikzf1* knock-out mice could be demonstrated to differ between human SLE cases and controls. If this were true, then pleiotropy between the effects of *IKZF1* on SLE and on IgG *N*-glycans in human plasma, revealed by independent GWA studies, would lead to a discovery of a novel class of biomarkers of SLE – IgG *N*-glycans – which could possibly extend their usefulness in prediction of other autoimmune disorders, cancer and neuropsychiatric disorders, through the same mechanism.

To test this hypothesis, we measured IgG *N*-glycans in 101 SLE cases and 183 matched controls (typically two controls per case), recruited in Trinidad (see materials and methods for further details). Table 4 shows the results of the measurements: for 10 of 12 *N*-glycan traits chosen on a basis of the experiments in mice (Table 3). The entire dataset for all glycans can be found in Table S8. There was a statistically significant difference ($p < 0.05$) between SLE cases and controls, which was generally not the case with other groups of *N*-glycans (data not shown). Moreover,

Table 3. Twelve groups of IgG N-glycans (of 77 measured) that showed nominally significant difference ($p<0.05$) in observed values between 5 mice that were heterozygous *Ikzf1* knock-outs (Neo) and 5 wild-type controls (wt).

Increased N-glycans					
N-glycan group code	**N-glycan trait**	**Mean (Neo)**	**Mean (wt)**	**Mean(Neo)/Mean(wt)**	**p-value***
IGP8	GP9 - FA2[3]G1	8.91	7.44	1.20	3.54E-03
IGP48	GP9n – GP9/GPn*100	11.71	10.34	1.13	1.41E-02
IGP64	% FG1n/G1n	98.47	97.53	1.01	2.63E-02
Decreased N-glycans					
N-glycan group code	**N-glycan trait**	**Mean (Neo)**	**Mean (wt)**	**Mean(Neo)/Mean(wt)**	**p-value***
IGP9	GP10 - FA2[6]BG1	0.13	0.17	0.76	2.32E-03
IGP10	GP11 - FA2[3]BG1	0.34	0.62	0.55	4.74E-02
IGP19	GP20 – (undetermined)	12.29	15.07	0.82	7.62E-03
IGP23	GP24 - FA2BG2S2	1.27	1.50	0.85	3.20E-02
IGP37	FBS1/FS1	0.12	0.18	0.67	4.93E-02
IGP38	FBS1/(FS1+FBS1)	0.10	0.15	0.67	4.84E-02
IGP49	GP10n – GP10/GPn*100	0.17	0.24	0.71	1.71E-03
IGP50	GP11n – GP11/GPn*100	0.44	0.87	0.51	4.27E-02
IGP68	% FBG1n/G1n	1.15	2.05	0.56	2.76E-02

The global difference test was significant ($p = 0.03$). *t-test for equality of means (2-tailed).
doi:10.1371/journal.pgen.1003225.t003

the significance of the difference was striking in some cases, e.g. $p<10^{-14}$ for IGP48, $p<10^{-13}$ for IGP8, and $p<10^{-6}$ for IGP64. Furthermore, the differences in the direction of effect in mice were strikingly preserved in humans (Table 4). The most significant differences observed across all 77 IgG *N*-glycans measurements between SLE cases and controls (Table 4) were overlapping well with the 12 *N*-glycan groups that were significantly changed in functional experiments in *Ikzf1* knock-out mice.

To strengthen our findings and control for possible bias, we repeated the analysis excluding all the cases on corticosteroid treatment at the time of interview (77/101) and subsequently all the cases that were not on corticosteroid treatment at the time of interview (24/101). Although the power of the analysis decreased due to reduced number of cases, the results did not change and they remained highly statistically significant. We also hypothesized that the observed glycan changes may not be specific to SLE, but

Table 4. Groups of IgG N-glycans from Table 3 that showed statistically significant difference in observed values (corrected by sex, age, and African admixture) between 101 Afro-Caribbean cases with SLE and 183 controls.

Decreased N-glycans					
N-glycan group code	**N-glycan trait**	**Mean (SLE)**	**Mean (controls)**	**Mean(SLE)/Mean(controls)**	**p-value***
IGP8	GP9 - FA2[3]G1	6.67	8.03	0.83	1.86E-14
IGP48	GP9n – GP9/GPn*100	9.09	11.06	0.82	6.72E-15
IGP64	% FG1n/G1n	80.93	83.22	0.97	5.07E-07
IGP19	GP20 – (undetermined)	0.73	0.80	0.91	4.87E-02
Increased N-glycans					
N-glycan group code	**N-glycan trait**	**Mean (SLE)**	**Mean (controls)**	**Mean(SLE)/Mean(controls)**	**p-value***
IGP9	GP10 - FA2[6]BG1	4.58	4.13	1.10	4.37E-04
IGP23	GP24 - FA2BG2S2	3.00	2.65	1.13	1.67E-03
IGP37	FBS1/FS1	0.20	0.18	1.11	5.08E-03
IGP38	FBS1/(FS1+FBS1)	0.17	0.15	1.13	4.83E-03
IGP49	GP10n – GP10/GPn*100	6.24	5.70	1.09	1.39E-03
IGP68	% FBG1n/G1n	17.34	15.30	1.13	2.19E-06

*t-test for equality of means (2-tailed).
doi:10.1371/journal.pgen.1003225.t004

may be caused by corticosteroid treatment, or secondary to any inflammatory process. For this reason, and in SLE cases only, we investigated whether corticosteroid treatments and/or CRP measurements, were associated with IgG *N*-glycan traits. Analysis for CRP was repeated with CRP treated as a binary variable (with cut-off value at 10 mg/L). In all these analyses, the initial results held and were not changed: the association of IgG *N*-glycans and SLE remained striking, while the association with corticosteroid treatment and CRP was not (Table S9). Finally, we also repeated the analysis adjusting for percent African admixture, as it has been reported that SLE in Afro-Caribbean population is associated with African admixture [69]. However, this adjustment only had a minor and non-systemic effect on the previous results, and the reported observations remained.

We then validated biomarker potential of IGP48, the IgG *N*-glycan trait most significantly associated with SLE status, in prediction of SLE in 101cases and 183 matched controls. We used the PredictABEL package for R (see materials and methods) [70]. As shown in Figure 3, age, sex and African admixture did not have any predictive power for this disease, but addition of IGP48 substantially increased sensitivity and specificity of prediction, with area under receiver-operator curve (AUC) increasing from 0.515 (95% confidence interval (CI): 0.441–0.590) to 0.842 (0.791–0.893). It is likely that further additions of other IgG *N*-glycans could provide even more accurate predictions. To cross-validate this result, we split our dataset with SLE cases and controls into a "training set" (2/3; 67 cases and 122 controls) and "test set" (1/3; 34 cases and 61 controls). Area under ROC-curve (AUC) was calculated for the test dataset. The whole process was repeated 1000 times, to allow computation of the mean AUC (and 95% CI) in the test datasets. Mean AUC was virtually unchanged compared to AUC obtained when using the complete dataset and no training, which suggests that the predictive power of IGP48 on SLE is very robust.

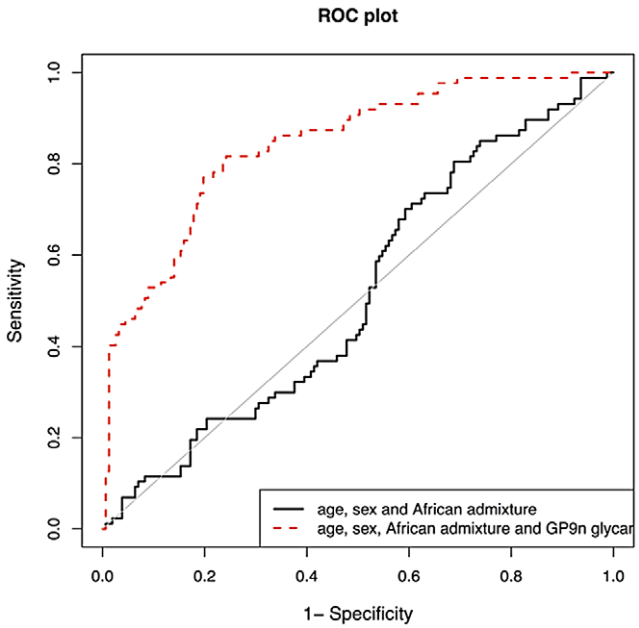


Figure 3. Validation of biomarker potential of IGP48 IgG N-glycan percentage in prediction of Systemic Lupus Erythematosus (SLE) in 101 Afro-Caribbean cases and 183 matched controls. As shown in the graph, age and sex do not have any predictive power for this disease, but addition of IGP48 substantially increases sensitivity and specificity of prediction, with area under receiver-operator curve increased to 0.828.
doi:10.1371/journal.pgen.1003225.g003

Discussion

This study clearly demonstrates that the recent developments in high-throughput glycomics and genomics now allow identification of genetic loci that control *N*-glycosylation of a single plasma protein using a GWAS approach. This progress should allow many similar follow-up studies of genetic regulation of *N*-glycosylation of other important plasma proteins, thus bringing unprecedented insights into the role of protein glycosylation in systems biology. As a prelude to this discovery, we recently reported the results of the first GWA study of the overall human plasma *N*-glycome using the HPLC method. Although the study was of a comparable sample size (N~2000), it only identified genome-wide associations with two glycosyltransferases and one transcription factor (*HNF1a*) [71]. We believe that the power of our initial study was reduced because *N*-glycans in human plasma originate from different glycoproteins where they have different functions and undergo protein-specific, or tissue-specific glycosylation. In this study the largest percentage of variance explained by a single association was 16–18% where as in the *N*-glycan study this was 1–6%. Furthermore, concentrations of individual glycoproteins in plasma vary in many physiological processes, introducing substantial "noise" to the quantitation of the whole-plasma *N*-glycome.

In this study we avoided both problems by isolating a single protein from plasma (IgG), which is produced by a single cell type (B lymphocytes), thus effectively excluding differential regulation of gene expression in different tissues, and the "noise" introduced by variation in plasma IgG concentration and by *N*-glycans on other plasma proteins. The only remaining "noise" in our system was the incomplete separation of some glycan structures (which co-eluted from the UPLC column) and the presence of Fab glycans on a subset of IgG molecules, but for the majority of glycan structures this "noise" was well below 10% [19]. We expected that the specificity of our phenotype and precision of the measurement provided by novel UPLC and MS methods should substantially increase the power of the study to detect genome-wide associations. Prior to analysis we could not predict which quantitation method would work better in GWA study design (UPLC vs. MS), so we used them both, each in one separate cohort of comparable sample size (N~2000).

The UPLC method yielded many more, and much stronger, genome-wide association signals in comparison to our previous study of the total plasma *N*-glycome in virtually same sample set of examinees [27,71]. Sixteen loci were identified in association with glycan traits with p-values$<5\times10^{-8}$ and nine reached the strict genome wide threshold of 2.27×10^{-9}. The parallel study in the LLS cohort using MS quantitation has independently identified two of those 16 loci, showing genome-wide association with *N*-glycan traits. MS quantitation also allowed us to replicate 6 further loci identified in the discovery analysis, using comparable *N*-glycan traits measured by the two methods. However, in this follow-up analysis we were unable to replicate associations for the remaining 8 loci. This was not unexpected, because those glycosylation traits correspond to different fucosylated glycans; since fucosylation was not quantified by MS, the association between glycans measured by MS and those regions should not be expected.

Among the nine loci that reached genome-wide statistical significance, four involved genes encoding glycosyltransferases known to glycosylate IgG (*ST6GALI, B4GALT1, FUT8, MGAT3*,). The enzyme beta1,4-galactosyltransferase 1 is responsible for the

addition of galactose to IgG glycans. Interestingly, variants in *B4GALT1* gene did not affect the main measures of IgG galactosylation, but rather differences in sialylation and the percentage of bisecting GlcNAc. These associations are still biologically plausible, because galactosylation is a prerequisite for sialylation, and enzymes which add galactose and bisecting GlcNAc compete for the same substrate [72]. A potential candidate for *B4GALT1* regulator is *IL6ST*, which codes for interleukin 6 signal transducer, because it showed stronger associations with the main measures of IgG galactosylation than *B4GALT1* itself. Molecular mechanisms behind this association remain elusive, but early work on *IL6* (then called *PHGF*) suggested that it may be relevant for glycosylation pathways in B lymphocytes [73].

Core-fucosylation of IgG has been intensively studied due to its role in antibody-dependent cell-mediated cytotoxicity (ADCC). This mechanism of killing is considered to be one of the major mechanisms of antibody-based therapeutics against tumours. Core-fucose is critically important in this process, because IgGs without core fucose on the Fc glycan have been found to have ADCC activity enhanced by up to 100-fold [74]. Alpha-(1,6)-fucosyltransferase (fucosyltransferase 8) catalyses the transfer of fucose from GDP-fucose to *N*-linked type complex glycopeptides, and is encoded by the *FUT8* gene. We found that SNPs located near this gene influenced overall levels of fucosylation. The directly measured IgG glycome traits most strongly associated with SNPs in the *FUT8* region consisted of A2, and, less strongly, A2G1 and A2G2. These associations are biologically plausible as these glycans serve as substrates for fucosyltransferase 8. Interestingly, SNPs located near the *IKZF1* gene influenced fucosylation of a specific subset of glycans, especially those without bisecting GlcNAc, and were also related to the ratio of fucosylated structures with and without bisecting GlcNAc. This suggests the *IKZF1* gene encoding Ikaros as a potential indirect regulator of fucosylation in B-lymphocytes by promoting the addition of bisecting GlcNAc, which then inhibits fucosylation. The analysis of IgG glycosylation in *Ikzf1* haplodeficient mice confirmed the postulated role of Ikaros in the regulation of IgG glycosylation (Table 3). The effect of *Ikzf1* haplodeficiency on IgG glycans manifested mainly in the decrease in bisecting GlcNAc on different glycan structures. The increase in fucose was observed only in a subset of structures, but since very high level of fucosylation was present in the wild type mouse (up to 99.8%), a further increase could not have been demonstrated.

Nearly all genome-wide significant loci in our study have already been clearly demonstrated to be associated with autoimmune diseases, haematologic cancers, and some of them also with chronic inflammation and/or neuropsychiatric disorders. Although the literature on those associations is extensive, we tried to highlight only those associations that were identified using genome-wide association studies in datasets independent from our study. We gave prominence to associations arising from GWA studies because they are typically replicable; GWA studies have sufficient power to detect true associations, and require stringent statistical testing and replication to avoid false positive results. They have been reviewed and summarized in Table 2. The table implies abundant pleiotropy between loci that control *N*-glycosylation (in this case, of IgG protein) and loci that have been implicated in many human diseases. Autoimmune diseases (including SLE, RA, UC and over 80 others) are generally thought to be triggered by aggressive responses of the adaptive immune system to self antigens, resulting in tissue damage and pathological sequelae [38]. Among other mechanisms, IgG autoantibodies are responsible for the chronic inflammation and destruction of healthy tissues by cross-linking Fc receptors on innate immune effector cells [75]. Class and glycosylation of IgG are important for pathogenicity of autoantibodies in autoimmune diseases (reviewed in [76]). Removal of IgG glycans leads to the loss of the proinflammatory activity, suggesting that *in vivo* modulation of antibody glycosylation might be a strategy to interfere with autoimmune processes [75]. Indeed, the removal of IgG glycans by injections of EndoS *in vivo* interfered with autoantibody-mediated proinflammatory processes in a variety of autoimmune models [75].

Results from our study suggest that IgG *N*-glycome composition is regulated through a complex interplay between loci affecting an overlapping spectrum of glycome measurements, and through interaction of genes directly involved in glycosylation and those that presumably have a "higher-level" regulatory function. SNPs at several different loci in this GWA study showed genome-wide significant associations with the same or similar IgG glycosylation traits. For example, SNPs at loci on chromosomes 9 (*B4GALT1* region) and 3 (*ST6GAL1* region) both influenced the percentage of sialylation of galactosylated fucosylated structures (without bisecting GlcNAc) in the same direction. SNPs at these loci also influenced the ratio of fucosylated monosialylated structures (with and without bisecting GlcNAc) in the opposite direction. SNPs at the locus on chromosome 9 (*B4GALT1*), and two loci on chromosome 22 (*MGAT3* and *SMARCB1-DERL3* region) simultaneously influenced the ratio of fucosylated disialylated structures with and without bisecting GlcNAc. SNPs at loci on chromosome 7 (*IKZF1* region) and 14 (*FUT8* region) influenced an overlapping range of traits: percentage of A2 and A2G1 glycans, and, in the opposite direction, the percentage of fucosylation of agalactosylated structures.

Finally, this study demonstrated that findings from "hypothesis-free" GWA studies, when targeted at a well defined biological phenotype of unknown relevance to human health and disease (such as *N*-glycans of a single plasma protein), can implicate genomic loci that were not thought to influence protein glycosylation. Moreover, unexpected pleiotropy of the implicated loci that linked them to diseases has changed this study from "hypothesis-free" to "hypothesis-driven" [77], and led us to explore biomarker potential of a very specific IgG *N*-glycan trait in prediction of a specific disease (SLE) with considerable success. To our knowledge, this is one of the first convincing demonstrations that GWA studies can lead to biomarker discovery for human disease. This study offers many additional opportunities to validate the role of further *N*-glycan biomarkers for other diseases implicated through pleiotropy.

Conclusions

A new understanding of the genetic regulation of IgG *N*-glycan synthesis is emerging from this study. Enzymes directly responsible for the addition of galactose, fucose and bisecting GlcNAc may not have primary responsibility for the final IgG *N*-glycan structures. For all three processes, genes that are not directly involved in glycosylation showed the most significant associations: *IL6ST-ANKRD55* for galactosylation; *IKZF1* for fucosylation; and *SMARCB1-DERL3* for the addition of bisecting GlcNAc. The suggested higher-level regulation is also apparent from the differences in IgG Fab and Fc glycosylation, observed in human IgG [78,79] and different myeloma cell lines [80], and further supported by recent observation that various external factors exhibit specific effects on glycosylation of IgG produced in cultured B lymphocytes [81].

Moreover, this study showed that it is possible to identify loci that control glycosylation of a single plasma protein using a

GWAS approach, and to develop a novel class of disease biomarkers. This should lead to large advances in understanding of the role of protein glycosylation in the future. This study identified 16 genetic loci that are likely to be part of a much larger genetic network that regulates the complex process of IgG *N*-glycosylation and several further loci that show suggestive association with glycan traits and merit further study. Genetic variants in several of these genes were previously associated with a number of inflammatory, neoplastic and neuropsychiatric diseases across ethnically diverse populations, all of which could benefit from earlier and more accurate diagnosis based on molecular biomarkers. Variations in individual SNPs have relatively small effects, but when several polymorphisms are combined in a complex pathway like *N*-glycosylation, the final product of the pathway - in this case IgG N-glycan - can be significantly different, with consequences for IgG function and possibly also disease susceptibility. Our results may also provide an explanation for the reported pleiotropy and antagonistic genetic effects of loci involved in autoimmune diseases and hematologic cancers [39,77].

Materials and Methods

Ethics statement

All research in this study that involved human participants has been approved by the appropriate ethics committees: the Ethics Committee of the University of Split Medical School for all Croatian examinees from Vis and Korcula islands; the Local Research Ethics Committees in Orkney and Aberdeen for the Orkney Complex Disease Study (ORCADES); the University of Uppsala (Dnr 2005:325) for all examinees from Northern Sweden; the Leiden University Medical Centre Ethical Committee for all participants in the Leiden Longevity Study (LLS); and the Ethics Committee of the London School of Hygiene and Tropical Medicine for all SLE cases and controls from Trinidad. All ethics approvals were given in compliance with the Declaration of Helsinki (World Medical Association, 2000). All human subjects included in this study have signed appropriate informed consent.

Study participants—discovery and replication cohorts

All population studies recruited adult individuals within a community irrespective of any specific phenotype. Fasting blood samples were collected, biochemical and physiological measurements taken and questionnaire data for medical history as well as lifestyle and environmental exposures were collected following similar protocols. Basic cohort descriptives are included in Table S11.

The CROATIA-Vis study includes 1008 Croatians, aged 18–93 years, who were recruited from the villages of Vis and Komiža on the Dalmatian island of Vis during 2003 and 2004 within a larger genetic epidemiology program [82]. The CROATIA-Korcula study includes 969 Croatians between the ages of 18 and 98 [83]. The field work was performed in 2007 and 2008 in the eastern part of the island, targeting healthy volunteers from the town of Korčula and the villages of Lumbarda, Žrnovo and Račišće.

The Orkney Complex Disease Study (ORCADES) was performed in the Scottish archipelago of Orkney and collected data between 2005 and 2011 [84]. Data for 889 participants aged 18 to 100 years from a subgroup of ten islands, were used for this analysis.

The Northern Swedish Population Health Study (NSPHS) is a family-based population study including a comprehensive health investigation and collection of data on family structure, lifestyle, diet, medical history and samples for laboratory analyses from peoples living in the north of Sweden [84]. Complete data were available from 179 participants aged 14 to 91 years.

DNA samples were genotyped according to the manufacturer's instructions on Illumina Infinium SNP bead microarrays (HumanHap300v1 for CROATIA-Vis, HumanHap300v2 for ORCADES and NSPHS and HumanCNV370v1 for CROATIA-Korcula). Genotypes were determined using Illumina BeadStudio software. Genotyping was successfully completed on 991 individuals from CROATIA-Vis, 953 from CROATIA-Korcula, 889 from ORCADES and 700 from NSPHS, providing a platform for genome-wide association study of multiple quantitative traits in these founder populations.

The Leiden Longevity Study (LLS) has been described in detail previously [85]. It is a family based study and consists of 1671 offspring of 421 nonagenarian sibling pairs of Dutch descent, and their 744 partners. 1848 individuals with available genotypic and IgG measurements data were included in the current analysis. Within the Leiden Longevity Study 1345 individuals were genotyped using Illumina660 W (Rotterdam, Netherlands) and 503 individuals were genotyped using Illumina OmniExpress (Estonian Biocentre, Genotyping Core Facility, Estonia).

Isolation of IgG and glycan analysis

In the discovery population cohorts (CROATIA-Vis, CROATIA-Korcula, ORCADES, and NSPHS), the IgG was isolated using protein G plates and its glycans analysed by UPLC in 2247 individuals, as reported previously [19]. Briefly, IgG glycans were labelled with 2-AB fluorescent dye and separated by hydrophilic interaction ultra-performance liquid chromatography (UPLC). Glycans were separated into 24 chromatographic peaks and quantified as relative contributions of individual peaks to the total IgG glycome. The majority of peaks contained individual glycan structures, while some contained more structures. Relative intensities of each glycan structure in each UPLC peak were determined by mass spectrometry as reported previously [19]. On the basis of these 24 directly measured "glycan traits", additional 54 "derived traits" were calculated. These include the percentage of galactosylation, fucosylation, sialylation, etc. described in the Table S1. When UPLC peaks containing multiple traits were used to calculate derived traits, only glycans with major contribution to fluorescence intensity were used.

In the replication population cohort (Leiden Longevity Study), the IgG was isolated from plasma samples of 1848 participants. Glycosylation patterns of IgG1 and IgG2 were investigated by analysis of tryptic glycopeptides using MALDI-TOF MS. Six glycoforms per IgG subclass were determined by MALDI-TOFMS. Since the intensities of all glycoforms were related to the monogalactosylated, core-fucosylated biantennary species (glycoform B), five relative intensities were registered per IgG subclass [86].

Genotype and phenotype quality control

Genotyping quality control was performed using the same procedures for all four discovery populations (CROATIA-Vis, CROATIA-Korcula, ORCADES, and NSPHS). Individuals with a call rate less than 97% were removed as well as SNPs with a call rate less than 98% (95% for CROATIA-Vis), minor allele frequency less than 0.02 or Hardy-Weinberg equilibrium p-value less than 1×10^{-10}. 924 individuals passed all quality control thresholds from CROATIA-Vis, 898 from CROATIA-Korcula, 889 from ORCADES and 656 from NSPHS.

Extreme outliers (those with values more than 3 times the interquartile distances away from either the 75th or the 25th percentile values) were removed for each glycan measure to

account for errors in quantification and to remove individuals not representative of normal variation within the population. After phenotype quality control the number of individuals with complete phenotype and covariate information for the meta-analysis was 2247, consisting of 906 men and 1341 women (802 from CROATIA-Vis, 851 from CROATIA-Korcula, 415 from ORCADES, 179 from NSPHS).

In Leiden Longevity Study, GenomeStudio was used for genotyping calling algorithm. Sample call rate was >95%, and SNP exclusions criteria were Hardy-Weinberg equilibrium p value$<10^{-4}$, SNP call rate<95%, and minor allele frequency <1%. The number of the overlapping SNPs that passed quality controls in both samples was 296,619.

To combine the data from the different array sets and to increase the overall coverage of the genome to up to 2.5 million SNPs, we imputed autosomal SNPs reported in the Haplotype Mapping Project (release #22, http://hapmap.ncbi.nlm.nih.gov) CEU sample. Based on the SNPs that were genotyped in all arrays and passed quality control, the imputation programmes MACH (http://www.sph.umich.edu/csg/abecasis/MACH/) or IMPUTE2 (http://mathgen.stats.ox.ac.uk/impute/impute_v2.html) were used to obtain ca. 2.5 million SNPs for further analysis.

For replication of genome-wide significant hits identified in the discovery meta-analysis, all SNPs listed in Table S6 were used and looked up in LLS. The only exception was rs11621121, which had low imputation accuracy and did not pass quality control criteria. For this SNP, a set of 11 proxy SNPs from HapMap r. 22 (all with $R^2>0.85$) was studied. All studied SNPs had imputation quality of 0.3 or greater.

Genome-wide association analysis

In the discovery populations, genome-wide association analysis was firstly performed for each population and then combined using an inverse-variance weighted meta-analysis for all traits. Each trait was adjusted for sex, age and the first 3 principal components obtained from the population-specific identity-by-state (IBS) derived distances matrix. The residuals were transformed to ensure their normal distribution using quantile normalisation. Sex-specific analyses were adjusted for age and principal components only. The residuals expressed as z-scores were used for association analysis. The "mmscore" function of ProbABEL [87] was used for the association test under an additive model. This score test for family based association takes into account relationship structure and allowed unbiased estimations of SNP allelic effect when relatedness is present between examinees. The relationship matrix used in this analysis was generated by the "ibs" function of GenABEL (using weight = "freq" option), which uses genomic data to estimate the realized pair-wise kinship coefficient. All lambda values for the population-specific analyses were below 1.05 (Table S4), showing that this method efficiently accounts for family structure.

Inverse-variance weighted meta-analysis was performed using the MetABEL package (http://www.genabel.org) for R. SNPs with poor imputation quality ($R^2<0.3$) were excluded prior to meta-analysis. Principal component analysis was performed using R to determine the number of independent traits used for these analyses (Table S10). 21 principal components explained 99% of the variance so an association was considered statistically significant at the genome-wide level if the p-value for an individual SNP was less than 2.27×10^{-9} ($5\times10^{-8}/22$ traits) [88]. SNPs were considered strongly suggestive with p-values between 5×10^{-8} and 2.27×10^{-9}. Regions of association were visualized using the web-based software LocusZoom [89] to display the linkage disequilibrium (LD) of the region based on hg18/1000 Genomes June 1010 CEU data. The effect of the most significant SNP in each gene region expressed as percentage of the variance explained was calculated for each glycan trait adjusted for sex, age and first 3 principal components in each cohort individually using the "polygenic" function of the GenABEL package for R. Conditional analysis was undertaken for all significant and suggestive regions. GWAS was performed as described above with the additional adjustment for the dosage of the top SNP in the region for only the chromosome containing the association. Subsequent meta-analysis was performed as described previously and the results visualised using LocusZoom to ensure that the association peak have been removed.

In LLS, all IgG measurements were log-transformed. The score statistic for testing for an additive effect of a diallelic locus on quantitative phenotype was used. To account for relatedness in offspring data we used the kinship coefficients matrix when computing the variance of the score statistic. Imputation was dealt with by accounting for loss of information due to genotype uncertainty [90]. For the association analysis of the GWAS data, we applied the score test for the quantitative trait correcting for sex and age using an executable C++ program QTassoc (http://www.lumc.nl/uh, under GWAS Software). For further details we refer to supplementary online information.

Experiments in Ikzf1 knockout mice

The *Ikzf1*$^{+/-}$ mice harbouring the Neo-PAX5-IRES-GFP knock in allele were obtained from Meinrad Busslinger (IMP, Vienna) and backcrossed to C57BL/6 mice. Both wild-type and *Ikzf1*Neo$^{+/-}$ animals at the age of about 8 months were subjected to retro-orbital puncture to collect blood in the presence of EDTA. Samples were centrifuged for 10 minutes at room temperature and plasma was harvested. IgG was isolated and subjected to glycan analyses.

Statistical significance of the difference in distributions of IgG glycome between wild type and the *Ikzf1*$^{+/-}$ mice was assessed using empirical version of the Hotelling's test. In brief, the empirical distribution of the Hotelling's T^2 statistics was worked out by permuting the group status of the animals at random without replacement 10,000 times. This empirical distribution was then contrasted with the original value of T^2, with the proportion of empirically observed T^2 values greater than or equal to the original T^2 regarded as the empirical p-value.

Dataset with SLE cases and matched controls

A total of 101 SLE cases and 183 controls from Trinidad were studied. The inclusion criteria for cases and controls in Trinidad were designed to restrict the sample to individuals without Indian or Chinese ancestry. Cases and controls were eligible to be included if they were resident in northern Trinidad (excluding the southern part of the island where Indians are in the majority) and they had Christian (rather than Hindu, Muslim or Chinese) first names. Identification of cases was carried out by contacting all physicians specializing in rheumatology, nephrology and dermatology at the two main public hospitals in northern Trinidad and asking for a list of all SLE patients from their out-patient clinics. At the main dermatology clinic a register of cases since 1992 was available. Furthermore, a systematic search of: (a) outpatient records at the two hospitals, (b) hospital laboratory test results positive for auto-antibodies (anti-nuclear or anti-double-stranded DNA antibody titre >1:256) and (c) histological reports of skin biopsy examination consistent with SLE was performed. Lastly, SLE cases were also identified through the Lupus Society of Trinidad and Tobago (90% of those patients were also identified through one of the two main public hospitals). For each case, randomly chosen households in the same neighbourhood were sampled by the field team to obtain (where possible) two controls, matched with the case for sex and for 20-year age group. Cases

and controls were interviewed at home or in the project office by using a custom made questionnaire.

The case definition of SLE was based on American Rheumatism Association (ARA) criteria [91], applied to medical records (available for more than 90% of cases), and to the medical history given by the patient. Informed consent for blood sampling and the use of the sample for genetic studies including estimation of admixture was obtained from each participant. Initial case ascertainment identified 264 possible cases of SLE. Of these, 72 (27%) were excluded either on the basis of their names or because their medical history did not meet ARA criteria for the diagnosis of SLE. Of the remaining 192 individuals, 54 had incomplete addresses or were not resident in northern Trinidad, four were too ill to be interviewed, eight were aged less than 18 years and two refused to participate. For 80% (99/124) of cases, two matched controls were obtained: the response rate from those invited to participate as controls was 70%. The total sample consisted of 124 cases and 219 controls aged over 20 years who completed the questionnaire. Blood samples were obtained from 122 cases and 219 controls and DNA was successfully extracted from 93% (317/341) of these. IgG glycans were successfully measured in 303 individuals. Age at sampling was not available for 17 individuals and 2 individuals were lost due to the ID mismatch.

To test predictive power of selected glycan trait, we fitted logistic regression models (including and excluding the glycan) and used predRisk function of PredictABEL package for R to evaluate the predictive ability.

Supporting Information

Figure S1 Forrest plots for associations of glycan traits measured by UPLC and genetic polymorphisms.
(PPT)

Table S1 The description of 23 quantitative IgG glycosylation traits measured by UPLC and 54 derived traits.
(XLS)

Table S2 Descriptive statistics of glycan traits in discovery cohorts.
(XLS)

Table S3 Summary data for all single-nucleotide polymorphisms and traits with suggestive associations ($p<1\times10^{-5}$) with glycans measured by UPLC.
(XLS)

Table S4 Population-specific and pooled genomic control (GC) factors for associations with UPLC glycan traits.
(XLS)

Table S5 Description of five glycan traits measured by MS and their descriptive statistic in the replication cohort.
(XLS)

Table S6 Summary data for all single-nucleotide polymorphisms with replicated in the LLS cohort.
(XLS)

Table S7 IgG glycans in 5 heterozygous *Ikzf1* knockout mice and 5 wild-type controls.
(XLS)

Table S8 Data for all IgG N-glycans measured in 101 Afro-Caribbean cases with SLE and 183 controls (Extended Table 4 from the main manuscript).
(XLS)

Table S9 Effects of corticosteroids on IgG glycans.
(XLS)

Table S10 Principal component analysis of IgG glycosylation traits.
(XLS)

Table S11 Description of the analysed populations.
(XLS)

Acknowledgments

The CROATIA-Vis and CROATIA-Korcula studies would like to acknowledge the invaluable contributions of the recruitment team (including those from the Institute of Anthropological Research in Zagreb) in Vis and Korcula, the administrative teams in Croatia and Edinburgh, and the people of Vis and Korcula. ORCADES would like to acknowledge the invaluable contributions of Lorraine Anderson, the research nurses in Orkney, and the administrative team in Edinburgh. SNP genotyping of the CROATIA-Vis samples and DNA extraction for ORCADES was carried out by the Genetics Core Laboratory at the Wellcome Trust Clinical Research Facility, WGH, Edinburgh, Scotland. SNP genotyping for CROATIA-Korcula and ORCADES was performed by Helmholtz Zentrum München, GmbH, Neuherberg, Germany. The authors are grateful to all patients and staff who contributed to the collection in Trinidad and the UK.

Author Contributions

Conceived and designed the experiments: GL MW AFW PMR CH YA HC IR. Performed the experiments: JEH MP BA AM MN JK TK BS LRR. Analyzed the data: LZ OP OG VV H-WU MM OS JJH-D PES MB YA. Contributed reagents/materials/analysis tools: ALP PM IK IKL FNvL AJMdC AMD QZ WW NDH UG JFW PMR. Wrote the paper: GL IR.

References

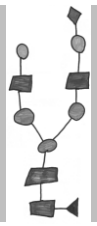
1. Opdenakker G, Rudd PM, Ponting CP, Dwek RA (1993) Concepts and principles of glycobiology. FASEB Journal 7: 1330–1337.
2. Skropeta D (2009) The effect of individual N-glycans on enzyme activity. Bioorg. Med.Chem. 17: 2645–2653.
3. Marek KW, Vijay IK, Marth JD (1999) A recessive deletion in the GlcNAc-1-phosphotransferase gene results in peri-implantation embryonic lethality. Glycobiology 9: 1263–1271.
4. Jaeken J, Matthijs G (2007) Congenital disorders of glycosylation: a rapidly expanding disease family. Annu Rev Genomics Hum Genet 8: 261–278.
5. Kobata A (2008) The N-linked sugar chains of human immunoglobulin G: their unique pattern, and their functional roles. Biochim Biophys Acta 1780: 472–478.
6. Jefferis R (2005) Glycosylation of recombinant antibody therapeutics. Biotechnol Prog 21: 11–16.
7. Zhu D, Ottensmeier CH, Du MQ, McCarthy H, Stevenson FK (2003) Incidence of potential glycosylation sites in immunoglobulin variable regions distinguishes between subsets of Burkitt's lymphoma and mucosa-associated lymphoid tissue lymphoma. Br J Haematol 120: 217–222.
8. Sutton BJ, Phillips DC (1983) The three-dimensional structure of the carbohydrate within the Fc fragment of immunoglobulin G. Biochem Soc Trans 11: 130–132.
9. Harada H, Kamei M, Tokumoto Y, Yui S, Koyama F, et al. (1987) Systematic fractionation of oligosaccharides of human immunoglobulin G by serial affinity chromatography on immobilized lectin columns. Anal Biochem 164: 374–381.
10. Parekh RB, Dwek RA, Sutton BJ, Fernandes DL, Leung A, et al. (1985) Association of rheumatoid arthritis and primary osteoarthritis with changes in the glycosylation pattern of total serum IgG. Nature 316: 452–457.
11. Kaneko Y, Nimmerjahn F, Ravetch JV (2006) Anti-inflammatory activity of immunoglobulin G resulting from Fc sialylation. Science 313: 670–673.
12. Anthony RM, Ravetch JV (2010) A novel role for the IgG Fc glycan: the anti-inflammatory activity of sialylated IgG Fcs. J Clin Immunol 30 Suppl 1: S9–14.
13. Nimmerjahn F, Ravetch JV (2008) Fcgamma receptors as regulators of immune responses. Nat Rev Immunol 8: 34–47.
14. Ferrara C, Stuart F, Sondermann P, Brunker P, Umana P (2006) The carbohydrate at FcgammaRIIIa Asn-162. An element required for high affinity binding to non-fucosylated IgG glycoforms. J Biol Chem 281: 5032–5036.

15. Mizushima T, Yagi H, Takemoto E, Shibata-Koyama M, Isoda Y, et al. (2011) Structural basis for improved efficacy of therapeutic antibodies on defucosylation of their Fc glycans. Genes to Cells 16: 1071–1080.
16. Shinkawa T, Nakamura K, Yamane N, Shoji-Hosaka E, Kanda Y, et al. (2003) The absence of fucose but not the presence of galactose or bisecting N-acetylglucosamine of human IgG1 complex-type oligosaccharides shows the critical role of enhancing antibody-dependent cellular cytotoxicity. J Biol Chem 278: 3466–3473.
17. Iida S, Misaka H, Inoue M, Shibata M, Nakano R, et al. (2006) Nonfucosylated therapeutic IgG1 antibody can evade the inhibitory effect of serum immunoglobulin G on antibody-dependent cellular cytotoxicity through its high binding to FcgammaRIIIa. Clin Cancer Res 12: 2879–2887.
18. Preithner S, Elm S, Lippold S, Locher M, Wolf A, et al. (2006) High concentrations of therapeutic IgG1 antibodies are needed to compensate for inhibition of antibody-dependent cellular cytotoxicity by excess endogenous immunoglobulin G. Mol Immunol 43: 1183–1193.
19. Pucic M, Knezevic A, Vidic J, Adamczyk B, Novokmet M, et al. (2011) High throughput isolation and glycosylation analysis of IgG-variability and heritability of the IgG glycome in three isolated human populations. Mol Cell Proteomics 10: M111 010090.
20. Kooner JS, Saleheen D, Sim X, Sehmi J, Zhang W, et al. (2011) Genome-wide association study in individuals of South Asian ancestry identifies six new type 2 diabetes susceptibility loci. Nat Genet 43: 984–989.
21. Franke A, McGovern DP, Barrett JC, Wang K, Radford-Smith GL, et al. (2010) Genome-wide meta-analysis increases to 71 the number of confirmed Crohn's disease susceptibility loci. Nat Genet 42: 1118–1125.
22. Mells GF, Floyd JA, Morley KI, Cordell HJ, Franklin CS, et al. (2011) Genome-wide association study identifies 12 new susceptibility loci for primary biliary cirrhosis. Nat Genet 43: 329–332.
23. Aouizerat BE, Vittinghoff E, Musone SL, Pawlikowska L, Kwok PY, et al. (2011) GWAS for discovery and replication of genetic loci associated with sudden cardiac arrest in patients with coronary artery disease. BMC Cardiovasc Disord 11: 29.
24. Baranzini SE, Srinivasan R, Khankhanian P, Okuda DT, Nelson SJ, et al. (2010) Genetic variation influences glutamate concentrations in brains of patients with multiple sclerosis. Brain 133: 2603–2611.
25. Dick DM, Aliev F, Krueger RF, Edwards A, Agrawal A, et al. (2011) Genome-wide association study of conduct disorder symptomatology. Mol Psychiatry 16: 800–808.
26. Pivac N, Knezevic A, Gornik O, Pucic M, Igl W, et al. (2011) Human plasma glycome in attention-deficit hyperactivity disorder and autism spectrum disorders. Mol Cell Proteomics 10: M110 004200.
27. Huffman JE, Knezevic A, Vitart V, Kattla J, Adamczyk B, et al. (2011) Polymorphisms in B3GAT1, SLC9A9 and MGAT5 are associated with variation within the human plasma N-glycome of 3533 European adults. Hum Mol Genet 20: 5000–5011.
28. Oda Y, Okada T, Yoshida H, Kaufman RJ, Nagata K, et al. (2006) Derlin-2 and Derlin-3 are regulated by the mammalian unfolded protein response and are required for ER-associated degradation. J Cell Biol 172: 383–393.
29. Pottier N, Cheok MH, Yang W, Assem M, Tracey L, et al. (2007) Expression of SMARCB1 modulates steroid sensitivity in human lymphoblastoid cells: identification of a promoter SNP that alters PARP1 binding and SMARCB1 expression. Hum Mol Genet 16: 2261–2271.
30. Chambers JC, Zhang W, Sehmi J, Li X, Wass MN, et al. (2011) Genome-wide association study identifies loci influencing concentrations of liver enzymes in plasma. Nat Genet 43: 1131–1138.
31. Sellars M, Reina-San-Martin B, Kastner P, Chan S (2009) Ikaros controls isotype selection during immunoglobulin class switch recombination. J Exp Med 206: 1073–1087.
32. Klug CA, Morrison SJ, Masek M, Hahm K, Smale ST, et al. (1998) Hematopoietic stem cells and lymphoid progenitors express different Ikaros isoforms, and Ikaros is localized to heterochromatin in immature lymphocytes. Proc Natl Acad Sci U S A 95: 657–662.
33. Trevino LR, Yang W, French D, Hunger SP, Carroll WL, et al. (2009) Germline genomic variants associated with childhood acute lymphoblastic leukemia. Nat Genet 41: 1001–1005.
34. Papaemmanuil E, Hosking FJ, Vijayakrishnan J, Price A, Olver B, et al. (2009) Loci on 7p12.2, 10q21.2 and 14q11.2 are associated with risk of childhood acute lymphoblastic leukemia. Nat Genet 41: 1006–1010.
35. Cunninghame Graham DS, Morris DL, Bhangale TR, Criswell LA, Syvanen AC, et al. (2011) Association of NCF2, IKZF1, IRF8, IFIH1, and TYK2 with Systemic Lupus Erythematosus. PLoS Genet 7: e1002341. doi:10.1371/journal.pgen.1002341
36. Han JW, Zheng HF, Cui Y, Sun LD, Ye DQ, et al. (2009) Genome-wide association study in a Chinese Han population identifies nine new susceptibility loci for systemic lupus erythematosus. Nat Genet 41: 1234–1237.
37. Gateva V, Sandling JK, Hom G, Taylor KE, Chung SA, et al. (2009) A large-scale replication study identifies TNIP1, PRDM1, JAZF1, UHRF1BP1 and IL10 as risk loci for systemic lupus erythematosus. Nat Genet 41: 1228–1233.
38. Davidson A, Diamond B (2001) Autoimmune diseases. N Engl J Med 345: 340–350.
39. Swafford AD, Howson JM, Davison LJ, Wallace C, Smyth DJ, et al. (2011) An allele of IKZF1 (Ikaros) conferring susceptibility to childhood acute lymphoblastic leukemia protects against type 1 diabetes. Diabetes 60: 1041–1044.
40. Barrett JC, Hansoul S, Nicolae DL, Cho JH, Duerr RH, et al. (2008) Genome-wide association defines more than 30 distinct susceptibility loci for Crohn's disease. Nat Genet 40: 955–962.
41. Gorlova O, Martin JE, Rueda B, Koeleman BP, Ying J, et al. (2011) Identification of novel genetic markers associated with clinical phenotypes of systemic sclerosis through a genome-wide association strategy. PLoS Genet 7: e1002178. doi:10.1371/journal.pgen.1002178
42. Jallow M, Teo YY, Small KS, Rockett KA, Deloukas P, et al. (2009) Genome-wide and fine-resolution association analysis of malaria in West Africa. Nat Genet 41: 657–665.
43. Ganesh SK, Zakai NA, van Rooij FJ, Soranzo N, Smith AV, et al. (2009) Multiple loci influence erythrocyte phenotypes in the CHARGE Consortium. Nat Genet 41: 1191–1198.
44. Stahl EA, Raychaudhuri S, Remmers EF, Xie G, Eyre S, et al. (2010) Genome-wide association study meta-analysis identifies seven new rheumatoid arthritis risk loci. Nat Genet 42: 508–514.
45. Gottardo L, De Cosmo S, Zhang YY, Powers C, Prudente S, et al. (2008) A polymorphism at the IL6ST (gp130) locus is associated with traits of the metabolic syndrome. Obesity (Silver Spring) 16: 205–210.
46. Birmann BM, Tamimi RM, Giovannucci E, Rosner B, Hunter DJ, et al. (2009) Insulin-like growth factor-1- and interleukin-6-related gene variation and risk of multiple myeloma. Cancer Epidemiol Biomarkers Prev 18: 282–288.
47. Barrett JC, Lee JC, Lees CW, Prescott NJ, Anderson CA, et al. (2009) Genome-wide association study of ulcerative colitis identifies three new susceptibility loci, including the HNF4A region. Nat. Genet. 41: 1330–1334.
48. Silverberg MS, Cho JH, Rioux JD, McGovern DP, Wu J, et al. (2009) Ulcerative colitis-risk loci on chromosomes 1p36 and 12q15 found by genome-wide association study. Nat Genet 41: 216–220.
49. McGovern DP, Gardet A, Torkvist L, Goyette P, Essers J, et al. (2010) Genome-wide association identifies multiple ulcerative colitis susceptibility loci. Nat Genet 42: 332–337.
50. Anderson CA, Boucher G, Lees CW, Franke A, D'Amato M, et al. (2011) Meta-analysis identifies 29 additional ulcerative colitis risk loci, increasing the number of confirmed associations to 47. Nat Genet 43: 246–252.
51. Asano K, Matsushita T, Umeno J, Hosono N, Takahashi A, et al. (2009) A genome-wide association study identifies three new susceptibility loci for ulcerative colitis in the Japanese population. Nat Genet 41: 1325–1329.
52. Kamio T, Toki T, Kanezaki R, Sasaki S, Tandai S, et al. (2003) B-cell-specific transcription factor BACH2 modifies the cytotoxic effects of anticancer drugs. Blood 102: 3317–3322.
53. Grant SF, Qu HQ, Bradfield JP, Marchand L, Kim CE, et al. (2009) Follow-up analysis of genome-wide association data identifies novel loci for type 1 diabetes. Diabetes 58: 290–295.
54. Cooper JD, Smyth DJ, Smiles AM, Plagnol V, Walker NM, et al. (2008) Meta-analysis of genome-wide association study data identifies additional type 1 diabetes risk loci. Nat Genet 40: 1399–1401.
55. Barrett JC, Clayton DG, Concannon P, Akolkar B, Cooper JD, et al. (2009) Genome-wide association study and meta-analysis find that over 40 loci affect risk of type 1 diabetes. Nat Genet 41: 703–707.
56. Plagnol V, Howson JM, Smyth DJ, Walker N, Hafler JP, et al. (2011) Genome-wide association analysis of autoantibody positivity in type 1 diabetes cases. PLoS Genet 7: e1002216. doi:10.1371/journal.pgen.1002216
57. Chu X, Pan CM, Zhao SX, Liang J, Gao GQ, et al. (2011) A genome-wide association study identifies two new risk loci for Graves' disease. Nat Genet 43: 897–901.
58. Dubois PC, Trynka G, Franke L, Hunt KA, Romanos J, et al. (2010) Multiple common variants for celiac disease influencing immune gene expression. Nat Genet 42: 295–302.
59. Sawcer S, Hellenthal G, Pirinen M, Spencer CC, Patsopoulos NA, et al. (2011) Genetic risk and a primary role for cell-mediated immune mechanisms in multiple sclerosis. Nature 476: 214–219.
60. Matsui T, Leung D, Miyashita H, Maksakova IA, Miyachi H, et al. (2010) Proviral silencing in embryonic stem cells requires the histone methyltransferase ESET. Nature 464: 927–931.
61. Igl W, Polasek O, Gornik O, Knezevic A, Pucic M, et al. (2011) Glycomics meets lipidomics-associations of N-glycans with classical lipids, glycerophospholipids, and sphingolipids in three European populations. Mol Biosyst 7: 1852–1862.
62. Cozen W, Li D, Best T, Van Den Berg DJ, Gourraud PA, et al. (2012) A genome-wide meta-analysis of nodular sclerosing Hodgkin lymphoma identifies risk loci at 6p21.32. Blood 119: 469–475.
63. Chung SA, Taylor KE, Graham RR, Nititham J, Lee AT, et al. (2011) Differential genetic associations for systemic lupus erythematosus based on anti-dsDNA autoantibody production. PLoS Genet 7: e1001323. doi:10.1371/journal.pgen.1001323
64. Hor H, Kutalik Z, Dauvilliers Y, Valsesia A, Lammers GJ, et al. (2010) Genome-wide association study identifies new HLA class II haplotypes strongly protective against narcolepsy. Nat Genet 42: 786–789.
65. Celestino-Soper PB, Shaw CA, Sanders SJ, Li J, Murtha MT, et al. (2011) Use of array CGH to detect exonic copy number variants throughout the genome in autism families detects a novel deletion in TMLHE. Hum Mol Genet 20: 4360–4370.
66. Yashin AI, Wu D, Arbeev KG, Ukraintseva SV (2010) Joint influence of small-effect genetic variants on human longevity. Aging (Albany NY) 2: 612–620.

67. Prasad RB, Hosking FJ, Vijayakrishnan J, Papaemmanuil E, Koehler R, et al. (2010) Verification of the susceptibility loci on 7p12.2, 10q21.2, and 14q11.2 in precursor B-cell acute lymphoblastic leukemia of childhood. Blood 115: 1765–1767.
68. Souabni A, Cobaleda C, Schebesta M, Busslinger M (2002) Pax5 promotes B lymphopoiesis and blocks T cell development by repressing Notch1. Immunity 17: 781–793.
69. Molokhia M, McKeigue P (2006) Systemic lupus erythematosus: genes versus environment in high risk populations. Lupus 15: 827–832.
70. Kundu S, Aulchenko YS, van Duijn CM, Janssens AC (2011) PredictABEL: an R package for the assessment of risk prediction models. Eur J Epidemiol 26: 261–264.
71. Lauc G, Essafi A, Huffman JE, Hayward C, Knežević A, et al. (2010) Genomics meets glycomics - The first GWAS study of human N-glycome identifies HNF1alpha as a master regulator of plasma protein fucosylation. PLoS Genet 6: e1001256. doi:10.1371/journal.pgen.1001256
72. Fukuta K, Abe R, Yokomatsu T, Omae F, Asanagi M, et al. (2000) Control of bisecting GlcNAc addition to N-linked sugar chains. J Biol Chem 275: 23456–23461.
73. Van Damme J, Opdenakker G, Simpson RJ, Rubira MR, Cayphas S, et al. (1987) Identification of the human 26-kD protein, interferon beta 2 (IFN-beta 2), as a B cell hybridoma/plasmacytoma growth factor induced by interleukin 1 and tumor necrosis factor. J Exp Med 165: 914–919.
74. Shields RL, Lai J, Keck R, O'Connell LY, Hong K, et al. (2002) Lack of fucose on human IgG1 N-linked oligosaccharide improves binding to human Fcgamma RIII and antibody-dependent cellular toxicity. J Biol Chem 277: 26733–26740.
75. Albert H, Collin M, Dudziak D, Ravetch JV, Nimmerjahn F (2008) In vivo enzymatic modulation of IgG glycosylation inhibits autoimmune disease in an IgG subclass-dependent manner. Proc Natl Acad Sci U S A 105: 15005–15009.
76. Baudino L, Azeredo da Silveira S, Nakata M, Izui S (2006) Molecular and cellular basis for pathogenicity of autoantibodies: lessons from murine monoclonal autoantibodies. Springer Semin Immunopathol 28: 175–184.
77. Sivakumaran S, Agakov F, Theodoratou E, Prendergast JG, Zgaga L, et al. (2011) Abundant pleiotropy in human complex diseases and traits. Am J Hum Genet 89: 607–618.
78. Youings A, Chang SC, Dwek RA, Scragg IG (1996) Site-specific glycosylation of human immunoglobulin G is altered in four rheumatoid arthritis patients. Biochem J 314 (Pt 2): 621–630.
79. Wormald MR, Rudd PM, Harvey DJ, Chang SC, Scragg IG, et al. (1997) Variations in oligosaccharide-protein interactions in immunoglobulin G determine the site-specific glycosylation profiles and modulate the dynamic motion of the Fc oligosaccharides. Biochemistry 36: 1370–1380.
80. Mimura Y, Ashton PR, Takahashi N, Harvey DJ, Jefferis R (2007) Contrasting glycosylation profiles between Fab and Fc of a human IgG protein studied by electrospray ionization mass spectrometry. J Immunol Methods 326: 116–126.
81. Wang J, Balog CI, Stavenhagen K, Koeleman CA, Scherer HU, et al. (2011) Fc-glycosylation of IgG1 is modulated by B-cell stimuli. Mol Cell Proteomics 10: M110 004655.
82. Vitart V, Rudan I, Hayward C, Gray NK, Floyd J, et al. (2008) SLC2A9 is a newly identified urate transporter influencing serum urate concentration, urate excretion and gout. Nat Genet 40: 437–442.
83. Zemunik T, Boban M, Lauc G, Jankovic S, Rotim K, et al. (2009) Genome-wide association study of biochemical traits in Korcula Island, Croatia. Croat Med J 50: 23–33.
84. McQuillan R, Leutenegger AL, Abdel-Rahman R, Franklin CS, Pericic M, et al. (2008) Runs of homozygosity in European populations. Am J Hum Genet 83: 359–372.
85. Schoenmaker M, de Craen AJ, de Meijer PH, Beekman M, Blauw GJ, et al. (2006) Evidence of genetic enrichment for exceptional survival using a family approach: the Leiden Longevity Study. Eur J Hum Genet 14: 79–84.
86. Ruhaak LR, Uh HW, Beekman M, Koeleman CA, Hokke CH, et al. (2010) Decreased levels of bisecting GlcNAc glycoforms of IgG are associated with human longevity. PLoS ONE 5: e12566. doi:10.1371/journal.pone.0012566
87. Aulchenko YS, Ripke S, Isaacs A, van Duijn CM (2007) GenABEL: an R library for genome-wide association analysis. Bioinformatics 23: 1294–1296.
88. Pe'er I, Yelensky R, Altshuler D, Daly MJ (2008) Estimation of the multiple testing burden for genomewide association studies of nearly all common variants. Genet Epidemiol 32: 381–385.
89. Pruim RJ, Welch RP, Sanna S, Teslovich TM, Chines PS, et al. (2010) LocusZoom: regional visualization of genome-wide association scan results. Bioinformatics 26: 2336–2337.
90. Uh HW, Deelen J, Beekman M, Helmer Q, Rivadeneira F, et al. (2011) How to deal with the early GWAS data when imputing and combining different arrays is necessary. Eur J Hum Genet 20: 572–576.
91. Tan EM, Cohen AS, Fries JF, Masi AT, McShane DJ, et al. (1982) The 1982 revised criteria for the classification of systemic lupus erythematosus. Arthritis Rheum 25: 1271–1277.

Changes in plasma and IgG *N*-glycome during childhood and adolescence

5



Glycobiology vol. 22 no. 7 pp. 975–982, 2012
doi:10.1093/glycob/cws062
Advance Access publication on March 16, 2012

Changes in plasma and IgG *N*-glycome during childhood and adolescence

Maja Pučić[2], Ana Mužinić[2], Mislav Novokmet[2], Marijana Škledar[3], Nela Pivac[4], Gordan Lauc[2,5], and Olga Gornik[1,5]

[2]Genos Ltd, Glycobiology Laboratory, 10000 Zagreb, Croatia; [3]Institute of Public Health of Zagreb County, Zaprešić, Croatia; [4]Division of Molecular Medicine, Ruđer Bošković Institute, 10000 Zagreb, Croatia; and [5]Faculty of Pharmacy and Biochemistry, University of Zagreb, Ante Kovačića 1, 10000 Zagreb, Croatia

Received on February 22, 2012; revised on March 13, 2012; accepted on March 13, 2012

Despite the importance of protein glycosylation in all physiological and pathological processes and their potential as diagnostic markers and drug targets, the glycome of children is still unexplored. We analyzed *N*-linked plasma and IgG glycomes in 170 children and adolescents between 6 and 18 years of age. The results showed large biological variability at the population level as well as a large number of associations between different glycans and age. The plasma *N*-glycome of younger children was found to contain a larger proportion of large complex glycan structures ($r = -0.71$ for tetrasialylated glycans; $r = -0.41$ for trisialylated glycans) as well as an increase in disialylated biantennary structures ($r = 0.55$) with age. Core fucosylation and the level of agalactosylated plasma and IgG glycans decreased while digalactosylated glycans increased with age. This pattern of age-dependent changes in children differs from changes reported in adult population in both, direction and the intensity of changes. Also, sex differences are much smaller in children than in adults and are present mainly during puberty. These important observations should be accounted for when glycan-based diagnostic tests or therapeutics are being developed or evaluated.

Keywords: ageing / children plasma and IgG glycome / glycan analysis / *N*-glycans / protein glycosylation

Introduction

Contrary to proteins which are defined by a single gene, glycans are a result of an interplay between hundreds of genes and their products and are therefore inherently sensitive to all changes in the physiology of the cell. Many pathological conditions are associated with changes in glycan structures and these changes are promising candidates for novel diagnostic and prognostic tools (Freeze 2006; Lebrilla and An 2009; Jaeken 2010; Reis et al. 2010). Understanding of the role of glycan in various pathological conditions is also a good foundation for the development of novel therapeutics. The knowledge about variability of glycosylation in healthy individuals and the influence of some normal physiological and environmental factors are an essential prerequisite for any further study. However, this knowledge is still limited and first comprehensive population studies of human plasma and IgG *N*-glycomes were performed only recently (Knežević et al. 2009; Ruhaak et al. 2010; Pučić et al. 2011). Large variability in glycome composition between individuals was observed in human populations, but within a single healthy individual, the composition of plasma glycome was found to be very stable (Gornik et al. 2009) and environmental factors had only a limited impact on the majority of glycans (Knežević et al. 2010). Specific altered individual glyco-phenotypes that can be associated with specific pathologies were also identified (Pučić et al. 2010).

Previous glycan analyses were performed exclusively on adult population and glycosylation in children, especially healthy children, is still nearly unexplored. In this work, we analyzed *N*-glycosylation profiles of plasma and IgG samples taken from children and adolescents of school age in order to determine the levels of specific glycans in healthy children and to establish sex- and age-dependent changes.

Results and discussion

Plasma *N*-glycans were quantified in plasma of 170 children (84 boys, 86 girls, median age 11 years) by hydrophilic interaction high-performance liquid chromatography (HPLC) and weak anion exchange (WAX) chromatography. Additionally, IgG was isolated and IgG *N*-glycans quantified in 164 children (68 boys, 96 girls, median age 13) by hydrophilic interaction ultra performance liquid chromatography. Main glycome features were deduced from ratios of different chromatographic peaks as described previously (Knežević et al. 2010; Pučić et al. 2011). Descriptives and sex differences of *N*-linked glycan structure characteristics of plasma proteins and IgG are shown in Table I (for full descriptive statistics, see Supplementary data, Table SI).

[1]To whom correspondence should be addressed: Tel: +385 1 639 4467; Fax: +385 1 639 4400; e-mail: ogornik@pharma.hr

Table I. Descriptive statistics and sex difference of the plasma and IgG *N*-glycomes in healthy children and adolescents (6–18 years of age)

	6–10 years			11–14 years			15–18 years		
	Girls [*N* = 45; median (IQR)]	Boys [*N* = 40; median (IQR)]	Sex difference (*P*-value)	Girls [*N* = 18; median (IQR)]	Boys [*N* = 24; median (IQR)]	Sex difference (*P*-value)	Girls [*N* = 23; median (IQR)]	Boys [*N* = 20; median (IQR)]	Sex difference (*P*-value)
Plasma glycan feature									
Total plasma glycans (neutral + charged)									
Sialylation									
Monosialylated	21.81 (2.91)	21.52 (3.76)	0.958	22.23 (3.89)	20.73 (3.15)	0.084	22.32 (3.89)	21.34 (2.82)	0.465
Disialylated	58.73 (3.42)	58.64 (2.44)	0.996	59.35 (1.87)	62.12 (2.5)	0.003	60.1 (2.98)	62.41 (1.97)	**<0.001**
Trisialylated	16.43 (2.53)	16.115 (2.34)	0.631	15.24 (3.35)	14.84 (2.46)	0.809	15.48 (3.92)	13.54 (2.75)	0.007
Tetrasialylated	3.41 (1.38)	3.54 (0.99)	0.311	2.63 (0.83)	2.43 (0.60)	0.500	2.33 (0.65)	2.23 (0.58)	0.488
Fucosylation									
Core fucose	22.77 (5.18)	24.81 (5.38)	0.100	21.24 (5.0)	20.67 (4.45)	0.170	22.67 (5.68)	21.58 (3.78)	0.032
Antennary fucose	2.57 (0.76)	2.52 (0.69)	0.579	3.18 (1.06)	2.64 (1.08)	0.147	2.87 (1.08)	2.81 (0.96)	0.422
Branching									
Biantennary	78.73 (2.75)	78.57 (2.55)	0.721	79.87 (3.49)	80.49 (2.68)	0.347	79.85 (3.98)	82.14 (1.76)	0.002
Triantennary	16.35 (1.4)	16.43 (1.93)	0.570	16.36 (2.96)	15.64 (2.00)	0.354	16.1 (2.78)	14.43 (1.99)	**0.001**
Tetraantennary	4.59 (1.48)	4.9 (1.01)	0.283	4.12 (1.01)	3.72 (0.84)	0.263	3.97 (0.82)	3.48 (0.85)	0.053
Sialylation of biantennary glycans									
Monosialylated	29.33 (2.56)	29.44 (2.77)	0.867	28.67 (3.21)	27.91 (3.15)	0.121	28.2 (2.37)	27.86 (2.63)	0.827
Disialylated	69.56 (3.67)	69.50 (3.85)	0.653	70.92 (5.02)	72.36 (3.49)	0.109	69.96 (4.29)	72.20 (3.95)	0.119
Galactosylation									
Agalactosylated (G0)	3.59 (1.27)	3.83 (1.23)	0.337	2.50 (1.11)	3.69 (1.37)	**0.001**	2.89 (1.08)	3.34 (1.69)	0.088
Monogalactosylated (G1)	8.87 (1.61)	8.88 (1.51)	0.535	8.47 (2.25)	9.02 (1.33)	0.360	9.05 (1.99)	9.32 (2.73)	0.527
Digalactosylated (G2)	66.01 (3.41)	65.74 (3.87)	0.546	68.12 (3.12)	67.79 (2.99)	0.559	68.28 (3.58)	69.27 (2.46)	0.111
Trigalactosylated (G3)	14.47 (2.55)	14.79 (2.30)	0.895	13.36 (3.36)	13.01 (2.25)	0.731	13.8 (2.91)	11.63 (2.17)	**0.001**
Tetragalactosylated (G4)	2.46 (0.87)	2.86 (0.59)	0.140	2.12 (0.79)	2.05 (0.23)	0.576	1.87 (0.58)	1.72 (0.55)	0.197
	Girls [*N* = 28; median (IQR)]	Boys [*N* = 14; median (IQR)]		Girls [*N* = 34; median (IQR)]	Boys [*N* = 28; median (IQR)]		Girls [*N* = 34; median (IQR)]	Boys [*N* = 26; median (IQR)]	
IgG glycan feature									
Total IgG glycans (neutral + charged)									
Sialylation									
FGS/(FG + FGS)	23.67 (25.56)	25.58 (13.69)	0.762	24.52 (20.81)	25.06 (24.61)	0.240	27.46 (25.14)	26.04 (16.63)	0.421
FBGS/(FBG + FBGS)	36.33 (36.15)	28.98 (20.0)	0.208	34.46 (36.66)	39.01 (36.21)	0.141	41.51 (39.07)	40.22 (31.52)	0.811
FGS/(F + FG + FGS)	15.79 (22.95)	18.02 (10.67)	0.607	16.98 (20.93)	18.14 (22.06)	0.343	22.68 (25.11)	19.07 (17.89)	0.239
FBGS/(FB + FBG + FBGS)	25.60 (31.62)	19.67 (17.36)	0.284	24.53 (31.13)	27.35 (32.93)	0.235	31.27 (35.14)	29.89 (27.38)	0.623
FG1S1/(FG1 + FG1S1)	8.78 (8.08)	8.83 (3.82)	0.989	8.44 (8.47)	8.97 (8.32)	0.641	9.2 (9.20)	9.49 (4.58)	0.502
FG2S1/(FG2 + FG2S1 + FG2S2)	36.95 (21.35)	38.54 (10.62)	0.553	35.7 (19.68)	37.2 (15.59)	0.315	37.0 (20.55)	38.13 (11.18)	0.395
FG2S2/(FG2 + FG2S1 + FG2S2)	8.83 (10.07)	7.17 (5.39)	0.308	6.89 (12.20)	10.08 (13.72)	0.014	8.36 (9.6)	9.67 (11.19)	0.071
FBG2S1/(FBG2 + FBG2S1 + FBG2S2)	36.12 (13.18)	33.11 (3.18)	0.004	34.38 (15.46)	34.49 (11.29)	0.955	34.27 (15.11)	35.84 (7.5)	0.007
FBG2S2/(FBG2 + FBG2S1 + FBG2S2)	35.77 (28.57)	30.5 (16.32)	0.484	33.74 (27.41)	37.89 (26.84)	0.045	37.46 (32.81)	38.70 (25.5)	0.363
$F^{total}S1/F^{total}S2$	3.4 (4.45)	4.42 (2.22)	0.085	4.11 (5.33)	3.31 (4.03)	0.087	3.32 (4.62)	3.32 (4.22)	0.387
FS1/FS2	5.32 (8.01)	6.85 (3.25)	0.452	7.18 (11.34)	5.11 (6.88)	0.010	5.51 (7.84)	5.07 (6.58)	0.200
FBS1/FBS2	0.97 (1.02)	1.02 (0.55)	0.927	1.22 (1.36)	1.0 (0.75)	0.049	0.96 (1.62)	0.96 (1.01)	0.754
Bisecting *N*-GlcNAc									
FBS^{total}/FS^{total}	0.27 (0.26)	0.18 (0.07)	**<0.001**	0.27 (0.27)	0.23 (0.19)	0.054	0.25 (0.31)	0.25 (0.2)	0.403
FBS1/FS1	0.16 (0.16)	0.1 (0.03)	**<0.001**	0.17 (0.17)	0.13 (0.13)	0.007	0.14 (0.18)	0.15 (0.12)	0.493
FBS1/(FS1 + FBS1)	0.14 (0.12)	0.09 (0.02)	**<0.001**	0.14 (0.12)	0.11 (0.1)	0.007	0.13 (0.13)	0.13 (0.08)	0.493
FBS2/FS2	0.93 (0.91)	0.65 (0.24)	0.005	0.94 (0.99)	0.73 (0.82)	**<0.001**	0.88 (0.94)	0.84 (0.49)	0.31
FBS2/(FS2 + FBS2)	0.48 (0.27)	0.39 (0.09)	0.005	0.49 (0.26)	0.42 (0.26)	**<0.001**	0.47 (0.24)	0.46 (0.14)	0.31

Neutral IgG glycans									
Galactosylation									
$G0^n$	38.53 (29.0)	37.68 (7.01)	0.571	35.14 (31.56)	35.77 (24.05)	0.601	29.04 (36.5)	32.96 (22.49)	0.021
$G1^n$	43.68 (15.93)	43.51 (3.09)	0.802	44.41 (12.19)	43.98 (11.99)	0.591	45.45 (12.28)	45.17 (8.39)	0.917
$G2^n$	17.83 (16.56)	19.44 (4.70)	0.284	19.94 (22.01)	20.13 (17.38)	0.621	24.92 (28.58)	20.68 (16.03)	0.008
Core fucosylation and bisecting GlcNAc									
$F^{n\ total}$	97.61 (3.60)	97.13 (1.37)	0.147	96.75 (4.14)	97.01 (4.3)	0.396	96.65 (5.07)	96.55 (3.68)	0.561
$FG0^{n\ total}/G0^n$	98.44 (3.40)	98.13 (1.16)	0.26	98.15 (5.02)	98.20 (5.56)	0.989	98.27 (3.75)	97.95 (3.57)	0.216
$FG1^{n\ total}/G1^n$	98.77 (1.68)	98.54 (0.67)	0.101	98.43 (3.03)	98.52 (2.97)	0.799	98.53 (3.51)	98.52 (1.96)	0.296
$FG2^{n\ total}/G2^n$	94.18 (8.02)	93.07 (4.06)	0.390	92.63 (13.16)	92.83 (6.51)	0.854	92.95 (11.83)	93.09 (6.33)	0.531
F^n	85.57 (12.51)	85.58 (5.03)	0.947	81.95 (10.29)	85.33 (10.33)	**<0.001**	83.79 (9.35)	83.59 (6.04)	0.665
$FG0^n/G0^n$	85.74 (17.26)	84.19 (7.55)	0.722	81.7 (12.78)	85.02 (13.58)	0.007	82.58 (15.57)	82.17 (8.80)	0.788
$FG1^n/G1^n$	85.51 (11.98)	86.15 (5.39)	0.802	83.1 (8.64)	86.57 (10.87)	**<0.001**	84.16 (9.53)	84.55 (5.48)	0.114
$FG2^n/G2^n$	84.91 (12.73)	84.91 (5.51)	0.885	82.99 (15.86)	84.03 (9.72)	0.095	83.85 (16.98)	84.52 (8.96)	0.823
FB^n	11.98 (10.86)	12.60 (3.84)	0.664	14.66 (7.3)	11.64 (8.93)	**<0.001**	13.39 (7.31)	13.1 (4.8)	0.387
$FBG0^n/G0^n$	12.89 (14.46)	14.32 (7.03)	0.762	16.18 (11.73)	13.13 (10.58)	0.003	15.72 (13.82)	15.11 (6.78)	0.355
$FBG1^n/G1^n$	13.25 (11.19)	13.0 (4.73)	0.518	15.32 (8.82)	12.06 (10.35)	**<0.001**	14.35 (7.98)	13.55 (6.15)	0.064
$FBG2^n/G2^n$	9.42 (7.31)	7.9 (1.52)	0.04	9.48 (7.84)	7.97 (5.86)	**0.001**	9.08 (6.89)	8.34 (5.56)	0.474
FB^n/F^n	0.14 (0.15)	0.15 (0.05)	0.782	0.18 (0.11)	0.14 (0.12)	**<0.001**	0.16 (0.1)	0.16 (0.06)	0.379
$FB^n/F^{n\ total}$	12.30 (11.24)	12.79 (4.08)	0.782	15.06 (7.92)	12.03 (9.32)	**<0.001**	13.7 (7.78)	13.52 (4.87)	0.379

Significant differences are highlighted in bold. For description of glycan features, see Supplementary data, Table SIII. IQR, interquartile range.

Significant sex associated differences in plasma glycan levels have been shown to exist for many glycan groups in the adult population (Knežević et al. 2010). However, in this study, we did not observe any significant differences in plasma glycans between boys and girls before puberty (Table I). Only with the onset of puberty boys and girls started to differ in the level of agalactosylated (G0) plasma glycans (Figure 1A). When the glycosylation of IgG alone was evaluated, sex differences were revealed even before puberty with girls showing higher levels of fucosylated sialylated structures with bisecting GlcNAc (Table I). At the onset of puberty boys and girls differed in several glycan features expressing mainly levels of fucosylated neutral IgG glycans with and without bisecting GlcNAc (Figure 1B–D). Compared with boys, girls showed higher levels of IgG glycans with bisecting GlcNAc which has been reported before in a large scale study of IgG glycosylation in adults where levels of glycans with bisecting GlcNAc were found to be more prevalent in females than in males (Ruhaak et al. 2010).

Changes in the glycosylation of plasma proteins with hormonal status were reported to be associated with pregnancy (van de Geijn et al. 2009), oral hormonal therapies (Brinkman-Van der Linden et al. 1996; Saldova et al. 2012) and menopausal age (Knežević et al. 2010); thus, the observed differences in children at the onset of puberty are not unexpected.

Since the majority of plasma glycans did not differ between boys and girls, further analyses of plasma glycosylation were performed on the entire study population. In contrast, since several IgG glycan features showed sex difference, all the analyses of IgG glycosylation were performed separately for boys and girls. A number of significant correlations between age and plasma and IgG glycans were observed (Table II).

The observed correlations in plasma glycome indicated that many glycan features change with child's growth and development, including the decrease in glycan branching, galactosylation (tri- and tetragalactosylated glycans) and sialylation (tri- and tetrasialylated glycans) with age (Table II). The highest correlation coefficient was observed for tetrasialylated glycans ($r = -0.71$); showing a significant decrease in the complexity of glycan structures in children with age which was accompanied by an increase in disialylated biantennary glycans ($r = 0.27$). Scatter diagrams together with regression lines and coefficients of correlations of glycan complexity with age are shown in Figure 2.

Core fucosylation ($r = -0.35$) and the level of agalactosylated plasma glycans ($r = -0.36$) significantly decreased while the level of digalactosylated glycans ($r = 0.59$) increased with age. Since core fucosylated biantennary glycans come predominantly from IgG (Arnold et al. 2008; Royle et al. 2008), it is not surprising that the core fucosylation of IgG also significantly decreased ($r = -0.34$) and galactosylation (digalactosylated: $r = 0.49$) increased with age as observed in girls (Table II). Age-dependence of IgG glycans in boys revealed a significant increase in the incidence of bisecting GlcNAc in core fucosylated sialylated structures with age ($r \geq 0.44$, Table II).

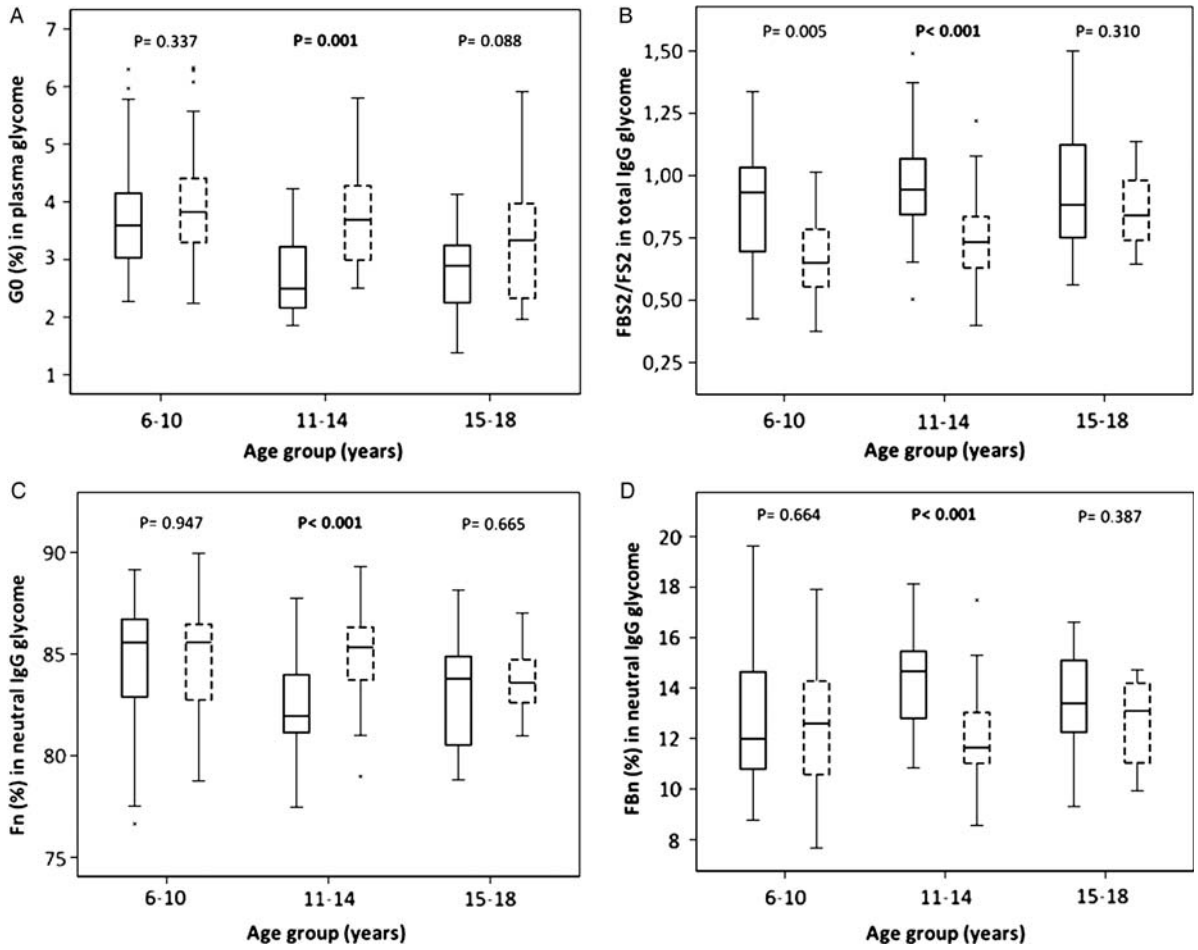


Fig. 1. Levels of plasma and IgG glycan features in three age groups (6–10, 11–14 and 15–18 years) in girls (continuous line) and boys (dashed line). G0 (%), the percentage of agalactosylated glycans in plasma glycome; FBS2/FS2, the ratio of fucosylated disialylated glycans with and without bisecting GlcNAc in the total IgG glycome; F^n (%), the percentage of fucosylated glycans without bisecting GlcNAc in the neutral IgG glycome; FB^n (%), the percentage of fucosylated glycans with bisecting GlcNAc in the neutral IgG glycome. Results are presented as box and whisker plots (median, middle lines; 25–75th percentiles, rectangles; range, lines; outliers, markers).

When the results of the plasma and IgG glycan changes during growing up in children were compared with those in adults, they differed significantly. The behavior of almost all glycan features changed trend or the rate of change in adulthood. An example of the opposite trends in plasma and IgG glycans between children and adults is presented in Figure 3. The increase in agalactosylated glycans with age is the most frequently reported change in glycans with age in adults (Yamada et al. 1997; Vanhooren et al. 2007; Klein 2008; Pučić et al. 2011). Our results revealed that agalactosylated glycans actually decrease with age in children and that the minimal level of agalactosylated glycans is reached in early adulthood (Figure 3). For many other glycan features that showed a significant correlation with age in children, in adulthood levels of glycans either stagnate or change with much lesser rate with age.

Children were then divided into three age groups of 6–10, 11–14 and 15–18 years. Results presented in Supplementary data, Table SII, show that plasma glycans in children significantly changed with the onset of puberty (differences between groups of 6–10 and 11–14 years). As for IgG glycans, significant difference related to the onset of puberty was only observed in boys for the levels of fucosylated sialylated structures with bisecting GlcNAc (Supplementary data, Table SII). In girls, the only significant difference was between the youngest (6–10 years) and oldest group (15–18 years) in the levels of galactosylation and core fucosylation (Supplementary data, Table SII).

Our study revealed that the compositions of the plasma and IgG *N*-glycome in childhood significantly differ from the plasma and IgG *N*-glycome in adulthood and, even more, changed with a great rate during child's growth. This is a very important observation, knowing the fact that changes in glycan structures are often studied as diagnostic biomarkers. Due to the involvement of glycans in many pathophysiological processes, carbohydrate-based therapeutics (such as Tamiflu, Relenza etc.) are being developed, aimed to interfere with, or modify, glycan–receptor bindings, which are parts of pathological cascades. However, the responsible use of these drugs in children requires a careful evaluation of different

Table II. Coefficients of correlation (*r*) between plasma and IgG glycans and age in healthy children and adolescents (6–18 years of age)

Plasma glycan feature	Both sex [*r* (*P*)]	IgG glycan feature	Girls [*r* (*P*)]	Boys [*r* (*P*)]	IgG glycan feature	Girls [*r* (*P*)]	Boys [*r* (*P*)]
Sialylation		*Sialylation*			*Galactosylation*		
Monosialylated	0.05 (0.514)	FGS/(FG + FGS)	0.17 (0.093)	0.01 (0.928)	$G0^n$	**−0.48** (<0.001)	−0.19 (0.130)
Disialylated	**0.55** (<0.001)	FBGS/(FBG + FBGS)	0.08 (0.445)	0.21 (0.081)	$G1^n$	0.28 (0.006)	0.28 (0.023)
Trisialylated	**−0.41** (<0.001)	FGS/(F + FG + FGS)	0.29 (0.005)	0.06 (0.649)	$G2^n$	**0.49** (<0.001)	0.11 (0.374)
Tetrasialylated	**−0.71** (<0.001)	FBGS/(FB + FBG + FBGS)	0.17 (0.099)	0.23 (0.064)	*Core fucosylation and bisecting GlcNAc*		
Fucosylation		FG1S1/(FG1 + FG1S1)	0.11 (0.281)	0.10 (0.400)	$F^{n\ total}$	**−0.34** (0.001)	−0.25 (0.039)
Core fucose	**−0.35** (<0.001)	FG2S1/(FG2 + FG2S1 + FG2S2)	0.02 (0.856)	−0.03 (0.797)	$FG0^{n\ total}/G0^n$	−0.14 (0.168)	−0.21 (0.082)
Antennary fucose	0.19 (0.014)	FG2S2/(FG2 + FG2S1 + FG2S2)	−0.08 (0.413)	0.13 (0.293)	$FG1^{n\ total}/G1^n$	−0.17 (0.097)	−0.18 (0.146)
Branching		FBG2S1/(FBG2 + FBG2S1 + FBG2S2)	−0.22 (0.029)	**0.42** (<0.001)	$FG2^{n\ total}/G2^n$	−0.07 (0.469)	−0.09 (0.452)
Biantennary	**0.48** (<0.001)	BG2S2/(FBG2 + FBG2S1 + FBG2S2)	0.00 (0.985)	0.10 (0.425)	F^n	−0.17 (0.102)	−0.31 (0.011)
Triantennary	**−0.32** (<0.001)	$F^{total}S1/F^{total}S2$	0.00 (0.986)	−0.22 (0.074)	$FG0^n/G0^n$	−0.26 (0.012)	−0.26 (0.034)
Tetraantennary	**−0.63** (<0.001)	FS1/FS2	0.03 (0.796)	−0.13 (0.307)	$FG1^n/G1^n$	−0.13 (0.208)	−0.21 (0.087)
Sialylation of biantennary glycans		FBS1/FBS2	−0.09 (0.402)	0.03 (0.805)	$FG2^n/G2^n$	−0.03 (0.805)	−0.17 (0.177)
Monosialylated	**−0.26** (0.001)	*Bisecting N-GlcNAc*			FB^n	0.08 (0.433)	0.23 (0.063)
Disialylated	**0.27** (<0.001)	FBS^{total}/FS^{total}	−0.04 (0.673)	**0.55** (<0.001)	$FBG0^n/G0^n$	0.25 (0.012)	0.21 (0.088)
Galactosylation		FBS1/FS1	−0.09 (0.381)	**0.59** (<0.001)	$FBG1^n/G1^n$	0.10 (0.339)	0.19 (0.120)
Agalactosylated (G0)	**−0.36** (<0.001)	FBS1/(FS1 + FBS1)	−0.09 (0.381)	**0.59** (<0.001)	$FBG2^n/G2^n$	−0.09 (0.397)	0.26 (0.032)
Monogalactosylated (G1)	0.00 (0.952)	FBS2/FS2	−0.02 (0.82)	**0.44** (<0.001)	FB^n/F^n	0.09 (0.366)	0.23 (0.055)
Digalactosylated (G2)	**0.59** (<0.001)	FBS2/(FS2 + FBS2)	−0.02 (0.82)	**0.44** (<0.001)	$FB^n/F^{n\ total}$	0.09 (0.366)	0.23 (0.055)
Trigalactosylated (G3)	**−0.47** (<0.001)						
Tetragalactosylated (G4)	**−0.68** (<0.001)						

Significant differences are highlighted in bold. For description of glycan features, see Supplementary data, Table SIII.

glycosylation profiles between children and adults, as well as the knowledge of the significant changes occurring in glycan structures with age in young population.

Materials and methods

Study population

Blood samples were collected from healthy children of preschool and school age during their regular medical examination at primary care units in Croatia. Blood samples were collected on anticoagulant; plasmas were immediately separated by centrifugation and stored at −20°C. The study was approved by the Medical School Osijek Ethical Committee and performed in conformance to the ethical guidelines of the Declaration of Helsinki. Plasma *N*-glycosylation was analyzed in 170 children (84 boys, 86 girls, median age 11 years, range 6–18 years) and IgG *N*-glycosylation was analyzed in 164 children (68 boys, 96 girls, median age 13, range 6–18 years).

IgG purification

Immunoglobulin G was isolated from plasma by affinity chromatography using 96-well protein G monolithic plates as described previously (Pučić et al. 2011). Briefly, 50 µL of plasma was diluted 10× with PBS, applied to the protein G plate and instantly washed. IgGs were eluted with 1 mL of 0.1 M formic acid and immediately neutralized with 1 M ammonium bicarbonate.

Glycan analysis

Glycan release and labeling was performed essentially as reported by Royle et al. (2008). Proteins were immobilized in a block of sodium dodecyl sulfate–polyacrylamide gel and *N*-glycans were released by digestion with recombinant *N*-glycosidase F. This was done in a 96-well microtiter plate to achieve the best throughput of sample preparation. After extraction, glycans were fluorescently labeled with 2-aminobenzamide.

Hydrophilic interaction chromatography

Released plasma glycans were separated by HPLC on a 250 × 4.6 mm i.d. 5 µm particle packed TSKgel Amide 80 column (Tosoh Bioscience, Stuttgart, Germany) at 30°C with 50 mM formic acid adjusted to pH 4.4 with ammonia solution as solvent A and acetonitrile as solvent B in a 48-min analytical run (Knežević et al. 2009). Released IgG glycans were separated by ultra performance liquid chromatography on a Waters BEH glycan column, 100 × 2.1 mm i.d., 1.7 µm BEH particles, at 60°C with 100 mM ammonium formate, pH 4.4, as solvent A and acetonitrile as solvent B in a 20-min analytical run (Pučić et al. 2011). In both cases, a fluorescence detector was set with excitation and emission wavelengths of 330 and 420 nm, respectively. The systems were calibrated using an external standard of hydrolyzed and 2AB-labeled glucose oligomers from which the retention times for the individual glycans were converted to glucose units (Royle et al. 2008).

Plasma chromatograms obtained were all separated in the same manner to 16 chromatographic peaks and 13 for

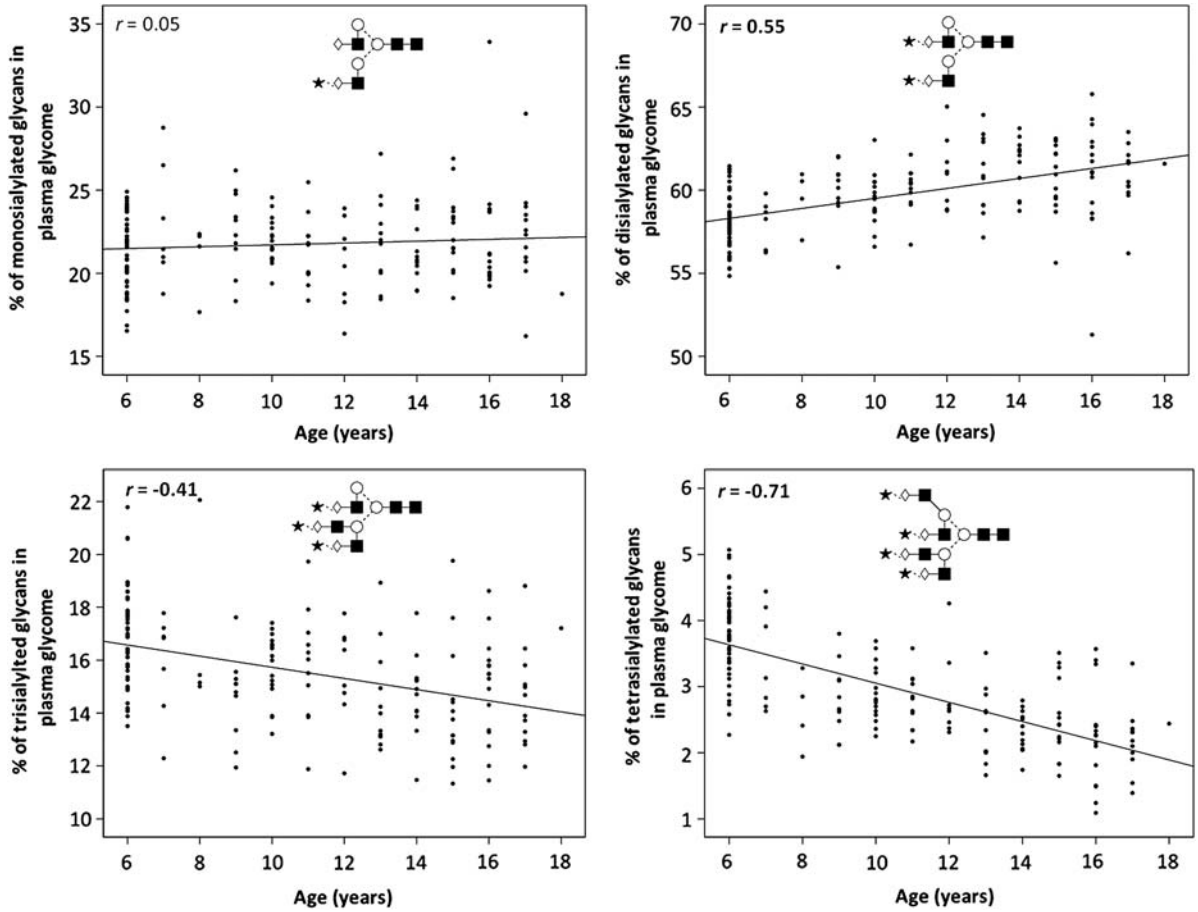


Fig. 2. Age-dependent decrease in branching and sialylation of plasma *N*-linked glycans in children (glycan scheme: square, *N*-acetylglucosamine; circle, mannose; romb, galactose; star, sialic acid). Significant correlation coefficients (*r*) in bold.

desialylated glycans, whereas the IgG chromatograms were separated to 24 peaks. The amounts of glycans present in each peak were expressed as % of the total integrated area.

Weak anion exchange HPLC

Additionally, plasma glycans were separated according to the number of sialic acids by WAX HPLC. The analysis was performed using a Prozyme GlycoSep C 75 mm × 7.5 mm column (Prozyme, Leandro, CA) at 30°C with 20% (v/v) acetonitrile in water as solvent A and 0.1 M acetic acid adjusted to pH 7.0 with ammonia solution in 20% (v/v) acetonitrile as solvent B. Compounds were retained on the column according to their charge density, the higher charged compounds being retained the longest. A fetuin *N*-glycan standard was used for calibration. Glycans were quantified from WAX profiles according to the level of sialylation (monosialylated, disialylated, trisialylated and tetrasialylated).

Sialidase digestion

Aliquots of the 2AB-labeled plasma glycan pool were dried down in 96-well polymerase chain reaction plates. To these, the following was added: 1 μL of 500 mM sodium acetate incubation buffer (pH 5.5), 1 μL (0.005 units) of *Arthrobacter ureafaciens* sialidase (releases α2-3,6,8 sialic acid, Prozyme) and H_2O to make up to 10 μL. This was incubated overnight (16–18 h) at 37°C and then passed through the AcroPrep™ 96 Filter Plates, 350 μL well, 10 K (Pall Corporation, Port Washington, NY) before applying to the HPLC.

Glycan structural features

Levels of glycans sharing the same structural features were approximated by adding the structures having the same structural characteristics (Supplementary data, Table SIII). Plasma glycans structural features were derived from either hydrophilic interaction chromatography [HILIC; total plasma glycans (GP) and total plasma glycans after sialidase treatment (DG)] or WAX integrated glycan profiles (Knežević et al. 2010). IgG glycan structural features were derived from HILIC integrated glycan profiles as described previously (Pučić et al. 2011). Individual glycan structures present in each plasma or IgG glycan peak were reported previously and are shown in Supplementary data, Table SIV.

Statistical analysis

The descriptive part of the statistical analysis was aimed at showing the basic characteristics of the population. The

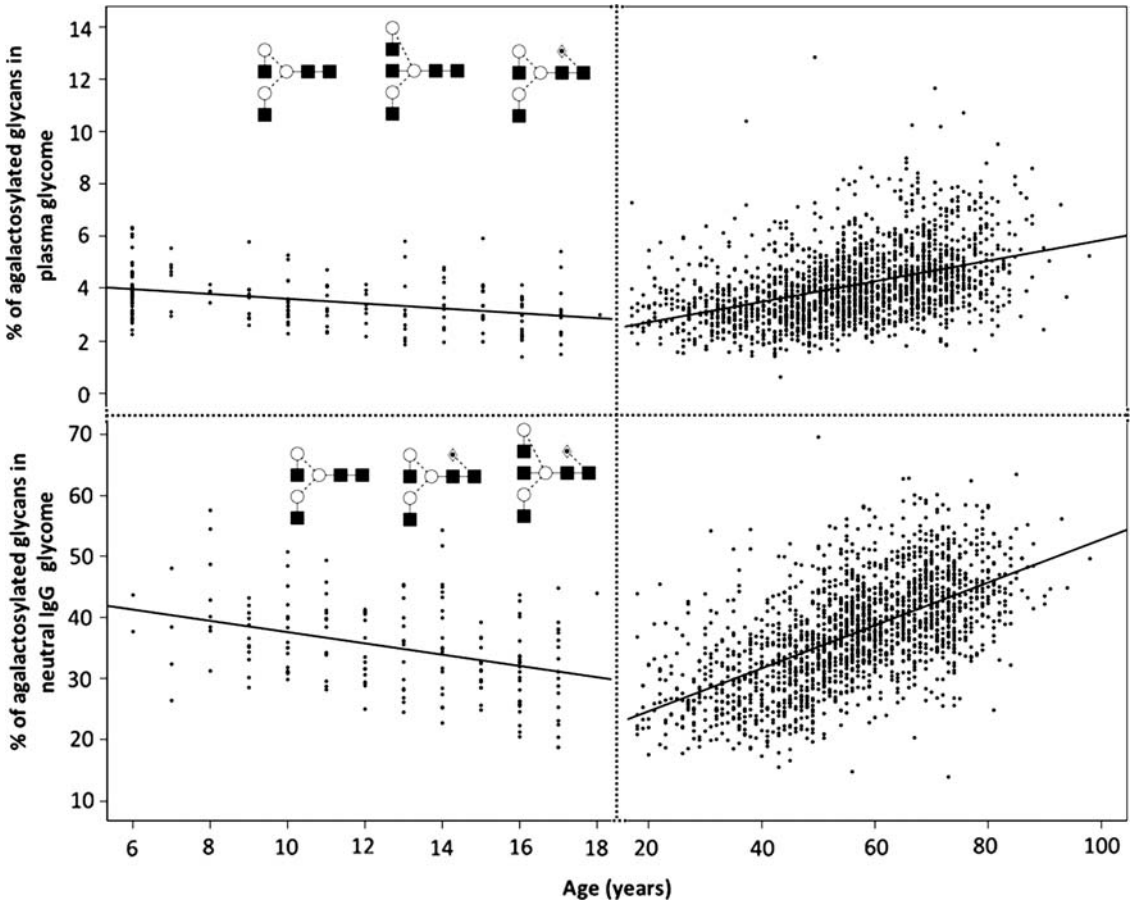


Fig. 3. Levels of agalactosylated glycans in plasma and IgG of children and adults. Changes in agalactosylated plasma and IgG glycans in adults were reported previously (Knežević et al. 2010; Pučić et al. 2011) and are here shown only for a comparison. Glycan scheme: square, *N*-acetylglucosamine; circle, mannose; romb with a dot, fucose.

population was then tested for normality using the Smirnov–Kolmogrov test and non-parametric statistical tests were further used. The Mann–Whitney test was used to analyze differences in levels of glycans between groups, whereas the Spearman's rank correlation was calculated for the association of glycan structures and age. The significance level was set at $P \leq 0.001$ to account for multiple testing. All reported *P*-values are two-tailed if not stated otherwise. Statistical analyses were performed with SPSS 13 (SPSS, Chicago, IL).

Supplementary data

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

Funding

The work is supported by grants #309–0061194-2023 (to G.L.) and #098-0982522-2455 (to N.P.) from the Croatian Ministry of Science, Education and Sport; and by European Commission GlycoBioM (contract #259869) and HighGlycan (contract #278535) grants.

Conflict of interest

None declared.

Abbreviations

G0, agalactosylated glycans; G1, monogalactosylated glycans; G2, digalactosylated glycans; G3, trigalactosylated glycan; G4, tetragalactosylated glycans; GlcNAc, *N*-acetylglucosamine; HILIC, hydrophilic interaction chromatography; HPLC, high-performance liquid chromatography; WAX, weak anion exchange.

References

Arnold JN, Saldova R, Hamid UM, Rudd PM. 2008. Evaluation of the serum *N*-linked glycome for the diagnosis of cancer and chronic inflammation. *Proteomics*. 8:3284–3293.

Brinkman-Van der Linden CM, Havenaar EC, Van Ommen CR, Van Kamp GJ, Gooren LJ, Van Dijk W. 1996. Oral estrogen treatment induces a decrease in expression of sialyl Lewis x on alpha 1-acid glycoprotein in females and male-to-female transsexuals. *Glycobiology*. 6:407–412.

Freeze HH. 2006. Genetic defects in the human glycome. *Nat Rev Genet*. 7:537–551.

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.

Jaeken J. 2010. Congenital disorders of glycosylation. *Ann N Y Acad Sci*. 1214:190–198.

Klein A. 2008. Human total serum *N*-glycome. *Adv Clin Chem*. 46:51–85.

Knežević A, Gornik O, Polašek O, Pučić M, Novokmet M, Redžić I, Rudd PM, Wright AF, Campbell H, Rudan I, et al. 2010. Effects of aging, body mass index, plasma lipid profiles, and smoking on human plasma *N*-glycans. *Glycobiology*. 20:959–969.

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.

Lebrilla CB, An HJ. 2009. The prospects of glycan biomarkers for the diagnosis of diseases. *Mol Biosyst*. 5:17–20.

Pučić M, Knežević A, Vidic J, Adamczyk B, Novokmet M, Polasek O, Gornik O, Supraha-Goreta S, Wormald MR, Redzic I, et al. 2011. High throughput isolation and glycosylation analysis of IgG-variability and heritability of the IgG glycome in three isolated human populations. *Mol Cell Proteomics*, 10:M111.010090.

Pučić M, Pinto S, Novokmet M, Knežević A, Gornik O, Polašek O, Vlahoviček K, Wei W, Rudd PM, Wright AF, et al. 2010. Common aberrations from normal human *N*-glycan plasma profile. *Glycobiology*. 20:970–975.

Reis CA, Osorio H, Silva L, Gomes C, David L. 2010. Alterations in glycosylation as biomarkers for cancer detection. *J Clin Pathol*. 63:322–329.

Royle L, Campbell MP, Radcliffe CM, White DM, Harvey DJ, Abrahams JL, Kim YG, Henry GW, Shadick NA, Weinblatt ME, et al. 2008. HPLC-based analysis of serum *N*-glycans on a 96-well plate platform with dedicated database software. *Anal Biochem*. 376:1–12.

Ruhaak LR, Uh HW, Beekman M, Koeleman CA, Hokke CH, Westendorp RG, Wuhrer M, Houwing-Duistermaat JJ, Slagboom PE, Deelder AM. 2010. Decreased levels of bisecting GlcNAc glycoforms of IgG are associated with human longevity. *PLoS One*. 5:e12566.

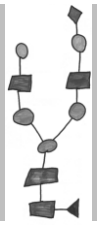
Saldova R, Huffman JE, Adamczyk B, Muzinic A, Kattla JJ, Pucic M, Novokmet M, Abrahams JL, Hayward C, Rudan I, et al. 2012. Association of medication and the human plasma *N*-glycans. *J Proteome Res*. 11:1821–1831.

van de Geijn FE, Wuhrer M, Selman MH, Willemsen SP, de Man YA, Deelder AM, Hazes JM, Dolhain RJ. 2009. Immunoglobulin G galactosylation and sialylation are associated with pregnancy-induced improvement of rheumatoid arthritis and the postpartum flare: Results from a large prospective cohort study. *Arthritis Res Ther*. 11:R193.

Vanhooren V, Desmyter L, Liu XE, Cardelli M, Franceschi C, Federico A, Libert C, Laroy W, Dewaele S, Contreras R, et al. 2007. *N*-Glycomic changes in serum proteins during human aging. *Rejuvenation Res*. 10:521–531a.

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

6 GENERAL DISCUSSION



IgG is one of the most studied glycoproteins. The attached glycans are essential for antibody's stability and efficacy, and even minor changes in glycan composition can have a profound influence on the biological activity of IgG by modulating binding to Fc receptors.[12, 13] IgG glycosylation has been shown to be dependent on various physiological parameters such as age, sex and pregnancy.[52-59] Moreover, it has long been known that IgG glycosylation patterns are skewed toward specific glycoforms in various diseases (tumors, infections, autoimmune diseases, etc.).[3] As well as in healthy individuals, galactosylation levels are a major source of IgG heterogeneity in diseased individuals.

Due to the undisputed importance of glycosylation for the function of IgGs, there was a need for a large-scale study which would identify the variability and heritability of IgG glycosylation in human population. Large-scale glycomic studies require the application of high-throughput methods that allow sensitive, robust and reliable glycoprofiling. Because of the structural complexity of glycans and technological limitations, until only a few years ago glycan analysis was extremely challenging and laborious. However, major progress has been made in the last few years resulting in several high-throughput analytical techniques for glycan analysis, such as high/ultra performance liquid chromatography, mass spectrometry and capillary electrophoresis.[56, 86-90] Hitherto several large population studies reported on high variability of the plasma glycome between individuals.[55, 135, 136] However, only a small fraction of the variability in plasma glycan levels could be explained by age and other measured environmental factors leaving the main source of glycome variation between individuals unknown.

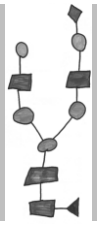
One of the major bottlenecks in large-scale proteomics and glycomics studies is protein purification from a large number of samples. In this thesis a development of a 96-well Protein G monolithic plate and its application for rapid isolation of IgG is described. Monoliths are continuous stationary phase with very large and highly interconnected pores which makes their dynamic binding capacity practically independent of the flow rate and enables fast purification of large volumes of complex biological compounds such as plasma.[97, 98] Optimized HILIC method on a novel Waters BEH Glycan chromatography column provided separation of IgG *N*-glycans into 24 chromatographic peaks and the identity of glycan structures in each peak was determined by exoglycosidase digestions and mass spectrometry. This analysis for the first time provided insight into the variability of IgG glycome within the human population, and for the first time estimated heritability of IgG glycosylation. In comparison to the total plasma glycome, the glycome of IgG, a single plasma protein, varied even more between individuals.[137] The average ratio between minimal and maximal values for all glycans in the IgG glycome was 17.2, what is nearly three times higher than the corresponding ratio in the plasma glycome.[136, 137] This difference suggests that the presence of a wide variety of glycoproteins, as in plasma, actually decreases the variability of glycosylation. By analyzing total plasma glycome, glycans are averaged across the proteome which introduces considerable noise to the quantitation and interpretation of plasma glycan levels.

In addition to adult cohort, we have also analyzed both IgG and total plasma glycomes of children and adolescents revealing again a high variability of glycosylation features. Interestingly, the behavior of almost all glycan features had different trend of change with age in childhood and in adulthood. Moreover, sex differences were shown to be much less pronounced in children than in adults and present mainly during puberty.

HILIC enabled a detailed analysis of a mixture of released Fc and Fab *N*-glycans of all four IgG subclasses. In order to obtain subclass- and site-specific IgG glycosylation profiling of the same adult cohort, we have analyzed IgG Fc *N*-glycopeptides by MALDI-TOF-MS. Fc *N*-glycosylation profiling confirmed most of the previously described age- and sex-related IgG glycosylation changes.[52-56, 139] Interestingly, the most prominent changes in glycosylation in females were observed around the age of 45 to 60 years when females usually enter the menopause. When we compared the results to the total IgG *N*-glycosylation of the same populations analyzed by HILIC, significant differences were observed in the levels of galactosylation, bisecting GlcNAc and particularly sialylation, which were shown to be higher in HILIC analysis. Age and sex associations of glycosylation features were, to a large extent, comparable between MALDI-TOF-MS and HILIC IgG glycosylation profiling. There are several important differences between HILIC and MALDI-TOF-MS IgG glycosylation profiling methods. Starting from the sample preparation, HILIC of fluorescently labeled glycans requires enzymatic glycan release, labeling and sample clean-up, while the glycopeptide profiling with MALDI-TOF-MS involves trypsin treatment and SPE. In general, sample preparation for MALDI-TOF-MS requires less time, labor and steps which can introduce variation during sample preparation. Secondly, these two glycosylation profiling methods differ in sample throughput. Even though HILIC has been notably speeded up by UPLC technology allowing the analysis of a couple of samples per hour, the speed of MALDI-TOF-MS is still much faster. Moreover, high sample purity is essential for HILIC of released glycans since glycans released from glycoprotein contaminants will interfere with the IgG glycosylation profile. On the other hand, the presence of low amounts of contaminating glycoproteins will, in most cases, not interfere with IgG Fc glycosylation profiling by MALDI-TOF-MS since this method allows distinction between IgG Fc glycopeptides and glycopeptides of other glycoproteins on the basis of the different masses of the peptide moieties. The most prominent advantage of MALDI-TOF-MS profiling is the assignment of glycans to the specific Fc glycosylation sites of IgG subclasses. IgG Fc glycans have very distinct functions as they modulate the interaction with Fc receptors[7, 22-24, 37] and other cell-surface receptors of immune cells.[46, 47] Since most of the modulating effects of IgG Fc glycans have been reported for IgG1[25, 37, 46] and may not apply to IgG2, subclass- and site-specific IgG glycosylation profiling as achieved by mass spectrometry of glycopeptides might be crucial for understanding functional implications of glycosylation features.

By combining high-throughput glycomics measurements with high-throughput genomics we performed a genome wide association study of the human IgG *N*-glycome. The first GWA study of the human *N*-glycome was done on the overall plasma *N*-glycans and it identified genome-wide associations with two glycosyltransferases and one transcription factor (HNF1α) with 1-6% of variance explained.[132] Plasma glycans originate from many different glycoproteins produced in different cell types with cell type-specific glycosylation which probably blurred regulation of glycosylation of individual glycoproteins and reduced the power of that study. In this study we excluded the influence of differential glycosylation of different plasma proteins by isolating a single plasma protein, i.e. IgG, produced by a single cell type (B lymphocytes). GWA studies of plasma and IgG glycome were of a comparable sample size, however, many more significant associations were revealed with the IgG glycans. In addition, the largest percentage of variance explained by a single association was substantially higher, between 16-18%. Among the nine loci that reached the strict genome-wide statistical significance, four involved genes encoding glycosyltransferases known to participate in IgG glycosylation (ST6GALI, B4GALT1, FUT8, MGAT3) by adding either sialic acid, galactose, fucose or bisecting GlcNAc. Thus these observation were biologically founded and served as a proof of principle that a single protein glycosylation GWAS approach can identify biologically important glycan pathways. The remaining five genetic loci (IKZF1, IL6ST-ANKRD55, ABCF2-SMARCD3, SUV420H1, and SMARCB1-DERL3) were not previously associated with protein glycosylation. Interestingly, these genes that are not directly involved in glycosylation showed the most significant associations with glycosylation processes such as the addition of galactose, fucose or bisecting GlcNAc. These findings suggest that the composition of IgG *N*-glycome is a result of a complex network of genes directly involved in glycosylation and those that apparently have some higher-level regulatory function. Nearly all genome-wide significant loci in our study have already been associated with autoimmune diseases and haematological cancers. Given that one of our identified genetic loci (IKZF1) was associated with systemic lupus erythematosus (SLE) in previous studies, we have investigated biomarker potential of a specific IgG *N*-glycan trait in predicting SLE with considerable success. This study demonstrated the possibility to identify genetic loci that control glycosylation of a single plasma protein and to generate guidelines for follow-up studies which could bring large advances in understanding the genetic regulation of IgG *N*-glycan synthesis.

7 CONCLUSIONS



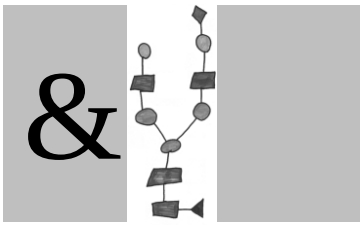
This thesis presents a development of a novel high-throughput approach for IgG purification and IgG glycan analysis and its application in understanding variability and heritability of IgG glycosylation in a human population.

A 96-well Protein G monolithic plate was successfully applied for the high-thoughput isolation of IgG from over 2000 plasma samples. HILIC method on a Waters BEH Glycan chromatography column was optimized to provide separation of IgG *N*-glycans into 24 chromatographic peaks and identity of glycan structures in each peak was determined by exoglycosidase digestions and mass spectrometry. Very high variability of IgG glycans between individuals was observed, while heritability was generally between 30 and 50%. The individual's age was associated with a significant decrease in galactose and increase of bisecting GlcNAc, whereas other functional elements of IgG glycosylation did not change much with age.

IgG Fc *N*-glycosylation profiling by MALDI-TOF-MS of glycopeptides of the same cohort described age- and sex-related IgG glycosylation changes. Levels of galactosylation and sialylation were found to decrease with increasing age and showed significant sex dependence. The most prominent changes in glycosylation in females were observed around the age of 45 to 60 years. The incidence of bisecting GlcNAc increased in younger individuals and reached a plateau at older age.

GWAS of IgG glycosylation traits identified nine genetic loci that control IgG glycosylation. Of these, four loci contained genes encoding glycosyltransferases (ST6GAL1, B4GALT1, FUT8 and MGAT3), while the remaining five genetic loci (IKZF1, IL6ST-ANKRD55, ABCF2-SMARCD3, SUV420H1, and SMARCB1-DERL3) were not previously implicated in protein glycosylation. Biomarker potential of *N*-glycans in systemic lupus erythematosus was explored and demonstrated substantial discriminative power.

Changes in plasma and IgG *N*-glycome during childhood and adolescence were described. High variability of glycosylation features and a large number of age-dependent glycans were observed. The plasma *N*-glycome of younger children was found to contain a higher proportion of large complex glycan structures which decreased with age. The behavior of almost all glycan features had different trend or the rate of change in childhood than in adulthood. Sex differences are much less pronounced in children than in adults and are present mainly during puberty.



REFERENCE LIST

1. Haltiwanger, R. S.; Lowe, J. B., Role of glycosylation in development. *Annu Rev Biochem* 2004, 73, 491-537.
2. Walt D, A.-K. K., Bendiak B, Bertozzi CR, Boons GJ, Darvill A, Hart G, Kiessling LL, Lowe J, Moon R, et al., Transforming Glycoscience: A Roadmap for the Future. *Washington, DC: The National Academies Press* 2012.
3. Arnold, J. N.; Wormald, M. R.; Sim, R. B.; Rudd, P. M.; Dwek, R. A., The impact of glycosylation on the biological function and structure of human immunoglobulins. *Annu Rev Immunol* 2007, 25, 21-50.
4. Nezlin, R.; Ghetie, V., Interactions of immunoglobulins outside the antigen-combining site. *Adv Immunol* 2004, 82, 155-215.
5. Schroeder, H. W., Jr.; Cavacini, L., Structure and function of immunoglobulins. *J Allergy Clin Immunol* 2010, 125, (2 Suppl 2), S41-52.
6. Jefferis, R., Isotype and glycoform selection for antibody therapeutics. *Arch Biochem Biophys* 2012, 526, (2), 159-66.
7. Huhn, C.; Selman, M. H.; Ruhaak, L. R.; Deelder, A. M.; Wuhrer, M., IgG glycosylation analysis. *Proteomics* 2009, 9, (4), 882-913.
8. Jefferis, R., Glycosylation as a strategy to improve antibody-based therapeutics. *Nat Rev Drug Discov* 2009, 8, (3), 226-34.
9. Deisenhofer, J., Crystallographic refinement and atomic models of a human Fc fragment and its complex with fragment B of protein A from Staphylococcus aureus at 2.9- and 2.8-A resolution. *Biochemistry* 1981, 20, (9), 2361-70.
10. Krapp, S.; Mimura, Y.; Jefferis, R.; Huber, R.; Sondermann, P., Structural analysis of human IgG-Fc glycoforms reveals a correlation between glycosylation and structural integrity. *J Mol Biol* 2003, 325, (5), 979-89.
11. Mimura, Y.; Church, S.; Ghirlando, R.; Ashton, P. R.; Dong, S.; Goodall, M.; Lund, J.; Jefferis, R., The influence of glycosylation on the thermal stability and effector function expression of human IgG1-Fc: properties of a series of truncated glycoforms. *Mol Immunol* 2000, 37, (12-13), 697-706.
12. Mimura, Y.; Sondermann, P.; Ghirlando, R.; Lund, J.; Young, S. P.; Goodall, M.; Jefferis, R., Role of oligosaccharide residues of IgG1-Fc in Fc gamma RIIb binding. *J Biol Chem* 2001, 276, (49), 45539-47.
13. Jefferis, R., Glycosylation of recombinant antibody therapeutics. *Biotechnol Prog* 2005, 21, (1), 11-6.
14. Holland, M.; Yagi, H.; Takahashi, N.; Kato, K.; Savage, C. O.; Goodall, D. M.; Jefferis, R., Differential glycosylation of polyclonal IgG, IgG-Fc and IgG-Fab isolated from the sera of patients with ANCA-associated systemic vasculitis. *Biochim Biophys Acta* 2006, 1760, (4), 669-77.
15. Stadlmann, J.; Pabst, M.; Kolarich, D.; Kunert, R.; Altmann, F., Analysis of immunoglobulin glycosylation by LC-ESI-MS of glycopeptides and oligosaccharides. *Proteomics* 2008, 8, (14), 2858-71.
16. Wormald, M. R.; Rudd, P. M.; Harvey, D. J.; Chang, S. C.; Scragg, I. G.; Dwek, R. A., Variations in oligosaccharide-protein interactions in immunoglobulin G determine the site-specific glycosylation profiles and modulate the dynamic motion of the Fc oligosaccharides. *Biochemistry* 1997, 36, (6), 1370-80.
17. Youings, A.; Chang, S. C.; Dwek, R. A.; Scragg, I. G., Site-specific glycosylation of human immunoglobulin G is altered in four rheumatoid arthritis patients. *Biochem J* 1996, 314 (Pt 2), 621-30.
18. Ceroni, A.; Dell, A.; Haslam, S. M., The GlycanBuilder: a fast, intuitive and flexible software tool for building and

isplaying glycan structures. *Source Code Biol Med* 2007, 2, 3.

19. Nimmerjahn, F.; Ravetch, J. V., Divergent immunoglobulin g subclass activity through selective Fc receptor binding. *Science* 2005, 310, (5753), 1510-2.
20. Sondermann, P.; Kaiser, J.; Jacob, U., Molecular basis for immune complex recognition: a comparison of Fc-receptor structures. *J Mol Biol* 2001, 309, (3), 737-49.
21. Matsumiya, S.; Yamaguchi, Y.; Saito, J.; Nagano, M.; Sasakawa, H.; Otaki, S.; Satoh, M.; Shitara, K.; Kato, K., Structural comparison of fucosylated and nonfucosylated Fc fragments of human immunoglobulin G1. *J Mol Biol* 2007, 368, (3), 767-79.
22. Shields, R. L.; Lai, J.; Keck, R.; O'Connell, L. Y.; Hong, K.; Meng, Y. G.; Weikert, S. H.; Presta, L. G., Lack of fucose on human IgG1 N-linked oligosaccharide improves binding to human Fcgamma RIII and antibody-dependent cellular toxicity. *J Biol Chem* 2002, 277, (30), 26733-40.
23. Ferrara, C.; Grau, S.; Jager, C.; Sondermann, P.; Brunker, P.; Waldhauer, I.; Hennig, M.; Ruf, A.; Rufer, A. C.; Stihle, M.; Umana, P.; Benz, J., Unique carbohydrate-carbohydrate interactions are required for high affinity binding between FcgammaRIII and antibodies lacking core fucose. *Proc Natl Acad Sci U S A* 2011, 108, (31), 12669-74.
24. Iida, S.; Misaka, H.; Inoue, M.; Shibata, M.; Nakano, R.; Yamane-Ohnuki, N.; Wakitani, M.; Yano, K.; Shitara, K.; Satoh, M., Nonfucosylated therapeutic IgG1 antibody can evade the inhibitory effect of serum immunoglobulin G on antibody-dependent cellular cytotoxicity through its high binding to FcgammaRIIIa. *Clin Cancer Res* 2006, 12, (9), 2879-87.
25. Davies, J.; Jiang, L.; Pan, L. Z.; LaBarre, M. J.; Anderson, D.; Reff, M., Expression of GnTIII in a recombinant anti-CD20 CHO production cell line: Expression of antibodies with altered glycoforms leads to an increase in ADCC through higher affinity for FC gamma RIII. *Biotechnol Bioeng* 2001, 74, (4), 288-94.
26. Umana, P.; Jean-Mairet, J.; Moudry, R.; Amstutz, H.; Bailey, J. E., Engineered glycoforms of an antineuroblastoma IgG1 with optimized antibody-dependent cellular cytotoxic activity. *Nat Biotechnol* 1999, 17, (2), 176-80.
27. Ferrara, C.; Brunker, P.; Suter, T.; Moser, S.; Puntener, U.; Umana, P., Modulation of therapeutic antibody effector functions by glycosylation engineering: influence of Golgi enzyme localization domain and co-expression of heterologous beta1, 4-N-acetylglucosaminyltransferase III and Golgi alpha-mannosidase II. *Biotechnol Bioeng* 2006, 93, (5), 851-61.
28. Shinkawa, T.; Nakamura, K.; Yamane, N.; Shoji-Hosaka, E.; Kanda, Y.; Sakurada, M.; Uchida, K.; Anazawa, H.; Satoh, M.; Yamasaki, M.; Hanai, N.; Shitara, K., The absence of fucose but not the presence of galactose or bisecting N-acetylglucosamine of human IgG1 complex-type oligosaccharides shows the critical role of enhancing antibody-dependent cellular cytotoxicity. *J Biol Chem* 2003, 278, (5), 3466-73.
29. Mori, K.; Iida, S.; Yamane-Ohnuki, N.; Kanda, Y.; Kuni-Kamochi, R.; Nakano, R.; Imai-Nishiya, H.; Okazaki, A.; Shinkawa, T.; Natsume, A.; Niwa, R.; Shitara, K.; Satoh, M., Non-fucosylated therapeutic antibodies: the next generation of therapeutic antibodies. *Cytotechnology* 2007, 55, (2-3), 109-14.
30. Preithner, S.; Elm, S.; Lippold, S.; Locher, M.; Wolf, A.; da Silva, A. J.; Baeuerle, P. A.; Prang, N. S., High concentrations of therapeutic IgG1 antibodies are needed to compensate for inhibition of antibody-dependent cellular cytotoxicity by excess endogenous immunoglobulin G. *Mol Immunol* 2006, 43, (8), 1183-93.
31. Boyd, P. N.; Lines, A. C.; Patel, A. K., The effect of the removal of sialic acid,

galactose and total carbohydrate on the functional activity of Campath-1H. *Mol Immunol* 1995, 32, (17-18), 1311-8.

32. Hodoniczky, J.; Zheng, Y. Z.; James, D. C., Control of recombinant monoclonal antibody effector functions by Fc N-glycan remodeling in vitro. *Biotechnol Prog* 2005, 21, (6), 1644-52.
33. Malhotra, R.; Wormald, M. R.; Rudd, P. M.; Fischer, P. B.; Dwek, R. A.; Sim, R. B., Glycosylation changes of IgG associated with rheumatoid arthritis can activate complement via the mannose-binding protein. *Nat Med* 1995, 1, (3), 237-43.
34. Arnold, J. N.; Dwek, R. A.; Rudd, P. M.; Sim, R. B., Mannan binding lectin and its interaction with immunoglobulins in health and in disease. *Immunol Lett* 2006, 106, (2), 103-10.
35. Sato, R.; Matsushita, M.; Miyata, M.; Sato, Y.; Kasukawa, R.; Fujita, T., Substances reactive with mannose-binding protein (MBP) in sera of patients with rheumatoid arthritis. *Fukushima J Med Sci* 1997, 43, (2), 99-111.
36. Nimmerjahn, F.; Anthony, R. M.; Ravetch, J. V., Agalactosylated IgG antibodies depend on cellular Fc receptors for in vivo activity. *Proc Natl Acad Sci U S A* 2007, 104, (20), 8433-7.
37. Karsten, C. M.; Pandey, M. K.; Figge, J.; Kilchenstein, R.; Taylor, P. R.; Rosas, M.; McDonald, J. U.; Orr, S. J.; Berger, M.; Petzold, D.; Blanchard, V.; Winkler, A.; Hess, C.; Reid, D. M.; Majoul, I. V.; Strait, R. T.; Harris, N. L.; Kohl, G.; Wex, E.; Ludwig, R.; Zillikens, D.; Nimmerjahn, F.; Finkelman, F. D.; Brown, G. D.; Ehlers, M.; Kohl, J., Anti-inflammatory activity of IgG1 mediated by Fc galactosylation and association of FcgammaRIIB and dectin-1. *Nat Med* 2012, 18, (9), 1401-6.
38. Katz, U.; Achiron, A.; Sherer, Y.; Shoenfeld, Y., Safety of intravenous immunoglobulin (IVIG) therapy. *Autoimmun Rev* 2007, 6, (4), 257-9.
39. Kuitwaard, K.; de Gelder, J.; Tio-Gillen, A. P.; Hop, W. C.; van Gelder, T.; van Toorenenbergen, A. W.; van Doorn, P. A.; Jacobs, B. C., Pharmacokinetics of intravenous immunoglobulin and outcome in Guillain-Barre syndrome. *Ann Neurol* 2009, 66, (5), 597-603.
40. Nimmerjahn, F.; Ravetch, J. V., The antiinflammatory activity of IgG: the intravenous IgG paradox. *J Exp Med* 2007, 204, (1), 11-5.
41. Debre, M.; Bonnet, M. C.; Fridman, W. H.; Carosella, E.; Philippe, N.; Reinert, P.; Vilmer, E.; Kaplan, C.; Teillaud, J. L.; Griscelli, C., Infusion of Fc gamma fragments for treatment of children with acute immune thrombocytopenic purpura. *Lancet* 1993, 342, (8877), 945-9.
42. Kaneko, Y.; Nimmerjahn, F.; Madaio, M. P.; Ravetch, J. V., Pathology and protection in nephrotoxic nephritis is determined by selective engagement of specific Fc receptors. *J Exp Med* 2006, 203, (3), 789-97.
43. Samuelsson, A.; Towers, T. L.; Ravetch, J. V., Anti-inflammatory activity of IVIG mediated through the inhibitory Fc receptor. *Science* 2001, 291, (5503), 484-6.
44. Anthony, R. M.; Nimmerjahn, F.; Ashline, D. J.; Reinhold, V. N.; Paulson, J. C.; Ravetch, J. V., Recapitulation of IVIG anti-inflammatory activity with a recombinant IgG Fc. *Science* 2008, 320, (5874), 373-6.
45. Kaneko, Y.; Nimmerjahn, F.; Ravetch, J. V., Anti-inflammatory activity of immunoglobulin G resulting from Fc sialylation. *Science* 2006, 313, (5787), 670-3.
46. Anthony, R. M.; Wermeling, F.; Karlsson, M. C.; Ravetch, J. V., Identification of a receptor required for the anti-inflammatory activity of IVIG. *Proc Natl Acad Sci U S A* 2008, 105, (50), 19571-8.
47. Anthony, R. M.; Kobayashi, T.; Wermeling, F.; Ravetch, J. V., Intravenous gammaglobulin suppresses inflammation through a novel T(H)2

pathway. *Nature* 2011, 475, (7354), 110-3.

48. Bayry, J.; Bansal, K.; Kazatchkine, M. D.; Kaveri, S. V., DC-SIGN and alpha2,6-sialylated IgG Fc interaction is dispensable for the anti-inflammatory activity of IVIg on human dendritic cells. *Proc Natl Acad Sci U S A* 2009, 106, (9), E24; author reply E25.
49. Stadlmann, J.; Weber, A.; Pabst, M.; Anderle, H.; Kunert, R.; Ehrlich, H. J.; Peter Schwarz, H.; Altmann, F., A close look at human IgG sialylation and subclass distribution after lectin fractionation. *Proteomics* 2009, 9, (17), 4143-53.
50. Guhr, T.; Bloem, J.; Derksen, N. I.; Wuhrer, M.; Koenderman, A. H.; Aalberse, R. C.; Rispens, T., Enrichment of sialylated IgG by lectin fractionation does not enhance the efficacy of immunoglobulin G in a murine model of immune thrombocytopenia. *PLoS One* 2011, 6, (6), e21246.
51. Kasermann, F.; Boerema, D. J.; Ruegsegger, M.; Hofmann, A.; Wymann, S.; Zuercher, A. W.; Miescher, S., Analysis and functional consequences of increased Fab-sialylation of intravenous immunoglobulin (IVIG) after lectin fractionation. *PLoS One* 2012, 7, (6), e37243.
52. Parekh, R.; Roitt, I.; Isenberg, D.; Dwek, R.; Rademacher, T., Age-related galactosylation of the N-linked oligosaccharides of human serum IgG. *J Exp Med* 1988, 167, (5), 1731-6.
53. Yamada, E.; Tsukamoto, Y.; Sasaki, R.; Yagyu, K.; Takahashi, N., Structural changes of immunoglobulin G oligosaccharides with age in healthy human serum. *Glycoconj J* 1997, 14, (3), 401-5.
54. Shikata, K.; Yasuda, T.; Takeuchi, F.; Konishi, T.; Nakata, M.; Mizuochi, T., Structural changes in the oligosaccharide moiety of human IgG with aging. *Glycoconj J* 1998, 15, (7), 683-9.
55. Ruhaak, L. R.; Uh, H. W.; Beekman, M.; Koeleman, C. A.; Hokke, C. H.; Westendorp, R. G.; Wuhrer, M.; Houwing-Duistermaat, J. J.; Slagboom, P. E.; Deelder, A. M., Decreased levels of bisecting GlcNAc glycoforms of IgG are associated with human longevity. *PLoS One* 2010, 5, (9), e12566.
56. Selman, M. H.; McDonnell, L. A.; Palmblad, M.; Ruhaak, L. R.; Deelder, A. M.; Wuhrer, M., Immunoglobulin G glycopeptide profiling by matrix-assisted laser desorption ionization Fourier transform ion cyclotron resonance mass spectrometry. *Anal Chem* 2010, 82, (3), 1073-81.
57. Rook, G. A.; Steele, J.; Brealey, R.; Whyte, A.; Isenberg, D.; Sumar, N.; Nelson, J. L.; Bodman, K. B.; Young, A.; Roitt, I. M.; et al., Changes in IgG glycoform levels are associated with remission of arthritis during pregnancy. *J Autoimmun* 1991, 4, (5), 779-94.
58. Selman, M. H.; Derks, R. J.; Bondt, A.; Palmblad, M.; Schoenmaker, B.; Koeleman, C. A.; van de Geijn, F. E.; Dolhain, R. J.; Deelder, A. M.; Wuhrer, M., Fc specific IgG glycosylation profiling by robust nano-reverse phase HPLC-MS using a sheath-flow ESI sprayer interface. *J Proteomics* 2011.
59. van de Geijn, F. E.; Wuhrer, M.; Selman, M. H.; Willemsen, S. P.; de Man, Y. A.; Deelder, A. M.; Hazes, J. M.; Dolhain, R. J., Immunoglobulin G galactosylation and sialylation are associated with pregnancy-induced improvement of rheumatoid arthritis and the postpartum flare: results from a large prospective cohort study. *Arthritis Res Ther* 2009, 11, (6), R193.
60. Williams, P. J.; Arkwright, P. D.; Rudd, P.; Scragg, I. G.; Edge, C. J.; Wormald, M. R.; Rademacher, T. W., Short communication: selective placental transport of maternal IgG to the fetus. *Placenta* 1995, 16, (8), 749-56.
61. Einarsdottir, H. K.; Selman, M. H.; Kapur, R.; Scherjon, S.; Koeleman, C. A.; Deelder, A. M.; van der Schoot, C. E.; Vidarsson, G.; Wuhrer, M.,

Comparison of the Fc glycosylation of fetal and maternal immunoglobulin G. *Glycoconj J* 2012, 30, (2), 147-57.

62. Parekh, R. B.; Dwek, R. A.; Sutton, B. J.; Fernandes, D. L.; Leung, A.; Stanworth, D.; Rademacher, T. W.; Mizuochi, T.; Taniguchi, T.; Matsuta, K.; et al., Association of rheumatoid arthritis and primary osteoarthritis with changes in the glycosylation pattern of total serum IgG. *Nature* 1985, 316, (6027), 452-7.
63. Alavi, A.; Arden, N.; Spector, T. D.; Axford, J. S., Immunoglobulin G glycosylation and clinical outcome in rheumatoid arthritis during pregnancy. *J Rheumatol* 2000, 27, (6), 1379-85.
64. Bond, A.; Alavi, A.; Axford, J. S.; Bourke, B. E.; Bruckner, F. E.; Kerr, M. A.; Maxwell, J. D.; Tweed, K. J.; Weldon, M. J.; Youinou, P.; Hay, F. C., A detailed lectin analysis of IgG glycosylation, demonstrating disease specific changes in terminal galactose and N-acetylglucosamine. *J Autoimmun* 1997, 10, (1), 77-85.
65. Bond, A.; Alavi, A.; Axford, J. S.; Youinou, P.; Hay, F. C., The relationship between exposed galactose and N-acetylglucosamine residues on IgG in rheumatoid arthritis (RA), juvenile chronic arthritis (JCA) and Sjogren's syndrome (SS). *Clin Exp Immunol* 1996, 105, (1), 99-103.
66. Ercan, A.; Cui, J.; Chatterton, D. E.; Deane, K. D.; Hazen, M. M.; Brintnell, W.; O'Donnell, C. I.; Derber, L. A.; Weinblatt, M. E.; Shadick, N. A.; Bell, D. A.; Cairns, E.; Solomon, D. H.; Holers, V. M.; Rudd, P. M.; Lee, D. M., Aberrant IgG galactosylation precedes disease onset, correlates with disease activity, and is prevalent in autoantibodies in rheumatoid arthritis. *Arthritis Rheum* 2010, 62, (8), 2239-48.
67. Gornik, I.; Maravic, G.; Dumic, J.; Flogel, M.; Lauc, G., Fucosylation of IgG heavy chains is increased in rheumatoid arthritis. *Clin Biochem* 1999, 32, (8), 605-8.
68. Parekh, R. B.; Roitt, I. M.; Isenberg, D. A.; Dwek, R. A.; Ansell, B. M.; Rademacher, T. W., Galactosylation of IgG associated oligosaccharides: reduction in patients with adult and juvenile onset rheumatoid arthritis and relation to disease activity. *Lancet* 1988, 1, (8592), 966-9.
69. van Zeben, D.; Rook, G. A.; Hazes, J. M.; Zwinderman, A. H.; Zhang, Y.; Ghelani, S.; Rademacher, T. W.; Breedveld, F. C., Early agalactosylation of IgG is associated with a more progressive disease course in patients with rheumatoid arthritis: results of a follow-up study. *Br J Rheumatol* 1994, 33, (1), 36-43.
70. Flogel, M.; Lauc, G.; Gornik, I.; Macek, B., Fucosylation and galactosylation of IgG heavy chains differ between acute and remission phases of juvenile chronic arthritis. *Clin Chem Lab Med* 1998, 36, (2), 99-102.
71. Dube, R.; Rook, G. A.; Steele, J.; Brealey, R.; Dwek, R.; Rademacher, T.; Lennard-Jones, J., Agalactosyl IgG in inflammatory bowel disease: correlation with C-reactive protein. *Gut* 1990, 31, (4), 431-4.
72. Parekh, R.; Isenberg, D.; Rook, G.; Roitt, I.; Dwek, R.; Rademacher, T., A comparative analysis of disease-associated changes in the galactosylation of serum IgG. *J Autoimmun* 1989, 2, (2), 101-14.
73. Perdivara, I.; Peddada, S. D.; Miller, F. W.; Tomer, K. B.; Deterding, L. J., Mass spectrometric determination of IgG subclass-specific glycosylation profiles in siblings discordant for myositis syndromes. *J Proteome Res* 2011, 10, (7), 2969-78.
74. Mehta, A. S.; Long, R. E.; Comunale, M. A.; Wang, M.; Rodemich, L.; Krakover, J.; Philip, R.; Marrero, J. A.; Dwek, R. A.; Block, T. M., Increased levels of galactose-deficient anti-Gal immunoglobulin G in the sera of hepatitis C virus-infected individuals with fibrosis and cirrhosis. *J Virol* 2008, 82, (3), 1259-70.

75. Moore, J. S.; Wu, X.; Kulhavy, R.; Tomana, M.; Novak, J.; Moldoveanu, Z.; Brown, R.; Goepfert, P. A.; Mestecky, J., Increased levels of galactose-deficient IgG in sera of HIV-1-infected individuals. *Aids* 2005, 19, (4), 381-9.

76. Saldova, R.; Royle, L.; Radcliffe, C. M.; Abd Hamid, U. M.; Evans, R.; Arnold, J. N.; Banks, R. E.; Hutson, R.; Harvey, D. J.; Antrobus, R.; Petrescu, S. M.; Dwek, R. A.; Rudd, P. M., Ovarian cancer is associated with changes in glycosylation in both acute-phase proteins and IgG. *Glycobiology* 2007, 17, (12), 1344-56.

77. Rademacher, T. W.; Parekh, R. B.; Dwek, R. A.; Isenberg, D.; Rook, G.; Axford, J. S.; Roitt, I., The role of IgG glycoforms in the pathogenesis of rheumatoid arthritis. *Springer Semin Immunopathol* 1988, 10, (2-3), 231-49.

78. Axford, J. S.; Mackenzie, L.; Lydyard, P. M.; Hay, F. C.; Isenberg, D. A.; Roitt, I. M., Reduced B-cell galactosyltransferase activity in rheumatoid arthritis. *Lancet* 1987, 2, (8574), 1486-8.

79. Scherer, H. U.; van der Woude, D.; Ioan-Facsinay, A.; el Bannoudi, H.; Trouw, L. A.; Wang, J.; Haupl, T.; Burmester, G. R.; Deelder, A. M.; Huizinga, T. W.; Wuhrer, M.; Toes, R. E., Glycan profiling of anti-citrullinated protein antibodies isolated from human serum and synovial fluid. *Arthritis Rheum* 2010, 62, (6), 1620-9.

80. Scherer, H. U.; Wang, J.; Toes, R. E.; van der Woude, D.; Koeleman, C. A.; de Boer, A. R.; Huizinga, T. W.; Deelder, A. M.; Wuhrer, M., Immunoglobulin 1 (IgG1) Fc-glycosylation profiling of anti-citrullinated peptide antibodies from human serum. *Proteomics Clin Appl* 2009, 3, (1), 106-15.

81. Selman, M. H.; Niks, E. H.; Titulaer, M. J.; Verschuuren, J. J.; Wuhrer, M.; Deelder, A. M., IgG fc N-glycosylation changes in Lambert-Eaton myasthenic syndrome and myasthenia gravis. *J Proteome Res* 2011, 10, (1), 143-52.

82. Wang, J.; Balog, C. I.; Stavenhagen, K.; Koeleman, C. A.; Scherer, H. U.; Selman, M. H.; Deelder, A. M.; Huizinga, T. W.; Toes, R. E.; Wuhrer, M., Fc-glycosylation of IgG1 is modulated by B-cell stimuli. *Mol Cell Proteomics* 10, (5), M110 004655.

83. Lastra, G. C.; Thompson, S. J.; Lemonidis, A. S.; Elson, C. J., Changes in the galactose content of IgG during humoral immune responses. *Autoimmunity* 1998, 28, (1), 25-30.

84. Selman, M. H.; de Jong, S. E.; Soonawala, D.; Kroon, F. P.; Adegnika, A. A.; Deelder, A. M.; Hokke, C. H.; Yazdanbakhsh, M.; Wuhrer, M., Changes in antigen-specific IgG1 Fc N-glycosylation upon influenza and tetanus vaccination. *Mol Cell Proteomics* 2012, 11, (4), M111 014563.

85. Rakus, J. F.; Mahal, L. K., New technologies for glycomic analysis: toward a systematic understanding of the glycome. *Annu Rev Anal Chem (Palo Alto Calif)* 2011, 4, 367-92.

86. Knežević, A.; Polašek, O.; Gornik, O.; Rudan, I.; Campbell, H.; Hayward, C.; Wright, A.; Kolčić, I.; O'Donoghue, N.; Bones, J.; Rudd, P. M.; Lauc, G., Variability, heritability and environmental determinants of human plasma N-glycome. *J Proteome Res* 2009, 8, (2), 694-701.

87. Pučić, M.; Knežević, A.; Vidič, J.; Adamczyk, B.; Novokmet, M.; Polašek, O.; Gornik, O.; Šupraha-Goreta, S.; Wormald, M. R.; Redžić, I.; Campbell, H.; Wright, A.; Hastie, N. D.; Wilson, J. F.; Rudan, I.; Wuhrer, M.; Rudd, P. M.; Josić, D.; Lauc, G., High Throughput Isolation and Glycosylation Analysis of IgG-Variability and Heritability of the IgG Glycome in Three Isolated Human Populations. *Mol Cell Proteomics* 2011, 10, (10), M111 010090.

88. Royle, L.; Campbell, M. P.; Radcliffe, C. M.; White, D. M.; Harvey, D. J.; Abrahams, J. L.; Kim, Y. G.; Henry, G. W.; Shadick, N. A.; Weinblatt, M. E.; Lee, D. M.; Rudd, P. M.; Dwek, R. A., HPLC-based analysis of serum N-glycans on a 96-well plate platform with

dedicated database software. *Anal Biochem* 2008, 376, (1), 1-12.

89. Ruhaak, L. R.; Hennig, R.; Huhn, C.; Borowiak, M.; Dolhain, R. J.; Deelder, A. M.; Rapp, E.; Wuhrer, M., Optimized workflow for preparation of APTS-labeled N-glycans allowing high-throughput analysis of human plasma glycomes using 48-channel multiplexed CGE-LIF. *J Proteome Res* 2010, 9, (12), 6655-64.
90. Schwarzer, J.; Rapp, E.; Reichl, U., N-glycan analysis by CGE-LIF: profiling influenza A virus hemagglutinin N-glycosylation during vaccine production. *Electrophoresis* 2008, 29, (20), 4203-14.
91. Omtvedt, L. A.; Royle, L.; Husby, G.; Sletten, K.; Radcliffe, C. M.; Harvey, D. J.; Dwek, R. A.; Rudd, P. M., Glycan analysis of monoclonal antibodies secreted in deposition disorders indicates that subsets of plasma cells differentially process IgG glycans. *Arthritis Rheum* 2006, 54, (11), 3433-40.
92. Wuhrer, M.; Stam, J. C.; van de Geijn, F. E.; Koeleman, C. A.; Verrips, C. T.; Dolhain, R. J.; Hokke, C. H.; Deelder, A. M., Glycosylation profiling of immunoglobulin G (IgG) subclasses from human serum. *Proteomics* 2007, 7, (22), 4070-81.
93. Barut, M.; Podgornik, A.; Urbas, L.; Gabor, B.; Brne, P.; Vidic, J.; Plevcak, S.; Strancar, A., Methacrylate-based short monolithic columns: enabling tools for rapid and efficient analyses of biomolecules and nanoparticles. *J Sep Sci* 2008, 31, (11), 1867-80.
94. Josic, D.; Buchacher, A., Application of monoliths as supports for affinity chromatography and fast enzymatic conversion. *J. Biochem. Biophys. Methods* 2001, 49, (1-3), 153-74.
95. Jungbauer, A.; Hahn, R., Polymethacrylate monoliths for preparative and industrial separation of biomolecular assemblies. *J. Chromatogr. A* 2008, 1184, (1-2), 62-79.
96. Roberts, M. W.; Ongkudon, C. M.; Forde, G. M.; Danquah, M. K., Versatility of polymethacrylate monoliths for chromatographic purification of biomolecules. *J. Sep. Sci.* 2009, 32, (15-16), 2485-94.
97. Podgornik, A.; Strancar, A., Convective Interaction Media (CIM)--short layer monolithic chromatographic stationary phases. *Biotechnol. Annu. Rev.* 2005, 11, 281-333.
98. Tennikova, T. B.; Reusch, J., Short monolithic beds: history and introduction to the field. *J. Chromatogr. A* 2005, 1065, (1), 13-7.
99. Anumula, K. R., Quantitative determination of monosaccharides in glycoproteins by high-performance liquid chromatography with highly sensitive fluorescence detection. *Anal Biochem* 1994, 220, (2), 275-83.
100. Anumula, K. R., Advances in fluorescence derivatization methods for high-performance liquid chromatographic analysis of glycoprotein carbohydrates. *Anal Biochem* 2006, 350, (1), 1-23.
101. Ruhaak, L. R.; Zauner, G.; Huhn, C.; Bruggink, C.; Deelder, A. M.; Wuhrer, M., Glycan labeling strategies and their use in identification and quantification. *Anal Bioanal Chem* 2010, 397, (8), 3457-81.
102. Guile, G. R.; Rudd, P. M.; Wing, D. R.; Prime, S. B.; Dwek, R. A., A rapid high-resolution high-performance liquid chromatographic method for separating glycan mixtures and analyzing oligosaccharide profiles. *Anal Biochem* 1996, 240, (2), 210-26.
103. Callewaert, N.; Geysens, S.; Molemans, F.; Contreras, R., Ultrasensitive profiling and sequencing of N-linked oligosaccharides using standard DNA-sequencing equipment. *Glycobiology* 2001, 11, (4), 275-81.
104. Laroy, W.; Contreras, R.; Callewaert, N., Glycome mapping on DNA sequencing equipment. *Nat Protoc* 2006, 1, (1), 397-405.
105. Clarke, A.; Harmon, B.; DeFelippis, M. R., Analysis of 3-(acetylamino)-6-

aminoacridine-derivatized oligosaccharides from recombinant monoclonal antibodies by liquid chromatography-mass spectrometry. *Anal Biochem* 2009, 390, (2), 209-11.

106. Wuhrer, M.; Deelder, A. M.; Hokke, C. H., Protein glycosylation analysis by liquid chromatography-mass spectrometry. *J Chromatogr B Analyt Technol Biomed Life Sci* 2005, 825, (2), 124-33.

107. Maslen, S.; Sadowski, P.; Adam, A.; Lilley, K.; Stephens, E., Differentiation of isomeric N-glycan structures by normal-phase liquid chromatography-MALDI-TOF/TOF tandem mass spectrometry. *Anal Chem* 2006, 78, (24), 8491-8.

108. Wuhrer, M.; de Boer, A. R.; Deelder, A. M., Structural glycomics using hydrophilic interaction chromatography (HILIC) with mass spectrometry. *Mass Spectrom Rev* 2009, 28, (2), 192-206.

109. Wuhrer, M.; Koeleman, C. A.; Deelder, A. M.; Hokke, C. H., Normal-phase nanoscale liquid chromatography-mass spectrometry of underivatized oligosaccharides at low-femtomole sensitivity. *Anal Chem* 2004, 76, (3), 833-8.

110. Harvey, D. J.; Hunter, A. P.; Bateman, R. H.; Brown, J.; Critchley, G., Relationship between in-source and post-source fragment ions in the matrix-assisted laser desorption (ionization) mass spectra of carbohydrates recorded with reflectron time-of-flight mass spectrometers. *International Journal of Mass Spectrometry* 1999, 188, (1-2), 131-146.

111. Kang, P.; Mechref, Y.; Novotny, M. V., High-throughput solid-phase permethylation of glycans prior to mass spectrometry. *Rapid Commun Mass Spectrom* 2008, 22, (5), 721-34.

112. Mechref, Y.; Kang, P.; Novotny, M. V., Solid-phase permethylation for glycomic analysis. *Methods Mol Biol* 2009, 534, 53-64.

113. Harvey, D. J., Analysis of carbohydrates and glycoconjugates by matrix-assisted laser desorption/ionization mass spectrometry: An update for 2003-2004. *Mass Spectrom Rev* 2009, 28, (2), 273-361.

114. Selman, M. H.; Hoffmann, M.; Zauner, G.; McDonnell, L. A.; Balog, C. I.; Rapp, E.; Deelder, A. M.; Wuhrer, M., MALDI-TOF-MS analysis of sialylated glycans and glycopeptides using 4-chloro-alpha-cyanocinnamic acid matrix. *Proteomics* 2012, 12, (9), 1337-48.

115. Behan, J. L.; Smith, K. D., The analysis of glycosylation: a continued need for high pH anion exchange chromatography. *Biomed Chromatogr* 2011, 25, (1-2), 39-46.

116. Wada, Y.; Azadi, P.; Costello, C. E.; Dell, A.; Dwek, R. A.; Geyer, H.; Geyer, R.; Kakehi, K.; Karlsson, N. G.; Kato, K.; Kawasaki, N.; Khoo, K. H.; Kim, S.; Kondo, A.; Lattova, E.; Mechref, Y.; Miyoshi, E.; Nakamura, K.; Narimatsu, H.; Novotny, M. V.; Packer, N. H.; Perreault, H.; Peter-Katalinic, J.; Pohlentz, G.; Reinhold, V. N.; Rudd, P. M.; Suzuki, A.; Taniguchi, N., Comparison of the methods for profiling glycoprotein glycans--HUPO Human Disease Glycomics/Proteome Initiative multi-institutional study. *Glycobiology* 2007, 17, (4), 411-22.

117. Cummings, R. D., The repertoire of glycan determinants in the human glycome. *Mol Biosyst* 2009, 5, (10), 1087-104.

118. Nairn, A. V.; York, W. S.; Harris, K.; Hall, E. M.; Pierce, J. M.; Moremen, K. W., Regulation of glycan structures in animal tissues: transcript profiling of glycan-related genes. *J Biol Chem* 2008, 283, (25), 17298-313.

119. Moremen, K. W.; Tiemeyer, M.; Nairn, A. V., Vertebrate protein glycosylation: diversity, synthesis and function. *Nat Rev Mol Cell Biol* 2012, 13, 448-462.

120. Zoldos, V.; Horvat, T.; Lauc, G., Glycomics meets genomics,

epigenomics and other high throughput omics for system biology studies. *Curr Opin Chem Biol* 2012.

121. Freeze, H. H., Genetic defects in the human glycome. *Nat Rev Genet* 2006, 7, (7), 537-51.

122. Knežević, A.; Gornik, O.; Polašek, O.; Pučić, M.; Redžić, I.; Novokmet, M.; Rudd, P. M.; Wright, A. F.; Campbell, H.; Rudan, I.; Lauc, G., Effects of aging, body mass index, plasma lipid profiles, and smoking on human plasma N-glycans. *Glycobiology* 2010, 20, (8), 959-69.

123. Gornik, O.; Wagner, J.; Pucic, M.; Knezevic, A.; Redzic, I.; Lauc, G., Stability of N-glycan profiles in human plasma. *Glycobiology* 2009, 19, (12), 1547-53.

124. Ovsyannikova, I. G.; Haralambieva, I. H.; Dhiman, N.; O'Byrne, M. M.; Pankratz, V. S.; Jacobson, R. M.; Poland, G. A., Polymorphisms in the vitamin A receptor and innate immunity genes influence the antibody response to rubella vaccination. *Journal of Infective Diseases* 2010, 201, (2), 207-13.

125. Dhiman, N.; Haralambieva, I. H.; Kennedy, R. B.; Vierkant, R. A.; O'Byrne, M. M.; Ovsyannikova, I. G.; Jacobson, R. M.; Poland, G. A., SNP/haplotype associations in cytokine and cytokine receptor genes and immunity to rubella vaccine. *Immunogenetics* 2010, 62, (4), 197-210.

126. Elliott, P.; Chambers, J. C.; Zhang, W.; Clarke, R.; Hopewell, J. C.; Peden, J. F.; Erdmann, J.; Braund, P.; Engert, J. C.; Bennett, D.; Coin, L.; Ashby, D.; Tzoulaki, I.; Brown, I. J.; Mt-Isa, S.; McCarthy, M. I.; Peltonen, L.; Freimer, N. B.; Farrall, M.; Ruokonen, A.; Hamsten, A.; Lim, N.; Froguel, P.; Waterworth, D. M.; Vollenweider, P.; Waeber, G.; Jarvelin, M. R.; Mooser, V.; Scott, J.; Hall, A. S.; Schunkert, H.; Anand, S. S.; Collins, R.; Samani, N. J.; Watkins, H.; Kooner, J. S., Genetic Loci associated with C-reactive protein levels and risk of coronary heart disease. *Jama* 2009, 302, (1), 37-48.

127. Altshuler, D.; Daly, M. J.; Lander, E. S., Genetic mapping in human disease. *Science* 2008, 322, (5903), 881-8.

128. Rudd, P. M.; Rudan, I.; Wright, A. F., High-throughput glycome analysis is set to join high-throughput genomics. *J Proteome Res* 2009, 8, (3), 1105.

129. Visscher, P. M.; Brown, M. A.; McCarthy, M. I.; Yang, J., Five years of GWAS discovery. *Am J Hum Genet* 2012, 90, (1), 7-24.

130. Huffman, J. E.; Knezevic, A.; Vitart, V.; Kattla, J.; Adamczyk, B.; Novokmet, M.; Igl, W.; Pucic, M.; Zgaga, L.; Johannson, A.; Redzic, I.; Gornik, O.; Zemunik, T.; Polasek, O.; Kolcic, I.; Pehlic, M.; Koeleman, C. A.; Campbell, S.; Wild, S. H.; Hastie, N. D.; Campbell, H.; Gyllensten, U.; Wuhrer, M.; Wilson, J. F.; Hayward, C.; Rudan, I.; Rudd, P. M.; Wright, A. F.; Lauc, G., Polymorphisms in B3GAT1, SLC9A9 and MGAT5 are associated with variation within the human plasma N-glycome of 3533 European adults. *Hum Mol Genet* 2011, 20, (24), 5000-11.

131. Kutalik, Z.; Benyamin, B.; Bergmann, S.; Mooser, V.; Waeber, G.; Montgomery, G. W.; Martin, N. G.; Madden, P. A.; Heath, A. C.; Beckmann, J. S.; Vollenweider, P.; Marques-Vidal, P.; Whitfield, J. B., Genome-wide association study identifies two loci strongly affecting transferrin glycosylation. *Hum Mol Genet* 2011, 20, (18), 3710-7.

132. Lauc, G.; Essafi, A.; Huffman, J. E.; Hayward, C.; Knezevic, A.; Kattla, J. J.; Polasek, O.; Gornik, O.; Vitart, V.; Abrahams, J. L.; Pucic, M.; Novokmet, M.; Redzic, I.; Campbell, S.; Wild, S. H.; Borovecki, F.; Wang, W.; Kolcic, I.; Zgaga, L.; Gyllensten, U.; Wilson, J. F.; Wright, A. F.; Hastie, N. D.; Campbell, H.; Rudd, P. M.; Rudan, I., Genomics meets glycomics-the first GWAS study of human N-Glycome identifies HNF1alpha as a master regulator of plasma protein fucosylation. *PLoS Genet* 2010, 6, (12), e1001256.

133. Mitsumoto, Y.; Oka, S.; Sakuma, H.; Inazawa, J.; Kawasaki, T., Cloning and chromosomal mapping of human glucuronyltransferase involved in biosynthesis of the HNK-1 carbohydrate epitope. *Genomics* 2000, 65, (2), 166-73.

134. Roxrud, I.; Raiborg, C.; Gilfillan, G. D.; Stromme, P.; Stenmark, H., Dual degradation mechanisms ensure disposal of NHE6 mutant protein associated with neurological disease. *Exp Cell Res* 2009, 315, (17), 3014-27.

135. Knežević, A.; Gornik, O.; Polašek, O.; Pučić, M.; Novokmet, M.; Redžić, I.; Rudd, P. M.; Wright, A. F.; Campbell, H.; Rudan, I.; Lauc, G., Effects of aging, body mass index, plasma lipid profiles, and smoking on human plasma N-glycans. *Glycobiology* 2010, 20, (8), 959-969.

136. Knežević, A.; Polašek, O.; Gornik, O.; Rudan, I.; Campbell, H.; Hayward, C.; Wright, A.; Kolčić, I.; O'Donoghue, N.; Bones, J.; Rudd, P. M.; Lauc, G., Variability, Heritability and Environmental Determinants of Human Plasma N-Glycome. *Journal of Proteome Research* 2009, 8, 694-701.

137. Pucic, M.; Knezevic, A.; Vidic, J.; Adamczyk, B.; Novokmet, M.; Polasek, O.; Gornik, O.; Supraha-Goreta, S.; Wormald, M. R.; Redzic, I.; Campbell, H.; Wright, A.; Hastie, N. D.; Wilson, J. F.; Rudan, I.; Wuhrer, M.; Rudd, P. M.; Josic, D.; Lauc, G., High throughput isolation and glycosylation analysis of IgG-variability and heritability of the IgG glycome in three isolated human populations. *Mol Cell Proteomics* 2011, 10, (10), M111 010090.

138. Sivakumaran, S.; Agakov, F.; Theodoratou, E.; Prendergast, J. G.; Zgaga, L.; Manolio, T.; Rudan, I.; McKeigue, P.; Wilson, J. F.; Campbell, H., Abundant pleiotropy in human complex diseases and traits. *Am J Hum Genet* 2011, 89, (5), 607-18.

139. Vanhooren, V.; Desmyter, L.; Liu, X. E.; Cardelli, M.; Franceschi, C.; Federico, A.; Libert, C.; Laroy, W.; Dewaele, S.; Contreras, R.; Chen, C., N-glycomic changes in serum proteins during human aging. *Rejuvenation Res* 2007, 10, (4), 521-531a.

SUMMARY

Immunoglobulin G (IgG) is the most abundant antibody class in the human blood (approx. 10 mg/ml) and a major effector molecule of the humoral immune response. IgG antibodies play an important role in defending the body against many different antigens by mediating activation of a wide range of effector functions which result in destruction and removal of the pathogen. Four subclasses of human IgG (IgG1, IgG2, IgG3 and IgG4) are glycoproteins composed of two heavy and two light chains linked together by interchain disulphide bonds. The two light chains together with the parts of the heavy chains (V_H and C_H1 domains) form two Fab moieties which are linked by a flexible hinge region to one Fc moiety formed by the remainders of the two heavy chains (C_H2 and C_H3 domains).[3] The length of the hinge region and the number of interchain disulphide bonds differs significantly between the IgG subclasses and influences mobility and conformation of the Fab and Fc moieties, with respect to each other.[4, 5] Each heavy chain in the Fc region carries a single covalently attached biantennary *N*-glycan at the highly conserved asparagine 297 in the C_H2 domain.[7] Fc glycans are essential structural components of the IgG molecule and even minor changes in glycan composition can have a profound influence on IgG effector functions by modulating binding to Fc receptors.[12, 13] The majority of IgG *N*-glycans are attached to the heavy chains of the Fc region, but cca. 20% of polyclonal human IgG molecules also contain *N*-glycans within the Fab regions of the light chain, the heavy chain or both.[13] Microheterogeneity of human IgG glycans is known to be dependent on various physiological parameters (age, sex, pregnancy) and pathological conditions (tumors, infections, autoimmune diseases, etc.).[3]

Due to the structural complexity of glycans and technological limitations, until only a few years ago glycan analysis was extremely laborious and complex, hampering large-scale studies of the glycome. However, major progress has been made in the last few years resulting in several high-throughput analytical techniques for glycan analysis.[56, 86-90]

The main objectives of this thesis were optimization and application of a new technology for understanding variability and heritability of IgG glycosylation in a human population. The first paper of this thesis (*High Throughput Isolation and Glycosylation Analysis of IgG – Variability and Heritability of the IgG Glycome in Three Isolated Human Populations*) describes a high-throughput quantitative glycan analysis method and its application in the first large scale population study of the IgG *N*-glycome. One of the major bottlenecks in a large scale proteomics and glycomics studies is protein purification from a large number of samples. A 96-well Protein G monolithic plate was developed and successfully applied for the high-thoughput isolation of IgG from over 2000 plasma samples. Optimized hydrophilic interaction liquid chromatography (HILIC) method on a Waters BEH Glycan chromatography column provided separation of IgG *N*-glycans into 24 chromatographic peaks. Identity of glycan structures in each chromatography peak was detemined by exoglycosidase

digestion and mass spectrometry. This analysis for the first time provided insight into the variability of IgG glycome within the human population, and for the first time estimated heritability of IgG glycosylation. Very high variability between individuals was observed, approximately three times higher than in the total plasma glycome.[86] Heritability of IgG glycans was generally between 30 and 50%. The individual's age was associated with a significant decrease in galactose and increase of bisecting *N*-acetylglucosamine (GlcNAc), whereas other functional elements of IgG glycosylation did not change much with age.

HILIC enabled a detailed analysis of a mixture of released Fc and Fab *N*-glycans of all four IgG subclasses. In order to obtain subclass- and site-specific IgG glycosylation profiling of the same cohort, we have analyzed IgG Fc *N*-glycopeptides by MALDI-TOF-MS as presented in the second paper of this thesis (*High-throughput IgG Fc N-glycosylation profiling by mass spectrometry of glycopeptides*). Fc *N*-glycosylation profiling confirmed most of the previously described age- and sex-related IgG glycosylation changes.[52-56, 139] Levels of galactosylation and sialylation decreased with increasing age and showed significant sex dependence. Interestingly, the most prominent changes in glycosylation in females were observed around the age of 45 to 60 years when females usually enter the menopause. The incidence of bisecting GlcNAc increased in younger individuals and reached a plateau at older age. In addition, we have compared two analytical approaches, HILIC and MALDI-TOF-MS, applied for the IgG glycosylation analysis of our cohort.

Third paper of this thesis (*Loci associated with N-glycosylation of human immunoglobulin G show pleiotropy with autoimmune diseases and haematological cancers*) shows the ability to identify genetic loci that control glycosylation of a single plasma protein using genome wide association study (GWAS) meta-analysis. Moreover, it represents one of the first convincing demonstrations that GWAS approach can lead to biomarker discovery for human disease. This approach enabled us to combine high-throughput glycomics measurements (HILIC and MS IgG glycan traits) with high-throughput genomics and resulted in identification of nine genetic loci that associate with glycans with genome-wide significance. Of these, four loci contained genes encoding glycosyltransferases (ST6GAL1, B4GALT1, FUT8 and MGAT3), thus the observed associations were biologically founded. The remaining five genetic loci (IKZF1, IL6ST-ANKRD55, ABCF2-SMARCD3, SUV420H1, and SMARCB1-DERL3) were not previously implicated in protein glycosylation, but the most of them have been reported to be relevant for autoimmune and inflammatory conditions and/or haematological cancers. A particularly interesting gene, IKZF1, was found to be associated with multiple IgG *N*-glycan traits. Since this gene has been involved in numerous diseases, including systemic lupus erythematosus (SLE), we explored biomarker potential of *N*-glycans in 101 cases with SLE and 183 matched controls and demonstrated substantial discriminative power. Results from this study suggest that IgG *N*-glycome is regulated through a complex interaction of genes which affect multiple glycan traits.

Fourth paper of this thesis (*Changes in plasma and IgG N-glycome during childhood and adolescence*) describes plasma and IgG *N*-glycome of 170 children and adolescents between 6 and 18 years of age. The results revealed a high variability of glycosylation features and a large number of age-dependent glycans. The plasma *N*-glycome of younger children was found to contain a higher proportion of large complex glycan structures (tri- and tetra-sialylated glycans) which decreased with age while disialylated glycans changed in the opposite direction. When the results of the plasma and IgG glycan changes during growing up in children were compared with those in adults, the behavior of almost all glycan features changed trend or the rate of change in adulthood. An example of the opposite trends is the level of agalactosylated glycans which decreased with age in children, reached the minimal level in early adulthood and then started to increase with age. Moreover, sex differences are much less pronounced in children than in adults and are present mainly during puberty.

This thesis presents a novel high-throughput approach for IgG purification and IgG glycan analysis and its application in understanding variability and heritability of IgG glycosylation in a human population. Moreover, the ability to identify genetic loci that control IgG glycosylation by combining high-throughput glycomics and genomics measurements has been demonstrated.

SAŽETAK

Imunoglobulin G (IgG) je najzastupljenije antitijelo u krvi čovjeka (prosječna konc. 10 mg/ml) i glavna efektorska molekula u humoralnom imunološkom odgovoru. IgG antitijela imaju važnu ulogu u obrani organizma od različitih antigena jer posreduju u aktivaciji niza efektorskih funkcija koje u konačnici dovode do uništavanja i uklanjanja patogena. Sve četiri potklase IgG-a (IgG1, IgG2, IgG3 i IgG4) su glikoproteini i sastoje se od dva teška i dva laka lanca međusobno povezana disulfidnim vezama. Dva laka lanca zajedno s dijelovima teških lanaca (V_H i C_H1 domene) čine dva Fab fragmenta povezana preko zglobne regije na jedan Fc fragment građen od ostataka teških lanaca (C_H2 i C_H3 domene).[3] Duljina zglobne regije i broj disulfidnih veza značajno se razlikuju među IgG potklasama te utječu na pokretljivost i konformaciju Fab i Fc fragmenata.[4, 5] Oba teška lanca Fc fragmenta imaju kovalentno vezan biantenarni *N*-glikan na Asn_{297} u visoko konzerviranom mjestu C_H2 domene.[7] Fc glikani su ključni za strukturu IgG molekule pa tako i male promjene u sastavu glikana mogu značajno utjecati na interakcije s Fc receptorima i time na efektorske funkcije IgG-a.[12, 13] Većina IgG *N*-glikana je vezana na teške lance Fc regije, međutim oko 20 % poliklonalnih IgG molekula u ljudi sadrži *N*-glikane i unutar Fab regija.[13] Kod ljudi IgG glikozilacija pokazuje ovisnost o raznim fiziološkim parametrima (dob, spol, trudnoća) i patološkim stanjima (tumori, infekcije, autoimune bolesti itd.).[3]

Zbog strukturalne složenosti glikana i tehnoloških ograničenja, analize glikana su do prije nekoliko godina bile iznimno zahtjevne i složene što je otežavalo velike studije glikoma. Međutim, velik napredak u analizi glikana napravljen je posljednjih godina što je dovelo do razvoja nekoliko visoko protočnih analitičkih metoda.[56, 86-90]

Glavni ciljevi ove doktorske disertacije su bili optimizacija i primjena nove tehnologije kako bi se razumjela varijabilnost i heritabilnost glikozilacije IgG-a u ljudskoj populaciji. Prvi rad ove disertacije (*High Throughput Isolation and Glycosylation Analysis of IgG – Variability and Heritability of the IgG Glycome in Three Isolated Human Populations*) opisuje visoko protočnu metodu za analizu glikana i njenu primjenu u prvoj velikoj populacijskoj studiji IgG *N*-glikoma. Jedno od glavnih uskih grla u proteomici i glikomici je pročišćavanje proteina iz velikog broja uzoraka. Razvijena je Protein G monolitna pločica s 96 jažica i uspješno primjenjena za visoko protočnu izolaciju IgG-a iz preko 2000 uzoraka plazme. Optimirana je kromatografska metoda za tekućinsku kromatografiju temeljenu na hidrofilnim interakcijama na Waters BEH Glycan kromatografskoj koloni što je omogućilo razdvajanje IgG *N*-glikana u 24 kromatografska vrška. Identitet glikanskih struktura u pojedinom kromatografskom vršku određen je egzoglikozidaznim digestijama i masenom spektrometrijom. Ova analiza je po prvi put omogućila uvid u varijabilnost IgG glikoma unutar populacije i procijenila heritabilnost glikozilacije IgG-a. Opažena je velika varijabilnost među pojedincima, prosječno tri puta veća nego u glikomu ukupne plazme.[86] Heritabilnost IgG glikana je bila između 30 i 50%. Dob pojedinca je povezana sa značajnim padom

galaktozilacije i povećanjem udjela račvajućeg *N*-acetilglukozamina (GlcNAc), dok se drugi funkcionalni elementi IgG glikozilacije nisu značajno mijenjali s godinama.

Tekućinska kromatografija temeljena na hidrofilnim interakcijama omogućila je detaljnju analizu ukupnih Fc i Fab *N*-glikana sa sve četiri potklase IgG-a. Drugi rad ove disertacije prikazuje analizu IgG Fc *N*-glikopeptida iste populacije ljudi MALDI-TOF masenom spektrometrijom (*High-throughput IgG Fc N-glycosylation profiling by mass spectrometry of glycopeptides*) pomoću koje je određena glikozilacija pojedinih potklasa IgG-a. Ovom metodom potvrdili smo većinu prethodno opisanih promjena glikozilacije IgG-a s dobi i spolom.[52-56, 139] Stupanj galaktozilacije i sijalinizacije se smanjivao s godinama i pokazao značajnu ovisnost o spolu. Najveće promjene u glikozilaciji opažene su kod žena između 45 i 60 godina starosti kad žene obično prolaze kroz menopauzu. Stupanj račvajućeg GlcNAc-a povećavao se kod mlađih pojedinaca i dosegnuo plato u starijoj dobi. Također, u sklopu ovog rada uspoređena su dva analitička pristupa primjenjena za analizu glikozilacije IgG-a populacija uključenih u ovu studiju, HILIC i MALDI-TOF-MS.

Treći rad ove disertacije (*Loci associated with N-glycosylation of human immunoglobulin G show pleiotropy with autoimmune diseases and haematological cancers*) prikazuje mogućnost identifikacije genskih lokusa koji kontroliraju glikozilaciju jednog proteina plazme pomoću cjelogenomske asocijacijske studije (engl. GWAS). Također, ovaj rad predstavlja jedan od prvih uvjerljivih prikaza učinkovitosti GWAS pristupa u otkrivanju biomarkera za određene ljudske bolesti. Ovaj pristup omogućio nam je kombiniranje visoko protočne glikomike (HILIC i MS analiza IgG glikozilacije) s visoko protočnom genomikom što je dovelo do identifikacije devet genskih lokusa koji su povezani s glikanima sa značajnošću na razini cjelokupnog genoma. Četiri lokusa su sadržavala gene koji kodiraju za glikoziltransferaze (ST6GAL1, B4GALT1, FUT8 i MGAT3) ukazujući na biološku utemeljenost naših opažanja. Preostalih pet genskih lokusa (IKZF1, IL6ST-ANKRD55, ABCF2-SMARCD3, SUV420H1, i SMARCB1-DERL3) dosad nisu povezivani s glikozilacijom proteina, ali većina njih je povezivana s autoimunim bolestima, upalnim procesima i/ili hematološkim tumorima. Posebno zanimljiv gen, IKZF1, povezan je s nekoliko glikanskih karakteristika IgG-a. Budući da je ovaj gen povezivan s nekoliko bolesti, uključujući i sistemski lupus eritematosus (SLE), istražili smo potencijal *N*-glikana kao biomarkera u 101 pacijentu oboljelom od lupusa i 183 kontrolna uzorka te pokazali značajnu diskriminativnu moć. Rezultati ove studije upućuju da je IgG *N*-glikom reguliran složenom interakcijom gena koji utječu na brojne glikanske karakteristike.

Četvrti rad ove disertacije (*Changes in plasma and IgG N-glycome during childhood and adolescence*) opisuje *N*-glikom plazme i IgG-a kod 170 djece i adolescenata između 6 i 18 godina starosti. Razultati su pokazali veliku varijabilnost glikozilacije i velik broj glikana ovisnih o dobi. Pokazano je da *N*-glikom plazme mlađe djece sadrži značajan udio velikih kompleksnih glikanskih struktura (tri- i tetrasijalinizirani glikani) koji se snižava s godinama dok se udio disijaliniziranih glikana mijenja u suprotnom smjeru. Usporedbom promjena u glikozilaciji plazme i IgG-a kod djece s promjenama u

odrasloj dobi, opaženo je drugačije ponašanje većine glikanskih karakteristika. Primjer suprotnih trendova je stupanj agalaktoziliranih glikana koji se kod djece smanjuje s godinama, doseže minimalne vrijednosti u ranoj odrasloj dobi nakon čega se počinje povećavati s godinama. Također, spolne razlike su puno manje izražene kod djece nego kod odraslih i većinom su prisutne tijekom puberteta.

Ova doktorska disertacija prikazuje novi visoko protočni pristup pročišćavanju IgG-a i analizi IgG glikana te njegovu primjenu u razumijevanju varijabilnosti i heritabilnosti glikazilacije IgG-a u ljudskoj populaciji. Također, prikazana je i mogućnost identifikacije genskih lokusa koji kontroliraju glikozilaciju IgG-a kombiniranjem visoko protočne glikomike i genomike.

ABBREVIATIONS

2-AA	2-aminobenzoic acid
2-AB	2-aminobenzamide
ABEE	*p*-aminobenzoic acid ethyl ester
ACN	acetonitrile
ACPA	anti-citrullinated protein antibodies
ADCC	antibody-dependent cellular cytotoxicity
APTS	1-aminopyrene-3,6,8-trisulfonic acid
Asn	asparagine
B3GAT1	β-1,3-glucuronyltransferase 1
BSA	bovine serum albumin
CDC	complement-dependent cytotoxicity
CE	capillary electrophoresis
CGE-LIF	capillary gel electrophoresis with laser induced fluorescence
Cl-CAA	α-cyano-4-hydroxycinnamic acid
DC	dendritic cell
DC-SIGN	Dendritic Cell-Specific Intercellular adhesion molecule-3-Grabbing Non-Integrin
DHB	dihydroxybenzoic acid
ESI	electrospray ionization
FA	formic acid
Fab	fragment antigen binding
Fc	fragment crystallizable
FcγR	Fcγ receptor
FUT3	α(1-3/4)-*L*-fucosyltransferase
FUT5	α(1,3)-fucosyltransferase-V
FUT6	α(1,3)-fucosyltransferase-VI
FUT8	α(1,6)-fucosyltransferase
GlcNAc	*N*-acetyl glucosamine
GnT-III	β-1,4-*N*-acetylglucosaminyltransferase III
GWAS	genome-wide assoaciation study
HILIC	hydrophilic interaction liquid chromatography
HNF1α	hepatocyte nuclear factor 1 α
HPAEC-PAD	high-performance anion-exchange chromatography with pulsed amperometric detection
HPLC	high-performance liquid chromatography
Ig	immunoglobulin
IgG	immunoglobulin G
IL-33	interleukin 33
IVIg	intravenous immunoglobulin

LC	liquid chromatography
LEMS	Lambert-Eaton myasthenic syndrome
m/z	mass over charge ratio
mAb	monoclonal antibody
MALDI	matrix asssisted laser desorption
MBL	mannose-binding lectin
MG	myasthenia gravis
MGAT5	mannosyl (α1,6)-glycoprotein beta-1,6-N-acetyl-glucosaminyltransferase
MS	mass spectrometry
NK cell	natural killer cell
PNGaseF	*N*-glycosidase F
RA	rheumatoid arthritis
RP	reverse phase
UPLC	ultra-performance liquid chromatography
SLC9A9	solute carrier family 9, subfamily A, member 9
SNA	Sambucus nigra agglutinin
SNP	single nucleotide polymorphism
SPE	solid phase extraction
TOF	time of flight
WAX	weak anion exchange

LIST OF PUBLICATIONS

Lauc G, Huffman J, **Pučić M**, Zgaga L, Adamczyk B, Mužinić A, Novokmet M, Polašek O, Gornik O, Krištić J, Keser T, Vitart V, Uh H-W, Molokhia M, Patrick AL, McKeigue P, Kolčić I, Lukić IK, Swann O, van Leeuwen FN, Ruhaak LR, Houwing-Duistermaat JJ, Slagboom PE, Beekman M, de Craen AJM, Deelder AM, Zeng Q, Wang W, Hastie ND, Gyllensten U, Wilson JF, Wuhrer M, Wright AF, Rudd PM, Hayward C, Aulchenko Y, Campbell H, Rudan I. LOCI ASSOCIATED WITH N-GLYCOSYLATION OF HUMAN IMMUNOGLOBULIN G SHOW PLEIOTROPY WITH AUTOIMMUNE DISEASES AND HAEMATOLOGICAL CANCERS. PLOS Genetics [accepted manuscript]

Pučić Baković M, Selman MH, Hoffmann M, Rudan I, Campbell H, Deelder AM, Lauc G, Wuhrer M. HIGH-THROUGHPUT IGG Fc N-GLYCOSYLATION PROFILING BY MASS SPECTROMETRY OF GLYCOPEPTIDES. J Proteome Res. 2013 Jan 8. [Epub ahead of print]

Thanabalasingham G, Huffman JE, Kattla JJ, Novokmet M, Rudan I, Gloyn AL, Hayward C, Adamczyk B, Reynolds RM, Mužinić A, Hassanali N, **Pučić M**, Bennett AJ, Essafi A, Polašek O, Mughal SA, Redžić I, Primorac D, Zgaga L, Kolčić I, Hansen T, Gasperikova D, Tjora E, Strachan MW, Nielsen T, Stanik J, Klimes I, Pedersen OB, Njølstad PR, Wild SH, Gyllensten U, Gornik O, Wilson JF, Hastie ND, Campbell H, McCarthy MI, Rudd PM, Owen KR, Lauc G, Wright AF. MUTATIONS IN HNF1A RESULT IN MARKED ALTERATIONS OF PLASMA GLYCAN PROFILE. Diabetes. 2012 Dec 28. [Epub ahead of print]

Pučić M, Mužinić A, Novokmet M, Škledar M, Pivac N, Lauc G, Gornik O. CHANGES IN PLASMA AND IgG N-GLYCOME DURING CHILDHOOD AND ADOLESCENCE. Glycobiology. 2012 Jul; 22(7):975-82.

Zoldoš V, Horvat T, Novokmet M, Cuenin C, Mužinić A, **Pučić M**, Huffman JE, Gornik O, Polašek O, Campbell H, Hayward C, Wright AF, Rudan I, Owen K, McCarthy MI, Herceg Z, Lauc G. EPIGENETIC SILENCING OF HNF1A ASSOCIATES WITH CHANGES IN THE COMPOSITION OF THE HUMAN PLASMA N-GLYCOME. Epigenetics. 2012 Feb 1;7(2):164-72.

Novokmet M, **Pučić M**, Redžić I, Mužinić A, Gornik O. ROBUSTNESS TESTING OF THE HIGH THROUGHPUT HPLC-BASED ANALYSIS OF PLASMA N-GLYCANS. Biochim Biophys Acta. 2012 Sep;1820(9):1399-404.

Saldova R, Huffman JE, Adamczyk B, Mužinić A, Kattla JJ, **Pučić M**, Novokmet M, Abrahams JL, Hayward C, Rudan I, Wild SH, Wright AF, Polašek O, Lauc G, Campbell H, Wilson JF, Rudd PM. ASSOCIATION OF MEDICATION WITH THE HUMAN PLASMA N-GLYCOME. J Proteome Res. 2012 Mar 2;11(3):1821-31.

Huffman JE, Knežević A, Vitart V, Kattla J, Adamczyk B, Novokmet M, Igl W, **Pučič M**, Zgaga L, Johannson A, Redzic I, Gornik O, Zemunik T, Polašek O, Kolčić I, Pehlic M, Koeleman CA, Campbell S, Wild SH, Hastie ND, Campbell H, Gyllensten U, Wuhrer M, Wilson JF, Hayward C, Rudan I, Rudd PM, Wright AF, Lauc G. POLYMORPHISMS IN B3GAT1, SLC9A9 AND MGAT5 ARE ASSOCIATED WITH VARIATION WITHIN THE HUMAN PLASMA N-GLYCOME OF 3533 EUROPEAN ADULTS. Hum Mol Genet. 2011 Dec 15;20(24):5000-11.

Pučić M, Knežević A, Vidič J, Adamcyzk B, Novokmet M, Polašek O, Gornik O, Šupraha-Goreta S, Wormald MR, Redžić I, Campbell H, Wright A, Hastie ND, Wilson JF, Rudan I, Wuhrer M, Rudd PM, Josić Dj, Lauc G. 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.

Igl W, Polašek O, Gornik O, Kneževic A, **Pučić M**, Novokmet M, Huffman J, Gnewuch C, Liebisch G, Rudd PM, Campbell H, Wilson JF, Rudan I, Gyllensten U, Schmitz G, Lauc G. GLYCOMICS MEETS LIPIDOMCS-ASSOCIATIONS OF N-GLYCANS WITH CLASSICAL LIPIDA, GLYCEROPHOSPHOLIPIDS, AND SPHINGOLIPIDS IN THREE EUROPEAN POPULATIONS. Mol Biosyst. 7(6):1852-62, 2011.

Pivac N, Knežević A, Gornik O, **Pučić M**, Igl W, Peeters H, Crepel A, Steyaert J, Novokmet M, Redžić I, Nikolac M, Hercigonja VN, Čurković KD, Čurković M, Nedic G, Muck-Seler D, Borovečki F, Rudan I, Lauc G. HUMAN PLASMA GLYCOME IN ATTENTION-DEFICIT HYPERACTIVITY DISORDER AND AUTISM SPECTRUM DISORDERS. Mol Cell Proteomics. 10(1):M110.004200, 2011.

Lauc G, Essafi A, Huffman JE, Hayward C, Knežević A, Kattla JJ, Polašek O, Gornik O, Vitart V, Abrahams JL, **Pučić M**, Novokmet M, Redžić I, Campbell S, Wild SH, Borovečki F, Wang W, Kolčić I, Zgaga L, Gyllensten U, Wilson JF, Wright AF, Hastie ND, Campbell H, Rudd PM, Rudan I. GENOMICS MEETS GLYCOMICS-THE FIRST GWAS STUDY OF HUMAN N-GLYCOME IDENTIFIES HNF1A AS A MASTER REGULATOR OF PLASMA PROTEIN FUCOSYLATION. PLoS Genet. 23;6(12):e1001256, 2010.

Pučić M, Pinto S, Novokmet M, Knežević A, Gornik O, Polašek O, Vlahoviček K, Rudd PM, Wright AF, Campbell H, Rudan I, Lauc G. COMMON ABBERATIONS FROM THE NORMAL HUMAN PLASMA N-GLYCANS PROFILE. Glycobiology. 20(8):970-5, 2010.

Knežević A, Gornik O, Polašek O, **Pučić M**, Redžić I, Novokmet M, Rudd PM, Wright AF, Campbell H, Rudan I, Lauc G. EFFECTS OF AGING, BODY MASS INDEX, PLASMA LIPID PROFILES, AND SMOKING OF HUMAN PLASMA N-GLYCANS. Glycobiology. 20(8):959-69, 2010.

Gornik O, Wagner J, **Pučić M**, Knežević A , Redžić I and Lauc G. STABILITY OF N-GLYCAN PROFILES IN HUMAN PLASMA. Glycobiology. 19(12):1547-53, 2009.

CURRICULUM VITAE

Maja Pučić Baković was born on 26th of August 1983 in Zagreb, Croatia, where she spent her childhood and high school days.

In 2002 she started studying Molecular Biology at the Faculty of Science in Zagreb. She made her graduation thesis entitled: „*Quantification of DNA from bone samples by real-time PCR*“ at the Department of Legal Medicine and Criminalistics of the Faculty of Medicine in Zagreb and graduated in April 2008. After graduation she started volunteering at the Laboratory of Molecular Virology and Bacteriology of the Ruđer Bošković Institute in Zagreb until October 2008 when she got employed as junior researcher at Genos Ltd., Zagreb. In 2009 she started PhD programme at the Faculty of Science of the University of Zagreb.

Due to collaboration with prof. dr. Djuro Josić in 2010 she spent three months at the Proteomics Core at COBRE Center for Cancer Research and Development of the Rhode Island Hospital in Providence, RI, USA, where she started her PhD entitled: „*Variability and heritability of immunoglobulin G glycosylation*“. Later that year she got the FEBS (Federation of European Bichemical Societies) scholarship and the opportunity to spend another three months at the Biomolecular Mass Spectrometry Unit, Department of Parasitology of the Leiden University Medical Center in Leiden, the Netherlands. In August 2012 she returned to Leiden and continued collaboration with dr. Manfred Wuhrer and his group. In 2012 she also went to BIA Separations Ltd. in Slovenia to work further on optimization of IgG purification. Maja received Young Scientist Award for 2011 from the Croatian Society of Biochemistry and Molecular Biology and Best Publication Award for 2011 from the Croatian Immunological Society.