Open access • Posted Content • DOI:10.1101/535237

Global variability of the human IgG glycome

Jerko Štambuk, Natali Nakić, Frano Vučković, Maja Pučić-Baković ...+43 more authors

Institutions: International School for Advanced Studies, University of Zagreb, Capital Medical University, Edith Cowan University ...+9 more institutions

Published on: 01 Feb 2019 - bioRxiv (Cold Spring Harbor Laboratory)

Topics: Immunoglobulin G and Population

Global variability of the human IgG glycome

Jerko Štambuk[1]; Natali Nakić[2]; Frano Vučković[1]; Maja Pučić-Baković[1]; Genadij Razdorov[1]; Irena Trbojević-Akmačić[1]; Mislav Novokmet[1]; Toma Keser[3]; Marija Vilaj[1]; Tamara Pavić[3]; Ivan Gudelj[1]; Mirna Šimurina[3]; Manshu Song[4,23]; Hao Wang[4]; Marijana Peričić Salihović[5], Harry Campbell[6]; Igor Rudan[6]; Ivana Kolčić[7]; Leigh Anne Eller[8,9]; Paul McKeigue[6]; Merlin L. Robb[8,9]; Jonas Halfvarson[10]; Metin Kurtoglu[11]; Vito Annese[12]; Tatjana Škarić-Jurić[5]; Mariam Molokhia[13]; Ozren Polašek[7]; Caroline Hayward[14]; Hannah Kibuuka[15]; Kujtim Thaqi[16]; Dragan Primorac[1], Christian Gieger[17]; Sorachai Nitayaphan[18]; Tim Spector[19]; Youxin Wang[4,23]; Therese Tillin[20]; Nish Chaturvedi[20]; James F. Wilson[6,14]; Moses Schanfield[21]; Maxim Filipenko[22]; Wei Wang[4,23]; Gordan Lauc[1,3*]

*e-mail: glauc@pharma.hr

[1]Genos Glycoscience Research Laboratory, Zagreb, Croatia.

[2]Department of Neuroscience, Scuola Internazionale Superiore di Studi Avanzati (SISSA), Trieste, Italy

[3]Faculty of Pharmacy and Biochemistry, University of Zagreb, Zagreb, Croatia.

[4]Beijing Key Laboratory of Clinical Epidemiology, School of Public Health, Capital Medical University, Beijing, China.

[5]Institute for Anthropological Research, Zagreb, Croatia.

[6]Centre for Global Health Research, Usher Institute of Population Health Sciences and Informatics, The University of Edinburgh, Edinburgh, United Kingdom.

[7]School of Medicine, University of Split, Split, Croatia.

[8]Walter Reed Army Institute of Research, Silver Spring, Maryland, USA.

[9]Henry M. Jackson Foundation for the Advancement of Military Medicine, Bethesda, Maryland, USA.

[10]Department of Gastroenterology, Faculty of Medicine and Health, Örebro University, Örebro, Sweden.

[11]Department of Oncology, Koç University School of Medicine, Istanbul, Turkey.

[12]Careggi University Hospital, Florence, Italy.

[13]School of Population Health and Environmental Sciences, King's College London, United Kingdom.

[14]MRC Human Genetics Unit, MRC Institute for Genetics and Molecular Medicine, University of Edinburgh, Edinburgh, United Kingdom.

[15]Makerere University Walter Reed Project, Kampala, Uganda.

[16]Institute of Clinical Biochemistry, Priština, Kosovo.

[17]Helmholtz Zentrum München - German Research Center for Environmental Health, Neuherberg, Germany.

[18]Armed Forces Research Institute of Medical Sciences, Bangkok, Thailand.

[19]Department of Twin Research & Genetic Epidemiology, King's College London, London, United Kingdom.

[20]Institute of Cardiovascular Science, Faculty of Population Health Sciences, London, United Kingdom.

[21]Department of Forensic Sciences, George Washington University, Washington, DC, USA.

[22]Institute of Chemical Biology and Fundamental Medicine, Novosibirsk, Russia.

[23]School of Medical and Health Sciences, Edith Cowan University, Perth, Australia.

Abstract

Immunoglobulin G (IgG) is the most abundant serum antibody and is a key determinant of the humoral immune response. Its structural characteristics and effector functions are modulated through the attachment of various sugar moieties called glycans. IgG N-glycome patterns change with the age of individual and in different diseases. Variability of IgG glycosylation within a population is well studied and is affected by a combination of genetic and environmental factors. However, global inter-population differences in IgG glycosylation have never been properly addressed. Here we present population-specific N-glycosylation patterns of whole IgG, analysed in 5 different populations totalling 10,482 IgG glycomes, and of IgG's fragment crystallisable region (Fc), analysed in 2,530 samples from 27 populations sampled across the world. We observed that country of residence associates with many N-glycan features and is a strong predictor of monogalactosylation variability. IgG galactosylation also strongly correlated with the development level of a country, defined by United Nations health and socioeconomic development indicators. We found that subjects from developing countries had low IgG galactosylation levels, characteristic for inflammation and ageing. Our results suggest that citizens of developing countries may be exposed to country-specific environmental factors that can cause low-grade chronic inflammation and the apparent increase in biological age.

Introduction

Immunoglobulin G is the most abundant glycoprotein and antibody class in human plasma[1]. It mediates interactions between antigens and the immune system[2]. There are four IgG subclasses present in plasma: IgG1, IgG2, IgG3 and IgG4[3]. Each subclass has distinctive functions, such as pronounced affinity for certain types of antigens, formation of immune complexes, complement activation, interactions with effector cells, half-life and placental transport[2].

Glycosylation is co- and post-translational modification which is orchestrated by a complex biosynthetic pathway[4]. IgG contains a conserved N-glycosylation site on Asn297 residue within its fragment crystallisable (Fc) region on each of the two heavy chains[5]. Glycans attached to IgG are a complex biantennary type, with core structure consisting of four *N*-acetylglucosamines and three mannoses. Different glycan moieties such as bisecting GlcNAc, galactose, sialic acid and fucose can be attached to this core[6]. IgG shows a high degree of glycosylation diversity. Each of four IgG subclasses displays a distinctive glycan profile[7]. Also, each of the heavy chains of the same molecule can carry different glycans, creating a large repertoire of possible glycan patterns[8]. Finally, in 15-20% of cases, an additional N-glycosylation site appears within a variable region of the antibody, as a result of sequence variation in the variable region[9].

The majority of IgG functions are achieved through interactions with receptors on immune cells and complement proteins. Fc glycans affect immunoglobulin conformation, which, in turn, defines binding affinity for Fc gamma receptors (FcγRs) on effector cells and complement, leading to alternations in effector functions[1,10]. IgG galactosylation has an extensive effect on its inflammatory potential[11]. Namely, agalactosylated IgG increases inflammation through activation of complement system[12,13]. Moreover, galactosylation was

found necessary for C1q complement component binding and activation of complement-dependent cytotoxicity (CDC)[14]. It was also required for increased binding of IgG to activating Fc gamma receptors and therefore activation of antibody-dependent cellular cytotoxicity (ADCC)[15]. On the other hand, presence of galactose is necessary for activation of anti-inflammatory cascade through interactions with FcγRIIB and inhibition of the inflammatory activity of C5a complement component[16,17]. Presence of fucose attached to the first N-acetylglucosamine, i.e. core fucose, decreases ADCC activity, while the presence of bisecting GlcNAc increases binding to activating Fcγ receptors[18]. Terminal sialic acid increases anti-inflammatory roles of IgG by decreasing ADCC. Sialylated IgG molecules are also recognized by lectin receptors and complement component C1q, but proposed mechanisms underlying anti-inflammatory functions initiated through these interactions are still subject of debates[19].

There is a prominent inter-individual variability of total IgG N-glycome[20]. Average IgG glycome heritability is approximately 50%, while the remaining variability can be mostly attributed to environmental factors[20,21,22]. The composition of IgG N-glycome gradually changes over the lifetime. It is strongly influenced by age, sex hormones and lifestyle[23]. Prominent changes in the IgG glycome were found in a number of diseases. In different autoimmune and alloimmune disorders, cancers and infectious diseases, IgG glycosylation changes reflect the increase inflammation, which is accompanying these conditions[11]. The impact of IgG glycosylation on its ability to modulate inflammation has been extensively studied as a potential biomarker for disease prognosis and therapy response, as well as for monoclonal antibody development[24,25].

Ageing is a process of damage aggregation in an organism, leading to the disruption of health. It is influenced by both genetic factors and environment/lifestyle. In a healthy individual, a gender-specific gradual change in IgG glycosylation can be observed with an

113 increase of chronological age. Namely, digalactosylated structures decrease with age, while
114 agalactosylation and bisecting GlcNAc increase in older individuals[23]. On the other hand,
115 changes in sialylation and core fucosylation displayed inconsistent trends in different
116 studies[23]. Proposed model to describe and explain age-related changes in glycosylation is "the
117 inflammageing model", which implicates that inflammation causes changes in IgG
118 glycosylation which, in turn, accelerate the process of ageing[26]. Consequently, changing
119 under influence of both genes and environment, IgG glycans predict biological age and
120 represent a measure of organisms health[11].

121 Despite the fact that structural and functional aspects of IgG glycosylation are intensively
122 studies and associated with predisposition and course of different diseases, little is known
123 about the regulation of IgG glycosylation or mechanisms that lead to extensive changes in
124 glycome composition after environmental challenge[11,27,28]. Therefore, the aim of this study
125 was to estimate and compare various IgG N-glycosylation patterns in populations across the
126 world as a result of their different genetic backgrounds and specific environmental
127 influences.

Results

Total IgG glycans change with chronological age

In the initial analysis, samples originating from 10,482 healthy individuals and 5 different populations were analysed. Fluorescently labelled N-glycans released from IgG were separated into 24 chromatographic peaks (Supplementary Table 1; Supplementary Figure 1). Additionally, derived glycan traits (galactosylation, core fucosylation, sialylation and presence of bisecting GlcNAc) were calculated, based on the initial 24 glycan measures. Derived glycan traits represent a portion of structurally similar glycan species in the total IgG glycome (Supplementary Table 2). In general, galactosylation showed the highest variability of all IgG glycan traits, which is in line with the previous studies (Supplementary Figure 2).

It is known that the chronological age of subject affects IgG glycosylation[23]. Age-related changes were observed in the levels of IgG glycan traits in all studied populations (Figure 1a). Agalactosylated species and glycans containing bisecting *N*-acetylglucosamine (GlcNAc) increased with the chronological age of the participant. The opposite trend was observed for the levels of digalactosylated and sialylated glycans, which were decreasing with chronological age. On the other hand, core fucosylation and monogalactosylation levels did not change consistently with age. Also, age-related changes in glycosylation displayed sex-specific patterns, where female participants displayed more rapid changes.

Age and country of residence can predict total IgG glycosylation

Although total IgG N-glycans showed similar age-related changes within each of the studied cohorts, every population displayed particular glycan patterns (Figure 1b). Again, the most pronounced differences between populations were observed in the levels of agalactosylated glycans, which increased with a median age of the analysed population. This glycan trait had the lowest median value in young Chinese cohort (20%), while the highest was observed in

Estonian cohort (36%), which contained the oldest population. Besides agalactosylation, pronounced differences between populations were also observed in the levels of digalactosylated and sialylated glycans (Supplementary Table 3). Relations of age, country of residence and sex with the total IgG glycans were evaluated to further investigate changes in IgG glycan traits in different populations. Chronological age appeared to be a good predictor of digalactosylation and agalactosylation variability (explaining 30% and 31% respectively), while it was able to describe 20% of bisecting GlcNAc variability. Contrary to age, participant's country of residence appeared to be the strongest predictor of core fucose levels ($P<6\times10^{-350}$, n=5), explaining 57% of the variability in this glycan trait. It was also a good predictor of monogalactosylation and sialylation variability. On the other hand, sex was able to explain less than 1% of the variability of any glycan trait (Supplementary Table 4).

Fc IgG glycan patterns in 14 countries

To validate observed diversity and unambiguously determine IgG N-glycosylation patterns in different populations, while eliminating potential batch effects, we compared glycan features derived from IgG subclass-specific Fc glycopeptides from 2,530 individuals (Supplementary Table 5, Supplementary Figure 3). This part of the study included 27 populations collected in 14 different countries (Supplementary Table 6). Subclass-specific glycopeptides were separated and accurate masses were measured for each glycoform. Calculated IgG Fc N-glycan derived traits displayed considerable variability between analysed populations. The Fc N-glycome composition is known to differ from the total IgG N-glycome, as a result of Fab N-glycome contribution to the total IgG glycome[29]. Again, the most prominent variation appeared to be in level of IgG1 galactosylation (Figure 2), although expected decrease of this glycan trait with the age of population was not observed. On the contrary, some populations appeared to have lower galactosylation than expected for the given chronological age. Population from Papua New Guinea, as the youngest one, surprisingly had the highest

177 median level of agalactosylation (45%), while the subjects from England exhibited the lowest
178 levels of this glycan trait (28%) on IgG1 subclass. The opposite effect was observed for
179 monogalactosylation levels - the subjects from Papua New Guinea had the lowest median
180 value of this glycan trait, while the highest levels were observed in the participants from
181 England. In a similar manner, participants from countries such as Germany and Italy had
182 higher monogalactosylation levels (comparable to subjects from England) than the ones from
183 countries such as Uganda, which were more similar to the subjects from Papua New Guinea
184 (Supplementary Table 7).

185 In the case of IgG2 and IgG4 subclasses, galactosylation levels displayed similar variation to
186 IgG1 subclass, although observed glycosylation patterns appeared to be subclass-specific,
187 especially in case of IgG4, which is the least abundant subclass in the human plasma
188 (Supplementary Tables 8 and 9).

189 Age and county of residence can predict IgG Fc glycosylation

190 To determine the source of variation in IgG Fc glycan profiles, we analysed relationship of
191 glycan traits with sex, chronological age and country of residence. Here as well,
192 chronological age was a predictor of agalactosylation and digalactosylation variability (Table
193 1). Interestingly, age was the best predictor of IgG2 digalactosylation. It was able to explain
194 28% of IgG2 agalactosylation variability compared to 22% in the case of IgG1.

195 Country of residence was the best predictor of IgG Fc monogalactosylation variability (Table
196 1). Namely, 38% of IgG1 Fc monogalactosylation variability was explained by the subject's
197 country of residence. Similar patterns were observed for IgG2 and IgG4 Fc glycosylation.
198 The same as in the case of total IgG glycans, sex was able to explain up to 1% of the IgG Fc
199 glycan variability. Therefore, chronological age and country of residence are good predictors
200 of IgG Fc glycosylation.

IgG Fc galactosylation correlates with the development level of a country

Development indicators measure the quality of specific life aspects. In order to resolve whether the observed associations between country of residence and studied IgG Fc glycan traits can be attributed to the development level of the country, we analysed relations between 45 development indicators and 5 derived glycan traits of each analysed IgG subclass (Supplementary Tables 10 - 12). The analysis resulted with 44 statistically significant correlations of IgG Fc monogalactosylation, digalactosylation and agalactosylation with 23 different development indicators. Majority of development indicators displayed significant positive correlations with IgG1 monogalactosylation (Supplementary Table 13). As for the subclasses IgG2 and IgG4, only monogalactosylation appeared to be significantly correlated with the studied indicators. On the other hand, we did not observe any significant correlations between any of the development indicators and sialylation or the incidence of bisecting GlcNAc on any of the IgG subclasses.

United Nation's Human development index (HDI) is a summary measure of the development level of a certain country. It represents three dimensions of life: economy, education and health quality. Among the studied populations, Western European nations (Germany, England, Scotland, Sweden) have the highest development level expressed through HDI, while Papua New Guinea and Uganda have the lowest HDI scores. We found a positive correlation between HDI and IgG1 Fc monogalactosylation (Figure 3a), while it negatively correlated with IgG1 agalactosylation. These observations replicated on IgG2 subclass, where HDI positively correlated with monogalactosylation levels. Therefore, participants from developing countries appear to have lower IgG galactosylation when compared to their counterparts from developed countries.

IgG Fc galactosylation as a marker of population's health status

To determine the impact of the quality of health on IgG glycans, correlations between the two were calculated. Population's health quality was expressed through overall health indices and specific health indicators. Countries with lower development level, in general, had also lower health-related indicators (Supplementary Table 12). Majority of health-related indicators appeared to be correlated with IgG monogalactosylation (Supplementary Table 13). Millennium development goals (MDG) index, which describes health-related indicators in MDG system, positively correlated with IgG1 and IgG2 monogalactosylation (r=0.97, $P=7.44\times10^{-6}$ and r=0.86, $P=4.59\times10^{-2}$ respectively) and IgG1 agalactosylation (r=-0.90, $P=8.16\times10^{-3}$; Figure 3a). In a similar fashion, positive correlation with IgG1 monogalactosylation displayed also the sustainable development goals (SDG) index, non-MDG index (health-related SDG indicators not included in MDG) and Health index, which like MDG index display overall health quality of a specific country (Supplementary Table 13, Figure 3a). SDG index was negatively correlated with IgG1 agalactosylation (Figure 3b).

Besides health-related indices, specific health-related indicators also correlated with IgG Fc galactosylation. Among all studied specific indicators, the decline in stunted growth prevalence demonstrated the strongest positive correlation with IgG1 monogalactosylation (r=0.97, $P=1.16\times10^{-5}$; $n=14$). Of the other studied indicators, universal health coverage and the decrease in occupational risk burden displayed substantial correlations with IgG Fc galactosylation. Life expectancy is also one of the most important indicators, used to describe life quality. Both female and male life expectancies were correlated with IgG Fc monogalactosylation. Exposure to various antigens was presented through indicators such as hygiene, water, WasH mortality and sanitation, which also displayed correlations with IgG Fc galactosylation. Of the infectious diseases, only hepatitis B showed a significant correlation with IgG Fc monogalactosylation (Supplementary Table 13).

Moreover, digalactosylation of IgG1 demonstrated five positive correlations with health-related development indicators, where skilled birth attendance and again, stunted growth, had the strongest associations with this glycan trait (Supplementary Table 13).

Although IgG Fc glycans showed the highest correlation coefficients with the health-related indicators, significant correlations between IgG1 Fc galactosylation features and the socioeconomic indicators such as education and economic development have also been determined. Education index was significantly correlated with both Fc monogalactosylation (r=0.94, P=0.0005, n=14) and agalactosylation (r=-0.90, P=0.010, n=14), while Gross Domestic Product (GDP) was significantly correlated only with monogalactosylation (r=0.89, P=0.01, n=14; Supplementary Table 13).

Discussion

IgG N-glycosylation varies between individuals within the same population as well as between different populations[20,30]. In this study, we compared glycan profiles of the whole IgG molecule in 10 482 subjects originating from 5 different populations and Fc glycan profiles of 2 530 subjects from 27 cohorts and 25 ethnicities. Furthermore, this study yielded valuable data on IgG glycan levels in healthy participants from 14 countries and 25 ethnicities. The observed changes in glycome composition of the analysed populations suggest country-specificity of IgG glycan profiles.

Besides genetics, environment plays a crucial role in IgG glycosylation[28,22]. Pathogens, stress and nutrition are possible players orchestrating non-genomic component of the inter-populational variation in IgG glycosylation patterns[21,31]. Development level of a country reflects human well-being in a specific community and thereby environmental impact on an individual. We found that the development level of a country of residence was positively correlated with IgG galactosylation level. Namely, IgG galactosylation was decreased in people from countries with lower development level, while people from highly developed countries had also the highest levels of IgG galactosylation. These associations were observed on all IgG subclasses, although the largest number of development indicators correlated with glycans originating from IgG1 subclass, probably due to the highest concentration of this subclass in plasma.

Besides overall development level of a country, different health-related indicators displayed associations with IgG galactosylation level, indicating the impact of health quality on IgG glycosylation. IgG glycosylation modulates antibody's pro- and anti-inflammatory actions, where a decline in galactosylation level, similar to the one observed in underdeveloped countries, was found in several inflammatory and autoimmune diseases, such as

283 inflammatory bowel disease (IBD), rheumatoid arthritis and systemic lupus
284 erythematosus[24,28,32]. Therefore, we speculated that populations with lower galactosylation
285 have higher inflammatory potential of IgG and a higher low-grade systematic inflammation.
286 Our findings are supported by a recent study on 773 children, which compared IgG
287 glycosylation in subjects from Gabon, Ghana, Ecuador, the Netherlands and Germany. The
288 increase in agalactosylated species was observed in individuals from Gabon, Ghana and
289 Ecuador, compared to participants from Netherlands and Germany. These changes were
290 correlated with the history of parasitic infections and generalised to immune activation[30].

291 In our study, the prevalence of stunted growth displayed a strong correlation with
292 glycosylation. This may be partially explained by the fact that stunted growth is caused by
293 environmental enteropathy, which is a chronic intestinal inflammation caused by
294 malnutrition, continuous bacterial exposure, repeated enteric infections and small intestinal
295 bacterial overgrowth[33]. Since many inflammatory conditions are associated with decreased
296 IgG galactosylation levels, environmental enteropathy could cause changes in IgG
297 galactosylation by inducing chronic subclinical inflammation in the gastrointestinal tract[34].
298 Hence, our results imply that country-specific differences in the quality of health influence
299 IgG inflammatory potential and cause specific glycosylation patterns in different countries.

300 Furthermore, we also observed significant correlations between the traits describing IgG Fc
301 galactosylation and the socioeconomic indicators. Economy and education quality are
302 reflected through country's development level and quality of the health system[35]. For that
303 reason, our findings are not surprising and further emphasize environmental influence on the
304 health of the certain population.

305 IgG glycosylation is known to change with chronological age of individual[23]. Within five
306 populations where the whole IgG glycan profiles were measured, we observed similar

307 decrease in galactosylation levels with the age of participant. This age-related decline in IgG
308 galactosylation is believed to be one of the causes of higher inflammation in older
309 individuals[23,36]. Although exact mechanisms are still unclear, there are several proposed
310 pathways which could explain underlying age-related changes in IgG glycosylation. Possible
311 mechanisms include various expression and/or activity of glycosylation-related enzymes,
312 selection of B-cell clones with specific IgG glycan patterns and B-cell independent
313 glycosylation. Furthermore, a decrease in IgG galactosylation was observed in the premature
314 ageing syndromes[36]. Through modulation of inflammation, IgG galactosylation, or, more
315 precisely, agalactosylation is proposed to contribute to biological ageing in a process of
316 inflammaging[37]. Since the decrease in IgG galactosylation is a hallmark of increased
317 biological age, the proinflammatory IgG Fc glycosylation profile in individuals from
318 developing countries may imply accelerated biological ageing in these populations, resulting
319 in a shorter expected lifespan.

320 In summary, we revealed that at a community level, immunoglobulin G glycosylation
321 patterns vary between different countries. We also correlated changes in galactosylation with
322 participant's chronological age and development level of a country. Constant environmental
323 pressure on the immune system in developing countries maintains IgG constantly in under-
324 galactosylated, proinflammatory state. As a consequence of this permanent low-degree IgG
325 Fc galactosylation, individuals from developing countries display premature populational
326 ageing and appear to be biologically "older" than residents of more developed countries.

Materials and methods

Study participants

Total IgG glycome analysis was based on 10,482 human participants from China[38,28], Croatia[39], Estonia and two cohorts from the United Kingdom (population of Scotland from Orkney Islands[20] and England from the TwinsUK cohort[40]) (Supplementary Table 1). Subclass specific analysis of IgG Fc glycosylation included 2,530 healthy individuals. Volunteers originated from 14 different countries and 25 different ethnic groups (Supplementary Table 6). For Kazak and English cohorts, we had two populations obtained from different medical centres. Samples were randomized across 96-well plates (31 in total), with 5 technical replicates of a standard sample and 1 blank, serving as a negative control. Development level of a country was assessed using three development indices: health-related Sustainable Development Goal index (SDG)[41], health-related Millennium Development Goal Index (MDG)[41] and United Nation Human Development Index (HDI)[42], while specific aspects of human life were assessed using other development indicators (Supplementary Table 13). Blocking was performed by equally distributing subjects of the same sex and age from all the cohorts across used plates. Plasma samples used as standards were obtained from Croatian National Institute of Transfusion Medicine. Study was performed in compliance with the Helsinki declaration and all participants gave written informed consent. Ethical approval was obtained by relevant ethics committees.

Immunoglobulin G isolation

Protein G affinity chromatography was used to isolate immunoglobulin G from human blood plasma as described previously[20]. In short, maximum volume of 100 µL of human peripheral blood plasma or serum were diluted with 1X phosphate buffer saline (PBS) and loaded onto protein G monolithic plate (BIA Separations, Ajovščina, Slovenia). Samples were washed

three times with 1X PBS and IgG was eluted using 0.1M formic acid (Merck, Darmstadt, Germany) followed by immediate neutralisation with 1M ammonium bicarbonate (Acros Organics, Pittsburgh, PA).

Immunoglobulin G trypsin digestion and purification

IgG glycopeptides were obtained and purified as described before[43]. Approximately 15 μg of isolated IgG was treated with 0.1 μg of sequencing grade trypsin (Promega, Fitchburg, WI) and incubated overnight at 37 °C. Reaction was stopped by dilution with 0.1% trifluoroacetic acid (TFA; Sigma-Aldrich, St. Louis, MI). Glycopeptides were purified using a solid phase extraction on Chromabond C-18 sorbent (Macherey-Nagel, Düren, Germany). Samples were loaded onto beads in 0.1% TFA, and washed three times using the same solvent. Glycopeptides were eluted from the phase with 20% LC-MS grade acetonitrile (ACN; Honeywell, Morris Plains, NJ). Eluted glycopeptides were vacuum-dried and reconstituted in 20 μL of ultrapure water prior to LC-MS analysis. All glycan analyses were performed at Genos laboratory.

Release and labelling of the total IgG N-glycans

Glycan release and labelling of Croatian samples was performed as previously described[23]. Briefly, IgG was incorporated into sodium dodecyl sulphate polyacrylamide gel and glycans were released from protein using an overnight incubation with PNGase F (ProZyme, Hayward, CA). Released glycans were labelled with 2-aminobenzamide (2-AB; Sigma-Aldrich) and purified on Whatman 3 mm chromatography paper. For cohorts from Scotland, England, China and Estonia, glycans were released as previously described[23,31]. Briefly, IgG was denatured using 1.33% (w/v) sodium dodecyl sulphate (Invitrogen, Carlsbad, CA) and samples were incubated at 65 °C for 10 minutes. Subsequently, 4% (v/v) Igepal CA-630 (Sigma–Aldrich) and 1.25 mU of PNGase F (ProZyme) were added to each sample and

375 incubated overnight at 37 °C. For glycan labelling, 48 mg/mL of 2-AB in dimethyl sulfoxide
376 (Sigma–Aldrich) and glacial acetic acid (Merck) (v/v 85:15) was mixed with reducing agent
377 (106.96 mg/mL of 2-picoline borane (Sigma–Aldrich) in dimethyl sulfoxide). Labelling
378 mixture was added to samples, followed by 2-hour incubation at 65 °C.

379 After incubation, Estonian and Chinese samples were brought to 96% ACN (J.T. Baker,
380 Phillipsburg, NJ) and applied to each well of a 0.2 μm GHP filter plate (Pall Corporation,
381 Ann Arbor, MI). Samples were subsequently washed five times using acetonitrile/water
382 (96:4, v/v). Glycans were eluted with water and stored at -20 °C until usage. Samples from
383 England and Scotland were purified using a solid-phase extraction on 200 μL of 0.1 g/L
384 microcrystalline cellulose suspension (Merck) in a 0.45 μm GHP filter plate (Pall
385 Corporation). Deglycosylation reaction was diluted four times with ACN loaded to cellulose.
386 Samples were washed three times with 80% ACN and eluted with ultrapure water.

387 HILIC-UPLC analysis of fluorescently labelled N-glycans

388 Fluorescently labelled N-glycans were separated by hydrophilic interaction liquid
389 chromatography (HILIC) on a Waters Acquity UPLC H-class instrument (Waters, Milford,
390 MA) equipped with FLR fluorescence detector set to 330 nm for excitation and 420 nm for
391 emission wavelength. Separation was achieved on a Waters ethylene bridged hybrid (BEH)
392 Glycan chromatography column, 100 × 2.1 mm i.d., 1.7 μm BEH particles with 100 mM
393 ammonium formate (pH 4.4) as a solvent A and ACN as a solvent B. Separation method used
394 linear gradient from 75% to 62% solvent B (v/v) at a flow rate of 0.4 mL/min in a 25-minute
395 analytical run. Column temperature was maintained at 60 °C. Obtained chromatograms were
396 manually separated into 24 peaks from which, using the total area normalisation, relative
397 abundances of 24 directly measured glycan traits were obtained.

398 LC-MS analysis of IgG Fc glycopeptides

Trypsin-digested, subclass-specific glycopeptides were separated and measured on nanoAcquity chromatographic system (Waters, Milford, MA) coupled to Compact mass spectrometer (Bruker, Bremen, Germany), equipped with Apollo II source as described previously with minor changes[44]. Samples (9 µL) were loaded onto PepMap 100 C8 trap column (5 mm × 300 µm i.d.; Thermo Fisher Scientific, Waltham, MA) at a flow rate of 40 µL/min of solvent A (0.1% TFA) and washed of salts and impurities for one minute. Subclass-specific glycopeptides were separated on C18 analytical column (150 mm × 100 µm i.d., 100 Å; Advanced Materials Technology, Wilmington, DE) in a gradient from 18% to 25% of solvent B (80% ACN) in solvent A. Column temperature was set to 30 °C and flow rate was 1 µl/min. NanoAcquity was coupled to mass spectrometer via capillary electrophoresis sprayer interface (Agilent, Santa Clara, CA), which allows mixing of analytical flow with sheath liquid (50% isopropanol, 20% propionic acid; Honeywell, Morris Plains, NJ).

Mass spectrometer was operated in a positive ion mode, with capillary voltage set to 4500 V, nebulizer pressure set to 0.4 bar and drying gas set to 4 l/min at 180 °C. Spectra were recorded in a *m/z* range of 600 - 1800. Collision energy was 4 eV.

Obtained raw data was converted to centroid mzXML files using ProteoWizard version 3.0.1. software. Samples were internally calibrated using defined list of IgG glycopeptides with highest signal-to-noise ratios and required isotopic patterns. After calibration, signals matching IgG Fc glycopeptides were extracted from data using 10 *m/z* extraction window. First four isotopic peaks of doubly and triply charged signals, belonging to the same glycopeptide species, were summed together, resulting in 20 glycopeptides per IgG subclass. Predominant allotype variant of IgG3 tryptic peptide carrying N-glycans in Caucasian population has the same amino acid sequence as IgG2. On the other hand, in Asian and African populations predominant variant of the same peptide has the same amino acid

424 composition as IgG4 making the separation of IgG3 from other subclasses impossible using
425 given separation methods[45]. Therefore, IgG glycopeptides were separated into three
426 chromatographic peaks labeled IgG1, IgG2 and IgG4. Signals of interest were normalised to
427 the total area of each IgG subclass.

428 Statistical analysis

429 Data analysis was performed using program R, version 3.0.1. with a ggplot2 package for
430 creation of visualisations. Derived glycan traits representing levels of galactosylation
431 (agalactosylation, monogalactosylation and digalactosylation), sialylation, core fucosylation
432 and incidence of bisecting GlcNAc were calculated from obtained data as described before
433 (Supplementary Tables 2 and 5)[43]. In short, traits were calculated as portion of glycans
434 containing common structural features in a total IgG glycome, or in a specific subclass of IgG
435 Fc glycopeptides. Core fucosylation was excluded from IgG Fc specific glycopeptide analysis
436 due to low data quality of non-fucosylated species. To remove experimental and batch
437 variation, batch correction was performed using a R ComBat package followed by log-
438 transformation of glycan or glycopeptide data. Linear mixed model was used to analyse
439 correlations between glycan traits and the subject's country of residence. In the model, sex
440 and age were fixed effects, while the country of residence was used as a random effect.
441 Likelihood ratio test was used to determine significance of country of origin variability in
442 glycan trait variability. Pearson's correlation coefficient was used to express relationships
443 between country-specific development indicator and level of glycan trait in participants from
444 the same country. *P* values were adjusted for multiple testing using Bonferroni correction.

445 Data availability

446 The data that support the findings of this study are available from the corresponding author
447 upon reasonable request.

448 References

449 1. Vidarsson, G., Dekkers, G. & Rispens, T. IgG subclasses and allotypes: From structure
450 to effector functions. *Front. Immunol.* **5**, 520 (2014).

451 2. Rispens, T. & Vidarsson, G. Human IgG Subclasses. *Antib. Fc Link. Adapt. Innate*
452 *Immun.* 159–177 doi:10.1016/B978-0-12-394802-1.00009-1 (2013).

453 3. Schur, P. H. IgG subclasses. A historical perspective. *Monogr. Allergy* **23**, 1–11
454 (1988).

455 4. Varki, A. & Gagneux, P. *Biological Functions of Glycans*. *Essentials of Glycobiology*
456 doi:10.1101/GLYCOBIOLOGY.3E.007 (Cold Spring Harbor Laboratory Press, 2015).

457 5. Arnold, J. N., Wormald, M. R., Sim, R. B., Rudd, P. M. & Dwek, R. A. The Impact of
458 Glycosylation on the Biological Function and Structure of Human Immunoglobulins.
459 *Annu. Rev. Immunol.* **25**, 21–50 (2007).

460 6. Kiyoshi, M., Tsumoto, K., Ishii-Watabe, A. & Caaveiro, J. M. M. Glycosylation of
461 IgG-Fc: a molecular perspective. *Int. Immunol.* **29**, 311–317 (2017).

462 7. Wuhrer, M. *et al.* Glycosylation profiling of immunoglobulin G (IgG) subclasses from
463 human serum. *Proteomics* **7**, 4070–4081 (2007).

464 8. Schroeder, H. W., Cavacini, L. & Cavacini, L. Structure and function of
465 immunoglobulins. *J. Allergy Clin. Immunol.* **125**, S41-52 (2010).

466 9. Potter, M. Structural correlates of immunoglobulin diversity. *Surv. Immunol. Res.* **2**,
467 27–42 (1983).

468 10. Pincetic, A. *et al.* Type i and type II Fc receptors regulate innate and adaptive
469 immunity. *Nat. Immunol.* **15**, 707–716 (2014).

11. Gudelj, I., Lauc, G. & Pezer, M. Immunoglobulin G glycosylation in aging and diseases. *Cellular Immunology* doi:10.1016/j.cellimm.2018.07.009 (2018).

12. Banda, N. K. *et al.* Initiation of the alternative pathway of murine complement by immune complexes is dependent on *N* -glycans in IgG antibodies. *Arthritis Rheum.* **58**, 3081–3089 (2008).

13. Malhotra, R. *et al.* Glycosylation changes of IgG associated with rheumatooid arthritis can activate complement via the mannose-binding protein. *Nat. Med.* **1**, 237–243 (1995).

14. Peschke, B., Keller, C. W., Weber, P., Quast, I. & Lünemann, J. D. Fc-Galactosylation of Human Immunoglobulin Gamma Isotypes Improves C1q Binding and Enhances Complement-Dependent Cytotoxicity. *Front. Immunol.* **8**, 646 (2017).

15. Thomann, M., Reckermann, K., Reusch, D., Prasser, J. & Tejada, M. L. Fc-galactosylation modulates antibody-dependent cellular cytotoxicity of therapeutic antibodies. *Mol. Immunol.* **73**, 69–75 (2016).

16. Karsten, C. M. *et al.* Anti-inflammatory activity of IgG1 mediated by Fc galactosylation and association of FcγRIIB and dectin-1. *Nat. Med.* **18**, 1401–1406 (2012).

17. Mihai, S. & Nimmerjahn, F. The role of Fc receptors and complement in autoimmunity. *Autoimmun. Rev.* **12**, 657–660 (2013).

18. Kurimoto, A. *et al.* The absence of core fucose up-regulates GnT-III and Wnt target genes: a possible mechanism for an adaptive response in terms of glycan function. *J. Biol. Chem.* **289**, 11704–14 (2014).

19. Jennewein, M. F. & Alter, G. The Immunoregulatory Roles of Antibody

493 Glycosylation. *Trends in Immunology* **38**, 358–372 (2017).

494 20. Pučić, M. *et al.* High Throughput Isolation and Glycosylation Analysis of IgG–
495 Variability and Heritability of the IgG Glycome in Three Isolated Human Populations.
496 *Mol. Cell. Proteomics* **10**, M111.010090 (2011).

497 21. Plomp, R. *et al.* Subclass-specific IgG glycosylation is associated with markers of
498 inflammation and metabolic health. *Sci. Rep.* **7**, (2017).

499 22. Nikolac Perkovic, M. *et al.* The association between galactosylation of
500 immunoglobulin G and body mass index. *Prog. Neuro-Psychopharmacology Biol.*
501 *Psychiatry* **48**, 20–25 (2014).

502 23. Krištić, J. *et al.* Glycans are a novel biomarker of chronological and biological ages.
503 *Journals Gerontol. - Ser. A Biol. Sci. Med. Sci.* **69**, 779–789 (2014).

504 24. Šimurina, M. *et al.* Glycosylation of Immunoglobulin G Associates With Clinical
505 Features of Inflammatory Bowel Diseases. *Gastroenterology* **154**, 1320–1333.e10
506 (2018).

507 25. Sha, S., Agarabi, C., Brorson, K., Lee, D. Y. & Yoon, S. N-Glycosylation Design and
508 Control of Therapeutic Monoclonal Antibodies. *Trends Biotechnol.* **34**, 835–846
509 (2016).

510 26. Prattichizzo, F. *et al.* Inflammageing and metaflammation: The yin and yang of type 2
511 diabetes. *Ageing Res. Rev.* **41**, 1–17 (2018).

512 27. Lauc, G., Pezer, M., Rudan, I. & Campbell, H. Mechanisms of disease: The human N-
513 glycome. *Biochim. Biophys. Acta - Gen. Subj.* **1860**, 1574–1582 (2016).

514 28. Vučković, F. *et al.* Association of systemic lupus erythematosus with decreased
515 immunosuppressive potential of the IgG glycome. *Arthritis Rheumatol. (Hoboken,*

516 *N.J.)* **67**, 2978–89 (2015).

517 29. Huffman, J. E. *et al.* Comparative Performance of Four Methods for High-throughput
518 Glycosylation Analysis of Immunoglobulin G in Genetic and Epidemiological
519 Research. *Mol. Cell. Proteomics* **13**, 1598–1610 (2014).

520 30. De Jong, S. E. *et al.* IgG1 Fc N-glycan galactosylation as a biomarker for immune
521 activation. *Sci. Rep.* **6**, 28207 (2016).

522 31. Menni, C. *et al.* Glycosylation of immunoglobulin g: role of genetic and epigenetic
523 influences. *PLoS One* **8**, e82558 (2013).

524 32. Alavi, A., Arden, N., Spector, T. D. & Axford, J. S. Immunoglobulin G glycosylation
525 and clinical outcome in rheumatoid arthritis during pregnancy. *J. Rheumatol.* **27**,
526 1379–85 (2000).

527 33. Vonaesch, P. *et al.* Stunted childhood growth is associated with
528 decompartmentalization of the gastrointestinal tract and overgrowth of oropharyngeal
529 taxa. *Proc. Natl. Acad. Sci.* 201806573 doi:10.1073/pnas.1806573115 (2018).

530 34. Dall'Olio, F. *et al.* N-glycomic biomarkers of biological aging and longevity: A link
531 with inflammaging. *Ageing Res. Rev.* **12**, 685–698 (2013).

532 35. Calixto, O. J. & Anaya, J. M. Socioeconomic status. The relationship with health and
533 autoimmune diseases. *Autoimmun. Rev.* **13**, 641–654 (2014).

534 36. Vanhooren, V. *et al.* Serum N-glycan profile shift during human ageing. *Exp.*
535 *Gerontol.* **45**, 738–743 (2010).

536 37. Ferrucci, L. & Fabbri, E. Inflammageing: chronic inflammation in ageing,
537 cardiovascular disease, and frailty. *Nature Reviews Cardiology* **15**, 505–522 (2018).

538 38. Wang, Y. *et al.* The Association between Glycosylation of Immunoglobulin G and
539 Hypertension. *Med. (United States)* **95**, e3379 (2016).

540 39. Rudan, I. *et al.* "10□001 Dalmatians:" Croatia Launches Its National Biobank. *Croat.*
541 *Med. J.* **50**, 4–6 (2009).

542 40. Moayyeri, A., Hammond, C. J., Valdes, A. M. & Spector, T. D. Cohort profile:
543 TwinsUK and healthy ageing twin study. *Int. J. Epidemiol.* **42**, 76–85 (2013).

544 41. Lim, S. S. *et al.* Measuring the health-related Sustainable Development Goals in 188
545 countries: a baseline analysis from the Global Burden of Disease Study 2015. *Lancet*
546 **388**, 1813–1850 (2016).

547 42. Human Development Data - Human Development Reports. Available at:
548 http://hdr.undp.org/en/data. (2017).

549 43. Keser, T. *et al.* Effects of statins on the immunoglobulin G glycome. *Biochim.*
550 *Biophys. Acta - Gen. Subj.* **1861**, 1152–1158 (2017).

551 44. Selman, M. H. J. *et al.* Fc specific IgG glycosylation profiling by robust nano-reverse
552 phase HPLC-MS using a sheath-flow ESI sprayer interface. *J. Proteomics* **75**, 1318–
553 1329 (2012).

554 45. Plomp, R., Bondt, A., de Haan, N., Rombouts, Y. & Wuhrer, M. Recent Advances in
555 Clinical Glycoproteomics of Immunoglobulins (Igs). *Mol. Cell. Proteomics* **15**, 2217–
556 2228 (2016).

557

558 Acknowledgements

559 This study was supported by European Structural and Investment Funds IRI grant
560 (#KK.01.2.1.01.0003) and Croatian National Centre of Research Excellence in Personalized
561 Healthcare grant (#KK.01.1.1.01.0010). Collection of Chinese cohorts was supported by
562 National Natural Science Foundation of China (81573215, 81773527, 81673247 and
563 81370083) and Australian-China Collaborative Grant (NH&MRC-APP1112767–NSFC
564 81561128020).

565 Author contributions

566 G.L. designed the study. J.Š., N.N., G.R. and M.N. carried out LC-MS analysis. F.V.
567 performed statistical analysis. M.P.B., I.T.A., T.K., T.P., M.Š., I.G., and M.V. performed
568 UPLC analysis. M.S., H.W., Y.W., W.W., M.P.S., T.Š.J., H.C., C.H., J.F.W., I.R., O.P., I.K.,
569 S.N., L.A.E., H.K., M.L.R., M.M., P.M., J.H., M.K., V.A., K.T., C.G., T.S., T.T., N.C., M.S.
570 and M.F. recruited participants and provided plasma samples. J.Š. drafted the manuscript. All
571 authors edited and approved the final version.

572 Conflict of interest

573 GL is founder and CEO of Genos – a private research organization that specializes in high-
574 throughput glycomic analysis and has several patents in this field. J.Š., F.V., M.P.B., G.R.,
575 I.T.A., I.G., M.V., and M.N. are employees of Genos.

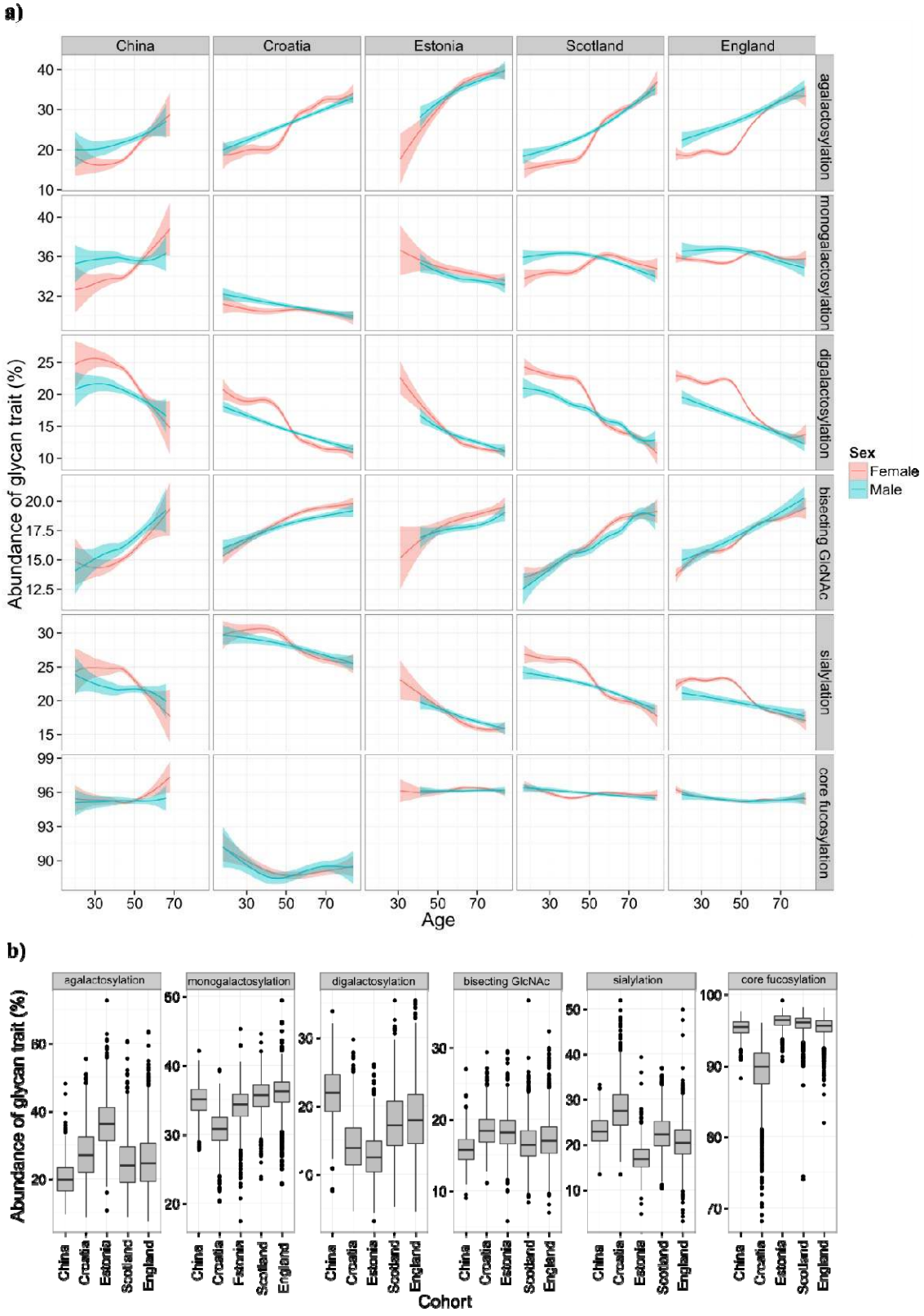
576 Correspondence and requests for materials should be addressed to G.L.

577 Table 1: Proportion of glycan feature variability in 14 countries explained by linear mixed
578 model with age and sex defined as fixed effects and country of residence as a random effect.
579 Displayed values represent percentage (%) glycan trait variability explained by country, age
580 and sex.

IgG subclass	Glycan feature	Percent of glycan trait variability explained by country of residence (%)	Percent of glycan trait variability explained by age (%)	Percent of glycan trait variability explained by sex (%)	Country of origin *P* value
IgG1	Agalactosylation	21.4	18.3	0.9	1.0×10^{-111}
	Monogalactosylation	38.0	2.6	0.1	7.7×10^{-193}
	Digalactosylation	18.6	21.7	1.1	3.5×10^{-98}
	Sialylation	10.8	13.2	0.6	1.3×10^{-54}
	Bisecting GlcNAc	18.6	17.5	0.0	9.3×10^{-110}
IgG2	Agalactosylation	12.8	23.9	0.9	7.2×10^{-62}
	Monogalactosylation	20.8	7.6	0.1	3.3×10^{-94}
	Digalactosylation	10.4	27.5	1.1	2.7×10^{-54}
	Sialylation	6.5	18.2	0.5	1.2×10^{-27}
	Bisecting GlcNAc	12.4	11.9	0.1	3.7×10^{-64}
IgG4	Agalactosylation	20.5	13.7	0.4	1.3×10^{-93}
	Monogalactosylation	20.8	2.6	0.1	1.3×10^{-78}
	Digalactosylation	18.2	14.1	0.8	1.3×10^{-86}
	Sialylation	15.4	9.7	0.5	6.4×10^{-75}
	Bisecting GlcNAc	7.6	15.7	0.6	2.0×10^{-37}

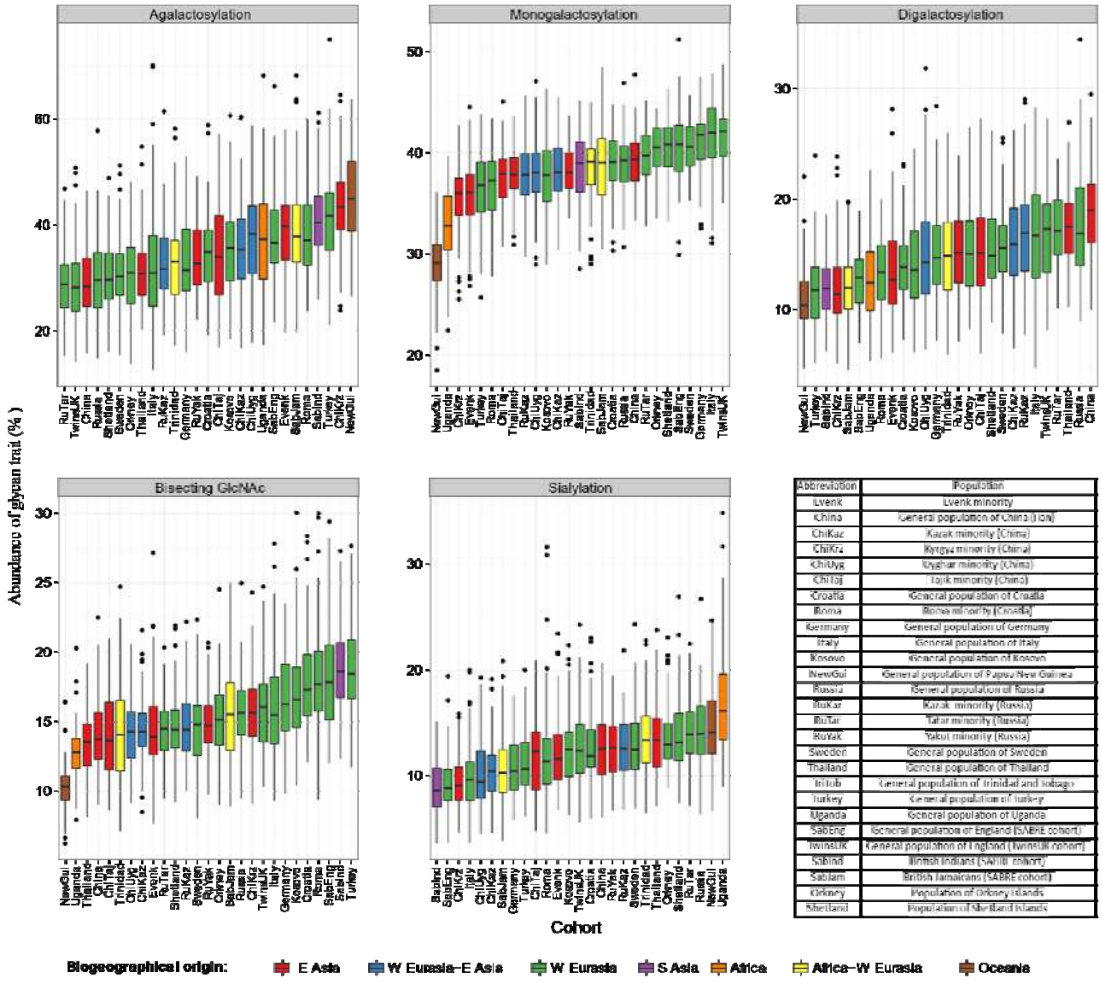
581

582 Figures



583

584 Figure 1: IgG glycan levels in five different populations. a, Relationship between age and
585 derived glycan trait. Plots describe associations between each of five glycan traits and
586 chronological age of participant. Blue and red curves represent fitted linear regression
587 models. The shaded region is the 95% confidence interval on the fitted values. b, Differences
588 in IgG glycosylation of participants from five populations. Each box represents interquartile
589 range (25^{th} to 75^{th} percentiles). Lines inside the boxes represent the median values, while
590 lines outside the boxes represent the 10^{th} and 90^{th} percentiles. Circles indicate outliers.



591

592 Figure 2: Levels of derived IgG1 glycan traits in 27 different populations. Each box

593 represents interquartile range (25th to 75th percentiles) with median values drawn as middle

594 line. Whiskers outside the boxes represent the 10th and 90th percentiles, while dots indicate

595 outliers.

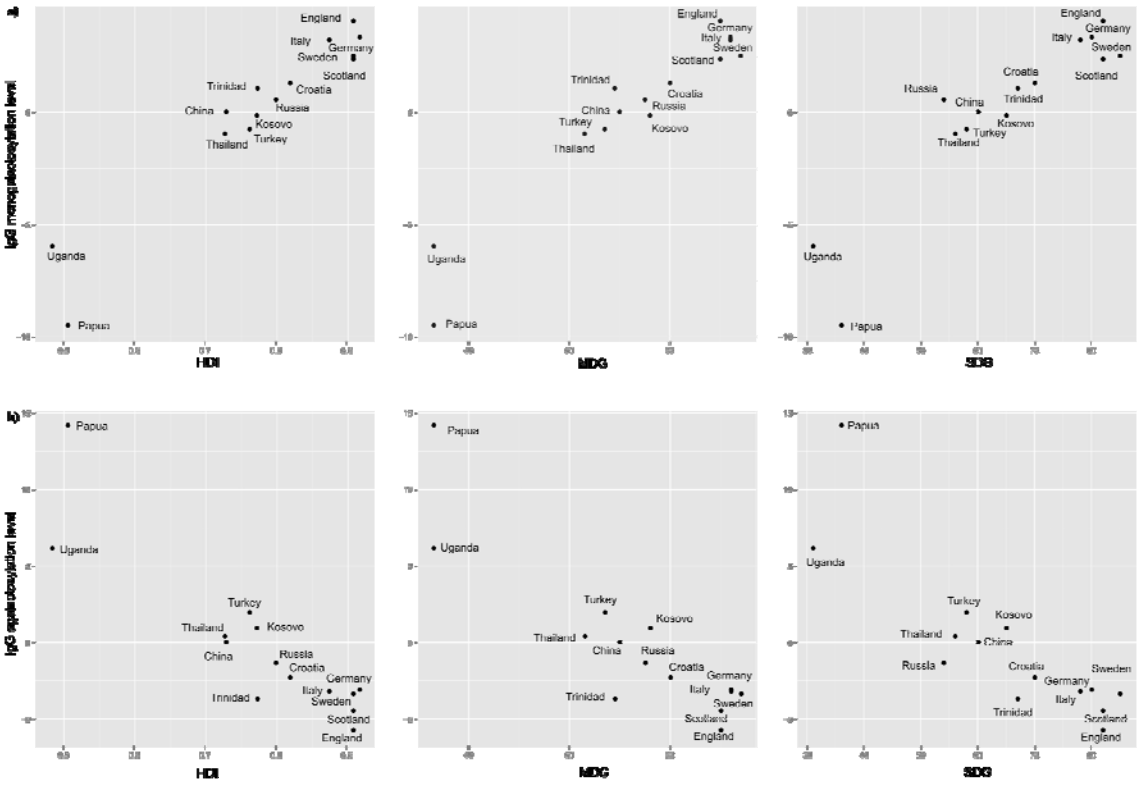


Figure 3: Relationship between IgG monogalactosylation level (a) and IgG agalactosylation level (b) with development indices in a specific country of residence. HDI = Human Development Index; SDI = health-related Sustainable Development Goals Index; MDG = health-related Millennium Development Goals Index.