IgG glycome in colorectal cancer

Immunoglobulin G glycome in colorectal cancer

Frano Vučković[1,*], Evropi Theodoratou[2,3,*], Kujtim Thaçi[1,*], Maria Timofeeva[3], Aleksandar Vojta[4], Jerko Štambuk[1], Maja Pučić-Baković[1], Pauline Rudd[5], Lovorka Đerek[6], Dražen Servis[7], Annika Wennerström[8], Susan M Farrington[3], Markus Perola[8], Yurii Aulchenko[9], Malcolm G Dunlop[3], Harry Campbell[2,3,*], Gordan Lauc[1,10,*]

1 Genos Glycoscience Research Laboratory, Zagreb, Croatia
2 Centre for Population Health Sciences, University of Edinburgh, Edinburgh, UK
3 Colon Cancer Genetics Group, Institute of Genetics and Molecular Medicine, University of Edinburgh and Medical Research Council Human Genetics Unit, Edinburgh, UK
4 University of Zagreb Faculty of Science, Zagreb, Croatia
5 National institute for bioprocessing research & training, Dublin, Ireland
6 Department of Medical Biochemistry and Laboratory Medicine, Clinical Hospital Merkur, Zagreb, Croatia
7 Allgemein- und Viszeralchirurgie, St. Anna Krankenhaus, Herne, Germany
8 Department of Health, THL, Finland
9 Polyomica, Groningen, The Netherlands
10 University of Zagreb Faculty of Pharmacy and Biochemistry, Zagreb, Croatia

*These authors contributed equally

Running title: IgG glycome in colorectal cancer

Correspondence: Professor Gordan Lauc, University of Zagreb Faculty of Pharmacy and Biochemistry, Zagreb, Croatia. glauc@pharma.hr

Grant support: Programme Grant No. C348/A12076 from Cancer Research UK, EU HighGlycans FP7 Grant, ET has a Chancellor's Fellowship from the University of Edinburgh. Glycan analysis was partly supported by European Commission GlycoBioM (contract #259869), IBD-BIOM (contract #305479), HighGlycan (contract #278535), MIMOmics (contract #305280), HTP-GlycoMet (contract #324400), IntegraLife (contract #315997) and GlyCoCan (contract #676412) grants.

Competing Interests Statement
Professor Gordan Lauc is founder and owner of Genos Ltd – a private research organization that specialises in high-throughput glycomic analysis. FV, JŠ, and MPB are employees of Genos Ltd. Yurii Aulchenko is a founder and owner of a private research organization, Maatschap PolyOmica, which specializes in omics data analysis.

Author contributions
Study design – HC, GL, ET,
Sample provision – HC, MGD, SMF, ET, LĐ, DS
Scientific analyses – FV, KT, JS, MPB, YA, MP, GL, AV
Writing of the manuscript – FV, ET, HC, GL

Statement of translational relevance

The observed differences in IgG glycome composition between patients and controls could potentially have relevance as a biomarker in the general population, especially if incorporated with other variables into risk profiling algorithms, have relevance to surveillance in subjects at high risk due to germline mutations, such as in Lynch Syndrome. Our findings also suggest that genetic factors influencing glycome composition could be explored as risk factors for colorectal cancer. Finally, glyco-modifications might have relevance to tumour immunosurveilance and in predicting response to monoclonal antibodies.

Abstract (250 words)

Purpose: Alternative glycosylation has significant structural and functional consequences on immunoglobulin G and consequently also cancer immunosurveilance. Due to technological limitations, the effects of highly heritable individual variations and the differences in the dynamics of changes in IgG glycosylation on colorectal cancer (CRC) were never investigated before.

Experimental design: Using recently developed high-throughput UPLC technology for IgG glycosylation analysis we analysed IgG glycome composition in 760 patients with CRC and 538 matching controls. Effects of surgery were evaluated in 28 patients sampled before and three times after surgery. A predictive model was built using regularized logistic regression and evaluated using a 10-cross validation procedure. Furthermore, IgG glycome composition was analysed in 39 plasma samples collected before initial diagnosis of CRC.

Results: We have found that CRC associates with decrease in IgG galactosylation, IgG sialylation and increase in core-fucosylation of neutral glycans with concurrent decrease of core fucosylation of sialylated glycans. While a model based on age and sex did not show discriminative power (AUC=0.499), the addition of glycan variables into the model considerably increased the discriminative power of the model (AUC=0.755). However, none of these differences were significant in the small set of samples collected before the initial diagnosis.

Conclusions: Considering the functional relevance of IgG glycosylation for both tumour immunosurveilance and clinical efficacy of therapy with monoclonal antibodies, individual variation in IgG glycosylation may turn out to be important for prediction of disease course or the choice of therapy, thus warranting further, more detailed studies of IgG glycosylation in CRC.

Introduction

The majority of extracellular and secreted proteins are post-translationally modified by the addition of glycans that are important structural and functional elements in the majority of biological processes (1). A recent comprehensive report endorsed by the US National Academies concluded that "*glycans are directly involved in the pathophysiology of every major disease*" and that "*additional knowledge from glycoscience will be needed to realize the goals of personalized medicine* and to take advantage of the substantial investments in human genome and proteome research and its impact on human health" (2). Glycans seem to have a particularly important role in the immune system and inter-individual variation in glycosylation may affect function of the immune system on multiple levels (3). The differences in an individual's immune repertoire and the capacity to process and present antigens are an important element of cancer immunosurveilance (4, 5). In addition to natural anti-cancer antibodies, therapeutic antibodies provide clinical benefit to patients with cancer and have been established as 'standard of care' therapy for several highly prevalent human cancers (6). While antigen specificity of antibodies is determined by nucleotide sequence of hypervariable regions, effector functions of antibodies are largely determined by alternative glycosylation of asparagine 297 in the Fc region of IgG (7, 8). Depending on the extent of galactosylation, sialylation and fucosylation of its glycans, IgG will activate complement, activate antibody-dependent cellular cytotoxicity (ADCC), or even have an anti-inflammatory action (9, 10). Both ADCC and complement-dependent cytotoxicity are important for function of anti-cancer antibodies (6), thus variation in IgG glycosylation may influence inter-individual variability in cancer immunosurveilance and/or response to therapeutic antibodies.

Our recent studies demonstrated that IgG glycosylation is very variable between individuals (11), but the functional relevance of this variation is not known. Since glycans do not have a direct genetic template, glycan structures that will be attached to an individual protein are determined by complex dynamic interactions between a number of genetic and environmental factors (12). Our recent studies have revealed that genetic loci associated with variation in IgG glycosylation are also known risk factors for several autoimmune diseases and cancers (13), indicating that variations in IgG function may be one of the mechanisms linking these genetic loci and diseases. Further support for this hypothesis came from recent large case / control studies that reported significantly reduced immunosuppressive features of IgG glycome in inflammatory bowel disease (14) and systemic lupus erythematous (15).

In this study we have compared IgG glycosylation in 760 patients with colorectal cancer (CRC) with that in 538 age and sex matched healthy controls. Effects of surgery on IgG glycome were evaluated in 28 patients sampled before surgery, and at three additional time points after surgery. Furthermore, we identified 39 historical samples of people who subsequently developed CRC and compared their IgG glycome composition to matched controls (80 individuals that did not develop CRC over the same period).

Materials and methods

Study population

The Study of Colorectal Cancer in Scotland (SOCCS) study (1999-2006) is a case-control study designed to identify genetic and environmental factors associated with non-hereditary colorectal cancer risk and survival outcomes. Approval for the study was obtained from the MultiCentre Research Ethics Committee for Scotland and Local Research Ethics committee, and all participants gave written informed consent. The study has been described in detail elsewhere(16).

The present study includes a subset of 760 patients with pathologically confirmed colorectal adenocarcinoma and 538 matching controls. IgG glycome composition was analysed in samples collected after CRC diagnosis or recruitment. Participants completed a detailed lifestyle questionnaire, a semi-quantitative food frequency (http://www.foodfrequency.org) and supplements questionnaire. Blood was collected and transferred to the research centre within 72 h of sampling. Plasma was prepared from whole blood by gentle centrifugation through a ficoll-hypaque gradient and 1.5 mL of each participant's plasma was stored at -80°C.

Effects of surgery on the IgG glycome were evaluated in 28 patients sampled before surgery, 24h, 48h and 7 days after surgery. Samples were collected at the Clinical Department for Laboratory Diagnostics at University Hospital Dubrava, The study was registratated at ClinicalTrials.gov, number NCT01244022 and was approved by the Ethics Committee of University Hospital Dubrava.

Patients who were sampled before the initial diagnosis of CRC, as well as matching controls that did not develop CRC during the same follow-up time were selected from the FINNRISK cohort (17).

Isolation of IgG from human plasma

The IgG was isolated using protein G monolithic plates (BIA Separations, Ajdovščina, Slovenia) as described previously (11).Briefly, 70 - 100 µl of plasma was diluted 8× with 1× PBS, pH 7.4, applied to the protein G plate and instantly washed with 1× PBS, pH 7.4 to remove unbound proteins. IgGs were eluted with 1 ml of 0.1 M formic acid (Merck, Darmstadt, Germany) and neutralized with 1 M ammonium bicarbonate (Merck, Darmstadt, Germany).

Glycan release

IgG samples were dried in a vacuum concentrator, dissolved in 30 µL 1.33% SDS (w/v) (Invitrogen, Carlsbad, CA, USA) and denatured by incubation at 65 °C for 10 minutes. After incubation, samples were left to cool down to room temperature for 30 minutes. Subsequently, 10 µL of 4% Igepal-CA630 (Sigma-Aldrich, St. Louis, MO, USA) were added to the samples and incubated on a shaker for 15 minutes. After shaking, 1.25 mU of PNGase F (ProZyme, Hayward, CA, USA) in 10 µL 5× PBS were added and incubated overnight at 37 °C for N-glycan release.

Glycan labelling

The released N-glycans were labelled with 2-aminobenzamide (2-AB). The labelling mixture was freshly prepared by dissolving 2-AB (Sigma-Aldrich, St. Louis, MO, USA) in DMSO (Sigma-Aldrich, St. Louis, MO, USA) and glacial acetic acid (Merck, Darmstadt, Germany) mixture (70:30, v/v) and by adding 2-picoline borane (Sigma-

Aldrich, St. Louis, MO, USA) to a final concentration of 19.2 mg/mL for 2AB and 44.8 mg/ml for 2-picoline borane. A volume of 25 µL of labelling mixture was added to each N-glycan sample in the 96-well plate and the plate was sealed using adhesive tape. Mixing was achieved by shaking for 10 minutes, followed by 2 hour incubation at 65 °C.

HILIC-SPE

Free label and reducing agent were removed from the samples using HILIC-SPE. Cellulose based HILIC SPE clean-up of the Hispanic cohort samples was performed as described previously (Kristic et al., 2014). The 2x 100 µL eluates were combined and stored at –20 °C until usage.

Hydrophilic Interaction Chromatography (HILIC)-UPLC

Fluorescently labelled N-glycans were separated by hydrophilic interaction chromatography on a Waters Acquity UPLC instrument (Milford, MA, USA) consisting of a quaternary solvent manager, sample manager and a FLR fluorescence detector set with excitation and emission wavelengths of 250 and 428 nm, respectively. The instrument was under the control of Empower 2 software, build 2145 (Waters, Milford, MA, USA). Labelled 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. Separation method used linear gradient of 75–62% acetonitrile (v/v) at flow rate of 0.4 ml/min in a 25 minute 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 labelled 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 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.

Statistical analysis

Clinical characteristics among patients and controls were compared using Wilcoxon and Fisher Exact tests. To make measurements across samples comparable, normalization by total area was performed where peak area of each of 24 glycan structures was divided by total area of corresponding chromatogram. Batch correction was performed on normalized log-transformed measurements using linear mixed models (R package lme4), where technical source of variation was modelled as random effect. In addition to 24 directly measured glycan structures, 12 derived traits were calculated from the directly measured glycans. These derived traits average particular glycosylation features (galactosylation, fucosylation, sialylation) across different individual glycan structures and consequently they are more closely related to individual enzymatic activities and underlying genetic polymorphisms.

Association analyses between disease status and glycan traits were performed using a logistic regression model with age and sex included as additional covariates. Prior to analyses, glycan variables were all transformed to standard Normal distribution

(mean=0,sd=1) by inverse transformation of ranks to Normality (R package "GenABEL", function rntransform). Using rank transformed variables in the analysis makes estimated odds ratios of different glycans comparable as transformed glycan variables have the same standardized variance, and in that case estimated odds ratios always correspond to one standard deviation change in the measured glycan trait. False discovery rate was controlled using Benjamini-Hochberg procedure (18).

For prediction of CRC status, regularized logistic (elastic net) regression model was applied (R package "glmnet"). For classification, only 24 initial glycan traits were used as predictors. Prior to model training and validation, elastic net regularization parameters (alpha and lambda) were tuned on 20% of samples (260 samples), and optimal parameters chosen by the tuning procedure (alpha = 0, lambda = 0.1) were used in further analysis. To evaluate performance of predictive model 10-cross validation procedure was used on remaining 80% of samples. Predictions from each validation round were merged into one validation set on which model performance was evaluated based on area under the receiver operating characteristic curve (AUC) criteria. The AUCs of different models were compared using a bootstrap test.

Data was analyzed and visualized using R programming language (version 3.0.1).

Expression datasets for colorectal carcinoma

Expression data for colorectal carcinoma were downloaded from the Gene Expression Omnibus (GEO) database (19). We picked datasets with a comparably large number of samples (49 to 148 samples total) that included both colorectal cancer and the matching non-cancerous tissue. The datasets are stored under the accession numbers GSE8671(20), GSE21510 (21) and GSE24514 (22).

Analysis of gene expression in colorectal carcinoma

Expression data analysis was conducted using the R language and environment for statistical computing (23), which was also used for preparation of figures. Data were retrieved with the R package “GEOquery”. Differences in expression levels with the corresponding *p*-values for each dataset containing cancer and non-cancer samples were calculated using the package “limma”. Genes were annotated according to the package and database “org.Hs.eg.db”. Results for the 20 selected candidate genes were visualized as volcano plots constructed with the aid of the package “ggplot2”. In these plots, the cut-off value for significance was set well above the level of p=0.05 adjusted using Bonferroni correction for multiple testing. The cut-off value for the fold change was picked for each individual dataset such that only the several most changed genes are selected.

Results

IgG glycome composition was analysed in 760 patients with CRC and 538 matching controls. Descriptive information on CRC patients and healthy controls are presented in Table 1. Total IgG glycome (combined Fc and Fab glycans) composition was determined by UPLC analysis of 2AB-labelled glycans as reported recently (13). A typical chromatogram showing separation of the IgG glycome into individual structures is shown in Fig 1. Many of these structures share the same structural features (galactose, sialic acid, core-fucose), thus additional derived traits were calculated that average these features across multiple glycans (Supplementary Table 1).

Significant differences were observed in several features of the glycome (Table 2, Fig 2), primarily reflecting decreased galactosylation (OR=2.35, p=2.39E-22 for G0 and OR=0.36, p=6.59E-29 for G2) and sialylation (OR=0.72, p=2.73E-05 for S total), as well as increased fucosylation of neutral IgG glycans (OR=1.24, p=3.57E-03 for F total) and decreased fucosylation of sialylated glycans (OR=0.72, p=5.85E-05 for F sialo). As many glycan structures showed strong association with CRC, we attempted to build a predictive model using regularized logistic regression. Only the 24 directly measured glycan traits were used as predictors in the model. Evaluation of model performance was conducted using a 10-cross validation procedure. While a model based on age and sex did not show significant discriminative power (AUC = 0.499), the addition of glycan variables into the model considerably increased the discriminative power of the model (AUC = 0.755, P < 1x10E-16)(Fig 3).

Glycome composition is known to change in acute inflammation (24) and to evaluate potential effects of surgery on the IgG glycome in CRC patients we analysed IgG glycome composition in 28 patients (i) before surgery, (ii) 24 after surgery), (iii) 48 h after surgery and (iv) 7 days after surgery. We did not observe any consistent and statistically significant changes in the IgG glycome that were caused by the surgery (data not shown).

To determine whether the observed changes were present before the disease onset, we identified 39 individuals from the FINNRISK cohort that were sampled before the initial diagnosis. However, when compared to matching controls, no statistically significant changes were found.

Core-fucose is added to glycans by the fucosyltransferase 8 enzyme (encoded by *FUT8* gene), which has recently been reported to be functionally relevant in some cancers (25-27). Aiming to investigate the potential molecular mechanisms underlying increased IgG core-fucosylation in CRC, we examined effects of single nucleotide polymorphisms (SNPs) that associated with fucose-related glycan traits on both glycans and CRC. A list of SNPs that showed the strongest associations with fucosylation in our recent IgG glycome GWAS are shown in Supplementary table 3 with effect sizes and p values for both glycans and CRC. Although SNPs that associate with fucosylation did not pass correction for multiple testing for association with CRC, it is of interest that we observed nominally significant associations with CRC risk for 96 out of 100 top glycan SNPs and that the direction of the effect was consistent for both core fucose and CRC risk in all these SNPs.

Furthermore, we examined the expression of all genes that were previously associated with the IgG glycome (13) in CRC and matching tissues in publicly available datasets. Significantly increased expression of *FUT8* was observed in all datasets (Fig 4).

Discussion

This study represents the first comprehensive analysis of IgG glycosylation in colorectal cancer. By applying the recently developed method for high-throughput glyco-profiling of IgG (11) on a well-characterized large cohort of CRC patients and matching controls we have filled this important gap in knowledge which may have relevance for understanding the immunosurveilance of CRC (4). More generally, given the established role of IgG glycan composition in determining IgG effector function, these data may be of importance in helping to understand determinants of success of anti-cancer therapy with monoclonal antibodies (6). By analysing 760 CRC patients and 538 matching controls we have found that CRC is associated with three major alterations in the IgG glycome composition: (i) decrease in IgG galactosylation, (ii) decrease in IgG sialylation and (iii) increase in core-fucosylation of neutral glycans with concurrent decrease of core fucosylation of sialylated glycans.

The addition of galactose and sialic acid to IgG glycans have been shown to result in the conversion of pro-inflammatory activity of IgG into an anti-inflammatory activity and are believed to be essential for the immunosuppressive activity of intravenously administered immunoglobulins (IVIG) (28). Galactosylation and sialylation of IgG decrease significantly with the age of an individual (29). The same pattern of changes in the IgG glycome have been reported in rheumatoid arthritis (30), osteoarthritis (30), inflammatory bowel disease (14, 31), systemic lupus erythematosus (15) and some other diseases (32). In this study we have observed consistent decreases in all structures with two terminal galactoses (A2BG2 (GP13), FA2G2 (GP14) and FA2BG2(GP15)) and an increase in structures without galactoses (A2 (GP2), A2B (GP3), FA2 (GP4) and FA2B (GP6)) (Supplementary Table 2). The decrease in galactosylation was also evident in the derived traits G0 and G2 that average galactosylation in several individual glycans (G0: OR=2,35; p =2,39E-22; G2: OR=0,36; p=6,59E-29). Several directly measured IgG glycans containing sialic acid, as well as the derived trait “S1 total” (measuring all monosialylated IgG glycans) were also decreased in CRC patients (Table 2).

Since decreases in galactosylation and sialylation have also been observed in a number of other diseases (32), this pattern of glycan changes (which are consistent with a decrease in the immunosuppressive potential of IgG) is not specific for CRC. Since little is known about the mechanisms of regulation of IgG glycosylation (12) it is very difficult to speculate about potential mechanisms and causes of these differences. The heritability of IgG glycosylation is relatively high. The heritability of galactosylation has been estimated to be between 40% and 70% and of sialylation between 30% and 60% (11, 33). However, the glycome has also been reported to be dynamic and capable of change (particularly in situations of disturbed homeostasis (24). Recently it was reported that the IgG glycome changes before the onset of rheumatoid arthritis (34, 35), indicating that they may be a part of disease pathophysiology. In an analogous attempt to address causality in this study we were able to identify 39 individuals that were sampled before the initial diagnosis of CRC. However, when compared to matching controls (individuals of the same age that did not develop CRC in the same period after recruitment) we did not identify any statistically significant differences. Since the sample size was very small it is hard to derive any conclusions from this exploratory part of the study, beside the fact that

we were not able to show any differences in IgG glycome composition before the onset of CRC.

Activation of ADCC is believed to be an important mechanism of therapeutic monoclonal antibodies, as indicated by the fact that a common single-nucleotide polymorphism in FcγRIIIa (V158F) is correlated with clinical responses to cetuximab (36) and rituximab (37). The attachment of fucose to the glycan core (core-fucose) is a mechanism that interferes with binding of IgG to FcγRIIIa and greatly diminishes its capacity to activate ADCC (38, 39). The removal of core-fucose from IgG glycans increases clinical efficacy of monoclonal antibodies, enhancing their therapeutic effect through ADCC mediated killing (40-42). Recently it has been shown that properly glyco-engineered antibodies (43) can efficiently elicit ADCC even in immunocompromised CRC patients (44). This observation indicates that efficient immunosurveilance of tumor cells depends on antibody/Fc receptor affinity. Thus, increased levels of core-fucose on neutral IgG glycans in CRC patients may influence disease risk and course by decreasing the ability of IgG to activate ADCC. Interestingly, core-fucose was increased in sialylated glycans. To the best of our knowledge this is the first report of different direction of changes in fucosylation in sialylated and neutral glycans. Recent studies clearly demonstrated that some antigen-specific antibodies can have significantly different levels of fucose (45), but it is hard to evaluate the importance in the different fucosylation of sialylated and neutral IgGs, since the relevance of sialylation on the impact of fucose on IgG function is currently not known (46).

This is further supported by the observation that polymorphisms in the *FUT8* gene seem to be associated with increased risk for CRC. In our relatively small dataset this association did not remain statistically significant after correction for multiple testing. Nevertheless, we suggest that this is an interesting and potentially important observation. It is supported by our finding that 96 out of 100 top glycan SNPs were also nominally significant for CRC risk, with directions of effect that were consistent (Supplementary Table 1) with our hypothesis. A larger study would be needed to confirm this association, but if replicated, this observation would indicate that individual variation in IgG fucosylation affects risk for the development of CRC.

Additional evidence supporting a possible functional importance of *FUT8* in colorectal cancer is the recent observation that micro RNA MiR-198 represses tumour growth and metastasis in colorectal cancer by targeting *FUT8* (47). Our analysis of transcriptional data in publicly available datasets also revealed significantly increased expression of *FUT8* in CRC (Fig 4). The finding that similar pattern of changes in *FUT8* expression can be observed or inferred in both the tumour tissue and the antibody-producing B lymphocytes points to the importance of general mechanisms controlling fucosylation in CRC and implies that the same features of genetic makeup influence glyco-gene expression (and thereby glycosylation profile) in both the tumour tissue and in B cells. We recommend that attempts should be made to replicate the findings of this “hypothesis-generating” study in other datasets. In particular, the association between increased levels of core-fucose on neutral IgG glycans and CRC patients warrants further investigation given their potential therapeutic implications and translational potential.

Decreases in galactosylation and sialylation have been previously observed in a number of diseases, including in our recent large studies of inflammatory bowel disease (14) and systemic lupus erythematous (15). Thus, these glycan changes do not appear to be specific for CRC. However, the increase in core fucose on neutral glycans, with concurrent decrease of core-fucose on sialylated glycans has not been previously reported.

In this study we have demonstrated significant differences in IgG glycome composition between CRC patients and controls. We were not able to detect these differences in historical samples (taken before CRC had developed in these patients). This could indicate that the changes are due to reverse causality (due to the disease process or treatment effects). However, it may also be due to inadequate study power in this small sub-study and so additional studies are required to investigate this further. Considering the functional relevance of IgG glycosylation for both tumour immunosurveilance and clinical efficacy of therapy with monoclonal antibodies, individual variation in IgG glycosylation may turn out to be important for prediction of disease course or the choice of therapy, thus warranting further, more detailed studies of IgG glycosylation in CRC.

References

1. Hudak JE, Bertozzi CR. Glycotherapy: new advances inspire a reemergence of glycans in medicine. Chem Biol. 2014;21:16-37.
2. Walt D, Aoki-Kinoshita KF, Bendiak B, Bertozzi CR, Boons GJ, Darvill A, et al. Transforming Glycoscience: A Roadmap for the Future. Washington: Nacional Academies Press; 2012.
3. Wolfert MA, Boons GJ. Adaptive immune activation: glycosylation does matter. Nat Chem Biol. 2013;9:776-84.
4. Vesely MD, Kershaw MH, Schreiber RD, Smyth MJ. Natural innate and adaptive immunity to cancer. Annu Rev Immunol. 2011;29:235-71.
5. Burnet M. Cancer; a biological approach. I. The processes of control. British medical journal. 1957;1:779-86.
6. Weiner LM, Surana R, Wang S. Monoclonal antibodies: versatile platforms for cancer immunotherapy. Nat Rev Immunol. 2010;10:317-27.
7. Kapur R, Einarsdottir HK, Vidarsson G. IgG-effector functions: "The Good, The Bad and The Ugly". Immunology letters. 2014.
8. Quast I, Lunemann JD. Fc Glycan-Modulated Immunoglobulin G Effector Functions. J Clin Immunol. 2014.
9. Nimmerjahn F, Ravetch JV. Translating basic mechanisms of IgG effector activity into next generation cancer therapies. Cancer immunity. 2012;12:13.
10. Gornik O, Pavic T, Lauc G. Alternative glycosylation modulates function of IgG and other proteins - Implications on evolution and disease. Biochim Biophys Acta. 2012;1820:1318-26.
11. Pucic M, Knezevic A, Vidic J, Adamczyk B, Novokmet M, Polasek O, et al. High throughput isolation and glycosylation analysis of IgG-variability and heritability of the IgG glycome in three isolated human populations. Mol Cell Proteomics. 2011;10:M111 010090.
12. Krištić J, Zoldoš V, Lauc G. Complex Genetics of Protein N-Glycosylation. In: Endo T, Seeberger PH, Hart GW, Wong C-H, Taniguchi N, editors. Glycoscience: Biology and Medicine: Springer Japan; 2014. p. 1-7.
13. Lauc G, Huffman JE, Pucic M, Zgaga L, Adamczyk B, Muzinic A, et al. Loci associated with N-glycosylation of human immunoglobulin G show pleiotropy with autoimmune diseases and haematological cancers. PLoS Genet. 2013;9:e1003225.
14. Trbojević Akmačić I, Ventham NT, Theodoratou E, Vučković F, Kennedy NA, Krištić J, et al. Inflammatory bowel disease associates with proinflammatory potential of the immunoglobulin g glycome. Inflamm Bowel Dis. 2015;21:1237-47.
15. Vučković F, Krištić J, Gudelj I, Artacho MT, Keser T, Pezer M, et al. Systemic lupus erythematosus associates with the decreased immunosuppressive potential of the IgG glycome. Arthritis & Rheumatology. 2015;67:2978-89.
16. Theodoratou E, Kyle J, Cetnarskyj R, Farrington SM, Tenesa A, Barnetson R, et al. Dietary flavonoids and the risk of colorectal cancer. Cancer Epidemiol Biomarkers Prev. 2007;16:684-93.

17. Vartiainen E, Laatikainen T, Peltonen M, Juolevi A, Mannisto S, Sundvall J, et al. Thirty-five-year trends in cardiovascular risk factors in Finland. Int J Epidemiol. 2010;39:504-18.
18. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society Series B. 1995:289-300.
19. Edgar R, Domrachev M, Lash AE. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. Nucleic Acids Res. 2002;30:207-10.
20. Sabates-Bellver J, Van der Flier LG, de Palo M, Cattaneo E, Maake C, Rehrauer H, et al. Transcriptome profile of human colorectal adenomas. Molecular cancer research : MCR. 2007;5:1263-75.
21. Tsukamoto S, Ishikawa T, Iida S, Ishiguro M, Mogushi K, Mizushima H, et al. Clinical significance of osteoprotegerin expression in human colorectal cancer. Clin Cancer Res. 2011;17:2444-50.
22. Alhopuro P, Sammalkorpi H, Niittymaki I, Bistrom M, Raitila A, Saharinen J, et al. Candidate driver genes in microsatellite-unstable colorectal cancer. Int J Cancer. 2012;130:1558-66.
23. R-Core-Team. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2013.
24. Novokmet M, Lukic E, Vuckovic F, Duric Z, Keser T, Rajsl K, et al. Changes in IgG and total plasma protein glycomes in acute systemic inflammation. Scientific reports. 2014;4:4347.
25. Chen CY, Jan YH, Juan YH, Yang CJ, Huang MS, Yu CJ, et al. Fucosyltransferase 8 as a functional regulator of nonsmall cell lung cancer. Proc Natl Acad Sci USA. 2013;110:630-5.
26. Wang X, Chen J, Li QK, Peskoe SB, Zhang B, Choi C, et al. Overexpression of alpha (1,6) fucosyltransferase associated with aggressive prostate cancer. Glycobiology. 2014;24:935-44.
27. Zhao YP, Xu XY, Fang M, Wang H, You Q, Yi CH, et al. Decreased core-fucosylation contributes to malignancy in gastric cancer. PLoS One. 2014;9:e94536.
28. Schwab I, Nimmerjahn F. Intravenous immunoglobulin therapy: how does IgG modulate the immune system? Nat Rev Immunol. 2013;13:176-89.
29. Kristic J, Vuckovic F, Menni C, Klaric L, Keser T, Beceheli I, et al. Glycans are a novel biomarker of chronological and biological ages. J Gerontol. 2014;69:779-89.
30. Parekh RB, Dwek RA, Sutton BJ, Fernandes DL, Leung A, Stanworth D, et al. Association of rheumatoid arthritis and primary osteoarthritis with changes in the glycosylation pattern of total serum IgG. Nature. 1985;316:452-7.
31. Dube R, Rook GA, Steele J, Brealey R, Dwek R, Rademacher T, et al. Agalactosyl IgG in inflammatory bowel disease: correlation with C-reactive protein. Gut. 1990;31:431-4.
32. Gornik O, Lauc G. Glycosylation of serum proteins in inflammatory diseases. Dis Markers. 2008;25:267-78.

33. Menni C, Keser T, Mangino M, Bell JT, Erte I, Akmačić I, et al. Glycosylation of Immunoglobulin G: Role of genetic and epigenetic influences. PLoS One. 2013;8:e82558.
34. Rombouts Y, Ewing E, van de Stadt LA, Selman MH, Trouw LA, Deelder AM, et al. Anti-citrullinated protein antibodies acquire a pro-inflammatory Fc glycosylation phenotype prior to the onset of rheumatoid arthritis. Annals of the rheumatic diseases. 2013.
35. Ercan A, Cui J, Chatterton DE, Deane KD, Hazen MM, Brintnell W, et al. Aberrant IgG galactosylation precedes disease onset, correlates with disease activity, and is prevalent in autoantibodies in rheumatoid arthritis. Arthritis and rheumatism. 2010;62:2239-48.
36. Zhang W, Gordon M, Schultheis AM, Yang DY, Nagashima F, Azuma M, et al. FCGR2A and FCGR3A polymorphisms associated with clinical outcome of epidermal growth factor receptor expressing metastatic colorectal cancer patients treated with single-agent cetuximab. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2007;25:3712-8.
37. Cartron G, Dacheux L, Salles G, Solal-Celigny P, Bardos P, Colombat P, et al. Therapeutic activity of humanized anti-CD20 monoclonal antibody and polymorphism in IgG Fc receptor FcgammaRIIIa gene. Blood. 2002;99:754-8.
38. Nimmerjahn F, Ravetch JV. Fcgamma receptors as regulators of immune responses. Nat Rev Immunol. 2008;8:34-47.
39. Ferrara C, Stuart F, Sondermann P, Brunker P, Umana P. The carbohydrate at FcgammaRIIIa Asn-162. An element required for high affinity binding to non-fucosylated IgG glycoforms. J Biol Chem. 2006;281:5032-6.
40. Shinkawa T, Nakamura K, Yamane N, Shoji-Hosaka E, Kanda Y, Sakurada M, et al. 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;278:3466-73.
41. Iida S, Misaka H, Inoue M, Shibata M, Nakano R, Yamane-Ohnuki N, et al. 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;12:2879-87.
42. Preithner S, Elm S, Lippold S, Locher M, Wolf A, da Silva AJ, et al. 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;43:1183-93.
43. Gerdes CA, Nicolini VG, Herter S, van Puijenbroek E, Lang S, Roemmele M, et al. GA201 (RG7160): a novel, humanized, glycoengineered anti-EGFR antibody with enhanced ADCC and superior in vivo efficacy compared with cetuximab. Clin Cancer Res. 2013;19:1126-38.
44. Oppenheim DE, Spreafico R, Etuk A, Malone D, Amofah E, Pena-Murillo C, et al. Glyco-engineered anti-EGFR mAb elicits ADCC by NK cells from colorectal cancer patients irrespective of chemotherapy. Br J Cancer. 2014;110:1221-7.
45. Kapur R, Kustiawan I, Vestrheim A, Koeleman CA, Visser R, Einarsdottir HK, et al. A prominent lack of IgG1-Fc fucosylation of platelet alloantibodies in pregnancy. Blood. 2014;123:471-80.

46. Biburger M, Lux A, Nimmerjahn F. How immunoglobulin G antibodies kill target cells: revisiting an old paradigm. Adv Immunol. 2014;124:67-94.
47. Wang M, Wang J, Kong X, Chen H, Wang Y, Qin M, et al. MiR-198 represses tumor growth and metastasis in colorectal cancer by targeting fucosyl transferase 8. Scientific reports. 2014;4:6145.

Table 1. Descriptive information on CRC patient and healthy controls

	Cases (n = 760)	Control (n = 538)	p-value
Age (median[IQR])	52 (48-56)	53 (48-56)	0.274
Sex (Men/Women)	415 (54.6%) / 345 (45.4%)	289 (53.7%) / 248 (46.2%)	0.821
Smoking status (Current/Ex/ Non/unknown)	133/172/271/184	94/139/211/103	0.836
BMI (median[IQR])	26.3 (23.4-26.8)	27.8 (25.8-28.6)	2.69E-10
Family History (Low/Medium or High/unknown)	524/197/39	517/6/15	6.50E-44

Table 2. IgG glycome composition in CRC patients and controls. Only the main derived traits describing glycome composition are shown. Directly measured glycan structures are available in Supplementary Table 2.

Glycan	Control (median[IQR])		Patient (median[IQR])		Odds Ratio	95% Confidence Interval	p-value	p- value adjusted
G0 total	25.35	(20.67-29.24)	29.24	(24.62-34.74)	2.35	(1.95 - 2.82)	2.65E-23	2.39E-22
G1 total	36.62	(34.87-38.09)	35.94	(34.49-37.39)	0.64	(0.55 - 0.74)	8.09E-10	3.64E-09
G2 total	16.65	(14.41-19.62)	14.12	(11.71-16.78)	0.36	(0.30 - 0.44)	1.83E-30	6.59E-29
F total	95.27	(94.42-95.97)	95.39	(94.40-96.11)	1.11	(0.97 - 1.28)	1.27E-01	1.53E-01
F neutral	97.07	(96.38-97.66)	97.32	(96.50-97.97)	1.24	(1.08 - 1.43)	1.98E-03	3.57E-03
F sialo	88.47	(86.69-89.79)	87.58	(85.39-89.29)	0.72	(0.62 - 0.84)	2.44E-05	5.85E-05
B total	18.25	(16.81-20.22)	18.11	(16.49-20.02)	0.88	(0.76 - 1.01)	7.68E-02	9.87E-02
B neutral	17.97	(16.35-20.54)	17.89	(16.07-19.93)	0.84	(0.73 - 0.97)	1.86E-02	2.91E-02
B sialo	18.64	(16.48-21.33)	19.19	(16.99-21.71)	1.17	(1.01 - 1.35)	3.85E-02	5.13E-02
S total	20.71	(18.64-23.62)	19.85	(17.39-21.96)	0.72	(0.62 - 0.83)	1.06E-05	2.73E-05
S1 total	15.9	(14.07-18.07)	14.97	(12.94-16.74)	0.64	(0.54 - 0.75)	9.16E-09	3.66E-08
S2 total	4.41	(3.66-5.32)	4.34	(3.65-5.08)	0.96	(0.84 - 1.11)	6.03E-01	6.03E-01

G – galactose; F – fucose; B – bisecting GlcNAc, S – sialic acid

False discovery rate was controlled using Benjamini-Hochberg procedure (p-value adjusted)

Fig 1. UPLC analysis of the IgG glycome. Representative chromatogram showing separation of labelled IgG glycans into 24 chromatographic peaks is shown.

Fig 2. IgG glycome composition in CRC patients and matching controls. IgG glycome was analysed in 760 patients with CRC and 538 matching controls. Main features of the IgG glycome are presented as box plots showing median values and 25% (box) and 75% (line) percentiles for patients and controls. Additional information is available in Table 2 and Supplementary Table 2.

Fig 3. Classification of CrC patients using IgG glycans. Fig 3. ROC curve illustrating the performance of regularized logistic regression model in discrimination between CRC patients and healthy controls from SOCCS retrospective study. While models based only on age and gender did not show predictive power (dotted line), addition of glycan traits increased predictive power of model (solid line).

Figure 4. Volcano plots showing changes in expression for the 20 candidate genes in three different datasets for colorectal carcinoma (see Materials and methods section). Names of the genes above cut-offs for the fold change and the statistical significance are indicated. One gene may correspond to multiple probe sets, each of them corresponding to a point on a graph. The $\log_2$ fold change refers to the ratio of expression in non-cancerous tissue to the expression in cancer tissue, which puts the genes with higher expression in cancer to the left and those with lower expression to the right. The position on the y-axis ($-\log_{10}$ *p*-value) corresponds to the statistical significance of the change, with higher significance positioned higher on the graph.

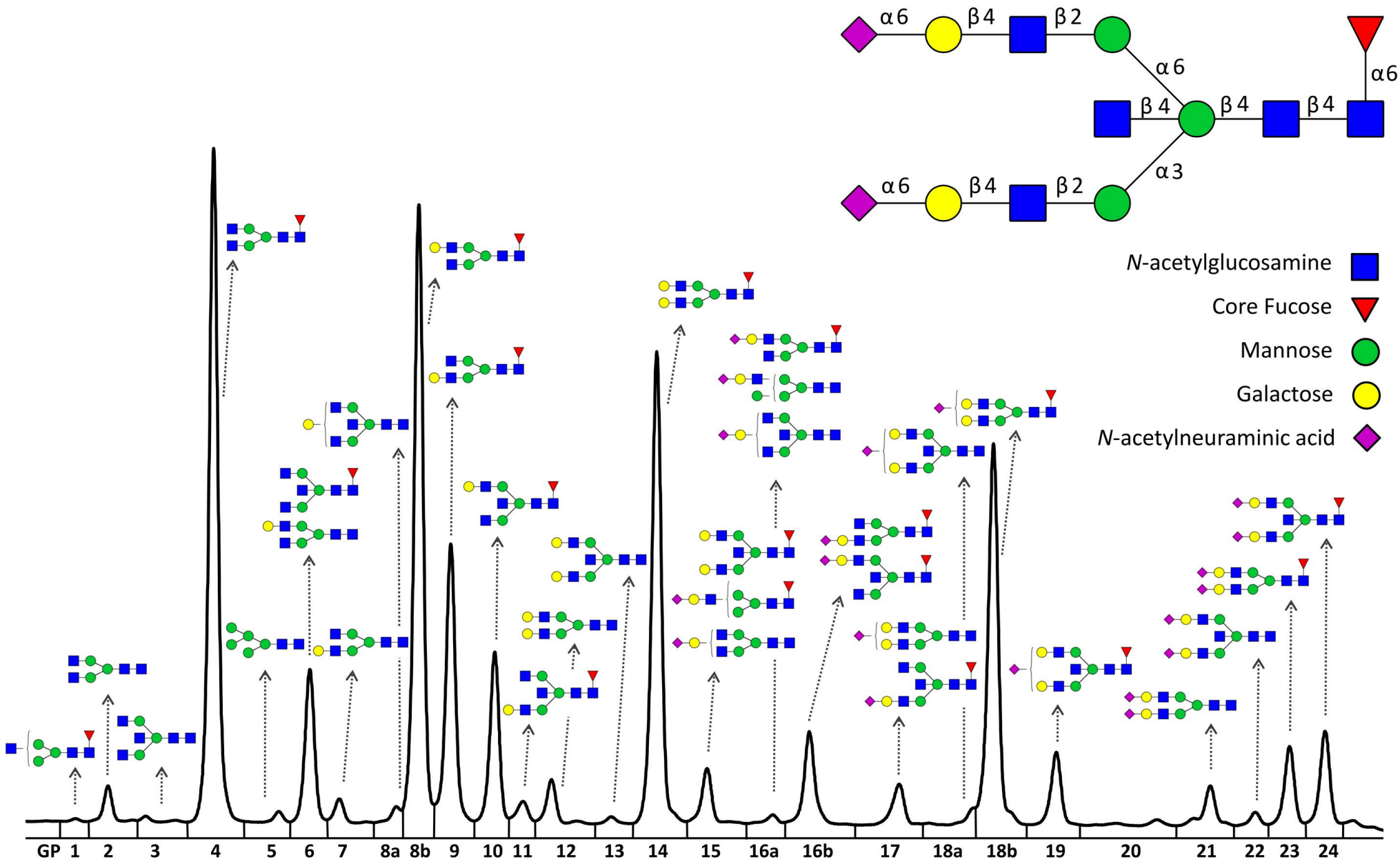
α6
β4
β2
α6
α6
β4
β4
β4
α6
α3
α6
β4
β2
N-acetylglucosamine
Core Fucose
Mannose
Galactose
N-acetylneuraminic acid
GP 1 2 3 4 5 6 7 8a 8b 9 10 11 12 13 14 15 16a 16b 17 18a 18b 19 20 21 22 23 24

Figure 1

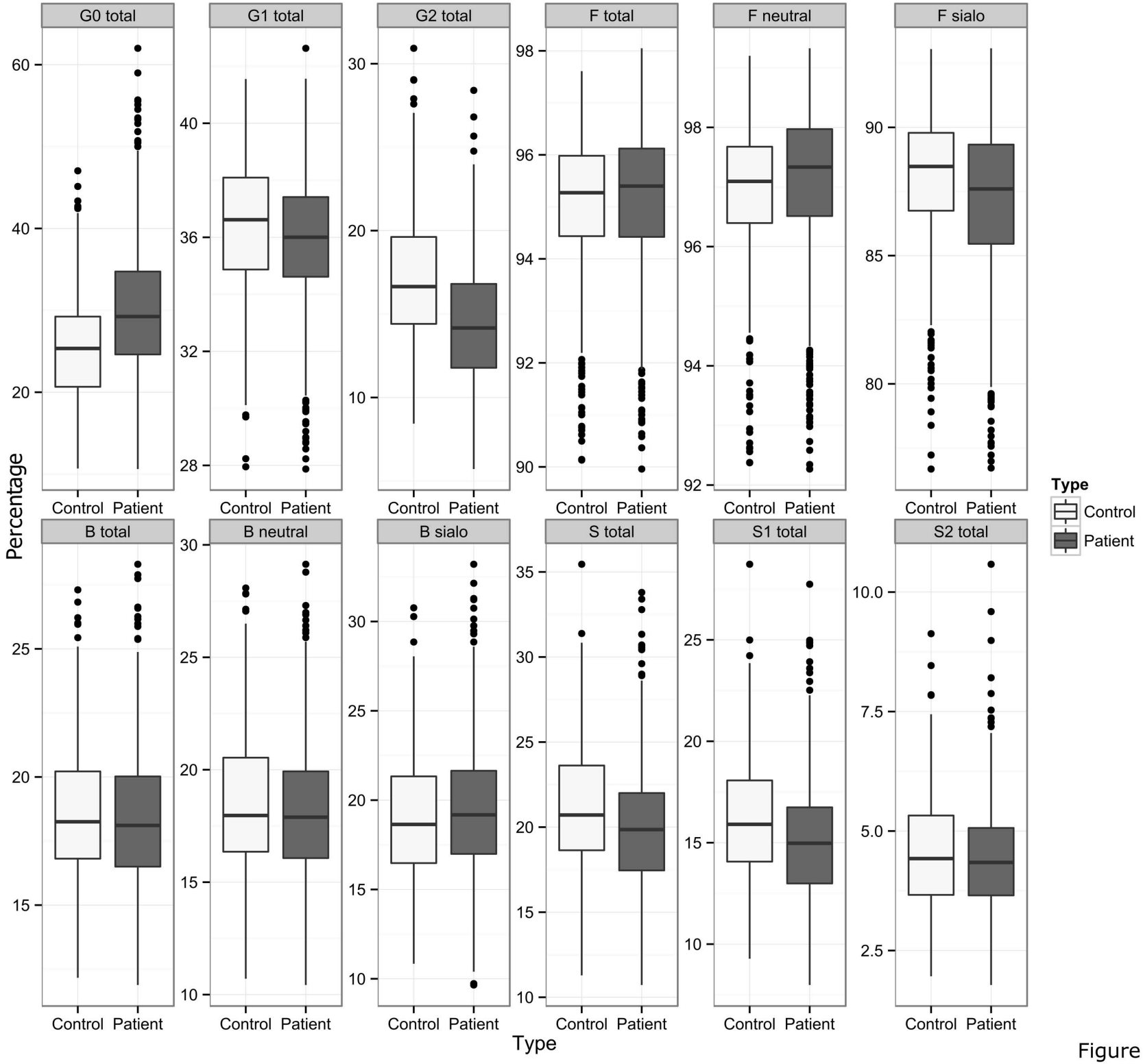
G0 total
G1 total
G2 total
F total
F neutral
F sialo
B total
B neutral
B sialo
S total
S1 total
S2 total
Percentage
Control
Patient
Type
Type
Control
Patient

Figure 2

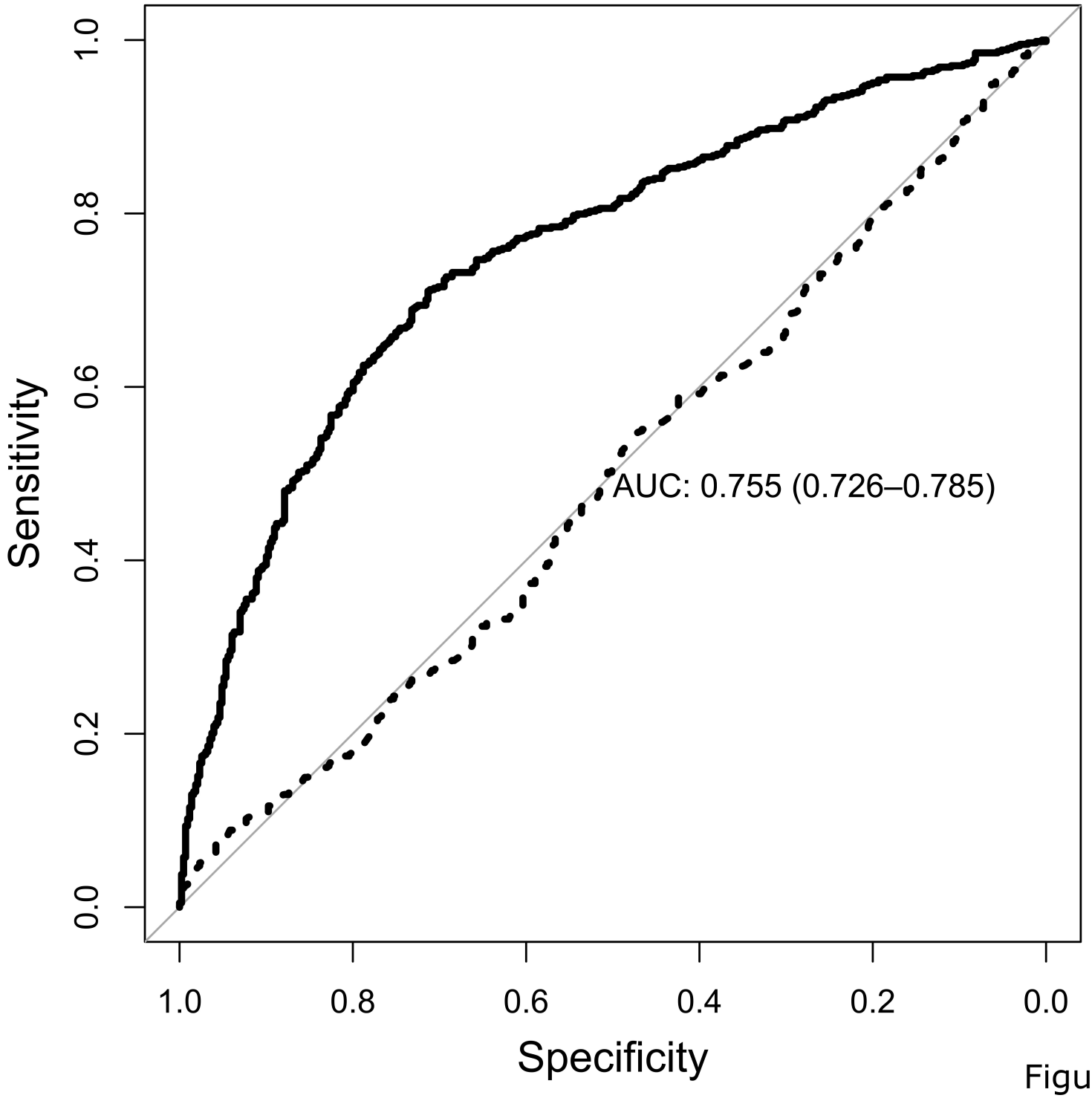
Sensitivity
1.0
0.8
0.6
0.4
0.2
0.0
AUC: 0.755 (0.726–0.785)
1.0
0.8
0.6
0.4
0.2
0.0
Specificity

Figure 3

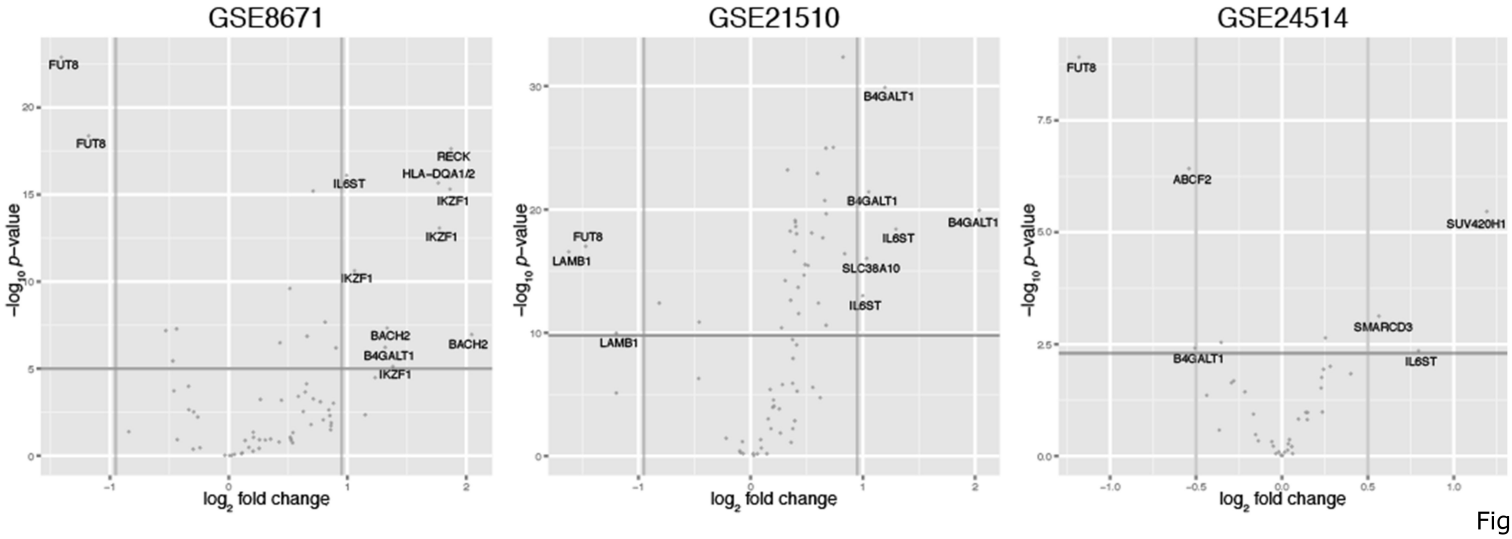
GSE8671
FUT8
FUT8
RECK
HLA-DQA1/2
IL6ST
IKZF1
IKZF1
IKZF1
BACH2
BACH2
B4GALT1
IKZF1
−log10 p-value
log2 fold change
GSE21510
B4GALT1
B4GALT1
B4GALT1
FUT8
IL6ST
LAMB1
SLC38A10
IL6ST
LAMB1
−log10 p-value
log2 fold change
GSE24514
FUT8
ABCF2
SUV420H1
SMARCD3
B4GALT1
IL6ST
−log10 p-value
log2 fold change

Figure 4