CELL LINES WITH INTEGRATED CRISPRa AND CRISPRi SYSTEM AND THEIR APPLICATION IN ANALYSIS OF IMMUNOGLOBULIN G GLYCOSYLATION

Mijakovac, Anika

University of Zagreb

FACULTY OF SCIENCE

DEPARTMENT OF BIOLOGY

Anika Mijakovac

CELL LINES WITH INTEGRATED CRISPRa AND CRISPRi SYSTEM AND THEIR APPLICATION IN ANALYSIS OF IMMUNOGLOBULIN G GLYCOSYLATION

DOCTORAL THESIS

Zagreb, 2023

Sveučilište u Zagrebu

PRIRODOSLOVNO-MATEMATIČKI FAKULTET

BIOLOŠKI ODSJEK

Anika Mijakovac

STANIČNE LINIJE S INTEGRIRANIM CRISPRa I CRISPRi SUSTAVOM I NJIHOVA PRIMJENA U ANALIZI GLIKOZILACIJE IMUNOGLOBULINA G

DOKTORSKI RAD

Zagreb, 2023

The work presented in this doctoral thesis was performed at the Division of Molecular Biology, Department of Biology, Faculty of Science, University of Zagreb, under the supervision of associate professor Aleksandar Vojta as a part of postgraduate program in Biology at the Department of Biology, Faculty of Science, University of Zagreb.

Ovaj doktorski rad je izrađen na Zavodu za molekularnu biologiju Biološkog odsjeka Prirodoslovno-matematičkog fakulteta Sveučilišta u Zagrebu pod vodstvom izv. prof. dr. sc. Aleksandra Vojte u sklopu Sveučilišnog poslijediplomskog studija Biologije pri Biološkom odsjeku Prirodoslovno-matematičkog fakulteta Sveučilišta u Zagrebu.

ZAHVALE

Iskreno se zahvaljujem mentoru izv. prof. dr. sc. Aleksandru Vojti na pruženoj podršci u trenutcima kada mi je bila najpotrebnija. Hvala vam na svim savjetima, motivacijskim razgovorima i humoru kojeg nikad nije nedostajalo.

Posebno zahvaljujem svojim znanstvenim roditeljima, prof. dr. sc. Vlatki Zoldoš i prof. dr. sc. Gordanu Laucu koji su prepoznali moju ljubav prema istraživanju i otvorili mi vrata u svijet znanosti. Vlatka, vi za mene niste samo iznimna profesorica i šefica grupe, već izuzetna žena koja me oblikovala u znanstvenicu i osobu koja sam danas. Ovi okviri su naprosto premali kako bi vam se zahvalila za sve što ste mi omogućili.

Veliko hvala mojoj znanstvenoj obitelji, Ani, Vedrani, Mariji, Niki, Goranu, Franu i Karlu. Zbog vas se ovo doktorsko istraživanje činilo kao zabava, a ne kao posao. Počašćena sam što sam imala priliku raditi s tako pametnim i izuzetnim ljudima. Ana, hvala ti što mi uvijek čuvaš leđa. Rijetko tko ima privilegiju raditi s najboljom prijateljicom. Također zahvaljujem svim kolegama na PMF-u i u Genos-u na nesebičnoj pomoći u svakom trenutku.

Hvala svim mojim prijateljima koji su me bodrili kad mi je bilo najteže. Posebno hvala Marinu koji danas zna puno više o epigenetici i glikobiologiji nego što bi jedan inženjer strojarstva trebao znati. Nisam znala što znači bezuvjetna podrška prije nego što sam te upoznala.

Za kraj zahvaljujem svojoj obitelji što su uvijek vjerovali u mene. Sve je lako pod okriljem vaše ljubavi.

Ovaj doktorat posvećujem baki Anki.

University of Zagreb
Faculty of Science
Department of Biology

Doctoral thesis

CELL LINES WITH INTEGRATED CRISPRa AND CRISPRi SYSTEM AND THEIR APPLICATION IN ANALYSIS OF IMMUNOGLOBULIN G GLYCOSYLATION

ANIKA MIJAKOVAC
Faculty of Science, University of Zagreb

Mechanisms regulating immunoglobulin G (IgG) glycosylation are largely unknown. To perform functional validation of genes implicated in IgG glycosylation we developed the HEK293-F system. Utilizing transposons, we integrated dCas9-VPR/dCas9-KRAB expression cassettes in the genome of FreeStyle™ 293-F cells for targeted manipulation (CRISPRa/CRISPRi) of selected genes. Transfection of HEK293-F cells with IgG-gRNA-bearing plasmid enabled simultaneous analysis of gene expression and IgG glycome. Utilizing this system, we validated the role of the *SPPL3* locus in IgG sialylation. Activation of estrogen-regulated genes, *RUNX3* and *SPINK4*, decreased the galactosylated IgG glycoforms, a phenotype characteristic of menopause and aging. To further examine the magnitude of genetic influences on IgG glycosylation in aging we performed a heritability analysis on the glycan clock marker of biological age. Heritability of the glycan clock, when corrected for the age of the individuals, proved to be high (~71%). This served to demonstrate the necessity to develop new technologies able to dissect the observed genetic factors regulating IgG glycosylation associated not only with disease but also with aging.

(102 pages, 20 figures, 1 table, 266 references, original in English)

Keywords: glycosylation, immunoglobulin G, CRISPR/dCas9, gene regulation, heritability

Supervisor: Aleksandar Vojta, PhD, Associate Professor

Reviewers: Vlatka Zoldoš, PhD, Professor

Ivana Ivančić Baće, PhD, Associate Professor

Gestur Vidarsson, PhD, Professor

Sveučilište u Zagrebu
Prirodoslovno-matematički fakultet
Biološki odsjek

Doktorski rad

STANIČNE LINIJE S INTEGRIRANIM CRISPRa I CRISPRi SUSTAVOM I NJIHOVA PRIMJENA U ANALIZI GLIKOZILACIJE IMUNOGLOBULINA G

ANIKA MIJAKOVAC

Prirodoslovno-matematički fakultet, Sveučilište u Zagrebu

Regulatorni mehanizmi glikozilacije imunoglobulina G (IgG) većinom su nepoznati. Kako bismo proveli funkcionalnu potvrdu uloge gena asociranih s IgG glikozilacijom razvili smo sustav HEK293-F. Korištenjem transpozona, ugradili smo kazete dCas9-VPR/dCas9-KRAB u genom stanica FreeStyle™ 293-F za ciljanu manipulaciju (CRISPRa/CRISPRi) odabranih gena. Transfekcija stanica HEK293-F s plazmidom IgG-gRNA omogućila nam je istovremenu analizu ekspresije gena i IgG glikoma. Korištenjem ovog sustava potvrdili smo ulogu lokusa *SPPL3* u IgG sijalinizaciji. Aktiviranjem gena reguliranih estrogenom, *RUNX3* i *SPINK4*, došlo je do smanjenja galaktoziliranih IgG glikoformi odnosno pojave fenotipa karakterističnog za menopauzu i starenje. Kako bismo dodatno istražili opseg genetičkih utjecaja na IgG glikozilaciju tijekom starenja, proveli smo analizu nasljednosti markera biološke dobi – glikanskog sata. Nasljednost glikanskog sata se pokazala visokom nakon korekcije za dob sudionika (~71%) čime smo pokazali kako je neophodno razvijati nove tehnologije koje će razjasniti genetičke faktore uključene u regulaciju glikozilacije IgG-a, povezane ne samo s bolestima već i sa starenjem.

(102 stranice, 20 slika, 1 tablica, 266 literaturnih navoda, jezik izvornika: engleski)

Ključne riječi: glikozilacija, imunoglobulin G, CRISPR/dCas9, regulacija gena, nasljednost

Mentor: Izv. prof. dr. sc. Aleksandar Vojta

Ocjenjivači: Prof. dr. sc. Vlatka Zoldoš

Izv. prof. dr. sc. Ivana Ivančić Baće

Prof. dr. sc. Gestur Vidarsson

Table of content

1. INTRODUCTION

1.1. Immunoglobulin G glycosylation: a general overview

Glycosylation is the most diverse protein modification present in all eukaryotic cells. It is defined as a sequence of enzymatic reactions resulting in the formation of complex sugars called glycans covalently attached to proteins (glycoproteins)[1–3]. The attachment of different glycans (alternative glycosylation) adds a new layer of complexity to proteins, extending the flexibility of their structure. The main difference between the protein primary sequence and glycosylation is that genes unambiguously determine the primary structure of a protein, while there is no simple genetic template for the glycans [4,5]. Instead, glycosylation is a complex trait, determined by the intricate interplay of many genes and their products, i.e. enzymes (glycosyltransferases), ion channels, transcription factors, etc[6–10]. It can affect the physical and biological properties of proteins to the same extent as genetic mutations [2,11,12]. Aberrant protein glycosylation has been observed in different autoimmune and inflammatory conditions, cancer and other age-related disorders [2,11,13–16]. It affects protein stability, solubility, folding, conformation and activity [17–20].

The most prevalent type of protein glycosylation is *N*-linked glycosylation, which involves the addition of glycans to the nitrogen atom of asparagine at the asparagine-X-serine/threonine motifs [21]. One of the most studied *N*-glycosylated proteins is the immunoglobulin G (IgG) [22]. IgG is an antibody produced by the B cells, with a central role in innate and adaptive immunity [23]. Alternative glycosylation of IgG can drastically change its function by acting as a switch between pro- and anti-inflammatory response [24]. Structurally, the IgG antibody has symmetrical conformation with two identical heavy (H) and two identical light (L) chains interconnected with disulfide bonds (Figure 1). Its function can be divided into two separate regions: Fab (fragment antigen binding) and Fc (fragment crystallizable). The Fab region is responsible for the binding of specific antigens, while the Fc region provides the effector function of this antibody. Each H chain of the Fc region has a conserved *N*-glycosylation (hereinafter referred to as glycosylation) site at the position asparagine (Asn) 297. Moreover, about 20% of IgG molecules have various glycosylation sites in the Fab region[25].

IgG Fc glycosylation modulates the immune response by affecting the binding of the Fc region to activating or inhibitory Fc receptors. These receptors are expressed on the surface of

various immune cells, which then mediate many immune processes such as antigen neutralization, complement activation, complement-dependent cytotoxicity (CDC), antibody-dependent cell-mediated cytotoxicity (ADCC), etc[26]. Considering this broad activity in the innate and adaptive immune response, IgG became the most studied glycoprotein up to date [22].

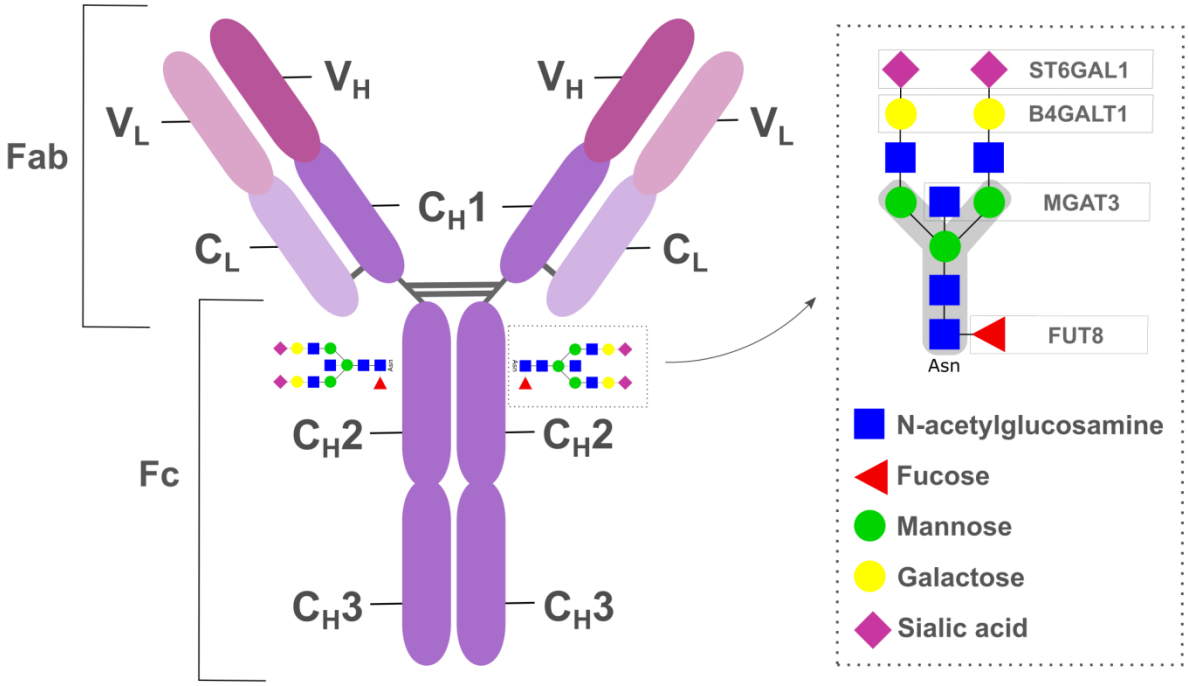

Figure 1. Schematic representation of the immunoglobulin G (IgG) antibody and the Asn297 attached *N*-glycan (hereinafter referred to as glycan). IgG is a homodimer composed of two heavy (H, colored in dark purple) and two light (L, colored in light purple) chains. Each heavy chain consists of three constant domains (C_H1, C_H2 and C_H3) and a variable domain (V_H) and each light chain consists of a constant domain (C_L) and a variable domain (V_L). Functionally, IgG is comprised of fragment antigen binding (Fab) and fragment crystallizable (Fc) domains. Each C_H2 domain has a conserved glycosylation site at the position asparagine 297 (Asn297). Some IgG molecules can have glycan attached to the Fab domain, which is not shown in the scheme. Fc attached glycan is shown on the right with the corresponding glycosyltransferase next to the monosaccharide unit it adds. Represented glycan is a diantennary core fucosylated digalactosylated and disialylated structure with a bisecting GlcNAc. The glycan core structure is highlighted in grey. ST6GAL1 - ST6 β-Galactoside α-2,6-Sialyltransferase 1; B4GALT1 – β-1,4-Galactosyltransferase 1; MGAT3 – β-1,4-Mannosyl-Glycoprotein 4-β-*N*-Acetylglucosaminyltransferase; FUT8 – α-(1,6)-fucosyltransferase.

1.1.1. Composition of IgG glycans

Glycans are composite oligosaccharides assembled from up to fifteen monosaccharide units. All IgG glycans have a conserved pentasaccharide core made of two *N*-acetylglucosamines (GlcNAc) and three mannose residues. IgG glycan core can be expanded by the addition of GlcNAc, core fucose, bisecting GlcNAc, galactose and sialic acid[27,28]. The most complex IgG glycan is a diantennary core fucosylated digalactosylated and disialylated structure with a bisecting GlcNAc, represented in Figure 1.

IgG isolated from healthy human plasma is predominantly core fucosylated (>90%), with moderate levels of galactose (~45%), low bisection (~15%) and low sialylation (~10%) [29,30]. Even though the pool of Fab and Fc glycans consists of the same structures, these two regions differ in their glycosylation patterns. Almost all Fab glycans are galactosylated with high levels of sialylation. In contrast, Fc glycans are rarely modified with sialic acid. Bisection of Fab glycans is threefold higher in the Fab region and core fucose is present in only about 70% of the structures, while almost all Fc glycans are core fucosylated[31–33].

1.1.2. Biological function of IgG glycosylation

Glycans attached to the Fc region are essential for the stability and activity of IgG. This is highlighted by the fact that IgG deglycosylation results in the complete loss of both pro- and anti-inflammatory activity of this antibody. Furthermore, alternative glycosylation alters Fc conformation which then modulates IgG's binding affinity to downstream effector cells of the immune system [34,35]. For example, the addition of bisecting GlcNac enhances ADCC by increasing IgG's affinity to FcγRIIIa receptor[36]. The same effect, but much more potent is exerted through the removal of core fucose which is exploited in the production of monoclonal antibodies (mAbs) that rely on ADCC[37,38]. It has been proposed that the addition of terminal sialic acid alters IgG's activity by directly changing the conformation of the Fc region. Due to this phenomenon, it was suggested that sialic acid acts as a switch between the inflammatory status of the IgG antibody and was considered to be an anti-inflammatory modification[39,40]. However, some recent studies have not been able to confirm that sialic acid has an anti-inflammatory nature demonstrating that more research is required on that subject[24,41]. Addition of galactose to the IgG glycan chain was also designated an anti-inflammatory modification. Few studies demonstrated that agalactosylated IgG glycans activate the complement through the lectin and alternative pathway while galactose inhibits C5a-dependant inflammation but the notion that agalactosylated glycans have a pro-inflammatory effect mostly came from

observational studies[11,42,43]. Recent research, on the other hand, points to a pro-inflammatory potential of galactosylated IgG glycans[24,41,44–46]. Latest studies even revealed a possible mechanism underlying the pro-inflammatory nature of IgG galactosylation where the addition of galactose promotes hexamerization of IgG necessary for complement activation[41,46]. These contradictions must be emphasized as they demonstrate that some established notions should be challenged and even more importantly, they serve to highlight the complexity of IgG biology. It must be acknowledged that classification of each glycan trait as pro- or anti-inflammatory is convenient but it represents an oversimplification of interdependent glycosylation pathways that regulate IgG activity.

1.2. Alternative IgG glycosylation in health and disease

Human IgG glycome is extremely heterogeneous on the population level mainly due to the variable monosaccharide composition and the presence of additional glycosylation sites in the Fab region. An example of this heterogeneity is the abundance of agalactosylated, core fucosylated IgG glycan that varies from 6% to 50% among healthy humans. On the contrary, IgG glycoprofile on an individual level is considered stable in homeostatic conditions[30,47]. However, when homeostasis is disturbed, IgG glycome composition can shift rapidly. The abundancies of different IgG glycoforms in healthy individuals are mainly influenced by age and sex hormone concentration. Moreover, aberrant IgG glycosylation is observed in various pathological conditions[11].

1.2.1. IgG glycome changes in various pathological conditions

Almost 40 years ago, Parekh *et al.* [48] observed an increase of IgG agalactosylation in patients with rheumatoid arthritis (RA). In the following years of research, IgG glycosylation was associated with a plethora of different conditions, including infectious, autoimmune and alloimmune diseases, cancer, cardiometabolic and neurodegenerative conditions and other age-related disorders. Still, the exact role IgG glycosylation has in all these different conditions remains elusive. It seems that a common mechanism for many autoimmune diseases may be the modulation of the general threshold of activation of immune effector cells[11,49]. But the question remains if the IgG glycosylation represents a functional effector, a predisposition or just a side-effect of certain disease pathology. This is a matter of ongoing debate.

1.2.2. IgG glycome changes with age

IgG glycosylation gradually changes with aging. Generally, the abundance of agalactosylated and bisected glycans increases while the abundance of sialylated glycans decreases from young adulthood to old age[11,27,30,49–52]. This decline of anti-inflammatory IgG glycans is accompanied by the process of inflammaging[53]. Inflammaging is defined as a chronic, low-grade inflammation designed to protect the organism from the lifetime of antigen load and immune triggering. Unfortunately, it results in a pro-inflammatory remodeling of the immune system characteristic for the elderly[54–56]. The decline of galactosylated IgG glycans is considered a hallmark of inflammaging, but it is also hypothesized that the supposed pro-inflammatory nature of agalactosylated glycans exacerbates inflammation. This creates a feedback loop where agalactosylated glycans serve not only as a biomarker of aging but also as functional effectors[53,57]. More research is needed to dissect the functional role of IgG glycans in aging, but their correlation with the biological age of an individual put them at the forefront of healthy aging research[51,58]. A study by Krištić *et al.*[51] conducted on 5117 individuals demonstrated that three IgG glycoforms can explain up to 58% of the variance in the chronological age, which is much more than all other markers of biological age such as telomere length can explain. More importantly, it turned out that the remaining variance strongly correlated with physiological parameters associated with biological aging. From these three IgG glycoforms, a biomarker of biological age, the glycan clock, was developed[51]. New studies also reported a possibility of the glycan clock reversal by healthy lifestyle changes which made IgG glycans one of the most alluring predictors of health status in modern medicine[59–61].

1.2.3. Sex hormones affect IgG glycosylation

Research on the effect of aging on IgG glycosylation revealed sex-specific differences between men and women[27,58,62,63]. Until the onset of menopause, women exhibit lower levels of agalactosylated glycans than men. These levels increase abruptly at the onset of menopause and are accompanied by the decrease of sialylated glycans. This is the reason why older men have slightly higher levels of galactosylated and sialylated glycoforms than women of the same age[63,64]. It seems that estrogen may be the main factor responsible for these sex-dependent differences[65]. In 2017, Ercan *et al.*[65] proved a causal relationship between sex hormones and IgG galactosylation. The study demonstrated that estrogen agonist therapy in postmenopausal women reduces agalactosylated IgG glycoforms compared to placebo. The same effect was observed in healthy women with therapeutically induced menopause after they were given

transdermal estradiol. This study was later expanded to analyze other derived traits, which demonstrated that sialylation of IgG is also mediated by estrogen[66]. Another recent study confirmed that estrogen supplementation affects IgG sialylation in female RA patients and postmenopausal mice models[67]. Pregnancy is also characterized by extensive changes in IgG glycosylation. An increase of IgG galactosylation and sialylation in the transition from 1st to 2nd trimester is associated with an increase of human chorionic gonadotropin (hCG)[68]. Remarkably, this pregnancy-associated increase of IgG galactosylation is accompanied by reduced disease severity in RA patients suggesting a functional role of galactosylation in RA pathology[69–71]. Unfortunately, the mechanisms underlying shifts in IgG glycome composition that are due to differential sex hormone signaling are mostly unknown and more research is required.

1.3. Regulation of IgG glycosylation

Hundreds of studies done in the past 30 years revealed the impact of IgG glycosylation on different human inflammatory pathways implicated in various pathologies. The focus has now turned to the regulation of IgG glycosylation – the largest unsolved problem of IgG biology up to date. Substantial inter-individual variability of IgG glycome composition indicates a very complex regulation of IgG glycosylation[30,47]. Glycan biosynthesis begins in the endoplasmic reticulum (ER) where sugar monomers assemble onto a dolichol pyrophosphate located on the cytoplasmic side of the ER membrane. This structure is then turned to the luminal side of the ER membrane and further extended to form a precursor glycan. This process is accompanied by the concurrent biosynthesis of IgG polypeptide. Precursor glycan is then transferred from dolichol-pyrophosphate to the Asn recognition site on the growing IgG polypeptide. After the removal of three sugar moieties, IgG is transferred from the ER to the Golgi apparatus where different glycosyltransferases and glycosyl hydrolases remodel branched glycan while IgG gets transferred from *cis* to *trans* Golgi compartments. Ultimately, the formed IgG glycoform is secreted from the B cell into the serum or is attached to the plasma membrane to form a B cell receptor (BCR)[72,73]. Variations in this process account for the observed heterogeneity of the IgG glycome. The expression and activity of the enzymatic machinery as well as the availability of the activated sugar donors results in the diversification of the IgG glycome composition. This is further complicated by the indirect upstream processes governing ER and Golgi function. For a long time, our understanding of mechanisms that regulate IgG glycosylation was very limited

due to their complexity and lack of appropriate technology. In the recent years, new advances have been made which will be discussed in the following sections.

1.4. Heritability of IgG glycans

Heritability is defined as the proportion of variation in a particular trait that is attributable to genetic factors. Heritability analyses are often the first step in a genetic study of a complex trait to determine if the trait is suitable for gene mapping[74]. The first heritability study on IgG glycans was conducted more than a decade ago to determine the proportion of genetic influence on each IgG glycoform[30]. It turned out that the heritability estimates greatly varied between different IgG glycoforms. Sialylation appeared to be most endogenously defined glycan trait with up to 60% of variance explained by genetic factors. In the case of core fucose, bisected GlcNac and galactose heritability estimates varied from 25% to 45% indicating a moderate influence of genetic factors. A few years later Menni *et al.*[75] performed another heritability study of IgG glycans on 530 twin pairs from the TwinsUK registry. They found that 51 out of the 76 analyzed glycan traits to be at least 50% heritable, indicating strong genetic control of IgG glycosylation pathways. Several other studies have demonstrated high heritability estimates for the whole plasma glycans as well[76,77]. Discovery that IgG glycans are under tight genetic control paved the way for future research into mechanisms that govern IgG glycosylation.

1.5. Role of glycosyltransferases in IgG glycosylation

The IgG glycan core is extended in the Golgi apparatus to form complex IgG glycoforms[72]. Heterogeneity of IgG glycome is partially due to the differential expression of four main glycosyltransferases (GTs) residing in the Golgi cisternae: FUT8, MGAT3, B4GALT1 and ST6GAL1[4,78] (Figure 1). The enzyme α-(1,6)-fucosyltransferase (FUT8) catalyzes the transfer of L-fucose from GDP-fucose to the most proximal GlcNac, a reaction termed core fucosylation[79]. The human genome codes for 13 different fucosyltransferases but only FUT8 performs core fucosylation[4]. This was demonstrated in different cell lines and knockout (KO) mouse models[80,81]. Disruption of both *Fut8* alleles in Chinese hamster ovary (CHO) cells resulted in the secretion of completely afucosylated mAbs with dramatically increased ADCC[81]. This phenomenon is exploited in cancer therapeutics with dozens of afucosylated mAbs currently in clinical trials[82,83]. Inactivation of FUT8 in KO mice models results in defective core fucosylation accompanied by a high mortality rate, severe respiratory

defects and growth restriction[80]. Patients with congenital FUT8 disorder of glycosylation (FUT8-CDG) also show growth retardation and respiratory deficiencies due to the complete loss of core fucosylation[84–86]. Furthermore, FUT8 is strongly linked to cancer cell invasion and metastasis, which can be partly explained by the alteration of ADCC[87,88]. ADCC is also affected by the addition of bisecting GlcNAc, a modification catalyzed by the β-1,4-Mannosyl-Glycoprotein 4-β-*N*-Acetylglucosaminyltransferase (MGAT3) enzyme[4,36]. Bisecting GlcNAc seems to have a regulatory role in the biosynthesis of complex glycans by inhibiting the action of other GTs such as FUT8, MGAT4, MGAT5 and ST6GAL1[89–91]. Garcia-Garcia *et al.*[89] have hypothesized that bisection precludes fucosylation by FUT8 due to significant steric hindrance exerted by this modification. Still, the inhibitory role of bisecting GlcNAc has yet to be demonstrated for IgG glycans. For example, Dekkers et *al.* upregulated MGAT3 to increase IgG bisection and no change in IgG fucosylation was observed[24]. Moreover, Kapur et al. reported a weak positive correlation between fucosylation and bisection on anti-HPA1 IgG antibodies[92]. The interdependence of glycosylation pathways is also highlighted through the addition of galactose and sialic acid catalyzed by β-1,4-Galactosyltransferase 1 (B4GALT1) and ST6 β-Galactoside α-2,6-Sialyltransferase 1 (ST6GAL1). These two GTs work in succession as ST6GAL1 can only sialylate galactosylated glycans[4]. More importantly, a recent study by Khoder-Agha *et al.*[93] demonstrated that B4GALT1 and ST6GAL1 form heteromers in the Golgi membranes. The study reports that the interaction of these two enzymes is a prerequisite for their full catalytic activity. Assembly of B4GALT1 homomers keeps this GT minimally active until it reaches the Golgi compartment where it binds ST6GAL1, thus revealing its active site. Many different KO and overexpression studies *in vitro* confirmed the role of B4GALT1 in IgG galactoslyation[94–97]. Moreover, a decrease of galactosylated IgG glycans was associated with the reduced activity of B4GALT1 in the peripheral B cells of RA patients[98–100] and myeloma cells[101]. However, other studies of RA patients found no association or even negative correlation between B4GALT1 mRNA expression and/or activity of this GT and the level of IgG galactosylation[102–104]. This lack of consistency in the expression data for B4GALT1 suggests a far more complex regulation of IgG galactosylation that has yet to be discovered. An even more complicated story surrounds IgG sialylation by ST6GAL1. Germ-line ablation of ST6GAL1 removes α2,6-linked sialic acid from almost all glycoproteins, including IgG[105]. In 2016, Jones *et al.*[106] generated B cell-specific ST6GAL1 KO mice, which exhibited completely normal IgG sialylation. They hypothesized that the hepatocyte-secreted plasma ST6GAL1 can sialylate IgG glycans extracellularly. In 2022, Oswald *et al.*[107] dismissed this hypothesis by the establishment of the

hepatocyte-specific ST6GAL1 KO mice that also exhibited normal IgG sialylation despite the complete lack of plasma ST6GAL1. Taken together, these results indicate that IgG sialylation can occur after secretion from the B cells, but the process is not only reliant on plasma ST6GAL1[108].

The complexity of mechanisms that regulate IgG glycosylation is additionally intensified by the epigenetic regulation of GTs activity, mainly through DNA methylation and microRNAs (miRNAs). Epigenetic regulation of MGAT3 expression has been extensively studied in the last decade. Treatment of hepatocellular carcinoma (HCC) and ovarian cells with 5-aza-2′-deoxycytidine (5-aza-dC), a nucleoside analogue that induces global hypomethylation, increased MGAT3 expression due to the decreased level of *MGAT3* promotor methylation. This intervention increased the ratio of bisecting GlcNAc on membrane glycoproteins in ovarian carcinoma cells and decreased the ratio of complex tetraantennary and fucosylated glycans in HCC cells[109–111]. An opposite effect was observed with targeted methylation of *MGAT3* promoter with CRISPR/dCas9-DNMT3A molecular tool, designed to specifically methylate a distinct region of interest. This led to MGAT3 downregulation in ovarian adenocarcinoma cell line and a concomitant decrease of bisected glycan residues[112]. Another study negatively correlated the *MGAT3* promoter methylation with bisected, galactosylated and sialylated IgG glycans in peripheral blood mononuclear cells (PBMCs) and CD19+ B cells isolated from patients suffering from inflammatory bowel diseases (IBD)[113]. Besides DNA methylation, MGAT3 appears to be regulated with miRNAs as demonstrated recently by Zohora and Mahal (2022)[114]. MiRNA regulation was most extensively studied in relation to core fucosylation. In HCC cells, four miRNAs regulate FUT8 activity and their overexpression leads to FUT8 downregulation and a concomitant decrease of core fucosylation[115,116]. A single miRNA molecule affects FUT8 activity in colorectal carcinoma (CRC) cells, as was demonstrated *in vitro* and *in vivo* using murine models[117]. New studies also link miRNA regulation with B4GALT1 and ST6GAL1 in chronic myeloid leukemia and HCC cell lines[118,119]. Taken together, this growing number of studies performed on GTs, point to the very intricate mechanisms that regulate IgG glycosylation.

1.6. Genome-wide association studies (GWAS) reveal complex gene networks regulating IgG glycosylation

The field of human genetics aims to discover the genetic factors underlying specific processes in the human body. A widely used approach are genome-wide association studies

(GWAS) that can identify common genetic variations associated with a certain phenotype[120,121]. The era of GWAS began in the 21st century when the Human Genome Project released the first human genome sequence[122]. This enabled the development of commercial Single Nucleotide Polymorphism (SNP) arrays indispensable for GWAS[122,123]. SNP is a single nucleotide change in the DNA, which occurs in the human genome with high frequency[124,125]. GWAS tests each SNP for an association with the phenotype of interest at extremely high resolution[120,121]. In the last decade, reliable high-throughput methods for glycan analysis became available and together with GWAS, enabled the identification of complex gene networks regulating IgG glycosylation[126–128].

The first GWA study of IgG glycosylation was performed by Lauc *et al.* in 2013[129] on 77 IgG glycome traits measured in 2247 individuals of European descent. The study revealed nine genomic loci associated with IgG glycosylation with four of them encoding the aforementioned glycosyltransferases. Five additional loci (*IKZF1, IL6ST-ANKRD55, SUV420H1, SMARCB1-DERL3* and *ABCF2-SMARCD3*) had no apparent role in IgG glycosylation. However, most of the discovered loci showed pleiotropy with autoimmune diseases and hematological cancers. To broaden the list of novel loci associated with IgG glycosylation, another GWAS was performed on the ORCADES (Orkney Complex Disease Study) cohort[130]. The study managed to replicate five previously associated loci and discovered five novel loci: *IGH, ELL2, HLA-B-C, FUT6-FUT3, AZI1* and *TMEM105* associated with IgG glycosylation. In 2018., Wahl *et al.*[131] augmented the list with one more region containing the *RUNX3* gene. The outcomes of different GWA studies indicated the complexity of IgG glycosylation and the polygenic nature of the process. In order to detect additional genomic loci with a smaller impact on IgG glycome composition a larger sample size was used in the GWAS study by Klarić *et al.*[10]. The authors tested 77 IgG glycan traits on 8090 individuals from four European cohorts. This was the largest GWAS of IgG glycans up to date and it helped to identify 14 novel loci and to replicate 13 previously discovered loci associated with IgG glycosylation (Figure 2). They also performed gene-set enrichment analysis which demonstrated the association of discovered GWAS loci with processes such as glycosylation, immune pathways and transcription. A putative functional network of IgG glycosylation with the strongest evidence for the regulation of *MGAT3* by chromatin remodeling factor SMARCB1 and transcription factors RUNX1 and RUNX3, was proposed. The second strongest evidence from the network analysis was for the regulation of *FUT8* by transcription factors IKZF1 and IKZF3. The latest GWAS on IgG glycans has found three novel associations (*TNFRSF13B,*

OVOL1/AP5B1, *RNF168*), which increased the list of genomic loci associated with IgG glycosylation to 29[132]. All these GWA studies revealed the complexity of pathways regulating IgG glycosylation, with possibly hundreds of genes operating together to fine-tune the immune response mediated by IgG antibodies. Yet, it is essential to perform *in vitro* and *in vivo* functional studies of novel loci to further confirm the role these genes have in IgG glycome composition.

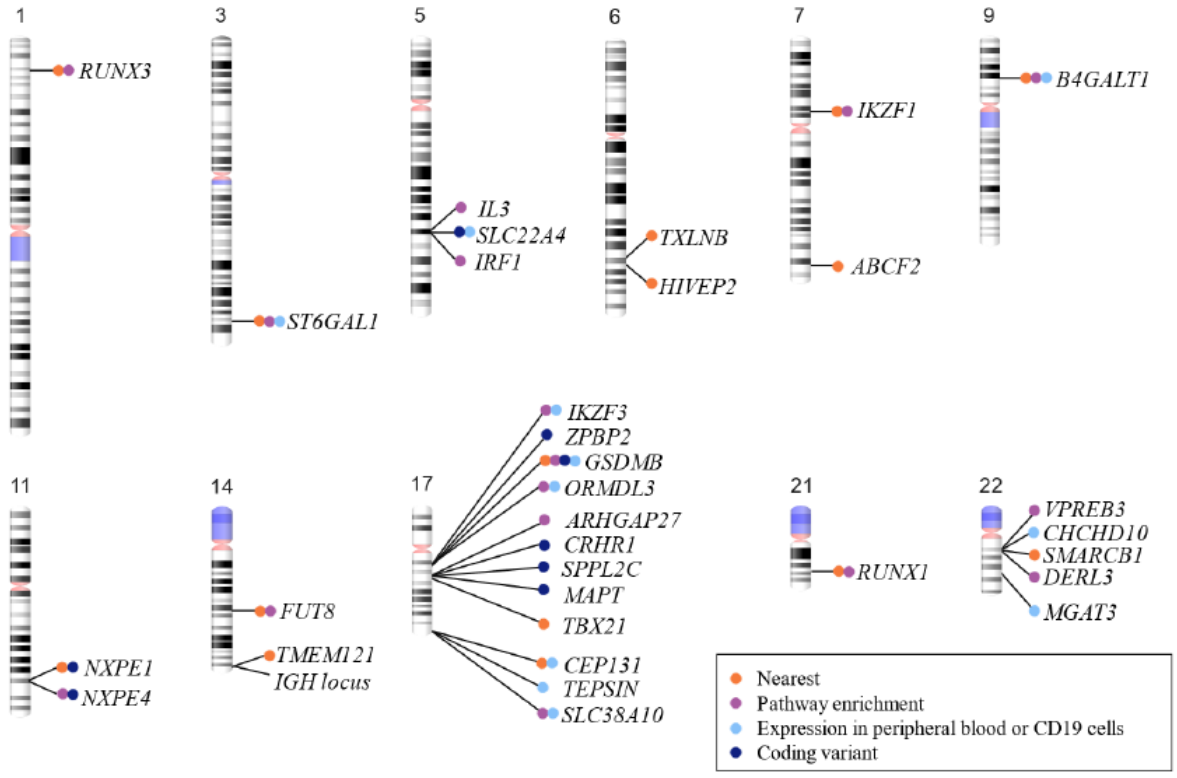

Figure 2. Chromosomal location of GWAS candidate genes associated with IgG glycosylation from the study by Klarić *et al*[10]. The gene prioritization was performed as follows: orange dots – nearest to the strongest association in the region; purple dots – biological pathway enrichment; light blue dots – pleiotropy with gene expression in peripheral blood or B cells; dark blue dots – missense mutation. The image was adapted from Pezer *et al*[133].

1.6.1. Functional follow-up of GWAS loci

Identification of novel loci through GWAS together with the results of heritability analysis on IgG glycans revealed the complexity of genetic control governing IgG glycosylation[10,30,75,129,131,132]. However, GWAS is a hypothesis generating method and every discovered locus needs to be functionally validated *in vitro* or *in vivo*. Klarić *et al*.[10] performed a functional follow-up of their GWA study in B cell-derived lymphoblastoid cell line MATAT6.

Knockdown of *IKZF1* resulted in the downregulation of *IKZF3* and upregulation of *FUT8* with a concomitant increase of IgG fucosylation. This was the only functional study performed on discovered GWAS hits with all of the remaining loci validated only with *in silico* methods. One of the main obstacles was the unavailability of methods for fast and efficient glycan analysis, though significant progress was made in the last decade with the development of high-throughput platforms for glycan analysis[126–128]. The lack of appropriate experimental models for *in vitro* or *in vivo* studies of IgG glycans is an issue to this day. Krištić *et al.*[134] report that common laboratory mouse strains are not optimal animal models to study the effects of IgG glycosylation due to very low intra-strain variability of IgG glycome composition. Moreover, most mouse strains lack bisecting GlcNAc in their IgG glycoprofile and their glycans terminate with *N*-glycolylneuraminic acid instead of *N*-acetylneuraminic acid (sialic acid)[135,136]. Mouse IgG glycans also harbor terminal α 1-3 bound galactose which is not present on human IgG glycans[135]. Problems arise with *in vitro* cell cultures as well. The main difficulty is the quantity of cell culture-produced IgG required for high-resolution glycan analysis. Another issue is the inability to transfect B cells with classical transfection methods. Recently, Vink *et al.*[137] developed a robust expression system for IgG production based on FreeStyle™ 293-F cells. They succeeded to obtain large quantities of glycosylated IgG in less than a week secreted from an easy to transfect FreeStyle™ 293-F cell line. Three years later, Dekkers *et al.*[94] expanded this IgG expression system with different glycoengineering tools to produce any desired IgG glycoform which enabled them to systematically explore the effect of certain glycosylation traits on IgG function. This system also holds promise for future functional validation of novel loci associated with IgG glycosylation due to its simplicity and efficacy in antibody production. Another, maybe the most important reason for the lack of functional follow-up studies was the unavailability of technology for targeted gene manipulation. This issue was resolved by the discovery of CRISPR/Cas9 technology for precise genome editing discussed in the next section.

1.7. CRISPR/Cas9 molecular tools

CRIPSR/Cas9 stands for Clustered Regularly Interspaced Short Palindromic Repeats/CRISPR Associated Protein 9, a bacterial immune system adapted as a highly efficient, RNA-guided molecular tool for precise genome editing and regulation[138]. The development of CRISPR/Cas9 technology revolutionized the whole field of molecular biology as it enabled simple and cost-efficient manipulation of desired loci. CRISPR/Cas9 system is based on two components: the Cas9 protein that can introduce a double-stranded break in targeted DNA and

the guide RNA molecule (gRNA) complementary to the target sequence (Figure 3a)[139]. The best characterized Cas9 protein is the one isolated from the bacterium *Streptococcus pyogenes* (SpCas9)[140]. Cas9 nuclease activity is provided by two distinct domains: RuvC and HNH each able to cut single-stranded DNA[138,140]. The PAM (protospacer adjacent motif) interacting domain initiates DNA binding by the Cas9 protein through the recognition of a specific PAM sequence (5'-NNG-3' for SpCas9) located downstream from the gRNA binding region[141]. gRNA molecule is 20 base pairs (bp) single-stranded RNA that forms a T-shaped loop and can be designed to have a 5' end complementary to the target sequence[138]. This way gRNA guides Cas9 protein to any target sequence that contains complementary DNA and PAM sequence specific for the desired Cas9 ortholog. Another commonly used Cas9 ortholog is SaCas9, isolated from the bacterium *Staphylococcus aureus*[142]. The main advantage of SaCas9 is its smaller size and higher specificity for target DNA conferred by the more complex PAM sequence 5'-NNGRRT-3'[143–145].

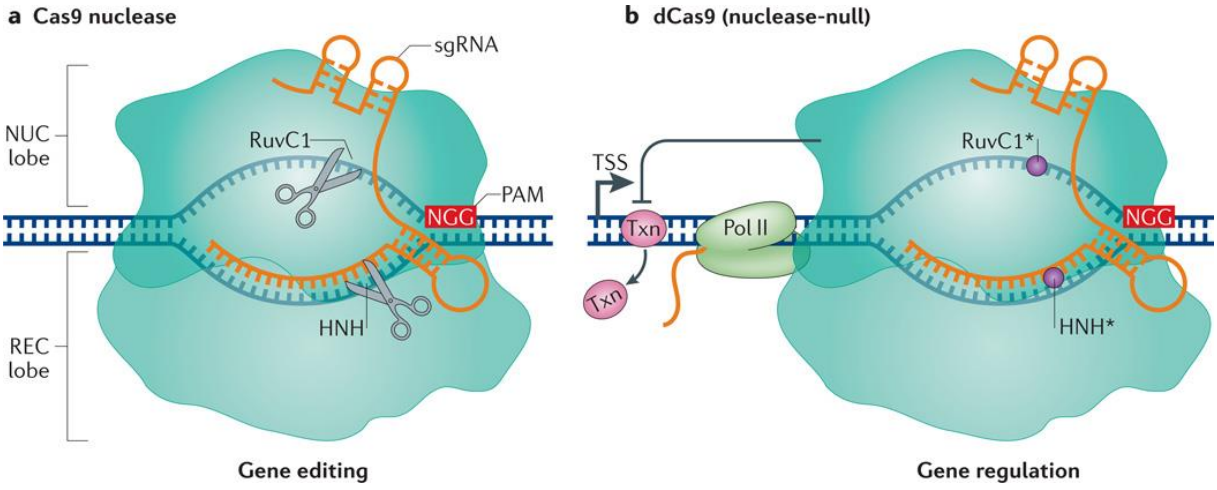

Figure 3. CRISPR/Cas9 vs. CRISPR/dCas9 technology. (a) CRISPR/Cas9 system for precise gene editing. The Cas9 protein isolated from *Streptococcus pyogenes* (SpCas9) consists of a nuclease lobe (NUC) and a recognition lobe (REC). The Cas9 is targeted to a specific DNA sequence by complementary single guide RNA (sgRNA or gRNA). For efficient binding, the target sequence has to contain a protospacer adjacent motif (PAM) specific for the Cas9 ortholog. Cas9 binding to the target sequence by gRNA complementarity is followed by DNA cleavage introduced by RuvC and HNH domains. **(b)** CRISPR/dCas9 system for precise gene regulation. Selective mutations D10A and H840A of the RuvC and HNH domains are introduced to Cas9 to create catalytically inactive dead Cas9 (dCas9) protein. The dCas9 protein still retains the ability to bind target sequence through gRNA and PAM but cannot initiate DNA cleavage. If dCas9 protein is targeted downstream of the transcription start site (TSS) it blocks

transcription by interfering with RNA polymerase (Pol II) and different transcription factors (Txn). Image adapted from Dominquez *et al*[139].

Applicability of the CRISPR/Cas9 technology substantially increased with the development of the catalytically inactive dead Cas9 (dCas9) protein (Figure 3b)[139,146]. This was accomplished by selective mutations D10A and H840A of RuvC and HNH domains, respectively. The dCas9 protein retains the ability to bind targeted DNA sequences as well as the gRNA molecule but cannot introduce double-stranded breaks[146]. Fusion of various epigenetic effector domains to dCas9 enabled site-specific epigenome modifications which greatly expanded the utility of CRISPR/dCas9 tools[147–157]. CRISPR/dCas9 system is most widely used for targeted transcriptional regulation: CRISPR interference (CRISPRi) and CRISPR-mediated gene activation (CRISPRa) for targeted downregulation and upregulation of genes, respectively[146,150,151,153,158] (Figure 4). When this technology was developed CRISPRi relied solely on dCas9 interference through steric blockage of various transcription factors and RNA polymerase. This strategy for gene silencing was very successful in prokaryotes but exerted only modest repression in eukaryotic cells[153,159]. The system was greatly advanced by fusing KRAB (Krüppel associated box) domain to inactivated Cas9 protein which improved gene repression from 50% to 99% depending on the locus and/or cell type[153,160]. The KRAB domain interacts with the scaffold protein KAP1 (KRAB associated protein 1), which recruits HP1 (heterochromatin protein 1). HP1 then recognizes and binds methylated lysine 9 of the histone H3 (H3K9me) recruiting additional histone methyltransferases, which leads to chromatin condensation and consequently gene repression[161,162]. Gene silencing using the dCas9-KRAB fusion protein was recently improved by tethering an additional domain, MeCP2 (methyl CpG binding protein 2), yielding a superior CRISPRi tool dCas9-KRAB-MeCP2, which demonstrated robust gene repression[152]. Many other effector proteins were fused to dCas9 to achieve gene silencing including DNA methyltransferases[149], histone methyltransferases[156], histone deacetylases[155], chromatin remodeling complexes[163], etc. Similar progress was made in the field of gene activation with different CRISPRa tools. In 2013. Gilbert *et al.*[153] reported targeted gene activation in the HEK293 cell line with dCas9-p65 and dCas9-VP64 fusion proteins. P65 is a large subunit of NFκB transcription factor known to induce strong gene transactivation in mammalian cells[164]. VP64 is also a potent transcriptional activator, a tetramer of the minimal activator VP16 (virion protein 16) isolated from the herpes simplex virus type 1[165]. Soon after the publication of Gilbert *et al*., another group demonstrated that dCas9-VP64 activity can further be improved by targeting desired loci with multiple

gRNAs[151]. The most important breakthrough came with the design of dCas9-VPR, a tripartite gene activator composed of VP64, p65 and Rta activator domains[150]. Chavez *et al.*[150] demonstrated that dCas9-VPR drastically outperforms dCas9-VP64 in robust multi-locus activation. Today, it is possible to achieve targeted gene upregulation with an array of different effector domains tethered to the dCas9 protein such as DNA demethylases[157], histone acetyltransferases[154], histone demethylases[166], etc. One of the finest examples on how this novel array of CRISPR/dCas9 based molecular tools can be exploited came from a recent study by Josipović *et al.*[112]. The authors established a modular and extensible CRISPR/dCas9 based molecular toolbox for epigenetic editing and direct gene regulation. Designed toolbox features the system for the expression of orthogonal dCas9 proteins fused to various effector domains and a multi-gRNA system for simultaneous targeting of up to six loci (Figure 5). The flexibility of this molecular toolbox enables an easy assembly of different promoters, dCas9 orthologs, effector domains, selection markers and gRNAs to suit any experimental design.

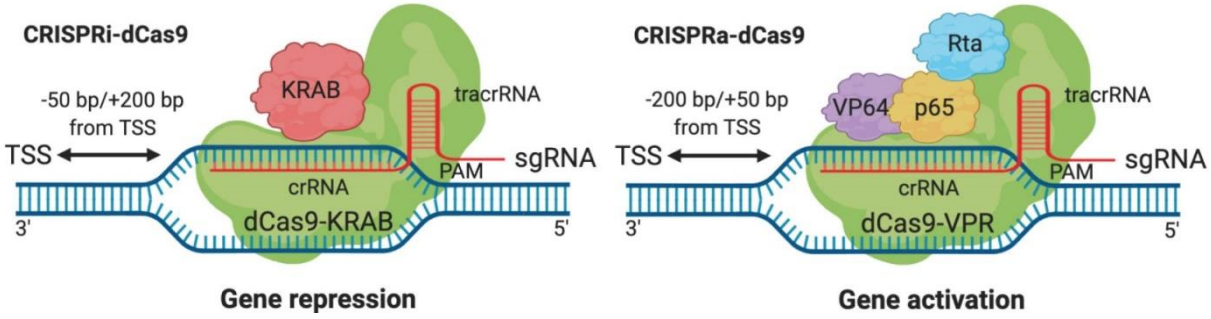

Figure 4. The dCas9-based molecular tools for gene repression (CRISPRi) and gene activation (CRISPRa). Catalytically inactive dCas9 protein can be fused with different effector domains for precise manipulation of gene expression through complementary single guide RNA (sgRNA or gRNA) and matching protospacer adjacent motif (PAM). The most widely used CRISPR/dCas9 tool for gene repression is dCas9-KRAB, which blocks gene expression if targeted around the transcription start site (TSS) in the range from -50 bp to +200 bp. Gene activation (CRISPRa) can be achieved with dCas9-VPR fusion protein. VPR is a tripartite activator comprised of VP64, p65 and Rta domains. dCas9-VPR upregulates gene expression if targeted around the TSS in the range from -200 bp to +50 bp. tracrRNA – transactivating RNA; crRNA – CRISPR RNA. Image adapted from Escobar *et al*[167].

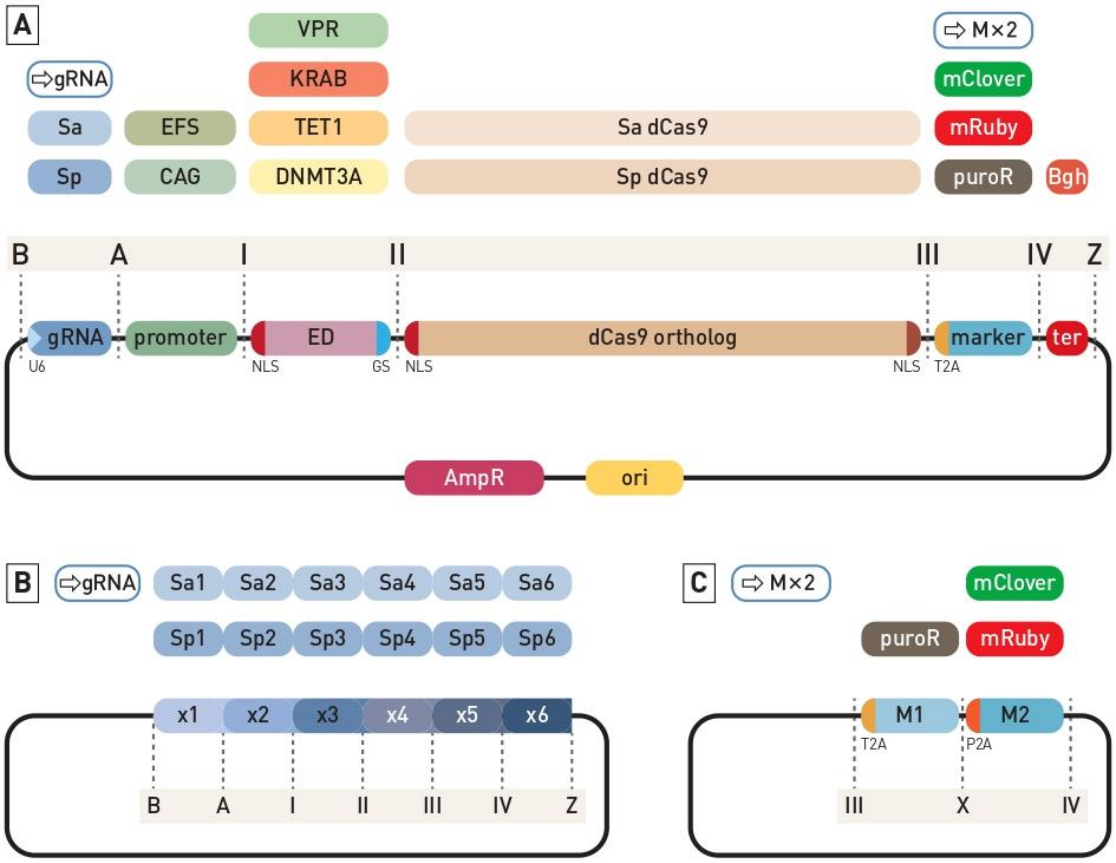

Figure 5. Modular CRISPR/dCas9 based molecular toolbox. (A) Individual modules required for the correct assembly of CRISPR/dCas9 expression cassette are cloned into a backbone plasmid using BsaI type IIS restriction enzyme (Golden Gate cloning). The ends of the backbone plasmid are designed to be compatible with B and Z type module ends. The first position (B-A) receives a multi-guide RNA module with SaCas9 or SpCas9 scaffold. The second position (A -I) receives a eukaryotic promoter followed by the desired effector domain (I -II) and desired dCas9 ortholog (II -III). The next position (III-IV) receives an antibiotic resistance or fluorescence marker which can be replaced with the dual marker system (described in C). The final position (IV-Z) receives a eukaryotic transcription terminator. **(B)** The multi-guide system accepts up to six gRNA modules for SaCas9 or SpCas9 targeting, each gRNA module with the appropriate ends (B through Z) required for the Golden Gate assembly facilitated with the type IIS restriction enzyme Esp3I. **(C)** The dual marker system enables simultaneous expression of both the fluorescent marker and antibiotic resistance. U6 – eukaryotic promoter for the expression of gRNAs; NLS – nuclear localization signal; ED – effector domain; GS – linker; T2A and P2A – self-cleaving peptides; ter – terminator; AmpR – bacterial ampicillin resistance; ori – bacterial origin of replication. Image adapted from Josipović *et al*[112].

Precise gene editing with CRISPR/Cas9 based molecular tools still has many hurdles. One of the major unresolved problems is the off-target effect. The Cas9 protein can bind and cleave unintended genomic sequences, potentially resulting in harmful genomic mutations and rearrangements[168–171]. Many strategies have been developed to mitigate Cas9 off-target effect, including the design of more specific gRNA molecules, tight control of gRNA and Cas9 expression, use of more specific Cas9 orthologues, engineering of mutant Cas9 proteins, etc[112,143,171–174]. Fortunately, the off-target effect is much less prominent in CRISPRa and CRISPRi technology when dCas9 is used for direct transcriptional regulation of gene expression. Chromatin immunoprecipitation and sequencing analysis (ChIP-seq) revealed that dCas9 binds to various off-target sites in the genome with different gRNA molecules[175–177]. Despite the non-specific binding of dCas9, dCas9-KRAB mediated gene expression turned out to be very specific[153,178]. Further studies reported that only three mismatches between gRNA and the targeted loci completely abolish CRISPRi gene repression[179]. In comparison, Cas9 can tolerate up to 5 mismatches.[168,180] CRISPRa tools also seem to be very specific[153,179,181,182]. Polstein *et al.*[181] demonstrated high specificity in targeted gene upregulation with dCas9-VP64 fusion protein. Another ChIP-seq analysis of the off-target effect, following dCas9-VP64 and dCas9-KRAB targeting with multiple gRNAs revealed a similarly high incidence of mismatched binding for fusion proteins compared to solo dCas9 binding. However, the on-target effect turned out to be very specific[176,178,181]. This is probably due to the fact that direct transcriptional regulators exert their function in a small region around the transcription start sites (TSS) in gene promoters where they interact with transcription factors and RNA polymerase[179]. It is important to mention that mere overexpression of catalytic domains fused to dCas9 protein can also contribute to the off-target effect, independently from dCas9 binding promiscuity. This can be resolved by performing dose-dependent experiments that limit the expression of the fusion proteins[147]. Even though the off-target effect of CRISPRa/CRISPRi tools is less prominent, researchers use different strategies to minimize it further to prevent any harmful consequence of non-specific gene deregulation.

1.8. *In vitro* delivery of CRISPR/Cas9 molecular tools

The specificity of CRISPR/Cas9 molecular tools in gene editing remains a significant obstacle, although to a lesser extent in analyses performed *in vitro*. For *in vitro* studies, the major issue is the delivery of large CRISPR/Cas9 cassettes to the cell type of interest[183–185]. This is particularly emphasized in the field of epigenome editing due to the size of different

effector domains fused to the dCas9 protein[148]. CRISPR/Cas9 cassettes can range in size from 9 to 19 kb, particularly when multiple gRNAs are used together with effector domains fused to the dCas9 protein[186]. This makes efficient delivery of CRISPR/Cas9 elements to the target cells notoriously difficult. Many studies also require the genomic integration of CRISPR/Cas9 transgenes for their long-term expression. Currently, CRISPR/Cas9 cargo is delivered *in vitro* using three main approaches: physical methods, non-viral vectors and viral vectors. Mostly used physical methods of delivery are microinjection and electroporation. Microinjection is not size limited and is highly effective but also technically complex and laborious. In contrast, electroporation is simple but not as effective as microinjection and can impact cell viability[185,187–189]. The downside of both methods is that they provide only transient expression of transgenes. Non-viral methods mainly encompass chemical reagents like lipid-based nanoparticles and non-lipid polymers[190]. Both reagents exploit their positively charged surface to encapsulate the negatively charged DNA and transport it across the cell membrane. Non-viral chemical methods of delivery are safe and simple to implement but are inefficient when compared to physical and viral methods of delivery[191–196]. Like physical methods, they cannot facilitate genomic integration of transgenes. On the other hand, mostly used viral vectors such as lentiviruses (LVs) and adeno-associated viruses (AAVs) have the capacity to insert their DNA cargo into the cell's genome[197–200]. Viral strategies are highly efficient but can be laborious and demand more strict laboratory safety measures compared to non-viral and physical methods. Also, AAVs are limited in their cargo capacity (~4.7 Kb)[201,202], while LVs have a high incidence of insertional mutagenesis[203]. Another CRISPR/Cas9-based delivery method that enables long-term transgene expression combines the simplicity of chemical non-viral methods with the capacity for genomic integration without the cargo limit. This method is based on DNA transposons, genetic elements that have the ability to change their position within the genome[204].

1.8.1. Transposon vectors

DNA transposons are mobile genetic elements that can change their position in the genome by *cut and paste* mechanism of transposition[205]. They code for the enzyme transposase, which acts on sequences flanked by inverted terminal repeats (ITRs). Transposase specifically recognizes ITRs and promotes the assembly of paired-end complex (PEC) to catalyze the excision of DNA transposon and its integration into a new location in the target DNA[206]. Transposase can also act *in trans*, facilitating excision and relocation of any transgene flanked

by ITRs and this discovery led to the development of transposon vectors which have the capability to integrate any expression cassette into the genome without the use of viral delivery methods[207,208]. The most widely used transposon vectors are derived from *Sleeping Beauty* (SB) and *PiggyBac* (PB) transposons. Both are dual vector systems comprised of a helper plasmid for transient transposase expression and a donor plasmid carrying an expression cassette flanked by ITRs[204,209] (Figure 6). The main advantage of SB vectors is their integration preference for intergenic regions, which lowers the risk of unwanted insertion events that may lead to harmful gene deregulation[210,211]. On the other hand, PB vectors have a large cargo capacity that can reach up to 200 kB[212], while SB integration efficiency decreases with the increasing length of transgenes[213]. Overall transposon vectors offer many advantages over viral vectors. Number one is the ease of their manufacturing and handling in comparison with the production and handling of viral vectors, which is often expensive and cumbersome[214]. Secondly, the cargo capacity of transposon vectors that substantially exceeds the cargo capacity of viral vectors makes them an excellent tool for the delivery of large CRISPR/Cas9 cassettes for gene editing[215]. In contrast to different physical and chemical non-viral methods of delivery, they enable stable genomic integration of transgenes. The stability of transgene expression can also be increased by flanking the vector with different insulator sequences that prevent silencing[216]. Sharma *et al.*[217] reported that cHS4 insulator increases transgene expression when coupled with SB and PB vector. Bire *et al.*[218] utilized another insulator type, CTF/NF1 and D4Z4, that enabled the long-term transgene expression of transgenes from the PB vector. Transposon vectors can be coupled with CRISPR/Cas9 molecular tools as well. The best example of the efficiency and simplicity of this technology comes from the latest study by Jiang *et al.*[219] where they generated a stable human pluripotent stem cell (hPSC) line utilizing PB vector that carried CRISPR/Cas9 tools. This system enabled robust genome and RNA editing to investigate gene functions in hPSC.

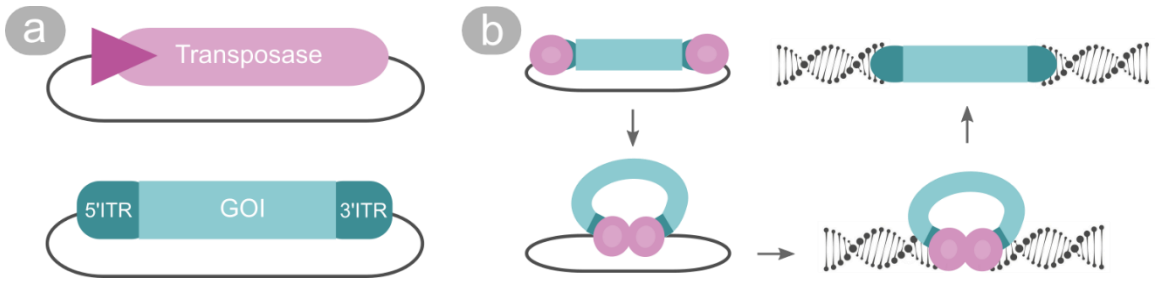

Figure 6. Transgene delivery using transposon vectors. (a) Transposon dual vector system consists of two plasmids. Pink colored: transposase expression plasmid; blue colored:

integrating plasmid with any gene of interest (GOI) flanked by inverted terminal repeats (ITRs). (b) Mechanism of transposon vector genomic integration. Following the delivery of the dual transposon vector system to the target cells, transposase (pink circles) is synthesized from the expression plasmid. Transposase binds ITRs (dark blue) and catalyzes the excision of the GOI (light blue) and its integration into the genome of the target cells.

1.9. CRISPR/Cas9 in post-GWAS analysis

Since GWAS is just a hypothesis generating method that associates a specific locus with a trait or a disease the next step is the validation of the functional relevance of the identified locus for a certain phenotype. Different statistical methods and genomic functional annotations are being applied to prioritize causal variants and their target genes, but *in vitro* functional studies still lag behind. However, with the progress of novel molecular tools for precise gene editing the post-GWAS analysis of gene function has been translated to *in vitro* setting[220–222]. Recently, Werder *et al*.[223] exploited CRISPRi to interrogate the function of 9 loci associated with chronic obstructive pulmonary disease (COPD) in induced pluripotent stem cell-derived type 2 alveolar epithelial cells (iAT2s). In this system, they demonstrated that the desmoplakin gene regulates cell-to-cell junctions, proliferation, mitochondrial function and cigarette smoke-induced injury. Another example of a post-GWAS functional study with CRISPRi comes from COVID-19 research. Targeting the two non-coding causal variants revealed their association with *LZTFL1* and *RAVER1* loci involved in airway cilia regulation and antiviral response, respectively[224]. The number of CRISPR-validated GWAS loci is expanding rapidly, with many examples within various fields of biology. However, this technology has not yet been employed in the validation of GWAS-identified regulatory networks governing IgG glycosylation.

1.10. Precise editing in glycoengineering

The development of CRISPR/Cas9 based molecular tools enabled precise genetic engineering of glycosylation in mammalian cells. This was immediately exploited for the production of therapeutic IgG antibodies with fine-tuned glycan patterns[225]. IgG antibodies with defined glycome composition are commonly produced in CHO and HEK293 cell lines as they are easy to handle and have similar glycosylation capacities as human B cells[226]. Simple combinations of CRISPR/Cas9-mediated knock-ins and knock-outs of main GTs enabled the production of different IgG glycoforms for therapeutic purposes[96,227,228]. One of the finest examples of this experimental approach is a knock-out of *FUT8* and *B4GALT1* genes and their

exchange with synthetic versions with inducible expression in CHO cell line[228]. This enabled the production of therapeutic IgG with the desired level of core fucosylation (0-97%) and galactosylation (0-87%). Currently, precise editing technologies are becoming a standard in recombinant antibody production. However, it is even more significant that precise glycoengineering offers a completely new platform for the dissection of biological interactions that govern protein glycosylation, opening a gateway for post-GWAS studies of gene networks associated with IgG glycosylation.

2. SCIENTIFIC PAPERS

2.1. A Transient Expression System with Stably Integrated CRISPR-dCas9 Fusions for Regulation of Genes Involved in Immunoglobulin G Glycosylation

The CRISPR Journal
Volume 5, Number 2, 2022
Mary Ann Liebert, Inc.
DOI: 10.1089/crispr.2021.0089

The CRISPR Journal

RESEARCH ARTICLE

A Transient Expression System with Stably Integrated CRISPR-dCas9 Fusions for Regulation of Genes Involved in Immunoglobulin G Glycosylation

Anika Mijakovac,[1] Karlo Miškec,[1] Jasminka Krištić,[2] Vedrana Vičić Bočkor,[1] Vanja Tadić,[3] Maria Bošković,[4] Gordan Lauc,[2,5] Vlatka Zoldoš,[1] and Aleksandar Vojta[1,*]

Abstract

Alternative glycosylation of immunoglobulin G (IgG) is functionally important in multiple human physiological and pathological states. Our understanding of molecular mechanisms that regulate IgG glycosylation is vague because of the complexity of this process, which involves hundreds of genes. Several genome-wide association (GWA) studies have revealed a network of genes associated with IgG glycosylation that are pleiotropic for a number of diseases. Here, we report a design of a versatile system for IgG production and gene manipulations that can be used for *in vitro* functional follow-up of GWA hits or any gene of interest. The system is based on CRISPR-dCas9, extended by a piggyBac integrase compatible vector, and drives IgG production in HEK-293F cells. We validated our systems that stably express VPR-dCas9 and KRAB-dCas9 by manipulation of four glyco-genes with a known role in IgG glycosylation, and then functionally validated three GWAS hits for IgG glycosylation with an as-yet-unknown role in this process.

Introduction

Immunoglobulin G (IgG), a glycoprotein secreted by plasma cells, is the most abundant antibody in the human plasma. It plays a role in multiple humoral immune processes, such as complement activation and complement-dependent cytotoxicity, antigen neutralization, antibody-dependent cell-mediated cytotoxicity, phagocytosis, and hypersensitivity reactions. The glycan part is essential for the antibody because alternative glycosylation influences its structural stability, conformation, half-life, as well as effector functions.[1–3] Alternative glycosylation can shift antibody activity from anti-inflammatory to proinflammatory and *vice versa*, and the complete removal of *N*-glycans from IgG results in the loss of both activities.[4,5] Proper glycosylation can significantly improve efficacy of therapeutic monoclonal antibodies, and glycoengineering has an important role in drug development.[6]

A conserved glycosylation site at Asn297 of the constant Fc region of IgG enables post-translational structural change of the Fc region required for optimal binding of antibody to different Fc gamma receptors (FcγR). Any change in glycan composition at this site can affect Fc structure and alter interaction with FcγR.[7] The IgG-Fc glycan is predominantly a bi-antennary complex structure composed of two *N*-acetylglucosamine (GlcNAc) and tri-mannose core, which can further be extended by addition of galactose, sialic acid, fucose, and bisecting GlcNAc (bisection).[8] Since the relative levels of these IgG glycoforms vary between individuals,[9] IgG glycome composition is extremely heterogenic in the human population. Several hundred differentially glycosylated IgG variants can be present in an individual at any given time. Furthermore, IgG glycome varies during life and in disturbed homeostasis. Differential IgG glycosylation is implicated in various physiological[10,11] and

[1]*Department of Biology, Division of Molecular Biology, Faculty of Science, University of Zagreb, Zagreb, Croatia;* [2]*Genos Glycoscience Research Laboratory, Zagreb, Croatia;* [3]*Division of Molecular Biology, Laboratory for Cell Biology and Signaling, Ruđer Bošković Institute, Zagreb, Croatia;* [4]*Laboratory for Cancer research, University of Split School of Medicine, Split, Croatia; and* [5]*Faculty of Pharmacy and Biochemistry, University of Zagreb, Zagreb, Croatia.*

*Address correspondence to: *leksandar Vojta, PhD, Department of Biology, Division of Molecular Biology, Faculty of Science, University of Zagreb, Horvatovac 102a, Zagreb 10000, Croatia, Email:* vojta@biol.pmf.hr

Correction added on January 18, 2022 after first online publication of January 12, 2021: The article reflects Open Access, with copyright transferring to the author(s), and a Creative Commons License (CC-BY) added (http://creativecommons.org/licenses/by/4.0).

pathological states[12–15] as well as in aging.[16,17] Indeed, IgG and its glycosylation are extensively studied in various diseases on an epidemiological scale because robust methods for high-throughput glycomics and glycoproteomics analyses are available.[18]

Molecular mechanisms by which alternative IgG glycosylation is regulated are only vaguely understood. Yet, they are important in disturbed homeostasis, different pathological states and aging. The reason for this is the complexity of the glycan biosynthetic pathways, which include genes for glycosyltransferases and glycosidases that add or remove monosaccharides to/from a glycan structure, but also other enzymes as well as various other proteins such as transporters, transcription factors, and so on. Several genome-wide association studies (GWAS), including the most recent one performed on more than 10,000 individuals, revealed a network of genes associated with IgG glycosylation, which are pleiotropic with a number of different diseases and pathophysiological states.[19–22] How these loci are functionally related to IgG glycosylation and diseases through the regulation of alternative IgG glycosylation remains to be resolved.

For *in vitro* functional follow-up of GWA hits, there is a need for an appropriate model cell system that would allow gene manipulations and analysis of the resulting glycan phenotype. Such a system must produce sufficiently high levels of IgG antibodies so that IgG glycosylation can be analyzed. On the other hand, it should enable reproducible and fast manipulation of gene expression. A highly efficient and simple transient expression system, capable of producing sufficient amounts of native secreted IgG antibodies, has already been developed by Vink *et al*. and is based on the suspension cell line FreeStyle™ 293-F (HEK-293F).[23] Importantly, IgG produced by this HEK-293F transient system shows glycosylation that resembles that of normal human plasma-derived IgG.[24] We used the system developed by Vink *et al*. as a foundation, and we upgraded it with a transposon vector combined with our CRISPR-dCas9-based modular system for orthogonal gene regulation.[25] As a result, we generated a transient expression system with stably integrated VPR-dCas9 and KRAB-dCas9 cassettes for up- and downregulation of any gene of interest only by co-transfection of cells with appropriate single guide RNAs (gRNA). As a proof of principle, we selected four glyco-genes coding for glycosyltransferases with defined function in IgG glycosylation and manipulated their transcriptional activity in our newly developed stable VPR-dCas9 and KRAB-dCas9 HEK-293F cell lines. Following transcriptional modulation, we analyzed changes in glycan phenotype on secreted IgG molecules. After successful validation of our system on these selected glyco-genes, we used it to study three GWAS hits for IgG glycosylation[19,22] with as-yet-unknown role in this process.

Methods

Construction of piggyBac-based vectors for genomic integration

Plasmid pBackBone-BZ described in Josipović *et al*. (see supplement therein)[25] was used as the backbone for piggyBac (PB)-based vector construction. All restriction endonucleases were from New England Biolabs (Ipswich, MA) unless indicated otherwise. First, 48 bp complementary oligonucleotides LeftH-Esp3I_sense and LeftH-Esp3I_antisense were designed to include *Esp*3I restriction site and *Acc*65I and *Sal*I cohesive ends. The oligonucleotides were annealed and cloned in pBackBone-BZ cut with *Acc*65I and *Sal*I. Next, 51 bp oligonucleotides RightH-BbsI_sense and RightH-BbsI_antisense containing *Bbs*I restriction site and *Xba*I and *Xho*I cohesive ends were annealed and cloned in the pBackBone-BZ with *Xba*I and *Xho*I to form a plasmid named pBB-ARMs.

Plasmid pUK21gg optimized for Golden Gate cloning from pUK21[25] was used as an intermediate to clone the insulator sequences and inverted terminal repeats (ITRs) of the PB transposon system. First, 84 bp complementary oligonucleotides MCS-assemble_sense and MCS-assemble_antisense including *Bsa*I and *Esp*3I restriction sites flanked by *Xba*I and *Xho*I cohesive ends were annealed and cloned in pUK21gg with *Xba*I and *Xho*I. Next, the 146 bp insulator sequence (U.S. patent no. US8133699B2, 2012) was custom-synthetized (Macrogen Europe BV, Amsterdam, The Netherlands) as two pairs of complementary oligonucleotides containing *Nhe*I and *Xho*I cohesive ends flanking the insulator sequence. They were annealed and cloned in pUK21gg cut with *Nhe*I and *Xho*I (New England Biolabs). This was repeated with *Xba*I and *Sal*I (New England Biolabs) to generate two tandem insulator sequences.

ITRs of the PB transposon were amplified by polymerase chain reaction (PCR) with PB3-FW and PB3-RE primers from plasmid pBCAG-eGFP (Addgene, Watertown, MA; plasmid #40973). Next, pUK21gg was cut with *Esp*3I, and the PCR products were cut with *Bsa*I and ligated to produce two plasmids, each containing left or right ITRs and two insulator sequences, named PB-ARM-L and PB-ARM-R. ITRs and insulator sequences were inserted in the backbone pBB-ARMs in two steps. First, PB-ARM-L and PB-ARM-R were cut with *Bsa*I, while pBB-ARMs were linearized with *Esp*3I to enable insertion of left ITRs and insulator sequences. The intermediary construct was linearized

with *Bpi*I (Thermo Fisher Scientific, Waltham, MA), and right ITRs and insulator sequences were inserted to obtained complete PB backbone (donor) vector optimized for Golden Gate cloning.

Two different Cas9 orthologs were used for the production of stable cell lines: dSaCas9 from *Staphylococcus aureus* and dSpCas9 from *Streptococcus pyogenes* (dCas9 denotes ''deactivated'' Cas9 with no nuclease activity but intact DNA-binding activity). To obtain PB-VPR-dSaCas9 vector for stable genomic integration, Golden Gate modules Ig-dual bridge oligonucleotide, EFS promotor, VPR domain, dSaCas9, T2A-X-puro, X-P2A-Clover, and bGH terminator were assembled using *Bsa*I in the PB backbone vector. To obtain PB-KRAB-dSpCas9 vector for stable genomic integration Golden Gate modules Ig-dual bridge oligonucleotide, EFS promotor, KRAB domain, dSpCas9, T2A-X-puro, X-P2A-Ruby, and bGH terminator were assembled with *Bsa*I in the PB backbone vector. The Golden Gate assembly system is described in Josipović *et al*. (see supplement therein).[25]

Hyperactive PB transposase that recognizes PB ITRs was designed according to the sequence described in Doherty *et al*.[26] and codon optimized for human cell lines. All unwanted restriction sites and potential splicing donors and acceptors were then removed. The PB transposase DNA sequence was custom synthesized (Integrated DNA Technologies, Coraville, IA) and amplified with primers C9_seq1 and C9_seq2 to produce I-IV module for Golden Gate assembly named HyPB-Xpress. The purified PCR product was cloned in pUK21gg with *Xba*I and *Xho*I. HyPB-Xpress together with SV40 promotor, SV40 terminator, and Ig-dual bridge oligonucleotide was assembled with *Bsa*I in pBackBone-BZ by Golden Gate cloning, forming a PBase expression vector.

To reduce unwanted recombination of repeating sequences, all plasmids were amplified in NEB® Stable Competent *Escherichia coli* bacteria (New England Biolabs). Sequences of the used oligonucleotides and PCR primers are listed in Supplementary Table S1. Details of plasmids and modules used in Golden Gate cloning are described in Supplementary Table S2.

Construction of plasmids encoding IgG heavy and light chains

The IgG light chain (LC) DNA sequence was kindly provided by Gestur Vidarsson (Sanquin Blood Supply Foundation, Amsterdam, The Netherlands) and amplified using Herculase II Fusion DNA Polymerase (Agilent Technologies, Santa Clara, CA) with Ig-PstI-tNS-1-FW and corr-Igk-tNS-4-XhoI-RE primers (Supplementary Table S3) flanked by *Pst*I and *Xho*I sites. The purified PCR product was cloned in pUK21gg using *Pst*I and *Xho*I. A *Bsa*I restriction site in the IgG LC sequence was removed by two different approaches: using the QuickChange Lightning Site-Directed Mutagenesis Kit (Agilent) with LC_noBsaI-S and LC_noBsaI-A oligonucleotides or by substituting the complete intron sequence with LC-BbsI_rm_Intron-S and LC-BbsI_rm_Intron-A annealed oligonucleotide containing *Bbs*I cohesive ends to obtain two different constructs. The IgG LC was then amplified in both constructs with LC-BsaI-2-NcoI-FW and LC-BsaI-2-KasI-RE using Herculase II Fusion DNA Polymerase (Agilent) to obtain *Bsa*I restriction sites, leaving *Nco*I and *Kas*I cohesive ends for the next step. The T2A module was linearized with *Nco*I and *Kas*I, and IgG LC PCR products were cut with *Bsa*I to produce Golden Gate module plasmids pLC_Int-T2A and pLC_noInt-T2A by ligation.

The IgG heavy chain (HC) was synthesized as 1,499 bp long gBlock (Integrated DNA Technologies) and cloned in pUK21gg with *Pst*I and *Sac*I, yielding the module plasmid pHC_1-f3.

pHC_1-f3 and pLC_Int-T2A or pLC-noInt-T2A modules were assembled with CbH promotor, bGH terminator, and short stuffer oligonucleotide Ig_B-A_dual into pBackBone plasmid using the Golden Gate method with *Bsa*I restriction enzyme to produce IgG-T2A-Int and IgG-T2A-noInt construct.

p3SVLT construction

Plasmid p3SVLT was used to enhance protein production in transfected cells. Cloning codon optimized version of SV40 large T antigen was obtained as a small HindIII-SV40-LT-EcoRI insert, which was kindly provided by Tom Vink. The insert was ligated in a linearized pCDNA3 (# V79020; Invitrogen, Carlsbad, CA) vector digested with *Eco*RI (New England Biolabs) and *Hind*III (New England Biolabs). The ligation reaction was done at a vector-to-insert ratio of 1:3 according to the Quick Ligation™ Kit, and the final construct was amplified in XL10-Gold® Ultracompetent Cells (StrateGene, Bellingham, WA).

pX_mCerulean3 construction

DNA sequence of mCerulean3[27] was amplified from mCerulean3-N1 (# 54730; Addgene) using FP_CC_NcoI_Fw and FP_KasI_Rev primers. The PCR product was cloned in pUK21gg with *Nco*I and *Kas*I, producing the pX-mCerulean3 plasmid, which was assembled with multiguide RNA module containing six random gRNA molecules, SV40 promoter, and SV40 terminator following the Golden Gate assembly protocol (*Bsa*I) to produce the mCerulean_test plasmid.

Construction of backbone for plasmids designed for magnetic-activated cell sorting

The magnetic-activated cell sorting (MACS) backbone plasmid was derived from a backbone plasmid with a secondary cassette described in Josipović *et al*.[25] First, the backbone plasmid was linearized with *Esp*3I (Thermo Fisher Scientific) and purified using a QIAquick PCR Purification Kit (Qiagen, Hilden, Germany). The 874 bp fragment containing the *LNGFR* gene was amplified from the pMACS-LNGFR plasmid (Miltenyi Biotec, Gladbach, Germany) using the primers LNGFR-FW and delta_LNGFR-RE. Promotor and terminator modules were amplified from plasmids described previously[25] using primer pairs pro_SV40_PstI-FW, pro_SV40_XhoI and ter_SV40_PstI-FW, ter_SV40_XhoI-RE, respectively. All modules were cut with *Bsa*I and purified using either the QIAquick Gel Extraction Kit or the QIAquick PCR Purification Kit (Qiagen). The backbone and the modules were ligated to obtain the pBB_NEW_LNGFR plasmid.

Construction of IgG-gRNA-bearing plasmids and plasmids for MACS

Plasmid constructs of IgG-gRNA-bearing plasmids were constructed in a modular system by the Golden Gate cloning method, as described in Josipović *et al*.[25] for targeting of genes *B4GALT1*, *ST6GAL1*, *FUT8*, *MGAT3*, *RUNX3*, *SPPL3*, and *GGA2*. Non-targeting gRNA was used as a control and was specifically designed for each dSpCas9 and dSaCas9 ortholog. Briefly, the final assembly was conducted with modules for antibody HC and LC, Cbh promoter and bGH terminator, and single guide or multiguide RNA molecule in a backbone pBackBone-BZ for IgG-gRNA plasmids or backbone pBB_NEW_LNGFR for plasmids designed for MACS with *Bsa*I restriction enzyme. The IgG expression cassette that drives expression of the antibody HC and LC linked by a 2A self-processing peptide and a furin cleavage site was designed based on the work of Fang *et al*.[28] Details of plasmids and modules used in Golden Gate cloning are described in Supplementary Table S2, and primers and custom oligonucleotides are listed in Supplementary Table S1. For glycosyltransferases *B4GALT1* and *FUT8*, we used a single gRNA for each gene. For *ST6GAL1*, we used all three gRNAs listed in the Supplementary Material on a single plasmid, since we could not achieve sufficient regulation using ST6GAL1_A-sg03 alone. For each of the other genes—*MGAT3*, *GGA2*, *RUNX3*, and *SPPL3*—we used three gRNAs on a single plasmid (Supplementary Table S1). When more than one gRNA (listed in the Supplementary Material) was used, we did not test them individually, but rather used them together in the multiguide system on a single plasmid in all experiments.

Dot blot assay

A dot blot assay was conducted to assess the expression levels of IgG from HEK-293F cells transfected with IgG-encoding plasmids. Briefly, the membrane was washed with 1×phosphate-buffered saline (PBS; 137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4, pH 7.3) and dried. One microliter of IgG samples collected on days 3 and 5 after cell transfection and IgG standards diluted at ratios of 1:8, 1:16, 1:32, 1:64, and 1:128 as well as bovine serum albumin (control) were applied onto the nitrocellulose membrane (GE Healthcare, Chicago, IL) and blocked in 5% skim milk powder dissolved in 1×PBST (1×PBS with Tween® 20 at a ratio of 1:1 v/v) for 30 min on a shaker at room temperature (RT). Next, the membrane was incubated with goat anti-human IgG H&L (horseradish peroxidase; HRP) antibody (Abcam, Cambridge, UK) diluted 1:10,000 in 5% fat-free milk for 1 h on a shaker at RT. Finally, the membrane was washed four times with PBST for 15 min, and the antigen-antibody complex was visualized by using Clarity Max ECL luminol-peroxide (1:1) solution (Bio-Rad, Hercules, CA).

Establishment of polyclonal and monoclonal cell lines

To integrate the dCas9-based expression cassette into cultured cells, we used a modified PB system consisting of a donor vector (described above) and a separate vector for expression of the transposase, which facilitates integration of the portion of the donor vector flanked by ITR sequences. FreeStyle™ 293-F cells (Gibco, Grand Island, NY; HEK-293F) were grown in 125 mL polycarbonate Erlenmeyer flasks with a filter cap (Nalgene, Rochester, NY), on a PSU-20i Multi-functional Orbital Shaker at 140 rpm at 37°C with 8% CO_2 and maintained in FreeStyle 293 medium (Gibco) in a total volume of 30 mL. When the cells reached the appropriate density, 500,000 cells/mL were plated on non-treated six-well plates. The cells were co-transfected with 100 ng of PB-VPR-dSaCas9 or PB-KRAB-dSpCas9 plasmid and 100 ng of PBase vector. The transfection was performed with 293fectin Transfection Reagent (Gibco) according to the manufacturer's protocol, which was adapted to the appropriate total volume of 2 mL per well. The cells were screened the next day for expression of fluorescent proteins mClover3 and mRuby3,[29] and were selected with 7 μg/mL of puromycin (Gibco) for 48 h. Cells were observed and passaged if needed until day 11

after transfection when a fraction of cells were frozen as polyclonal cell lines. Meanwhile, conditioned FreeStyle 293 medium was prepared by seeding 500,000 HEK-293F cells/mL in a 125 mL flask. The conditioned medium was harvested after 24 h by filtering the cell suspension through a 0.22 μm syringe filter. Eleven days after transfection, the cells were diluted to reach a density of 5 cells/mL and seeded on 96-well plates for suspension cells in a volume of 100 μL per well. One hour after seeding, the cells were screened by fluorescent microscopy, and wells with one cell were marked for further observation. Fifteen days after seeding, three PB-dSaCas9-VPR monoclonal cell lines and five PB-dSpCas9-KRAB were scaled up to a 24-well plate for suspension cells (CytoOne) and 5 days later to non-treated six-well plates for suspension cells. When they reached a density of 2.5×10^6, two aliquots of each cell line were frozen according to the HEK-293FreeStyle manufacturer's protocol.

Puromycin validation

Monoclonal cell lines PB-dSaCas9-VPR-1 and PB-dSpCas9 and HEK-293F cells were thawed following the protocol from Vink *et al.*[23] The cells were subcultured three times before seeding them on non-treated six-well plates, 500,000 cells/mL. The next day, the cells were treated with 7 μg/mL of puromycin, which was removed by centrifugation (5 min, 100 *g*, 4°C) after 48 h. On day 3 after selection, the cells were counted using an EVE™ Automated Cell Counter (NanoEntek, Seoul, Korea).

Western blot

To verify the expression of dSaCas9 or dSpCas9 in established monoclonal and polyclonal cell lines, we performed Western blotting. FreeStyle™ 293-F cells (Gibco) served as a negative control. Approximately 5×10^6 cells for each cell line were pelleted by centrifugation (5 min, 500 *g*) and washed twice in 700 μL 1 × PBS. Cells were lysed in RIPA buffer supplemented with Protease Inhibitor Cocktail (cOmplete™ ULTRA Tablets; Roche, Basel, Switzerland) and sonicated using Sonorex Super Ultrasonic bath (Bandelin, Berlin, Germany). Protein concentration was determined with BCA protein assay (Pierce™ BCA Protein Assay Kit; Thermo Fisher Scientific) according to the manufacturer's protocol. Sixty micrograms of protein was used for each sample, and β-actin served as a loading control. Rabbit anti-SaCas9 1:5,000 (ab203943; Abcam), mouse anti-SpCas9 1:1,000 (ab191468; Abcam), mouse anti-β-actin 1:2,000 (sc-69879; Santa Cruz Biotechnology, Santa Cruz, CA), goat anti-rabbit IgG H&L (HRP) 1:200,000 (ab6721; Abcam), and goat anti-mouse IgG H&L (HRP) 1:200,000 (ab205719; Abcam) antibodies were used. Specific proteins were visualized by Clarity™ Western ECL Substrate (Bio-Rad) according to the manufacturer's protocol using the Alliance Q9 Advanced (UVITECTM Cambridge) imaging system.

Transfection efficiency test

FreeStyle™ 293-F cells were plated on non-treated six-well plates at a density of 500,000 cells/mL. They were transfected with 500, 700, 900, 1,200, 1,600, 2,000, 2,400, and 2,800 ng of mCerulean_test plasmid with 293fectin Transfection reagent according to the manufacturer's protocol. The transfection efficiency was analyzed with fluorescence microscopy 24 h after transfection.

Transfection of stable polyclonal and monoclonal cell lines

Suspension-adapted monoclonal cell lines PB-dSaCas9-VPR-1 and PB-dSpCas9-KRAB-3 as well as polyclonal cell lines PB-dSaCas9-VPR-p and PB-dSpCas9-KRAB-p were maintained in the same manner as HEK-293F cells. The cell lines were grown until they reached appropriate density with ≥90 viability and were plated at a concentration of 500,000 cells/mL on non-treated six-well plates. Transfections were performed with 293fectin Transfection Reagent according to the manufacturer's protocol. Briefly, 2 μg of plasmids were diluted in Opti-MEM I Reduced Serum Medium (Gibco) in the final volume of 80 mL. The protocol from Vink *et al.*[23] was used for improved expression of IgG optimized for an ideal ratio of plasmids containing gRNA and IgG chains, p3SVLT, pORF-hp21 (InvivoGen, San Diego, CA), and pORF-hp27 (InvivoGen; mass ratio: 0.69/0.01/0.05/0.25). During vigorous vortexing, 80 mL of diluted 293-fectin was added dropwise to the diluted mixture of plasmids, and cells were allowed to grow for 5 days as described. After 5 days, cells were separated from supernatant (enriched with secreted IgG) by centrifugation (5 min, 4,000 *g*, 4°C). Cell pellets were used for analysis of gene expression, while supernatants were used for glycan analysis.

Reverse transcription and quantitative PCR

Isolation of total RNA was performed with the RNeasy Mini Kit (Qiagen) from cell pellets obtained 5 days after transfection according to the manufacturer's instructions. The concentration of RNA was determined by spectrophotometry at 260 nm, and equal amounts of RNA (50 ng) were subjected to reverse transcription using PrimeScript RTase (Takara, Shiga, Japan) and random hexamer primers (Invitrogen) for TaqMan Gene Expression Assay. Samples for detection of *MGAT3* and *RUNX3* transcript were treated with TURBO

DNase (Invitrogen) to remove residual DNA, considering that primers detecting *MGAT3* are able to detect genomic DNA as well as cDNA obtained from RNA transcripts. Reverse transcription quantitative PCR (RT-qPCR) was performed using the 7500 Fast Real-Time PCR System using TaqMan Gene Expression Master Mix according to the manufacturer's instructions in a total volume of 20 μL with the following assays: Hs00155245_m1 (*B4GALT1*), Hs00949382_m1 (*ST6GAL1*), Hs02379589_s1 (*MGAT3*), Hs00189535_m1 (*FUT8*), and Hs02800695_m1 (*HPRT1*). The gene *HPRT1* was used as a reference gene to normalize gene expression in each sample where gene expression was manifested as a mean value of nine replicates using the $\Delta\Delta$Ct method.[30] Fold change (FC) was presented in comparison to gene expression in cells transfected with non-targeting gRNA control plasmid. To maintain easy comparability of up- and downregulated transcripts, fold regulation (FR) was calculated by taking the reciprocal value for every FC <1 and making it a negative number (thus, the FC of 0.5 is FR of −2; values for FC ≥1 equal FR).

MACS

Monoclonal cell lines PB-dSaCas9-VPR-1 and PB-dSpCas9-KRAB-3 transfected with plasmids designed for MACS were sorted 5 days after transfection according to the manufacturer's protocol with a MACS MicroBead Technology (Miltenyi Biotec) sorting kit. Gene expression of three biological replicates that went through the sorting procedure was compared to gene expression of three biological replicates that were not sorted using the $\Delta\Delta$Ct method (normalized to *HPRT1* expression). FC was calculated as stated before in comparison to the mean of gene expression in cells (sorted and unsorted cells combined) transfected with non-targeting gRNA control plasmid.

Determination of MACSelect™ enrichment rate

Five days after transfection of polyclonal cell lines PB-dSaCas9-VPR-p and PB-dSpCas9-KRAB-p with plasmids designed for MACS, the enrichment rate was determined by comparing magnetically labeled cells before and after MACS to fluorescent staining. The cells were centrifugated and re-suspended in PBE buffer according to the manufacturer's protocol (MACS® MicroBead Technology sorting kit; Miltenyi Biotec). Control FITC antibody was used specifically to stain cells that express LNGFR membrane protein and was added to cells at a ratio 1:10. The cells were counted on a fluorescent microscope before and after separation to determine the enrichment rate calculated by the formula given in the manufacturer's protocol.

Isolation of IgG secreted from FreeStyle™ 293-F cells

IgG was isolated from HEK-293F monoclonal and polyclonal cell culture supernatants using Protein G Agarose fast flow beads (Merck, Darmstadt, Germany). The beads were prewashed three times with 10×bead volume of 1×PBS. In each washing step, the beads were re-suspended in 1×PBS and centrifugated at 150 *g* for 10 s, and the supernatant was removed. After the last wash, prewashed beads were re-suspended in 1×PBS to make a 50:50 (v/v) bead slurry. Approximately 2 mL of HEK cell culture supernatant was mixed with an equal volume of 1×PBS and 40 μL of prepared 50% bead slurry in a 5 mL tube. The samples were re-suspended by pipetting and were incubated for 1 h at RT with gentle shaking to allow IgG to bind to the beads. After incubation, the samples were centrifugated at 150 *g* for 10 s, and the supernatants were then carefully removed and discarded. The beads were washed three times with 300 μL of 1×PBS and three times with 300 μL of ultrapure water to remove nonspecifically bound proteins. After the washing steps, bound IgG was eluted by incubating the beads in 100 μL of 0.1 M formic acid (Merck) for 15 min at RT with gentle shaking. Eluted IgG was neutralized with 17 μL of 1 M ammonium bicarbonate (Merck). IgG concentration in the eluate was measured using Nanodrop 8000 (Thermo Fisher Scientific). Samples were subsequently dried in a vacuum concentrator.

Glycan analysis

Release of *N*-glycans, glycan labeling, and clean-up of glycans were performed according to a previously established protocol[31] with some modifications. Briefly, dried IgG was denatured with sodium dodecyl sulfate (SDS; Invitrogen) and heated at 65°C. The excess SDS was neutralized with Igepal-CA630 (Merck), and *N*-glycans were released by 18 h incubation with PNGaseF (Promega, Madison, WA). The released glycans were fluorescently labeled with procainamide in a two-step reaction. In the first step, 25 μL of freshly prepared labeling solution, containing 172.8 mg/mL of procainamide hydrochloride in a mixture of dimethyl sulfoxide (DMSO) and glacial acetic acid (70:30, v/v), was added to each sample followed by incubation for 1 h at 65 °C. Then, in the next step, 25 μL of freshly prepared solution, containing 179.2 mg/mL of 2-picoline borane as a reducing agent in a mixture of DMSO and acetic acid (70:30, v/v), was added to each sample followed by incubation for 1.5 h at 65°C. Free label and reducing agent were removed from the samples using hydrophilic interaction liquid chromatography (HILIC) solid-phase extraction on a 0.2 μm GHP filter plate (Pall Corporation, Ann Arbor, MI). Glycans were eluted with ultrapure water and stored at −20°C until use.

Fluorescently labeled *N*-glycans were separated by hydrophilic interaction chromatography on a Waters Acquity ultra-performance liquid chromatography (UPLC) instrument (Waters, Milford, MA) consisting of a quaternary solvent manager, sample manager, and a FLR fluorescence detector set with excitation and emission wavelengths of 310 nm and 370 nm, respectively. The instrument was under the control of Empower 3 software, build 3471 (Waters). Labeled *N*-glycans were separated on a Waters BEH Glycan chromatography column, 100×2.1 mm i.d., 1.7 μm BEH particles, with 100 mM ammonium formate, pH 4.4, as solvent A and ACN as solvent B. The separation method used linear gradient of 75–62% ACN (v/v) at a flow rate of 0.4 mL/min over 31 min. Samples were maintained at 10°C before injection, and the separation temperature was 60°C. The system was calibrated using an external standard of hydrolyzed and procainamide-labeled glucose oligomers from which the retention times for the individual glycans were converted to glucose units (GU). 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 separated in the same manner as chromatograms of human plasma-derived IgG glycans, into 24 peaks, and the amount of glycans in each peak was expressed as a percentage of total integrated area. The structural assignment of the glycans present in the chromatographic peaks was done based on overlay with the chromatogram of human plasma IgG glycans for which structures corresponding to each peak had been previously determined[32] as well as GU values of the glycan peaks using the GlycoStore database. The presence of glycans that contained terminal mannose and sialic acid residues in the chromatographic peaks was determined with exoglycosidase digestions. The enzymes α2-3,6,8,9 neuraminidase A and α1-2,3,6 mannosidase (New England Biolabs) were used for digestions. Aliquots of the procainamide-labeled glycan pool were digested according to the manufacturer's protocol. After overnight incubation at 37°C, the enzymes were removed by filtration through AcroPrep 96 Filter Plates, 10K (Pall Corporation). Digested glycans were then separated by HILIC-UHPLC for comparison against undigested glycans.

Statistical analysis

All statistical analyses were performed using R (The R Foundation for Statistical Computing, Vienna, Austria). Raw figures were prepared with R, with glycan structures added later using the standard symbols.[33] RT-qPCR data points were averages of two technical replicates; biological replicates were pooled from independent experiments. Differences between groups were assessed using the two-tailed Mann–Whitney *U*-test on ΔΔCt data, which were transformed and plotted as fold difference.

Glycans were measured as percentage of area of identified chromatographic peak in relation to total glycans (all glycan peak areas). Derived traits were calculated by summing the appropriate peaks. As each experimental batch of biological replicates contained a different baseline, differences between treatments and controls were united using ''metafor'' in R for the meta-analysis approach. To achieve normal distribution of glycan data, Box–Cox transformations (from ''EnvStats'' in R) were used where appropriate. The ability of the Box–Cox transformation to produce normally distributed data was confirmed on a larger data set from an earlier study using the same methodology[34] with help of ''nortest'' in R. After using either normally distributed data or the appropriate transformation, UPLC data for glycans were analyzed using Student's *t*-test either explicitly or implicitly by running meta-analysis (''metafor'' in R) to combine multiple experimental runs.

Results

Generation of polyclonal and monoclonal VPR-dCas9 or KRAB-dCas9 FreeStyle™ 293-F cell lines

FreeStyle™ 293-F (HEK-293F) cells were used due to their ability to grow as suspension culture in a chemically defined medium at high cell densities, their ease of transfection, and their ability to produce high quantities of recombinant proteins. To establish HEK-293F cell lines with stably integrated VPR-dCas9 or KRAB-dCas9, we constructed a modular Golden Gate adapted plasmid vector based on the PB integration system (Fig. 1). Essentially, we augmented our existing modular system[25] with a new backbone plasmid designed for PB integration. The plasmid pBackBone-BZ was modified to contain PB inverted repeats recognized by the PB transposase, insulator sequences to prevent silencing upon genomic integration, and *Bsa*I restriction sites for Golden Gate assembly, compatible with our previously constructed system. Expression of VPR-dCas9 and KRAB-dCas9 fusion proteins was driven by the weak EFS promotor to minimize off-target effect. Reporter genes *mClover3* (for VPR-dCas9) and *mRuby3* (for KRAB-dCas9) as well as puromycin resistance gene were fused with T2A and P2A self-cleaving peptides, respectively. HEK-293F cells were co-transfected with 100 ng PB VPR-dCas9 or KRAB-dCas9 and a minimal quantity (100 ng) of plasmid encoding PB integrase to achieve a low number of genomic integrations. Following puromycin selection, the efficiency of integration was assessed by fluorescence microscopy. Monoclonal cell lines were obtained by limiting dilution from VPR-dCas9 and KRAB-dCas9 polyclonal HEK-293F cells. To

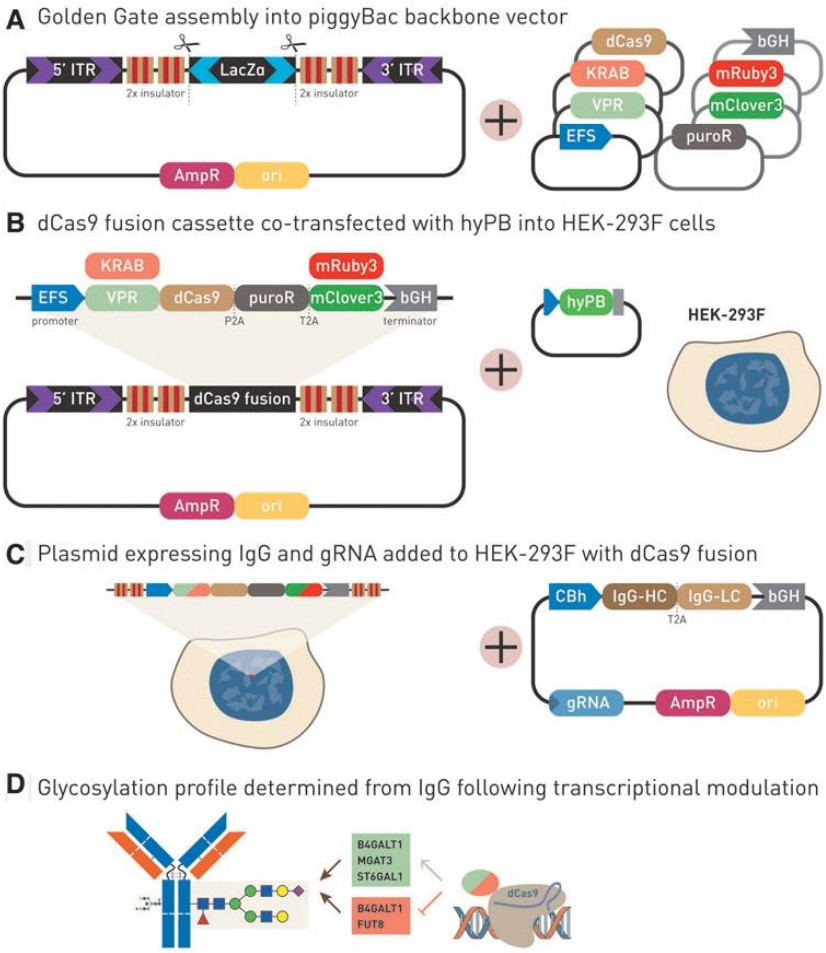

FIG. 1. Generation of the HEK-293F transient expression system for the production of IgG with stably integrated VPR-dCas9 and KRAB-dCas9 cassettes for the up- and downregulation of glyco-genes. **(A)** The expression cassette, encoding dCas9 fused to effector domain, fluorescent marker, and antibiotic resistance, is assembled using the Golden Gate method into the backbone plasmid that includes ITR elements for genomic integration and two repeats of insulator sequence on each side. **(B)** The assembled construct is co-transfected with hyPB into the HEK-293F cell line. **(C)** After antibiotic selection and optional limiting dilution (for creation of monoclonal cell lines), cells with an integrated dCas9 fusion construct are transfected with a plasmid encoding heavy and light chains of IgG and guide RNA (gRNA). **(D)** As specific gRNAs target appropriate dCas9 fusion to a candidate gene, transcriptional expression is modulated, which affects glycosylation on IgG secreted from HEK-293F cells. IgG, immunoglobulin G; IgG-HC, IgG heavy chain; IgG-LC, IgG light chain; hyPB, hyperactive piggyBac integrase; AmpR, ampicillin resistance cassette (bacterial); ori, pUC origin of replication. Color images are available online.

validate genomic stability of the expression cassette integrated by the PB system, another round of puromycin selection was performed after limiting dilution, and no decrease in cell viability or number was observed (Supplementary Table S4). Additionally, we verified continued stable expression of dCas9 fusions by observing fluorescence in monoclonal cell lines (mClover3 for VPR-dCas9 and mRuby3 for KRAB-dCas9). Finally, we verified expression and stability of dCas9 by Western blotting (Supplementary Fig. S1).

Transfection efficiency in FreeStyle™ 293-F cells

In order to determine the optimal amount of plasmid DNA for highest transfection efficiency in our system, a multiguide RNA plasmid with the *mCerulean3* reporter gene was designed, which is similar in size to IgG-gRNA-bearing plasmid (Supplementary Table S2). Cells were transfected with increasing amounts of plasmid DNA, and the optimal amount of plasmid DNA was determined to be 2 μg per million cells, with a transfection efficiency of 69% (Supplementary Fig. S2).

IgG glycan profiles of FreeStyle™ 293-F, and stable VPR-dCas9 and KRAB-dCas9 HEK-293F cell lines

The HEK-293F transient system for a high yield of IgG antibodies was first developed by Vink *et al.*,[23] and we modified it as follows to match our experimental setup. We adapted the IgG expression system by cloning both IgG chains in one plasmid able to receive up to six gRNA molecules by *Bsa*I Golden Gate cloning. IgG antibody quantities were measured using dot blots 3 and 5 days after transfection (Supplementary Fig. S3). The plasmid with an intron in IgG light chain (IgG-T2A-Int) produced higher quantities of IgG (15.1 mg/L) than the plasmid without the intron (IgG-T2A-noInt), which had a yield of 12.6 mg/L, peaking on day 5 following transfection (Supplementary Fig. S3). Therefore, we used the plasmid with IgG-T2A-Int and collected cells 5 days following transfection in all further experiments. IgG glycan profiles from HEK-293F cells were compared to the glycan profiles of IgG produced from the stable VPR-dCas9 and KRAB-dCas9 HEK-293F cell lines. The IgG glycan profiles from all three cell lines were nearly identical (Fig. 2 and Supplementary Fig. S4). Glycan structures were assigned to UPLC peaks according to our previous study.[35] All three profiles were characterized by four dominant peaks and a greater number of very small peaks. In all three profiles, the peak of the highest signal intensity, and thus the highest abundance, corresponded to the biantennary core-fucosylated *N*-glycan without galactose (FA2 glycan). The second most intense peak in all three profiles corresponded to the biantennary core-fucosylated *N*-glycan with one galactose bound on α-6 mannose arm (FA2(6)G1 glycan), while the third largest peak corresponded to the biantennary core-fucosylated *N*-glycan with one galactose bound on α-3 mannose arm (FA2(3)G1 glycan). The fourth main peak in the IgG *N*-glycan profiles of HEK-293F cells, VPR-dCas9, and KRAB-dCas9 cell lines corresponded to the biantennary core-fucosylated *N*-glycan with two galactoses (FA2G2 glycan). The relative abundances (average of three replicates) of the four dominant *N*-glycans were comparable between all three compared profiles. Glycans with bisecting GlcNAc and glycans with sialic acid as well as glycans without core fucose were generally present at low abundance in *N*-glycan profiles of IgG derived from all three cell lines (Fig. 2).

Manipulation of candidate glyco-genes in VPR-dCas9 and KRAB-dCas9 HEK-293F monoclonal cell lines

For testing the newly designed VPR-dCas9 and KRAB-dCas9 HEK-293F transient expression systems, we chose four glyco-genes (*B4GALT1*, *FUT8*, *MGAT3*, and *ST6GAL1*), encoding glycosyltransferases involved in

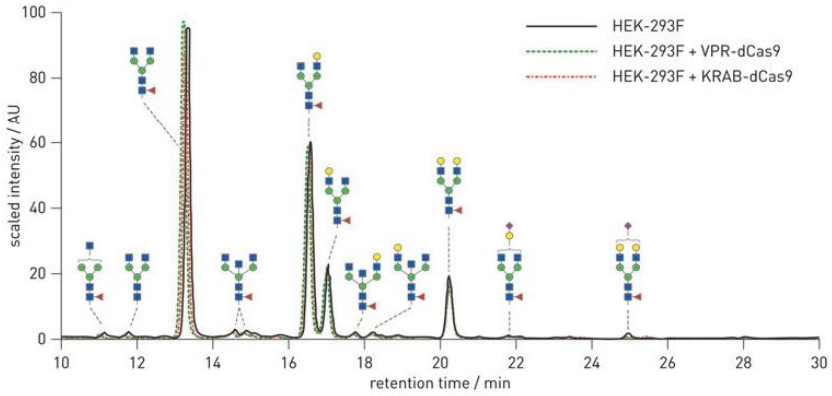

FIG. 2. Comparison of *N*-glycan profiles of IgG produced in HEK-293F cells and in VPR-dCas9 and KRAB-dCas9 cell lines. The three profiles are scaled and overlapped, showing their similarity. Glycan structures contributing to each peak are shown using standard symbols. Color images are available online.

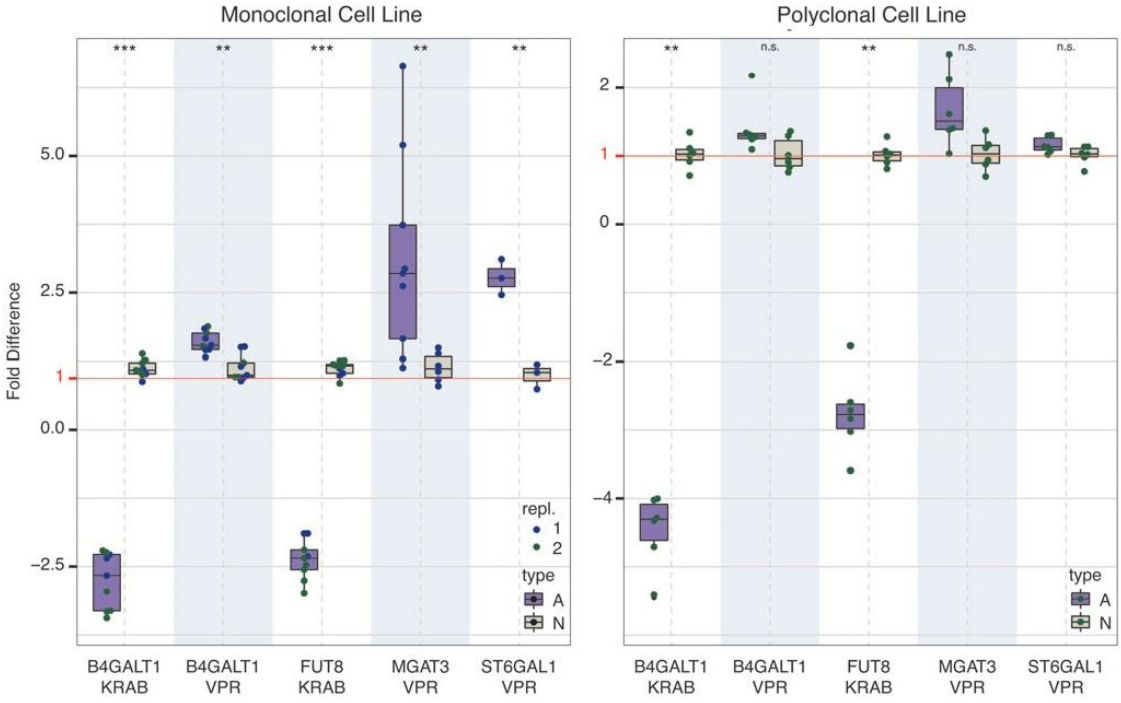

FIG. 3. Changes in transcription of targeted genes as measured by RT-qPCR in monoclonal (*left panel*) and polyclonal (*right panel*) stable VPR-dCas9 and KRAB-dCas9 HEK-293F cell lines. Targeted glyco-genes and the type of stable HEK-293F cell line (VPR-dCas9 or KRAB-dCas9) used are indicated on the *x*-axis. Expression levels are shown as fold difference relative to the non-targeting control. The fold difference equals fold change for values >1; for fold change <1, a negative inverse is taken for better visual representation. For the monoclonal cell line, "repl" denotes the experimental replicate (independent experimental runs). Bars denoted with "A" (*lavender*) show experimental treatment, while "N" (*tan*) denotes non-targeting control. ***$p < 0.001$; **$p < 0.01$; *$p < 0.05$; n.s., not significant. RT-qPCR, reverse transcription quantitative polymerase chain reaction. Color images are available online.

biosynthetic pathways of the glycan structures found on IgG. We targeted these candidate genes using specific gRNAs. All four glyco-genes were detectably expressed in HEK-293F cells. Changes in gene expression as measured by RT-qPCR are given in Figure 3 (left panel). Overview of the relevant metabolic pathways is given in Figure 4. Changes relative to the reference housekeeping gene *HPRT1* for all RT-qPCR data analyzed are given in Supplementary Table S5, indicating their baseline expression level in HEK-293F cells. Additional RT-qPCR raw data including those for *RUNX3*, *GGA2*, *SPPL3*, *FUT8*, and *ST6GAL1* are given in Supplementary Table S6. Transcription levels were successfully modified for all four selected glyco-genes using VPR-dCas9 and the selected gRNAs. Following manipulation of gene expression, we observed changes in the glycosylation profile of IgG expressed in both HEK-293F stable cell lines (Fig. 5). Upon upregulation of *B4GALT1* (coding for the enzyme that adds galactose to IgG) with VPR-dCas9, we observed an increase in galactosylated structures with a concomitant decrease in agalactosylated structures. Downregulation of *B4GALT1* with KRAB-dCas9 showed the opposite effect on IgG glycosylation: a decrease in galactosylation, with a concomitant increase in agalactosylation. Decreased expression of *FUT8* (coding for the enzyme that adds core fucose to IgG) using KRAB-dCas9 led to a decrease in fucosylation, while increased expression of *MGAT3* (coding for the enzyme that adds bisecting GlcNAc to IgG) using VPR-dCas9 led to an increase in structures with bisecting GlcNAc. The upregulation of *ST6GAL1* (encoding the enzyme that adds sialic acid to galactose containing glycan chain), using VPR-dCas9 and appropriate gRNAs, resulted in an increase in sialylated structures on IgG.

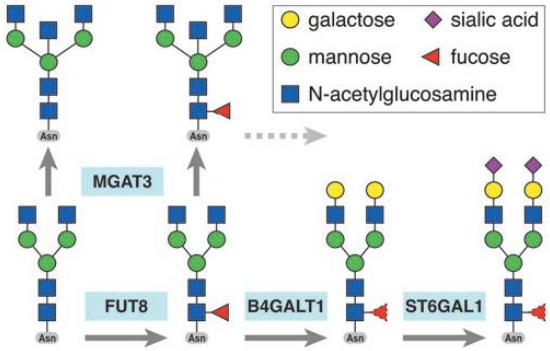

FIG. 4. Simplified IgG glycosylation pathway showing the relevant genes encoding glycosyltransferases and the reactions they catalyze. Fucosyltransferase FUT8 adds a fucose to the core glycan structure, glycosyltransferase MGAT3 adds a bisecting *N*-acetylglucosamine to trimannose core of the glycan chain, while galactosyltransferase B4GALT1 and sialyltransferase ST6GAL1 extend the glycan structure with addition of galactose and sialic acid, respectively. Color images are available online.

Manipulation of candidate glyco-genes in VPR-dCas9 and KRAB-dCas9 HEK-293F polyclonal cell lines and enrichment of transfected cells using MACS

In order to validate the data obtained using our monoclonal cell line–based system, the same experimental setup was repeated for all four genes using polyclonal VPR-dCas9 and KRAB-dCas9 HEK-293F cell lines. These polyclonal stable cell lines were also obtained by integration of dCas9 fusions and puromycin selection but without the limiting dilution step. Gene expression levels were comparable to those from monoclonal cell lines (Fig. 3, right panel). Transcription level of *B4GALT1* (KRAB-dCas9), *FUT8*, and *MGAT3* was significantly altered in the expected direction. Transcription level of *B4GALT1* increased slightly (FC = 1.36), but the change was not statistically significant ($p = 0.132$). Changes in IgG glycosylation profile were similar to those measured in monoclonal cell lines (Fig. 5). The most robust change was related to galactosylation upon either up- or downregulation of *B4GALT1*, despite the fact that the change in transcription level of this gene (using VPR-dCas9) lacked statistical significance. However, no changes of sialylation and fucosylation could be observed following manipulation of *ST6GAL1* and *FUT8*, and the modest relative increase (14.85%) in bisecting structures upon

	G0 agalactosylated	G galactosylated	S sialylated	F fucosylated	B bisected
B4GALT1 ↑	**–13.9 % [–28 %]** –9.1 % [–16 %]	**+11.7 % [+25 %]** +8.5 % [+24 %]	**+2.1 % [+76 %]** n.s.	**–0.7 % [–1 %]** n.s.	n.s. n.s.
B4GALT1 ↓	**+20.2 % [+35 %]** +22.0 % [+50 %]	**–16.9 % [–45 %]** –19.9 % [–42 %]	n.s. n.s.	n.s. n.s.	n.s. –2.0 % [–17 %]
FUT8 ↓	n.s. +4.3 % [+10 %]	**–3.5 % [–9 %]** n.s.	**+2.1 % [+74 %]** n.s.	**–1.9 % [–2 %]** n.s.	**+2.1 % [+30 %]** n.s.
ST6GAL1 ↑	**+1.5 % [+3 %]** n.s.	**–3.5 % [–10 %]** n.s.	**+1.8 % [+46 %]** n.s.	n.s. n.s.	n.s. n.s.
MGAT3 ↑	n.s. n.s.	n.s. n.s.	n.s. n.s.	n.s. n.s.	**+4.7 % [+100 %]** n.s.

FIG. 5. Effect of glyco-gene expression manipulation on IgG glycosylation profile. Right of the gene labels, the arrows indicate the direction of change in gene transcription (*green up arrow*: activation with VPR-dCas9; *red down arrow*: silencing with KRAB-dCas9). Glycan derived traits, as indicated, are shown for statistically significant ($p < 0.05$) changes as absolute change in percentage and as relative change (with non-targeting control as baseline) in square brackets. Upper line (*bold typeface*) shows data for monoclonal cell lines (nine biological replicates—independent samples, except for ST6GAL1, which was done in six biological replicates), while the lower line (*regular typeface*) shows results for the polyclonal cell line (six biological replicates). Color images are available online.

upregulation of *MGAT3* was not statistically significant. Although results from the polyclonal cell lines largely matched those from the experiment in monoclonal lines, some parameters were not statistically significant, probably due to the decreased statistical power in the smaller sample (six replicates vs. nine in monoclonal cell lines).

After MACS enrichment of transfected cells, results of gene regulation, as measured by RT-qPCR, were similar to those before enrichment (Supplementary Fig. S5). The absolute rate of enrichment was 19–67% (48% difference) for *B4GALT1*/VPR, 38–77% (39%) for *MGAT3*, 31–75% (44%) for *ST6GAL1*, 19–66% (47%) for *FUT8*, and 33–70% (37%) for *B4GALT1*/KRAB.

Functional validation of genes associated with IgG glycosylation in VPR-dCas9 and KRAB-dCas9 HEK-293F monoclonal cell lines

As a proof of principle, we manipulated glycosyltransferases with a known role in IgG glycosylation and obtained expected glycan changes. Thus, we confirmed the usefulness of HEK-293F transient expression system, with stably integrated CRISPR-dCas9 fusions, for studying a potential role of any gene of interest in the process of IgG glycosylation. The next step was functional validation of the three genes associated with IgG glycosylation with hitherto unknown function in this process: *RUNX3*, *GGA2*, and *SPPL3*. The *RUNX3* (runt-related transcription factor 3) and the *GGA2* loci (Golgi associated gamma adaptin ear containing, ARF binding protein 2) were GWAS hits for IgG galactosylation and sialylation, respectively.[19,22]

SPPL3 protein (signal peptide peptidase-like 3) acts as an intramembrane-cleaving aspartyl protease of GxGD type, and it was demonstrated that its expression strongly affects intracellular glycosyltransferase levels and alters cellular *N*-glycosylation.[36] However, its role in glycosylation of IgG protein is not known. The three candidate genes were targeted using appropriate gRNAs in both VPR-dCas9 and KRAB-dCas9 HEK-293F monoclonal cell lines. Transcriptional level of these genes as well as *N*-glycosylation on IgG secreted from both cell lines were analyzed subsequently. All three genes were successfully up- or downregulated. However, IgG glycosylation was changed only in the case of *RUNX3* and *SPPL3* (Fig. 6). Upregulation of *RUNX3* resulted in significant decrease of galactosylated structures with consequent increase in agalactosylated structures. Downregulation of *SPPL3* was accompanied with hyperglycosylation phenotype, as sialylated and galactosylated structures were significantly increased with a concomitant decrease of agalactosylated glycans. Even though the transcriptional

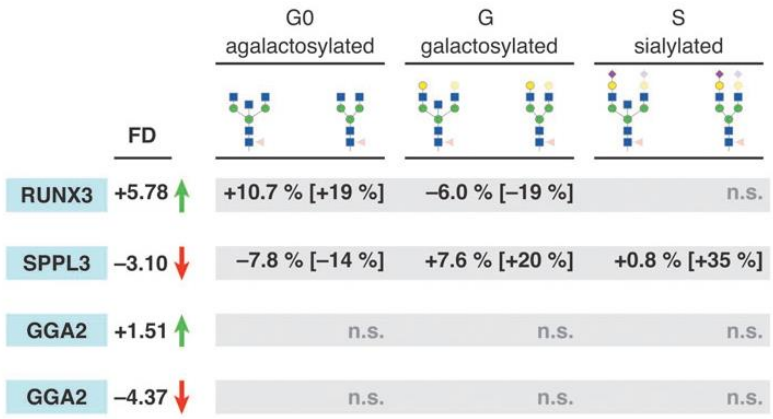

FIG. 6. Some GWAS hits show effect on protein glycosylation. After changing expression of glycosyltransferases, we targeted GWAS hits implicated in various stages of protein glycosylation with dCas9-based effectors for transcriptional activation (VPR, *green up arrow*) or silencing (*red down arrow*). Changes in transcription are shown as fold difference (FD) and were significant to the level $p < 0.05$. Genes *RUNX3* and *SPPL3* were implicated in galactosylation and sialylation, respectively, and show differences in abundance of glycan structures accordingly. Relative differences (in *square brackets*) show a more profound change than absolute changes in percentage. All displayed changes are statistically significant ($p < 0.05$) unless denoted otherwise (n.s.). GWAS, genome-wide association studies. Color images are available online.

level of *GGA2* was up- and downregulated in VPR-dCas9 and KRAB-dCas9 HEK-293F monoclonal cell lines, we could not observe any change in the IgG glycosylation profile. Additional data supporting the glycan analysis, with emphasis on raw UPLC chromatograms and their comparison, are given in Supplementary Figures S6–S21.

Statistical properties of IgG glycan data

In order to validate our statistical methods applied on glycan data, we used a previously published large data set (Klasić *et al.*; Edinburgh cohort, healthy controls, 120 observations)[34] fully compatible with our data. We found that glycan data can be transformed to normal distribution using the Box–Cox transformation for derived glycan traits ''agalactosylated,'' ''galactosylated,'' ''monogalactosylated,'' ''digalactosylated,'' ''sialylated,'' ''monosialylated,'' ''disialylated,'' ''fucosylated,'' and ''bisected'' (Supplementary Table S7). All derived traits passed multiple normality tests: Anderson–Darling, Cramer–von Mises, Lilliefors (Kolmogorov–Smirnov), Shapiro, and Shapiro–Francia tests (Supplementary Table S8). Thus, the usage of parametric statistical tests, which is implicitly necessary when pooling experiments using meta-analysis, is justified.

Discussion

Glycosylation is the most diverse post-translational modification of proteins, which increases the complexity of the human proteome by several orders of magnitude.[37] Alternative glycosylation is functionally important in multiple biological and pathophysiological processes,[38] but our understanding of molecular mechanisms regulating this template-free process is limited. The present knowledge is that glycome composition is inherited as a complex trait[39] regulated by a network of gene loci,[19] but mechanisms and roles of individual genes in this process are mostly unknown. Here, we report a design of a new and versatile transient system for IgG production based on an animal cell model, which can be used for functional validation of genes potentially involved in IgG glycosylation, identified by GWAS.[19–22]

This system is based on the CRISPR-dCas9 modular system for orthogonal gene manipulations,[25] extended by a PB integrase compatible vector, along with the system for IgG production in the HEK-293F cell line[24] (Fig. 1). While each individual component is useful on its own, the complete system represents a unique approach for studying effects of candidate gene up- or downregulation on IgG glycan phenotype. Stable integration of dCas9 fusions into the genome of HEK-293F cells facilitates the key step in gene manipulations, since cells do not have to be transfected with large plasmids encoding dCas9 fusions (typically 14 kbp and more) in each experiment. Since only the plasmid encoding IgG and gRNA needs to be transfected into a cell line with the dCas9 fusion already in place, the notoriously challenging step is fundamentally simplified, enabling high-throughput experimental setup, saving time and reagents costs. Different Cas9 orthologs (dSaCas9 and dSpCas9) were used in conjunction with effector domains conferring opposite activities (VPR, KRAB) to allow for antagonistic gene regulation if desired, although this configuration was not used in our experimental setup. A particularly elegant aspect of this system is colocalization of the IgG expression cassette (modified for monocistronic expression) and up to six simultaneously expressed gRNAs, which ensures that IgG is secreted only from the cells where targeted gene regulation also takes place. In cases when transcription needs to be analyzed in transfected cells, the system can be configured to include a MACS-based enrichment of transfected cells. We hypothesized that for small and moderately sized plasmids, this costly step could be omitted because of outstanding transfection efficiency (Supplementary Fig. S2). To be certain, we performed MACS-based enrichment on VPR-dCas9 and KRAB-dCas9 polyclonal cell lines after targeting *B4GALT1*, *ST6GAL1*, *MGAT3*, and *ST6GAL1* and observed no statistically significant change in expression of these genes with or without MACS (Supplementary Fig. S5). This experiment further demonstrates the simplicity of our system, as there is no need for any type of selection of transfected cells.

The existing flexible platform for Golden Gate assembly[25] was extended with a new ''backbone'' (donor) plasmid, which enables easy genomic integration using the PB system. Essentially, the assembled cassette is flanked by appropriate ITR sequences, with additional insulator sequences that facilitate stable long-term expression following genomic integration. This method for stable integration is surprisingly efficient and much less technically demanding than lentiviral delivery. This approach is possible with cell lines, while lentiviral delivery is indispensable only when working with primary cells. Finally, if there is no need to have a cell line with strictly defined integrations, a polyclonal cell line shows sufficient reproducibility that it can be used instead of a monoclonal line. So, the time-consuming step of limiting dilution can be omitted. Our data set confirms that polyclonal cell lines can be used for most purposes. For all practical purposes, the answers from experiments with monoclonal and polyclonal cell lines were qualitatively identical, although polyclonal cell lines show somewhat weaker dCas9-based activity and glycosylation phenotype. To validate our newly developed system, we chose to manipulate

four glyco-genes with known roles in IgG glycosylation: *FUT8*, encoding for alpha-(1,6)-fucosyltransferase, which adds fucose on a core of bi-antennary structure; *B4GALT1*, encoding for beta-1,4-galactosyltransferase 1, which transfers galactose to *N*-acetylglucosamine (GlcNAc); *MGAT3*, encoding for beta-1,4-mannosyl-4-beta-*N*-acetylglucosamintransferase, which produces bisecting GlcNAc; and *ST6GAL1*, encoding for beta-galactosidase alpha-2,6-sialyltransferase, which adds sialic acid to galactose containing glycan chain. Indeed, IgG glycosylation is dependent on expression levels of these glycosyltransferases in the secretion pathway of plasma B cells, whose function is mimicked by our transient system based on HEK-293F cells. Since the glycan profile of HEK-293F cells showed a low abundance of glycans with bisecting GlcNAc and sialic acids, expression of *ST6GAL1* and *MGAT3* was manipulated in order to increase the relative abundance of those structures. Likewise, *FUT8* was only manipulated with KRAB in order to downregulate its expression and increase the abundance of glycans without a core fucose, which are ordinarily underrepresented in the glycan profile of HEK-293F cells. When *FUT8*, *B4GALT1*, *MGAT3*, and *ST6GAL1* were up- and downregulated with VPR-dCas9 and KRAB-dCas9 fusions, we obtained expected changes in glycan abundances in the monoclonal cell line. Most of the changes were replicated in the polyclonal cell line. When upregulating *B4GALT1* using VPR-dCas9, we observed not only the expected increase in the level of glycans with galactose, but also a small but statistically significant increase in the level of glycans with sialic acid. Similar results have already been reported previously[24] and can be explained simply by the presence of increased levels of glycans with galactose(s), which actually serve as acceptor substrates for sialyltransferase ST6GAL1. However, it has also been demonstrated that the enzymes B4GALT1 and ST6GAL1 interact with each other, and that the activity of both enzymes is higher when they act together in a complex.[40,41] Therefore, it is possible that following VPR-dCas9-based transcriptional activation of the *B4GALT1* gene, the increased level of B4GALT1 enzyme led to increased formation of B4GALT1/ST6GAL1 heterodimers, which then increased the catalytic activity of ST6GAL1 and resulted in a slight increase in the level of glycans with sialic acid. Silencing of the *FUT8* gene with KRAB-dCas9 resulted in an approximately 60% decrease in *FUT8* mRNA, but such a decrease in mRNA level was accompanied by only a very small decrease in the level of fucosylated IgG glycans (around 90% of total glycans that originated from IgG obtained from HEK cells after KRAB-dCas9-based silencing of the *FUT8* gene corresponded to glycans with core fucose). In humans, around 95% of all glycans attached to blood plasma IgG contain core fucose.[42] It has also been observed that the level of total (summed) IgG glycans with core fucose does not vary much between individuals from the general population (the level of fucosylated IgG glycans is generally high in individuals from the general population and ranges from 87% to 97%, and the minimum level observed in the general population was 68%).[42] Moreover, it has been demonstrated that the degree of protein fucosylation generated by the action of the FUT8 enzyme is dependent on the context of the protein.[43] Our results indicate that the residual expression level of fucosyltransferase 8 after KRAB-dCas9-based silencing of the *FUT8* gene was still sufficient to enable efficient fucosylation of IgG glycans.

It should be noted that the expression of glycosyltransferases is generally low in HEK293 cells (and derived cell lines). For this reason, it is possible that even small changes in expression (especially as seen in polyclonal cell lines) result in detectable glycan changes. In the context of *SPPL3* and its role in protein galactosylation, it is possible that even small changes in *B4GALT1* expression compensate for SPPL3 activity. Along with a generally low abundance of glyco-gene products (enzymes) in cells, their known activity as homo- and heterodimers might exacerbate the effect of small transcriptional changes due to kinetics more sensitive to the abundance of the enzyme, with a critical concentration required for dimerization. This is a highly speculative but plausible explanation.

Results for gene expression and glycan change were comparable for mono- and polyclonal cells, which validates both systems as fit for reproducible functional analysis of candidate genes potentially involved in IgG glycosylation. Also, it should be noted that the percentage of transfected cells after MACS enrichment was about equal to transfection efficiency with the smaller plasmid without LNGFR. Therefore, no difference in change in expression of targeted genes between gRNA-only and gRNA-LNGFR plasmids was observed, since a higher transfection efficiency exactly balanced the level of enrichment by MACS. Still, the MACS option is potentially useful in cases of transfection with larger constructs, where the size of the LNGFR marker would be a small percentage of the total plasmid size (therefore having only a minor influence on transfection efficiency). In such scenario, one or two rounds of MACS enrichment process could dramatically increase the percentage of transfected cells, thus removing background noise or obviating the need for antibiotic selection.

After a successful validation of our system using glyco-genes with known function in IgG glycosylation as a proof of principle, we chose three gene loci associated with IgG glycosylation with yet unknown exact role in this process. Two of the loci—*RUNX3* and *GGA2*—are recently published GWA hits for IgG glycosylation,[19,22] while *SPPL3* was shown to liberate *medial/trans*-Golgi glycosyltransferases from their N-terminal membrane anchors to regulate the intracellular pool of active Golgi glycosyltransferases and the extent of *N*-glycosylation of cellular glycoproteins.[44] However, its role in glycosylation of a single plasma protein such as IgG has not yet been studied. We were able to confirm the role of the two loci, *RUNX3* and *SPPL3*, in IgG glycosylation, while manipulation of the *GGA2* gene did not show any effect on IgG glycans. Upregulation of *RUNX3* in the VPR-dCas9 monoclonal cell line resulted in a decrease in galactosylated structures with a concomitant increase in agalactolysed structures (Fig. 6). To explore the pathway involving *RUNX3* and its effect on IgG galactosylation further, we measured the transcriptional activity of *B4GALT1* following *RUNX3* upregulation. However, no significant changes in *B4GALT1* transcript levels were observed. This might suggest that *RUNX3* does not suppress *B4GALT1* directly but rather through some indirect pathway that should be investigated further. Indeed, Klarić *et al.*[19] proposed a functional network of loci associated with IgG glycosylation where most of the nodes clustered around a group of TFs known to have central roles in B-cell maturation and differentiation, and *RUNX3* was one of them. This clustering suggests that the regulatory network acting in B lymphocytes is driven by transcription factors and affects the expression of key glycosyltransferase enzymes either directly or through a complicated TF-driven network.

Upregulation of *SPPL3* in the VPR-dCas9 monoclonal cell line was followed by hyperglycosylation of plasma IgG antibodies, which is in line with the model of SPPL3 activity on cellular glycoproteins proposed by Voss *et al.*[36] One of the first identified SPPL3 substrates was MGAT5 (alpha-1,6-mannosylglycoprotein 6-beta-N-acetylglucosaminyltransferase) responsible for biosynthesis of branched, complex-type *N*-glycans,[36] which are, however, not present in the IgG glycan profile. Recently, secretome analysis of *Sppl3*$^{-/-}$ mouse embryonic fibroblasts (MEFs) identified B4GALT1 and ST6GAL1 as possible SPPL3 substrates.[45] The significant increases in galactosylation and sialylation that we observed after manipulation of *SPPL3* (Fig. 6) are in accordance with secretome studies in MEFs, and they highlight the importance of SPPL3 in the regulation of B4GALT1 and ST6GAL1 activity in IgG glycosylation.

In conclusion, we demonstrated that our newly developed system with stably integrated CRISPR-dCas9 fusions for transient expression of IgG is an excellent tool for studying role of genes associated with IgG glycosylation by GWAS. Although the system was designed specifically for studying IgG glycosylation, it can easily be repurposed to serve as a model for other proteins and their posttranslational modifications, with appropriate targeting via gRNA. Further, many technologies developed in this study can be used alone or as an extension of the modular dCas9-based system for orthogonal gene regulation and epigenetic manipulation.[25]

Author Disclosure Statement

G.L. is the founder and owner of Genos Ltd., a private research organization that specializes in high-throughput glycomic analyses and has several patents in this field. J.K. is an employee of Genos Ltd. None of the other authors has any conflicts of interest to declare.

Funding Information

The work on design and generation of the CRISPR-dCas9-based transient system for IgG production and gene manipulations performed in the Laboratory of Epigenetics, Faculty of Science, University of Zagreb, was supported by Centre of Competence in Molecular Diagnostics grant (#KK.01.2.2.03.0006), Croatian National Centre of Research Excellence in Personalized Healthcare grant (#KK.01.1.1.01.0010), and European Regional Development Fund Grant, project CasMouse (#KK.01.1.1.04.0085). Glycosylation analysis was performed in the Genos Glycoscience Research Laboratory and was partly supported by the European Union's Horizon 2020 grant IMForFuture (grant #721815), Centre of Competence in Molecular Diagnostics grant (#KK.01.2.2.03.0006), and Croatian National Centre of Research Excellence in Personalized Healthcare grant (#KK.01.1.1.01.0010).

Supplementary Material

Supplementary Figure S1
Supplementary Figure S2
Supplementary Figure S3
Supplementary Figure S4
Supplementary Figure S5
Supplementary Figure S6
Supplementary Figure S7
Supplementary Figure S8
Supplementary Figure S9
Supplementary Figure S10
Supplementary Figure S11
Supplementary Figure S12
Supplementary Figure S13
Supplementary Figure S14
Supplementary Figure S15
Supplementary Figure S16

Supplementary Figure S17
Supplementary Figure S18
Supplementary Figure S19
Supplementary Figure S20
Supplementary Figure S21
Supplementary Table S1
Supplementary Table S2
Supplementary Table S3
Supplementary Table S4
Supplementary Table S5
Supplementary Table S6
Supplementary Table S7
Supplementary Table S8

References

1. Gornik O, Pavic T, Lauc G. Alternative glycosylation modulates function of IgG and other proteins—implications on evolution and disease. *Biochim Biophys Acta* 2012;1820:1318–1326. DOI: 10.1016/j.bbagen.2011.12.004.
2. Krapp S, Mimura Y, Jefferis R, et al. Structural analysis of human IgG-Fc glycoforms reveals a correlation between glycosylation and structural integrity. *J Mol Biol* 2003;325:979–989. DOI: 10.1016/s0022-2836(02)01250-0.
3. Mimura Y, Church S, Ghirlando R, et al. The influence of glycosylation on the thermal stability and effector function expression of human IgG1-Fc: properties of a series of truncated glycoforms. *Mol Immunol* 2000;37:697–706. DOI: 10.1016/s0161-5890(00)00105-x.
4. Arnold JN, Wormald MR, Sim RB, et al. The impact of glycosylation on the biological function and structure of human immunoglobulins. *Annu Rev Immunol* 2007;25:21–50. DOI: 10.1146/annurev.immunol.25.022106.141702.
5. Schwab I, Nimmerjahn F. Intravenous immunoglobulin therapy: how does IgG modulate the immune system? *Nat Rev Immunol* 2013;13:176–189. DOI: 10.1038/nri3401.
6. Ma B, Guan X, Li Y, et al. Protein glycoengineering: an approach for improving protein properties. *Front Chem* 2020;8:622. DOI: 10.3389/fchem.2020.00622.
7. Ferrara C, Grau S, Jager C, et al. Unique carbohydrate–carbohydrate interactions are required for high affinity binding between FcgammaRIII and antibodies lacking core fucose. *Proc Natl Acad Sci U S A* 2011;108:12669–12674. DOI: 10.1073/pnas.1108455108.
8. Parekh RB, Dwek RA, Sutton BJ, et al. Association of rheumatoid arthritis and primary osteoarthritis with changes in the glycosylation pattern of total serum IgG. *Nature* 1985;316:452–457. DOI: 10.1038/316452a0.
9. Bakovic MP, Selman MH, Hoffmann M, et al. High-throughput IgG Fc N-glycosylation profiling by mass spectrometry of glycopeptides. *J Proteome Res* 2013;12:821–831. DOI: 10.1021/pr300887z.
10. Chen G, Wang Y, Qiu L, et al. Human IgG Fc-glycosylation profiling reveals associations with age, sex, female sex hormones and thyroid cancer. *J Proteomics* 2012;75:2824–2834. DOI: 10.1016/j.jprot.2012.02.001.
11. Kapur R, Kustiawan I, Vestrheim A, et al. A prominent lack of IgG1-Fc fucosylation of platelet alloantibodies in pregnancy. *Blood* 2014;123:471–480. DOI: 10.1182/blood-2013-09-527978.
12. Alter G, Ottenhoff THM, Joosten SA. Antibody glycosylation in inflammation, disease and vaccination. *Semin Immunol* 2018;39:102–110. DOI: 10.1016/j.smim.2018.05.003.
13. Gudelj I, Lauc G, Pezer M. Immunoglobulin G glycosylation in aging and diseases. *Cell Immunol* 2018;333:65–79. DOI: 10.1016/j.cellimm.2018.07.009.
14. Seeling M, Bruckner C, Nimmerjahn F. Differential antibody glycosylation in autoimmunity: sweet biomarker or modulator of disease activity? *Nat Rev Rheumatol* 2017;13:621–630. DOI: 10.1038/nrrheum.2017.146.
15. Stowell SR, Ju T, Cummings RD. Protein glycosylation in cancer. *Annu Rev Pathol* 2015;10:473–510. DOI: 10.1146/annurev-pathol-012414-040438.
16. Kavur M, Lauc G, Pezer M. Systems glycobiology: immunoglobulin G glycans as biomarkers and functional effectors in aging and diseases. In: Comprehensive Glycoscience - Reference Module in Chemistry, Molecular Sciences and Chemical Engineering. (Barchi, JJ; ed). Amsterdam, Netherlands: Elsevier, 2021; pp. 1–98.
17. Kristic J, Vuckovic F, Menni C, et al. Glycans are a novel biomarker of chronological and biological ages. *J Gerontol A Biol Sci Med Sci* 2014;69:779–789. DOI: 10.1093/gerona/glt190.
18. Huffman JE, Pucic-Bakovic M, Klaric L, et al. Comparative performance of four methods for high-throughput glycosylation analysis of immunoglobulin G in genetic and epidemiological research. *Mol Cell Proteomics* 2014;13:1598–1610. DOI: 10.1074/mcp.M113.037465.
19. Klaric L, Tsepilov YA, Stanton CM, et al. Glycosylation of immunoglobulin G is regulated by a large network of genes pleiotropic with inflammatory diseases. *Sci Adv* 2020;6:eaax0301. DOI: 10.1126/sciadv.aax0301.
20. Lauc G, Huffman JE, Pucic M, et al. Loci associated with N-glycosylation of human immunoglobulin G show pleiotropy with autoimmune diseases and haematological cancers. *PLoS Genet* 2013;9:e1003225. DOI: 10.1371/journal.pgen.1003225.
21. Shen X, Klaric L, Sharapov S, et al. Multivariate discovery and replication of five novel loci associated with immunoglobulin G N-glycosylation. *Nat Commun* 2017;8:447. DOI: 10.1038/s41467-017-00453-3.
22. Wahl A, van den Akker E, Klaric L, et al. Genome-wide association study on immunoglobulin G glycosylation patterns. *Front Immunol* 2018;9:277. DOI: 10.3389/fimmu.2018.00277.
23. Vink T, Oudshoorn-Dickmann M, Roza M, et al. A simple, robust and highly efficient transient expression system for producing antibodies. *Methods* 2014;65:5–10. DOI: 10.1016/j.ymeth.2013.07.018.
24. Dekkers G, Plomp R, Koeleman CA, et al. Multi-level glyco-engineering techniques to generate IgG with defined Fc-glycans. *Sci Rep* 2016;6:36964. DOI: 10.1038/srep36964.
25. Josipovic G, Tadic V, Klasic M, et al. Antagonistic and synergistic epigenetic modulation using orthologous CRISPR-dCas9-based modular system. *Nucleic Acids Res* 2019;47:9637–9657. DOI: 10.1093/nar/gkz709.
26. Doherty JE, Huye LE, Yusa K, et al. Hyperactive piggyBac gene transfer in human cells and *in vivo*. *Hum Gene Ther* 2012;23:311–320. DOI: 10.1089/hum.2011.138.
27. Markwardt ML, Kremers GJ, Kraft CA, et al. An improved cerulean fluorescent protein with enhanced brightness and reduced reversible photoswitching. *PLoS One* 2011;6:e17896. DOI: 10.1371/journal.pone.0017896.
28. Fang J, Yi S, Simmons A, et al. An antibody delivery system for regulated expression of therapeutic levels of monoclonal antibodies *in vivo*. *Mol Ther* 2007;15:1153–1159. DOI: 10.1038/sj.mt.6300142.
29. Bajar BT, Wang ES, Lam AJ, et al. Improving brightness and photostability of green and red fluorescent proteins for live cell imaging and FRET reporting. *Sci Rep* 2016;6:20889. DOI: 10.1038/srep20889.
30. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(–Delta Delta C(T)) method. *Methods* 2001;25:402–408. DOI: 10.1006/meth.2001.1262.
31. Trbojevic-Akmacic I, Ugrina I, Lauc G. Comparative analysis and validation of different steps in glycomics studies. *Methods Enzymol* 2017;586:37–55. DOI: 10.1016/bs.mie.2016.09.027.
32. Keser T, Pavic T, Lauc G, et al. Comparison of 2-aminobenzamide, procainamide and RapiFluor-MS as derivatizing agents for high-throughput HILIC-UPLC-FLR-MS N-glycan analysis. *Front Chem* 2018;6:324. DOI: 10.3389/fchem.2018.00324.
33. Varki A, Cummings RD, Aebi M, et al. Symbol nomenclature for graphical representations of glycans. *Glycobiology* 2015;25:1323–1324. DOI: 10.1093/glycob/cwv091.
34. Klasic M, Markulin D, Vojta A, et al. Promoter methylation of the *MGAT3* and *BACH2* genes correlates with the composition of the immunoglobulin G glycome in inflammatory bowel disease. *Clin Epigenetics* 2018;10:75. DOI: 10.1186/s13148-018-0507-y.
35. Pucic M, Knezevic A, Vidic J, 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. DOI: 10.1074/mcp.M111.010090.
36. Voss M, Kunzel U, Higel F, et al. Shedding of glycan-modifying enzymes by signal peptide peptidase-like 3 (SPPL3) regulates cellular N-glycosylation. *EMBO J* 2014;33:2890–2905. DOI: 10.15252/embj.201488375.
37. Aebersold R, Agar JN, Amster IJ, et al. How many human proteoforms are there? *Nat Chem Biol* 2018;14:206–214. DOI: 10.1038/nchembio.2576.
38. Smith BAH, Bertozzi CR. The clinical impact of glycobiology: targeting selectins, Siglecs and mammalian glycans. *Nat Rev Drug Discov* 2021;20:217–243. DOI: 10.1038/s41573-020-00093-1.

39. Kristic J, Zaytseva OO, Ram R, et al. Profiling and genetic control of the murine immunoglobulin G glycome. *Nat Chem Biol* 2018;14:516–524. DOI: 10.1038/s41589-018-0034-3.
40. Hassinen A, Pujol FM, Kokkonen N, et al. Functional organization of Golgi N- and O-glycosylation pathways involves pH-dependent complex formation that is impaired in cancer cells. *J Biol Chem* 2011;286:38329–38340. DOI: 10.1074/jbc.M111.277681.
41. Khoder-Agha F, Harrus D, Brysbaert G, et al. Assembly of B4GALT1/ST6GAL1 heteromers in the Golgi membranes involves lateral interactions via highly charged surface domains. *J Biol Chem* 2019;294:14383–14393. DOI: 10.1074/jbc.RA119.009539.
42. Stambuk J, Nakic N, Vuckovic F, et al. Global variability of the human IgG glycome. *Aging (Albany NY)* 2020;12:15222–15259. DOI: 10.18632/aging.103884.
43. Yang Q, Wang LX. Mammalian alpha-1,6-fucosyltransferase (FUT8) is the sole enzyme responsible for the N-acetylglucosaminyltransferase I-independent core fucosylation of high-mannose N-glycans. *J Biol Chem* 2016;291:11064–11071. DOI: 10.1074/jbc.M116.720789.
44. Voss M, Schroder B, Fluhrer R. Mechanism, specificity, and physiology of signal peptide peptidase (SPP) and SPP-like proteases. *Biochim Biophys Acta* 2013;1828:2828–2839. DOI: 10.1016/j.bbamem.2013.03.033.
45. Kuhn PH, Voss M, Haug-Kroper M, et al. Secretome analysis identifies novel signal Peptide peptidase-like 3 (Sppl3) substrates and reveals a role of Sppl3 in multiple Golgi glycosylation pathways. *Mol Cell Proteomics* 2015;14:1584–1598. DOI: 10.1074/mcp.M115.048298.

Received: July 18, 2021
Accepted: November 5, 2021
Online Publication: January 12, 2022
Issue Publication: April 20, 2022

2.2. Effects of Estradiol on Immunoglobulin G Glycosylation: Mapping of the Downstream Signaling Mechanism

frontiers in Immunology

ORIGINAL RESEARCH
published: 25 May 2021
doi: 10.3389/fimmu.2021.680227

Effects of Estradiol on Immunoglobulin G Glycosylation: Mapping of the Downstream Signaling Mechanism

Anika Mijakovac[1†], Julija Jurić[2†], Wendy M. Kohrt[3,4], Jasminka Krištić[2], Domagoj Kifer[5], Kathleen M. Gavin[3,4], Karlo Miškec[1], Azra Frkatović[2], Frano Vučković[2], Marija Pezer[2], Aleksandar Vojta[1], Peter A. Nigrović[6,7], Vlatka Zoldoš[1] and Gordan Lauc[2,5*]*

[1] Department of Molecular Biology, University of Zagreb Faculty of Science, Zagreb, Croatia, [2] Genos Glycoscience Research Laboratory, Zagreb, Croatia, [3] Division of Geriatric Medicine, School of Medicine, University of Colorado Anschutz Medical Campus, Aurora, CO, United States, [4] Eastern Colorado VA Geriatric Research, Education and Clinical Center, Aurora, CO, United States, [5] Faculty of Pharmacy and Biochemistry, University of Zagreb, Zagreb, Croatia, [6] Division of Rheumatology, Immunology and Allergy, Brigham and Women's Hospital, Boston, MA, United States, [7] Division of Immunology, Boston Children's Hospital, Boston, MA, United States

Glycans attached to immunoglobulin G (IgG) directly affect this antibody effector functions and regulate inflammation at several levels. The composition of IgG glycome changes significantly with age. In women, the most notable change coincides with the perimenopausal period. Aiming to investigate the effect of estrogen on IgG glycosylation, we analysed IgG and total serum glycomes in 36 healthy premenopausal women enrolled in a randomized controlled trial of the gonadotropin-releasing hormone analogue ($GnRH_{AG}$) leuprolide acetate to lower gonadal steroids to postmenopausal levels and then randomized to transdermal placebo or estradiol (E_2) patch. The suppression of gonadal hormones induced significant changes in the IgG glycome, while E_2 supplementation was sufficient to prevent changes. The observed glycan changes suggest that depletion of E_2 primarily affects B cell glycosylation, while liver glycosylation stays mostly unchanged. To determine whether previously identified IgG GWAS hits *RUNX1*, *RUNX3*, *SPINK4*, and *ELL2* are involved in downstream signaling mechanisms, linking E_2 with IgG glycosylation, we used the FreeStyle 293-F transient system expressing IgG antibodies with stably integrated CRISPR/dCas9 expression cassettes for gene up- and downregulation. *RUNX3* and *SPINK4* upregulation using dCas9-VPR resulted in a decreased IgG galactosylation and, in the case of *RUNX3*, a concomitant increase in IgG agalactosylation.

Keywords: immunoglobulin G glycosylation, estradiol, CRISPR, inflammation, Runx3

OPEN ACCESS

Edited by:
Mohamed Abdel-Mohsen, Wistar Institute, United States

Reviewed by:
Adam Barb, University of Georgia, United States
Amit Kumar Singh, National Institute on Aging, United States

***Correspondence:**
Vlatka Zoldoš
vzoldos@biol.pmf.hr
Gordan Lauc
glauc@pharma.hr

†These authors have contributed equally to this work

Specialty section:
This article was submitted to B Cell Biology, a section of the journal Frontiers in Immunology

Received: *18 March 2021*
Accepted: *06 May 2021*
Published: *25 May 2021*

Citation:
Mijakovac A, Jurić J, Kohrt WM, Krištić J, Kifer D, Gavin KM, Miškec K, Frkatović A, Vučković F, Pezer M, Vojta A, Nigrović PA, Zoldoš V and Lauc G (2021) Effects of Estradiol on Immunoglobulin G Glycosylation: Mapping of the Downstream Signaling Mechanism. Front. Immunol. 12:680227. doi: 10.3389/fimmu.2021.680227

INTRODUCTION

Most proteins in human serum are glycosylated by the covalent addition of diverse glycan structures that fine-tune their function. The regulatory role of glycans has been most extensively explored on immunoglobulin G (IgG) antibodies, where different glycoforms regulate the immune response on multiple levels (1). Glycans attached to the Fc part of the IgG molecule affect interactions with different Fc receptors, which is why changes in glycosylation have direct effects on the immune system at multiple levels (2). IgG glycosylation is altered in many diseases (3), and glycan changes can even appear before the onset of disease symptoms (4–6). In some cases, changes in IgG glycans were shown to be a causative element contributing to the disease development (7–9). IgG glycans that associate with age are known functional effectors of inflammation, and changes in IgG glycosylation seem to be an important factor contributing to ageing at the molecular level (10–12) that can also be used as a biomarker to track individual trajectories of biological ageing (13).

In women, the most prominent change of glycan age coincides with the perimenopausal period (10). A recent intervention study demonstrated that estrogen regulates IgG glycosylation (14), which may explain why perimenopausal females undergo significant changes in the IgG glycome composition. Unfortunately, limitations of the glycoprofiling method used in that study, *i.e.* only IgG galactosylation was estimated from the total plasma glycome profile, prevented us from the detailed characterization of the estrogen effect on IgG glycosylation. In the present study, we aimed at a better understanding of the estrogen role in the regulation of IgG glycosylation, therefore we reanalyzed samples from the previous intervention study (14) using state-of-the-art glycoprofiling technologies (15). We first defined the components of IgG glycome affected by estradiol (E_2). We then used data from our recent large genome-wide association study (GWAS) of the IgG glycome (16) to identify candidate genes possibly involved in mediating effects of E_2 on IgG glycosylation. We selected four gene loci, *RUNX1–RUNX3*, *SPINK4*, and *ELL2*, involved in E_2 downstream signaling mechanisms, assuming that these loci represent a part of the molecular pathway linking E_2 to IgG glycosylation. *In vitro* system used in this study was based on a FreeStyle 293-F (HEK-293FS) transient expression system optimized for secreting a high quantity of native IgG antibodies (16). The system was modified by stable integration of CRISPR/dCas9 expression cassette containing either VPR (for gene upregulation) or KRAB (for gene downregulation). Using this system, we were able to demonstrate the effects of selected genes on specific IgG glycans which were previously associated with biological ageing.

METHODS

Institutional Approval

This study was conducted at the University of Colorado Anschutz Medical Campus (CU-AMC). All procedures were performed in accordance with the ethical standards and approved by the Colorado Multiple Institutional Review Board (COMIRB) and the Scientific Advisory and Review Committee at the University of Colorado Anschutz Medical Campus (CU-AMC). The study was registered on ClinicalTrials.gov (NCT00687739) on May 28, 2008.

Participants and Screening Procedures

Participants were healthy eumenorrheic premenopausal women who volunteered to take part in the study. All volunteers underwent screening procedures, as described previously (17). The main inclusion criteria were age (25 to 49 years) and regular menstrual cycle function [no missed cycles in the previous year, cycle length 28 ± 5 days and confirmation of ovulatory status (ClearPlan Easy, Unipath Diagnostics, Waltham, MA)]. Exclusion criteria were pregnancy or lactation, hormonal contraception, oral glucocorticoids or diabetes medications, smoking, and body mass index (BMI) >39 kg/m^2. Following the Declaration of Helsinki, all volunteers provided written informed consent to participate, with the knowledge that the risks of the study included menopause-like effects (e.g., weight gain, bone loss, menopausal symptoms).

Experimental Design and Study Procedures

The parental trial was a randomized, double-blinded, placebo-controlled trial to determine the effects of estradiol (E_2) deficiency on body composition, bone mineral density, components of energy expenditure and physical activity in premenopausal women (17, 18). In short, all participants underwent suppression of ovarian sex hormones with gonadotropin-releasing hormone agonist therapy ($GnRH_{AG}$, leuprolide acetate 3.75 mg, Lupron; TAP Pharmaceutical Products, Inc; Lake Forest, IL) in the form of monthly intramuscular injections. A single injection of leuprolide acetate produces an initial stimulation (for 1 to 3 weeks) followed by a prolonged suppression of pituitary gonadotropins FSH and LH, while repeated monthly dosing suppresses ovarian hormone secretion (19). A urine pregnancy test confirmed the absence of pregnancy before each dosing. After completing the screening procedures, eligible volunteers underwent baseline testing during the early follicular phase (days 2 to 6 after the onset of menses) of the menstrual cycle. At the beginning of the following menstrual cycle, participants began 5-months of $GnRH_{AG}$ therapy to suppress ovarian function. Participants were randomized to receive either transdermal E_2 0.075 mg/d (Bayer HealthCare Pharmaceuticals, Berkeley, CA) or placebo patches ($GnRH_{AG}$ + E_2, n = 15; $GnRH_{AG}$ + PL, n = 21). The E_2 regimen kept serum E_2 concentrations in the mid-to-late follicular phase range (100 to 150 pg/ml). To reduce the risk of endometrial hyperplasia and minimize exposure to progesterone, women randomized to E_2 received medroxyprogesterone acetate (5 mg/d, as a pill) for 12 days every other month (end of months 2 and 4, and after completion of follow-up testing). During these monthly visits, participants were under supervision of the research nurse practitioner. Participants

were asked to report health and medication use changes (e.g., doctor visits, hospitalizations), as well as any study-related problems/concerns over the past 4 weeks.

Sample Collection

Blood samples were collected at three timepoints: during baseline testing (T1), during week 20 of the hormonal intervention (T2), and at the spontaneous recovery of the normal menstrual cycle function, approximately 4-months after completion of the drug intervention (T3). A single sample (~5 ml) was obtained in the morning (~8 AM), after an overnight fast (at least 10 h). Baseline samples were obtained immediately before the first $GnRH_{AG}$ injection. Serum was separated from each collected sample upon blood withdrawal and stored at −80°C until analysis.

Sex Hormone Concentration

Collected sera were analyzed for numerous sex hormones. Estrone (E1), estradiol (E2) and progesterone (P) concentrations were determined by radioimmunoassay (RIA, Diagnostic Systems Lab, Webster, TX). Total testosterone (T) concentration was determined by chemiluminescence immunoassay (Beckman Coulter, Inc. Fullerton, CA), and sex hormone-binding globulin (SHBG) concentration was determined by immunoradiometric assay (Diagnostic Systems Laboratory).

Isolation of Immunoglobulin G, Release and Labeling of N-Glycans From IgG

The whole procedure was performed according to the already published protocol (20). In short, IgG was isolated from sera (100 µl) by affinity chromatography using a 96-well plate with protein G coupled to a monolithic stationary phase (BIA Separations, Slovenia). The isolated IgG was denatured with the addition of SDS (Invitrogen, USA) and incubation at 65°C, after which the excess of SDS was neutralized with Igepal CA-630 (Sigma-Aldrich, USA). N-glycans were released from IgG with the addition of PNGase F (Promega, USA) in a PBS buffer during the overnight incubation at 37°C. The released glycans were fluorescently labelled with 2-AB dye (Merck, Germany) in the 2 h incubation at 65°C. Free label and reducing agent were removed from the samples by hydrophilic interaction liquid chromatography solid phase extraction (HILIC-SPE). IgG N-glycans were eluted with ultrapure water and stored at −20°C.

Release and Labeling of N-Glycans From Total Serum Proteins

The whole procedure was performed as described previously (4). In short, serum proteins (10 µl) were denatured by SDS and incubated at 65°C. Excess SDS was neutralized by Igepal CA-630 (Sigma-Aldrich, USA). Serum proteins were deglycosylated by PNGase F (Promega, USA) in a PBS buffer during the overnight incubation at 37°C. Released glycans were fluorescently labelled with 2-AB dye (Merck, Germany) in the 2h incubation time at 65°C. Excess of reagents and proteins from previous steps was removed by hydrophilic interaction liquid chromatography solid phase extraction (HILIC-SPE). Serum N-glycans were eluted with ultra-pure water and stored at −20°C.

Hydrophilic Interaction Chromatography (HILIC)-UPLC Analysis of Labeled Glycans

Fluorescently labeled N-glycans were separated by ultra-performance liquid chromatography (UPLC) on a Waters Acquity UPLC *H*-Class Instrument consisting of a sample manager, quaternary solvent manager, and a fluorescence (FLR) detector set with excitation and emission wavelengths at 250 and 428 nm, respectively. The UPLC system was under the control of Empower 3 software, build 3471 (Waters, USA). Labeled N-glycans were separated on an amide ACQUITY UPLC® Glycan BEH chromatography column (Waters, USA), 100 × 2.1 mm i.d. for IgG glycans and 150 × 2.1 mm for glycans from total serum proteins, 1.7 µm BEH particles, with 100 mM ammonium formate pH 4.4 as solvent A, and 100% acetonitrile as solvent B. The separation method used a linear gradient of 75–62% acetonitrile at a flow rate of 0.40 ml/min in a 27 min analytical run for IgG glycans and a linear gradient of 70–53% acetonitrile at a flow rate of 0.561 ml/min in a 23 min analytical run for glycans from total serum proteins. Samples were kept at 10°C before injection onto the column. The separation temperature of the column was 60°C for the IgG glycans and 25°C for glycans from serum proteins. Data processing included an automatic integration method that was manually corrected to maintain the same intervals of chromatographic integration across all samples. Chromatograms were separated in the same manner into 24 peaks for IgG N-glycans and 39 peaks for N-glycans from total serum proteins. The abundance of glycans in each chromatographic peak was expressed as a percentage of the total integrated area (% area).

Plasmid Constructs

Plasmid constructs pORF-hp21 and pORF-hp27 used to enhance protein production were obtained from Invivogen, while p3SVLT was constructed by cloning a codon-optimized version of the SV40 large T antigen coding region (16) in pcDNA3 (Addgene). Unwanted BsaI restriction sites in IgG heavy and light chain (kindly provided by Gestur Vidarsson, Sanquin, Amsterdam) were removed using QuikChange Lightning Site-Directed Mutagenesis Kit (Agilent). IgG chains were then cloned into pUK21gg for the subsequent Golden Gate cloning step. Expression plasmids encoding gene-specific guide RNA (gRNA) molecules were constructed in the multi-guide system described by Josipović et al. (21). Three gRNA molecules were cloned individually in a backbone plasmid pSgMx-A or pSgMx-G (where x represents the order of gRNA molecules; 1,2,3 and A or G represents Cas9 ortholog (dSaCas9 or dSpCas9 respectively) it recognizes) (22) for each gene: *RUNX1, RUNX3, SPINK4* and *ELL2*. gRNA molecules for *RUNX1, RUNX3*, *ELL2* and *SPINK4* were then cloned in pSgx3 as modular "multiguide" molecules. Two non-targeting gRNA molecules and one gRNA molecule targeting *B4GALT1* for dSaCas9 and dSpCas9 were cloned in the same way described above. Together with modules for antibody heavy chain (HC) and light chain (LC), Cbh promoter and bGH terminator, single guide or multiguide RNA molecules were cloned in a backbone pBackBone-BZ by modular Golden Gate cloning method described in Josipović

et al. (21). Sequences of gRNA molecules and details of plasmids/modules used in Golden Gate cloning are given in **Supplementary Tables 6, 7**.

Cell Culture and Transfections

Stable cell lines PB-dSaCas9-VPR-1 and PB-dSpCas9-KRAB-3 were established from FreeStyle™ 293-F cells (Gibco) with piggyBac transposon system by limiting dilution method (unpublished data) and maintained in FreeStyle™ 293 Expression Medium (Gibco) in 125 ml Erlenmeyer flasks (Nalgene) and cultivated at 37°C in the atmosphere with 8% CO_2 on PSU-20i Multi-functional Orbital Shaker at 140 rpm according to the protocol from Vink et al. (16). Transfections of stable cell lines were done using 293fectin Transfection Reagent (Gibco) according to the manufacturer's protocol optimized for 2 ml per well. When cells reached ≥90 viability, they were plated in non-treated 6-well plates at a concentration of 500,000 cells/ml and were transfected with 2 µg of plasmids diluted in Opti-MEM I Reduced Serum Medium (Gibco) to reach a volume of 80 ml. For enhanced expression of immunoglobulin G, cells were transfected with a plasmid containing gRNA and IgG heavy and light chain, p3SVLT, pORF-hp21 and pORF-hp27 in the ratio: 0.69/0.01/0.05/0.25 (16). Cells were collected 5 days after transfection by centrifugation at 4,000*g* (5 min). The cell pellet was used for gene expression profiling, while the supernatant was used for glycan analysis.

Quantitative Real-Time PCR (qPCR)

For gene expression profiling, total RNA was extracted with RNeasy Mini Kit (Qiagen) from cell pellets collected five days after transfection. Reverse transcription was done on 50 ng of total isolated RNA using the PrimeScript RTase (TaKaRa) and random hexamer primers (Invitrogen) for TaqMan Gene Expression Assay or on 5 ng of total isolated RNA (pretreated with TURBO DNase (Invitrogen)) for SYBR Green Gene Expression Assay. Both variants of RT-qPCR were performed according to the manufacturer's protocol using the 7500 Fast Real-Time PCR System using TaqMan Gene Expression Master Mix with the following assays: Hs01021970_m1 (RUNX1), Hs00205508_m1 (SPINK4), Hs01023022_m1 (ELL2), Hs00155245_m1 (B4GALT1) and Hs02800695_m1 (HPRT1) or PowerUp SYBR Green Master Mix with primer sequences listed in **Supplementary Table 8**. The mean value of 12 replicates was normalized to the expression of the *HPRT1* gene as endogenous control and was analysed using the ΔΔCt method (23). Fold change (FC) was shown relative to gene expression in cells transfected with a plasmid expressing non-targeting gRNA.

IgG Isolation From FreeStyle™ 293-F Cells, N-Glycan Release, Labeling and HILIC-UPLC Analysis

IgG was isolated from FreeStyle 293-F cell culture supernatants using Protein G Agarose fast flow beads (Merck, Germany). The beads were prewashed three times with 10× bead volume of 1× PBS. In each washing step, beads were resuspended in 1× PBS, centrifugated at 150×*g* for 10 s, and the supernatant was removed. After the last wash, prewashed beads were resuspended in 1× PBS to make a 50:50 (v/v) beads slurry. Approximately 2 ml of FreeStyle 293-F cell culture supernatant were mixed with an equal volume of 1× PBS and 40 µl of prepared 50% bead slurry in a 5 ml tube. The samples were resuspended by pipetting action and incubated 1 h at room temperature with gentle shaking to allow IgG to bind to the beads. During the incubation period, the samples were resuspended twice by pipetting action. After incubation, the samples were centrifugated at 150×*g* for 10 s, and the supernatants were then carefully removed and discarded. The beads were washed three times with 300 µl of 1× PBS and three times with 300 µl of ultrapure water to remove non-specifically bound proteins. After washing steps, bound IgG was eluted by incubating the beads in 100 µl of 0.1 M formic acid (Merck) for 15 min at room temperature with gentle shaking. Eluted IgG was neutralised with 17 µl of 1M ammonium bicarbonate (Merck, Germany). IgG concentration in the eluate was measured using Nanodrop 8000 (Thermo Scientific, USA). Samples were subsequently dried in a vacuum concentrator.

N-glycan release, glycan labeling, clean-up of glycans and separation of glycans by HILIC-UHPLC were performed according to a previously established protocol (20) with some modifications. Briefly, dried IgG was denatured by SDS (Invitrogen, USA) and heated at 65°C. The excess of SDS was neutralised with Igepal CA-630 (Merck, Germany), and N-glycans were released by 18 h of incubation with PNGaseF (Promega, USA). The released glycans were fluorescently labeled with procainamide in a two-step reaction. In the first step, 25 µl of freshly prepared labeling solution, containing 172.8 mg/ml of procainamide hydrochloride (Thermo Fisher Scientific, USA) in a mixture of DMSO (Merck, Germany) and glacial acetic acid (Merck, Germany) (70:30, v/v), was added to each sample followed by incubation for 1 h at 65°C. Then in the next step, 25 µl of freshly prepared solution, containing 179.2 mg/ml of 2-picoline borane as a reducing agent in a mixture of DMSO and acetic acid (70:30, v/v), was added to each sample followed by incubation for 1.5 h at 65°C. Free label and reducing agent were removed from the samples using hydrophilic interaction liquid chromatography solid-phase extraction (HILIC-SPE) on a 0.2 µm GHP filter plate (Pall Corporation, USA). Glycans were eluted with ultrapure water and stored at −20°C until use. Fluorescently labeled N-glycans were separated by hydrophilic interaction chromatography on a Waters Acquity UPLC instrument (Waters, USA) consisting of a quaternary solvent manager, sample manager and FLR fluorescence detector set with excitation and emission wavelengths of 310 and 370 nm, respectively. The instrument was under the control of Empower 3 software, build 3471 (Waters, USA). Labeled N-glycans were separated on a Waters BEH Glycan chromatography column, 100 × 2.1 mm i.d., 1.7 µm BEH particles, with 100 mM ammonium formate, pH 4.4, as solvent A and ACN as solvent B. The separation method used a linear gradient of 75–62% ACN (v/v) at a flow rate of 0.4 ml/min over 31 min. Samples were maintained at 10°C before injection, and the separation temperature was 60°C. The system was calibrated using an external standard of hydrolyzed and procainamide-labeled glucose oligomers from which the retention

times for the individual glycans were converted to glucose units (GU). 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 separated in the same manner as chromatograms of human plasma/serum-derived IgG glycans into 24 peaks, and the abundance of glycans in each peak was expressed as a percentage of the total integrated area. The structural assignment of the glycans present in the chromatographic peaks was done based on i) overlay with the chromatogram of human plasma IgG glycans for which structures corresponding to each peak had been previously determined (24) and ii) GU values of the glycan peaks using the GlycoStore database (www.glycostore.org).

Statistical Analysis

The area under chromatogram peaks was normalized to total chromatogram area, then each glycan peak was logit transformed, and batch corrected using ComBat method (R package 'sva') (25). Data were back transformed, and derived glycan traits were calculated as a sum or ratio of selected directly measured glycan peaks based on particular glycosylation features (i.e. sialylation or fucosylation).

Mixed models were used to estimate the effect of the intervention (R package 'lme4') (26). Hormone concentration or particular glycan level was set as dependent variable and timepoint (with levels: *baseline*, *after intervention* and *after recovery*) nested within the treatment group (*placebo* and *estradiol*) as independent variables. Also, the model was age-adjusted, and the subject's ID was included as a random intercept to account for variation between the subjects.

Change in estradiol and change in glycan abundance were calculated by subtracting values of consecutive timepoints. Mixed models were used to estimate the relationship between the change in glycans and the change in estradiol concentration. mixed models were used. The change in glycan abundance was defined as a dependent variable, while the change in estradiol concentration was defined as a fixed effect. Group and time period nested within the group were defined as random factors. Both change in glycan abundance and change in estradiol concentration were transformed to a standard normal distribution by inverse transformation of ranks to normality.

Prior modeling, glycan levels were transformed to a standard normal distribution by inverse transformation of ranks to normality (R package 'GenABEL') (27), while hormone concentrations were log transformed. Based on fitted models, changes of dependent variables after intervention or recovery (relative to baseline) were compared between the groups (placebo vs estradiol) using *post-hoc* t-test. False discovery rate was controlled using the Benjamini–Hochberg method at a significance level of 0.05.

Differences between groups for gene expression and glycan levels following CRISPR/dCas9 manipulations were tested using the non-parametric Mann–Whitney test. Results with $p < 0.05$ were considered statistically significant. All statistical analyses were performed in R programming software (version 3.6.3) (28).

RESULTS

IgG Glycome Is Affected by Estradiol

Thirty-six healthy premenopausal women were enrolled in a randomized controlled trial of the gonadotropin-releasing hormone analogue ($GnRH_{AG}$) leuprolide acetate to lower gonadal steroids to postmenopausal levels and then randomized to transdermal placebo (PL) or estradiol (E_2) patch (**Figure 1**) (17). In order to analyse total serum and IgG glycomes, serum samples were collected: at baseline (Sampling 1); after five months of $GnRH_{AG}$ administration with concurrent supplementation with either E_2 or placebo (Sampling 2); and four

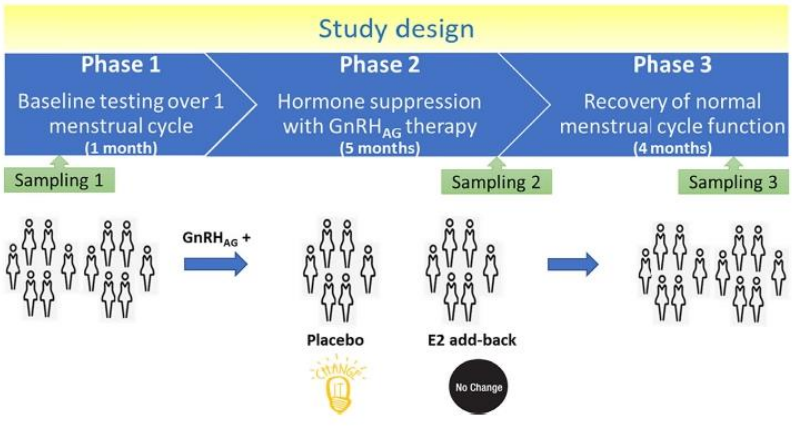

FIGURE 1 | Design of the study.

months after the end of the intervention, when natural hormonal cycling was restored (Sampling 3). **Figure 2** shows representative UHPLC chromatograms of the total serum (**Figure 2A**) and IgG (**Figure 2B**) glycomes and the direction of glycan changes after the suppression of gonadal hormones. IgG glycosylation analysis revealed significant changes in IgG glycome composition after gonadal hormone suppression (time point at the end of Phase 2), while E_2 supplementation was sufficient to prevent changes in the IgG glycome composition (**Figure 3**, **Table 1** and **Supplementary Table 1**). After four months of the recovery period (Sampling 3 after the end of Phase 3), the IgG glycome composition returned to nearly pre-intervention values in the placebo group. Galactosylation was the most affected IgG glycome feature, with a significant decrease of digalactosylated glycans (G2) and an increase of monogalactosylated (G1) and agalactosylated (G0) glycans. The level of sialylated glycans (S) and the ratio of sialylation and galactosylation (S/G) of IgG significantly decreased, while the abundance of glycans with bisecting GlcNAc (B) increased. Only the abundance of core-fucosylated (F) glycans did not change by depletion of E_2. IgG glycosylation traits related to galactosylation and sialylation, which were particularly affected by the suppression of E_2, are also the main components of the glycan age clock of the biological age. This index has initially been developed to predict chronological age (10) but was subsequently converted into the test of biological age (29). Suppression of E_2 resulted in a median increase of GlycanAge by 9.1 years, which was completely attenuated by E_2 add-back (**Supplementary Figure 1**). At the individual level, the extent of changes in hormone concentration (**Supplementary Table 2**) correlated moderately with the extent of changes in individual IgG glycans (**Supplementary Table 3**), suggesting that other factors (beside gonadal hormones) also strongly affect the composition of the IgG glycome.

To determine whether the effects of E_2 on glycosylation were restricted to IgG, we also analysed total serum protein N-glycome in the same samples. Changes observed in the total serum N-glycome (**Figure 4**) were restricted only to some neutral glycans and core-fucosylated sialylated biantennary glycans known to originate nearly exclusively from immunoglobulins (30). This suggests that depletion of E_2 affects B cell (IgG) glycosylation, while liver glycosylation does not seem to be affected, at least not in a way that would alter proportions of individual non-immunoglobulin *N*-glycans in the total serum glycome (**Figure 2A**).

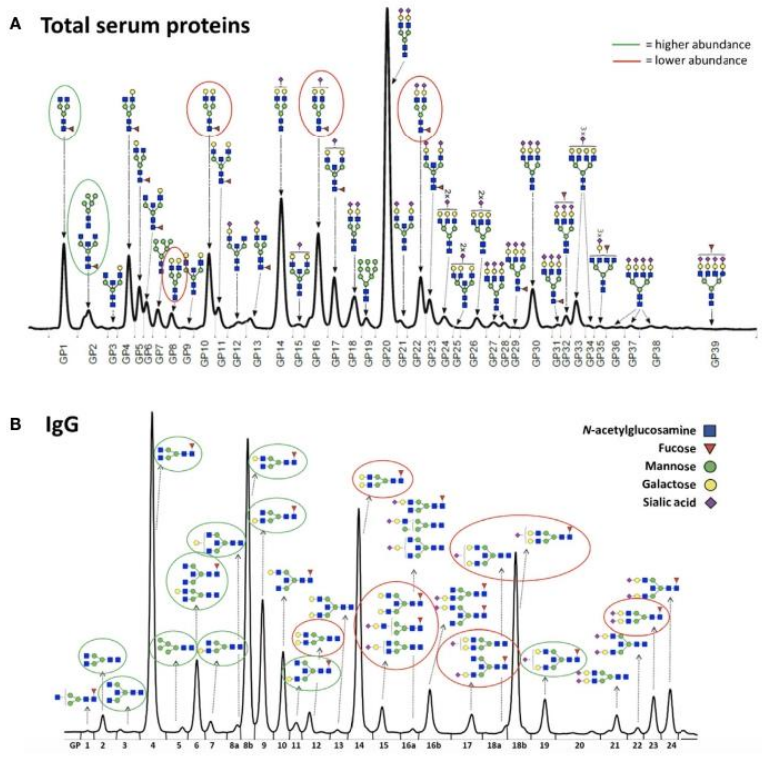

FIGURE 2 | Representative chromatograms of **(A)** the total serum glycome and **(B)** the IgG glycome. Glycans that decreased after the gonadal hormone suppression with gonadotropin-releasing hormone analogue leuprolide acetate ($GnRH_{AG}$) therapy are circled in red, and those that increased are circled in green.

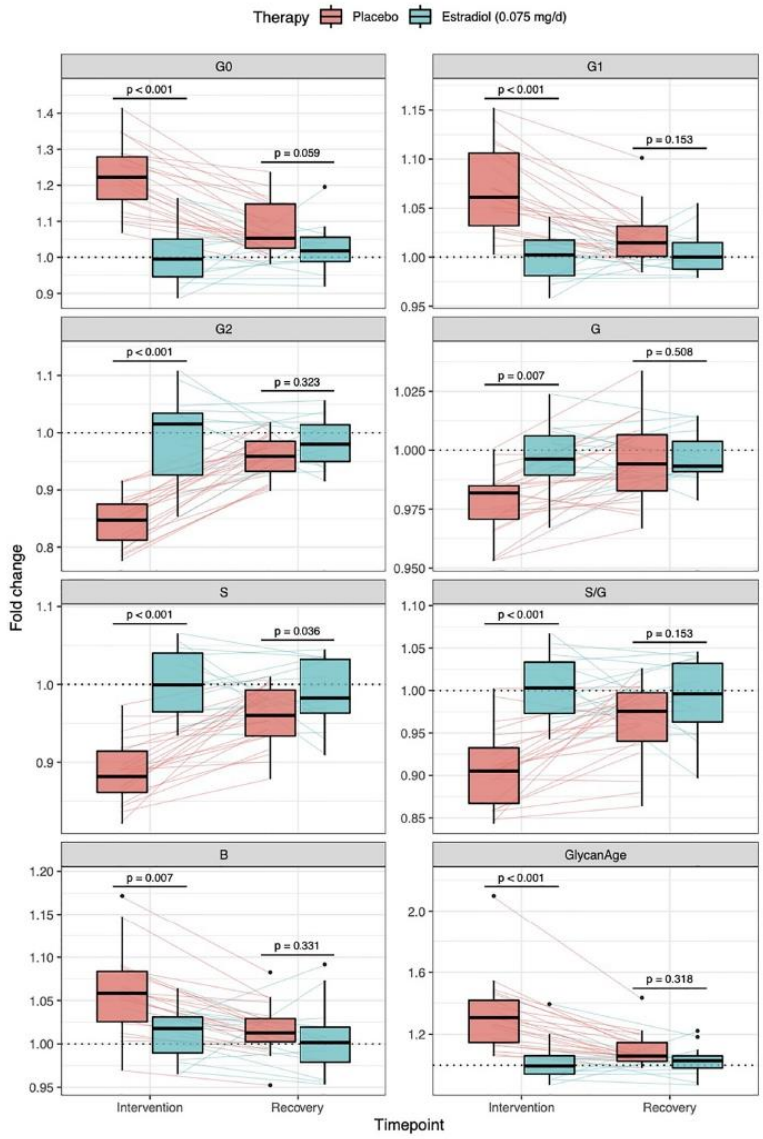

FIGURE 3 | Effects of gonadal hormone suppression on IgG glycosylation. Gonadotropin-releasing hormone analogue leuprolide acetate ($GnRH_{AG}$) was used to lower gonadal steroids to postmenopausal levels in healthy premenopausal women (n = 36) that were then randomized to transdermal placebo (n = 21) or estradiol patch (n = 15). Changes in the IgG glycome composition after five months of $GnRH_{AG}$ (Intervention) with supplementation of E_2 (transdermal estradiol supplementation) or without supplementation of E_2 (supplementation with placebo) and four months after the end of the intervention (Recovery) are shown on the graph. G2, digalactosylated glycans; G1, monogalactosylated glycans; G0, agalactosylated glycans; S, sialylated glycans; S/G, ratio of sialylation and galactosylation; B, glycans with bisecting GlcNAc; G, all glycans with galactose.

Downstream Signaling Mechanisms Linking Estradiol With IgG Glycan Traits

Glycans are inherited as complex traits defined by multiple genes (31, 32) which play a role in the synthesis and variation of individual glycan structures. Through a series of GWAS papers in the last decade (32–35), we mapped an extensive network of genes that potentially regulate the glycosylation of IgG. Using the Signalling Pathways Project (SPP) web knowledgebase (36), we explored the effects of E_2 on GWAS hits for IgG galactosylation and sialylation (35). Results presented in **Supplementary Figure 2** indicate that E_2

TABLE 1 | Glycan abundances (%) of directly measured IgG glycan traits at the baseline and deviations from the baseline after intervention and after recovery timepoint.

Glycan	Intervention	Glycan Abundance (%) at Baseline median (IQR)	Difference in glycan abundance (%) relative to baseline. sampling after:			
			Intervention median (IQR)	p_I	Recovery median (IQR)	p_R
GP1	Placebo	0.049 (0.040–0.051)	0.010 (0.008–0.012)	1.95×10^{-1}	0.002 (0.000–0.010)	1.48×10^{-1}
	Estradiol	0.069 (0.041–0.091)	0.009 (0.000–0.010)		−0.001 (−0.007–0.006)	
GP2	Placebo	0.269 (0.199–0.654)	0.102 (0.082–0.148)	$\mathbf{3.53 \times 10^{-6}}$	0.027 (−0.010–0.042)	1.48×10^{-1}
	Estradiol	0.361 (0.261–0.495)	−0.003 (−0.047–0.031)		0.005 (−0.026–0.028)	
GP3	Placebo	0.059 (0.059–0.070)	0.011 (0.009–0.013)	$\mathbf{4.08 \times 10^{-4}}$	0.003 (−0.001–0.011)	7.56×10^{-2}
	Estradiol	0.070 (0.061–0.096)	−0.006 (−0.010–0.004)		−0.001 (−0.008–0.003)	
GP4	Placebo	14.0 (12.4–17.6)	3.08 (2.35–3.93)	$\mathbf{2.46 \times 10^{-8}}$	0.869 (0.345–1.771)	6.58×10^{-2}
	Estradiol	17.3 (14.0–19.2)	−0.152 (−1.003–1.031)		0.187 (−0.374–0.847)	
GP5	Placebo	0.158 (0.143–0.180)	0.021 (0.011–0.028)	$\mathbf{2.56 \times 10^{-4}}$	0.009 (−0.004–0.015)	$\mathbf{3.22 \times 10^{-2}}$
	Estradiol	0.162 (0.155–0.194)	−0.004 (−0.010–0.006)		−0.004 (−0.013–0.005)	
GP6	Placebo	3.44 (2.98–4.28)	0.673 (0.567–0.824)	$\mathbf{1.17 \times 10^{-7}}$	0.200 (0.124–0.301)	1.48×10^{-1}
	Estradiol	3.71 (3.28–4.36)	0.067 (−0.114–0.253)		0.126 (−0.112–0.198)	
GP7	Placebo	0.363 (0.240–0.466)	0.039 (0.020–0.056)	$\mathbf{3.51 \times 10^{-3}}$	0.010 (−0.013–0.026)	4.56×10^{-1}
	Estradiol	0.394 (0.289–0.435)	−0.010 (−0.029–0.019)		0.000 (−0.016–0.014)	
GP8	Placebo	18.5 (17.5–19.9)	0.990 (0.604–1.246)	$\mathbf{2.22 \times 10^{-5}}$	0.139 (−0.189–0.474)	4.91×10^{-1}
	Estradiol	20.0 (18.5–21.0)	−0.030 (−0.351–0.208)		0.032 (−0.316–0.207)	
GP9	Placebo	9.66 (8.31–10.73)	0.843 (0.560–1.274)	$\mathbf{1.14 \times 10^{-7}}$	0.137 (−0.015–0.405)	3.64×10^{-1}
	Estradiol	9.69 (9.09–10.31)	−0.024 (−0.285–0.221)		0.069 (−0.050–0.313)	
GP10	Placebo	4.99 (4.69–5.38)	0.177 (0.127–0.358)	5.28×10^{-1}	0.087 (0.029–0.175)	3.50×10^{-1}
	Estradiol	4.65 (4.37–5.52)	0.125 (0.020–0.167)		−0.040 (−0.104–0.047)	
GP11	Placebo	0.636 (0.587–0.696)	0.051 (0.030–0.092)	$\mathbf{2.46 \times 10^{-4}}$	0.018 (0.000–0.031)	5.77×10^{-1}
	Estradiol	0.665 (0.555–0.715)	0.000 (−0.020–0.036)		0.015 (−0.004–0.030)	
GP12	Placebo	1.010 (0.621–1.333)	−0.178 (−0.256 to −0.129)	$\mathbf{1.19 \times 10^{-5}}$	−0.021 (−0.060–0.000)	9.52×10^{-1}
	Estradiol	0.781 (0.596–0.984)	0.000 (−0.095–0.020)		−0.020 (−0.062–0.010)	
GP13	Placebo	0.249 (0.211–0.281)	−0.010 (−0.029–0.011)	2.57×10^{-1}	−0.009 (−0.019–0.010)	4.90×10^{-1}
	Estradiol	0.230 (0.210–0.245)	0.008 (−0.020–0.019)		0.001 (−0.015–0.011)	
GP14	Placebo	20.1 (16.6–21.9)	−2.95 (−3.93 to −2.11)	$\mathbf{4.71 \times 10^{-7}}$	−0.735 (−1.383 to −0.315)	3.12×10^{-1}
	Estradiol	16.3 (14.8–18.8)	0.212 (−0.987–0.708)		−0.469 (−0.977–0.253)	
GP15	Placebo	1.98 (1.76–2.39)	−0.250 (−0.337 to −0.129)	$\mathbf{5.39 \times 10^{-7}}$	−0.031 (−0.107–0.018)	3.68×10^{-1}
	Estradiol	1.90 (1.55–2.05)	0.002 (−0.040–0.075)		−0.041 (−0.059–0.053)	

(Continued)

TABLE 1 | Continued

Glycan	Intervention	Glycan Abundance (%) at Baseline median (IQR)	Difference in glycan abundance (%) relative to baseline. sampling after:			
			Intervention median (IQR)	p_I	Recovery median (IQR)	p_R
GP16	Placebo	3.09 (2.54–3.38)	0.010 (–0.059–0.122)	4.78×10^{-1}	–0.019 (–0.088–0.039)	7.14×10^{-2}
	Estradiol	3.05 (2.89–3.26)	0.041 (–0.010–0.073)		0.026 (–0.002–0.067)	
GP17	Placebo	1.043 (0.835–1.086)	–0.100 (–0.177 to –0.060)	**2.75×10^{-5}**	–0.034 (–0.060 to –0.011)	1.95×10^{-1}
	Estradiol	0.895 (0.761–0.990)	–0.010 (–0.037–0.031)		0.008 (–0.048–0.016)	
GP18	Placebo	12.9 (10.3–14.1)	–2.42 (–2.99 to –1.92)	**9.92×10^{-9}**	–0.585 (–1.129 to –0.156)	3.23×10^{-1}
	Estradiol	11.0 (10.1–12.3)	–0.009 (–0.717–0.823)		–0.383 (–0.513–0.248)	
GP19	Placebo	1.87 (1.72–1.99)	0.062 (0.000–0.089)	**9.19×10^{-3}**	–0.043 (–0.064–0.039)	9.72×10^{-1}
	Estradiol	1.88 (1.61–2.18)	0.000 (–0.034–0.050)		–0.001 (–0.065–0.017)	
GP20	Placebo	0.418 (0.381–0.443)	–0.010 (–0.041–0.010)	**1.75×10^{-2}**	–0.014 (–0.048–0.007)	3.68×10^{-1}
	Estradiol	0.389 (0.346–0.418)	0.010 (–0.010–0.048)		0.006 (–0.014–0.019)	
GP21	Placebo	0.851 (0.762–0.974)	–0.029 (–0.051–0.022)	3.98×10^{-1}	–0.007 (–0.062–0.073)	7.58×10^{-1}
	Estradiol	0.790 (0.725–0.815)	0.010 (–0.032–0.036)		0.000 (–0.034–0.027)	
GP22	Placebo	0.12 (0.11–0.16)	0.001 (–0.020–0.010)	2.81×10^{-1}	–0.010 (–0.010–0.001)	8.83×10^{-1}
	Estradiol	0.120 (0.111–0.135)	0.001 (0.000–0.010)		0.001 (–0.010–0.010)	
GP23	Placebo	1.90 (1.70–2.09)	–0.178 (–0.218 to –0.100)	**5.17×10^{-3}**	–0.034 (–0.157–0.028)	1.48×10^{-1}
	Estradiol	1.91 (1.57–2.21)	–0.032 (–0.115–0.014)		0.028 (–0.112–0.061)	
GP24	Placebo	1.89 (1.65–2.10)	0.050 (–0.059–0.108)	1.53×10^{-1}	–0.060 (–0.089–0.027)	6.09×10^{-1}
	Estradiol	1.92 (1.62–2.22)	–0.008 (–0.035–0.045)		–0.020 (–0.068–0.028)	

p values describe statistical significance of difference between estradiol and placebo group after intervention (p_I) and recovery (p_R). p values smaller than 0.05 are bolded. IQR, limits of the interquartile range (1st–3rd quartile); GP, glycan peak.

affects the expression of B4GALT1, glycosyltransferase which adds galactose to IgG glycans, but also the genes which are not directly involved in IgG glycosylation, such as the *RUNX1–RUNX3* loci (Runt-related transcription factor 1 and RUNX family transcription factor 3, found in many promoters and enhancers, which can either activate or suppress transcription) and the *SPINK4* locus (serine peptidase inhibitor, also known as PEC-60). These genes were identified as GWAS hits for IgG galactosylation. In addition, *ELL2* (Elongation Factor for RNA Polymerase 2), another GWAS hit for IgG glycosylation, more specifically sialylation, also appears to be strongly regulated by E_2. On the other hand, there were no conclusive results on ST6GAL1, the enzyme that adds sialic acid to IgG.

To determine the involvement of the *RUNX1–RUNX3*, *SPINK4*, and *ELL2* loci in downstream signaling mechanisms linking E_2 with IgG glycosylation, we directly manipulated their transcriptional activity in the *in vitro* IgG expression system HEK-293FS using CRISPR/dCas9 molecular tools and subsequently analysed IgG glycan phenotype. A recently developed HEK-293FS transient system for IgG secretion with stably integrated CRISPR/dCas9 expression cassette for gene upregulation (dCas9-VPR) and downregulation (dCas9-KRAB) was used for this purpose. These cells were transfected with a plasmid containing genes for IgG heavy and light chains aiming to induce the production and secretion of IgG antibodies. Described plasmid also contains specifically designed gRNAs targeting the appropriate fusion constructs (either dCas9-VPR or dCas9-KRAB) to candidate genes *RUNX1*, *RUNX3*, *SPINK4*, and *ELL2*. As a proof of concept (*i.e.*, positive control), we targeted dCas9-KRAB to the promoter region of the *B4GALT1* gene, coding for a glycosyltransferase responsible for IgG galactosylation. We observed a significant decrease in the *B4GALT1* gene expression level and subsequent decrease in galactosylated glycans, with a concomitant increase in

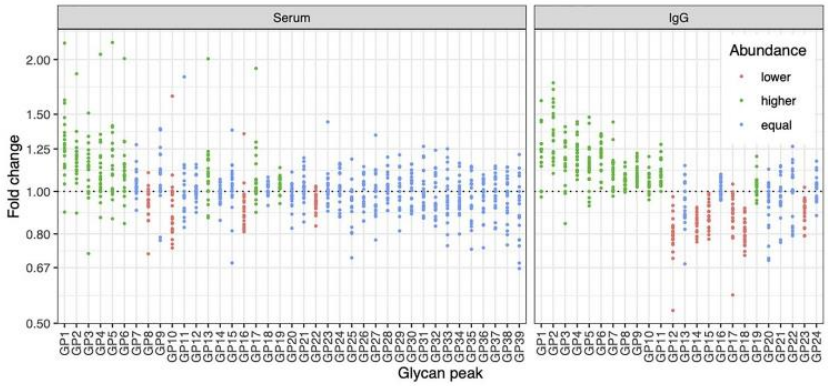

FIGURE 4 | Effects of gonadal hormone depletion on total plasma glycans and IgG glycans. Gonadotropin-releasing hormone analogue leuprolide acetate ($GnRH_{AG}$) was used to lower gonadal steroids to postmenopausal levels in healthy premenopausal women (n = 36) that were then randomized to transdermal placebo (n = 21) or estradiol patch (n = 15). Changes in the total plasma glycome and IgG glycome composition after five months of $GnRH_{AG}$ without supplementation of E_2 (supplementation with placebo) are shown. Each dot is a change in a single individual. Changes that are statistically significant after correction for multiple testing are shown in red (statistically significant decrease) or in green (statistically significant increase).

agalactosylated glycans, as expected (**Figure 5A**). Subsequently, we upregulated *RUNX1*, *RUNX3*, and *SPINK4* genes using dCas9-VPR and downregulated *RUNX1*, *RUNX3* and *ELL2* genes using dCas9-KRAB. Using specific gRNA for our targets, we found significant changes in the expression of *RUNX1*/VPR, *RUNX3*/VPR, *RUNX3*/KRAB, and *SPINK4*/VPR. The changes in gene expression were replicated in two independent sets of experiments (**Supplementary Table 4**). However, only the changes in *RUNX3*/VPR and *SPINK4*/VPR were accompanied by significant change in IgG glycosylation profile, in both cases related to the level of IgG galactosylation (**Figure 5B**). We found a decrease in galactosylated glycans upon upregulation of *RUNX3* and *SPINK4* and a concomitant increase in agalactosylated glycans in the case of *RUNX3* (**Figure 5B**). The dataset for all glycan traits is available in **Supplementary Table 5**.

DISCUSSION

The composition of IgG glycome is an essential aspect in the regulation of the immune system (1). However, molecular mechanisms contributing to changes of the IgG glycome composition are only vaguely understood. Here we show that estradiol is an important factor in regulating IgG glycosylation in women and that its effects on N-glycosylation are limited explicitly to B cells, as depletion of E_2 did not cause N-glycosylation changes of other serum proteins. Previous analysis of the same cohort of patients demonstrated that depletion of E_2 decreases galactosylation of IgG (14), and here we expand this finding to the decrease of sialylation and an increase of bisecting GlcNAc. Particularly interesting is the change in the ratio of sialylation and galactosylation (S/G, **Figure 3**, **Supplementary Table 1**), suggesting that the depletion of gonadal hormones directly affects sialylation and that the decrease in sialylation is not just a reflection of decreased galactose levels (needed for subsequent sialylation). It was previously reported that estrogen affects the expression of the *ST6GAL1* gene in both mice and humans (37). Unfortunately, we were not able to prove this using ours *in vitro* transient expression system. Therefore mechanistic aspects of this association remain to be demonstrated.

When mechanisms regulating protein glycosylation are investigated, the focus is always on the expression of glycosyltransferases, the enzymes that synthesize glycans (38). Nevertheless, in general, there is a slight correlation between glycosyltransferase expression levels and levels of glycans it synthesizes (39), which indicates that regulatory mechanisms may be more complex than a simple change in the expression of glycosyltransferases. Indeed, in a series of GWAS studies performed in the last decade, we identified a network of at least 30 genes associated with and potentially involved in regulating IgG glycosylation (32–35).

One of the potential mechanisms by which E_2 could increase IgG galactosylation is through direct activation of the B4GALT1 galactosyltransferase, which adds galactose to IgG. *In vitro* studies showed that both overexpression of estrogen receptor (40) and treatment of cells with E_2 leads to increased expression of the *B4GALT1* gene (41, 42). Our study decreased *B4GALT1* expression using CRISPR/dCas9-KRAB fusion in a unique IgG-secreting model cell system resulting in the expected change of the IgG glycome composition. It confirmed the importance of this critical biosynthetic enzyme for IgG galactosylation and proved the efficacy of our FreeStyle 293-F cell system, containing stably integrated dCas9-VPR and -KRAB and secreting IgG, for functional validation of GWAS hits for IgG glycosylation.

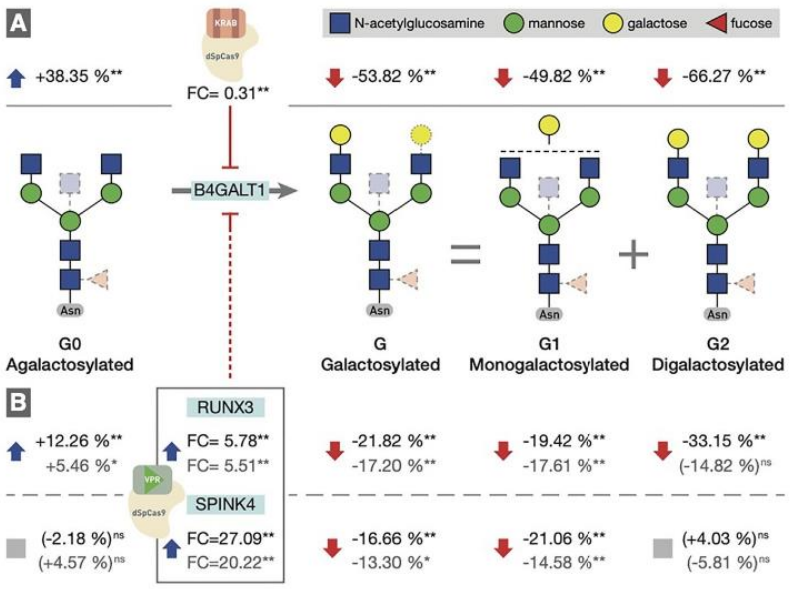

FIGURE 5 | **(A)** Downregulation of the *B4GALT1* gene by dCas9-KRAB induced changes in IgG galactosylation. Fold change (FC) between cells in which *B4GALT1* was directly downregulated by dCas9/KRAB and control cells (containing non-targeting gRNA) was 0.31, and subsequent change was recorded in IgG glycan phenotype: an increase of agalactosylated glycan structures appeared with a concomitant decrease in mono- and digalactosylated glycans (G1, G2). Corresponding changes in glycan structures are given as a relative change with non-targeting gRNA glycan levels as a baseline. Agalactosylated glycan structures (G0) are converted to galactosylated structures (G) by the enzymatic activity of B4GALT1. **(B)** Changes in IgG glycosylation resulting from upregulation of RUNX3 and SPINK4 by dCas9/VPR. FC values between cells in which *RUNX3* and *SPINK4* were directly upregulated by dCas9/VPR and control cells (containing non-targeting gRNA) are given for the first experiment (indicated in black) and the replicate (indicated in gray). The resulting putative inhibition of *B4GALT1* was confirmed indirectly by the effect on the glycosylation profile. The indirect and speculative nature of RUNX3/SPINK4 effect on galactosylation is indicated by the dashed red line. Statistical significance: *<0.05; **<0.01; ns, not significant.

By analysing the Signalling Pathways Project (SPP) web knowledgebase (36), we found that several other GWAS hits for the IgG glycosylation, which are not glycosyltransferases but genes with other functions, such as transcription factors *RUNX1* and *RUNX3*, as well as *SPINK4* and *ELL2*, also seem to be regulated by estradiol. However, the fact that these genes were both involved in the regulation of IgG glycosylation and affected by E_2 is not strong enough to prove their direct involvement in B cell IgG glycosylation because their effects could also be through an indirect mechanism. Using CRISPR/dCas9 molecular tools, we were able to increase the expression of *RUNX3*, which resulted in lower levels of IgG galactosylation (**Figure 5**).

With this experiment, we confirmed that *RUNX3* is involved in regulating IgG glycosylation in our *in vitro* model system of B cells (16). The first experimental validation of this GWAS hit positions *RUNX3* as a potential target for pharmacological interventions. *RUNX3* gene downregulation may improve IgG galactosylation and sialylation and may have potential anti-inflammatory effects. Activation of *SPINK4* had similar effects on the IgG glycome composition as activation of *RUNX3*. However, because the basal expression of *SPINK4* in HEK-293FS cells was low, we could not confirm if its suppression would have the opposite effect. Furthermore, *SPINK4* is located in relative proximity to *B4GALT1* (i.e., 50 kb distance). Therefore, although we did not observe a statistically significant change in transcript levels at the time of analysis, we cannot exclude the possibility that the binding of the dCas9-VPR fusion construct, used for the activation of SPINK4 in the region between *SPINK4* and *B4GALT1*, negatively affected *B4GALT1* expression. One putative mechanism could be spurious upregulation of *B4GALT1* antisense RNA 1 (*B4GALT1*-AS1), which is also located in this region, although we have not verified this hypothesis experimentally. We did not observe any statistically significant effects of *ELL2* on IgG glycosylation. This was not surprising because *ELL2* is a GWAS hit for sialylation and our *in vitro* expression system produces IgG antibodies with very low levels of sialic acid which presents a difficulty for evaluation of effects on sialylation. Therefore, the role of *ELL2* in the regulation of IgG glycosylation by estrogen still needs further exploration.

For the first time, the molecular mechanism through which E2 could regulate IgG glycosylation has been identified and functionally validated in the present study. Considering multiple functional roles of IgG glycans in balancing the immune system, this pathway may be a target for the future

development of a new class of anti-inflammatory drugs acting downstream of E_2 and having only a subset of the molecular consequences of hormone therapy.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Scientific Advisory and Review Committee at the University of Colorado Anschutz Medical Campus. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

GL, VZ, and PN designed the study. JJ, JK, and MP performed glycosylation analysis and interpreted glycan data. AM and KM performed CRISPR/dCas9 gene manipulations in HEK-293FS cells, collected IgG antibodies, and interpreted the data. WK and KG performed the intervention study. DK, AF, FV and AV analysed the data. GL wrote the initial draft of the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

Glycosylation analysis was performed in Genos Glycoscience Research Laboratory and partly supported by the European Union's Horizon 2020 grant IMForFuture (grant #721815), European Structural and Investment Funds grants "Centre of Competence in Molecular Diagnostics grant" (#KK.01.2.2.03.0006), and "Croatian National Centre of Research Excellence in Personalized Healthcare" (#KK.01.1.1.01.0010) and European Regional Development Fund, under grant agreement No. KK.01.1.1.04.0085, project "Genomic engineering and gene regulation in cell lines and model organisms by CRISPR/Cas9 technology—CasMouse".

ACKNOWLEDGMENTS

We acknowledge the members of our research groups who carried out the day-to-day activities for the project. Finally, we want to thank the women who volunteered to participate in the study for their time and efforts.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fimmu.2021.680227/full#supplementary-material

REFERENCES

1. Seeling M, Brückner C, Nimmerjahn F. Differential Antibody Glycosylation in Autoimmunity: Sweet Biomarker or Modulator of Disease Activity? *Nat Rev Rheumatol* (2017) 13:621–30. doi: 10.1038/nrrheum.2017.146
2. Nimmerjahn F, Ravetch JV. Fcgamma Receptors as Regulators of Immune Responses. *Nat Rev Immunol* (2008) 8:34–47. doi: 10.1038/nri2206
3. Gudelj I, Lauc G, Pezer M. Immunoglobulin G Glycosylation in Aging and Diseases. *Cell Immunol* (2018) 333:65–79. doi: 10.1016/j.cellimm.2018.07.009
4. Wittenbecher C, Štambuk T, Kuxhaus O, Rudman N, Vučković F, Štambuk J, et al. Plasma N-Glycans as Emerging Biomarkers of Cardiometabolic Risk: A Prospective Investigation in the EPIC-Potsdam Cohort Study. *Diabetes Care* (2020) 43:661–8. doi: 10.2337/dc19-1507
5. Gudelj I, Salo PP, Trbojević-Akmačić I, Albers M, Primorac D, Perola M, et al. Low Galactosylation of IgG Associates With Higher Risk for Future Diagnosis of Rheumatoid Arthritis During 10 Years of Follow-Up. *Biochim Biophys Acta - Mol Basis Dis* (2018) 1864:2034–39. doi: 10.1016/j.bbadis.2018.03.018
6. 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 Rheum* (2010) 62:2239–48. doi: 10.1002/art.27533
7. Tanigaki K, Sacharidou A, Peng J, Chambliss KL, Yuhanna IS, Ghosh D, et al. Hyposialylated IgG Activates Endothelial IgG Receptor Fcγriib to Promote Obesity-Induced Insulin Resistance. *J Clin Invest* (2017) 128:309–22. doi: 10.1172/JCI89333
8. Peng J, Vongpatanasin W, Sacharidou A, Kifer D, Yuhanna IS, Banerjee S, et al. Supplementation With the Sialic Acid Precursor N-Acetyl-D-Mannosamine Breaks the Link Between Obesity and Hypertension. *Circulation* (2019) 140:2005–18. doi: 10.1161/CIRCULATIONAHA.119.043490
9. Hafkenscheid L, Moel E, Smolik I, Tanner S, Meng X, Jansen BC, et al. N-Linked Glycans in the Variable Domain of IgG Anti–Citrullinated Protein Antibodies Predict the Development of Rheumatoid Arthritis. *Arthritis Rheumatol* (2019) 71:1626–33. doi: 10.1002/art.40920
10. Krištić J, Vučković F, Menni C, Klarić L, Keser T, Beceheli I, et al. Glycans are a Novel Biomarker of Chronological and Biological Ages. *J Gerontol - Ser A Biol Sci Med Sci* (2014) 69:779–89. doi: 10.1093/gerona/glt190
11. De Haan N, Reiding KR, Driessen G, Van Der Burg M, Wuhrer M. Changes in Healthy Human IgG Fc-Glycosylation After Birth and During Early Childhood. *J Proteome Res* (2016) 15:1853–61. doi: 10.1021/acs.jproteome.6b00038
12. Cheng HD, Tirosh I, de Haan N, Stöckmann H, Adamczyk B, McManus CA, et al. Igg Fc Glycosylation as an Axis of Humoral Immunity in Childhood. *J Allergy Clin Immunol* (2020) 145:710–3.e9. doi: 10.1016/j.jaci.2019.10.012
13. Dall'Olio F, Vanhooren V, Chen CC, Slagboom PE, Wuhrer M. Franceschi C. N-Glycomic Biomarkers of Biological Aging and Longevity: A Link With Inflammaging. *Ageing Res Rev* (2013) 12:685–98. doi: 10.1016/j.arr.2012.02.002
14. Ercan A, Kohrt WM, Cui J, Deane KD, Pezer M, Yu EW, et al. Estrogens Regulate Glycosylation of IgG in Women and Men. *JCI Insight* (2017) 2: e89703. doi: 10.1172/jci.insight.89703
15. Hanić M, Lauc G, Trbojević-Akmačić I. N-Glycan Analysis by Ultra-Performance Liquid Chromatography and Capillary Gel Electrophoresis With Fluorescent Labeling. *Curr Protoc Protein Sci* (2019) 97:1–21. doi: 10.1002/cpps.95
16. Vink T, Oudshoorn-Dickmann M, Roza M, Reitsma JJ, de Jong RN. A Simple, Robust and Highly Efficient Transient Expression System for Producing Antibodies. *Methods* (2014) 65:5–10. doi: 10.1016/j.ymeth.2013.07.018
17. Shea KL, Gavin KM, Melanson EL, Gibbons E, Stavros A, Wolfe P, et al. Body Composition and Bone Mineral Density After Ovarian Hormone Suppression With or Without Estradiol Treatment. *Menopause* (2015) 22:1045–52. doi: 10.1097/GME.0000000000000430
18. Melanson EL, Lyden K, Gibbons E, Gavin KM, Wolfe P, Wierman ME, et al. Influence of Estradiol Status on Physical Activity in Premenopausal Women. *Med Sci Sports Exerc* (2018) 50:1704–9. doi: 10.1249/MSS.0000000000001598

19. Belchetz PE, Plant TM, Nakai Y, Keogh EJ, Knobil E. Hypophysial Responses to Continuous and Intermittent Delivery of Hypothalamic Gonadotropin-Releasing Hormone. *Sci (80-)* (1978) 202:631–3. doi: 10.1126/science.100883
20. Trbojević-Akmačić I, Ugrina I, Lauc G. Comparative Analysis and Validation of Different Steps in Glycomics Studies. *Methods Enzymol* (2017) 586:37–55. doi: 10.1016/bs.mie.2016.09.027
21. Josipović G, Tadić V, Klasić M, Zanki V, Bečeheli I, Chung F, et al. Antagonistic and Synergistic Epigenetic Modulation Using Orthologous CRISPR/dCas9-based Modular System. *Nucleic Acids Res* (2019) 47:9637–57. doi: 10.1093/nar/gkz709
22. Josipović G, Zoldoš V, Vojta A. Active Fusions of Cas9 Orthologs. *J Biotechnol* (2019) 301:18–23. doi: 10.1016/j.jbiotec.2019.05.306
23. Livak KJ, Schmittgen TD. Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* (2001) 25:402–8. doi: 10.1006/meth.2001.1262S1046-2023(01)91262-9[pii
24. Keser T, Pavic T, Lauc G, Gornik O. Comparison of 2-Aminobenzamide, Procainamide and RapiFluor-MS as Derivatizing Agents for High-Throughput HILIC-UPLC-FLR-MS N-Glycan Analysis. *Front Chem* (2018) 6:324. doi: 10.3389/fchem.2018.00324
25. Leek JT, Johnson WE, Parker HS, Jaffe AE, Storey JD. The Sva Package for Removing Batch Effects and Other Unwanted Variation in High-Throughput Experiments. *Bioinformatics* (2012) 28:882–3. doi: 10.1093/bioinformatics/bts034
26. Bates D, Mächler M, Bolker B, Walker S. Fitting Linear Mixed-Effects Models Using Lme4. *J Stat Softw* (2015) 67:1–48. doi: 10.18637/jss.v067.i01
27. Karssen LC, van Duijn CM, Aulchenko YS. The GenABEL Project for statistical genomics. *F1000Res* (2016) 5:914. doi: 10.12688/f1000research.8733.1
28. Team RDC. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: The R Foundation (2009). Available at: http://www.r-project.org.
29. Vilaj M, Gudelj I, Trbojević-Akmačić I, Lauc G, Pezer M. Igg Glycans as a Biomarker of Biological Age. In: A Moskalev, editor. *Biomarkers of Human Aging. Healthy Ageing and Longevity*. Cham: Springer (2019). p. 81–99. doi: 10.1007/978-3-030-24970-0_7
30. Clerc F, Reiding KR, Jansen BC, Kammeijer GSM, Bondt A, Wuhrer M. Human Plasma Protein N-Glycosylation. *Glycoconj J* (2016) 33:309–43. doi: 10.1007/s10719-015-9626-2
31. Pučić M, Knežević A, Vidič J, Adamczyk B, Novokmet M, Polašek 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. doi: 10.1074/mcp.M111.010090
32. Krištić J, Zaytseva OO, Ram R, Nguyen Q, Novokmet M, Vučković F, et al. Profiling and Genetic Control of the Murine Immunoglobulin G Glycome. *Nat Chem Biol* (2018) 14:516–24. doi: 10.1038/s41589-018-0034-3
33. Lauc G, Huffman JE, Pučić M, Zgaga L, Adamczyk B, Mužinić 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. doi: 10.1371/journal.pgen.1003225
34. Shen X, Klarić L, Sharapov S, Mangino M, Ning Z, Wu D, et al. Multivariate Discovery and Replication of Five Novel Loci Associated With Immunoglobulin G N-Glycosylation. *Nat Commun* (2017) 8:447. doi: 10.1038/s41467-017-00453-3
35. Klarić L, Tsepilov YA, Stanton CM, Mangino M, Sikka TT, Esko T, et al. Glycosylation of Immunoglobulin G Is Regulated by a Large Network of Genes Pleiotropic With Inflammatory Diseases. *Sci Adv* (2020) 6:eaax0301. doi: 10.1126/sciadv.aax0301
36. Ochsner SA, Abraham D, Martin K, Ding W, McOwiti A, Kankanamge W, et al. The Signaling Pathways Project, an Integrated 'Omics Knowledgebase for Mammalian Cellular Signaling Pathways. *Sci Data* (2019) 6:252. doi: 10.1038/s41597-019-0193-4
37. Engdahl C, Bondt A, Harre U, Raufer J, Pfeifle R, Camponeschi A, et al. Estrogen Induces St6gal1 Expression and Increases IgG Sialylation in Mice and Patients With Rheumatoid Arthritis: A Potential Explanation for the Increased Risk of Rheumatoid Arthritis in Postmenopausal Women. *Arthritis Res Ther* (2018) 20(1):84. doi: 10.1186/s13075-018-1586-z
38. Moremen KW, Tiemeyer M, Nairn AV. Vertebrate Protein Glycosylation: Diversity, Synthesis and Function. *Nat Rev Mol Cell Biol* (2012) 13:448–62. doi: 10.1038/nrm3383
39. Nairn AV, Aoki K, dela Rosa M, Porterfield M, Lim JM, Kulik M, et al. Regulation of Glycan Structures in Murine Embryonic Stem Cells: Combined Transcript Profiling of Glycan-Related Genes and Glycan Structural Analysis. *J Biol Chem* (2012) 287:37835–56. doi: 10.1074/jbc.M112.405233
40. Nott SL, Huang Y, Li X, Fluharty BR, Qiu X, Welshons WV, et al. Genomic Responses From the Estrogen-Responsive Element-Dependent Signaling Pathway Mediated by Estrogen Receptor α Are Required to Elicit Cellular Alterations. *J Biol Chem* (2009) 284:15277–88. doi: 10.1074/jbc.M900365200
41. Hah N, Danko CG, Core L, Waterfall JJ, Siepel A, Lis JT, et al. A Rapid, Extensive, and Transient Transcriptional Response to Estrogen Signaling in Breast Cancer Cells. *Cell* (2011) 145:622–34. doi: 10.1016/j.cell.2011.03.042
42. Coser KR, Chesnes J, Hur J, Ray S, Isselbacher KJ, Shioda T. Global Analysis of Ligand Sensitivity of Estrogen Inducible and Suppressible Genes in MCF7/BUS Breast Cancer Cells by DNA Microarray. *Proc Natl Acad Sci USA* (2003) 100:13994–9. doi: 10.1073/pnas.2235866100

Conflict of Interest: GL is the founder and owner of Genos Ltd, a private research organization that specializes in high-throughput glycomic analyses and has several patents in this field. JJ, JK, AF, FV and MP are employees of Genos Ltd.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The handling editor declared a past co-authorship with one of the authors GL.

2.3. Heritability of the Glycan Clock of Biological Age

frontiers | Frontiers in Cell and Developmental Biology

TYPE Original Research
PUBLISHED 22 December 2022
DOI 10.3389/fcell.2022.982609

OPEN ACCESS

EDITED BY
Anja Lux,
University of Erlangen Nuremberg, Germany

REVIEWED BY
Ulrike Steffen,
University of Erlangen Nuremberg, Germany
Ozren Polašek,
University of Split, Croatia

*CORRESPONDENCE
Gordan Lauc,
glauc@genos.hr

SPECIALTY SECTION
This article was submitted to Molecular and Cellular Pathology, a section of the journal Frontiers in Cell and Developmental Biology

RECEIVED 30 June 2022
ACCEPTED 24 November 2022
PUBLISHED 22 December 2022

CITATION
Mijakovac A, Frkatović A, Hanić M, Ivok J, Martinić Kavur M, Pučić-Baković M, Spector T, Zoldoš V, Mangino M and Lauc G (2022), Heritability of the glycan clock of biological age.
Front. Cell Dev. Biol. 10:982609.
doi: 10.3389/fcell.2022.982609

Heritability of the glycan clock of biological age

Anika Mijakovac[1], Azra Frkatović[2], Maja Hanić[2], Jelena Ivok[2], Marina Martinić Kavur[2], Maja Pučić-Baković[2], Tim Spector[3], Vlatka Zoldoš[1], Massimo Mangino[3,4] and Gordan Lauc[2,5]*

[1]Division of Molecular Biology, Department of Biology, Faculty of Science, University of Zagreb, Zagreb, Croatia, [2]Genos Glycoscience Research Laboratory, Zagreb, Croatia, [3]Department of Twin Research and Genetic Epidemiology, King's College London, London, United Kingdom, [4]NIHR Biomedical Research Centre at Guy's and St Thoma's Foundation Trust, London, United Kingdom, [5]Department of Biochemistry and Molecular Biology, Faculty of Pharmacy and Biochemistry, University of Zagreb, Zagreb, Croatia

Immunoglobulin G is posttranslationally modified by the addition of complex N-glycans affecting its function and mediating inflammation at multiple levels. IgG glycome composition changes with age and health in a predictive pattern, presumably due to inflammaging. As a result, a novel biological aging biomarker, glycan clock of age, was developed. Glycan clock of age is the first of biological aging clocks for which multiple studies showed a possibility of clock reversal even with simple lifestyle interventions. However, none of the previous studies determined to which extent the glycan clock can be turned, and how much is fixed by genetic predisposition. To determine the contribution of genetic and environmental factors to phenotypic variation of the glycan clock, we performed heritability analysis on two TwinsUK female cohorts. IgG glycans from monozygotic and dizygotic twin pairs were analyzed by UHPLC and glycan age was calculated using the glycan clock. In order to determine additive genetic, shared, and unique environmental contributions, a classical twin design was applied. Heritability of the glycan clock was calculated for participants of one cross-sectional and one longitudinal cohort with three time points to assess the reliability of measurements. Heritability estimate for the glycan clock was 39% on average, suggesting a moderate contribution of additive genetic factors (A) to glycan clock variation. Remarkably, heritability estimates remained approximately the same in all time points of the longitudinal study, even though IgG glycome composition changed substantially. Most environmental contributions came from shared environmental factors (C), with unique environmental factors (E) having a minor role. Interestingly, heritability estimates nearly doubled, to an average of 71%, when we included age as a covariant. This intervention also inflated the estimates of unique environmental factors contributing to glycan clock variation. A complex interplay between genetic and environmental factors defines alternative IgG glycosylation during aging and, consequently, dictates the glycan clock's ticking. Apparently, environmental factors (including lifestyle choices) have a strong impact on the biological age measured with the glycan clock, which additionally clarifies why this aging clock is one of the most potent biomarkers of biological aging.

KEYWORDS

heritability, glycan clock, aging biomarker, biological age, IgG glycosylation

Introduction

Glycosylation is a series of enzymatic reactions in which carbohydrates are attached to other molecules (e.g., proteins or lipids) resulting in the formation of complex carbohydrates and glycoconjugates commonly referred to as "glycans." Protein glycosylation is one of the most frequent secondary modifications (Mariño et al., 2010). The addition of different glycan extensions (alternative glycosylation) greatly affects the structure and function of glycoproteins and it can be compared to changes in protein sequences (Dalziel et al., 2014). The main difference is that genes unquestionably determine the protein sequence, while there is no genetic template for the glycans (Taniguchi et al., 2014). Instead, glycosylation is controlled by many genes and their products, interacting in complex networks that are furthermore influenced by epigenetic modifications and the environment (Freeze, 2006; Abbott et al., 2008; Nairn et al., 2008; Lauc et al., 2014; Klarić et al., 2020). Heritability analysis of plasma glycans revealed that the majority of traits have high heritability estimates, indicating a tight genetic control of glycosylation (Zaytseva et al., 2020). Similar results were observed for glycans attached to immunoglobulin G (IgG), where only a minority of traits exhibited higher environmental contribution (Menni et al., 2013). Glycosylation of IgG antibody is especially interesting as it dramatically affects its function and acts as a molecular switch between pro- and anti-inflammatory immune responses (Gornik et al., 2012; Li et al., 2017; Alter et al., 2018). Aberrant IgG glycosylation is commonly observed in various pathological states, including autoimmune and inflammatory disorders, but the largest change of IgG glycome composition occurs during aging and in age-related conditions (Gudelj et al., 2018).

Aging is defined as the accumulation of molecular, cellular and organ damage over time that leads to loss of function and, consequently, increased vulnerability to disease (Fontana et al., 2010). This age-associated physiological decline is termed biological aging and it separates individuals of the same chronological age based on their health and functionality (Kavur et al., 2021). The accurate prediction of chronological and biological age from biochemical parameters became a priority in the aging field (Mitnitski and Rockwood, 2014; Jylhävä et al., 2017; Porter et al., 2021). It is hypothesized that biological age, influenced by different molecular hallmarks such as telomere shortening, genomic instability and cellular senescence, gives rise to age-related disease risk. Therefore, biological age is a much more potent prognostic tool for health outcomes than chronological age, and more importantly, it can be reversed (Jurić et al., 2020; Greto et al., 2021; Macdonald-Dunlop et al., 2022). Since this notion has been proposed, different predictors of biological age, termed aging clocks, were constructed using various methods (Horvath et al., 2020; Macdonald-Dunlop et al., 2022).

One of the most prominent aging clocks are based on DNA methylation, which is strongly associated with chronological age (Hannum et al., 2013; Horvath, 2013; Porter et al., 2021). Unlike DNA methylation, glycosylation of IgG does not predict chronological age with high accuracy. IgG glycosylation changes with aging in a predictive pattern which lead to the development of the glycan clock, based entirely on N-linked glycans attached to IgG. The glycan clock predicts chronological age with an estimated error of 9.7 years, but its acceleration associates with various biochemical and physiological traits related to inflammation and poor metabolic health reflective of biological aging (Krištić et al., 2014). The high plasticity of IgG glycome in response to environmental stimuli and its unique role in the immune response make the glycan clock one of the best predictors of biological age (Le Couteur et al., 2014; Russell et al., 2018). New studies also showed a possibility of glycan clock reversal with simple lifestyle changes such as weight loss, dietary supplements and exercise (Peng et al., 2019; Tijardović et al., 2019; Greto et al., 2021; Deriš et al., 2022b) as well as medical interventions such as bariatric surgery and regulation of sex hormones (Jurić et al., 2020; Greto et al., 2021).

Nevertheless, no research was done to determine to which extent the glycan clock can be modified and how much it depends on fixed genetic information. To answer this question, we performed a heritability analysis on the glycan clock data from TwinsUK (Verdi et al., 2019). A classical twin design (Visscher, 2004) enabled us to differentiate the contribution of genetic, shared environmental and unique environmental factors to phenotypic variation of the glycan clock. To shed more light on the effect of chronological age on the glycan clock variation, we performed the same analysis including age as a covariate. With this study, we gave an answer to an intriguing question on how much our environment and lifestyle choices influence the biological age measured by the glycan clock and to what proportion genes determine the glycan age.

Materials and methods

Study participants

Participants of this study were monozygotic (MZ) and dizygotic (DZ) adult female twins from the TwinsUK cohort. The TwinsUK is a nationwide same-sex twin registry based in the United Kingdom hosted by the Department of Twin Research and Genetic Epidemiology at King's College London (Verdi et al., 2019). All twins were recruited as volunteers and were not selected by any particular trait. In this study, we included

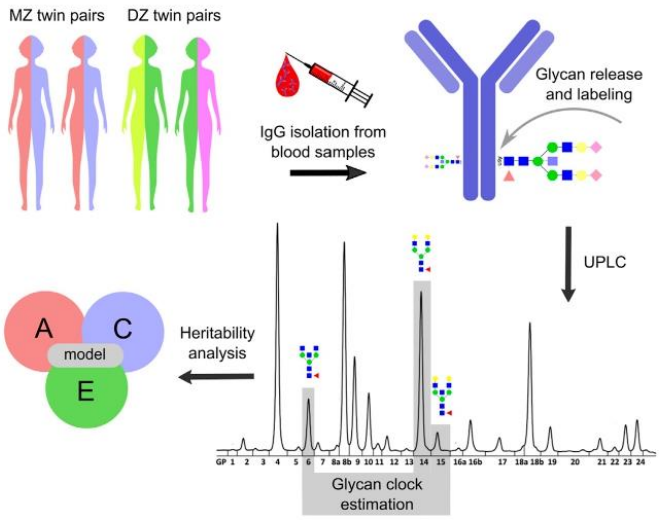

FIGURE 1
Study design. Heritability analysis was performed on the glycan clock data from the TwinsUK registry. The study included 479 monozygotic (MZ) and 1,193 dizygotic (DZ) female twin pairs. Blood samples were taken from twin pairs involved in one cross-sectional and one longitudinal cohort with three time points. IgG was isolated from collected blood samples which was followed by N-glycan release and labeling. IgG N-glycome composition was analyzed for the cross-sectional cohort by ultra-high-performance liquid chromatography (UHPLC) and the resulting IgG N-glycan chromatograms were separated into 24 distinct peaks. The glycan clock was estimated from GP6, GP14 and GP15 peaks as described in Krištić et al., 2014. Utilizing ACE model the heritability of the glycan clock was calculated. From this model, the contribution of genetic (A), shared environmental (C) and unique environmental (E) factors on the phenotypic variation of the glycan clock was estimated. Note that the IgG N-glycan analysis was slightly different for the longitudinal cohort (described in Materials and methods) but was not included for the simplicity of the display.

4,282 female twins for whom the IgG glycome measurements were obtained as a part of a cross-sectional study (Freidin et al., 2016). Among these, 1,598 female twins were later included in a longitudinal study where the IgG glycome was assessed in three time points. The number of male twins was small in both cohorts to allow for accurate estimation of model for the glycan clock calculation so all male participants were excluded from the study.

Immunoglobulin G isolation and glycan analysis from plasma samples

IgG was isolated from plasma samples obtained from MZ and DZ twin pairs (Figure 1). IgG isolation procedure was carried out using protein G monolithic plates (BIA Separations, Ajdovščina, Slovenia) as described previously (Pučić et al., 2011). Briefly, 50–100 µl of plasma was diluted with 1x PBS (pH = 7.4) in a 1:7 ratio and applied to the protein G plate. The plate was then washed with 1x PBS (pH = 7.4) in order to remove unbound proteins. Purified IgG was eluted with 0.1 M formic acid (Merck, Darmstadt, Germany) to a final volume of 1 ml and neutralized with 1 M ammonium bicarbonate (Merck, Darmstadt, Germany). Purified IgG was dried in a vacuum centrifuge. IgG N-glycan release and purification for samples from the longitudinal cohort was done with GlycoWorks RapiFluor-MS N-Glycan Kit obtained from Waters Corporation (United States) as described previously (Deriš et al., 2022a). IgG samples from the cross-sectional cohort were first denatured with 1.33% SDS sodium dodecyl sulfate (Invitrogen, Carlsbad, CA) and deglycosylated with PNGase F (ProZyme) overnight at 37°C. Released N-glycans were labeled with 2-aminobenzamide (Sigma-Aldrich) as described previously (Freidin et al., 2016). All labeled N-glycans were then purified by hydrophilic interaction chromatography solid phase extraction (HILIC-SPE). Both RapiFlour-MS labeled IgG N-glycans and 2-aminobenzamide labeled IgG N-glycans were analyzed using ultra-high-performance liquid chromatography based on hydrophilic interactions with fluorescence detection (HILIC-UHPLC-FLD) on Waters Acquity UPLC H-class instruments as described previously (Freidin et al., 2016; Deriš et al., 2022a). Obtained chromatograms of RapiFlour-MS labeled IgG N-glycans were separated into 22 peaks by automated integration (Agakova et al., 2017), while chromatograms of 2-aminobenzamide labeled IgG N-glycans were separated into

24 peaks using a traditional integration algorithm and manual correction (Supplementary Figure S1).

Glycan clock estimation

Total area normalization was applied to the area under the chromatogram peaks, followed by log-transformation and batch correction using the ComBat method as implemented in R package "sva" (version 3.30.1) (Leek et al., 2012). The glycan peak values were transformed back to the original scale prior to glycan clock estimation.

Glycan clock values were calculated according to Krištić et al., 2014 (Figure 1). First, clock model coefficients were estimated in the female samples in cross-sectional and longitudinal cohorts separately by including only one twin from each twin pair. For the cross-sectional cohort, glycan clock values were calculated using the following formula: $53.83 + (5.24 \times GP6) - (0.29 \times GP6^2) - (1.57 \times GP14) + (1.76 \times GP15)$. Taking into the consideration the difference in glycan peak annotation due to the difference in peak integration for the longitudinal cohort, formula for the glycan clock was as follows: $50.29 + (5.11 \times GP4) - (0.32 \times GP4^2) - (1.24 \times GP12) + (1.49 \times GP13)$. Data preprocessing and analysis were performed in R software (version 3.5.1.) (R Core Team, 2018).

Heritability analysis

Heritability of the glycan clock was estimated using the structural equation modeling (SEM) to decompose the observed phenotypic variance into three latent sources of variation: A—additive genetic variance, C—shared/common environment variance and E—unique environment variance (Neale and Cardon, 1992). Taking into account that the heritability studies require twin pairs, all participants without a co-twin were excluded thus leaving a total of 3,344 females separated into 479 MZ and 1193 DZ twin pairs in the cross-sectional cohort. The same thing was done for the measurements in the longitudinal cohort leaving a total of 549 monozygotic and 1,201 dizygotic twin pairs. Additive genetic effects (A) represent the cumulative impact of genes and they are indicated when the intrapair phenotypic correlation for monozygotic twins (rMZ) is greater than the intrapair phenotypic correlation for dizygotic twins (rDZ). Shared environmental effects (C) result from influences to which both members of a twin pair are exposed regardless of zygosity and contribute equally to rMZ and rDZ, thus, increasing twin similarity. Unique environmental effects (E) are events occurring to one twin but not the other and serve to decrease twin similarity. Unique environmental effects also include the measurement error. Utilizing listed factors (A, C and E), the best model was determined (Figure 1; Supplementary Tables S1, S2). That was done by sequential removal of each factor from the full model (ACE) followed by the likelihood ratio test ($p = 0.05$) to check the deterioration in fit of the various nested models. The best fitting model was selected using the Akaike information criterion (AIC). Lastly, the heritability of the glycan clock was estimated using the most parsimonious model as a proportion of the observed phenotypic variation attributable to genetic factors. Heritability analyses were performed using the package METs (version 1.2.7.1) (Scheike et al., 2014) in R (version 4.0.2).

Results

Glycan clock heritability was estimated in a sample of 3,344 female adult twins from the TwinsUK registry that were a part of one cross-sectional and one longitudinal cohort with three time points. Blood samples from 479 MZ and 1193 DZ twin pairs in the cross-sectional cohort and blood samples from 549 MZ and 1201 DZ twin pairs in the longitudinal cohort were used for the analysis of IgG glycome composition (Figure 1). Representative IgG glycan chromatograms are given in Supplementary Figure S1. Mean time difference was 7.5 years for points 1 and 2 (sd = 3.27) and 6 years for points 2 and 3 (sd = 2.5) in the longitudinal cohort. Glycan clock was estimated from the released IgG glycans and used to calculate heritability. The best fitting model for both cohorts was the full ACE model encompassing genetic (A), shared environmental (C) and unique environmental (E) components (Supplementary Table S1). Heritability estimation for the glycan clock in the cross-sectional study was 38.6%, which suggests a moderate contribution of additive genetic factors to this biomarker of aging (Figure 2A; Supplementary Table S2). The heritability of the glycan clock in the longitudinal cohort was 40%, 33%, and 43% for the first, second and third time point, respectively. The contribution of the environmental factors fell largely on shared environmental variance. The shared environmental influence on the glycan clock variation was 44.6% on average with a minor deviation between the cohorts and different time points. The smallest contribution came from unique environmental factors with an average of 16.8% (Figure 2A).

To better understand the impact of aging on heritability estimates of the glycan clock we performed the same analysis including the age of the participants as a covariate (Supplementary Table S3). This intervention largely inflated the additive genetic influences to an average of 65.5% in the cross-sectional cohort using ACE modeling (Figure 2B). The same phenomenon occurred in the longitudinal cohort but the best fitting model was AE (Supplementary Table S4) with an average genetic contribution of 72.9%. When we compared the results of both cohorts using ACE models, we observed very similar heritability estimates across all datasets averaging at 65.9%. Estimates of shared environmental influence dropped to 10.7%, and unique environmental contribution increased to

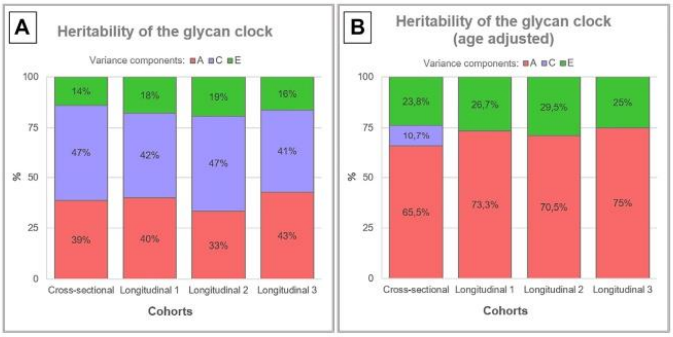

FIGURE 2
Heritability analysis of the glycan clock. Heritability of the glycan clock was calculated for the cross-sectional and longitudinal cohort with three time points. **(A)** The best fitting model was the full ACE model that includes additive genetic (A), shared environmental (C) and unique environmental (E) variance. **(B)** After correction for the age of the participants the full model for the longitudinal cohort was reduced to AE model.

23.8% in the cross-sectional cohort. The increase of the unique environmental component was also observed in the longitudinal cohort with all three time points averaging at 27.1% with minor deviations (Figure 2B).

Discussion

The glycan clock of age, based entirely on IgG N-glycans, can predict biological age with high accuracy (Krištić et al., 2014). To better understand genetic and environmental factors influencing the glycan clock variation, we performed a heritability analysis on the data from two cohorts included in the TwinsUK registry (Figure 1). This was the largest dataset for IgG glycan heritability analysis up to date which gave us the statistical power to detect the heritability of the glycan clock with 95% probability and enabled us to replicate the results to boost the reliability of the data. However, it is important to highlight that heritability studies do come with certain limitations, which can decrease the accuracy of acquired data. They are inherently restricted because all estimates heavily depend on environmental variance. Moreover, twin design assumes that correlations and interactions of genes and environment are minimal and cannot take the effect of epigenetics into account. All of these factors can falsely attribute the contribution of environmental factors to genetics leading to artificially inflated heritability estimates. However, classic heritability studies currently offer the best approximation of environmental and genetic contributions to the analyzed phenotype (Rijsdijk and Sham, 2002; Kaminsky et al., 2009; Manolio et al., 2009; Mayhew and Meyre, 2017).

Previous heritability studies conducted on IgG glycans demonstrated that IgG glycome has a variable heritability depending on the exact glycan structure analyzed. Most of the glycan traits turned out to be at least 50% heritable, with only a few having a low genetic contribution (Pučić et al., 2011; Menni et al., 2013). This tight genetic control was later explained by complex gene networks discovered through numerous GWA (genome-wide association) studies conducted on IgG glycan traits (Huffman et al., 2011; Lauc et al., 2013; Shen et al., 2017; Wahl et al., 2018b; Klarić et al., 2020). One of the first GWA studies on IgG glycan phenotype revealed that IgG glycosylation is not only regulated through the expression of glycosyltransferases that add specific sugars to the growing IgG glycan but through different genes with a previously unknown role in IgG glycosylation (Lauc et al., 2013). A recent GWAS discovered a gene network involving 27 loci implicated in the process of IgG glycosylation with most of them associated with various autoimmune and inflammatory conditions (Klarić et al., 2020). Furthermore, part of this gene network was functionally validated *in vitro* utilizing a recently developed transient expression system based on CRISPR tools (Mijakovac et al., 2021, 2022). Despite this tight genetic control of the IgG glycome that is currently being unraveled, heritability analysis of the glycan clock revealed only a moderate genetic contribution averaging around 39% for both cohorts. In the longitudinal cohort, IgG glycome composition changed substantially due to aging but the heritability estimates remained fixed. The contribution of the environmental factors fell largely on shared environmental variance with only a proportion coming from the unique environment.

Considering the fact that aging has the largest impact on IgG glycosylation, we included the age of the individuals as a covariate in the heritability analysis of the glycan clock. This intervention largely deflated the contribution of shared environmental variance to the point where AE model became the best fitting one for the longitudinal cohort. As a consequence,

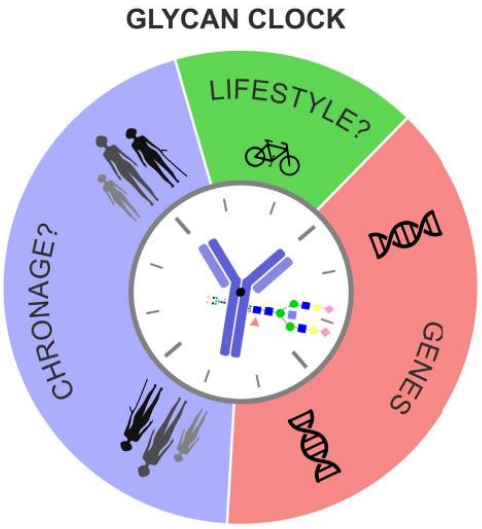

FIGURE 3
Genes, chronological age and lifestyle influence biological age measured by the glycan clock. Biological age, measured by the glycan clock is determined by a complex interplay between genetic and environmental factors. Even though the genetic template partly explains the phenotypic variation of the glycan clock this heritability analysis showed how big of a role environment has. Chronological age (ChronAge), that causes biological fatigue, together with fixed genetic information and unique lifestyle choices dictates the glycan clock ticking.

the heritability estimates almost doubled to an average of 71%. The observed increase in the genetic component could be a consequence of chronological age as a shared environmental variance characteristic for every individual and determined by their genetic makeup and epigenetic regulation. Aging in general leads to epigenetic mediated deregulation of genes so it is safe to assume that the glycan clock heritability estimates are not devoid of this effect (Saldanha and Watanabe, 2015). Different studies reported that the process of protein glycosylation is under strong epigenetic control (Horvat et al., 2013; Agrawal et al., 2014; Greville et al., 2016, 2021; Indellicato and Trinchera, 2021). We hypothesize that aging related epigenetic changes such as altered DNA methylation, abnormal chromatin state, altered histone modifications and deregulation of non-coding RNAs (Saldanha and Watanabe, 2015) also impact the glycan clock variation which is reflected in the inflated heritability estimates after age correction.

The glycan clock is a very powerful predictor of chronological and biological age but more importantly, it can be turned by simple lifestyle changes as many recent studies have reported. Lifestyle decisions such as exercise and nutrition can have an immense impact on biological age measured by the glycan clock (Peng et al., 2019; Tijardović et al., 2019; Greto et al., 2021). But how much can the glycan clock actually be turned remains an open question. We argue that the contribution of the unique environmental variance to the phenotypic variation of the glycan clock is comparable to the contribution of lifestyle decisions. After we corrected the data for the age of the individuals, estimates for the unique environmental variance averaged at 26%. This result emphasizes the high plasticity of the IgG glycome in response to environmental stimuli and in part supports the notion that the glycan clock can be rejuvenated by simple lifestyle choices. Mechanisms by which the lifestyle choices affect the glycan clock are still unknown, but most of the recent studies point to the epigenetic regulation of IgG glycosylation with an accent on DNA methylation (Menni et al., 2013; Lauc et al., 2014; Wahl et al., 2018a; Klasić and Zoldoš, 2021).

In conclusion, we propose that the biological age, measured by the glycan clock, is determined by a complex interplay of fixed genetic information, chronological age (ChronAge) of the individual and unique lifestyle choices, mediated by the plasticity of the human epigenome (Figure 3). Responsiveness of the glycan clock to a healthier lifestyle and its potential to integrate genetic, epigenetic and environmental cues puts this biomarker as one of the most alluring predictors of biological age in modern personalized medicine.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding author.

Ethics statement

The studies involving human participants were reviewed and approved by National Research Ethics Committees of the UK National Health Service (NHS) under the reference number 16/YH/0247. The TwinsUK study was approved by NRES Committee London–Westminster. The patients/participants provided their written informed consent to participate in this study. Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

Author contributions

GL and VZ designed the study. TS collected and processed the data from the TwinsUK registry. MP-B, MH, and JI performed IgG isolation and glycan analysis. AF performed

the glycan clock analysis. MM analyzed the heritability of the glycan clock. AM interpreted the data. AM and MM wrote the manuscript. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by the European Structural and Investment Funds grant for the Croatian National Centre of Competence in Molecular Diagnostics (Grant No. KK.01.2.2.03.0006); Croatian National Centre of Research Excellence in Personalized Healthcare grant (Grant No. KK.01.1.1.01.0010) and European Regional Development Fund Grant, project CasMouse (#KK.01.1.1.04.0085). TwinsUK receives funding from the Wellcome Trust (212904/Z/18/Z), Medical Research Council (AIMHY; MR/M016560/1) and European Union (H2020 contract #733100). TwinsUK and MM are supported by the National Institute for Health Research (NIHR)-funded BioResource, Clinical Research Facility and Biomedical Research Centre based at Guy's and St Thomas' NHS Foundation Trust in partnership with King's College London.

Acknowledgments

We are grateful to the twins who took part in TwinsUK and the whole TwinsUK team, which includes academic researchers, clinical staff, laboratory technicians, administrative staff and research managers.

Conflict of Interest

AF, MH, JI, MMK, MP-B, and GL are employed by the Genos Glycoscience Research Laboratory.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fcell.2022.982609/full#supplementary-material

References

Abbott, K. L., Nairn, A. V., Hall, E. M., Horton, M. B., McDonald, J. F., Moremen, K. W., et al. (2008). Focused glycomic analysis of the N-linked glycan biosynthetic pathway in ovarian cancer. *Proteomics* 8, 3210–3220. doi:10.1002/pmic.200800157

Agakova, A., Vučković, F., Klarić, L., Lauc, G., and Agakov, F. (2017). Automated integration of a UPLC glycomic profile. *Methods Mol. Biol.* 1503, 217–233. doi:10.1007/978-1-4939-6493-2_17

Agrawal, P., Kurcon, T., Pilobello, K. T., Rakus, J. F., Koppolu, S., Liu, Z., et al. (2014). Mapping posttranscriptional regulation of the human glycome uncovers microRNA defining the glycocode. *Proc. Natl. Acad. Sci. U. S. A.* 111, 4338–4343. doi:10.1073/pnas.1321524111

Alter, G., Ottenhoff, T. H. M., and Joosten, S. A. (2018). Antibody glycosylation in inflammation, disease and vaccination. *Semin. Immunol.* 39, 102–110. doi:10.1016/j.smim.2018.05.003

Dalziel, M., Crispin, M., Scanlan, C. N., Zitzmann, N., and Dwek, R. A. (2014). Emerging principles for the therapeutic exploitation of glycosylation. *Science* 80, 1235681. doi:10.1126/science.1235681

Deriš, H., Kifer, D., Cindrić, A., Petrović, T., Cvetko, A., Trbojević-Akmačić, I., et al. (2022a). Immunoglobulin G glycome composition in transition from premenopause to postmenopause. *iScience* 25, 103897. doi:10.1016/j.isci.2022.103897

Deriš, H., Tominac, P., Vučković, F., Briški, N., Astrup, A., Blaak, E. E., et al. (2022b). Effects of low-calorie and different weight-maintenance diets on IgG glycome composition. *Front. Immunol.* 13, 995186–995187. doi:10.3389/fimmu.2022.995186

Fontana, L., Partridge, L., and Longo, V. D. (2010). Dietary restriction, growth factors and aging: From yeast to humans. *Science* 328, 321–326. doi:10.1126/science.1172539

Freeze, H. H. (2006). Genetic defects in the human glycome. *Nat. Rev. Genet.* 7, 537–551. doi:10.1038/nrg1894

Freidin, M. B., Keser, T., Gudelj, I., Stambuk, J., Vucenovic, D., Allegri, M., et al. (2016). The association between low back pain and composition of IgG glycome. *Sci. Rep.* 6, 26815. doi:10.1038/srep26815

Gornik, O., Pavić, T., and Lauc, G. (2012). Alternative glycosylation modulates function of IgG and other proteins - implications on evolution and disease. *Biochim. Biophys. Acta* 1820, 1318–1326. doi:10.1016/j.bbagen.2011.12.004

Greto, V. L., Cvetko, A., Štambuk, T., Dempster, N. J., Kifer, D., Deriš, H., et al. (2021). Extensive weight loss reduces glycan age by altering IgG N-glycosylation. *Int. J. Obes.* 45, 1521–1531. doi:10.1038/s41366-021-00816-3

Greville, G., Llop, E., Howard, J., Madden, S. F., Perry, A. S., Peracaula, R., et al. (2021). 5-AZA-dC induces epigenetic changes associated with modified glycosylation of secreted glycoproteins and increased EMT and migration in chemo-sensitive cancer cells. *Clin. Epigenetics* 13, 34–16. doi:10.1186/s13148-021-01015-7

Greville, G., McCann, A., Rudd, P. M., and Saldova, R. (2016). Epigenetic regulation of glycosylation and the impact on chemo-resistance in breast and ovarian cancer. *Epigenetics* 11, 845–857. doi:10.1080/15592294.2016.1241932

Gudelj, I., Lauc, G., and Pezer, M. (2018). Immunoglobulin G glycosylation in aging and diseases. *Cell. Immunol.* 333, 65–79. doi:10.1016/j.cellimm.2018.07.009

Hannum, G., Guinney, J., Zhao, L., Zhang, L., Hughes, G., Sadda, S., et al. (2013). Genome-wide methylation profiles reveal quantitative views of human aging rates. *Mol. Cell* 24, 359–367. doi:10.1016/j.molcel.2012.10.016

Horvat, T., Deželjin, M., Redžić, I., Barišić, D., Herak Bosnar, M., Lauc, G., et al. (2013). Reversibility of membrane N-glycome of HeLa cells upon treatment with epigenetic inhibitors. *PLoS One* 8, e54672–e54679. doi:10.1371/journal.pone.0054672

Horvath, S. (2013). DNA methylation age of human tissues and cell types. *Genome Biol.* 14, R115. doi:10.1186/gb-2013-14-10-r115

Horvath, S., Singh, K., Raj, K., Khairnar, S., Sanghavi, A., Shrivastava, A., et al. (2020). Reversing age: Dual species measurement of epigenetic age with a single clock. *bioRxiv* 2020, 082917. doi:10.1101/2020.05.07.082917

Huffman, J. E., Knežević, A., Vitart, V., Kattla, J., Adamczyk, B., Novokmet, M., et al. (2011). Polymorphisms in B3GAT1, SLC9A9 and MGAT5 are associated with variation within the human plasma N-glycome of 3533 European adults. *Hum. Mol. Genet.* 20, 5000–5011. doi:10.1093/hmg/ddr414

Indellicato, R., and Trinchera, M. (2021). Epigenetic regulation of glycosylation in cancer and other diseases. *Int. J. Mol. Sci.* 22, 2980. doi:10.3390/ijms22062980

Jurić, J., Kohrt, W. M., Kifer, D., Gavin, K. M., Pezer, M., Nigrovic, P. A., et al. (2020). Effects of estradiol on biological age measured using the glycan age index. *Aging (Albany. NY)* 12, 19756–19765. doi:10.18632/aging.104060

Jylhävä, J., Pedersen, N. L., and Hägg, S. (2017). Biological age predictors. *EBioMedicine* 21, 29–36. doi:10.1016/j.ebiom.2017.03.046

Kaminsky, Z. A., Tang, T., Wang, S.-C., Ptak, C., Oh, G. H. T., Wong, A. H. C., et al. (2009). DNA methylation profiles in monozygotic and dizygotic twins. *Nat. Genet.* 41, 240–245. doi:10.1038/ng.286

Kavur, M. M., Lauc, G., and Pezer, M. (2021). in *Systems glycobiology: Immunoglobulin G glycans as biomarkers and functional effectors in aging and DiseasesComprehensive glycoscience*. Editors J. J. BarchiJr., S. Vidal, S. J. Williams, R. Narain, and N. Taniguchi (Elsevier), 439–478. doi:10.1016/B978-0-12-819475-1.00086-9

Klarić, L., Tsepilov, Y. A., Stanton, C. M., Mangino, M., Sikka, T. T., Esko, T., et al. (2020). Glycosylation of immunoglobulin G is regulated by a large network of genes pleiotropic with inflammatory diseases. *Sci. Adv.* 6, eaax0301. doi:10.1126/sciadv.aax0301

Klasić, M., and Zoldoš, V. (2021). "Epigenetics of immunoglobulin G glycosylation," in *Antibody glycosylation. Experientia supplementum* (Cham: Springer), 289–301. doi:10.1007/978-3-030-76912-3_9

Krištić, J., Vučković, F., Menni, C., Klarić, L., Keser, T., Beceheli, I., et al. (2014). Glycans are a novel biomarker of chronological and biological ages. *J. Gerontol. A Biol. Sci. Med. Sci.* 69, 779–789. doi:10.1093/gerona/glt190

Lauc, G., Huffman, J. E., Pučić, M., Zgaga, L., Adamczyk, B., Mužinić, A., et al. (2013). Loci associated with N-glycosylation of human immunoglobulin G show pleiotropy with autoimmune diseases and haematological cancers. *PLoS Genet.* 9, e1003225. doi:10.1371/journal.pgen.1003225

Lauc, G., Vojta, A., and Zoldoš, V. (2014). Epigenetic regulation of glycosylation is the quantum mechanics of biology. *Biochim. Biophys. Acta* 1840, 65–70. doi:10.1016/j.bbagen.2013.08.017

Le Couteur, D. G., Simpson, S. J., and De Cabo, R. (2014). Are glycans the holy grail for biomarkers of aging? (Comment on: Glycans Are a Novel Biomarker of Chronological and Biological Age by Kristic et al.). *J. Gerontol. A Biol. Sci. Med. Sci.* 69, 777–778. doi:10.1093/gerona/glt202

Leek, J. T., Johnson, W. E., Parker, H. S., Jaffe, A. E., and Storey, J. D. (2012). The SVA package for removing batch effects and other unwanted variation in high-throughput experiments. *Bioinformatics* 28, 882–883. doi:10.1093/bioinformatics/bts034

Li, T., DiLillo, D. J., Bournazos, S., Giddens, J. P., Ravetch, J. V., and Wang, L. X. (2017). Modulating IgG effector function by Fc glycan engineering. *Proc. Natl. Acad. Sci. U. S. A.* 114, 3485–3490. doi:10.1073/pnas.1702173114

Macdonald-Dunlop, E., Taba, N., Klarić, L., Frkatović, A., Walker, R., Hayward, C., et al. (2022). A catalogue of omics biological ageing clocks reveals substantial commonality and associations with disease risk. *Aging (Albany. NY)* 14, 623–659. doi:10.18632/aging.203847

Manolio, T. A., Collins, F. S., Cox, N. J., Goldstein, D. B., Hindorff, L. A., Hunter, D. J., et al. (2009). Finding the missing heritability of complex diseases. *Nature* 461, 747–753. doi:10.1038/nature08494

Mariño, K., Bones, J., Kattla, J. J., and Rudd, P. M. (2010). A systematic approach to protein glycosylation analysis: A path through the maze. *Nat. Chem. Biol.* 6, 713–723. doi:10.1038/nchembio.437

Mayhew, A. J., and Meyre, D. (2017). Assessing the heritability of complex traits in humans: Methodological challenges and opportunities. *Curr. Genomics* 18, 332–340. doi:10.2174/1389202918666170307161450

Menni, C., Keser, T., Mangino, M., Bell, J. T., Erte, I., Akmačić, I., et al. (2013). Glycosylation of immunoglobulin G: Role of genetic and epigenetic influences. *PLoS One* 8, e82558. doi:10.1371/journal.pone.0082558

Mijakovac, A., Jurić, J., Kohrt, W. M., Krištić, J., Kifer, D., Gavin, K. M., et al. (2021). Effects of estradiol on immunoglobulin G glycosylation: Mapping of the downstream signaling mechanism. *Front. Immunol.* 12, 680227. doi:10.3389/fimmu.2021.680227

Mijakovac, A., Miškec, K., Krištić, J., Vičić Bočkor, V., Tadić, V., Bošković, M., et al. (2022). A transient expression system with stably integrated CRISPR-dCas9 fusions for regulation of genes involved in immunoglobulin G glycosylation. *Cris. J.* 5, 237–253. doi:10.1089/crispr.2021.0089

Mitnitski, A., and Rockwood, K. (2014). Biological age revisited. *J. Gerontol. A Biol. Sci. Med. Sci.* 69, 295–296. doi:10.1093/gerona/glt137

Nairn, A. V., York, W. S., Harris, K., Hall, E. M., Pierce, J. M., and Moremen, K. W. (2008). Regulation of glycan structures in animal tissues: Transcript profiling of glycan-related genes. *J. Biol. Chem.* 283, 17298–17313. doi:10.1074/jbc.M801964200

Neale, M. C., and Cardon, L. R. (1992). *Methodology for genetic studies of twins and families*. 1st ed. Dordrecht: Springer.

Peng, J., Vongpatanasin, W., Sacharidou, A., Kifer, D., Yuhanna, I. S., Banerjee, S., et al. (2019). Supplementation with the sialic acid precursor N-acetyl-D-mannosamine breaks the link between obesity and hypertension. *Circulation* 140, 2005–2018. doi:10.1161/CIRCULATIONAHA.119.043490

Porter, H. L., Brown, C. A., Roopnarinesingh, X., Giles, C. B., Georgescu, C., Freeman, W. M., et al. (2021). Many chronological aging clocks can be found throughout the epigenome: Implications for quantifying biological aging. *Aging Cell* 20, e13492. doi:10.1111/acel.13492

Pučić, M., Knežević, A., Vidič, J., Adamczyk, B., Novokmet, M., Polašek, O., et al. (2011). High throughput isolation and glycosylation analysis of IgG-variability and heritability of the IgG glycome in three isolated human populations. *Mol. Cell. Proteomics* 10, M111.010090. doi:10.1074/mcp.M111.010090

R Core Team (2018). *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing.

Rijsdijk, F. V., and Sham, P. C. (2002). Analytic approaches to twin data using structural equation models. *Brief. Bioinform.* 3, 119–133. doi:10.1093/bib/3.2.119

Russell, A., Adua, E., Ugrina, I., Laws, S., and Wang, W. (2018). Unravelling immunoglobulin G Fc N-glycosylation: A dynamic marker potentiating predictive, preventive and personalised medicine. *Int. J. Mol. Sci.* 19, 390. doi:10.3390/ijms19020390

Saldanha, S. N., and Watanabe, L. P. (2015). Epigenetics and aging. *Epigenetics Dermatology* 4, 379–406. doi:10.1016/B978-0-12-800957-4.00018-7

Scheike, T. H., Holst, K. K., and Hjelmborg, J. B. (2014). Estimating heritability for cause specific mortality based on twin studies. *Lifetime Data Anal.* 20, 210–233. doi:10.1007/s10985-013-9244-x

Shen, X., Klarić, L., Sharapov, S., Mangino, M., Ning, Z., Wu, D., et al. (2017). Multivariate discovery and replication of five novel loci associated with Immunoglobulin G N-glycosylation. *Nat. Commun.* 8, 447. doi:10.1038/s41467-017-00453-3

Taniguchi, N., Honke, K., Fukuda, M., Narimatsu, H., Yamaguchi, Y., and Angata, T. (2014). in *Handbook of glycosyltransferases and related genes*. Editors N. Taniguchi, K. Honke, M. Fukuda, H. Narimatsu, Y. Yamaguchi, and T. Angata. second editionSecond edi (Springer). doi:10.1007/978-4-431-54240-7

Tijardović, M., Marijančević, D., Bok, D., Kifer, D., Lauc, G., Gornik, O., et al. (2019). Intense physical exercise induces an anti-inflammatory change in IgG N-glycosylation profile. *Front. Physiol.* 10, 1522. doi:10.3389/fphys.2019.01522

Verdi, S., Abbasian, G., Bowyer, R. C. E., Lachance, G., Yarand, D., Christofidou, P., et al. (2019). TwinsUK: The UK adult twin registry update. *Twin Res. Hum. Genet.* 22, 523–529. doi:10.1017/thg.2019.65

Visscher, P. M. (2004). Power of the classical twin design revisited. *Twin Res.* 7, 505–512. doi:10.1375/1369052042335250

Wahl, A., Kasela, S., Carnero-Montoro, E., van Iterson, M., Štambuk, J., Sharma, S., et al. (2018a). IgG glycosylation and DNA methylation are interconnected with smoking. *Biochim. Biophys. Acta. Gen. Subj.* 1862, 637–648. doi:10.1016/j.bbagen.2017.10.012

Wahl, A., van den Akker, E., Klaric, L., Štambuk, J., Benedetti, E., Plomp, R., et al. (2018b). Genome-wide association study on immunoglobulin G glycosylation patterns. *Front. Immunol.* 9, 277. doi:10.3389/fimmu.2018.00277

Zaytseva, O. O., Freidin, M. B., Keser, T., Štambuk, J., Ugrina, I., Šimurina, M., et al. (2020). Heritability of human plasma N-glycome. *J. Proteome Res.* 19, 85–91. doi:10.1021/acs.jproteome.9b00348

3. DISCUSSION

The complexity of the human proteome is amplified by the addition of glycans, compound sugars posttranslationally attached to proteins by the enzymatic process of glycosylation. The addition of glycans to proteins has a functional significance as it affects their structural and biological properties[1–3,12,229]. This is exemplified by the addition of glycans to the immunoglobulin G (IgG) antibody, creating hundreds of different IgG glycoforms that control the immune response on multiple levels[24,230]. Alternative IgG glycosylation is observed in various pathological states (e.g. autoimmune and inflammatory conditions, cancer, cardiometabolic diseases, etc.), but the most striking changes in healthy individuals occur as a consequence of aging[11]. This knowledge was exploited to create a novel marker of biological age, based exclusively on glycans attached to IgG, termed the glycan clock[51]. The glycan clock can predict chronological age with moderate accuracy, but its acceleration associates with various physiological and biochemical features related to inflammation, which is reflective of biological aging. A striking increase in the biological age, measured by the glycan clock, occurs in women entering the menopausal period[66]. The abundance of anti-inflammatory IgG glycans drops suddenly, switching the antibody pool to a pro-inflammatory status. This seems to be a consequence of estrogen depletion characteristic for the menopausal period, as demonstrated in the study by Ercan *et al.*[65]. This study confirmed that estrogen regulates IgG glycosylation *in vivo,* defining a pathway by which sex modulates the immune response. However, the mediators of this pathway are completely unknown. Moreover, the entire process that regulates IgG glycosylation in aging and disease has been puzzling scientists for decades.

Even though the addition of different glycans can impact the function of glycoproteins to the same extent as genetic mutations, the process of glycosylation is not template-driven. On the contrary, glycosylation is a complex trait, determined by a sophisticated interplay of many genes and their products, further influenced by epigenetic modifications and the environment[9]. The intricacy of these mechanisms made the research of IgG glycosylation regulation that more complicated. The largest progress in the field made in the last several years came from the genome-wide association (GWA) studies. GWAS and post-GWAS *in silico* analyses revealed novel gene networks potentially implicated in the regulation of IgG glycosylation, but the functional *in vitro* follow-up studies are still missing[10,129–132]. In order to perform functional studies of mechanisms governing IgG glycosylation, we developed a ground-breaking *in vitro* system based on the CRISPRa/CRISPRi technology. We stably integrated dCas9-VPR (CRISPRa) and dCas9-KRAB (CRISPRi) expression cassettes into the genome of FreeStyle™

293-F cells using transposon technology. Following genomic integration, we obtained monoclonal cell cultures expressing dCas9-VPR or dCas9-KRAB fusion proteins for targeted upregulation or downregulation of selected genes. By transient transfection of established monoclonal cell lines with a plasmid co-expressing specific gRNA molecules and recombinant IgG, we can manipulate candidate genes potentially implicated in IgG glycosylation and analyze the glycan phenotype of IgG in one step. We validated the developed system by targeting four main glycosyltransferases (GTs) with a known role in IgG glycosylation (FUT8, MGAT3, B4GALT1 and ST6GAL1)[4]. This intervention resulted in the expected change of IgG glycome composition, validating our cell lines for further research. To demonstrate that this system can be used for *in vitro* functional follow-up of GWA studies, we manipulated two GWAS candidate genes with a potential role in IgG glycosylation: SPPL3 and GGA2. Targeted downregulation of SPPL3 resulted in an increase of all complex glycan traits found on IgG, while the role of GGA2 in the process of IgG glycosylation could not be validated in our cell lines. To address a direct biological issue, we decided to utilize our system to discover the mediators of estrogen signaling responsible for a sudden drop of anti-inflammatory IgG glycoforms in aging women entering the menopausal period. Targeted CRISPR activation of estrogen-regulated loci: *RUNX3* and *SPINK4* resulted in a decrease of galactosylated IgG glycans, a phenotype characteristic for menopausal women. This was the first *in vitro* validation of an estrogen signaling pathway implicated in IgG glycosylation that positioned *RUNX3* and *SPINK4* loci as potential targets for pharmacological interventions. Targeting the *RUNX3* and *SPINK4* to silence their expression could potentially ameliorate the unfavorable functional changes of IgG glycome composition in aging women entering menopause. This study raised another important question regarding the magnitude of genetic influences on alternative IgG glycosylation characteristic for human aging. Various heritability studies performed in the last several years demonstrated that the process of IgG glycosylation operates under tight genetic control[30,75]. We wondered if it is possible that the environment plays a larger role when it comes to age-related changes in IgG glycome composition. Recent research demonstrated that lifestyle choices like exercise, dietary supplements and weight loss can have an extensive impact on IgG glycosylation[59–61,66]. Moreover, healthy lifestyle choices can reverse the process of biological aging measured by the glycan clock biomarker. To analyze the contribution of genetic and environmental factors on age-related changes in IgG glycome composition, we conducted the largest heritability study up to date on the glycan clock data from 3344 female twins enlisted in the TwinsUK registry[231]. We demonstrated a substantial contribution of unique lifestyle choices (~26%) to the phenotypic variation of the glycan clock. This discovery highlighted the plasticity

of IgG glycans in response to various environmental stimuli and positioned the glycan clock as one of the most promising markers of biological age. On the other hand, the heritability of the glycan clock turned out to be rather high (~39%) which was even more significant for our research. It emphasized the necessity to develop new technologies in order to decode the genetic mechanisms behind the process of IgG glycosylation.

The *in vitro* technology for functional validation of novel candidate genes implicated in the regulation of IgG glycosylation must satisfy two main criteria. This *in vitro* system must be able to secrete IgG in sufficient quantity for glycan analysis but also enable fast, efficient and reproducible manipulation of gene expression. With that in mind, we developed our transient expression system based on CRISPRa/CRISPRi tools for precise gene regulation in FreeStyle™ 293-F cells (HEK-293F). FreeStyle™ 293-F-based system for the production of IgG antibodies was developed by Vink *et al.* in 2014[137] and upgraded by Dekkers *et al.*[94] two years later to enable the secretion of IgG with defined Fc glycosylation. Both studies verified the FreeStyle™ 293-F cells to be an excellent system for the efficient production of high quantities of IgG with human-like glycosylation pattern. The introduction of this cell line into our laboratory confirmed the claims made by Vink *et al*. Another cell line widely used for high-throughput IgG production is the Chinese hamster ovary (CHO). Moreover, CHO-produced monoclonal antibodies account for the majority of antibody therapeutics[96,225,232–234]. However, CHOs are non-human mammalian cells unable to fully replicate the glycosylation patterns of human cells because they lack some GTs like MGAT3 and ST6GAL1, and have additional non-human posttranslational modifications[235–237]. Currently, one of the better models for human B cells are the lymphoblastoid cell lines (LCLs). These cells are easily established from B cells by Epstein-Bar virus (EBV) infection and can be maintained in cell culture for long periods of time[238]. Unfortunately, they do not provide a platform for high-throughput experimental approaches due to their low ability to accept transgenes either by common transfection or transduction methods[239,240]. On the other hand, FreeStyle™ 293-F cells are easy to transfect, produce large quantities of IgG and exhibit human-like glycosylation patterns making them the best option for fast and efficient analysis of mechanisms regulating IgG glycosylation[94,137,241].

In the first phase of our research, we upgraded the FreeStyle™ 293-F system by altering the IgG-coding plasmid to be compatible with our modular molecular toolbox based on the Golden Gate cloning method (Figure 5)[112,242]. This toolbox enables us to assemble up to seven functional units in a defined order in a single Golden Gate reaction with the BsaI enzyme. For the purpose of FreeStyle™ 293-F system validation, we assembled five functional units into an

expression plasmid: a strong promoter, IgG heavy chain, IgG light chain, a fluorescent marker and a transcription terminator. The fluorescent marker was assembled only to quantify the transfection efficiency and will be swapped with the specific gRNAs for CRISPRa/CRISPRi targeting discussed later. Lipid-based delivery (293fectin™) of our IgG expression vector to FreeStyle™ 293-F cells was exceptionally high, considering the plasmid size of approximately 6 kbs. The efficiency of the delivery approached 70% obviating the need for antibiotic selection. The highest yield of the secreted IgG was obtained on the fifth day after transfection when the IgG expression vector is co-transfected with three small plasmids for cell cycle arrest. These small plasmids increase the yield of secreted proteins and a complete protocol for their co-transfection in FreeStyle™ 293-F cells was obtained from Vink *et al*[137]. After the isolation of secreted IgG and *N*-linked glycan release, we analyzed IgG glycosylation by HILIC-UPLC (hydrophilic interaction liquid chromatography – ultraperformance liquid chromatography) in Genos Ltd. The obtained chromatograms contained only the Fc glycans similar to the ones attached to IgG isolated from human plasma. The Fab glycosylation site was not introduced in our IgG expression plasmid since we wished to analyze only the regulation of the Fc glycosylation as it directly affects the interaction of IgG with other cells of the immune system[230]. The acquisition of clear chromatographic peaks from the IgG antibodies secreted from FreeStyle™ 293-F cells marked the end of the first phase.

Next, we moved to the most delicate part of the project: the generation of monoclonal FreeStyle™ 293-F cell lines that stably express dCas9-VPR or dCas9-KRAB fusion proteins for targeted upregulation or downregulation of selected loci. The delivery of CRISPR/Cas9 cassettes to the cells of interest is notoriously difficult due to their large size, reaching up to 19 kb[183,185,186]. Classical transfection methods are extremely inefficient in the delivery of large plasmids and viral vectors usually cannot form functional viral particles with sequences of this length[243,244]. To circumvent this challenging step, we decided to split the CRISPR/dCas9 cassette into two parts: (1) dCas9-VPR/dCas9-KRAB expression cassette integrated into the genome of FreeStyle™ 293-F cells; (2) IgG + gRNAs expression vector for the transient transfection. To achieve stable genomic integration of dCas9-VPR/dCas9-KRAB cassettes, we opted for the transposon technology due to its simplicity and ease of handling. We constructed a PiggyBac (PB) based dual vector system comprised of an expression plasmid coding the hyperactive PB transposase and a transposon plasmid able to integrate any DNA sequence of interest into the host genome[245]. We extended the transposon plasmid with two insulator sequences at each side of the excision site to enable long-term stable expression of our

CRISPRa/CRISPRi cassettes. The novel transposon vector was also adapted for our modular molecular toolbox, enabling the assembly of any selected functional unit (Figure 5). For this experimental setup, we assembled two different dCas9 orthologs, dSaCas9 and dSpCas9, with the transcriptional activator VPR and the transcriptional repressor KRAB, respectively. The co-transfection of these dCas9 orthologs would allow for the antagonistic gene regulation if desired, but this approach was not used as a part of this doctoral thesis. The dCas9 effector proteins were assembled with a weak EFS promoter to mitigate the possible off-target effects together with a fluorescent marker and a puromycin resistance gene. Following the co-transfection of FreeStyle™ 293-F cells with the transposon dual vector carrying dCas9-VPR or dCas9-KRAB cassettes, we selected the cells with puromycin to generate the polyclonal FreeStyle™ 293-F cell culture with dCas9 tools integrated in the genome. Monoclonal FreeStyle™ 293-F cell lines were obtained by the method of limiting dilution. To verify that monoclonal FreeStyle™ 293-F cells still stably express dCas9-VPR or dCas9-KRAB protein we subjected them to another round of puromycin selection that had no effect on their viability. We also confirmed the presence of both dCas9 orthologs by the western blot method. To verify that the obtained monoclonal cells express functional dCas9 fusion proteins, we transiently transfected them with already available gRNAs tested in HEK293 cells for the manipulation of *HNF1α, MGAT3, IL6ST* and *BACH2* loci. As a negative control, we used two non-targeting gRNAs, one for each dCas9 ortholog. Non-targeting gRNAs form complexes with dCas9 protein but are not complementary to any genomic region, making them non-functional for dCas9 targeting. Targeted manipulation of the aforementioned loci in the monoclonal cell lines resulted in the same change of gene expression observed previously in HEK293 cells, further validating the stable expression of dCas9 fusion proteins[112].

To validate our novel HEK293-F system, based on monoclonal FreeStyle™ 293-F cells stably expressing dCas9-VPR/dCas9-KRAB fusion proteins, for the study of the regulation of IgG glycosylation we decided to target four main GTs with a known role in this process. The first step was to assemble specific gRNAs targeting each GT (FUT8, MGAT3, B4GALT1 and ST6GAL1) with the genes coding IgG heavy and light chains. This configuration needs to be emphasized as it enables the production of IgG only from the cells where the targeted manipulation takes place, obviating the need for any type of selection. The decision whether to activate or silence each GT was made based on the previously acquired IgG glycan chromatograms. The secreted IgG was highly fucosylated and moderately galactosylated with low levels of bisecting GlcNAc and sialylation. Therefore, we downregulated *FUT8* and

upregulated *MGAT3* and *ST6GAL1* loci. Considering the moderate levels of glycans containing galactose, we decided to target the *B4GALT1* locus in both monoclonal cell lines, facilitating its up- and downregulation. As a negative control, we assembled genes coding for IgG with two non-targeting gRNAs (described above). This plasmid served as a negative control in all experiments conducted in the HEK293-F system. The results of these interventions were in line with our hypothesis. Silencing of FUT8 resulted in a decrease of core fucosylation. Upregulation of MGAT3 and ST6GAL1 was followed by an increase of bisection and sialylation, respectively. Interestingly, no change in the ratio of fucosylated IgG glycans was observed after the upregulation of MGAT3 challenging the hypothesis from Garcia-Garica et *al.*[89] that bisection precludes fucosylation. Bidirectional manipulation of B4GALT1 resulted in the expected change of IgG galactosylation. Moreover, the glycome composition of IgG isolated from monoclonal cells transfected with non-targeting gRNAs was identical to the glycome of IgG isolated from native FreeStyle™ 293-F cells. This result together with a successful manipulation of every targeted GT and the consequent change in IgG glycome composition was a definite proof-of-concept experiment that validated the potential our HEK293-F system holds in studies of IgG glycosylation regulation.

The most challenging step in the generation of our system was the limiting dilution method. To omit this time-consuming step, we decided to manipulate the same GTs in the FreeStyle™ 293-F polyclonal cell lines. The results qualitatively matched those obtained in the monoclonal cells but some experiments did not reach statistical significance. Thus, we concluded that the polyclonal cell lines can be used when the resulting phenotype is robust but when the expected changes are small the monoclonal cells are a better option. Considering that IgG glycosylation is a complex trait, the manipulation of one gene in the molecular network would most likely exert only a small effect on the IgG glycome composition. For that reason, we decided to proceed with the monoclonal FreeStyle™ 293-F cells in all future experiments. To further upgrade our HEK293-F system, we extended the IgG-bearing plasmid with the magnetic-activated cell sorting technology (MACS). Regardless of the high transfection efficiencies obtained when transfecting the FreeStyle™ 293-F cells with IgG expression plasmid we wanted to validate our selection-free approach. We performed MACS-based enrichment of transfected cells after the manipulation of four main GTs but observed no difference in expression levels before and after the enrichment, validating our selection-free approach. This was the final experiment in the optimization of our novel HEK293-F system for the studies of gene networks that regulate IgG glycosylation.

One of the most pronounced advantages of the HEK293-F transient expression system is its simplicity and efficiency. When the monoclonal FreeStyle™ 293-F cells expressing the CRISPR/dCas9 tools are established they can be stored in liquid nitrogen indefinitely. Upon the demand for targeted transcriptional modulation of desired loci, it is possible to design and assemble specific gRNAs with IgG coding genes while the monoclonal cells are recovering from the thawing procedure. When the cells have adjusted to the culture conditions, they are co-transfected with the gRNA/IgG plasmid and three small plasmids for cell cycle arrest. IgG is collected on the fifth day after transfection and can be isolated for glycan analysis immediately. To verify transcriptional regulation of selected loci, whole RNA can be isolated in parallel and analyzed by RT-qPCR (reverse transcription quantitative PCR). The complete workflow lasts only five weeks and it is possible to target many loci at once enabling easy and high-throughput studies of gene functions in the process of IgG glycosylation. Another important characteristic of the HEK293-F system is its modularity. IgG-coding genes can be easily swapped for sequences coding any other protein of interest. Different fluorescent and selection markers can be assembled with up to 6 different gRNA molecules. Monoclonal cell lines can be generated to express other desired CRISPR/Cas9 tools, from classical Cas9 to dCas9 fused to various epigenetic modulators. We assembled a transposon vector for genomic integration of dCas9-DNMT3A and dCas9-TET1 for targeted methylation or demethylation of selected loci. However, the described configuration was not used as a part of this project and serves only as an example of the modularity of the HEK293-F system.

The idea behind the development of the HEK293-F system was to have a simple high-throughput *in vitro* system for the discovery of novel loci implicated in the regulation of IgG glycosylation. Considering that GWA studies made the largest progress in this field, the logical next step was to utilize the HEK293-F system for functional validation of GWAS candidate genes potentially involved in IgG glycosylation. We decided to target two novel Golgi-localized GWAS candidate genes associated with IgG sialylation: *GGA2* and *SPPL3*. GGA2 (Golgi Associated, Gamma Adaptin Ear Containing, ARF Binding protein 2) is a part of the ubiquitous coat protein family that regulates the trafficking of proteins between the *trans* Golgi network and endosomes[246]. Among its many roles is the regulation of retrograde transport of the phosphorylated form of BACE1 (β-secretase 1) from endosomes to the *trans* Golgi network[247]. BACE1 is a transmembrane aspartic protease that cleaves ST6GAL1 from the Golgi membranes to initiate the secretion of this GT into the plasma[248,249]. Jones *et al.*[106] hypothesized that IgG sialylation occurs in plasma and is mediated by the hepatocyte-secreted ST6GAL1,

cleaved by BACE1 protease. This hypothesis arose after B cell-specific KO of ST6GAL1 that resulted in no change of IgG sialylation, indicating that this process occurs outside of B cells. This hypothesis was recently rejected by Oswald *et al.*[107] where they showed that IgG sialylation remains unchanged in the hepatocyte-specific KO of ST6GAL1 or BACE1, B cell-specific KO of ST6GAL1 or any combination of these KOs. This experiment demonstrated that the addition of sialic acid on IgG glycans does not rely on plasma ST6GAL1 or Golgi-resident, B cell-specific ST6GAL1. When we directly upregulated ST6GAL1 in our HEK293-F transient expression system, IgG sialylation did increase, although only slightly indicating that some other pathway could be responsible for alternative IgG sialylation. Considering these recent findings, we wondered if GGA2 manipulation would result in the change of IgG sialylation. We found that neither upregulation nor downregulation of GGA2 expression resulted in the change of IgG sialylation. If IgG sialylation indeed occurs in the plasma, our HEK293-F expression system is not a good model for this type of research. A follow-up study should be performed *in vivo,* using mouse models, but verification of that approach would exceed the scope of this doctoral thesis. Similar reasoning led us to SPPL3 (Signal Peptide Peptidase Like 3) protease, also a GWAS hit for IgG sialylation[250,251]. A known substrate of SPPL3 is MGAT5 (Alpha-1,6-Mannosylglycoprotein 6-Beta-*N*-Acetylglucosaminyltransferase), a GT responsible for the biosynthesis of complex, branched glycans, usually not present on IgG[252]. However, beside MGAT5, SPLL3 can also cleave B4GALT1 as demonstrated by Voss *et al*. in 2014[250]. SPPL3 KO resulted in reduced B4GALT1 secretion and consequent intracellular accumulation of this GT. Considering that SPPL3 can cleave different GTs, we wondered if ST6GAL1 could also act as its substrate. Targeted downregulation of SPPL3 in the HEK293-F system increased all complex glycans on IgG, a process known as hyperglycosylation. Decreased expression of SPPL3 probably resulted in the accumulation of GTs on the Golgi membranes, which in turn increased the ratio of complex glycans on IgG. A significant increase in ratio was observed for IgG galactosylation, confirming that SPPL3 can cleave B4GALT1 from Golgi membranes. More importantly, we observed an increase of IgG sialylation, demonstrating for the first time that ST6GAL1 is indeed a substrate for SPPL3 in our model system. Moreover, we functionally validated the role of SPPL3 in the process of IgG sialylation. At the same time, it is important to consider a possible antagonistic role of the SPPL3 protease. If sialylation occurs in the plasma, targeted downregulation of SPPL3 could result in a decrease of IgG sialylation because this intervention would reduce the secretion of ST6GAL1. To test this hypothesis, a follow-up study should be performed in *in vivo* models, which would exceed the scope of this thesis. If a consensus that IgG sialylation is a plasma-localized process is reached, our HEK293-F

expression system should not be used as a model for IgG sialylation. The results obtained in this system up to now do point to a cell-specific pathway that regulates IgG sialylation, which does not exclude the possibility of another, plasma-specific mechanism. The success of our novel HEK293-F expression system in the functional validation of SPPL3 protease verified this approach as suitable for further validation of candidate genes with a potential role in IgG glycosylation.

IgG glycome undergoes extensive alterations during the human aging process[11,27,30,49–52]. These alterations are accompanied by a chronic, low-grade inflammation, known as inflammaging. Inflammaging is characterized by an increase in various inflammatory markers, conferring a significant susceptibility to chronic morbidity[54–56,253]. Age-related changes of IgG glycome composition are reminiscent of the ones occurring in various inflammatory and autoimmune conditions, pointing toward an intimate link between inflammaging and IgG glycosylation[53,133,254,255]. Still, the mediators of this molecular link remain largely unknown. However, a significant advance was made in the field of female aging. Women around the age of 45–60 years suffer abrupt age-related changes of the IgG glycome composition[27,47,58,65,133]. This time frame coincides with the perimenopausal period marked by a drop in estrogen production. The causal relationship between estrogen and IgG glycosylation was revealed in the study by Ercan *et al.*[65]. First, they demonstrated that menopause associates with an increase of agalactosylated IgG glycans, particularly with the increase of fucosylated non-bisected agalactosylated glycoform in two independent cohorts. Then they obtained the blood samples from 58 postmenopausal women randomized to receive placebo, conjugated estrogens or raloxifene (an estrogen-receptor binding substance that mimics estrogen signaling). Both conjugated estrogen and raloxifene therapy resulted in a decrease of agalactosylated IgG glycans. To confirm these findings, they obtained samples from 21 healthy premenopausal women treated with gonadotropin-releasing hormone agonist therapy to lower their gonadal steroid production to postmenopausal levels. This intervention resulted in an increase of agalactosylated IgG glycans, the same phenotype observed in women entering the menopause naturally. The women were then randomized to receive either transdermal estradiol or placebo. The application of transdermal estradiol reversed the effect of menopause induction on IgG glycosylation, further validating their previous findings. Recently, a study by Deriš *et al.*[64] offered novel insights into the menopause-related changes of IgG glycosylation on multiple samples from 1940 females from the TwinsUK registry. Beside the increase of agalactosylated glycans, they observed elevated levels of bisected structures and a decrease in IgG

digalactosylation and monosialylation, all considered hallmarks of inflammation. Several years earlier Jurić *et al*.[66] also demonstrated that the biological age measured by the glycan clock accelerates in menopausal women, an effect that can be ameliorated by estrogen supplementation.

All the studies described above established the role of estrogen as the mediator of menopause-related changes of IgG glycosylation. However, how do the estrogen signaling pathways regulate IgG glycosylation remains an open question. In an attempt to address this, we utilized our novel HEK293-F expression system. To prioritize the candidate genes potentially involved in this process, we analyzed the effects of estrogen on GWAS hits for IgG glycosylation using the Signaling Pathways Project (SPP) knowledgebase. An association was observed for B4GALT1, a GT that catalyzes the addition of galactose to the growing IgG glycan. As the role of B4GALT1 in IgG glycosylation is well established and further confirmed in our HEK293-F system, we focused on other significantly associated loci with the currently unknown role in this process: *RUNX1* (Runt-related transcription factor 1), *RUNX3* (Runt-related transcription factor 3), *SPINK4* (Serine Peptidase Inhibitor Kazal Type 4) and *ELL2* (Elongation Factor For RNA Polymerase II 2). *RUNX1*, *RUNX3* and *SPINK4* are GWAS hits for IgG galactosylation, while *ELL2* associates mainly with IgG sialylation[10,130,131,256]. Ratios of galactose and sialic acid on IgG are most significantly affected by estrogen and are decreased in menopausal women so we ought to confirm these findings in our HEK293-F system[64]. We manipulated all four loci in dCas9-VPR and dCas9-KRAB monoclonal cell lines and observed a major effect on IgG galactosylation following the upregulation of RUNX3 and SPINK4. The role of RUNX1 and ELL2 in IgG glycosylation could not be confirmed in our HEK293-F model system. Targeted upregulation of RUNX3 resulted in an increase of agalactosylated IgG glycans, the same phenotype observed in menopausal women. This was accompanied by a decrease of both mono- and digalactosylated IgG glycans. The first report of RUNX3 as a potential mediator of IgG glycosylation came from the GWA study by Wahl *et al*[131]. They associated this transcription factor with decreased IgG galactosylation, the same phenotype we observed in our HEK293-F system after RUNX3 upregulation. In the GWA study by Klarić *et al*.[10] *RUNX3* was associated with the glycosyltransferase MGAT3 but we could not confirm this result in our HEK293-F system as no change in IgG bisection was observed. Considering that RUNX3 acts as a transcription factor, we hypothesized that *B4GALT1* could be a target for RUNX3 transcriptional repression[257]. To test this, we analyzed B4GALT1 expression after RUNX3 upregulation, but no change in B4GALT1 mRNA levels was observed, disproving our

hypothesis. Currently, the exact mechanism of the RUNX3-mediated regulation of IgG galactosylation remains unknown, but many different studies in the last two decades demonstrated that RUNX3 acts downstream of estrogen[258–264]. This positioned the RUNX3 protein as a potential target for pharmacological interventions. RUNX3 inhibition could have beneficial, anti-inflammatory effects on the phenotype with only a subset of molecular consequences of estrogen therapy. The same is true for the SPINK4 peptidase as its upregulation in our HEK293-F system resulted in a decrease of galactosylated glycans. Interestingly, the observed change came only from the monogalactosylated IgG glycans with digalactosylated and agalactosylated structures remaining unchanged. Following this intervention, we discovered that the targeted promoter of the *SPINK4* gene is located in the same region as the promoter of *B4GALT1*[265]. We wondered if the VPR-targeting of the *SPINK4* promoter inadvertently affected the expression of B4GALT1 even though the expected effect of this off-target would be the upregulation of B4GALT1 and consequent increase of galactosylated IgG glycoforms. Our intervention did not unintentionally change the expression of B4GALT1, validating the role of SPINK4 in IgG glycosylation. Our results were confirmed in a recent study by Klarić *et al*.[10] where the authors reported the strongest correlation between *SPINK4* and *B4GALT1* loci. Furthermore, the observed effect on IgG galactosylation for the upregulation of both RUNX3 and SPINK4 loci mimicked the pro-inflammatory IgG glycan phenotype of menopausal women. Taken together, the results obtained from our HEK293-F system point to a novel mechanism where RUNX3 and SPINK4 act as mediators of the estrogen signaling pathway that regulates the glycome composition of IgG antibodies.

This study also served to demonstrate that age-related alternative IgG glycosylation operates under the control of distinct genetic factors, at least in women entering the menopause. However, the exact magnitude of genetic and environmental influences on age-related changes of IgG glycome composition was never analyzed. To tackle this question, we performed a heritability analysis on the glycan clock measured from blood samples of female twins enrolled in the TwinsUK registry. Previous heritability studies revealed that the impact of genetic factors on IgG glycosylation varies depending on the analyzed IgG glycoform[30,75,77]. Nevertheless, most of the IgG glycans turned out to be at least half heritable with only a few having a low genetic contribution. This strong genetic control of IgG glycosylation as a complex trait was later confirmed in numerous GWA studies revealing intricate gene networks regulating this process[10,129–132]. GWAS were recently followed by the *in vitro* validation of IgG glycosylation-associated loci in our HEK293-F expression system. As the novel loci emerged, it became

obvious that some of them were associated with age-related changes of IgG glycome composition. To gain a better understanding of the genetics of IgG glycosylation in aging, we decided to analyze heritability of three IgG glycans strongly correlated with the age of an individual, commonly referred to as the glycan clock. The glycan clock was established in 2014 as a potent marker of biological age based exclusively on IgG glycans[51]. In our study, the glycan clock served as a model of age-related changes in IgG glycosylation, upon which we can analyze the influence of genetic and environmental factors. We calculated the glycan clock from blood samples of 479 monozygotic (MZ) and 1193 dizygotic (DZ) female twin pairs enrolled in the cross-sectional cohort and blood samples of 549 MZ and 1201 DZ female twin pairs enrolled in the longitudinal cohort with three defined time points. To estimate the contribution of genetic and environmental factors on the phenotypic variation of the glycan clock, we applied a classical twin design[266]. This type of modeling enabled us to decompose the observed variance into three latent sources of variation: additive genetic (A), shared environmental (C) and unique environmental variance (E). The contribution of additive genetic factors to the phenotypic variation of the glycan clock averaged at 39% for both cohorts. Despite the substantial shift of IgG glycome composition in the three time points of the longitudinal cohort, the heritability estimates remained stable. Interestingly, the majority of the variation came from shared environmental factors (~45%). We hypothesized that the shared environmental factors in this heritability study represent the contribution of chronological age as a type of a cellular environment determined by the genetic makeup of an individual. To test this hypothesis, we replicated the heritability study with the age of the participants included as a covariate. This intervention deflated the shared environmental estimates to the point where the full ACE model became unfitting for the longitudinal cohort. When the AE model was applied, it revealed a high heritability of the glycan clock, averaging at 71% for both cohorts. The contribution of the unique environmental factors increased from an average of 17% to an average of 27%. These results demonstrated that the change of IgG glycome composition observed in aging is strongly influenced by the genetic predisposition of an individual. The high heritability of the glycan clock also emphasized the need to develop novel molecular tools that would validate the genetic loci potentially involved in age-related but also disease-related changes of IgG glycosylation. Moreover, the fairly high contribution of the unique environmental factors clarified the potential of the glycan clock to be reversed. Various recent studies have shown that the glycan clock can be turned by simple lifestyle choices like exercising and dieting[59–61,64]. The first report linking exercise and IgG glycosylation came from a study by Tijardović *et al.*[60]. They examined the effect of intensive exercise on IgG glycans to

expand the understanding of the link between physical activity and inflammation. Remarkably, increased physical activity resulted in an increase of digalactosylated and monosialylated IgG glycoforms with anti-inflammatory properties. The ratio of agalactosylated IgG glycans decreased mimicking a phenotype of young individuals. A year later, Greto *et al.*[61] demonstrated that a low-calorie diet, bariatric surgery and a general decrease in body mass index all led to the reduction of the biological age measured with the glycan clock. In the light of these studies, we argue that the contribution of unique environmental factors to age-related IgG glycosylation changes equals the contribution of lifestyle to the variation of the glycan clock. Therefore, the high contribution of unique lifestyle choices supports the notion that the glycan clock can be "rewound" by the rejuvenating effect of the lifestyle interventions listed above. The responsiveness of the glycan clock to a healthier lifestyle positioned this biomarker as the most promising predictor of biological age in modern personalized medicine. Moreover, the potential of the glycan clock to integrate genetic and environmental cues highlighted the plasticity of the IgG glycome in response to various external and internal stimuli. However, the high contribution of the genetic factors to the glycan clock variation was the most notable discovery of this heritability study in the light of our research on the regulation of IgG glycosylation. This observation served as an affirmation to continue our efforts to discover the mechanisms behind alternative IgG glycosylation associated not only with various pathological states but also with the process of aging.

To summarize, we developed a simple and efficient *in vitro* cell system (HEK293-F) to study the mechanisms regulating IgG glycosylation, based on CRISPRa/CRISPRi technology and FreeStyle™ 293-F cell lines. Utilizing the HEK293-F system, we were the first to functionally validate the role of the GWAS candidate gene SPPL3 in IgG sialylation. Then we exploited the newly established system to map the estrogen signaling pathway responsible for altered IgG glycosylation in women entering the menopause. By targeted activation of two estrogen-regulated and IgG glycosylation-associated loci *RUNX3* and *SPINK4*, we imitated the pro-inflammatory IgG glycome composition observed in menopausal women. This was the first report of a potential signaling pathway that links estrogen and IgG glycosylation, opening a possibility for the discovery of novel therapeutics acting downstream of this hormone. The surprising success of the HEK293-F system to robustly validate the genetic loci associated not only with pathologies but also with aging raised the question of the magnitude of genetic influences regulating alternative IgG glycosylation in the aging processes. The heritability study on the glycan clock data demonstrated that environmental factors determine the

phenotypic variation of the glycan clock to some extent, but the contribution of genetic factors turned out to be fairly high. This served as a proof that alternative IgG glycosylation observed in aging is under relatively tight genetic control emphasizing the need to develop novel technologies able to dissect the observed genetic influences. Our HEK293-F system is a pioneering technology for this type of studies, as it enables functional validation of different loci associated with IgG glycosylation. More than a decade of GWAS left us with dozens of novel loci waiting to be validated in our HEK293-F *in vitro* system. Still, we have to admit certain shortcomings of this technology. First, the selected model FreeStyle™ 293-F cell line is a cell type unrelated to B cells that secrete native IgG[94,137]. Clearly, FreeStyle™ 293-F cells have a completely different regulatory profile than B lymphocytes, though some core mechanisms are most likely shared between all cell types. Protein glycosylation certainly fits that paradigm as we confirmed that FreeStyle™ 293-F cells secrete recombinant IgG with a surprisingly human-like glycosylation pattern on the Fc region. Still, our HEK293-F system would not be a suitable model for the functional validation of B cell-specific mechanisms of IgG glycosylation. In the future, we hope to translate the developed technology to LCL cell lines that naturally secrete IgG and are derived from B lymphocytes. Currently, no B cell-derived line could match the simplicity of the HEK293-F system and the power of this technology to allow for a close to high-throughput approach in the validation of novel loci implicated in IgG glycosylation. But as the methodology in glycan analysis advances and new techniques for gene transfer emerge, the translation of the HEK293-F technology to a better model cell line is becoming an achievable goal[128]. However, as was mentioned previously in relation to IgG sialylation, no *in vitro* system could integrate the multicellular effect on IgG glycosylation when the organism is analyzed as a whole. Currently, the best *in vivo* models of human IgG glycosylation are different murine strains, which also come with various limitations, mainly related to the non-human-like glycosylation patterns on IgG[134]. Nonetheless, the research of IgG glycosylation in all these different models is imperative as the variations in this process have an immense impact on the function of IgG and potentially act as the drivers of various human pathologies. The knowledge we obtained using the novel HEK293-F system serves as a foundation for future studies of the mechanisms regulating IgG glycosylation.

4. CONCLUSION

1. The HEK293-F *in vitro* system for the functional validation of genes associated with IgG glycosylation was successfully established.

- Transposon-based piggyBac constructs, upgraded with insulators and adapted for the Golden Gate modular toolbox, enabled genomic integration of the dCas9-VPR (CRISPRa) and dCas9-KRAB (CRISPRi) expression cassettes in the FreeStyle™ 293-F cells.
- The obtained monoclonal and polyclonal FreeStyle™ 293-F cell lines demonstrated stable and long-term expression of dCas9-VPR and dCas9-KRAB fusion proteins.
- Transient co-transfection of CRISPRa/CRISPRi FreeStyle™ 293-F cell lines with recombinant IgG/gRNA expression plasmid and three small plasmids for cell cycle arrest resulted in IgG yield sufficient for glycan analysis. The glycosylation profile of the Fc region exhibited a human-like pattern: highly fucosylated, moderately galactosylated with low levels of bisection and sialylation.
- Targeted manipulation of four main GTs with a known role in IgG glycosylation resulted in the expected change of the IgG glycan phenotype: FUT8 downregulation – decreased IgG fucosylation; MGAT3 upregulation – increased IgG bisection; B4GALT upregulation/downregulation – increased/decreased IgG galactosylation; ST6GAL1 upregulation – increased IgG sialylation.
- Targeting the GTs in both monoclonal and polyclonal cell-based HEK293-F system resulted in the same IgG glycan phenotype, although to a lesser extent in the polyclonal cells.
- Magnetic-activated cell sorting after transient transfection of the monoclonal and polyclonal cell lines resulted in no observed enrichment of IgG/gRNA positive cells, thus validating the selection free-approach

2. The HEK293-F *in vitro* system proved to be a powerful tool for functional validation of genes associated with IgG glycosylation

- Targeted downregulation of the *SPPL3* locus resulted in an increase of all complex glycan structures on IgG, including IgG sialylation. This was the first *in vitro* report of the involvement of the SPPL3 enzyme in the process of IgG sialylation.

- Targeted upregulation of two estrogen regulated loci, *RUNX3* and *SPINK4*, resulted in the decrease of galactosylated glycans on IgG, a phenotype that characterizes menopause and aging in general. This was the first report of RUNX3 and SPINK4 as mediators of estrogen signaling pathway that triggers the change in IgG glycosylation observed in menopause.

3. The heritability of the glycan clock marker of biological aging proved to be high, averaging around 71% when corrected for the age of the individuals. The high heritability of the glycan clock demonstrated that the age-related changes of IgG glycosylation operate under tight genetic control. This served to demonstrate the need to develop novel technologies, like the HEK293-F system, able to dissect the mechanism regulating alternative IgG glycosylation associated not only with pathologies but also with aging.

5. REFERENCES

1. Schjoldager, K. T., Narimatsu, Y., Joshi, H. J. & Clausen, H. Global view of human protein glycosylation pathways and functions. *Nat. Rev. Mol. Cell Biol.* **21**, 729–749 (2020).

2. Reily, C., Stewart, T. J., Renfrow, M. B. & Novak, J. Glycosylation in health and disease. *Nat. Rev. Nephrol.* **15**, 346–366 (2019).

3. Moremen, K. W., Tiemeyer, M. & Nairn, A. V. Vertebrate protein glycosylation: diversity, synthesis and function. *Nat. Rev. Mol. Cell Biol.* **13**, 448–462 (2012).

4. Taniguchi, N. *et al. Handbook of glycosyltransferases and related genes, second edition*. *Handbook of Glycosyltransferases and Related Genes, Second Edition* **1**–**2**, (Springer Tokyo, 2014).

5. Dalziel, M., Crispin, M., Scanlan, C. N., Zitzmann, N. & Dwek, R. A. Emerging principles for the therapeutic exploitation of glycosylation. *Science (80-.).* **343**, 1235681 (2014).

6. Freeze, H. H. Genetic defects in the human glycome. *Nat. Rev. Genet.* **7**, 537–551 (2006).

7. Abbott, K. L. *et al.* Focused glycomic analysis of the N-linked glycan biosynthetic pathway in ovarian cancer. *Proteomics* **8**, 3210–3220 (2008).

8. Nairn, A. V. *et al.* Regulation of glycan structures in animal tissues: Transcript profiling of glycan-related genes. *J. Biol. Chem.* **283**, 17298–17313 (2008).

9. Lauc, G., Vojta, A. & Zoldoš, V. Epigenetic regulation of glycosylation is the quantum mechanics of biology. *Biochim. Biophys. Acta - Gen. Subj.* **1840**, 65–70 (2014).

10. Klarić, L. *et al.* Glycosylation of immunoglobulin G is regulated by a large network of genes pleiotropic with inflammatory diseases. *Sci. Adv.* **6**, eaax0301 (2020).

11. Gudelj, I., Lauc, G. & Pezer, M. Immunoglobulin G glycosylation in aging and diseases. *Cell. Immunol.* **333**, 65–79 (2018).

12. Alavi, A. & Axford, J. S. Sweet and sour: The impact of sugars on disease. *Rheumatology* **47**, 760–770 (2008).

13. Alter, G., Ottenhoff, T. H. M. & Joosten, S. A. Antibody glycosylation in inflammation, disease and vaccination. *Semin. Immunol.* **39**, 102–110 (2018).

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

15. Pinho, S. S. & Reis, C. A. Glycosylation in cancer: Mechanisms and clinical implications. *Nat. Rev. Cancer* **15**, 540–555 (2015).

16. Stowell, S. R., Ju, T. & Cummings, R. D. Protein glycosylation in cancer. *Annu. Rev. Pathol. Mech. Dis.* **2**, 473–510 (2015).

17. Lee, H. S., Qi, Y. & Im, W. Effects of N-glycosylation on protein conformation and dynamics: Protein Data Bank analysis and molecular dynamics simulation study. *Sci. Rep.* **5**, 8926 (2015).

18. Goettig, P. Effects of glycosylation on the enzymatic activity and mechanisms of proteases. *Int. J. Mol. Sci.* **17**, 1969 (2016).

19. Mikolajczyk, K., Kaczmarek, R. & Czerwinski, M. How glycosylation affects glycosylation: The role of N-glycans in glycosyltransferase activity. *Glycobiology* **30**, 941–969 (2020).

20. Shental-Bechor, D. & Levy, Y. Effect of glycosylation on protein folding: A close look at thermodynamic stabilization. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 8256–8261 (2008).

21. Schwarz, F. & Aebi, M. Mechanisms and principles of N-linked protein glycosylation. *Curr. Opin. Struct. Biol.* **21**, 576–582 (2011).

22. Lv, J. *et al. Knowledge Domain and Emerging Trends of IgG Glycosylation: A Bibliometric Study Based on CiteSpace*. https://doi.org/10.21203/rs.3.rs-74133/v1 (2020).

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

24. Dekkers, G. *et al.* Decoding the human immunoglobulin G-glycan repertoire reveals a spectrum of Fc-receptor- and complement-mediated-effector activities. *Front. Immunol.* **8**, 877 (2017).

25. Yagi, H., Yanaka, S. & Kato, K. Structure and dynamics of immunoglobulin G

glycoproteins. *Adv. Exp. Med. Biol.* **1104**, 219–235 (2018).

26. Lux, A., Aschermann, S., Biburger, M. & Nimmerjahn, F. The pro and anti-inflammatory activities of immunoglobulin G. *Ann. Rheum. Dis.* **69**, i92 LP-i96 (2010).

27. Baković, M. P. *et al.* High-throughput IgG Fc N-glycosylation profiling by mass spectrometry of glycopeptides. *J. Proteome Res.* **12**, 821–831 (2013).

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

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

30. Pučić, M. *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* **10**, M111.010090 (2011).

31. Bondt, A. *et al.* Immunoglobulin G (IgG) fab glycosylation analysis using a new mass spectrometric high-throughput profiling method reveals pregnancy-associated changes. *Mol. Cell. Proteomics* **13**, 3029–3039 (2014).

32. Holland, M. *et al.* Differential glycosylation of polyclonal IgG, IgG-Fc and IgG-Fab isolated from the sera of patients with ANCA-associated systemic vasculitis. *Biochim. Biophys. Acta - Gen. Subj.* **1760**, 669–677 (2006).

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

34. Wada, R., Matsui, M. & Kawasaki, N. Influence of N-glycosylation on effector functions and thermal stability of glycoengineered IgG1 monoclonal antibody with homogeneous glycoforms. *MAbs* **11**, 350–372 (2019).

35. Zheng, K., Bantog, C. & Bayer, R. The impact of glycosylation on monoclonal antibody conformation and stability. *MAbs* **3**, 568–576 (2011).

36. Davies, J. *et al.* Expression of GnTIII in a recombinant anti-CD20 CHO production cell line: Expression of antibodies with altered glycoforms leads to an increase in ADCC

through higher affinity for FcγRIII. *Biotechnol. Bioeng.* **74**, 288–294 (2001).

37. Masuda, K. *et al.* Enhanced binding affinity for FcγRIIIa of fucose-negative antibody is sufficient to induce maximal antibody-dependent cellular cytotoxicity. *Mol. Immunol.* **44**, 3122–3131 (2007).

38. Shields, R. L. *et al.* Lack of fucose on human IgG1 N-linked oligosaccharide improves binding to human FcγRIII and antibody-dependent cellular toxicity. *J. Biol. Chem.* **277**, 26733–26740 (2002).

39. Sondermann, P., Pincetic, A., Maamary, J., Lammens, K. & Ravetch, J. V. General mechanism for modulating immunoglobulin effector function. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 9868–9872 (2013).

40. Pincetic, A. *et al.* Type-I and Type-II Fc receptors regulate innate and adaptive immunity. *Nat. Immunol.* **15**, 707–716 (2014).

41. van Osch, T. L. J. *et al.* Fc Galactosylation Promotes Hexamerization of Human IgG1, Leading to Enhanced Classical Complement Activation. *J. Immunol.* **207**, 1545–1554 (2021).

42. 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).

43. Arnold, J. N., Dwek, R. A., Rudd, P. M. & Sim, R. B. Mannan binding lectin and its interaction with immunoglobulins in health and in disease. *Immunol. Lett.* **106**, 103–110 (2006).

44. Dekkers, G., Rispens, T. & Vidarsson, G. Novel Concepts of Altered Immunoglobulin G Galactosylation in Autoimmune Diseases. *Front. Immunol.* **9**, 553 (2018).

45. 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).

46. Wei, B. *et al.* Fc galactosylation follows consecutive reaction kinetics and enhances immunoglobulin G hexamerization for complement activation. *MAbs* **13**, 1893427 (2021).

47. Štambuk, J. *et al.* Global variability of the human IgG glycome. *Aging (Albany. NY).* **12**, 15222–15259 (2020).

48. Parekh, R. B. *et al.* Association of rheumatoid arthritis and primary osteoarthritis with changes in the glycosylation pattern of total serum IgG. *Nature* **316**, 452–457 (1985).

49. Kavur, M. M., Lauc, G. & Pezer, M. Systems Glycobiology: Immunoglobulin G Glycans as Biomarkers and Functional Effectors in Aging and Diseases. in *Comprehensive Glycoscience* (eds. Barchi Jr., J. J., Vidal, S., Williams, S. J., Narain, R. & Taniguchi, N.) 439–478 (Elsevier, 2021). doi:10.1016/B978-0-12-819475-1.00086-9.

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

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

52. Vanhooren, V. *et al.* N-glycomic changes in serum proteins during human aging. *Rejuvenation Res.* **10**, 521–531 (2007).

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

54. Franceschi, C. *et al.* Inflamm-aging. An evolutionary perspective on immunosenescence. *Ann. N. Y. Acad. Sci.* **908**, 244–254 (2000).

55. De Martinis, M., Franceschi, C., Monti, D. & Ginaldi, L. Inflamm-ageing and lifelong antigenic load as major determinants of ageing rate and longevity. *FEBS Lett.* **579**, 2035–2039 (2005).

56. Monti, D., Ostan, R., Borelli, V., Castellani, G. & Franceschi, C. Inflammaging and human longevity in the omics era. *Mech. Ageing Dev.* **165**, 129–138 (2017).

57. Franceschi, C., Garagnani, P., Parini, P., Giuliani, C. & Santoro, A. Inflammaging: a new immune–metabolic viewpoint for age-related diseases. *Nat. Rev. Endocrinol.* **14**, 576–590 (2018).

58. Yu, X. *et al.* Profiling IgG N-glycans as potential biomarker of chronological and

biological ages: A community-based study in a Han Chinese population. *Med. (United States)* **95**, e4112 (2016).

59. Peng, J. *et al.* Supplementation with the Sialic Acid Precursor N-acetyl-D-Mannosamine Breaks the Link Between Obesity and Hypertension. *Circulation* **140**, 2005–2018 (2019).

60. Tijardović, M. *et al.* Intense Physical Exercise Induces an Anti-inflammatory Change in IgG N-Glycosylation Profile. *Front. Physiol.* **10**, 1522 (2019).

61. Greto, V. L. *et al.* Extensive weight loss reduces glycan age by altering IgG N-glycosylation. *Int. J. Obes.* **45**, 1521–1531 (2021).

62. Ruhaak, L. R. *et al.* Decreased levels of bisecting GLcNAc glycoforms of IgG are associated with human longevity. *PLoS One* **5**, e12566 (2010).

63. Chen, G. *et al.* Human IgG Fc-glycosylation profiling reveals associations with age, sex, female sex hormones and thyroid cancer. *J. Proteomics* **75**, 2824–2834 (2012).

64. Deriš, H. *et al.* Immunoglobulin G glycome composition in transition from premenopause to postmenopause. *iScience* **25**, 103897 (2022).

65. Ercan, A. *et al.* Estrogens regulate glycosylation of IgG in women and men. *JCI Insight* **2**, e89703 (2017).

66. Jurić, J. *et al.* Effects of estradiol on biological age measured using the glycan age index. *Aging (Albany. NY).* **12**, 19756–19765 (2020).

67. Engdahl, C. *et al.* Estrogen induces St6gal1 expression and increases IgG sialylation in mice and patients with rheumatoid arthritis: A potential explanation for the increased risk of rheumatoid arthritis in postmenopausal women. *Arthritis Res. Ther.* **20**, 84 (2018).

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

69. Bondt, A. *et al.* ACPA IgG galactosylation associates with disease activity in pregnant patients with rheumatoid arthritis. *Ann. Rheum. Dis.* **77**, 1130–1136 (2018).

70. Bondt, A. *et al.* Association between galactosylation of immunoglobulin G and

improvement of rheumatoid arthritis during pregnancy is independent of sialylation. *J. Proteome Res.* **12**, 4522–4531 (2013).

71. Van de Geijn, F. E. *et al.* Immunoglobulin G galactosylation and sialylation are associated with pregnancy-induced improvement of rheumatoid arthritis and the postpartum flare: Results from a large prospective cohort study. *Arthritis Res. Ther.* **11**, R193 (2009).

72. Russell, A., Adua, E., Ugrina, I., Laws, S. & Wang, W. Unravelling immunoglobulin G Fc N-glycosylation: A dynamic marker potentiating predictive, preventive and personalised medicine. *Int. J. Mol. Sci.* **19**, 390 (2018).

73. Trombetta, E. S. The contribution of N-glycans and their processing in the endoplasmic reticulum to glycoprotein biosynthesis. *Glycobiology* **13**, 77–91 (2003).

74. Visscher, P. M., Hill, W. G. & Wray, N. R. Heritability in the genomics era - Concepts and misconceptions. *Nat. Rev. Genet.* **9**, 255–266 (2008).

75. Menni, C. *et al.* Glycosylation of immunoglobulin G: Role of genetic and epigenetic influences. *PLoS One* **8**, 6–13 (2013).

76. Knežević, A. *et al.* Variability, heritability and environmental determinants of human plasma n-glycome. *J. Proteome Res.* **8**, 694–701 (2009).

77. Zaytseva, O. O. *et al.* Heritability of Human Plasma N-Glycome. *J. Proteome Res.* **19**, 85–91 (2020).

78. Benedetti, E. *et al.* Network inference from glycoproteomics data reveals new reactions in the IgG glycosylation pathway. *Nat. Commun.* **8**, 1483 (2017).

79. Boruah, B. M. *et al.* Characterizing human α-1,6-fucosyltransferase (FUT8) substrate specificity and structural similarities with related fucosyltransferases. *J. Biol. Chem.* **295**, 11727–11745 (2020).

80. Wang, X. *et al.* Dysregulation of TGF-β1 receptor activation leads to abnormal lung development and emphysema-like phenotype in core fucose-deficient mice. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 15791–15796 (2005).

81. Yamane-Ohnuki, N. *et al.* Establishment of FUT8 knockout Chinese hamster ovary cells: An ideal host cell line for producing completely defucosylated antibodies with

enhanced antibody-dependent cellular cytotoxicity. *Biotechnol. Bioeng.* **87**, 614–622 (2004).

82. Subramaniam, J. M., Whiteside, G., McKeage, K. & Croxtall, J. C. Mogamulizumab. *Drugs* **72**, 1293–1298 (2012).

83. Mössner, E. *et al.* Increasing the efficacy of CD20 antibody therapy through the engineering of a new type II anti-CD20 antibody with enhanced direct and immune effector cell - mediated B-cell cytotoxicity. *Blood* **115**, 4393–4402 (2010).

84. Park, J. H. *et al.* L-Fucose treatment of FUT8-CDG. *Mol. Genet. Metab. Reports* **25**, 100680 (2020).

85. Ng, B. G. *et al.* Expanding the molecular and clinical phenotypes of FUT8-CDG. *J. Inherit. Metab. Dis.* **43**, 1871–879 (2020).

86. Ng, B. G. *et al.* Biallelic Mutations in FUT8 Cause a Congenital Disorder of Glycosylation with Defective Fucosylation. *Am. J. Hum. Genet.* **102**, 188–195 (2018).

87. Pereira, N. A., Chan, K. F., Lin, P. C. & Song, Z. The "less-is-more" in therapeutic antibodies: Afucosylated anti-cancer antibodies with enhanced antibody-dependent cellular cytotoxicity. *MAbs* **10**, 693–711 (2018).

88. Liao, C. *et al.* FUT8 and protein core fucosylation in tumours: From diagnosis to treatment. *J. Cancer* **12**, 4109–4120 (2021).

89. García-García, A. *et al.* Structural basis for substrate specificity and catalysis of α1,6-fucosyltransferase. *Nat. Commun.* **11**, 973 (2020).

90. Nakano, M. *et al.* Bisecting GlcNAc is a general suppressor of terminal modification of N-glycan. *Mol. Cell. Proteomics* **18**, 2044–2057 (2019).

91. Nagae, M. *et al.* Atomic visualization of a flipped-back conformation of bisected glycans bound to specific lectins. *Sci. Rep.* **6**, 22973 (2016).

92. Kapur, R. *et al.* A prominent lack of IgG1-Fc fucosylation of platelet alloantibodies in pregnancy. *Blood* **123**, 471–480 (2014).

93. Khoder-Agha, F. *et al.* Assembly of B4GALT1/ST6GAL1 heteromers in the Golgi membranes involves lateral interactions via highly charged surface domains. *J. Biol. Chem.* **294**, 14383–14393 (2019).

94. Dekkers, G. *et al.* Multi-level glyco-engineering techniques to generate IgG with defined Fc-glycans. *Sci. Rep.* **6**, 36964 (2016).

95. Nguyen, N. T. B. *et al.* Multiplexed engineering glycosyltransferase genes in CHO cells via targeted integration for producing antibodies with diverse complex-type N-glycans. *Sci. Rep.* **11**, 12969 (2021).

96. Yang, Z. *et al.* Engineered CHO cells for production of diverse, homogeneous glycoproteins. *Nat. Biotechnol.* **33**, 842–844 (2015).

97. Sha, J. *et al.* The B-Cell-Specific Ablation of B4GALT1 Reduces Cancer Formation and Reverses the Changes in Serum IgG Glycans during the Induction of Mouse Hepatocellular Carcinoma. *Cancers (Basel).* **14**, 1333 (2022).

98. Keusch, J., Lydyard, P. M., Berger, E. G. & Delves, P. J. B lymphocyte galactosyltransferase protein levels in normal individuals and in patients with rheumatoid arthritis. *Glycoconj. J.* **15**, 1093–1097 (1998).

99. Axford, J. S. *et al.* Reduced B-Cell Galactosyltransferase Activity in Rheumatoid Arthritis. *Lancet* **330**, 1486–1488 (1987).

100. Alavi, A. & Axford, J. Evaluation of β1,4-galactosyltransferase in rheumatoid arthritis and its role in the glycosylation network associated with this disease. *Glycoconj. J.* **12**, 206–210 (1995).

101. Nishiura, T. *et al.* Carbohydrate analysis of immunoglobulin G myeloma proteins by lectin and high performance liquid chromatography: role of glycosyltransferases in the structures. *Cancer Res.* **50**, 5345–5350 (1990).

102. Delves, P. J. *et al.* Polymorphism and expression of the galactosyltransferase-associated protein kinase gene in normal individuals and galactosylation-defective rheumatoid arthritis patients. *Arthritis Rheum.* **33**, 1655–1664 (1990).

103. Jeddi, P. A. *et al.* Agalactosyl IgG and beta-1,4-galactosyltransferase gene expression in rheumatoid arthritis patients and in the arthritis-prone MRL lpr/lpr mouse. *Immunology* **87**, 654–659 (1996).

104. Catera, M. *et al.* Identification of novel plasma glycosylation-associated markers of aging. *Oncotarget* **7**, 7455–7468 (2016).

105. Hennet, T., Chui, D., Paulson, J. C. & Marth, J. D. Immune regulation by the ST6Gal sialyltransferase. *Proc. Natl. Acad. Sci. U. S. A.* **95**, 4504–4509 (1998).

106. Jones, M. B. *et al.* B-cell-independent sialylation of IgG. *Proc. Natl. Acad. Sci. U. S. A.* **113**, 7207–7212 (2016).

107. Oswald, D. M. *et al.* ST6Gal1 in plasma is dispensable for IgG sialylation. *Glycobiology* **32**, 803–813 (2022).

108. Werner, A. & Nimmerjahn, F. Response to Oswald et al.: ST6Gal1 in plasma is dispensable for IgG sialylation. *Glycobiology* **32**, 917–918 (2022).

109. Klasić, M. *et al.* DNA hypomethylation upregulates expression of the MGAT3 gene in HepG2 cells and leads to changes in N-glycosylation of secreted glycoproteins. *Sci. Rep.* **6**, 24363 (2016).

110. Anugraham, M. *et al.* Specific glycosylation of membrane proteins in epithelial ovarian cancer cell lines: glycan structures reflect gene expression and DNA methylation status. *Mol. Cell. proteomics* **13**, 2213–2232 (2014).

111. Kohler, R. S. *et al.* Epigenetic activation of MGAT3 and corresponding bisecting GlcNAc shortens the survival of cancer patients. *Oncotarget* **7**, 51674–51686 (2016).

112. Josipović, G. *et al.* Antagonistic and synergistic epigenetic modulation using orthologous CRISPR/dCas9-based modular system. *Nucleic Acids Res.* **47**, 9637–9657 (2019).

113. Klasić, M. *et al.* Promoter methylation of the MGAT3 and BACH2 genes correlates with the composition of the immunoglobulin G glycome in inflammatory bowel disease. *Clin. Epigenetics* **10**, 75 (2018).

114. Zohora, F. T. & Mahal, L. K. Comprehensive Analysis of miRNA Regulation of MGAT3 Using the miRFluR Assay. *FASEB J.* **36**, (2022).

115. Bernardi, C., Soffientini, U., Piacente, F. & Tonetti, M. G. Effects of MicroRNAs on Fucosyltransferase 8 (FUT8) Expression in Hepatocarcinoma Cells. *PLoS One* **8**, 6–11 (2013).

116. Cheng, L. *et al.* Comprehensive N-glycan profiles of hepatocellular carcinoma reveal association of fucosylation with tumor progression and regulation of FUT8 by

microRNAs. *Oncotarget* **7**, 61199–61214 (2016).

117. Wang, M. *et al.* MiR-198 represses tumor growth and metastasis in colorectal cancer by targeting fucosyl transferase 8. *Sci. Rep.* **4**, 6145 (2014).

118. Han, Y. *et al.* MiR-9 inhibits the metastatic ability of hepatocellular carcinoma via targeting beta galactoside alpha-2,6-sialyltransferase 1. *J. Physiol. Biochem.* **74**, 491–501 (2018).

119. Liu, Y. X. *et al.* MIR-124-3p/B4GALT1 axis plays an important role in SOCS3-regulated growth and chemo-sensitivity of CML. *J. Hematol. Oncol.* **9**, 69 (2016).

120. Bush, W. S. & Moore, J. H. Chapter 11: Genome-Wide Association Studies. *PLoS Comput. Biol.* **8**, e1002822 (2012).

121. Uffelmann, E. *et al.* Genome-wide association studies. *Nat. Rev. Methods Prim.* **1**, 1–21 (2021).

122. Lander, E. S. *et al.* Initial sequencing and analysis of the human genome. *Nature* **409**, 860–921 (2001).

123. Belmont, J. W. *et al.* A haplotype map of the human genome. *Nature* **437**, 1299–1320 (2005).

124. Altshuler, D. L. *et al.* A map of human genome variation from population-scale sequencing. *Nature* **467**, 1061–1073 (2010).

125. The International HapMap 3 Consortium. Integrating common and rare genetic variation in diverse human populations: The International HapMap 3 Consortium. *Nature* **467**, 52–58 (2010).

126. De Haan, N. *et al.* Developments and perspectives in high-throughput protein glycomics: enabling the analysis of thousands of samples. *Glycobiology* **32**, 651–663 (2022).

127. Huffman, J. E. *et al.* Comparative performance of four methods for high-throughput glycosylation analysis of immunoglobulin G in genetic and epidemiological research. *Mol. Cell. Proteomics* **13**, 1598–1610 (2014).

128. Trbojević-Akmačić, I. *et al.* High-Throughput Glycomic Methods. *Chem. Rev.* **122**, 15865–15913 (2022).

129. Lauc, G. *et al.* Loci Associated with N-Glycosylation of Human Immunoglobulin G Show Pleiotropy with Autoimmune Diseases and Haematological Cancers. *PLoS Genet.* **9**, e1003225 (2013).

130. Shen, X. *et al.* Multivariate discovery and replication of five novel loci associated with Immunoglobulin G N-glycosylation. *Nat. Commun.* **8**, 447 (2017).

131. Wahl, A. *et al.* Genome-wide association study on immunoglobulin G glycosylation patterns. *Front. Immunol.* **9**, 277 (2018).

132. Shadrina, A. S. *et al.* Multivariate genome-wide analysis of immunoglobulin G N-glycosylation identifies new loci pleiotropic with immune function. *Hum. Mol. Genet.* **30**, 1259–1270 (2021).

133. Pezer, M. *Antibody Glycosylation*. (Springer Cham, 2021).

134. Krištić, J. *et al.* Profiling and genetic control of the murine immunoglobulin G glycome. *Nat. Chem. Biol.* **14**, 516–524 (2018).

135. Blomme, B. *et al.* Alterations of serum protein N-glycosylation in two mouse models of chronic liver disease are hepatocyte and not B cell driven. *Am. J. Physiol. - Gastrointest. Liver Physiol.* **300**, 833–842 (2011).

136. Maresch, D. & Altmann, F. Isotype-specific glycosylation analysis of mouse IgG by LC-MS. *Proteomics* **16**, 1321–1330 (2016).

137. Vink, T., Oudshoorn-Dickmann, M., Roza, M., Reitsma, J. J. & de Jong, R. N. A simple, robust and highly efficient transient expression system for producing antibodies. *Methods* **65**, 5–10 (2014).

138. Jinek, M. *et al.* A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science (80-.).* **337**, 816–821 (2012).

139. Dominquez, A. A., Lim, W. A. & Qi, L. S. Beyond editing: repurposing CRISPR–Cas9 for precision genome regulation and interrogation. *Nat. Rev. Mol. Cell Biol.* **17**, 5–15 (2016).

140. Tang, H. *et al.* Active-Site Models of Streptococcus pyogenes Cas9 in DNA Cleavage State. *Front. Mol. Biosci.* **8**, 653262 (2021).

141. Globyte, V., Lee, S. H., Bae, T., Kim, J. & Joo, C. CRISPR /Cas9 searches for a

protospacer adjacent motif by lateral diffusion. *EMBO J.* **38**, e99466 (2019).

142. Friedland, A. E. *et al.* Characterization of Staphylococcus aureus Cas9: A smaller Cas9 for all-in-one adeno-associated virus delivery and paired nickase applications. *Genome Biol.* **16**, 257 (2015).

143. Ran, F. A. *et al.* In vivo genome editing using Staphylococcus aureus Cas9. *Nature* **520**, 186–191 (2015).

144. Tabebordbar, M. *et al.* In vivo gene editing in dystrophic mouse muscle and muscle stem cells. *Science (80-.).* **351**, 407–411 (2016).

145. Nelson, C. E. *et al.* In vivo genome editing improves muscle function in a mouse model of Duchenne muscular dystrophy. *Science (80-.).* **351**, 403–407 (2016).

146. Qi, L. S. *et al.* Repurposing CRISPR as an RNA-Guided Platform for Sequence-Specific Control of Gene Expression. *Cell* **152**, 1173–1183 (2013).

147. Tadić, V., Josipović, G., Zoldoš, V. & Vojta, A. CRISPR/Cas9-based epigenome editing: An overview of dCas9-based tools with special emphasis on off-target activity. *Methods* **164**–**165**, 109–119 (2019).

148. Brocken, D. J. W., Tark-Dame, M. & Dame, R. T. dCas9: A versatile tool for epigenome editing. *Curr. Issues Mol. Biol.* **26**, 15–32 (2018).

149. Vojta, A. *et al.* Repurposing the CRISPR-Cas9 system for targeted DNA methylation. *Nucleic Acids Res.* **44**, 5615–5628 (2016).

150. Chavez, A. *et al.* Highly efficient Cas9-mediated transcriptional programming. *Nat. Methods* **12**, 326–328 (2015).

151. Perez-Pinera, P. *et al.* RNA-guided gene activation by CRISPR-Cas9-based transcription factors. *Nat. Methods* **10**, 973–976 (2013).

152. Yeo, N. C. *et al.* An enhanced CRISPR repressor for targeted mammalian gene regulation. *Nat. Methods* **15**, 611–616 (2018).

153. Gilbert, L. A. *et al.* CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. *Cell* **154**, 442 (2013).

154. Hilton, I. B. *et al.* Epigenome editing by a CRISPR/Cas9-based acetyltransferase activates genes from promoters and enhancers. *Nat. Biotechnol.* **33**, 510–517 (2015).

155. Kwon, D. Y., Zhao, Y. T., Lamonica, J. M. & Zhou, Z. Locus-specific histone deacetylation using a synthetic CRISPR-Cas9-based HDAC. *Nat. Commun.* **8**, 15315 (2017).

156. Chen, X. *et al.* Construction and validation of the CRISPR/dCas9-EZH2 system for targeted H3K27Me3 modification. *Biochem. Biophys. Res. Commun.* **511**, 246–252 (2019).

157. Choudhury, S. R., Cui, Y., Lubecka, K., Stefanska, B. & Irudayaraj, J. CRISPR-dCas9 mediated TET1 targeting for selective DNA demethylation at BRCA1 promoter. *Oncotarget* **7**, 46545–46556 (2016).

158. Didovyk, A., Borek, B., Tsimring, L. & Hasty, J. Transcriptional Regulation with CRISPR-Cas9: Principles, Advances, and Applications. *Curr. Opin. Biotechnol.* **40**, 177–184 (2017).

159. Bikard, D. *et al.* Programmable repression and activation of bacterial gene expression using an engineered CRISPR-Cas system. *Nucleic Acids Res.* **41**, 7429–7437 (2013).

160. Zalatan, J. G. *et al.* Engineering complex synthetic transcriptional programs with CRISPR RNA scaffolds. *Cell* **160**, 339–350 (2015).

161. Pengue, G. & Lania, L. Krüppel-associated box-mediated repression of RNA polymerase II promoters is influenced by the arrangement of basal promoter elements. *Proc. Natl. Acad. Sci. U. S. A.* **93**, 1015–1020 (1996).

162. Groner, A. C. *et al.* KRAB-zinc finger proteins and KAP1 can mediate long-range transcriptional repression through heterochromatin spreading. *PLoS Genet.* **6**, e1000869 (2010).

163. Kubo, N. *et al.* Promoter-proximal CTCF-binding promotes long-range-enhancer dependent gene activation. *Nat. Struct. Mol. Biol.* **28**, 152–161 (2021).

164. Schmitz, M. L. & Baeuerle, P. A. The p65 subunit is responsible for the strong transcription activating potential of NF-χB. *EMBO J.* **10**, 3805–3817 (1991).

165. Beerli, R. R., Segal, D. J., Dreier, B. & Barbas, C. F. Toward controlling gene expression at will: Specific regulation of the erbB-2/HER-2 promoter by using polydactyl zinc finger proteins constructed from modular building blocks. *Proc. Natl. Acad. Sci. U. S. A.* **95**, 14628–14633 (1998).

166. Kearns, N. A. *et al.* Functional annotation of native enhancers with a Cas9 -histone demethylase fusion. *Nat. Methods* **12**, 401–403 (2015).

167. Martinez-Escobar, A., Luna-Callejas, B. & Ramón-Gallegos, E. CRISPR-dCas9-Based Artificial Transcription Factors to Improve Efficacy of Cancer Treatment With Drug Repurposing: Proposal for Future Research. *Front. Oncol.* **10**, 604948 (2021).

168. Fu, Y. *et al.* High frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nat. Biotechnol.* **31**, 822–826 (2013).

169. Pattanayak, V. *et al.* High-throughput profiling of off-target DNA cleavage reveals RNA-programmed Cas9 nuclease specificity. *Nat. Biotechnol.* **31**, 839–843 (2013).

170. Lessard, S. *et al.* Human genetic variation alters CRISPR-Cas9 on- and off-targeting specificity at therapeutically implicated loci. *Proc. Natl. Acad. Sci. U. S. A.* **114**, E11257–E11266 (2017).

171. Cho, S. W. *et al.* Analysis of off-target effects of CRISPR/Cas-derived RNA-guided endonucleases and nickases. *Genome Res.* **24**, 132–141 (2014).

172. Naeem, M., Majeed, S., Hoque, M. Z. & Ahmad, I. Latest Developed Strategies to Minimize the Off-Target Effects in CRISPR-Cas-Mediated Genome Editing. *Cells* **9**, 1608 (2020).

173. Doench, J. G. *et al.* Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. *Nat. Biotechnol.* **34**, 184–191 (2016).

174. Vakulskas, C. A. *et al.* A high-fidelity Cas9 mutant delivered as a ribonucleoprotein complex enables efficient gene editing in human hematopoietic stem and progenitor cells. *Nat. Med.* **24**, 1216–1224 (2018).

175. Wu, X. *et al.* Genome-wide binding of the CRISPR endonuclease Cas9 in mammalian cells. *Nat. Biotechnol.* **32**, 670–676 (2014).

176. O'Geen, H., Henry, I. M., Bhakta, M. S., Meckler, J. F. & Segal, D. J. A genome-wide analysis of Cas9 binding specificity using ChIP-seq and targeted sequence capture. *Nucleic Acids Res.* **43**, 3389–3404 (2015).

177. Kuscu, C., Arslan, S., Singh, R., Thorpe, J. & Adli, M. Genome-wide analysis reveals characteristics of off-target sites bound by the Cas9 endonuclease. *Nat. Biotechnol.* **32**,

677–683 (2014).

178. Thakore, P. I. *et al.* Highly Specific Epigenome Editing by CRISPR/Cas9 Repressors for Silencing of Distal Regulatory Elements. *Nat. Methods* **12**, 1143–1149 (2015).

179. Gilbert, L. A. *et al.* Genome-Scale CRISPR-Mediated Control of Gene Repression and Activation. *Cell* **159**, 647–661 (2014).

180. Hsu, P. D. *et al.* DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat. Biotechnol.* **31**, 827–832 (2013).

181. Polstein, L. R. *et al.* Genome-wide specificity of DNA binding, gene regulation, and chromatin remodeling by TALE- and CRISPR/Cas9-based transcriptional activators. *Genome Res.* **25**, 1158–1169 (2015).

182. Perez-Pinera, P. *et al.* RNA-guided gene activation by CRISPR-Cas9–based transcription factors. *Nat. Methods* **10**, 973–976 (2013).

183. Yip, B. H. Recent advances in CRISPR/Cas9 delivery strategies. *Biomolecules* **10**, 839 (2020).

184. Liu, Z. *et al.* Application of Various Delivery Methods for CRISPR/dCas9. *Mol. Biotechnol.* **62**, 355–363 (2020).

185. Lino, C. A., Harper, J. C., Carney, J. P. & Timlin, J. A. Delivering crispr: A review of the challenges and approaches. *Drug Deliv.* **25**, 1234–1257 (2018).

186. Søndergaard, J. N. *et al.* Successful delivery of large-size CRISPR/Cas9 vectors in hard-to-transfect human cells using small plasmids. *Commun. Biol.* **3**, 319 (2020).

187. Horii, T. *et al.* Validation of microinjection methods for generating knockout mice by CRISPR/Cas-mediated genome engineering. *Sci. Rep.* **4**, 4513 (2014).

188. Le, Q. A. *et al.* Comparison of the effects of introducing the CRISPR/Cas9 system by microinjection and electroporation into porcine embryos at different stages. *BMC Res. Notes* **14**, 7 (2021).

189. Laustsen, A. & Bak, R. O. Electroporation-Based CRISPR/Cas9 Gene Editing Using Cas9 Protein and Chemically Modified sgRNAs. *Methods Mol. Biol.* **1961**, 127–134 (2019).

190. Li, L., Hu, S. & Chen, X. Non-viral delivery systems for CRISPR/Cas9-based genome

editing: challenges and opportunities. *Biomaterials* **171**, 207–218 (2018).

191. Xu, Y. & Szoka, F. C. Mechanism of DNA Release from Cationic Liposome/DNA Complexes Used in Cell Transfection,. *Biochemistry* **35**, 5616–5623 (1996).

192. Zhu, L. & Mahato, R. I. Lipid and polymeric carrier-mediated nucleic acid delivery. *Expert Opin. Drug Deliv.* **7**, 1209–1226 (2010).

193. Whitt, M., Buonocore, L. & Rose, J. K. Liposome-mediated transfection. *Curr. Protoc. Immunol.* **Chapter 10**, Unit 10.16 (2001).

194. Chen, K. *et al.* Cationic polymeric nanoformulation: Recent advances in material design for CRISPR/Cas9 gene therapy. *Prog. Nat. Sci. Mater. Int.* **29**, 617–627 (2019).

195. Zhang, L. *et al.* Lipid nanoparticle-mediated efficient delivery of CRISPR/Cas9 for tumor therapy. *NPG Asia Mater.* **9**, 3–10 (2017).

196. Van Der Aa, M. A. E. M. *et al.* Cellular uptake of cationic polymer-DNA complexes via caveolae plays a pivotal role in gene transfection in COS-7 cells. *Pharm. Res.* **24**, 1590–1598 (2007).

197. Xu, C. L., Ruan, M. Z. C., Mahajan, V. B. & Tsang, S. H. Viral delivery systems for crispr. *Viruses* **11**, 28 (2019).

198. Ciuffi, A. Mechanisms Governing Lentivirus Integration Site Selection. *Curr. Gene Ther.* **8**, 419–429 (2008).

199. Deyle, D. R. & Russell, D. W. Adeno-associated virus vector integration. *Curr. Opin. Mol. Ther.* **11**, 442–447 (2009).

200. Recchia, A. *et al.* Site-specific integration mediated by a hybrid adenovirus/adenoassociated virus vector. *Proc. Natl. Acad. Sci. U. S. A.* **96**, 2615–2620 (1999).

201. Hirsch, M. L., Sonya, W. J. & Samulski, R. J. Delivering Transgenic DNA Exceeding the Carrying Capacity of AAV Vectors. *Methods Mol. Biol.* **1382**, 21–39 (2016).

202. Grieger, J. C. & Samulski, R. J. Packaging Capacity of Adeno-Associated Virus Serotypes: Impact of Larger Genomes on Infectivity and Postentry Steps. *J. Virol.* **79**, 9933–9944 (2005).

203. Schlimgen, R. *et al.* Risks associated with lentiviral vector exposures and prevention

strategies. *J. Occup. Environ. Med.* **58**, 1159–1166 (2016).

204. Sandoval-Villegas, N., Nurieva, W., Amberger, M. & Ivics, Z. Contemporary transposon tools: A review and guide through mechanisms and applications of sleeping beauty, piggybac and tol2 for genome engineering. *Int. J. Mol. Sci.* **22**, 5084 (2021).

205. Wells, J. N. & Feschotte, C. A Field Guide to Eukaryotic Transposable Elements. *Annu. Rev. Genet.* **54**, 539–561 (2020).

206. Hickman, A. B. & Dyda, F. Mechanisms of DNA Transposition. *Microbiol. Spectr.* **3**, MDNA3-0034–2014 (2015).

207. Ivics, Z., Hackett, P. B., Plasterk, R. H. & Izsvák, Z. Molecular reconstruction of sleeping beauty, a Tc1-like transposon from fish, and its transposition in human cells. *Cell* **91**, 501–510 (1997).

208. Elick, T. A., Bauser, C. A. & Fraser, M. J. Excision of the piggyBac transposable element in vitro is a precise event that is enhanced by the expression of its encoded transposase. *Genetica* **98**, 33–41 (1996).

209. Grabundzija, I. *et al.* Comparative analysis of transposable element vector systems in human cells. *Mol. Ther.* **18**, 1200–1209 (2010).

210. Miskey, C. *et al.* Engineered Sleeping Beauty transposase redirects transposon integration away from genes. *Nucleic Acids Res.* **50**, 2807–2825 (2022).

211. Zhou, Y., Ma, G., Yang, J., Gao, Z. & Guo, Y. The Integration Preference of Sleeping Beauty at Non-TA Site Is Related to the Transposon End Sequences. *Front. Genet.* **12**, 639125 (2021).

212. Li, R., Zhuang, Y., Han, M., Xu, T. & Wu, X. piggyBac as a high-capacity transgenesis and gene-therapy vector in human cells and mice. *Dis. Model. Mech.* **6**, 828–833 (2013).

213. Izsvák, Z., Ivics, Z. & Plasterk, R. H. Sleeping Beauty, a wide host-range transposon vector for genetic transformation in vertebrates. *J. Mol. Biol.* **302**, 93–102 (2000).

214. Hudecek, M. & Ivics, Z. Non-viral therapeutic cell engineering with the Sleeping Beauty transposon system. *Curr. Opin. Genet. Dev.* **52**, 100–108 (2018).

215. Rostovskaya, M. *et al.* Transposon-mediated BAC transgenesis in human ES cells.

Nucleic Acids Res. **40**, e150 (2012).

216. Emery, D. W. The use of chromatin insulators to improve the expression and safety of integrating gene transfer vectors. *Hum. Gene Ther.* **22**, 761–774 (2011).

217. Sharma, N. *et al.* The Impact of cHS4 Insulators on DNA Transposon Vector Mobilization and Silencing in Retinal Pigment Epithelium Cells. *PLoS One* **7**, e48421 (2012).

218. Bire, S. *et al.* Optimization of the piggyBac transposon using mRNA and insulators: Toward a more reliable gene delivery system. *PLoS One* **8**, e82559 (2013).

219. Jiang, Y. *et al.* Robust genome and RNA editing via CRISPR nucleases in PiggyBac systems. *Bioact. Mater.* **14**, 313–320 (2022).

220. Rao, S., Yao, Y. & Bauer, D. E. Editing GWAS: experimental approaches to dissect and exploit disease-associated genetic variation. *Genome Med.* **13**, 41 (2021).

221. Gallagher, M. D. & Chen-Plotkin, A. S. The Post-GWAS Era: From Association to Function. *Am. J. Hum. Genet.* **102**, 717–730 (2018).

222. Lichou, F. & Trynka, G. Functional studies of GWAS variants are gaining momentum. *Nat. Commun.* **11**, 2–5 (2020).

223. Werder, R. B. *et al.* CRISPR interference interrogation of COPD GWAS genes reveals the functional significance of desmoplakin in iPSC-derived alveolar epithelial cells. *Sci. Adv.* **8**, eabo6566 (2022).

224. Fink-Baldauf, I. M., Stuart, W. D., Brewington, J. J., Guo, M. & Maeda, Y. CRISPRi links COVID-19 GWAS loci to LZTFL1 and RAVER1. *eBioMedicine* **75**, 103806 (2022).

225. Narimatsu, Y. *et al.* Genetic glycoengineering in mammalian cells. *J. Biol. Chem.* **296**, 100448 (2021).

226. Narimatsu, Y. *et al.* An Atlas of Human Glycosylation Pathways Enables Display of the Human Glycome by Gene Engineered Cells. *Mol. c* **75**, 394–407 (2019).

227. Schulz, M. A. *et al.* Glycoengineering design options for IgG1 in CHO cells using precise gene editing. *Glycobiology* **28**, 542–549 (2018).

228. Chang, M. M. *et al.* Small-molecule control of antibody N-glycosylation in engineered

mammalian cells. *Nat. Chem. Biol.* **15**, 730–736 (2019).

229. Bieberich, E. Synthesis, Processing, and Function of N-glycans in N-glycoproteins. *Adv. Neurobiol.* **9**, 47–70 (2014).

230. Wang, T. T. IgG Fc Glycosylation in Human Immunity. *Curr. Top. Microbiol. Immunol.* **423**, 63–75 (2019).

231. Verdi, S. *et al.* TwinsUK: The UK Adult Twin Registry Update. *Twin Res. Hum. Genet.* **22**, 523–529 (2019).

232. Kaneyoshi, K., Yamano-Adachi, N., Koga, Y., Uchiyama, K. & Omasa, T. Analysis of the immunoglobulin G (IgG) secretion efficiency in recombinant Chinese hamster ovary (CHO) cells by using Citrine-fusion IgG. *Cytotechnology* **71**, 193–207 (2019).

233. Sharker, S. M. & Rahman, A. A Review on the Current Methods of Chinese Hamster Ovary (CHO) Cells Cultivation for the Production of Therapeutic Protein. *Curr. Drug Discov. Technol.* **18**, 354–364 (2021).

234. Walsh, G. Biopharmaceutical benchmarks 2018. *Nat. Biotechnol.* **36**, 1136–1145 (2018).

235. Malm, M. *et al.* Harnessing secretory pathway differences between HEK293 and CHO to rescue production of difficult to express proteins. *Metab. Eng.* **72**, 171–187 (2022).

236. Butler, M. & Spearman, M. The choice of mammalian cell host and possibilities for glycosylation engineering. *Curr. Opin. Biotechnol.* **30**, 107–112 (2014).

237. Goh, J. B. & Ng, S. K. Impact of host cell line choice on glycan profile. *Crit. Rev. Biotechnol.* **38**, 851–867 (2018).

238. Jeon, J. P. *Human Lymphoblastoid Cell Lines in Pharmacogenomics*. *Handbook of Pharmacogenomics and Stratified Medicine* (Elsevier Inc., 2014). doi:10.1016/B978-0-12-386882-4.00006-2.

239. Jolly, L. A., Sun, Y., Carroll, R., Homan, C. C. & Gecz, J. Robust imaging and gene delivery to study human lymphoblastoid cell lines. *J. Hum. Genet.* **63**, 945—955 (2018).

240. Winiarska, M. *et al.* Selection of an optimal promoter for gene transfer in normal B cells. *Mol. Med. Rep.* **16**, 3041–3048 (2017).

241. Malm, M. *et al.* Evolution from adherent to suspension : systems biology of HEK293 cell line development. *Sci. Rep.* 1–15 (2020) doi:10.1038/s41598-020-76137-8.

242. Engler, C. & Marillonnet, S. Golden Gate Cloning. *Methods Mol. Biol.* **1116**, 119–131 (2014).

243. Kreiss, P. *et al.* Plasmid DNA size does not affect the physicochemical properties of lipoplexes but modulates gene transfer efficiency. *Nucleic Acids Res.* **27**, 3792–3798 (1999).

244. Sweeney, N. P. & Vink, C. A. The impact of lentiviral vector genome size and producer cell genomic to gag-pol mRNA ratios on packaging efficiency and titre. *Mol. Ther. Methods Clin. Dev.* **21**, 574–584 (2021).

245. Doherty, J. E. *et al.* Hyperactive *piggyBac* Gene Transfer in Human Cells and *In Vivo*. *Hum. Gene Ther.* **23**, 311–320 (2012).

246. Hirst, J. *et al.* A Family of Proteins with Gamma-Adaptin and VHS Domains that Facilitate Trafficking between the Trans-Golgi Network and the Vacuole/Lysosome. *J. Cell Biol.* **149**, 67–79 (2000).

247. He, X., Li, F., Chang, W.-P. & Tang, J. GGA proteins mediate the recycling pathway of memapsin 2 (BACE). *J. Biol. Chem.* **280**, 11696–11703 (2005).

248. Kitazume, S. *et al.* Alzheimer's beta-secretase, beta-site amyloid precursor protein-cleaving enzyme, is responsible for cleavage secretion of a Golgi-resident sialyltransferase. *Proc. Natl. Acad. Sci.* **98**, 13554–13559 (2001).

249. Sugimoto, I. *et al.* Beta-Galactoside Alpha-2,6-Sialyltransferase I Cleavage by BACE1 Enhances the Sialylation of Soluble Glycoproteins. A Novel Regulatory Mechanism for Alpha2,6-Sialylation. *J. Biol. Chem.* **282**, 34896–34903 (2007).

250. Voss, M. *et al.* Shedding of glycan-modifying enzymes by signal peptide peptidase-like 3 (SPPL3) regulates cellular N-glycosylation. *EMBO J.* **33**, 2890–2905 (2014).

251. Kuhn, P.-H. *et al.* Secretome Analysis Identifies Novel Signal Peptide Peptidase-Like 3 (Sppl3) Substrates and Reveals a Role of Sppl3 in Multiple Golgi Glycosylation Pathways. *Mol. Cell. Proteomics* **14**, 1584–1598 (2015).

252. Nagae, M. *et al.* Structure and mechanism of cancer-associated N-

acetylglucosaminyltransferase-V. *Nat. Commun.* **9**, 3380 (2018).

253. Ferrucci, L. & Fabbri, E. Inflammageing: chronic inflammation in ageing, cardiovascular disease, and frailty. *Nat. Rev. Cardiol.* **15**, 505–522 (2018).

254. Plomp, R. *et al.* Subclass-specific IgG glycosylation is associated with markers of inflammation and metabolic health. *Sci. Rep.* **7**, 1–10 (2017).

255. Dall'Olio, F. & Malagolini, N. Immunoglobulin G Glycosylation Changes in Aging and Other Inflammatory Conditions. *Exp. Suppl.* **112**, 303–340 (2021).

256. Sharapov, S. Z. *et al.* Replication of 15 loci involved in human plasma protein N-glycosylation in 4802 samples from four cohorts. *Glycobiology* **31**, 82–88 (2021).

257. Korinfskaya, S., Parameswaran, S., Weirauch, M. T. & Barski, A. Runx Transcription Factors in T Cells—What Is Beyond Thymic Development? *Front. Immunol.* **12**, 701924 (2021).

258. Wang, D. *et al.* Oestrogen up-regulates DNMT1 and leads to the hypermethylation of RUNX3 in the malignant transformation of ovarian endometriosis. *Reprod. Biomed. Online* **44**, 27–37 (2022).

259. Wang, X. *et al.* Genome-wide interaction analysis of menopausal hormone therapy use and breast cancer risk among 62,370 women. *Sci. Rep.* **12**, 6199 (2022).

260. Liu, H. *et al.* RUNX3 Epigenetic Inactivation Is Associated With Estrogen Receptor Positive Breast Cancer. *J. Histochem. Cytochem.* **66**, 709–721 (2018).

261. Song, X. Y., Li, B. Y., Zhou, E. X. & Wu, F. X. The clinicopathological significance of RUNX3 hypermethylation and mRNA expression in human breast cancer, a meta-analysis. *Onco. Targets. Ther.* **9**, 5339–5347 (2016).

262. Guo, C. *et al.* RUNX3 is inactivated by promoter hypermethylation in malignant transformation of ovarian endometriosis. *Oncol. Rep.* **32**, 2580–2588 (2014).

263. Huang, B. *et al.* RUNX3 acts as a tumor suppressor in breast cancer by targeting estrogen receptor α. *Oncogene* **31**, 527–534 (2012).

264. Sakuma, A. *et al.* Loss of Runx3 affects ovulation and estrogen-induced endometrial cell proliferation in female mice. *Mol. Reprod. Dev.* **75**, 1653–1661 (2008).

265. Kent, W. J. *et al.* The human genome browser at UCSC. *Genome Res.* **12**, 996–1006

(2002).

266. Visscher, P. M. Power of the classical twin design revisited. *Twin Res.* **7**, 505–512 (2004).

6. BIOGRAPHY

Anika Mijakovac was born on August 18, 1995 in Zadar, Croatia. She grew up in Knin, Croatia where she finished the secondary school of languages in 2014. Three years later she received a Bachelor's degree in Biology at the Faculty of Science, University of Zagreb. She obtained the Master's degree in Molecular biology at the same university, graduating in 2019, first in her generation.

In 2019, Anika started working in the group for Epigenetics led by professor Vlatka Zoldoš at the Faculty of Science, University of Zagreb as a project associate (IRI - The New Generation of High-throughput Glycoservices). In the same research group she started her Biology PhD under the mentorship of associate professor Aleksandar Vojta. Her main research focus is the regulation of immunoglobulin G glycosylation.

Anika participated in several international conferences with poster presentations. She is the first author of three original research articles published in highly-ranked scientific journals. She mentored several undergraduate and graduate students in laboratory skill training and co-mentored one graduate thesis. She also works as a Genetics teaching assistant at the Faculty of Science, University of Zagreb for the last four years.