Fucosylation of anti-dsDNA IgG1 correlates with disease activity of treatment-naïve systemic lupus erythematosus patients

Jing Han,[a,†] Zhuochao Zhou,[b,†] Rongrong Zhang,[a,†] Yijun You,[b,†] Zizhen Guo,[c] Jinyan Huang,[d] Fan Wang,[b] Yue Sun,[b] Honglei Liu,[b] Xiaobing Cheng,[b] Yutong Su,[b] Hui Shi,[b] Qiongyi Hu,[b] Jialin Teng,[b,*] Chengde Yang,[b,*] Shifang Ren,[a,**] and Junna Ye [b,*]

[a]NHC Key Laboratory of Glycoconjugates Research, Department of Biochemistry and Molecular Biology, School of Basic Medical Sciences, Fudan University, Shanghai, China
[b]Department of Rheumatology and Immunology, Ruijin Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China
[c]Department of Plastic and Reconstructive Surgery, Shanghai Ninth People's Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China
[d]Biomedical Big Data Center, The First Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou, China

Summary

Background Systemic Lupus Erythematosus (SLE) is a complex and heterogeneous autoimmune disease mediated by quantities of autoantibodies in which anti-double-stranded DNA (anti-dsDNA) antibodies are important. Besides, glycosylation is one of the most commonly post-translational modifications of antibodies. The association of anti-dsDNA antibodies glycosylation and SLE disease activity is still unknown.

Methods We enrolled 101 consecutive treatment-naÿve SLE patients with positive anti-dsDNA antibodies from the Department of Rheumatology and Immunology at Ruijin Hospital, Shanghai, between 2017 and 2019. Serum samples were used in this study. We analysed the glycosylation of anti-dsDNA IgG and total IgG subclasses according to systemic lupus erythematosus disease activity index (SLEDAI) scores. Statistical analysis and machine learning were performed to assess the correlation between glycosylation of anti-dsDNA IgG and total IgG with disease activity.

Findings Serum samples from 86 patients could be detected with anti-dsDNA IgG glycopeptide and subclass of IgG glycoform. Cluster analysis showed that glycosylation of anti-dsDNA IgG and total IgG subclasses were different in SLE patients. Fucosylation, galactosylation, and sialylation levels of anti-dsDNA IgG1 were increased with SLEDAI scores (all $p<0.05$). The results of machine learning showed that all the glycoforms of anti-dsDNA IgG1 had better performance with lower standardised square error (SSE) than that of total IgG1, with anti-dsDNA IgG1 fucosylation level having the lowest SSE (0.009).

Interpretation Our study indicated that glycosylation of anti-dsDNA IgG was different from that of total IgG and fucosylation of anti-dsDNA IgG1 correlated best with SLE disease activity.

*Corresponding authors at: Department of Rheumatology and Immunology, Ruijin Hospital, Shanghai Jiao Tong University School of Medicine, No. 197 Ruijin Second Road, Huangpu District, Shanghai 200025, China.
**NHC Key Laboratory of Glycoconjugates Research, Department of Biochemistry and Molecular Biology, School of Basic Medical Sciences, Fudan University, Shanghai 200032, China.

E-mail addresses: tengteng8151@sina.com (J. Teng), yangchengde@sina.com (C. Yang), renshifang@fudan.edu.cn (S. Ren), yjn0912@qq.com (J. Ye).

† These authors make the same contributions to this paper.

Articles

Research in context

Evidence before this study

Systemic Lupus Erythematosus (SLE) is a complex and heterogeneous autoimmune disease mediated by quantities of autoantibodies in which anti-double-stranded DNA (anti-dsDNA) antibodies are important. Numerous researches reported that autoantibody modifications played a pivotal role in different autoimmune diseases such as SLE. Glycosylation is one of the most commonly post-translational modifications of antibodies. The association of anti-dsDNA antibodies glycosylation and SLE disease activity is unknown.

Added value of this study

The glycosylation of specific antibodies in SLE has not been evaluated, especially anti-dsDNA autoantibodies, which are the most critical antibodies in SLE. Our study compared the glycosylation of anti-dsDNA IgG with total IgG purified from treatment-naïve SLE patients' serum, including their subclasses in terms of different systemic lupus erythematosus disease activity index (SLEDAI) scores. Cluster analysis showed that glycosylation of anti-dsDNA IgG and total IgG subclasses were different in SLE patients. Fucosylation, galactosylation, and sialylation levels of anti-dsDNA IgG1 were increased with SLEDAI scores (all $p<0.05$). Machine learning was further performed and found that all the glycoforms of anti-dsDNA IgG1 had better performance with lower standardised square error (SSE) than that of total IgG1, with anti-dsDNA IgG1 fucosylation level having the lowest SSE (0.009), indicating that glycosylation of anti-dsDNA IgG was different from that of total IgG and fucosylation of anti-dsDNA IgG1 correlated best with SLE disease activity.

Implications of all the available evidence

The previous study only compared the glycosylation changes in total IgG of SLE with healthy controls. In our study, we compared the association of glycosylation of specific purified anti-dsDNA IgG and total IgG for the first time, including their subclasses, with disease activity of SLE. Fucosylation of anti-dsDNA IgG1 was found to be correlated best with SLE disease activity, which might be a novel way to assess disease activity and contribute to pathogenesis of SLE.

Introduction

Systemic Lupus Erythematosus (SLE) is a complex and heterogeneous autoimmune disease. Autoantibodies such as antinuclear antibodies (ANA) including anti-double-stranded DNA (anti-dsDNA) antibodies play an important role in SLE. Anti-dsDNA antibodies could contribute to organ damage, which subsequently influence patients' quality of life.[1] The association of anti-dsDNA antibodies with organ damage has acquired much attention.[2] It's noteworthy that renal involvement is significantly more frequent in patients with positive anti-dsDNA antibodies than negative ones.[3] Anti-dsDNA antibodies are part of systemic lupus erythematosus disease activity index (SLEDAI) score, a well-defined tool reflecting the disease activity of SLE, which includes rash, mucous ulcers, alopecia, proteinuria and so on.[4,5] Recently, numerous researches reported that autoantibody modifications played a pivotal role in different diseases (e.g., SLE, Rheumatoid Arthritis (RA), Anti-phospholipid Syndrome).[6-8] Among them, glycosylation, one of the post-translational modifications of immunoglobulin (IgG), was attached importance in the pathogenesis of autoimmune diseases. The interactions between protein and glycan residues are critical for the IgG molecule's structural stability and functional activity, influencing the outcome of the immune response.[9] For example, one study found that in anti-citrullinated protein antibody (ACPA) positive RA, the pregnancy-induced changes in galactosylation of ACPA-IgG, not that of total IgG, were associated with disease activity.[10] Another study suggested that abnormal total IgG glycome composition or changes in total IgG glycosylation may be an important molecular mechanism in SLE. The most significant changes in total IgG glycosylation included decreased galactosylation, sialylation, core fucose, and increased bisecting N-acetylglucosamine.[11] Native circulating total IgG complexes from active SLE patients exposed fucosyl residues, and their glycan core was accessible to soluble lectins.[12] Besides, in a mice study, mutation of a single gene encoding α-mannosidase II, which regulated the hybrid to the complex branching pattern of extracellular asparagine (N)-linked oligosaccharide chains (N-glycans), resulted in a systemic autoimmune disease similar to SLE.[13] Thus, glycosylation of total IgG might be crucial in the pathogenesis of SLE. However, total IgG is an aggregate of all immunoglobulins. The glycosylation of specific antibodies in SLE has not been evaluated, especially anti-dsDNA IgG, which is the most critical antibody in SLE. Therefore, we aimed to compare the glycosylation of anti-dsDNA IgG with total IgG, including their subclasses in treatment-naÿve SLE patients in terms of different SLEDAI scores.

Methods

Study population

We enrolled 101 consecutive treatment-naÿve SLE patients with positive anti-dsDNA antibody from the Department of Rheumatology and Immunology at Rui-jin Hospital, Shanghai, from 2017 to 2019. Patients with acquired immunodeficiency syndrome, cancer or individuals who were less than 18 years old were not included in this study. The diagnosis of SLE was confirmed by two qualified rheumatologists (Jialin Teng and Chengde Yang) using the 2019 European League Against Rheumatism/American College of

Rheumatology classification criteria and its 2017 version.[14] Demographic and clinical characteristics (e.g., fever, pericarditis, pleuritis, mouth ulcers, rash, arthritis, vasculitis, and nephritis), as well as laboratory tests (e.g., anti-dsDNA IgG levels measured by enzyme-linked immunosorbent assay (ELISA), erythrocyte sedimentation rate (ESR), white blood cell counts (WBC), haemoglobin (Hb), platelets (PLT), C-reactive protein (CRP), IgG, complement 3 (C3) and complement 4 (C4)), were collected from patients' medical records. Nephritis was defined as having casts, haematuria, proteinuria, and pyuria, excluding other renal diseases. SLEDAI scores were used to measure disease activity according to the following four segments: 0-4: not active, 4-9: mildly active, 9-14: moderately active, >14: severely active. Serum was obtained and then stored at -80°C before use.

Ethics The study was approved by the Institutional Research Ethics Committee of Ruijin Hospital (ID: 2016-62), Shanghai, China. The study followed the ethical standards for human experimentation established in the Declaration of Helsinki and all patients provided their informed consent to participate in this study.

Isolation of total and anti-dsDNA IgG from human serum

Immunoglobulins were isolated from human serum using Protein G according to previous reports with mild modifications.[15] Isolation was performed at room temperature; human serum samples were diluted with 1X PBS (pH=7.0). The diluted serum was loaded on protein G Agarose Prepacked Column (Beyotime Biotechnology, Shanghai) by centrifugation and washed three times with 1X PBS and H_2O, respectively. The bound IgGs were eluted with 100mM formic acid, concentrated on a CentriVap Centrifuge to dryness, stored at -20°C before the following step. Then, anti-dsDNA IgG was purified from total IgG by a dsDNA affinity column (Sigma, USA) as described previously.[16]

Trypsin digestion of IgGs

The captured and dried IgGs were reconstituted in 50mM ammonium bicarbonate (Sigma Aldrich, Steinheim, Germany) and then reduced by 2 μL of 550 mM DTT and alkylated by 4 μL of 450 mM IAA (both solvents are 50mM NH_4HCO_3, freshly made) sequentially. Next, the samples were digested with Trypsin at 37°C overnight. Tryptic peptides were subjected to LC/MS analysis without further processing to ensure reliable quantification as described before.[17]

UPLC-ESI-QqQ MS analysis

An ultra performance liquid chromatography system Nexera UPLC LC-30A (Shimadzu Corporation, Kyoto, Japan) equipped with a ZORBAX RRHD Eclipse Plus C18 column (1.8 μm, 2.1 mm × 100 mm, Agilent

Technologies, Santa Clara, CA) was coupled online to a 6500 plus Qtrap Mass Spectrometer (AB Sciex, CA, USA). Mobile phase A was 0.1% FA and 3% acetonitrile in nanopure water (v/v/v), and mobile phase B was 0.1% FA in 90% acetonitrile. Peptides and glycopeptides were separated by a binary gradient at 40°C with 0.5 mL/min flow rate consisting of (1) 0 min at 2.0% B; 0.5-1.0 min at 5.0% B; 6.0 min at 30.0% B; (2) 6.1-8.0 min at 100.0% B), (3) 8.1-10.0 min at 2.0% B. The MS was operated in the positive mode. Curtain gas: 30.0 psi, collision gas: high, ion spray voltage: 5500.0 V, temperature: 300.0°C, ion source gas 1: 55.0 psi, ion source gas 2: 60.0 psi, declustering potential: 20.0 V, entrance potential: 10.0 V, and collision cell exit potential: 14 V. The scheduled MRM mode was used. Q1 and Q3 were set to unit resolution. The cycle time was fixed to 500 ms, the MRM detection window was set to 30.0 s, and the target scan time was 0.5 s. CE for each MRM transition was optimized by a 5 V step followed by a 2 V step fine-tuning. The analysis was controlled using the AB Sciex Analyst software (version 1.6.3).

Data processing and statistical analysis

Data processing. Raw LC-MS-MRM data were converted to mzXML and intensities of peak area using AB Sciex MultiQuant software (version 3.0.2). Peak intensities were used for the quantification of each subclass-specific glycans. The limit of quantification was set as the signal-to-noise ratio (S/N) of 10. To achieve IgG subclass-specific quantification of glycoforms, we have isolated total IgG and targeted identification based on the specific amino acid sequences. Compared with the previous report, we reconfirm the fragments ion m/z of specific-subclass glycopeptides based on analysis of our experimental data achieving the strongest detectable signal (details in Supplementary Table S1).[18] We have selected a list of 26 glycoforms initially according to the prior analysis of glycans from reports in the literature.[19] Still, in fact, six glycoforms (A2G1S, FA2BG1S, FA2BG2S2, FA2G2S2, A2BG1S, A2BG2S) were not detectable in any subclass by our current method. Since IgG 1-4 subclasses vary among the samples, the intensity of glycopeptides was normalized to the intensity of the unique peptide corresponding to each IgG subclass. Glycopeptides from IgG3 and IgG4 could not be exactly separated, as their specific peptide (IgG3, EEQYNSTFR; IgG4, EEQFNSTYR) share the same amino acid composition and therefore have identical masses.[20] Some glycans were detectable, but below the quantification limit. Some glycans were quantifiable but in a limited subset of the samples. Finally, we compared 20 (IgG1), 16 (IgG2), 7 (IgG3/4) glycoforms detectable in all samples (bolded entries in Supplementary Table S1-S2).

Statistical analysis

Statistical analysis was performed using R (version 4.0.0). In descriptive statistics, data was expressed in the form of numbers (percentages) for categorical variables, and medians (Q1-Q3) or means ± standard deviations (SD) for continuous variables. Chi-squared test was used to compare categorical variables, and Wilcoxon test was used to compare continuous variables after exploring the normality of data distribution using the Shapiro-Wilk test. An overview of IgG in each SLEDAI group was performed by cluster analysis. The data were divided into four groups according to the SLEDAI scores as described above, and then cluster analysis was conducted using R package 'pheatmap' in each group. Specific cluster analysis code is provided as follows:

```
Pheatmap (negSLE1mt, cluster_cols = FALSE, show_colnames = FALSE,
         annotation_col = annota_col, border_color = NA,
         breaks = seq(0, 0.45, length.out = 100)).
```

Ordinary one-way ANOVA (analysis of variance) was performed to compare the concentration of subclass glycoforms of anti-dsDNA IgG and total IgG in four SLEDAI groups. T-test was used for comparisons between every two SLEDAI groups. These glycoforms data was explored normal distribution using Anderson-Darling test, Shapiro-Wilk test, Kolmog'orov-Smirnov test, and D'Agostino-Pearson test (GraphPad Prism 8.0). The level of statistical significance was set at a two-tailed α-value of 0.05 by default.

Gradient Boost Decision Tree (GBDT) machine learning

Machine learning was performed using the GBDT method through xgboost module by python (version 3.9). 75% of the samples were used as training data sets, 25% of the samples were used as validation data sets. All of the demographic and clinical characteristics (i.e., gender, age, disease duration (in months), photosensitivity, Raynaud's phenomenon, fever, pericarditis, pleurisy, mouth ulcers, rash, alopecia, myositis, arthritis, vasculitis, seizures, epilepsy, cerebrovascular accident, lupus headache, organic encephalopathy, visual impairment, cranial neuropathy) and laboratory indicators (i.e., leukocyturia, proteinuria, hematuria, cylinderuria, WBC, PLT, C3, C4, anti-dsDNA antibody) involved in SLEDAI rating scale were listed as training objectives. Different derived glycosylation traits: galactosylation (Gal), sialylation (Sia), bisection (Bis), and fucosylation (Fuc) with IgG subclasses (IgG1, IgG2, IgG3/4) (details in Supplementary Table S3) were used as test objectives. Predict value result from the module and real value of each sample was presented by a line plot. Standardized square error (SSE) (equation 1) of each test sample (after log-transformed

normalization) was calculated to evaluate the performance of machine learning.

Standardized square error

$$= \frac{\sum_{i-n}(\text{predict value}(i) - \text{real value}(i)) \wedge 2}{n} \quad (1)$$

where n is the total number of test samples.

Role of funding source

The funders of the manuscript had no role in the study design, data collection, data analysis, data interpretation, or writing of the manuscript.

Results

Study population and clinical characteristics

The workflow of this study was shown in Figure 1. Serum samples from 101 treatment-naïve SLE patients with positive anti-dsDNA antibodies were used in this study, and anti-dsDNA IgG was purified, of which 86 patients could be detected with anti-dsDNA IgG glycopeptide and subclass of IgG glycoform. The clinical characteristics of patients involved in this study were shown in Table 1. Besides, we set up Ruijin SLE cohort consisting of 200 SLE patients, and compared the demographics of patients in this study with patients from Ruijin SLE cohort. It showed that there were no difference between them in terms of age, gender, disease duration (months), anti-dsDNA IgG levels and SLEDAI score (all $p > 0.05$) (Supplementary Table S4). Then, the comparisons of glycosylation subclasses between anti-dsDNA IgG and total IgG in treatment-naïve SLE patients were shown in Table 2.

Cluster analysis of glycosylation of anti-dsDNA IgG and total IgG

For the first time, the differences of subclass-specific glycoforms and derived traits of SLE patients were analysed. Firstly, cluster analysis was used to analyse galactosylation, sialylation, bisection, and fucosylation of purified anti-dsDNA IgG and total IgG. The glycosylation level of IgG was shown in blue colour as low level and red colour as high level (Figure 2). A similar level of glycosylation was clustered together and one stripe represented one patient. The patients were presented from left to right with increasing SLEDAI scores as four categories (not active; mildly active; moderately active; severely active). An increase of core fucosylation was observed in both anti-dsDNA IgG and total IgG, but not in the same subclasses of IgG. For example, high-level fucosylation was seen in anti-dsDNA IgG1 and IgG3/4, while in total IgG1. Comparing the relative levels of
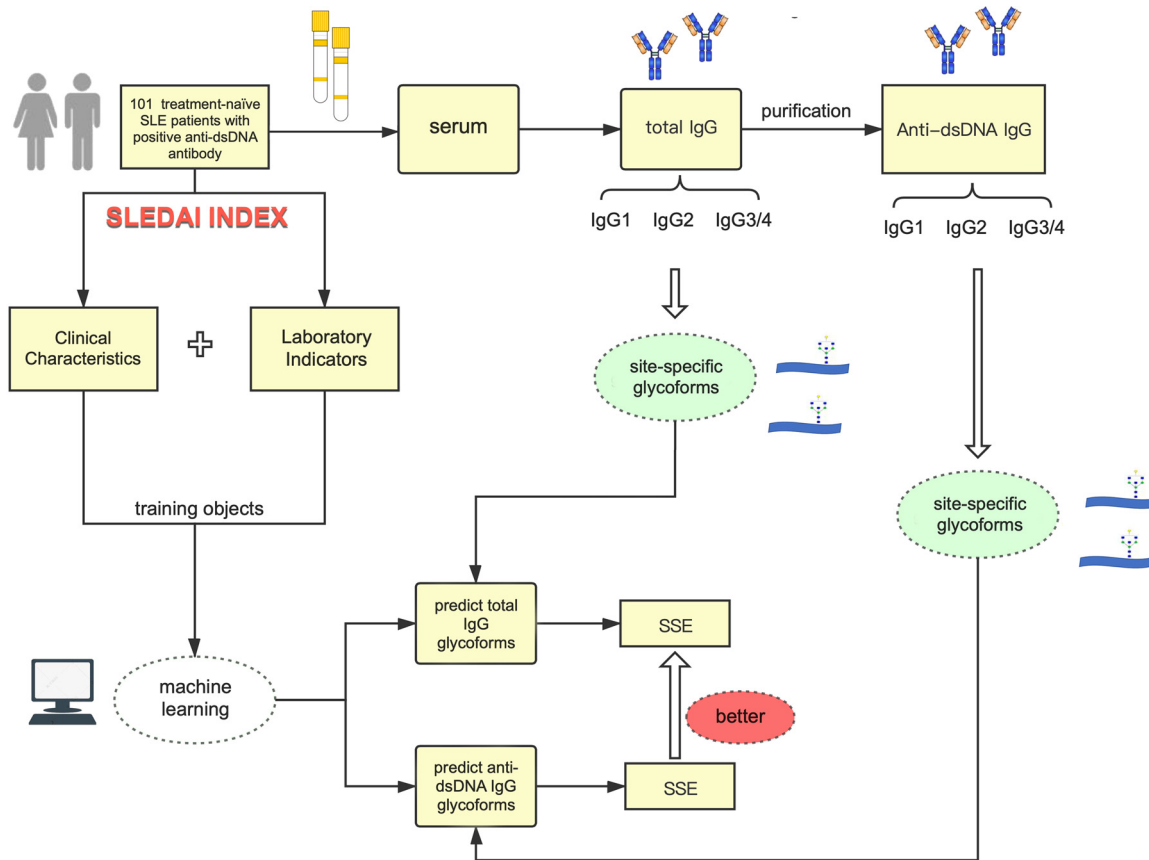
Figure 1. Workflow of comparison of glycosylation of anti-dsDNA IgG and total IgG in treatment-naïve SLE patients. 101 consecutive treatment-naïve SLE patients with positive anti-dsDNA antibodies were included. The demographic data, clinical characteristics, and laboratory findings were collected. Then, IgG was isolated from serum by using Protein G column and anti-dsDNA IgG was purified from total IgG. IgG subtype-specific glycoforms were quantified by UPLC-ESI-QqQ MS analysis. Statistical analysis and machine learning were then performed. SSE of predicted glycoform value and real glycoform value for each sample were calculated to assess the performance of machine learning.

Abbreviations: IgG: Immunoglobulin; SLEDAI: systemic lupus erythematosus disease activity index; SLE: systemic lupus erythematosus; SSE: Standardised square error.

sialylation in the two groups, interestingly, we found that each subclass of IgG in both groups has low-level sialylation. Conversely, the galactosylation showed a difference in each group: low level was in anti-dsDNA IgG2 and high level in total IgG1, indicating that the relative level of galactosylation was not identical. Levels of bisection for anti-dsDNA IgG1 and IgG2, total IgG3/4 were relatively low. A difference in glycosylation of specific subclasses of IgG between the two groups was observed.

SLE disease activity was associated with different subclasses of IgG glycosylation between anti-dsDNA IgG and total IgG

In Figure 2, the heatmap showed the difference of subclass-specific glycosylation of anti-dsDNA IgG and total IgG. To further analyse this result, the levels of glycosylation of anti-dsDNA IgG and total IgG were analysed according to four SLEDAI groups (group 1: SLEDAI scores ranged from 0-4; group 2: SLEDAI scores ranged from 5-9; group 3: SLEDAI scores ranged from 10-14; group 4: SLEDAI scores were>14). One-way ANOVA analysis and pairwise t-test among four groups were performed and the results demonstrated that anti-dsDNA IgG1 Fuc, IgG1 Gal, and IgG1 Sia increased with the trend of SLE disease activity ($p<0.05$) (Figure 3), while the rest of the glycoform showed no difference (including anti-dsDNA IgG2, anti-dsDNA IgG3/4, total IgG1, total IgG2 and total IgG3/4) (all $P > 0.05$) (Supplementary Figure S1-S2). Besides, we also analysed the correlation between glycoform of anti-dsDNA IgG and protein-uria (Supplementary Figure S3).

	Anti-dsDNA IgG (median (quartile25%, 75%)) or N (%)	Total IgG (median (quartile25%, 75%)) or N (%)	P value
N	86	101	
Age	32 (24, 46.75)	32 (24, 47)	0.92
Gender (female, N (%))	76 (88.37)	89 (88.12)	0.96
Anti-dsDNA antibody	707.3 (226.2, 1527.9)	693.4 (226.6, 1527.9)	0.74
Duration (month)	3 (1, 12)	3 (1, 12)	0.88
CRP (mg/L)	0.355 (0.16, 1.18)	0.330 (0.14, 1.14)	0.80
ESR (mm/h)	47 (24.5, 80)	46 (26, 80)	0.99
SLEDAI	11 (8.25, 17)	12 (9, 17)	0.71
Fever, N (%)	33 (38.37)	40 (39.60)	0.98
Pericarditis, N (%)	11(12.79)	13 (12.87)	0.99
Pleuritis, N (%)	10 (11.62)	11 (10.89)	0.87
Mouth ulcers, N (%)	14 (16.27)	18 (17.82)	0.93
Rash, N (%)	37 (43.02)	45 (44.55)	0.95
Arthritis, N (%)	58 (67.44)	70 (69.30)	0.90
Vasculitis, N (%)	12 (13.95)	15 (14.85)	0.86
24 hour urine protein (mg)	264.50 (122, 688.75)	276.50 (124, 801.60)	0.69
C3 (g/L)	45 (33, 64)	46 (35, 64)	0.77
C4 (g/L)	9 (3.25, 13)	9 (4, 13)	0.92
White blood cell (*10^9/L)	2.99 (2.32, 4.41)	3.0 (2.40, 4.50)	0.84
Hemoglobin (mg/L)	105 (87.25, 119)	106 (90, 119)	0.83
Platelet (*10^9/L)	132 (90.50, 179)	138 (90, 188)	0.96

Table 1: **Clinical characteristics of anti-dsDNA IgG and total IgG in treatment-naïve SLE patients.**
Abbreviations: N: number; CRP: C-reactive protein; ESR: erythrocyte sedimentation rate; IgG: immunoglobulin; SLEDAI: systemic lupus erythematosus disease activity index; C: complement.

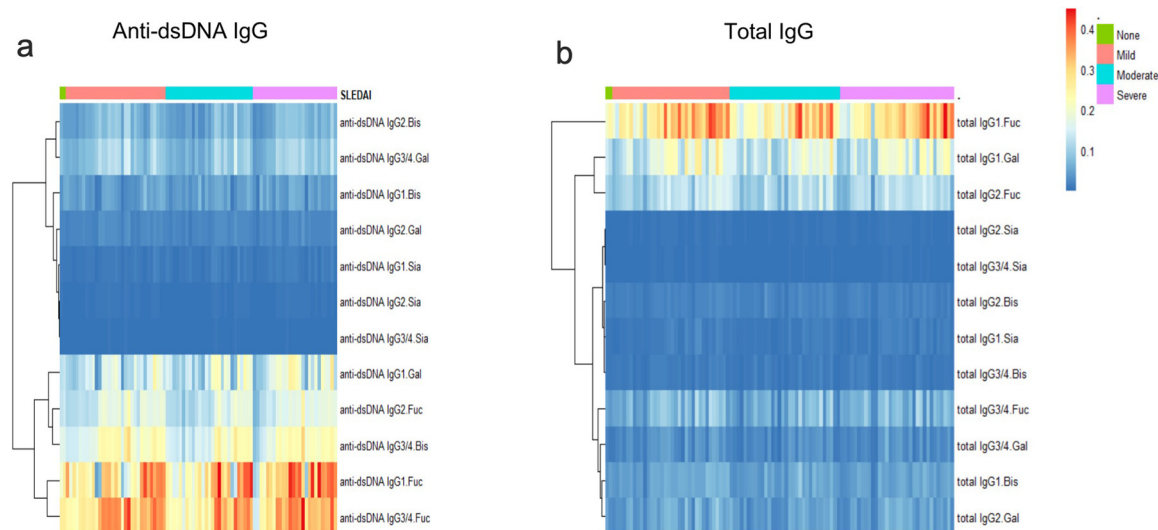|	Anti-dsDNA IgG			Total IgG			
	N	Mean	SD	N	Mean	SD	P value
Fuc							
IgG1	86	0.272	0.094	101	0.277	0.075	0.034
IgG2	86	0.148	0.035	101	0.113	0.028	<0.001
IgG3/4	86	0.295	0.065	101	0.065	0.028	<0.001
Gal							
IgG1	86	0.141	0.054	101	0.158	0.051	0.022
IgG2	86	0.023	0.005	101	0.046	0.018	<0.001
IgG3/4	86	0.087	0.026	101	0.028	0.014	<0.001
Bis							
IgG1	86	0.044	0.014	101	0.049	0.014	0.013
IgG2	86	0.064	0.023	101	0.018	0.005	<0.001
IgG3/4	86	0.191	0.046	101	0.011	0.005	<0.001
Sia							
IgG1	86	0.010	0.005	101	0.011	0.006	0.068
IgG2	86	0.005	0.002	101	0.005	0.002	0.019
IgG3/4	86	0.003	0.001	101	0.003	0.002	0.006

Table 2: **The comparison of glycosylation between anti-dsDNA IgG and total IgG in treatment-naïve SLE patients.**
Abbreviations: Fuc: fucosylation; Gal: galctosylation; Bis: bisecting GlcNAc; Sia: sialylation; N: number; SD: standard deviation; IgG: immunoglobulin.

Machine learning of anti-dsDNA IgG and total IgG

SLEDAI score was based on the summary of several items, which patients fit in, but the weight of each item was inconsistent. For example, the same SLEDAI score could be calculated in a sum up of different items in a variety of ways. Thus, a high-weight item or several low-weight items might result in the same score. So linear correlation

Figure 2. Cluster analysis of glycosylation of anti-dsDNA IgG (a) and total IgG (b) in SLE patients with different SLEDAI scores. The glycosylation level of IgG was shown in blue colour as low level and red colour as high level. Colour depth increased with the concentration of glycosylation level. A similar level of glycosylation was clustered together and one stripe represented one patient. The patients were presented from left to right with increasing SLEDAI score categories (green colour representing not active, orange colour mildly active, blue colour moderately active and purple colour severely active).
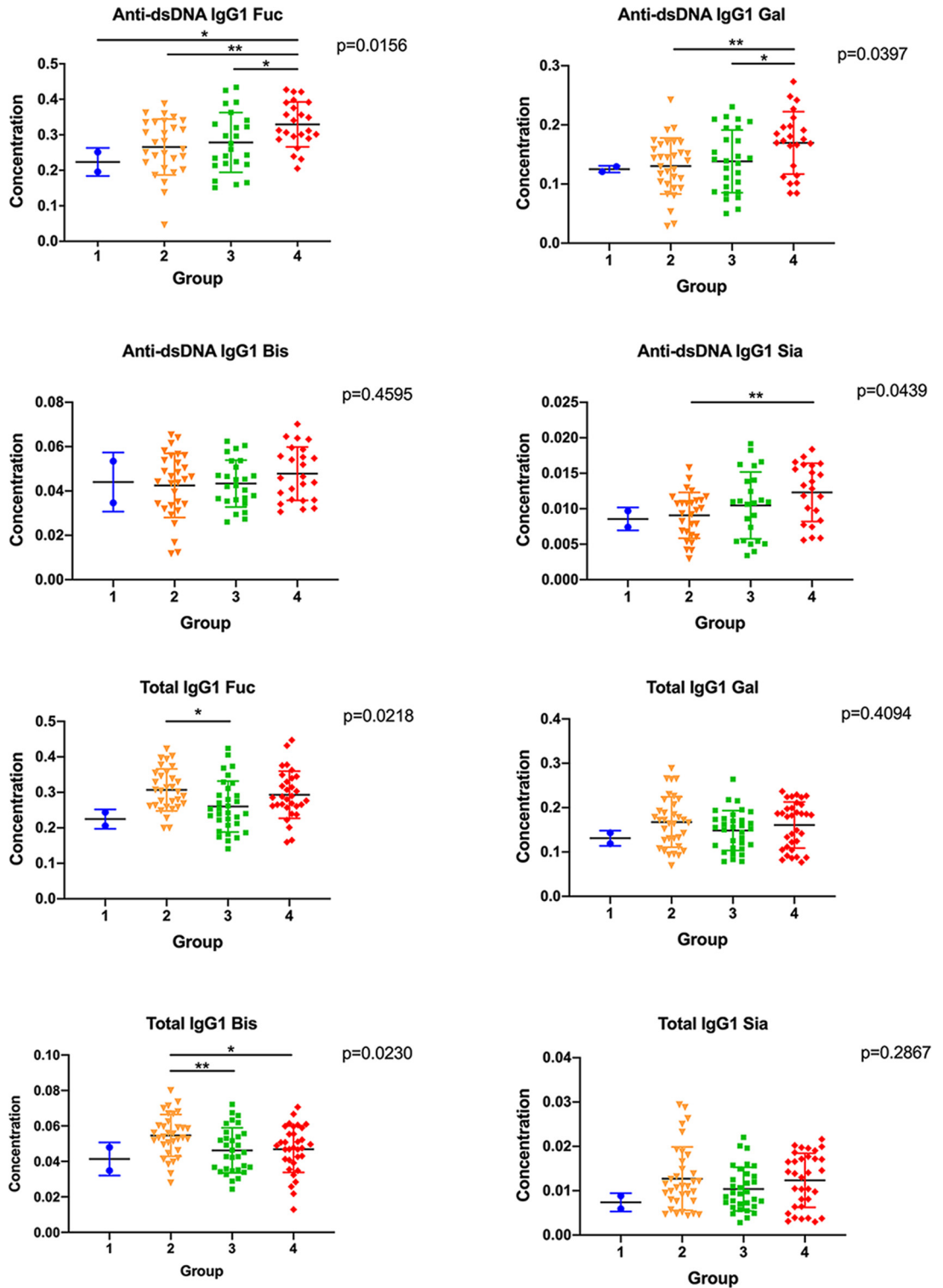
Abbreviations: IgG: Immunoglobulin; SLEDAI: systemic lupus erythematosus disease activity index; SLE: Systemic Lupus Erythematosus; Fuc: fucosylation; Gal: galctosylation; Bis: bisecting GlcNAc; Sia: sialylation.

analysis was not appropriate for analyzing the association between SLEDAI and glycoform concentrations. To remove the influence of weight, we took the items in the SLEDAI rating scale as the training objectives of our machine learning to establish a predictive model. Besides, factors that might affect the patient's condition (such as age, gender, disease duration) were also included in the list of training objectives. In the predictive model, we could acquire a predicted glycoform concentration value (predicted value). SSE was used to compare the difference between the predicted value and the true value (data in our study). We found that all the glycoforms of anti-dsDNA IgG1 showed better performance with lower SSE (Figure 4, Supplementary Table S5) than that of total IgG1, which also suggested that anti-dsDNA IgG1 were more relevant to the SLEDAI than total IgG1. Among all the subclasses of IgG1, fucosylation had the lowest SSE (SSE=0.009). Consistent with previous results (Figure 2, 3), it was further verified that anti-dsDNA IgG1 Fuc had the best performance correlated with disease activity of SLE. Besides, anti-dsDNA IgG2, anti-dsDNA IgG3/4 also performed better than that of total IgG2, total IgG3/4 (Supplementary Figure S4 and S5). These results also indicated that anti-dsDNA IgG glycosylation correlated better with disease activity of SLE than total IgG.

Discussion

The role of glycosylation of IgG has long been discussed. In autoimmune diseases, glycosylation of IgG in RA was widely studied, in which ACPA-IgG harboured N-glycans in their domains.[21] ACPA was significantly changed in Fc galactosylation and fucosylation before the onset of RA, which meant a more pro-inflammatory phenotype.[22] Besides, sialylated IgG reduced the phagocytosis by macrophages and switched the cytokine secretion from interleukin (IL)-6/IL-8 to tumor necrosis factor (TNF)-α/IL-1β.[23] Despite RA, the role of glycosylation of IgG in Antiphospholipid Syndrome (APS) was also reported. A significantly lower sialylation of IgG against β2GP1 of patients with APS was observed when compared to IgG of asymptomatic carriers.[24] Moreover, a large scale, multi-institute study revealed that IgG Gal-ratio could distinguish 12 types of cancers from non-cancer controls.[25] These studies indicated that glycosylation of IgG could be used as a biomarker in different diseases.

Glycosylation in SLE has also acquired much attention in recent years. Glycosylation of H2A resulted in the generation of neo-epitopes on H2A histone, which were preferably bound by anti-DNA autoantibodies, which implied that deoxyribose-modified-H2A might trigger immune response resulting in the generation of anti-glycated H2A antibodies with DNA cross-reacting properties.[26] Moreover, it was reported that lectins were

Figure 3. SLE patients with positive anti-dsDNA IgG were divided into four groups according to SLEDAI scores (group 1: SLEDAI scores ranged from 0-4 (n=2); group 2: SLEDAI scores ranged from 4-9 (n=31); group 3: SLEDAI scores ranged from 9-14 (n=27); group 4: SLEDAI scores were>14 (n=26)). (a)-(h) suggested four subtype glycoform concentrations of IgG1 in anti-dsDNA IgG and total IgG, respectively. (a), (b), (d) showed the concentration of fucosylation, galctosylation, sialylation levels of anti-dsDNA IgG1 by SLEDAI score category, respectively (all $P <0.05$). *$P <0.05$, **$P <0.01$.

Note: IgG: Immunoglobulin; SLEDAI: systemic lupus erythematosus disease activity index; SLE: Systemic Lupus Erythematosus; Fuc: fucosylation; Gal: galctosylation; Bis: bisecting GlcNAc; Sia: sialylation.
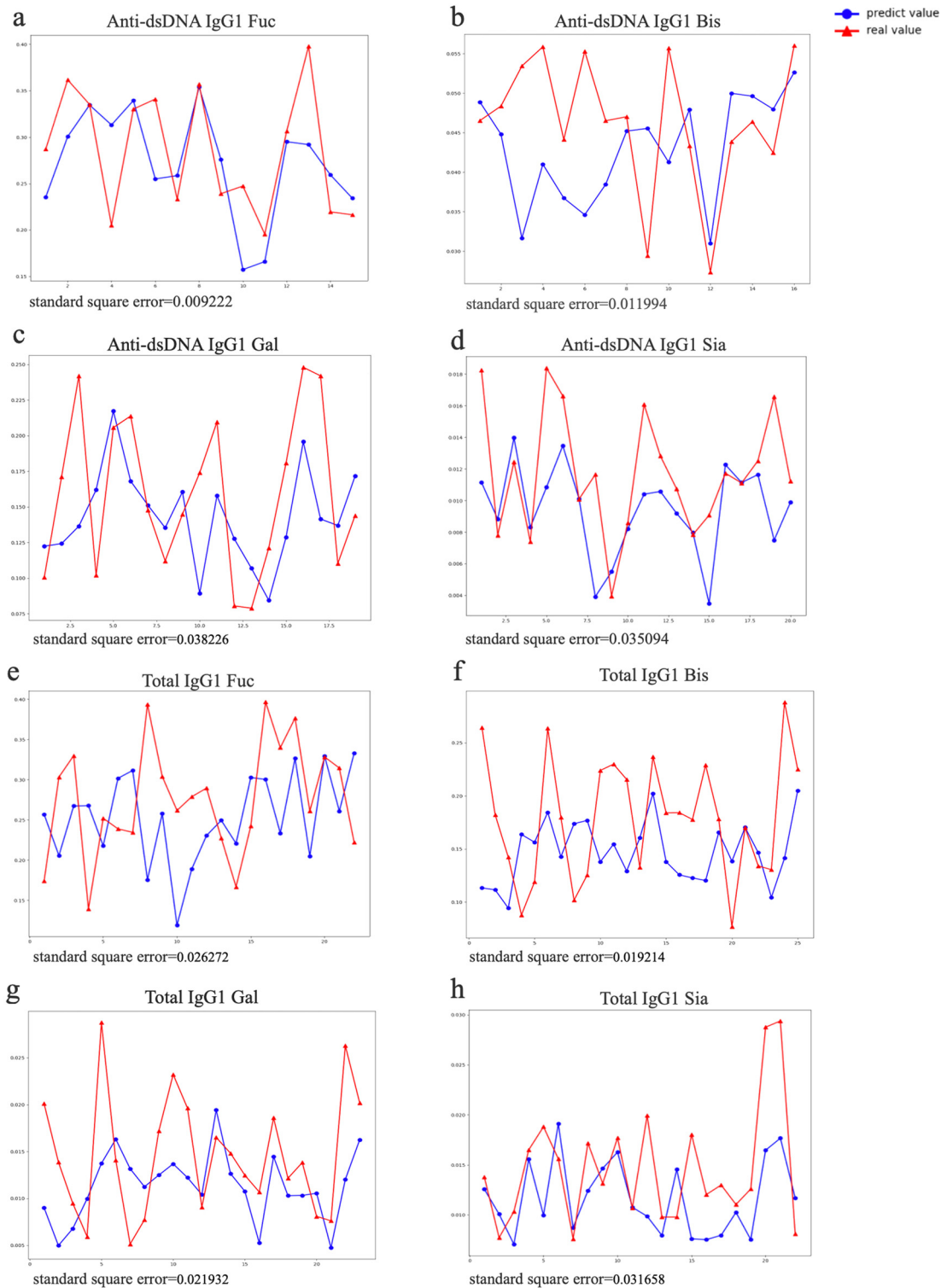
Figure 4. Machine learning model of anti-dsDNA IgG and total IgG. (a)-(h) showed the predicted glycoform value (predicted value) and true value (data from our study) of anti-dsDNA IgG1 and total IgG1. The blue line represented the predicted value, and the red line represented the real value. SSE of each sample was used to compare the difference between the predicted value and the true value. (a)-(d) suggested that anti-dsDNA IgG1 had a better performance than that of total IgG1 with smaller SSE.

Abbreviations: IgG: Immunoglobulin; SLEDAI: systemic lupus erythematosus disease activity index; SLE: Systemic Lupus Erythematosus; SSE: Standardised square error.

specific for the disaccharide Gal-GalNAc, such as amaranthus leucocarpus lectin (ALL), which could be considered as a marker to determine the activity of the disease.[27] The previous study only compared the glycosylation changes in total IgG of SLE with healthy controls.[27] In this study, we compared the association of glycosylation of specific purified anti-dsDNA IgG and total IgG for the first time, including their subclasses with disease activity of SLE.

Glycosylation was synthesized in the endoplasmatic reticulum. However, the complex processing of the glycan occurs mainly in the Golgi apparatus.[28] The heavy-chain locus encodes multiple constant regions, including four IgG isotypes arranged in the following order: IgG3, IgG1, IgG2, and then IgG4. IgG1 is the main subclass consisting of 60% of the serum IgG, with IgG2 at 32% and IgG3 and Ig4 at 4%.[29] The addition and removal of the variably added sugars (galactose, fucose, β-GlcNAc, or sialic acid) have been linked to altered antibody functionality. In this study, cluster analysis showed that anti-dsDNA IgG and total IgG had different glycoform levels. However, fucosylation of both anti-dsDNA IgG1 and total IgG1 possessed higher levels than other subclasses. This may be due to the reason that IgG1 is the main proportion of serum IgG.[29] Furthermore, we used supervised GBDT and unsupervised machine learning (cluster analysis) methods to analyse the potential relationship between glycosylation and disease activity, which showed that fucosylation of anti-dsDNA IgG1 was better than that of total IgG1 in reflecting the disease activity of SLE.

Fucosylated glycan structures occur commonly on cell surfaces and play an important role in a variety of biological and pathological processes in eukaryotic organisms, including tissue development, cell adhesion, infection, angiogenesis, and tumour metastasis.[30-32] Fucosylation of CD4+ T cells was significantly increased in SLE.[33] T cell receptor complex are highly core-fucosylated glycoproteins, which play an important role in T cell activation.[33] Core fucosylation is likely to be important in three stages involved in the T cell activation. Firstly, it is essential for the T cell receptor (TCR) structural formation. Secondly, core fucosylation of TCR could regulate the recognition of peptide-major histocompatibility complex (pMHC) and affect the T cell activation threshold. Thirdly, fucose-specific lectins,[34] might participate in the events in the T-B cell interaction. However, altered core fucosylation's role in the SLE remains unclear.[33] More studies are needed to elaborate on it.

There were several limitations in our study. Firstly, the number of patients was relatively small. The accuracy of the prediction model would be increased with the number of samples. Secondly, the number of patients in the SLEDAI score 0-4 was also small. More patients are needed to increase the accuracy of the analysis. Thirdly, we did not collect blood samples during

the follow-up of these patients and measure the fluctuation of glycosylation as the disease progresses. More samples could be included, and additional follow-up information could be added in our future study. Fourthly, it was a single-centre study. Multi-centre studies could be considered in the future.

In conclusion, our study indicated that glycosylation of anti-dsDNA IgG was different from that of total IgG and fucosylation of anti-dsDNA IgG1 correlated best with SLE disease activity.

Declaration of interests
None declared.

Contributors
JT, SR, and JY designed the study. FW, Y Sun, H-L L, XC, Y Su, HS and QH were involved in sample preparation and processing. ZG, J Huang and YY conducted statistical analyses and interpreted the results. J Han, ZZ, and RZ drafted the manuscript. CY contributed to the discussion and revised the manuscript. All authors reviewed the manuscript and gave their approval for submission. ZZ, YY, and JY have verified the underlying data.

Data Sharing Statement
The main data supporting the findings of this study are available within the manuscript and its Supplementary materials. All data generated, both raw images and analysed datasets, for the figures in this study were uploaded with the manuscript.

Supplementary materials
Supplementary material associated with this article can be found in the online version at doi:10.1016/j. ebiom.2022.103883.

References
1 Wang X, Xia Y. Anti-double Stranded DNA Antibodies: Origin, Pathogenicity, and Targeted Therapies. *Front Immunol.* 2019;10:1667.
2 Gheita TA, Abaza NM, Hammam N, Mohamed AAA, El II G, Eissa AH. Anti-dsDNA titre in female systemic lupus erythematosus patients: relation to disease manifestations, damage and antiphospholipid antibodies. *Lupus.* 2018;27(7):1081–1087.
3 Conti F, Ceccarelli F, Perricone C, et al. Systemic Lupus Erythematosus with and without Anti-dsDNA Antibodies: Analysis from a Large Monocentric Cohort. *Mediators Inflamm.* 2015;2015: 328078.
4 Andrejevic S, Jeremic I, Sefik-Bukilica M, Nikolic M, Stojimirovic B, Bonaci-Nikolic B. Immunoserological parameters in SLE: high-avidity anti-dsDNA detected by ELISA are the most closely

associated with the disease activity. *Clin Rheumatol.* 2013;32 (11):1619–1626.

5 Gladman DD, Ibanez D, Urowitz MB. Systemic lupus erythematosus disease activity index 2000. *J Rheumatol.* 2002;29(2):288–291.

6 Vasudevan V, Agnihotri P, Biswas S. Post Translational Modification and Its Pathologic Association in Rheumatoid Arthritis: A Brief Perspective. *Curr Protein Pept Sci.* 2021;22(7):548–558.

7 Lande R, Pietraforte I, Mennella A, et al. Complementary Effects of Carbamylated and Citrullinated LL37 in Autoimmunity and Inflammation in Systemic Lupus Erythematosus. *Int J Mol Sci.* 2021;22(4).

8 Anunciacion-Llunell A, Miro-Mur F, Esteve-Valverde E, Marques-Soares J, Pardos-Gea J, Alijotas-Reig J. Proteomics and enriched biological processes in Antiphospholipid syndrome: A systematic review. *Autoimmun Rev.* 2021;20:(12) 102982.

9 Jefferis R. Glycosylation as a strategy to improve antibody-based therapeutics. *Nat Rev Drug Discov.* 2009;8(3):226–234.

10 Bondt A, Hafkenscheid L, Falck D, et al. ACPA IgG galactosylation associates with disease activity in pregnant patients with rheumatoid arthritis. *Ann Rheum Dis.* 2018;77(8):1130–1136.

11 Vuckovic F, Kristic J, Gudelj I, et al. Association of systemic lupus erythematosus with decreased immunosuppressive potential of the IgG glycome. *Arthritis Rheumatol.* 2015;67(11):2978–2989.

12 Sjowall C, Zapf J, von Lohneysen S, et al. Altered glycosylation of complexed native IgG molecules is associated with disease activity of systemic lupus erythematosus. *Lupus.* 2015;24(6):569–581.

13 Chui D, Sellakumar G, Green R, et al. Genetic remodeling of protein glycosylation in vivo induces autoimmune disease. *Proc Natl Acad Sci U S A.*. 2001;98(3):1142–1147.

14 Aringer M, Costenbader K, Daikh D, et al. 2019 European League Against Rheumatism/American College of Rheumatology classification criteria for systemic lupus erythematosus. *Ann Rheum Dis.* 2019;78(9):1151–1159.

15 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(10). M111 010090.

16 Zhang W, Reichlin M. A peptide DNA surrogate that binds and inhibits anti-dsDNA antibodies. *Clin Immunol.* 2005;117(3):214–220.

17 Sanda M, Pompach P, Brnakova Z, Wu J, Makambi K, Goldman R. Quantitative liquid chromatography-mass spectrometry-multiple reaction monitoring (LC-MS-MRM) analysis of site-specific glycoforms of haptoglobin in liver disease. *Mol Cell Proteomics.* 2013;12 (5):1294–1305.

18 Yuan W, Sanda M, Wu J, Koomen J, Goldman R. Quantitative analysis of immunoglobulin subclasses and subclass specific glycosylation by LC-MS-MRM in liver disease. *J Proteomics.* 2015;116:24–33.

19 Alvarez-Manilla Jr GA, Guo Y, Warren NL, Orlando R, Pierce M. Tools for glycoproteomic analysis: size exclusion chromatography facilitates identification of tryptic glycopeptides with N-linked glycosylation sites. *J Proteome Res.* 2006 Mar;5(3):701–708.

20 Hong Q, Lebrilla CB, Miyamoto S, Ruhaak LR. Absolute quantitation of immunoglobulin G and its glycoforms using multiple reaction monitoring. *Anal Chem.* 2013;85(18):8585–8593.

21 Rombouts Y, Willemze A, van Beers JJ, et al. Extensive glycosylation of ACPA-IgG variable domains modulates binding to citrullinated antigens in rheumatoid arthritis. *Ann Rheum Dis.* 2016;75 (3):578–585.

22 Rombouts Y, Ewing E, van de Stadt LA, et al. Anti-citrullinated protein antibodies acquire a pro-inflammatory Fc glycosylation phenotype prior to the onset of rheumatoid arthritis. *Ann Rheum Dis.* 2015;74(1):234–241.

23 Magorivska I, Munoz LE, Janko C, et al. Sialylation of anti-histone immunoglobulin G autoantibodies determines their capabilities to participate in the clearance of late apoptotic cells. *Clin Exp Immunol.* 2016;184(1):110–117.

24 Fickentscher C, Magorivska I, Janko C, et al. The Pathogenicity of Anti-beta2GP1-IgG Autoantibodies Depends on Fc Glycosylation. *J Immunol Res.* 2015;2015: 638129.

25 Ren S, Zhang Z, Xu C, et al. Distribution of IgG galactosylation as a promising biomarker for cancer screening in multiple cancer types. *Cell Res.* 2016;26(8):963–966.

26 Alam S, Arif Z, Alam K. Glycated-H2A histone is better bound by serum anti-DNA autoantibodies in SLE patients: glycated-histones as likely trigger for SLE? *Autoimmunity.* 2015;48(1):19–28.

27 Ramos-Martinez E, Lascurain R, Tenorio EP, et al. Differential Expression of O-Glycans in CD4(+) T Lymphocytes from Patients with Systemic Lupus Erythematosus. *Tohoku J Exp Med.* 2016;240 (1):79–89.

28 Jennewein MF, Alter G. The Immunoregulatory Roles of Antibody Glycosylation. *Trends Immunol.* 2017;38(5):358–372.

29 Vidarsson G, Dekkers G, Rispens T. IgG subclasses and allotypes: from structure to effector functions. *Front Immunol.* 2014;5:520.

30 Miyoshi E, Moriwaki K, Nakagawa T. Biological function of fucosylation in cancer biology. *J Biochem.* 2008;143(6):725–729.

31 Ma B, Simala-Grant JL, Taylor DE. Fucosylation in prokaryotes and eukaryotes. *Glycobiology.* 2006;16(12):158R.. -84R.

32 Li J, Hsu HC, Mountz JD, Allen JG. Unmasking Fucosylation: from Cell Adhesion to Immune System Regulation and Diseases. *Cell Chem Biol.* 2018;25(5):499–512.

33 Liang W, Mao S, Sun S, et al. Core Fucosylation of the T Cell Receptor Is Required for T Cell Activation. *Front Immunol.* 2018;9:78.

34 Lehrman MA, Hill RL. The binding of fucose-containing glycoproteins by hepatic lectins. Purification of a fucose-binding lectin from rat liver. *J Biol Chem.* 1986;261(16):7419–7425.