





## ORIGINAL ARTICLE

# Validation of diagnostic and predictive biomarkers for hereditary angioedema via plasma *N*-glycomics

Zejian Zhang<sup>1</sup>  | Xue Wang<sup>2</sup>  | Jianqing Gu<sup>2</sup> | Jianqiang Wu<sup>1</sup> | Yang Cao<sup>2</sup>  | Yingyang Xu<sup>2</sup>  | Lisha Li<sup>2</sup> | Kai Guan<sup>2</sup> | Peng Liu<sup>1</sup> | Jia Yin<sup>2</sup> | Yuxiang Zhi<sup>2</sup> | Shuyang Zhang<sup>3</sup>

<sup>1</sup>Department of Medical Research Center, State Key Laboratory of Complex Severe and Rare Diseases, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China

<sup>2</sup>Department of Allergy & Clinical Immunology, National Clinical Research Center for Immunologic Diseases, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China

<sup>3</sup>Department of Cardiology, State Key Laboratory of Complex Severe and Rare Diseases, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China

## Correspondence

Yuxiang Zhi, Department of Allergy & Clinical Immunology, National Clinical Research Center for Immunologic Diseases, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, No. 1 Shuaifuyuan Wangfujing, Dongcheng District, Beijing 100730, China.  
Email: [yuxiang\\_zhi@126.com](mailto:yuxiang_zhi@126.com)

Shuyang Zhang, Department of Cardiology, State Key Laboratory of Complex Severe and Rare Diseases, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, No. 1 Shuaifuyuan Wangfujing, Dongcheng District, Beijing, 100730, China.  
Email: [shuyangzhang103@nrdrs.org](mailto:shuyangzhang103@nrdrs.org)

## Funding information

National Natural Science Foundation of China, Grant/Award Numbers: 81472870, 31901041; National Key Research and Development Program of China, Grant/Award Number: 2016YFC0901501; the CAMS Innovation Fund for Medical Science (CIFMS), Grant/Award Number: CIFMS 2021-I2M-1-003

## Abstract

**Background:** Hereditary angioedema (HAE) is a rare disease with heterogeneous clinical symptoms. It is vitally important to predict whether an HAE patient will develop severe symptoms in clinical practice, but there are currently no predictive biomarkers for HAE stratification. Plasma *N*-glycomes are disease-specific and have great potential for the discovery of non-invasive biomarkers. In this study, we profiled the plasma *N*-glycome of HAE patients from two independent cohorts to identify candidate biomarkers.

**Methods:** Linkage-specific sialylation derivatization combined with matrix-assisted laser desorption/ionization time-of-flight mass spectrometry detection and automated data processing was employed to analyze the plasma *N*-glycome of two independent type-1 HAE cohorts.

**Results:** HAE patients had abnormal glycan complexity, galactosylation, and  $\alpha$ 2,3- and  $\alpha$ 2,6-linked sialylation compared to healthy controls (HC). The classification models based on dysregulated glycan traits could successfully discriminate between HAE and HC with area under the curves (AUCs) being greater than 0.9. Some of the aberrant glycans showed response to therapy. Moreover, we identified a series of glycan traits with strong associations with the occurrence of laryngeal or gastrointestinal angioedema or disease severity score. Predictive models based on these traits could be used to predict disease severity (AUC > 0.9). These results were replicated in an independent cohort.

**Conclusions:** We reported the full plasma *N*-glycomic signature of HAE for the first time, and identified potential biomarkers. These findings may play a critical role in

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2021 The Authors. Clinical and Translational Allergy published by John Wiley & Sons Ltd on behalf of European Academy of Allergy and Clinical Immunology.

predicting disease severity and guide the treatment of HAE in clinical practice. Further protein-specific and prospective studies are needed to validate our findings.

**KEYWORDS**

biomarker, disease severity, galactosylation, plasma *N*-glycome, sialylation

**SCHLÜSSELWÖRTER**

biomarker, galaktosylierung, plasma-*N*-glykom, schwere der erkrankung, sialylierung

## 1 | INTRODUCTION

Hereditary angioedema (HAE) is a rare autosomal dominant disease that is primarily characterized by unpredictable and potentially life-threatening subcutaneous or submucosal edema, which can affect the extremities, face, larynx, abdomen, or genitalia, with an incidence of 1/50,000 to 1/100,000.<sup>1–3</sup> Typically, HAE is caused by C1 inhibitor (C1-INH) deficiency (reduced level or abnormal function), and mutations in the *SERPING1* gene.<sup>1,3</sup> HAE attacks can be spontaneous or provoked by emotional fluctuation, menstruation, fever, and trauma.<sup>1,2</sup> Currently, HAE is clinically diagnosed by combining family history, clinical symptoms, and laboratory tests (primarily complement tests).<sup>1</sup> As HAE is a rare disease, and sometimes presents similar symptoms to other diseases, HAE patients are often misdiagnosed and treated incorrectly. For example, gastrointestinal angioedema episodes present clinical symptoms similar to those of acute intestinal obstruction, which often leads to unnecessary surgical intervention.<sup>3</sup> Delaying diagnosis and treatment can have fatal consequences for HAE patients. Laryngeal angioedema attacks cause the obstruction of the upper airways, for which ineffective treatments (antihistamines, adrenaline, and corticosteroids) are often misused, and suffocation may ensue. However, specific target therapies, such as bradykinin B2 receptor antagonists, C1 inhibitor concentrate, kallikrein inhibitor, and sometimes tracheotomy, can prevent death in HAE patients.<sup>4</sup> The situation is complicated by the heterogeneity of HAE clinical manifestations, even in individuals from the same family. Management strategies for different HAE patients differ greatly and may be adjusted based on the variable frequency and types of HAE attacks experienced by individual patients. However, there are currently no available methods to stratify HAE patients. Currently, danazol is the primary treatment option for short- and long-term prophylaxis in China. However, danazol is associated with serious side effects, such as virilization, sequelae of obesity, and increased hepatic enzyme levels. Therefore, it is vital to develop tools used for patient stratification to guide individualized disease management.

*N*-glycosylation is a common and functionally relevant post-translational modification involving the synergistic action of multiple enzymes and transporters.<sup>5–7</sup> Glycosylation is not template-driven, but rather introduces variability in proteins that is independent of their corresponding DNA sequences.<sup>7</sup> *N*-glycans are highly diverse, and regulate multiple biological processes, from

protein folding and stability to receptor-ligand interactions and immune responses.<sup>8</sup> Specific glycoforms depend on genetic, pathophysiological, and environmental factors.<sup>9</sup> Plasma or serum protein *N*-glycosylation in a given physiological state is stable, but can dramatically change in response to pathological conditions. Specific glycosylation features of the plasma *N*-glycome are related to various pathological conditions, such as Down syndrome,<sup>10</sup> rheumatoid arthritis,<sup>11</sup> inflammatory bowel disease,<sup>12</sup> and cancers.<sup>13,14</sup> Importantly, the plasma *N*-glycome is often disease-specific. The plasma *N*-glycome may represent a non-invasive biomarker, and provide a more thorough understanding of specific disease mechanisms. Currently, little is known regarding *N*-glycosylation in HAE.

In this study, we investigated the plasma protein *N*-glycomic features in HAE patients, HAE patients after treatment, and healthy controls (HCs). Thereafter, we validated our findings in an independent cohort. As the function of sialylation depends on the linkage type, we employed the linkage-specific sialic acid derivatization method developed by Reiding et al.<sup>15</sup> to distinguish between the  $\alpha$ 2,3- and  $\alpha$ 2,6-linked sialic acids on the non-reducing end of plasma glycans. We used matrix-assisted laser desorption/ionization time-of-flight (TOF) mass spectrometry (MALDI-TOF MS) to profile the *N*-glycome of the cohort.<sup>15</sup> Our primary objective was to reveal the plasma *N*-glycomic features that are HAE-specific, and to discover and validate putative glycan biomarkers for HAE diagnosis, stratification, and monitoring.

## 2 | METHODS

### 2.1 | Study design and participants

In the present study, the discovery cohort enrolled 21 drug-naïve type-1 HAE patients, 12 type-1 HAE patients after treatment, and 50 HCs. The independent validation cohort consisted of 18 drug-naïve type-1 HAE patients, 30 type-1 HAE patients after treatment, and 40 HCs. The plasma of each individual from the two cohorts was collected from the Peking Union Medical College Hospital (Beijing, China). The age and sex of the three subgroups in each cohort were matched where possible. Plasma samples were stored at  $-80^{\circ}\text{C}$ . HAE patients were diagnosed on the basis of medical history, clinical symptoms, and clinical parameters from the

laboratory: (1) a history of recurrent episodes of subcutaneous and/or submucosal edema (such as edema affecting the extremities, face, larynx, abdomen, or genitalia) and (2) confirmed lower levels of clinical parameters, including C4 and C1-INH antigen. This study was approved by the regional ethics committee of Peking Union Medical College Hospital (No. HS-2402). All study participants gave written informed consent.

## 2.2 | Plasma *N*-glycome detection and MS data (pre-)processing

Plasma samples were enzymatically treated to release *N*-glycans according to a previously reported method.<sup>15</sup> Briefly, 10  $\mu$ L of 2% SDS was added to 5  $\mu$ L of plasma, and the mixture was incubated for 15 min at 65°C. The *N*-glycans were then released by the addition of 10  $\mu$ L of release mixture (1 U PNGase F, 2% NP-40, and 2.5  $\times$  PBS), followed by overnight incubation at 37°C. To derivatize sialic acids,  $\alpha$ 2,3-linked sialic acids were lactonized, and  $\alpha$ 2,6-linked sialic acids were ethyl-esterified, allowing mass-based distinction between the sialic-acid linkage variants.<sup>15</sup> Released *N*-glycans were then enriched by hydrophilic interaction liquid chromatography solid-phase extraction (HILIC-SPE) micro-tips using cotton thread as the stationary phase packed in the pipette tips and enriched glycans were eluted with Milli-Q water according to the method described previously.<sup>13,14</sup> HILIC-SPE cotton-tips allowed the removal of salts, most deglycosylated peptides, and detergents from glycoconjugate samples. In addition, subsequent MALDI-TOF MS glycan profiles were very repeatable with different tips.<sup>15</sup> The sialylated *N*-glycans were detected simultaneously with non-sialylated *N*-glycans using MALDI-TOF MS in positive-ion mode as previously described, with minor modification.<sup>14</sup> Briefly, 1  $\mu$ L of purified glycans was mixed with 1  $\mu$ L of matrix (5 mg/mL super-2,5-dihydroxybenzoic acid with 1 mM NaOH in 50% acetonitrile) on the target plate. Mass spectra were obtained using a rapiflexXtreme MALDI-TOF MS (Bruker Daltonics). The instrument was equipped with a Smartbeam-3D laser and was controlled using flexControl 4.0 (Bruker Daltonics). The laser was fired 5000 times per spot in a random walking pattern at a frequency of 5000 Hz. The mass range was set to 1000–5000 *m/z*. The instrument was calibrated using external calibrants (Bruker Peptide Calibration Standard II).

Raw MS data was baseline-subtracted and smoothed. The MS data was transformed to .xy files and re-calibrated using selected glycan signals as calibrants (Table S1), using the in-house developed software by Jansen et al. (MassyTools, version 0.1.8.1.2).<sup>16</sup> Thereafter, summed spectra were generated for each biological group (untreated HAE, treated HAE, and HC) and the quality control group (randomly distributed plasma standard). For each summed spectrum, mono-isotopic peaks with good signal to noise (*S/N*; >3), good relative intensity (>0.1%), and good isotopic patterns were filtered for further analysis, and 90 mono-isotopic peaks were assigned to *N*-glycan structures using Glycoworkbench<sup>17</sup> as well as previously confirmed glycan compositions.<sup>15</sup> Finally, an *N*-glycan

composition list was generated for use in subsequent targeted data extraction. The peak intensities of the putative *N*-glycans were extracted as peak area (background-corrected) for all samples using the *N*-glycan composition list and MassyTools. Further processing of the extracted data was done in Microsoft Excel. Glycan structures were excluded after applying cut-offs for *S/N* (>9), mass accuracy (ppm error < |20|), isotopic pattern quality (QC score < 25%), and the minimum percentage (>50%) of presence in all spectra of HAE, HC, or quality control plasma samples.<sup>14</sup> After curation, 59/90 *N*-glycans (Table S2) were included for quantitative analysis. The sum of the areas of these remaining 59 compositions per spectrum was normalized to one. To combine the exact effects of individual glycans sharing similar glycan structures, and to enable interpretation of their biological effects, 82 derived *N*-glycan traits were calculated from the directly detected *N*-glycan traits based on their common structural characteristics (Table S3). The formulas used for the calculation of the derived glycan traits are given in Table S3, and calculations were performed in RStudio. The subject of the calculation is represented by the last letter, for example, sialylation (S), and the group on which it is calculated by the preceding letters, for example, triantennary fucosylated species (A3F). This, for instance, translates A3FS into the sialylation within triantennary fucosylated species.<sup>14,15,18</sup> Differential derived *N*-glycan traits demonstrate that changes to glycosylation are shared by a series of structurally related *N*-glycans.<sup>14</sup>

## 2.3 | Statistical analysis

Comparisons were made for derived *N*-glycan traits between two subgroups in each cohort (untreated HAE vs. HC, treated HAE vs. HC) using the Mann-Whitney-Wilcoxon test (because the data was non-normally distributed). Multiple testing correction was performed to adjust the significance threshold ( $p = 0.05/82$ ). Regression analysis was performed in RStudio. The diagnostic/predictive potential of the individual *N*-glycan traits was further evaluated based on receiver operator characteristics (ROC). Classification/prediction biomarker models were constructed using multivariate algorithms for support vector machines (SVMs) based on the differentially expressed derived *N*-glycan traits or glycan traits strongly associated with clinical symptoms. ROC curves were obtained by Monte Carlo cross validation (MCCV). In each MCCV, two-thirds of the samples were used to assess the importance of each glycan trait, and the remaining one-third was used to validate the biomarker models generated in the first step. The top-ranking important glycan traits were subsequently used to construct predictive biomarker models. These steps were repeated, and the performance of each model was calculated and compared. The area under the curve (AUC) of ROC curves and predictive accuracy were used to assess the performance of the output models. We considered  $AUCs \geq 0.9$  to represent highly accurate tests, whereas  $0.8 \leq AUCs < 0.9$  represented accurate tests, and  $0.7 \leq AUCs < 0.8$  represented moderately accurate tests.

TABLE 1 Clinical characteristics of all participants arranged by subgroups

	Discovery cohort			Validation cohort		
	Untreated HAE	Treated HAE	HC	Untreated HAE	Treated HAE	HC
Sample size	<i>n</i> = 21	<i>n</i> = 12	<i>n</i> = 50	<i>n</i> = 18	<i>n</i> = 30	<i>n</i> = 40
Age (years), mean (SD)	33.1 (11.0)	41.5 (13.4)	35.6 (7.7)	43.3 (11.9)	42.6 (12.9)	42.5 (9.9)
Gender, male (%)	10 (47.6%)	4 (33.3%)	25 (50.0%)	5 (27.8%)	9 (30.0%)	12 (30.0%)
Percentage of patients who have experienced laryngeal edema (%)	84.2%	75.0%	NA	55.6%	73.3%	NA
Percentage for patients who have experienced abdominal pain (%)	68.4%	66.7%	NA	77.8%	76.7%	NA
No. with HAE severity score (0 point)	0	0	NA	0	8	NA
No. with HAE severity score (1 point)	0	0	NA	2	1	NA
No. with HAE severity score (2 points)	2	2	NA	2	2	NA
No. with HAE severity score (3 points)	2	1	NA	1	7	NA
No. with HAE severity score (4 points)	9	5	NA	5	2	NA
No. with HAE severity score (5 points)	5	3	NA	6	5	NA
No. with HAE severity score (6 points)	0	0	NA	1	1	NA
No. with HAE with severity score (NA)	3	1	NA	1	4	NA
C4 (g/l), normal reference value: 0.100–0.400 g/l	0.078 (0.056–0.129)	0.098 (0.062–0.159)	NA	0.032 (0.014–0.066)	0.062 (0.052–0.095)	NA
Cl-INH (g/l), normal reference value: 0.21–0.39 g/l	0.080 (0.050–0.110)	0.075 (0.058–0.123)	NA	0.050 (0.040–0.060)	0.060 (0.055–0.095)	NA

Note: Clinical parameters are represented as median and interquartile range. The clinical severity score (0–7 points) system was performed as follow: frequency of attacks >1/month, 1'; skin edema of any area ever, 1'; single submucous site involvement, 1'; multiple submucous sites involvement, 2'; emergency visit ever, 1'; improper abdominal operation ever, 1'; tracheotomy, 1'.

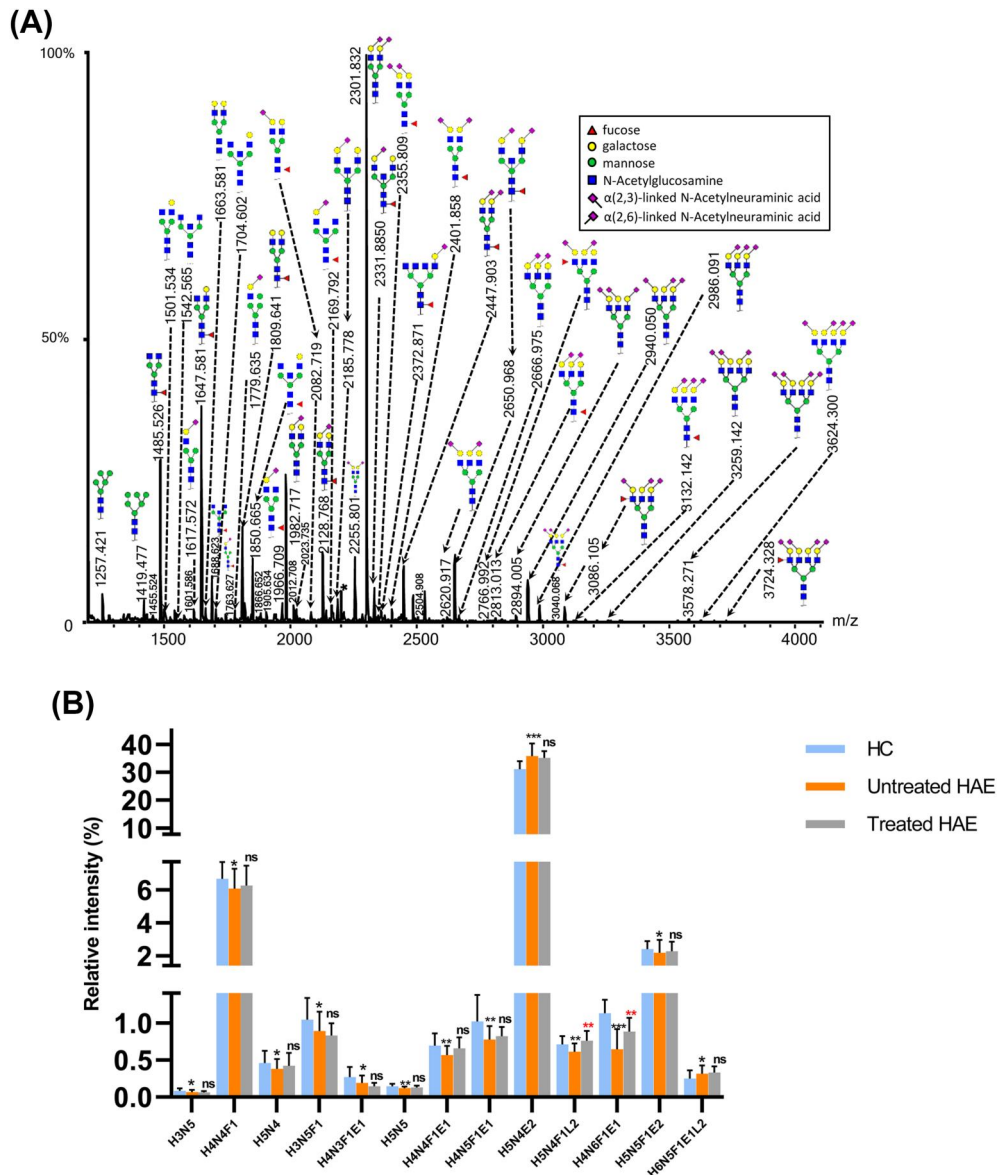
Abbreviations: HAE, hereditary angioedema; HC, healthy control; NA, not available.

### 3 | RESULTS

The clinical characteristics of two cohorts are presented in Table 1. Fifty-nine directly detected *N*-glycan compositions passed our quality criteria for subsequent quantification. Eighty-two derived *N*-glycan traits were generated based on the structural features of the directly detected glycans. Features of the derived *N*-glycan traits include the number of antennae (A), fucosylation (F), bisection (B), galactosylation (G),  $\alpha$ 2,3-linked (L), and  $\alpha$ 2,6-linked sialylation (E) (Figure 1; Tables S2–S3). The consistent quality and overall method repeatability of the data was assessed by including plasma standards (Table S4). The average RSD of the directly detected *N*-glycan traits (top 20) and all 82 derived traits was 3.93% and 2.32%, respectively (Table S4). Derived traits appear to have better technical robustness than directly detected glycans. When doing analysis, we found 13 directly detected glycan traits were significantly changed in HC compared with untreated HAE (Figure 1B). As derived *N*-glycan traits could combine the exact effects of individual glycans sharing similar glycan structures, facilitate interpretation of biological effects, and have higher repeatability than the individual glycan traits from which they were calculated,<sup>19</sup> we subsequently focused on the derived *N*-glycan traits for the comparison and analysis in the present study.

#### 3.1 | Identification of plasma *N*-glycomic features in HAE patients

Eight replicated derived *N*-glycan traits differed significantly between patients with drug-naïve type 1 HAE and HCs in the two independent cohorts (Figure 2; Table 2; Figure S1). *N*-glycomes in HAE patients had different antennarity within the complex-type glycans (CA) than in HCs. The level of tetra-antennary glycans within complex type (CA4) in HAE patients decreased compared with that in HCs (Figure 2; Table 2; Figure S1). The levels of galactosylation of tetra-antennary glycans (A4G) and galactosylation of sialylated diantennary glycans (A2SG) in HAE patients were higher than that in HCs (Figure 2; Table 2; Figure S1). Patterns of sialylation also differed between HAE patients and HCs. The level of sialylation of tetra-antennary glycans (A4S) was higher in subjects with HAE than in HCs, which was primarily caused by the increase in  $\alpha$ 2,3-linked sialylation of tetra-antennary glycans (A4L; Figure 2; Table 2; Figure S1). Increased  $\alpha$ 2,6-linked sialylation of non-fucosylated tetra-antennary glycans (A4FOE) was also observed in HAE (Figure 2; Table 2; Figure S1). In contrast,  $\alpha$ 2,6-linked sialylation of fucosylated tetra-antennary glycans (A4FE) and  $\alpha$ 2,6-linked sialylation per galactose of tetra-antennary glycans (A4GE) decreased in HAE patients compared to that in HCs (Figure 2; Table 2; Figure S1). No differences



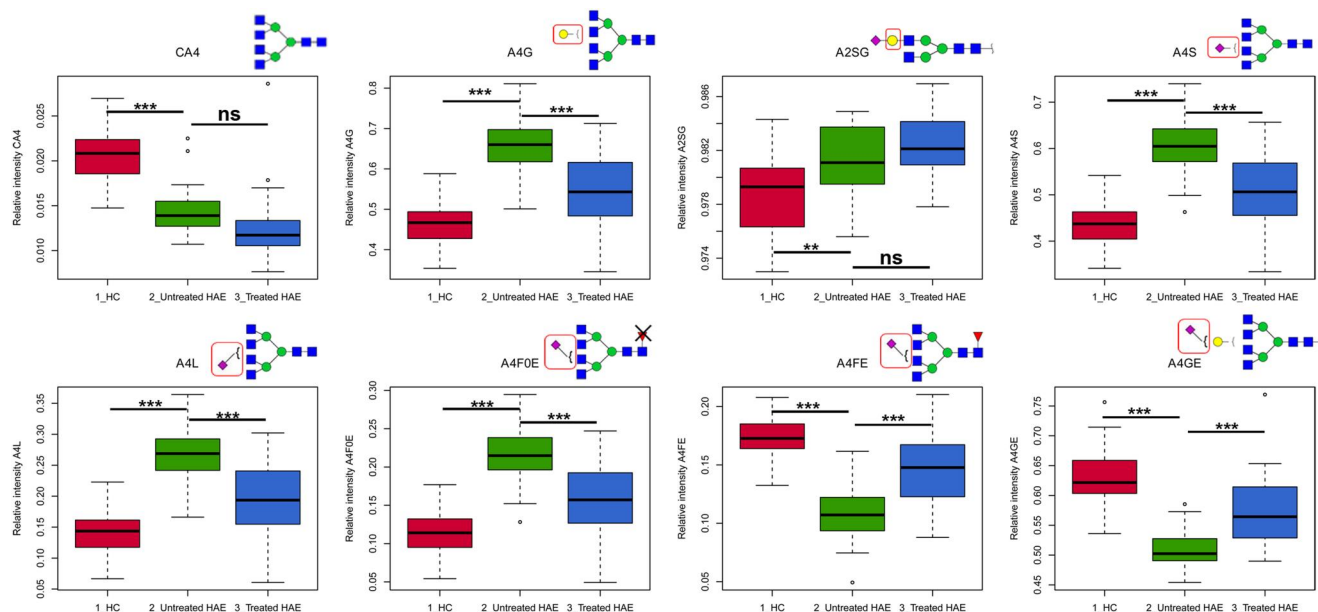
**FIGURE 1** (A) A typical matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF-MS) spectrum of the plasma protein *N*-glycome in hereditary angioedema patients. Spectra were recorded in positive-ion reflectron mode on a Bruker rapifleXtreme mass spectrometer. Major *N*-glycan peaks were annotated and assigned to compositions. The presence of structural isomers cannot be excluded. Asterisks (\*) indicate by-products. (B) The bar plot with standard deviation bars for the directly detected glycan traits changing in untreated hereditary angioedema (HAE) patients compared to that in healthy controls of the cohort. \*\*\*:  $p$ -value < 0.001, \*\*:  $p$ -value < 0.01, \*:  $p$ -value < 0.05, ns: not significant. The "\*" or "ns" on the top of the histogram of the treated HAE represent the statistical significance for "HC versus untreated HAE," on the top of treated HAE represent the statistical significance for "untreated HAE versus treated HAE." HAE, hereditary angioedema; HC, healthy control; H, hexose; N, N-acetylhexosamine; F, deoxyhexose (fucose); L, acetonized N-acetylneuraminic acid ( $\alpha$ 2,3-linked); E, ethyl esterified N-acetylneuraminic acid ( $\alpha$ 2,6-linked)

were found for fucosylation (F), bisection (B), hybrid (Thy), and high-mannose structures (TM) between HAE patients and HCs.

### 3.2 | Performance of plasma glycan traits for diagnosing HAE

The top seven potential classification biomarker models were constructed using multivariate algorithms for SVMs, based on the

eight replicated differentially expressed derived *N*-glycan traits in the two independent cohorts (Figure 3; Figure S2). For each model, two to eight derived *N*-glycan traits were selected via automated important feature identification (Figure 3D; Figure S2D). ROC curves were generated for each model. The performance of the seven models was highly accurate, with AUCs ranging from 0.927 to 0.931, and predictive accuracies from 87.5% to 90.0%, for discriminating between HAE patients and HCs in the discovery cohort (Figure S2A–S2C). The performance of the



**FIGURE 2** Boxplots of derived glycan traits changing in untreated hereditary angioedema (HAE) patients compared to that in healthy controls in the validation cohort. The boxplot for the group of HAE patients after receiving treatment is also shown. \*\*\*:  $p$ -value < 0.001, \*\*:  $p$ -value < 0.01, \*:  $p$ -value < 0.05, ns: not significant (after multiple testing correction). HAE, hereditary angioedema; HC, healthy control

models was validated in the validation cohort, with AUCs ranging from 0.959 to 0.999, and predictive accuracies from 89.6% to 97.9% (Figure 3A–3C). Of note, the diagnostic performance of these biomarker models was better than that of individual derived *N*-glycan traits (Figures S3 and S4).

### 3.3 | Plasma *N*-glycomic changes in HAE patients after treatment

Among the eight replicated abnormal plasma *N*-glycomic features identified in HAE patients, A4G, A4S, A4L, A4F0E, A4FE, and A4GE showed responses to treatment (Figure 2; Figure S1). After treatment, these derived *N*-glycan traits showed significant differences between treated and untreated HAE patient groups and returned to near normal levels (Figure 2; Figure S1). In contrast, derived *N*-glycan traits CA4 and A2SG did not change after treatment (Figure 2; Figure S1). In addition, medication control can reduce the frequency of edema attacks of the patients (data not shown). To investigate whether these glycan traits could be used as biomarkers for disease monitoring/prognosis, we attempted to construct predictive models using multivariate algorithms based on the derived glycan traits that showed a response to treatment. The top five models included two to six of the glycan traits with “accurate” performance based on their AUCs, which ranged from 0.809 to 0.868 in predicting the response to treatment in the discovery cohort (Figure S5A). The results were validated in the validation cohort, with AUCs ranging from 0.831 to 0.852 (Figure S5B). These models may have applications as biomarkers in HAE disease monitoring and prognosis.

### 3.4 | Association between plasma *N*-glycomes and clinical features and parameters

The association between plasma *N*-glycomes and the occurrence of laryngeal angioedema, gastrointestinal angioedema, and disease severity score in HAE patients was explored by logistic regression. A severity scoring system mainly based on the edema frequency and locations was used to assess the disease severity of HAE patients, as previously described.<sup>20</sup> Specifically, the clinical severity score (0–7 points) system was performed as follow: frequency of attacks >1/month, 1 point; skin edema of any area ever, 1 point; single submucous site involvement, 1 point; multiple submucous sites involvement, 2 points; emergency visit ever, 1 point; improper abdominal operation ever, 1 point; tracheotomy, 1 point.<sup>20</sup> Galactosylation of diantennary glycans (A2G) was positively associated with the occurrence of laryngeal angioedema in HAE patients (Table S5).  $\alpha$ 2,3-linked sialylation (A2FOL, A2F0GL, and A3L) was negatively associated with gastrointestinal angioedema occurrence (Table S5). Galactosylation of diantennary glycans (A2G) was found to be positively associated with disease severity score (Table S5). Several glycan traits (A3F, A3LF, A3EF, A3F0S, and A4GE) were associated with plasma levels of CI-INH (Table S5). The associations of these glycan traits were validated in the validation cohort (Table S5).

### 3.5 | Performance of plasma glycan traits in predicting the occurrence of laryngeal and gastrointestinal angioedema

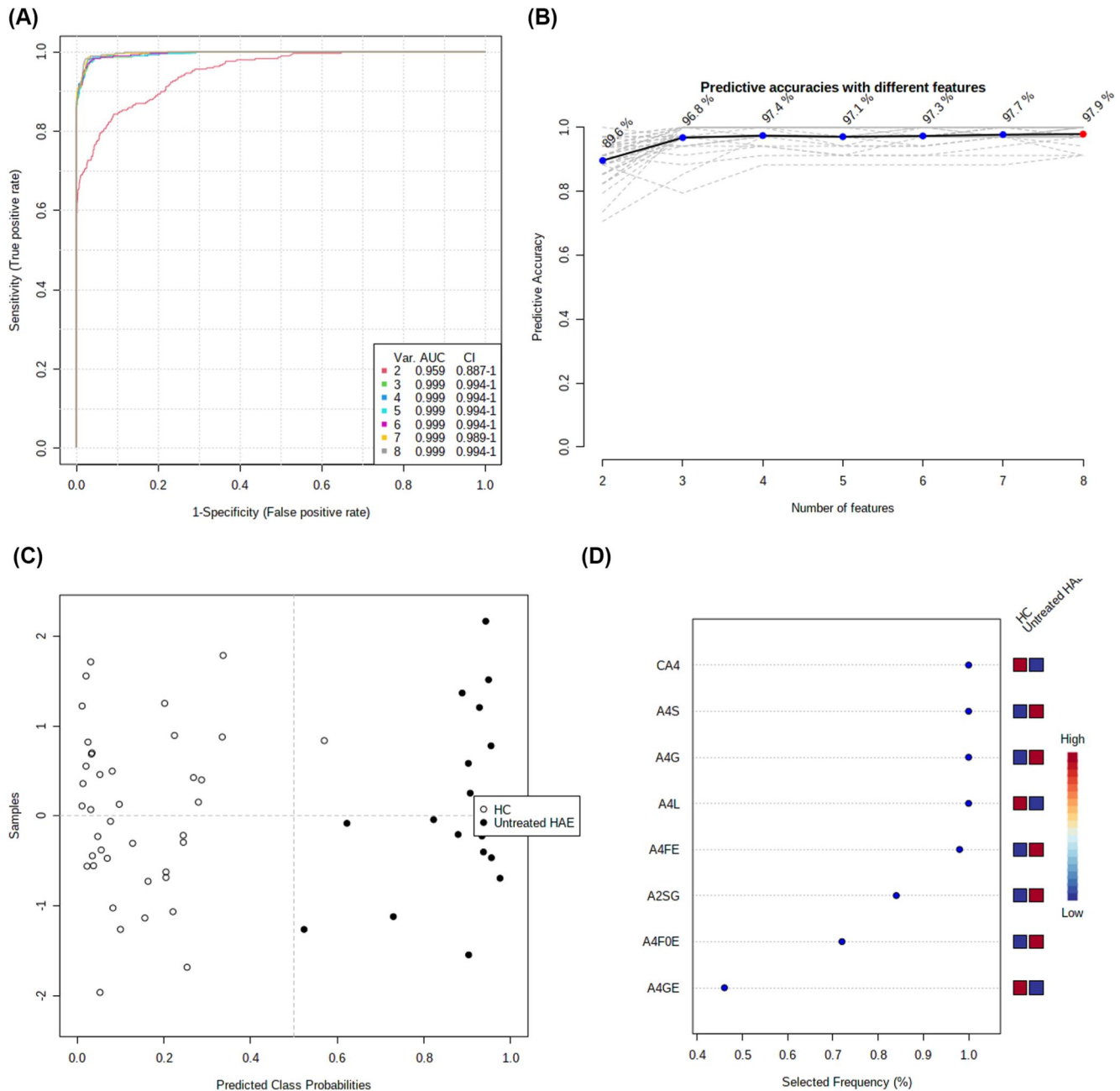
The predictive efficacy of glycan traits with strong associations with laryngeal or gastrointestinal angioedema was evaluated by plotting

**TABLE 2** The dysregulated derived N-glycan traits in hereditary angioedema (HAE) patients compared to healthy controls in the discovery and validation cohort

	Description	Discovery cohort					Validation cohort				
		Median (IQR) in untreated HAE	Median (IQR) in HC	Median (IQR) in treated HAE	<i>p</i> -value	Response to therapy	Median (IQR) in untreated HAE	Median (IQR) in HC	Median (IQR) in treated HAE	<i>p</i> -value	Response to therapy
					Untreated HAE versus HC					Untreated HAE versus HC	
<b>Glycan traits-general</b>											
CA4	Tetraantennary species (A4) in complex glycans	0.0112 (0.0097–0.0133)	0.0200 (0.0180–0.0224)	0.0123 (0.0117–0.0154)	<0.001	No	0.0139 (0.0125–0.0157)	0.0208 (0.0185–0.0224)	0.0117 (0.0105–0.0135)	<0.001	No
<b>Glycan traits-galactosylation(G)</b>											
A4G	In tetraantennary (A4)	0.6132 (0.5427–0.6425)	0.4920 (0.4594–0.5354)	0.4909 (0.4512–0.5331)	<0.001	Yes	0.6601 (0.6123–0.7047)	0.4668 (0.4274–0.4950)	0.5430 (0.4835–0.6162)	<0.001	Yes
A2SG	In sialylated diantennary (A2)	0.9813 (0.9766–0.9825)	0.9753 (0.9694–0.9793)	0.9789 (0.9739–0.9806)	<0.001	No	0.9811 (0.9795–0.9837)	0.9793 (0.9763–0.9807)	0.9821 (0.9808–0.9842)	<0.001	No
<b>Glycan traits-sialylation (S)</b>											
A4S	In tetraantennary (A4)	0.5806 (0.5105–0.6015)	0.4698 (0.4375–0.5070)	0.4569 (0.4188–0.4974)	<0.001	Yes	0.6048 (0.5668–0.6472)	0.4371 (0.4036–0.4632)	0.5063 (0.4550–0.5688)	<0.001	Yes
<b>Glycan traits-<math>\alpha</math>2,3-sialylation (L)</b>											
A4L	In tetraantennary (A4)	0.2457 (0.1933–0.2676)	0.1559 (0.1319–0.1879)	0.1544 (0.1343–0.1762)	<0.001	Yes	0.2688 (0.2383–0.2972)	0.1436 (0.1174–0.1623)	0.1936 (0.1547–0.2408)	<0.001	Yes
<b>Glycan traits-<math>\alpha</math>2,6-sialylation (E)</b>											
A4FOE	In non-fucosylated tetraantennary (A4)	0.2052 (0.1596–0.2270)	0.1326 (0.1092–0.1645)	0.1268 (0.1008–0.1564)	<0.001	Yes	0.2148 (0.1933–0.2406)	0.1140 (0.0946–0.1323)	0.1571 (0.1263–0.1930)	<0.001	Yes
A4FE	In fucosylated tetraantennary (A4)	0.1242 (0.1158–0.1454)	0.1614 (0.1477–0.1719)	0.1740 (0.1555–0.1919)	<0.001	Yes	0.1071 (0.0908–0.1241)	0.1727 (0.1636–0.1852)	0.1477 (0.1227–0.1673)	<0.001	Yes
A4GE	Per galactose in tetraantennary (A4)	0.5455 (0.5084–0.5905)	0.6348 (0.5975–0.6593)	0.6345 (0.6007–0.6587)	<0.001	Yes	0.5024 (0.4872–0.5301)	0.6216 (0.6034–0.6590)	0.5643 (0.5281–0.6159)	<0.001	Yes

Note: Table showing the descriptions of the derived traits, median values (IQR), *p*-values (Mann-Whitney-Wilcoxon test), and the response to therapy. The significance threshold was corrected for multiple tests. Differences were considered significant if  $p < 6.10E-4$  (*p*-value of 0.05/82). Dark and light gray shading indicate up- and down-regulation in untreated HAE compared with in HCs, respectively.

Abbreviations: HAE, hereditary angioedema; HC, healthy control.

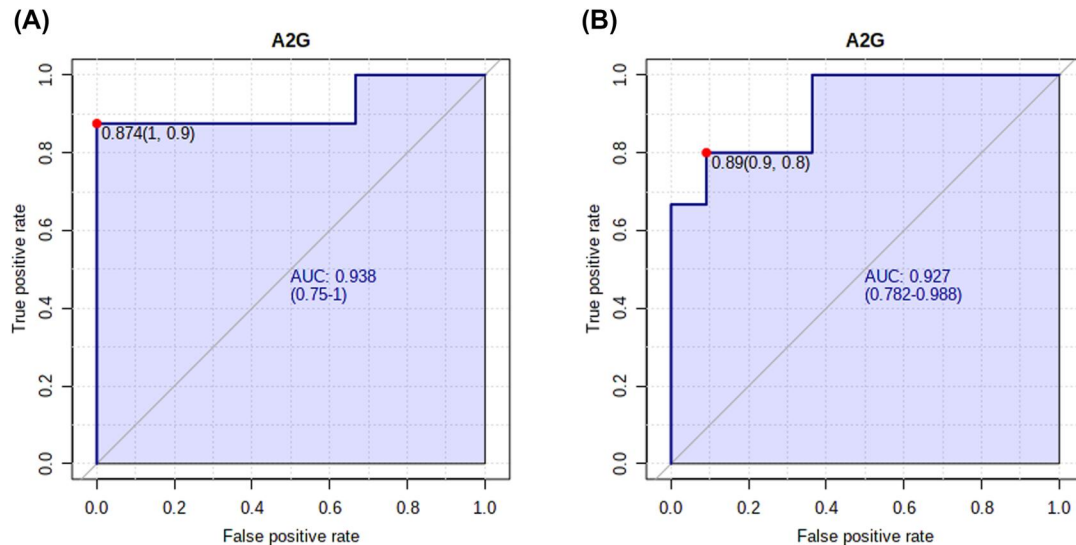


**FIGURE 3** Classification models for differentiating between hereditary angioedema (HAE) and healthy controls based on the eight differentially expressed glycan traits in the validation cohort. (A) Receiver operator characteristics (ROC) curves for the top seven putative biomarker models, based on each model's average performance across all Monte Carlo cross validation (MCCV) runs. (B) Plot of the classification accuracy of the classification models with an increased number of glycan traits. The most accurate model is marked with a red dot. (C) Plot of the classification between case and control using a single biomarker model. Due to the equilibrium of the subsampling, the classification boundary is at the center ( $x = 0.5$ ). Selected model: seven. (D) Plot of the most important glycan traits according to the classification models (from most to least important). HAE, hereditary angioedema; HC, healthy control

ROC curves. In the discovery cohort, the AUC of A2G was 0.938 for distinguishing between HAE groups in which laryngeal angioedema had ever occurred versus not occurred (Figure 4A). The performance of A2G in predicting the occurrence of laryngeal angioedema was validated in the validation cohort with an AUC of 0.927 (Figure 4B), suggesting that A2G has good predictive ability

for laryngeal angioedema occurrence. Based on the three derived glycan traits that showed strong associations with gastrointestinal angioedema, two potential predictive biomarker models were constructed (Figure S6). For each model, two to three derived N-glycan traits were selected through automated important feature identification (Figure S6). ROC curves were generated for each model. The





**FIGURE 4** Performance of glycan trait A2G in predicting the occurrence of laryngeal angioedema (A) in the discovery cohort and (B) in the validation cohort. The red dot represents the optimal cut-off value

models were moderately accurate ( $0.7 < \text{AUC} < 0.8$ ) in predicting whether HAE patients will ever develop gastrointestinal angioedema (Figure S6).

## 4 | DISCUSSION

Accurate, non-invasive biomarkers for the diagnosis, prediction of disease severity, and disease monitoring of HAE are currently unmet clinical needs. Recently, plasma/serum protein *N*-glycomes have been identified as biologically significant, and have emerged as a showcase of non-invasive biomarkers for various diseases.<sup>10-14,18</sup> The objective of this study was to reveal the disease-specific *N*-glycome phenotype of HAE to discover potential biomarkers for the diagnosis, prediction of disease severity, and monitoring of HAE progression. Primary dysregulation of plasma *N*-glycosylation was identified in HAE patients, particularly in the antennarity of glycans, galactosylation, and ( $\alpha$ 2,3- and  $\alpha$ 2,6-linked) sialylation. Eight derived glycan traits were abnormal in HAE patients, from which seven classification models with potential as biomarkers were constructed. Several of these dysregulated glycan traits in HAE returned to near normal levels after treatment, showing potential in monitoring this disease. Predictive models based on the glycan traits that were strongly associated with clinical symptoms had good predictive ability for disease severity. To our knowledge, this study is the first report of the plasma protein *N*-glycome of HAE, and provides putative glycan-based biomarkers with validation in an independent cohort. In previous studies of plasma/serum *N*-glycomes for diseases other than HAE, the differentially expressed *N*-glycan traits were different from those identified for HAE in the present study,<sup>10-14</sup> confirming a disease-specific *N*-glycome signature, and suggesting that the plasma *N*-glycome can serve as disease-specific biomarkers.

Here, we found that the levels of galactosylation of tetra-antennary glycans (A4G) and galactosylation of sialylated diantennary glycans (A2SG) were increased in HAE compared to that in HCs. Higher galactosylation level was previously linked with increased risk of developing type 2 diabetes.<sup>21</sup> Moreover, the levels of tri- and tetra-galactosylated *N*-glycans were previously reported to be increased with age in patients with Down syndrome, but not in HCs.<sup>10</sup> In humans, the glycan traits A2SG/A4G are primarily derived from alpha1-antitrypsin, alpha1B-glycoprotein, fibrinogen, haptoglobin, hemopexin, serotransferrin, and other acute-phase proteins produced in the liver.<sup>18</sup> Evidence for altered hepatic synthesis has been found in HAE patients.<sup>22</sup> Furthermore, some of these liver-derived proteins are related to the pathophysiology of HAE. Acute-phase proteins, such as fibrinogen, are risk factors for angioedema induced by bradykinin.<sup>23</sup> In addition, the production of the inflammatory peptide bradykinin by the contact system can be inhibited by alpha1-antitrypsin variants.<sup>24</sup> Galactosylation occurs via a set of beta-1,4-galactosyltransferases, the activity of which in plasma is associated with inflammatory diseases and aging.<sup>10,25</sup> Decreased IgG-derived galactosylation has been linked to inflammation, immune disorders, cancers, and aging.<sup>26-28</sup> We also observed decreased IgG-galactosylation in HAE patient plasma, although this decrease was not significant after multiple testing correction (TA2FS0). Though the role of increased non-IgG-derived galactosylation, which was linked to acute-phase liver proteins in the present study, has not yet been established, we assume that abnormal non-IgG-derived galactosylation is involved in the causal mechanisms of HAE, and thus warrants further investigation.

Sialylation directly participates in the activation of the immune system, which depends on sialic acid linkage types.<sup>29,30</sup> For example, siglecs on immune cells specifically recognize non-fucosylated  $\alpha$ 2,6-sialic acid epitopes.<sup>29</sup> Our novel MS-based approach enables the

discrimination between  $\alpha$ 2,3-linked and  $\alpha$ 2,6-linked sialylation. Overall, we found sialylation of tetra-antennary glycans (A4S) was higher in HAE patients than in HCs. Interestingly, sialylation within non-fucosylated glycans (A4FOE) was increased, whereas sialylation within fucosylated glycans (A4FE) was decreased in HAE patients. The decrease in  $\alpha$ 2,6-sialylation with a fucose (A4FE) may derive from immunoglobulins, whereas the increased glycan traits are primarily derived from a mixture of glycoproteins containing highly sialylated glycans produced by the liver (e.g.,  $\alpha$ 1-acid glycoprotein and  $\alpha$ 1-antitrypsin),<sup>18</sup> as immunoglobulins primarily carry the  $\alpha$ 2,6-sialylation with core-fucosylation.<sup>31,32</sup> The observed increase in sialylation within non-fucosylated liver-derived tetra-antennary glycans might reflect glycosylation and abundance changes in acute-phase proteins. Abnormal sialylation has been found previously in autoimmune diseases, multiple cancer types, and inflammation.<sup>33,34</sup> Furthermore,  $\alpha$ 2,6-sialylation of non-fucosylated multiple antennary glycans (A(3–4)FOGE) is associated with inflammatory markers in inflammatory bowel disease.<sup>12</sup> Although most glycan traits containing sialylation were altered during HAE, our data suggest that these changes in sialylation were partially driven by  $\alpha$ 2,6-sialyltransferases. ST6Gal1 is the main sialyltransferase, which attaches  $\alpha$ 2,6-linked sialic acids to *N*-glycans.<sup>25</sup> Circulatory ST6Gal1 is a systemic regulator, and can directly influence the number of inflammatory cells.<sup>35,36</sup> We hypothesize that abnormal  $\alpha$ 2,6-linked sialylation of tetra-antennary glycans (A4FOE, A4FE, and A4GE) in HAE may be implicated in the HAE-associated inflammatory changes.

We found a series of glycan traits that showed associations with clinical features or parameters. Considering the stability of glycans during storage, convenience to detect, low amount of plasma required, and low cost, the glycan traits related to clinical parameter (CI-INH) may serve as biomarkers for HAE diagnosis. Of note, as excellent diagnostic tests (high sensitivity/specificity) for HAE-C1INH deficiency exist, the value of the potential glycan biomarkers screened out in the present study may be less for diagnosis of HAE-C1INH deficiency. However, the potential value of the glycan biomarker candidates for the diagnosis of HAE-normal C1INH (where diagnostics are sorely lacking) would be a strong consideration for future study and work. On the other hand, the glycan traits associated with laryngeal or gastrointestinal angioedema occurrence or disease severity (represented by severity score in the present study) may play a critical role in predicting the disease severity of HAE, and guide disease prevention and treatment. The predictive and prognostic value of these glycan traits for anatomical location of attacks and/or attack frequency would represent an advance since it largely lacks any predictors of phenotype or clinical course in this unpredictable condition in the clinical setting. In the present study, “response to treatment” means glycans returned to (near) normal levels after treatment. Among the eight replicated abnormal plasma *N*-glycomic features identified in HAE patients, five *N*-glycan traits showed responses to treatment. After treatment,

these derived *N*-glycan traits showed significant differences between treated and untreated HAE patient groups and returned to near normal levels. Additionally, HAE patients with regular use of danazol >1 month were included in the subgroup of treated HAE. Medication control can reduce the frequency of edema attacks of these patients (data not shown). Considering both the glycan changes and reduction of attacks after treatment, we propose that the *N*-glycan traits which showed response to treatment may have potential as predictive or monitoring biomarkers. However, this needs further long-term follow-up and validation studies in large cohorts.

The present study has some limitations. First, isomers may exist for our assigned glycan structures. Second, MS-based *N*-glycome profiling provides relative quantification, and is influenced to some extent by plasma/serum levels of the related glycoproteins. Quantifying protein-specific glycosylation in combination with measuring the levels of glycoproteins will give in-depth insights into the mechanisms underpinning the *N*-glycome, but remains a challenge in this field. Third, because of the rarity of this disease, the untreated and treated HAE groups could not be paired, which could affect the resulting data. Despite these limitations, our research serves as a starting point for future validation and mechanistic studies.

In conclusion, using comprehensive *N*-glycomic profiling methods, we analyzed the plasma *N*-glycome for two independent well-characterized type 1 HAE cohorts, and for the first time reported the full plasma *N*-glycomic signature of HAE. We identified plasma *N*-glycosylation changes specific to HAE, and observed major disease-specific dysregulated glycosylation, namely branching of complex (CA), galactosylation, and sialylation. Novel associations between clinical symptoms/parameters of HAE and glycosylation were found. The combination of altered glycan traits was used to generate classification/prediction models that could be used for the diagnosis, prediction of disease severity, and monitoring of HAE. All the results were validated in an independent validation cohort. These results may play a critical role in predicting/assessing the disease severity of HAE in the future. Further studies are needed to improve our understanding of the role of glycosylation in the pathology of HAE.

## ACKNOWLEDGMENTS

We thank all members of our group involved in this study. We thank the patients and volunteers who provided blood samples for the present study. This research was supported by grants from the National Natural Science Foundation of China (No. 81472870, to Yuxiang Zhi, No. 31901041, to Zejian Zhang), the CAMS Innovation Fund for Medical Sciences (CIFMS 2021-I2M-1-003) (to Yuxiang Zhi), and the National Key Research and Development Program of China (No. 2016YFC0901501, to Yuxiang Zhi).

## CONFLICT OF INTEREST

The authors declare no conflict of interest.

## AUTHOR CONTRIBUTIONS

**Zejian Zhang:** Conceptualization; Data curation; Formal analysis; Funding acquisition; Investigation; Methodology; Resources; Software; Validation; Visualization; Writing – original draft; Writing – review & editing. **Xue Wang:** Formal analysis; Investigation; Resources; Validation; Visualization; Writing – original draft. **Jianqing Gu:** Formal analysis; Investigation; Resources; Writing – original draft. **Jianqiang Wu:** Formal analysis; Investigation; Resources; Writing – original draft. **Yang Cao:** Investigation; Resources; Writing – original draft. **Yingyang Xu:** Formal analysis; Resources; Writing – original draft. **Lisha Li:** Investigation; Resources; Writing – original draft. **Kai Guan:** Formal analysis; Resources; Writing – original draft. **Peng Liu:** Resources; Writing – original draft. **Jia Yin:** Resources; Writing – original draft. **Yuxiang Zhi:** Conceptualization; Formal analysis; Funding acquisition; Project administration; Resources; Supervision; Visualization; Writing – original draft. **Shuyang Zhang:** Conceptualization; Formal analysis; Project administration; Resources; Supervision; Writing – original draft.

## ORCID

Zejian Zhang  <https://orcid.org/0000-0002-2731-8348>

Xue Wang  <https://orcid.org/0000-0002-5612-7735>

Yang Cao  <https://orcid.org/0000-0002-8144-0889>

Yingyang Xu  <https://orcid.org/0000-0002-0583-1925>

## REFERENCES

- Busse PJ, Christiansen SC. Hereditary angioedema. *N Engl J Med*. 2020;382(12):1136-1148.
- Eyice Karabacak D, Demir S, Yegit OO, et al. Impact of anxiety, stress and depression related to COVID-19 pandemic on the course of hereditary angioedema with C1 inhibitor deficiency. *Allergy*. 2021;76(8):2535-2543. <https://doi.org/10.1111/all.14796>
- Cao Y, Liu S, Zhi Y. Recurrent and acute abdominal pain as the main clinical manifestation in patients with hereditary angioedema. *Allergy Asthma Proc*. 2021;42(2):131-135.
- Balla Z, Ignacz B, Varga L, Kohalmi KV, Farkas H. How angioedema quality of life questionnaire can help physicians in treating C1-inhibitor deficiency patients? *Clin Rev Allergy Immunol*. 2021; 61(1):50-59.
- Hart GW, Copeland RJ. Glycomics hits the big time. *Cell*. 2010; 143(5):672-676.
- Moremen KW, Tiemeyer M, Nairn AV. Vertebrate protein glycosylation: diversity, synthesis and function. *Nat Rev Mol Cell Biol*. 2012;13(7):448-462.
- Ohtsubo K, Marth JD. Glycosylation in cellular mechanisms of health and disease. *Cell*. 2006;126(5):855-867.
- Varki A. Biological roles of glycans. *Glycobiology*. 2017;27(1):3-49.
- Knezevic A, Polasek O, Gornik O, et al. Variability, heritability and environmental determinants of human plasma N-glycome. *J Proteome Res*. 2009;8(2):694-701.
- Borelli V, Vanhooren V, Lonardi E, et al. Plasma N-glycome signature of Down syndrome. *J Proteome Res*. 2015;14(10):4232-4245.
- Reiding KR, Vreeker GCM, Bondt A, et al. Serum protein N-glycosylation changes with rheumatoid arthritis disease activity during and after pregnancy. *Front Med (Lausanne)*. 2017;4:241.
- Clerc F, Novokmet M, Dotz V, et al. Plasma N-glycan signatures are associated with features of inflammatory bowel diseases. *Gastroenterology*. 2018;155(3):829-843.
- Vreeker GCM, Hanna-Sawires RG, Mohammed Y, et al. Serum N-glycome analysis reveals pancreatic cancer disease signatures. *Cancer Med*. 2020;9(22):8519-8529.
- Zhang Z, Westhryn M, Bondt A, Wuhrer M, Standal T, Holst S. Serum protein N-glycosylation changes in multiple myeloma. *Biochim Biophys Acta Gen Subj*. 2019;1863(5):960-970.
- Reiding KR, Blank D, Kuijper DM, Deelder AM, Wuhrer M. High-throughput profiling of protein N-glycosylation by MALDI-TOF-MS employing linkage-specific sialic acid esterification. *Anal Chem*. 2014;86(12):5784-5793.
- Jansen BC, Reiding KR, Bondt A, et al. MasyTools: a high-throughput targeted data processing tool for relative quantitation and quality control developed for glycomic and glycoproteomic MALDI-MS. *J Proteome Res*. 2015;14(12):5088-5098.
- Ceroni A, Maass K, Geyer H, Geyer R, Dell A, Haslam SM. Glyco-Workbench: a tool for the computer-assisted annotation of mass spectra of glycans. *J Proteome Res*. 2008;7(4):1650-1659.
- Clerc F, Reiding KR, Jansen BC, Kammeijer GS, Bondt A, Wuhrer M. Human plasma protein N-glycosylation. *Glycoconj J*. 2016;33(3): 309-343.
- Bladergroen MR, Reiding KR, Hipgrave Ederveen AL, et al. Automation of high-throughput mass spectrometry-based plasma N-glycome analysis with linkage-specific sialic acid esterification. *J Proteome Res*. 2015;14(9):4080-4086.
- Xu YY, Zhi YX, Yin J, et al. Mutational spectrum and geno-phenotype correlation in Chinese families with hereditary angioedema. *Allergy*. 2012;67:1430-1436.
- Keser T, Gornik I, Vuckovic F, et al. Increased plasma N-glycome complexity is associated with higher risk of type 2 diabetes. *Diabetologia*. 2017;60(12):2352-2360.
- Johnson AM, Alper CA, Rosen FS, Craig JM. C1 inhibitor: evidence for decreased hepatic synthesis in hereditary angioneurotic edema. *Science*. 1971;173(3996):553-554.
- Bas M, Kirchhartz N, Hochfeld J, et al. Potential role of vasomotor effects of fibrinogen in bradykinin-induced angioedema. *J Allergy Clin Immunol*. 2008;121(4):969-975.
- de Maat S, Sanrattana W, Mailer RK, et al. Design and characterization of alpha1-antitrypsin variants for treatment of contact system-driven thromboinflammation. *Blood*. 2019;134(19):1658-1669.
- Catera M, Borelli V, Malagolini N, et al. Identification of novel plasma glycosylation-associated markers of aging. *Oncotarget*. 2016;7(7):7455-7468.
- 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(2): 685-698.
- van de Geijn FE, Wuhrer M, Selman MH, 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*. 2009;11(6):1-10.
- 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.
- Varki A, Gagneux P. Multifarious roles of sialic acids in immunity. *Ann N Y Acad Sci*. 2012;1253:16-36.
- McCurdy TR, Bhakta V, Eltringham-Smith LJ, Gataiance S, Fox-Robichaud AE, Sheffield WP. In vivo clearance of alpha-1 acid glycoprotein is influenced by the extent of its N-linked glycosylation and by its interaction with the vessel wall. *J Biomed Biotechnol*. 2012;2012:292730.
- Arnold JN, Wormald MR, Suter DM, et al. Human serum IgM glycosylation: identification of glycoforms that can bind to mannan-binding lectin. *J Biol Chem*. 2005;280(32):29080-29087.

32. Saldova R, Stockmann H, O'Flaherty R, Lefeber DJ, Jaeken J, Rudd PM. N-glycosylation of serum IgG and total glycoproteins in MAN1B1 deficiency. *J Proteome Res.* 2015;14(10):4402-4412.
33. Zhang Z, Wuhrer M, Holst S. Serum sialylation changes in cancer. *Glycoconj J.* 2018;35(2):139-160.
34. Novokmet M, Lukic E, Vuckovic F, et al. Changes in IgG and total plasma protein glycomes in acute systemic inflammation. *Sci Rep.* 2014;4:4347.
35. Dougher CWL, Buffone A, Jr, Nemeth MJ, et al. The blood-borne sialyltransferase ST6Gal-1 is a negative systemic regulator of granulopoiesis. *J Leukoc Biol.* 2017;102(2):507-516.
36. Abe Y, Smith CW, Katkin JP, et al. Endothelial alpha 2,6-linked sialic acid inhibits VCAM-1-dependent adhesion under flow conditions. *J Immunol.* 1999;163(5):2867-2876.

## SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

**How to cite this article:** Zhang Z, Wang X, Gu J, et al. Validation of diagnostic and predictive biomarkers for hereditary angioedema via plasma N-glycomics. *Clin Transl Allergy.* 2021;e12090. <https://doi.org/10.1002/ctt2.12090>