Glycomics based biomarkers of the rate of aging : development and applications of high-throughput N-glycan analysis

Ruhaak, L.R.

Chapter 5.
Plasma protein N-glycan profiles are associated with calendar age, familial longevity and health

L.Renee Ruhaak[1], Hae-Won Uh[2], Marian Beekman[3], Cornelis H. Hokke[1], Rudi G.J. Westendorp[4], Jeanine Houwing-Duistermaat[2], Manfred Wuhrer[1], André M. Deelder[1], P.Eline Slagboom[3,5]

[1] Dept. of Parasitology, Biomolecular Mass Spectrometry Unit, Leiden University Medical Center, Leiden, The Netherlands

[2] Dept. of Medical Statistics and Bioinformatics, section Medical Statistics, Leiden University Medical Center, Leiden, The Netherlands

[3] Dept. of Medical Statistics and Bioinformatics, section Molecular Epidemiology, Leiden University Medical Center, Leiden, The Netherlands

[4] Dept. of Gerontology and Geriatrics, Leiden University Medical Center, Leiden, The Netherlands

[5] Netherlands Consortium of Healthy Aging

J. Proteome Res., in press

Abstract

Background: The development of medical interventions resulting in preservation of disease-free longevity would be greatly facilitated by markers that predict healthy aging. Protein N-glycosylation has recently been recognized as an important reservoir of potential biomarkers associated with a range of physiological processes. Changes in N-glycosylation patterns have been found to be associated with calendar age. We recently reported that offspring of nonagenarian siblings in the Leiden Longevity Study (LLS) have decreased levels of di-antennary non-galactosylated glycans on plasma immunoglobulin G (IgG) compared to their partners. Here we investigate whether glycans derived from the total glycoprotein pool in plasma also mark familial longevity and distinguish healthy from unhealthy aging in the LLS.

Methods and results: Total plasma N-glycan profiles of 2396 middle aged participants in the LLS were obtained by glycan release, labeling and subsequent HPLC analysis with fluorescence detection. After normalization and batch correction, several regression strategies were applied to evaluate associations between glycan patterns, familial longevity and healthy aging. Two N-glycan features (LC-7 and LC-8) could be identified to be more abundant in plasma of the offspring of long-lived individuals as compared to their partners as controls. These results were not confounded by the altered lipid status or glucose homeostasis of the offspring. The glycan features could be associated with CRP levels, indicating that glycosylation probably also marks health status in this cohort. Furthermore, a decrease in levels of LC-8 was associated with the occurrence of myocardial infarction, indicating that plasma glycosylation patterns do not only mark familial longevity, but may also reflect healthy aging.

Conclusion: We here describe two glycan features, of which increased levels in human plasma mark familial longevity while decreased levels mark the presence of cardiovascular disease. The two N-glycan features do not interfere with previously reported markers for longevity, and may thus be considered a novel group of markers.

Introduction

Glycosylation is the enzymatic addition of oligosaccharides (also known as glycans) to proteins and lipids. In N-glycosylation, the glycans are attached to the asparagine residues in the protein. N-Glycans have important functions in many biological processes such as protein folding [1], protein clearance [2], cell adhesion [3;4], receptor binding and receptor activation [5;6]. Protein N-glycosylation may be very diverse, and is a dynamic equilibrium: in a given physiological state, the glycan signature is highly reproducible [7;8]; however, when the physiological state changes, e.g. due to aging or disease, the glycosylation machinery of affected cells in an organism may be altered, and the glycan pattern can change dramatically [7]. Therefore, protein N-glycosylation patterns may represent an important group of potential biomarkers of health and disease [9].

The analysis methods required for the evaluation of protein glycosylation patterns in larger sample sets have only recently been developed. The variability and heritability of plasma N-glycosylation, the age-dependence as well as the influence of some behavioral parameters on plasma glycosylation have been reported. Very large biological variability in glycosylation was observed in 1008 healthy Croatian family members [21], with coefficients of variation ranging from 6.4 to 50 %. Moreover a broad range of variation in heritability could be observed for the different N-glycan features, ranging from insignificant to high, indicating that some glycans are mostly controlled by genetic factors, wile other glycans are mainly under environmental control [21].

Total plasma N-glycosylation patterns were found to be associated with calendar age in a study population of 100 Belgian individuals. It was reported that elderly individuals above 50 years of age showed increased levels of non-galactosylated glycans, while the levels of galactosylated glycan structures decreased with increasing calendar age [20]. N-glycosylation patterns were also monitored in a group of Italian centenarians: In this exceptionally high-age group associations were observed between glycosylation and calendar age that were similar to the associations observed in the Belgian population [20]. In a more recent study [24], N-glycans

were analyzed from plasma of a larger cohort of 1914 unrelated Croatian individuals. Changes in levels of glycan features were observed with increasing age, and were sex specific. In general, females showed more profound associations between glycan patterns and age than males, while some glycans may showed opposite associations in males compared to females. Interestingly, glycosylation patterns of women changed most dramatically between the age-groups 40-49 years and 50-59 years, suggesting an influence of the hormonal changes associated with entrance of the menopause.

Irrespective of being influenced by calendar age, plasma N-glycosylation was shown to be associated with body fat parameters as well as lipid status (total cholesterol, LDL-cholesterol, HDL-cholesterol and triglyceride levels) in the same study [24]. Several glycan features mainly including tetra- and trisialylated compounds correlated positively with cholesterol and lipoproteins. Changes in glycosylation could also be observed in smoking individuals [24].

While the above described changes in glycosylation mark calendar age and lipid phenotypes, parameters that mark mechanisms of healthy aging in humans would be even more desirable, as such markers might provide targets for specific interventions aimed at preservation of disease-free longevity. To this purpose, the Leiden Longevity Study (LLS) was designed: long-lived siblings were recruited together with their offspring and the partners thereof. Families were recruited if at least two long-lived siblings were alive and fulfilled the age-criterion of 89 years or older for males and 91 year or older for females, representing less than 0.1% of the Dutch population in 2001. In total 956 long-lived siblings were included with a mean age of 94, 1750 offspring with a mean age 61 and 758 partners with a mean age of 60. The exceptional aging process in these long-lived subjects and their family members is evident by their relatively low mortality [160], and at middle age by the lower prevalence of diabetes, myocardial infarction and hypertension [159]. Beneficial metabolic profiles were observed in family members as compared to controls such as lower glucose levels [161], larger LDL particle sizes and lower triglyceride levels [162].

To identify parameters that mark familial longevity, the offspring is compared to their partners. In such a comparison, the offspring is regarded to represent individuals with a higher susceptibility to become long-lived, while their partners, representing the general population, serve as controls. In this study, N-glycosylation patterns of the offspring and their partners were generated to allow identification of alterations in total plasma N-glycosylation that are associated with and could potentially mark familial longevity in the LLS and mark the physiological distinction between healthy and unhealthy aging individuals. First, the associations of glycosylation with age and sex are evaluated. Next, alterations in total plasma N-glycosylation that mark familial longevity are identified and the role of lipid profiles is investigated, given the previously described associations between body fat parameters, lipid status and glycosylation profiles [24] and between familial longevity and glucose homeostasis, as higher glucose levels were previously observed in the LLS long-lived families [161]. Finally, we assess whether glycan patterns are a marker of disease by testing association with the prevalence of diabetes mellitus, cerebrovascular accident (CVA) and myocardial infarction (MI).

Materials and Methods

Participants

In the Leiden Longevity study, Caucasian families were recruited if at least two long-lived siblings were alive and fulfilled the age-criterion of 89 years or older for males and 91 year or older for females, representing less than 0.1 % of the Dutch population in 2001 [160]. In total, 944 long-lived proband siblings were included, 1671 offspring with a mean age of 59 (st.dev 6.5) and 744 partners with a mean age of 58 (st.dev. 7.5).

The study protocol was approved by the Leiden University Medical Centre ethical committee and an informed consent was signed by all participants prior to participation in the study.

Phenotypic parameters

All serum measurements were performed with fully automated equipment. For insulin, the Immulite 2500 from DPC (Los Angeles, CA, USA) was applied. For glucose, total cholesterol, HDL-cholesterol (HDL-C) and triglycerides the Hitachi Modular or the Cobas Integra 800, both from Roche, Almere, the Netherlands were applied. LDL-cholesterol level (LDL-C) was calculated using the Friedewald formula (LDL-C = total cholesterol - HDL-C - (triglycerides/2.2); unit mmol/l) and set to missing if plasma triglyceride concentration exceeded 4.52 mmol/l.

N-glycan preparation

N-glycans from the total protein pool in plasma from the offspring as well as their partners were released, labeled with 2-aminobenzoic acid and purified using hydrophilic interaction liquid chromatography (HILIC) -SPE as previously described [97]. Briefly, 20 µl of 2% SDS were added to 10 µl of plasma, randomly distributed in 28 96-well plates. After protein denaturation and subsequent addition of NP-40, N-glycans were released overnight using PNGase F. Without intermediate purification, the N-glycans were labeled with 2-aminobenzoic acid (2-AA) in the presence of $NaBH_3CN$ and acetic acid for 2 hours at 65˚C. HILIC-SPE was performed using 40 mg cellulose in 96-well 0.45 µm GHP-filter plates (Pall Corporation, Ann Arbor, MI). All wells of the filter plate were washed using water and subsequently equilibrated using acetonitrile (ACN):water (80:20 v/v). The labeled N-glycan samples were then applied to the wells in 80% ACN, and the wells were washed using ACN:water (80:20 v/v). Purified 2-AA-labeled N-glycans were eluted into 0.8 ml deep well collection plates (Abgene via Westburg, Leusden, The Netherlands) using 400 µl water.

HILIC-HPLC analysis

Purified 2-AA-labeled N-glycans were separated using hydrophilic interaction high performance liquid chromatography (HILIC-HPLC) with trapping columns in dual mode as previously described [97]. In the Ultimate LC system (Dionex, Sunnyvale, CA) a Famos autosampler, a Switchos module with a loading pump and an Ultimate module with two pumps were connected. A nanovalve was used to connect the sys-

tem to the detector. Chromeleon software (Dionex) was used to control the system. Two 2.0 mm x 10 mm TSK gel Amide-80 trapping columns (Tosoh Biosciences, Stuttgart, Germany) and two 2.0 mm x 250 mm TSK gel Amide-80 analytical columns (Tosoh Biosciences) were used for the separation of 2-AA-labeled N-glycans, while a fluorescence detector (FP-2020 plus; Jasco, Easton, MD) was used for detection. Briefly, 50 µl of aqueous eluate from the HILIC-SPE were mixed with 150 µl ACN in a 96-deep-well plate. After a 20 µl injection, the 2-AA-labeled N-glycans were trapped on the trapping column and washed using ACN:50 mM ammonium formate (80:20, v/v; pH 4.4). Subsequently the 2-AA-labeled N-glycans were separated on the analytical column using a linear gradient of ACN (solvent A) and 50 mM ammonium formate (pH 4.4; solvent B), resulting in a total analysis time of 106 min.

Data pre-processing

HILIC-HPLC chromatograms were exported from Chromeleon wp 6.50 as ASCII files and were loaded into Matlab (version 2007a) software (The Mathworks, Inc., Natick, MA). The data were normalized and subsequently pre-aligned using the peak of highest intensity. After data reduction by cropping of the data to the range of 30 to 80 min, alignment of the data was performed by correlation optimized warping (COW) according to the method described by Skov et al.[170], which included reference sample generation, segment length and slack size optimization and alignment. Manual peak picking was performed and resulted in 26 areas under the curve.

Statistics

The samples of 2395 participants were divided over 28 individual plates to record plasma N-glycosylation patterns. To correct for batch effects, the 26 plasma N-glycosylation values were regressed on the categorical variable batch memberships. The standardized residuals of this model were used for further statistical analysis.

Since we have multiple offspring from the same family, the sandwich estimator was used to obtain valid standard errors. P-values <0.05 were regarded statistically significant. First, linear regression was used to explore relationships between each of the response variables and covariates - age and sex - adjusting for the family status

(offspring or partner of the offspring). To determine potential biomarkers for longevity, logistic regression was applied to investigate whether a glycosylation feature (independent variable) was predictive in classifying the family status after adjustment for age, sex, and their interaction. In that respect, the response variable (the family status) is coded as 0 (= partner) and 1 (= offspring of long-lived sibling).
Then, linear regression was used to explore relationships between LC-7 and LC-8 and covariates – BMI, levels of cholesterol, HDL-cholesterol, LDL-cholesterol, triglyceride and glucose as well as insulin activity - adjusting for the family status (offspring or partner of the offspring) age, sex and the age x sex interaction. To evaluate whether the association between LC-7 and LC-8 was independent of these covariates, the covariates were included in the logistic regression model for classification of family status. Finally, linear regression was used to explore relationships between LC-7 and LC-8 and disease status – incidence of MI, CVA or diabetes- adjusting for age, sex and their interaction.

Analyses were performed using STATA 10 (StataCorp LP, College Station, Texas, USA) and R version 2.9.0 (R Development Core Team).

Results

Analysis

Glycans were released from plasma proteins using PNGase F, labeled with 2-AA, and subsequently analyzed by HILIC-HPLC-FL as previously reported [97]. A typical HILIC-HPLC-FL chromatogram of 2-AA-labeled N-glycans from plasma is depicted in Figure 5 1, where only high-abundant glycans have been annotated. To allow statistical evaluation, chromatograms were normalized to the highest peak, aligned, and subsequently integrated according to the 26 intervals indicated. The repeatability of the sample preparation and profiling procedure using HILIC-HPLC-FL was reported previously, and was shown to remain below 9% for all glycan features [97].

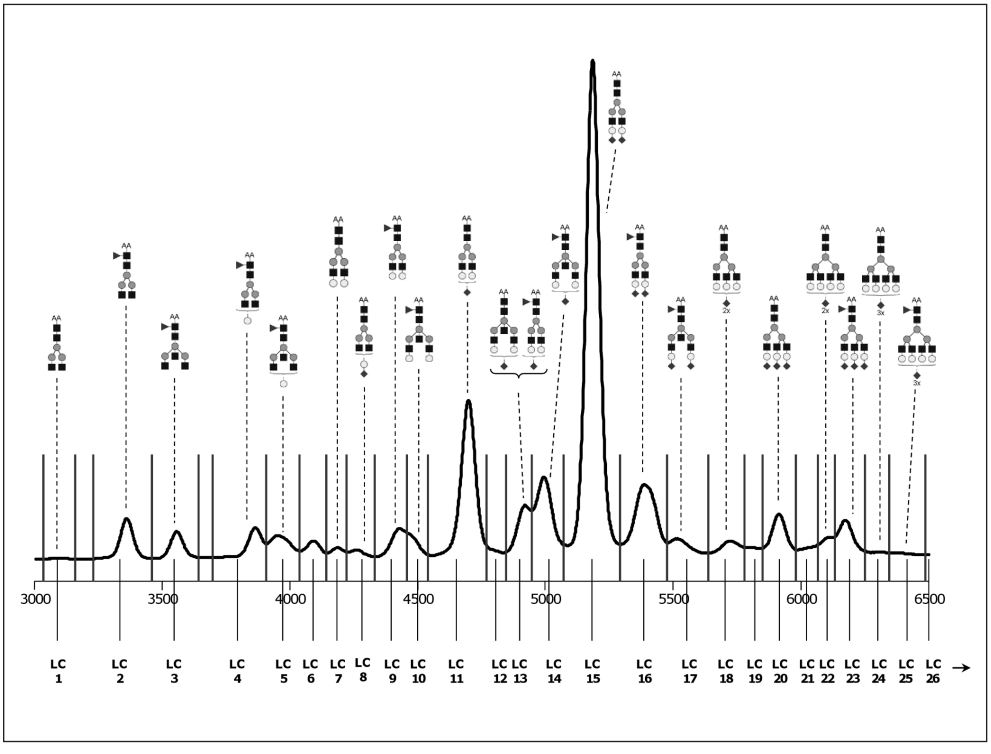

Figure 5 1. *A typical HILIC-HPLC-FL chromatogram of plasma derived N-glycans labeled with 2-AA. Only high-abundant glycans are annotated. Glyan compositions and structural schemes are given in terms of N-acetylglucosamine (square), mannose (dark circle), galactose (light circle), sialic acid (diamond) and fucose (triangle).*

Glycan integrals could be obtained for 2395 individuals from the LLS (1662 Offspring and 733 partners) between 30 and 80 years of age. The average age was 59.3 years. A summary of the numbers of individuals, age and gender are depicted in Table 5 1.

	total		offspring		partners		
No. of individuals	2395		1662		733		
	mean	Std / (95% C.I.)	mean	Std / (95% C.I.)	mean	Std./.(95% C.I.)	P
Age	59.3	6.8	59.4	6.6	58.9	7.3	0.079
% of male individuals	45%	50%	47%	50%	42%	49%	0.028
Body Mass Index	25.42	3.6	25.33	3.6	25.61	3.61	0.117
Totaal Cholesterol (mmol/L)	5.59	1.19	5.57	1.2	5.62	1.16	0.425
HDL cholesterol (mmol/L)	1.44	0.45	1.45	0.45	1.42	0.47	0.151
LDL cholesterol (mmol/L)	3.33	0.98	3.33	0.99	3.34	0.95	0.829
Triglyceride (mmol/L)	1.56	(0.68-3.65)	1.52	(0.66-3.63)	1.64	(0.72-4.14)	0.002
Glucose (mmol/L)	5.87	1.54	5.79	1.37	6.04	1.86	0.000
Insuline (mU/L)	16.52	(4.00-64.00)	16.01	(4.00-61.00)	17.71	(4.00-69.60)	0.008
CRP (mg/L)	1.42	(0.31-9.16)	1.41	(0.31-9.26)	1.46	(0.32-8.55)	0.506

Table 5 1 Descriptives of measures within the LLS. Descriptives are depicted for the total cohort as well as for offspring and partners of the offspring separately. P-values were generated by comparison of the offspring with their partners using univariate linear regression.

Plasma N-glycosylation is gender-specific and changes with calendar age

Several groups previously published on altered glycosylation patterns of plasma glycoproteins (both total plasma as well as specific IgG) with chronological age and for the different sexes [24;40-42;58;185]. A decreased galactosylation was observed with increasing calendar age, which was more profound in females than in males. To assess the effects of chronological age and sex in the current dataset, a linear regression model was used with the glycan feature as outcome. The results are shown in Table 5 2.

Clearly, several glycan features are strongly associated with calendar age. Non-galactosylated glycans in features LC-1 to LC-3 increase with age, while core fucosylated non- and mono-sialylated bigalactosylated glycans in features LC-9, LC-10, LC-13 and LC-14 were lowered with increasing age, thus corroborating the previsouly found degalactosylation with increasing calendar age [24]. Interestingly, the triantennary trisialylated glycan in feature LC-20 decreased with calendar age, while its fucosylated counterpart in LC-23 was positively correlated with calendar age.

Levels of several tetraantennary glycans (in glycan features LC-22 and LC-24) increased with calendar age.

Several glycan features showed different levels in males compared to females (see Table 5 2): the levels of a number of monosialylated glycans (in features LC-8, LC-11 and LC-14) decreased with age in males. Others, like LC-4, LC-16, LC-19 and LC-23 increased with age in males. Since the cohort has a wide age-range, we tested the interaction between age and sex in the plasma N-glycosylation features. Several glycan features (LC-7, LC-9, LC-10 and LC-13) showed significant positive interaction between age and sex, indicating that the decrease in the levels of these glycan features is more profound in females than in males. Therefore, age, sex and the age and sex interaction were included in all further analyses.

	Change with increasing calendar age	P	Sex-related differencee[a]	P	Sex-related change with calendar age[a]	P
LC-01	+	**0.000**	-	0.879	+	0.407
LC-02	+	**0.000**	+	0.022	-	0.013
LC-03	+	**0.000**	-	0.018	+	0.815
LC-04	-	0.127	+	**0.000**	-	0.590
LC-05	-	0.459	+	0.007	+	0.951
LC-06	-	0.442	-	0.039	+	0.432
LC-07	-	0.137	-	0.004	+	**0.001**
LC-08	+	0.010	-	**0.000**	+	0.013
LC-09	-	**0.000**	+	0.019	+	**0.000**
LC-10	-	**0.000**	+	0.107	+	**0.000**
LC-11	-	0.741	-	**0.000**	+	0.126
LC-12	+	0.917	-	0.735	+	0.020
LC-13	-	**0.000**	+	0.005	+	**0.000**
LC-14	-	**0.000**	-	**0.002**	+	0.005
LC-16	+	0.890	+	**0.000**	+	0.633
LC-17	-	0.045	-	**0.000**	+	0.113
LC-18	-	0.147	-	**0.000**	+	0.132
LC-19	+	0.145	+	**0.000**	+	0.022

LC-20	-	**0.001**	-	**0.000**	+	0.295
LC-21	+	0.113	-	**0.000**	+	0.172
LC-22	+	**0.001**	-	**0.000**	+	0.515
LC-23	+	**0.000**	+	**0.000**	+	0.110
LC-24	+	**0.000**	+	0.463	+	0.108
LC-25	+	0.098	-	**0.000**	+	0.262
LC-26	+	0.013	-	0.124	+	0.111
a Female =0, male=1						

Table 5 2. Associations of total plasma N-glycosylation with age, sex and the age x sex interaction. As normalization was performed using feature LC-15, no values are obtained for LC-15. Significant (P≤0.002, after Bonferroni correction) results are highlighted in bold.

Plasma N-glycosylation features are associated with longevity

The observation that N-glycosylation patterns change with age is not new and does not explain to which extent it marks physiological aging or subclinical and clinically diagnosed disease, in which case N-glycosylation patterns would provide biomarkers of health. A first indication of the value of the glycan features as biomarkers comes from a comparison of offspring of long-lived subjects and their partners. Thus, to identify N-glycan features that associate with familial longevity, for each glycan feature a logistic regression model was fitted with partner status as outcome. The results are depicted in Table 5 3, with significant (p≤0.002, after correction for multiple testing) results highlighted in bold.

Glycan feature	Odds Ratio	P	95% C.I.	
LC-01	1.096	0.041	1.004	1.196
LC-02	1.038	0.457	0.941	1.144
LC-03	1.046	0.378	0.947	1.154
LC-04	1.018	0.72	0.925	1.12
LC-05	1.045	0.37	0.949	1.151
LC-06	1.017	0.744	0.921	1.122
LC-07	**1.209**	**0.001**	**1.078**	**1.357**
LC-08	**1.174**	**0.001**	**1.068**	**1.291**
LC-09	1.035	0.503	0.936	1.144
LC-10	1.019	0.718	0.922	1.125

LC-11	1.085	0.087	0.988	1.191
LC-12	1.018	0.766	0.907	1.142
LC-13	0.996	0.93	0.902	1.099
LC-14	1.06	0.34	0.941	1.193
LC-15	-	-	-	-
LC-16	0.935	0.182	0.847	1.032
LC-17	1.038	0.461	0.94	1.145
LC-18	1.013	0.798	0.92	1.114
LC-19	1.079	0.111	0.982	1.186
LC-20	0.989	0.827	0.893	1.094
LC-21	1.014	0.773	0.922	1.115
LC-22	0.939	0.208	0.85	1.036
LC-23	1.027	0.596	0.931	1.133
LC-24	0.989	0.827	0.899	1.088
LC-25	1.046	0.336	0.955	1.146
LC-26	1.051	0.271	0.962	1.148

Table 5 3. *The associations between glycosylation and longevity. For all glycan features odds ratios are reported together with their 95% confidence interval and their respective p-values. Significant results are highlighted in bold (p≤0.002, after Bonferroni correction).*

The levels of two glycan features, both with non-fucosylated glycans as their major glycan (LC-7 and LC-8) are significantly different between the partners and the offspring. The odds ratios for LC-7 and LC-8 depicted in Table 5 3, are above 1, indicating that elevated levels of LC-7 and LC-8 could be observed in the offspring, relative to their partners. This is confirmed in Figure 5 2, where averages of the standardized values, adjusted for age, sex and the age x sex interaction are plotted. Thus increased levels of these N-glycan features reflect the propensity for longevity. Interestingly, these N-glycan features are not significantly associated to calendar age but they do have the tendency to change with age in a sex-specific manner .

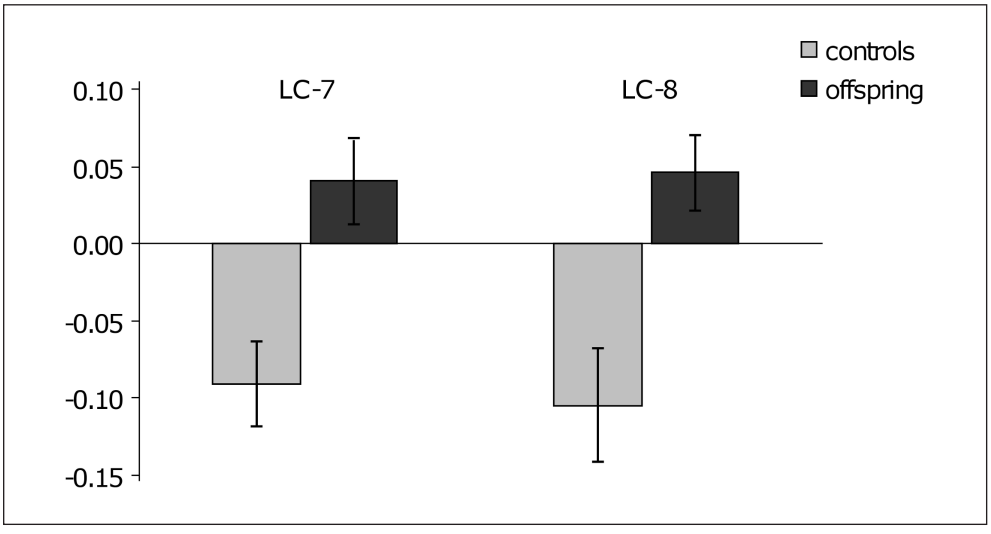

Figure 5 2. *Glycan features LC-7 and LC-8 are associated with familial longevity. Mean values for partners (light grey) and offspring (dark grey) are depicted with their S.E.M.*

Association of glycan features with other BMI and plasma lipid status parameters as well as glucose homeostasis

As the high-throughput analysis of total plasma N-glycans has only recently become feasible, not much is known regarding biochemical parameters that are associated with plasma glycosylation. However, in one recent study the association of plasma glycosylation patterns with body fat parameters and plasma lipid status (total cholesterol, HDL-cholesterol, LDL-cholesterol and triglyceride levels) has been evaluated [24], and several groups of glycans were found to be associated with such parameters. Moreover, one of the major marks of healthy aging in the LLS is the preserved glucose homeostasis and insulin sensitivity [161], illustrated by the fact that plasma levels of glucose and insulin activity are significantly lower in the offspring compared to the partners in the current study (see Table 5 1).

The plasma concentrations of glucose, total cholesterol, HDL-cholesterol, LDL-cholesterol and triglycerides as well as BMI and insulin activity have been obtained

from the participants from the LLS, and the association of these parameters with the glycosylation features was evaluated. Table 5 4 summarizes the results for LC-7 and LC-8. Interestingly, all effect sizes are larger for LC-8 than for LC-7. We did not find significant associations of the two glycan features with cholesterol, LDL-cholesterol and glucose levels. However, high levels of LC-8 associate with high HDL-cholesterol levels and low triglyceride levels and both glycan features are negatively correlated with BMI. These are all strong predicting phenotypes of cardiovascular health. The association of LC-8 levels follows the well known negative correlation of HDL-cholesterol levels on one hand and BMI and triglyceride levels on the other hand. High levels of LC-8 further associate with higher insulin activity.

	LC-7	LC-8
BMI	**-0.027 (0.000)**	**-0.066 (0.000)**
Cholesterol	0.010 (0.645)	-0.036 (0.062)
HDL-cholesterol	0.084 (0.106)	**0.371 (0.000)**
LDL-cholesterol	0.016 (0.586)	-0.040 (0.084)
Triglyceride	-0.080 (0.050)	**-0.392 (0.000)**
Glucose	-0.015 (0.151)	-0.034 (0.004)
Insulin activity	-0.063 (0.004)	**0.183 (0.000)**

Table 5 4. *Regression coefficients of the association of BMI and lipid parameters with glycosylation. P-values are depicted between parenthesis and significant results are highlighted in bold (p≤0.002, after Bonferroni correction).*

Next we evaluated whether the association of LC-7 and LC-8 with longevity was actually based on association with the metabolic parameters tested. Upon inclusion of these parameters, the association between longevity and glycosylation features remained significant: for LC-7 the P-values were 0.002, 0.006, 0.008 and 0.008 upon inclusion of BMI, HDL concentration, triglyceride levels and insulin activity, respectively. For LC-8 the P-values were 0.001, 0.005, 0.014 and 0.006 upon inclusion of BMI, HDL concentration, triglyceride levels and insulin activity, respectively. Therefore, the observed associations of LC-7 and LC-8 with familial longevity may be regarded independent of the metabolic parameters in Table 5 4. We also conclude

that LC-8 is the parameter showing the most prominent effects.

Plasma N-glycosylation features and their association to healthy aging

Then we tested whether LC-7 and LC-8 are indicative parameters of the health status of the individuals in the whole study. Given the results in Table 5 4 we would expect the features to mark cardiovascular health. A high level of plasma C-reactive protein (CRP), indicating an acute inflammatory response, is a marker for decreased health [193]. N-glycosylation has previously been reported to be correlated to CRP levels [33]. The associations between glycosylation and CRP-levels in total plasma were evaluated for LC-7 and LC-8. Both glycans were negatively associated with CRP levels (both $p<0.001$). This indicates that the glycosylation features are not only altered by familial longevity and metabolic health, but may also mark an individual's inflammatory status.

This leads us directly to the question whether LC-7 and LC-8 could be markers for metabolic health and cardiovascular disease. Even though the number of diagnosed individuals in the LLS study cohort is limited, we tested for association of the glycan features and metabolic disease: myocardial infarction (MI), cerebrovascular accident (CVA) and diabetes in the complete cohort (Table 5 5). No significant associations could be observed for CVA and diabetes. For myocardial infaction a boarderline significant negative association was observed, indicating that high levels of LC-8 are related to decreased incidence of MI.

	LC-7		LC-8	
	Coef.	P	Coef.	P
MI	-0.028	0.406	**-0.065**	**0.049**
CVA	0.019	0.679	0.013	0.753
diabetes	-0.004	0.828	0.005	0.824

Table 5-5. *Association between glycan fractions and health parameters. Significant results are highlighted in bold ($p<0.05$).*

As the numbers of individuals to which MI has occurred is low in the study population (2.8 %) up to now, it could well be that this association may become stronger over time, when more individuals have been diagnosed with MI.

Discussion

In this study we aimed to identify total plasma protein N-glycosylation based markers that reflect familial longevity and healthy aging in the LLS. We detected many N-glycan features that associated with calendar age. These observations were not new, as such relations were recently reported in a study in Healthy Croatian individuals [24]: In this study, the most considerable correlations with calendar age were observed for the biantennary nongalactosylated glycan, core-fucosylated glycans, disialylated forms of biantennary glycans, as well as nongalactosylated and digalactosylated glycans. Upon comparison of the middle aged offspring of nonagenarian siblings, representing 'healthy agers'and their partners as controls, however, two N-glycan features (LC-7 and LC-8) were identified that are more abundant in plasma of the offspring. These glycans did not show a significant association with calendar age, but they do have a tendency to change with age in a sex-specific manner.

Glucose homeostasis and insulin sensitivity were previously found to be associated with familial longevity in the LLS [161] and BMI and lipid profile parameters have previously been reported to be associated with plasma N-glycosylation patterns [24]. Thus we tested whether the glycosylation features LC-7 and LC-8 were associated with these metabolic phenotypes. LC-8 was found to be associated with BMI, levels of HDL-cholesterol and triglycerides and insulin activity, while LC-7 could only be associated with BMI. The association of LC-7 and LC-8 with familial longevity was, however, found to be independent of these metabolic phenotypes.

To allow the development of interventions aimed at disease-free aging, medical science would benefit from having markers that reflect healthy aging. The blood CRP-level has been described as a marker for acute inflammatory response, and thus reflects decreased health. As the two glycosylation features LC-7 and LC-8 are

both negatively correlated with CRP-levels, it may be concluded that these glycan-features also reflect an individual's inflammatory health status. LC-7 and LC-8 associated with several metabolic parameters, which indicated that these glycan features could also reflect metabolic health and/or cardiovascular disease. The offspring and partners have an average age of around 59 years (Table 5 1), and the prevalence of disease is still rather low in these individuals, however, a slightly significant association between incidence of MI and lower levels of LC-8 could be observed. Upon follow-up of the individuals of the LLS, this association might become stronger, as it is expected that more individuals will be diagnosed with metabolic diseases.

In the current study, features of glycosylation could be associated with familial longevity, independent of glucose homeostasis and insulin sensitivity. As the glycosylation features could also not be associated with blood glucose levels, it may be speculated that the regulation of glycosylation and glucose homeostasis in the LLS is controlled at different levels in the biochemical pathway.

Methods for high-throughput analysis of total plasma N-glycosylation patterns have only recently become available (e.g. [80;97]), and the implications of this approach are only starting to emerge. Some considerations have to be taken into account when interpreting data from such analyses. Interestingly, glycan features LC-7 and LC-8 show very similar behavior in their association with familial longevity, and these features elute subsequently during analysis. This raises the question whether these features are overlapping. However, there are also differences between the two features e.g. the association between glycosylation and insulin activity is not observed for LC-7, but is highly significant for LC-8, indicating that LC-7 and LC-8 do not completely overlap. Currently, the glycan annotation of the features is performed using fractionation of one standard plasma sample from a healthy individual [97], and only the most abundant glycans in the fractions are annotated. LC-7 is annotated as a di-antennary non-fucosylated and non-sialylated glycan, while LC-8 is annotated as a non-fucosylated di-antennary glycan with one truncated antenna and one sialylated antenna. Based on previous work and the GlycoBase database [135], which was

recently developed for the annotation of AB-labeled glycans in HILIC-LC analysis, it is clear that also less abundant glycans elute in glycan fractions LC-7 and LC-8. Therefore, there is a need to further identify the exact composition of the associating glycan features LC-7 and LC-8. To do so, it would be important to study the composition in a larger set of samples, derived from both offspring and partners. The most suitable technique for such analyses would be LC-MS, as the availability of mass spectrometric data largely facilitates the confirmation of the annotation.

As the current method comprises release of N-glycans from all plasma glycoproteins, the N-glycosylation pattern reflects N-glycan on this total glycoprotein pool. Altered glycosylation patterns may, therefore, be due to changes in plasma glycoprotein concentrations, but may also be caused by altered glycosylation of one or more glycoproteins. It could thus be beneficial to generate in parallel quantitative protein profiles, to allow identification of altered protein expression. Alternatively, the glycosylation pattern of a given specific glycoprotein may be monitored, similarly to our previous study on IgG glycosylation [185].

Given that the plasma profiles generated in this study reflect the glycosylation of a pool of glycoproteins, and assuming that the changes in glycosylation are caused by altered glycosylation and not by protein expression, it would be interesting to investigate whether the changes in the levels of these glycan(s) are caused by the attachment to a single glycoprotein, a group or even several groups of glycoproteins. Such future studies would need the use of extensive lectin-, antibody- or chromatographic- purification steps.

The regulation of glycosylation is a very complex cellular process, and the biological pathways involved in longevity and healthy aging have only started to become unraveled. Currently, there is no clear regulator that would link familial longevity with the regulation of the plasma N-glycan profile. To allow further in-depth analysis, the glycosylation in long-lived mice, such as the growth hormone receptor knock-out (GHRKO) mouse described in [194], could be studied. As mice can be bred under

standardized conditions, and additional parameters could be analysed more easily than in humans, such a model would facilitate the search for alterations associatied with longevity that could regulate protein N-glycosylation.

In conclusion, we found two glycan features, of which increased levels in human plasma mark familial longevity and healthy aging. The two N-glycan features are not correlated with glucose level, a previously found marker for human longevity [161], and may therefore be considered a novel group of markers. Further studies are needed to identify regulatory pathways that cause such altered glycosylation.