Supplementary Information (SI) for RSC Advances.

Supplemental Information

LOESS-Based Normalization Workflow for Targeted HDL Glycoproteomics in an Alzheimer's Disease Cohort

Brian V. Hong[1], Yiyun Liu[2], Armin Oloumi[2], Joanne K, Agus[1], Yasmine Bouchibti[2], Lee-Way Jin[3], Izumi Maezawa[3], Danielle J Harvey[4], Carlito B Lebrilla[2], Angela M Zivkovic[1]

[1]Department of Nutrition, University of California-Davis, Davis, CA 95616, USA

[2]Department of Chemistry, University of California-Davis, Davis, CA 95616, USA

[3]Department of Pathology and Laboratory Medicine, School of Medicine, University of California-Davis, Davis, CA 95817, USA

[4]Department of Public Health Sciences, University of California-Davis, Davis, CA 95616, USA

Supplementary Table 1. Participant characteristics. Summary of demographic and clinical variables for the 194 participants included in the HDL glycoproteomic analysis.

Characteristics	**Control**	**MCI**	**AD**	***F*(*df*, *df* Error)** **or χ^2(*df*)**	***p*-Value**
***n* combined**	83	40	71	n/a	n/a
APOE3E3	44	22	37	n/a	n/a
APOE3E4	39	18	34	n/a	n/a
Sex proportion, (male/female), combined	35:48:00	23:17	27:44:00	$\chi^2(2) = 4.10$	0.129
APOE3E3	18:26	12:10	15:22	$\chi^2(2) = 1.34$	0.511
APOE3E4	17:22	11:07	12:22	$\chi^2(2) = 3.18$	0.203
Age, years, mean ± SD, (*n*), combined	75.5 ± 7.0	78.0 ± 7.0	78.9 ± 7.2	*F*(2, 191) = 4.82	0.009 [ab]
	83	40	71		
APOE3E3	75.9 ± 7.0	78.6 ± 7.7	79.7 ± 8.4	*F*(2, 100) = 2.55	0.08
	44	22	37		
APOE3E4	75.0 ± 7.1	77.3 ± 6.2	78.1 ± 5.7	*F*(2, 88) = 2.28	0.108
	39	18	34		
BMI, kg/m^2, mean ± SD, (*n*), combined	28.1 ± 4.6	27.1 ± 4.2	26.6 ± 4.6	*F*(2, 181) = 2.00	0.138
	81	39	64		
APOE3E3	28.6 ± 4.5	27.8 ± 5.1	26.2 ± 4.7	*F*(2, 93) = 2.49	0.089
	43	21	32		
APOE3E4	27.5 ± 4.6	26.3 ± 2.7	27.0 ± 4.9	*F*(2, 85) = 0.42	0.658
	38	18	32		

Ethnicity proportion, (African American/Asian/Hispanic/White), combined	18:2:19:44	4:1:4:31	7:1:11:52	$\chi^2(6) = 10.97$	0.089
Diabetes at sample collection, %, (*n*), combined	34%	7.90%	4.90%	$\chi^2(2) = 22.94$	$<0.001^{abc}$
	76	38	61		
APOE3E3	39%	9.50%	0%	$\chi^2(2) = 19.30$	$<0.001^{abc}$
	41	21	32		
APOE3E4	29%	5.90%	10%	$\chi^2(2) = 5.64$	0.06
	35	17	29		
History of diabetes, %, (*n*), combined	39%	10%	5.60%	$\chi^2(2) = 28.50$	$<0.001^{abc}$
	83	39	71		
APOE3E3	39%	14%	0%	$\chi^2(2) = 19.51$	$<0.001^{abc}$
	44	21	37		
APOE3E4	38%	5.60%	12%	$\chi^2(2) = 8.52$	0.014^{abc}
	39	18	34		
History of hypercholesterolemia, %, (*n*), combined	74%	60%	54%	$\chi^2(2) = 6.70$	0.035^{ab}
	81	40	70		
APOE3E3	62%	64%	47%	$\chi^2(2) = 2.21$	0.331
	42	22	36		
APOE3E4	87%	56%	62%	$\chi^2(2) = 8.52$	0.014^{abc}
	39	18	34		

History of hypertension, %, (*n*), combined	72%	65%	76%	$\chi^2(2) = 1.56$	0.459
	83	40	71		
APOE3E3	75%	77%	76%	$\chi^2(2) = 0.04$	0.98
	44	22	37		
APOE3E4	69%	50%	76%	$\chi^2(2) = 3.84$	0.147
	39	18	34		

Abbreviations: AD = Alzheimer's disease dementia, BMI = body mass index, MCI = mild cognitive impairment. ANOVA were performed for continuous variables (age and BMI), and chi-squared (χ2) were performed for categorical variables (sex proportion, ethnicity proportion, diabetes at sample collection, history of diabetes, history of hypercholesterolemia, and history of hypertension). If significance is reached by ANOVA or χ2, Tukey's or Bonferroni's post hoc comparison was carried out, respectively. [a]Significance across all three groups. [b]Significance between the control compared with AD. [c]Significance between the control compared with MCI.

Supplementary Table 2. Coefficients of variation before and after normalization. Summary of QC mean intensity, standard deviation, and pre- and post-normalization coefficients of variation (CV Pre and CV Post) for all peptide and glycopeptide features exceeding the QC intensity threshold in the pooled HDL sample-preparation QC. The pooled QC, generated from all 194 participant HDL isolates, was used to assess analytical reproducibility and to select the optimal Sigma internal-standard peptide for each feature. CV Pre reflects raw measurement variability, and CV post reflects variability after LOESS drift correction and feature-specific internal-standard normalization.

Feature	Mean Intensity	SD Intensity	CV Pre (%)	CV Post (%)
A1AT_107_5401	936543.97	636107.52	67.9	42.5
A1AT_107_5411	445945.23	126646.67	28.4	27.4
A1AT_107_6501/6520	152797.93	78474.28	51.4	43.8
A1AT_107_6503	41242.52	12406.98	30.1	32.3
A1AT_107_6513	56467.15	22995.47	40.7	33.1
A1AT_271MC_5402	809341.26	221403.71	27.4	28.5
A1AT_70_5402	9206793.26	5913704.31	64.2	46.5
A1AT_AVLTIDEK	431194.24	371528.46	86.2	74.3
APOA1_DLATVYVDVLK	6880019.88	6751924.78	98.1	84.7
APOA1_LAEYHAK	101797993.98	40088933.41	39.4	27.5
APOA2_AGTELVNFLSYFVELGTQPATQ_z2	9385904.72	7770244.42	82.8	61.9
APOA2_SKEQLTPLIK_z2	1527049.00	1489372.85	97.5	92.2
APOA4_LAPLAEDVR_z2	590185.32	322556.43	54.7	39.0
APOA4_LGEVNTYAGDLQK_z2	20385.34	24291.29	119.2	89.3
APOA5_AQLLGGVDEAWALLQGLQSR_z3	407080.33	115234.44	28.3	27.8
APOB_2982_5421	56552.80	19095.44	33.8	34.7
APOB_GFEPTLEALFGK	61313.57	78551.32	128.1	122.7
APOC1_QSELSAK	185208.91	218737.51	118.1	81.7
APOC2_TYLPAVDEK_z2	15070397.38	10151103.22	67.4	43.1
APOC3_94AMC_11000_z3	36410.13	35145.45	96.5	92.1
APOC3_94A_11000_z2	109054.87	35012.28	32.1	30.2
APOC3_94Aoff_1101	33040.23	13528.89	40.9	34.6
APOC3_94MC_1101	23108.42	12569.18	54.4	46.3
APOC3_94MC_1102	133178.97	97501.50	73.2	62.7
APOC3_94_0300	1416295.65	572517.51	40.4	32.6
APOC3_94_0310	374451.92	98072.71	26.2	24.2
APOC3_94_1101	958903.81	337507.67	35.2	32.0
APOC3_94_1102	209663.63	99002.02	47.2	44.5
APOC3_DALSSVQESQVAQQAR	2881933.52	3055871.51	106.0	71.6
APOC3_GWVTDGFSSLK	1806719.18	1265961.77	70.1	51.3
APOD_65_6610	34148.58	32011.25	93.7	89.1
APOD_VLNQELR	17290155.66	10194224.89	59.0	55.2

APOE3/4_LAVYQAGAR	3173489.00	3152575.74	99.3	62.0
APOE_307/308_11010_z3	52857.52	54114.45	102.4	66.8
APOE_AATVGSLAGQPLQER_z2	8877758.47	5285755.11	59.5	37.6
APOE_LGPLVEQGR	11565889.58	7487096.80	64.7	37.4
APOF_273/274_11010_z3	98530.80	120561.91	122.4	94.2
APOF_273/274_11020_z3	22798.95	26473.86	116.1	94.1
APOF_SGVQQLIQYYQDQK_z2	152651.54	168002.82	110.1	77.2
APOL1_VTEPISAESGEQVER_z2	503533.43	677499.43	134.5	81.4
APOM_AFLLTPR	34597743.56	8896514.63	25.7	27.4
APOM_SLTSCLDSK	5555061.99	2468016.50	44.4	41.2
C1S_LLEVPEGR	158862.30	181946.89	114.5	74.3
CLUS_374_5420/5401	34791.92	34644.17	99.6	88.7
CLUS_ASSIIDELFQDR	58649.40	82672.46	141.0	101.6
FETUA_176_6501	34924.36	33197.39	95.1	88.8
FETUA_AHYDLR	124398.68	83943.16	67.5	39.7
LCAT_LEPGQQEEYYR_z2	34925.43	38463.38	110.1	106.1
LCAT_SSGLVSNAPGVQIR_z2	30944.12	37595.27	121.5	112.4
PLTP_245_5412	450420.05	161363.88	35.8	35.7
PLTP_GAFFPLTER_z2	222318.57	231741.42	104.2	65.0
PON1_253_43010_z3	61232.13	42310.35	69.1	65.0
PON3_29_76210_z3	35448.39	32257.73	91.0	84.7
SAA1_FFGHGAEDSLADQAANEWGR_z3	92799.44	154625.71	166.6	89.6
SAA2_GPGGAWAAEVISNAR_z3	28581.88	34207.97	119.7	115.2
SAA4_94_5401_z4	615936.51	271964.14	44.2	35.8
SAA4_94_5420_z4	627466.16	297024.22	47.3	34.5
SAA4_94_6522	153721.53	49618.63	32.3	26.7
SAA4_FRPDGLPK_z2	120591.26	120565.86	100.0	75.0

Supplementary Table 3. Pooled HDL QC performance for the 21 analytes included in downstream statistical modeling. For each feature, the table reports mean intensity, standard deviation, and pre- and post-normalization coefficients of variation (CV) in the pooled HDL QC samples.

Feature	Mean Intensity	SD Intensity	CV Pre (%)	CV Post (%)
A1AT_107_5411	445945.23	126646.67	28.4	27.4
A1AT_271MC_5402	809341.26	221403.71	27.4	28.5
A1AT_70_5402	9206793.26	5913704.31	64.2	46.5
APOA1_LAEYHAK	101797993.98	40088933.41	39.4	27.5
APOA4_LAPLAEDVR_z2	590185.32	322556.43	54.7	39.0
APOA5_AQLLGGVDEAWALLQGLQSR_z3	407080.33	115234.44	28.3	27.8
APOC2_TYLPAVDEK_z2	15070397.38	10151103.22	67.4	43.1
APOC3_94A_11000_z2	109054.87	35012.28	32.1	30.2
APOC3_94_0300	1416295.65	572517.51	40.4	32.6
APOC3_94_0310	374451.92	98072.71	26.2	24.2
APOC3_94_1101	958903.81	337507.67	35.2	32.0
APOC3_94_1102	209663.63	99002.02	47.2	44.5
APOC3_GWVTDGFSSLK	1806719.18	1265961.77	70.1	51.3
APOD_VLNQELR	17290155.66	10194224.89	59.0	55.2
APOE_AATVGSLAGQPLQER_z2	8877758.47	5285755.11	59.5	37.6
APOE_LGPLVEQGR	11565889.58	7487096.80	64.7	37.4
APOM_AFLLTPR	34597743.56	8896514.63	25.7	27.4
APOM_SLTSCLDSK	5555061.99	2468016.50	44.4	41.2
FETUA_AHYDLR	124398.68	83943.16	67.5	39.7
PLTP_245_5412	450420.05	161363.88	35.8	35.7
SAA4_94_5420_z4	627466.16	297024.22	47.3	34.5

Supplementary Table 4. Internal standard assignment for each peptide and glycopeptide feature. Sigma peptides serving as internal standards selected for normalization of each HDL-associated peptide or glycopeptide. Selection was performed using pooled HDL QC injections by identifying, for each feature, the Sigma peptide that minimized post-normalization coefficient of variation (CV). Only features passing the QC intensity threshold (mean QC intensity > 20,000) and detected in the participant dataset were included. The table reports the feature name and assigned internal standard.

Sigma	Feature
sigma_1	CLUS_291_5421/5402
sigma_3	APOA1_LAEYHAK
sigma_3	FETUA_AHYDLR
sigma_4	APOD_VLNQELR
sigma_5	APOM_SLTSCLDSK
sigma_6	APOE_LGPLVEQGR
sigma_7	APOA4_LAPLAEDVR_z2
sigma_7	APOC2_TYLPAVDEK_z2
sigma_8	APOE_AATVGSLAGQPLQER_z2
sigma_9	APOM_AFLLTPR
sigma_10	APOC3_GWVTDGFSSLK
sigma_12	APOC3_94_0310
sigma_12	APOC3_94_0300
sigma_12	APOC3_94A_11000_z2
sigma_12	APOC3_94_1102
sigma_13	A1AT_107_5411
sigma_13	A1AT_271MC_5402
sigma_13	SAA4_94_5420_z4
sigma_13	A1AT_70_5402
sigma_13	APOA5_AQLLGGVDEAWALLQGLQSR_z3
sigma_13	PLTP_245_5412

Supplementary Table 5. APOE genotype-associated differences in HDL glycopeptide and peptide features. Unadjusted Wilcoxon rank-sum p values, Benjamini-Hochberg FDR-adjusted q values, and covariate-adjusted p values from linear models including age, ethnicity, and clinical status. Direction reflects the median abundance in APOE3/E4 compared with APOE3/E3 carriers. Six peptides showed nominal genotype-associated differences; five remained significant after FDR correction ($q < 0.05$), and all six remained significant after covariate adjustment.

Feature	p (Unadjusted)	q (FDR)	p (Covariate-Adjusted)	Direction (E4 vs E3)
A1AT_271MC_5402	0.021	0.075	0.019	Lower in E3/E4
APOD_VLNQELR	0.011	0.046	0.011	Lower in E3/E4
APOE_AATVGSLAGQPLQER_z2	2.57×10^{-4}	0.003	2.81×10^{-4}	Lower in E3/E4
APOE_LGPLVEQGR	4.93×10^{-5}	0.0010	5.33×10^{-5}	Lower in E3/E4
APOM_AFLLTPR	0.011	0.046	0.010	Lower in E3/E4
APOM_SLTSCLDSK	0.002	0.013	0.002	Lower in E3/E4

Supplementary Table 6. Summary of MRM LC-MS acquisition and run-design parameters used in the current study.

Category	Parameter	Value
Sample preparation before injection	Post-digest handling before LC-MRM	Proteolysis was quenched by acidification with formic acid prior to LC-MRM analysis.
LC system	Liquid chromatography platform	Agilent 1290 Infinity LC system
MS system	Mass spectrometer	Agilent 6490 triple quadrupole mass spectrometer (Agilent Technologies)
Acquisition mode	Targeted MS mode	Multiple reaction monitoring (MRM)
Analytical column	LC column	Agilent Eclipse Plus C18, RRHD 1.8 µm, 2.1 × 100 mm
Guard column	Precolumn	Agilent Eclipse Plus C18 precolumn, RRHD 1.8 µm, 2.1 × 5 mm
Injection volume	Sample injected per run	1.0 µL
Run time	LC gradient duration	25 min binary gradient
Mobile phase A	Solvent A	3% acetonitrile, 0.1% formic acid in nanopure water (v/v)
Mobile phase B	Solvent B	90% acetonitrile, 0.1% formic acid in nanopure water (v/v)
Flow rate	LC flow rate	0.5 mL/min
Transition design	Transition requirements	Monitored transitions required predetermined LC retention time and collision-induced dissociation behavior from prior glycopeptide characterization
Peak integration software	Data system	Agilent MassHunter Quantitative Analysis B.5.0
Targeted panel size	Total targeted transitions in current study	164
Internal standards	Synthetic peptide internal standards	12 Sigma peptides
Study sample run design	Participant samples analyzed	194 participant HDL samples
Plate structure	Sample layout	3 × 96-well plates
Acquisition window	Total analytical sequence	20 days (20 September to 9 October 2023)
Instrument QC	Daily system-stability injection	Commercial human HDL preparation injected at the start of each acquisition day

Pooled sample-preparation QC	QC material	QC generated by pooling equal-volume aliquots of HDL from all 194 study participants
QC frequency	QC injection schedule	Injected at the beginning of each plate-row block, approximately every 10 study samples
Total pooled QC injections	Number of QC injections	20