

RESEARCH

Open Access



# Comparison of intact protein and digested peptide techniques for high throughput proteotyping of ApoE

Anthony Maus<sup>\*</sup>, Dan Figdore, Dragana Milosevic, Alicia Algeciras-Schimmich and Joshua Bornhorst

## Abstract

**Introduction:** Apolipoprotein E (ApoE) genotyping has been shown to have diagnostic value in the evaluation of cardiovascular diseases and neurodegenerative disorders such as Alzheimer's disease. Although genetic testing is well established for this application, liquid chromatography-mass spectrometry (LC-MS) has the potential to provide a high throughput, low-cost alternative for ApoE evaluation.

**Methods:** Serum samples were analyzed by peptide, intact protein, and genomic techniques. For peptide analysis, samples were digested with trypsin followed by liquid chromatography-tandem mass spectrometry analysis (LC-MS/MS) using a high-throughput multichannel LC system coupled to a Sciex 7500 mass spectrometer. For intact protein analysis, ApoE was immuno-purified using a monoclonal antibody immobilized on magnetic beads followed by high-resolution LC-MS analysis using an Exploris 480. DNA was extracted and evaluated using Sanger sequencing as a reference method.

**Results and discussion:** The peptide measurement method produced one discrepant result when compared to genomic sequencing (out of 38 sequenced samples), whereas the intact protein analysis followed by deconvolution resulted in two discrepant results and when the intact protein data was processed with chromatographic integration there were three discrepant results. Therefore, the intact protein method proved slightly less accurate, required longer analysis time, and is substantially more costly, while providing only a 30 min improvement in sample preparation time.

**Conclusions:** With current MS technology clinical laboratories appear to be better served to utilize trypsin digest sample preparation and LC-MS/MS as opposed to high-resolution LC-MS intact protein analysis techniques for evaluation of ApoE proteotype. Peptide analysis methods are capable of producing accurate results with high throughput and minimal cost.

**Keywords:** Apolipoprotein E, Alzheimer's disease, Tandem mass spectrometry, High-resolution mass spectrometry

## Introduction

Apolipoprotein E (ApoE) is a 34 kDa protein that is a major component in cholesterol-rich very low density lipoprotein (VLDL) and is of relatively high abundance in some subclasses of high density lipoproteins (HDL), such as HDL<sub>1</sub> and HDL<sub>c</sub> [1]. There are three common isoforms of ApoE, known as ApoE2, ApoE3, and ApoE4, corresponding to the respective alleles in the ApoE gene, which have frequencies in the global population of approximately 8%, 78% and 14%, respectively.[2, 3].

\*Correspondence: Maus.anthony@mayo.edu

Department of Laboratory Medicine and Pathology, Division of Clinical Biochemistry and Immunology, Mayo Clinic, 200 First Street SW, Rochester, MN 55905, USA



© The Author(s) 2022. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

These genetic differences result in different amino acids at two positions in the sequence, with ApoE2 possessing cysteine in 112 and 158, ApoE3 has cysteine in position 112 and arginine in 158, and ApoE4 has arginine in both the 112 and 158 positions. The combinations of these alleles give rise to 6 possible genotypes, with both homozygous (E2/E2, E3/E3, E4/E4) and heterozygous (E2/E3, E3/E4, E2/E4) being prevalent at various frequencies in the global population [4].

The ApoE genotype has implications for cardiovascular diseases and neurodegenerative disorders [1, 5]. ApoE2 homozygosity has been shown to be related to type III hyperlipoproteinemia, which ultimately causes premature atherosclerosis [6, 7]. The ApoE4 genotype is associated with increased abundance of low-density lipoproteins and risk of atherosclerosis [7, 8]. The relationship between ApoE genotype and neurodegenerative disorders has also been extensively explored. The presence of ApoE4 protein has been associated with negative ramifications in regards to traumatic brain injury, stroke, frontotemporal dementia, Down syndrome, Parkinson's disease, and Lewy body disease [1, 9], in addition to having been shown to increase the risk of Alzheimer's disease (AD) three–fourfold in heterozygotes and about 9–15 fold in ApoE4 homozygotes compared to non-carriers of the ApoE4 protein [2, 10, 11]. In contrast, several studies show that the ApoE2 genotype reduces the risk of cognitive impairment and AD [10, 12]. One notable example from Reiman et al. indicated that ApoE2 homozygotes have a 66% decrease in AD risk compared to ApoE2/ApoE3 heterozygotes, an 87% reduction relative to ApoE3 homozygotes, and a 99.6% reduction compared to ApoE4 homozygotes [10, 13].

Although the relationship between ApoE genotype, cardiovascular diseases, and neurodegenerative disorders is well established, the diagnostic role of quantitative plasma ApoE concentration measurements are more controversial [5]. Increased ApoE concentrations have generally been shown to be related to increased levels of triglycerides and cholesterol, which are well known drivers of cardiovascular diseases [14–16]. However, some studies indicate a relationship between ApoE concentrations and neurocognitive decline, whereas others do not suggest such a relationship [17–21]. Therefore, at this time no clear clinical benefit for quantitative ApoE testing has been established.

Genetic testing has been utilized as the standard for ApoE classification determination for decades [22, 23]; however, liquid chromatography-mass spectrometry (LC-MS) has the potential to provide a low cost, high throughput alternative to genetic testing with the additional potential benefit of being capable of producing precise and accurate quantitative results in cases where both protein identity and concentration are desired.

Several investigations predicated on measurement of trypsin digested peptides have been published [3, 4, 24, 25]. These methods rely on the detection of four target peptides derived from the ApoE protein to determine the proteotype. A particularly relevant example of this methodology was recently published by Brkovic et al. [26]. This peptide analysis method used online sample purification, a cornerstone of clinical LC-tandem mass spectrometry (LC-MS/MS) analyses, and had an analysis time of only 6.5 min. However, the method did use overnight digestion, which would negatively impact turnaround time in a clinical testing environment. Additionally, this method utilized microflow flow rates, which are not widely employed in routine clinical laboratories.

Given the relatively moderate size of ApoE (34 kDa) analysis of the intact protein is an attractive alternative to trypsin digestion. Hu et al. demonstrated this when they performed immuno-purification of ApoE, followed by analysis using matrix-assisted laser desorption/ionization-time of flight-mass spectrometry (MALDI-TOF-MS) [27]. This intact protein analysis technique is easy to perform and high-throughput; however, the lack of mass or chromatographic resolution of different ApoE isoforms would make interpretation in a routine clinical laboratory challenging.

In this investigation, we aimed to explore the potential of trypsin digestion followed by analysis of peptides using LC separation and a triple quadrupole mass spectrometer, as well as immuno-purification (IP) followed by analysis of the intact ApoE protein using LC separation and a high-resolution Exploris 480 Orbitrap mass spectrometer.

## Materials and methods

### Chemicals

Water was purified using a Barnstead Nanopure system (ThermoFisher Scientific, Waltham, MA). Phosphate Buffered Saline (PBS), tris base, hydrochloric acid, and trifluoroacetic acid (TFA) were purchased from Fisher Scientific. LC-MS grade acetonitrile (ACN), formic acid, isopropanol (IPA), and bovine serum albumin (BSA), were purchased from MilliporeSigma (Burlington, MA). Isotopically labeled peptides used as internal standards (IS) for the four target peptides were synthesized by the Proteomics Core at Mayo Clinic using standard 9-fluorenylmethoxycarbonyl (Fmoc) chemistry on a Liberty Blue (CEM Corp. Matthews, NC) peptide synthesizer with methods suggested by the manufacturer.

### Trypsin Digest and low-resolution LC–MS/MS Analysis

First, 25  $\mu\text{L}$  of serum sample was put in a 96 well plate, followed by 25  $\mu\text{L}$  of internal standard, 50  $\mu\text{L}$  of 1 M tris–HCl (pH 8) buffer, and 400  $\mu\text{L}$  of water. Proteins were reduced by adding 50  $\mu\text{L}$  of 100 mM dithiothreitol (MilliporeSigma) and incubated at 48 °C for 45 min. Next, proteins were alkylated by adding 125  $\mu\text{L}$  of 100 mM iodoacetamide (MilliporeSigma) and incubating in the dark for 30 min. Proteins were digested for 1 h at 37 °C following the addition of 100  $\mu\text{L}$  of 1 mg/mL trypsin (Worthington, Lakewood, NJ). The digestion was stopped by the addition of TFA to a final concentration of 0.2%.

LC separation was performed using a Thermo Transcend TLX-4 TurboFlow system. Digests (10  $\mu\text{L}$ ) were injected onto a  $\text{C}_8$  cartridge (Phenomenex, Torrance, CA, Item number: AJO-6073) with an internal diameter of 2 mm and a length of 4 mm using the loading pump to deliver 2% B at 1 mL/min for 1 min. Next, the sample was eluted off the cartridge at 100  $\mu\text{L}/\text{min}$  for 1.25 min using 40% B, and this eluate was mixed via a tee with the eluting pump flowing 2% B at 400  $\mu\text{L}/\text{min}$  prior to loading onto the analytical column. The analytical column was a Kinetex PS- $\text{C}_{18}$  with an inner diameter of 3 mm, a length of 50 mm, and 2.6  $\mu\text{m}$  particles (Phenomenex). Separation was performed at 600  $\mu\text{L}/\text{min}$  using a gradient from 5%B to 13% B over 6 min. The column was then washed for 3 min and equilibrated at starting conditions for 1 min. Therefore, the total LC method time is 12.5 min, but when multiplexed across the 4 channels, results are produced effectively every 3.125 min.

Tandem mass spectrometry analysis was performed using a Sciex 7500 mass spectrometer. Source conditions and MS/MS parameters can be seen in Additional file 1: Tables S1 and S2, respectively. After the LC–MS/MS measurement, chromatographic peak areas for two fragments from the IS and analyte signals were integrated in Sciex OS. The retention times corresponding to the IS were integrated regardless of analyte signal intensity. The resulting integrated areas were then exported to Microsoft Excel for further processing. The analyte signals were corrected by dividing by the IS peak areas. This IS corrected result was then used to determine positive/negative for a given peptide, which was ultimately used to determine the proteotype.

### Immuno-purification of ApoE and high-resolution LC–MS analysis

Antibody purification was performed using a rabbit monoclonal antibody purchased from ThermoFisher Scientific (catalog #-701,241) coupled to tosyl activated magnetic beads (ThermoFisher Scientific, catalog#-14,204).

For coupling, beads are washed with a 0.1 M boric acid (MilliporeSigma) buffer pH 9.5. Antibody was then incubated overnight with the beads at 37 °C in a solution of 500  $\mu\text{L}$  of 0.1 M boric acid buffer pH 9.5 and 500  $\mu\text{L}$  of 3 M ammonium sulphate (MilliporeSigma) in 0.1 M boric acid buffer pH 9.5. After overnight incubation, remaining active sites are blocked using a solution of 5 mg/mL BSA in PBS and incubating for 1 h at 37 °C. Next, the beads were washed three times with 0.1% Tween-20 (MilliporeSigma, catalog #-P-1379) in PBS. After final washing, beads were resuspended in 0.1% Tween-20 in PBS and refrigerated for future use.

When performing immuno-purification of ApoE, 10  $\mu\text{g}$  of bead coupled antibody was added to 500  $\mu\text{L}$  of patient serum followed by a 1.5 h incubation at room temperature. The beads were then washed twice with PBS, twice with water, and eluted with 50  $\mu\text{L}$  of 20% acetonitrile, 0.4% TFA in water. This eluate was transferred to autosampler vials and injected without further manipulation.

LC separation was conducted using a ThermoFisher Scientific Vanquish Duo, which allows for separation across two channels; however, the method could be easily transferred to a TLX-4 if throughput improvement were necessary. Sample (20  $\mu\text{L}$ ) was injected onto an Agilent (Santa Clara, CA) Poroshell 300SB  $\text{C}_3$  column with a 2.1 mm ID and 75 mm in length (catalog #660,750–909) with the pump delivering 20% B at 400  $\mu\text{L}/\text{min}$ . Starting conditions were held and separation was performed using a gradient up to 50% B over 20 min. The column was then washed for 7 min and reequilibrated at starting conditions for 3 min. Therefore, the total LC method time is 30 min, but when multiplexed across the 4 channels, individual results are produced effectively every 7.5 min.

High-resolution mass spectrometry analysis was performed using an Exploris 480 mass spectrometer. The instrument scanned from  $m/z$  1000–3000 at a resolution of 240,000. Additional mass spectrometry parameters can be seen in Additional file 1: Table S3. The resulting mass spectra were then analyzed using extracted ion chromatogram (XIC) integration and deconvolution of the spectra. XICs were produced in TraceFinder (ThermoFisher Scientific) by summing 6 isotopes from the 8 most abundant charge states (48 total  $m/z$ ), integrated, and exported to Microsoft Excel. We determined this to be the optimal number of isotopes, and adding additional lower intensity isotope signals either did not benefit or negatively impacted S/N of the chromatograms. Spectral deconvolution was performed using Thermo BioPharma Finder. The ApoE proteoforms were chromatographically separated; therefore, the deconvoluted spectra were produced using the “average over selected retention time” function in the software and retention times were set

as mean retention times for a given protein  $\pm 0.3$  min. Deconvoluted spectral intensities were also exported to Microsoft Excel for box-plot generation and proteotype determination.

### DNA sequencing

Cell free DNA was extracted from residual serum samples with the Qiagen Circulating Nucleic Acid kit (Qiagen, Valencia, CA) using the manufacturer's protocol. The extraction took approximately 120 min. The status of the sequence variants in the DNA sequence (NM\_000041.4) corresponding to ApoE protein was evaluated by Sanger sequencing. A targeted PCR reaction was performed to amplify region containing both SNP's (rs429358 and rs7412). Universal primers were then used to sequence these regions using ABI BigDye BigDyeTerminators v1.1 (ThermoFisher Scientific). The PCR and sequencing required 155 min. The sequencing traces were analyzed manually with Mutation Surveyor software (SoftGenetics, LLC, State College, PA) (Table 1).

### Samples and human subjects

Random deidentified clinical residual serum samples ( $n=276$ ) were obtained for this study. After initial screening by the peptide analysis method as described, 41 of these samples were analyzed using intact protein analysis mass spectrometry-based techniques and by cell-free DNA sequencing. This selected subset was chosen to disproportionately represent atypical ApoE genotypes (see Table 2). The Mayo Clinic Rochester Institutional Review Board approved this study as exempt.

### Results and discussion

A diagram summarizing workflows for the two different mass spectrometry-based approaches used in this investigation is shown in Fig. 1 (created in BioRender). Initial screening for a suitable sample set for intact protein analysis was done by performing trypsin digestion on 276 residual serum samples and the resulting peptides were measured using a Thermo TLX-4 coupled to a Sciex 7500. From this, a subset of 41 samples was selected to undergo intact protein analysis and genomic sequencing and the

results are shown in Table 2. Example chromatograms for three patients with ApoE proteotypes of E3/E4 (patient 1), E2/E3 (patient 11), and E2/E4 (patient 16) are shown in Fig. 2. Box plots showing the resulting distribution of IS corrected peak areas for the 4 target peptides are shown in Fig. 3. Based on these plots, thresholds for determining positive and negative were set between these apparent distributions to assign positive and negative status for a given peptide and ultimately determine proteotype. When performing peptide analysis, differences in IS corrected intensity were very large between samples assigned as positive and negative. The closest difference between a positive and negative assignment was a factor of 8 in IS corrected area for the total set, and a factor of 25 for the smaller subset of 41 samples. The wide discrepancy is further evidenced by the very small p-values shown in Fig. 3. This large discrepancy between the positive and negative samples allowed for easy differentiation between the two distributions and ultimate assignment of the ApoE proteotype. All the samples exhibited a signal with a S/N ratio greater than 50 for peptide LAVYQAGAR; therefore, all samples were considered positive for this peptide. This peptide would only be absent in the case of ApoE2 homozygotes, which is the rarest possible genotype [2, 3]. Thus, our conclusion that our screening sample set possessed no ApoE2 homozygotes was plausible.

The selected subset of 41 samples then underwent IP followed by high-resolution LC-MS analysis on an Exploris 480. Signals from the intact protein analyses were processed using traditional integration of XICs and spectral deconvolution in BioPharma Finder. Example chromatograms from the aforementioned E3/E4, E2/E3, and E2/E4 patients are shown in Additional file 1: Fig. S2. These chromatograms are of relatively low S/N. The box plots of peak areas and coinciding p-values shown in Additional file 1: Fig. S3 also indicate a relatively small separation between the positive and negative samples, with only approximately a factor of 2 separating the positives and negatives for the ApoE3 and E4 proteins. These factors made accurately assigning the correct proteotype by this method more difficult.

The benefits of using spectral deconvolution for detection of ApoE proteins were also explored. Deconvolution was done by averaging spectra in the chromatographic window corresponding to the three target proteins, which are chromatographically separated. Representative spectra from the aforementioned E3/E4, E2/E3, and E2/E4 patients are shown in Fig. 4. The deconvoluted spectra at the established retention times are compared to the theoretical average mass to determine the ApoE proteotype. Box plots of the deconvoluted spectral signal intensities are shown in Additional file 1: Fig. S4. Similar to the peptide approach described above, intensity thresholds

**Table 1** Peptide detection criteria for ApoE proteotyping

Peptides Detected	ApoE Proteotype
CLAVYQAGAR, LGADMEDVCGR	E2/E2
CLAVYQAGAR, LAVYQAGAR, LGADMEDVCGR	E2/E3
LAVYQAGAR, LGADMEDVCGR	E3/E3
LAVYQAGAR, LGADMEDVCGR, LGADMEDVR	E3/E4
CLAVYQAGAR, LAVYQAGAR, LGADMEDVCGR, LGADMEDVR	E2/E4
LAVYQAGAR, LGADMEDVR	E4/E4

**Table 2** Results from the various proteotyping techniques utilized in this work compared to genomic sequencing

Patient #	Intact Protein Deconvolution	Intact Protein Chromatogram	Peptide	Genomic Sequencing
1	E3/E4	E3/E4	E3/E4	E3/E4
2	E3/E4	E3/E4	E3/E4	E3/E4
3	E3/E3	E3/E3	E3/E3	E3/E3
4	E3/E4	E3/E4	E3/E4	E3/E4
5	E3/E3*	E3/E3*	E3/E3*	E3/E4*
6	E3/E4	E3/E4	E3/E4	E3/E4
7	E3/E3	E3/E3	E3/E3	E3/E3
8	E3/E3	E3/E3	E3/E3	E3/E3
9	E3/E3	E3/E3	E3/E3	E3/E3
10	E3/E4	E3/E4	E3/E4	E3/E4
11	E2/E3	E2/E3	E2/E3	E2/E3
12	E3/E3	E3/E3	E3/E3	Unable Seq*
13	E3/E4	E3/E4	E3/E4	Unable Seq*
14	E3/E3	E3/E3	E3/E3	E3/E3
15	E3/E3	E3/E3	E3/E3	E3/E3
16	E2/E4	E2/E4	E2/E4	E2/E4
17	E3/E3	E3/E3	E3/E3	E3/E3
18	E3/E3	E3/E3	E3/E3	E3/E3
19	E3/E3	E3/E3	E3/E3	E3/E3
20	E3/E4	E3/E4	E3/E4	E3/E4
21	E2/E3	E2/E3	E2/E3	E2/E3
22	E3/E3	E3/E3	E3/E3	E3/E3
23	E3/E3	E3/E3	E3/E3	E3/E3
24	E3/E3	E3/E3	E3/E3	E3/E3
25	E3/E4	E3/E4	E3/E4	E3/E4
26	E2/E3	E2/E3	E2/E3	E2/E3
27	E2/E3	E2/E3	E2/E3	E2/E3
28	E2/E3	E2/E3	E2/E3	E2/E3
29	E4/E4	E4/E4	E4/E4	Unable Seq*
30	E2/E3	E2/E3	E2/E3	E2/E3
31	E3/E3	E3/E3	E3/E3	E3/E3
32	E3/E3	E3/E3	E3/E3	E3/E3
33	E3/E3	E3/E3	E3/E3	E3/E3
34	E3/E3	E3/E3	E3/E3	E3/E3
35	E3/E3	E3/E3	E3/E3	E3/E3
36	E2/E4	E2/E4	E2/E4	E2/E4
37	E4/E4	E4/E4	E4/E4	E4/E4
38	E2/E4	E4/E4*	E2/E4	E2/E4
39	E4/E4	E4/E4	E4/E4	E4/E4
40	E3/E3	E3/E3	E3/E3	E3/E3
41	E3/E3*	E3/E3*	E3/E4	E3/E4

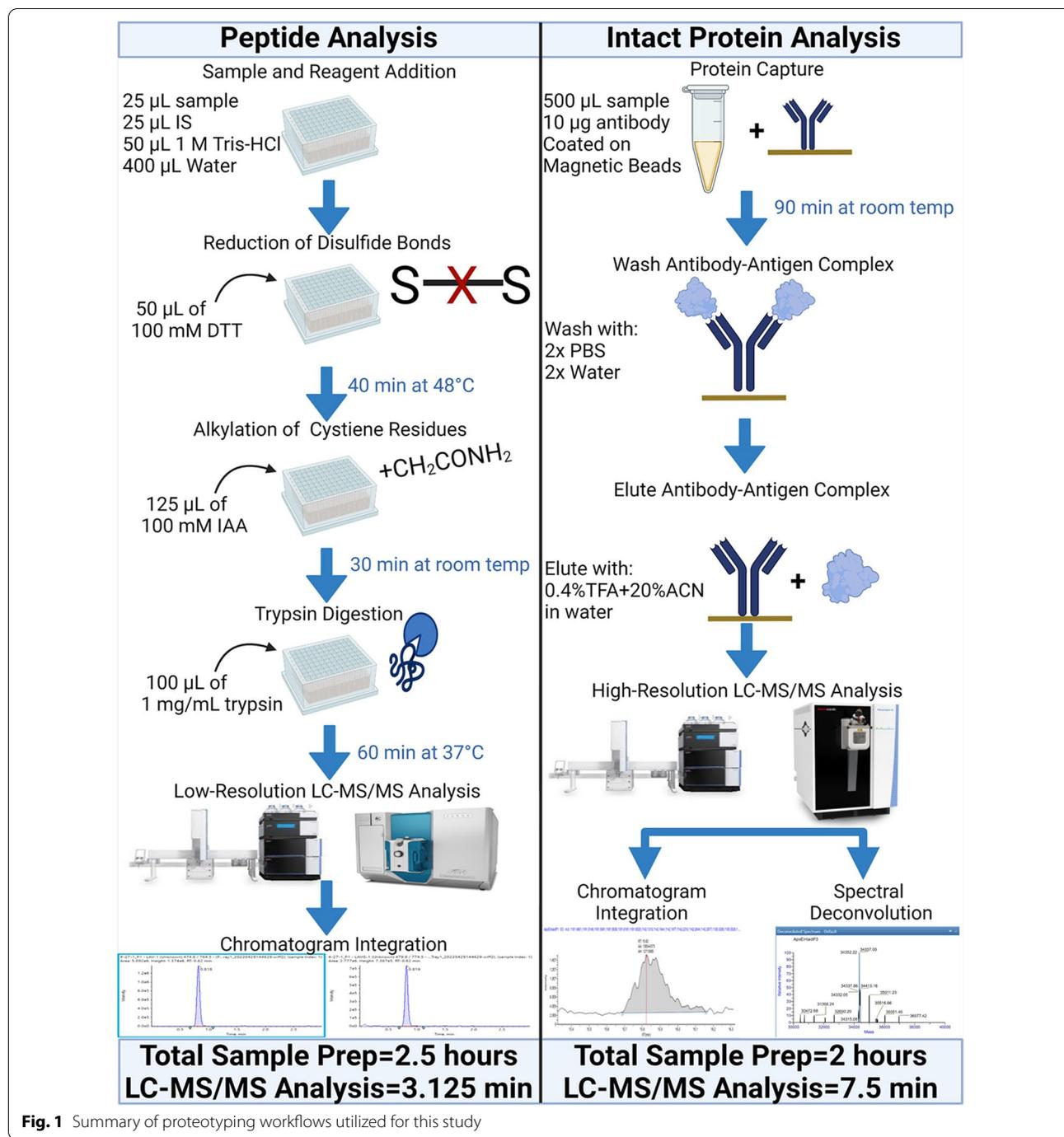
Three samples were unable to yield sufficient sequencing data. The results from the mass spectrometry-based techniques disagreed with sequencing results for patient 5. The intact methods failed to detect the E4 protein for patient 41, and the chromatogram data processing method did not detect E2 for patient 38

\* indicates the Discrepant results

were set between the apparent positive and negative distributions for the respective proteins. The minimum separation between positive and negative assignments

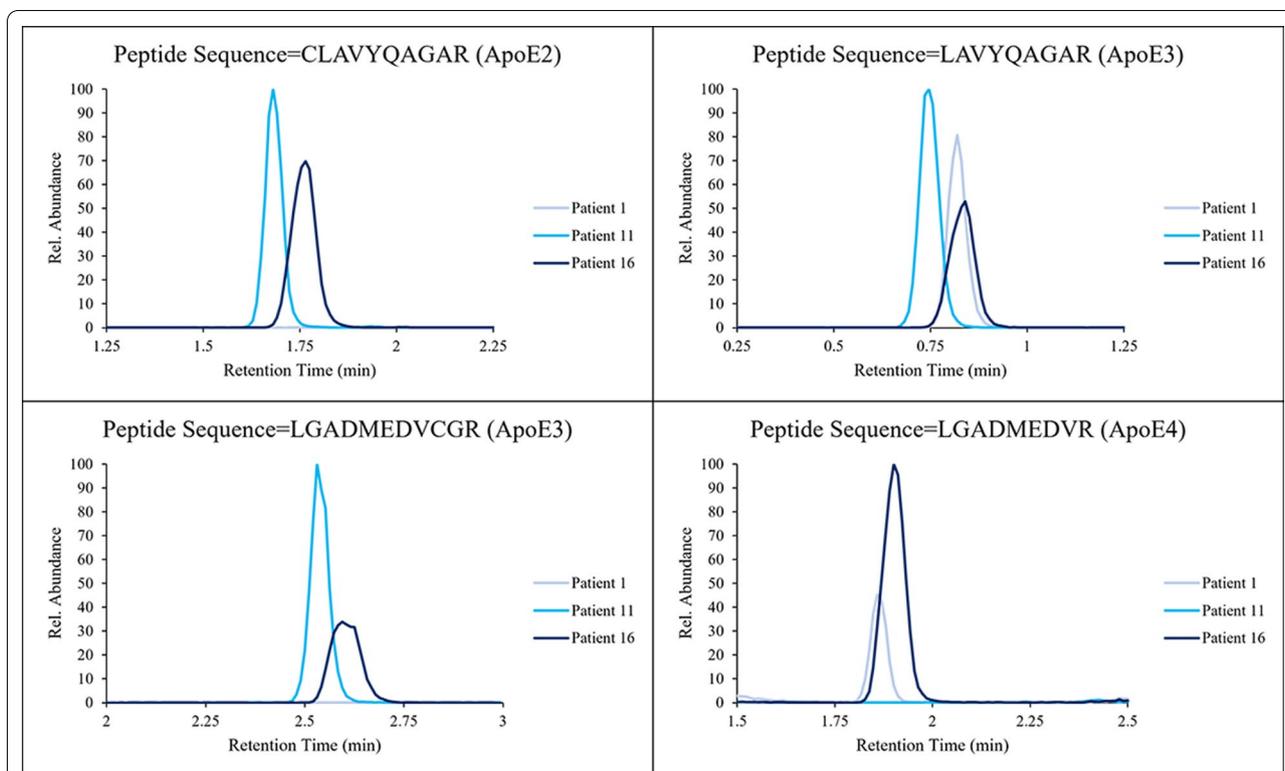
was observed for E4, which differed by a factor of 3.7 and resulted in the relatively large p-value of  $1.88 \times 10^{-2}$ .

The subset of 41 samples that underwent both trypsin digestion and subsequent intact protein characterization



were also analyzed by ApoE genomic sequencing in order to compare assessment by genotyping and proteotyping methods (Table 2). The accuracy of these mass spectrometry-based techniques compared to the genomic sequencing are shown in Table 3. The chromatographic integration of LC-MS signals was the least accurate methodology with three discrepant results compared

to genomic sequencing likely due to the factors discussed previously. Spectral deconvolution resolved one of the discrepancies of the intact method in comparison to sequencing results. Peptide analysis resulted in only a single discrepancy when compared to the genomic sequencing results. The results for this discrepant sample matched for all of the mass spectrometry-based



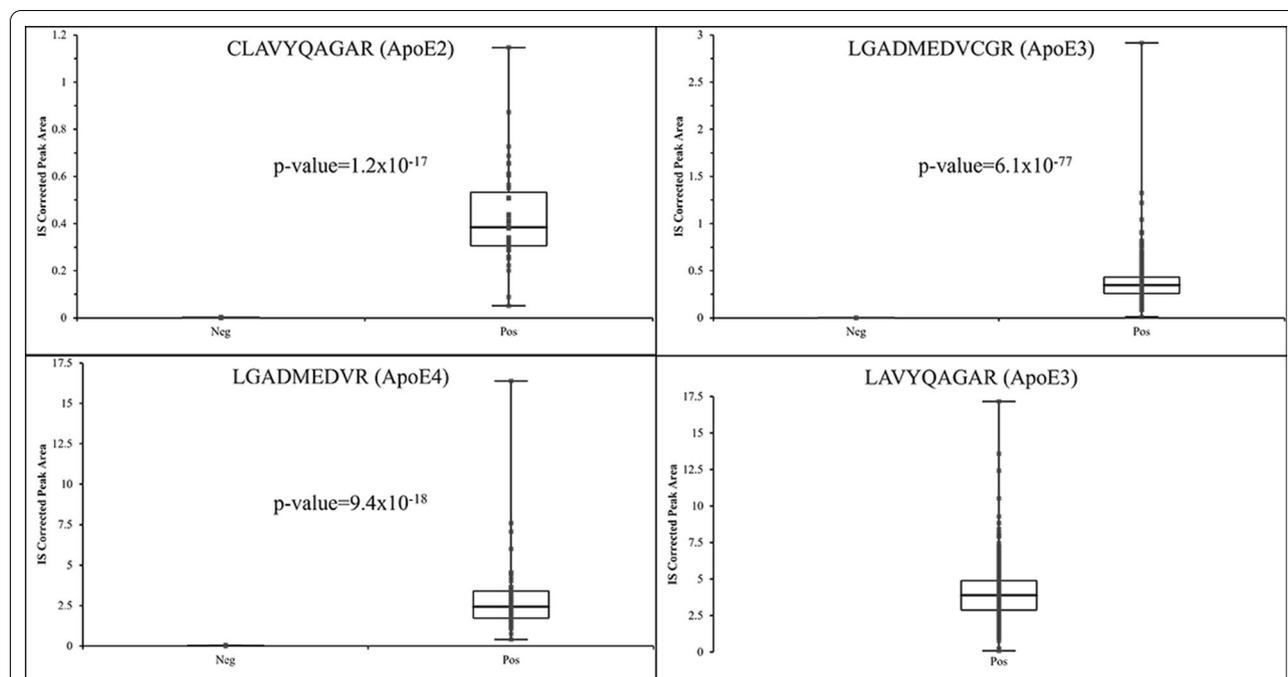
**Fig. 2** Representative chromatograms of the 4 target peptides from three patients. Patient 1 expressed the ApoE3 and ApoE4 proteins, patient 11 expressed the ApoE2 and ApoE3 proteins, and patient 16 expressed the ApoE2 and ApoE4 proteins. Note that the slight difference in retention time is due to the slight variation across the four lc channels on the multiplex system

techniques, but were contradictory to the sequencing results. The sample was reanalyzed by the peptide technique and yielded the same result, but there was insufficient sample for repeat genomic sequencing. Therefore, the reason for the discrepancy is unknown. If deconvolution is utilized to process the intact protein spectra, both the peptide and the intact method mischaracterized the same E3/E4 genotype sample as an E3/E3 and the intact method mischaracterized an additional E3/E4 genotype sample as an E3/E3 proteotype sample.

Overall, this work demonstrates the viability of intact protein workflows for LC-MS proteotyping of ApoE. Additionally, the comparison of chromatographic and deconvolution-based approaches is relevant to clinical laboratories developing tests targeting intact proteins. Most clinical mass spectrometry tests are predicated on the integration and quantitation of chromatograms; whereas the field of intact proteomics widely relies on deconvolution for qualitative analyses. As quantitative clinical analyses of larger proteins become more prevalent, the benefits, precision, and accuracy of deconvolution must be carefully vetted and clearly demonstrated. Our data indicates that deconvolution improves the S/N and coincidingly the ability to characterize the proteins,

which aligns with research applications. Assessing the long-term quantitative performance was beyond the scope of this work.

Intact protein analysis clearly has several limitations. Most importantly, in the small sample set analyzed in this investigation, both data processing techniques resulted in more discrepancies compared to genomic sequencing than did the peptide analysis method. This can be attributed to the lower sensitivity to the presence of ApoE of the intact protein technique compared to the peptide detection method. Intact protein analysis yielded lower S/N, as evidenced by the reduced difference between the positive and negative samples, generally lower p-values, and chromatograms. To overcome the S/N challenges with intact protein analysis, a large quantity of antibody was used to increase the signal. We initially tried 5  $\mu$ g as described by Hu and coworkers [27], but found our methodology benefited from the approximately  $2 \times$  gain in S/N when using 2 times more antibody (data not shown). However, the cost of this strategy would likely be prohibitive in a routine testing environment. A state-of-the-art high-resolution mass spectrometer was also used for the intact protein analysis. Although high-resolution mass spectrometers are becoming more common in clinical



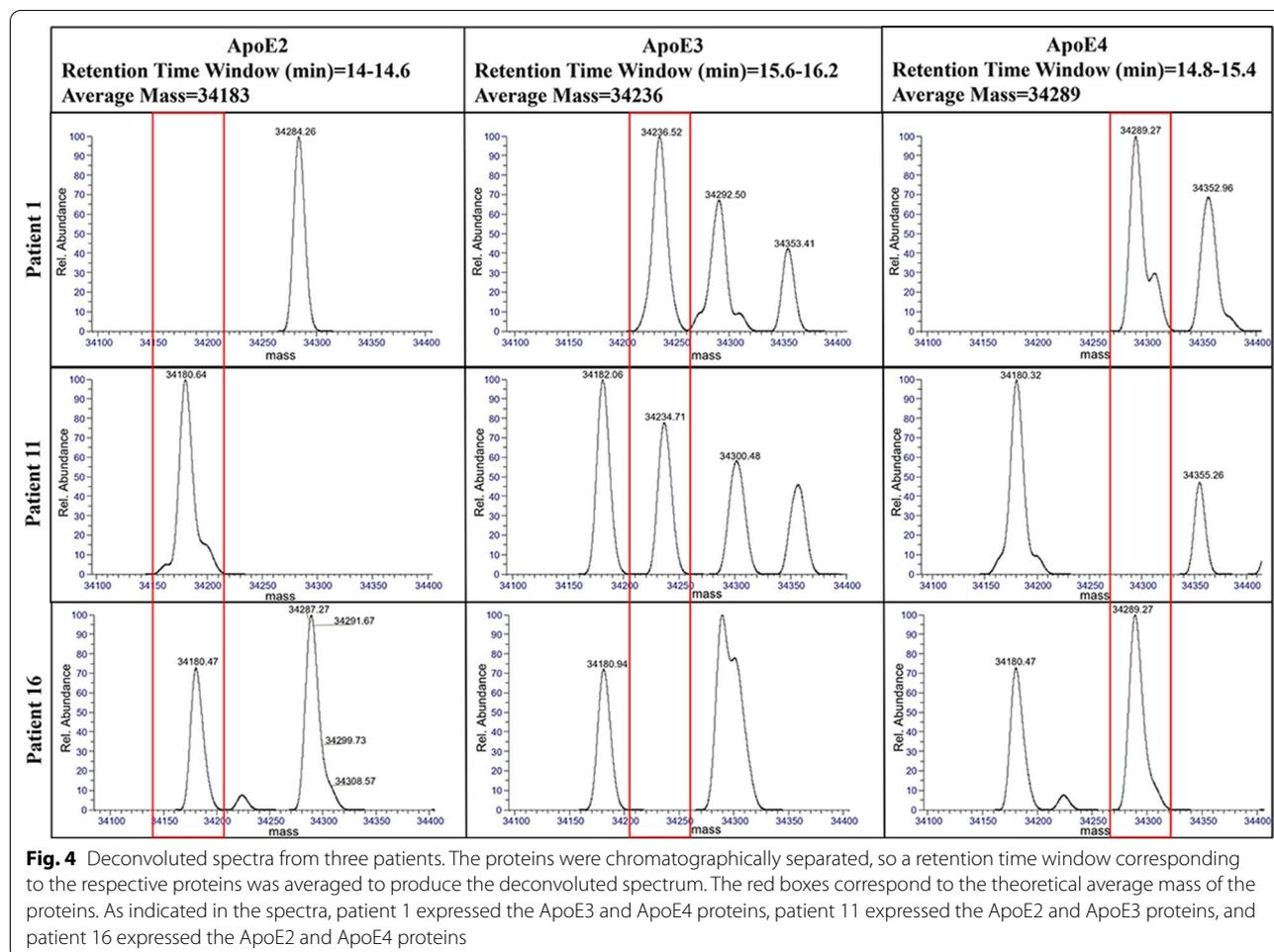
**Fig. 3** Box plots of the distribution of IS corrected peak areas derived from LS-MS/MS measurement of the ApoE peptides necessary for proteotyping. The p-value from a Welch's t-test comparing the positive and negative peak area distributions is inset. No samples without LAVYQAGAR peptide were found in this investigation. As indicated by the p-value and box plots, there was a large discrepancy between signals assigned as positive and negative enabling differentiation. Box plots represent medians and quartiles with whiskers extending across the result distribution

labs, they are not nearly as well established as the triple quadrupole mass spectrometer used for peptide analysis, and therefore, not as preferable.

Potential differences in the run-time and approximate cost should be considered when selecting a viable ApoE proteotyping method. Intact protein analysis requires IP from 500  $\mu$ L of sample and two hours of sample preparation. This process can be automated, but was quite costly as 10  $\mu$ g of antibody was required to yield a sufficient signal-to-noise (S/N) ratio. Peptide analysis requires only 25  $\mu$ L of sample and is relatively inexpensive. The sample preparation is slightly more time consuming and laborious; however, the higher S/N of the resulting peptide signals allowed for a  $2\times$  decrease in chromatographic separation time and several commercial platforms are available for automation of these types of procedures. A summary comparison of the materials cost, sample preparation time, and analysis time can be seen in Table 4. This clearly demonstrates the financial advantages of the peptide analysis method, as the list price of the materials for protein measurements cost \$55.00 and sequencing was \$36.02, whereas the peptide measurement materials were only \$0.08. Based on the publication by Hu et al., MALDI-TOF-MS yields higher S/N, but would still require expensive antibody-based purification and the

resulting spectra would be difficult to interpret in the routine clinical laboratory due to the lack of chromatographic and mass resolution.

Another significant challenge when performing analysis of intact proteins such as ApoE is finding an appropriate internal standard to compensate for variation in instrument performance and sample preparation. Isotopically labeled peptides are easily synthesized for relatively low cost. However, isotopically labeled proteins are often not commercially available or are prohibitive from a cost perspective. This investigation provides a quintessential example of these challenges. The ApoE peptide analysis utilized isotopically labeled peptides, which makes the analysis more reproducible overall. These implications are magnified when setting thresholds for positive and negative as done in this work. The use of isotopically labeled internal standards would make the use of a positive/negative intensity threshold more robust over the short and long term. In contrast, no feasible internal standard material was found for the intact protein analysis. Even if an isotopically labeled protein existed, if one added it prior to the antibody capture this would reduce the capacity for the target protein and lower the LC-MS signal from the protein. Therefore, a myriad of factors (primarily the lack of availability) made the use of an internal standard



**Table 3** Percent accuracy of the mass spectrometry-based techniques relative to genomic sequencing

Genotype	Number of samples (based on genomic sequencing)	% Accuracy intact protein deconvolution	% Accuracy intact protein chromatogram	% Accuracy peptide
E2/E2	0	100%	100%	100%
E2/E3	6	100%	100%	100%
E3/E3	18	100%	100%	100%
E3/E4	9	78%	78%	89%
E2/E4	3	100%	66%	100%
E4/E4	2	100%	100%	100%

impractical for the intact protein analysis measurement. This likely contributes to the comparatively lower performance of the protein analysis observed herein, and would likely negatively impact the ability to apply positive/negative thresholds over the long term.

The peptide analysis methodology presented within this study has several advantages over methods described in previous publications. The sample preparation is completed in approximately 2.5 h and is highly amenable to

automation. The multiplex analytical flow LC and triple quadrupole mass spectrometer used are widely available in routine clinical laboratories today, well suited for online sample purification, and are capable of producing results approximately every 3 min. Although quantitation was not a goal of our work, with the addition of an external calibration curve this methodology would be well-suited to produce accurate and precise quantitative results.

**Table 4** Comparison of the materials cost based on list prices, sample preparation time, and analysis time estimates when performing peptide, protein, and genomic analyses

	Peptide analysis	Intact protein analysis	Genomic sequencing
Materials Cost per Sample (\$)	0.08	55.00	36.02
Sample Preparation Time (H)	2.5	2	2
Instrument Analysis Time (min)	3.125	7.5	155

This investigation has several limitations. Due to the relatively high cost of genomic and intact protein analyses, only a subset of the initial 276 samples could be analyzed by these techniques. Based on the initial screening, a sample set was selected that disproportionately represented the lower frequency genotypes, but it would be ideal to have a larger sample set with a wider variety of genotypes, including an ApoE2 homozygous sample, one of which was not identified in the initial screening. It would also be advantageous to obtain another sample from the subject that produced the discrepancy between the mass spectrometry and genomic methodologies. Performing a longer-term study to assess the performance of these methodologies and the applicability of the thresholds established herein over time would also be beneficial, but such an effort is beyond the scope of this work. Use of an external calibrator to yield quantitative results and to establish thresholds as opposed to signal intensity metrics would also be a possible improvement for future work.

## Conclusions

Our results demonstrate that measurement of intact ApoE is a viable means of proteotyping. We used immuno-purification followed by high-resolution LC-MS analysis and spectral deconvolution to yield results that were in over 90% agreement with established genomic sequencing techniques. However, when compared to the high throughput, inexpensive peptide analysis method developed on the triple quadrupole mass spectrometer, the intact protein method is less accurate, more costly, more laborious, less robust, and requires more advanced instrumentation. Therefore, until LC-MS technology advances clinical laboratories are much better served to perform these types of analyses on digested peptides.

## Abbreviations

ApoE: Apolipoprotein E; VLDL: Very low density lipoprotein; HDL: High density lipoproteins; AD: Alzheimer's disease; LC-MS: Liquid chromatography mass spectrometry; LC-MS/MS: Liquid chromatography-tandem mass spectrometry; PBS: Phosphate Buffered Saline; ACN: Acetonitrile; IPA: Isopropanol; BSA: Bovine serum albumin; IS: Internal standards; FMOC: Fluorenylmethoxycarbonyl; TFA: Trifluoroacetic acid; XIC: Extracted ion chromatogram; S/N: Signal-to-noise.

## Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12014-022-09379-5>.

**Additional file 1: Table S1.** Sciex 7500 source conditions. **Table S2.** Sciex 7500 parameters. **Table S3.** Exploris 480 parameters. **Fig. S1.** Box plots of the distribution of IS corrected peak areas derived from LC-MS/MS measurement of the ApoE peptides necessary for proteotyping for only the subset of samples selected for subsequent analyses. **Fig. S2.** Representative chromatograms of the intact protein XICs. **Fig. S3.** Box plots of the distribution of peaks areas from the integration of chromatographic peaks when performing LC-MS measurement of intact ApoE. **Fig. S4.** Box plots of the distribution of deconvoluted spectral signal intensities when performing LC-MS measurement of intact ApoE.

## Author contributions

AM, DF, and DM performed the analysis; AM and JB wrote the main manuscript text. All authors read and approved the final manuscript.

## Funding

Not applicable.

## Availability of data and materials

The mass spectrometry data have been deposited in the PeptideAtlas SRM Experiment Library with the data set identifier PASS01770.

## Declarations

### Ethics approval and consent to participate

This study has been approved by the Mayo Clinic Institutional Review Board.

### Competing interests

The authors have no competing interests as defined by BMC, or other interests that might be perceived to influence the results and/or discussion reported in this paper.

Received: 22 July 2022 Accepted: 11 October 2022

Published online: 15 November 2022

## References

- Mahley RW. Apolipoprotein E: from cardiovascular disease to neurodegenerative disorders. *J Mol Med (Berl)*. 2016;94:739–46. <https://doi.org/10.1007/s00109-016-1427-y>.
- Farrer LA, Cupples LA, Haines JL, Hyman B, Kukull WA, Mayeux R, Myers RH, Pericak-Vance MA, Risch N, van Duijn CM. Effects of age, sex, and ethnicity on the association between apolipoprotein E Genotype and Alzheimer disease: a meta-analysis. *JAMA*. 1997;278(16):1349–56. <https://doi.org/10.1001/jama.1997.03550160069041>.
- Minta K, Brinkmalm G, Janelidze S, Sjödin S, Portelius E, Stomrud E, Zetterberg H, Blennow K, Hansson O, Andreasson U. Quantification of total Apolipoprotein E and its isoforms in cerebrospinal fluid from patients with neurodegenerative diseases. *Alzheimer's Res Ther*. 2020;12(1):19. <https://doi.org/10.1186/s13195-020-00585-7>.
- Blanchard V, Ramin-Mangata S, Billon-Crossouard S, Aguesse A, Durand M, Chemello K, Nativel B, Flet L, Chétiveaux M, Jacobi D, Bard J-M, Ouguerram K, Lambert G, Krempf M, Croyal M. Kinetics of Plasma

- Apolipoprotein E Isoforms by LC-MS/MS: a pilot study. *J Lipid Res.* 2018;59(5):892–900. <https://doi.org/10.1194/jlr.P083576>.
5. Rasmussen KL. Plasma levels of Apolipoprotein E, APOE genotype and risk of dementia and ischemic heart disease: a review. *Atherosclerosis.* 2016;255:145–55. <https://doi.org/10.1016/j.atherosclerosis.2016.10.037>.
  6. Mahley RW, Huang Y, Rall SC. Pathogenesis of Type III Hyperlipoproteinemia (Dysbetalipoproteinemia): Questions, Quandaries, and Paradoxes. *J Lipid Res.* 1999;40(11):1933–49. [https://doi.org/10.1016/S0022-2275\(20\)32417-2](https://doi.org/10.1016/S0022-2275(20)32417-2).
  7. Mahley RW, Rall SC. Apolipoprotein E: far more than a lipid transport protein. *Annu Rev Genomics Hum Genet.* 2000;1:507–37. <https://doi.org/10.1146/annurev.genom.1.1.507>.
  8. Bennet AM, Di Angelantonio E, Ye Z, Wensley F, Dahlin A, Ahlborn A, Keavney B, Collins R, Wiman B, de Faire U, Danesh J. Association of Apolipoprotein E genotypes with lipid levels and coronary risk. *JAMA.* 2007;298(11):1300–11. <https://doi.org/10.1001/jama.298.11.1300>.
  9. Mahley RW, Huang Y. Apolipoprotein e sets the stage: response to injury triggers neuropathology. *Neuron.* 2012;76(5):871–85. <https://doi.org/10.1016/j.neuron.2012.11.020>.
  10. Husain MA, Laurent P, Plourde M. APOE and Alzheimer's Disease: from lipid transport to physiopathology and therapeutics. *Front Neurosci.* 2021;15:67.
  11. Neu SC, Pa J, Kukull W, Beekly D, Kuzma A, Gangadharan P, Wang L-S, Romero K, Arneric SP, Redolfi A, Orlandi D, Frisoni GB, Au R, Devine S, Auerbach S, Espinosa A, Boada M, Ruiz A, Johnson SC, Kosciak R, Wang J-J, Hsu W-C, Chen Y-L, Toga AW. Apolipoprotein E Genotype and Sex risk factors for Alzheimer Disease: a meta-analysis. *JAMA Neurol.* 2017;74(10):1178–89. <https://doi.org/10.1001/jamaneurol.2017.2188>.
  12. Li Z, Shue F, Zhao N, Shinohara M, Bu G. APOE2: Protective mechanism and therapeutic implications for Alzheimer's disease. *Mol Neurodegener.* 2020;15(1):63. <https://doi.org/10.1186/s13024-020-00413-4>.
  13. Reiman EM, Arboleda-Velasquez JF, Quiroz YT, Huentelman MJ, Beach TG, Caselli RJ, Chen Y, Su Y, Myers AJ, Hardy J, Paul Vonsattel J, Younkin SG, Bennett DA, De Jager PL, Larson EB, Crane PK, Keene CD, Kamboh MI, Kofler JK, Duque L, Gilbert JR, Gwirtsman HE, Buxbaum JD, Dickson DW, Frosch MP, Ghetti BF, Lunetta KL, Wang L-S, Hyman BT, Kukull WA, Foroud T, Haines JL, Mayeux RP, Pericak-Vance MA, Schneider JA, Trojanowski JQ, Farrer LA, Schellenberg GD, Beecham GW, Montine TJ, Jun GR. Exceptionally low likelihood of Alzheimer's dementia in APOE2 homozygotes from a 5,000-person neuropathological study. *Nat Commun.* 2020;11(1):667. <https://doi.org/10.1038/s41467-019-14279-8>.
  14. Mooijaart SP, Berbée JFP, Heemst D, van Havekes L M., Craen A. J. M. de; Slagboom P E., Rensen P. C. N., Westendorp R. G. J. ApoE plasma levels and risk of cardiovascular mortality in old age. *PLoS Med.* 2006;3(6): e176. <https://doi.org/10.1371/journal.pmed.0030176>.
  15. Boerwinkle E, Utermann G. Simultaneous effects of the apolipoprotein E polymorphism on Apolipoprotein E, apolipoprotein B, and cholesterol metabolism. *Am J Hum Genet.* 1988;42(1):104–12.
  16. Haddy N, Bacquer DD, Chemaly MM, Maurice M, Ehnholm C, Evans A, Sans S, Martins MC, Backer GD, Siest G, Visvikis S. The Importance of Plasma Apolipoprotein E concentration in addition to its common polymorphism on inter-individual variation in lipid levels: results from Apo Europe. *Eur J Hum Genet.* 2002;10(12):841–50. <https://doi.org/10.1038/sj.ejhg.5200864>.
  17. Siest G, Bertrand P, Qin B, Herbeth B, Serot J-M, Masana L, Ribalta J, Passmore AP, Evans A, Ferrari M, Franceschi M, Shepherd J, Cuchel M, Beisiegel U, Zuchowsky K, Rukavina AS, Sertic J, Stojanov M, Kostic V, Mitrevski A, Petrova V, Sass C, Merched A, Salonen JT, Tiret L, Visvikis S. Group13, the A Apolipoprotein E Polymorphism and Serum Concentration in Alzheimer's Disease in Nine European Centres: The ApoEurope Study. *Disease.* 2000;38(8):721–30. <https://doi.org/10.1515/CCLM.2000.102>.
  18. Gupta VB, Laws SM, Villemagne VL, Ames D, Bush AI, Ellis KA, Lui JK, Masters C, Rowe CC, Szeoke C, Taddei K, Martins RN. Plasma Apolipoprotein E and Alzheimer Disease Risk: The AIBL Study of Aging. *Neurology.* 2011;76(12):1091–8. <https://doi.org/10.1212/WNL.0b013e318211c352>.
  19. Slooter AJC, de Knijff P, Hofman A, Cruts M, Breteler MMB, Van Broeckhoven C, Havekes LM, van Duijn CM. Serum Apolipoprotein E Level Is Not Increased in Alzheimer's Disease: The Rotterdam Study. *Neurosci Lett.* 1998;248(1):21–4. [https://doi.org/10.1016/S0304-3940\(98\)00339-5](https://doi.org/10.1016/S0304-3940(98)00339-5).
  20. Soares HD, Potter WZ, Pickering E, Kuhn M, Immermann FW, Shera DM, Ferm M, Dean RA, Simon AJ, Swenson F, Siuciak JA, Kaplow J, Thambisetty M, Zagouras P, Koroshetz WJ, Wan HJ, Trojanowski JQ, Shaw LM. Biomarkers Consortium Alzheimer's Disease Plasma Proteomics Project, for the Plasma Biomarkers Associated With the Apolipoprotein E Genotype and Alzheimer Disease. *Arch Neurol.* 2012;69(10):1310–7. <https://doi.org/10.1001/archneurol.2012.1070>.
  21. Scacchi R, Gambina G, Ruggeri M, Martini MC, Ferrari G, Silvestri M, Schiavon R, Corbo RM. Plasma Levels of Apolipoprotein E and genetic markers in elderly patients with Alzheimer's Disease. *Neurosci Lett.* 1999;259(1):33–6. [https://doi.org/10.1016/S0304-3940\(98\)00889-1](https://doi.org/10.1016/S0304-3940(98)00889-1).
  22. McConnell LM, Koenig BA, Greely HT, Raffin TA. Genetic Testing and Alzheimer disease: has the time come? *Nat Med.* 1998;4(7):757–9. <https://doi.org/10.1038/nm0798-757>.
  23. Zick CD, Mathews CJ, Roberts JS, Cook-Deegan R, Pokorski RJ, Green RC. Genetic testing for Alzheimer's Disease and its impact on insurance purchasing behavior. *Health Aff.* 2005;24(2):483–90. <https://doi.org/10.1377/hlthaff.24.2.483>.
  24. Simon R, Girod M, Fonbonne C, Salvador A, Clément Y, Lantéri P, Amouyel P, Lambert JC, Lemoine J. Total ApoE and ApoE4 Isoform Assays in an Alzheimer's Disease case-control study by targeted mass spectrometry (n = 669): a pilot assay for methionine-containing proteotypic peptides \*. *Mol Cell Proteomics.* 2012;11(11):1389–403. <https://doi.org/10.1074/mcp.M112.018861>.
  25. Kirmess KM, Meyer MR, Holubasch MS, Knapik SS, Hu Y, Jackson EN, Harpstrite SE, Verghese PB, West T, Fogelman I, Braunstein JB, Yarasheski KE, Contois JH. The PrecivityAD™ Test: accurate and reliable LC-MS/MS assays for quantifying plasma amyloid beta 40 and 42 and Apolipoprotein E proteotype for the assessment of brain amyloidosis. *Clin Chim Acta.* 2021;519:267–75. <https://doi.org/10.1016/j.cca.2021.05.011>.
  26. Begcevic Brkovic I, Zöhner B, Scholz M, Reinicke M, Ditttrich J, Kamalsada S, Baber R, Beutner F, Teren A, Engel C, Wirkner K, Thiele H, Löffler M, Riedel-Heller SG, Ceglarek U. Simultaneous mass spectrometry-based apolipoprotein profiling and apolipoprotein E phenotyping in patients with ASCVD and mild cognitive impairment. *Nutrients.* 2022;14(12):2474. <https://doi.org/10.3390/nu14122474>.
  27. Hu Y, Meuret C, Go S, Yassine HN, Nedelkov D. Simple and Fast Assay for Apolipoprotein E Phenotyping and Glycotyping: Discovering Isoform-Specific Glycosylation in Plasma and Cerebrospinal Fluid. *J Alzheimers Dis.* 2020;76(3):883–93. <https://doi.org/10.3233/JAD-200203>.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

### Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

