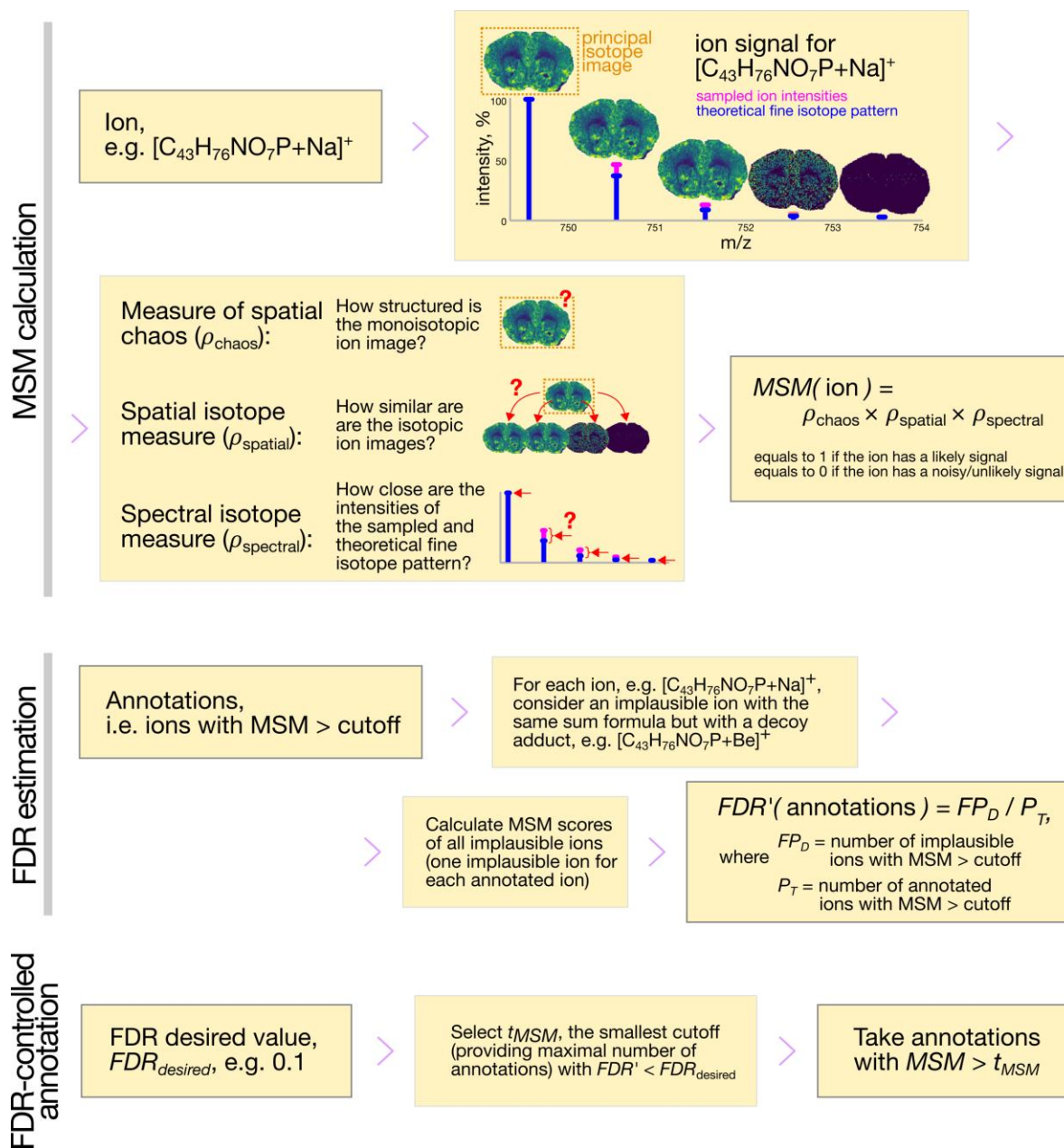


### **Supplementary Figure S1**

#### ***Illustration of an ion signal in the dataset a1s1 for $[M+Na]^+$***

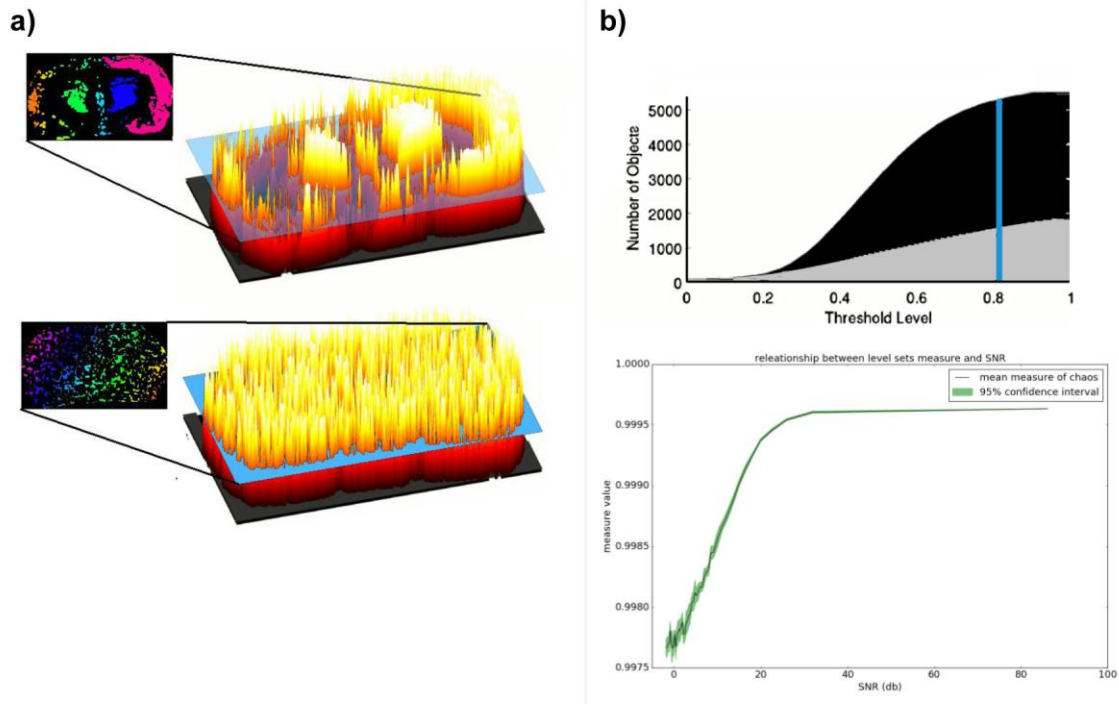
*Illustration of an ion signal in the dataset a1s1 corresponding to the sum formula  $C_{43}H_{76}NO_7P$  corresponding to PE(P-38:5) in HMDB, and the +Na ion adduct. The theoretical isotope pattern (blue) was predicted at the resolving power of  $R=100000$  at  $m/z$  400 and ion images of the predicted peaks were generated with the tolerance of  $\pm 2.5$  ppm. The ion was annotated at the desired FDR level of 0.1 and was validated by LC-MS/MS as PE(P-18:1\_20:4); see Section S11.*



Supplementary Figure S2

**Detailed scheme of the proposed framework**

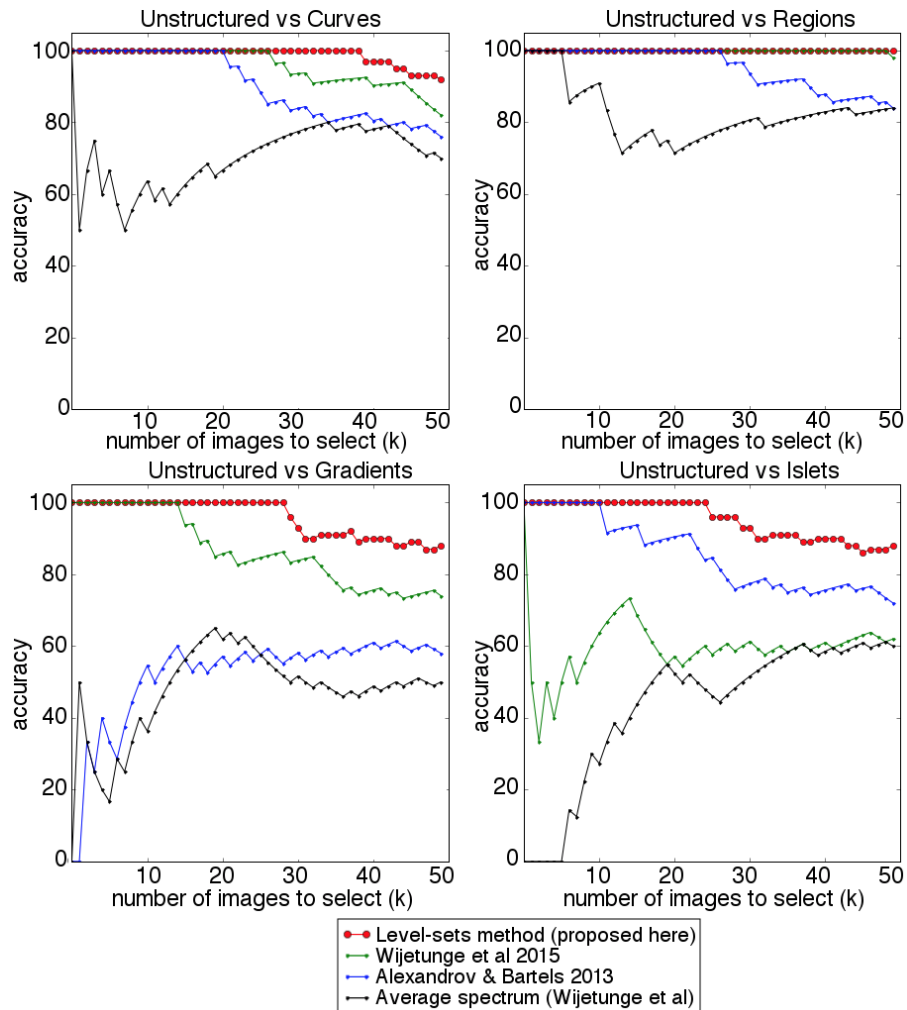
The framework consists of three parts: calculation of the Metabolite-Signal-Match score, FDR estimation, and FDR-controlled annotation. This scheme provides more comments as compared to Figure 1. The ion  $[C_{43}H_{76}NO_7P+Na]^+$  was annotated at the desired FDR level of 0.1 and was validated by LC-MS/MS as PE(P-18:1\_20:4); see Section S11.



### Supplementary Figure S3

**Detail on the novel measure of spatial chaos used to quantify the informativeness of a principal ion image in the ion signal.**

**a)** The principle of the novel measure of spatial chaos is counting the number of objects in an ion image at a particular intensity level and measuring the statistical distribution of objects amongst a set of levels. Ion image intensities are shown as a 3D landscape with the heat colormap (from black to red to yellow from intensities from low to high). The level-set cutting plane is shown in half-transparent blue. Objects from one level are shown on the left with connected pixels from the same object shown in the same color; pixels below the level-set-intensity shown in black. Top: an exemplary structured ion image; Bottom: an exemplary unstructured noisy ion image. **b)** Illustrations demonstrating behavior of the measure of spatial chaos. Top: Dependence of number of objects in an ion image at a particular level-set on the threshold level (intensity value used for the level set) for the structured (grey) and unstructured (black) images shown in the panel a). For a threshold level, the measure of spatial chaos is equal to the area under the curve for the range from 0 to the given level. Bottom: Dependence of the statistic on the signal-to-noise ratio (SNR) for a structured ion image with added noise.

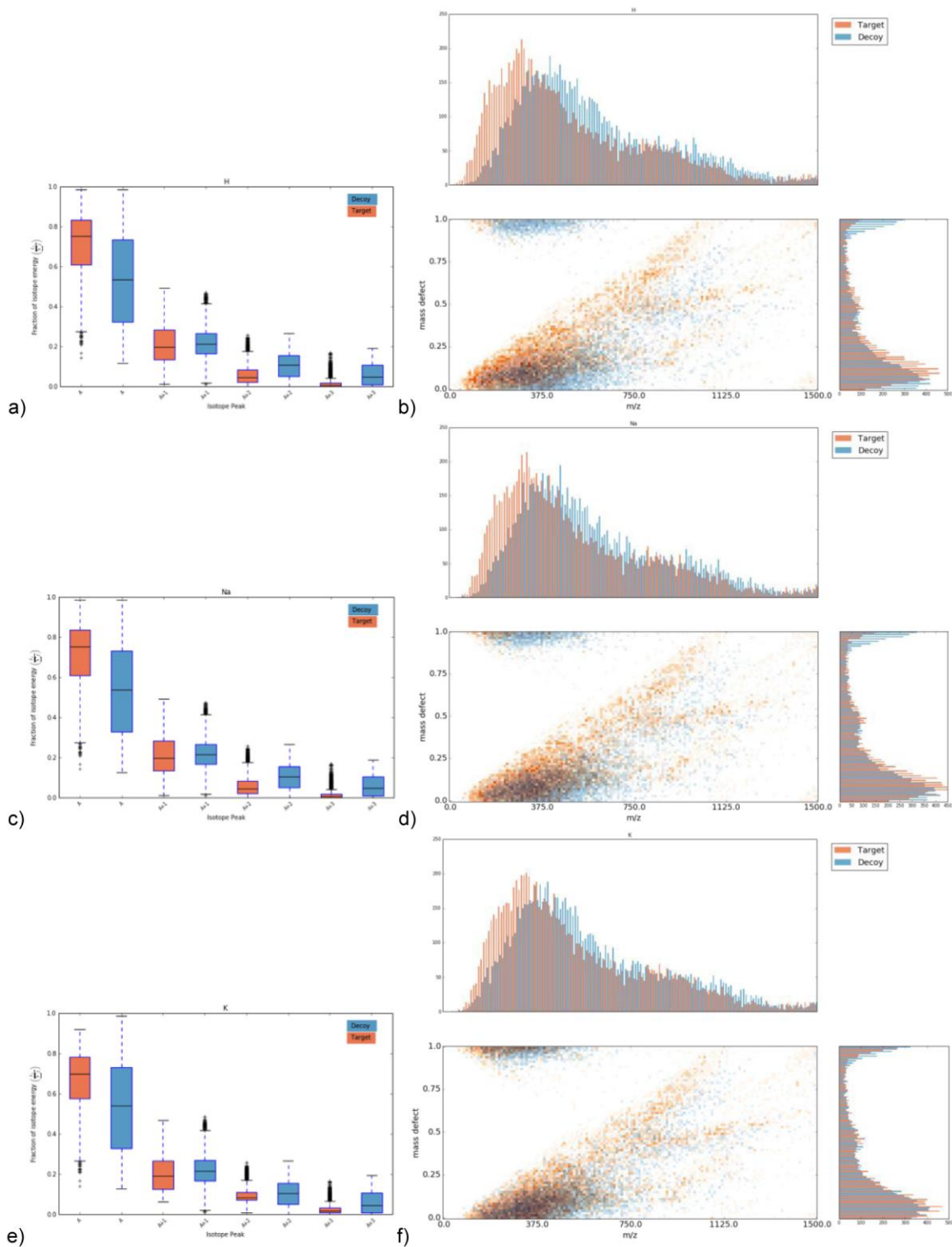


**Supplementary Figure S4**

**Evaluation of the proposed level-sets-based measure of spatial chaos**

Evaluation of the proposed level-sets-based measure of spatial chaos as compared to other existing measures of spatial chaos: from (Alexandrov & Bartels, 2013)<sup>1</sup> and (Wijetunge et al., 2015)<sup>2</sup>. The evaluation is performed according to the statistical framework originally proposed by us in (Alexandrov & Bartels, 2013)<sup>1</sup>.

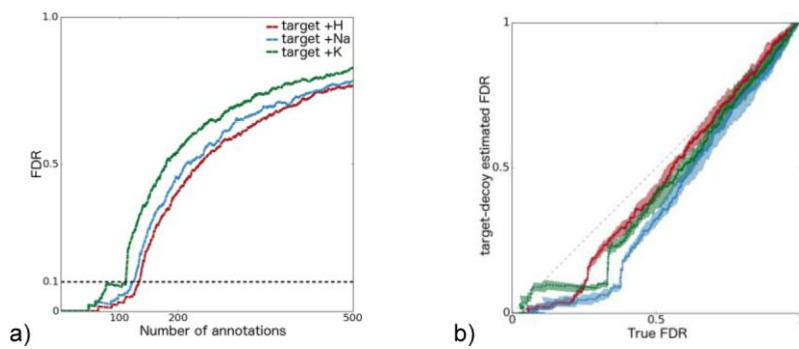




**Supplementary Figure S5**

***Evaluation of the similarity between the target and decoy sets of ions for metabolites from HMDB.***

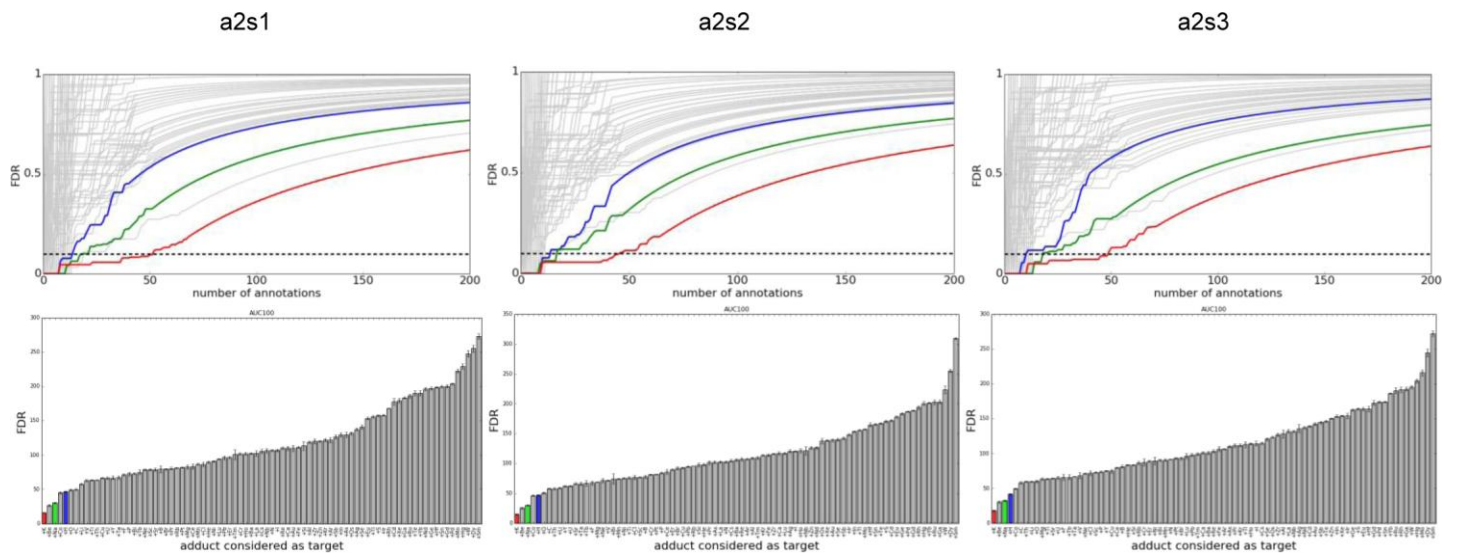
*Left (a, c, e): Box plots of the four highest isotopic peaks. Right (b, d, f): Kendrick plot with projections along either m/z or mass defect. The decoy set is selected as a random sampling from the set of decoy ions.*



**Supplementary Figure S6**

**Evaluation of the false-discovery-rate estimation on simulated data.**

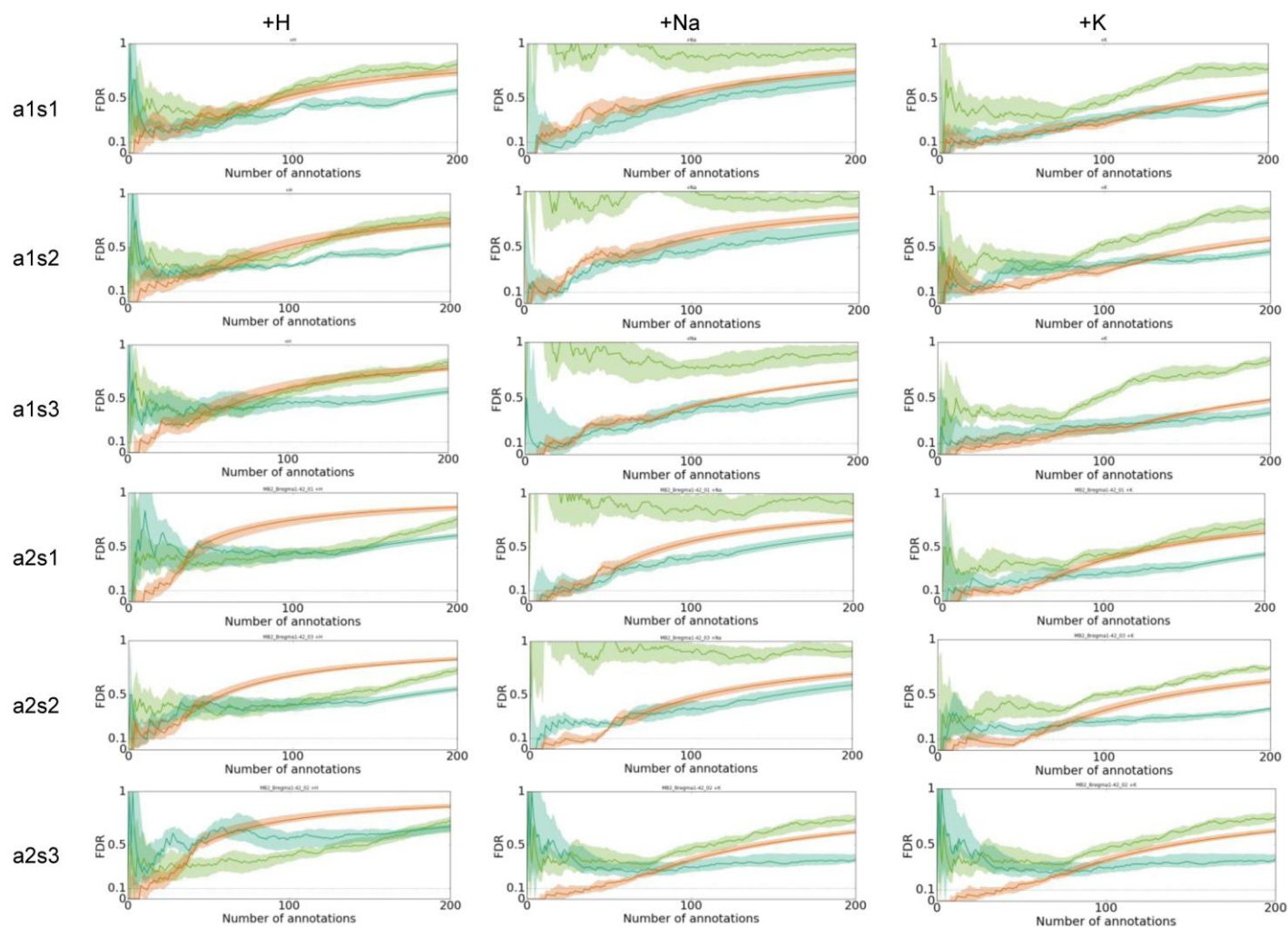
a) FDR curves for the target adducts +H, +Na, and +K. b) Relation between the true FDR and the FDR for the target adducts +H (red), +Na (blue), and +K (green). Shaded area shows confidence limits (5-95) from 20 independent replicates.



### Supplementary Figure S7

#### Negative control experiments using implausible adducts as the 'target'.

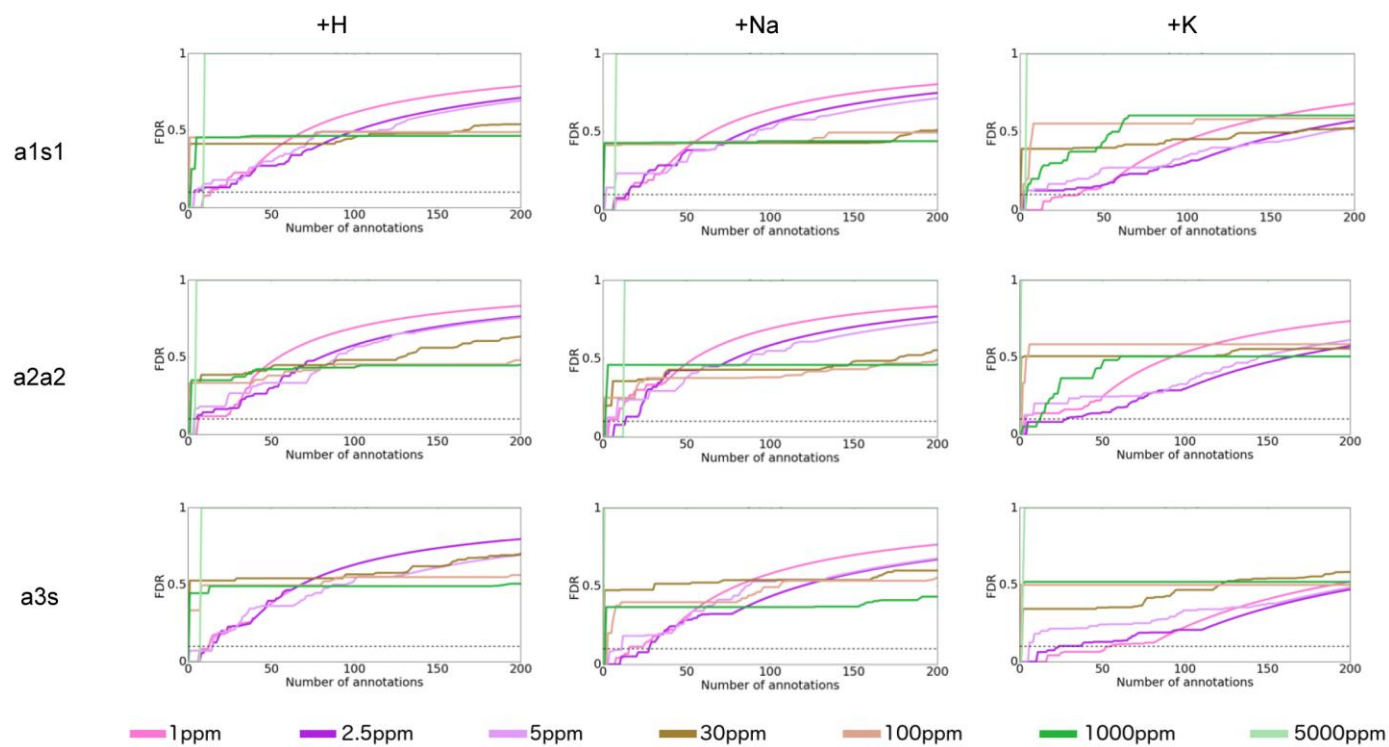
Top row: FDR curves per an implausible adduct Bottom row:  $AUC_{100}$  values. As baselines, the FDR curves and  $AUC_{100}$  values for the plausible adducts are shown for +H (blue), +Na (green), and +K (red). The curves are shown for 200 first annotations. Error bars in the  $AUC_{100}$  diagrams show standard deviation across 10 sampled decoy sets.



**Supplementary Figure S8**

**FDR-curves produced when using different scoring measures**

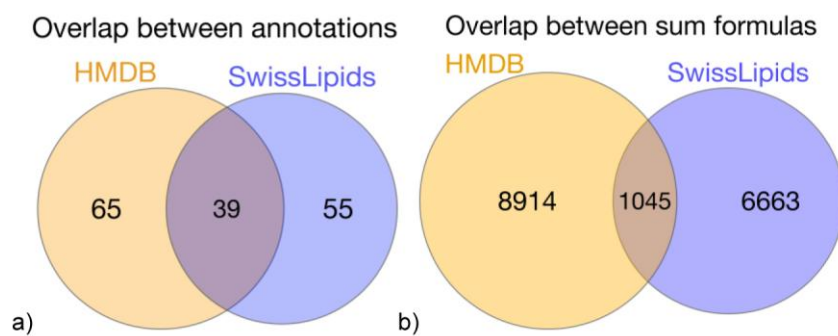
Three measures were tested: exact mass matching (light green), measure of chaos (blue green) and MSM (orange). The FDR-curves show that the MSM measure outperforms other measures (has lowest FDR-curve) across all datasets and adducts.



### Supplementary Figure S9

FDR curves showing the effect of changing the m/z tolerance for generating ion images

FDR-curves for different m/z-tolerance values used to sample ion images (2.5 ppm was used in the paper), for different datasets and ion adducts. The change of m/z-tolerance was used to evaluate effect of mass accuracy and resolving power.



**Supplementary Figure S10**

**Comparing annotations using the HMDB and SwissLipids molecular databases.**

a) Overlap in the molecular annotations for all 10 datasets from the paper with the desired FDR of 0.1. b) Overlap between sum formulas from HMDB and SwissLipids



# Supplementary Tables

Property	Dataset				
	a1s1	a1s1	a1s3	a1s4	a1s5
Specimen type	Mouse brain	Mouse brain	Mouse brain	Mouse brain	Mouse brain
Section direction	Coronal	Coronal	Coronal	Coronal	Coronal
Bregma distance (mm)	1.42	1.42	1.42	-1.46	-3.88
Pixel size ( $\mu\text{m}$ )	50	50	50	50	50
Number of pixels	12816	12614	11007	17809	16804
Number of peaks per spectrum	375 +/- 17	375 +/- 17	401 +/- 14	424 +/- 19	408 +/- 17
Raw data size* (GB)	23	23	21	32	30
m/z acquisition range	100-1200	100-1200	100-1200	100-1200	100-1200
Mass resolving power (at m/z 400)	130000	130000	130000	130000	130000

**Supplementary Table 1.** Properties of the high-resolution imaging MS datasets collected from mouse brain of the animal 1 (a1). \*Raw profile data in the .sl format, SCiLS Lab software, version 2016a (SCiLS, Bremen, Germany).

Property	Dataset				
	a2s1	a2s1	a2s3	a2s4	a2s5
Specimen type	Mouse brain	Mouse brain	Mouse brain	Mouse brain	Mouse brain
Section direction	Coronal	Coronal	Coronal	Coronal	Coronal
Bregma distance (mm)	1.42	1.42	1.42	-1.46	-3.88
Pixel size ( $\mu\text{m}$ )	50	50	50	50	50
Number of pixels	12078	19249	17829	11866	12088
Number of peaks per spectrum	495 +/- 13	474 +/- 54	388 +/- 137	491 +/- 22	495 +/- 11
Raw data size* (GB)	22	22	23	35	34
m/z acquisition range	100-1200	100-1200	100-1200	100-1200	100-1200
Mass resolving power (at m/z 400)	130000	130000	130000	130000	130000

**Supplementary Table 2.** Properties of the high-resolution imaging MS datasets collected from mouse brain of the animal 2 (a2). \*Raw profile data in the .sl format, SCiLS Lab software, version 2016a (SCiLS).

# Supplementary Note 1

## 1. Individual measures used in the MSM score

### 1.1 Evaluation of the proposed measure of spatial chaos

Discriminating images containing spatial structure (informative) from noise images (non-informative) based on an internal statistical measure was proposed by us in (Alexandrov & Bartels, 2013)<sup>1</sup>. However, the original measure<sup>1</sup> poorly performed on images of the type “gradient” (see the original paper for details). Since 2013, another algorithm was proposed<sup>2</sup> that showed better accuracy than the original algorithm on all but “islets” images.

To evaluate the novel measure of spatial chaos (Online Methods) and to compare it to the original measure and a measure proposed in (Wijetunge et al., 2015)<sup>2</sup>, we followed the framework proposed in (Alexandrov & Bartels, 2013)<sup>1</sup>. For the evaluation, the following ground-truth sets of manually selected ion images were used: one unstructured and four structured (‘regions’, ‘curves’, ‘gradients’, and ‘islets’). We quantified the accuracy of recognizing structured images vs. unstructured images with the help of the measure of spatial chaos as follows. The unstructured set was mixed with a structured set (one of four) and the values of a measure of spatial chaos were calculated for all images. The images were sorted by their values. For  $k$  from 1 to the size of the mixture set (50), the first images with the highest values are selected and the accuracy of detecting structured images (number of structured divided by  $k$ ) is calculated. Supplementary Figure 3 shows the curves of the accuracy values plotted against  $k$ . One can see that the proposed measure of spatial chaos outperforms earlier measures<sup>1,2</sup> and certainly the average spectrum approach from (Wijetunge et al., 2015)<sup>2</sup>, achieving the classification accuracy for  $k=50$  (all structured vs. all unstructured) of 92.0%, 100.0%, 88.0% and 88.0% for the ‘regions’, ‘curves’, ‘gradients’ and ‘islets’ types of structure respectively.

Moreover, the measure is computationally efficient, taking 0.031s per an ion image from the ground truth sets (121x202 pixels), twice as fast as the published implementation from (Alexandrov and Bartels 2013) as compared on the same computer (MacBook Pro, 3 GHz Intel Core i7, 16 GB 1600 MHz DDR3 RAM).

Thus, we conclude that our new measure outperforms earlier measures of spatial chaos both in accuracy and speed.

### 1.2 Calculating the spectral isotope measure

For a metabolite ion, the spectral isotope measure quantifies the spectral similarity between the theoretical isotope pattern and the sampled isotope pattern from an imaging MS dataset (Online Methods).

For defining the measure, we compared two ways of calculating the sampled isotope pattern motivated by the following consideration. The most common approach for isotope pattern matching in high-resolution mass spectrometry is using high-quality individual spectra<sup>3</sup>.

With a sufficient mass resolution and quality of mass spectra, one can fit a chemical formula directly to an isotope pattern<sup>4</sup>. However, the quality of peaks within individual spectra in imaging MS is hampered because of no separation used to produce a spectrum, normally complex samples like tissues analyzed, and lower dynamic range as compared to other state-of-the-art MS techniques. This effect is especially prominent for low-abundant molecules. This makes fitting an isotope pattern for each molecule to every spectrum unappealing. Another approach is to calculate a dataset-average sampled isotope pattern but this approach may lack sensitivity due to averaging.

We compared both approaches: (1) single-spectrum calculation of the match between the theoretical isotope pattern and a spectrum from a single pixel, (2) dataset-average calculation of the match between the theoretical pattern and the average spectrum from all pixels where the principal peak is present. The comparison was done on the a2s2 dataset for several selected sum formulas (Table SN1.1). For each formula, we found a gold-standard spectrum in the dataset using either of two ways: (1.1) a high-intensity spectrum from the pixel where the principal peak was maximal and (1.2) a low-intensity spectrum from a pixel with the principal peak intensity equal to the 10%-quantile of all pixels where the principal peak is present. The low-intensity spectrum was considered to obtain baseline (relatively low) values.

Table SN1.1 shows that calculating the isotope pattern match on dataset-average intensities performs comparably well to the calculation on the highest-intensity signal spectrum and better than the baseline-case of taking a low-intensity single spectrum. This motivated us to employ the dataset-average strategy rather than doing single-spectrum pixel-by-pixel calculations and to formulate the spectral isotope measure according to Online Methods, Equation OM2.

Metabolite ion	Principal m/z	$\rho_{spectral}$		
		Single-spectrum approach		Dataset-average approach
		Low-intensity spectrum	Max-intensity spectrum	
$[C_{47}H_{93}N_2O_6P + K]^+$	851.64026	0.88	0.97	0.95
$[C_{48}H_{91}NO_8 + Na]^+$	832.66369	0.85	0.94	0.94
$[C_{40}H_{77}NO_3 + H]^+$	620.59762	0.88	0.99	0.96
$[C_{42}H_{82}NO_8P + H]^+$	760.58508	0.97	0.97	0.99
$[C_{40}H_{80}NO_8P + Na]^+$	756.55138	0.64	0.99	0.97
$[C_{41}H_{82}NO_8P + Na]^+$	770.56703	0.94	0.92	0.92
$[C_{40}H_{80}NO_8P + H]^+$	734.56943	0.97	0.99	0.99

**Table SN1.1.** Comparing two approaches of calculating the spectral isotope match ( $\rho_{spectral}$  as defined in Online Methods, Equation OM2) on seven molecular annotations from HMDB: the single-spectrum approach (two versions, selecting a high-intensity spectrum and a low-intensity spectrum, the latter used for getting baseline values) and the dataset-average approach.

## 2. Evaluation of the proposed FDR approach

An essential part of the proposed molecular annotation approach is the ability to both quantify and control the FDR of the resulting annotations. We propose to calculate FDR by using a target-decoy approach (Online Methods). Here we provide results of evaluating the proposed FDR approach.

Since currently there exist no accepted approaches for calculating FDR in metabolomics, we developed our own evaluation framework with the following steps: 1) evaluation of the proposed decoy set (Supplementary Section 3.1), 2) comparison of the estimated FDR with true FDR on simulated data (Supplementary Section 3.2), and 3) negative control experiment when using an implausible adduct as the target (Supplementary Section 3.3).

### 2.1 Evaluation of the decoy database

In a target-decoy approach, a decoy must satisfy the following criteria <sup>5</sup>: 1) follow the same statistical distributions as the target, 2) produce a similar response to noise in the real data, and 3) be compatible with the search pipeline presented in this work.

We evaluated compared the statistical distribution of elemental frequency, mass distribution and chemical space of the decoy and target. Supplementary Figure 6 shows the following properties of the target and decoy sets generated from HMDB for each of the target ions +H, +Na, +K: isotopic patterns,  $m/z$ 's and mass defects with Kendrick plots.

The isotope ratios (Supplementary Figures 6a, 6c, 6e) are similar between the target and decoy with the decoy having higher intensities for heavier isotopic peaks (as expected from adding elements with higher atomic number and more complex isotope patterns as decoy adducts). This effect is moderate and can be explained by the decoy adducts having higher atomic numbers and more complex isotope patterns than a target adduct. An increased variance in the isotope abundances can be observed for all but the second isotopic peak.

Examining the comparative Kendrick plots with histograms showing the distributions projected along both the  $m/z$  and mass defect axes (Supplementary Figures 6b, 6d, 6f), there is a large overlap between the clouds of points, particularly along the mass defect projection. The systematic shift of the distribution (as expected when generally heavier elements as “adducts”) may limit the FDR to ions above  $m/z$  100. This is not critical, as the mass-accuracy of high-resolution instruments is typically sufficient to unambiguously determine the molecular formula of mass spectral peaks below  $m/z$  100 <sup>6</sup>.

### 2.2 Evaluation of the FDR estimate on simulated data

We calculated FDR curves for the MSM-based molecular annotation for each of the target adducts +H, +Na, and +K (Supplementary Figure 6a). The number of formulas per adduct was 121, 90, 88 for +H, +Na, +K respectively. At the desired FDR level of 0.1, there were 136, 119, and 112 annotations returned for +H, +Na, +K respectively of which 101, 78, and 75 were present in the ground truth (see Table SN1.2). This corresponds well to the expected true FDR values (calculated as  $FP_{target}/P_{target}$ ) for such sets of molecular annotations which are equal to 0.25, 0.36 and 0.33 for +H, +Na, and +K respectively. Next,

we calculated the relation between the true FDR and our FDR estimate for all values of FDR, visualized in Supplementary Figure 6b, with one-sigma confidence intervals from 20 random iterations used in Supplementary Note 3, Algorithm 3. The estimated FDR follows the true FDR well, with most difference in the low FDR range where an underestimation is visible. Importantly, one can see the low variance of the estimate with the largest width of the confidence interval to be 0.10, 0.15, 0.11 for +H, +Na and +K respectively.

	FDR target value											
	0.01			0.05			0.1			0.2		
Adduct	+H	+Na	+K	+H	+Na	+K	+H	+Na	+K	+H	+Na	+K
Recall (%)	72	55	44	80	83	77	83	87	85	84	91	86
True FDR	0.09	0.07	0.03	0.21	0.16	0.06	0.25	0.36	0.33	0.27	0.38	0.33

**Table SN1.2.** Statistics for the annotation of the simulated dataset at different FDR target values.

### 2.3 Negative control experiment

In the negative control experiment, we used in turn each implausible adduct as the target instead of either of the common adducts (+H, +Na, and +K). The implausible adduct considered as the target was taken out from the decoy adducts. Supplementary Figure 8 shows the obtained FDR curves (in grey) as well as the FDR curves for the common adducts +H (blue), +Na (green), and +K (red). As expected, most of the negative-control FDR curves have higher FDR values than the curves for +H, +Na, +K. To summarize the information, for each FDR curve we calculated the value  $AUC_{100}$  of the area under curve for the first 100 annotations. As expected,  $AUC_{100}$  values for most of the implausible adducts are higher than for plausible ones. Only +Be and +F have  $AUC_{100}$  lower than the common adducts.

## 3. Comparing alternative scoring measures

The typical approach for metabolite annotation in imaging mass spectrometry is assignment based on exact mass, which does not take any spatial information into account. We recently proposed an approach for detecting a signal that incorporates spatial information<sup>1</sup>. We compared these two approaches against the MSM score proposed in this paper using the FDR calculation. It can be seen in Supplementary Figure 4 that across all datasets that incorporating spatial information provides better FDR curves (i.e. the area under the curve is lower) and in almost all cases the MSM score is the best performing approach.

## 4. Simulating lower mass resolving power

To estimate the effect of a reduced resolving power, or lower mass accuracy, on the FDR-controlled annotation pipeline we repeatedly ran the pipeline using different  $m/z$  tolerances for generating ion images. The effect on the FDR curves can be seen in Supplementary Figure 5 where increasing the  $m/z$  tolerance can be seen to have a detrimental effect. This highlights the importance of using high-resolving power instrumentation for spatial metabolomics experiments.

## 5. Runtime

The time for FDR-molecular annotation of a dataset in this paper was under one hour (MacBook Pro, 3 GHz Intel Core i7, 16 GB 1600 MHz DDR3 RAM).

The three most costly computational operations within the framework are: 1) isotope pattern prediction, 2) sampling ion image signals from a dataset, 3) MSM-scoring of the signals. Our efficient implementation of the isotope pattern prediction according to (Ipsen, 2014)<sup>7</sup> takes approximately 300 millisecond per sum formula (MacBook Pro, 3 GHz Intel Core i7, 16 GB 1600 MHz DDR3 RAM). The total time to generate the isotope patterns for sum formulas from the HMDB (7708 unique carbon-containing sum formulas) for the 3 main adducts and 85 decoy adducts is approximately 5 minutes. For FDR-controlled molecular annotation of a dataset, in the order of 1 million of ion images need to be sampled from the dataset when the following parameters are used: 4 ion images per molecular formula, 20 decoy replicates for each target ion. For datasets in this paper, sampling an ion image takes approximately 200 microsecond, sampling all ion images takes 3 minutes.

## 6. Annotations using the SwissLipids database

FDR controlled annotation at the desired FDR level of 0.1 was performed for molecules from the SwissLipids database<sup>8</sup>, downloaded on June 1st, 2016. Overall, 94 sum formulas were annotated compared to 104 for HMDB, with 39 sum formulas contained both in HMDB and SwissLipids. Despite the overlap between the databases being around 10% the overlap between annotations is around 40%, that is well in line with our results that most of the annotated metabolites in the considered datasets are lipids.

Please note that calculation of the MSM scores is not affected by the choice of the database. The differences for specific sum formulas which are present in both databases (HMDB and SwissLipids) can be only at the stage of FDR calculation. The differences can be due to different coverage of  $m/z$ -values space by molecules in the database but they should not be significant for relatively large databases created for similar biological systems, such as HMDB and SwissLipids.

## References

1. Alexandrov, T. & Bartels, A. Testing for presence of known and unknown molecules in imaging mass spectrometry. *Bioinformatics* **29**, 2335–2342 (2013).



2. Wijetunge, C. D. *et al.* EXIMS: an improved data analysis pipeline based on a new peak picking method for EXploring Imaging Mass Spectrometry data. *Bioinformatics* **31**, 3198–3206 (2015).
3. Pluskal, T., Uehara, T. & Yanagida, M. Highly accurate chemical formula prediction tool utilizing high-resolution mass spectra, MS/MS fragmentation, heuristic rules, and isotope pattern matching. *Anal. Chem.* **84**, 4396–4403 (2012).
4. Qian\*, K. *et al.* Resolution and Identification of Elemental Compositions for More than 3000 Crude Acids in Heavy Petroleum by Negative-Ion Microelectrospray High-Field Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *Energy Fuels* **15**, 1505–1511 (2001).
5. Storey, J. D. & Tibshirani, R. Statistical significance for genomewide studies. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 9440–9445 (2003).
6. Kind, T. & Fiehn, O. Metabolomic database annotations via query of elemental compositions: mass accuracy is insufficient even at less than 1 ppm. *BMC Bioinformatics* **7**, 234 (2006).
7. Ipsen, A. Efficient calculation of exact fine structure isotope patterns via the multidimensional Fourier transform. *Anal. Chem.* **86**, 5316–5322 (2014).
8. Aimo, L. *et al.* The SwissLipids knowledgebase for lipid biology. *Bioinformatics* **31**, 2860–2866 (2015).



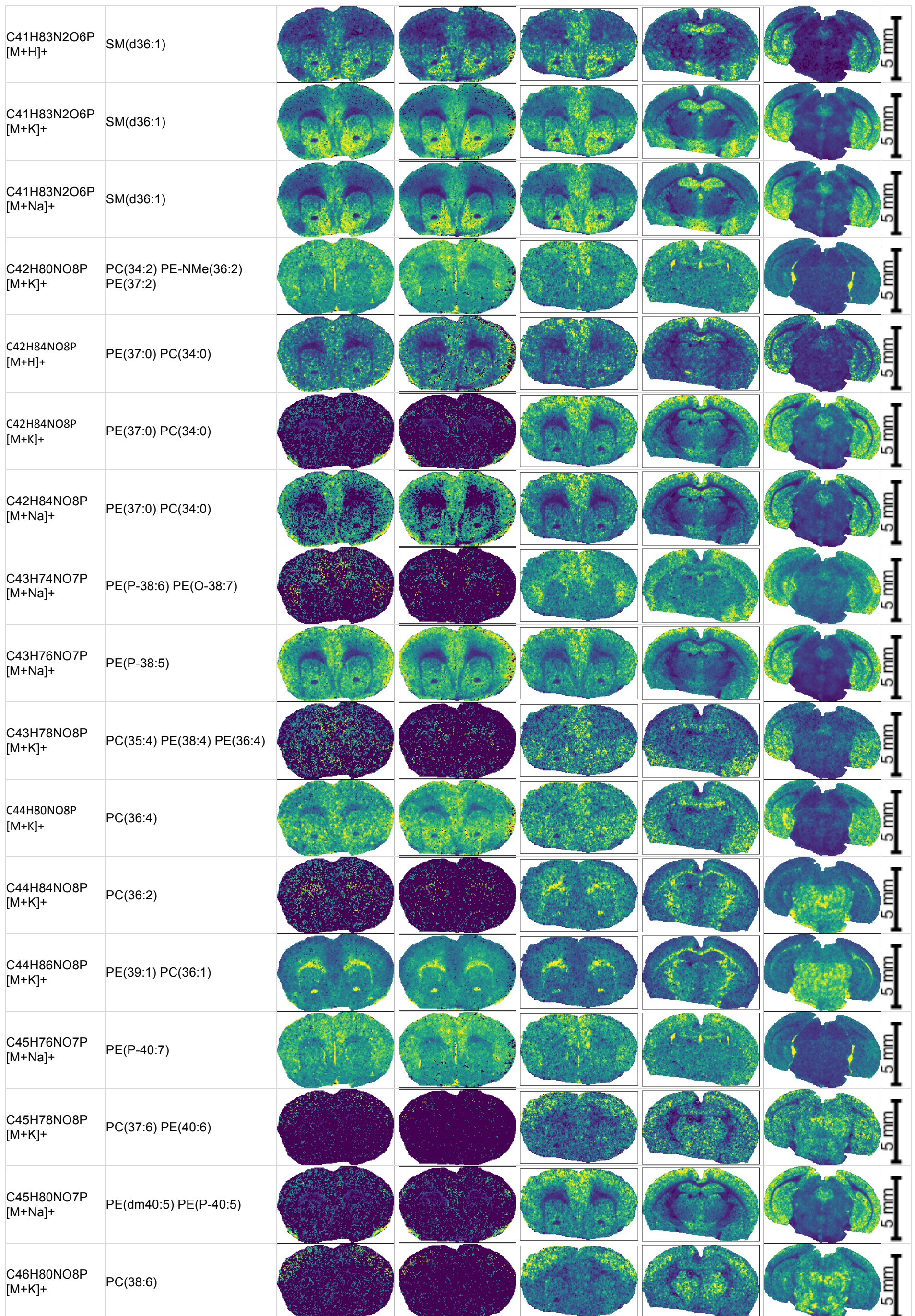
# Supplementary Note 2

## 1. Ion images for all annotations from animal 1

The table below shows ion images for all annotations from the datasets a1s1, a1s2, a1s3 from the three serial sections for the Bregma distance 1.42 mm from the animal a1. The parameters were: FDR target value 0.1, HMDB version 2.5, m/z tolerance 2.5 ppm. Colors are assigned to each image individually using the parula colormap (blue indicating low intensity, yellow high intensity).

Annotated sum formulas, ion adduct	Metabolites from HMDB corresponding to the annotated sum formula	a1s1	a1s2	a1s3	a1s4	a1s5
C28H33O14 [M+Na] <sup>+</sup>	Pelargonidin 3-sophoroside					
C35H66O4 [M+H] <sup>+</sup>	Artemoin D Artemoin C Artemoin B Artemoin A					
C37H68O4 [M+H] <sup>+</sup>	Cohibin C Cohibin D					
C37H71O8P [M+K] <sup>+</sup>	9-Octadecenoic acid 1-[[[(1-oxohexadecyl)oxy)methyl]-2-(phosphonoxy)ethyl ester PA(34:1)					
C37H71O8P [M+Na] <sup>+</sup>	9-Octadecenoic acid 1-[[[(1-oxohexadecyl)oxy)methyl]-2-(phosphonoxy)ethyl ester PA(34:1)					
C39H73O8P [M+K] <sup>+</sup>	PA(36:2) 9-Octadecenoic acid 1-[[[(phosphonoxy)methyl]-12-ethanediy] ester					
C39H73O8P [M+Na] <sup>+</sup>	PA(36:2) 9-Octadecenoic acid 1-[[[(phosphonoxy)methyl]-12-ethanediy] ester					
C39H79N2O6P [M+K] <sup>+</sup>	SM(d34:1)					
C40H78NO8P [M+K] <sup>+</sup>	PE-NMe(34:1) PE(35:1) PC(32:1)					
C40H80NO8P [M+H] <sup>+</sup>	PC(32:0) PE(35:0)					
C40H80NO8P [M+K] <sup>+</sup>	PC(32:0) PE(35:0)					
C40H80NO8P [M+Na] <sup>+</sup>	PC(32:0) PE(35:0)					
C41H82NO8P [M+H] <sup>+</sup>	PC(33:0) Phosphatidylethanolamine PE(36:0)					







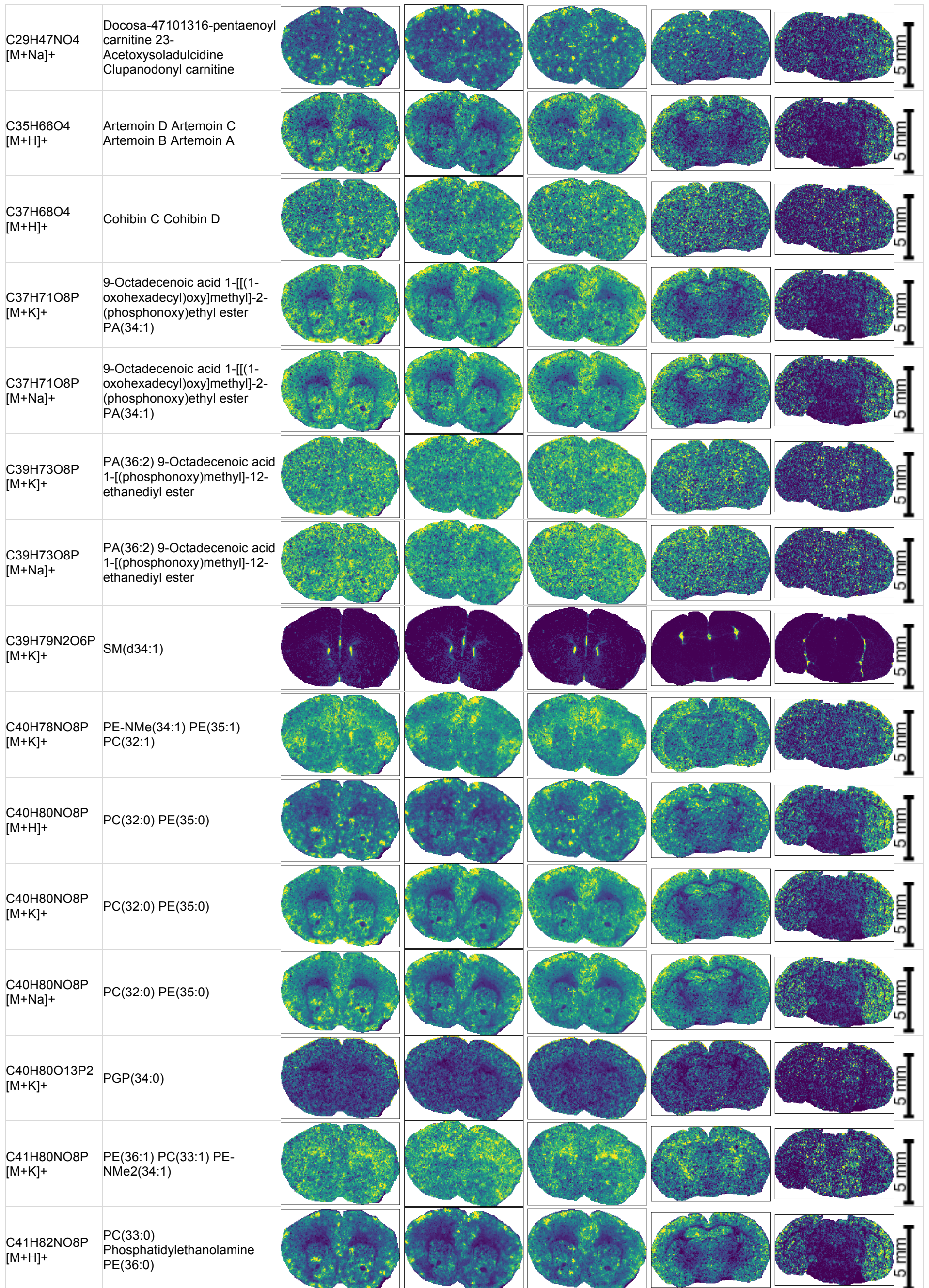
C46H82NO8P [M+K] <sup>+</sup>	PC(38:5)						5 mm
C46H84NO8P [M+K] <sup>+</sup>	PC(38:4)						5 mm
C46H84NO8P [M+Na] <sup>+</sup>	PC(38:4)						5 mm
C48H84NO8P [M+K] <sup>+</sup>	PC(40:6)						5 mm
C48H91NO8 [M+K] <sup>+</sup>	Galactosylceramide (d42:2) GlcCeramide (d42:2)						5 mm

## 2. Ion images for all annotations from animal 2

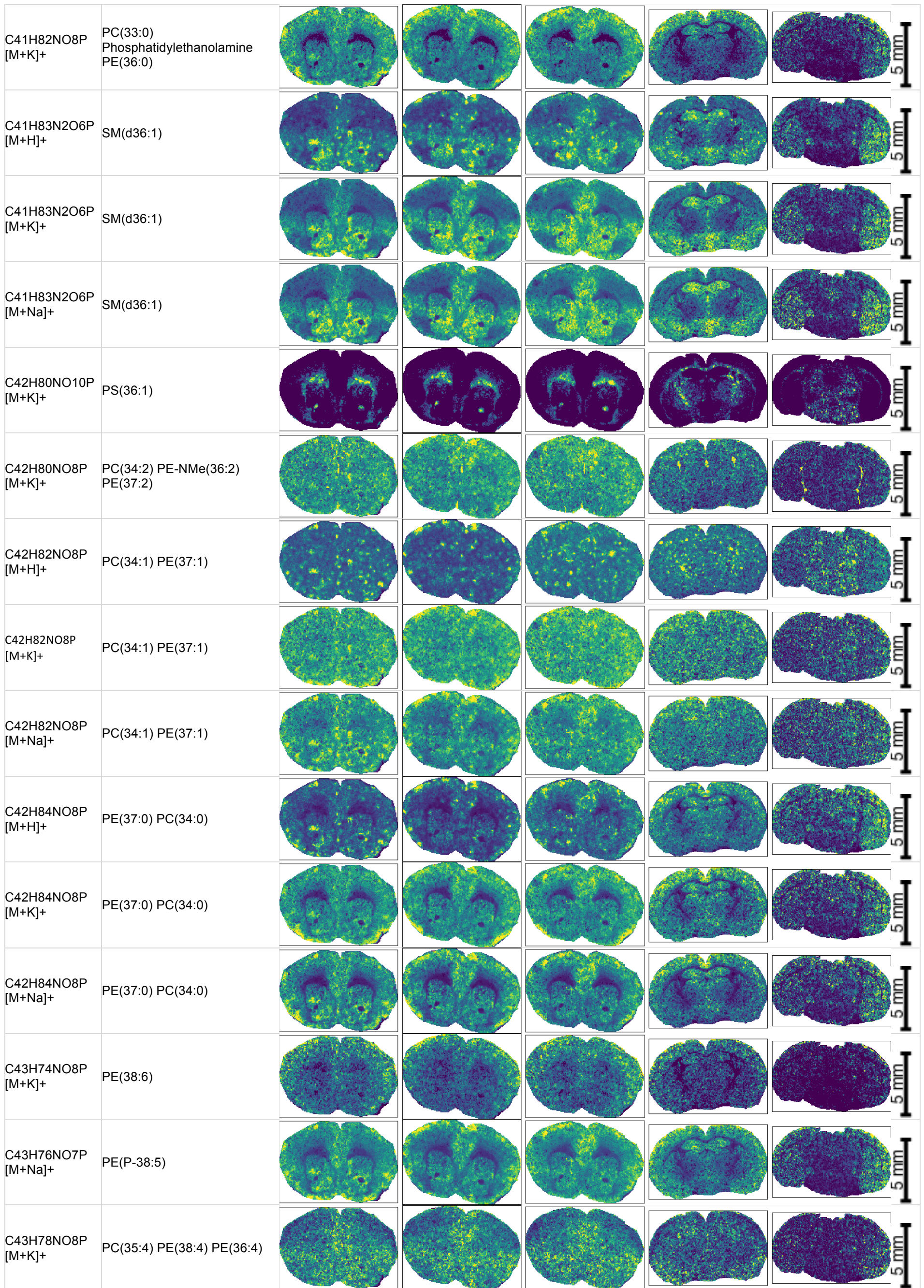
The table below shows ion images for all annotations from the datasets a2s1, a2s2, a2s3 from the three serial sections for the Bregma distance 1.42 mm from the animal a2. The parameters were: FDR target value 0.1, HMDB version 2.5, m/z tolerance 2.5 ppm. Colors are assigned to each image individually using the parula colormap (blue indicating low intensity, yellow high intensity).

Annotated sum formulas, ion adduct	Metabolites from HMDB corresponding to the annotated sum formula	a2s1	a2s2	a2s3	a2s4	a2s5
C10H7NO3 [M+K] <sup>+</sup>	1-Nitronaphthalene-56-oxide 1-Nitronaphthalene-78-oxide Kynurenic acid					
C23H46NO7P [M+K] <sup>+</sup>	LysoPE(18:1)					
C24H50NO7P [M+H] <sup>+</sup>	LysoPC(16:0)					
C26H52NO6P [M+H] <sup>+</sup>	LysoPC(P-18:1)					
C26H52NO7P [M+K] <sup>+</sup>	LysoPC(18:1)					
C27H44NO7P [M+K] <sup>+</sup>	LysoPE(22:6)					
C28H33O14 [M+Na] <sup>+</sup>	Pelargonidin 3-sophoroside					

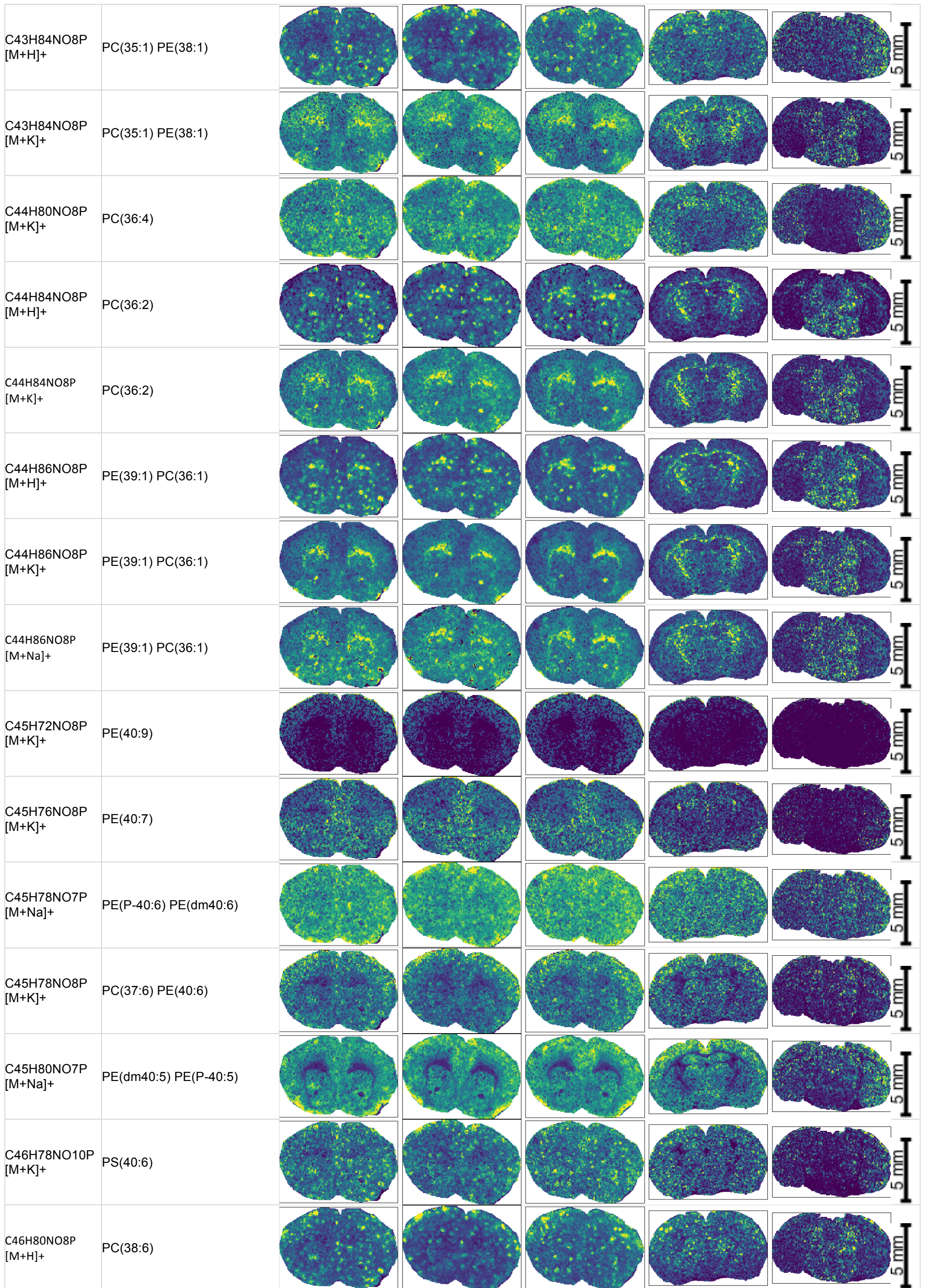




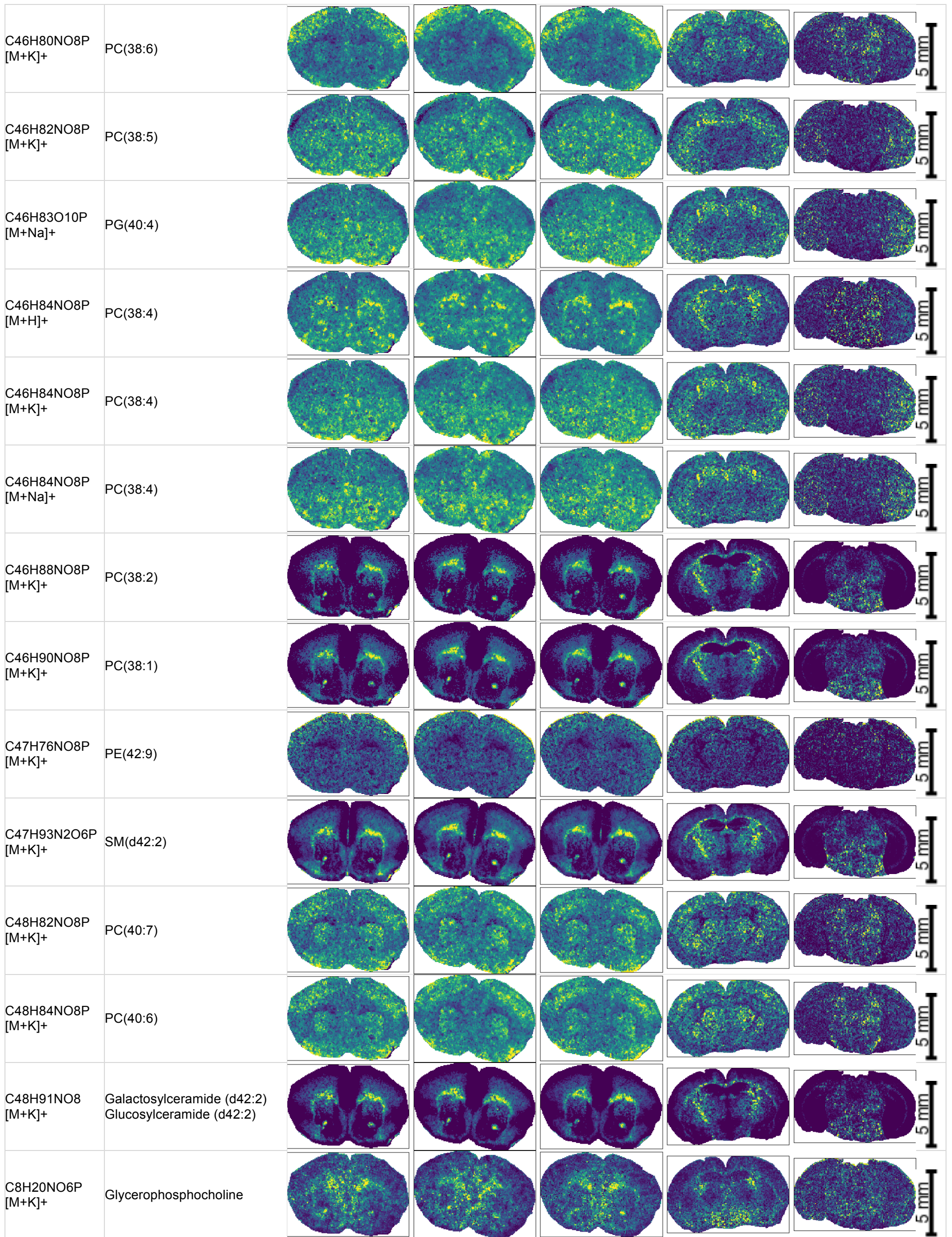














# Supplementary Note 3

**Input:** Real-valued image  $B$ , number of levels  $n_{levels}$

**Output:** measure of spatial chaos  $\rho_{chaos}$

**Algorithm:**

- ```
// scale image intensity range to [0 1]
1.  $B' = \frac{B - \min(B)}{\max(B) - \min(B)}$ 

// main part
2. For  $n$  in  $[0, 1, \dots, n_{levels}]$ :
    // threshold image at a current level
3.  $t_n = \frac{n}{n_{levels}}$ 
4.  $B_{bw} = B' > t_n$ 
    // fill single-pixel holes
5.  $\widehat{B}_{bw} = \psi_s(B_{bw})$ 
    // count separate objects with 4-connectivity
6.  $n_{objects_n} = \text{imlabel}(\widehat{B}_{bw})$ 
7.  $\rho_{chaos} = 1 - \frac{\sum_{n=1}^{n_{levels}} n_{objects_n}}{n_{pixels} * n_{levels}}$ 
8. return  $m_c$ 
```

**Algorithm 1.** The level-sets based algorithm for calculating the measure of spatial chaos of an ion image.  $\psi_s(A)$  is a hole-filling operation to ‘fill in’ isolated missing pixels that can happen in HR imaging MS (and to avoid overestimating the number of objects). It consists of a sequence of morphological operation:  $\psi_s(A) = (A \ominus S) \oplus S'$  with structuring elements  $S' = [[010], [111], [010]]$ ,  $S = [[111], [111], [111]]$ .  $\text{imlabel}$  uses the label function from scipy (<http://www.scipy.org/>) with 4-connectivity and returns the number of disconnected objects in an image.

**Input:** Metabolite sum formula, adduct, charge, resolving power of the spectra, imaging MS dataset, MSM threshold  $t_{MSM}$

**Output:** Decision whether the ion  $[sum\ formula + adduct]^{charge}$  is present in the dataset

**Algorithm:**

*// Predict isotopic patterns*

1. Predict the isotope envelope\* for  $[sum\ formula + adduct]^{charge}$  at the resolving power
2. Detect centroids of the isotope envelope\*, exact  $m/z$ 's and relative intensities  $(m, a)$   
*// Generate and score signals from the dataset*

3. For  $m_i$  in  $m$ :

4. Generate an ion image  $B_i$  for the  $i$ 'th isotopic peak at  $m/z\ m_i$

5. Calculate  $\rho_{chaos}$  from  $B_1$  and  $\rho_{spatial}, \rho_{spectral}$  from  $B$  and  $a$  according to Algorithm 1 as well as Online Methods, Equations OM1 and OM2, respectively

6. Calculate the *MSM* score according to Online Methods, Equation OM3

*// Annotate the data*

7. If  $MSM > t_{MSM}$  :

8. the ion  $[sum\ formula + adduct]^{charge}$  is annotated as being present in the dataset

**Algorithm 2.** *MSM-based molecular annotation determining whether a metabolite ion is present in an imaging MS dataset.*

**Input:** Metabolite database, resolving power of the mass spectrometer used, imaging MS dataset, ion charge, target adduct, decoy adducts, desired FDR level  $FDR_{desired}$ , number of decoy samplings

**Output:** A set of molecular annotations (ions from the metabolite database detected as present in the dataset)

**Algorithm:**

- ```
// Predict and score all metabolite signals
1. For sum formula in database:
2.    $ion_{target} = [sum\ formula + target\ adduct]^{charge}$ 
3.   Calculate  $MSM_{ion_{target}}$  according to Algorithm 2.(1-3)
4.    $ion_{decoy} = [sum\ formula + decoy\ adduct]^{charge}$ , where decoy adduct is randomly chosen from the list of decoy adducts
5.   Calculate  $MSM_{ion_{decoy}}$  according to Algorithm 2.(1-3)

// Calculate the MSM cutoff corresponding to the desired FDR level
6. Form a combined vector of MSM values  $s = MSM_{ion_{target}} \cup MSM_{ion_{decoy}}$ 
// Find the maximal number of annotations providing FDR below  $FDR_{desired}$ 
7. Sort  $s$  in descending order.
8.  $i = 2 * n_{molecules}$ 
9. While  $FDR > FDR_{desired}$ :
10.   $n_{target} = \sum_{j=0}^{n_{molecules}} p_j$  where  $p_j = 1$  if  $MSM_{ion_{targetj}} < s_i$ , otherwise  $p_j = 0$ 
11.   $n_{decoy} = \sum_{j=0}^{n_{molecules}} p_j$  where  $p_j = 1$  if  $MSM_{ion_{decoy}j} < s_i$ , otherwise  $p_j = 0$ 
12.  Calculate FDR according to Online Methods, Equation OM4
13.   $i := i - 1$ 
14.  $t_{MSM} = s_i$ 

15. Repeat steps 1-11 according to the number of decoy samplings,  $n$ 
     $t = [t_{MSM_1} t_{MSM_2} \dots t_{MSM_n}]$ 
16.  $t_{MSM} = median(t)$ 

// Perform the MSM-based molecular annotation with the calculated cutoff
17. For sum formula in database:
    a. If  $MSM_{ion_{target}} > t_{MSM}$  then add  $ion_{target}$  into the list of molecular annotations
```

**Algorithm 3.** FDR-controlled molecular annotation that screens for metabolite ions present in an imaging MS dataset, with the desired FDR level.

# Supplementary Data 1

1. LC-MS/MS validation of annotations using authentic standards



# Metabolite annotations with authentic standards available

Annotation	Authentic standard	Sum formula	Validated as	Validation summary
Glycerophosphocholine	L- $\alpha$ -Glycerophosphorylcholine, Sigma G5291 [HMDB00086]	C8H20NO6P	Glycerophosphorylcholine	<b>HD:</b> ChoP, phosphoglycerol, Choline
LysoPC(16:0)	LysoPC(16:0/0:0), Sigma L5254 [HMDB61702]	C24H50NO7P	LysoPC(16:0/0:0)	<b>HD:</b> ChoP, phosphoglycerol, Choline <b>SD:</b> 16:0 FA
PC(32:0)	PC(16:0/16:0), Avanti 850355 [HMDB00564]	C40H80NO8P	PC(16:0/16:0)	<b>HD:</b> ChoP <b>SD:</b> 16:0 FA
PC(34:1)	PC(16:0/18:1), Avanti 850457P [HMDB07972]	C42H82NO8P	PC(16:0/18:1)	<b>HD:</b> ChoP <b>SD:</b> 16:0 FA, 18:1 FA
PC(36:2)	PC(18:1(9Z)/18:1(9Z)), Avanti 850375C [HMDB00593]	C44H84NO8P	PC(18:1/18:1)	<b>HD:</b> ChoP <b>SD:</b> 18:1 FA
PC(36:4)	PC(18:2(9Z,12Z)/18:2(9Z,12Z)), Sigma P0537 [HMDB08138]	C44H80NO8P	PC(18:2/18:2)	<b>HD:</b> ChoP <b>SD:</b> 18:2 FA
PA(36:2)	L-Phosphatidic acid, PA(18:1/18:1), Sigma P2767 [HMDB07865]	C39H73O8P	PA (18:1/18:1)	<b>HD:</b> Glycerol phosphate-H2O <b>SD:</b> 18:1 FA
PE(36:0)	Phosphatidylethanolamine PE(18:0/18:0), Sigma P3531 [HMDB08991]	C41H82NO8P	PE(18:0/18:0)	<b>HD:</b> 36:0 FA (NL) <b>SD:</b> 18:0 FA
SM(d36:1)	SM(d18:1/18:0), Sigma S0756 [HMDB12099]	C41H83N2O6P	SM(d36:1)	<b>HD:</b> ChoP
SM(d34:1)	SM(d18:1/16:0), Sigma 91553 [HMDB10169]	C39H79N2O6P	SM(d34:1)	<b>HD:</b> ChoP

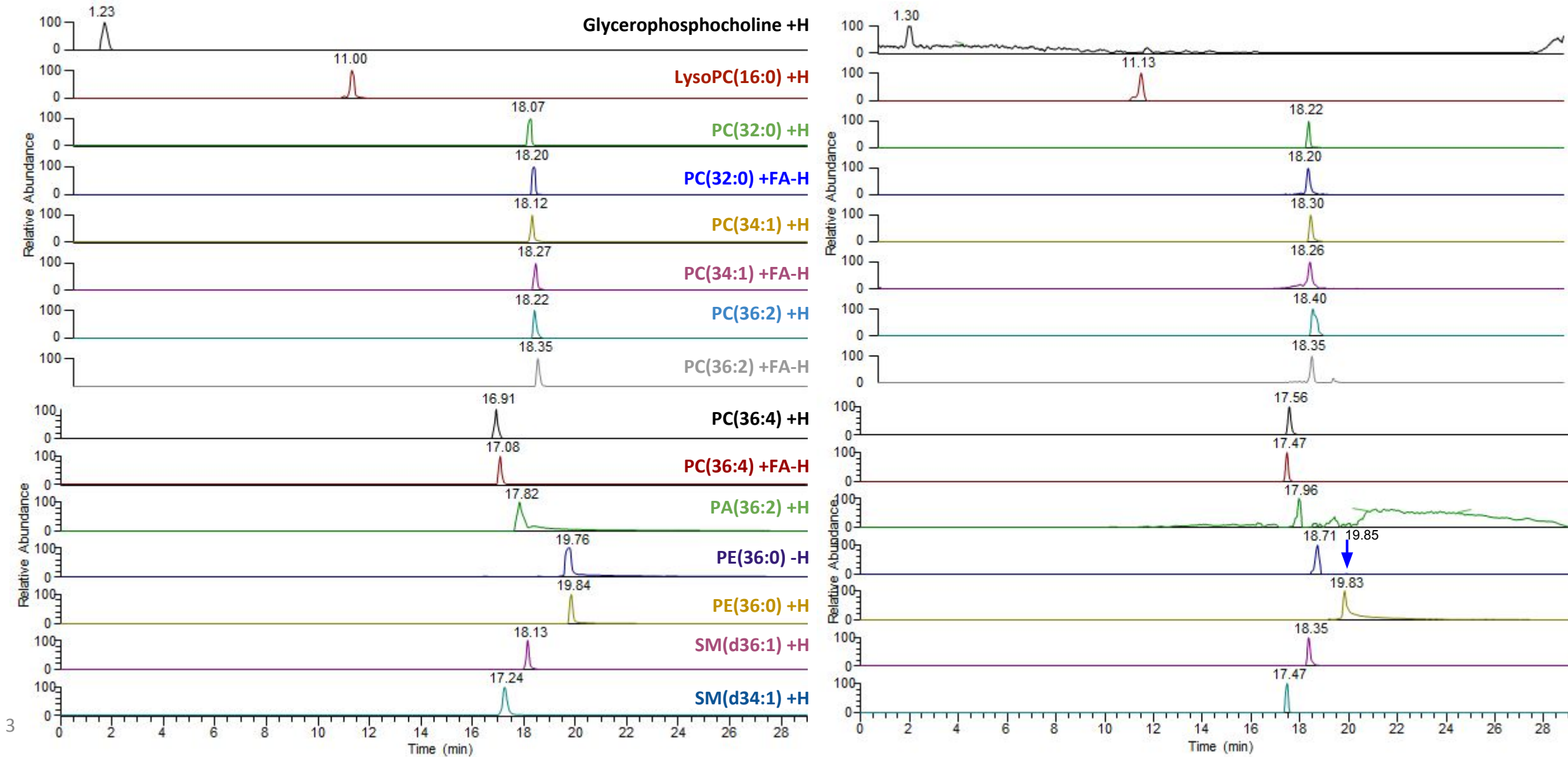
**HD:** Head-group  
**SD:** Side chain

**FA:** Fatty acid, **ChoP:** Phosphocholine, **Cer:** Ceramide, **Sph:** Sphingosine, **PEA:** Phosphorylethanolamine  
**NL :** neutral loss of head-group

# Extracted ion chromatograms (XIC)

Standard mixture XICs ( $\pm 5$  ppm)

Sample XICs ( $\pm 5$  ppm)

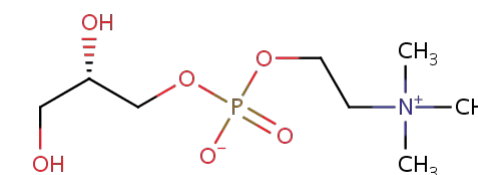


3

**Molecular annotation: C<sub>8</sub>H<sub>20</sub>NO<sub>6</sub>P, +H, m/z 258.1101, Glycerophosphocholine**

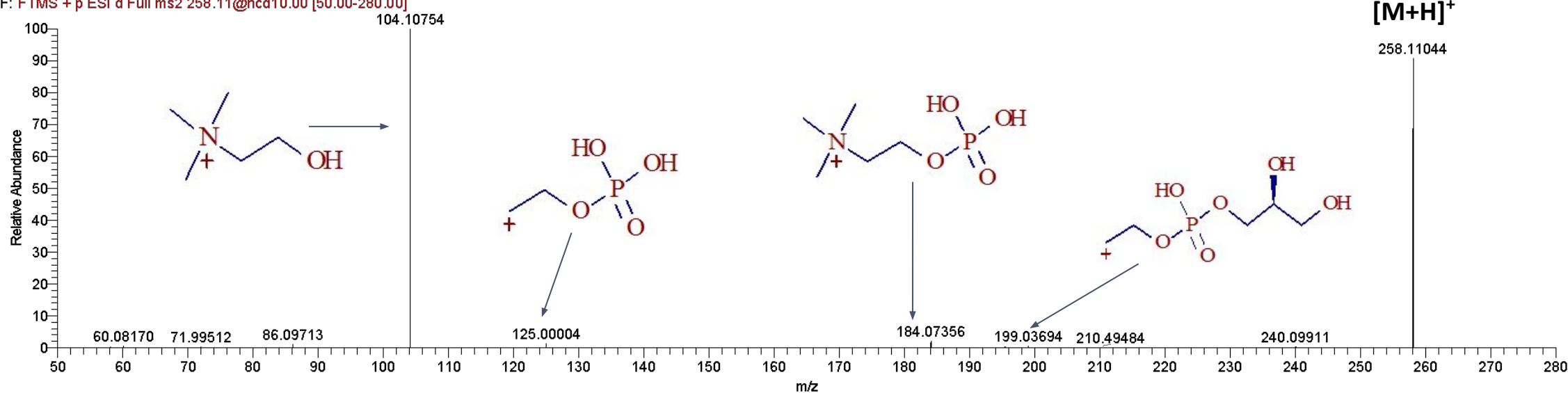
**Standard used : L-Glycerophosphorylcholine from Sigma G5291 [HMDB00086]**

**Validated as : Glycerophosphorylcholine using LC-MS/MS method 'Positive'**



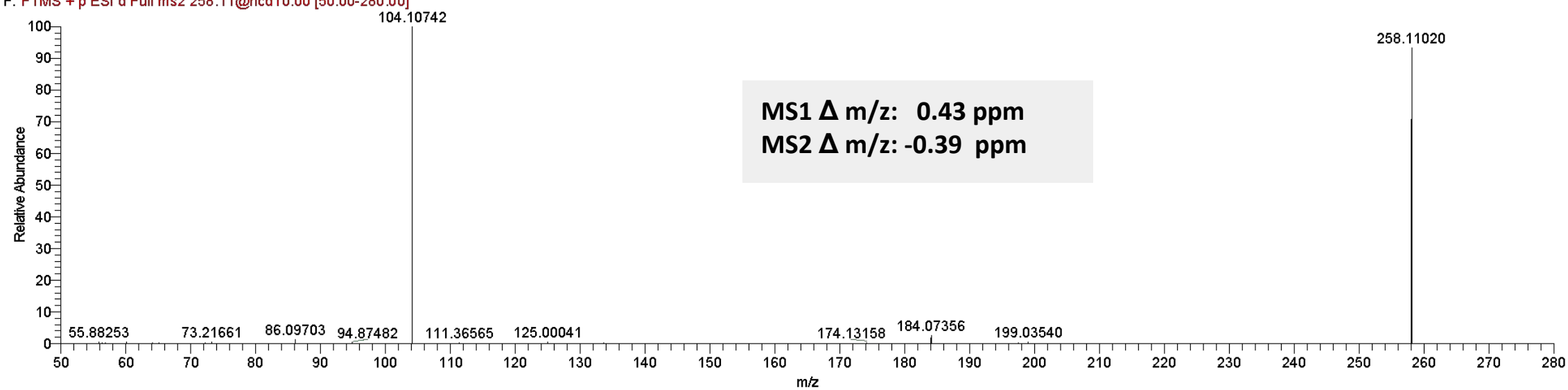
Standard

LIPID\_STD\_MIX\_FDR\_1\_2mix\_all\_10\_CE10 #565 RT: 1.44 AV: 1 NL: 2.65E7  
F: FTMS + p ESI d Full ms2 258.11@hcd10.00 [50.00-280.00]



Sample

Mouse\_Whole\_Brain\_CHCl3\_Lip\_ext\_3Xdil\_Pos\_incl\_list\_CE10 #583 RT: 1.45 AV: 1 NL: 5.83E5  
F: FTMS + p ESI d Full ms2 258.11@hcd10.00 [50.00-280.00]

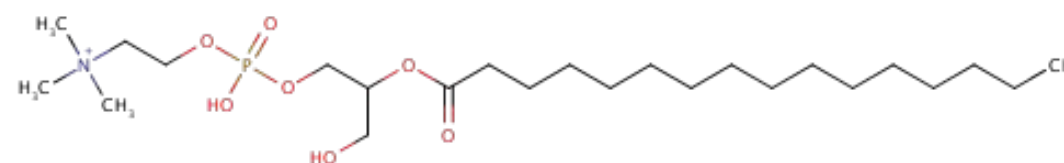


MS1  $\Delta$  m/z: 0.43 ppm  
MS2  $\Delta$  m/z: -0.39 ppm

**Molecular annotation: C<sub>24</sub>H<sub>50</sub>NO<sub>7</sub>P, +H, m/z 496.3397, LysoPC(16:0)**

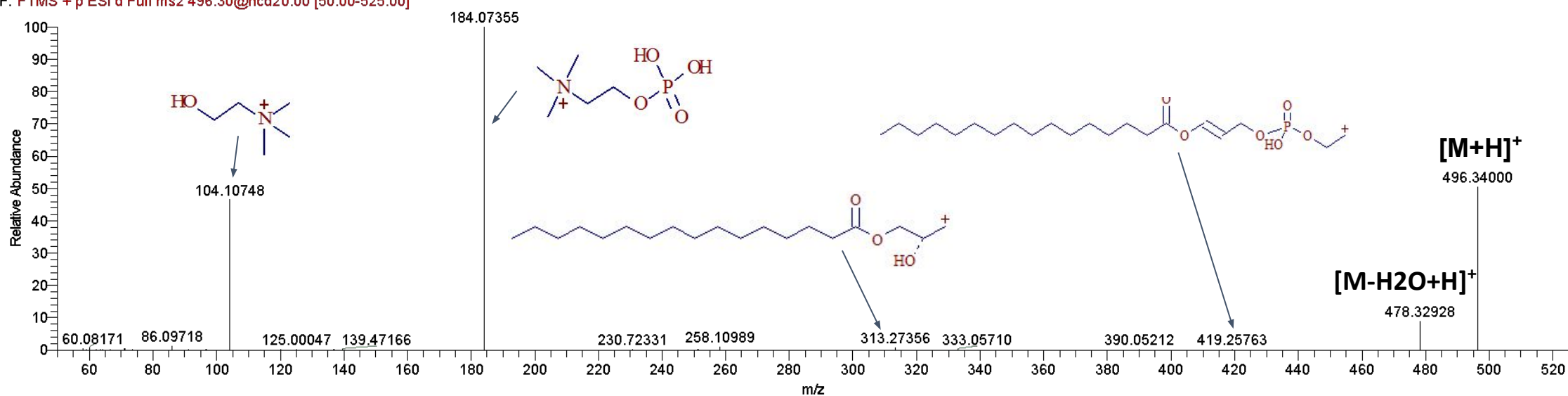
**Standard used : LysoPC(16:0/0:0) from Sigma L5254 [HMDB10382]**

**Validated as : LysoPC(16:0/0:0) using LC-MS/MS method 'Positive'**



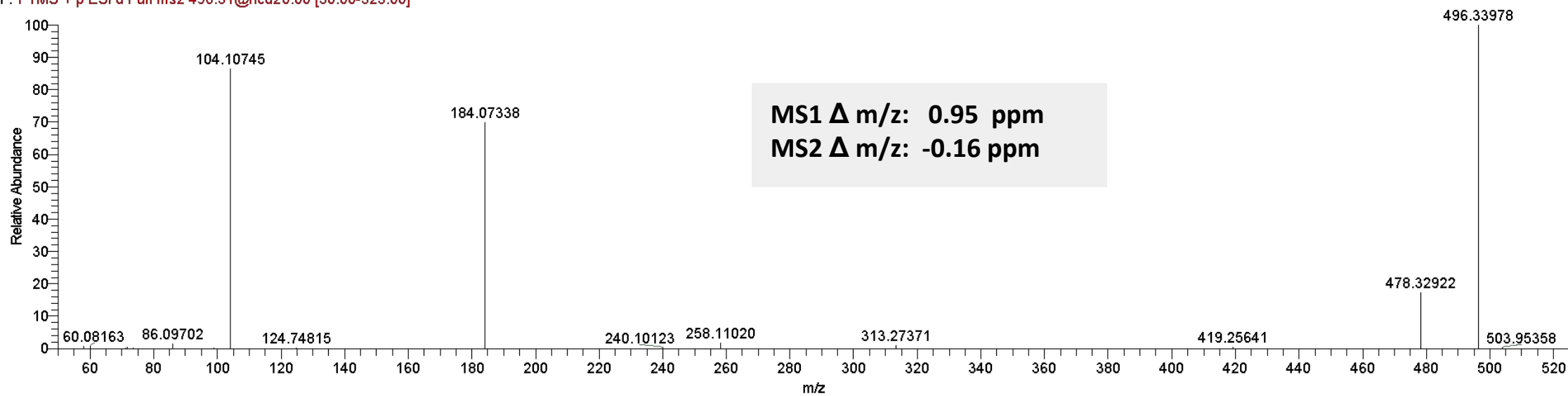
Standard

Lipid\_std\_3\_mix\_FDR3 #6347 RT: 10.86 AV: 1 NL: 2.64E9  
F: FTMS + p ESI d Full ms2 496.30@hcd20.00 [50.00-525.00]



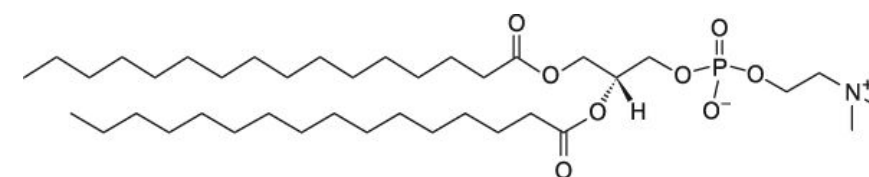
Mouse\_Whole\_Brain\_3Xdil\_CHCl3\_Lip\_ext\_Pos\_incl\_list\_CE20\_1 #5400 RT: 11.04 AV: 1 NL: 3.03E7  
F: FTMS + p ESI d Full ms2 496.51@hcd20.00 [50.00-525.00]

Sample

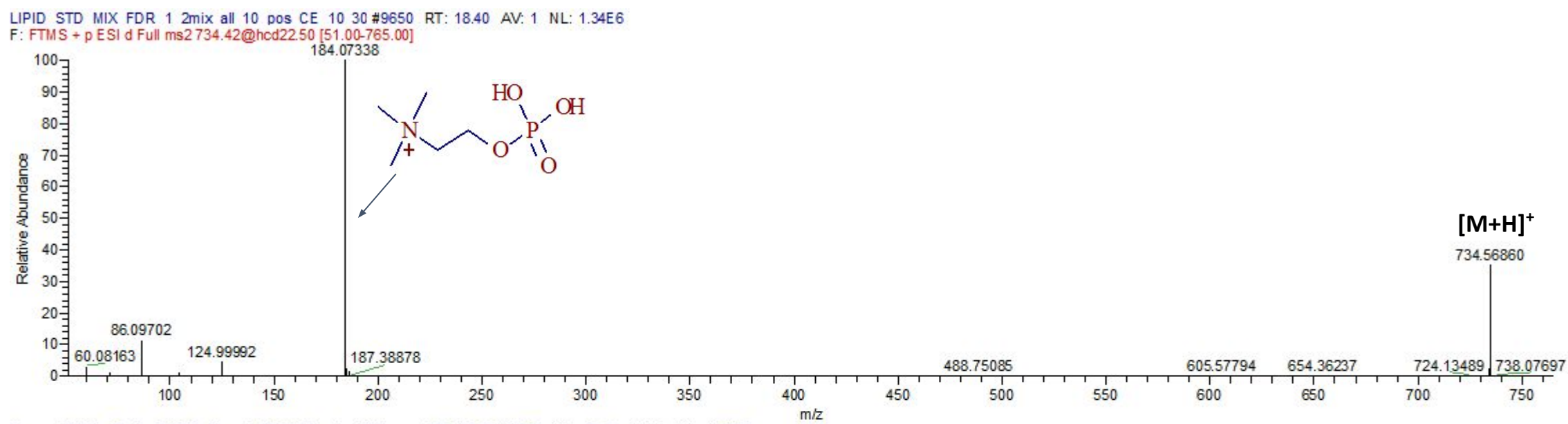


5

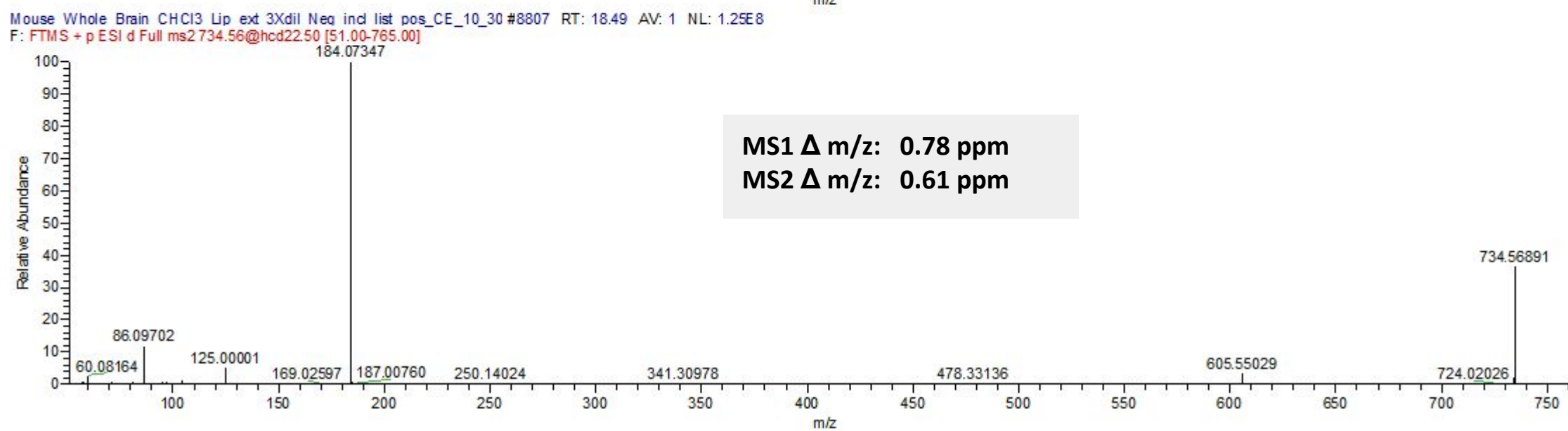
**Molecular annotation: C<sub>40</sub>H<sub>80</sub>NO<sub>8</sub>P, +H, m/z 734.56936; PC(32:0)**  
**Standard used : PC(16:0/16:0) from Avanti-ID 850355 [HMDB00564]**  
**Validated as : PC(16:0/16:0) using LC-MS/MS method 'Positive'**



Standard

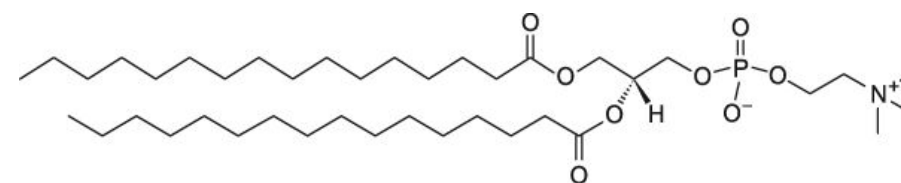


Sample

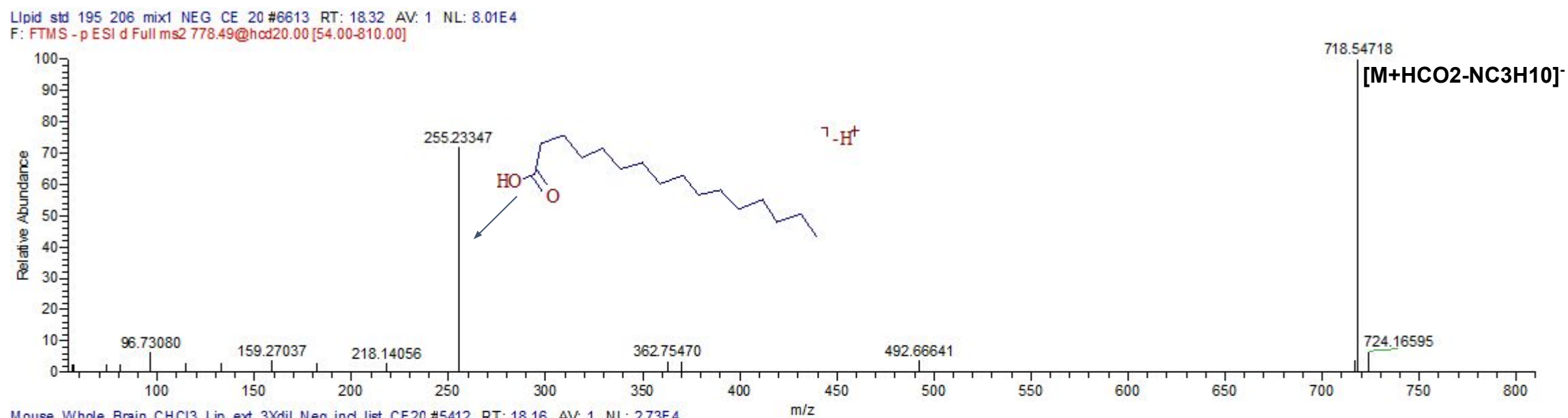


6

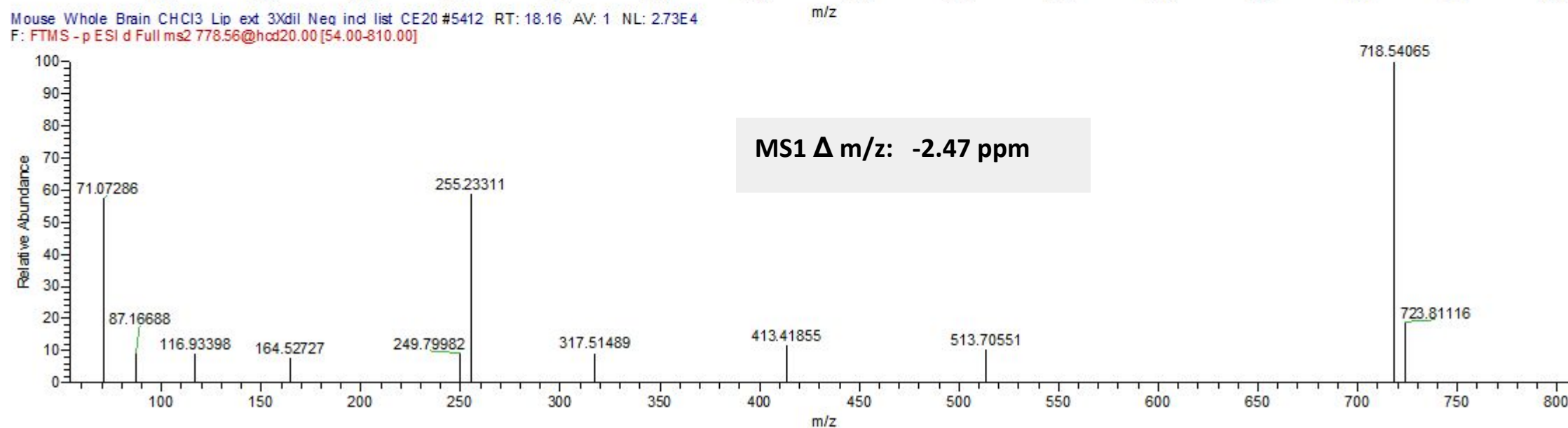
**Molecular annotation** : C<sub>40</sub>H<sub>80</sub>NO<sub>8</sub>P, +FA-H, m/z 778.55973; PC(32:0)  
**Standard used** : PC(16:0/16:0) from Avanti-ID 850355 [HMDB00564]  
**Validated as** : PC(16:0/16:0) using LC-MS/MS method 'Negative 2'



Standard



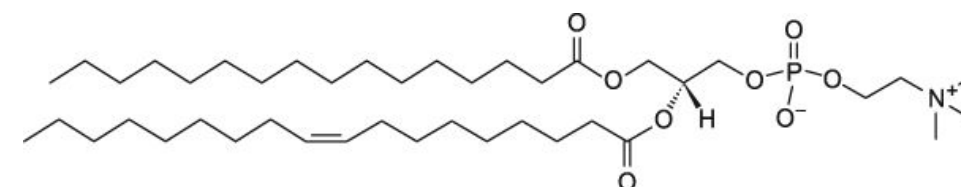
Sample



7

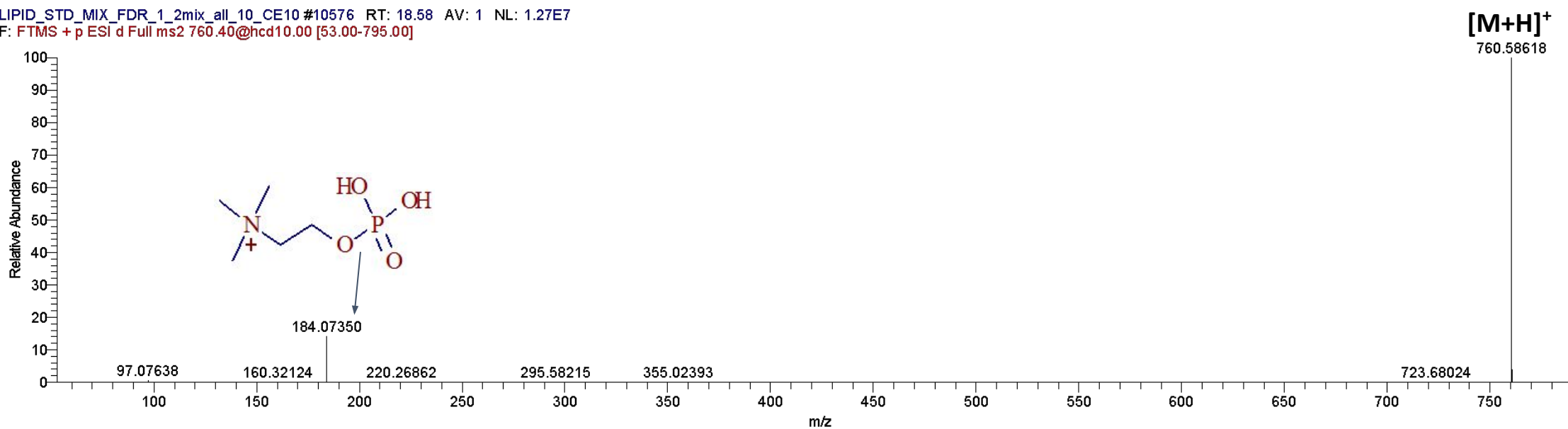


**Molecular annotation** : C<sub>42</sub>H<sub>82</sub>NO<sub>8</sub>P, +H, m/z 760.58500, PC(34:1)  
**Standard used** : PC(16:0/18:1) from Avanti-ID 850457P [HMDB07972]  
**Validated as** : PC(16:0/18:1) using LC-MS/MS method 'Positive'



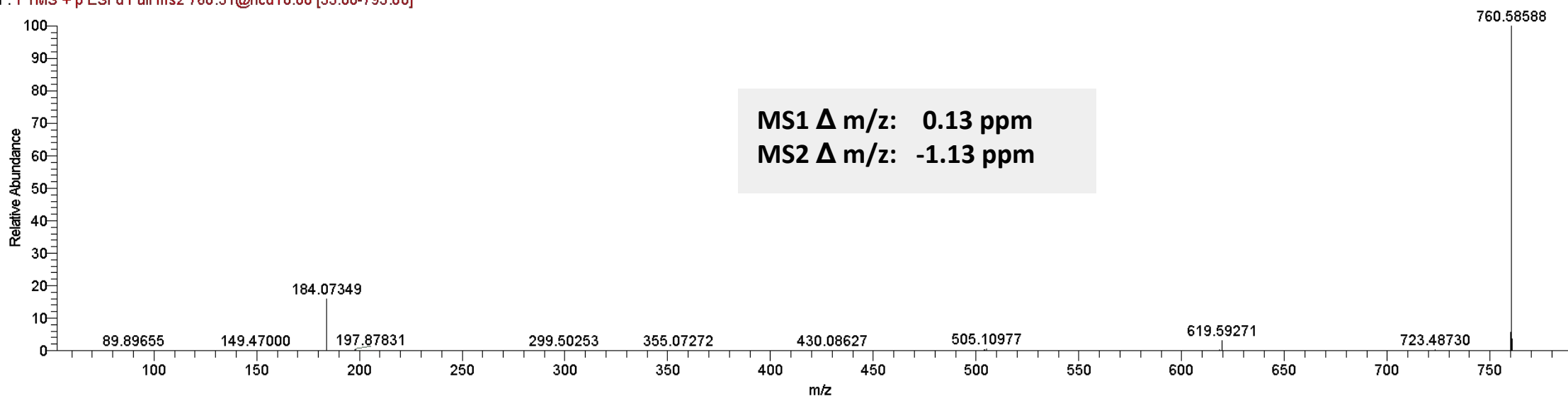
Standard

LIPID\_STD\_MIX\_FDR\_1\_2mix\_all\_10\_CE10 #10576 RT: 18.58 AV: 1 NL: 1.27E7  
 F: FTMS + p ESI d Full ms2 760.40@hcd10.00 [53.00-795.00]



Sample

Mouse Whole Brain\_CHCl3\_Lip\_ext\_3Xdil\_Pos\_incl\_list\_CE10 #9607 RT: 19.71 AV: 1 NL: 2.39E6  
 F: FTMS + p ESI d Full ms2 760.51@hcd10.00 [53.00-795.00]

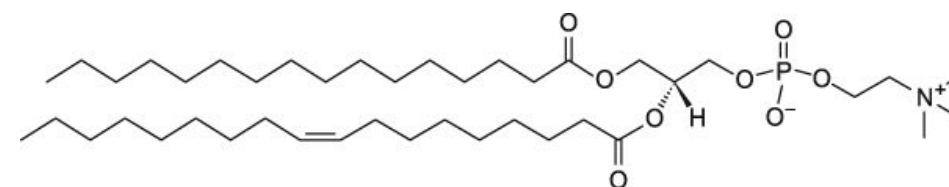


8

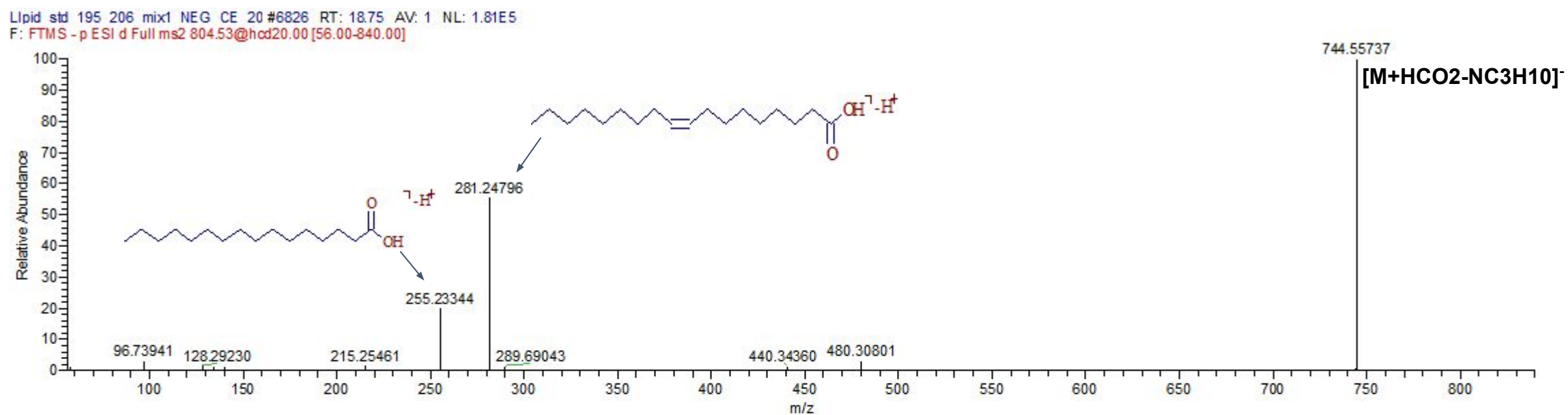
**Molecular annotation: C<sub>42</sub>H<sub>82</sub>NO<sub>8</sub>P, +FA-H, m/z 804.57538, PC(34:1)**

**Standard used : PC(16:0/18:1) from Avanti-ID 850457P [HMDB07972]**

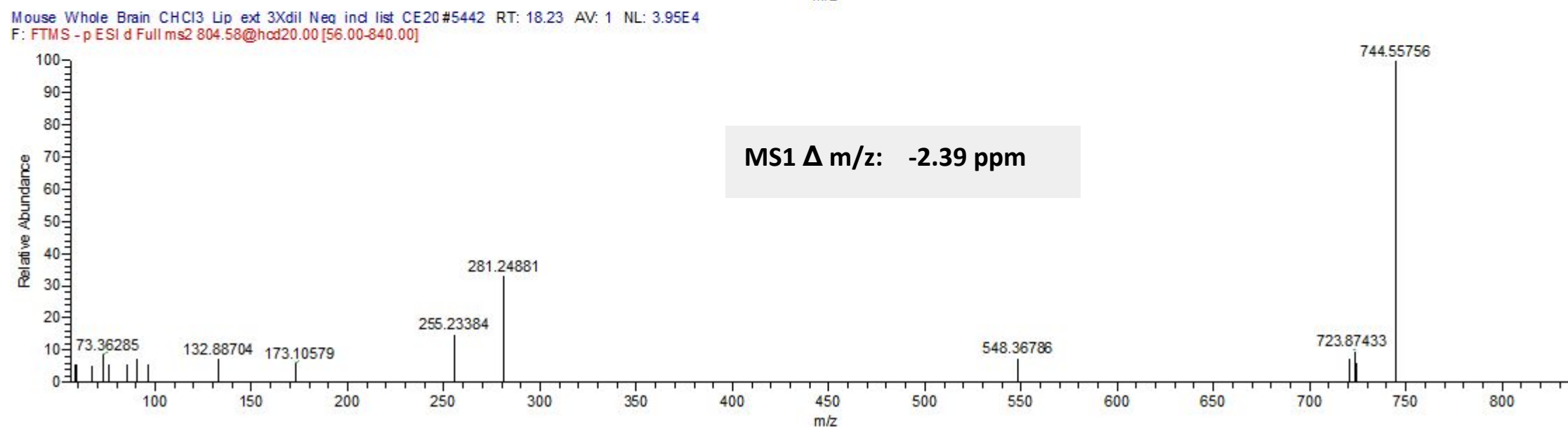
**Validated as : PC(16:0/18:1) using LC-MS/MS method 'Negative 2'**



Standard



Sample



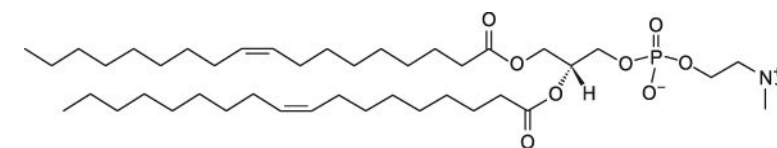
9



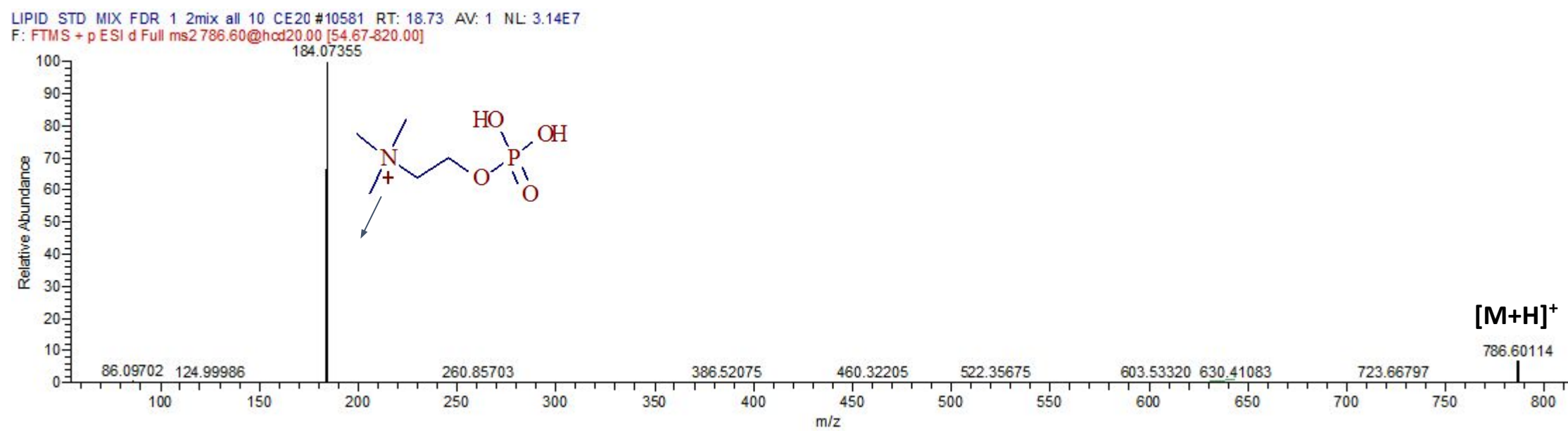
**Molecular annotation: C<sub>44</sub>H<sub>84</sub>NO<sub>8</sub>P, +H, m/z 786.6006, PC(36:2)**

**Standard used : PC(18:1(9Z)/18:1(9Z)) from Avanti-ID 850375C [HMDB00593]**

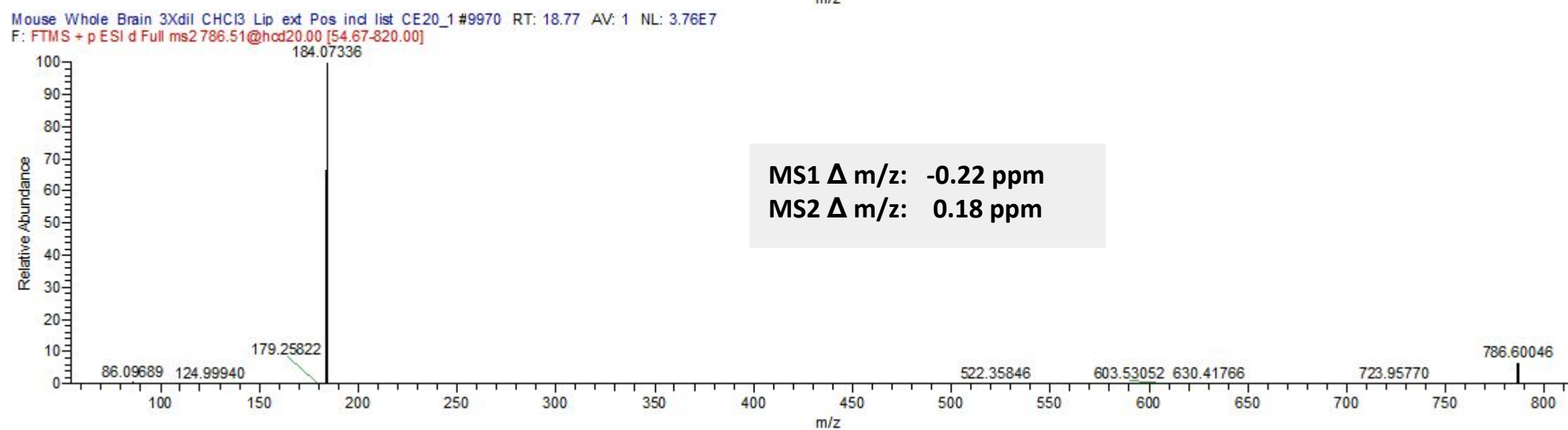
**Validated as : PC(18:1/18:1) using LC-MS/MS method 'Positive'**



Standard



Sample

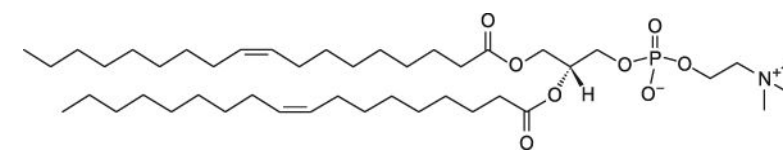


10

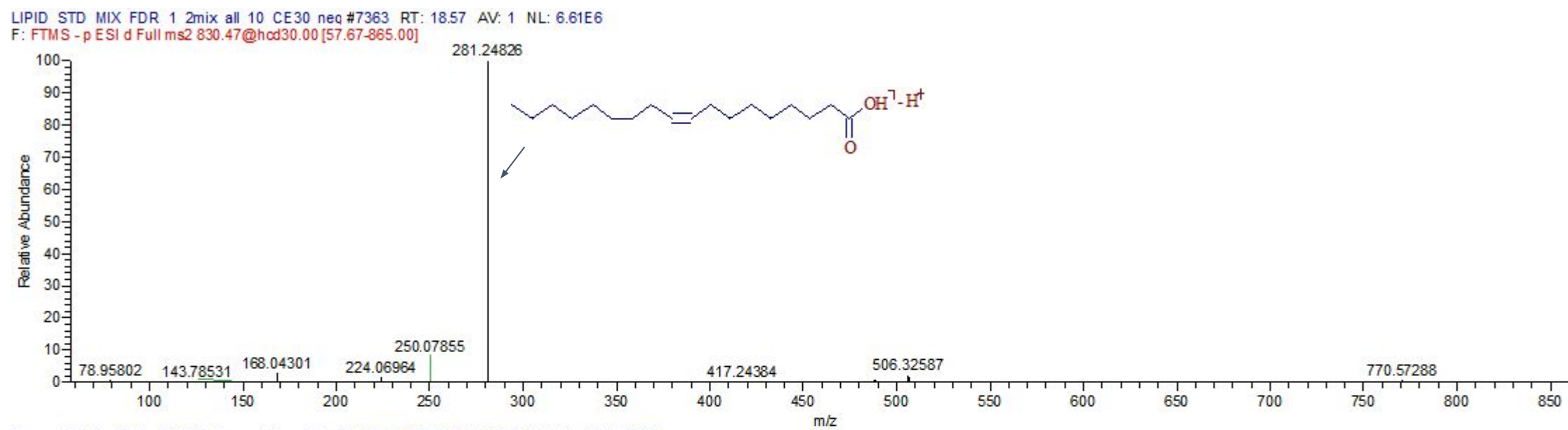
**Molecular annotation: C<sub>44</sub>H<sub>84</sub>NO<sub>8</sub>P, +FA-H, m/z 830.5910, PC(36:2)**

**Standard used : PC(18:1(9Z)/18:1(9Z)) from Avanti-ID 850375C [HMDB00593]**

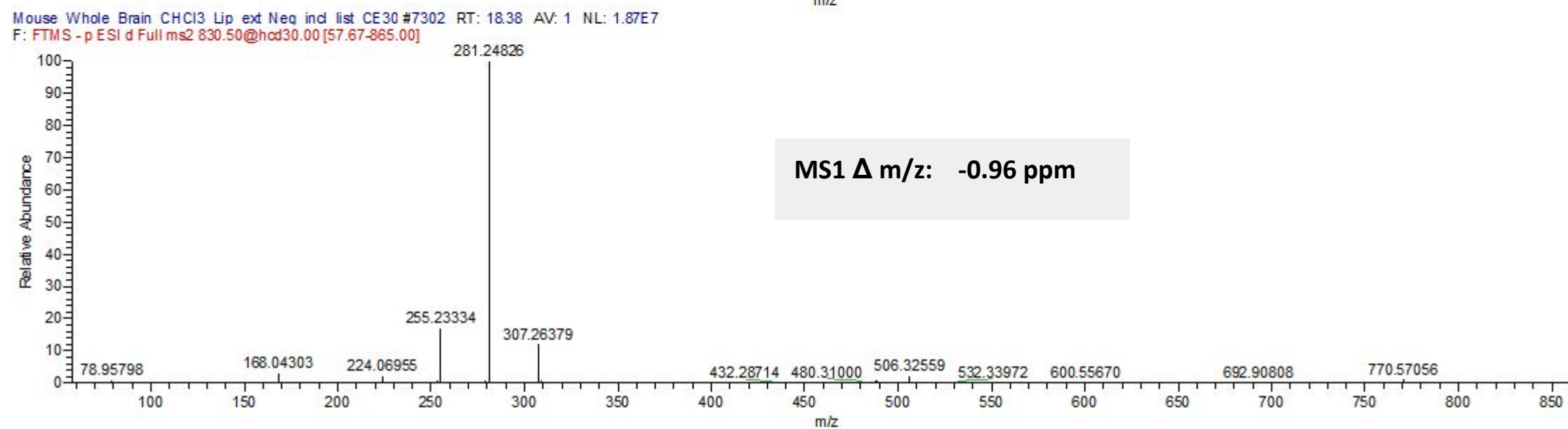
**Validated as : PC(18:1/18:1) using LC-MS/MS method 'Negative 2'**



Standard



Sample

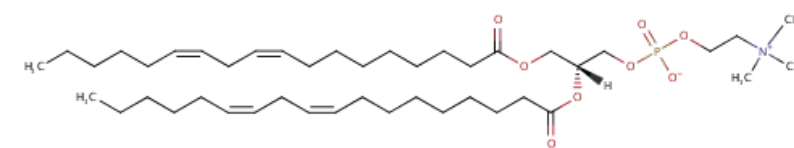


11

**Molecular annotation: C<sub>44</sub>H<sub>80</sub>NO<sub>8</sub>P, +H, m/z 782.5694, PC(36:4)**

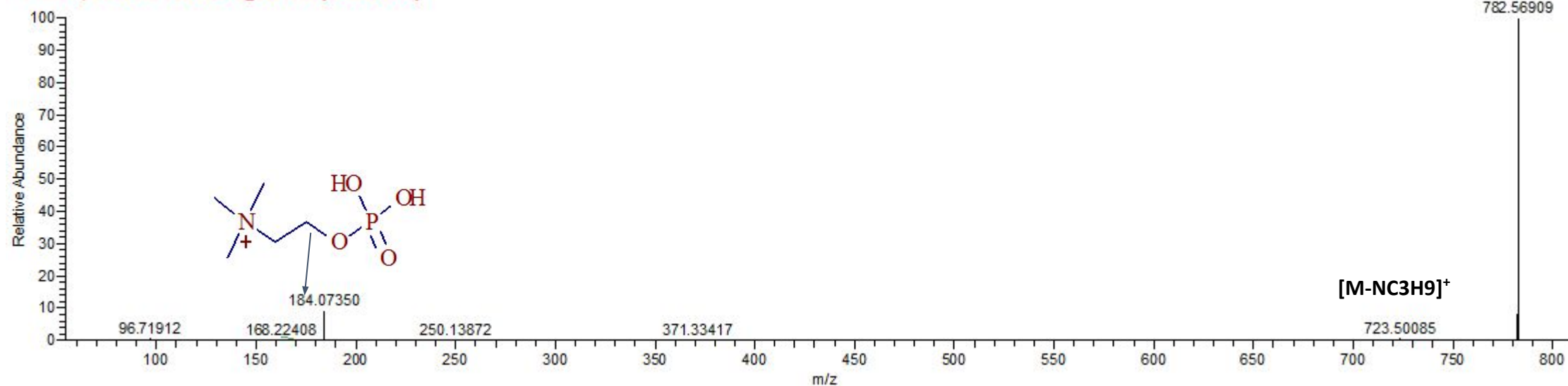
**Standard used : PC(18:2(9Z,12Z)/18:2(9Z,12Z)) from Sigma P0537 [HMDB08138]**

**Validated as : PC(18:2/18:2) using LC-MS/MS method 'Positive'**



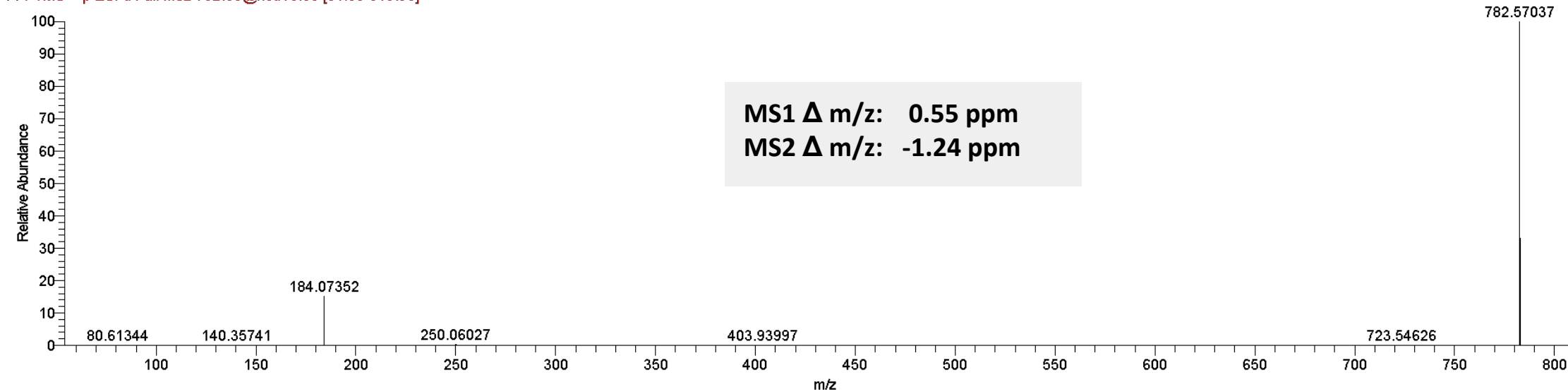
Standard

LIPID STD MIX FDR 1 2mix all 10 CE10 #10321 RT: 18.04 AV: 1 NL: 1.20E6  
F: FTMS + p ESI d Full ms2 782.44@hcd10.00 [54.33-815.00]



Sample

Mouse Whole Brain CHCl<sub>3</sub> Lip\_ext\_3Xdil\_Pos\_incl\_list\_CE10 #8518 RT: 17.49 AV: 1 NL: 2.54E8  
F: FTMS + p ESI d Full ms2 782.50@hcd10.00 [54.33-815.00]

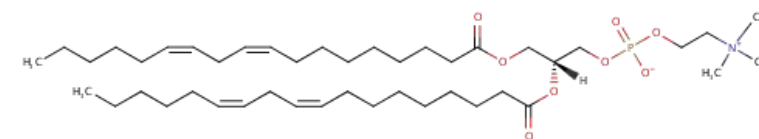


12

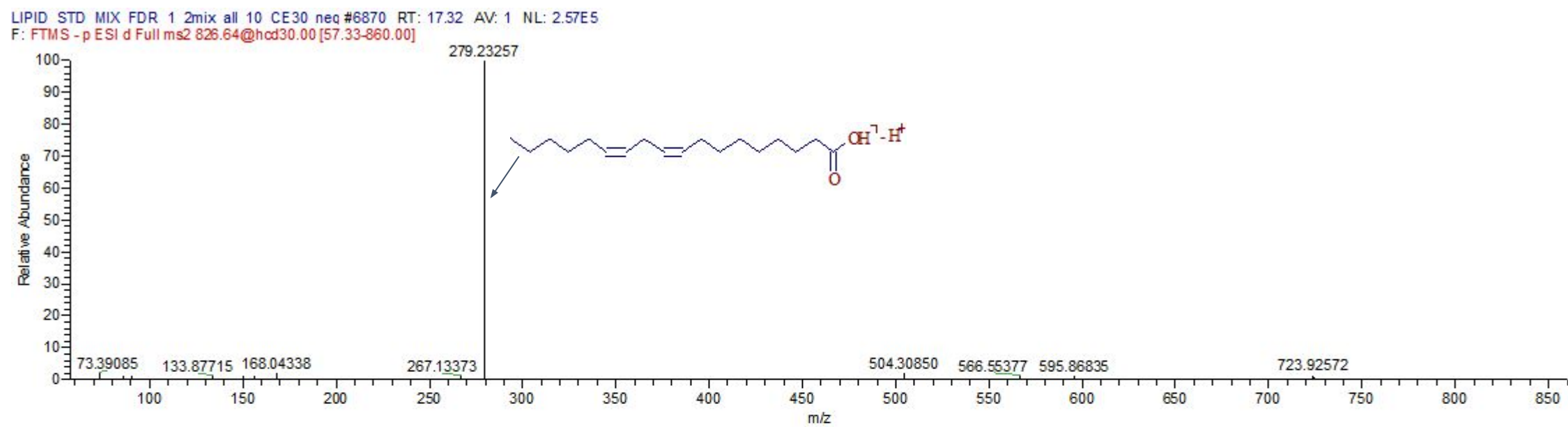
**Molecular annotation: C<sub>44</sub>H<sub>80</sub>NO<sub>8</sub>P, +FA-H, m/z 826.55973, PC(36:4)**

**Standard used : PC(18:2(9Z,12Z)/18:2(9Z,12Z)) from Sigma P0537 [HMDB08138]**

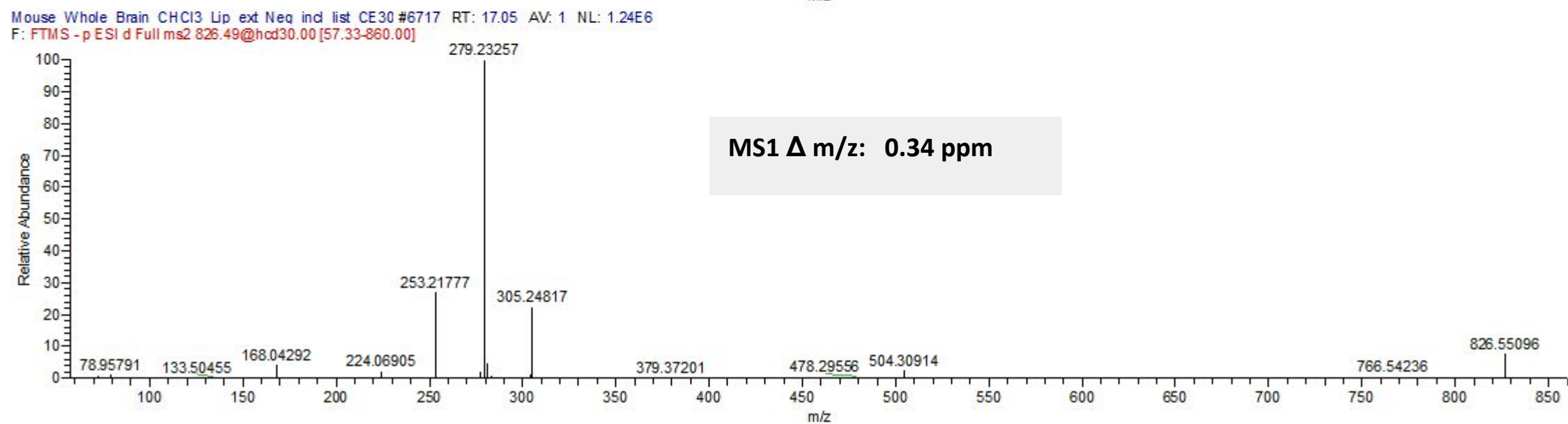
**Validated as : PC(18:2/18:2) using LC-MS/MS method 'Negative 2'**



Standard



Sample

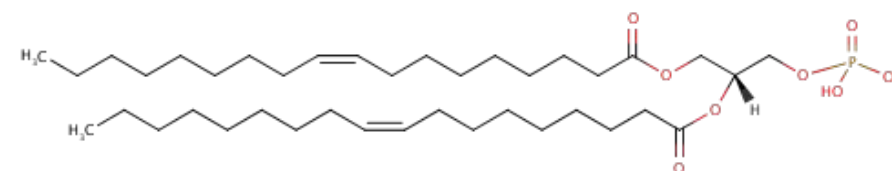


13

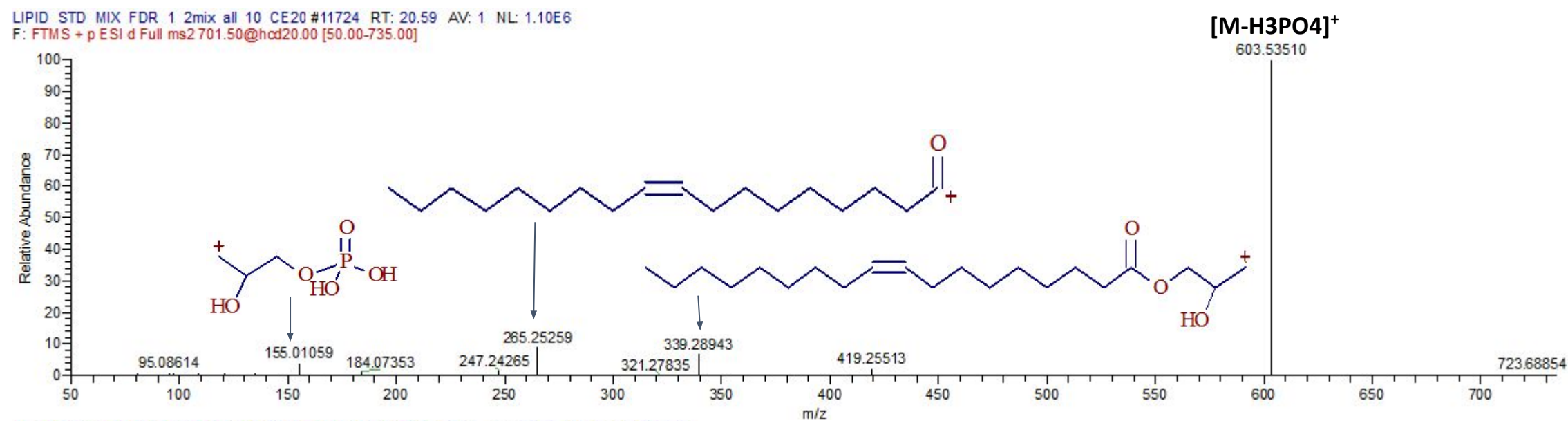
**Molecular annotation: C<sub>39</sub>H<sub>73</sub>O<sub>8</sub>P, +H, m/z 701.5115, PA(36:2)**

**Standard used : L-Phosphatidic acid, PA(18:1/18:1) from Sigma P2767 [HMDB07865]**

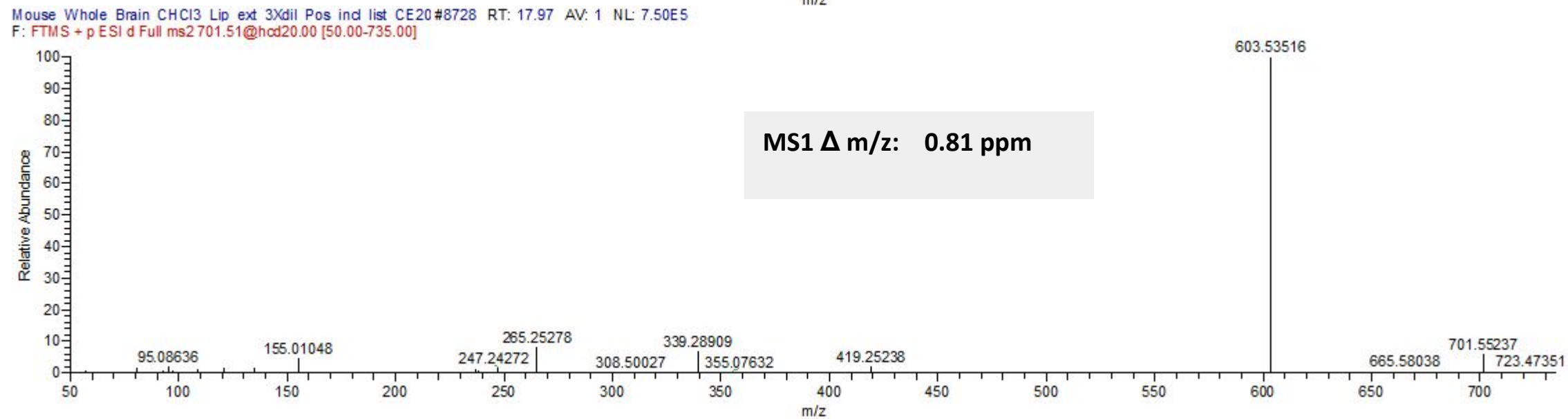
**Validated as : PA (18:1/18:1) using LC-MS/MS method 'Positive'**



Standard



Sample



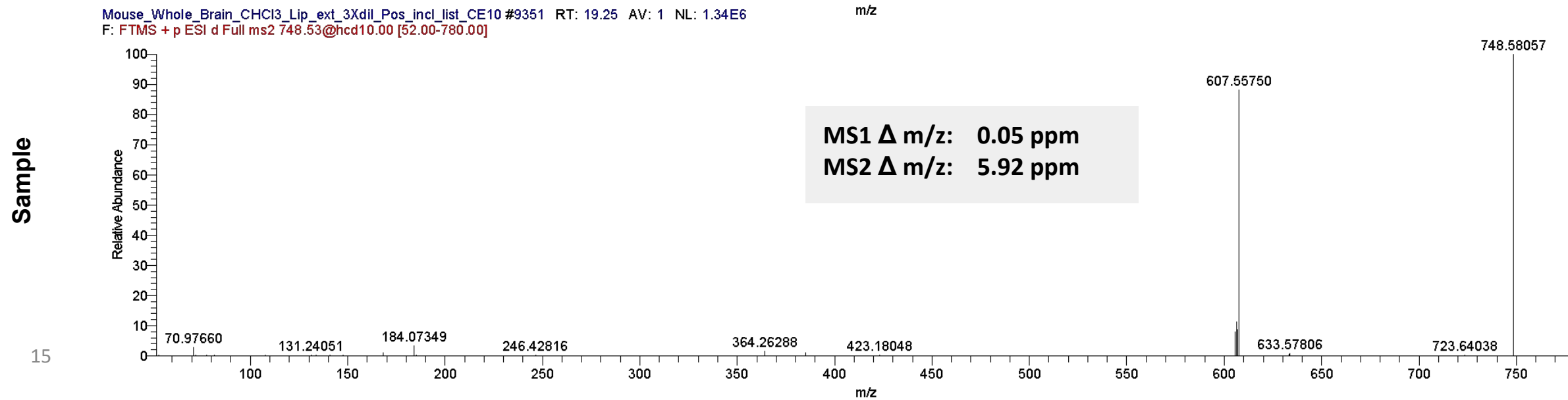
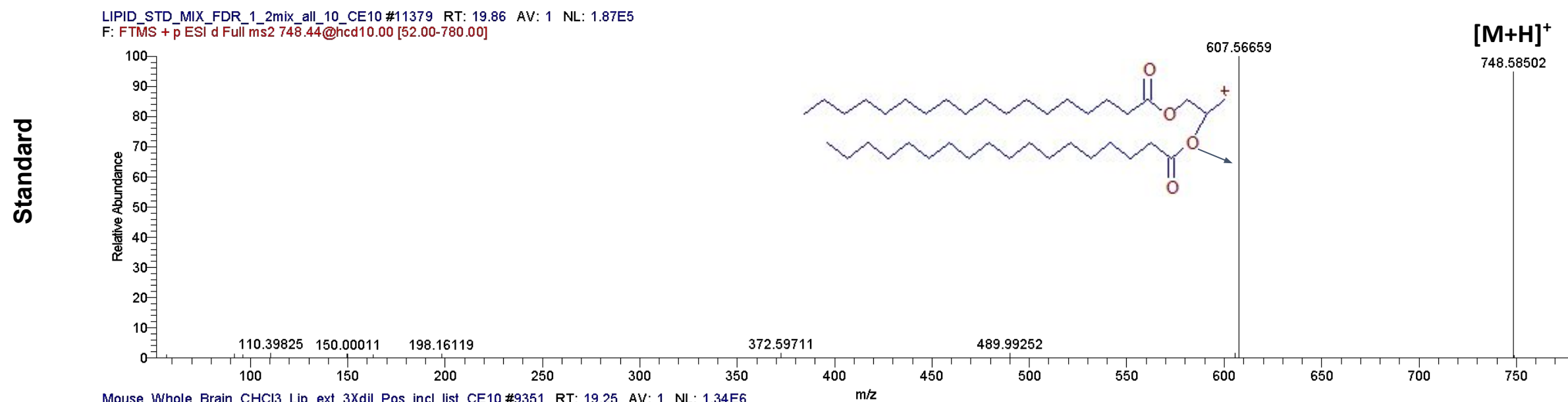
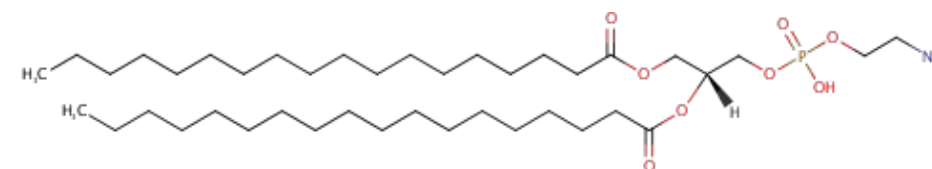
14



**Molecular annotation: C<sub>41</sub>H<sub>82</sub>NO<sub>8</sub>P, +H, m/z 748.58500, PE(36:0)**

**Standard used : Phosphatidylethanolamine from Sigma P3531[HMDB08991]**

**Validated as : PE(18:0/18:0) using LC-MS/MS method 'Positive'**

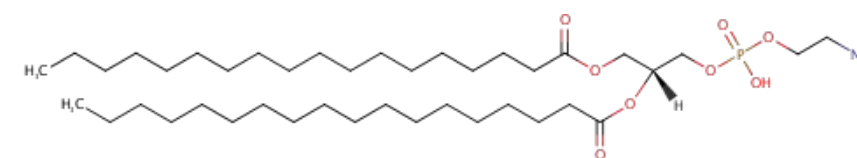


15

**Molecular annotation: C41H82NO8P, -H, m/z 746.56998, PE(36:0)**

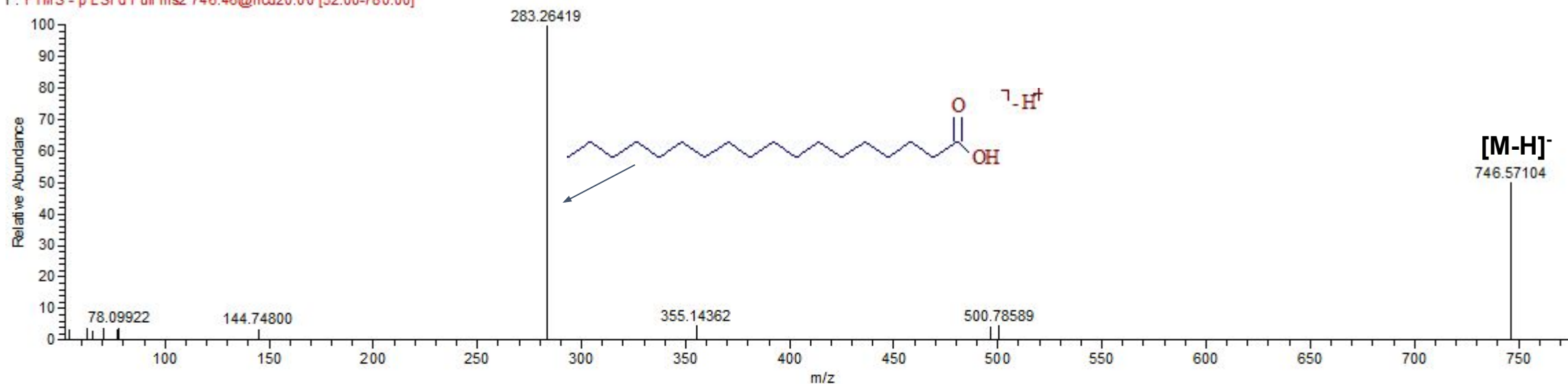
**Standard used : Phosphatidylethanolamine from Sigma P3531 [HMDB08991]**

**Validated as : PE(18:0/18:0) using LC-MS/MS method 'Negative 1'**



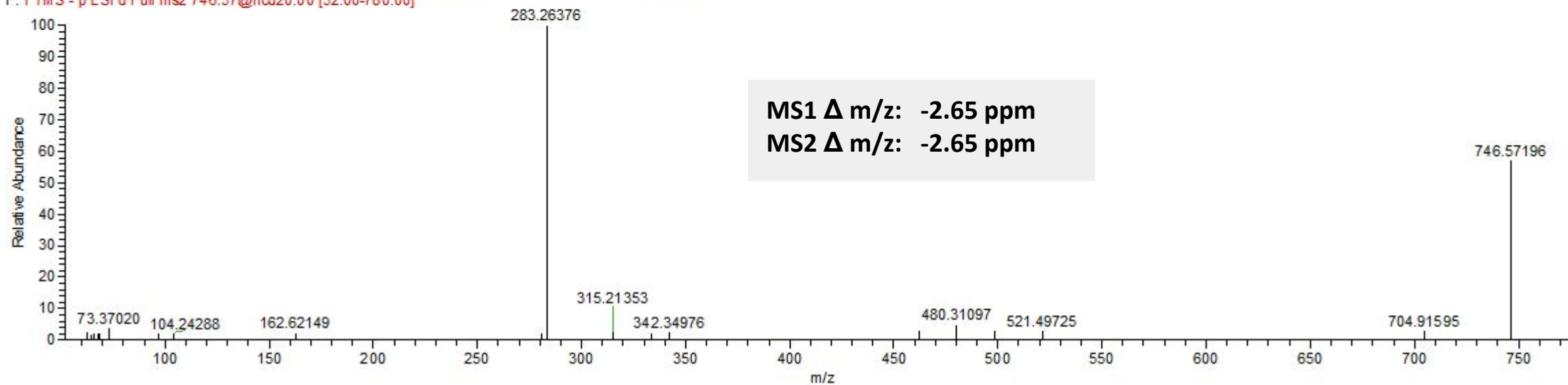
**Standard**

Lipid std 195 206 mix1 NEG CE 20 #7200 RT: 19.71 AV: 1 NL: 6.10E4  
F: FTMS - p ESI d Full ms2 746.46@hcd20.00 [52.00-780.00]



**Sample**

Mouse Whole Brain CHCl3 Lip ext 3Xdil Neg ind list CE20 #5984 RT: 19.77 AV: 1 NL: 1.01E5  
F: FTMS - p ESI d Full ms2 746.57@hcd20.00 [52.00-780.00]

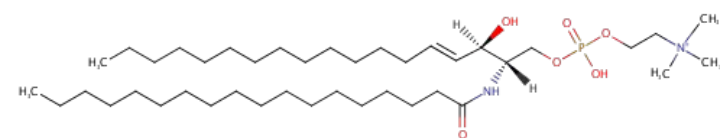


16

**Molecular annotation: C<sub>41</sub>H<sub>83</sub>N<sub>2</sub>O<sub>6</sub>P, +H, m/z 731.60608, SM(d36:1)**

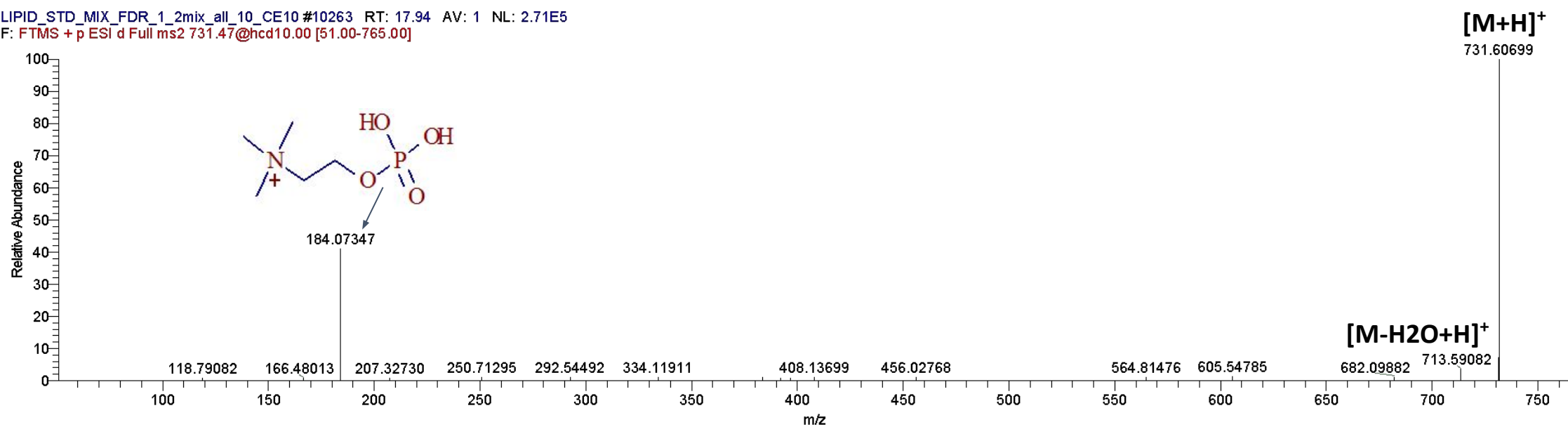
**Standard used : SM(d18:1/18:0) from Sigma S0756 [HMDB12099]**

**Validated as : SM(d36:1) using LC-MS/MS method 'Positive'**



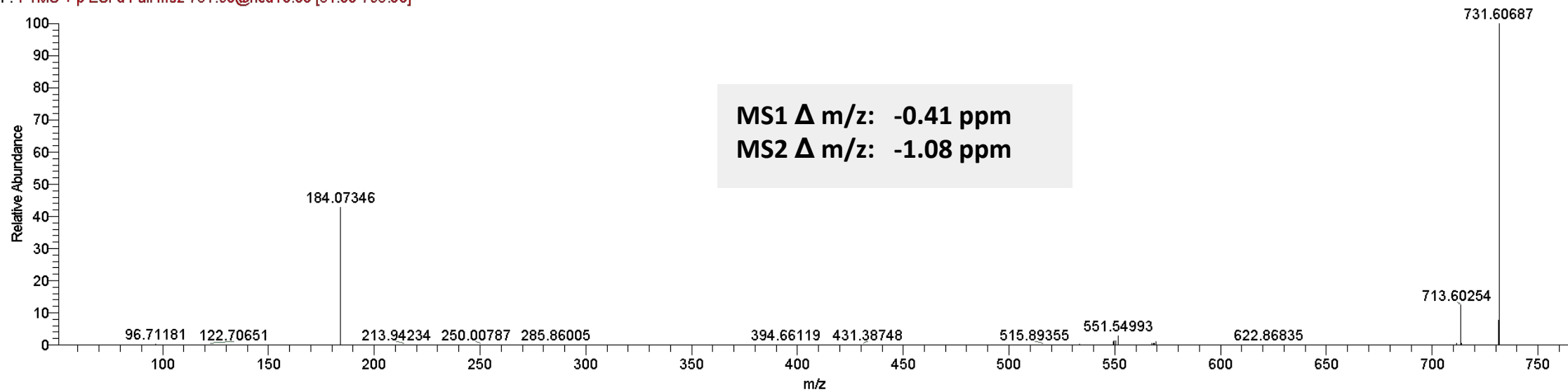
**Standard**

LIPID\_STD\_MIX\_FDR\_1\_2mix\_all\_10\_CE10 #10263 RT: 17.94 AV: 1 NL: 2.71E5  
F: FTMS + p ESI d Full ms2 731.47@hcd10.00 [51.00-765.00]



**Sample**

Mouse\_Whole\_Brain\_CHCl3\_Lip\_ext\_3Xdil\_Pos\_incl\_list\_CE10 #8944 RT: 18.43 AV: 1 NL: 1.11E7  
F: FTMS + p ESI d Full ms2 731.60@hcd10.00 [51.00-765.00]



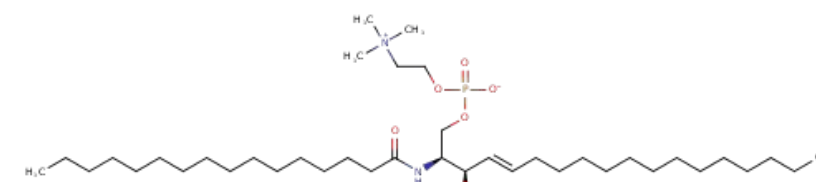
17



**Molecular annotation: C<sub>39</sub>H<sub>79</sub>N<sub>2</sub>O<sub>6</sub>P, +H, m/z 703.5748, SM(d34:1)**

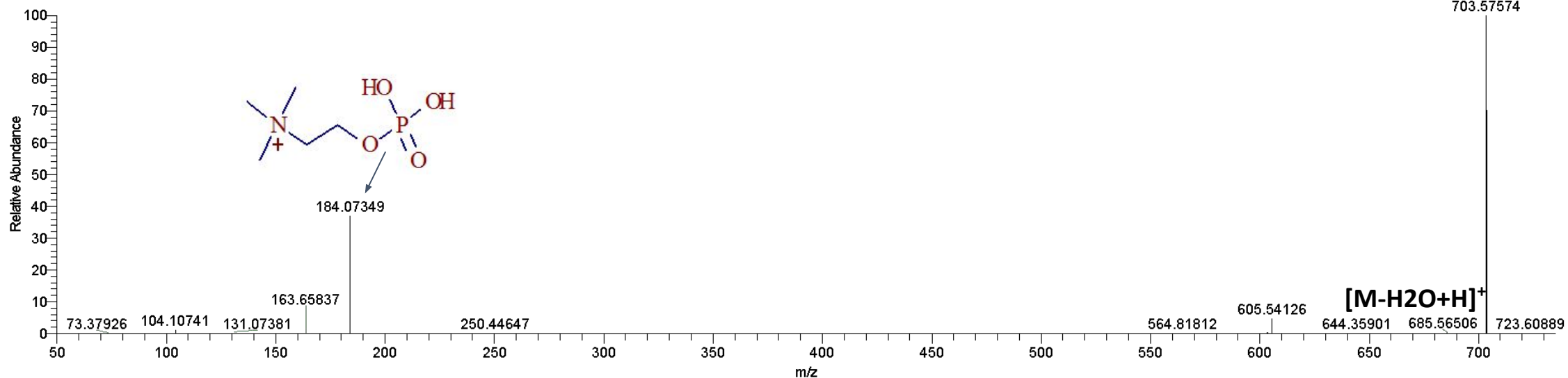
Standard used : SM(d18:1/16:0) from Sigma 91553 [HMDB61712]

Validated as : SM(d34:1) using LC-MS/MS method 'Positive'



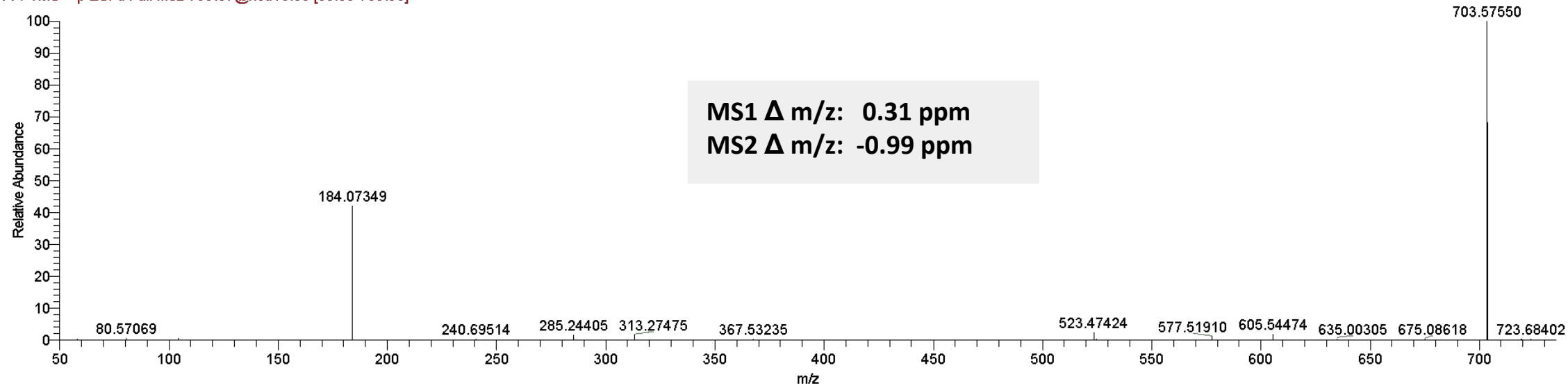
Standard

LIPID\_STD\_MIX\_FDR\_1\_2mix\_all\_10\_CE10 #10203 RT: 17.84 AV: 1 NL: 4.38E7  
F: FTMS + p ESI d Full ms2 703.56@hcd10.00 [50.00-735.00]



Sample

Mouse\_Whole\_Brain\_CHCl3\_Lip\_ext\_3Xdil\_Pos\_incl\_list\_CE10 #8664 RT: 17.78 AV: 1 NL: 8.80E5  
F: FTMS + p ESI d Full ms2 703.57@hcd10.00 [50.00-735.00]



18

## 2. LC-MS/MS validation of annotations with no authentic standards available

## Metabolite annotations with no standards were available

Annotation	Sum formula	Validated as	Validation summary
PC(34:0)	C42H84NO8P	PC(16:0_18:0)	<b>HD:</b> ChoP <b>SD:</b> 16:0 FA, 18:0 FA
PE(P-38:5)	C43H76NO7P	PE(P-18:1_20:4)	<b>HD:</b> 38:5 FA (NL) <b>SD:</b> 20:4 FA, 18:1 FA+PEA
PE(40:6)	C45H78NO8P	PE(18:0_22:6)	<b>HD:</b> 40:6 FA (NL) <b>SD:</b> 18:0 FA, 22:6 FA
PS(36:1)	C42H80NO10P	PS(18:0_18:1)	<b>HD:</b> 36:1 FA <b>SD:</b> 18:0 FA, 18:1 FA
Galactosylceramide(d42:2)	C48H91NO8	Galactosylceramide or Glucosylceramide (d18:1_24:1)	<b>HD:</b> Cer, Cer-H2O, Cer-2H2O (NL) <b>SC:</b> 18:0 FA-amide, Sph-H2O, Sph-2H2O according to <i>Boutin et al. (2016) Anal Chem. 2016 Feb 2;88(3):1856-63</i>
PC(34:2)	C42H80NO8P	PC(16:0_18:2)	<b>HD:</b> ChoP <b>SD:</b> 16:0 FA, 18:02 FA

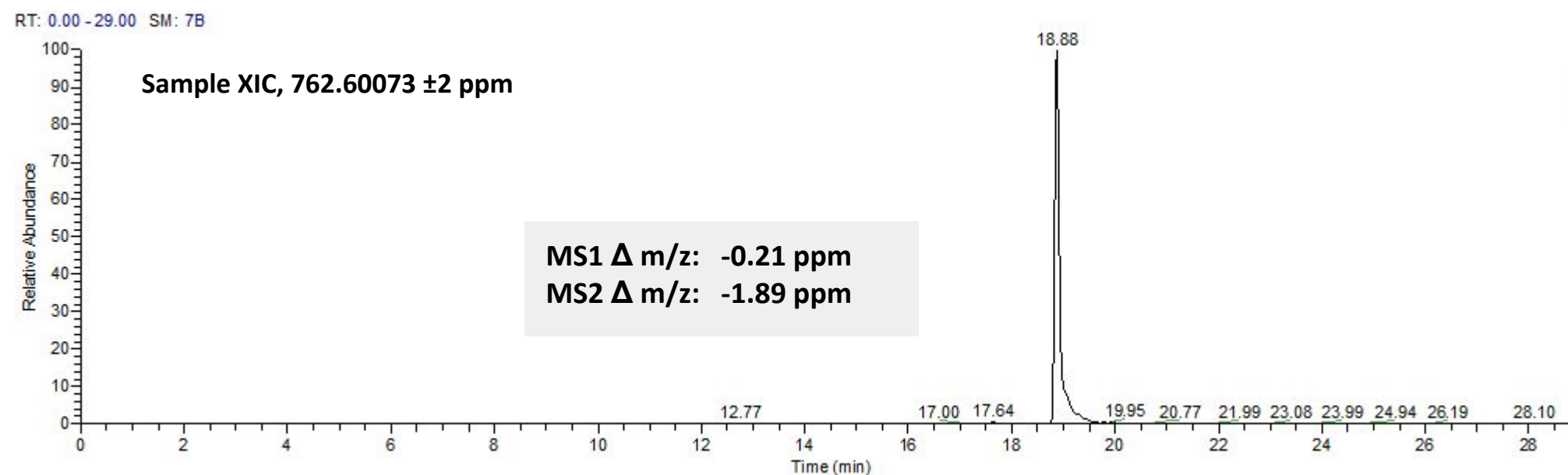
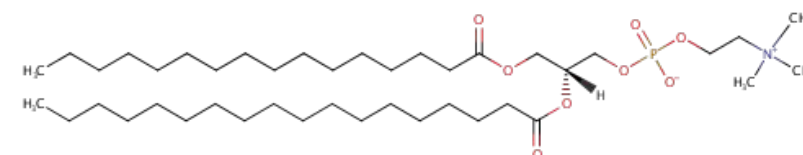
**HD:** Head-group  
**SD:** Side chain

**FA:** Fatty acid, **ChoP:** Phosphocholine, **Cer:** Ceramide, **Sph:** Sphingosine, **PEA:** Phosphorylethanolamine  
**NL :** neutral loss of head-group

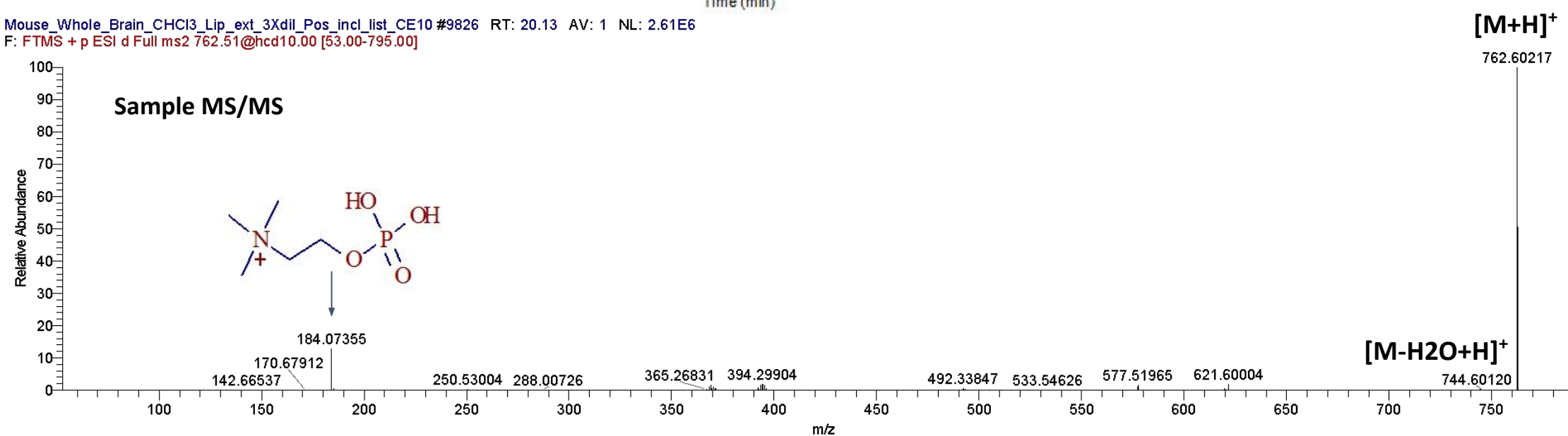
**Molecular annotation** : C<sub>42</sub>H<sub>84</sub>NO<sub>8</sub>P, +H, m/z 762.60073, PC(34:0)

**No standard**

**Validated as** : PC(34:0) [HMDB07970] using LC-MS/MS method 'Positive'



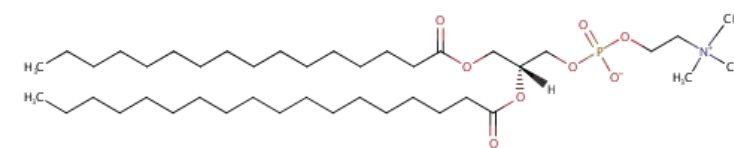
Mouse\_Whole\_Brain\_CHCl3\_Lip\_ext\_3Xdil\_Pos\_incl\_list\_CE10 #9826 RT: 20.13 AV: 1 NL: 2.61E6  
F: FTMS + p ESI d Full ms2 762.51@hcd10.00 [53.00-795.00]



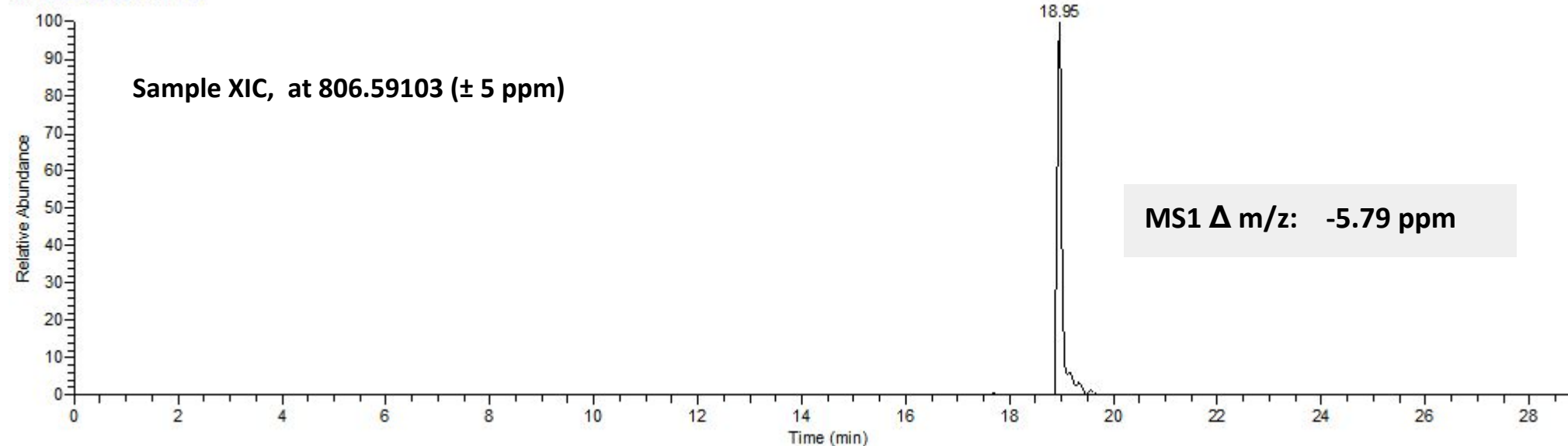
**Molecular annotation** : C<sub>42</sub>H<sub>84</sub>NO<sub>8</sub>P, +FA-H, m/z 806.59103, PC(34:0)

**No standard**

**Validated as** : PC(16:0\_18:0) [HMDB07865] using LC-MS/MS method 'Negative 2'



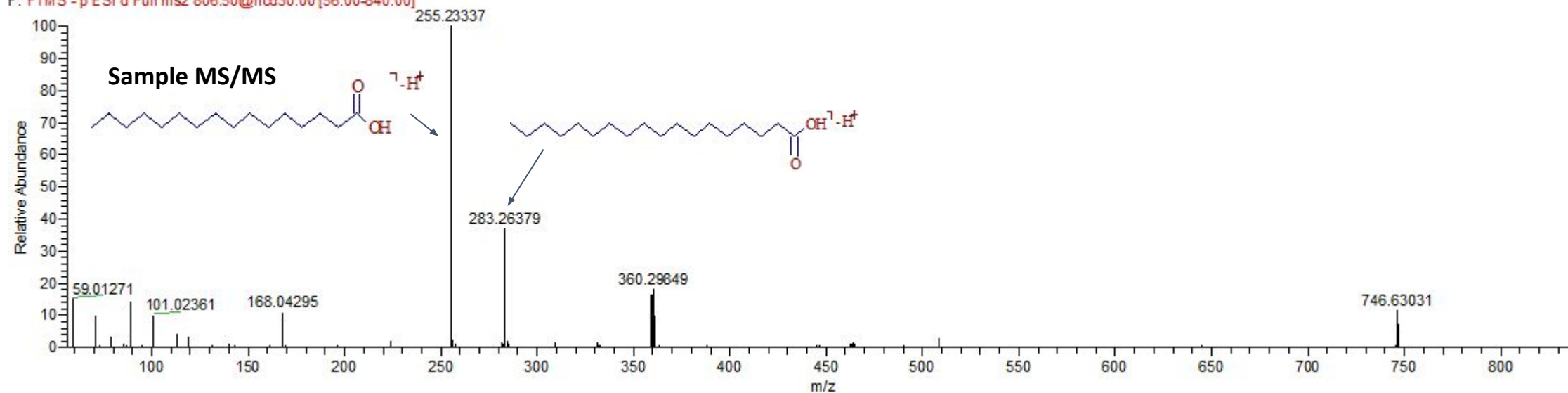
RT: 0.00 - 29.00 SM: 7B



NL: 2.62E8

Base Peak m/z=  
806.58700-806.59506 F:  
ms MS  
Mouse\_Whole\_Brain\_CH  
Cl3\_Lip\_ext\_Neg\_incl\_lis  
t\_CE30

Mouse Whole Brain CHCl<sub>3</sub> Lip ext Neg incl list CE30 #7756 RT: 19.41 AV: 1 NL: 3.60E6  
F: FTMS - p ESI d Full ms2 806.50@hcd30.00 [56.00-840.00]



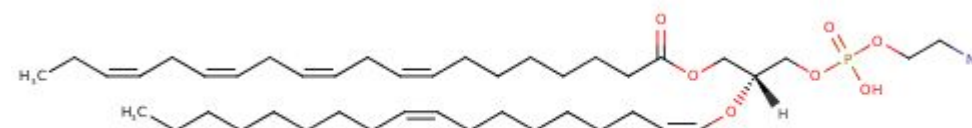




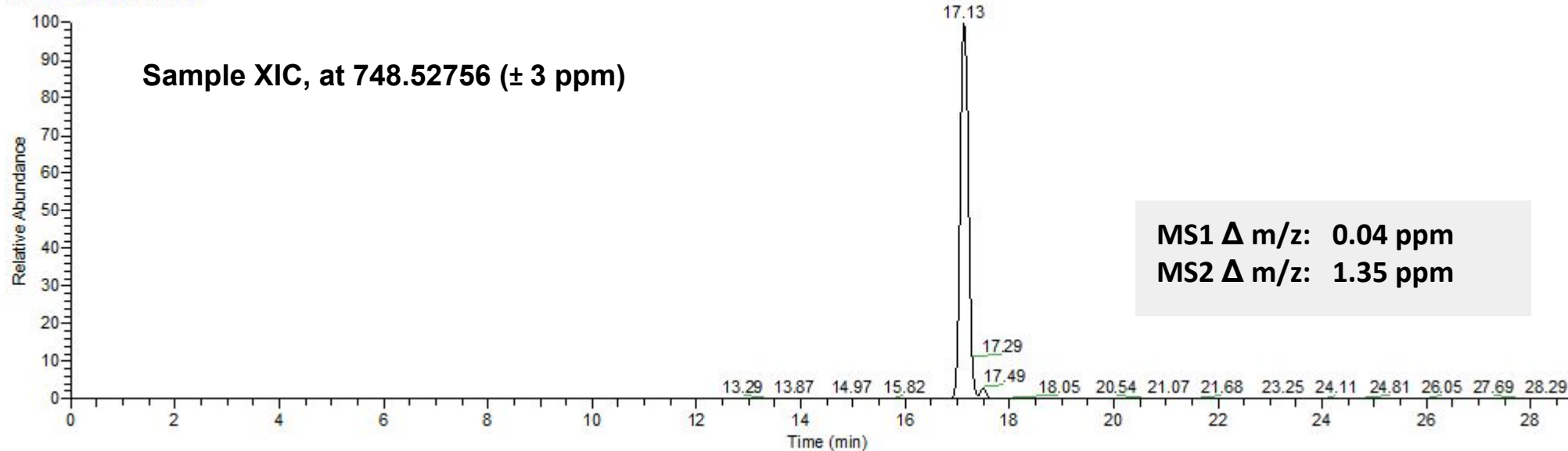
**Molecular annotation** : C<sub>43</sub>H<sub>76</sub>NO<sub>7</sub>P, -H, m/z 748.52756, PE(P-38:5)

**No standard**

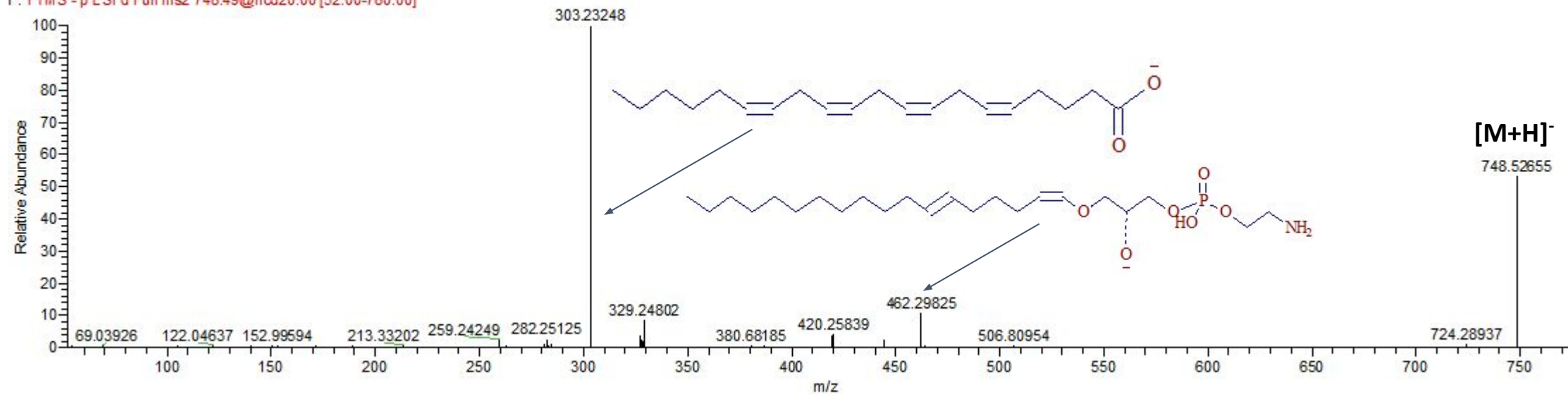
**Validated as** : PE(P-18:1\_20:4) using LC-MS/MS method 'Negative 1'



RT: 0.00 - 29.00 SM: 7B



Mouse Brain2 CHCL3 Lip old ext Neg New incl list CE20 #8963 RT: 18.02 AV: 1 NL: 4.39E5  
F: FTMS - p ESI d Full ms2 748.49@hcd20.00 [52.00-780.00]



24

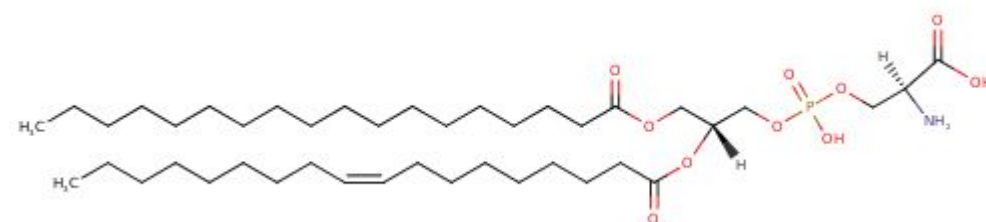




**Molecular annotation** : C<sub>42</sub>H<sub>80</sub>NO<sub>10</sub>P, +H, m/z 790.55918, PS(36:1)

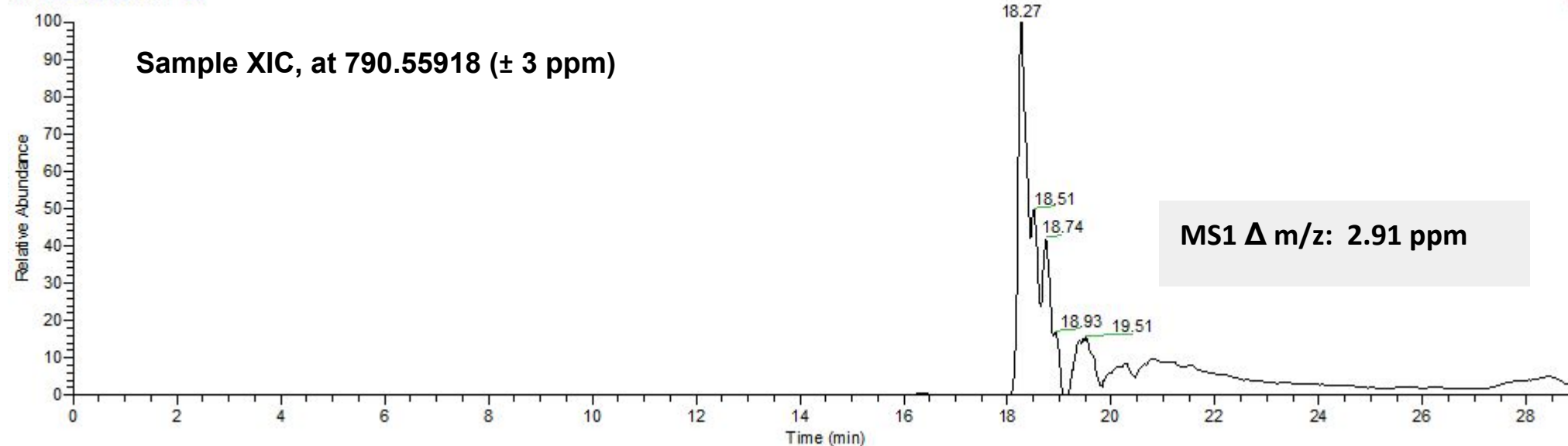
**No standard**

**Validated as** : PS(36:1) using LC-MS/MS method 'Positive'

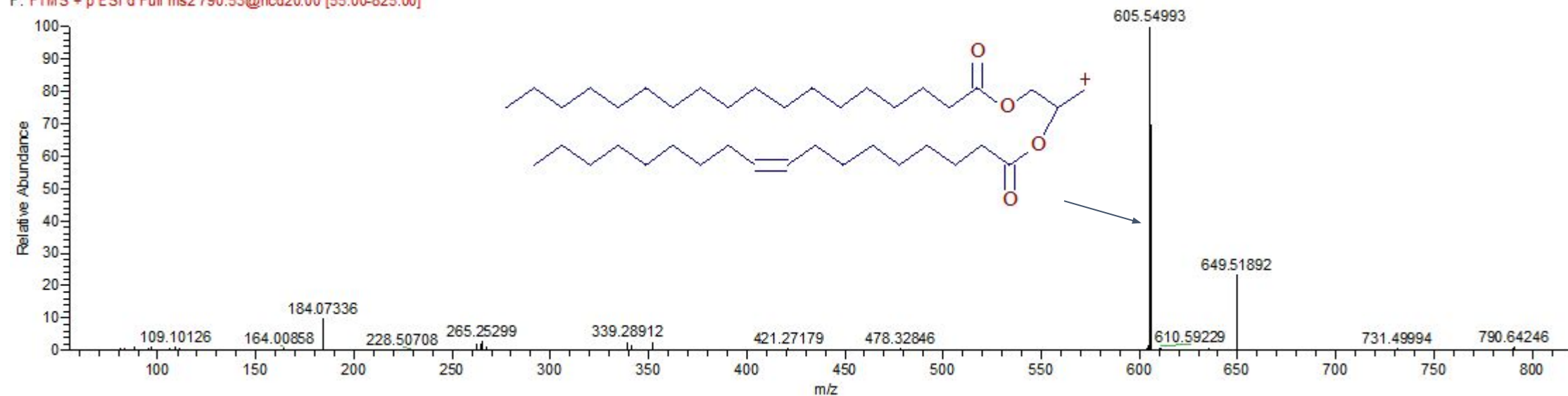


m/z=  
790.55523-790.56313  
F: ms MS  
Mouse\_Whole\_Brain\_3X  
dil\_CHCl3\_Lip\_ext\_Pos\_  
incl\_list\_CE20\_1

RT: 0.00 - 29.00 SM: 15B



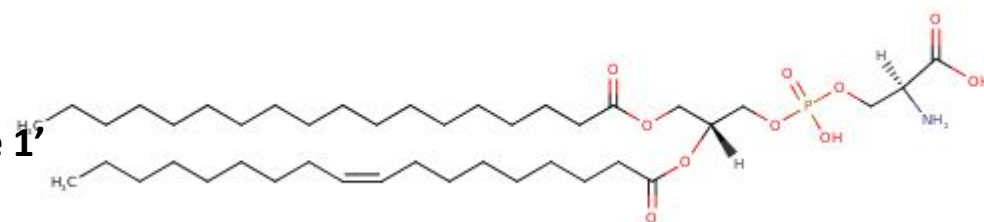
Mouse Whole Brain 3Xdil CHCl3 Lip ext Pos incl list CE20\_1 #9718 RT: 18.17 AV: 1 NL: 1.88E6  
F: FTMS + p ESI d Full ms2 790.53@hcd20.00 [55.00-825.00]



Molecular annotation: C<sub>42</sub>H<sub>80</sub>NO<sub>10</sub>P, -H, m/z 788.54353, PS(36:1)

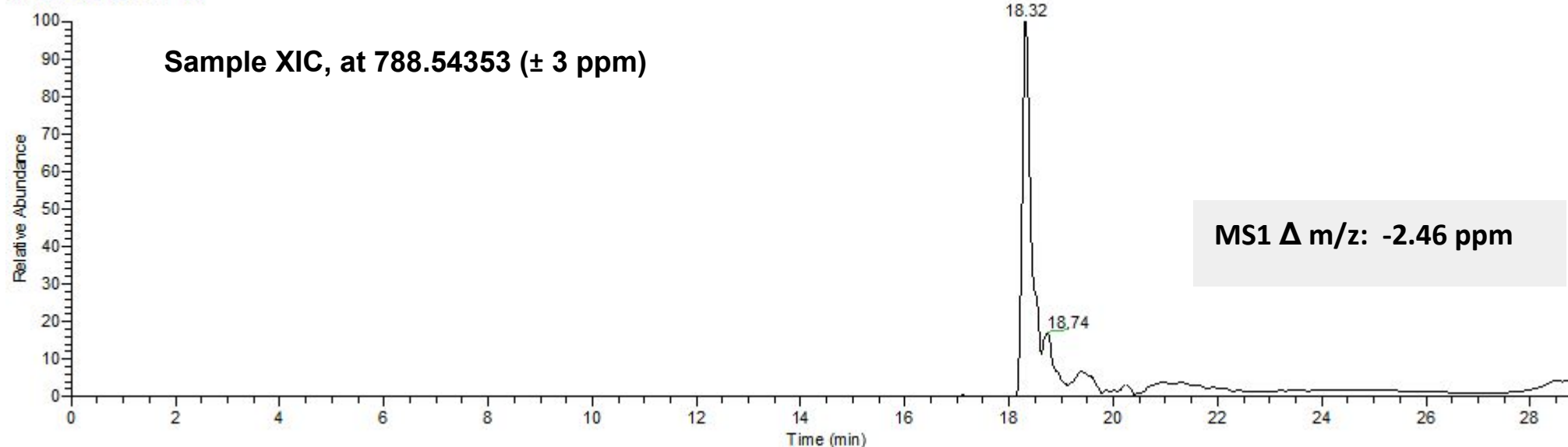
No standard

Validated as : PS(18:0\_18:1) [HMDB10163] using LC-MS/MS method 'Negative 1'

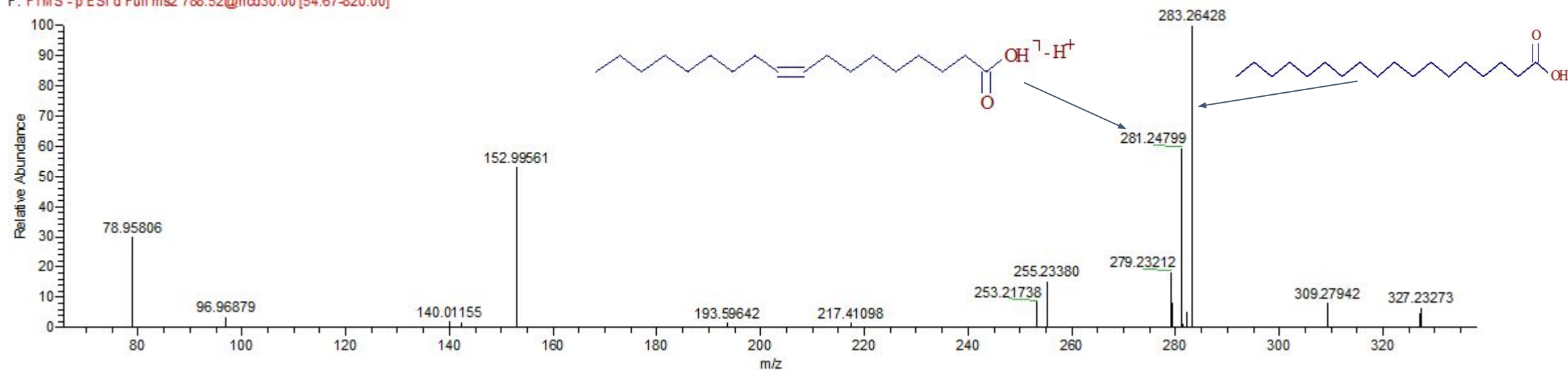


NL: 0.0427  
Base Peak m/z=  
788.53959-788.54747 F:  
ms MS  
Mouse\_Whole\_Brain\_CH  
Cl3\_Lip\_ext\_Neg\_incl\_lis  
t\_CE30

RT: 0.00 - 29.00 SM: 15B



Mouse Whole Brain CHCl<sub>3</sub> Lip ext Neg incl list CE30 #7222 RT: 18.19 AV: 1 NL: 1.75E5  
F: FTMS -p ESI d Full ms2 788.52@hcd30.00 [54.67-820.00]

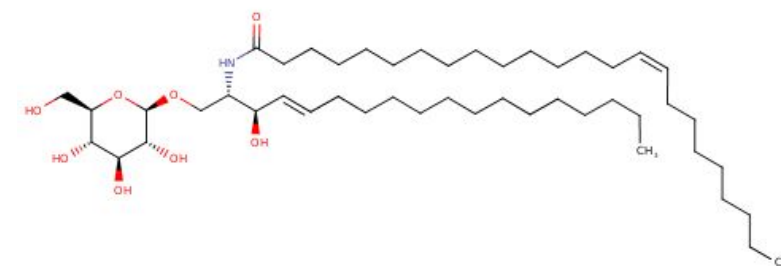


**Molecular annotation: C<sub>48</sub>H<sub>91</sub>NO<sub>8</sub>, +H, m/z 810.68166, Galactosylceramide(d42:2)**

**No standard**

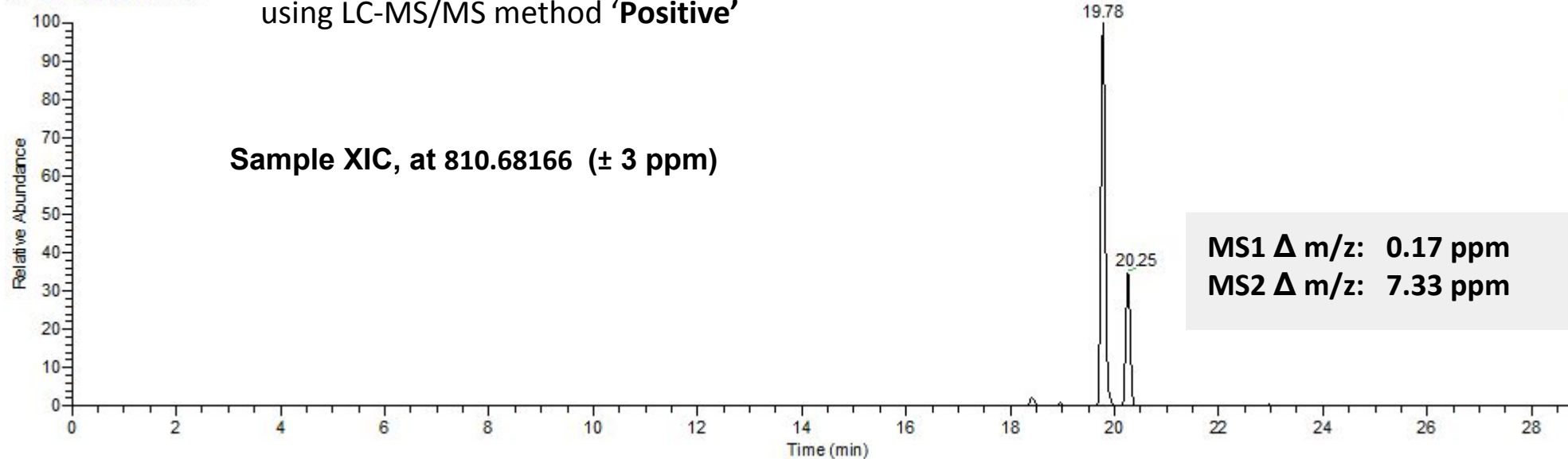
**Validated as**

**: Galactosylceramide(d18:1\_24:1) [HMDB10712] or  
Glucosylceramide(d18:1\_24:1) [HMDB04975]  
using LC-MS/MS method 'Positive'**

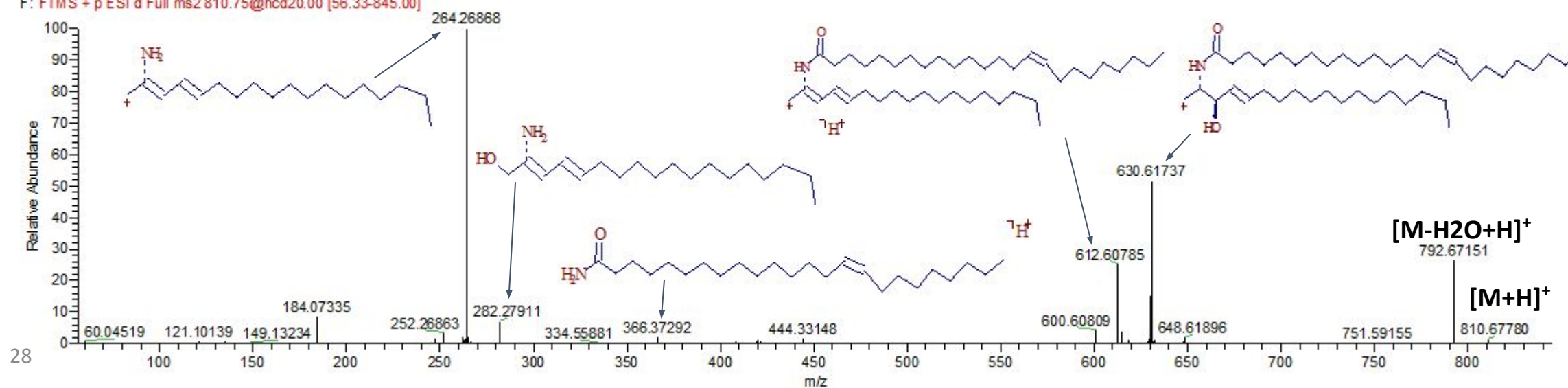


Mouse Whole Brain\_3Xdil  
CHCl3\_Lip\_ext\_Pos\_incl\_list  
\_CE30\_800\_900mz

RT: 0.00 -29.00 SM: 7B



Mouse Whole Brain\_3Xdil\_CHCl3\_Lip\_ext\_Pos\_incl\_list\_CE30\_800\_900mz#5063 RT: 19.70 AV: 1 NL: 1.54E7  
F: FTMS + p ESI d Full ms2 810.75@hcd20.00 [56.33-845.00]

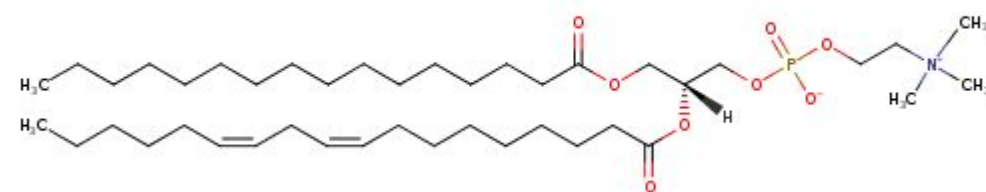




**Molecular annotation: C<sub>42</sub>H<sub>80</sub>NO<sub>8</sub>P, +H, m/z 758.56935, PC(34:2)**

**No standard**

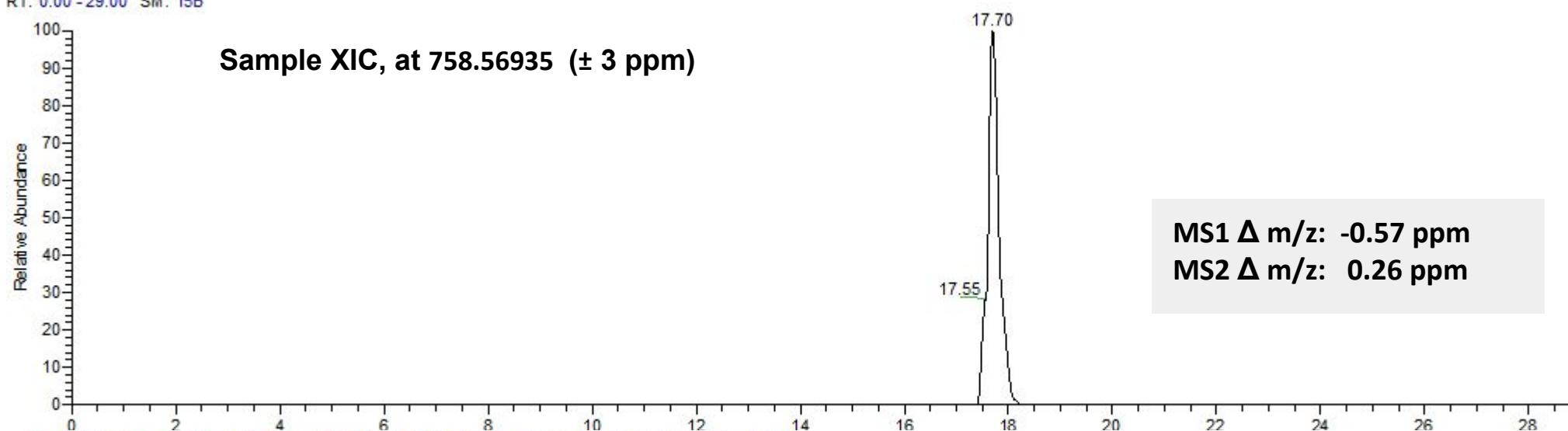
**Validated as : PC(34:2) using LC-MS/MS method 'Positive'**



758.56707-758.57163  
F: ms MS  
Mouse Whole\_Brain\_3X  
dil\_CHCl3\_Lip\_ext\_Pos\_  
incl\_list\_CE20\_1

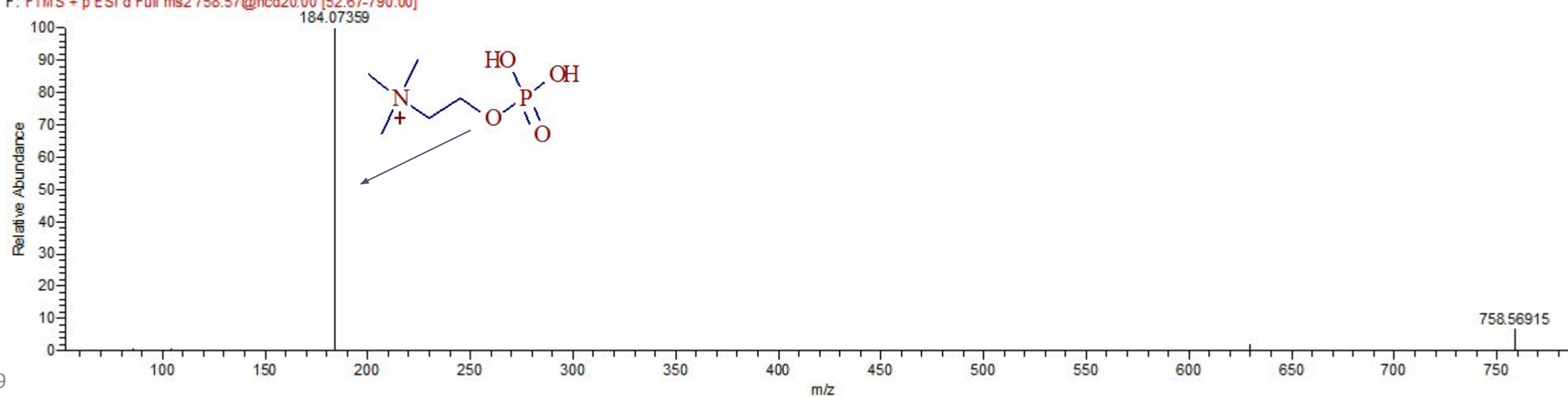
RT: 0.00 - 29.00 SM: 15B

**Sample XIC, at 758.56935 (± 3 ppm)**



**MS1 Δ m/z: -0.57 ppm**  
**MS2 Δ m/z: 0.26 ppm**

Mouse Whole Brain 3Xdil CHCl3 Lip ext Pos incl list CE20\_1 #9707 RT: 18.15 AV: 1 NL: 5.57E6  
F: FTMS + p ESI d Full ms2 758.57@hcd20.00 [52.67-790.00]

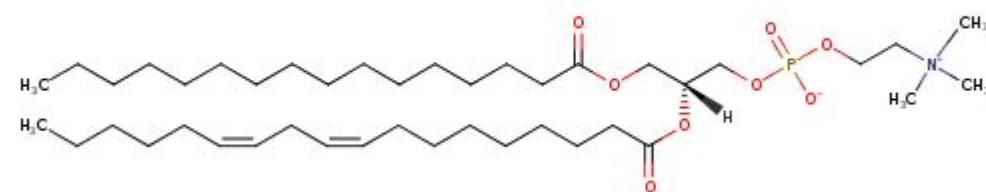


**Molecular annotation: C<sub>42</sub>H<sub>80</sub>NO<sub>8</sub>P, -H+FA, m/z 802.55973, PC(34:2)**

**No standard**

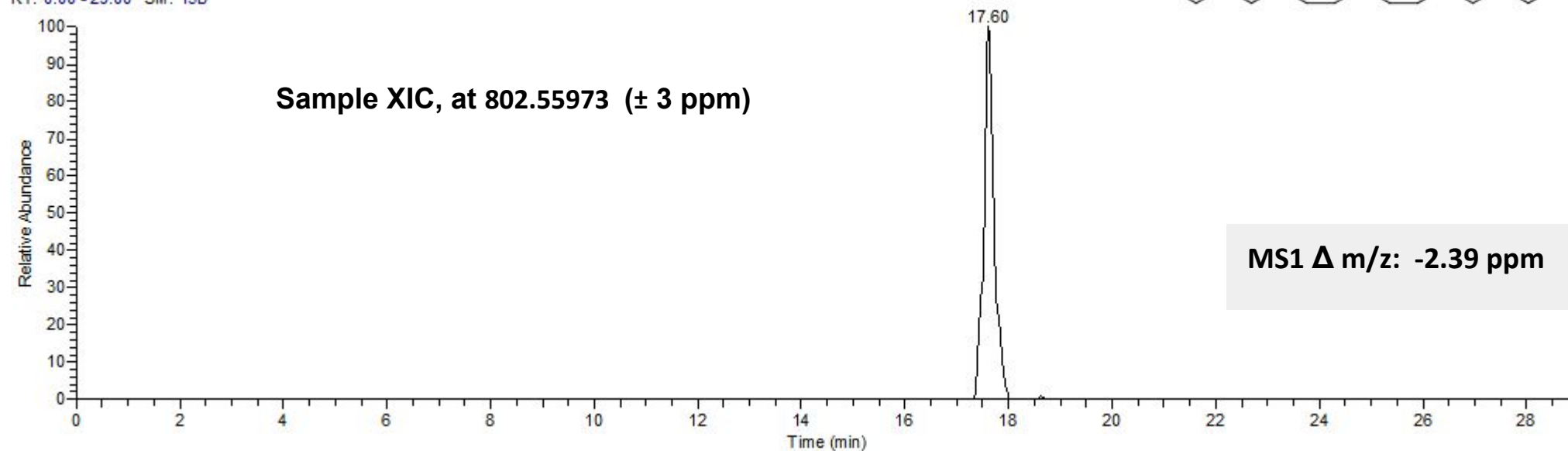
**Validated as**

**: PC(16:0\_18:2) [HMDB07973] using LC-MS/MS method 'Negative 2'**



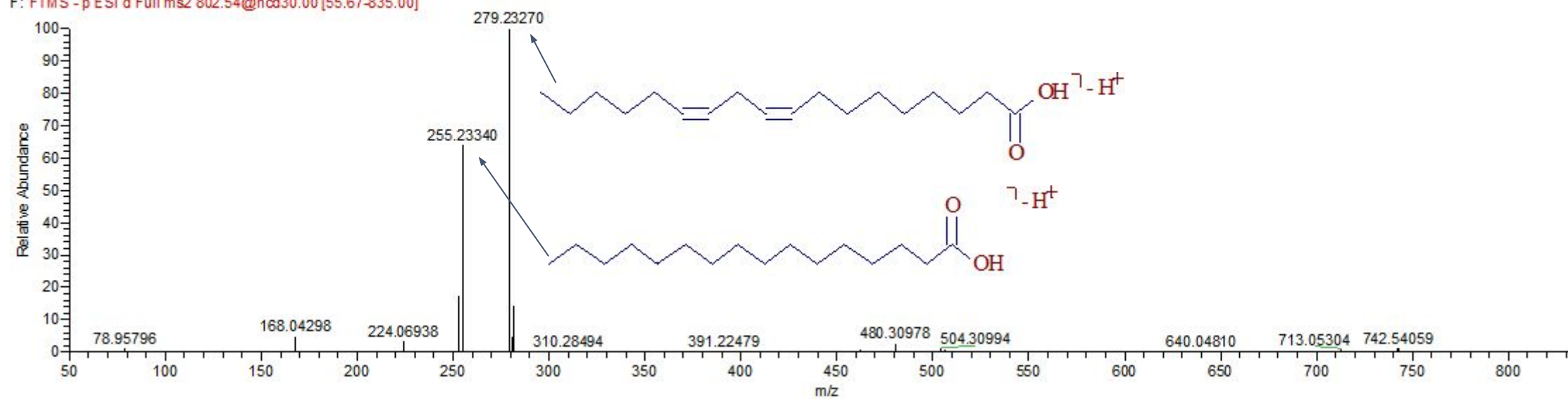
ms MS  
Mouse\_Whole\_Brain\_CH  
Cl3\_Lip\_ext\_Neg\_incl\_lis  
t\_CE30

RT: 0.00 - 29.00 SM: 15B



**MS1 Δ m/z: -2.39 ppm**

Mouse Whole Brain CHCl<sub>3</sub> Lip ext Neg incl list CE30 #7041 RT: 17.77 AV: 1 NL: 6.43E6  
F: FTMS - p ESI d Full ms2 802.54@hcd30.00 [55.67-835.00]



# Supplementary Data 2

## Analysis of rat brain datasets

### Methods

#### Imaging mass spectrometry

Two wild-type mouse brains (a1, a2) were obtained from Envigo RMS GmbH (Rossdorf, Germany). Care and handling of all animals complied with EU directive 2010/63/EU on the protection of animals used for scientific purposes. 12  $\mu\text{m}$  thick sections were collected on a cryomicrotome (Leica CM1950) and thaw-mounted onto indium tin oxide (ITO) coated glass slides (Bruker Daltonics, Bremen, Germany) and immediately desiccated. 2,5-dihydroxybenzoic acid (DHB) MALDI-matrix was applied by sublimation using a custom-made sublimation chamber.

For MALDI-MS measurements the prepared slides were mounted into a slide adapter (Bruker Daltonics) and loaded into the dual source of a 12T FTICR mass spectrometer solariX (Bruker Daltonics). The laser was running at 500 *Hz* and the ions were accumulated externally (hexapole) before being transferred into the ICR cell for a single scan. External calibration was carried out using arginine clusters in the electrospray mode.

A MALDI image from the animal a1 (dataset **RBa1s1**) was acquired with x-y-raster widths of 50  $\mu\text{m}$  using smartbeam™ II laser optics with the laser focus setting minimum ( $\approx 10 \mu\text{m}$ ). For a pixel a single scan was recorded using the ions generated by 30 laser shots. Every spectrum was internally calibrated by single point correction using matrix cluster of DHB [ $\text{C}_{14}\text{H}_9\text{O}_6$ ,  $m/z=273.039364$ ] if present. Data was acquired for the mass range  $150 < m/z < 3,000$  followed by a single zero filling and a sin-apodization.

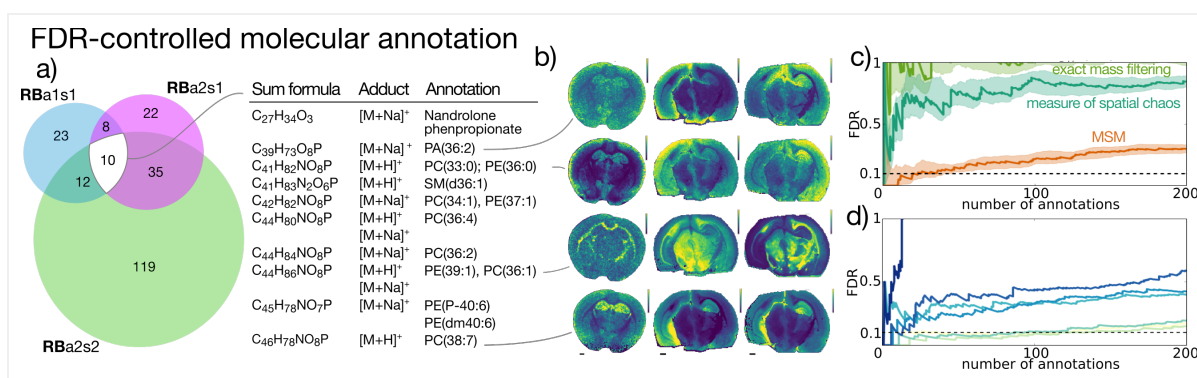
MALDI images from the animal a2 (datasets **RBa2s1** and **RBa2s2**) were acquired with x-y-raster widths of 100  $\mu\text{m}$  using smartbeam™ II laser optics with the laser focus setting medium ( $\approx 40 \mu\text{m}$ ). For a pixel, a single scan was recorded using the ions generated by 50 laser shots. Every spectrum was internally calibrated by single point correction using a matrix cluster of DHB [ $\text{C}_{14}\text{H}_9\text{O}_6$ ]<sup>+</sup> at  $m/z$  273.039364 and [Phosphatidylcholine+H]<sup>+</sup> at  $m/z$  760.585082. Data was acquired for the mass range  $200 < m/z < 3,000$  followed by a single zero filling and a sin-apodization.

Data was exported from SCiLS Lab software (SCiLS, Bremen) and in the imzML format. Centroid detection was performed and the centroided data was stored again as imzML. Ion images were generated with a window of  $\pm 4 \text{ ppm}$ . A hot-spot removal was performed for each image independently by setting the value of 1% highest-intensity pixels to the value of the 99'th percentile followed by an edge-preserving denoising using a median filter with the window 3x3.



# Results

## Summary



**Figure SD1.1.** Showcasing the FDR-controlled molecular annotation for three MALDI-FTICR imaging MS datasets from rat brain coronal sections (HMDB,  $FDR_{desired}=0.1$ ): a) Venn diagrams showing numbers of molecular annotations and a list of ten annotations from all three datasets, b) exemplary ion images of four annotations (see Table SD1.2 for detailed information about all ten annotations), as well as FDR curves illustrating: f) superiority of MSM as compared to individually considered exact mass filtering and measure of spatial chaos, and g) decrease of reliability of molecular annotation while simulating deteriorated mass accuracy and resolution by taking signals with a larger  $m/z$  tolerance.

## Properties of the collected datasets

Three high-mass resolution imaging MS datasets were collected from rat brains from two animals (a1 and a2). One section (RBa1s1) was collected from the first specimen (RBa1) and two consecutive sections were collected from the second specimen (RBa2). All three sections were prepared and imaged within the same day. These datasets can be considered as two biological and two technical replicates.

Table SD1.1 shows detailed properties of the datasets. Briefly, each of the collected spectra contained approximately two million  $m/z$  channels over the  $m/z$  range 100-3000. For the considered mass resolving power (FWHM for a peak at  $m/z$  760.5543 was 0.0038), the amount of potential peaks (the peak capacity of each spectrum) was 470000. Across the datasets a typical peak capacity filling of 2% was observed.

## Data availability

The imaging mass spectrometry data is publicly available at the MetaboLights repository under the accession numbers MTBLS313.

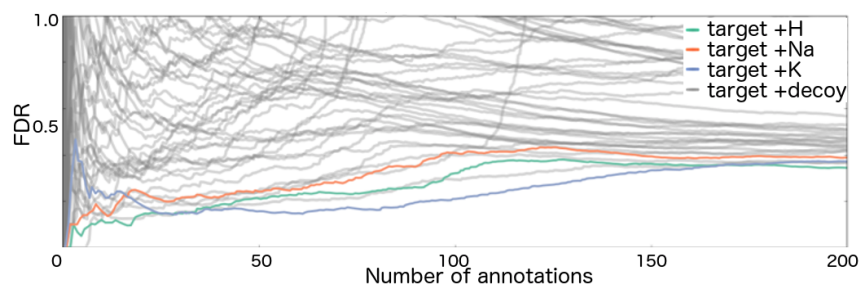
Property	Dataset		
	<i>RBa1s1</i>	<i>RBa2s1</i>	<i>RBa2s2</i>
Specimen type	Rat brain	Rat brain	Rat brain
Section direction	Coronal	Coronal	Coronal
Pixel size ( $\mu\text{m}$ )	50	100	100
Number of pixels	44674	10727	10916
Number of m/z bins	2019850	1955436	1955436
Number of peaks per spectrum	7200 +/- 700	11000 +/- 800	11000 +/- 800
Raw data size [compressed]* (GB)	340 [40]	80 [10]	80 [10]
Collected m/z range <sup>^</sup>	150-3000	200-3000	200-3000
Mass resolving power (at m/z 700)	70000	80000	80000

**Table SD1.1.** Technical summary of the collected high-resolution imaging MS datasets. \*Raw data stored as 32-bit floats; compressed data stored in the proprietary .h5 (SCiLS Lab 2014b) format. <sup>^</sup>The m/z range of 100-1000 was included in analysis.

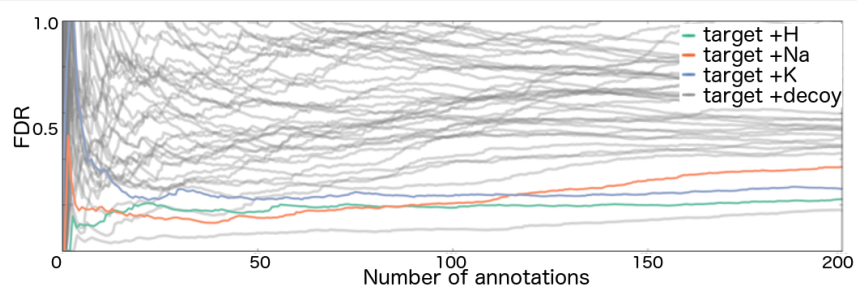
## Negative control experiment

As the last step of the evaluation of the proposed FDR approach, we performed a negative control experiment when using an implausible adduct as the target. For this, instead either of the common adducts (+H, +Na, and +K), we considered either of the implausible adducts as the target. The list of possible decoy adducts was reduced by taking out the considered implausible adduct.

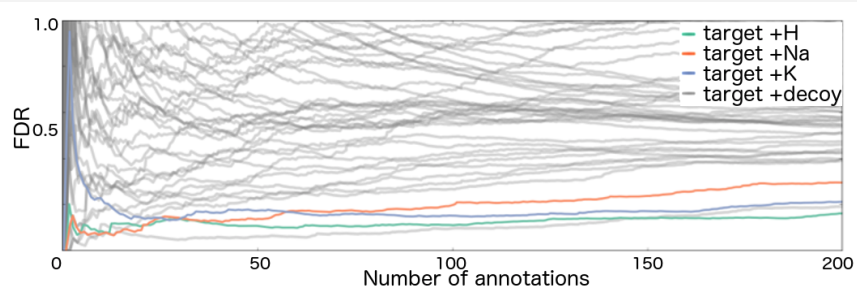
Figure SD1.2 shows the obtained negative-control FDR curves (in grey) as well as the FDR curves for the common adducts +H, +Na, and +K (in green, orange, blue). Most of the negative-control curves are higher that confirms the soundness of our FDR approach (implausible adducts should not be annotated). For better understanding of the difference between the individual negative-control FDR curves, we calculated an Area Under Curve value for the first 500 annotations ( $AUC_{500}$ ). As expected,  $AUC_{500}$  values (Figur SD1.3) for most of the implausible adducts are higher than for plausible ones. A few implausible adducts (+Be, +F) consistently have  $AUC_{500}$  comparable to plausible adducts. Although we looked into this, so far we have no definite explanation.



a) RBA1s1



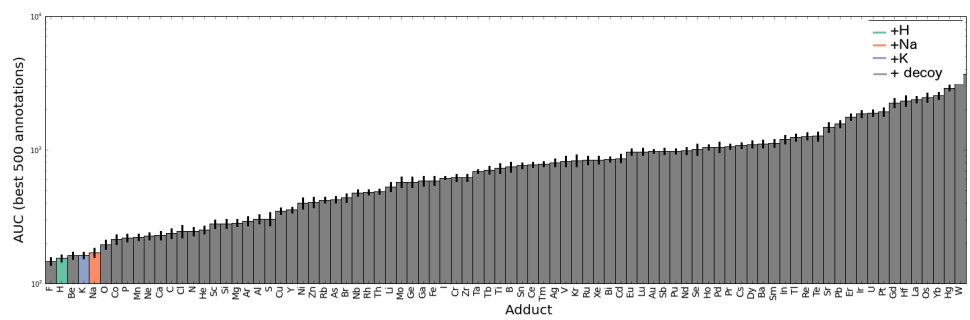
b) RBA2s1



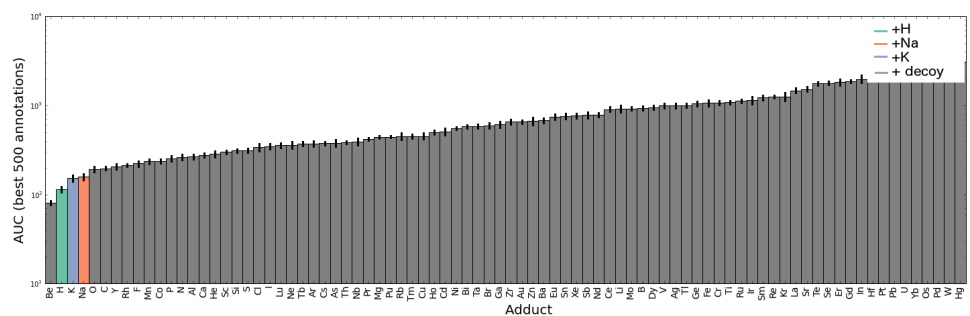
c) RBA2s2

**Figure SD1.2.** FDR curves for the negative-control experiment when using an implausible adduct as the target (in grey, one curve per an implausible adduct). For comparison, the FDR curves for the common adducts (+H, +Na, +K) are shown. The curves are shown for only 200 first molecular annotations to better see the region of interest.

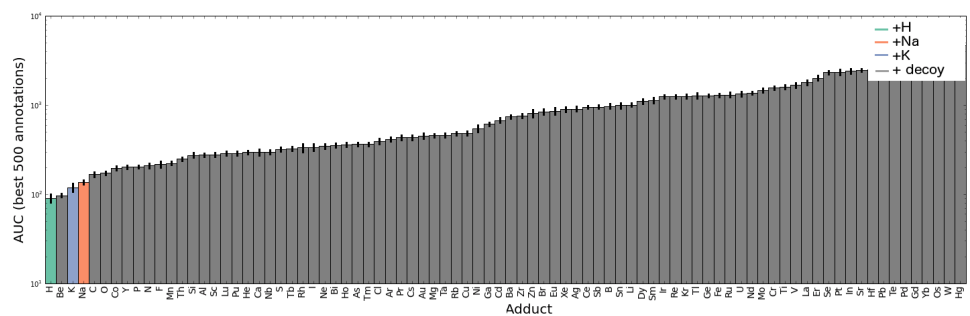




**a) RBa1s1**



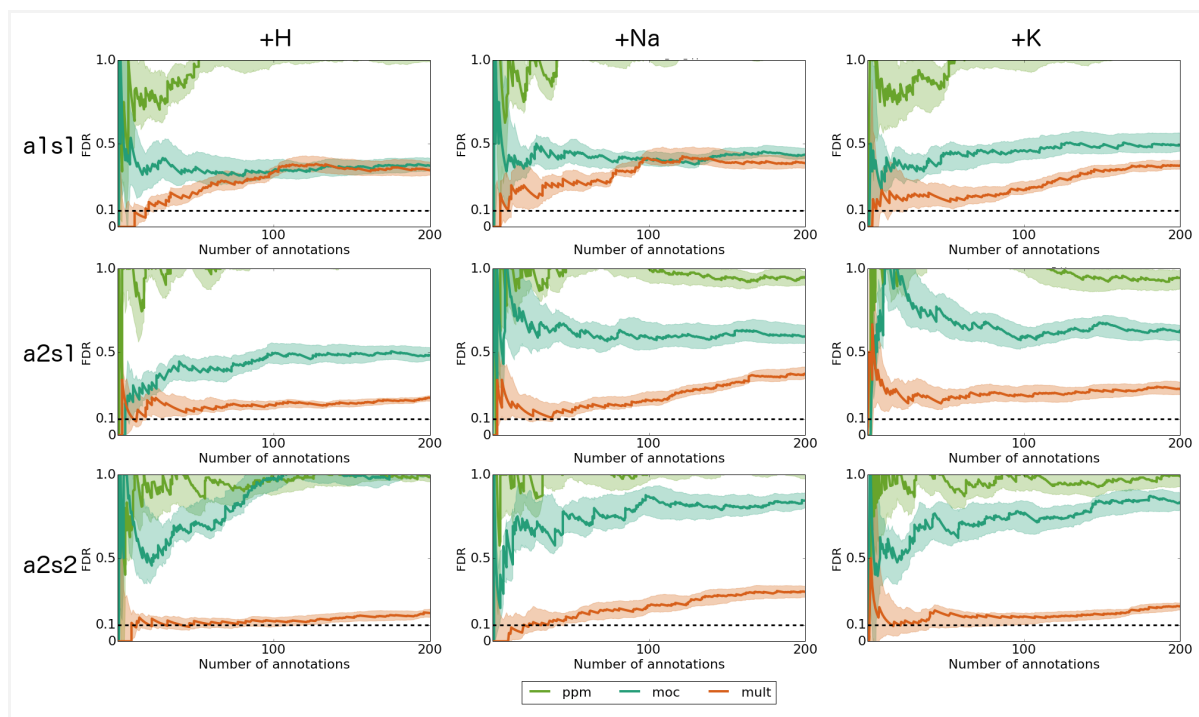
**b) RBa2s1**



**c) RBa2s2**

**Figure SD1.3.** Summary of the negative control experiment, showing an Area Under Curve value (for the first 500 annotations) for each FDR curve from Figure SD1.2.

## Comparing alternative scoring measures

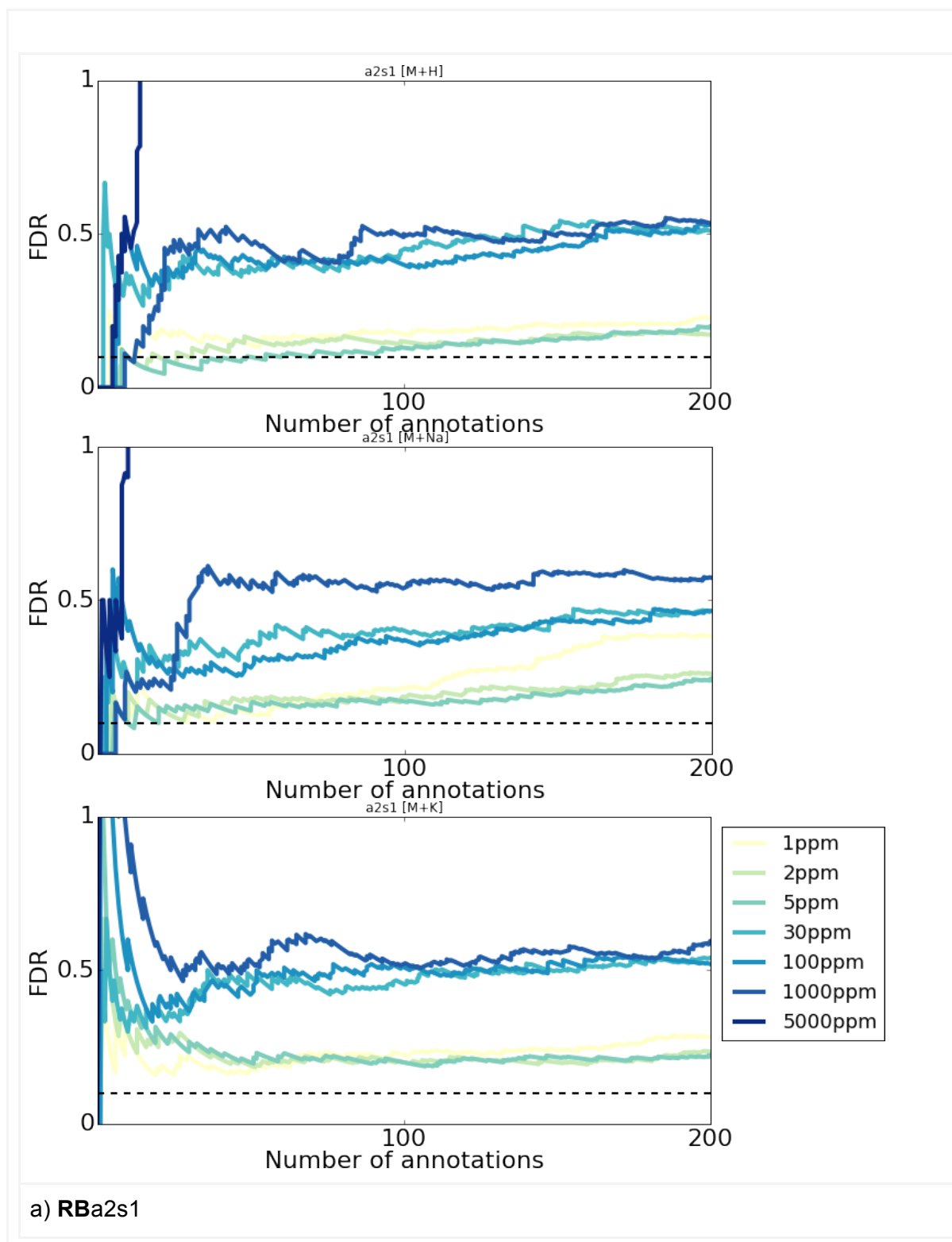


**Figure SD1.4.** FDR-curves when using different scoring measures: exact mass matching (light green), measure of chaos (dark green) and MSM (orange). The FDR-curves show that the MSM measure outperforms other measures (has lowest FDR-curve).

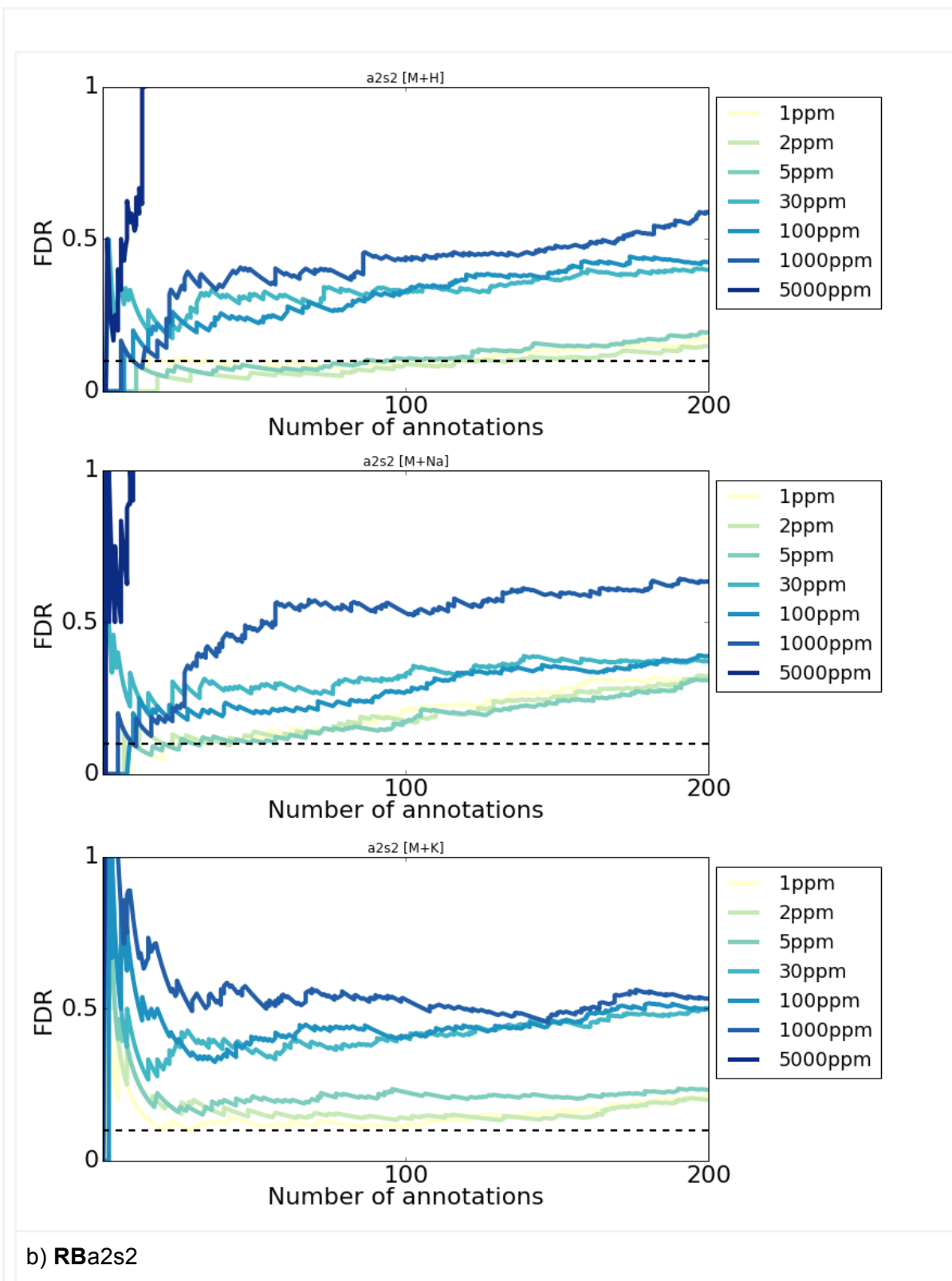
FDR-curves plot the FDR against the number of annotations to illustrate how many annotations would be made at a fixed FDR, or what the FDR at a particular MSM threshold is. Figure SD1.4 shows the FDR-curves when using the MSM score (proposed annotation approach) as well as when using either exact mass matching or measure of chaos along as scoring measures instead of MSM. One can see the using MSM outperforms other measures (providing the lowest FDR-curve). They also help quantify how unspecific exact mass filtering is (with the FDR exceeding 0.75 after 10 annotations) and highlight the need for an advanced score like our proposed MSM.

## Simulating different mass resolving power

The effect of decreasing the resolving power was estimated by changing the  $m/z$ -tolerances used when sampling the ion signal from the imaging MS dataset. The FDR curves for each of the technical replicate datasets (**RBa2s1**, **RBa2s2**) for each of the considered  $m/z$ -tolerances are shown in Figure SD1.5.

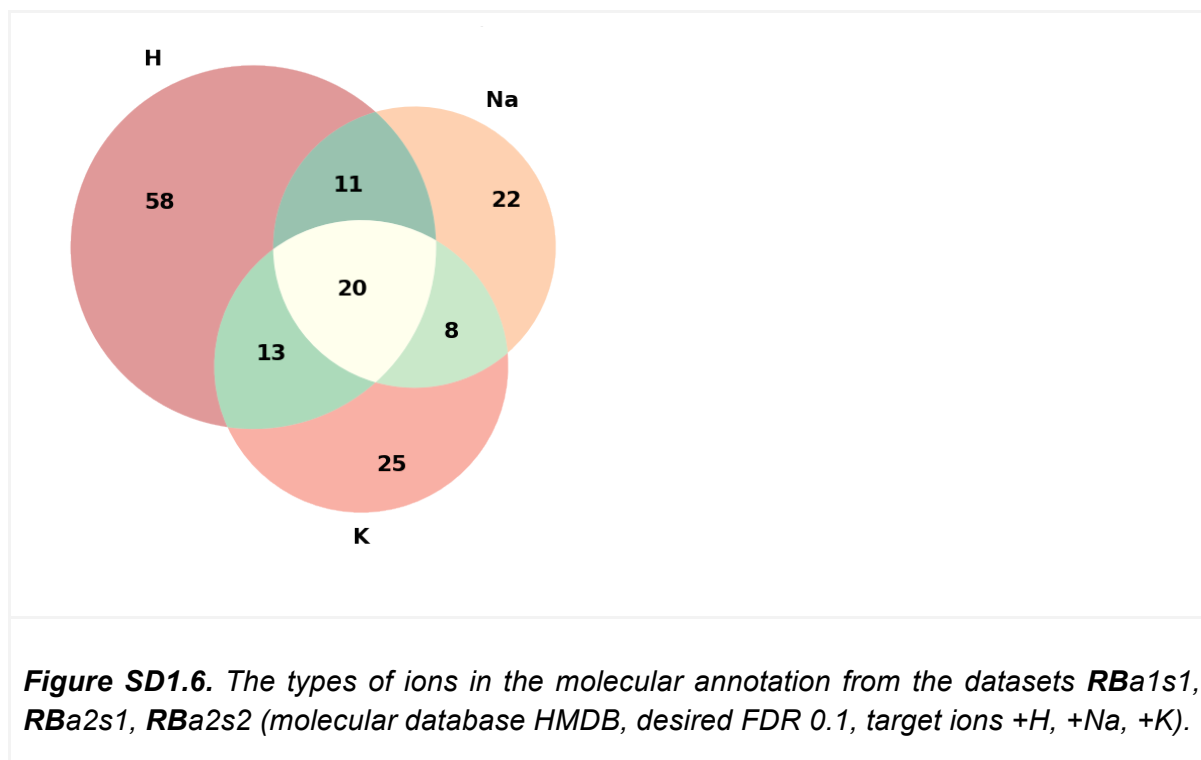




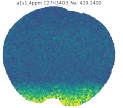
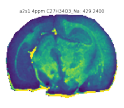
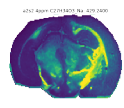
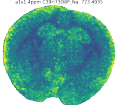
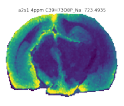
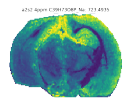
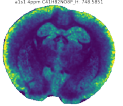
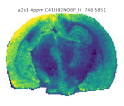
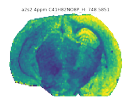
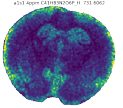
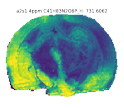
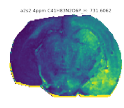


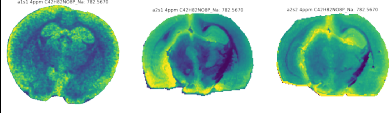
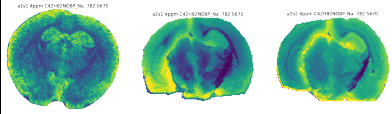
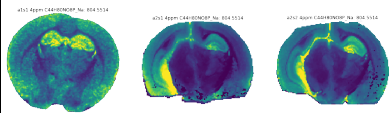
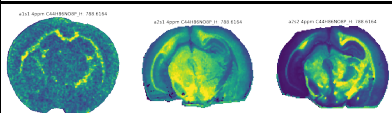
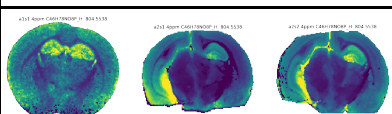
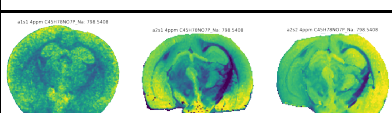
**Figure SD1.5.** FDR-curves when simulating the change of the resolving power by increasing the  $m/z$ -tolerance used to sample the ion images, for the datasets **RBa2s1** and **RBa2s2**, for the ion adducts +H, +Na, +K.

## Types of ion adducts in the molecular annotations



## Detailed information on the ions annotated in all three rat brain datasets

Sum formula	Adduct	Principal <i>m/z</i>	Name	MSM-score value			Ion images		
				RBa1s1	RBa2s1	RBa2s2	RBa1s1	RBa2s1	RBa2s2
C27H34O3	[M+Na] <sup>+</sup>	429.2400	Nandrolone phenpropionate	0.69	0.63	0.44			
C39H73O8P	[M+Na] <sup>+</sup>	723.4935	PA(36:2)	0.64	0.41	0.52			
C41H82NO8P	[M+H] <sup>+</sup>	748.5851	PC(33:0) PE(36:0)	0.68	0.59	0.32			
C41H83N2O6P	[M+H] <sup>+</sup>	731.6062	SM(d36:1)	0.58	0.46	0.56			

C42H82NO8P	[M+Na] <sup>+</sup>	782.5670	PC(34:1) PE(37:1)	0.57 0.49 0.63	
C44H80NO8P	[M+H] <sup>+</sup>	782.5694	PC(36:4)	0.48 0.55 0.30	
	[M+Na] <sup>+</sup>	804.5514	PC(36:4)	0.55 0.51 0.67	
C44H86NO8P	[M+H] <sup>+</sup>	788.6164	PE(39:1) PC(36:1)	0.58 0.80 0.54	
	[M+Na] <sup>+</sup>	810.5983	PE(39:1) PC(36:1)	0.47 0.63 0.58	
C45H78NO7P	[M+Na] <sup>+</sup>	798.5408	PE(P-40:6) PE(dm40:6)	0.60 0.38 0.67	

**Table SD1.2.** Detailed information about the molecular annotations in all three rat brain datasets (brief information was shown in Figure SD1.1).