Electronic Supplementary Material (ESI) for Analyst.

Utility of neurological smears for intrasurgical brain cancer diagnostics and tumour cell percentage by DESI-MS

Valentina Pirro[1], Alan K. Jarmusch[2], Clint M. Alfaro[1], Eyas M. Hattab[2,*], Aaron A. Cohen-Gadol[3,*], R. Graham Cooks[1,*]

[1] Department of Chemistry, Purdue University, West Lafayette, IN, USA

[2] Department of Pathology and Laboratory Medicine, University of Louisville, Louisville, KY, USA

[3] Department of Neurological Surgery, Indiana University School of Medicine, Indianapolis, IN, USA

***Corresponding authors**

R Graham Cooks, cooks@purdue.edu

Eyas M. Hattab, eyas.hattab@louisville.edu

Aaron A. Cohen-Gadol, acohenmd@gmail.com

Supplementary Information

Supplementary Materials and Methods

Specimens

Cryopreserved human neurological specimens were obtained from 29 patients from the Biorepository of the Methodist Research Institute (Purdue IRB # 1410015344). Additional details are reported in Table S1. Tissue smears were prepared from the specimens by removing a small portion of frozen tissue (about 10-50 mm^3) placing it onto a glass microscope slide allowing the tissue to come to room temperature (approximately 30 s), and then smearing it using a custom 3D printed device (Video S1). The 3D printed device was designed to smear the tissue along the longest dimension of the glass slide, and confine it to the center area with an appropriate thickness (~100 μm) in order to avoid excessive dilution of the cellular material, which would negatively affect the MS signal intensity. Indeed, the absolute ion abundance in the mass spectra proved directly related to the quantity and cellularity of the tissue smeared; however, the mass spectral profiles (i.e. relative ion intensities used as chemical signature of the tissue) are grossly uninfluenced by differences in smearing [1]. Confinement of the material to a consistent area of the slide facilitated also the standardization of the moving stage and data acquisition programs. All tissue smears were stored at −80 °C prior to DESI-MS analysis. Smears were H&E stained subsequently to DESI-MS analysis for blind pathological evaluation. The H&E protocol is described in Jarmusch et al. [1] To increase confidence in histopathology, tissue smears and adjacent tissue sections were evaluated always in parallel.

DESI-MS Imaging

DESI-MS imaging of tissue smears were performed using a linear ion trap mass spectrometer (Finnigan LTQ, Thermo Scientific) coupled with a custom DESI source (similar to Prosolia, Inc. commercial 2D source) equipped with a precision 2D moving stage. DESI-MS was performed using DMF – ACN (1:1 v/v), which preserves tissue morphology for subsequent histopathology. Tissue smears were analyzed twice by DESI-MS imaging in the negative ion mode with automatic gain control (AGC) disabled. The first DESI-MS acquisition (which we will refer to as the lipid profile) included data over *m/z* 700 - 1000. The second acquisition (the metabolite profile) included data over *m/z* 80 - 200. Details are reported in Table S2. The moving stage was reset to the X,Y origin in between the lipid and metabolite profile acquisitions in order to image the same area twice. Images were collected as rows by coordinating linear motion of the moving stage with MS acquisition rate, defining resolution of 250 μm in both X

and Y axes (Figure S1). An area of 60 mm^2 (12 x 5 mm, 47 x 20 pixels along the X-axis and Y-axis respectively, 250µm pixel size in X and Y) was imaged for each smear.

Chemical Diagnosis

Data Handling

The MATLAB software (The MathWorks, Inc. Natick, USA) was used to process all data unless otherwise specified. MS data were exported from XCalibur 2.0 and converted into .mzXML files using MSConvertGUI (http://proteowizard.sourceforge.net/index.shtml), and then imported into MATLAB using the mzxmlread.mat function. In-house MATLAB routines were used to process the data unless otherwise specified. DESI-MS images were reconstructed as datacubes that are composites of the spatial domain (X and Y) and the spectral domain (*m/z* values and corresponding ion intensities). The *m/z* domain was comprised of 3,600 datapoints for the lipid profile (*m/z* 700-1000) and 1,400 for the metabolite profile (*m/z* 80-200), both using an acquisition step of *m/z* 0.0833. DESI-MS images of tissue smears were used to predict for their tissue state (e.g. glioma) using statistical methods for recognition of mass spectral patterns. Pixels related to the background of the glass slide, where no tissue was smeared, were excluded a priori based on a threshold of signal intensity. Absolute intensities of ions *m/z* 788, 794, 834, 885, and 888 – glycerophospholipids selectively present in the brain tissue but absent in the glass slide – were summed and then subtracted to the ion intensities of *m/z* 349 and 363 detected from the background, to define the threshold. Only pixels with mass spectra having the summed ion intensity greater than the threshold were used for chemical predictions. The selected pixels were plotted in red in the 2D spatial domain to double-check the spatial distribution compared to H&E stains (Figure S1). The same pixels were excluded from the lipid and the metabolite profile images. The AGC was deactivated for the acquisition of DESI-MS images to synchronize the duty cycle of the instrument with the position of the moving stage.

PCA-LDA Prediction of Brain Tumors

Tissue smears were chemically diagnosed as grey matter, white matter or glioma using a linear discriminant analysis (LDA) with data compression by principal component analysis (PCA). The mass spectral patterns detected for the tissue smears were tested against a reference library of spectra that was built using histologically-defined

tissue sections analyzed by DESI-MS imaging. Note that discriminant classification techniques such as LDA normally attribute unknown samples uniquely to one of the categories being modelled with a rigid delineation between categories. DESI-MS images of tissue smears were predicted pixel by pixel, with prior exclusion of the background pixels (i.e. glass slide) where no tissue was smeared. Predictions were shown in the corresponding 2D DESI spatial domain using false-colors (Figure S1). Images serve to provide average values and estimate variance due to any incomplete homogenization of the tissue. Note that pathology evaluation is given for the overall tissue smear consistent with prior tissue mixing. All predictions were validated against pathology done on the same smears subsequently to DESI-MS analysis.

Estimation of Tumor Cell Percentage

In parallel with PCA-LDA predictions of tissue state, the ion intensity of NAA (*m/z* 174) - normalized to the total ion count (TIC) - was used to estimate the percentage of tumor cells via linear regression. The TCP predictions based solely on NAA's intensity stand in contrast to the chemical predictions based on both lipid and metabolite signatures in evaluation of their respective contributions to chemical diagnosis. The linear regression model used to predict TCP in tissue smears is described in Jarmusch et al. [1] TCP was estimated pixel by pixel in the DESI-MS smear images. Values of 0 and 100% were assigned to all samples whom predictions exceeded those values (< 0% and > 100%). Predictions were color-coded and shown in the 2D DESI spatial domain (Figure S1). Images of tissue smears served the purpose to estimate variance of TCP prediction, with averages compared to the pathological evaluation.

Univariate Statistics

One-sample t-test was used to compare the average prediction for the DESI-MS images to the 'hypothesized' value provided by histopathology. One-way ANOVA was used to test if population means were equal. The median test was used to test if population medians were equal. The null hypothesis H_0 affirms that there are no significant differences. A significance level (a two-tailed p-value) of 0.05 (CI = 95%) was chosen for the statistical test. When the experimental p-value is lower than the critical p-value, the hypothesis H_0 is rejected. The statistical tests were performed in OriginPro 2015 using built-in statistics package.

Table S1. Human neurological specimen information

	Pathological Evaluation on Tissue Sections Adjacent to the Smeared Tissue				DESI-MS Smear Image
Specimen	**Diagnosis**	**Grade**	**Density**	**TCP**	**# Pixels**
1	Normal Grey matter Normal White Matter	- -	- -	< 5% < 5%	704
2	Glioma	High	High	100*	453
3	Normal Grey Matter Normal White Matter	- -	- -	< 5% < 5%	597
4	Glioma	High	high	100*	531
5	Normal Grey Matter	-	-	< 5%	397
6	Normal Grey Matter	-	-	< 5%	474
7	Normal Grey Matter Normal White Matter Glioma	- - High	- - High	< 5% < 5% 100*	584
8	Normal Grey Matter	-	-	< 5%	460
9	Glioma	High	High	100*	654
10	Infiltrated Grey Matter Infiltrated White Matter	- -	- -	10% 5%	646
11	Infiltrated Grey Matter	-	-	20%	283
12	Infiltrated White Matter Infiltrated Grey Matter	- -	- -	20% 25%	523
13	Glioma	High	High	100*	365
14	Infiltrated Grey Matter Infiltrated White Matter	- -	- -	20% 10%	554
15	Infiltrated Grey Matter Infiltrated White Matter	- -	- -	15% 15%	681
16	Normal Grey Matter Normal White Matter	- -	- -	< 5% < 5%	775
17	Infiltrated Grey Matter Infiltrated white Matter	- -	- -	10 10	673
18	Normal Grey Matter Glioma	- High	- High	< 5% 100*	571
19	Infiltrated White Matter Infiltrated Grey Matter	- -	- -	20% 20%	443
20	Infiltrated Grey Matter Infiltrated White Matter	- -	- -	15% 15%	663
21	Infiltrated Grey Matter	-	-	10%	609
22	Infiltrated Grey Matter	-	-	10%	638
23	Normal Grey Matter Infiltrated Grey Matter Normal White Matter	- - -	- - -	< 5% 10% < 5%	468
24	Glioma	High	Medium	66.7*	608

25	Glioma	High	High	100*	455
26	Infiltrated Grey Matter Normal White Matter	- -	- -	10 % < 5%	369
27	Infiltrated Grey Matter Infiltrated White Matter	- -	- -	40% 40%	759
28	Infiltrated Grey Matter Glioma	- High	- Medium	15% 66.7*	776
29	Infiltrated Grey Matter	-	-	10%	793
30	Glioma	High	Medium	66.7*	Excluded**
31	Normal Grey Matter Glioma	- Low	- Medium	< 5% 66.7*	Excluded**

*Values arbitrarily assigned to categorize medium and high density

**Excluded because of low signal.

Table S2. Instrumental parameters for DESI-MS imaging

Full Scan MS	**Lipid Profile *m/z* 700-1000**	**Metabolite Profile *m/z* 80-200**
Solvent flow rate	1.0 µL/min	
Pressure of nitrogen gas	160 PSI	
Applied high voltage	−5 kV	
Incident angle	52°	
Spray-to-surface distance	2-3 mm	
Spray-to-MS distance	5-7 mm	
Capillary voltage	−9 V	−30 V
Tube lenses potential	−150 V	−30 V
Tuned mass	*m/z* 786	*m/z* 174
Injection time	350 ms	125 ms
Microscans	2	4
Speed of the moving stage	294.117 µm/s	403.225 µm/s
Pixel size	250 µm	

Table S3. DESI-MS and pathological evaluations for the tissue smears

	Chemical Diagnosis					Pathological Evaluation			
Specimens	**PCA-LDA Diagnosis**	**# of Pixels**	**% of Pixels**	**Average TCP (%)**	**Standard Deviation in TCP (%)**	**Diagnosis**	**Grade**	**Density**	**TCP (%)**
1	GM	0	0	9	12	WM	*n/a*	*n/a*	< 5
	WM	703	99.9						
	GL	0	0						
	M	0	0						
	P	1	0						
2	GM	17	3.8	33	18	GL	High	High	100*
	WM	39	8.6						
	GL	395	87.2						
	M	0	0.0						
	P	2	0.4						
3	GM	504	84.4	28	20	GM	*n/a*	*n/a*	< 5
	WM	9	1.5						
	GL	84	14.1						
	M	0	0.0						
	P	0	0.0						
4	GM	0	0.0	99	3	GL	High	Medium	66.7*
	WM	1	0.2						
	GL	530	99.8						
	M	0	0.0						
	P	0	0.0						
5	GM	370	93.2	18	14	GM	*n/a*	*n/a*	< 5
	WM	0	0.0						
	GL	27	6.8						
	M	0	0.0						
	P	0	0.0						
6	GM	384	81.0	11	8	GM	*n/a*	*n/a*	< 5
	WM	80	16.9						
	GL	10	2.1						
	M	0	0.0						
	P	0	0.0						
7	GM	291	49.8	13	10	GM	*n/a*	*n/a*	< 5
	WM	73	12.5						
	GL	220	37.7						
	M	0	0.0						
	P	0	0.0						
8	GM	322	69.8	22	10	GM	*n/a*	*n/a*	< 5
	WM	0	0.0						
	GL	138	29.9						
	M	0	0.0						
	P	1	0.2						

9	GM	0	0.0	96	6	GL	High	High	100*
	WM	0	0.0						
	GL	654	100.0						
	M	0	0.0						
	P	0	0.0						
10	GM	194	30.0	19	11	GM	*n/a*	*n/a*	20
	WM	218	33.8						
	GL	234	36.2						
	M	0	0.0						
	P	0	0.0						
11	GM	19	6.7	28	12	GM	*n/a*	*n/a*	20
	WM	0	0.0						
	GL	264	93.3						
	M	0	0.0						
	P	0	0.0						
12	GM	523	100.0	2	6	GM	*n/a*	*n/a*	25
	WM	0	0.0						
	GL	0	0.0						
	M	0	0.0						
	P	0	0.0						
13	GM	0	0.0	98	5	GL	High	High	100*
	WM	0	0.0						
	GL	365	100.0						
	M	0	0.0						
	P	0	0.0						
14	GM	0	0.0	28	14	*n/a*	*n/a*	*n/a*	10
	WM	546	98.6						
	GL	8	1.4						
	M	0	0.0						
	P	0	0.0						
15	GM	648	95.2	6	8	GM	*n/a*	*n/a*	10
	WM	0	0.0						
	GL	33	4.8						
	M	0	0.0						
	P	0	0.0						
16	GM	129	16.6	39	14	GM	*n/a*	*n/a*	10
	WM	0	0.0						
	GL	646	83.4						
	M	0	0.0						
	P	0	0.0						
17	GM	225	33.4	50	16	GM	*n/a*	*n/a*	10
	WM	0	0.0						
	GL	448	66.6						
	M	0	0.0						
	P	0	0.0						

18	GM	178	31.2	31	24	GL	High	Medium	66.7*
	WM	0	0.0						
	GL	393	68.8						
	M	0	0.0						
	P	0	0.0						
19	GM	68	15.3	63	13	GM	*n/a*	*n/a*	10
	WM	0	0.0						
	GL	375	84.7						
	M	0	0.0						
	P	0	0.0						
20	GM	565	85.2	15	12	GM	*n/a*	*n/a*	10
	WM	0	0.0						
	GL	98	14.8						
	M	0	0.0						
	P	0	0.0						
21	GM	455	74.5	15	15	GM	*n/a*	*n/a*	10
	WM	2	0.3						
	GL	154	25.2						
	M	0	0.0						
	P	0	0.0						
22	GM	234	36.7	20	14	GM	*n/a*	*n/a*	10
	WM	0	0.0						
	GL	404	63.3						
	M	0	0.0						
	P	0	0.0						
23	GM	385	82.3	3	8	GM	*n/a*	*n/a*	10
	WM	3	0.6						
	GL	80	17.1						
	M	0	0.0						
	P	0	0.0						
24	GM	1	0.2	87	15	GL	High	Low	*n/a*
	WM	2	0.3						
	GL	589	96.1						
	M	19	3.1						
	P	2	0.3						
25	GM	0	0.0	99	3	GL	High	Medium	66.7*
	WM	0	0.0						
	GL	18	4.0						
	M	9	2.0						
	P	428	94.1						
26	GM	0	0.0	56	14	GL	High	Medium	66.7*
	WM	366	99.2						
	GL	3	0.8						
	M	0	0.0						
	P	0	0.0						

27	GM	131	17.3	46	15	GL	High	Medium	66.7*
	WM	0	0.0						
	GL	628	82.7						
	M	0	0.0						
	P	0	0.0						
28	GM	5	0.6	69	11	GL	High	Medium	66.7*
	WM	0	0.0						
	GL	771	99.4						
	M	0	0.0						
	P	0	0.0						
29	GM	633	79.8	6	9	GM	*n/a*	*n/a*	< 5
	WM	114	14.4						
	GL	46	5.8						
	M	0	0.0						
	P	0	0.0						

n/a = not provided.
*Values arbitrary assigned to categorize medium and high density.

Table S4. Descriptive statistics for TCP distribution between DESI-MS smear images. Medians are statistically different (p-value <0.05)

	Low TCP	Medium TCP	High TCP
Minimum	0	0	0
Median	15	71	96
Interquartile Range	28	55	49
Maximum	100	100	100
Number of Pixels*	10968	4074	1472

*Total number of pixels selected for chemical diagnosis from 29 images

Table S5. Row-by-row predictions for DESI-MS smear images of the heterogeneous specimens # 10 and 18.

	Specimen # 10		Specimen # 18	
	# pixels	**TCP**	**# pixels**	**TCP**
Average DESI-MS image	646	19±11*	571	31±24
Line 1	34	8±10**	32	13±17
Line 2	34	14±11	31	19±21
Line 3	32	14±8	29	25±25
Line 4	32	16±5	28	25±24
Line 5	31	15±7	28	28±23
Line 6	31	14±7	29	31±25
Line 7	31	15±7	28	29±20
Line 8	31	15±8	29	33±23
Line 9	32	19±12	29	29±22
Line 10	32	17±8	28	30±23
Line 11	33	18±7	29	29±24
Line 12	32	18±12	29	33±25
Line 13	36	23±14	28	31±23
Line 14	33	19±7	28	34±22
Line 15	32	19±6	28	30±23
Line 16	32	20±7	28	31±23
Line 17	33	22±9	28	35±24
Line 18	32	23±9	28	42±24
Line 19	33	29±11	27	51±22**
Line 20	30	38±13**	27	51±21**

*Mean ± Standard Deviation

**Line averages statistically different from the other lines with 95% IC (p-value<0.05)

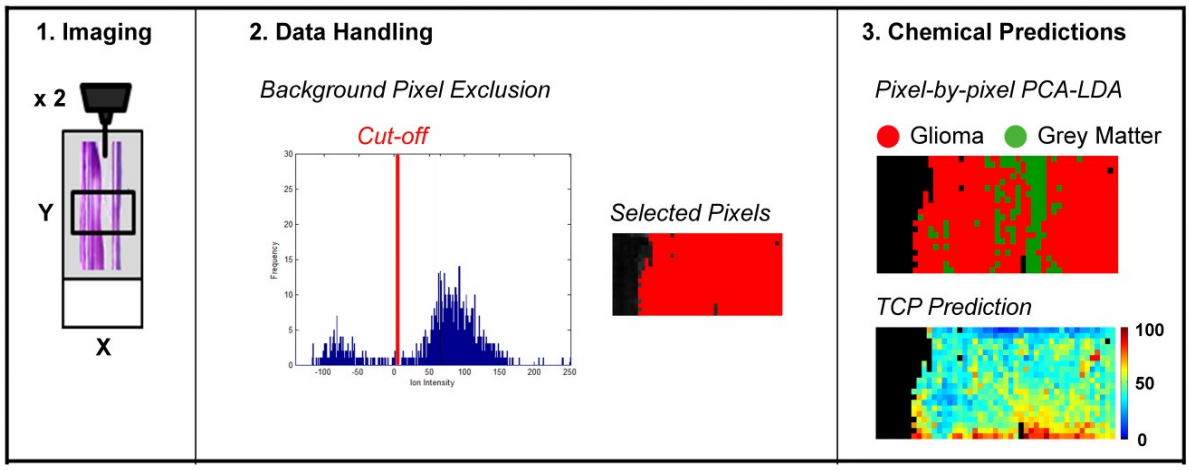

Fig. S1. Schematic description of DESI-MS imaging and data analysis workflow. **1.** For each smear, an area of 60 mm^2 was imaged by DESI-MS. **2.** Based on a threshold of ion intensities, background areas were excluded from data analysis. Pixels with ion intensities greater than the threshold were selected for data analysis. **3.** Individual pixels were chemically predicted by PCA-LDA. TCP values were estimated by linear regression of NAA's ion intensity. Predictions were color-coded and shown in the 2D DESI-MS spatial domain.

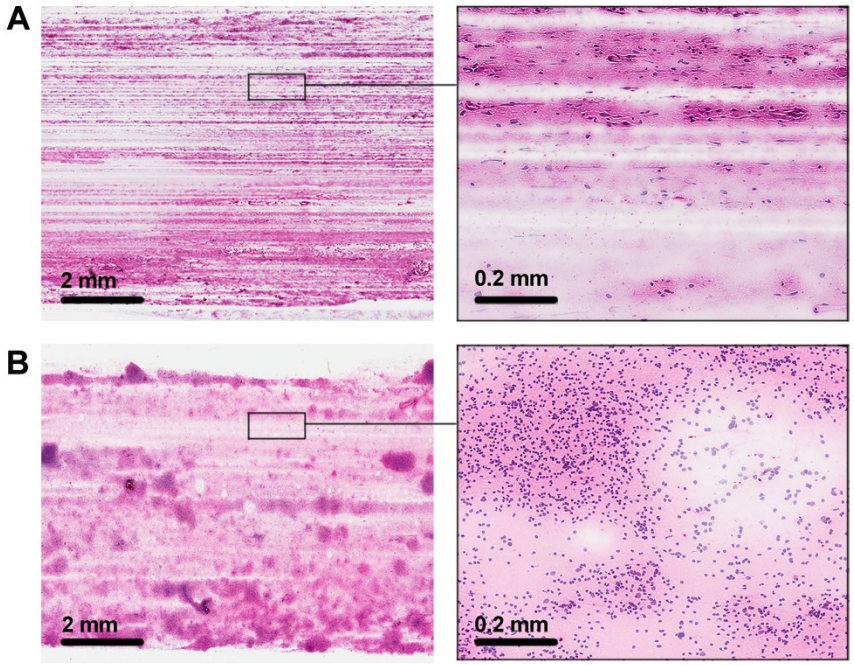

Fig. S2 (a) H&E stain of specimen # 18 with magnified region inset. **(b)** H&E stain of specimen # 28 with magnified region inset.

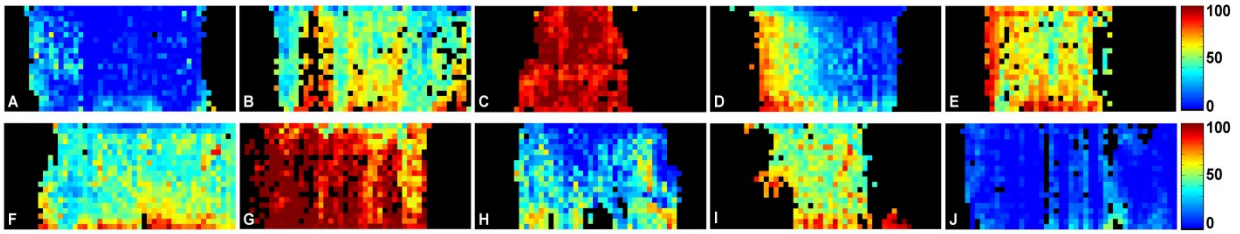

Fig. S3 Pixel-by-pixel TCP predictions based on NAA intensity of **(a)** specimen # 1 (infiltrated white matter with < 5% TCP by histopathology), **(b)** specimen # 17 (infiltrated grey matter with 10% TCP), **(c)** specimen # 4 (high grade medium density glioma), **(d)** specimen # 18 (high grade medium density glioma, **(e)** specimen # 19 (infiltrated grey with 20% TCP), **(f)** specimen # 27 (high grade medium density glioma), **(g)** specimen # 24 (high grade glioma), **(h)** specimen # 3 (infiltrated grey matter with < 5% TCP), **(i)** specimen # 26 (high grade medium density glioma), **(j)** specimen # 29 (infiltrated grey matter with < 5% TCP).

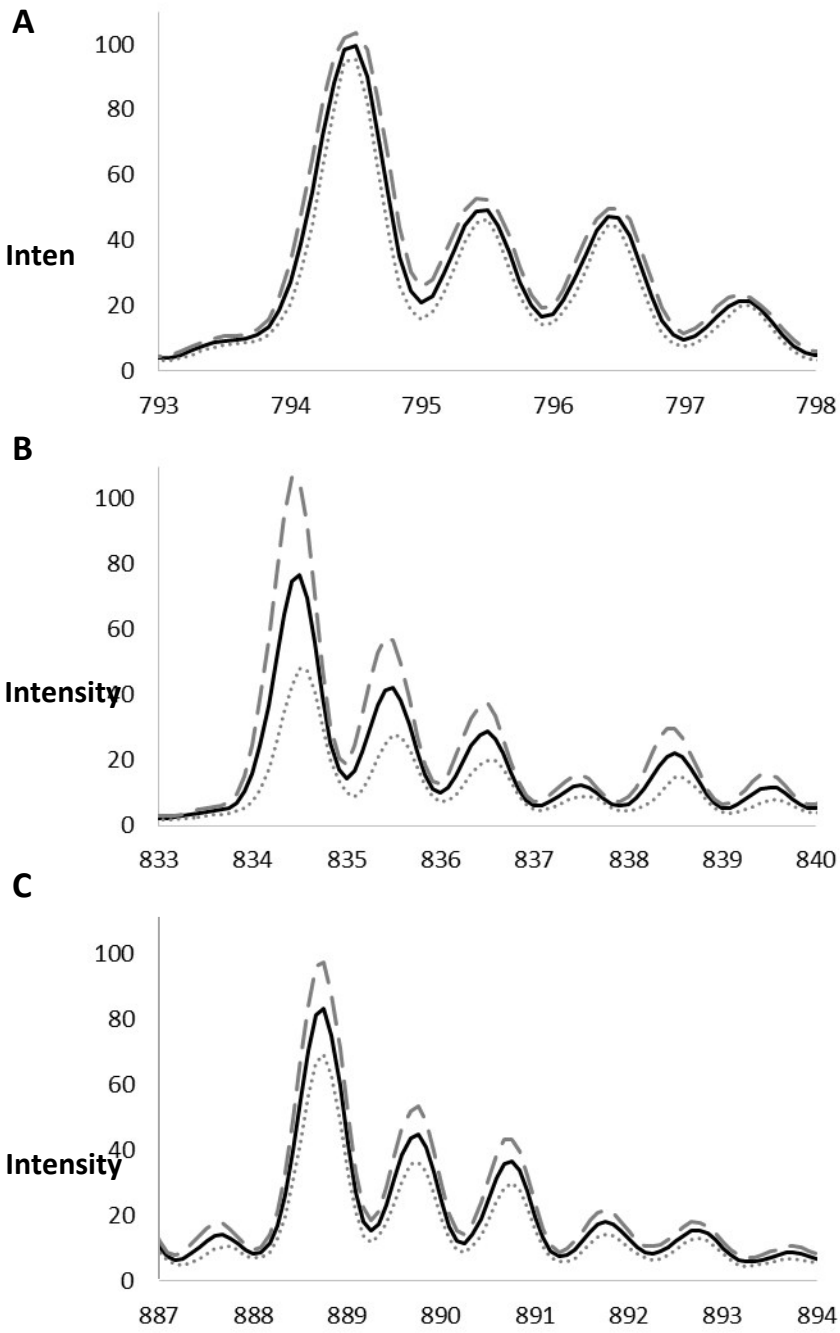

Fig. S4. Plot of line-by-line average peaks with standard deviation. **(a)** *m/z* 794, PC 34:1+Cl-, **(b)** *m/z* 834, PS 18:0_22:6, and **(c)** *m/z* 888, (3'-sulfo)GalCer 24:1 of specimen # 10. Mean denoted by solid black line; mean + three times the standard deviation illustrated by the mashed line; mean – three times the standard deviation illustrated by the dotted line.