Single Cell Profiling using Ionic Liquid Matrix-Enhanced Secondary Ion Mass Spectrometry for Neuronal Cell Type Differentiation

Thanh D. Do[┼], *Troy J. Comi*[┼], *Sage J. B. Dunham, Stanislav S. Rubakhin, Jonathan V. Sweedler*[*]

Department of Chemistry and the Beckman Institute

University of Illinois, Urbana-Champaign, IL 61801

* Corresponding author. Email: jsweedle@illinois.edu. Tel: 217-244-7359

┼ These authors contributed equally to this work.

SUPPORTING INFORMATION

Table of Contents

Sample Preparation

Aplysia californica. Two *Aplysia californica* (100–250 g body weight) were purchased from the National Resource for *Aplysia* (Rosenstiel School of Marine and Atmospheric Science University of Miami, FL). The mollusks were kept in aerated, circulated, filtered and chilled to 20 ^{o}C sea water prepared from Instant Ocean Sea Salt (Instant Ocean, Aquarium Systems Inc., Mentor, OH) dissolved in purified water. Animals were anesthetized by injection of isotonic $MgCl_2$ (~30% to 50% of body weight) into the body cavity. Central nervous system ganglia were dissected and placed in artificial sea water (ASW) containing 460 mM NaCl, 10 mM KCl, 10 mM $CaCl_2$, 22 mM $MgCl_2$, 26 mM $MgSO_4$ and 10 mM HEPES in Milli-Q water (Millipore, Billerica, MA), with the pH adjusted to 7.8 using 1 M NaOH in Milli-Q water. Ganglia were treated with enzyme solution consisting of 1% (wt/vol) protease type IX (Sigma Aldrich, St. Louis, MO) in ASW supplemented with 100 units/mL penicillin G, 100 µg/mL streptomycin, and 100 µg/mL gentamicin for 45 min at 34.4 ^{o}C. The connective tissue surrounding neurons and neuropil was surgically removed and multiple, individual neurons were mechanically isolated. Cells were deposited on indium-titanium oxide (ITO)-coated glass slides, 70–100 Ω/sq, 25 × 75 × 1.1 mm (Delta Technologies, Loveland, CO), which were washed with Milli-Q water and ethanol before use. The ITO-coated glass slides were placed in ASW and cells were allowed to adhere for 30 min. Next, ASW was replaced with a 33% glycerol 67% ASW (v/v) solution that was decanted after a 5 min incubation. Cells were left to dry overnight.

Rattus norvegicus. Seven 2.5–3 month old male Sprague-Dawley outbred rats (*Rattus norvegicus*) (www.envigo.com) were housed on a 12-h light cycle and fed ad libitum. Animal euthanasia was performed in accordance with the appropriate institutional animal care guidelines (the Illinois Institutional Animal Care and Use Committee), and in full compliance with federal guidelines for the humane care and treatment of animals. The studies were planned in accordance with the ARRIVE guidelines.[1]

Rats were killed by quick decapitation using a sharp guillotine. Rat trunks were placed on ice, where all surgical procedures were performed. Dorsal root ganglia (DRG) were surgically isolated during the ~10-min dissection procedure and placed into ~5 mL of cold Modified Gey's balanced salt solution (mGBSS) containing 1.5 mM $CaCl_2$, 4.9 mM KCl, 0.2 mM KH_2PO_4, 11 mM $MgCl_2$, 0.3 mM $MgSO_4$, 138 mM NaCl, 27.7 mM $NaHCO_3$, 0.8 mM NaH_2PO_4, and 25 mM HEPES dissolved in Milli-Q water, with the pH adjusted to 7.2 using 1 M NaOH in Milli-Q water.

To remove the surrounding connective tissue and isolate individual neurons, the DRG were incubated in 2.5% collagenase in oxygenated mGBSS for 25 min. All steps in the protocol were carried out at 37 ^{o}C, and oxygenated mGBSS was used in all cases. The DRG were then washed with 1% bovine serum albumin (BSA) in mGBSS for 7 min. The BSA solution was replaced with mGBSS and the DRG were incubated for an additional 20 min. Next, the DRG were treated with 0.65% trypsin in mGBSS for 20 min, followed by a 1% BSA solution wash for 7 min. The BSA solution was replaced with mGBSS containing Hoechst 33342 nuclear stain (Thermo Fisher Scientific, Waltham, MA) (1 mg/mL Hoechst 33342 stock solution, stored at 14 ^{o}C, diluted 1:500 in mGBSS) in which the DRG were incubated for 10 min. The Hoechst nuclear stain solution was then replaced with mGBSS and the DRG were mechanically dissociated by trituration. Cells were stabilized using a 40% glycerol, 60% mGBSS (v/v) mixture and after 5–10 min, plated on ITO-coated glass slides. The samples were then stored in the dark overnight to allow the cells to adhere onto the glass surface. On the next day, excess glycerol-containing media was removed from the preparations. At this stage, the cells were either subjected to immediate SIMS or MALDI analysis, or left to adhere to the ITO-coated glass surface in a nitrogen-purged dry box for 24 h before the MS measurements. Before analysis, the sample slides were rinsed with 2 mL of 150 mM ammonium acetate buffer (pH 10). This step helped remove the excess glycerol and did not induce observable damage to the cells.[2-4]

C_{60} SIMS Instrument Modifications

Several modifications were made to the cell coordinate registration protocol and hardware of the previously-reported C_{60} SIMS instrument.[5] For the registration protocol, three types of coordinate systems were utilized to interface with the instrument and optical images. The first system was in pixel coordinates in the microscopy image such that each dispersed cell had a unique (X, Y) pixel location. The second coordinate system was the stepper motor location (X_{motor}, Y_{motor}) of the instrument *x*,*y*-translation stage. When the ITO-coated glass slide was loaded into the instrument stage, the fiducial markers were used to calculate the transformation from (X, Y) to (X_{motor}, Y_{motor}) for each cell. These steps were identical to the previously reported coordinate registration protocol.[3,4] Unfortunately, the oMALDI Server software (AB Sciex, Framingham, MA) does not accept motor coordinates as direct inputs. The “Search Pattern” feature directed the stage to dwell at a series of points relative to a pre-set origin dictated by a separate coordinate system, which we refer to as the “Pattern” (PTN) coordinate

system. One PTN unit is equivalent to 0.5 mm. The origin was set at the point $X_{motor,}$ Y_{motor} = (8,000, 17,600) which was equal to X_{PTN}, Y_{PTN} = (5, 5) in the PTN coordinate system. Through two coordinate transformations, each cell from the optical image was mapped to a PTN coordinate readable by oMALDI Server. Cells were then analyzed with the 20kV C_{60}^{+} ion beam as the stage traveled through the locations in the "Search Pattern" inputs.

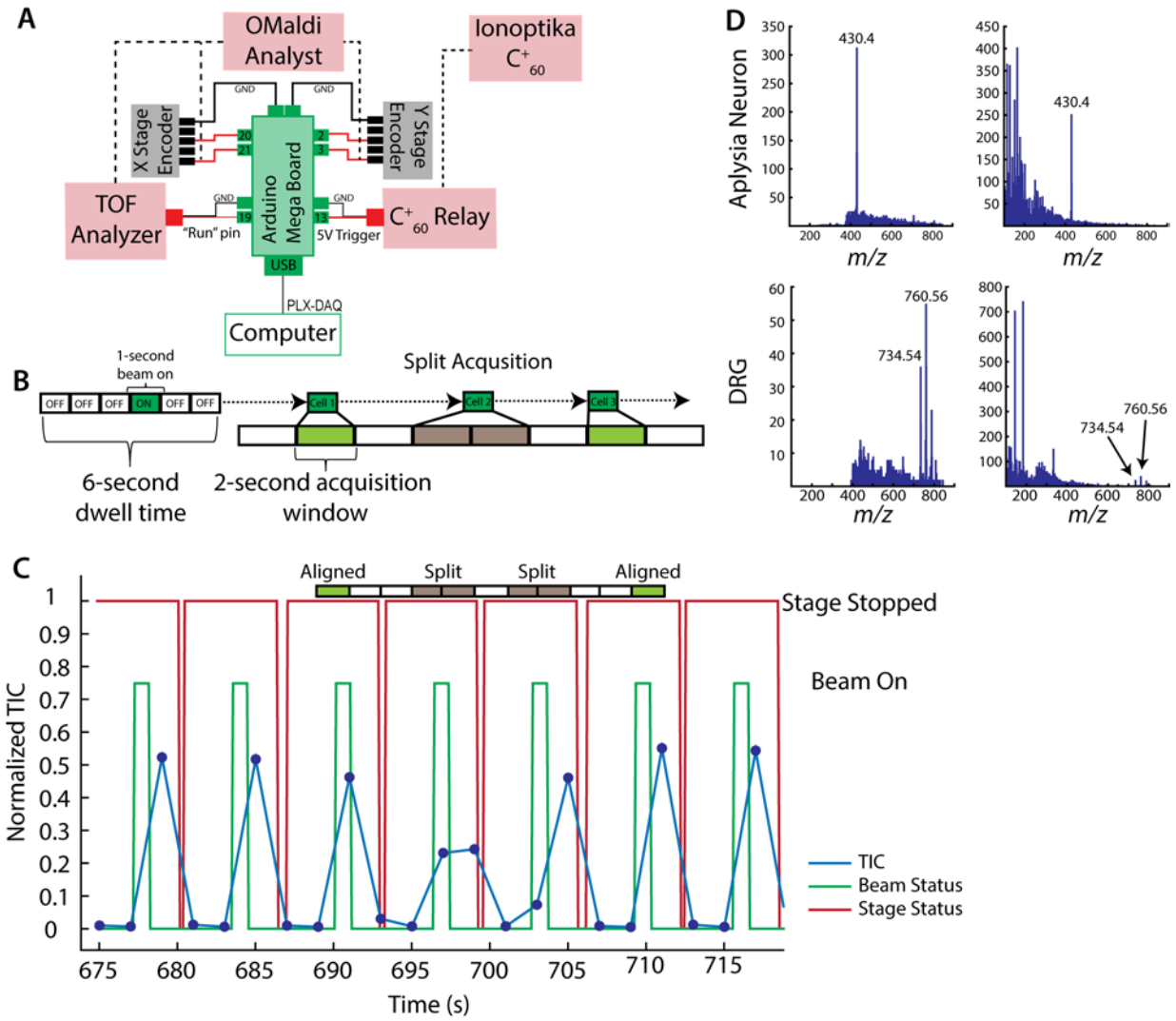

Figure S1. (A) Schematic of the microcontroller integration to synchronize motor stage movements, mass spectrometer acquisition, and C_{60}^{+} ion beam status. (B) An illustration of a 6-second dwell time and a "split" cell with a 2 s acquisition time. (C) Overlay of total ion current, stage motion, and beam state as a function of time for a representative sample. Instances of aligned and split spectra are highlighted above. (D) Representative examples of "split" cells. In the top two mass spectra, the signals of vitamin E (*m/z* 430.4) are split between the two acquisition windows. In the bottom two mass spectra, the lipid signals are also split.

An Arduino Atmega 2560 board (https://www.arduino.cc/) was programmed by the Arduino Software, v1.6.9, to do the described tasks (Figure S1A). Since each stepper motor was tracked by a linear encoder, its signals from the quadrature channels A/B were tapped out and connected to two interrupt pins on the Arduino board, which provide fast response times. A 4x decoding method was used to determine when the stage was moving (i.e., changes in signals of either encoder) and when the stage stopped (i.e., no changes in both encoders). As the data acquisition on the mass spectrometer was initiated by Analyst (AB Sciex), a 5 V signal from the mass spectrometer control board was received by interrupt pin 19 on the Arduino board to activate stage monitoring and synchronize the start times of the stage and mass spectrometer. When the stage stopped at a cell, the Arduino board would delay for 3 s and then send a 5 V signal through pin 13 to a relay controlling the C_{60}^+ beam, initiating desorption for 1 s (see Figure S1B). Longer beam "On" times cause damage on the sample surface and generate more chemical noise. The beam "Off" delays ensured no contamination occurred between cells of interest and surrounding neighbors. Data from the Arduino, including clock time, elapsed time, encoder position and ion beam status, were recorded with the PLX-DAQ add-on for Excel 2010 (https://www.parallax.com/downloads/plx-daq).

Instrument Parameters

MALDI/C_{60}-SIMS Q-TOF instrument parameters. The 40-μm diameter, 20 kV C_{60}^+ ion beam (Ionoptika, Ltd., Hampshire, UK) was operated in continuous mode with 500 pA sample current to yield an ion dose of 2.5×10^{14} ions/cm^2. Positive secondary ions were collected from *m/z* 60–850 using a Q1 bias of 5%, 5%, 20%, 35% and 35% at *m/z* 100, 180, 300, 500 and 700, respectively. The signal accumulation time was set to 2 s. The time-bins-to-sum was set to 10. Tandem MS spectra were collected in product-ion mode with the argon collision gas and collision-induced dissociation energy set at 35 eV.

Bruker ultrafleXtreme MALDI TOF/TOF instrument parameters. The molecular mass scan window was set to *m/z* 400−8000 and the laser was operated in the "Ultra" mode, producing a ~100-μm diameter footprint. The ultrafleXtreme AutoXecute feature (Bruker Daltonics, Billerica, MA) was utilized with the custom geometry file as previously reported.[4] Each spectrum represents the summed signals acquired during 1,000 laser shots fired at 1 kHz.

Figures S2–10; Table S1

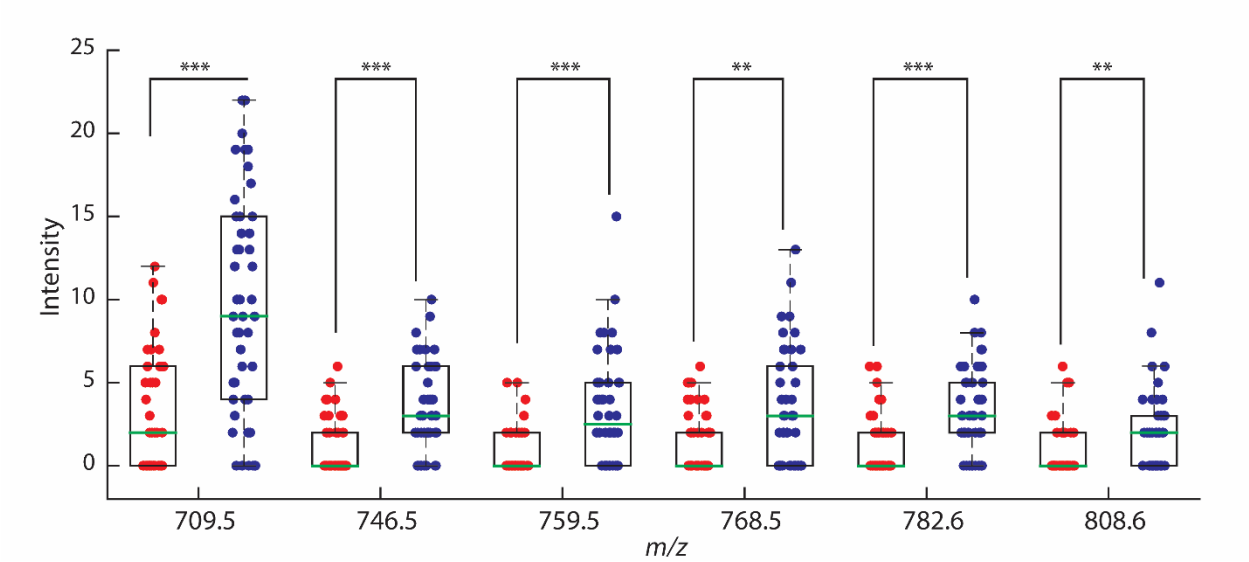

Figure S2. Box plot graph summarizing data for no-matrix (red) and matrix-assisted SIMS (blue) for lipid signals identified in a previous SIMS imaging report.[8] The statistical significance in each comparison was determined by a Wilcoxon rank sum test, ** p <0.005, *** p <0.0005, n = 47 cells from two *Aplysia*.

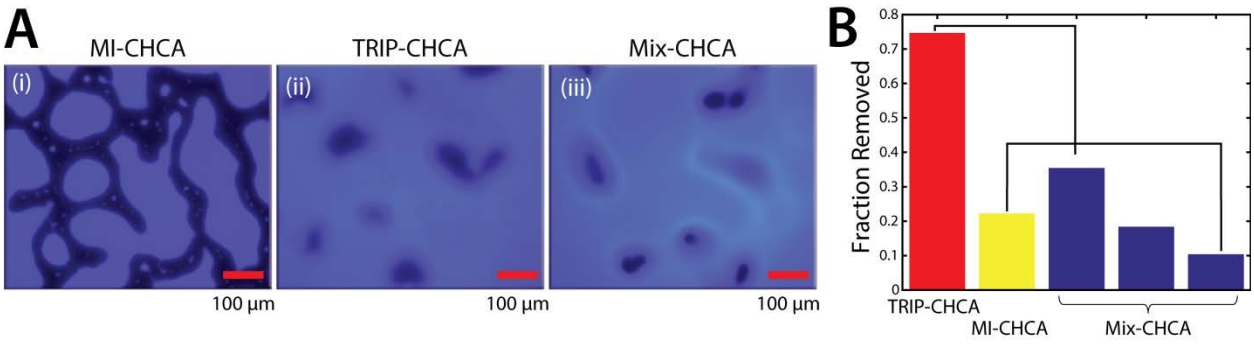

Figure S3. (A) Representative fluorescent micrograph of ITO slides coated with (i) MI-CHCA, (ii) TRIP-CHCA, and (iii) Mix-CHCA. (B) The fraction of data acquired from individual DRG cells was removed due to unacceptably low lipid signals. The black lines indicate samples from the same animals.

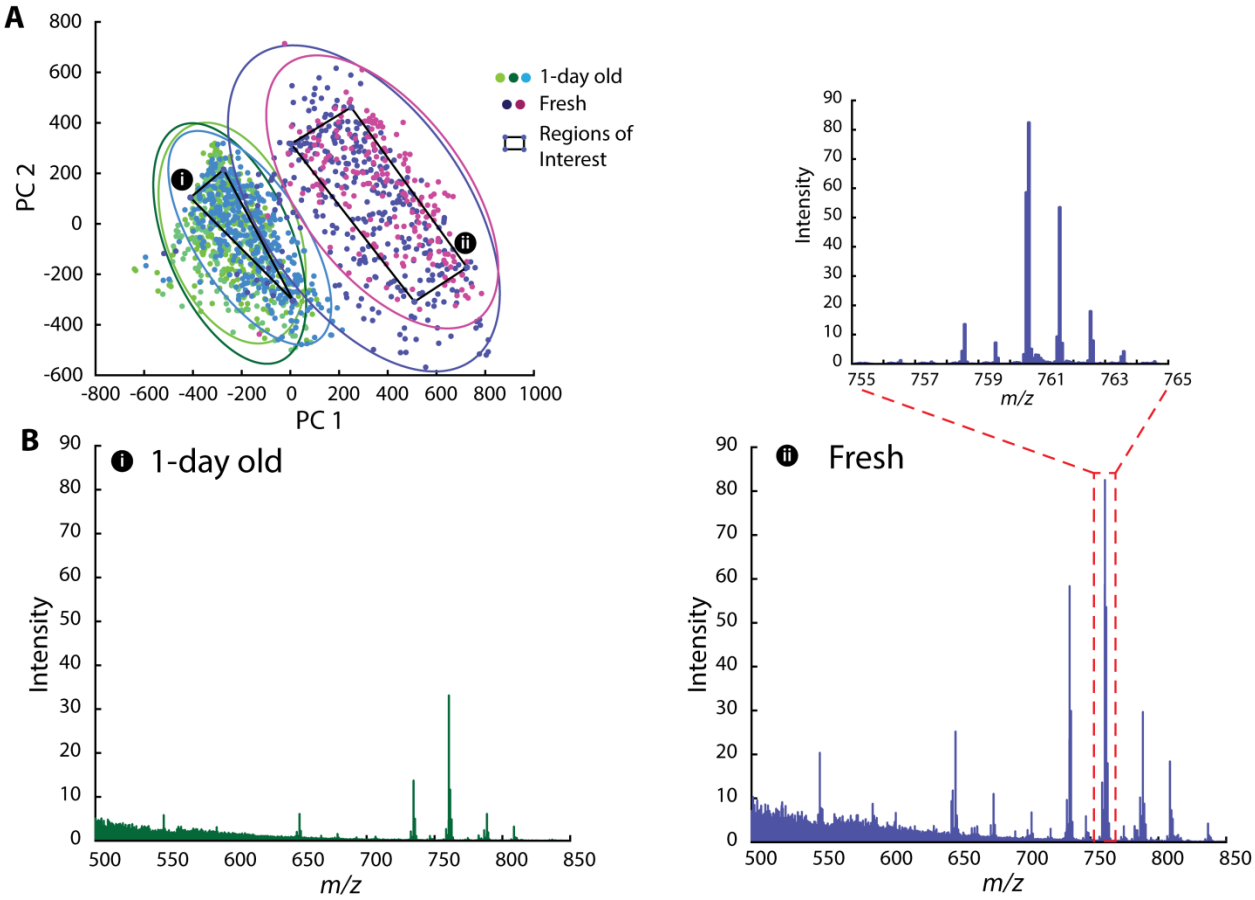

Figure S4. (A) PCA score plot of 1-day old and fresh DRG samples. (B) Averaged mass spectra for each sample type, (i) 1-day old samples, (ii) fresh samples.

Table S1. Tentative assignments of lipids observed in SIMS single cell profiling of DRG.

Species	$[M+H]^+$ (Da)	Observed
PC(40:5)	836.54	836.60
PC(38:1)	816.64	816.60
PC(38:5)	808.60	808.60
PC(38:6)	806.56	806.54
PC(36:1)	788.61	788.60
PC(36:2)	786.64	786.60
PC(36:3)	784.58	784.56
PC(36:4)	782.56	782.56
PC(35:1)	774.60	774.58
PC(35:2)	772.58	772.56
PC(34:0)	762.60	762.60
PC(34:1)	760.59	760.56
PC(34:2)	758.56	758.56
PC(33:0)	748.58	748.60
PC(33:1)/PC(O-34:1)	746.56	746.60
PC(32:0)	734.57	734.54
PC(32:1)	732.55	732.50
PC(26:0)/PC(O-28:0)	650.47	650.40
SM(18:0)	731.60	731.60
PC(31:0)	720.55	720.50
PC(30:0)	706.53	706.50
PC(32:0) Frag	478.32	478.30
LysoPC(16:0)/PC(16:0)	496.33	496.40
PC(34:1) Frag/LysoPE(20:3)	504.40	504.30
PC(O-18:1)/PC(18:0)/LysoPC(18:1)	522.34	522.30
PC(O-20:1)/PC(19:1)/	550.35	550.30

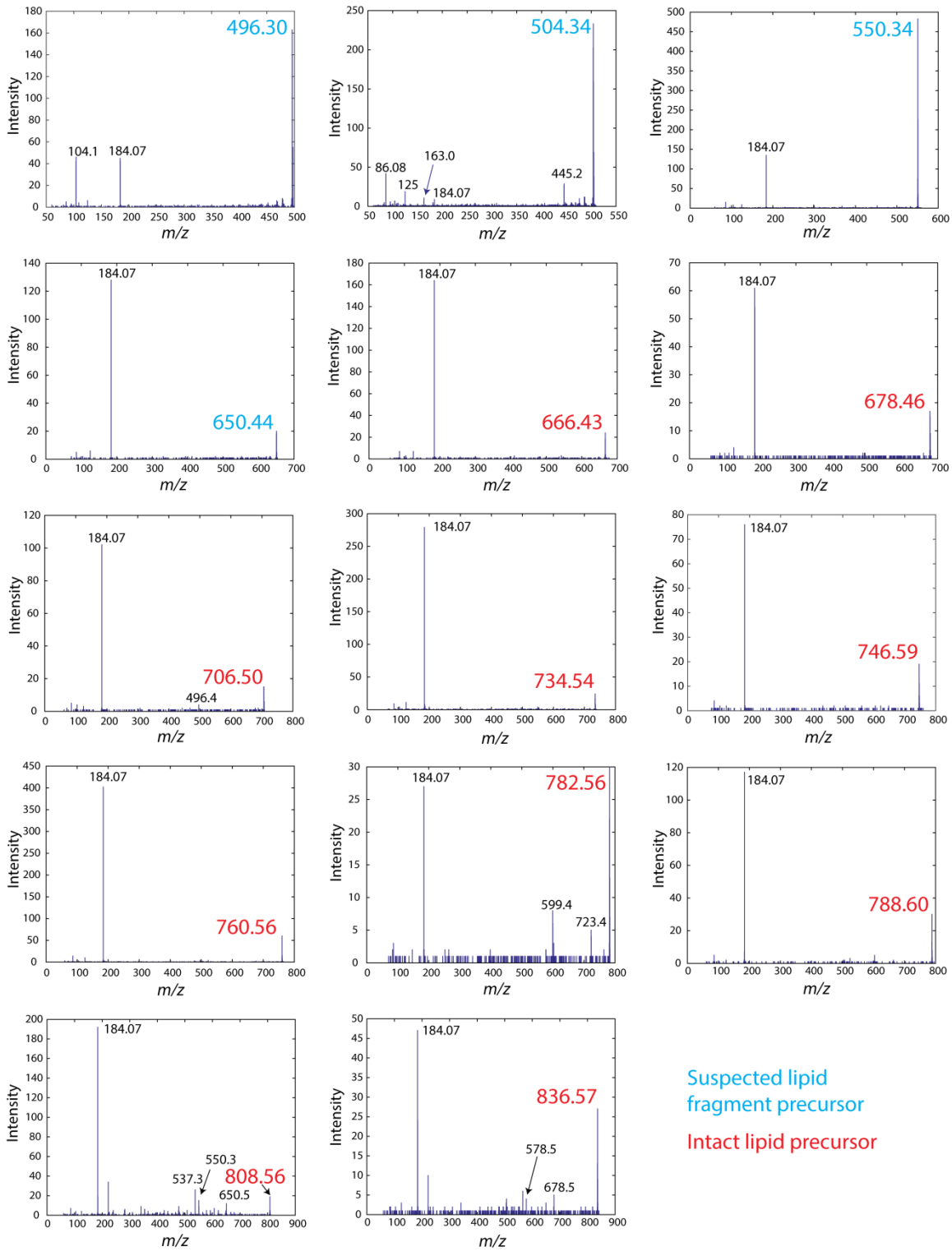

Figure S5. Tandem mass spectra of major lipids obtained from SIMS single cell profiling. The tentative identity of each lipid is provided in Table S1.

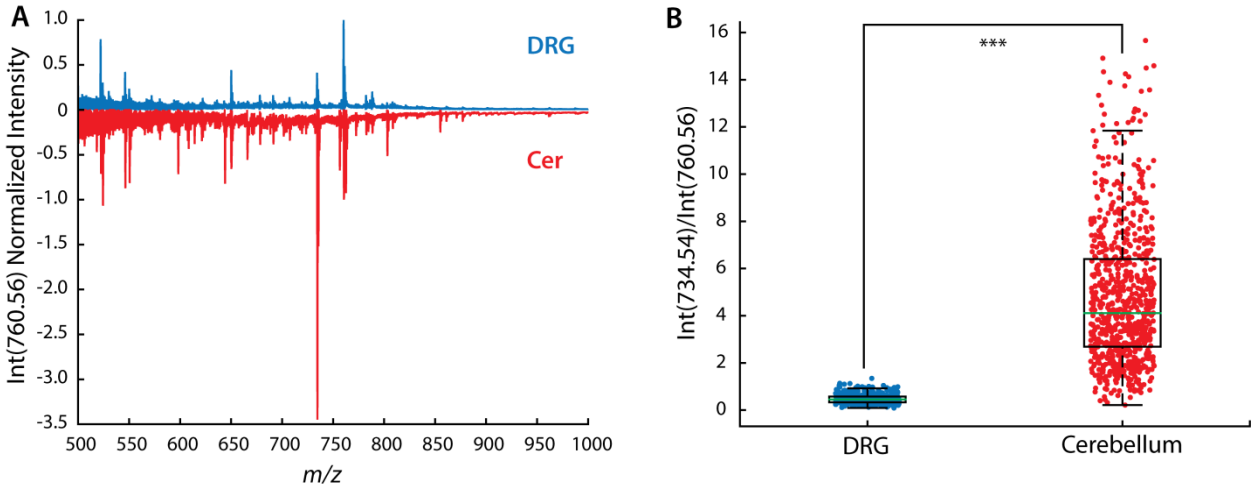

Figure S6. (A) Comparison of lipid profiles obtained from MALDI MS of single DRG cells (top) and cerebellum (Cer) cells (bottom) using Mix-CHCA. (B) Quantitative comparison of the ratio of two dominant lipids, PC(32:0) and PC(34:1) at *m/z* 734.54 and 760.56, respectively, between the two cell types, *** p <0.005 by a Wilcoxon rank sum test.

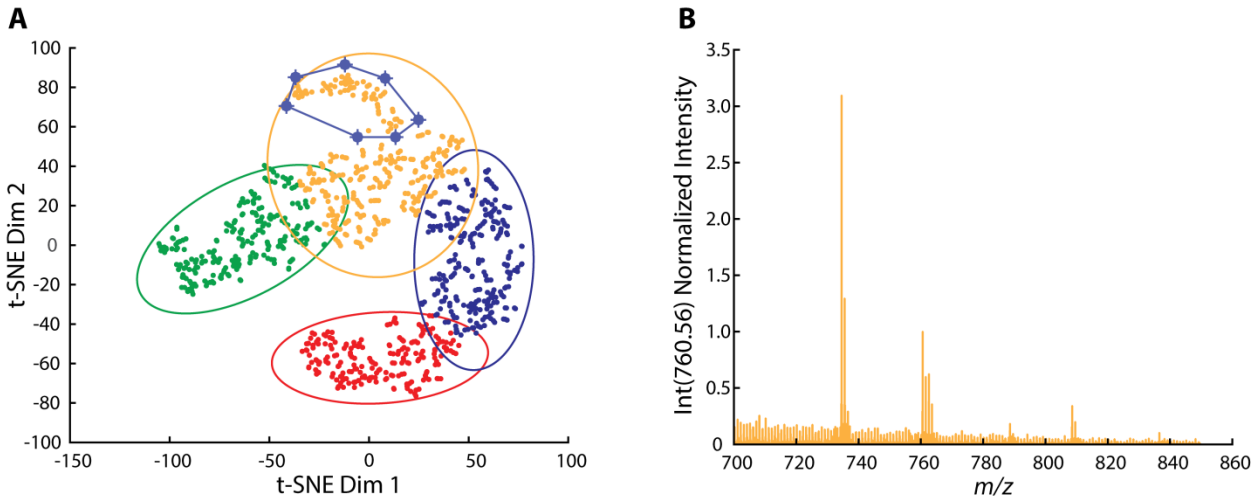

Figure S7. (A) Multivariate analysis of single cerebellum cells including *k*-mean clustering reproduced from Figure 6 in the main text. A subsection of the yellow cluster is annotated. (B) The averaged mass spectrum of single cerebellum cells in the subsection highlighted in (A) showing low S/N.

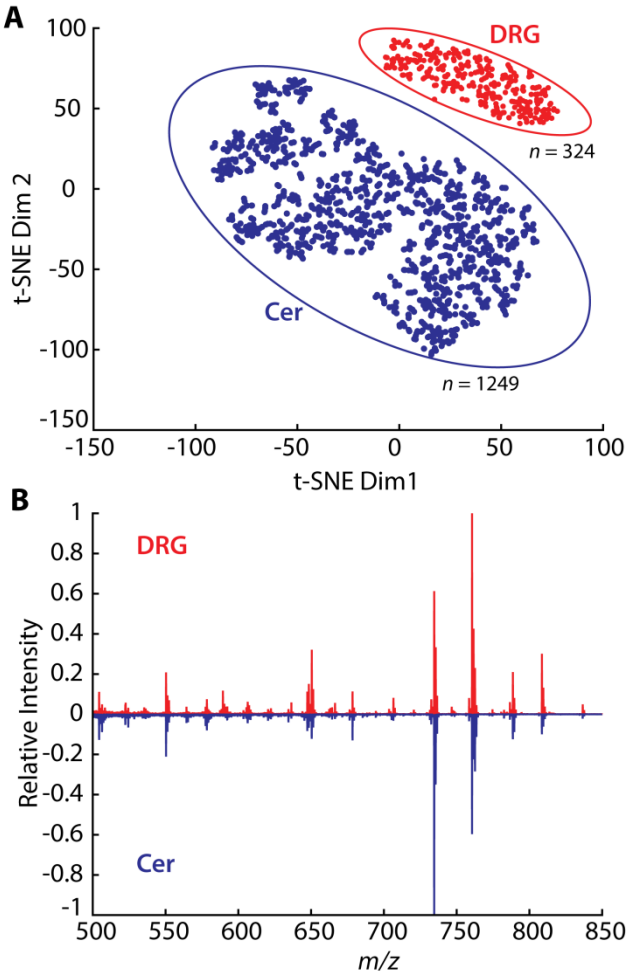

Figure S8. Differentiation of two neuronal cell types based on single cell SIMS profiling. (A) t-SNE plot of intact lipids from DRG and cerebellar (Cer) cells with 95% confidence ellipses. (B) The averaged SIMS spectra of single DRG (red) and cerebellar cells (blue).

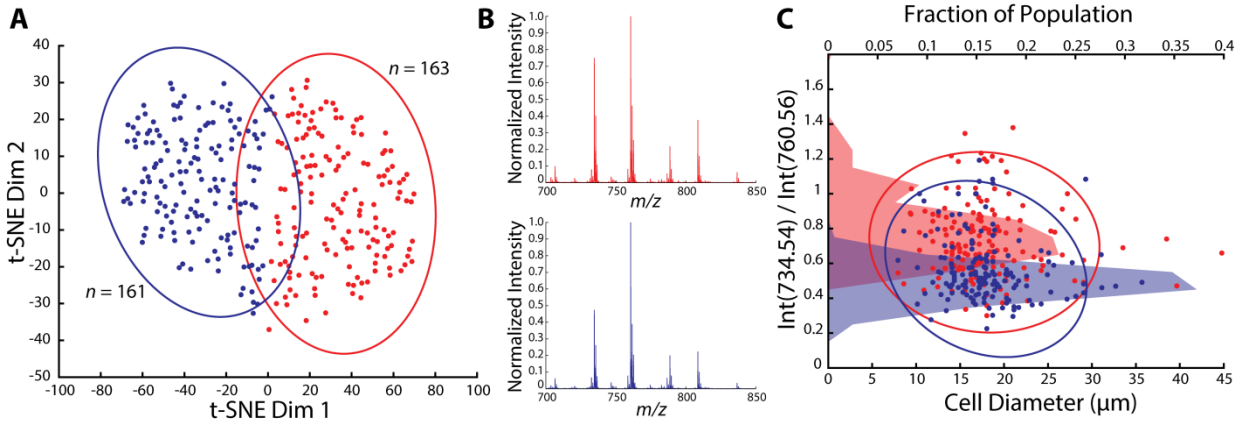

Figure S9. Identification of subpopulations within single DRG neurons. (A) *k*-means clusters based on t-SNE distribution of single DRG cells. (B) Averaged spectra of each cluster. The averaged normalized intensities of PC(32:0) are 0.80 and 0.49 cluster 1 (red) and cluster 2 (blue), respectively. (C) Plot of Int(734.56)/Int(760.56) as a function of cell diameter colored based on the *k*-means clusters in panel A. Overlay: a histogram of the relative intensities for each cluster. The 95% confidence ellipses are shown in A and C.

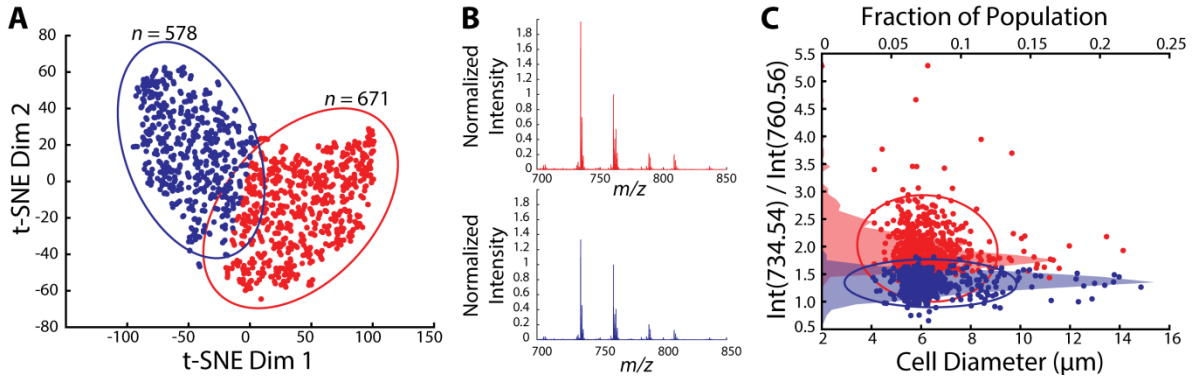

Figure S10. Multivariate analysis of single cerebellum cells. (A) *k*-means clustering of *t*-SNE distributions. (B) Averaged spectra of each cluster. The averaged normalized intensities of PC(32:0) are 2.0 and 1.3 for cluster 1 (red) and cluster 2 (blue), respectively. (C) Plot of cell diameter versus Int(734.54)/Int(760.56) showing the relative distribution of cells within each cluster in terms of size. Overlay: histograms of the lipid ratio for each cluster.

Threshold Limit for Mass Spectral Filtering

For the data sets used to generate Figures 4-6 and Supporting Information Figures S8-S10, the cutoff threshold of 250 counts for *m/z* 184.07 is quite low as few DRG cells (< 5 cells) failed to pass this filter (see Figure S11, panel A). For the cerebellar samples, the larger population required additional time to analyze than for DRG samples. During acquisition, the ionic liquid matrix degraded which reduced the quality of the mass spectra over the course of several hours. As shown in Figure S11, panel B, the data obtained from batch 03 of dataset 01 (which were collected after batches 01 and 02; ~2.5 hours in high vacuum) had many cells whose intensity of *m/z* 184.07 was below the threshold limit. Removing these cells ensured the highest quality spectra were compared between batches.

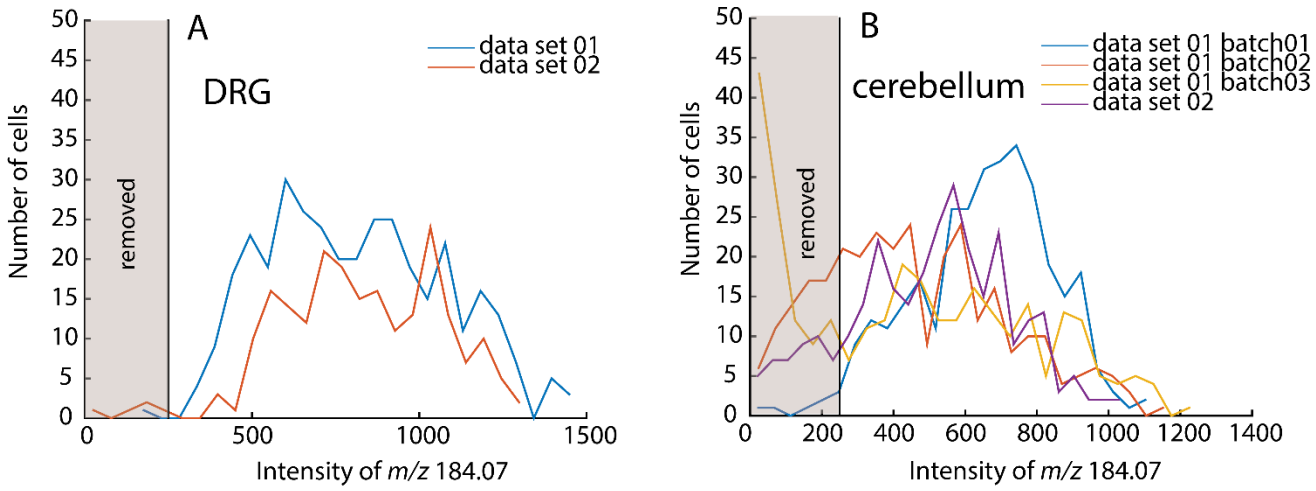

Figure S11. Histograms of the intensities of *m/z* 184.07 for (A) DRG cells and (B) cerebellar cells. The data of these cells were used to generate Figures 4 to 6 in the manuscript.

The population of the filtered cells did not appear to represent a separate population of cells. The averaged spectrum of all the removed cerebellar cells are shown in Figure S12, panel A, which displays strong chemical noise in the *m/z* 200–600 region and intense signals of *m/z* 172, a suspected degradation product of α-CHCA. Panel B shows the distribution of Int(734.54)/Int(760.56) for the cells that were removed. Panel C shows the overlaid mass spectra of 8 cells that have Int(734.54)/Int(760.56) > 5.

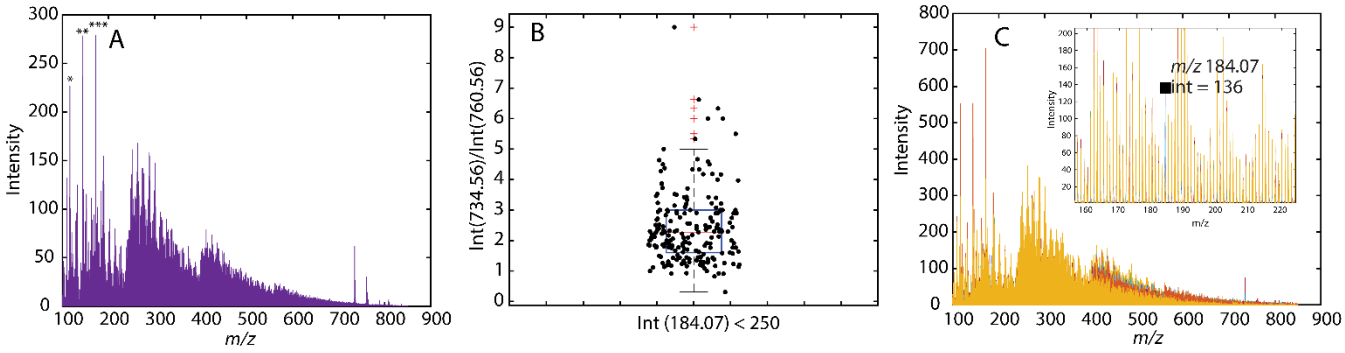

Figure S12. (A) The averaged mass spectrum of cerebellar cells removed prior to analysis. (*) *m/z* 116, (**) *m/z* 144 (Indium), (***) *m/z* 172. (B) Box plot of Int(734)/Int(760). (C) Overlaid mass spectra of single cerebellum cells with Int(734)/Int(760) > 5 showing low intensities of lipids and *m/z* 184.07 suggesting the cells were likely missed.

Multivariate Statistical Analysis

Data analysis was performed with custom scripts written in MATLAB (R2015b). MALDI MS data were read directly with the readbrukermaldi function (https://github.com/AlexHenderson/readbrukermaldi), resampled to 10,000 *m/z* values in the range *m/z* 500–1000, background corrected, smoothed, and normalized by standardizing the area under each spectrum to the median of the dataset. SIMS data was first converted from the native wiff format into mzXML with msconvert[6] for import into MATLAB. The cell coordinates, diameters, and the corresponding log file from the instrument microcontroller were also utilized to parse continuous SIMS acquisition files. As the Arduino monitors when an acquisition begins, the start time for mass spectral acquisition and stage movement were synchronized. In MATLAB, the *x,y* translational stage stop events were recorded and used to delineate mass spectra corresponding to target cell positions. Stage dwell events shorter than 3 s or longer than 7 s were discarded as noise in the stage or encoders. As the ion beam is only “On” for 1 s within the 6 s dwell time, the spectrum with the highest intensity phosphocholine head group (*m/z* 184.07) and PC(34:1) (*m/z* 760.56) signal was selected as the single cell spectrum. Dwell events of single cell signals that occurred in two adjacent acquisition windows were discarded as “split cells”. Finally, all mass spectra acquired from DRG and cerebellar cells were filtered for an intensity of the *m/z* 184.07 signal greater than 250 counts.

Statistical significance was established with a Wilcoxon rank sum test as intensity distributions were non-normal by a Kolmogrov-Smirnov test. Initial mass spectral data visualization was performed with

principal component analysis (PCA) to evaluate the effects of different ionic liquid formulations. Unsupervised cell classification for DRG and cerebellum samples was performed with *t*-SNE[7] to reduce and group the data in two dimensions, followed by *k*-means clustering with a Euclidean distance metric. The number of clusters was determined by the mean silhouette value as a function of *k*.

For the PCA, each cell spectrum was considered as an independent sample with the different *m/z* channels as the measured variables. The intensities of each *m/z* value for each cell produced a two-dimensional matrix which was decomposed to score and loading values with the built-in pca function in MALTAB. Principal component (PC) scores for 1 and 2 were displayed on a scatter plot to assess data grouping based on ionic liquid formulation. The loading plot of PC1 was also displayed to determine the cause of data spread in PC1 axis. In this plot, negative loading for a specific *m/z* indicates the specie is present in high abundance in cells found in the negative PC1 range and similar for positive values. Based on this interpretation and the assistance of average spectra, it was determined negative PC scores correspond largely to cells with high chemical noise while positive values contained biologically relevant lipid and phosphocholine signals.

References

(1) Kilkenny, C.; Browne, W. J.; Cuthill, I. C.; Emerson, M.; Altman, D. G. *PLoS Biol.* **2010**, *8*, e1000412.
(2) Berman, E. S.; Fortson, S. L.; Checchi, K. D.; Wu, L.; Felton, J. S.; Wu, K. J.; Kulp, K. S. *J. Am. Soc. Mass Spectrom.* **2008**, *19*, 1230-1236.
(3) Jansson, E. T.; Comi, T. J.; Rubakhin, S. S.; Sweedler, J. V. *ACS Chem. Biol.* **2016**, *11*, 2588-2595.
(4) Ong, T. H.; Kissick, D. J.; Jansson, E. T.; Comi, T. J.; Romanova, E. V.; Rubakhin, S. S.; Sweedler, J. V. *Anal. Chem.* **2015**, *87*, 7036-7042.
(5) Lanni, E. J.; Dunham, S. J.; Nemes, P.; Rubakhin, S. S.; Sweedler, J. V. *J. Am. Soc. Mass Spectrom.* **2014**, *25*, 1897-1907.
(6) Chambers, M. C.; Maclean, B.; Burke, R.; Amodei, D.; Ruderman, D. L.; Neumann, S.; Gatto, L.; Fischer, B.; Pratt, B.; Egertson, J.; Hoff, K.; Kessner, D.; Tasman, N.; Shulman, N.; Frewen, B.; Baker, T. A.; Brusniak, M. Y.; Paulse, C.; Creasy, D.; Flashner, L., et al. *Nat. Biotechnol.* **2012**, *30*, 918-920.
(7) Maaten, L. J. P. v. d.; Hinto, G. E. *J. Mach. Learn. Res.* **2008**, *9*, 2579-2605.
(8) Passarelli, M. K.; Ewing, A. G.; Winograd, N. *Anal. Chem.* **2013**, *85*, 2231-2238.