

## Supplementary Information

### ***In situ* tissue pathology from spatially encoded mass spectrometry classifiers visualized in real time through augmented reality**

Michael Woolman<sup>a,b,\*</sup>, Jimmy Qiu<sup>a,\*</sup>, Claudia M. Kuzan-Fischer<sup>c,f</sup>, Isabelle Ferry<sup>c,f</sup>, Delaram Dara<sup>a</sup>, Lauren Katz<sup>a,b</sup>, Fowad Daud<sup>a,b</sup>, Megan Wu<sup>c,h,i</sup>, Manuela Ventura<sup>a</sup>, Nicholas Bernards<sup>a</sup>, Harley Chan<sup>a</sup>, Inga Fricke<sup>a</sup>, Mark Zaidi<sup>a</sup>, Brad G. Wouters<sup>a,b</sup>, James T. Rutka<sup>c,d,f</sup>, Sunit Das<sup>c,d,f</sup>, Jonathan Irish<sup>a</sup>, Robert Weersink<sup>a</sup>, Howard J. Ginsberg<sup>a,d,e</sup>, David A. Jaffray<sup>a,b</sup> and Arash Zarrine-Afsar<sup>a,b,d,e,\*\*</sup>

<sup>a</sup> Techna Institute for the Advancement of Technology for Health, University Health Network, 100 College Street, Toronto, ON, M5G 1P5, Canada

<sup>b</sup> Department of Medical Biophysics, University of Toronto, 101 College Street, Toronto, ON, M5G 1L7, Canada

<sup>c</sup> Peter Gilgan Centre for Research and Learning, Hospital for Sick Children, 686 Bay Street, Toronto, ON, M5G 0A4, Canada

<sup>d</sup> Department of Surgery, University of Toronto, 149 College Street, Toronto, ON, M5T 1P5, Canada

<sup>e</sup> Keenan Research Center for Biomedical Science & the Li Ka Shing Knowledge Institute, St. Michael's Hospital, 30 Bond Street, Toronto, ON, M5B 1W8, Canada

<sup>f</sup> Arthur and Sonia Labatt Brain Tumor Research Centre, The Hospital for Sick Children, Toronto, ON, M5G 1X8, Canada

\* These authors contributed equally to this work

\*\*Corresponding author: Arash Zarrine-Afsar

Email: arash.zarrine.afsar@utoronto.ca

101 College Street, Room 7-207,

MaRS Building, Princess Margaret Cancer Research Tower, 7th floor (STTARR)

Toronto, ON M5G 1L7

TEL (Office) : ++1-416-581-8473

#### **Author Contributions**

Designed research: DAJ, AZ-A, JQ, RW, M. Woolman, CMKF, I. Ferry

Performed research: JQ, MW, DD, FD, CMKF, I. Ferry

Created and provided reagents: M. Wu, MV, NB, HC, I. Fricke

Analyzed results: AZ-A, M. Woolman, JQ, FD, DD, LK, MZ

Provided supervisory and support: AZ-A, JI, BGW, SD, JTR

Wrote the manuscript: AZ-A, M. Woolman, CMKF

## Table of Contents

Table S1-----	Page 3
Figure S1-----	Page 4-5
Figure S2-----	Page 6
Figure S3-----	Page 7
Figure S4-----	Page 8
Experimental Methods (additional information)-----	Page 9-11
Determination positional accuracy-----	Pages 12-14
References-----	Page 15

Med8A/DAOY/Muscle						Correct Classification Rate	
	Data point groups					Per classifiable	Per all (classifiable and unclassifiable)
	PIRL-MS spectra data points	Correctly classified data points	Misclassified data points	Unclassifiable data points	Classifiable data points	data points	data points
Total	149	146	0	3	146	100.00%	97.99%
Confusion matrix							
	Med8A	DAOY	Muscle	Unclassifiable data points		Total	
Med8A	51	0	0	0		51	
DAOY	0	44	0	2		46	
Muscle	0	0	51	1		52	
Total	51	44	51	3		149	

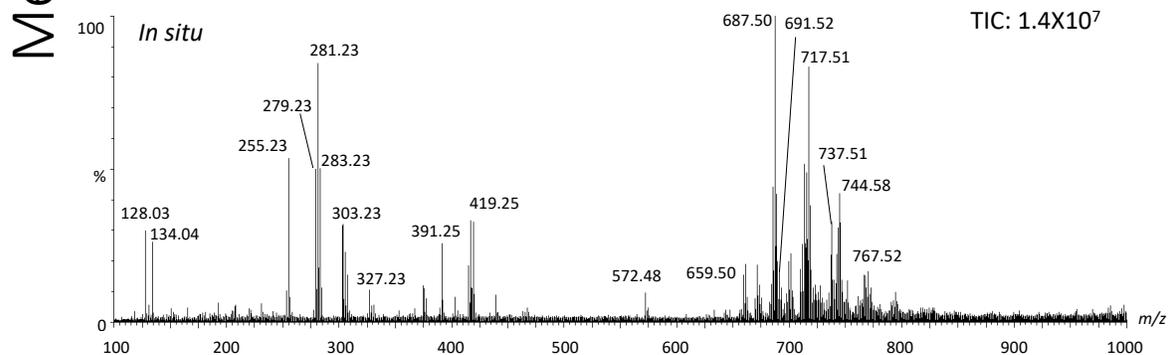
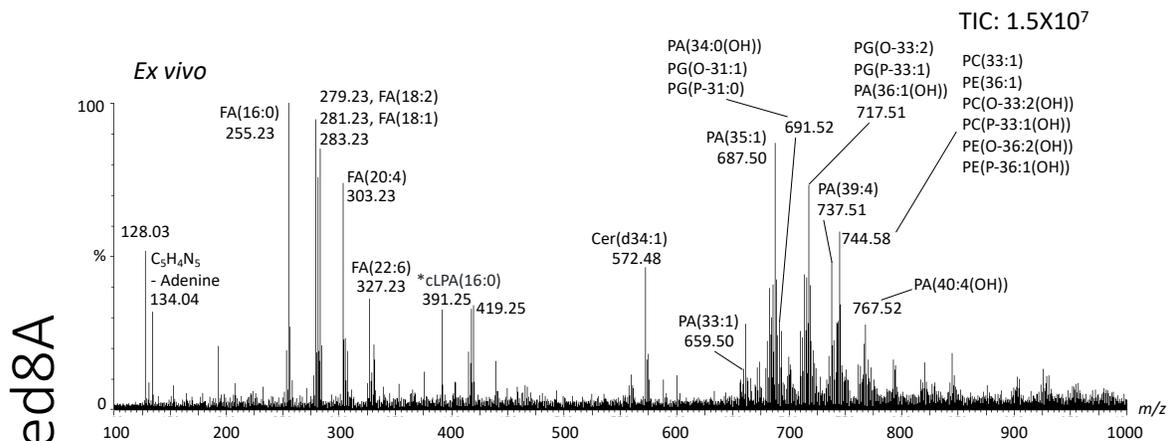
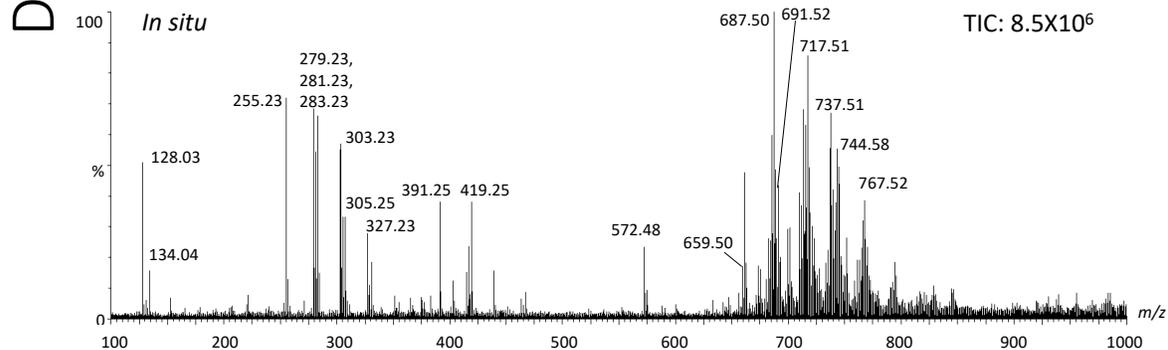
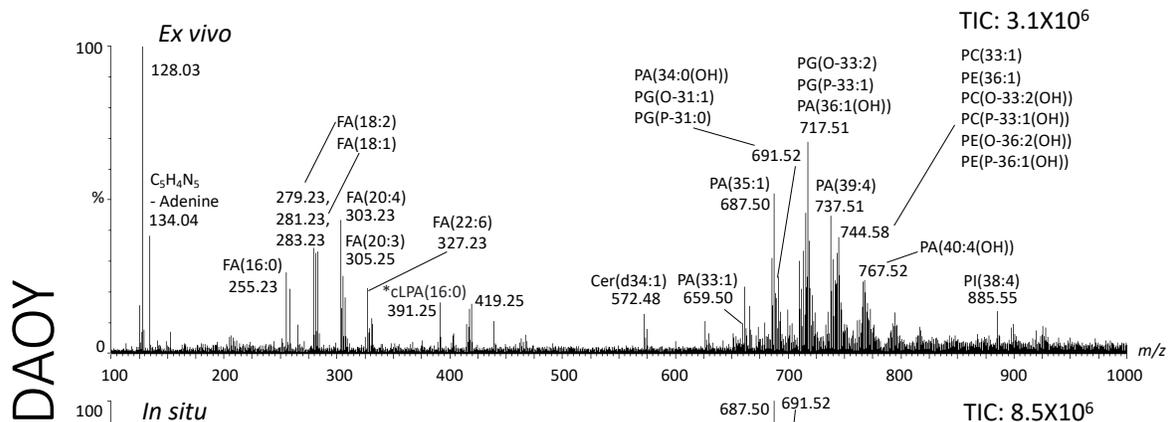
  

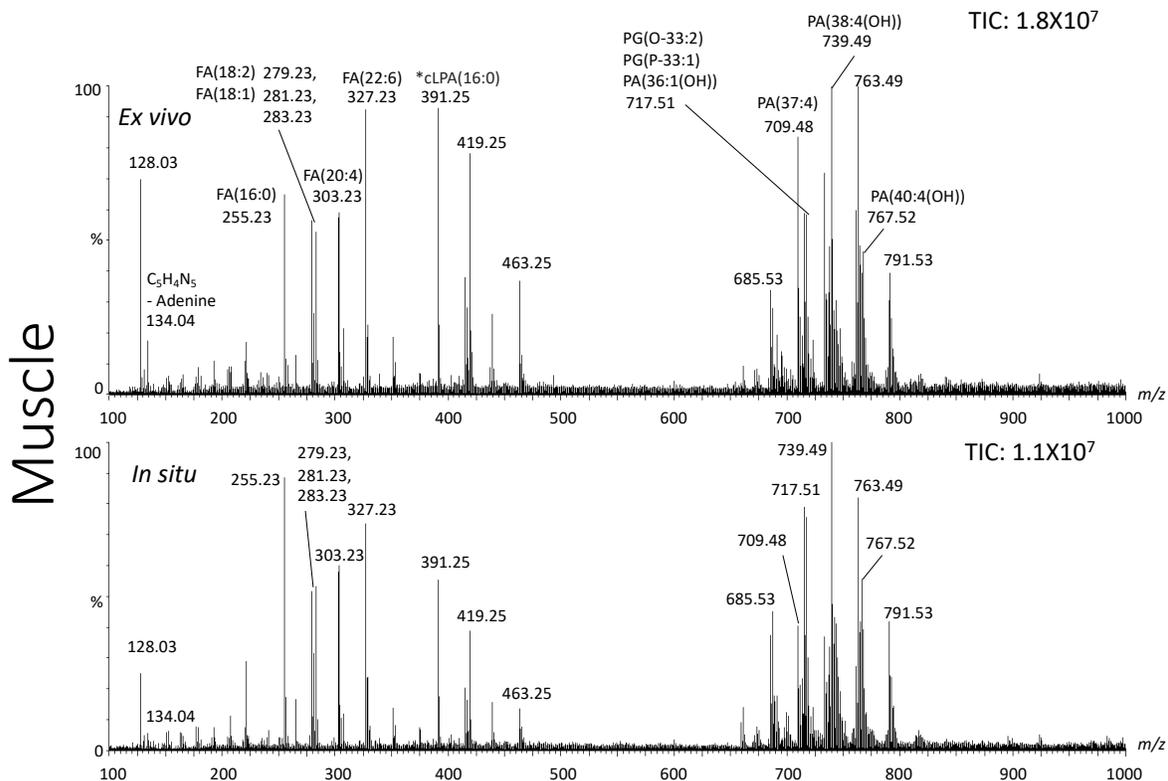
LM2-4/MDA-MB-231/Muscle						Correct Classification Rate	
	Data point groups					Per classifiable	Per all (classifiable and unclassifiable)
	PIRL-MS spectra data points	Correctly classified data points	Misclassified data points	Unclassifiable data points	Classifiable data points	data points	data points
Total	126	120	1	5	121	99.17%	95.24%
Confusion matrix							
	LM2-4	MDA	Muscle	Unclassifiable data points		Total	
LM2-4	35	0	0	1		36	
MDA-MB-231	0	41	1	3		45	
Muscle	0	0	44	1		45	
Total	35	41	45	5		126	

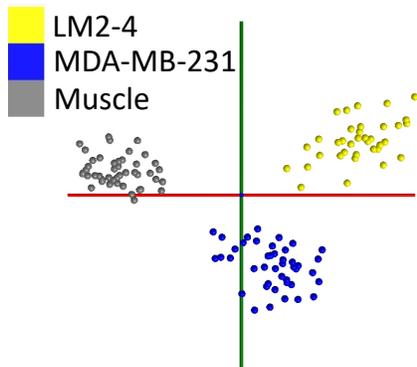
Cal-33/Muscle						Correct Classification Rate	
	Data point groups					Per classifiable	Per all (classifiable and unclassifiable)
	PIRL-MS spectra data points	Correctly classified data points	Misclassified data points	Unclassifiable data points	Classifiable data points	data points	data points
Total	62	61	0	1	61	100.00%	98.39%
Confusion matrix							
	Cal-33	Muscle	Unclassifiable data points		Total		
Cal-33	15	0	1		16		
Muscle	0	46	0		46		
Total	15	46	1		62		

**Table S1. Robustness of the *ex vivo* tissue PCA-LDA models used to classify *in situ* data from a 20% leave out and remodel test.** In this table we report confusion matrices for each of the *ex vivo* tissue PCA-LDA models listed in Table 1 as performed with Abstract Model Builder (AMX) and published previously by our group<sup>1</sup>. The total sampling events for each *ex vivo* tissue piece is listed. Unclassifiable data points refer to those that cannot be classified into either of the model groups upon the leave out test, and classified data points represent those that can be group into one of the model groups (either correctly or incorrectly in the case of misclassified points). Correct classification rates are reported per classifiable (those that classified to a model group, irrespective of whether the classification was correct or incorrect) and per all data points. As can be seen, all of the three Med8A/DAOY/Muscle, LM2-4/MDA-MB-231/muscle and Cal-33/muscle models established from 10-second PIRL-MS sampling of previously frozen *ex vivo* PIRL-MS. Some of the datapoints re-analyzed here with PCA-LDA correspond to previously published results<sup>2</sup>. We used subcutaneous xenograft models for our Table 1 assessments, thus we chose muscle tissue which alongside fat from the skin were close to the site of cancer growth as a proxy for 'healthy' tissue. This model, however, has little relevance for clinical use. Overall, little to no misclassifications in PCA-LDA *ex vivo* tissue model data was seen indicative of robustness. Fig. S2 presents the PCA-LDA scores and loading plots for these models.

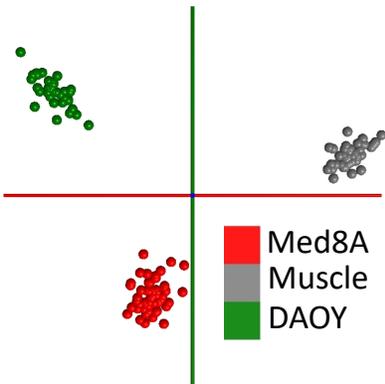
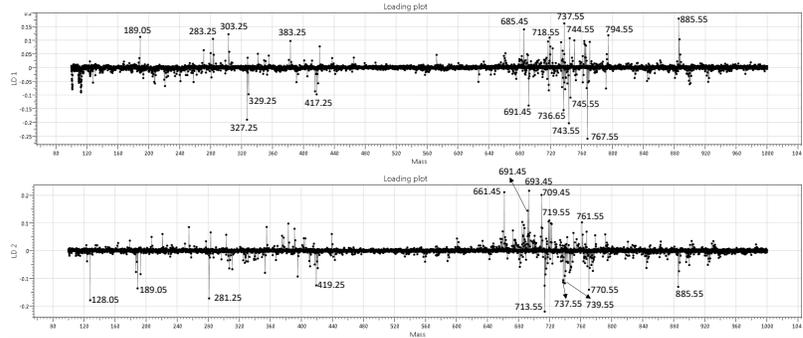




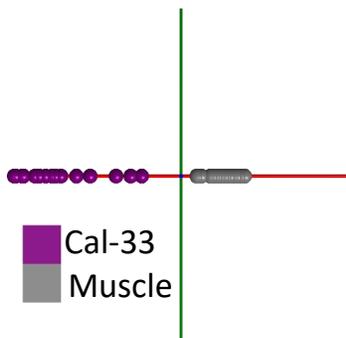
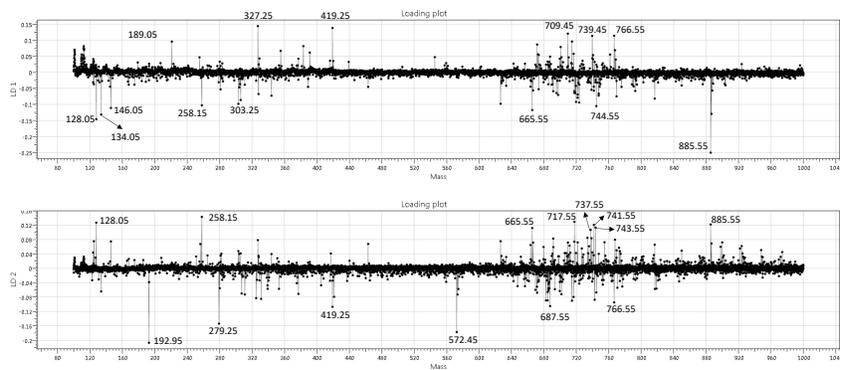
**Figure S1. Comparison of the PIRL-MS spectra of *ex vivo* and *in situ* brain cancers and muscle tissue.** In this figure we have included the PIRL-MS spectra of subcutaneous Med8A and DAOY xenograft *ex vivo* tissues previously analyzed by our group<sup>2</sup> and further compared them with *in situ* spectra obtained in this work and used for the classifications presented in Fig. 2. We have also included tentative identity assignments for some of the ions from a combination of previous works on mouse xenografts<sup>2-5</sup> as well as human tissue<sup>1</sup>. These assignments are gathered from collective HPLC-MS or HPLC-MS/MS analysis of captured laser ablation plume on high resolution Synapt instrument (Waters)<sup>1,2</sup> or extrapolated from DESI-MS analysis of same samples<sup>3</sup>, or extracts thereof<sup>4</sup>. As illustrated here, PIRL-MS spectra of *in situ* and *ex vivo* tissue are similar in terms of spectral complexity and strengths over comparable 10-15 second acquisition periods at the rate of 1 Hz (1 scan/ second). The *m/z* values only assigned from LipidMaps database<sup>6</sup> search as close match for data resolved on Xevo-TOF (Waters) are marked with (\*).



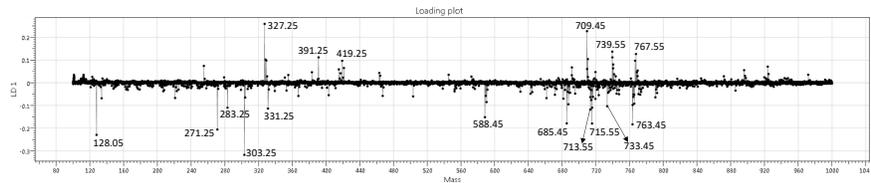
Loading plots



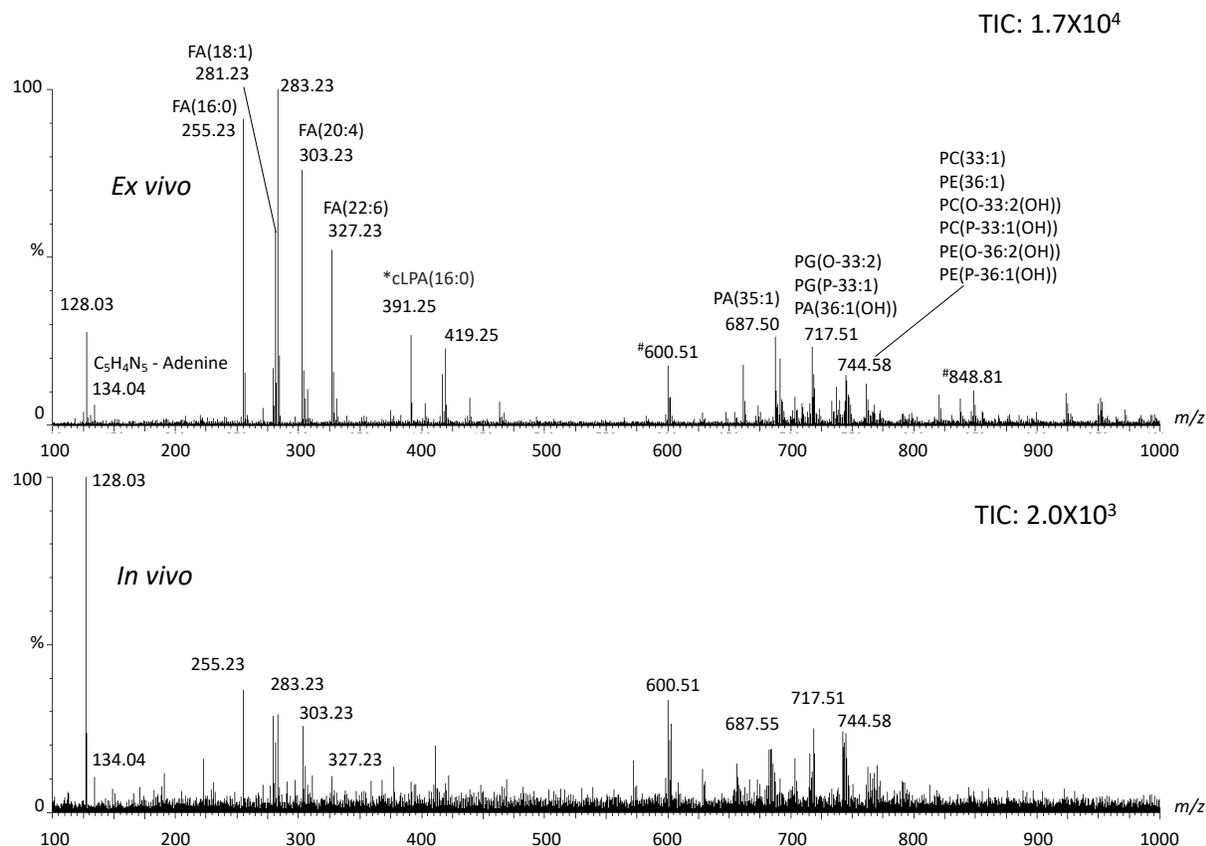
Loading plots



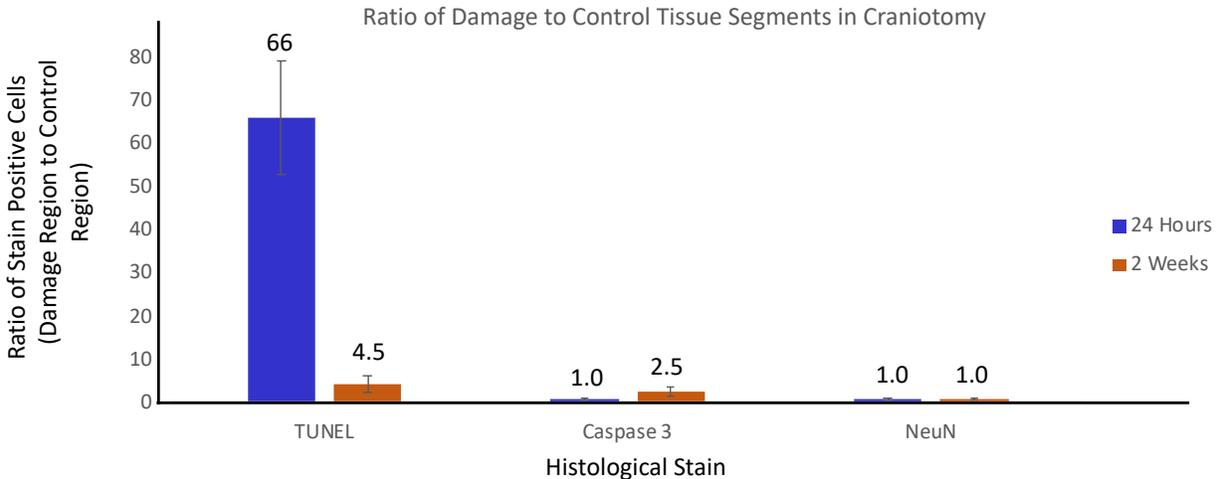
Loading plot



**Figure S2. PIRL-MS PCA-LDA models of *ex vivo* tissue used for Table 1 classification of *in situ* tissue sampling events.** This figure shows PCA-LDA scores plots as well as the loading plots of the three models Med8A/DAOY/muscle, LM2-4/MDA-MB-231/muscle and Cal-33/muscle developed from *ex vivo* tissue. Clear separation between data groups (See Table S1 for the number of data points used) is seen where each loading plot suggests contributions of many *m/z* values to the statistical discrimination used. These models are subjected to 20% leave out and remodel test in Table 1 where they have been shown to be robust. PCA-LDA modeling was performed as described in the methods section using the program AMX<sup>7</sup>. For three component models two loading plots along each of the LD is shown.



**Figure S3. PIRL-MS spectrum of *in vivo* brain tissue sampled under anesthesia.** PIRL-MS spectra of *in vivo* sampled mouse brain under anesthesia compared to a PIRL-MS spectrum of previously frozen *ex vivo* brain tissue<sup>4</sup>. This representative spectrum is classified against *ex vivo* mouse organ model<sup>4</sup> in Fig. 3. Despite presence of blood and body fluids during craniotomy, the *in vivo* spectrum is classifiable and contains similar complexity and is only slightly weaker compared to its *ex vivo* tissue counterpart. Both measurements compare a ~10-second acquisition event. The *m/z* values marked with (\*) are tentatively assigned from LipidMaps<sup>6</sup>, the rest are taken from previous works of xenograft or human brain tissue/cancer extracts analyzed on high resolution Synapt (Waters) either in form of tissue extract, captured laser ablation plume after extraction using either HPLC-MS<sup>2</sup> or HPLC-MS/MS<sup>1</sup>. The *m/z* values marked with (#) have been observed previously in PIRL-MS analysis of orthotopic murine brain cancers and the surrounding healthy mouse brain tissue<sup>5</sup> yet we do not have an identity assignment for.



**Fig. S4. Estimation of the craniotomy damage in control mice across histological stains used to examine brain tissue damage upon spatially encoded PIRL-MS sampling.** Assessment of the neuronal tissue damage by craniotomy (as control) with results evaluated using the same three TUNEL, Caspase 3 and NeuN stains (n=10 mice through the same number of 4 histological slides per stain as detailed in the main text) was performed, aimed at quantifying the extent of damage caused by the craniotomy process in the absence of laser or scalpel probe interaction with the brain tissue. The ratio of positive staining in the craniotomy only (damage) compared to control quadrant (Fig. 3C) is provided. Errors bars represent standard error of the mean and mean values are provided for each measurement. For the NeuN stain, the damage to control ratio of craniotomy insult in both immediate (24 hour) and long term (two week) assessment periods were both very close to 1 indicative of little influence on probe damage results reported in Fig. 4. For Caspase 3 stain, craniotomy insult resulted in elevated (> 1) damage ratio only in the long term period, suggesting that the estimates provided in Fig. 4 for laser or scalpel probe damage may be an upper bound compounded by the long term neuronal damage imparted by the craniotomy process itself. TUNEL stain, on the other hand, results suggest that significant DNA damage insult in immediate 24 hour period shown in Fig. 4 for the laser or scalpel probes are significantly convoluted by the damage caused by the craniotomy itself and must be interpreted with caution. Overall, findings presented in this figure further contextualize Fig. 4 results in keeping with the previous reports that PIRL sampling does not result in significant long term damage to biological tissues<sup>1,8-12</sup>. Consistent with this observation, a three-component (single factor) ANOVA test was also performed for TUNEL staining that confirms no statistical difference between craniotomy alone, laser and scalpel damage in both 24 hour and two week time points.

## Experimental methods (additional information)

### *Generation of animal model xenografts*

Xenografts of DAOY, and Med8A medulloblastoma (brain) were created as described previously under institutional authorization<sup>4</sup>. LM2-4 and MDA-MB-231 (breast) xenografts were produced as per protocol previously used by our group<sup>3,4,13</sup>. The injection protocol and amounts of cells are as described previously for both brain and breast models above, injected in the flank of SCID (breast)<sup>3,4,13</sup> and NSG (brain) models. The Cal-33 (tongue squamous cell carcinoma) as head & neck model tumors were produced in athymic nude mice (female, 20g, 6-8 weeks from Envigo) by injecting five million (5M cells) into the thigh using a 28G needle. Anesthesia induction used 5% isoflurane and maintenance was performed at 1%-2%. Mice were sacrificed by CO<sub>2</sub> asphyxiation (Cal-33, Medulloblastoma), and by cervical dislocation (breast). Tumors were then exposed surgically for sampling.

The performance of the spatially encoded *in situ* pathology platform was evaluated using 75 independent spatially encoded measurements performed on 14 mice bearing 5 different cancer cell lines as subcutaneous murine xenografts (uni/bilateral tumors in the flank). These corresponded to pathologies in the brain (n=9; 5 Med8A and 4 DAOY medulloblastoma grown bilaterally, 1 DAOY tumor failed to develop), breast (n=6; 3 LM2-4 and 3 MDA-MB-231) and head & neck (n=3 from Cal-33 tongue squamous cell carcinoma) as well as healthy tissue (n=14 from muscle in flank) with statistical models described in Table 1. The PCA-LDA models as listed in Table 1 are 3-component (DAOY, Med8A, muscle for brain; LM2-4, MDA-MB-231, muscle for breast) and 2-component (Cal-33, muscle for head & neck).

### *PIRL-MS analysis*

The description of the PIRL-MS setup (operating at 1 kHz, 240 mW at 3,000 nm with pulse duration of ~300 picoseconds, fiber diameter of 0.425 mm) is as reported previously<sup>4</sup>. principal component analysis linear discriminant analysis (PCA-LDA) on AMX<sup>7</sup> (version 1.0.1360.0) was performed using settings described previously<sup>1</sup> with the following modifications to the reported values where PCA dimensions in each of the Table 1 models were set to ~1/5<sup>th</sup> of the total number of data points in the model and LDA dimensions were set to the number of groups in the model minus one (i.e. to 2 dimensions for a 3-component model, and to 1 for a 2-component model). On average each model consisted of 40 data points per group with the exception of Cal-33 where only 16 data points comprised the cancer group of the model. The 'number of scans per spectrum' was set to 10 scans, but the 'wait for good spectrum timeout' value was kept at 12 seconds<sup>1</sup>. For our analysis, the

signal intensity threshold of  $5 \times 10^5$  was used. An exception to these setting is provided in Fig. 2A (and supplementary video file 1) where 5 seconds of averaging in continuous motion was used to provide point sampling over tissue surface. AMX (version 1.0.1360.0), previously used for 10-second classification of cancer types by our group<sup>1</sup>, contains two modules; a 'model building' module where principal component analysis linear discriminant analysis (PCA-LDA) modeling of the classifiable groups are made; and a 'recognition module' where real-time PCA-LDA decomposition of the mass spectra of a sampling event and comparison thereof to the model results produces a classifier<sup>7</sup>. For example, in a 'two-component' PCA-LDA model of cancer versus healthy, real-time PCA-LDA comparison of the PIRL-MS spectrum of an unknown specimen with those of the model will result in three potential classifiers of healthy, cancer or outlier (i.e. unclassifiable). The latter classifier is produced when within a given confidence interval no appreciable similarity to the model data points is detected. Model robustness was examined using AMX and a 20% leave out test as completed previously<sup>1</sup>. The same platform was used to generate the loading plots for PCA-LDA assessments.

### *Craniotomy and histopathologic analysis*

32 black-6 mice (Charles River, Canada) were divided into three cohorts of 'laser', 'scalpel' and 'control' (craniotomy only) and subjected to craniotomy under isoflurane anesthesia (maintained at 2% with medical air) with stereotactic immobilization (Fig. 3B). A  $\sim 3 \text{ mm}^2$  opening in the skull was created as illustrated in said figure through creating perforations followed with the removal of the skull section using surgical tweezers. Once the skull was removed, the PIRL or scalpel 'incisions' were superficially made on the surface of the cerebral hemispheres in the frontal lobe area (affected brain surface area per laser/scalpel method of  $\sim 1 \text{ mm}^3$ , in the form of a straight line 2 mm long,  $\sim 1 \text{ mm}$  wide, and  $\sim 0.5 \text{ mm}$  deep). PIRL-MS spectra were collected from the brain of 8 mice as per previous setup used for *ex vivo* tissue examinations<sup>1,2</sup>. The overlaying skin was then closed with surgical glue. Mice were monitored closely post surgery for behavioral changes and analgesic was administered as approved by institutional animal used protocol (AUP 4135.10, University Health Network, Toronto, Canada). No behavioral changes were noticed and only three mice did not survive the first 24 hours post surgery (2 for the laser group, and 1 control). At end points (24 hours and two weeks) mice were sacrificed with  $\text{CO}_2$  asphyxiation, and subjected to dissection to remove the brains. The brains were transferred to 10% formaldehyde for 72 hours and then exchanged for 70% ethanol. In total, the extracted brains of 29 mice were subjected to sectioning and staining after embedding as detailed below. Brains were laterally halved across the craniotomy region and 4 histologic ( $4 \mu\text{m}$  each) slides were cut for each stain of TUNEL (for DNA damage), Caspase-3 (for cell death) and NeuN (for neuronal viability), totaling 12 histologic sections per mouse, resulting in a grand total of 348 sections. Immunohistochemistry using the above stains were performed using standard protocols from manufacture after deparaffinization (TUNEL, Roche, after Proteinase K treatment; NeuN, Cell Signaling (Cat. # 24307), 1:400 dilution in Dako with

secondary biotinylated anti-rabbit Ig, Vector Labs; Caspase-3, Cell Signaling (Cat. # 9661), 1:200 dilution in Dako with secondary goat-anti-rabbit Ig antibody 1:200, Vector Labs. 3,3'-Diaminobenzidine (DAB) was purchased from Abcam. All primary antibodies were reactive to mouse proteins.

Stained slides were digitized at 20x magnification and imported for image segmentation and cell classification into Definiens TissueStudio 4.3 (Definiens Inc, Munich Germany). In brief: tissue was distinguished from background using an automated threshold, with the entire brain section detected. The brain region of interest was divided into quadrants manually, using morphological / structural landmarks (See Fig 3C). Following stain separation to isolate the hematoxylin from 3,3'-Diaminobenzidine (DAB) signal cell segmentation was performed on each quadrant, utilizing the hematoxylin signal to identify the locations of the nuclei. A threshold was determined manually to separate positive from negative TUNEL, Caspase-3 and NeuN DAB signal. The number of positive cells detected within each quadrant was summed and reported as a percentage of all cells (positive and negative) detected in that quadrant, was reported. Standard error of the mean is reported, and statistical significance (compared to control; marked with asterisk on Fig. 4) was calculated using two-sample student's t-test ( $p < 0.01$ ).

## Determination of positional accuracy

### *Projected pixel errors from the camera*

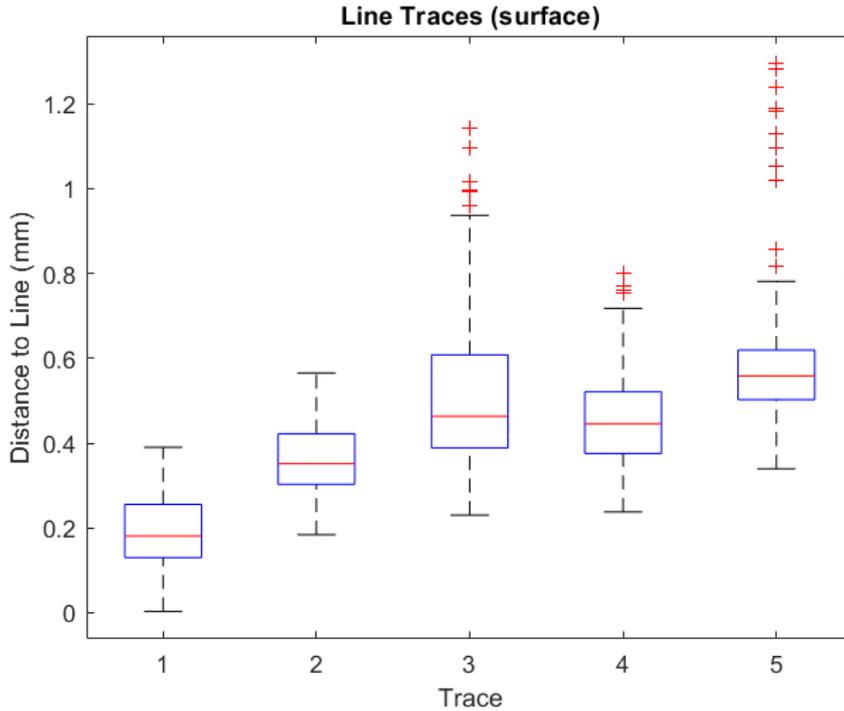
Estimation of the Logitech C920 camera intrinsic and extrinsic parameters were computed using the Camera Calibration Toolbox for MATLAB and GTxEyes software using 11 checkerboard images captured with optical tracking data (position and orientation of IR sensor attached to the Logitech C920 and corners of the checkerboard picked by a tracked pointer). Post calibration, the mean translation error between the IR sensor and camera center is 0.48 mm with a pixel re-projection error of 0.20 for each independent run.

### *Determination of line tracing uncertainties*

A pre-calibrated pointer was used to define a planar surface using 5 picked points, 2 of which formed a straight 5 cm line on the plane. Two sets of line tracing measurements were performed following the line, on the plane surface in two modes of (1) contact mode where pointer tip touched the surface and was supported by the surface, and (2) above the surface in air, hand held ~2 mm away from the surface. Each set contained 5 traces. The tables and box plots of the line tracing attempts are provided below for each mode.

<b>Trace # (Contact mode)</b>	<b>Samples</b>	<b>Mean Distance to Line (mm)</b>	<b>Standard Deviation (mm)</b>
1	224	0.19	0.08
2	151	0.36	0.08
3	143	0.53	0.21
4	136	0.46	0.12
5	139	0.60	0.19

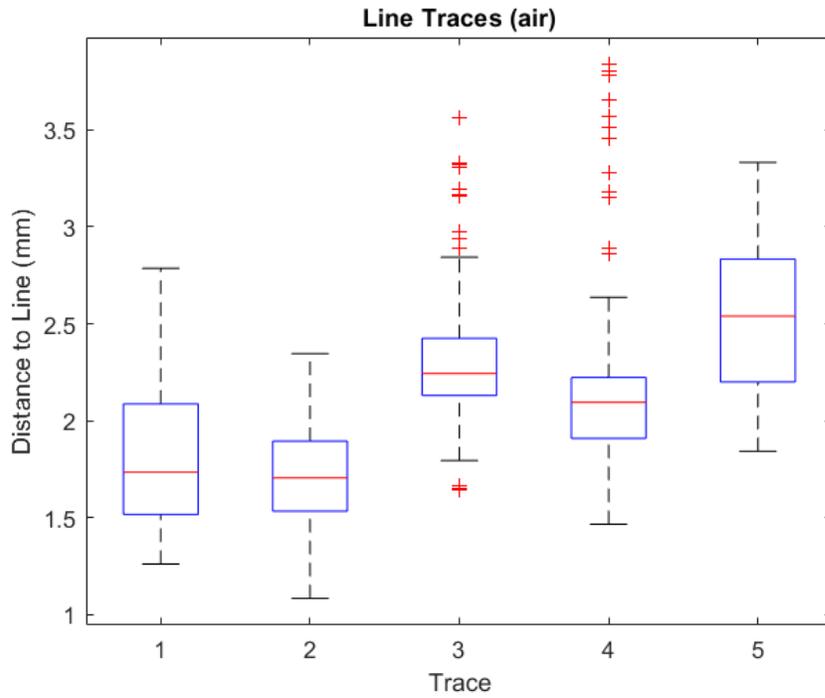
Line trace measurements using pre-calibrated pointer on surface of the plane in the contact mode with pointer tip being supported by the surface. Standard deviation values are provided for each measurement.



Box plot illustrating distance to the traced line for supported contact mode for each tracing attempt.

<b>Trace # Hand-held mode</b>	<b>Samples</b>	<b>Mean Distance to Line (mm)</b>	<b>Standard Deviation (mm)</b>
1	146	1.82	0.39
2	124	1.71	0.28
3	121	2.33	0.37
4	119	2.18	0.54
5	121	2.55	0.39

Line trace measurements using pre-calibrated pointer held without support ~2 mm off above the surface of the plane. Standard deviation values are provided for each measurement.



Box plot illustrating distance to the traced line for non-supported virtual pointer mode, hand held ~ 2mm above the surface of the plane for each tracing attempt.

## References

- 1 Woolman, M. *et al.* Picosecond Infrared Laser Desorption Mass Spectrometry Identifies Medulloblastoma Subgroups on Intrasurgical Timescales. *Cancer Res* **79**, 2426-2434, doi:10.1158/0008-5472.CAN-18-3411 (2019).
- 2 Woolman, M. *et al.* Rapid determination of medulloblastoma subgroup affiliation with mass spectrometry using a handheld picosecond infrared laser desorption probe. *Chem Sci* **8**, 6508-6519, doi:10.1039/c7sc01974b (2017).
- 3 Tata, A. *et al.* Rapid Detection of Necrosis in Breast Cancer with Desorption Electrospray Ionization Mass Spectrometry. *Sci Rep* **6**, 35374, doi:10.1038/srep35374 (2016).
- 4 Woolman, M. *et al.* Optimized Mass Spectrometry Analysis Workflow with Polarimetric Guidance for ex vivo and in situ Sampling of Biological Tissues. *Sci Rep* **7**, 468, doi:10.1038/s41598-017-00272-y (2017).
- 5 Katz, L. *et al.* Dual Laser and Desorption Electrospray Ionization Mass Spectrometry Imaging Using the Same Interface. *Anal Chem* **92**, 6349-6357, doi:10.1021/acs.analchem.9b05340 (2020).
- 6 O'Donnell, V. B., Dennis, E. A., Wakelam, M. J. O. & Subramaniam, S. LIPID MAPS: Serving the next generation of lipid researchers with tools, resources, data, and training. *Sci Signal* **12**, doi:10.1126/scisignal.aaw2964 (2019).
- 7 Bodai, Z. *et al.* Effect of Electrode Geometry on the Classification Performance of Rapid Evaporative Ionization Mass Spectrometric (REIMS) Bacterial Identification. *J Am Soc Mass Spectrom* **29**, 26-33, doi:10.1007/s13361-017-1818-5 (2018).
- 8 Amini-Nik, S. *et al.* Ultrafast mid-IR laser scalpel: protein signals of the fundamental limits to minimally invasive surgery. *PLoS One* **5**, doi:10.1371/journal.pone.0013053 (2010).
- 9 Hess, M. *et al.* Picosecond infrared laser (PIRL): an ideal phonomicrosurgical laser? *Eur Arch Otorhinolaryngol* **270**, 2927-2937, doi:10.1007/s00405-013-2561-6 (2013).
- 10 Jowett, N. *et al.* Heat generation during ablation of porcine skin with erbium:YAG laser vs a novel picosecond infrared laser. *JAMA Otolaryngol Head Neck Surg* **139**, 828-833, doi:10.1001/jamaoto.2013.3974 (2013).
- 11 Linke, S. J. *et al.* [Perspectives of laser-assisted keratoplasty: current overview and first preliminary results with the picosecond infrared laser ( $\lambda = 3$  microm)]. *Ophthalmologie* **111**, 523-530, doi:10.1007/s00347-013-2995-7 (2014).
- 12 Jowett, N. *et al.* Bone ablation without thermal or acoustic mechanical injury via a novel picosecond infrared laser (PIRL). *Otolaryngol Head Neck Surg* **150**, 385-393, doi:10.1177/0194599813517213 (2014).
- 13 Tata, A. *et al.* Wide-field tissue polarimetry allows efficient localized mass spectrometry imaging of biological tissues. *Chem Sci* **7**, 2162-2169, doi:10.1039/c5sc03782d (2016).