Image fusion of mass spectrometry and microscopy: a multimodality paradigm for molecular tissue mapping

Raf Van de Plas[1–3], Junhai Yang[1,2], Jeffrey Spraggins[1,2] & Richard M Caprioli[1,2,4–6]

We describe a predictive imaging modality created by 'fusing' two distinct technologies: imaging mass spectrometry (IMS) and microscopy. IMS-generated molecular maps, rich in chemical information but having coarse spatial resolution, are combined with optical microscopy maps, which have relatively low chemical specificity but high spatial information. The resulting images combine the advantages of both technologies, enabling prediction of a molecular distribution both at high spatial resolution and with high chemical specificity. Multivariate regression is used to model variables in one technology, using variables from the other technology. We demonstrate the potential of image fusion through several applications: (i) 'sharpening' of IMS images, which uses microscopy measurements to predict ion distributions at a spatial resolution that exceeds that of measured ion images by ten times or more; (ii) prediction of ion distributions in tissue areas that were not measured by IMS; and (iii) enrichment of biological signals and attenuation of instrumental artifacts, revealing insights not easily extracted from either microscopy or IMS individually.

Biology and medicine are experiencing an unprecedented level of information acquisition, fueled by technological advancements that deliver enormous amounts of data with ever-increasing instrumental precision. As a result, the integration of information across different data types is a crucial challenge. This is apparent in the molecular imaging field, where a progressively heterogeneous set of imaging modalities and sensor types delivers a wide range of information on the molecular processes taking place in living cells[1,2]. Each imaging technology has its own molecular targets, advantages and constraints. Although the use of multiple imaging modalities toward answering a single biomedical question is not uncommon[3–5], most multimodal studies treat their different image types as separate entities. Different modalities are commonly registered and overlaid to generate a single display, but true integration of data across technologies is largely left to human interpretation. Even though the potential of multimodal integration is well recognized, much work has focused on developing instrumental[6–14] and chemical answers[15,16], whereas broad computational approaches capable of handling the heterogeneity and multiresolution challenges have been largely lacking. As a result, biological insights can be segregated along technological borders, and important structural information may be overlooked.

To help resolve this problem, we employ the concept of image fusion[17,18] and demonstrate the power of cross-modality modeling on tissue samples, with data obtained by mass spectrometry and microscopy (**Fig. 1**). Image fusion refers to the generation of a single image from several source images, and it typically aims to provide a more accurate description of the sample or combine information toward a particular human or machine perception task. Substantial development in recent years, particularly in multisensor image fusion in which the source images are of different sensor types[19,20], has led to applications in fields as diverse as satellite-based remote sensing[21], clinical diagnostics[22–24] and concealed-weapon detection[25].

We develop and extend the concept of multisensor image fusion to study protein, peptide, lipid, small-metabolite and drug distributions in tissue. We introduce a two-part method that first models detectable relationships between the source modalities and then uses that model for prediction (**Supplementary Fig. 1**). Matrix-assisted laser desorption/ionization (MALDI) IMS[26–30] is employed as a source modality that delivers chemically specific information over a wide mass range, beyond 100 kDa, and that is applicable to a variety of biomolecules in living cells and tissues[31–35]. It is particularly well suited for image fusion because it is information rich, mapping the spatial distributions of many hundreds of biomolecules throughout a tissue section. Optical microscopy of stained tissue is used as the second source modality, providing the fine-grained textural information that IMS typically does not supply. Our computational approach is broadly applicable and can perform fusion with other imaging technologies as well (for example, MRI, CT, PET and others).

Images of the same subject, acquired using different modalities, often exhibit correspondences. In a brain study, for example, one might see the corpus callosum outlined both in a microscopy image and in an ion image obtained through IMS.

[1]Mass Spectrometry Research Center, Vanderbilt University, Nashville, Tennessee, USA. [2]Department of Biochemistry, Vanderbilt University, Nashville, Tennessee, USA. [3]Delft Center for Systems and Control, Delft University of Technology, Delft, the Netherlands. [4]Department of Chemistry, Vanderbilt University, Nashville, Tennessee, USA. [5]Department of Pharmacology, Vanderbilt University, Nashville, Tennessee, USA. [6]Department of Medicine, Vanderbilt University, Nashville, Tennessee, USA. Correspondence should be addressed to R.M.C. (richard.m.caprioli@vanderbilt.edu).

RECEIVED 5 MARCH 2014; ACCEPTED 8 DECEMBER 2014; PUBLISHED ONLINE 23 FEBRUARY 2015; DOI:10.1038/NMETH.3296

Typically, microscopy will deliver a spatially fine-grained outline of the corpus callosum without telling us much about its chemical content, whereas IMS will characterize that same tissue structure in a chemically very specific but spatially coarse or pixelated manner (**Fig. 1**). Although these technologies sample the same biological structure at different spatial resolutions, the visual trace present in both modalities indicates a correspondence between the IMS and microscopy measurements for this region. Such correspondences that tie observations in one modality to observations in another modality are used in image fusion for integrative and predictive purposes. Some cross-modality correspondences (as in the example above) are straightforward, can be described by a simple model (for example, correlation) and are abundant enough to be visually recognized across image types. Most correspondences, however, tend to be more nuanced and complex. They typically involve derivatives of the variables natively provided by a sensor, often require advanced multivariate models to be described accurately and usually fall below the level of detection for human eyes.

We integrate IMS and microscopy by capturing IMS-microscopy relationships in a model. As this enables prediction of observations in one modality on the basis of measurements in the other modality, the model sets the stage for a multitude of predictive fusion applications. One example is an upsampling application, known in remote sensing as sharpening[36–38], which can predict the distribution of ion m/z 778.5 (identified as PE(P-40:4)) to a spatial resolution that exceeds that of the native IMS measurements by tenfold (**Fig. 2**). We demonstrate IMS-microscopy fusion for different tissue types, different target molecules and various histological staining protocols, and at different scales. In the latter case, we generate ion predictions at the nanometer range, below the resolution achievable with current MALDI IMS instrumentation.

RESULTS

We first report the fundamentals of our image fusion method and then demonstrate IMS-microscopy fusion with three different applications: (i) predicting ion distributions to a spatial resolution that exceeds that of the measured ion images (sharpening); (ii) predicting ion distributions in tissue areas that were not measured by IMS (out-of-sample prediction); and (iii) discovery of biological patterns that are difficult to retrieve from the source modalities separately (enrichment). In total we cover seven distinct case studies (**Supplementary Table 1**), each describing a multimodal tissue imaging experiment with microscopy and MALDI IMS from the same or an adjacent tissue section.

Image fusion method

We approached the challenge of modeling across modalities as a massive multivariate regression task[39] involving variables derived from IMS measurements and variables derived from microscopy measurements (**Supplementary Fig. 2**). Our goal was to link each ion intensity variable to a combination of photon-based variables by modeling the distribution of spatially paired measurements from both images using partial least-squares (PLS) regression[40] (**Supplementary Fig. 3**). The resulting model is a set of slopes and intercepts that, when combined with a microscopy measurement, outputs a prediction for the IMS variables. In the sharpening application, for example, the model is applied to each microscopy pixel, effectively predicting ion intensities at spatial resolutions that exceed the native resolution of the IMS measurements.

The two-phase nature of the image fusion method consists of a generic multimodal modeling phase followed by an application-specific prediction phase (Online Methods). Each of these phases consists of several modular substeps (**Supplementary Figs. 2** and **3** and **Supplementary Notes 1–3**). **Supplementary Note 4** enumerates assumptions that are specific to the applications in which image fusion is employed. The performance of image fusion-driven prediction, and the performance indicators we have developed to accompany such predictions, are examined in the Discussion and in **Supplementary Note 2**. In the **Supplementary Discussion**, we address the concept of prediction beyond the IMS measurement resolution, and we discuss how image fusion uses measurements from another modality to outperform methods that have no other-modality information to drive their predictions. The **Supplementary Discussion** further highlights how the empirical nature of our image fusion method, namely mining for cross-modality relationships rather than predefining them, ensures broad applicability to imaging modalities and sensor types beyond the data sources illustrated here. We also briefly discuss how the fusion process enables bidirectional communication and corroboration between different data sources and how fusion with IMS can contribute to microscopy interpretation and the pathological recognition process.

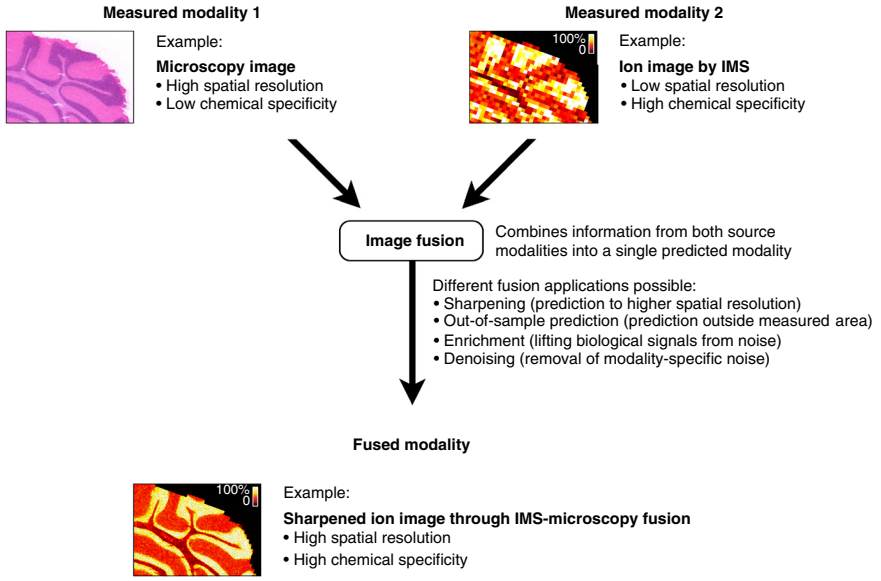


Figure 1 | Concept of image fusion of imaging mass spectrometry (IMS) and microscopy. Image fusion generates a single image from two or more source images, combining the advantages of the different sensor types. The integration of IMS and optical microscopy is given as an example. The IMS-microscopy fusion image is a predictive modality that delivers both the chemical specificity of IMS and the spatial resolution of microscopy in one integrated whole. Each source image measures a different aspect of the content of a tissue sample. The fused image predicts the tissue content as if all aspects were observed concurrently.

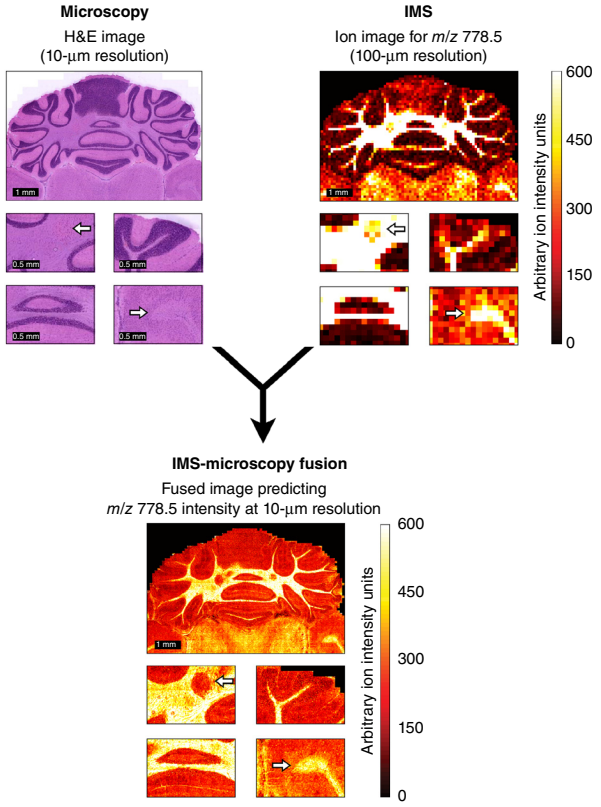


Figure 2 | Example of IMS-microscopy fusion. An ion image measured in mouse brain, describing the distribution of *m/z* 778.5 (identified as lipid PE(P-40:4)) at 100-µm spatial resolution (top right), is integrated with an H&E microscopy image measured from the same tissue sample at 10-µm resolution (top left). By combining the information from both image types, the image fusion process can predict the ion distribution of *m/z* 778.5 at 10-µm resolution (bottom).

To enable the readers to experiment on their own data sets, we provide a full implementation of the image fusion framework as a command line utility available for download at http://fusion.vueinnovations.com/. This package includes an example IMS-microscopy multimodal data set.

Sharpening ion distributions by microscopy

The sharpening of ion images provides a means of predicting molecular tissue content to a higher spatial resolution and serves as a concrete case to introduce the general modeling procedure that underlies all our fusion applications. In case study 1, the tissue distribution of ion *m/z* 762.5 (identified as PE(16:0/22:6)) was measured through an IMS experiment performed at 100-µm spatial resolution (**Fig. 3a**). A microscopy image at 10-µm resolution was taken from the same tissue section stained with hematoxylin and eosin (H&E) after the IMS acquisition. Our two-phase fusion method (Online Methods and **Supplementary Figs. 1–3**) combined the information from both modalities, built a model to capture any cross-modality relationships it could detect, and employed that model and the microscopy measurements to predict the distribution for all ions at 10-µm resolution. The spatial prediction for ion *m/z* 762.5 is shown (**Fig. 3a**). Additionally, we compared the 100-µm measurement and the fusion-based 10-µm prediction to an ion image for *m/z* 762.5 actually acquired at 10 µm on an adjacent tissue section (**Fig. 3b**). The IMS acquisition parameters for the measured 100-µm and 10-µm ion images were identical except for the increased laser power required to compensate for reduced signal intensities at 10-µm pixel widths. An example by an alternative upsampling approach, bilinear interpolation (**Supplementary Fig. 4**), was clearly inferior. (See Online Methods and **Supplementary Note 2** for details on the model-building and evaluation process.)

The IMS-microscopy model obtained by PLS regression between IMS and microscopy-derived variables contains a submodel for each IMS variable (**Supplementary Fig. 5**). Model performance is therefore specific to each IMS variable and needs to be evaluated to ascertain for which ion peaks good prediction (and useful fusion) is possible. We introduce a 'reconstruction score' as a measure of predictive power (**Supplementary Fig. 6**). This percentage score summarizes how well the measured ion distribution can be predicted using the given microscopy observations. A value of 100% indicates that the cross-modality relationships allow complete reconstruction of the measured intensity distribution using variables from the other technology only. A value close to 0% signifies that no cross-modality relationship could be found or modeled for this ion and that fusion-driven prediction is not an option for it. Most measured ion images can be predicted at least partially using microscopy, and in our examples we relied on fusion only for ions with a score above 75%. We also introduce the absolute residuals image and the 95% confidence interval (CI) image as measures of prediction performance in function of tissue location (**Supplementary Figs. 6–9**). Whereas *m/z* 762.5

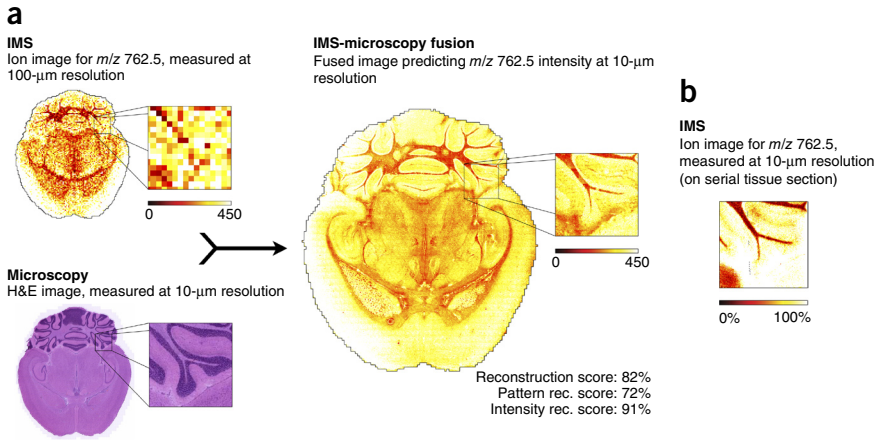


Figure 3 | Prediction of the ion distribution of *m/z* 762.5 in mouse brain at 10-µm resolution from 100-µm IMS and 10-µm microscopy measurements (sharpening). (**a**) This example in mouse brain fuses a measured ion image for *m/z* 762.5 (identified as lipid PE(16:0/22:6)) at 100-µm spatial resolution with a measured H&E microscopy image at 10-µm resolution, predicting the ion distribution of *m/z* 762.5 at 10-µm resolution. Reconstruction (rec.) scores are shown. Color scales encode arbitrary ion intensity units. (**b**) For comparison, a measured ion image for *m/z* 762.5 at 10-µm spatial resolution is shown, acquired from a neighboring tissue section.

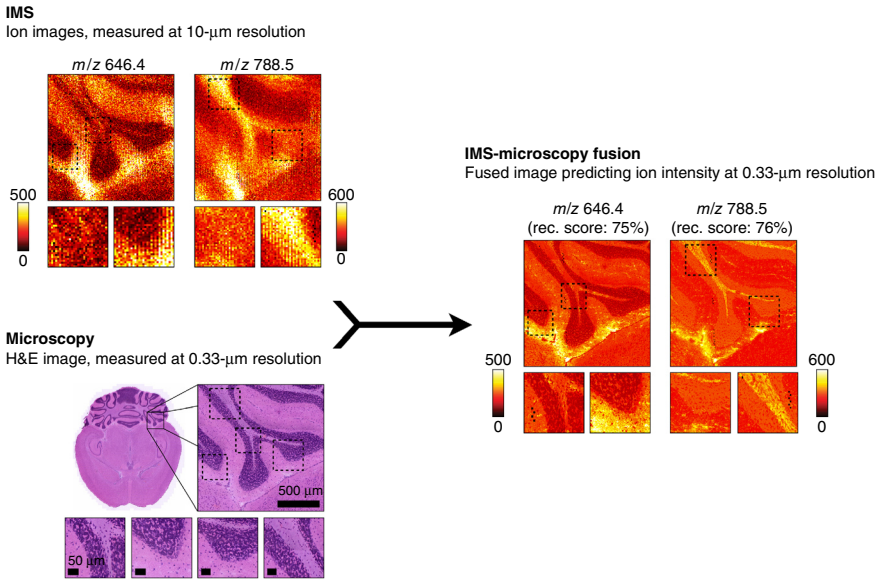
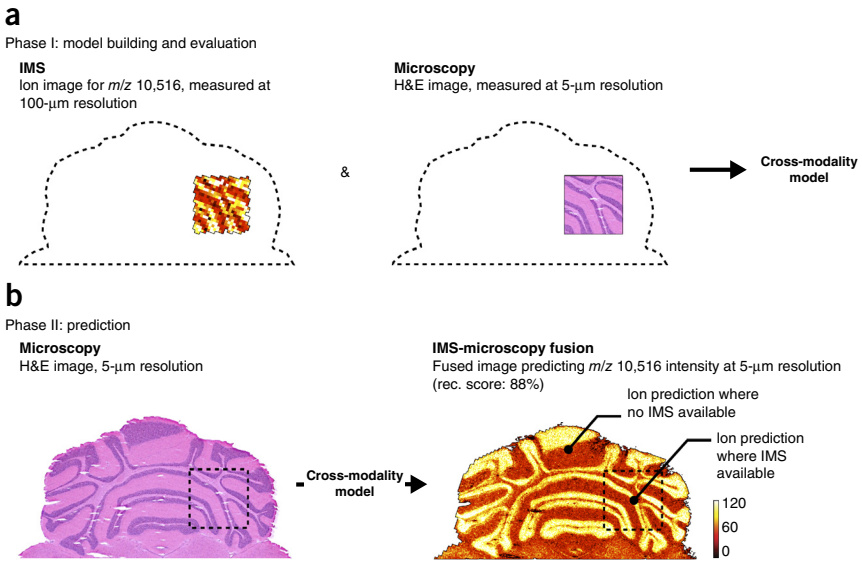


Figure 4 | Prediction of the ion distributions of *m/z* 646.4 and 788.5 in mouse brain at 330-nm resolution from 10-µm IMS and 330-nm microscopy measurements (sharpening). Measured ion images acquired in mouse brain for *m/z* 646.4 and *m/z* 788.5 at 10-µm spatial resolution are fused with an H&E microscopy image measured at 0.33-µm resolution. The resulting IMS-microscopy model is combined with the microscopy measurements to predict the ion distributions of *m/z* 646.4 and *m/z* 788.5 at 330-nm resolution with overall reconstruction scores of, respectively, 75% and 76%. Color scales encode arbitrary ion intensity units.

(**Supplementary Fig. 6**) and 747.5 (a combination of three nominally isobaric lipids; **Supplementary Fig. 9**) exhibited a strong cross-modality relationship to H&E microscopy, *m/z* 766.5 (PE(18:0/20:4)) and 715.6 (PE-Cer(d16:1/22:0)) from the same IMS experiment are examples of ions for which no such relationships could be found and thus prediction was not recommended (**Supplementary Figs. 7** and **8**).

The access to information from different technologies allows the fusion method to predict beyond hard constraints inherent to a particular sensor type. This advantage, for example, enables the image sharpening application to push beyond the physical limits of the IMS laser by predicting at a spatial resolution below the wavelength of the laser. As an example, we applied the fusion procedure to a mouse brain sample on which both source modalities were pushed to their current state of the art in spatial resolution for the instrumentation used in this experiment (**Fig. 4**). Lipid distributions for *m/z* 646.4 (nominal isobars Cer(d18:1/24:1) and second ^{13}C isotope of CerP(d18:1/18:0)) and 788.5 (nominal isobars PS(18:0/18:1) and PE(40:7)) were successfully predicted with 75% and 76% scores, respectively, at 330-nm resolution, using a 355-nm wavelength laser to acquire the source IMS measurements. The ion images were acquired at 10-µm resolution; the textural information was encoded at 330-nm resolution in an H&E microscopy image of an adjacent section (**Fig. 4**).

We examined other aspects of image fusion through additional examples (**Supplementary Results** and **Supplementary Table 1**), using case study 1 (**Supplementary Figs. 1–11**), case study 2 (**Supplementary Fig. 12**), case study 4 (**Supplementary Figs. 13–16**), case study 5 (**Supplementary Fig. 17**), case study 6 (**Supplementary Figs. 18–20**) and case study 7 (**Supplementary Fig. 21**). Image fusion was able to use any spatial resolution between the native IMS and microscopy resolutions as its target for prediction (**Supplementary Figs. 10**, **11** and **19**). It worked independently of the tissue type or the molecule type measured (**Supplementary Figs. 12** and **18**) and could be used with tissue stains other than H&E (**Supplementary Figs. 13–17**). Fusion provided useful integration and prediction even for ion peaks that reported multiple unresolved ion species (**Supplementary Fig. 9**). Finally, using a synthetic multimodal data set, we verified the method against a gold standard of known cross-modal and modality-specific patterns (**Supplementary Fig. 17**).



Predicting molecule distributions in non-IMS-measured tissue

Image fusion makes it possible to predict ion distributions in areas not measured by IMS; this is achievable using a model built from areas in which data from both modalities are available. As an example, we predicted the ion distribution for *m/z*

Figure 5 | Prediction of *m/z* 10,516 distribution in mouse brain areas not measured by IMS (out-of-sample prediction). (**a**) An IMS-microscopy model is built on a tissue subarea for which IMS is available at 100-µm resolution and an H&E microscopy image is available at 5-µm resolution. (**b**) The model is then used to predict the distribution of *m/z* 10,516 in areas where no IMS was acquired and only microscopy is available. Color scales encode arbitrary ion intensity units. (A non-sharpened version is available in **Supplementary Fig. 21**.)

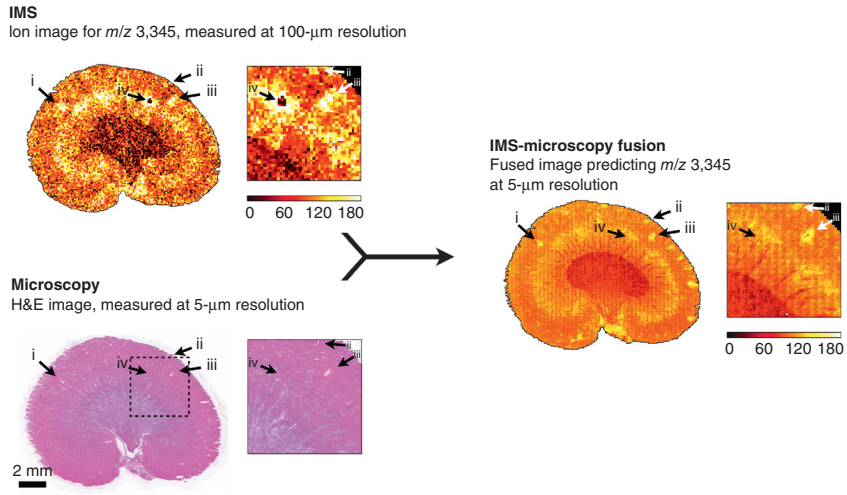


Figure 6 | Discovery of tissue features through multimodal enrichment. An ion image for *m/z* 3,345 measured by IMS at 100-µm resolution in a rat kidney section is fused with an H&E microscopy image acquired at 5-µm resolution to produce an ion distribution prediction at 5-µm resolution (reconstruction score: 85%). Annotations i–iii demonstrate multimodal enrichment. Their successful propagation through the fusion process and their presence in the final fused image confirms that they are genuine tissue features that are corroborated by another technology (in this case, microscopy). Annotation iv demonstrates multimodal attenuation. The lack of cross-modal support for this localized drop in ion intensity reduces confidence in the biological nature of this feature and at least labels it as an IMS-specific observation. Color scales encode arbitrary ion intensity units.

10,516 in non-IMS-measured mouse brain areas (**Fig. 5**). First, an IMS-microscopy model was built on a subarea of the tissue where IMS measurements (in the molecular weight range of interest) were available at 100-µm resolution and H&E-stained tissue measurements were available at 5-µm resolution. Subsequently, the model was used to predict both inside the IMS-measured area and the area outside, where only microscopy data were available. The predicted pattern for *m/z* 10,516 outside the modeled area was successfully confirmed via IMS measurements external to this case study. Although prediction at the native IMS resolution of 100 µm was possible (**Supplementary Fig. 21**), the availability of microscopy at 5-µm resolution enabled additional sharpening of the ion predictions to 5 µm both inside and outside the modeled area (**Fig. 5**).

Discovery through multimodal enrichment

The discovery of cross-modality relationships enables the separation of a measured tissue signal into a part that can be predicted by the other technology and a part that is modality specific. For measured variables and tissue areas with strong cross-modal support (for example, variables with high reconstruction scores and tissue areas marked by low values in the absolute residuals image), fusion thus enables enrichment of genuine tissue signal and removal of modality-specific noise.

Multimodal enrichment can cause a feature to be discovered or recognized as biological or sample based in cases when in a single-modality analysis it would be mistaken for noise or go unnoticed. We demonstrated this in a sharpened ion image of rat kidney (**Fig. 6**). The propagation of features (i–iii) through the fusion process into the predictions indicated that these patterns were corroborated by measurements in both modalities. If only the ion image were considered, these features could be mistaken for 'matrix hotspot' noise (a type of noise specific to IMS, no matrix involved in microscopy). However, the presence of the features in the final fused image meant that they were corroborated by measurements from another technology (in this case microscopy). As technology-specific noise rarely correlates across different measurement principles, a multimodal corroboration of a pattern through fusion tends to confirm that the pattern is not noise but rather a genuine tissue feature that influenced the measurements of both sensor types. If only the microscopy image were considered, those particular features might be so faint that they would go undetected. However, when the microscopy measurements are combined with another data source (in this case IMS), microscopy patterns reporting these features are lifted from almost invisible intensity variations to clearly distinguishing contrast levels. Overall, from a microscopy viewpoint, fusion enables discovery of patterns that may otherwise go undetected. From an IMS viewpoint, fusion provides an increased confidence in the genuine nature of these features so that they may be recognized from faint signals in the noise.

Fusion separates features with support across technologies from those supported by only a single sensor type. One annotation (iv; **Fig. 6**) is an example of a single-modality feature and demonstrated multimodal attenuation. In IMS, feature iv showed a localized intensity drop, which, when we consider only the ion image, might be perceived as biological because it covered multiple pixels. However, in the microscopy image (acquired post-IMS) there was no visual trace at this location, and even the fusion process found no cross-modal evidence to support this drop as a real tissue feature. Although lack of cross-modal corroboration does not necessarily mean a feature is noise—a biological feature detectable by only one of the technologies is still possible—it does reduce confidence in the feature being genuine and suggests here an IMS acquisition anomaly.

Fusion-based filtering can also be used as a global method for denoising ion images with strong cross-modal support. As an example, we fused an ion image for *m/z* 5,666 measured at 100-µm resolution in mouse brain with an H&E microscopy image at 5-µm resolution (**Supplementary Fig. 20**). The predicted ion distribution showed a clear decrease in speckled noise variation. A similar cleanup effect and increase in signal-to-noise ratio (SNR) was also observed in predictions for other ions from this data set (**Supplementary Figs. 18** and **19**).

DISCUSSION

The image fusion process combines different image types, each measuring a different aspect of the content of a tissue sample, and predicts the tissue content in an integrated way as if all aspects were observed concurrently. However, it is important to

have those predictions accompanied by indicators of prediction performance so that the fidelity of the fusion results can be properly assessed. In our image fusion method, the modeling between technologies is addressed as a multivariate regression analysis in which IMS-derived variables (peak intensities) are designated response variables and microscopy-derived variables (for example, hue, texture and entropy) are used as predictor variables. The resulting mathematical model attempts to approximate each ion-peak distribution in the IMS source as a linear combination of microscopy-derived patterns. Because each IMS variable has its own relationship to the microscopy patterns, the fusion model contains a submodel for each variable (**Supplementary Fig. 5**) and predictive performance will be variable specific. Building the best possible model does not ensure a good prediction in all cases, as there may not be a cross-modality relationship to exploit for a particular peak, or the complexity of the relationship cannot be described accurately using the present model. It is therefore imperative that the model-building step is followed by an evaluation step that assesses for each m/z its prediction reliability.

We provide three indicative approaches. The first indicator is the reconstruction score, which reports how close a prediction using microscopy measurements gets to the measured ion distribution, given the model. The second indicator is the absolute residuals image, which shows where in the tissue microscopy-driven prediction approximates the IMS measurements well and where it does not. A third indicator conveys prediction robustness as a function of location by calculating a 95% CI image using bootstrapping (Online Methods). This CI image highlights in which tissue areas the prediction is more robust and in which areas the prediction might be model specific and less trustworthy. Further details on these indicators are available in **Supplementary Note 2 (Supplementary Figs. 6–9)**.

The prediction of molecular distributions in tissue areas for which no IMS is available is a form of out-of-sample prediction (**Fig. 5**). In molecular imaging studies where comprehensive measurement of the entire tissue sample is not practical or economical, or in studies where the number of samples is very large, out-of-sample prediction can provide a remarkable alternative. It can also provide an evidence-supported view into the probable content of tissue areas or samples that need to be saved for other analysis techniques, for example, in multibranched drug-discovery workflows. Further, it can perform 'anomaly detection', in which the measured content of a diseased tissue sample is directly compared to the content predicted to be there on the basis of a model trained on normal tissue. In high-dimensional data sets, this would immediately highlight differences and provide a facile path to pathology-derived anomalies. However, it is important to keep the conditions for the (microscopy) measurements between these different areas as similar as possible and to keep the model from overfitting on the training data.

Fusion-based predictions show an enrichment of patterns supported by multiple technologies and an attenuation of modality-specific patterns. The inherent separation of each measured tissue signal into a part that can be predicted by the other technology and a part that is modality-specific lies at the basis of this aspect of image fusion. For measured signals with strong cross-modal support, this separation can be used to filter technology-specific noise from genuine tissue signal. This process can substantially increase the SNR of observations and enables discovery of patterns that might otherwise be missed. Each ion image measured by IMS can be considered a superposition of noise variation on top of a biological or sample signal pattern. The fusion model tries to write each ion image as a linear combination of microscopy-derived patterns, and for ion images with strong cross-modal support, the model succeeds in reproducing most if not all of the biological or sample signal pattern that way. Because the noise tends to be sensor specific, it is highly unlikely that fusion finds a link to an identical pattern in the other technology. As a result, modality-specific noise variation will often not survive the modeling and prediction process, effectively denoising the predictions. Because the biological signal often reports an underlying tissue structure that can modulate measurements in both modalities, finding a relationship across technologies is not uncommon; and when that happens, the presence of real tissue variation in the predictions is more readily observed. It is not necessarily true that all technology-specific variation is noise, as a measurement might be reporting a biological feature that can be detected by only one of the sensors. However, that situation can be detected by a reduced reconstruction score and can even be spatially confined to a tissue subarea using, for example, the absolute residuals image. The multimodal enrichment effect is inherently present in fusion-based predictions and can be used to increase confidence in observations by cross-modal corroboration (**Supplementary Figs. 18–20**). The ability of fusion to aid in separating true signal variation from instrumental noise variation by integrating with another data source has immediate value for increasing measurement sensitivity and SNR without the need to physically adjust the instrument.

In conclusion, image fusion enables the creation of novel predictive imaging modalities that combine the advantages of different sensor types to deliver insights that cannot be normally obtained from the separate technologies alone. The modeling of cross-modality relationships provides predictive paths to new biological understanding through the integration of observations from different measurement principles. It also allows these fused modalities to circumvent sensor-specific limitations (such as IMS resolution via sharpening), to attenuate technology-specific noise sources (such as matrix artifact attenuation by multimodal signal enrichment) and to predict observations in the absence of measurements (such as prediction of ion intensity in non-IMS-measured areas). The predictive applications of image fusion can contribute in those cases where conducting a physical measurement is impractical (for example, owing to long measurement time, varying instrument stability or drift), uneconomical (laser wear, detector deterioration) or even infeasible because of instrumental limitations (spatial resolution beyond laser capabilities, low SNR). This study shows that the fusion of microscopy and mass spectrometry can provide remarkable results, and it lays the groundwork for more advanced modeling across technologies. These prediction methods, and the cross-technology pattern discovery that drives them, can fulfill an instrumental role in realizing the full potential of multimodal imaging studies, particularly in the molecular mapping of tissue.

METHODS

Methods and any associated references are available in the online version of the paper.

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

ACKNOWLEDGMENTS

This work was supported by the US National Institutes of Health grants NIH/NIGMS R01 GM058008-15 and NIH/NIGMS P41 GM103391-04. R.V.d.P. thanks E. Waelkens for his encouragement and support.

AUTHOR CONTRIBUTIONS

R.V.d.P. conceived of and developed the methodology, designed experiments, analyzed and interpreted data and wrote the manuscript; J.Y. designed experiments, acquired data and edited the manuscript; J.S. designed experiments, acquired data, performed identifications and edited the manuscript; R.M.C. designed experiments, interpreted data, revised the manuscript and is principal investigator for the grants that fund this research.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

1. Weissleder, R. Scaling down imaging: molecular mapping of cancer in mice. *Nat. Rev. Cancer* **2**, 11–18 (2002).
2. Massoud, T.F. & Gambhir, S.S. Molecular imaging in living subjects: seeing fundamental biological processes in a new light. *Genes Dev.* **17**, 545–580 (2003).
3. Jahn, K.A. *et al.* Correlative microscopy: providing new understanding in the biomedical and plant sciences. *Micron* **43**, 565–582 (2012).
4. Jacobs, R.E. & Cherry, S.R. Complementary emerging techniques: high-resolution PET and MRI. *Curr. Opin. Neurobiol.* **11**, 621–629 (2001).
5. Chughtai, S. *et al.* A multimodal mass spectrometry imaging approach for the study of musculoskeletal tissues. *Int. J. Mass Spectrom.* **325–327**, 150–160 (2012).
6. Smith, C. Two microscopes are better than one. *Nature* **492**, 293–297 (2012).
7. Caplan, J., Niethammer, M., Taylor, R.M. II. & Czymmek, K.J. The power of correlative microscopy: multi-modal, multi-scale, multi-dimensional. *Curr. Opin. Struct. Biol.* **21**, 686–693 (2011).
8. Modla, S. & Czymmek, K.J. Correlative microscopy: a powerful tool for exploring neurological cells and tissues. *Micron* **42**, 773–792 (2011).
9. Townsend, D.W. A combined PET/CT scanner: the choices. *J. Nucl. Med.* **42**, 533–534 (2001).
10. Townsend, D.W., Beyer, T. & Blodgett, T.M. PET/CT scanners: a hardware approach to image fusion. *Semin. Nucl. Med.* **33**, 193–204 (2003).
11. Masyuko, R., Lanni, E.J., Sweedler, J.V. & Bohn, P.W. Correlated imaging - a grand challenge in chemical analysis. *Analyst* **138**, 1924–1939 (2013).
12. Bocklitz, T.W. *et al.* Deeper understanding of biological tissue: quantitative correlation of MALDI-TOF and Raman imaging. *Anal. Chem.* **85**, 10829–10834 (2013).
13. Clarke, F.C. *et al.* Chemical image fusion. The synergy of FT-NIR and Raman mapping microscopy to enable a more complete visualization of pharmaceutical formulations. *Anal. Chem.* **73**, 2213–2220 (2001).
14. Judenhofer, M.S. *et al.* Simultaneous PET-MRI: a new approach for functional and morphological imaging. *Nat. Med.* **14**, 459–465 (2008).
15. Glenn, D.R. *et al.* Correlative light and electron microscopy using cathodoluminescence from nanoparticles with distinguishable colours. *Sci. Rep.* **2**, 865 (2012).
16. Josephson, L., Kircher, M.F., Mahmood, U., Tang, Y. & Weissleder, R. Near-infrared fluorescent nanoparticles as combined MR/optical imaging probes. *Bioconjug. Chem.* **13**, 554–560 (2002).
17. Blum, R.S. & Liu, Z. *Multi-sensor Image Fusion and Its Applications* (CRC Press, 2005).
18. Bretschneider, T. & Kao, O. in *Proc. 1st Online Symp. Electron. Eng.* 1–8 (2000).
19. Pohl, C. & Van Genderen, J.L. Multisensor image fusion in remote sensing: concepts, methods and applications. *Int. J. Remote Sens.* **19**, 823–854 (1998).
20. Price, J.C. Combining multispectral data of differing spatial resolution. *IEEE Trans. Geosci. Rem. Sens.* **37**, 1199–1203 (1999).
21. Simone, G., Farina, A., Morabito, F.C., Serpico, S.B. & Bruzzone, L. Image fusion techniques for remote sensing applications. *Inf. Fusion* **3**, 3–15 (2002).
22. Gaemperli, O. *et al.* Cardiac image fusion from stand-alone SPECT and CT: clinical experience. *J. Nucl. Med.* **48**, 696–703 (2007).
23. Li, H. *et al.* Object recognition in brain CT-scans: knowledge-based fusion of data from multiple feature extractors. *IEEE Trans. Med. Imaging* **14**, 212–229 (1995).
24. Yang, L., Guo, B.L. & Ni, W. Multimodality medical image fusion based on multiscale geometric analysis of contourlet transform. *Neurocomputing* **72**, 203–211 (2008).
25. Varshney, P.K. *et al.* in *Proc. Int. Conf. Image Proc.* **3**, 532–536 (IEEE, 1999).
26. Caprioli, R.M., Farmer, T.B. & Gile, J. Molecular imaging of biological samples: localization of peptides and proteins using MALDI-TOF MS. *Anal. Chem.* **69**, 4751–4760 (1997).
27. Stoeckli, M., Chaurand, P., Hallahan, D.E. & Caprioli, R.M. Imaging mass spectrometry: a new technology for the analysis of protein expression in mammalian tissues. *Nat. Med.* **7**, 493–496 (2001).
28. Amstalden van Hove, E.R., Smith, D.F. & Heeren, R. A concise review of mass spectrometry imaging. *J. Chromatogr. A* **1217**, 3946–3954 (2010).
29. Chaurand, P. Imaging mass spectrometry of thin tissue sections: a decade of collective efforts. *J. Proteomics* **75**, 4883–4892 (2012).
30. Norris, J.L. & Caprioli, R.M. Analysis of tissue specimens by matrix-assisted laser desorption/ionization imaging mass spectrometry in biological and clinical research. *Chem. Rev.* **113**, 2309–2342 (2013).
31. Murphy, R.C., Hankin, J.A. & Barkley, R.M. Imaging of lipid species by MALDI mass spectrometry. *J. Lipid Res.* **50**, S317–S322 (2009).
32. Van de Plas, R. *Tissue Based Proteomics and Biomarker Discovery – Multivariate Data Mining Strategies for Mass Spectral Imaging*. PhD thesis, KU Leuven (2010).
33. Andersson, M., Andren, P. & Caprioli, R.M. in *Neuroproteomics* (ed. Azalte, O.) Ch. 7, 115–134 (CRC Press, 2009).
34. Franck, J. *et al.* MALDI mass spectrometry imaging of proteins exceeding 30,000 daltons. *Med. Sci. Monit.* **16**, BR293–BR299 (2010).
35. Bradshaw, R., Bleay, S., Wolstenholme, R., Clench, M.R. & Francese, S. Towards the integration of matrix assisted laser desorption ionisation mass spectrometry imaging into the current fingermark examination workflow. *Forensic Sci. Int.* **232**, 111–124 (2013).
36. Chavez, P.S. Jr., Sides, S.C. & Anderson, J.A. Comparison of three different methods to merge multiresolution and multispectral data- Landsat TM and SPOT panchromatic. *Photogramm. Eng. Remote Sensing* **57**, 295–303 (1991).
37. Garguet-Duport, B., Girel, J., Chassery, J.-M. & Patou, G. The use of multiresolution analysis and wavelets transform for merging SPOT panchromatic and multispectral image data. *Photogramm. Eng. Remote Sensing* **62**, 1057–1066 (1996).
38. Lee, J. & Lee, C. Fast and efficient panchromatic sharpening. *IEEE Trans. Geosci. Remote Sens.* **48**, 155–163 (2010).
39. Draper, N.R., Smith, H. & Pownell, E. *Applied Regression Analysis* 1st edn. (Wiley, 1966).
40. Wold, S., Sjöström, M. & Eriksson, L. PLS-regression: a basic tool of chemometrics. *Chemom. Intell. Lab. Syst.* **58**, 109–130 (2001).

ONLINE METHODS

Image fusion method. A multimodal study of tissue typically provides two or more image types covering the same or related tissue samples. The measured images can be spatially registered to each other to provide multimodal overlays and to enable human interpretation. However, image registration does not entail mining the registered image data for mutually informative relationships. Such multimodal data mining is part of image fusion. In its most basic form, image fusion can be seen as a process of combining two or more images into a single image such that

$$\text{image}_{\text{modality } A} \,\&\, \text{image}_{\text{modality } B} \rightarrow \text{image}_{\text{fusion of modalities } A \text{ and } B}$$

Although we do describe this scenario in our sharpening examples, we prefer a broader definition of image fusion in order to also capture predictive applications that do not necessarily output an image:

$$\text{image}_{\text{modality } A} \,\&\, \text{image}_{\text{modality } B} \rightarrow \text{function } q$$

with function q being a cross-modality modeling function that ties observations in modality A to observations in modality B and vice versa.

Some studies have suggested implementing function q by calculating the correlation between image patterns of different origin as a computational follow-up to acquisition and registration[3,6–8,11,12]. However, our data show that for modalities such as microscopy and IMS, which employ substantially different measurement principles, it is quite rare to have two natively measured variables from two different technologies report the same pattern (heterogeneity problem). This suggests that in a heterogeneous imaging environment, the insight provided by a univariate measure of correspondence (for example, correlation) will often be limited. Also, calculating any measure of cross-modal correspondence is often nontrivial with modalities acquired at different spatial resolutions (multiresolution problem). The image fusion method developed in this work specifically addresses both the heterogeneity and multiresolution challenges inherent to multimodal imaging and focuses specifically on IMS-microscopy fusion.

The objective of image fusion is to merge information from several source images into a single image or to deliver a more accurate description of an imaged sample[17] (**Fig. 1**). The integration of information from different sensor types generally requires some form of model (implementing function q) to guide how observations of distinct origin are combined[18]. Such a fusion model can be defined a priori on the basis of technology considerations or assumptions. However, this assumes an exhaustive understanding of the measurement types, their distinctive noise sources, and where the connections between their observations lie. In an environment of heterogeneous and high-dimensional image types, this is often hard to specify beforehand. Our method avoids these difficulties by building its fusion model in an unsupervised way from a set of registered multimodal measurements. Although the model structure needs to be specified (for example, linear versus nonlinear, stochastic versus deterministic), the method learns the optimal model parameters empirically and without prior knowledge by mining them from a set of example observations. By establishing a model empirically rather than through prior specification, the method reduces human bias in the fusion and prediction results. Additionally, the empirical approach can reveal relationships between the sensor types that are not known before the analysis or that are specific to the case study at hand.

The fusion model and the multimodal data mining that leads up to it are some of the primary aspects in which image fusion differentiates itself from image registration. Image registration is the process of spatially aligning source images to each other, and it commonly leads to co-represented data rather than data mining across the modalities. Although image registration is not the focus of this paper, it is a prerequisite for image fusion, and it is available in most state-of-the-art multimodal experiments.

The field of image fusion encapsulates a broad variety of methods and approaches, stemming from its widely diverse application areas. In order to introduce image fusion to molecular imaging at the micrometer scale, we employ a relatively basic modeling method based on regression analysis to deliver our proof-of-concept examples. We approach the modeling challenge as a massive multivariate regression problem[39] between variables derived from measurements in one technology (for example, IMS) and variables derived from measurements in another technology (for example, microscopy). Regression analysis is a staple of analytical chemistry. A simple mass spectrometry example is the estimation of a calibration curve for quantitation, where an ion intensity variable is linked to an abundance variable by modeling their relationship as a curve through a point cloud of paired measurements. The curve, a simple model, can then be used to predict abundance for an experimentally measured ion intensity value. In our work, this concept is expanded to a multivariate setting using partial least-squares (PLS) regression[40]. The goal is to link multiple ion intensity variables to multiple photon-based variables by modeling the distribution of spatially paired measurements from both sides. The resulting model is not a single curve but rather a set of slopes and intercepts that, when fed a microscopy measurement, outputs a prediction for the IMS variables. Once a model is built, it can be used in a variety of predictive scenarios depending on the particular task required by the fusion application.

The fusion method we developed consists of two distinct phases (**Supplementary Fig. 1**). The first phase is independent of the application being pursued and focuses on building and evaluating a cross-modality model. Given a set of registered measurements provided by different sensor types, this phase seeks to detect as many cross-modality relationships as possible, attempts to capture them in a mathematical model, and evaluates the fidelity of predictions resulting from that model. For the examples outlined in this paper, this first phase entails finding IMS-microscopy correspondences and capturing them as best as possible in a linear model structure. The second phase entails using the cross-modality model in a prediction scenario. Its implementation is thus specific to the fusion application that is being pursued. For sharpening, the model is applied to microscopy measurements available at a spatial resolution that exceeds the IMS resolution. For out-of-sample prediction, the model is applied to microscopy measurements in tissue areas where no IMS is measured to begin with. For multimodal enrichment, prediction at the native IMS resolution is sufficient because only the filtering of observations

by going through the model is pursued. Other fusion applications can require entirely different implementations of the second phase. As such, this is an area in which future research might be able to extend from the method described here. We illustrate both phases using a sharpening example (**Supplementary Fig. 1**).

In **Supplementary Notes 1–3** we give a mathematical description of the developed fusion framework. We provide mathematical definitions to establish the general framework steps, independent of which image sources are being considered. These are followed by IMS and microscopy-specific definitions that incorporate modality-specific considerations. The function definitions can be readily implemented using standard methods available in most algebraic environments such as Matlab (The Mathworks Inc.). Where a nonstandard implementation is used, we provide pseudocode. A full implementation of the image fusion framework is provided as a command-line utility that can be downloaded at http://fusion.vueinnovations.com/. Additional material related to case study 1 is provided in the **Supplementary Data**. **Supplementary Figure 3** provides an overview that ties the different mathematical steps together.

Case-study materials and methods. Tissue samples (purchased from Peel-Freez Biologicals (Rogers)) were cryo-sectioned (12 µm thickness), thaw mounted onto ITO-coated microscope slides (Delta Technologies), and dried in a vacuum desiccator (~2 h) before preparation for MALDI analysis. For protein imaging experiments, tissue sections were rinsed with 70% ethanol, 100% ethanol, Carnoy's fluid, 100% ethanol, water, and 100% ethanol. (Except for Carnoy's fluid for 2 min, every other rinsing was 30 s). Sinapinic acid was applied using sublimation at ~0.2–0.3 mg/cm^2. The sublimated sections were then recrystallized under 85 °C for 3.5 min with 1 mL of water and 50 µL of acetic acid. For lipid analysis, 1,5-diaminonaphthalene (DAN) was applied to unwashed tissue sections using sublimation at ~0.15–0.20 mg/cm^2. Sinapinic acid and DAN were purchased from Acros-Organics. Acetic acid, ethanol, chloroform, and acetone were purchased from Sigma-Aldrich.

The optical images of the stained sections were acquired by first removing the MALDI matrix by rinsing IMS-analyzed sections with acetone or staining neighboring sections. Hematoxylin and eosin (H&E) and Nissl staining were performed using standard protocols, and optical images were acquired with a Mirax scanner (Carl Zeiss) at 0.3-µm resolution.

The MALDI-TOF IMS analyses were performed on a Bruker Autoflex Speed TOF mass spectrometer in positive linear mode (protein-focused imaging) and a Bruker Autoflex Speed TOF/TOF mass spectrometer in negative reflector mode (lipid-focused imaging) using FlexControl 3.3 software. The mass ranges and spatial resolutions for the different case studies are specified in **Supplementary Table 1**. Approximately 100 shots per spot were acquired at a 1-kHz repetition rate using a Smartbeam II Nd:YAG laser. Image acquisition was carried out using FlexImaging 3, and further processing took place in Matlab. The spectra were normalized on the basis of their common ion current, disregarding differential peaks[32]. They were baseline corrected using a spline approximation of the baseline at the 10% quantile of ion intensities and employing window sizes of 500 and 50 and step sizes of 250 and 25 for the protein and lipid-focused spectra, respectively. The spectra were also optimally aligned along the m/z axis to reduce peak drift, allowing a maximum m/z shift of 12 in the case of the protein-focused spectra and 0.5 in the case of the lipid-focused spectra. Both steps were performed using the Bioinformatics Toolbox of Matlab (MathWorks).

For evaluation of the predicted ion distributions, high mass resolution imaging data were collected using a MALDI FTICR mass spectrometer (9.4 Tesla solariX, Bruker Daltonics) equipped with an Apollo II dual ion source and a Nd:YAG frequency-tripled 355-nm solid-state laser (2 kHz). Images were collected at 50-µm or 10-µm lateral resolution using 1,000 shots per pixel and the minimum laser setting (10-µm laser spot size). The instrument was tuned and calibrated to obtain high mass accuracy (<1 p.p.m.) for broadband measurement over m/z 400–2,000 with a 1 M transient and a single scan per pixel.

Lipid identification was performed using FTICR mass spectrometry to provide both accurate mass measurements and fragmentation data. Precursor ions were mass selected in the source region of the instrument using a linear quadrupole (m/z window: 1.5). The selected ions were fragmented in the ICR cell using sustained off-resonance irradiation collision-induced dissociation (SORI-CID). SORI power (1.0–2.1%) was tuned to maximize fragmentation for each analyte (pulsed argon, 0.25-s 500-Hz irradiation). Identifications were confirmed by comparing accurate masses (<1 p.p.m.) of the precursor ions and fragment ions to theoretical values in the LIPID MAPS database (Nature Lipidomics Gateway, http://www.lipidmaps.org/) and manual interpretation of the fragmentation patterns.

doi:10.1038/nmeth.3296