

# Computational Methods for Mass Spectrometry Imaging: Challenges, Progress, and Opportunities

Chanchala D. Kaddi and May D. Wang

**Abstract** Mass spectrometry imaging (MSI) is a rapidly growing field of research, with applications in proteomics, lipidomics, and metabolomics. The benefit of MSI is its capacity to measure spatially resolved molecular information. Computational methods are important to extracting information from MSI data for basic and translational research. In this chapter, we examine current and emerging methods for analysis of MSI data, and highlight associated challenges and opportunities in computational research for MSI.

## Introduction

Mass spectrometry imaging (MSI) is a large-scale experimental technique that can yield spatially resolved information about the molecular composition of a biological sample. MSI datasets are generated by acquiring the complete mass spectrum at multiple points across the sample surface, yielding a three-dimensional ( $x, y$ : spatial, e.g., tissue, and  $z$ : spectral or  $m/z$ ) dataset as shown in Fig. 1.

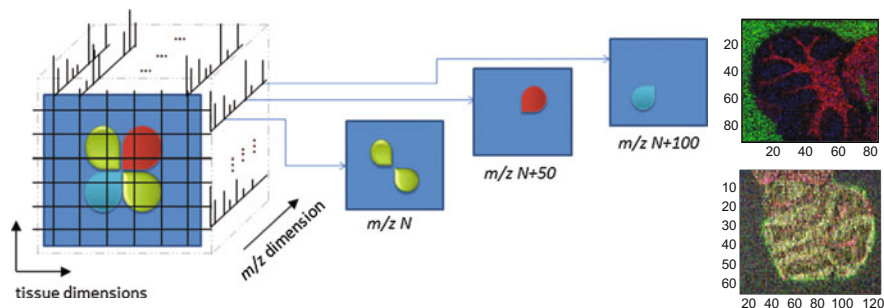
The MSI dataset includes valuable information which is not obtainable through similar analyses using immunohistochemistry staining or non-imaging mass spectrometry. In traditional histological analysis, tissue is typically stained for a small number of molecular targets; in contrast, MSI is capable of simultaneously tracking thousands of  $m/z$  (mass-to-charge ratio) values. Depending on the MSI acquisition modality, each  $m/z$  value can be interpreted as a molecule or molecular fragment. Additionally, staining can only identify known molecular targets, while the large-scale data acquired by MSI enables discovery of sample components (and hence, potential biomarkers). Compared to mass spectrometry alone, MSI preserves

---

C.D. Kaddi · M.D. Wang (✉)

Georgia Institute of Technology and Emory University, Atlanta, Georgia, USA

e-mail: maywang@bme.gatech.edu



**Fig. 1** (Left) Three-dimensional structure of MSI data. (Right) False-color visualizations of multiple  $m/z$  values from MSI datasets of mouse models of Tay-Sachs/Sandhoff disease

the sample's spatial and morphological information. Thus, spectra corresponding to different regions of organs, or to tumor, marginal, or normal sections of biopsies, can be differentiated, enabling more detailed and target-specific analysis. Due to these benefits, MSI is emerging as a popular experimental technique in proteomics [1], lipidomics [2], and metabolomics [3] research.

Because MSI is spatially resolved, it is particularly relevant for research into diseases which have spatially localized characteristics—particularly cancer. Recent MSI studies have investigated cancers of the head and neck [4], brain [5], breast [6], renal [7], stomach [8], prostate [9], colon [10], pancreas [11], and bladder [12]. Other recent MSI studies have targeted diseases including Tay-Sachs/Sandhoff disease [13], Behçet disease [14], Parkinson's disease [15, 16], Alzheimer's disease [17], Duchenne muscular dystrophy [18, 19], Fabry disease [20], atherosclerosis [21], and stroke and ischemic injury [22–24]. In addition, MSI has been used to study bio-implant interfaces [25, 26] and drug distribution within tissues [27–32].

The spectral dimension of MSI data can be very large (e.g., tens of thousands of  $m/z$  values), making computational analysis essential to interpretation. It is critical to identify and to develop effective analytical methods for large-scale data mining, pattern recognition, and exploration. This chapter begins by discussing several key open challenges in MSI research. Next, the current state-of-the-art in MSI analysis will be described, including techniques such as principal component analysis, clustering, and classification. Additionally, several emerging methods for MSI analysis, such as non-negative matrix factorization, will be introduced. All methods are discussed in the context of recent MSI studies which apply them, spanning several MSI modalities, such as Matrix-Assisted Laser Desorption/Ionization (MALDI)-MSI and Desorption Electrospray Ionization (DESI)-MSI. Finally, a case study in applying unsupervised analysis methods for pattern detection in MSI will be provided.

## Challenges

As will be described in greater detail in the following sections, much progress has been made in identifying and developing analytical methods for pattern detection in MSI data. Research on this topic is still highly active. In particular, we highlight three new areas of interest in computational MSI research.

### ***Challenge 1: Integration of MSI Data with Complementary Imaging Modalities***

An emerging area of interest in MSI data analysis is the integration of MSI data to other images, such as those acquired via different MSI modalities or non-MSI data types. One example is the combination of MALDI-MSI data with magnetic resonance images [33]; another is integration of DESI-MSI data with histology images [34]. This type of integration harnesses the different strengths of the data types—for example, MRI imaging yields much higher resolution data than MSI, but does not measure molecular information like MSI. Similar reasoning is behind efforts to combine different MSI modalities: PCA and CCA have also been used to link low-mass SIMS-MSI data with high-mass MALDI-MSI data from the same brain tissue sample [35]. Because of the inherent differences in imaging modalities, several important computational challenges are in the image processing domain: for example, image alignment algorithms used to ‘stitch together’ multiple MALDI-MSI images obtained from a large sample [36], and registration algorithms to map consecutive optical images in order to construct MSI datasets for three-dimensional samples [37].

### ***Challenge 2: Movement Toward MSI from Three-Dimensional Samples***

The movement toward MSI analysis of three-dimensional samples, as just mentioned, is an important development in the field. The studies described thus far have all implemented MSI on two-dimensional samples, e.g., very thin slices of tissue. If multiple spatially consecutive slices are taken from an organ or tumor,  $m/z$  images from multiple MSI datacubes can be stitched together to track the spatial distribution and expression of an  $m/z$  value through the original three-dimensional sample. We refer readers to the recent publication [4–32] for an example of this technique and discussion of the computational challenges associated with MSI for three-dimensional data.

### ***Challenge 3: Reproducibility, Data Standardization, and Community Resources***

MSI also provides a rich opportunity for biomarker identification. However, reproducibility of results is a major challenge. Many MSI studies consider a small number of samples, making it difficult to generalize the suitability of the analytical methods used. It would be valuable to examine alternative analytical pipelines for MSI data in a systematic manner on a variety of different MSI datasets, similar to the MAQC-II study conducted for microarray-based predictive modeling [38]. However, in addition to the obstacle of scale, such efforts are hindered because unlike microarray and RNA sequencing data, MSI data is not readily shared in public repositories. The development of community resources and infrastructure—as well as standards for quality control and transparency like Minimum Information protocols (<http://mibbi.sourceforge.net/>)—would facilitate this process. The recent release of mzML [39], a standardized data format for mass spectrometry, and the PRIDE proteomics data repository from the European Bioinformatics Institute for MS/MS proteomics datasets (<http://www.ebi.ac.uk/pride/archive/>), are therefore encouraging developments.

### **Current Techniques in MSI Analysis**

Analytical methods for data and knowledge mining are divided into two main classes: supervised and unsupervised learning. In supervised learning, a predictive model is constructed from annotated training data, such that when a new sample is provided, the model can correctly predict the annotation of the sample. Supervised methods are further divided into two main categories: classification, in which the predicted annotation is a group label (e.g., “healthy” or “diseased”), or regression, in which it is a numerical value. For example, classification models have been developed using MALDI-MSI data to distinguish HER2 positive and HER2-negative tissues [6]; to distinguish cancerous and non-cancerous prostate tissue [40]; and to classify breast cancer sample regions as necrotic, viable/active tumor, or tumor interface region, while distinguishing them from embedding gelatin and glass or holes, using SIMS-MSI data [41]. In contrast, unsupervised learning requires no annotation or prior knowledge of the data structure. Unsupervised methods are used for exploration of the data and the identification of potential patterns; the results of these analyses can be a precursor to supervised analysis. Common unsupervised methods include dimensionality reduction and clustering. The remainder of this section will introduce unsupervised methods for MSI data analysis.

## A. Dimensionality reduction

### *Principal component analysis*

Principal component analysis (PCA) is currently one of the most popular techniques for exploratory data analysis in MSI. The utility of PCA is in its ability to highlight different spatial patterns present in the data, and the  $m/z$  values which contribute to them. Given a data matrix  $X$  of dimensions  $M \times N$  (i.e.,  $M$  mass spectra each containing  $N$   $m/z$  values), PCA performs a linear transformation that projects the data into a different, potentially more meaningful, spectral coordinate space. The axes directions in this transformed space are defined by a set of orthogonal  $M$ -dimensional basis vectors (the principal components), and are related to the variance in  $X$ . The first principal component is the direction in which the variance of the data is maximized, and can be interpreted as the most prominent pattern in the data. The second principal component corresponds to the direction of the second highest variance in the data, and so on. After performing PCA, the  $p$  principal components which contribute to the majority of the variance in the dataset are retained. Since  $p$  is typically chosen to be  $\ll N$ , PCA produces a dimensionally reduced  $M \times p$  dataset.

### *Non-negative matrix factorization*

Non-negative matrix factorization (NMF) is another method that can be used for dimensionality reduction. In NMF, a data matrix  $X$  of dimensions  $(M \times N)$  is factored into two matrices  $W$  ( $M \times k$ ) and  $H$  ( $k \times N$ ), such that  $X \approx WH$ . This is done iteratively by minimizing the residual  $\|X - WH\|^2$  such that  $W, H \geq 0$  [42]. The user-selected parameter  $k$  is the number of components into which the data is separated. The matrix  $W$  is a set of basis vectors describing the  $m/z$  values which comprise each component. The  $k$  columns of matrix  $W$  can be interpreted as groupings of  $m/z$  values corresponding to prominent spatial patterns in the data. NMF has been assessed on both MALDI and DESI-MSI data [43–46]. A web-based tool, omniSpect, is also available for performing NMF on MSI data [47].

### *Other Dimensionality Reduction Methods*

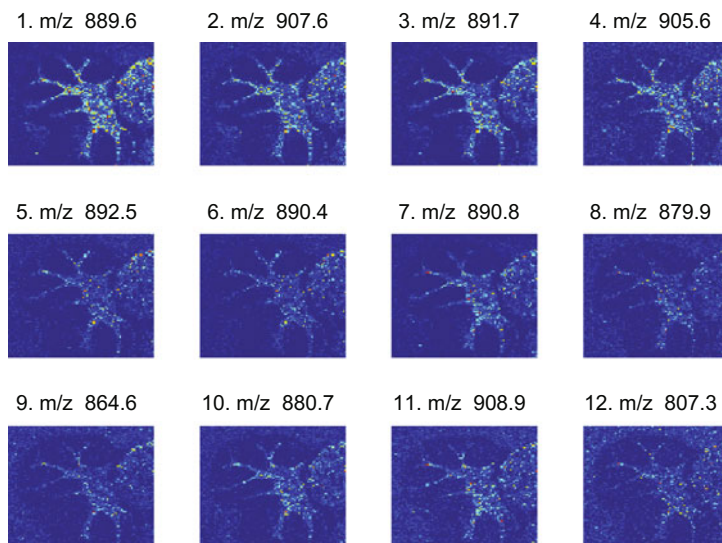
Independent component analysis (ICA) separates a mixture into components, based on the assumption that the mixture is a linear combination of statistically independent components with non-Gaussian distributions [48]. ICA has been assessed for studying intratumor heterogeneity via MSI data [46] and compared with PCA and NMF on MALDI-MSI data [43]. Like PCA, ICA presents an obstacle in terms of interpretation of the components, which can be negative. In canonical correlation analysis (CCA), two datasets, each with different dimensions, may be projected onto the feature space of the other such that the project data has maximum correlation [49]. CCA has also been used to correlate low-mass SIMS-MSI data with high-mass MALDI-MSI data, which improved image contrast and interpretation of the data [35]. Parallel factor analysis (PARAFAC) is another method often used in chemometrics for decomposing high-dimensional data; it has been tested on SIMS and LDI MSI datasets [50].

## B. Clustering

Clustering is an unsupervised method for data analysis in which a sample (e.g., an individual mass spectrum) is allocated into a specific group (e.g., a cluster) based on a quantitative measure of the similarity. Clustering can reveal potentially meaningful structures and patterns within the data. For example, using MALDI-MSI data of gastric cancer, hierarchical clustering was used both to cluster spectra within a single tissue dataset, and also to cluster tumors from different patients [51]. LA-ICP-MSI data of rat brain tissue was analyzed by  $k$ -means clustering, revealing meaningful patterns in which clusters corresponded to known anatomical features [52]. Similarly, high dimensional discriminant clustering (HDDC) was applied to MALDI-MSI data of a rat brain and an intestinal-invading neuroendocrine tumor to find spectral clusters corresponding to morphological structures [53]. This type of mapping, termed spatial segmentation, is discussed further in a recent review [54]. While numerous algorithms exist for separating data into clusters [55], two of the most commonly applied methods are hierarchical and  $k$ -means approaches. In hierarchical clustering, a dataset  $X$  is separated into different levels of clusters, culminating in a dendrogram or cluster tree. The terminal leaf nodes each correspond to a single sample. In  $k$ -means, the dataset is separated into a predefined number  $k$  of clusters.

## C. Spatial similarity

While similarity measures (in the spectral dimension) are utilized in clustering, another independent application of similarity measures is to identify  $m/z$  images with similar expression patterns in the spatial dimension. Figure 2 shows an



**Fig. 2** Highly similar  $m/z$  images, as identified by the multivariate hypergeometric similarity measure [58]

example of similarity measure-based retrieval of  $m/z$  images with a spatial pattern similar to that of the  $m/z$  value of interest. For MALDI-MSI data, the similarity of  $m/z$  images to each other has been assessed using Pearson correlation [56] and using similarity measures based on the hypergeometric and multivariate hypergeometric distributions [57, 58]. Multivariate least-squares-based query has also been applied for this task [59].

## Case Study

The following case study compares and contrasts PCA and NMF for finding potentially relevant patterns in MSI data. The data used in this example is MALDI-MSI, from a mouse model of Tay-Sachs disease [17].

**Example 1** A step-by-step implementation of PCA in MATLAB R2014a to find patterns in MSI data.

- Step 1: After loading the three-dimensional MSI dataset into MATLAB, restructure it into a two-dimensional matrix. Here,  $a$  and  $b$  are the spatial dimensions (i.e., the number of pixels in the horizontal and vertical directions), and  $c$  is the spectral dimension (i.e., the number of  $m/z$  values). In the restructured  $M \times N$  matrix,  $M$  is the number of spectra ( $M = a \times b$ ) and  $N = c$ .

```
[a,b,c] = size(data);
reshaped = double(reshape(data,a*b,c));
[M,N] = size(reshaped);
```

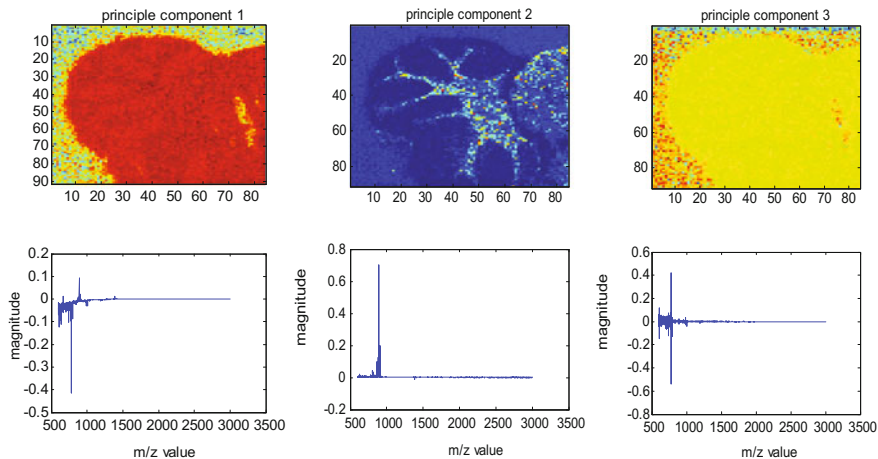
- Step 2: Mean-center the data by subtracting the mean in the spectral dimension (i.e., the average of each  $m/z$  value).

```
mean_reshaped = mean(reshaped,1);
reshaped = reshaped - repmat(mean_reshaped,M,1);
```

- Step 3: Calculate the covariance matrix and the find its eigenvectors and eigenvalues; these are the principal components and their weights. Sort the principal components in order of descending eigenvalue magnitude.

```
covariance_matrix = (1 / (N-1)) * (reshaped' *
reshaped);
[PC,V] = eig(covariance_matrix);
V = diag(V);
[~,indices] = sort(-1*V);
V = V(indices); PC = PC(:,indices);
```

- Step 4: Retain the top 3 principal components, and project the original data onto these components. The reduced dataset will have dimensions  $a \times b \times 3$ .



**Fig. 3** The first three principal components reveal different spatial patterns and associated  $m/z$  values in an MSI dataset (false-color visualization)

```
projected = reshaped * PC(:, 1:3);
PCA_datacube = reshape(projected, a, b, 3);
```

The plots in Fig. 3 show the three images in the PCA-processed datacube (top row), and the principal components themselves (bottom row). PCA reveals different structures within the data—primarily the tissue versus non-tissue regions in PC 1 and PC 3, and cerebellum structure in PC 2.

**Example 2** A step-by-step implementation of NMF in MATLAB to find patterns in MSI data.

For brevity, the MATLAB function ‘nnmf’ is used to perform the analysis in this example.

Step 1: Load and reshape the datacube into an  $M \times N$  matrix as before.

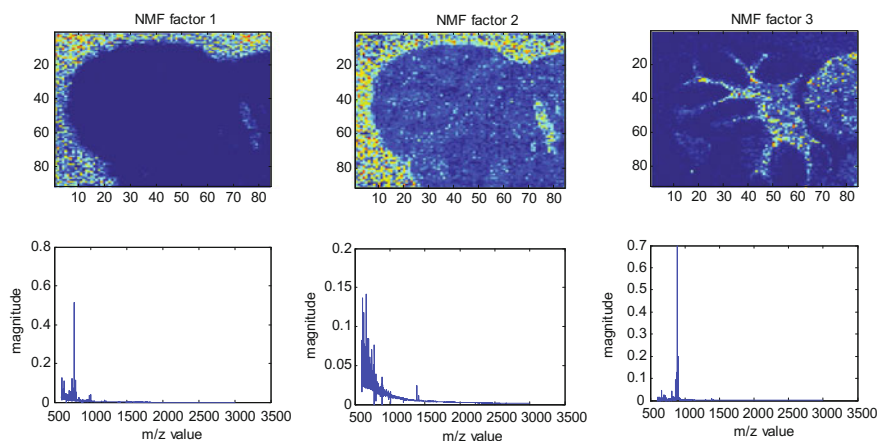
```
[a,b,c] = size(data);
reshaped = double(reshape(data, a*b, c));
```

Step 2: Define  $k$ , the number of components, and perform NMF:

```
k = 3;
[w,h] = nnmf(reshaped, k);
NMF_datacube = reshape(w, a, b, k);
```

Figure 4 shows the three images in the NMF-processed data, i.e., the component matrix  $w$  (top row), and the corresponding row of the weight matrix  $h$  (bottom row). Similar to PCA, NMF reveals tissue versus. non-tissue patterns in the Factor 1 and Factor 2, and the cerebellum structure in Factor 3. Factor 2 also contains some information on the tissue interior. Unlike in PCA, the numerical labeling of NMF





**Fig. 4** The first three NMF factors reveal different spatial patterns and associated  $m/z$  values in an MSI dataset (false-color visualization)

factors is arbitrary. Additionally, as its name indicates, the spectral profiles found by NMF are constrained to be non-negative. This is a useful property for MSI, since the data is also non-negative. In contrast, the negative values in the PCA-generated components in Fig. 3 can be difficult to interpret in terms of biology or chemistry.

## Conclusion

MSI is a rapidly developing area of research with exciting implications for our understanding of numerous biological processes and diseases. In this chapter, we have described the key role of computational analyses in extracting meaningful information from MSI data. State-of-the-art techniques were described and a case study was examined. Finally, we have highlighted several challenges for computational research for MSI. In conclusion, there exist numerous opportunities for researchers to become involved in the development of computational methods and tools for MSI. Ongoing research into more effective and informative analytical techniques will help to harness the power of MSI for accelerating both basic and translational research.

**Acknowledgements** We thank Dr. M. Cameron Sullards and Dr. Yanfeng Chen for sharing the MSI data used in the examples in this review. This research is supported by NIH grant U01 CA151802, Georgia Cancer Coalition distinguished cancer scholar award to Professor May D. Wang, and NSF graduate fellowship to Ms. Chanchala Kaddi.

## References

1. L. MacAleese, J. Stauber, R.M.A. Heeren, Perspectives for imaging mass spectrometry in the proteomics landscape. *Proteomics* **9**, 819–834 (2009)
2. N. Goto-Inoue, T. Hayasaka, N. Zaima, M. Setou, Imaging mass spectrometry for lipidomics. *Biochimica Et Biophysica Acta-Molecular and Cell Biology of Lipids* **1811**, 961–969 (2011)
3. Y. Sugiura, M. Setou, Imaging mass spectrometry for visualization of drug and endogenous metabolite distribution: toward in situ pharmacometabolomes. *J. Neuroimmune Pharmacol.* **5**, 31–43 (2010)
4. S.A. Patel, A. Barnes, N. Loftus, R. Martin, P. Sloan, N. Thakker et al., Imaging mass spectrometry using chemical inkjet printing reveals differential protein expression in human oral squamous cell carcinoma. *Analyst* **134**, 301–307 (2009)
5. N.Y.R. Agar, J.G. Malcolm, V. Mohan, H.W. Yang, M.D. Johnson, A. Tannenbaum et al., Imaging of meningioma progression by matrix-assisted laser desorption ionization time-of-flight mass spectrometry. *Anal. Chem.* **82**, 2621–2625 (2010)
6. S. Rauser, C. Marquardt, B. Balluff, S.O. Deininger, C. Albers, E. Belau et al., Classification of HER2 receptor status in breast cancer tissues by MALDI imaging mass spectrometry. *J. Proteome Res.* **9**, 1854–1863 (2010)
7. S.R. Oppenheimer, D.M. Mi, M.E. Sanders, R.M. Caprioli, Molecular analysis of tumor margins by MALDI mass spectrometry in renal carcinoma. *J. Proteome Res.* **9**, 2182–2190 (2010)
8. Y. Morita, K. Ikegami, N. Goto-Inoue, T. Hayasaka, N. Zaima, H. Tanaka et al., Imaging mass spectrometry of gastric carcinoma in formalin-fixed paraffin-embedded tissue microarray. *Cancer Sci.* **101**, 267–273 (2010)
9. L.H. Cazares, D. Troyer, S. Mendrinos, R.A. Lance, J.O. Nyalwidhe, H.A. Beydoun et al., Imaging mass spectrometry of a specific fragment of mitogen-activated protein kinase/extracellular signal-regulated kinase kinase 2 discriminates cancer from uninvolved prostate tissue. *Clin. Cancer Res.* **15**, 5541–5551 (2009)
10. J.W. Park, H.K. Shon, B.C. Yoo, I.H. Kim, D.W. Moon, T.G. Lee, Differentiation between human normal colon mucosa and colon cancer tissue using ToF-SIMS imaging technique and principal component analysis. *Appl. Surf. Sci.* **255**, 1119–1122 (2008)
11. M.C. Djidja, E. Claude, M.F. Snel, P. Scriven, S. Francese, V. Carolan et al., MALDI-Ion mobility separation-mass spectrometry imaging of glucose-regulated protein 78 kDa (Grp78) in human formalin-fixed, paraffin-embedded pancreatic adenocarcinoma tissue sections. *J. Proteome Res.* **8**, 4876–4884 (2009)
12. A.L. Dill, D.R. Ifa, N.E. Manicke, A.B. Costa, J.A. Ramos-Vara, D.W. Knapp et al., Lipid profiles of canine invasive transitional cell carcinoma of the urinary bladder and adjacent normal tissue by desorption electrospray ionization imaging mass spectrometry. *Anal. Chem.* **81**, 8758–8764 (2009)
13. Y.F. Chen, J. Allegood, Y. Liu, E. Wang, B. Cachon-Gonzalez, T.M. Cox et al., Imaging MALDI mass spectrometry using an oscillating capillary nebulizer matrix coating system and its application to analysis of lipids in brain from a mouse model of Tay-Sachs/Sandhoff disease. *Anal. Chem.* **80**, 2780–2788 (2008)
14. M. Aranyosiova, M. Michalka, M. Kopani, B. Rychly, J. Jakubovsky, D. Velic, Microscopy and chemical imaging of Behcet brain tissue. *Appl. Surf. Sci.* **255**, 1584–1587 (2008)
15. D. Hare, B. Reedy, R. Grimm, S. Wilkins, I. Volitakis, J.L. George et al., Quantitative elemental bio-imaging of Mn, Fe, Cu and Zn in 6-hydroxydopamine induced Parkinsonism mouse models. *Metallomics* **1**, 53–58 (2009)
16. K. Skold, M. Svensson, A. Nilsson, X.Q. Zhang, K. Nydahl, R.M. Caprioli et al., Decreased striatal levels of PEP-19 following MPTP lesion in the mouse. *J. Proteome Res.* **5**, 262–269 (2006)

17. R.W. Hutchinson, A.G. Cox, C.W. McLeod, P.S. Marshall, A. Harper, E.L. Dawson et al., Imaging and spatial distribution of beta-amyloid peptide and metal ions in Alzheimer's plaques by laser ablation-inductively coupled plasma-mass spectrometry. *Anal. Biochem.* **346**, 225–233 (2005)
18. N. Tahallah, A. Brunelle, S. De La Porte, O. Laprevote, Lipid mapping in human dystrophic muscle by cluster-time-of-flight secondary ion mass spectrometry imaging. *J. Lipid Res.* **49**, 438–454 (2008)
19. D. Touboul, A. Brunelle, F. Halgand, S. De La Porte, O. Laprevote, Lipid imaging by gold cluster time-of-flight secondary ion mass spectrometry: application to Duchenne muscular dystrophy. *J. Lipid Res.* **46**, 1388–1395 (2005)
20. D. Touboul, S. Roy, D.P. Germain, P. Chaminade, A. Brunelle, O. Laprevote, MALDI-TOF and cluster-TOF-SIMS imaging of Fabry disease biomarkers. *Int. J. Mass Spectrom.* **260**, 158–165 (2007)
21. N.E. Manicke, M. Neffliu, C. Wu, J.W. Woods, V. Reiser, R.C. Hendrickson et al., Imaging of lipids in atheroma by desorption electrospray ionization mass spectrometry. *Anal. Chem.* **18**, 8702–8707 (2009)
22. J.H. Kim, B.J. Ahn, J.H. Park, H.K. Shon, Y.S. Yu, D.W. Moon et al., Label-free calcium imaging in ischemic retinal tissue by TOF-SIMS. *Biophys. J.* **94**, 4095–4102 (2008)
23. S. Koizumi, S. Yamamoto, T. Hayasaka, Y. Konishi, M. Yamaguchi-Okada, N. Goto-Inoue et al., Imaging mass spectrometry revealed the production of lyso-phosphatidylcholine in the injured ischemic rat brain. *Neuroscience* **168**, 219–225 (2010)
24. S.X. Jiang, S. Whitehead, A. Aylsworth, J. Slinn, B. Zurakowski, K. Chan et al., Neuropilin 1 directly interacts with Fer Kinase to mediate Semaphorin 3A-induced death of cortical neurons. *J. Biol. Chem.* **285**, 9908–9918 (2010)
25. C. Eriksson, K. Borner, H. Nygren, K. Ohlson, U. Bexell, N. Billerdahl, et al., Studies by imaging TOF-SIMS of bone mineralization on porous titanium implants after 1 week in bone, (2006), pp. 6757–6760
26. H. Nygren, C. Eriksson, K. Hederstierna, P. Malmberg, TOF-SIMS analysis of the interface between bone and titanium implants-Effect of porosity and magnesium coating, (2008), pp. 1092–1095
27. E. Acquadro, C. Cabella, S. Ghiani, L. Miragoli, E.M. Bucci, D. Corpillo, Matrix-assisted laser desorption ionization imaging mass spectrometry detection of a magnetic resonance imaging contrast agent in mouse liver. *Anal. Chem.* **81**, 2779–2784 (2009)
28. S.J. Atkinson, P.M. Loadman, C. Sutton, L.H. Patterson, M.R. Clench, Examination of the distribution of the bioreductive drug AQ4N and its active metabolite AQ4 in solid tumours by imaging matrix-assisted laser desorption/ionisation mass spectrometry. *Rapid Commun. Mass Spectrom.* **21**, 1271–1276 (2007)
29. L. Signor, E. Varesio, R.F. Staack, V. Starke, W.F. Richter, G. Hopfgartner, Analysis of erlotinib and its metabolites in rat tissue sections by MALDI quadrupole time-of-flight mass spectrometry. *J. Mass Spectrom.* **42**, 900–909 (2007)
30. P.J. Trim, C.M. Henson, J.L. Avery, A. McEwen, M.F. Snel, E. Claude et al., Matrix-assisted laser desorption/ionization-ion mobility separation-mass spectrometry imaging of vinblastine in whole body tissue sections. *Anal. Chem.* **80**, 8628–8634 (2008)
31. J.M. Wiseman, D. R. Ifa, Y. X. Zhu, C. B. Kissinger, N. E. Manicke, P.T. Kissinger et al., *Desorption electrospray ionization mass spectrometry: Imaging drugs and metabolites in tissues, Proceedings of the National Academy of Sciences of the United States of America*, vol. 105 (2008), pp. 18120–18125
32. M. Zoriy, A. Matusch, T. Spruss, J.S. Becker, Laser ablation inductively coupled plasma mass spectrometry for imaging of copper, zinc, and platinum in thin sections of a kidney from a mouse treated with cis-platin. *Int. J. Mass Spectrom.* **260**, 102–106 (2007)
33. T.K. Sinha, S. Khatib-Shahidi, T.E. Yankeelov, K. Mapara, M. Ehtesham, D.S. Cornett et al., Integrating spatially resolved three-dimensional MALDI IMS with in vivo magnetic resonance imaging. *Nat. Methods* **5**, 57–59 (2008)

34. K.A. Veselkov, R. Mirnezami, N. Strittmatter, R.D. Goldin, J. Kinross, A.V. Speller, et al., Chemo-informatic strategy for imaging mass spectrometry-based hyperspectral profiling of lipid signatures in colorectal cancer. *Proc. Natl. Acad. Sci. USA* **111**, 1216–21 (2014)
35. G.B. Eijkel, B.K. Kaletas, I.M. van der Wiel, J.M. Kros, T.M. Luider, R.M.A. Heeren, Correlating MALDI and SIMS imaging mass spectrometric datasets of biological tissue surfaces. *Surf. Interface Anal.* **41**, 675–685 (2009)
36. A. Broersen, R. van Lier, A.F.M. Altelaar, R.M.A. Heeren, L.A. McDonnell, Automated, feature-based image alignment for high-resolution imaging mass spectrometry of large biological samples. *J. Am. Soc. Mass Spectrom.* **19**, 823–832 (2008)
37. A.C. Crecelius, D.S. Cornett, R.M. Caprioli, B. Williams, B.M. Dawant, B. Bodenheimer, Three-dimensional visualization of protein expression in mouse brain structures using imaging mass spectrometry. *J. Am. Soc. Mass Spectrom.* **16**, 1093–1099 (2005)
38. L. Shi, G. Campbell, W.D. Jones, F. Campagne, Z. Wen, S.J. Walker et al., The MicroArray Quality Control (MAQC)-II study of common practices for the development and validation of microarray-based predictive models. *Nat. Biotechnol.* **28**, 827–838 (2010)
39. L. Martens, M. Chambers, M. Sturm, D. Kessner, F. Levander, J. Shofstahl, et al., mzML-a Community Standard for Mass Spectrometry Data. *Mol. Cell. Proteomics*. **10** (2011)
40. K. Schwamborn, R.C. Krieg, M. Reska, G. Jakse, R. Knuechel, A. Wellmann, Identifying prostate carcinoma by MALDI-imaging. *Int. J. Mol. Med.* **20**, 155–159 (2007)
41. M. Hanselmann, U. Kothe, M. Kirchner, B.Y. Renard, E.R. Amstalden, K. Glunde et al., Toward digital staining using imaging mass spectrometry and random forests. *J. Proteome Res.* **8**, 3558–3567 (2009)
42. D.D. Lee, H.S. Seung, Algorithms for non-negative matrix factorization. *Adv. Neural. Inf. Process. Syst.* **13**, 556–562 (2002)
43. P.W. Siy, R.A. Moffitt, R.M. Parry, Y.F. Chen, Y. Liu, M.C. Sullards et al., *Matrix Factorization Techniques for Analysis of Imaging Mass Spectrometry Data*, (2008)
44. X.C. Xiong, X. Fang, Z. Ouyang, Y. Jiang, Z.J. Huang, Y.K. Zhang, Feature extraction approach for mass spectrometry imaging data using non-negative matrix factorization. *Chin. J. Anal. Chem.* **40**, 663–669 (2012)
45. E.A. Jones, R. Shyti, R.J.M. van Zeijl, S.H. van Heiningen, M.D. Ferrari, A.M. Deelder et al., Imaging mass spectrometry to visualize biomolecule distributions in mouse brain tissue following hemispheric cortical spreading depression. *Journal of Proteomics* **75**, 5027–5035 (2012)
46. E.A. Jones, A. van Remoortere, R.J.M. van Zeijl, P.C.W. Hogendoorn, J. Bovee, A.M. Deelder et al., Multiple statistical analysis techniques corroborate intratumor heterogeneity in imaging mass spectrometry datasets of myxofibrosarcoma. *Plos One* **6** (2011)
47. R.M. Parry, A.S. Galhena, C.M. Gamage, R.V. Bennett, M.D. Wang, F.M. Fernandez, omniSpect: an open MATLAB-based tool for visualization and analysis of matrix-assisted laser desorption/ionization and desorption electrospray ionization mass spectrometry images. *J. Am. Soc. Mass Spectrom.* **24**, 646–649 (2013)
48. A. Hyvärinen, E. Oja, Independent component analysis: algorithms and analysis. *Neural Networks* **13**, 411–430 (2000)
49. C.J.C. Burges, Dimension reduction: a guided tour. *Found. Trends® Mach. Learn.* **2**, 275–365 (2010)
50. L.A. Klerk, A. Broersen, I.W. Fletcher, R. van Lier, R.M.A. Heeren, Extended data analysis strategies for high resolution imaging MS: new methods to deal with extremely large image hyperspectral datasets. *Int. J. Mass Spectrom.* **260**, 222–236 (2007)
51. S.O. Deininger, M.P. Ebert, A. Futterer, M. Gerhard, C. Rocken, MALDI imaging combined with hierarchical clustering as a new tool for the interpretation of complex human cancers. *J. Proteome Res.* **7**, 5230–5236 (2008)
52. A.M. Oros-Peusquens, A. Matusch, J.S. Becker, N.J. Shah, Automatic segmentation of tissue sections using the multielement information provided by LA-ICP-MS imaging and k-means cluster analysis. *Int. J. Mass Spectrom.* **307**, 245–252 (2011)

53. T. Alexandrov, M. Becker, S.O. Deininger, G. Ernst, L. Wehder, M. Grasmair et al., Spatial segmentation of imaging mass spectrometry data with edge-preserving image denoising and clustering. *J. Proteome Res.* **9**, 6535–6546 (2010)
54. T. Alexandrov, MALDI imaging mass spectrometry: statistical data analysis and current computational challenges. *BMC Bioinf.* **13**, S11 (2012)
55. P. Berkhin, Survey of clustering data mining techniques. *Accrue Software, Inc. Technical Report*, (2002)
56. L.A. McDonnell, A. van Remoortere, R.J.M. van Zeijl, A.M. Deelder, Mass spectrometry image correlation: quantifying colocalization. *J. Proteome Res.* **7**, 3619–3627 (2008)
57. C.D. Kaddi, R.M. Parry, M.D. Wang, Hypergeometric similarity measure for spatial analysis in tissue imaging mass spectrometry. *Proceedings of IEEE BIBM 2011*, pp. 604–607, (2012)
58. C.D. Kaddi, R.M. Parry, M.D. Wang, Multivariate hypergeometric similarity measure. *IEEE/ACM Trans. Comput. Biol. Bioinform.* **10**, 1505–16 (2013)
59. R. van de Plas, K. Pelckmans, B. De Moor, W. E., Spatial querying of imaging mass spectrometry data: a nonnegative least squares approach. *Neural Inf. Proc. Syst. Workshop Mach. Learn. Comput. Biol.* (2007)

Health Informatics Data Analysis

Methods and Examples

Xu, D.; Wang, M.D.; Zhou, F.; Cai, Y. (Eds.)

2017, X, 210 p. 54 illus., Hardcover

ISBN: 978-3-319-44979-1