

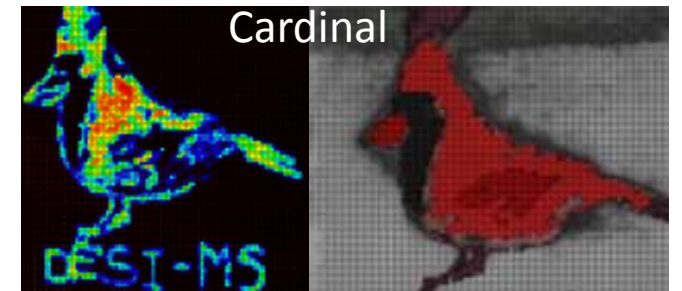
Mass spectrometry imaging analysis with Cardinal

Melanie C. Föll

Northeastern University
Khoury College of Computer Sciences

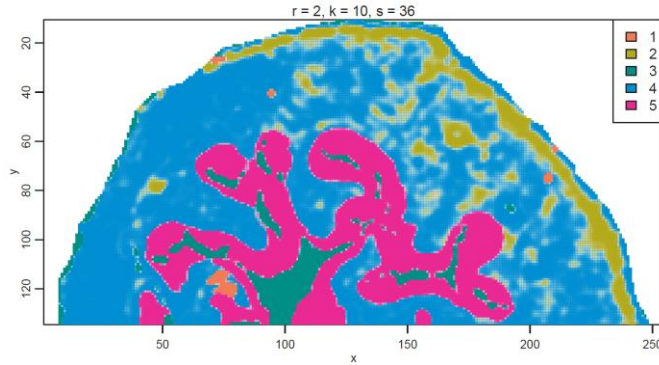


Northeastern University

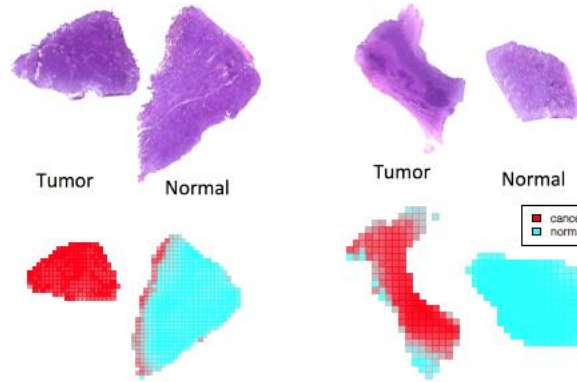


Mass spectrometry imaging

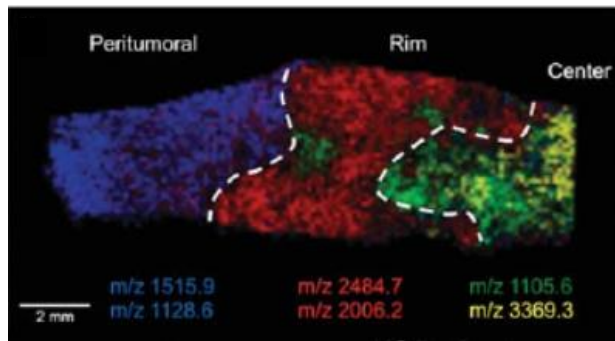
Tissue histology



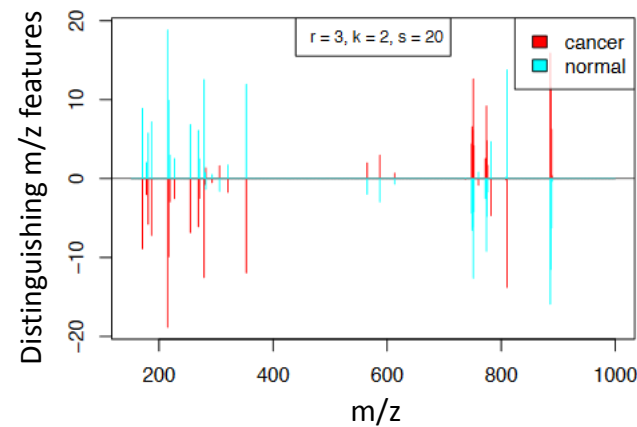
Biomarker discovery



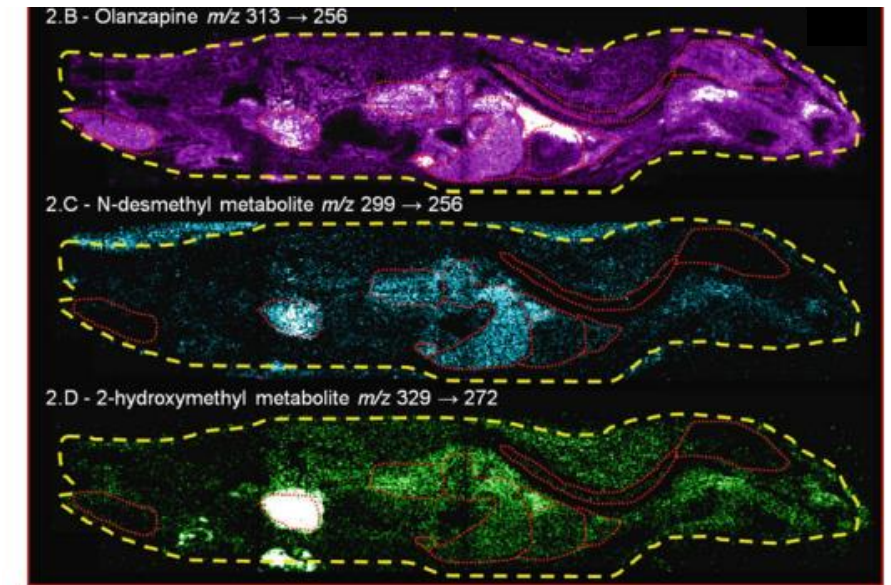
Tissue heterogeneity



Turtoi 2014, Hepatology

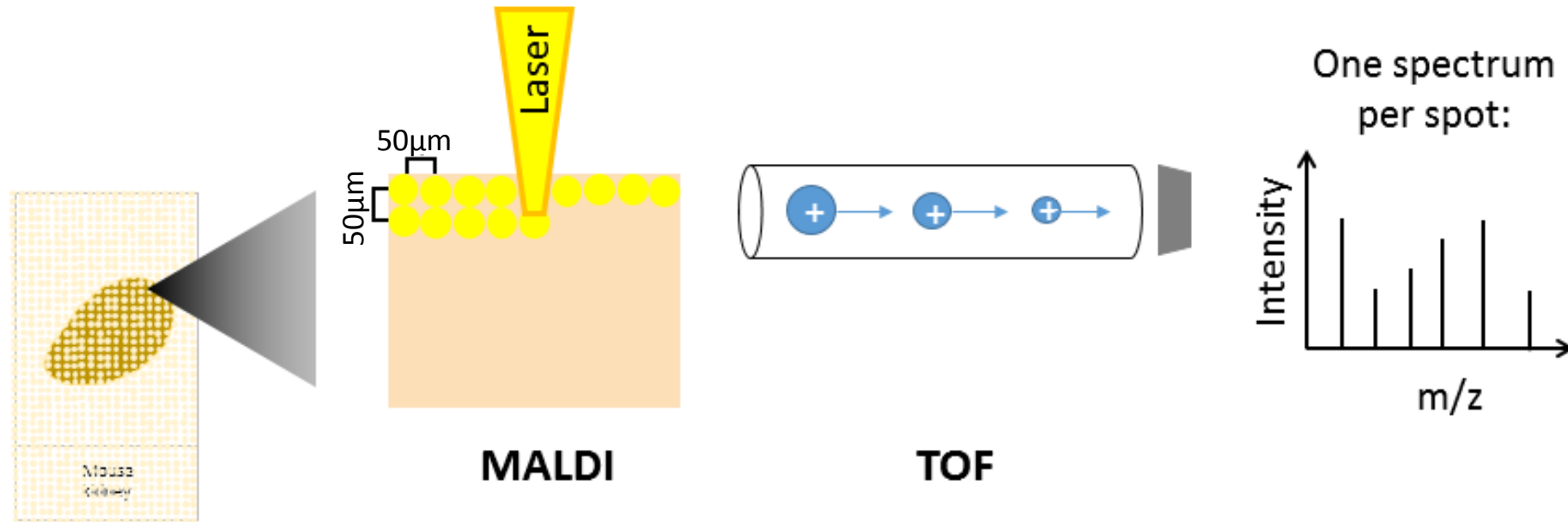


Drug imaging

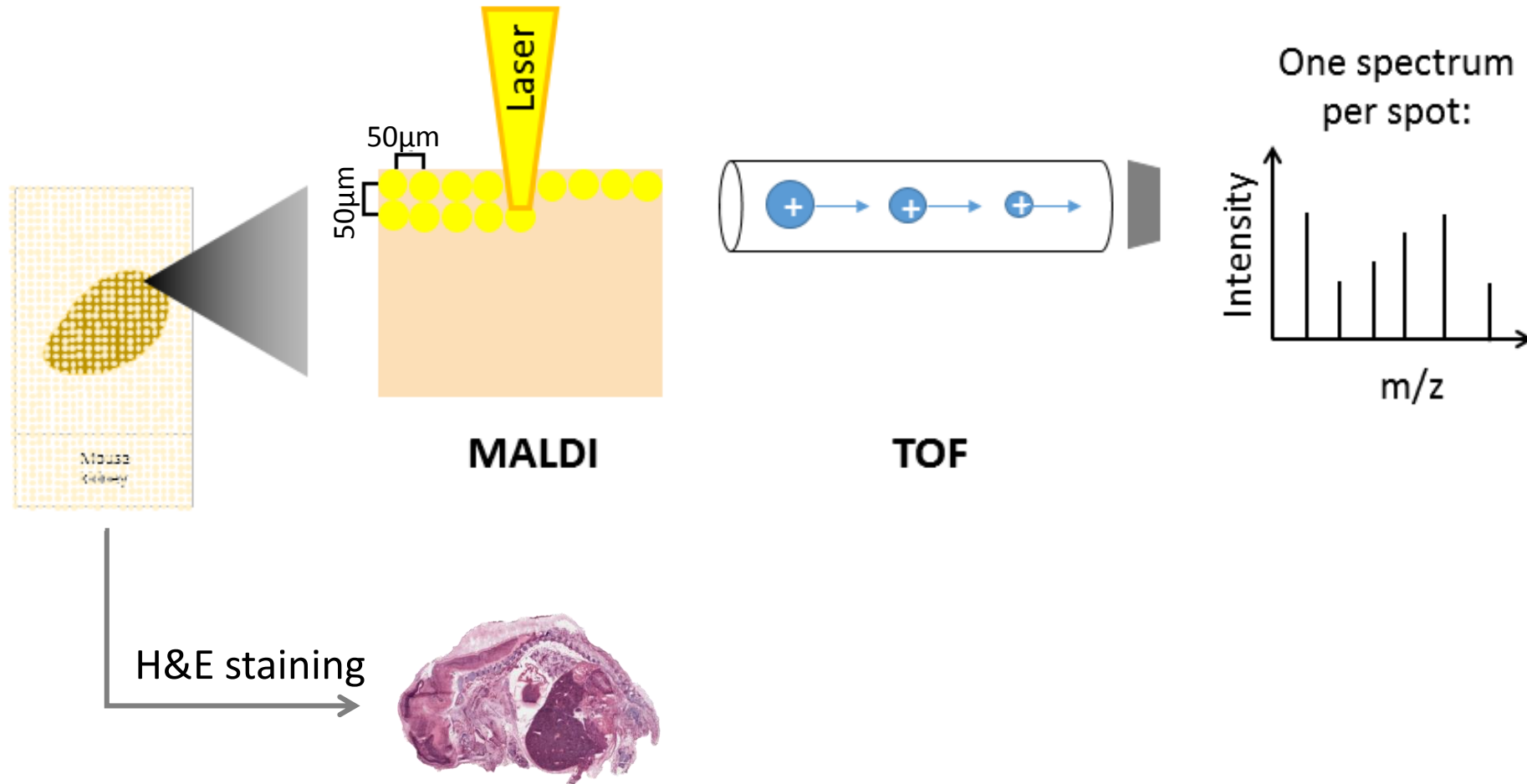


Khatib-Shahidi 2006, Anal. Chem.

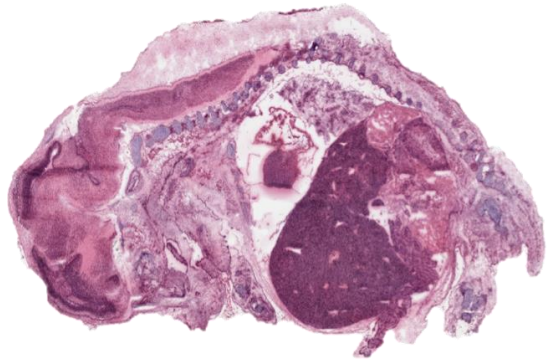
Data acquisition



Data acquisition



From spectra to distribution profiles

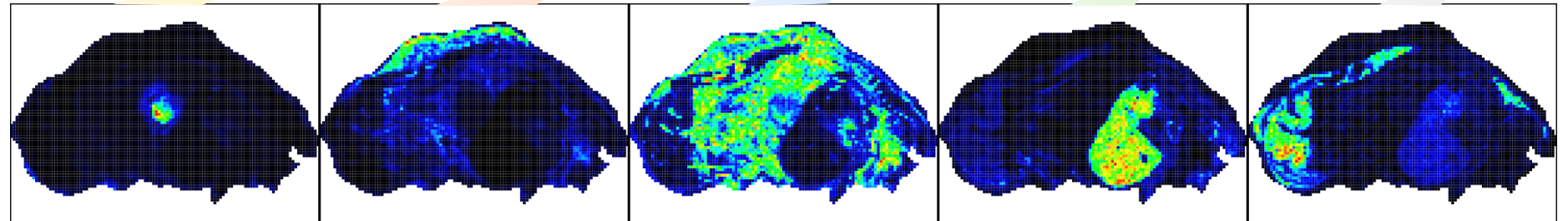
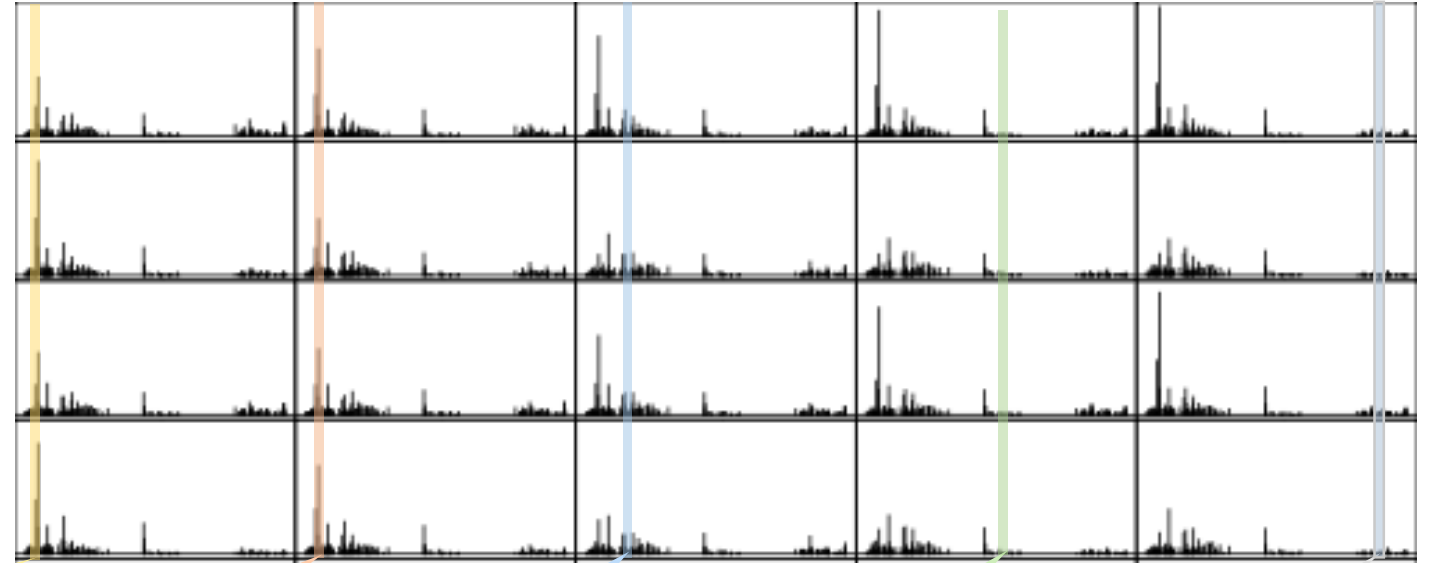


Spectra 1

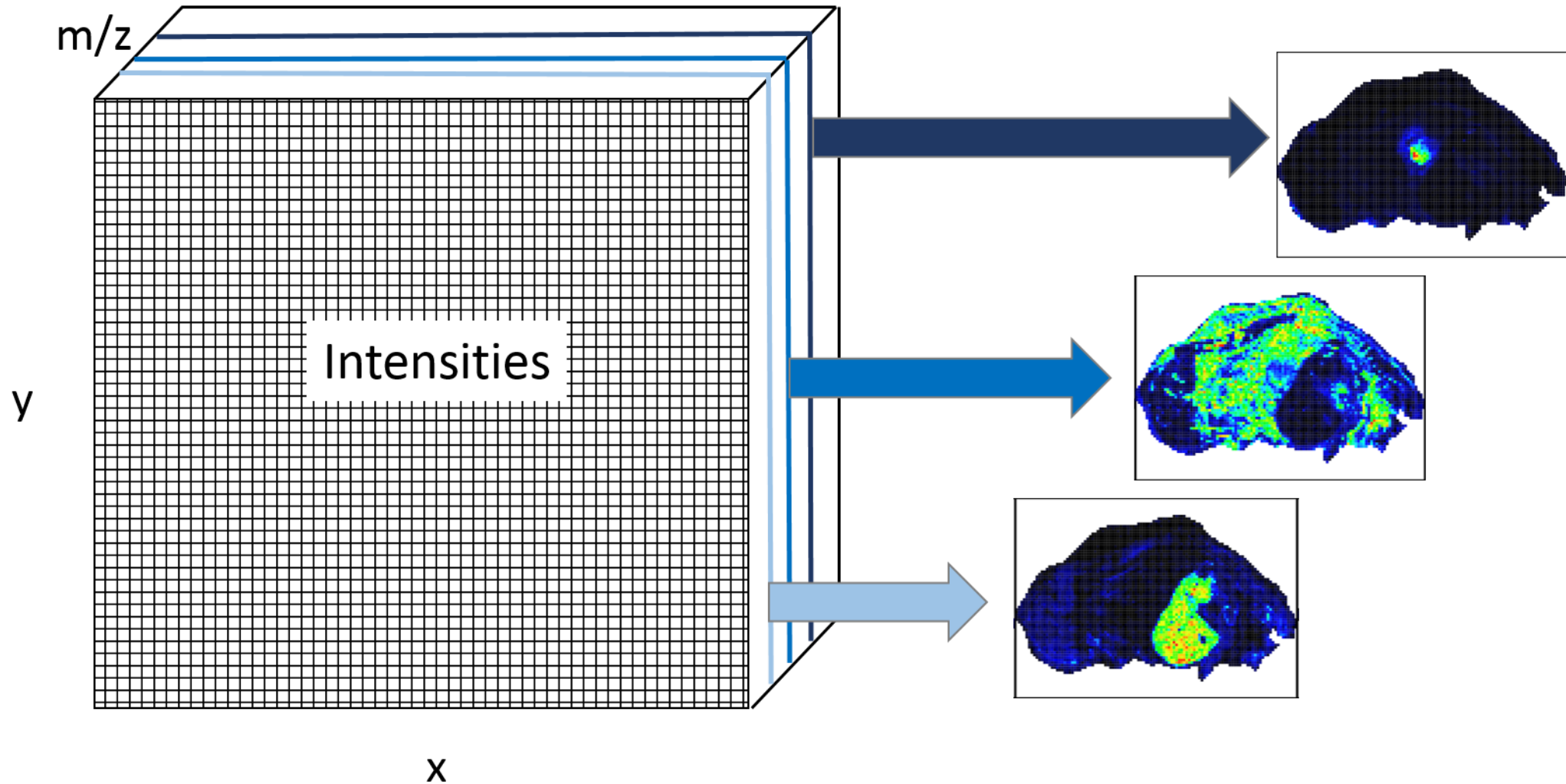
Spectra 2

Spectra 3

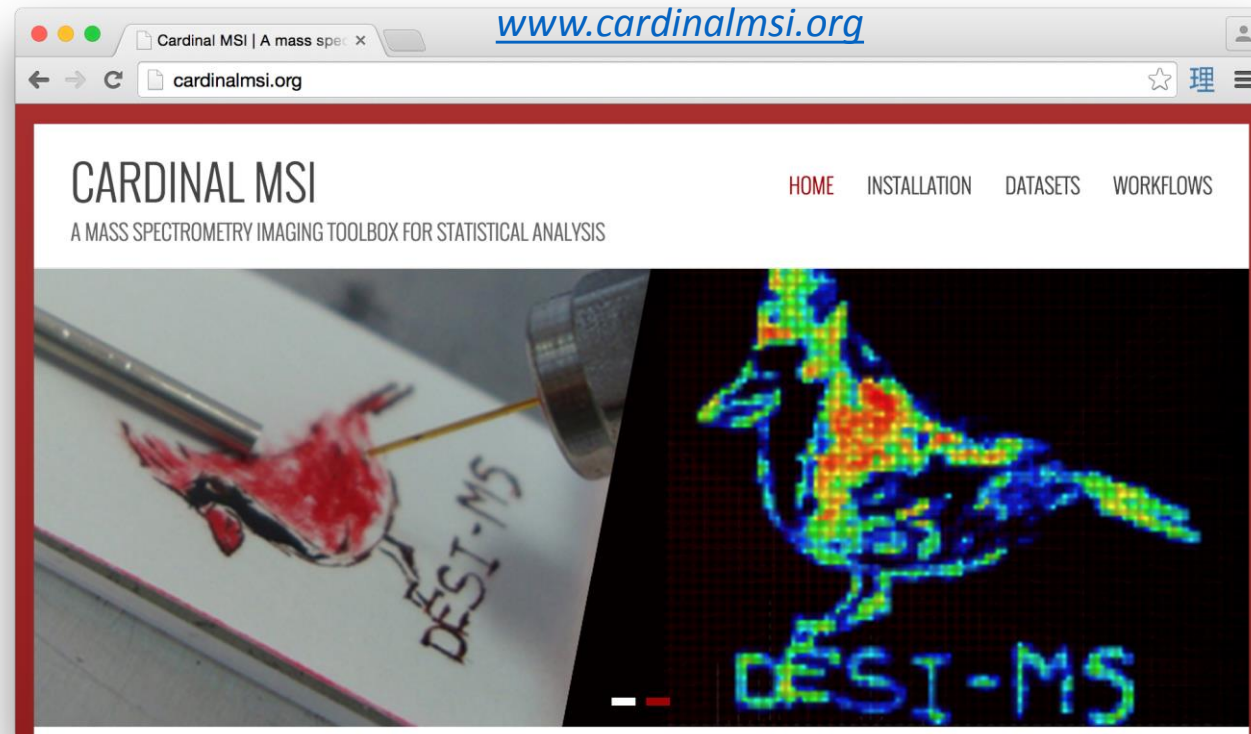
Spectra n



From spectra to distribution profiles



Cardinal: open-source statistical software for MSI



- Free, open-source
- R-based
- Rich methods portfolio
- Compatible with larger than memory datasets
- Extensive documentation
- Active mailing list

- Over 7,000 unique downloads since public release in 2015
- 2015 John M. Chambers Statistical Software Award

Bioconductor



Software

K. Bemis et al. Bioinformatics, 2015.

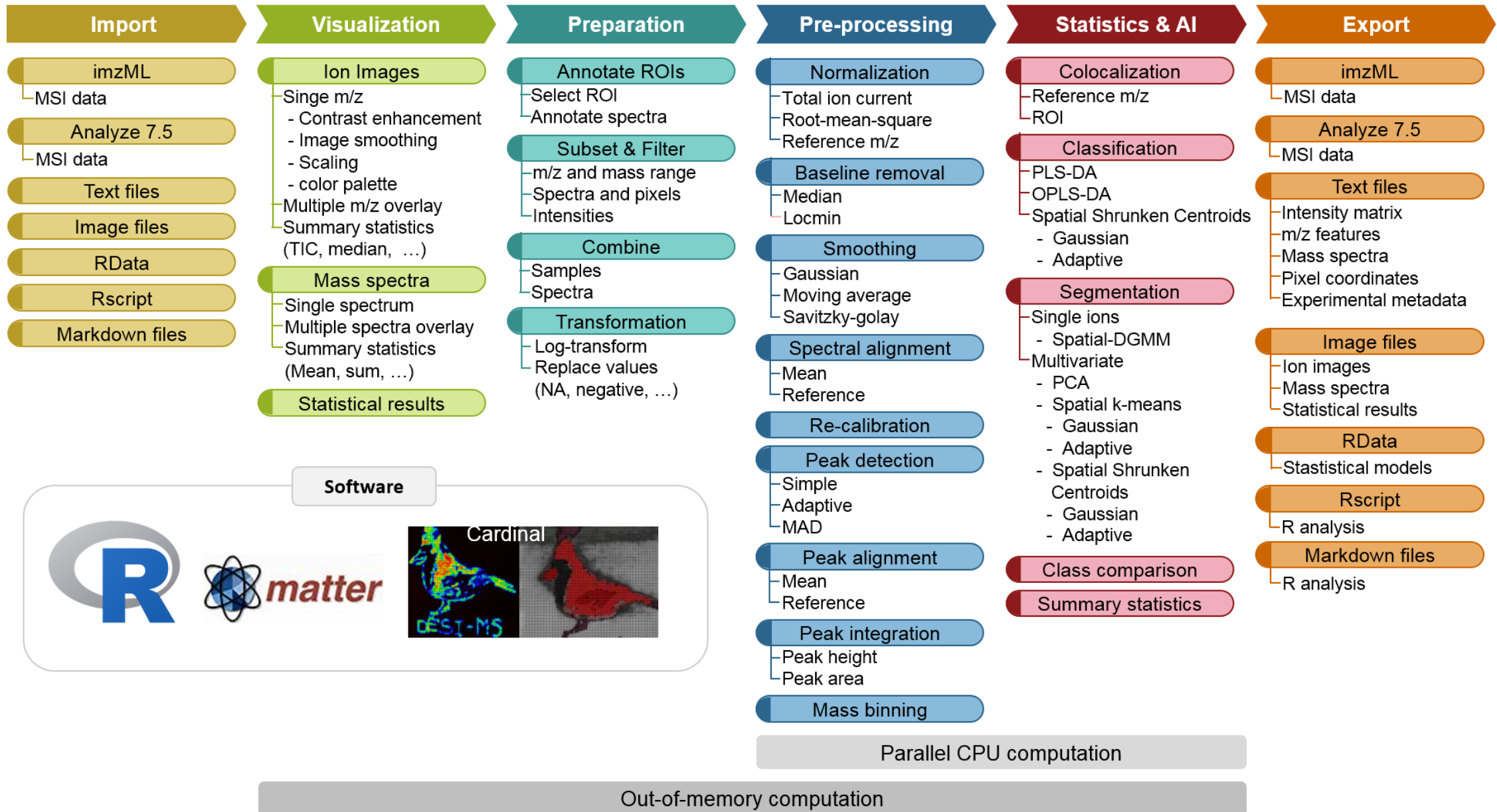
Statistical methods

K. Bemis et al. Molecular and Cellular Proteomics, 2016.

Large-than memory data

K. Bemis et al. Bioinformatics, 2017.

Cardinal: Methods portfolio

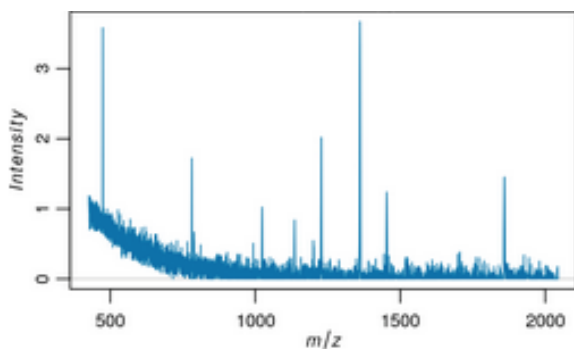


Visualizations

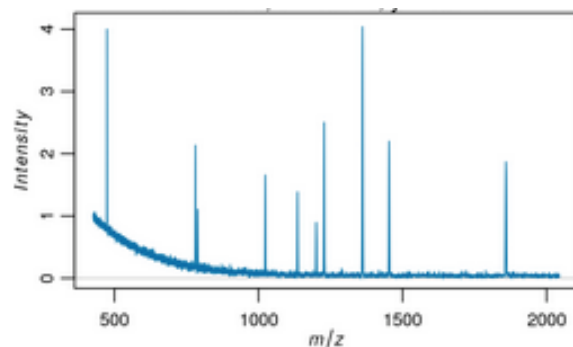
Mass spectra

`plot(mse, ...)`

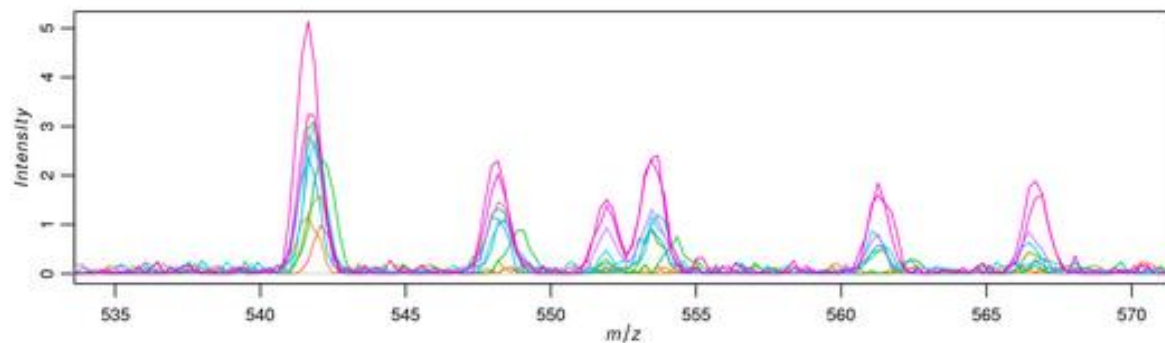
Single mass spectra



Average mass spectra



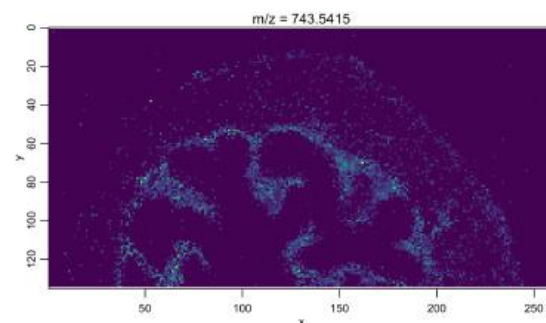
Zoomed in & overlaid mass spectra



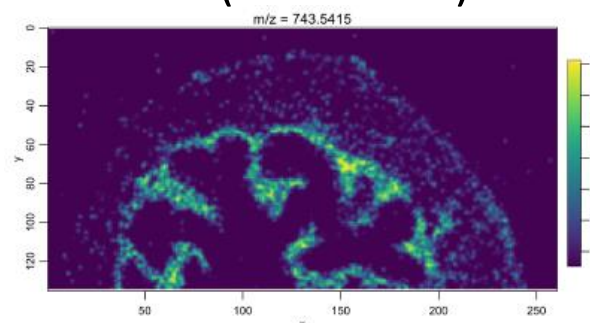
Ion images

`image(mse, ...)`

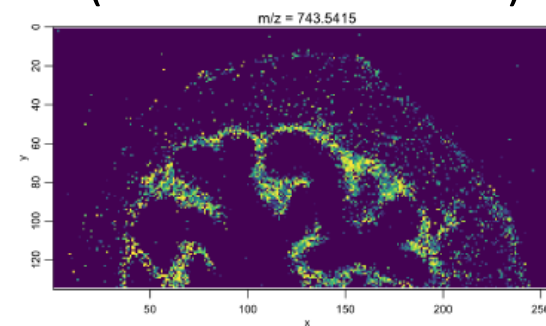
Ion image



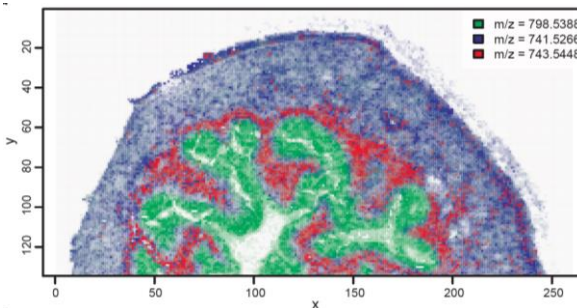
Ion image (smoothed)



Ion image (contrast enhanced)

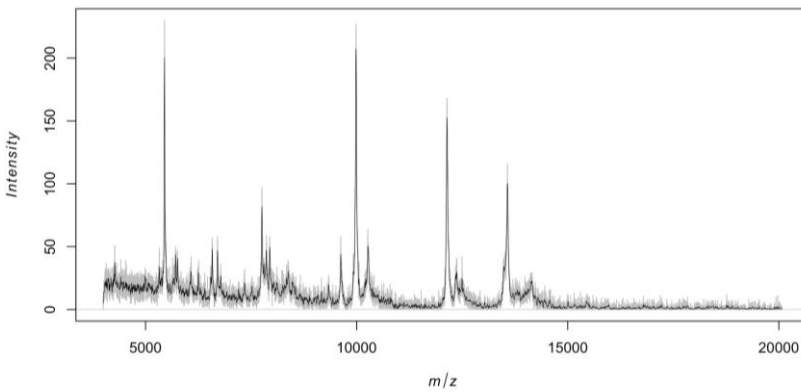


Ion image overlay of three m/z



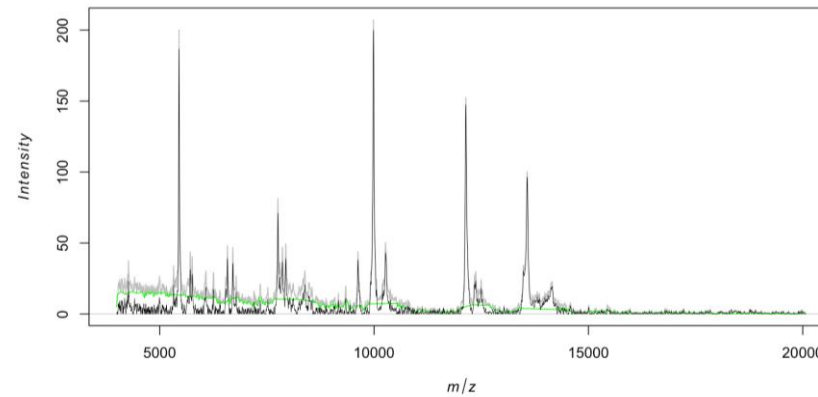
Pre-processing

Smoothing



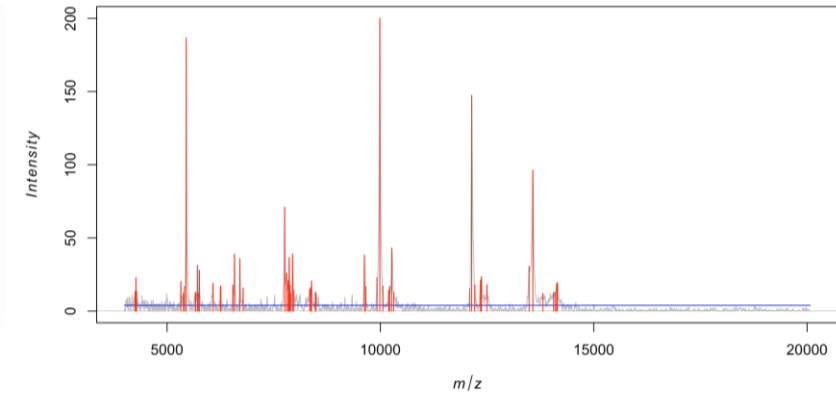
`smoothSignal(mse, ...)`

Baseline removal



`reduceBaseline(mse, ...)`

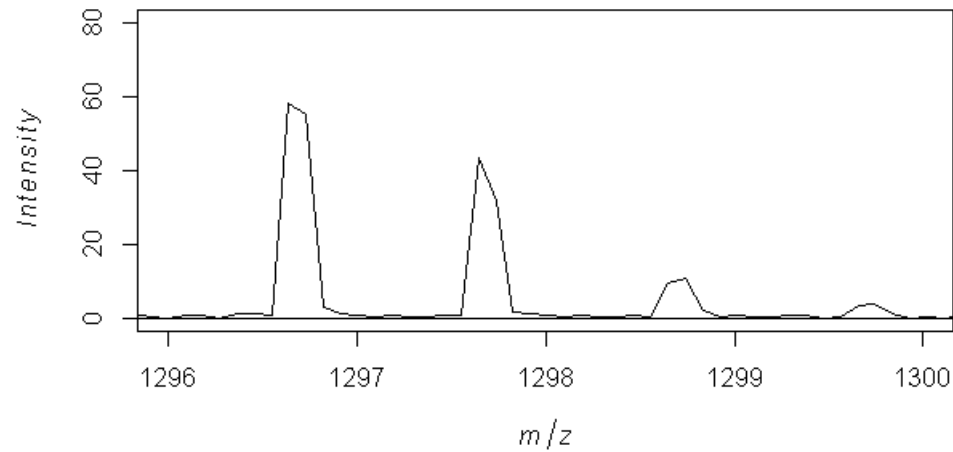
Peak detection



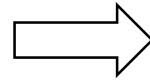
`peakPick(mse, ...)`

Pre-processing

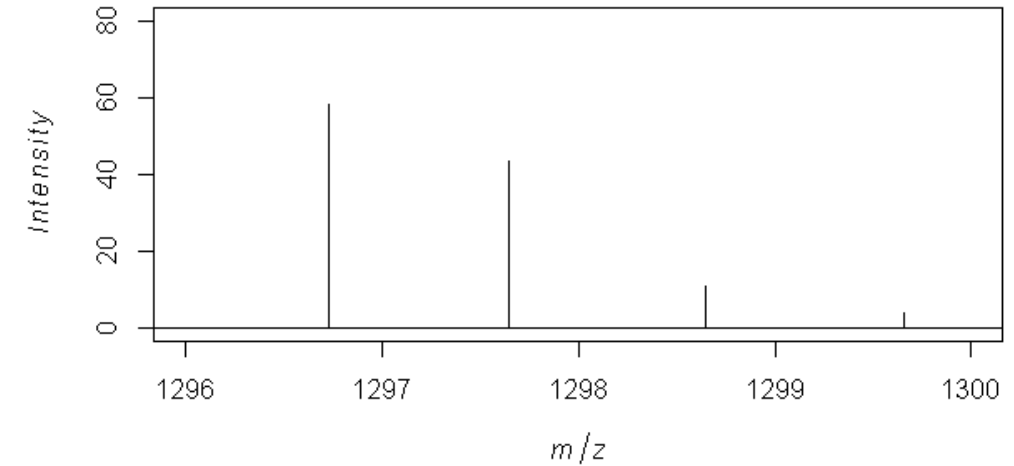
Profile spectrum



`centroided(mse) = FALSE`

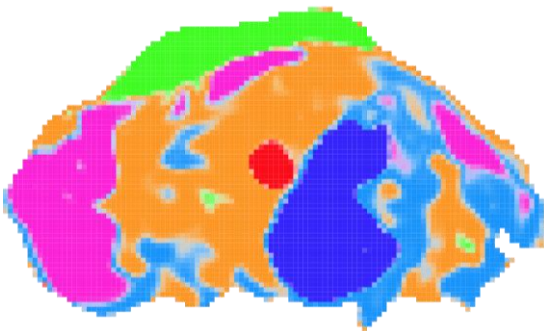
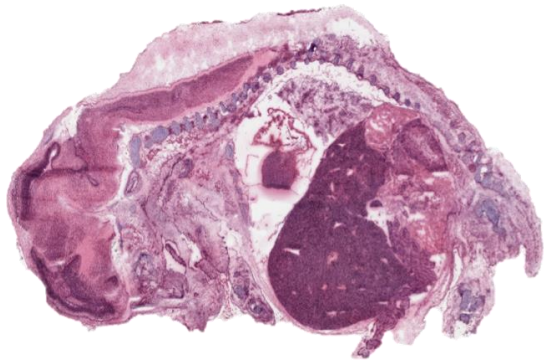


Centroided spectrum



`centroided(mse) = TRUE`

1. Class discovery (segmentation)

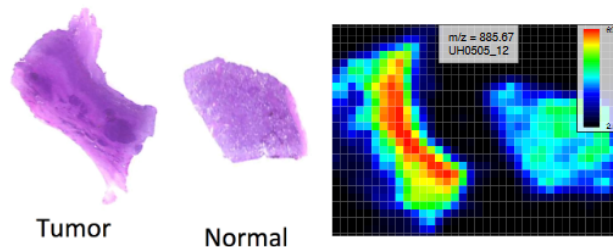
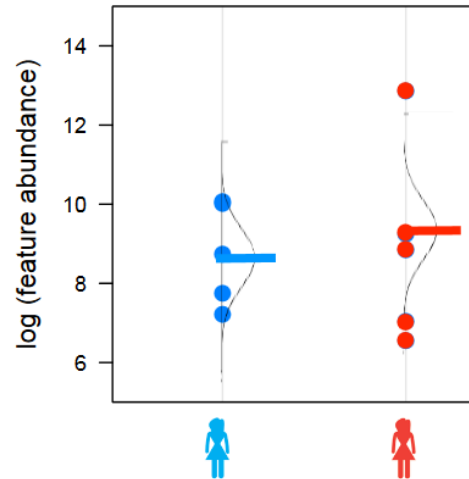


`spatialShrunkenCentroids(mse, ...)`

`spatialDGMM(mse, ...)`

....

2. Class comparison (linear models)

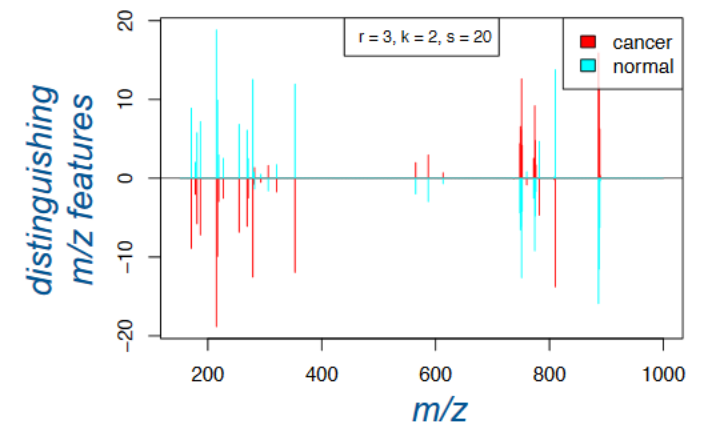
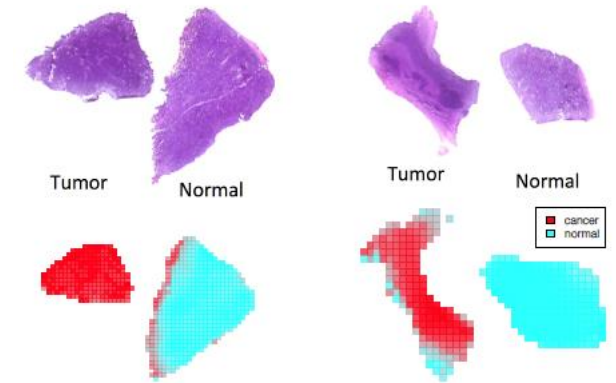


`meansTest(mse, ...)`

`segmentationTest(mse, ...)`

....

3. Class prediction (classification)



`spatialShrunkenCentroids(mse, ...)`

....

MSI data import

& export

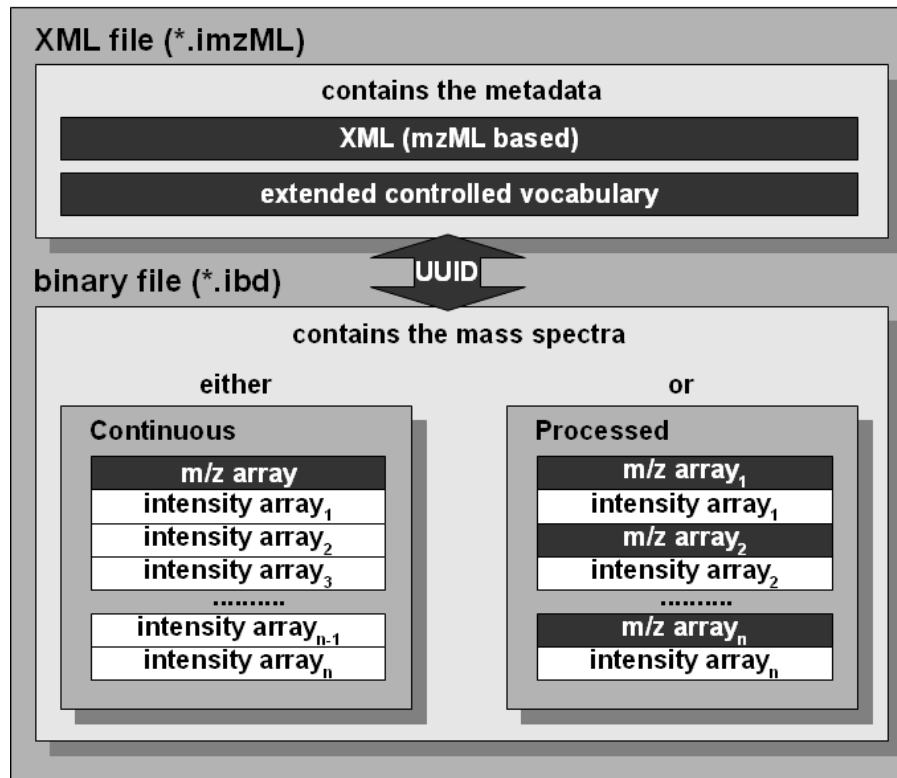
readImzML(mse, ...)

writeImzML(mse, ...)

imzML: open standard file format for MSI data

File1: .imzML

File2: .ibd



MSContinuous

MSProcessed

ImagingExperiment

ImagingExperiment

imzML converter:

<https://www.cs.bham.ac.uk/~ibs/imzMLConverter>

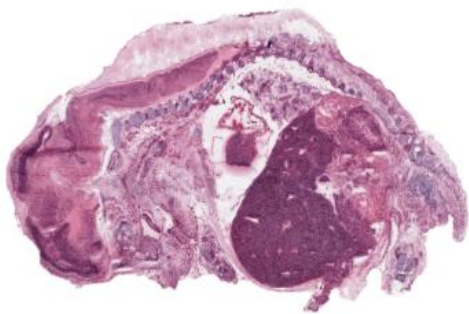
Race, Alan M., Iain B. Styles, and Josephine Bunch. "Inclusive sharing of mass spectrometry imaging data requires a converter for all." *Journal of proteomics* 75.16 (2012): 5111-5112.

imzML validator:

<https://gitlab.com/imzML/imzMLValidator>

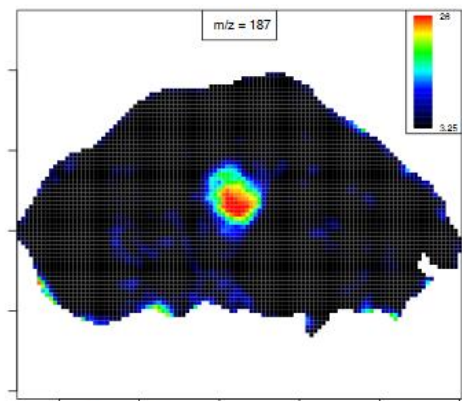
Race, Alan M., and Andreas Römpf. "Error-Free Data Visualization and Processing through imzML and mzML Validation." *Analytical chemistry* 90.22 (2018): 13378-13384..

Today's dataset & analysis

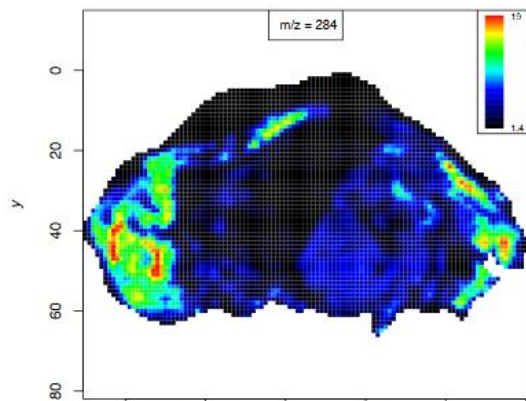


Optical image

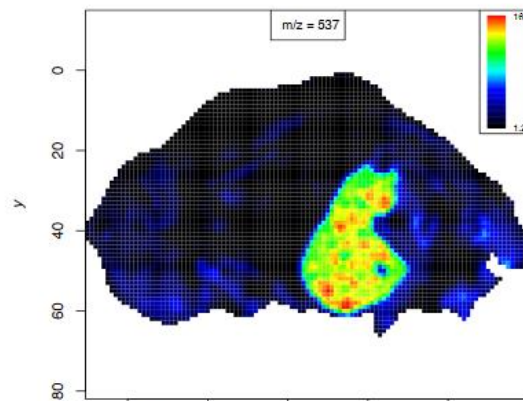
- Pig fetus section
- DESI LTQ mass spectrometer
- m/z range 150-1000



Peak associated
with heart



Peak associated
with brain



Peak associated
with liver

Loading data

Exploring data

Pre-processing

Segmentation

Visualization

Vignette:

<http://bioconductor.org/packages/release/data/experiment/vignettes/CardinalWorkflows/inst/doc/MSI-segmentation.html>

R

Part I: Loading & exploring MSI data

Loading: 75%



Loading data

Exploring data

Pre-processing

Segmentation

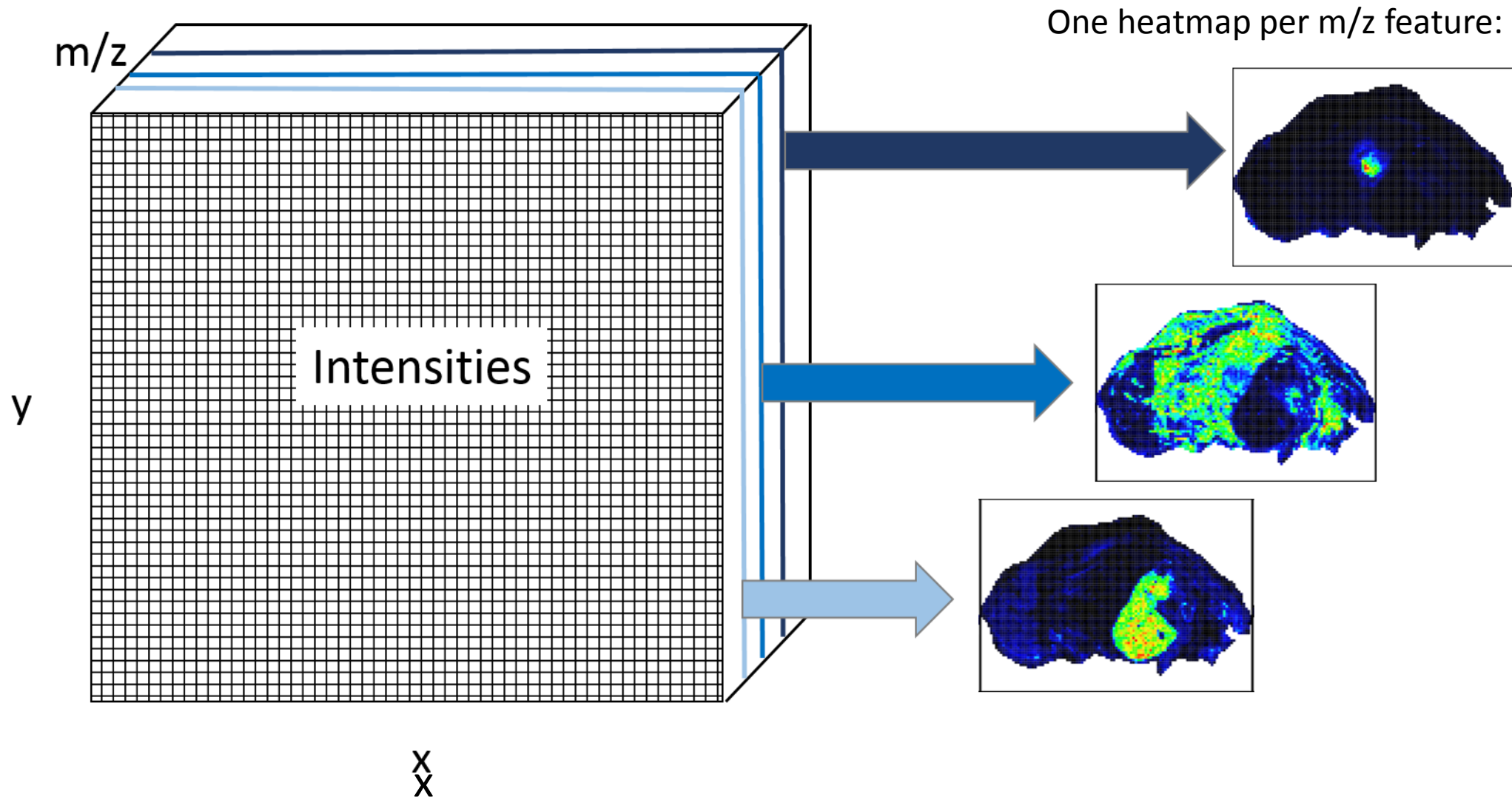
Visualization

Importing MSI data

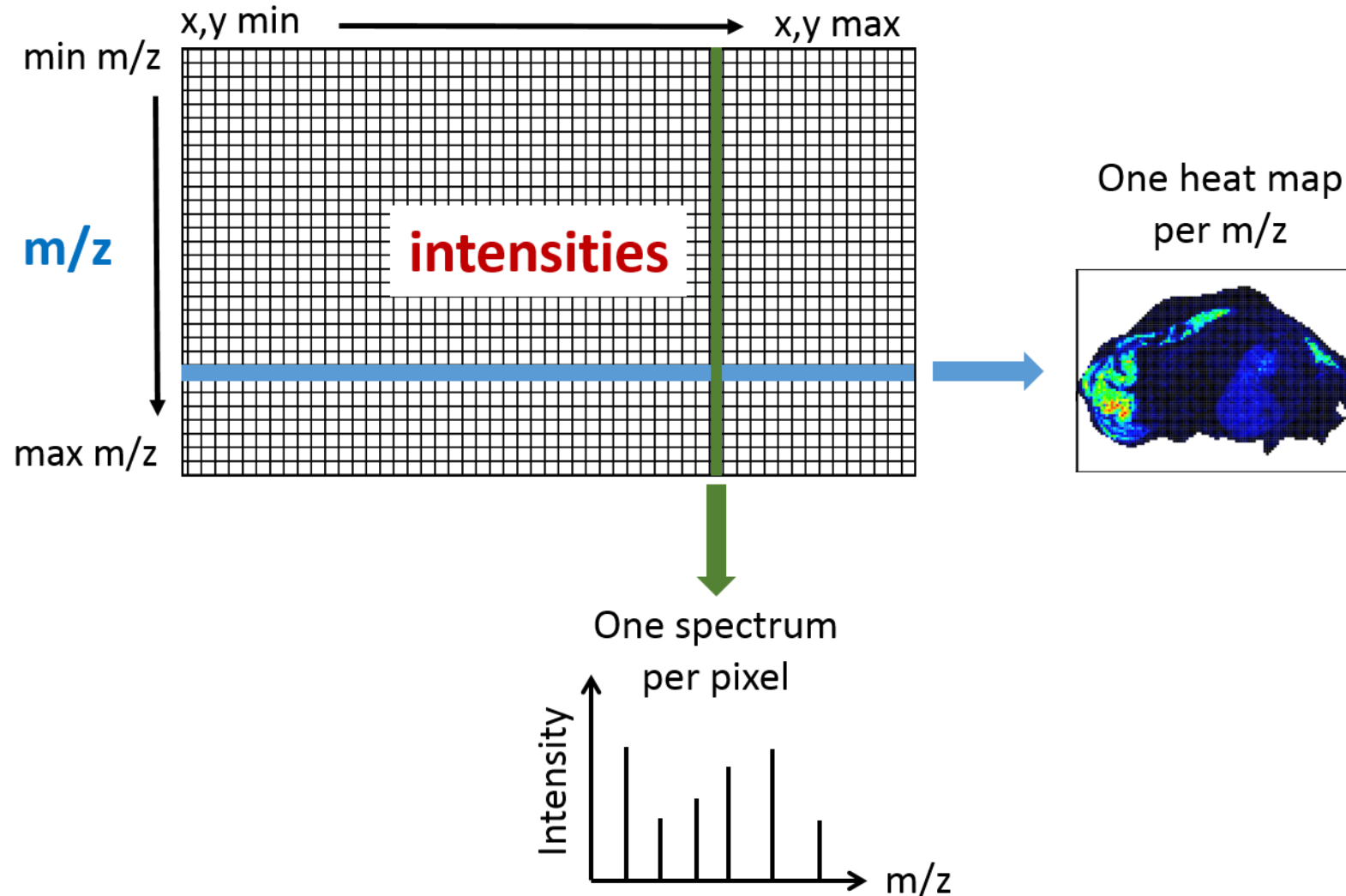
- From CardinalWorkflows package
 - Object type: MSIImageSet (old format)
 - Coercion to MSIImagingExperiment (current format)



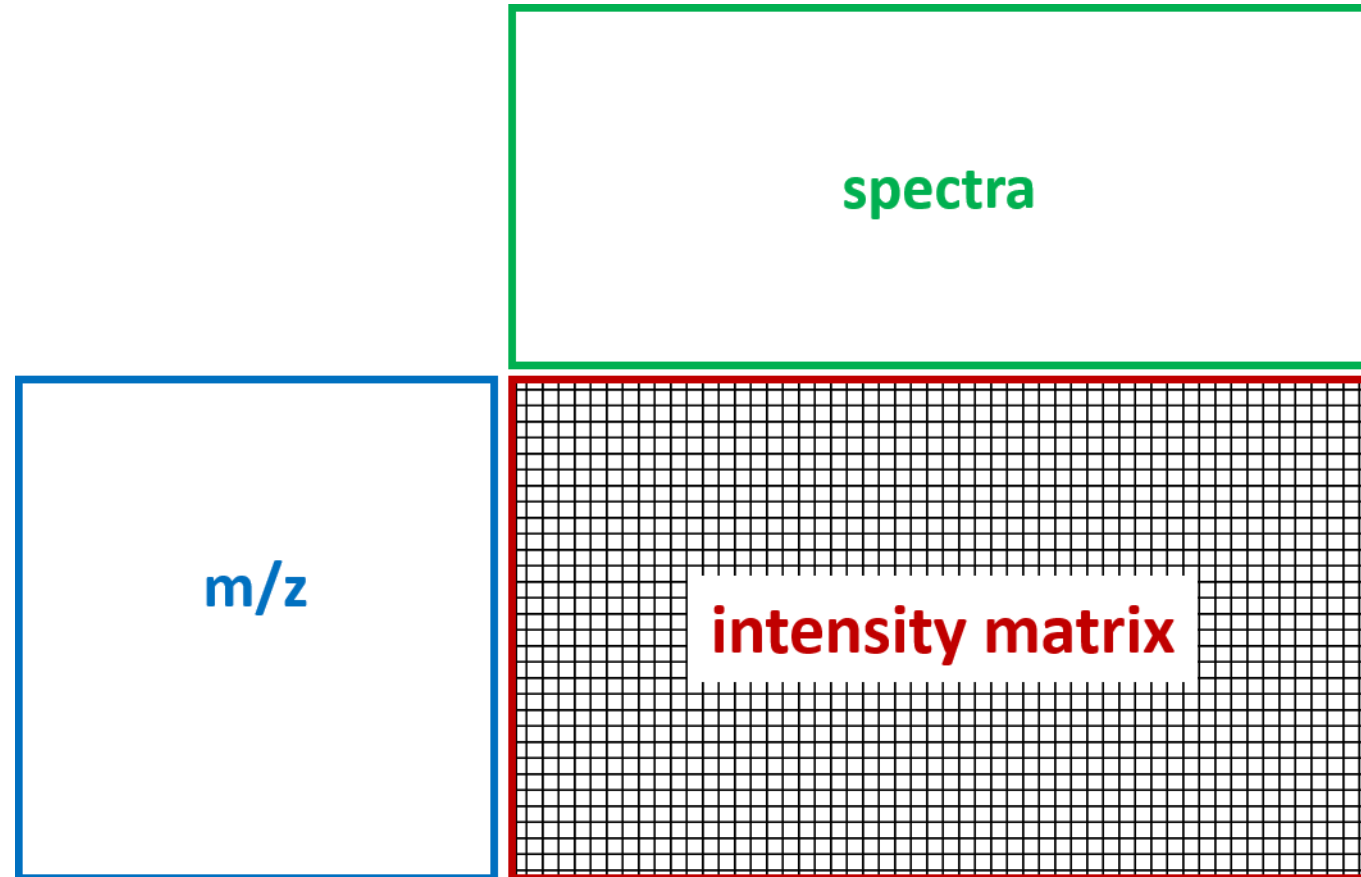
MSI data structure – 3D presentation



MSI data structure – 2D presentation



MSImagingExperiment – basic data structure



MSImagingExperiment – basic data structure

spectra

```
pixelData(mse) = pData(mse)  
coord(mse)  
run(mse)
```

m/z

```
mz(mse)  
featureData(mse)  
= fData(mse)
```

intensity matrix

```
imageData(mse)[[1]]  
= iData(mse)  
spectraData(mse)[[1]] =  
spectra(mse)
```

Part II: Pre-processing



Loading data



Exploring data



Pre-processing

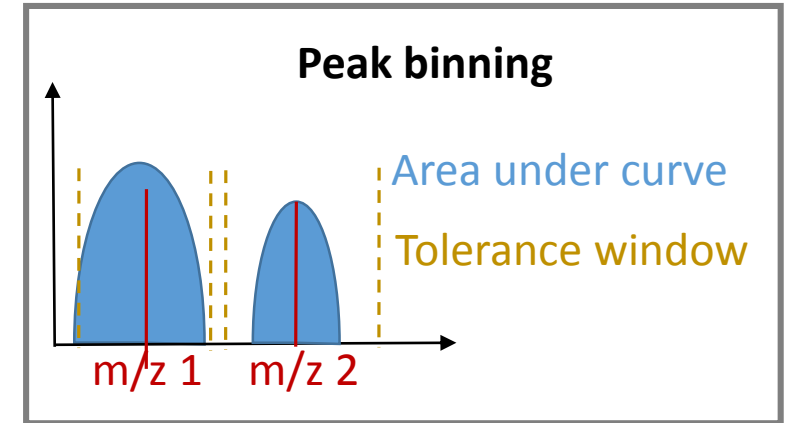
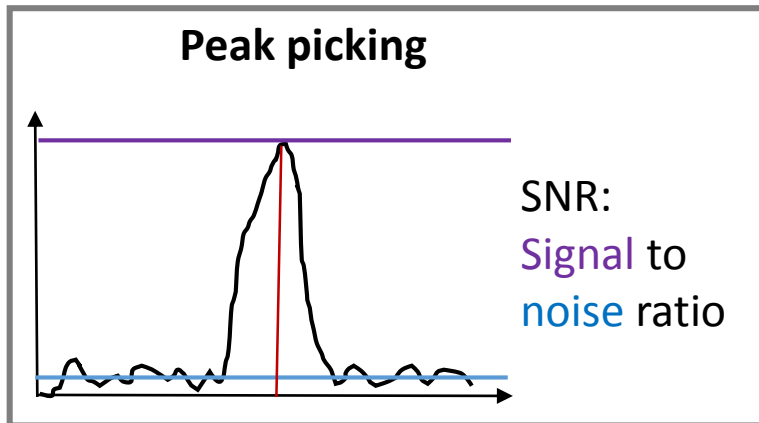
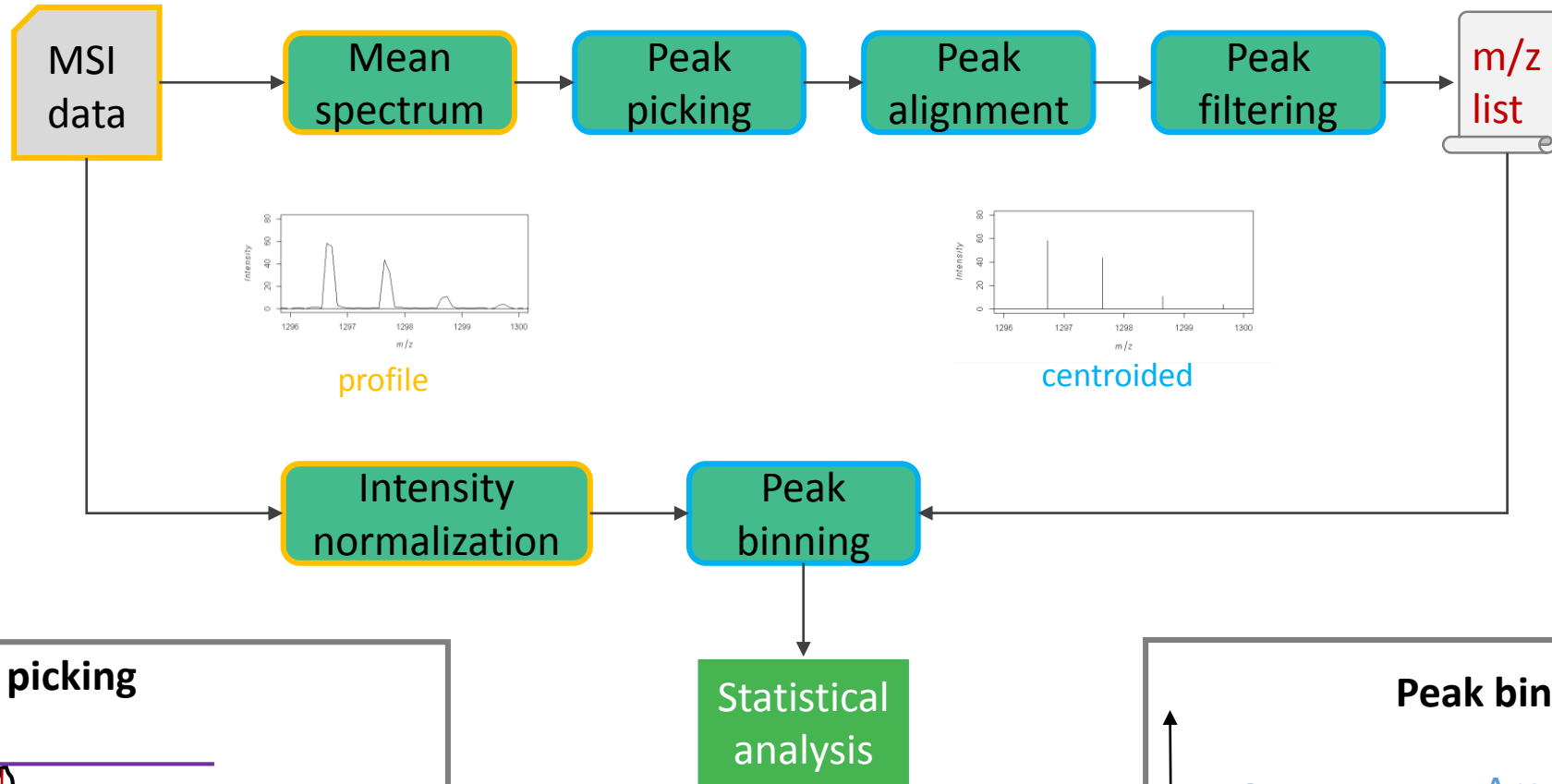


Segmentation



Visualization

Pre-processing workflow



Part III: Segmentation & visualization



Loading data



Exploring data



Pre-processing



Segmentation



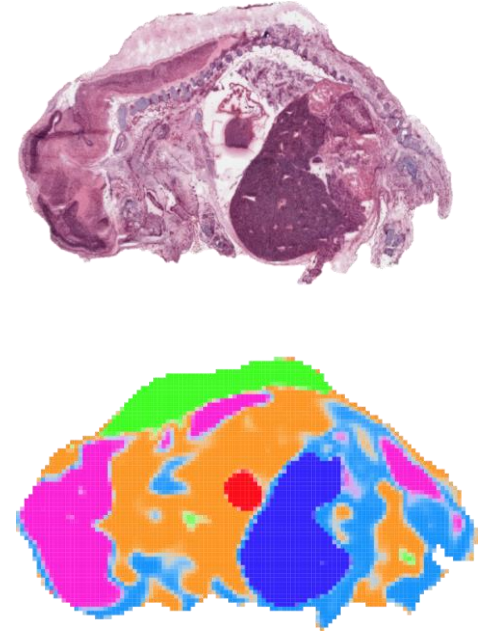
Visualization

Unsupervised segmentation

Spatial shrunken centroids:

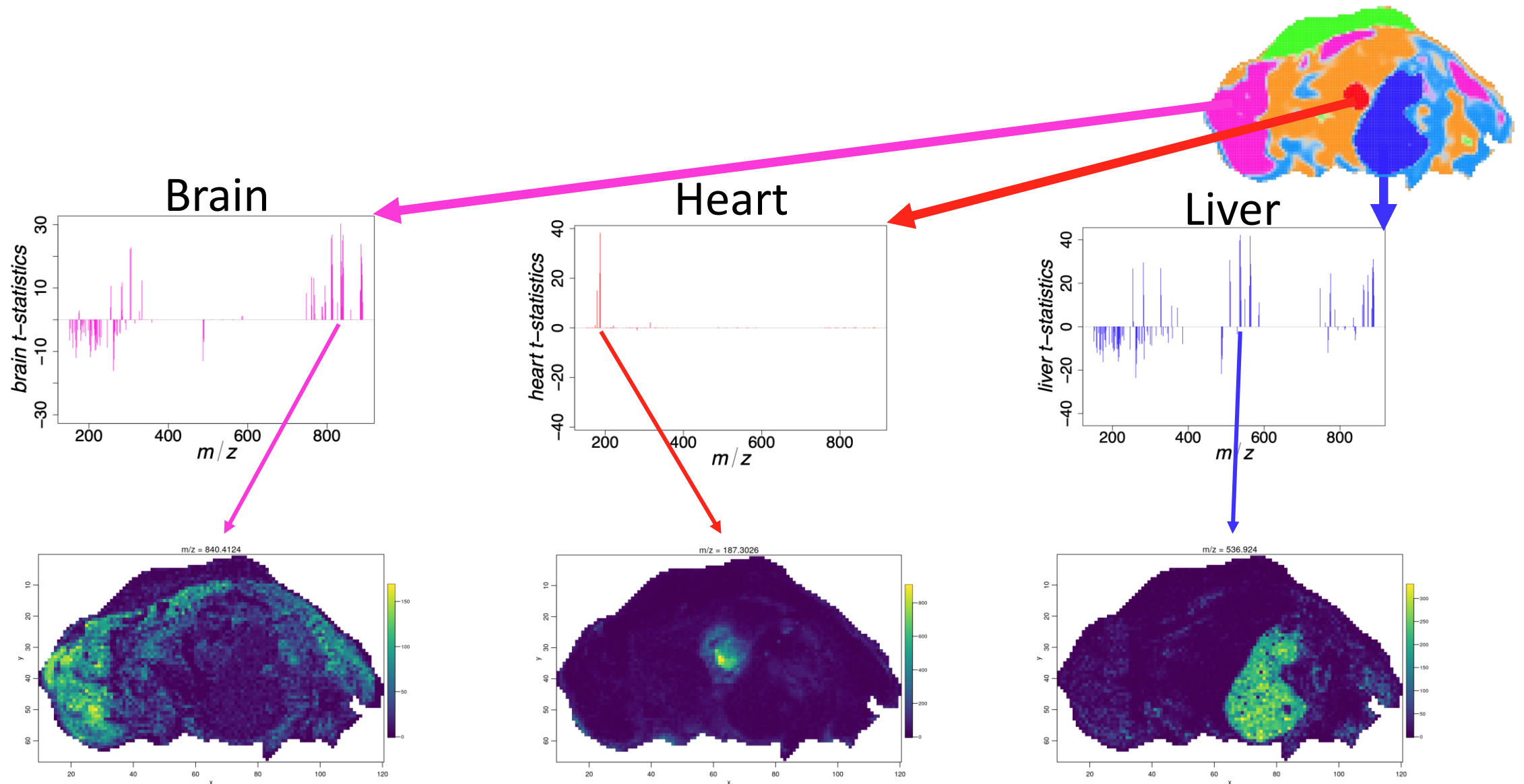
spatially-aware segmentation with feature selection

- Data-driven selection of appropriate number of segments
- Selects most important m/z for distinguishing segments
- Probability model characterizes uncertainty

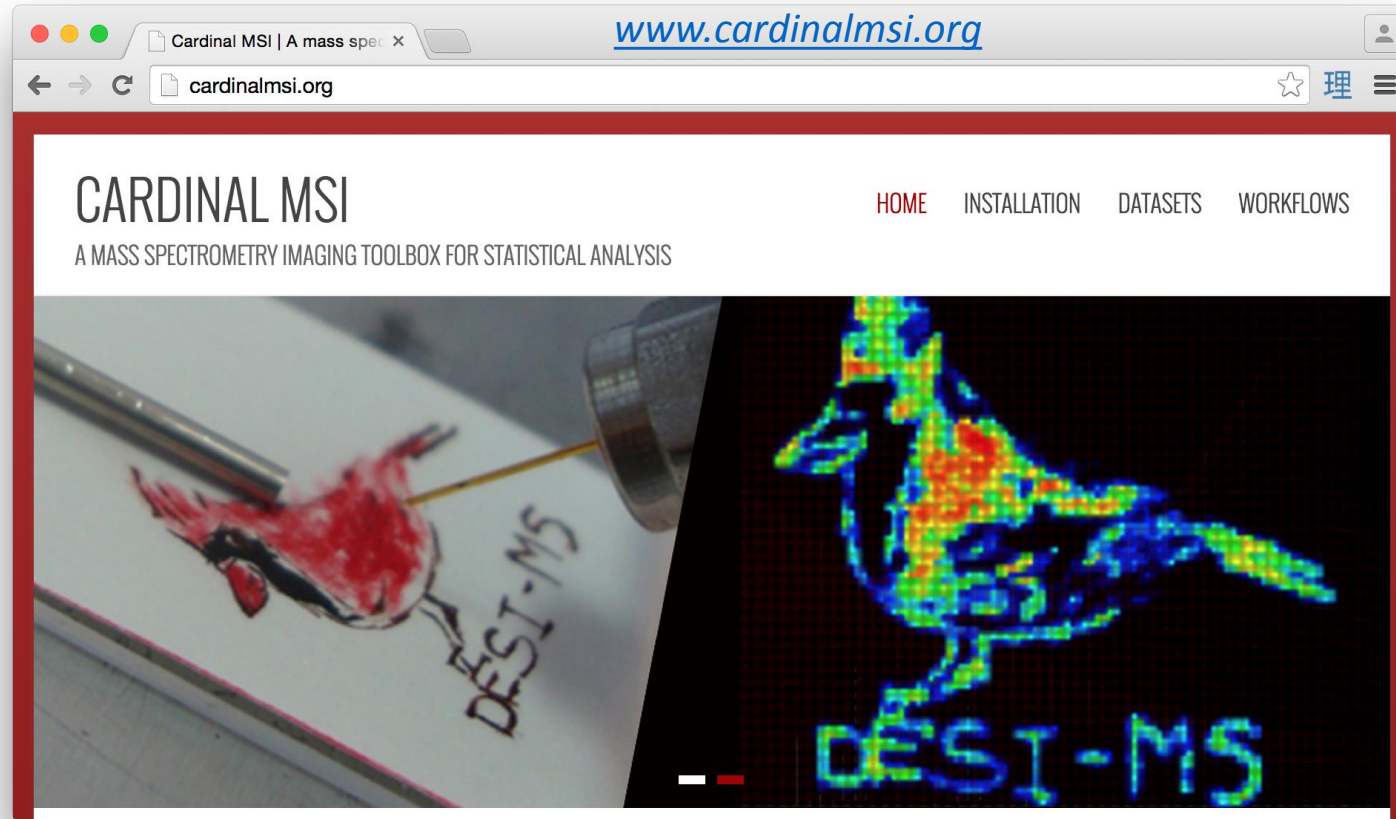


K. Bemis, et al. "Probabilistic segmentation of mass spectrometry (MS) images helps select important ions and characterize confidence in the resulting segments." *Molecular & Cellular Proteomics* 15.5 (2016): 1761-1772.

Visualization of most important m/z



Cardinal: resources & thank you



Kylie Bemis



Olga Vitek



Dan Guo



Sai Lakkimsetty

User guides/vignettes

Videos May Institute

Help forum

<http://cardinalmsi.org/workflows/>

<https://www.youtube.com/c/MayInstituteNEU/playlists>

<http://groups.google.com/d/forum/CardinalMSI>