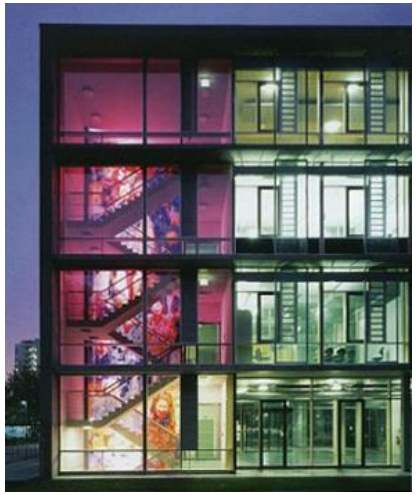
Galaxy wishlist from the mass spectrometry imaging (MSI) community

Melanie Föll

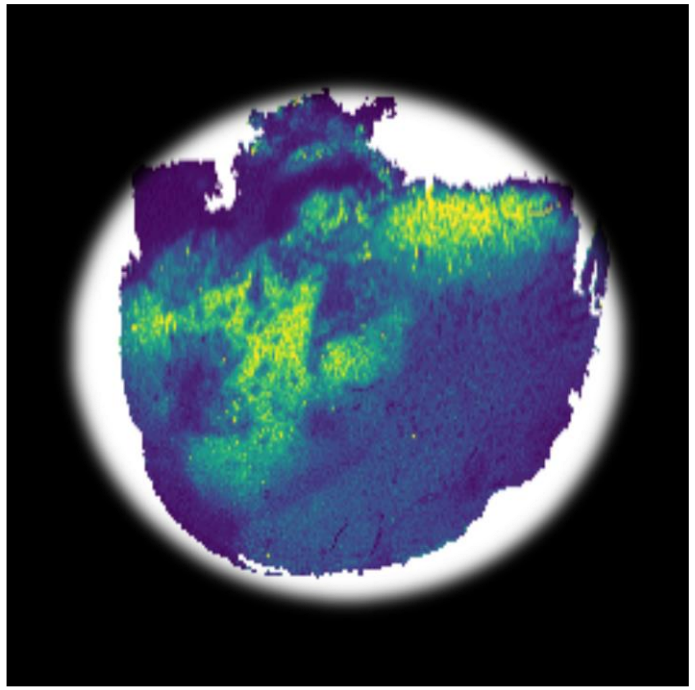
melanie.foell@mol-med.uni-freiburg.de

@MCFoell

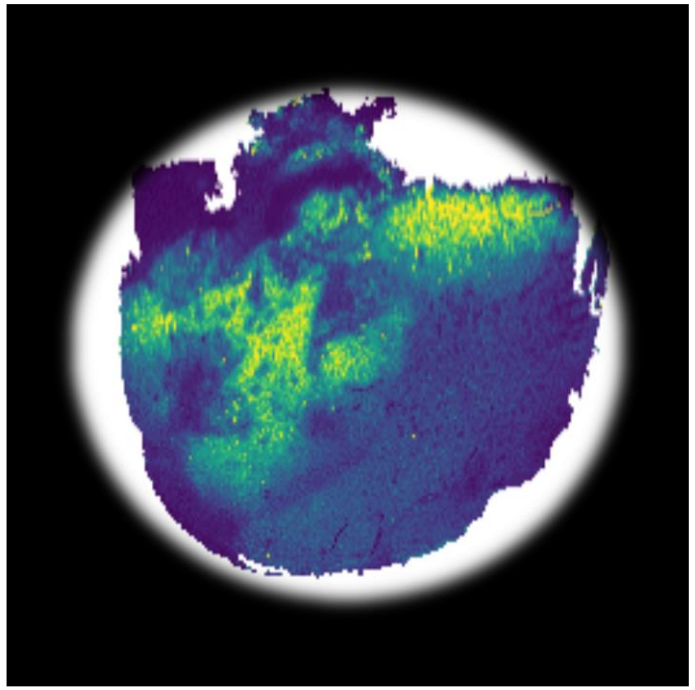
Institute for Surgical Pathology
University Medical Center Freiburg, Germany
&
Khoury College of Computer Sciences
Northeastern University, USA



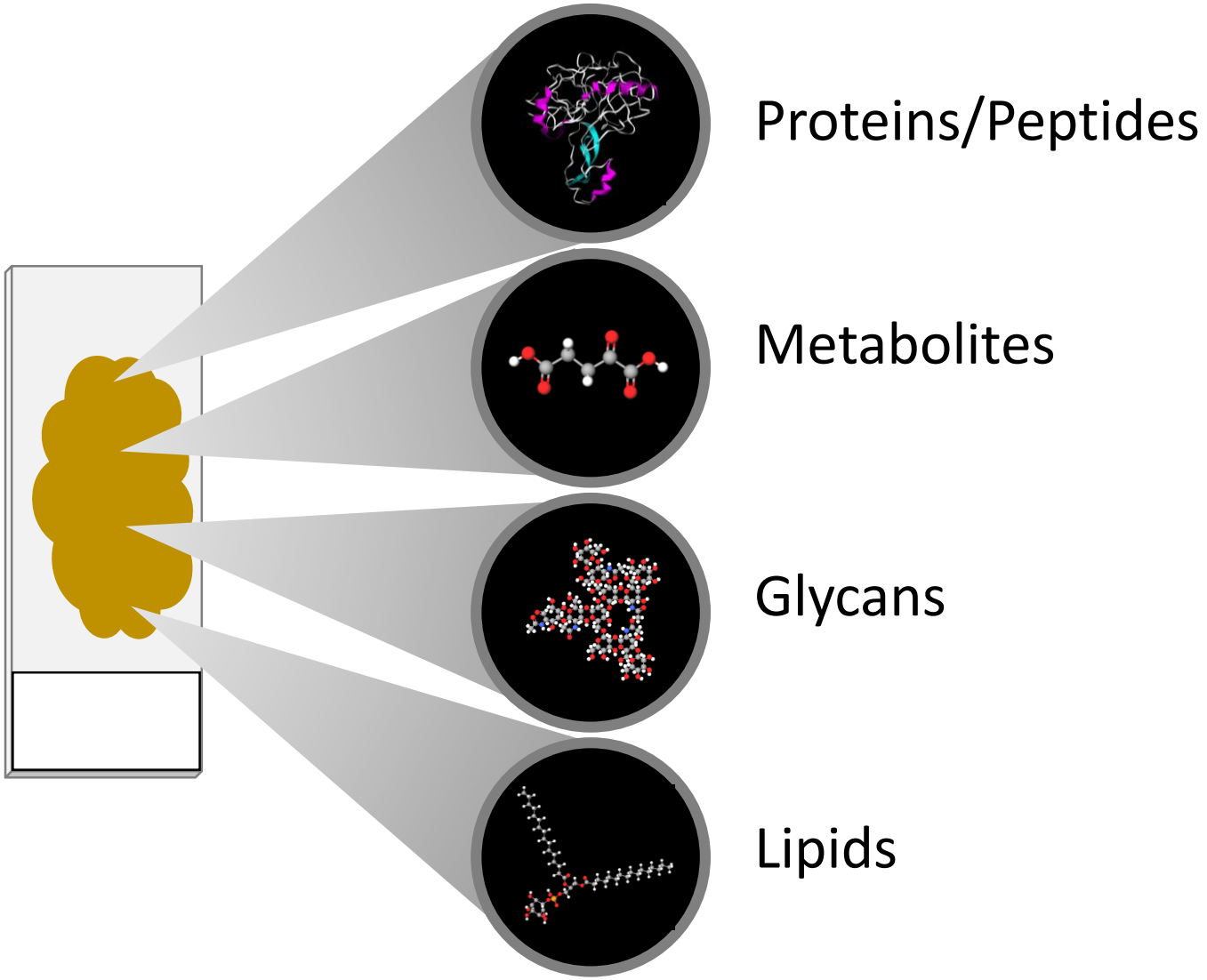
Mass spectrometry imaging enables molecular microscopy



Mass spectrometry imaging enables molecular microscopy

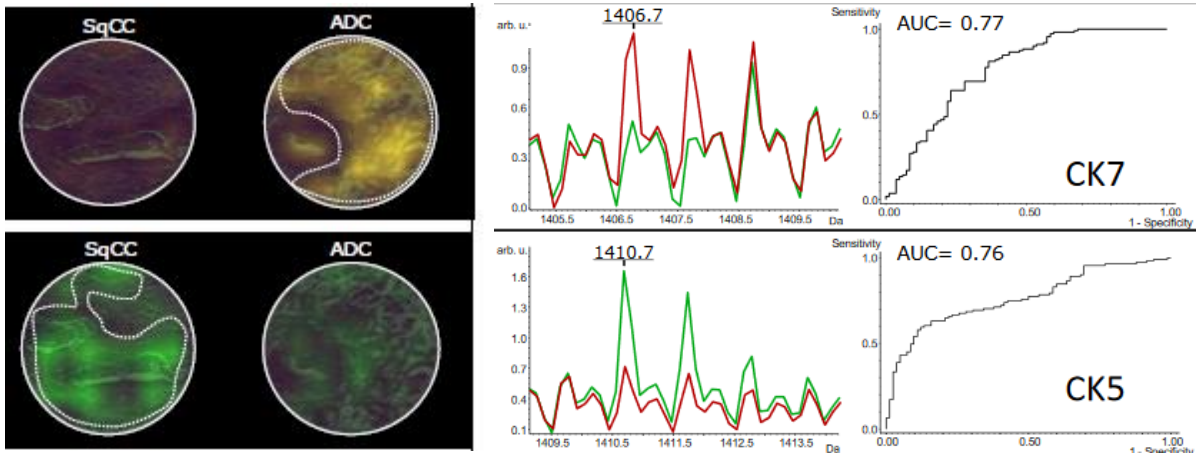


Endogeneous and exogeneous analytes



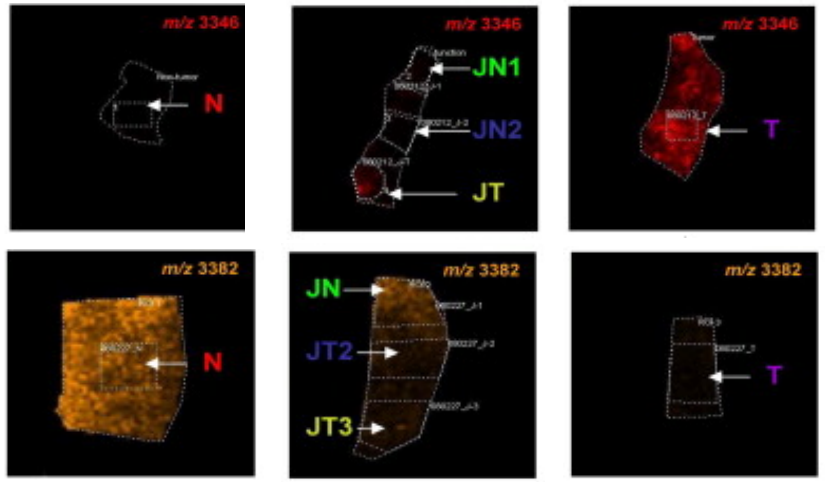
Mass spectrometry imaging in tumor research

Tumor subtyping



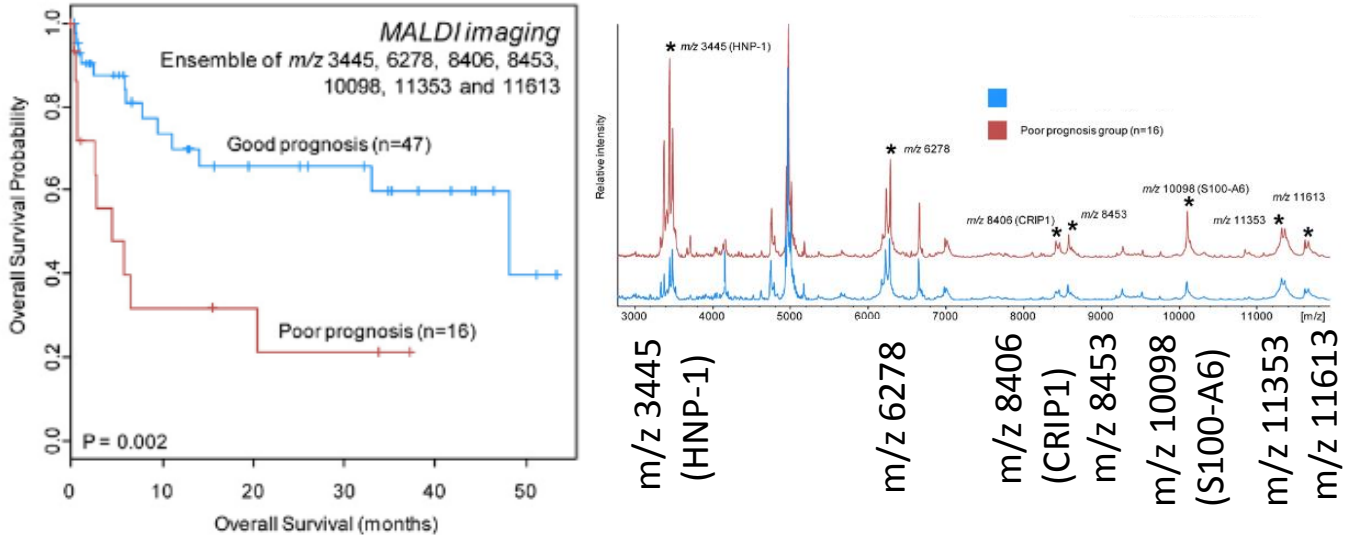
Kriegsmann 2016, Molecular & Cellular Proteomics

Resection margins / heterogeneity



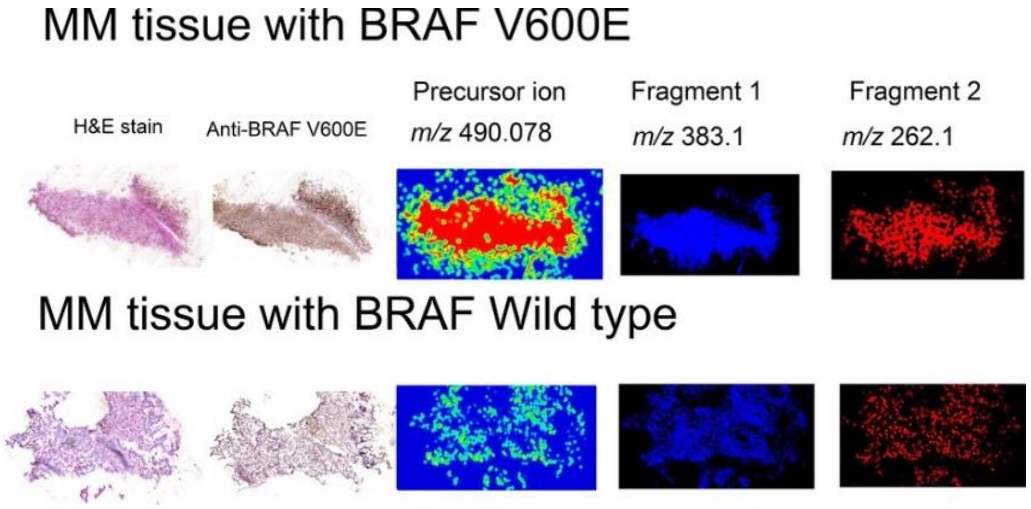
Han et al. 2011, Clinica Chimica Acta

Risk stratification



Balluff et al. 2011, The American journal of pathology

Drug imaging



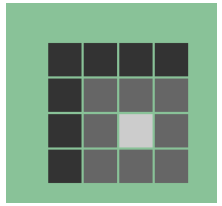
Sugihara et al. 2014, Proteomics

Challenges in MSI data analysis

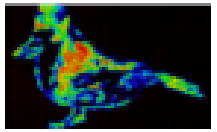
- Large datasets

Our solutions

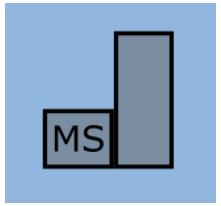
Galaxy PROJECT

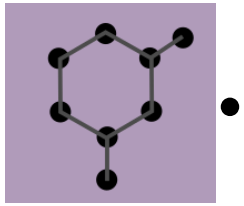


- Spatial dependence between spectra

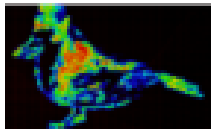


Cardinal

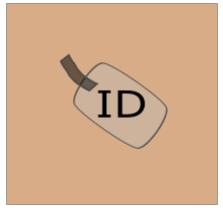




- Diverse mass spectrometers and analytes

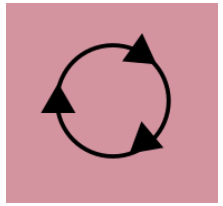


Cardinal



- Analyte identification

Only partially solved in the field, common option: LC-MS/MS



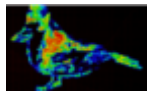
- Reproducibility via shared data and code

Galaxy PROJECT

MSI data analysis tools in Galaxy

Quality Control	File Handling	Prepro-cessing	Annota-tions	Statistics	Visuali-zation	Identifi-cation
MSI Qualitycontrol	MSI combine	MSI preprocessing	Scale image	MSI segmentation	MSI mz images	Join two files
	MSI filtering	MALDIquant preprocessing	Landmark registration	MSI single ion segmentation	MSI plot spectra	
	MSI data exporter	MALDIquant peak detection	Overlay	MSI classification		
			Coordinates of ROIs	MSI colocalization		
			Projective transformation			
			Switch axis coordinates			

Tools based on


Cardinal

MALDI quant MALDIquant


scikit-image

Availability

Galaxy EUROPE
- https://usegalaxy.eu

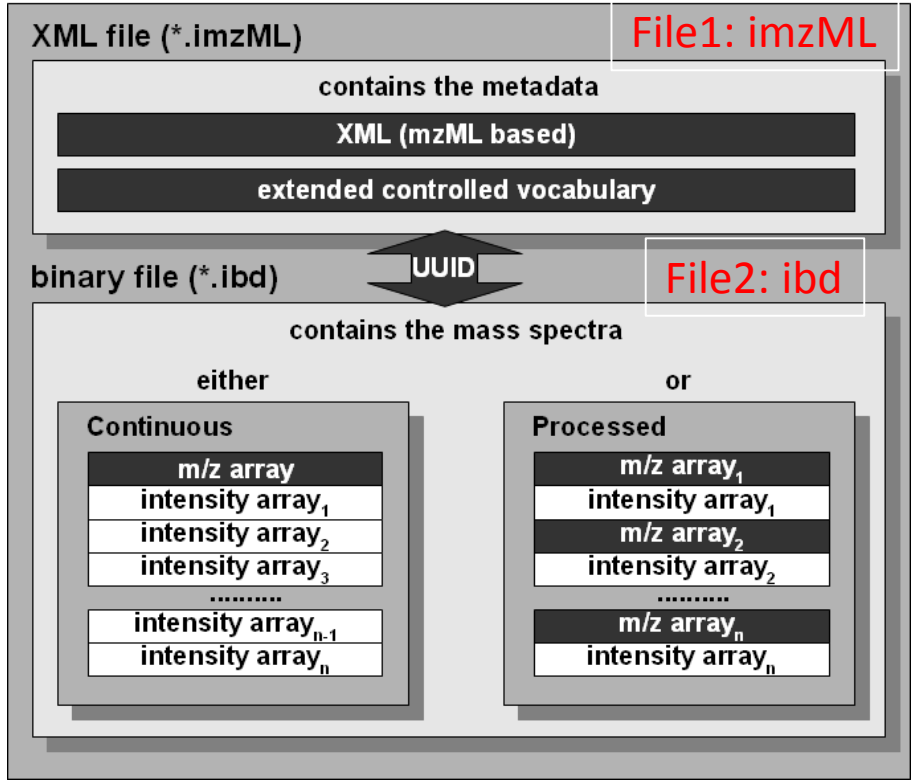
Galaxy ToolShed
- https://toolshed.g2.bx.psu.edu

docker
- https://github.com/foellmelanie/docker-galaxy-msi

Föll, M.C. *et al.* Accessible and reproducible mass spectrometry imaging data analysis in Galaxy. *Gigascience.* 8, 628719 (2019)

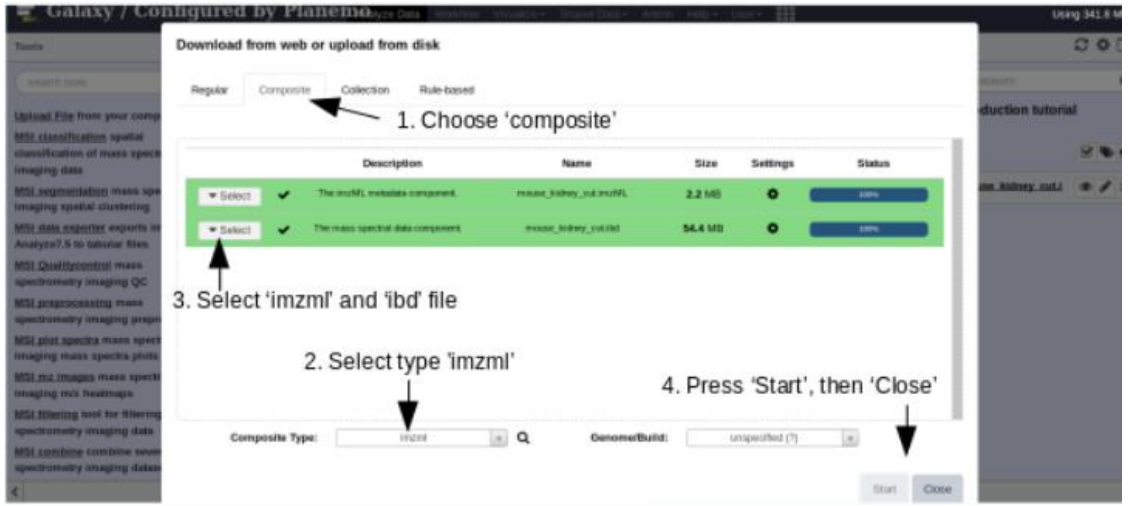
File formats: composite

imzML open standard file format



Alternative file format:
Analyze 7.5 = hdr + img + t2m

Composite upload in Galaxy



wishlist

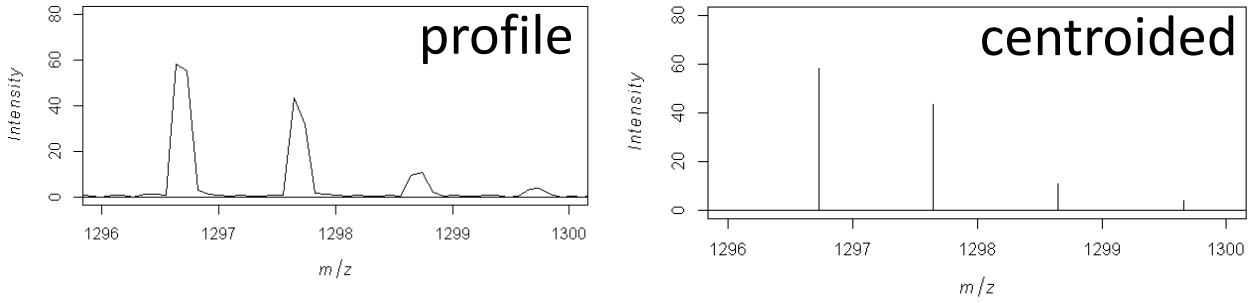
- Tests for composite upload (imzML and Analyze 7.5) for new Galaxy versions
- Parallel upload of multiple composite files (collection)

Schramm, Thorsten, et al. "imzML—a common data format for the flexible exchange and processing of mass spectrometry imaging data." *Journal of proteomics* 75.16 (2012): 5106-5110.

File sizes

Depend on:

- Size of the tissue / sample
- Type of mass spectrometer
- Options set in the mass spectrometer
  - Spatial resolution
  - Mass resolution
  - Ion mobility
- Profile or centroided data

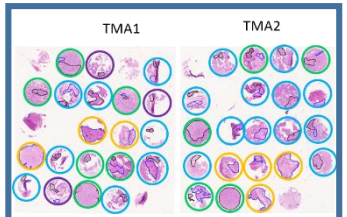


File sizes of our raw files:
0.5 - 800 GB (single file)

wishlist

- Ideas / Suggestions for resulting problems:
- Need >>> 250 GB available per user in EU
- Need this for >>> 1 year
- Adjustment of cores and RAM for MSI tools according to input file sizes?

Urothelial carcinoma cohort - Workflow



39 urothelial tissue cores

Sample preparation

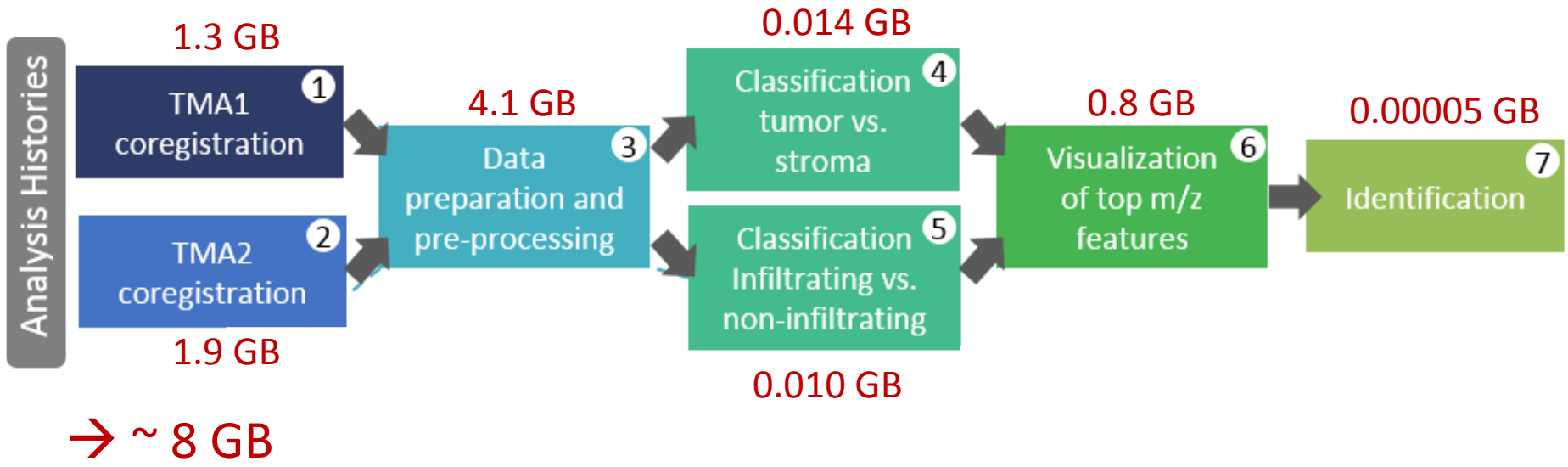
Data acquisition

Galaxy EUROPE

Data analysis

Galaxy EUROPE

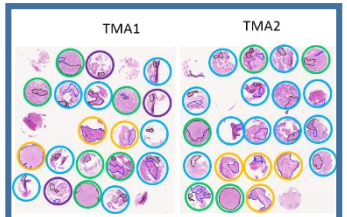
Analysis histories and workflows shared in Galaxy



→ ~ 8 GB

Föll, M.C. et al.,
Clinical Proteomics 2022

Urothelial carcinoma cohort - Workflow



39 urothelial tissue cores

Sample preparation

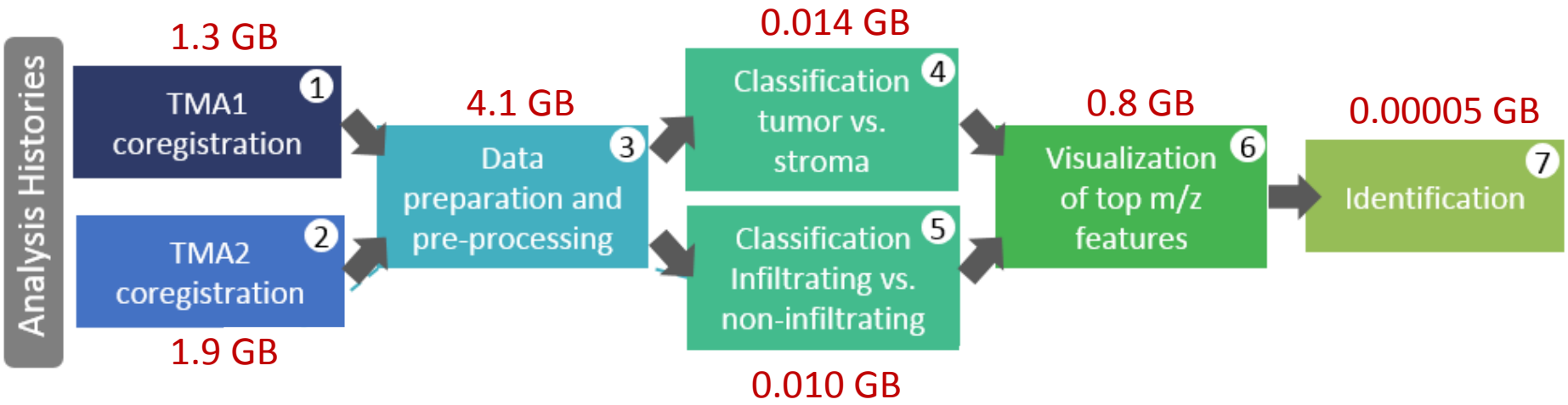
Data acquisition

Galaxy EUROPE

Data analysis

Galaxy EUROPE

Analysis histories and workflows shared in Galaxy

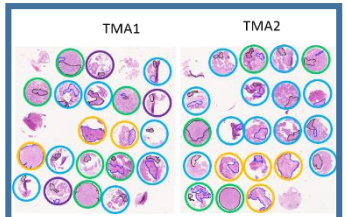


→ ~ 8 GB

wishlist

- Galaxy „publication service“ → Histories shared with publications in peer-reviewed journals are removed from a users quota, stored for X years
- Option to „lock“ published histories

Urothelial carcinoma cohort - Workflow



39 urothelial tissue cores

Sample preparation

Data acquisition

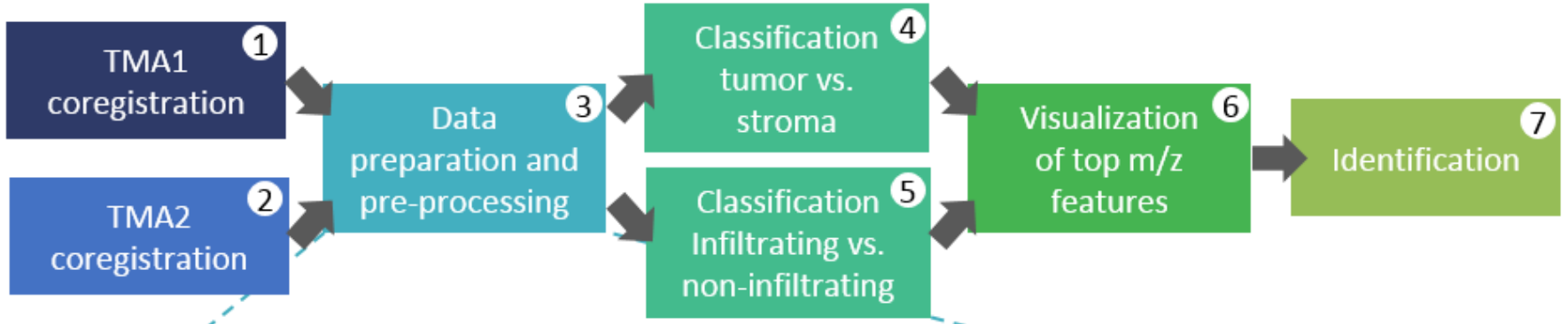
Galaxy EUROPE

Data analysis

Analysis Histories

Galaxy EUROPE

Pre-processing Workflow

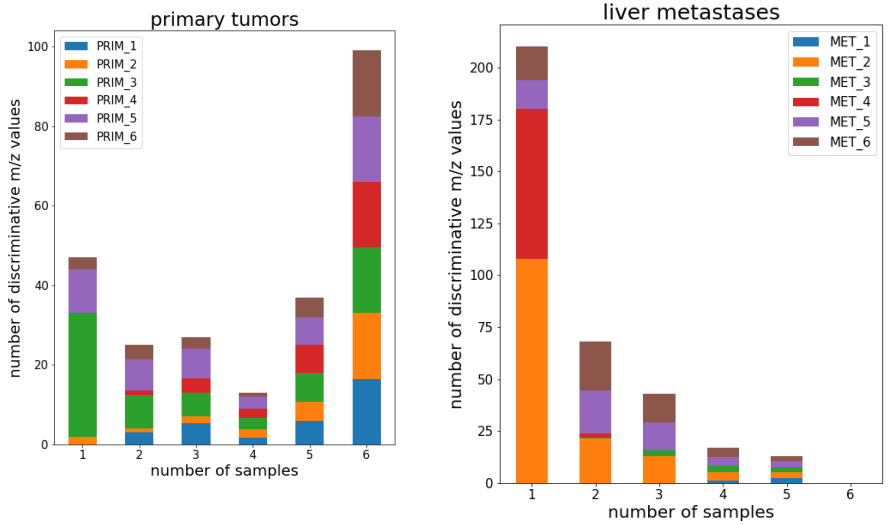


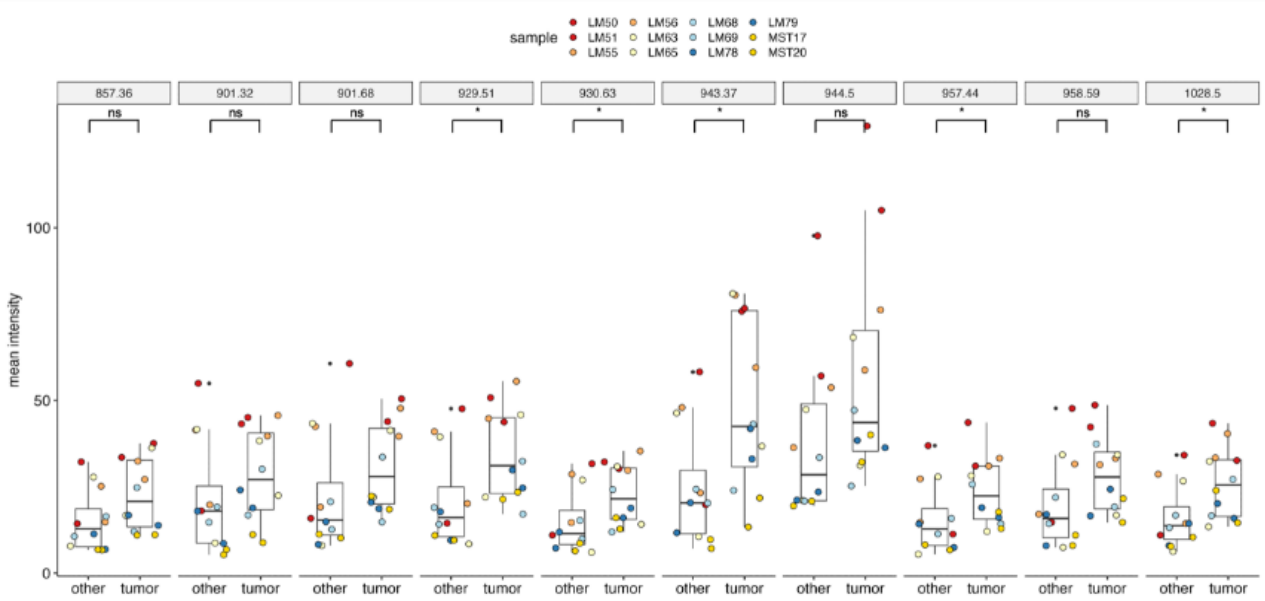
wishlist

- Colouring of each tool box in workflow editor
- Drawing colourful boxes / annotations

wishlist continued…

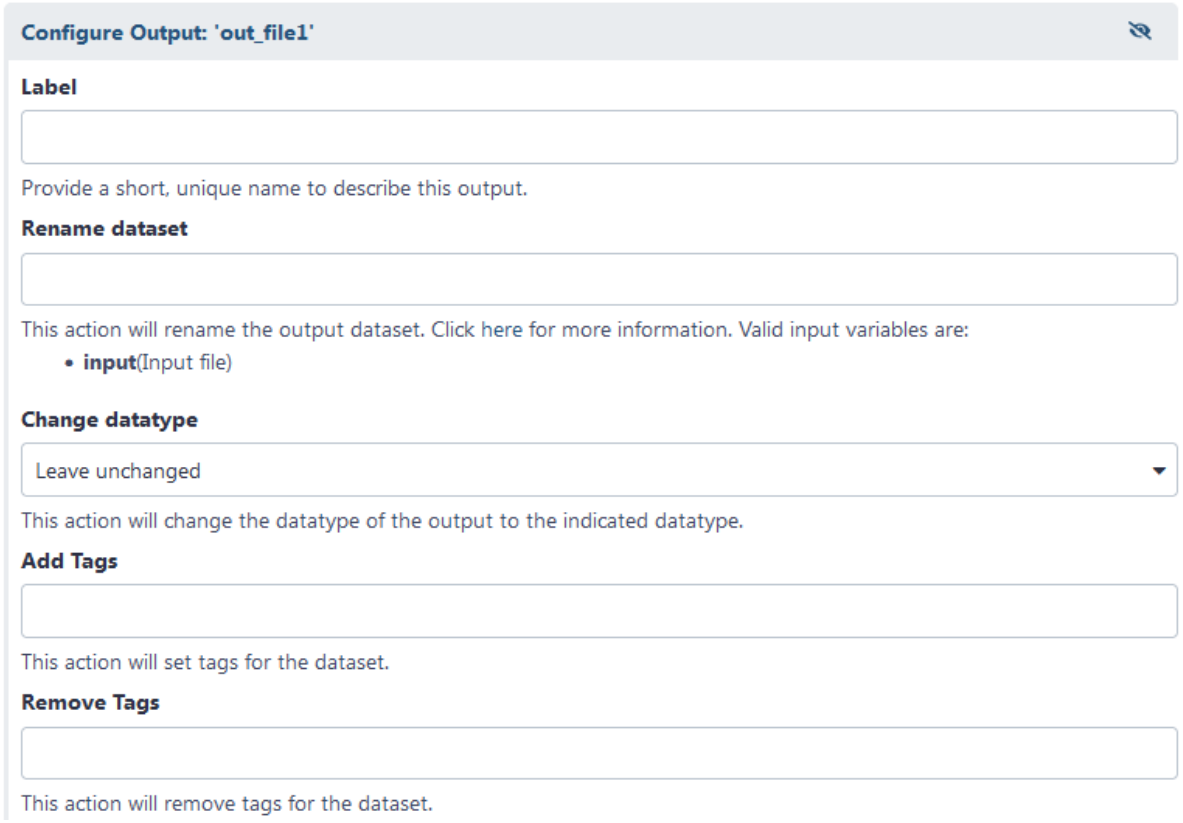
- Basic visualizations in publishable quality for tabular files
  - Barplots
  - Boxplots
  - Scatter plots
  - Line plots
  - Heatmaps





wishlist continued...

- Option to confiugure output file from normal tool run
  - Dataset name
  - Tags



Just like in the workflow editor…

Acknowledgement

Faculty of Medicine

MEDICAL SCIENTIST PROGRAMME

MALDI imaging group

Bálint Barta
Nadine Meier
Larissa Meyer
Johanna Thiery
Andreas Weber
Flo Wiedemann

Schilling Lab Freiburg

Oliver Schilling
Maren Stillger
Niko Pinter
Jenny Li
Lennart Moritz
Veronika Volkmann

Institute of Surgical Pathology Freiburg

Kathrin-Enderle-Ammour
Peter Bronsert
Anna Frey
Martin Werner

Galaxy Team Freiburg

Björn Grüning

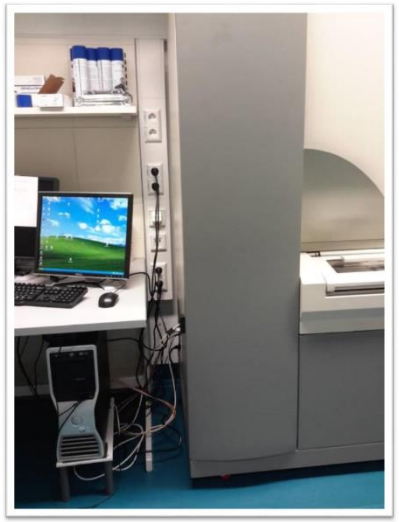
Galaxy Community

Northeastern University Boston

Olga Vitek
Kylie Bemis
Dan Guo
Sai Lakkimsetty

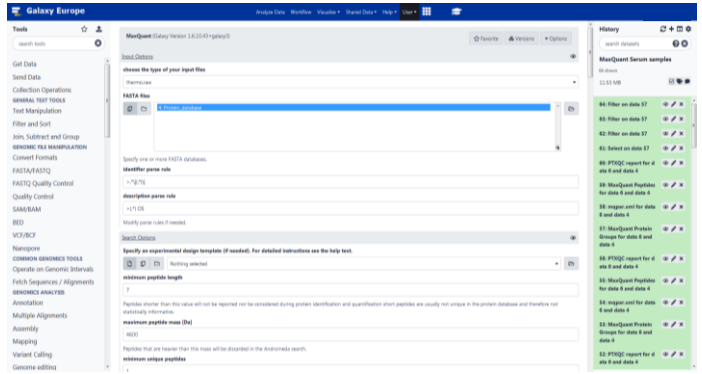
Mass spectrometry imaging @pathology in Freiburg

Sprayer



MALDI-TOF/TOF

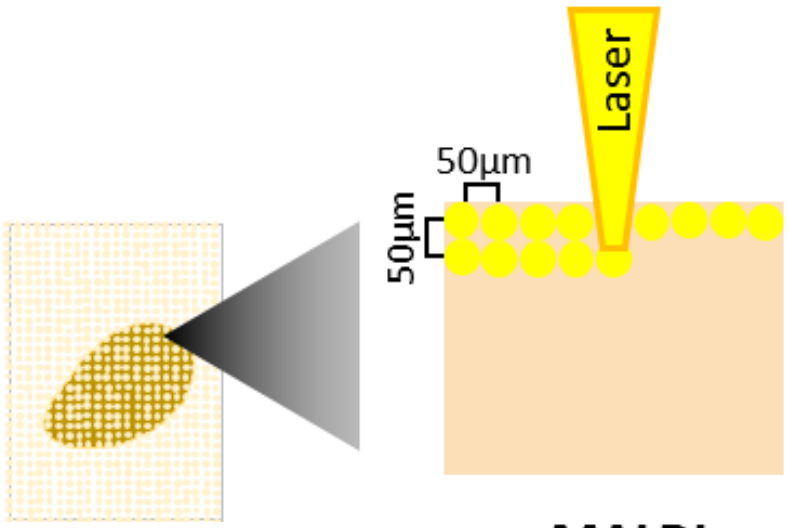
Galaxy EUROPE

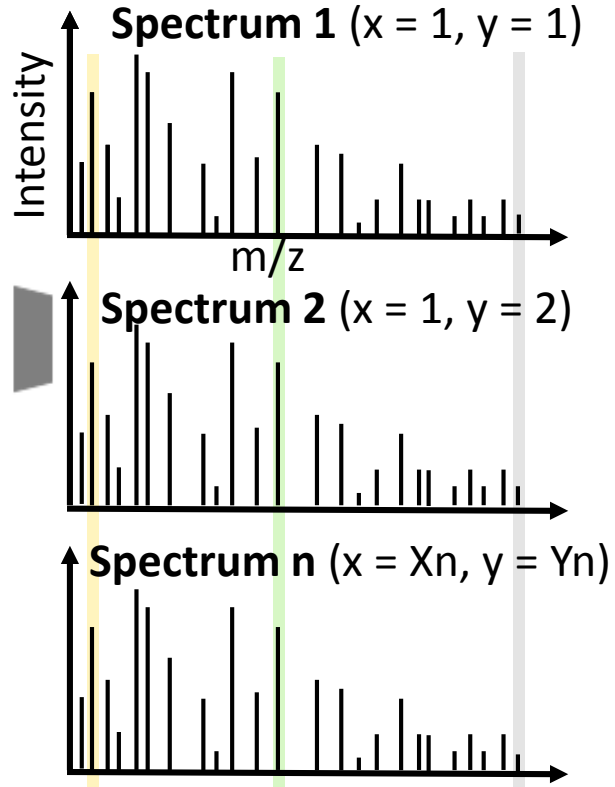


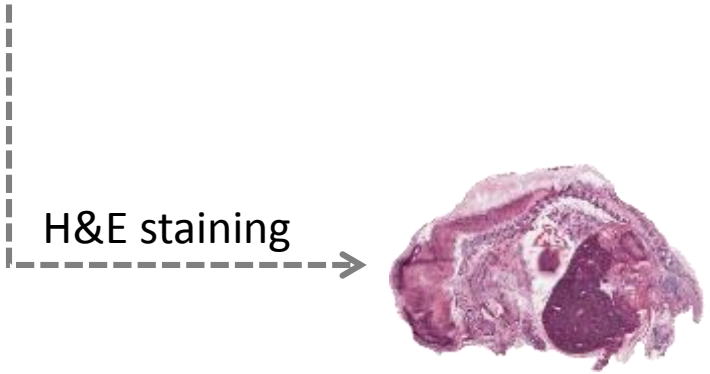
Analysis Software

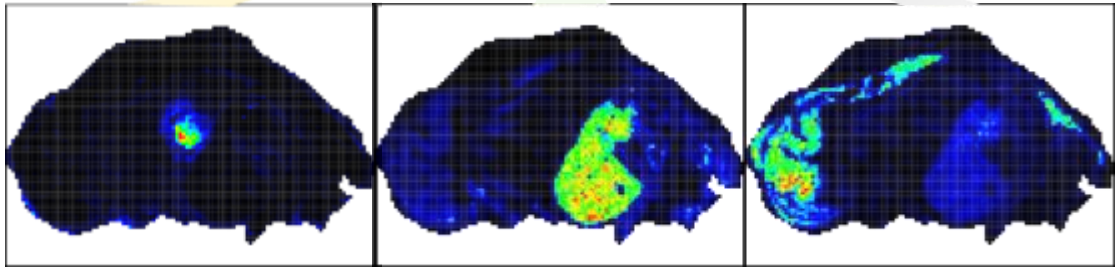
Sample preparation → Data acquisition → Data analysis

Mass spectrometry imaging





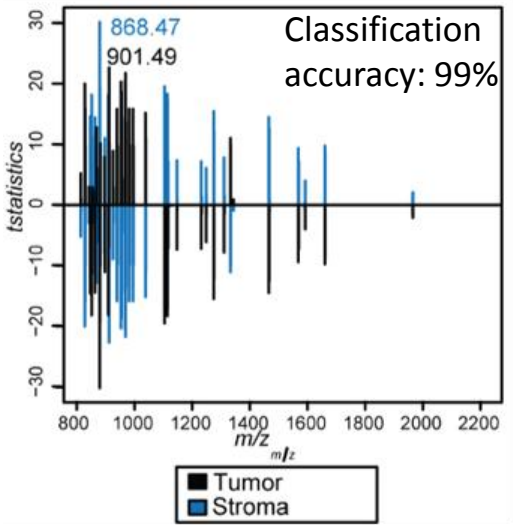


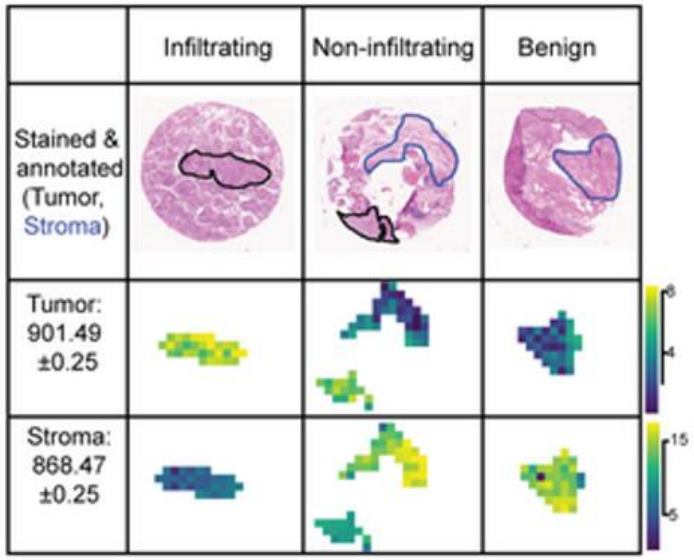


Pig fetus, R. Graham Cooks lab

1) Urothelial carcinoma classification & IHC

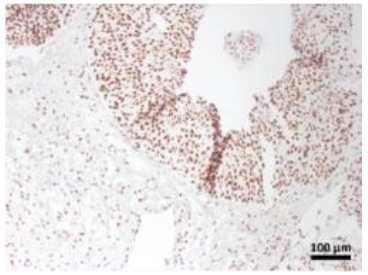
Tumor vs. stroma



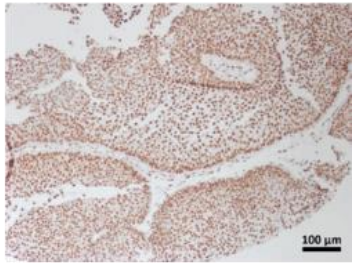


Immunohistochemistry of Histone H2A

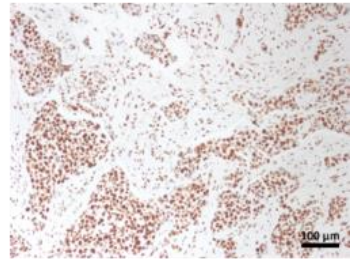
Non-muscle infiltrating low grade



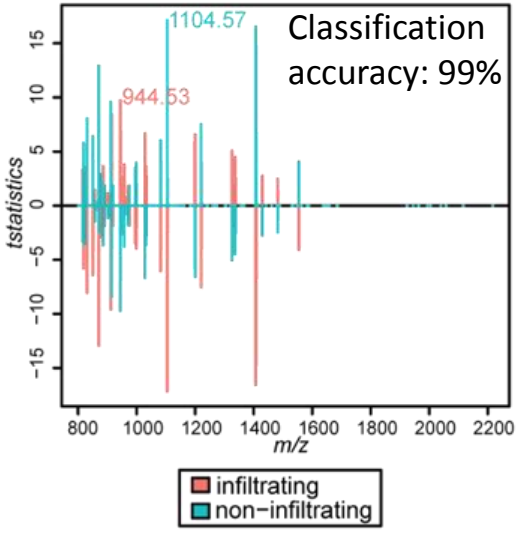
Non-muscle infiltrating high grade

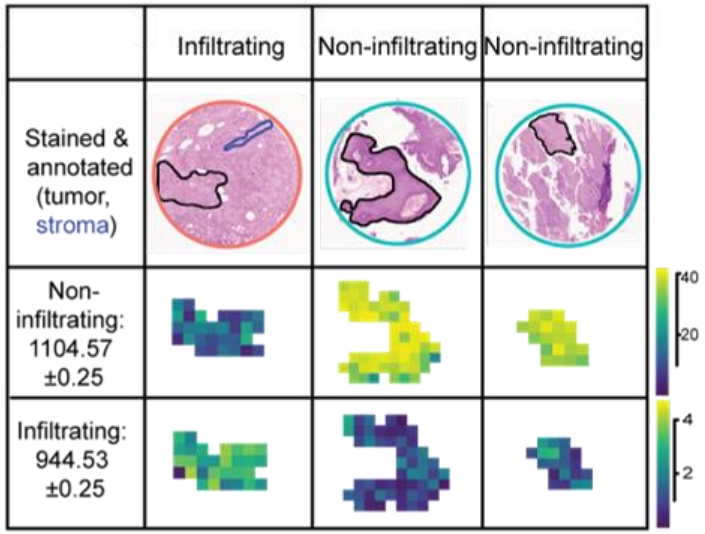


Muscle infiltrating

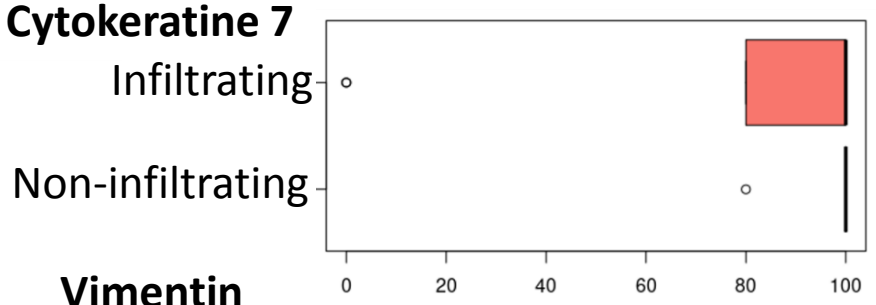


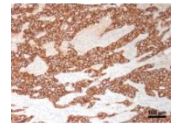
Muscle-infiltration



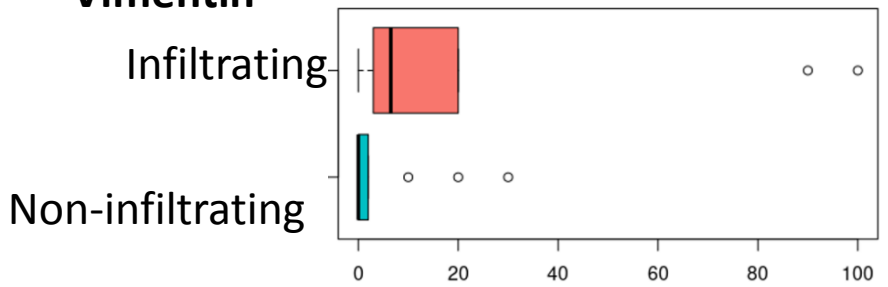


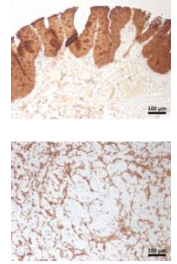
Cytokeratine 7





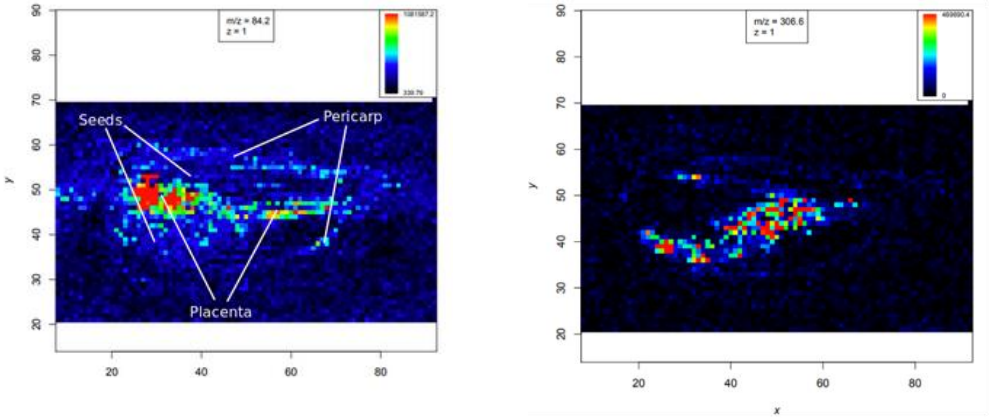
Vimentin



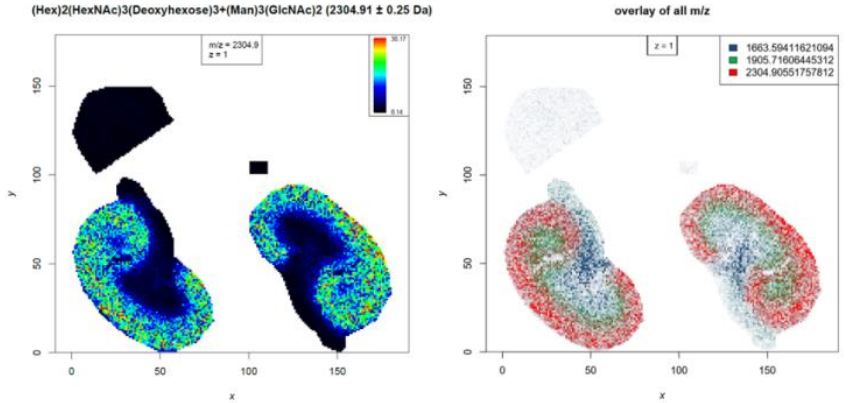


Galaxy MSI tutorials

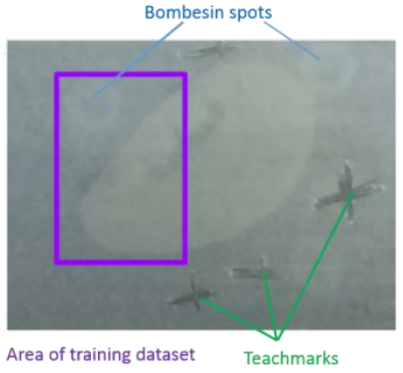
1) Distribution of volatile metabolites in a chilli section

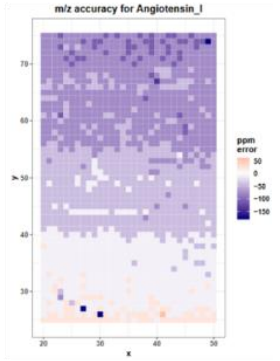


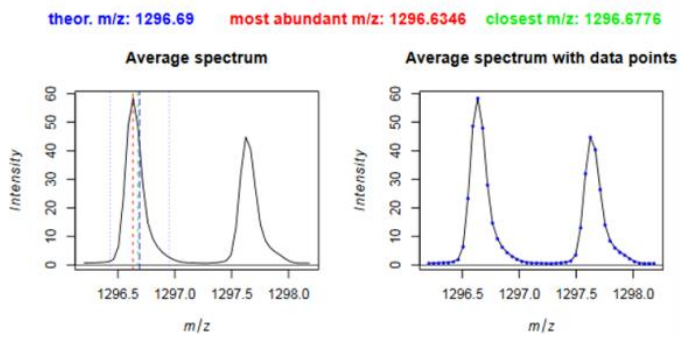
2) N-linked glycan distribution in mouse kidney tissue



3) Quality control of peptide distribution in mouse kidney tissue







https://galaxyproject.github.io/training-material/