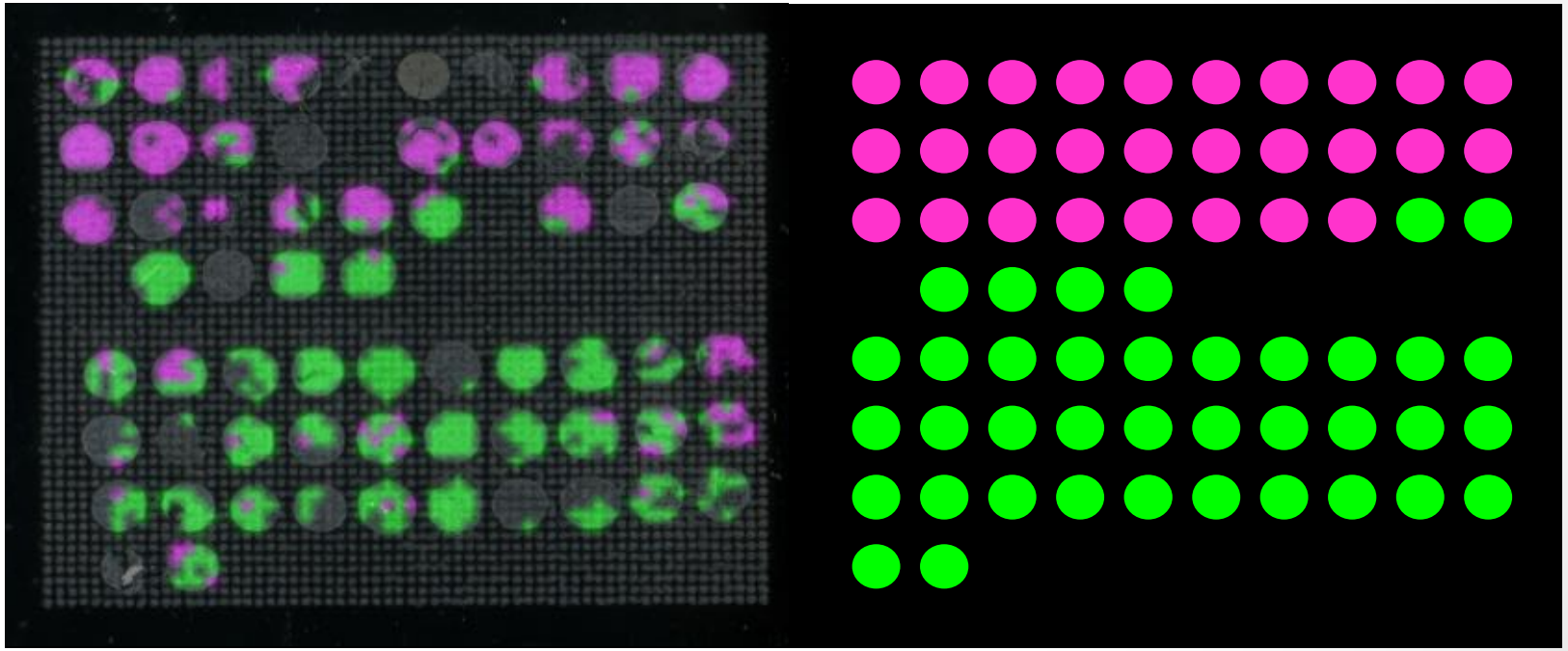
Practical Considerations for Clinical Mass Spectrometry Imaging Studies

Erin H Seeley, PhD



Fata, C., *et al.* Hum Pathol. **2015**, *46*, 1427-36.

The Three Keys to Success

Reproducibility

Reproducibility

Reproducibility

Clinical MS Studies

- May take weeks to years to complete
- May involve tens to thousands of samples
- Generally detecting intensity differences between disease states, rarely all-or-nothing
- Reproducibility is of utmost importance to obtain accurate results
- Automation and robotics help to ensure consistent results

How Many Samples Do I Need?

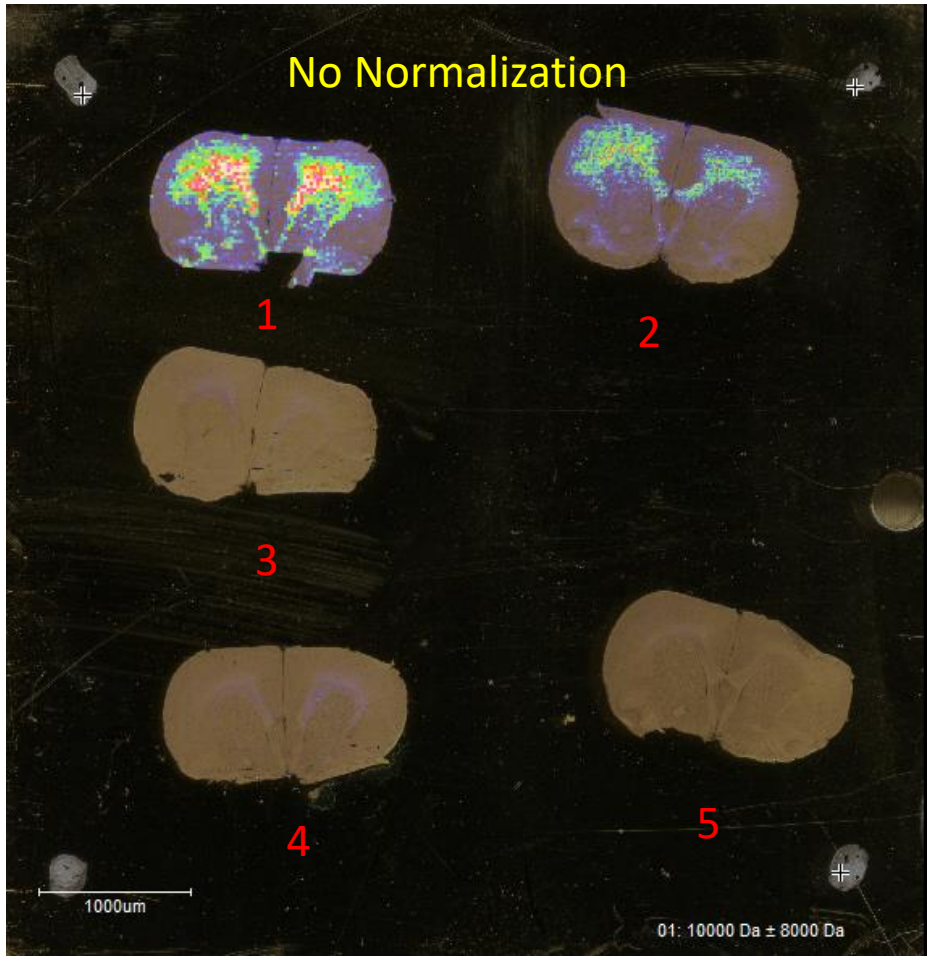
- Multiple sets needed
  - Training – creation of a classification algorithm
  - Validation – evaluation of classification algorithm accuracy
  - Testing – evaluation of algorithm performance in a clinical setting
- Power calculations have shown that typically a minimum of 25 human samples per cohort are needed for statistical significance
- More than that are preferred for a good representation of the expected patient population
- Are there multiple phenotypes/subtypes I need to account for?
  - Distribution of phenotypes/subtypes should be balanced
- What is my end goal?
  - Publication – may need 50-100 samples
  - Clinical Assay – may need 300-1000+ samples

Do I Need to Care about Ambient Conditions?

- YES!!!
- Changes in temperature and humidity may affect digestion efficiency or evaporation time
- How to standardize
  - Temperature and humidity control for the lab
  - Well regulated and calibrated ovens
  - Principle of deliquescence for digestion*
  - Heated tray in HTX Sprayer for constant evaporation
  - etc.

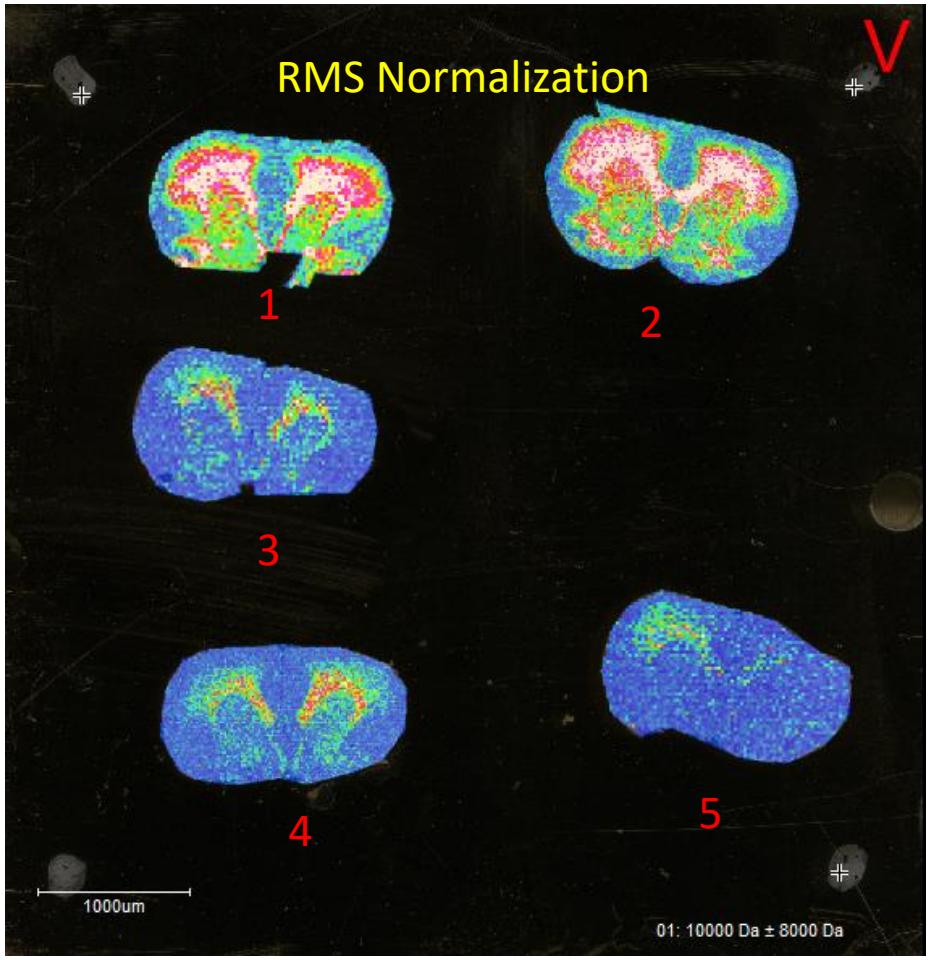
*Ly, A., *et al. Proteomics Clin Appl*. **2019**, *13(1)*, e1800029.

What About Instrument Performance?

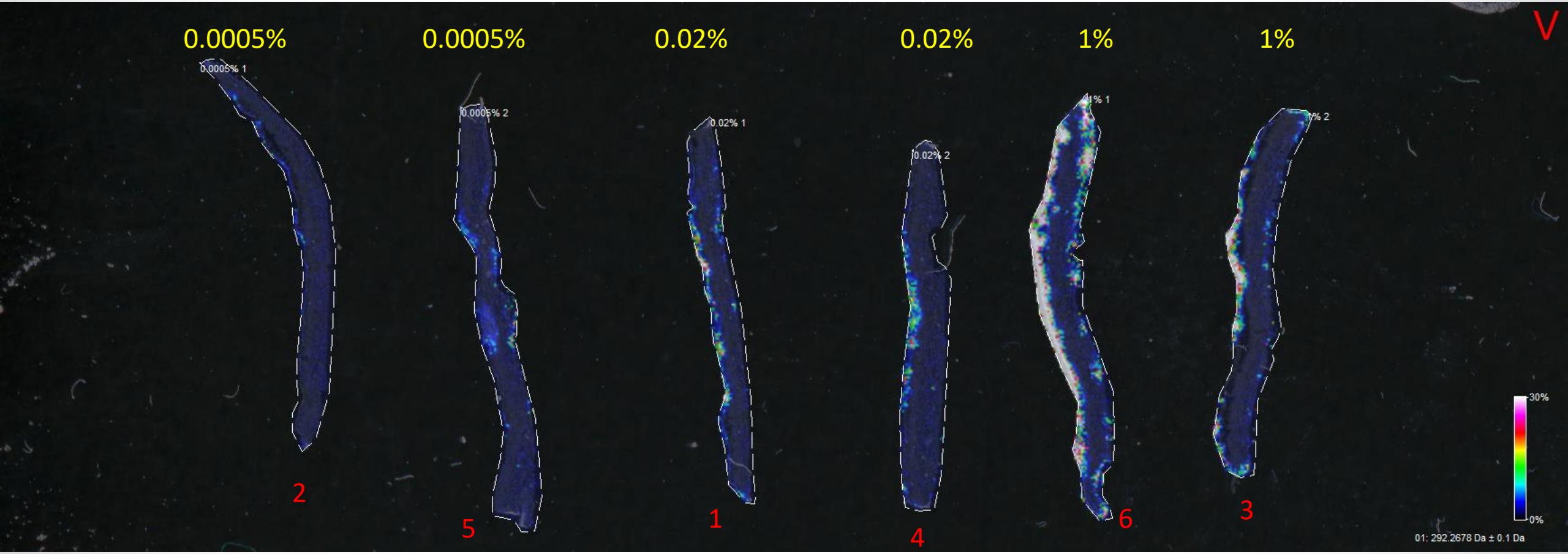


~22,000 pixels, 100 shots per pixel = 2.20 million shots

But I Can Fix That With Normalization, Right?

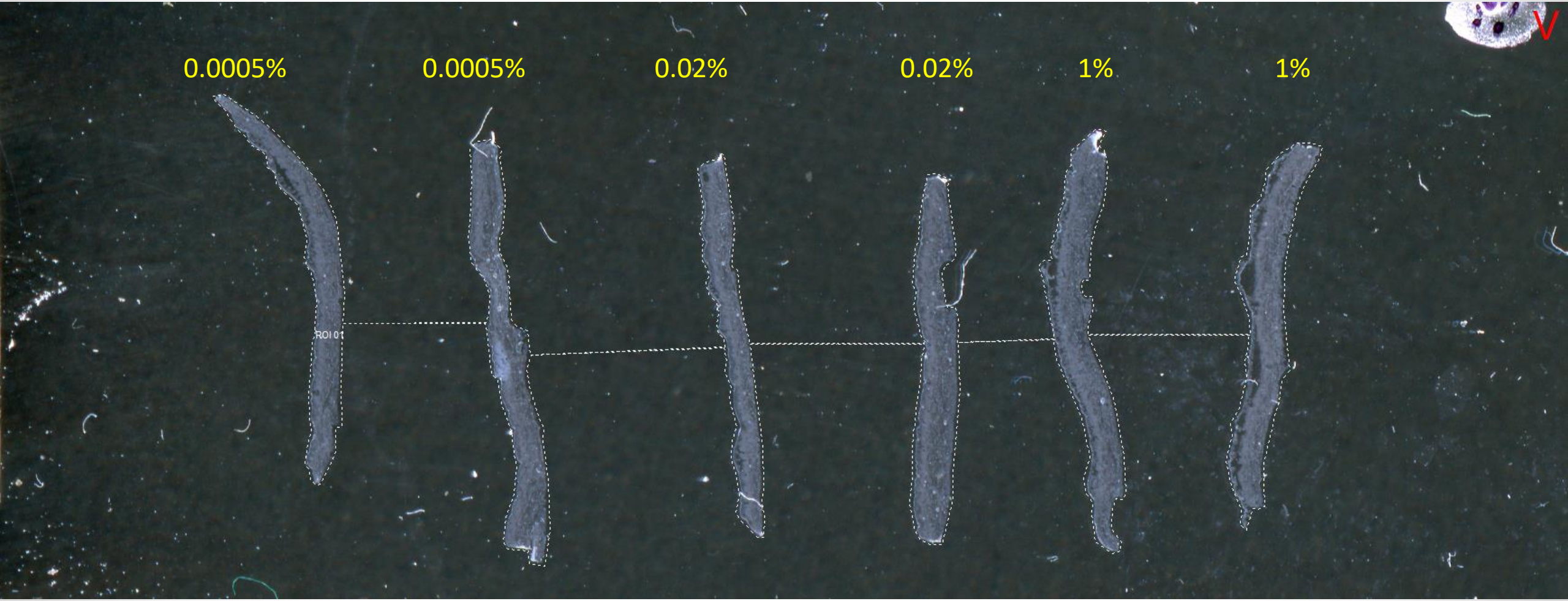


Randomize Collection Order



Skin treated with varying concentrations of an antifungal compound

Collect All Samples as One Region



Minimizing Sampling Bias

- Verify instrument performance each day with a known standard
  - Particularly important after instrument maintenance is performed
  - Different service engineers may tune differently
- Include a “standard sample” with each analysis
- Ensure that each batch of samples analyzed contains a mix of diagnostic states
- Mix data collection so that types to be compared are not collected sequentially
- Perform regular source cleaning/maintenance so that signal is not degraded over time
- Minimize data volume
  - Lower spatial resolution
  - Limited regions for data collection
  - Histology-guided profiling
  - Use of Tissue MicroArrays (TMAs)
- Carry out data analysis in a single batch

Histology-Guided Mass Spectrometry Profiling

- A targeted approach in which only discrete areas within a tissue section are analyzed
- Histological staining is used to guide the acquisition of spectra
- Each analyzed spot is enriched for a single cell type
  - Ensure high quality/purity of annotations
- Conducive to statistical analysis and classification algorithm generation
- Can provide biological insight not attainable by standard histology (disease outcome, improved diagnostics, treatment response, etc.)

Melanocytic Skin Lesions

Histology-Guided Mass Spectrometry Profiling – FFPE Tissue

In Collaboration with Dr. Rossitza Lazova, California Skin Institute

Melanoma Diagnosis Discordance

- 4+ M biopsies performed per year in the US to rule out melanoma.[†] Of these, over 10% cannot be definitively classified using routine histopathology.[‡]
- There can be considerable disagreement among dermatopathologists in the diagnosis of melanocytic lesions
- In a 1996 study*, 8 pathologists reviewed the same 37 slides
  - 13 cases had complete agreement
  - 10 cases had one discordance
  - 14 cases had 2 or more discordances, 8 of these had 3 or more
- More objective diagnosis is needed to help with difficult cases

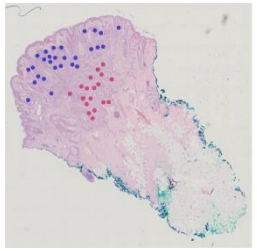
[†] *Br J Dermatol,* **2017**, *176,* 949-954
[‡] *J Am Acad Dermatol*, **2012**, *67*, 727-735
Human Pathol*. **1996, *27*, 528-531

Study Parameters

- Data collected over a period of 4 years
- 756 specimens, mixed between biopsies and tissue microarrays (TMAs)
- Age of specimens 5-40+ years
- Two different sprayer platforms [SunChrom SunCollect and HTX M5(2)]
- Two different mass spectrometers [Bruker ultrafleXtreme and rapifleX(2)]
- Three different personnel collected data

Histology-Guided MS Profiling Workflow

Tissue section a MALDI target

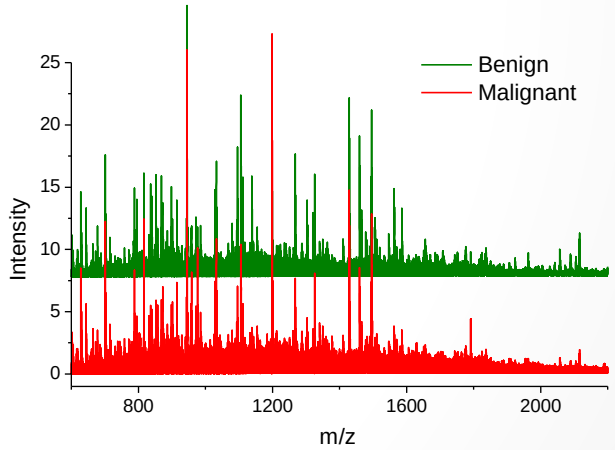


Stained serial section annotated

Many tissues in one experiment

Images of stained and unstained sections merged

Trypsin and matrix applied



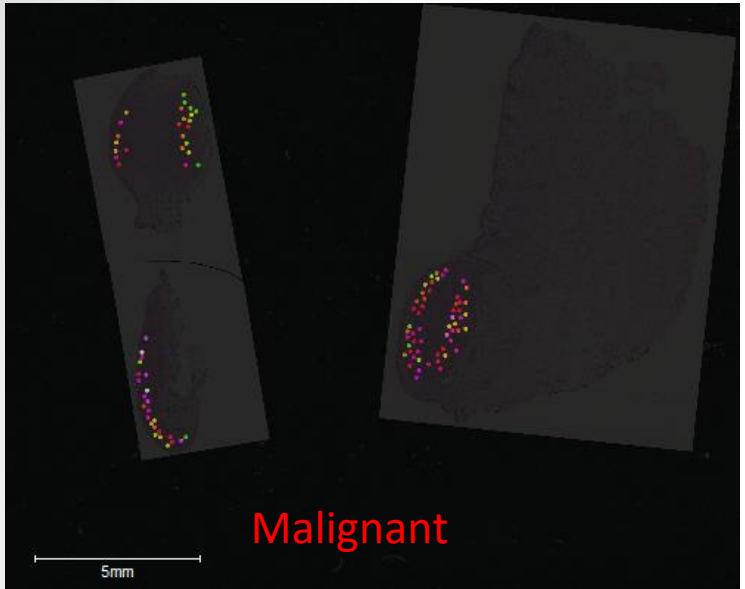
Mass spectra collected

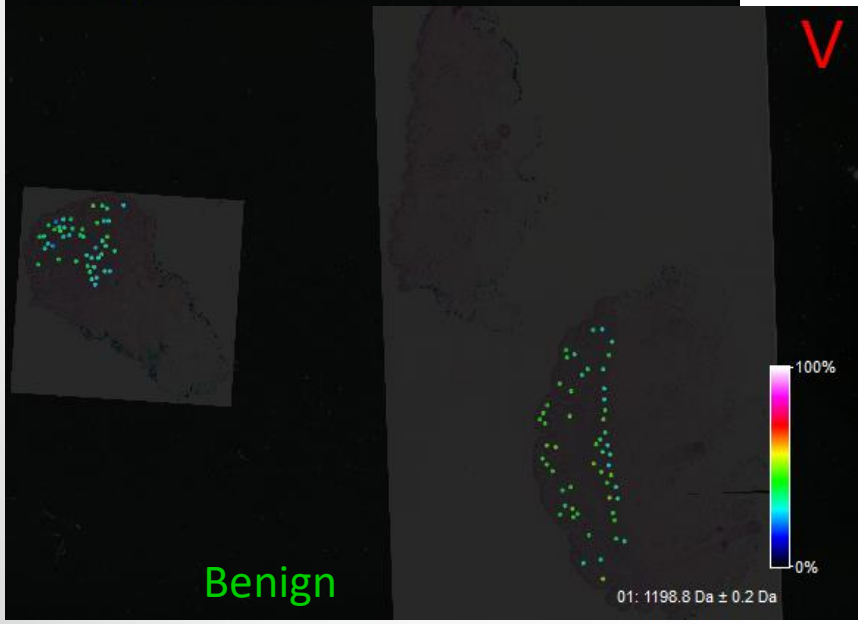
Statistical Analysis Methods

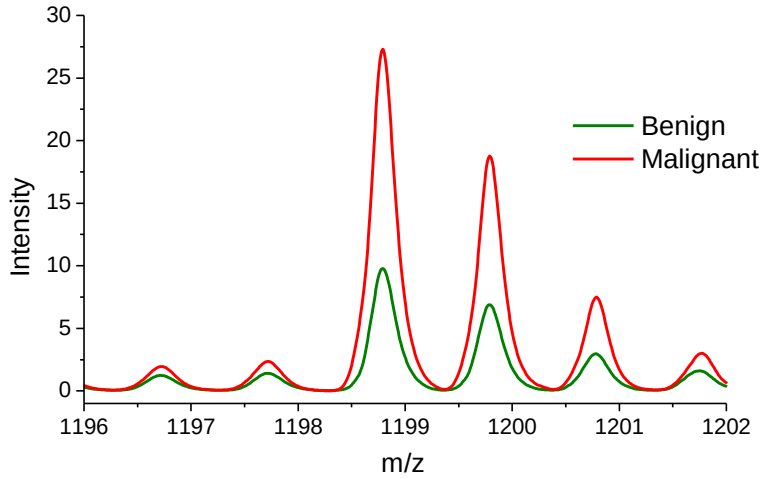
- Spectra loaded into SCiLS Lab software and baseline corrected, normalized, and peaks picked
- Samples sorted into Training and Validation/Testing sets
- Hypothesis testing carried out on training set
- Linear discriminant analysis (LDA) classification algorithm generated and validated

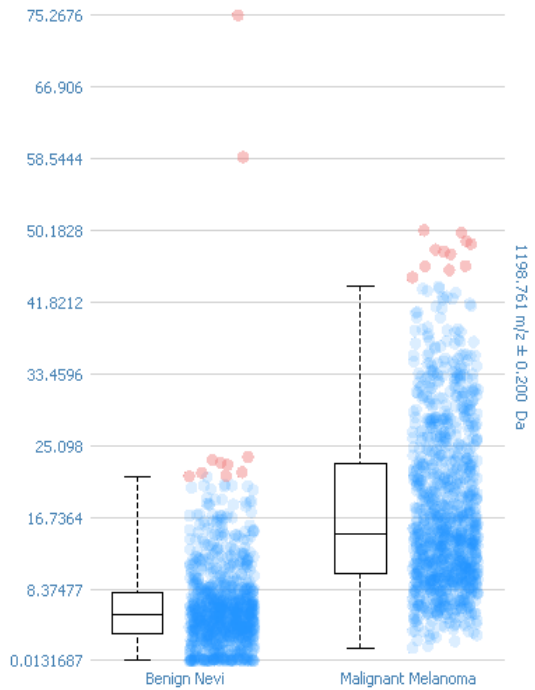
327/1075 peaks were significant after Bonferroni correction

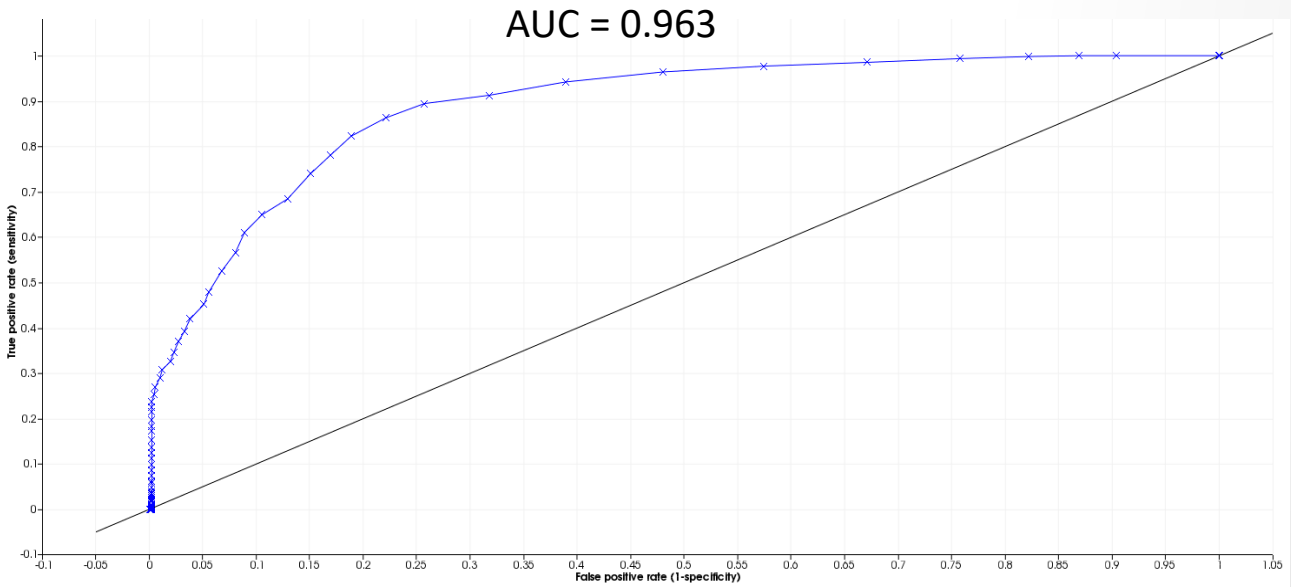
Significant Peak – m/z 1198.80



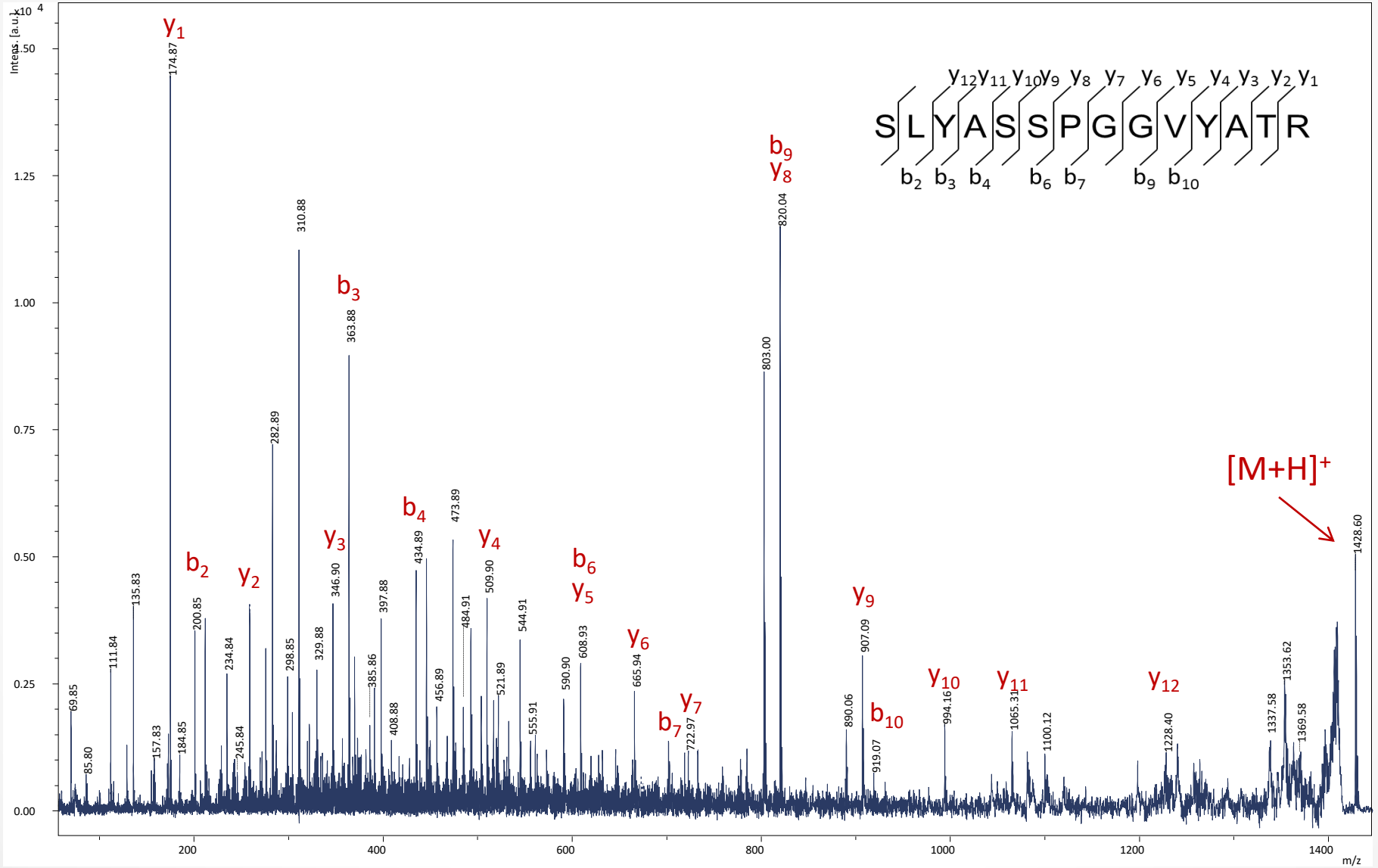








Identification of a Peptide from Vimentin



Classification Algorithm Generation

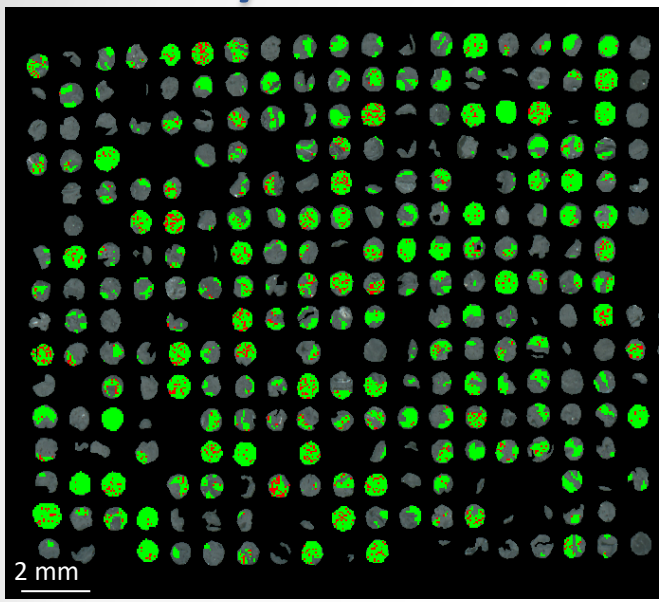
- Melanoma subtypes included: acral, desmoplastic, lentigo maligna, nevoid, nodular, Spitzoid, and superficial spreading. Nevi subtypes included: acral, conventional, and Spitz
- Linear discriminant analysis (LDA) classification algorithm built using leave-10%-out repeated random sub-sampling cross validation

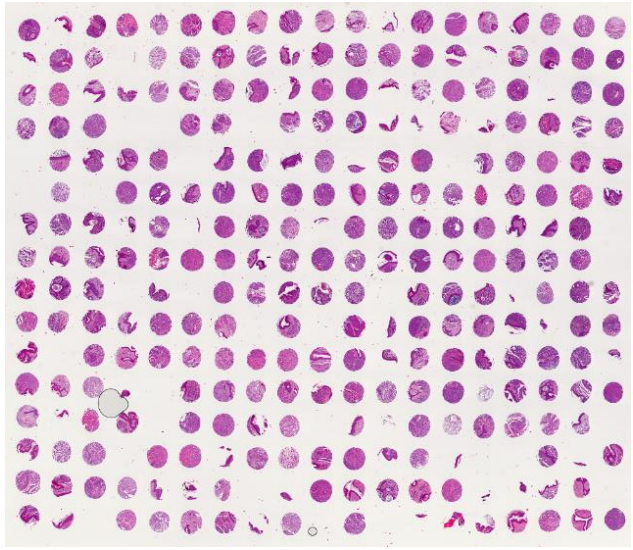
	Number of Patients	Internal Cross Validation Spectral Accuracy
Malignant Melanoma	100	94.1%
Benign Nevi	111	

- Optimized classifier applied to independent set of unambiguous malignant melanoma and benign nevi specimens – traditional biopsies and TMAs
  - **Malignant Melanoma – 288 patients**
  - **Benign Nevi – 257 patients**
  - **93% Sensitivity, 98% Specificity***

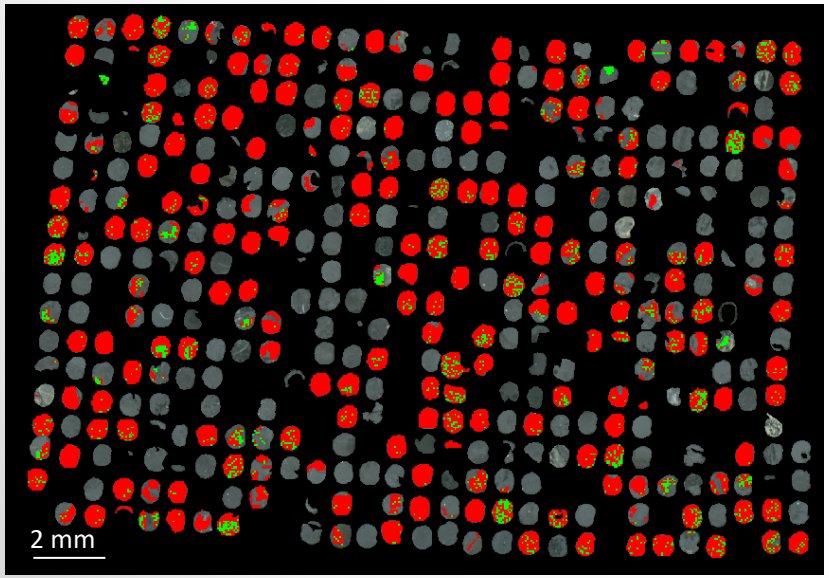
*A total of 20 samples (15 melanoma, 5 nevi; 3.6%) classified as indeterminant, included with incorrect samples

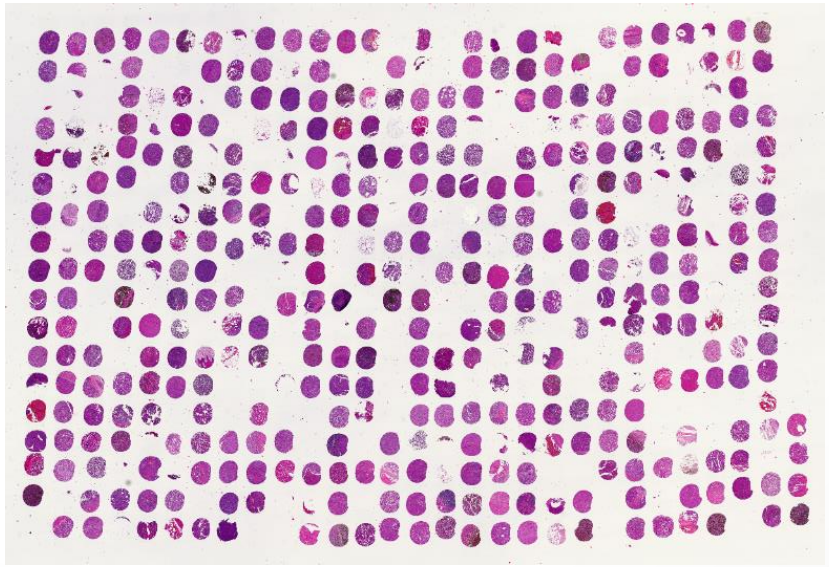
Analysis of Melanocytic Lesion TMAs





- 380 samples originated from 6 total TMAs (172 malignant, 208 benign)
- Only melanocytic component evaluated
- Core with no data – inappropriate for testing





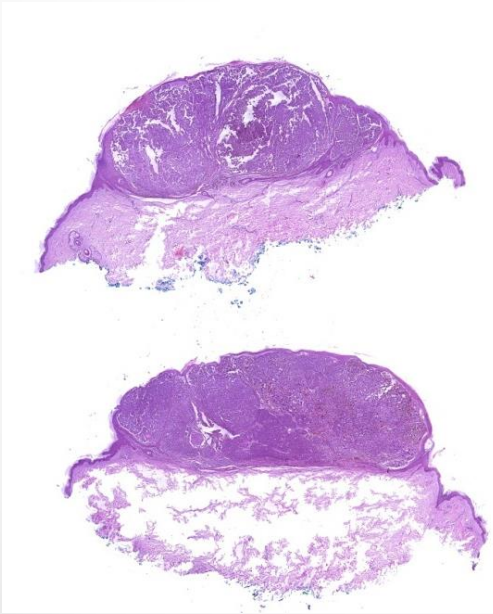
Benign

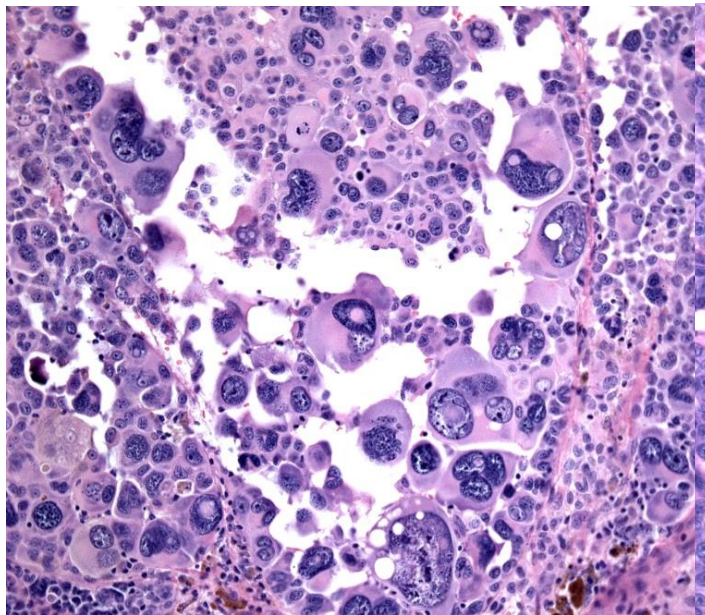
Malignant

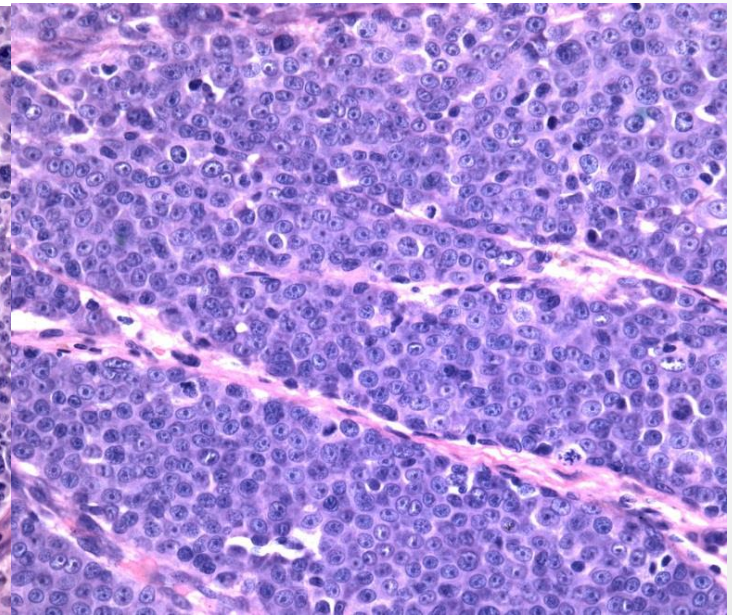
Non-Melanocytic Lesion

Case Study

- 37 year old pregnant woman presents with lesion on upper arm
- Excisional biopsy performed and determined to be malignant
- Insufficient margins taken for size of lesion
- No further treatment during pregnancy

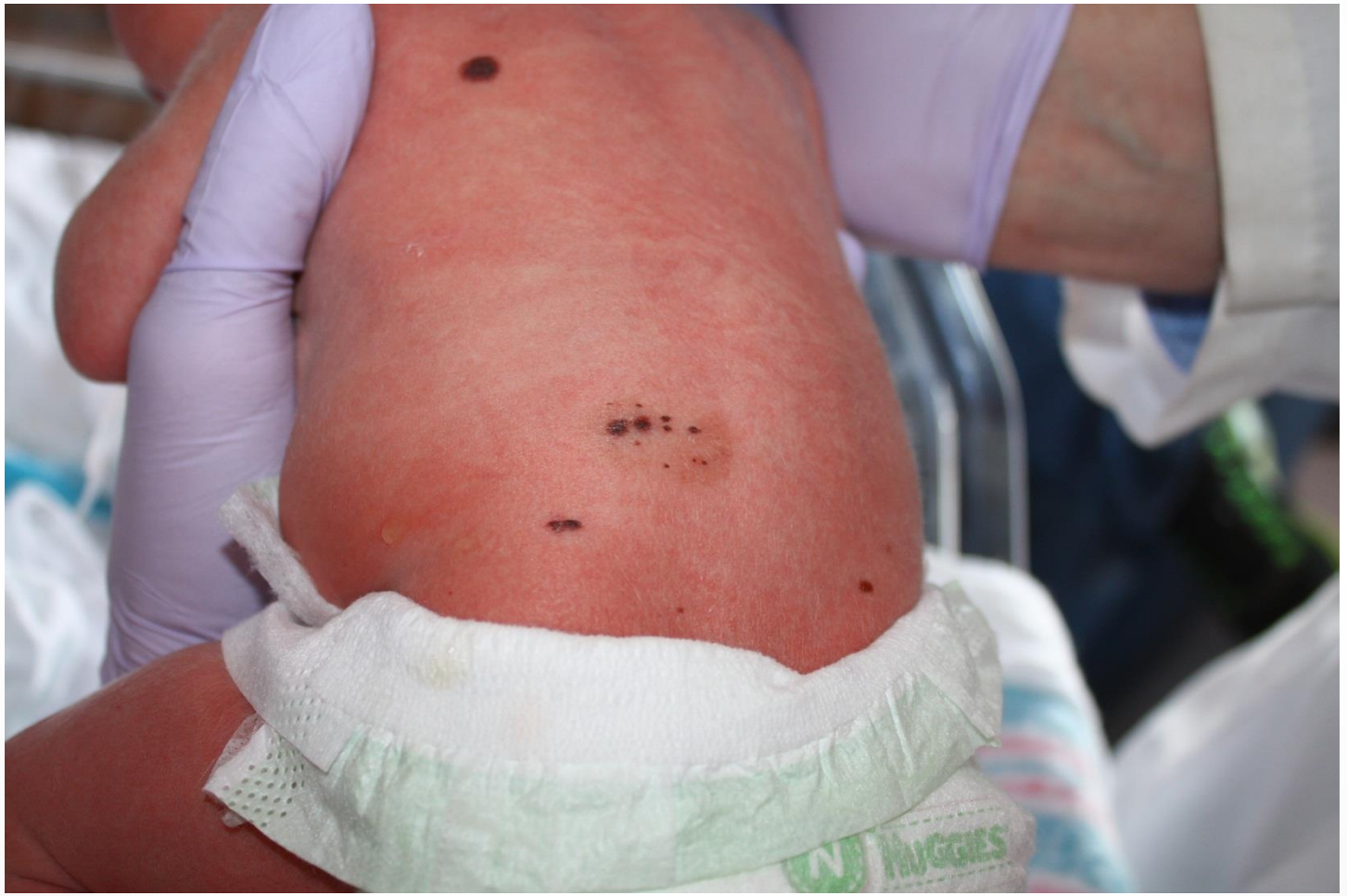






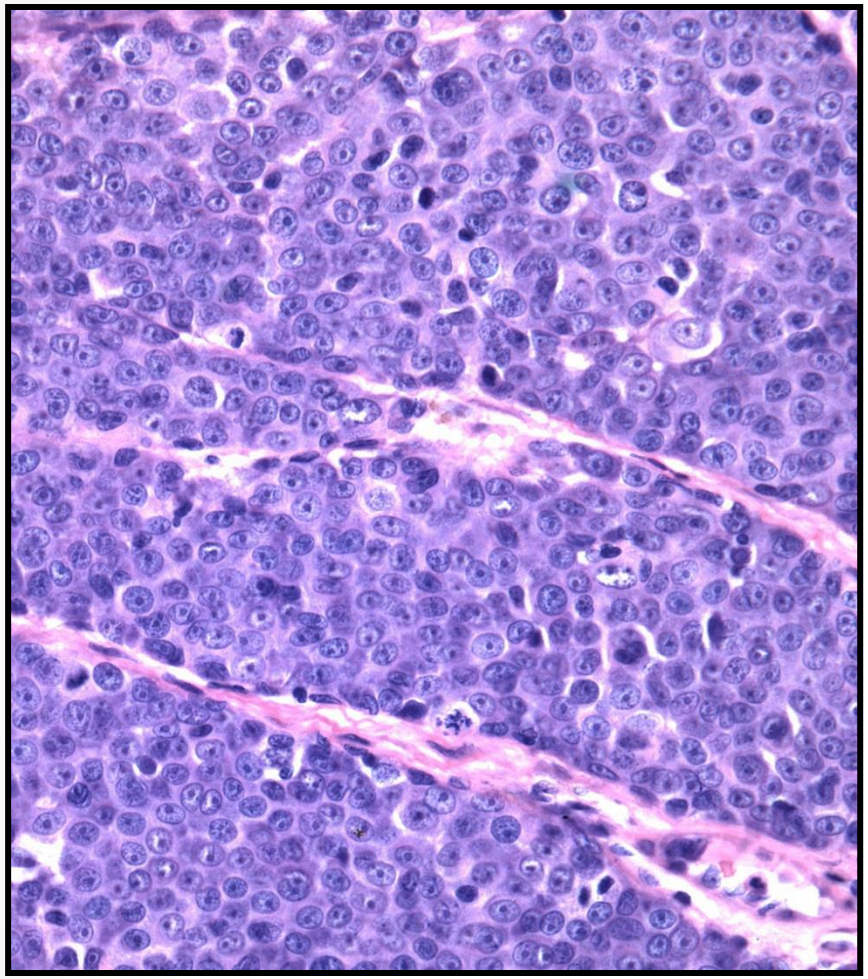
Case Study

- Two months later, male baby born with melanocytic lesions

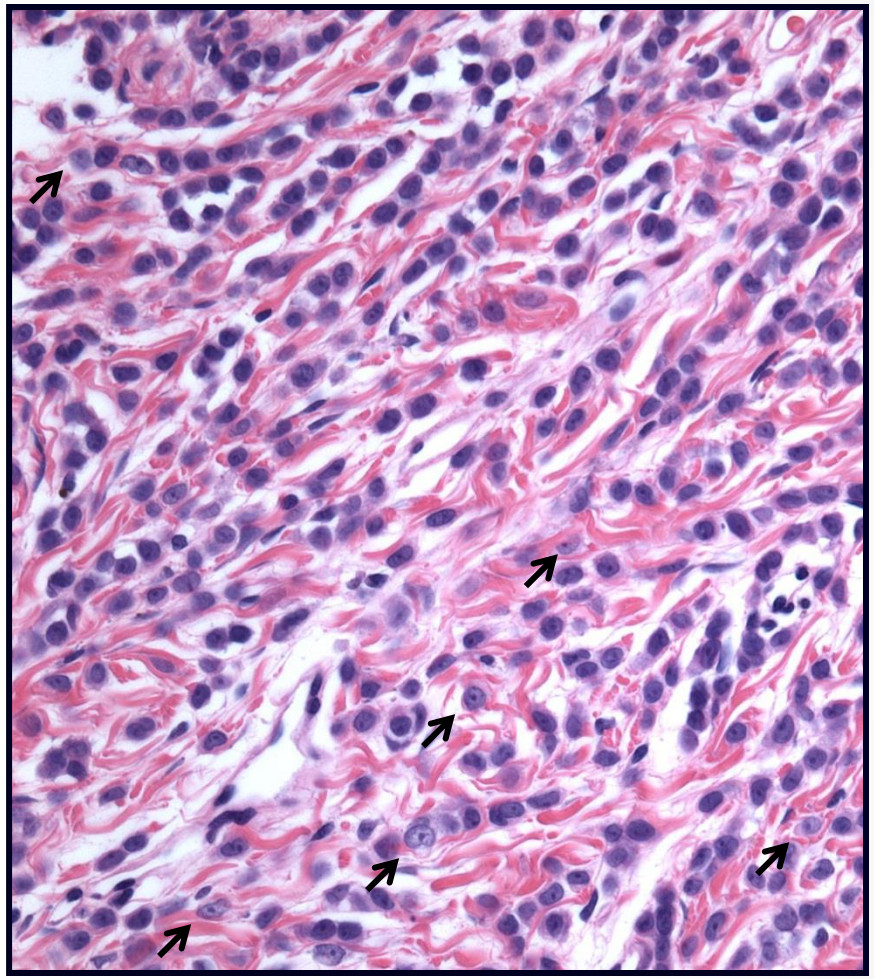


Alomari A, Glusac EJ, Choi J, Hui P, Seeley EH, Caprioli RM, Watsky KL, Urban J, Lazova R. *J Cutan Pathol.* **2015,** *42,* 757-64

Mother



Baby



Metastases or Congenital Nevi?

Mass Spectrometry Analysis

Mother	
Histological Diagnosis	Mass Spectrometry (number of spectra)
Malignant Melanoma	Malignant Melanoma – 29/29

Baby Skin Lesions	
Histological Diagnosis	Mass Spectrometry (number of spectra)
Biopsy A – Indeterminant	Benign Nevus – 23/23
Biopsy B – Indeterminant	Benign Nevus – 9/9

Cells within lesions on baby contained *y* chromosome

Summary

- Mass spectrometry imaging is a powerful technology for clinical applications
- The use of machine learning allows for development of classification algorithms for diagnostic and prognostic assays
- Provides biological insight not available by other techniques
- Helps to provide solutions to clinically challenging problems
- Reproducibility is key to successful clinical studies

Acknowledgements

- Rossitza Lazova
- Heather Anderson
- Katy Smoot