Optimized data analysis pipeline for MALDI MSI based tumor typing from FFPE tissue samples evaluated on six benchmark classification tasks

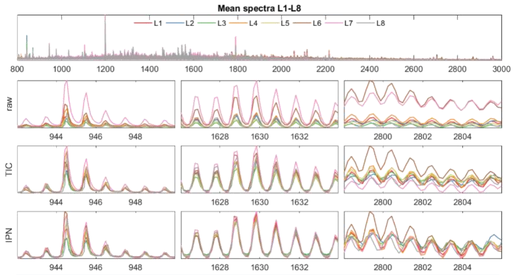
D. Lachmund[1], J. von Schröder[1], T. Boskamp[1,2], L. Hauberg-Lotte[1], J.H. Kobarg[2], S.O. Deininger[3], K. Kriegsmann[4], M. Kriegsmann[4], R. Casadonte[5], J. Kriegsmann[5], P. Maaß[1,2]

ASMS 2019, MP 340

[1]University of Bremen, Bremen, Germany [2]SCiLS, Bremen, Germany [3]Bruker Daltonik, Bremen, Germany [4]University of Heidelberg, Heidelberg, Germany [5]Proteopath, Trier, Germany

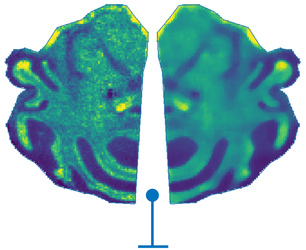
Goals

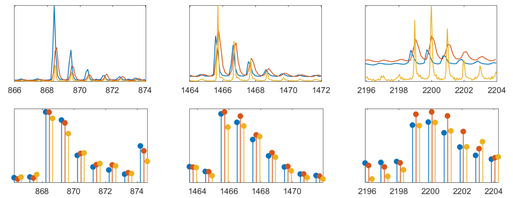
- Develop optimized pre-processing pipeline for MALDI MSI based tumor typing
- Consider different clinical tumor typing and subtyping tasks
- Consider intra- and inter-lab scenarios and different instrument types



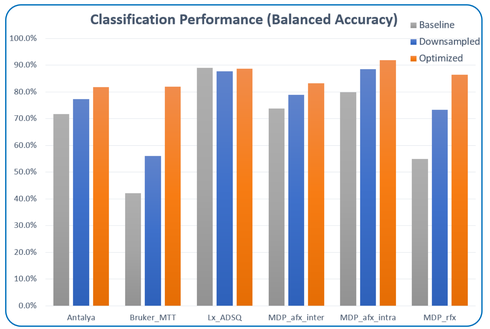
Non-linear **intensity profile normalization** (IPN, bottom row) improves comparability across different acquisitions *(Boskamp et al, ASMS 2018)*

Mild Gaussian kernel **spatial denoising** (right) increases signal-to-noise ratio as compared to original data (left)





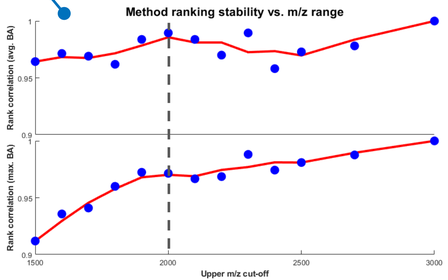
Dimensionality reduction by downsampling to peak areas over 0.4 Da intervals *(Boskamp et al, ASMS 2018)*



Data → **m/z range** → **Normalization** → **Alignment** → **Spatial denoising** → **Downsampling** → **Normalization** → **Transformation** → **Classification**

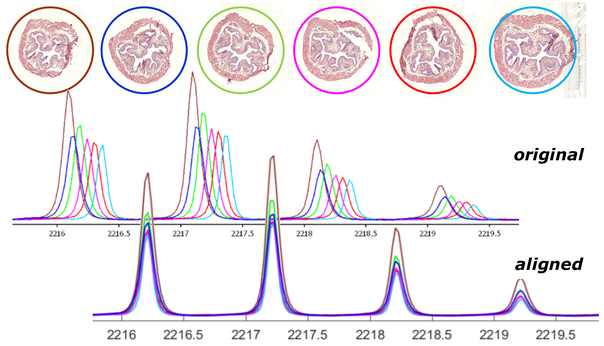
- Benchmark panel acquired from **25 TMAs**, **2031 cores** and **1410 patients** total

Task	Instrument	Description
Antalya	autoflex	• Four tumor entities, 8 TMAs • Lung, pancreas, colon, breast
Bruker MTT	rapiflex	• Six tumor entities on one TMA • Five measurements in four labs • Training and test data from different SOP's
Lx ADSQ	autoflex	• Eight TMAs with mix of adeno- and squamous cell carcinoma, afx
MDP afx inter	autoflex	• Breast, ovary tumors, 5 TMAs • Measured in two labs • Inter-lab cross-validation
MDP afx intra	autoflex	• Same as above, but intra-lab cross-validation
MDP rfx	rapiflex	• Breast, ovary tumors, 5 TMAs • Single lab

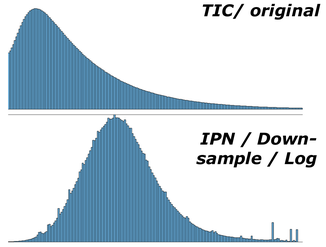


Reducing mass range for feature selection down to 700 … 2000 m/z increases robustness and speed without affecting accuracy

Mass calibration based on statistical peptide mass model reduces misalignment *(Boskamp et al, ASMS 2018)*



Logarithmic transform with appropriate scaling results in more symmetric intensity distributions – beneficial for subsequent LDA classification



- Balanced accuracy 82% and 92%
- **Performance gain** over baseline (TIC only) **9.5 … 39.8% pts.** for five of six tasks
- Mass alignment / downsampling alone yields 5 … 18.5% pts. for five of six tasks

Conclusion

- Systematic investigation of six benchmark problems yields an **optimized pre-processing pipeline** for MALDI MSI tumor typing applications
- **Significant performance gains** achieved in intra- and inter-lab scenarios
- **Improved robustness** towards SOP variations and technical variability

Imaging MS