Cross-normalization for MALDI MSI data improves site-to-site reproducibility

Tobias Boskamp*[1,2], Rita Casadonte[3], Lena Hauberg-Lotte[2], Soeren-Oliver Deininger[1], Jörg Kriegsmann[3], Peter Maass[2]

[1]Bruker Daltonics GmbH & Co. KG, Bremen, Germany; [2]Center for Industrial Mathematics, University of Bremen, Bremen, Germany;
[3]Proteopath GmbH, Trier, Germany; *tobias.boskamp@bruker.com

Introduction

- Applications of MALDI imaging in larger studies often suffer from limited reproducibility, caused by batch effects as well as technical and protocol variation.
- We evaluated different data normalization methods for their capability to reduce technical variation without adversely impacting biological information.
- Our proposed cross-normalization aligns statistical intensity distributions across spectra and samples (intensity profile normalization, IPN), and does a model-driven resampling particularly designed for peptide imaging (peptide mass resampling, PMR) [1].

Methods

- A human teratoma sample (**Fig. 1**) was used, as well as five TMAs of breast and ovarian tumors [2].
- MALDI MSI was done at two labs according to three different protocols and instrumentation (**Fig. 2 left**).
- Different preprocessing pipelines were applied, including TIC and MFC normalization, as well as our proposed cross-normalization (**Table 1, Fig. 3**).
- Different classification tasks were executed, covering intra-sample, inter-patient / slide-to-slide, inter-lab, and cross-protocol scenarios (**Fig. 2 right**)

Results

- In terms of AUC classification accuracy, the proposed cross-normalization scheme outperforms conventional normalization methods (**Fig. 4**).
- Particularly high AUC gains are observed in inter-lab and cross-protocol tasks, but also in some slide-to-slide tasks where stronger batch effects are assumed.
- Both intensity normalization as well as peptide mass resampling are essential for successful cross-normalization.

Fig. 1: MALDI MSI data from a human teratoma. Histological annotations of seven different tissue phenotypes were provided. ▼

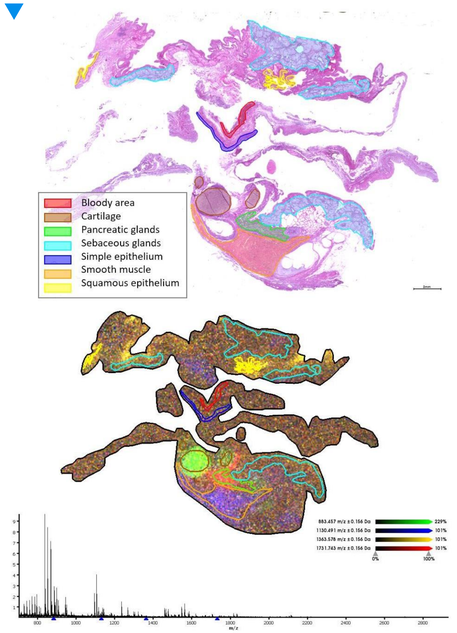


Fig. 2: Experiment design and classification tasks used on the teratoma and tumor TMA datasets. Different configurations of training and test data, from intra-sample to cross-protocol scenarios, all tasks including 2-fold (A vs. B) cross validation. ▼

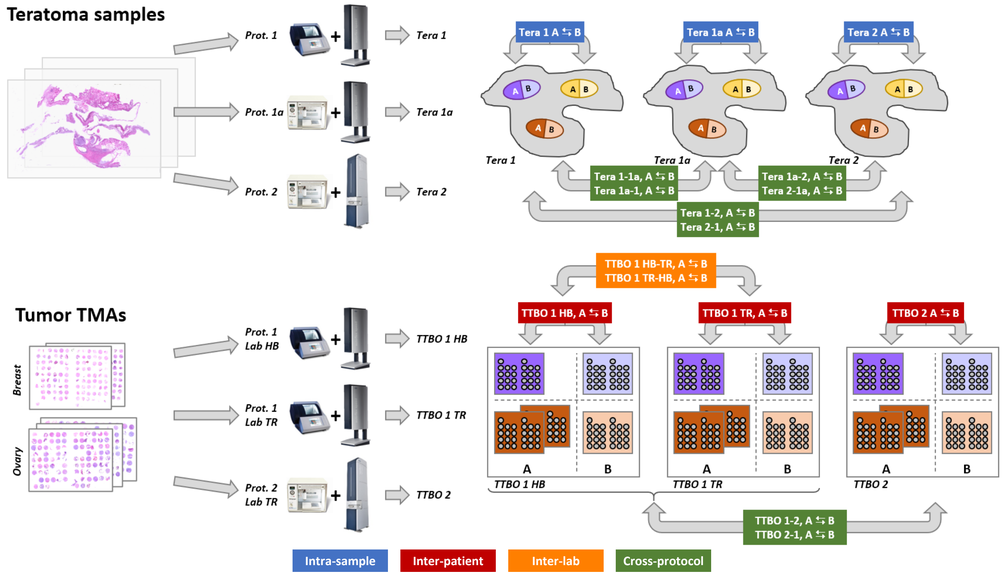


Table 1: Preprocessing methods used in this study ▶

Method	Description
tic	Total ion count normalization
mfc	Inter-sample median fold change normalization
ipn	Intensity profile normalization [1]
cal	Mass recalibration [3]
pmr	Peptide mass resampling [1]

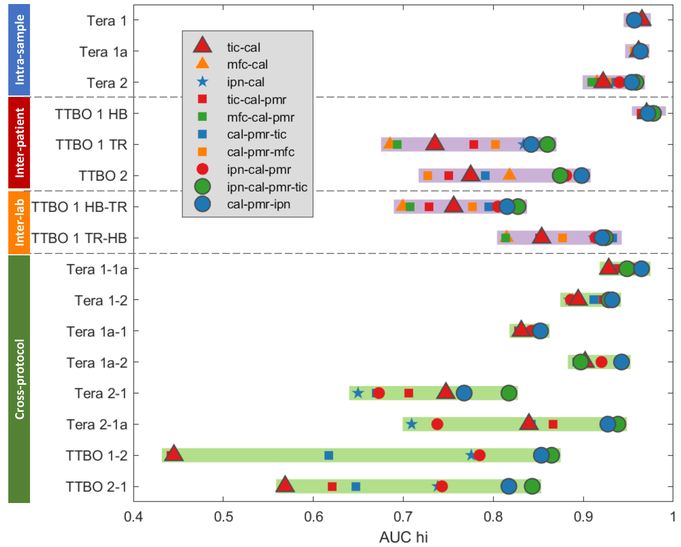
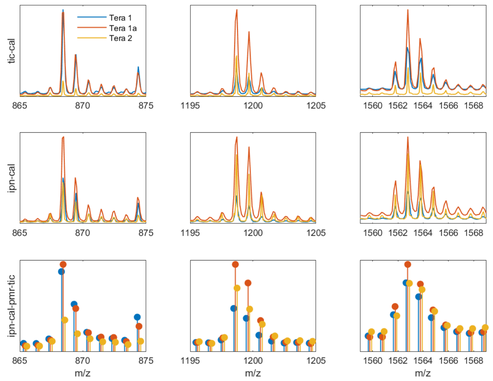


Fig. 4: AUC performance for ▲ different preprocessing methods, grouped by classification task

Fig. 3: Close-ups of spectral peaks after applying different preprocessing methods ▼



Conclusions

- Proposed cross-normalization is highly effective in reducing technical variation, preserves biological information.
- Joint evaluation of multiple experiments is feasible even under conditions where preparation and acquisition protocols are subject to variation.
- Conventional normalization methods fail to compensate batch effects in slide-to-slide, inter-lab or cross-protocol scenarios.

References

[1] Boskamp et al.; Anal. Chem. 2021, DOI: 10.1021/acs.analchem.1c01792
[2] Cordero et al.; Prot. Clin. Appl. 2019, DOI: 10.1002/prca.201700168.
TMA tissue samples kindly provided by Dr. Mark Kriegsmann, Univ. of Heidelberg, Germany
[3] Boskamp et al.; Anal. Chem. 2020, DOI: 10.1021/acs.analchem.9b04473

For research use only. Not for use in clinical diagnostic procedures.

Innovation with Integrity