

Practical Considerations on Normalization in MALDI Imaging

Soeren Deininger¹; Michael Becker¹;
 Eryk Wolski¹; Herbert Kaminski¹;
 Rainer Paape¹; Shannon Cornett²
¹Bruker Daltonik GmbH, Bremen,
 Germany
²Bruker Daltonics, Billerica, USA

Introduction

Normalization of MALDI imaging data is a topic of debate, as it can lead to artifacts. The use of normalization is always based on certain assumptions, which are not always met on real world data. We present here datasets that lead to artifacts with the commonly used normalization on TIC or the vector norm and propose ways to detect and avoid these.

Methods

A mass spectrum can be seen as a vector of intensity values: $\vec{s} = y_1, y_2, \dots, y_n$
 For normalization, the spectrum is divided by a factor: $\vec{s}_{normalized} = \frac{1}{f} \vec{s}$

The normalization factors used in this work were:

$$TIC (T): f = \sum_i |y_i| \quad \text{Vector norm (V): } f = \sqrt{\sum_i y_i^2}$$

$$\text{Noise (N): } d_i = y_i - y_{i-1} \quad f = \text{median}(d_i - \text{median}(d))$$

$$\text{Median (M): } f = \text{median}(y_i)$$

TIC with exclusion masses (TX), here prior to TIC normalization user defined mass ranges are excluded from the mass spectrum: $\hat{y}_i = \begin{cases} 0, & i_{lower} < i < i_{upper} \\ y_i, & \text{else} \end{cases}$

Results

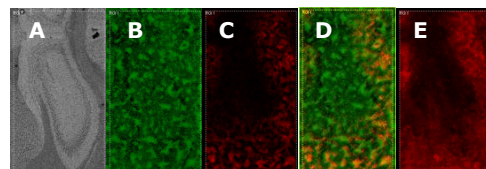


Fig. 1 20µm imaging on rat hippocampus: **A)** Optical image of tissue, **B)** Optical image of HCCA matrix layer (false color, matrix in green) **C)** Mass spectrometric image. **D)** overlay of B and C, **E)** Normalized mass spectrometric image from C), vector norm. **Only the normalized image shows a result in good agreement with the histology.**

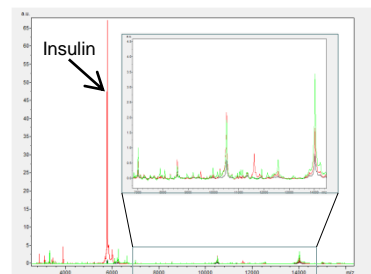


Fig. 2 Average spectra from a mouse pancreas imaging dataset. Red: Average spectra of islets of Langerhans Green: Average spectrum from "normal" region. **In the islet region, there is a very intensive signal from insulin**

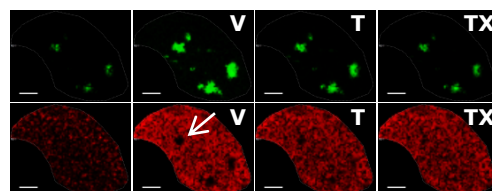


Fig. 3 MALDI images of insulin signal (upper panel) and "housekeeping" signal (lower panel) with no normalization, vector norm (V), TIC (T) and TIC with exclusion of the insulin signal (TX). **Vector norm leads to artifacts: Enlarged islets of langerhans and "holes" in background**

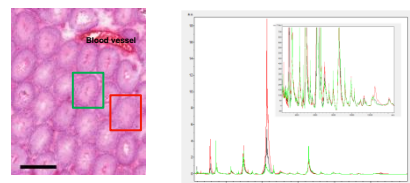


Fig. 4 H&E image of rat testis and averaged mass spectra of selected seminiferous tubuli. Tubuli differ in maturation state of spermatides. (Scalebar 500µm). This part of the testis section is used in figures 5 and 6. **Some tubuli show an intensive signal with a large area**

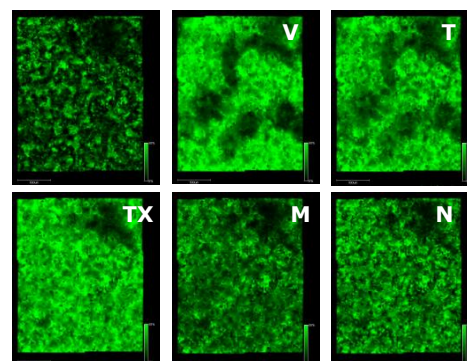


Fig. 5 Testis, MALDI images (20µm resolution, HCCA matrix) of a "housekeeping" protein that is present everywhere except the blood vessel with no normalization, normalization on the vector norm (V), TIC (T), TIC with exclusion of the intense 6kD signal (TX), median (M) and noise (N). **Both vector norm and TIC lead to artifacts due to the 6kD signal with a large area. Interpretation of the TIC (or V) image would lead to the wrong conclusion that the displayed mass is down-regulated in certain spermatide development stages. TIC with exclusion, median and noise normalization lead to a correct interpretation.**

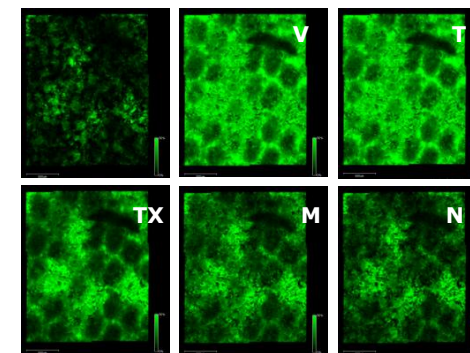


Fig. 5 Testis, MALDI images (20µm resolution, HCCA matrix) of a selected no normalization, normalization on the vector norm (V), TIC (T), TIC with exclusion of the intense 6kD signal (TX), median (M) and noise (N). **Both normalization on vector norm and TIC lead to artifacts and the wrong conclusion that the protein is most abundant in the interstitium. Without normalization, the presence of the signal in the interstitium is not seen. Only TIC with exclusion, median and noise normalization lead to a correct interpretation.**

The assumptions regarding decisions on normalizations are:

- No normalization: The matrix layer and endogenous effects such as spatially different salt or lipid content do not influence the result.
 - Vector norm: The overall intensities of the signals have a uniform spatial distribution.
 - TIC: The overall areas of the signals have a uniform spatial distribution.
 - Median: The baseline of all spectra in the dataset shows a uniform spatial distribution.
 - Noise: The noise level of all spectra in the dataset shows a uniform spatial distribution.
- If these assumptions are not met, the corresponding normalization can lead to artifacts.

● ASMS 2010, TP32-542

Summary

The normalization is sometimes necessary to see the true distribution of the signals. If signals with a very large area are present in certain regions it can lead to artifacts. Exclusion of these signals for the calculation of the TIC solves this. The comparison of data after TIC and median or noise normalization can be used to detect the inapplicability of a TIC normalization. The median/noise or TIC with exclusion of mass ranges can be used in such cases.

Acknowledgement

We thank Axel Walch and Sandra Rauser from the Helmholtz-Centre in Munich, Germany for the pancreas sample and Charles Pineau and Régis Lavigne from the University of Rennes, France for the testis sample.

Conclusions

- Normalization should be applied with care
- Normalization on TIC with the exclusion of selected mass ranges and on median or noise should be considered as part of the "standard toolbox" for MALDI imaging