A novel software for the characterization of microorganisms using lipid phenotyping and statistical analysis of MALDI-MS data

Gema Méndez-Cervantes[1], Luis Mancera[1], Gerald Stübiger[2]

[1]Clover Bioanalytical Software, [2]Medical University of Vienna

Introduction

MALDI-TOF MS has shown a high potential for microbial identification, typically using intact cells or cell extracts. The main limitation of the technique is that the current capacity of spectral databases is not enough for other microbes such as yeasts and fungi. Lipids show the potential to serve as additional molecular parameters for biotyping purposes.

We present a novel software for the classification of yeasts and filamentous fungi based on analysis of the lipid composition.

Our method is based on a customized pre-processing step followed by unsupervised machine learning. It is database-free and identifies which lipid parameters are important for discrimination.

Besides, we show the use of the software to perform a reproducibility study. Having two individuals (1, 21) and two samples (a, b) each corresponding to a different treatment with placebo and drug. The matching between samples and treatments is unknown. We study the reproducibility of the analysis and show that both individual treatments lead to a difference at the molecular level.

Material and methods

The yeasts were grown on malt extract agar plates for two days at 30°C. Filamentous fungi were grown on the same agar for one week at room temperature.

For lipid extraction, samples were harvested and transferred to 1.5 mL reaction tubes, washed with UHQ and centrifuged.

The supernatants containing lipids were transferred to 300 µL glass vials and stored at -30°C for further analysis.

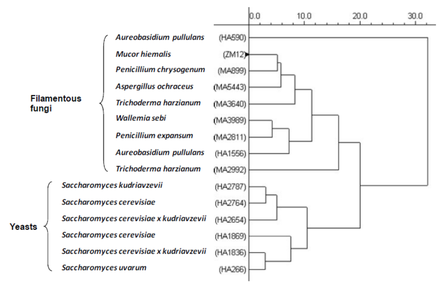
For MALDI-TOF analysis we used THAP (10 mg/mL) dissolved in AcOH: MeOH = 70:30 (v/v) for positive mode and 9AA matrix (10 mg/mL) dissolved in ISO: CAN = 60:40 for negative mode.

Samples were analyzed using an AXIMA-series (Shimadzu, Manchester, UK) curved-field reflectron TOF mass spectrometer, and data acquired using Launchpad 2.9.3 (Shimadzu, UK).

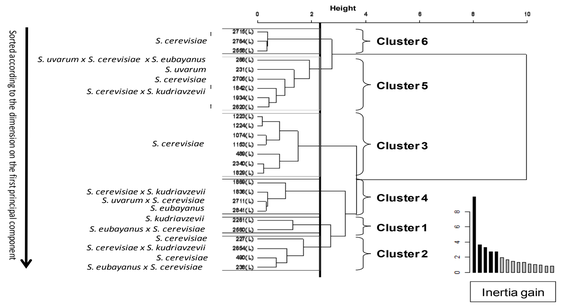
Processing and classification was done using Clover MS Software (Clover Bioanalytical Software, Spain).

Differentiation of yeasts and fungi

- We use MALDI-TOF based lipid phenotyping.
- We analysed lipid extracts of a selection of 15 different species of prominent genera of yeasts (e.g. *Saccharomyces*) and filamentous fungi (e.g. *Penicillium*, *Aspergillus*, *Trichoderma*).



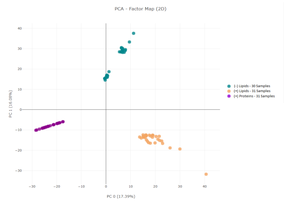
- **Yeasts-fungi separation**: Based on MALDI spectra of lipid extracts. Using information in the range m/z 400-900.
- Good separation, but some closely related strains were found in different clusters → indicates subtle difference in phenotype of closely related strains.



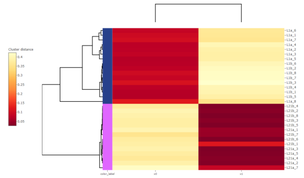
- ***Saccharomyces***: We analysed 26 *S. cerevisiae* strains and interspecies hybrids of genotypically very closely related yeasts.
- Reference peaks were aligned and blank was removed. PCA was used to reduce dimensionality, and hierarchical clustering applied to the first two PCs.
- Thereby, the clustering leads to a distinct separation of certain *S. cerevisiae* strains and interspecies hybrids with other *Saccharaomyces* species (e.g. *S. kudriavzevii*, *S. eubayanus*, *S. uvarum*).

Reproducibility study

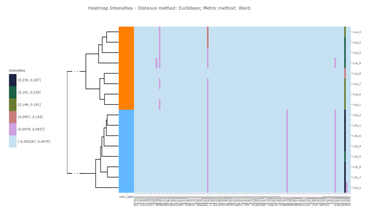
- We have three datasets: (+)Lipids, mass range 360-850 Da); (-)Lipids (530-950 Da), (+)Proteins (2000-18000 Da).
- We performed 6-8 measurements of each sample.
- We submitted all relevant peaks in the mass range for analysis.



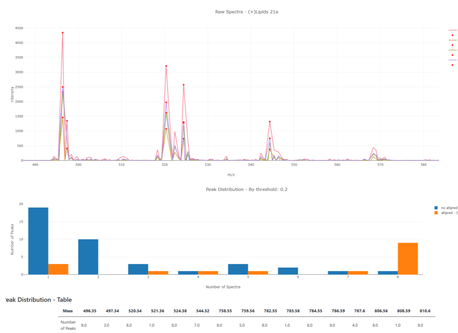
- **PCA**: Used for clustering the three datasets. All spectra peaks were aligned with a mass accuracy of +/- 1000 ppm. Factor Map of (PC0, PC1 – 9.72% of the variance) represented three well defined clusters.



- **Heatmap**: Used for clustering -Lipids dataset. Performed PCA for obtaining two clusters. Calculated the Euclidean distance between each sample and each cluster centre. Performed a Heatmap with these distances showing how individuals (1, 21) clustered in separated groups.



- **Clustering**: Used for -Lipids individual 1 dataset. Heatmap obtained from the Peak Matrix resulted of the spectra alignment with a threshold of 3% for detecting peaks. Treatments (a, b) were well differentiated.



- **Peak Distribution:** Used for study the reproducibility of the sample +Lipids21a measurements. Threshold of 20% for detecting peaks. The Histogram represented the distribution of these peaks before and after alignment. The number of peaks detected at the same position for all sample before alignment was 1/40 and after, 9/16.

Conclusion

- We have shown the influence of the lipid extraction method on the quality of the mass spectra and the reproducibility of the lipid profiles.
- We have shown the usefulness of Clover MS software to differentiate 26 different strains of fungi, as *Penicillium*, *Aspergillus* or *Thricoderma*, and yeasts as *Saccharomyces*.
- Our classification results show difference between the lipid phenotypes of even closely related species.
- We also show a reproducibility study where we extract two samples corresponding to a different treatment in a set of individuals. Our software could separate the data according to sample type, patient and treatment without any prior labelling to a reference mass list.

References

G. Stübiger, M. Wuczkowski, L. Mancera, K. Lopandic, K. Sterflinger and O. Belgacem. Characterization of Yeasts and Filamentous Fungi using MALDI Lipid Phenotyping. Journal of Microbiological Methods. Vol **130**, pp. 27-37. November 2016.