Comparative analysis of ROI-MCR and Compounds Discover protocols for preprocessing LC-MS signals in untargeted metabolomic studies of Parmigiano Reggiano cheese

Caterina Durante, Samuele Pellacani, Marina Cocchi, Lorenzo Strani

University of Modena and Reggio Emilia, Dept. Chemical and Geological Sciences, via Campi 103, Modena, 41125, Italy

caterina.durante@unimore.it

In recent years, the advancement of high-resolution mass spectrometry technologies has enabled the development of untargeted approaches for food product analysis [1,2], which do not require prior knowledge of the compounds being investigated, leading to the discovery of unexpected or unidentified compounds. However, in the use of liquid chromatography-mass spectrometry (LC-MS) or hyphenate techniques in general, the produced signals can reach sizes up to gigabytes and for their analysis, it is necessary to reduce the file size and compress the information to avoid memory issues on limited computers. Compression must be carefully studied to prevent the loss of useful information, eliminating possible sources of error or spurious variability. Furthermore, it is of utmost importance to use methodologies able to resolve overlapped signals and depict 'identitation' features.

In this study, based on an untargeted analysis, LC-MS signals, obtained from the metabolic profile analysis of Parmigiano Reggiano DOP cheese samples, were preprocessed using two different methodologies, namely the ROIMCR protocol [3] and the Compounds discover software, highlighting the strengths and weaknesses of both procedures. Finally, the characteristic features extracted from both procedures were subjected to Principal Component Analysis to differentiate Parmigiano Reggiano samples based on their geographic origin.

References

[1] L. Laghi et al. *TrAC* 59 (2014) 93–102
[2] M. Herrero et al. *Mass Spectrometry Reviews*, 2012, 31, 49–69
[4] Gorrochategui et al. *BMC Bioinformatics* (2019) 20:256