

MetProc: separating measurement artifacts from true metabolites in an untargeted metabolomics experiment.

Mark D. Chaffin^{1,2,†}, Liu Cao^{1,2,17,†}, Amy A. Deik³, Clary B. Clish³, Frank B. Hu^{1,4}, Miguel A. Martínez-González^{4,5,6}, Cristina Razquin^{5,6}, Monica Bullo^{6,7}, Dolores Corella^{6,8}, Enrique Gómez-Gracia^{6,9}, Miquel Fiol^{6,10}, Ramon Estruch^{6,11}, José Lapetra^{6,12}, Montserrat Fitó^{6,13}, Fernando Arós^{6,14}, Lluís Serra-Majem^{6,15}, Emilio Ros^{5,16}, and Liming Liang^{1,2,}*

*Contact E-mail for Corresponding Author: lliang@hsph.harvard.edu

† These authors contributed the same

¹Department of Epidemiology, Harvard TH Chan School of Public Health, Boston, MA 02115, USA,

²Department of Biostatistics, Harvard TH Chan School of Public Health, Boston, MA 02115, USA,

³Metabolomics Platform, Broad Institute of MIT and Harvard, Cambridge, MA 02142, USA,

⁴Department of Nutrition, Harvard TH Chan School of Public Health, Boston, MA 02115, USA,

⁵Department of Preventive Medicine and Public Health, School of Medicine, University of Navarra, 31009 Pamplona, Spain,

⁶Ciber Fisiopatología de la Obesidad y Nutrición (CIBEROBN), Instituto de Salud Carlos III (ISCIII), 28029 Madrid, Spain,

⁷Human Nutrition Department, Hospital Universitari Sant Joan, Institut d'Investigació Sanitària Pere Virgili, Universitat Rovira i Virgili, E-43204 Reus, Spain,

⁸Department of Preventive Medicine, University of Valencia, 46010 Valencia, Spain,

⁹Department of Preventive Medicine, University of Malaga, 29071 Malaga, Spain,

¹⁰Institute of Health Sciences, Instituto de Investigación Sanitaria de Palma, 07120 Palma de Mallorca, Spain,

¹¹Department of Internal Medicine, IDIBAPS, Hospital Clínic, 08036 Barcelona, Spain,

¹²Department of Family Medicine, Primary Care Division of Sevilla, San Pablo Health Center, 41007 Sevilla, Spain,

¹³Cardiovascular and Nutrition Research Group, IMIM, Institut de Recerca Hospital del Mar, Parc de Salut Mar, 08003 Barcelona, Spain,

¹⁴Department of Cardiology, University Hospital of Alava, 01009 Vitoria, Spain,

¹⁵Research Institute of Biomedical and Health Sciences, University of Las Palmas de Gran Canaria, 35016 Las Palmas, Spain,

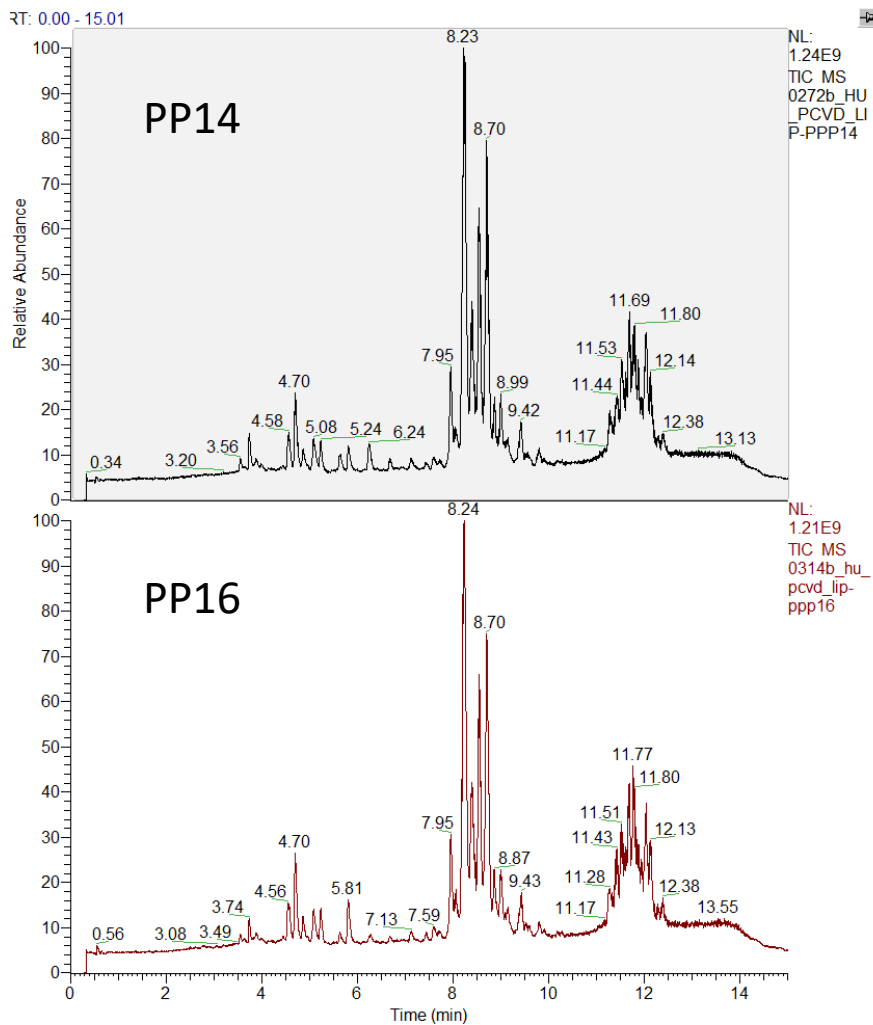
¹⁶Lipid Clinic, Endocrinology and Nutrition Service, Hospital Clínic, Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), 08036 Barcelona, Spain,

¹⁷Computational Biology Department, School of Computer Science, Carnegie Mellon University, Pittsburgh, PA 15213

Table of content

Figure S1. Panel 1-20. We randomly took 20 metabolic features MetProc had determined to reject, and inserted those m/z and retention times into a targeted software program called Trace Finder (Thermo Fisher Scientific; Waltham, MA). We were able to visually inspect every PREDIMED sample and confirm the abundance of each rejected metabolite, in comparison to MetProc.

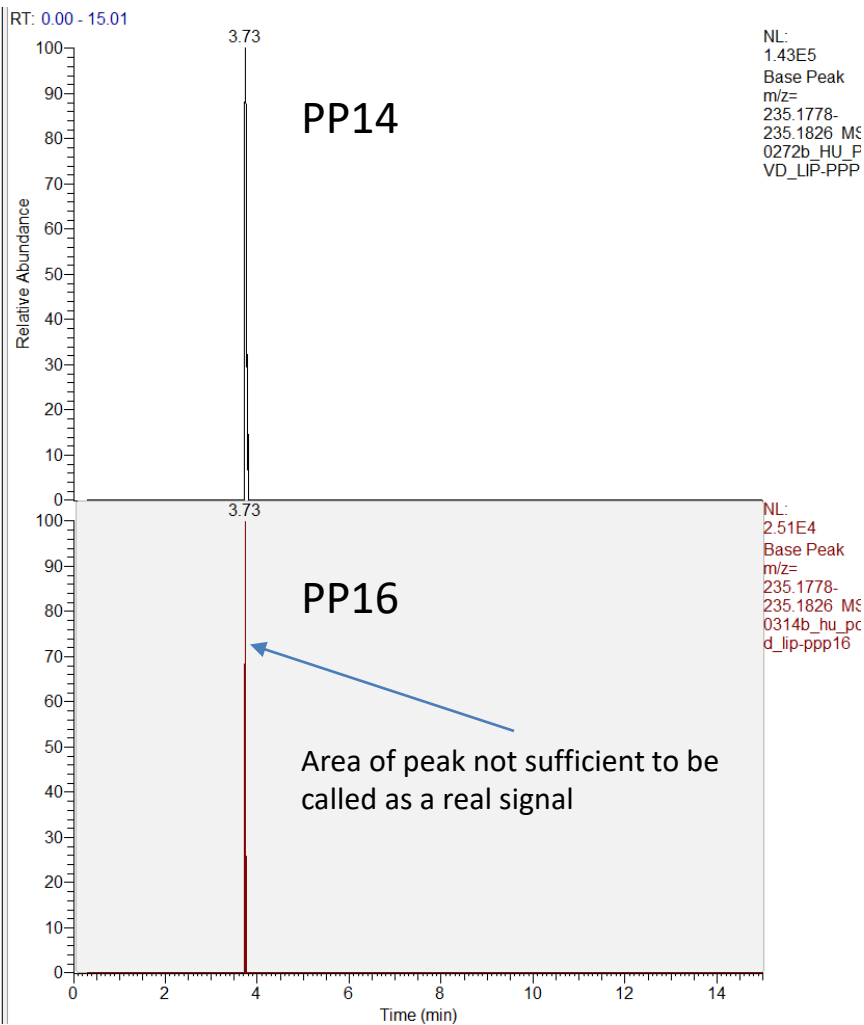
a. Total Ion Chromatograms



b. Extracted Ion Chromatograms

Q1569 m/z 235.1802; RT 3.74

Matches missingness pattern of MetProc

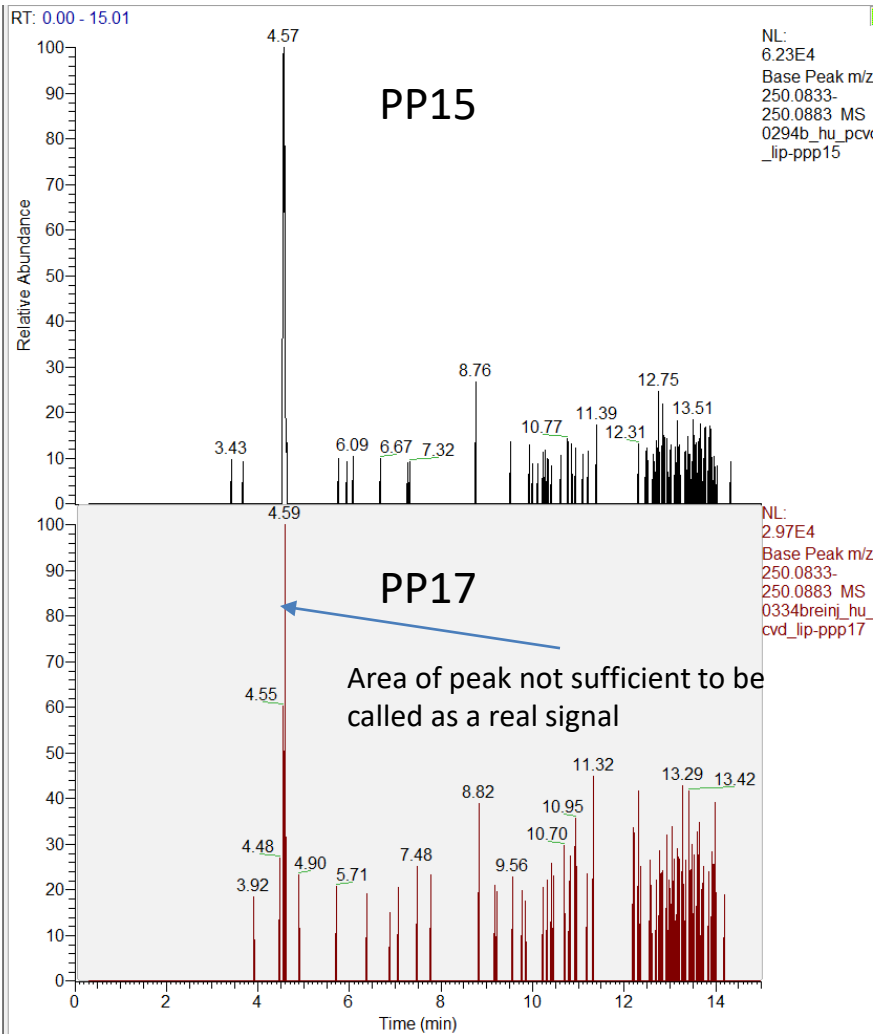
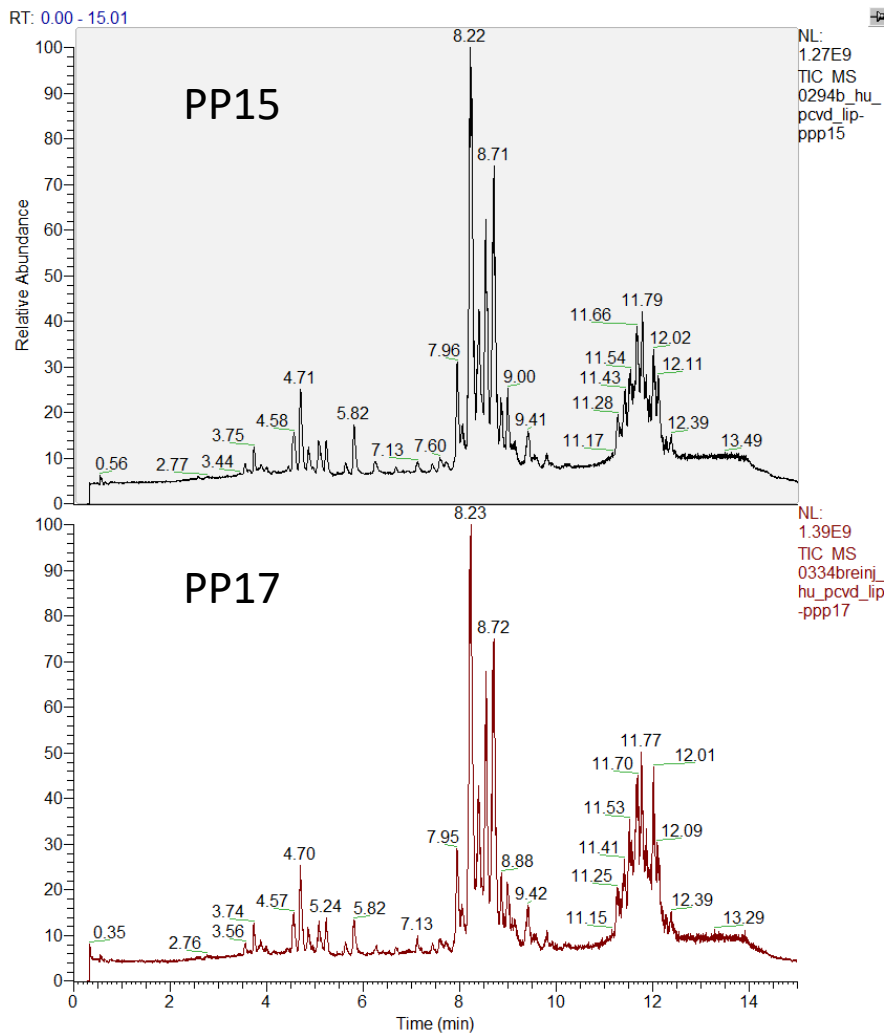


a. Total Ion Chromatograms

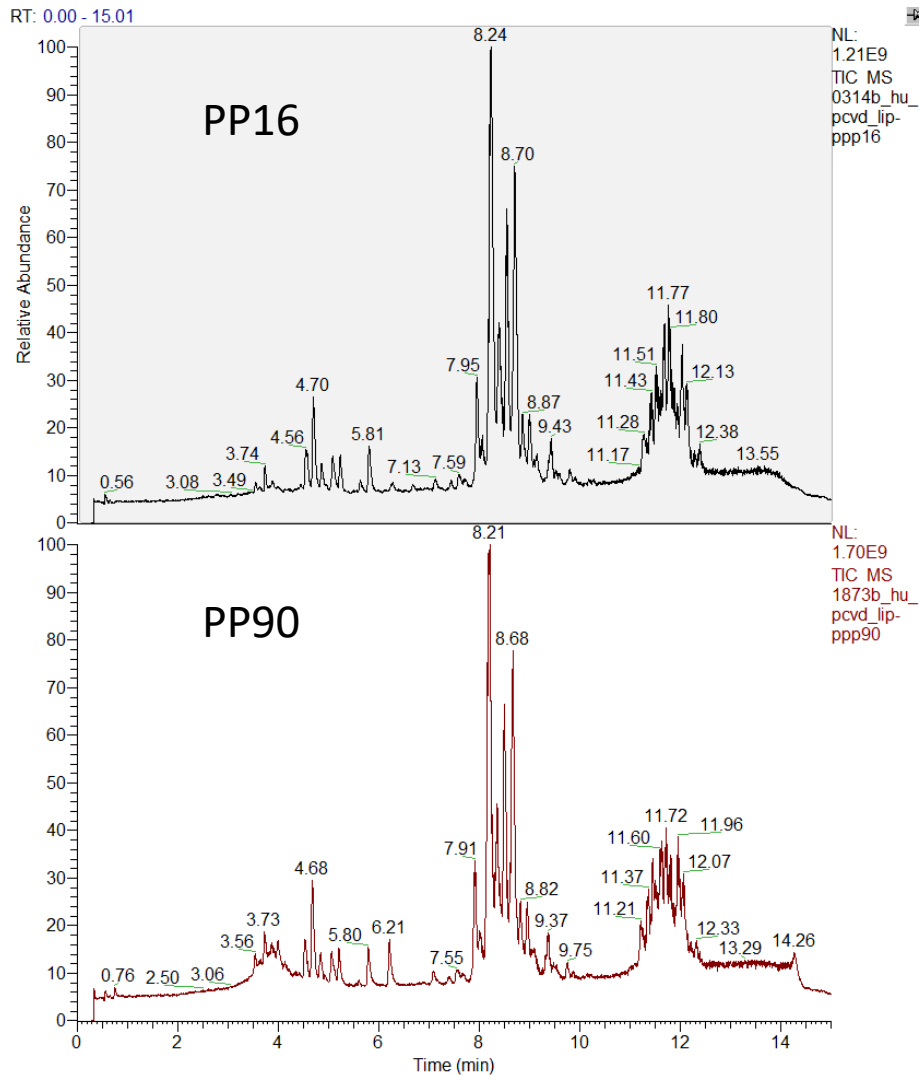
b. Extracted Ion Chromatograms

Q1642 m/z 250.0858; RT 4.6

Matches missingness pattern of MetProc



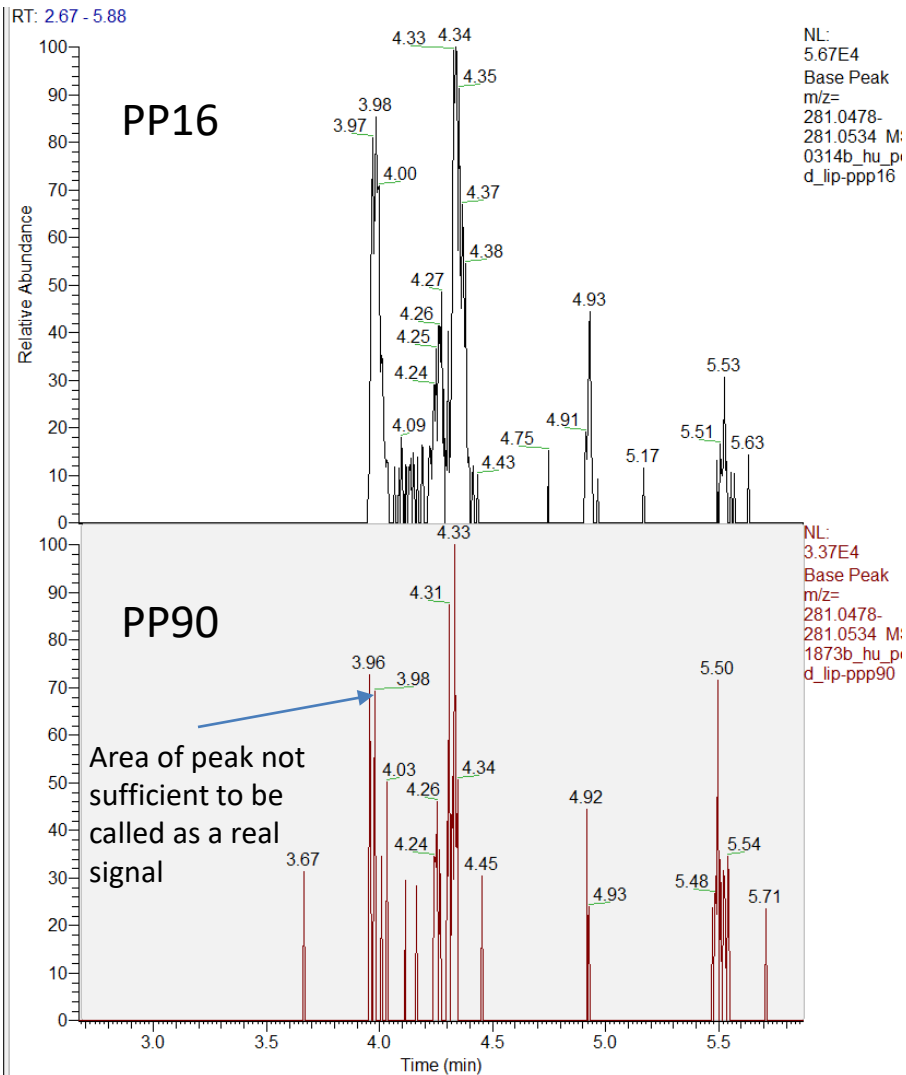
a. Total Ion Chromatograms



b. Extracted Ion Chromatograms

Q1797 m/z 281.0506; RT 3.98

Matches missingness pattern of MetProc



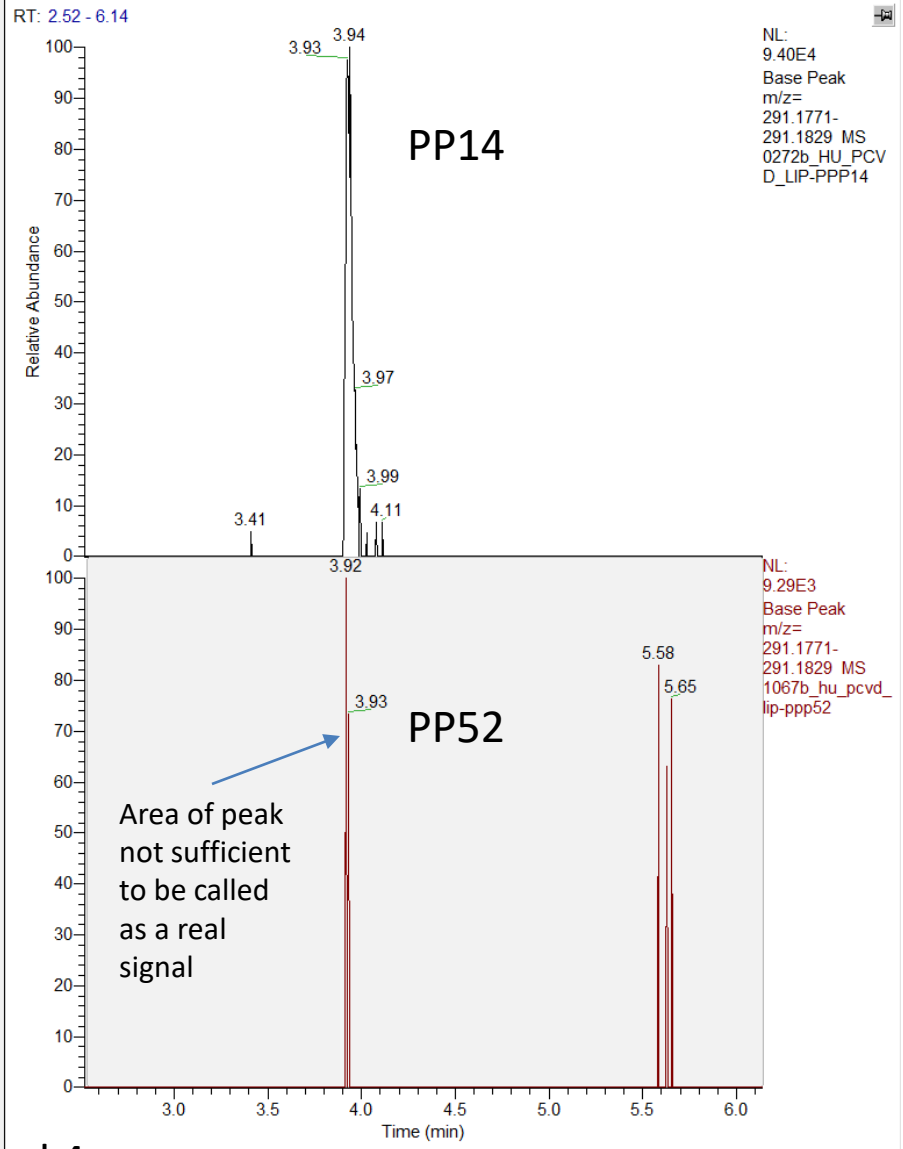
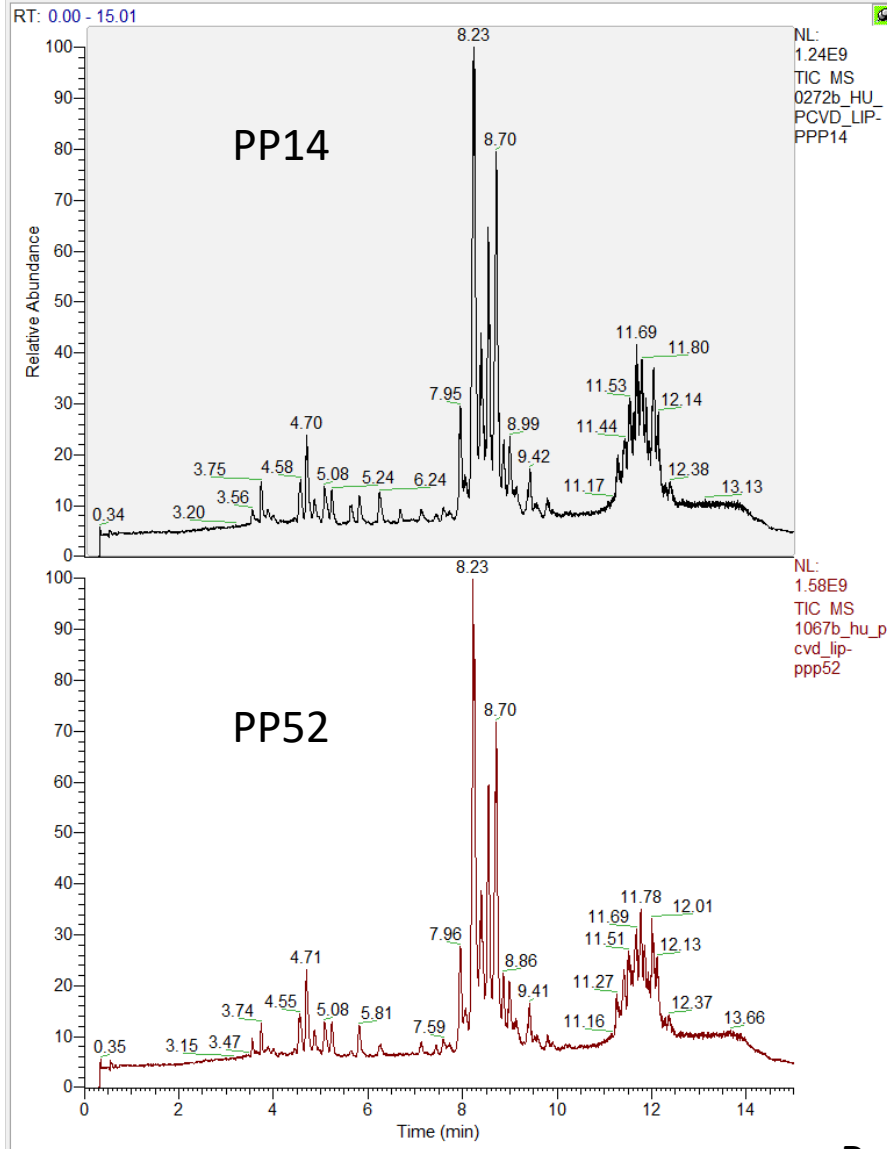
Panel 3

a. Total Ion Chromatograms

b. Extracted Ion Chromatograms

QI860 m/z 291.1800; RT 3.94

Matches missingness pattern of MetProc



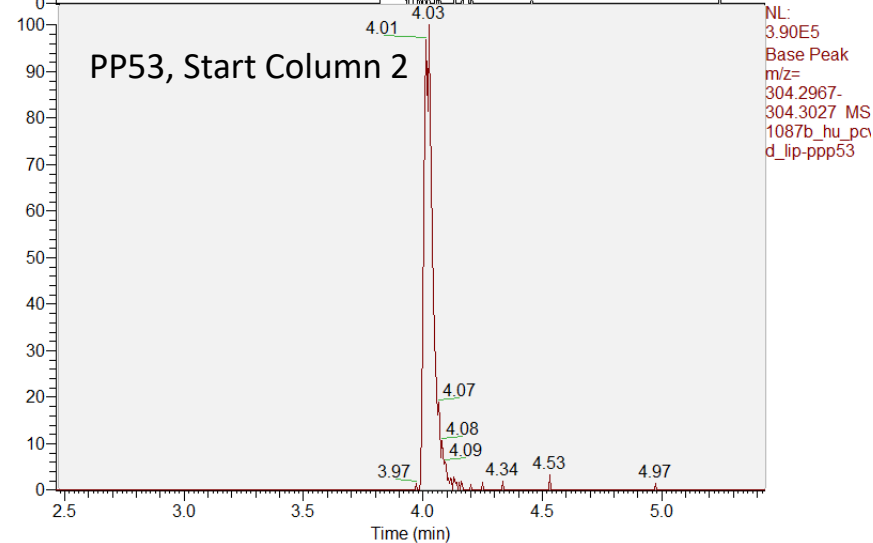
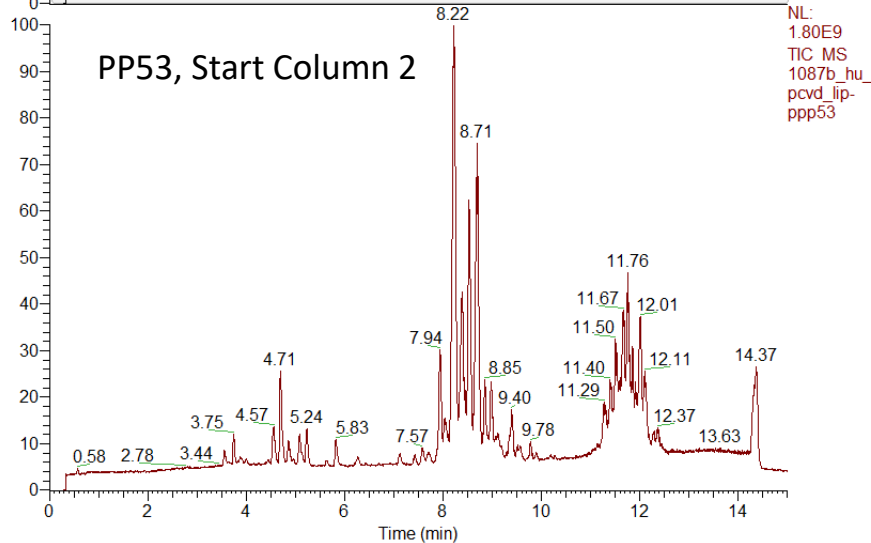
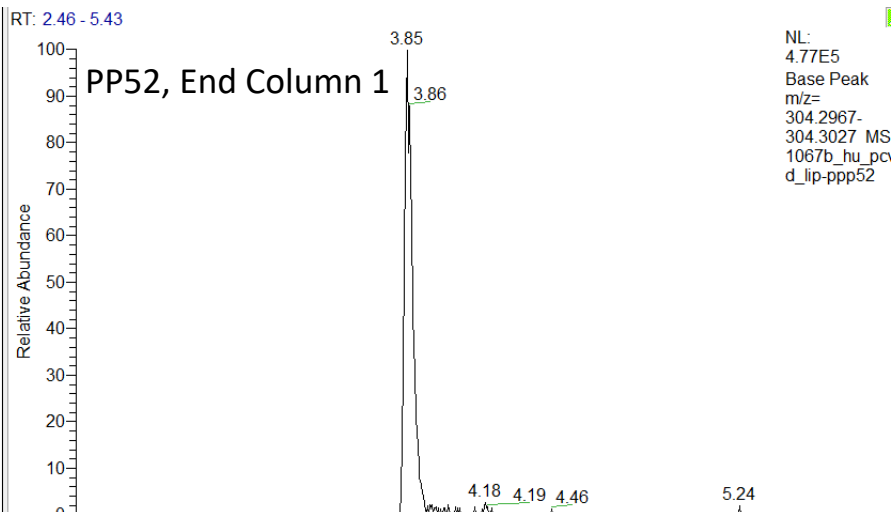
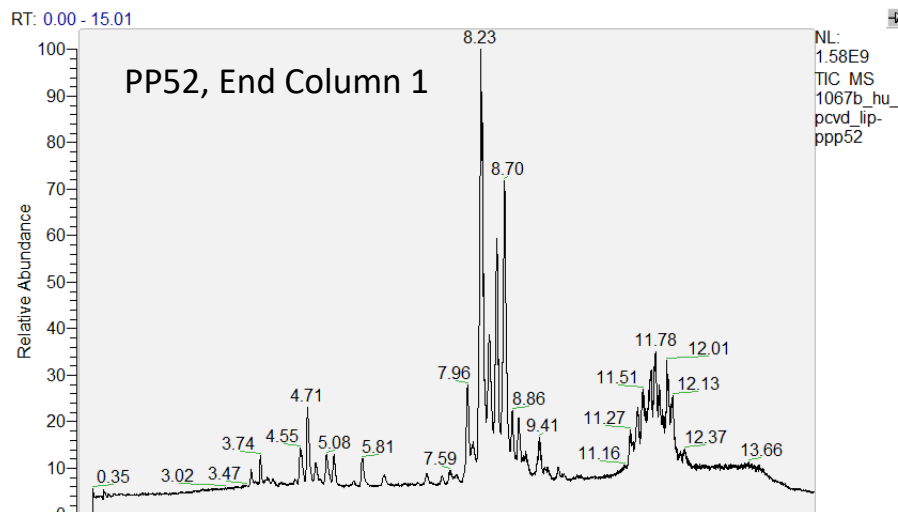
Panel 4

a. Total Ion Chromatograms

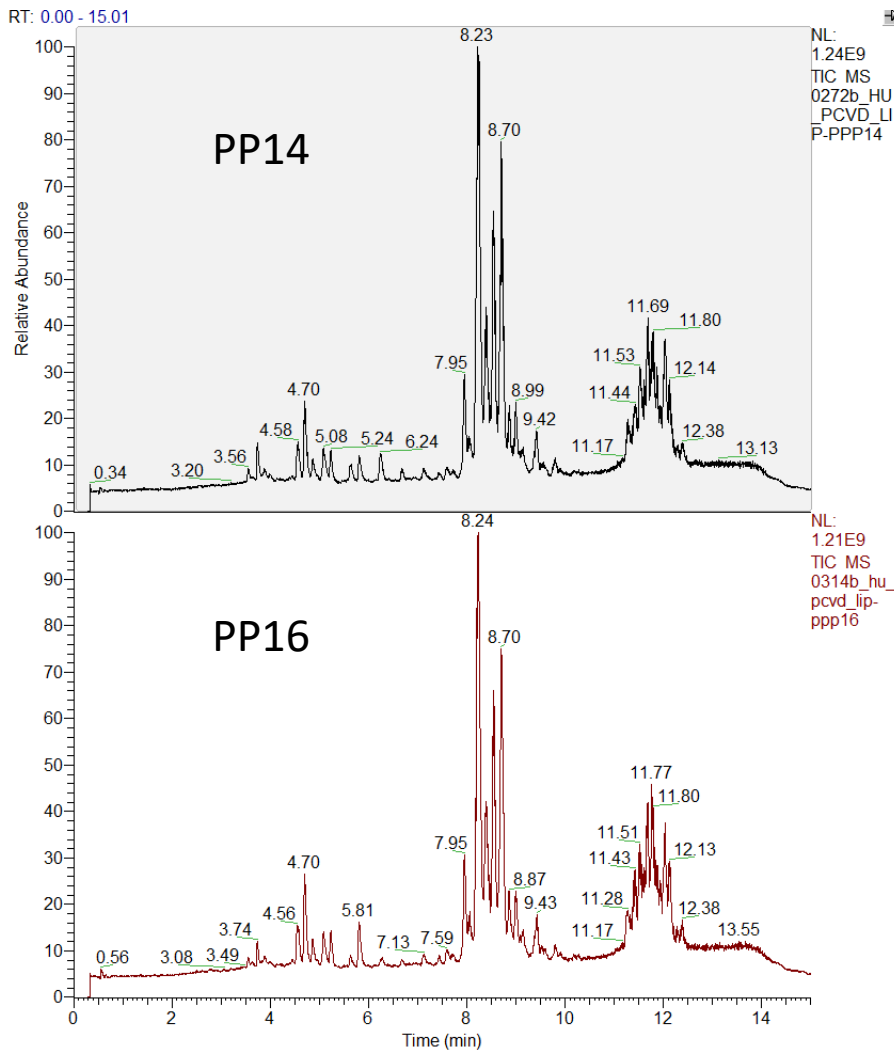
b. Extracted Ion Chromatograms

QI975 m/z 304.2997; RT Col 1 3.85; Col2 4.03

QI did not align the RT correctly between columns



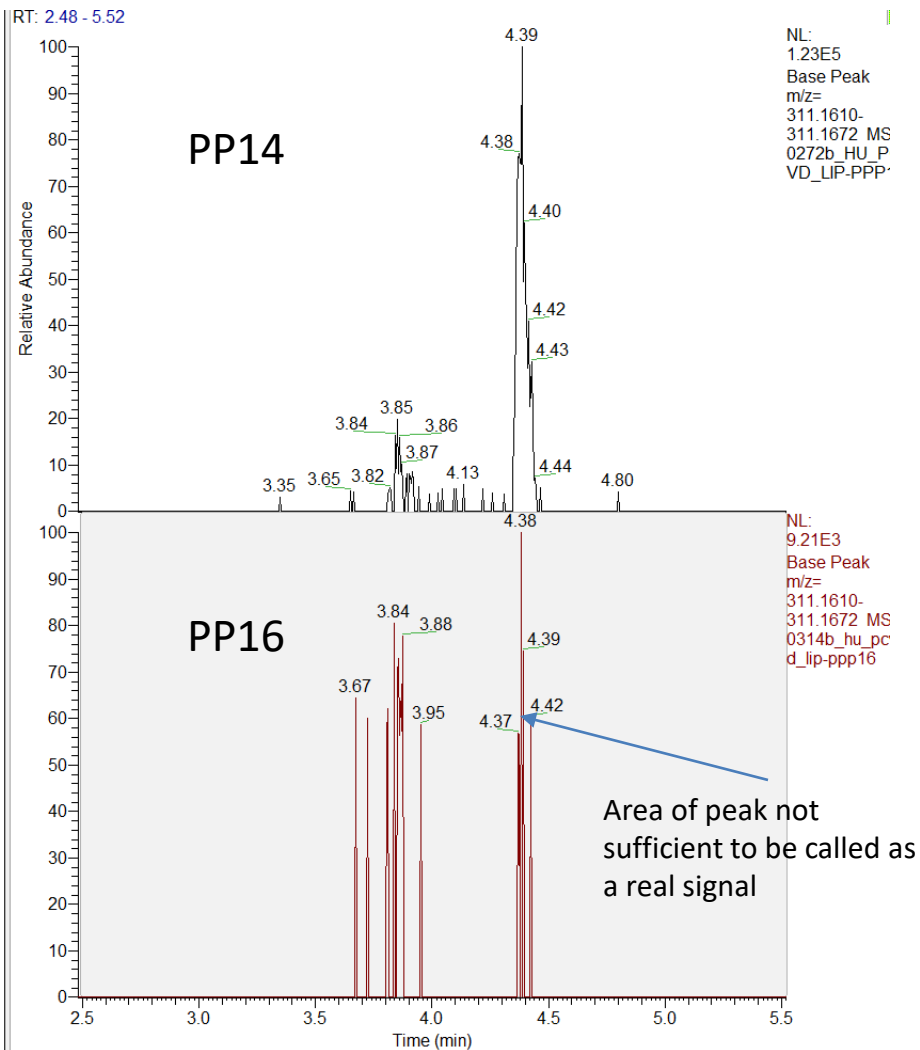
a. Total Ion Chromatograms



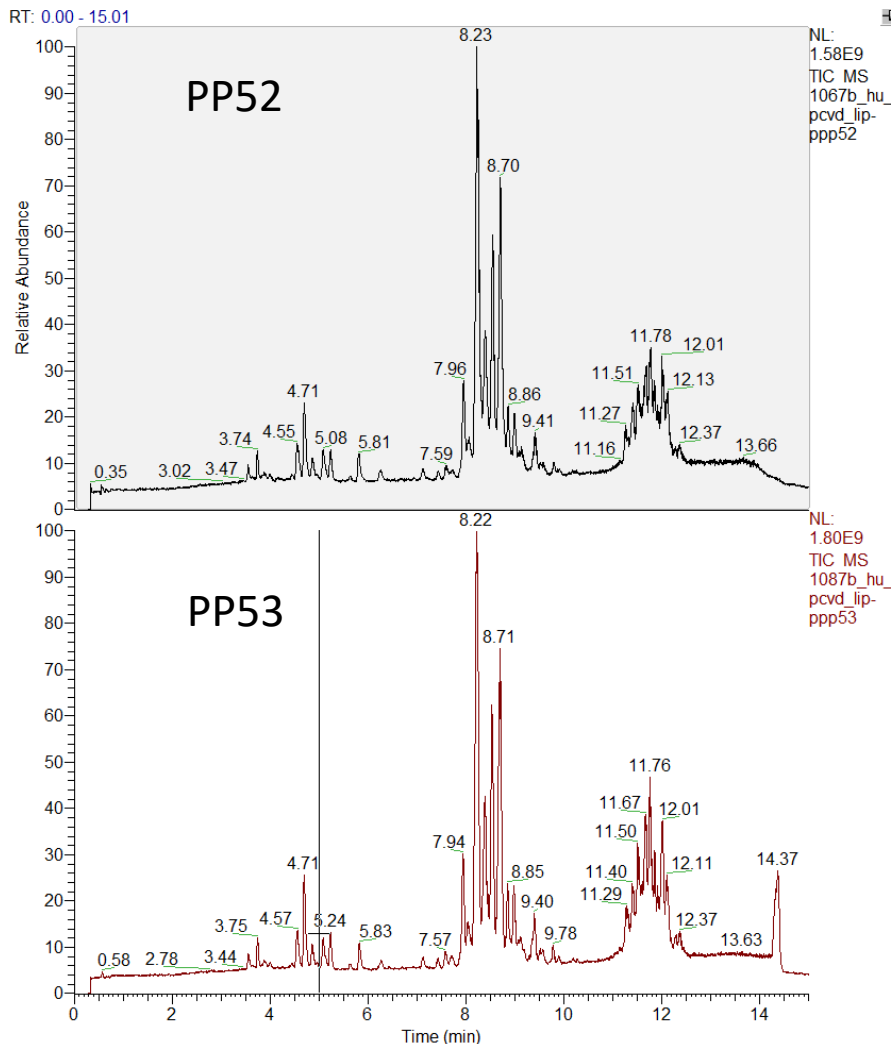
b. Extracted Ion Chromatograms

QI1002 m/z 311.1641; RT 4.40

Matches missingness pattern of MetProc

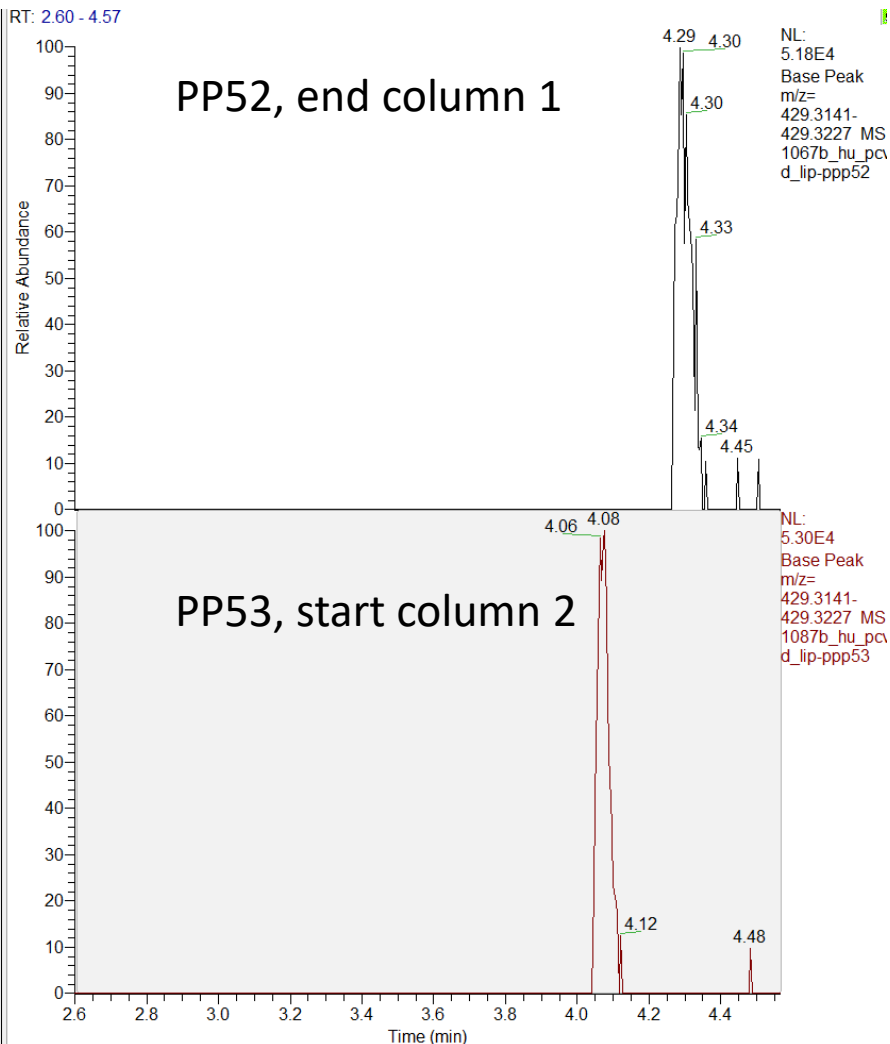


a. Total Ion Chromatograms



b. Extracted Ion Chromatograms

QI1869 m/z 429.31841; RT Col1 4.3, Col2 4.06
 QI did not align the RT correctly between columns

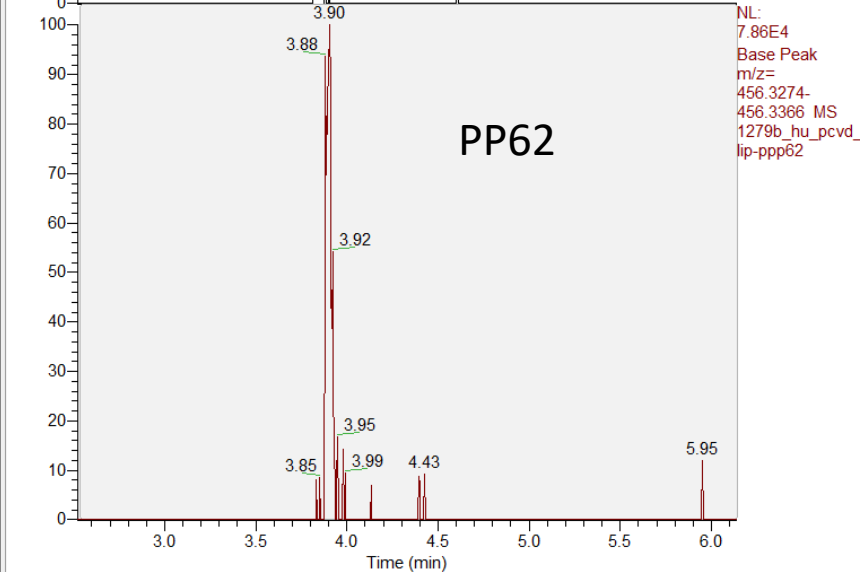
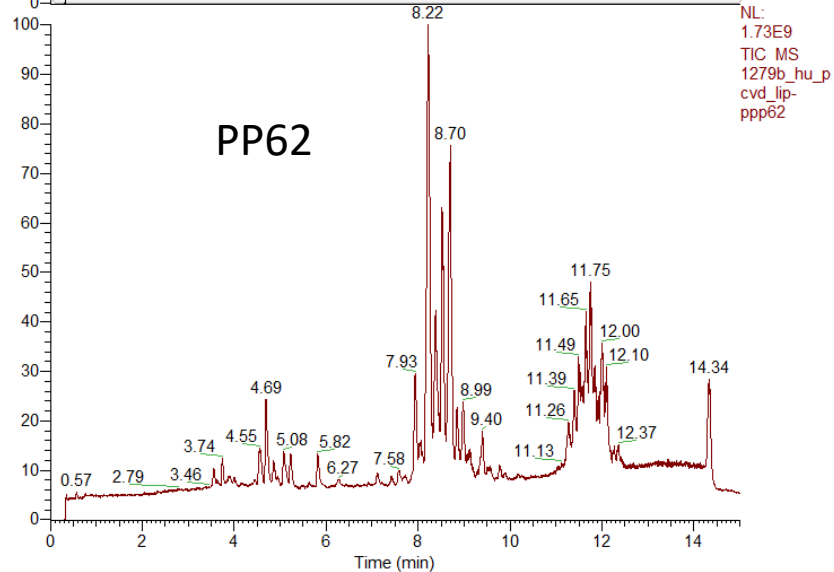
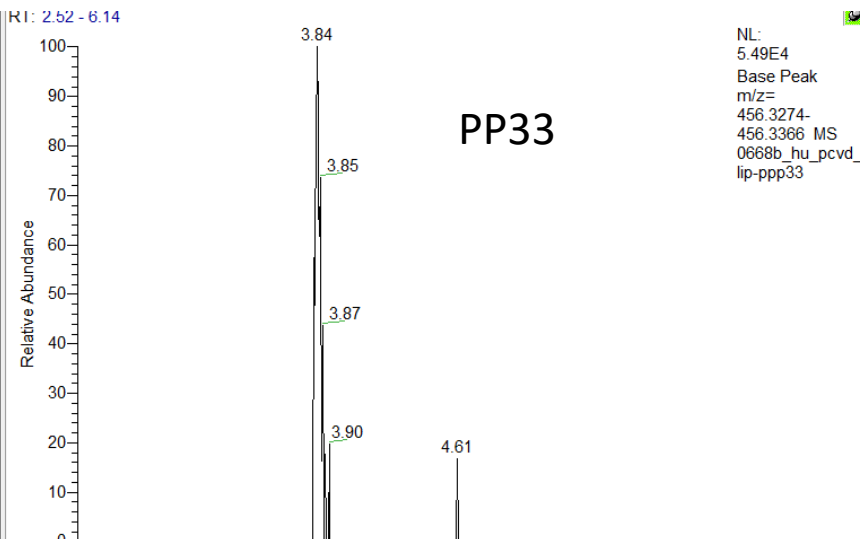
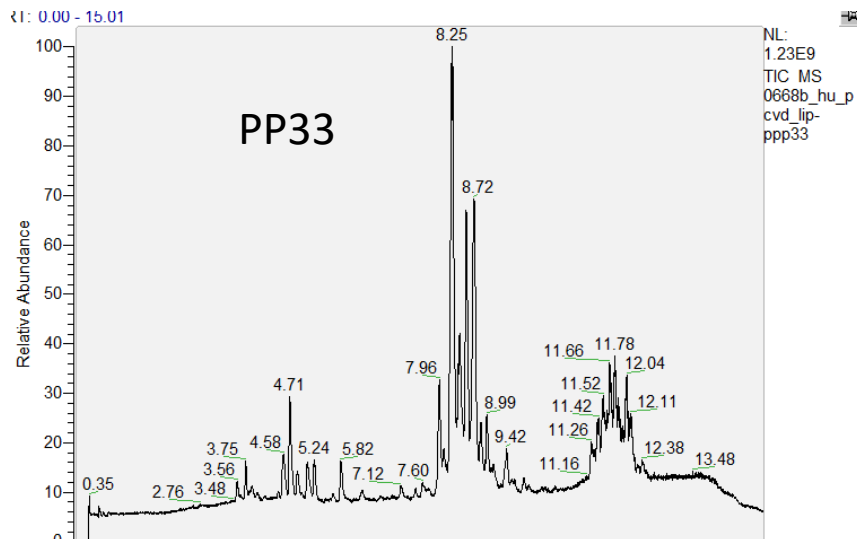


a. Total Ion Chromatograms

b. Extracted Ion Chromatograms

QI2116 m/z 456.3320; RT 3.85

QI did not detect the peak in Col 2, while Visual inspection shows there was a peak



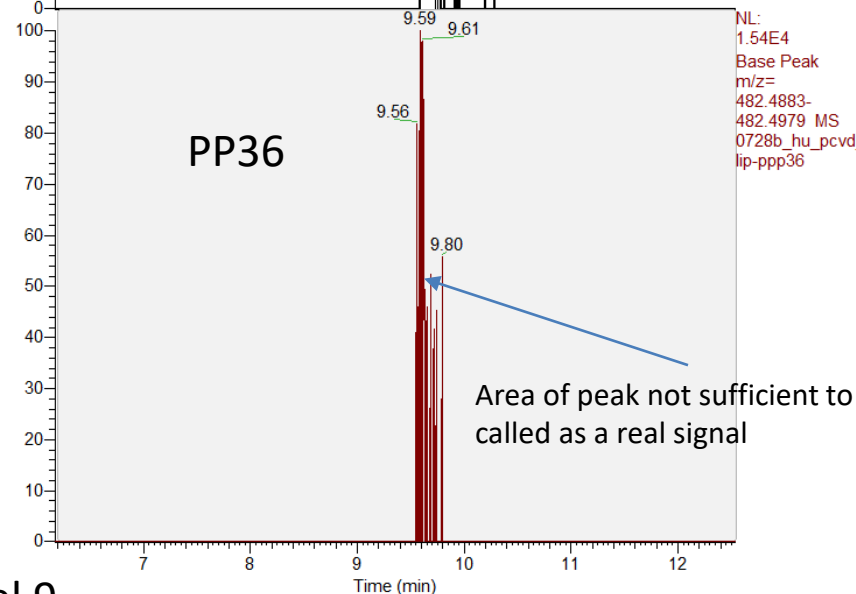
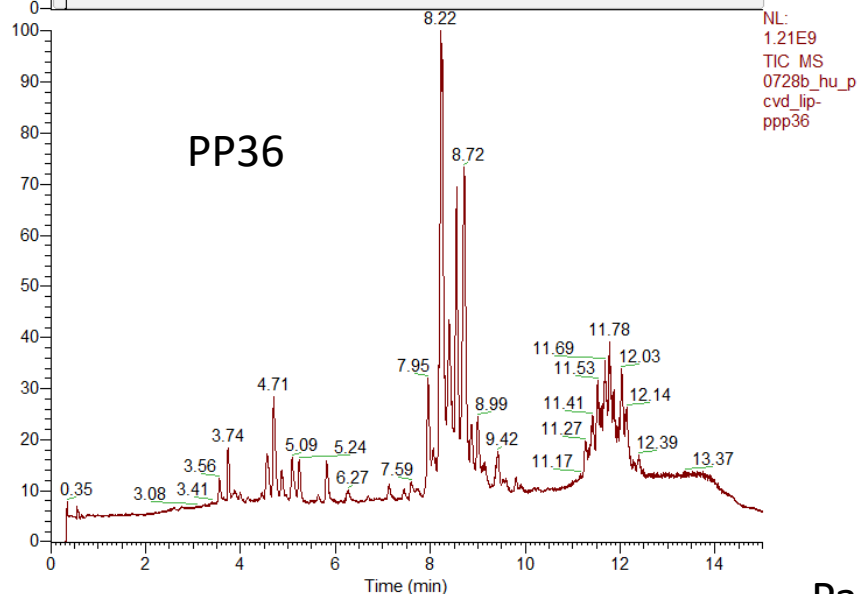
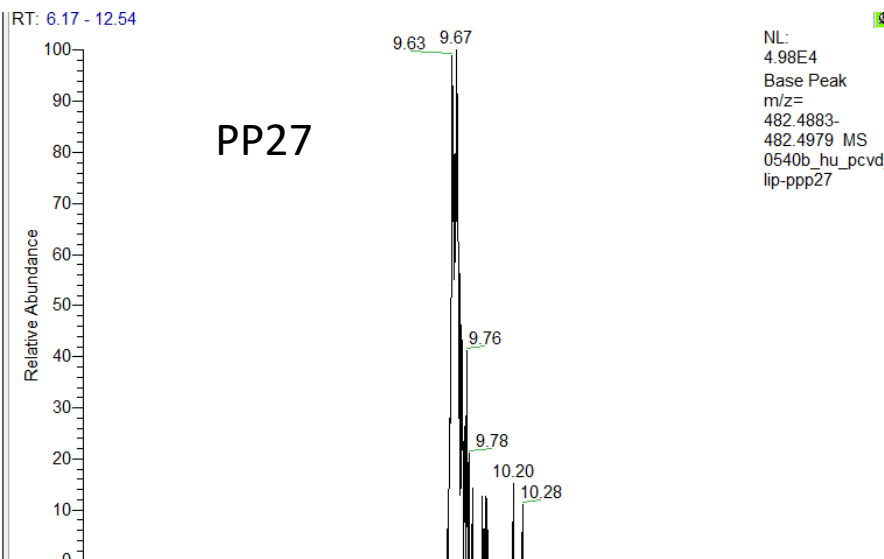
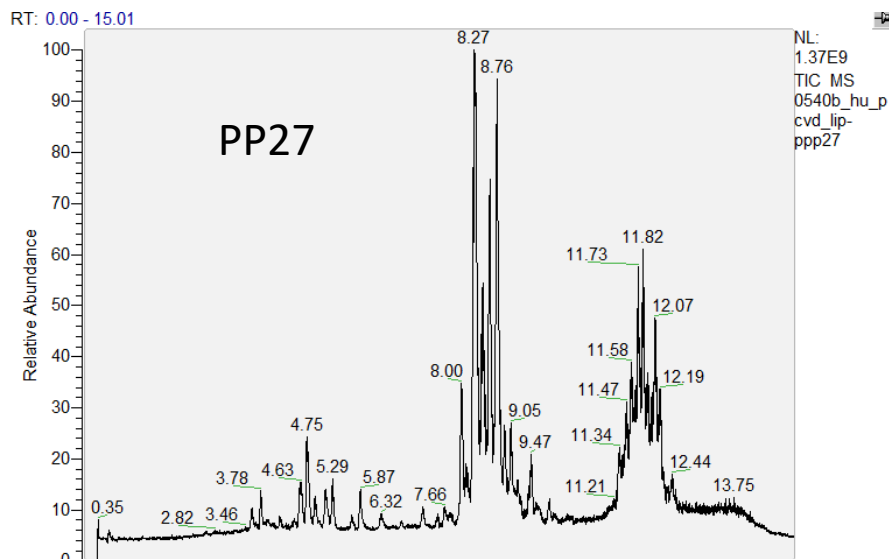
Panel 8

a. Total Ion Chromatograms

b. Extracted Ion Chromatograms

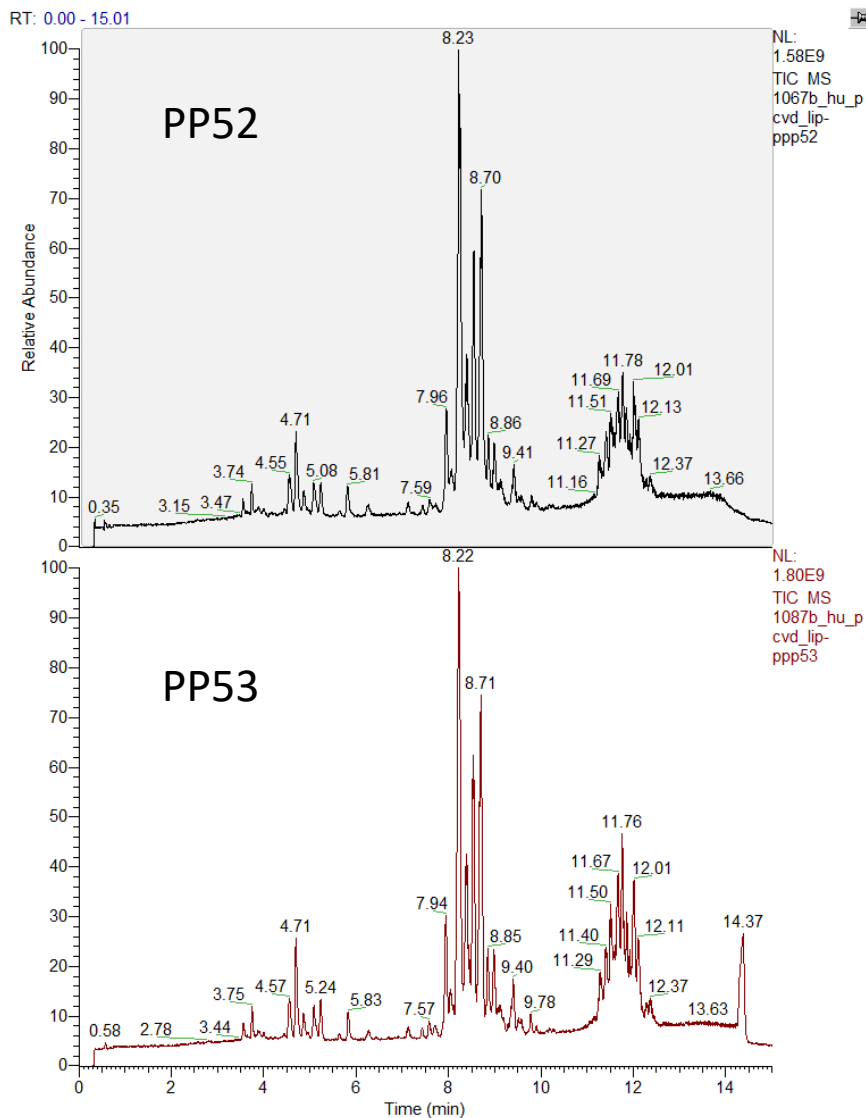
QI2331 m/z 482.4931 RT 9.65

Matches missingness pattern of MetProc



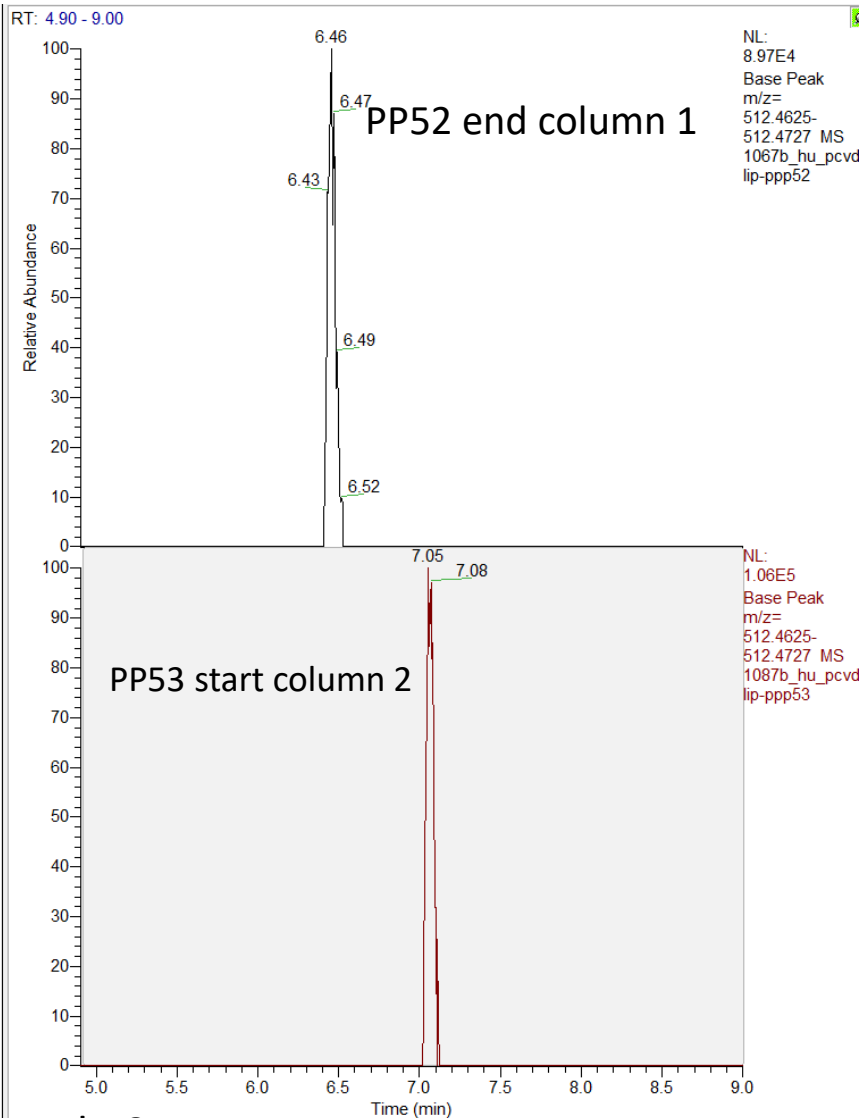
Panel 9

a. Total Ion Chromatograms



b. Extracted Ion Chromatograms

QI2502 m/z 512.4676 RT Col1: 6.45, Col2: 7.05
QI did not align the RT correctly between columns

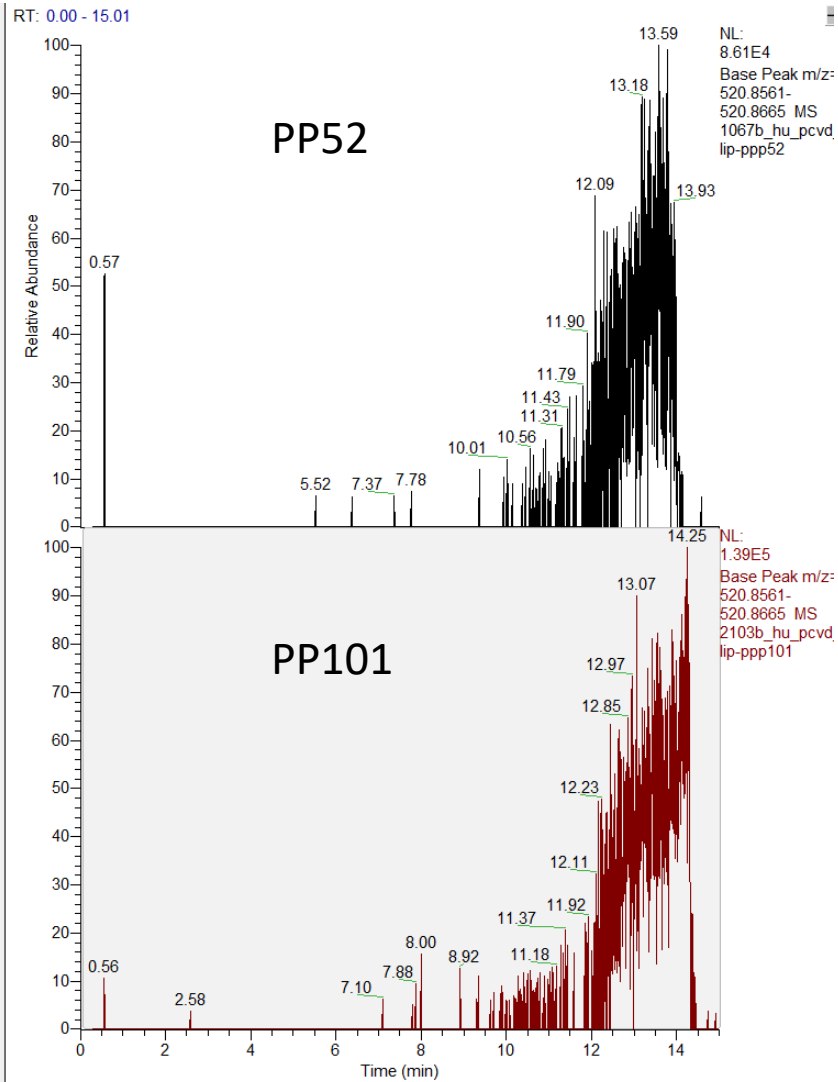
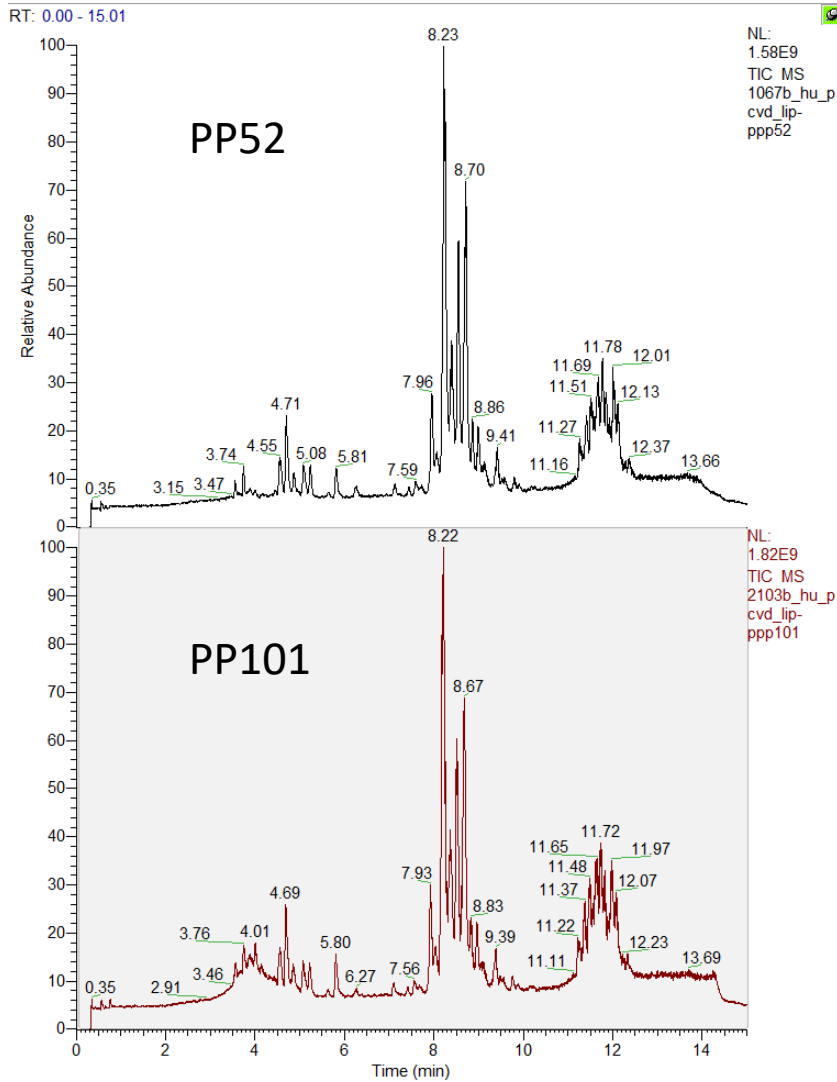


a. Total Ion Chromatograms

b. Extracted Ion Chromatograms

Q12543 m/z 520.8613 RT 13.74

Peaks detected in Q1 proved to be background noise



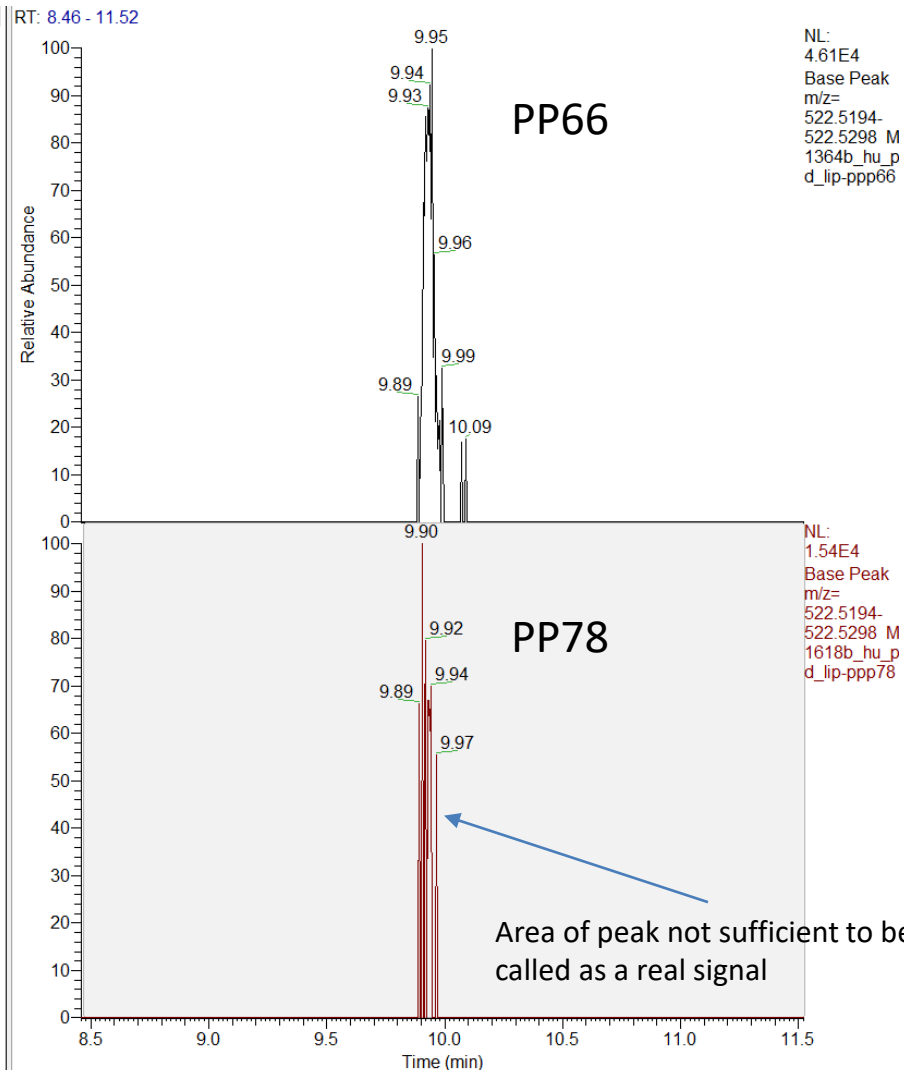
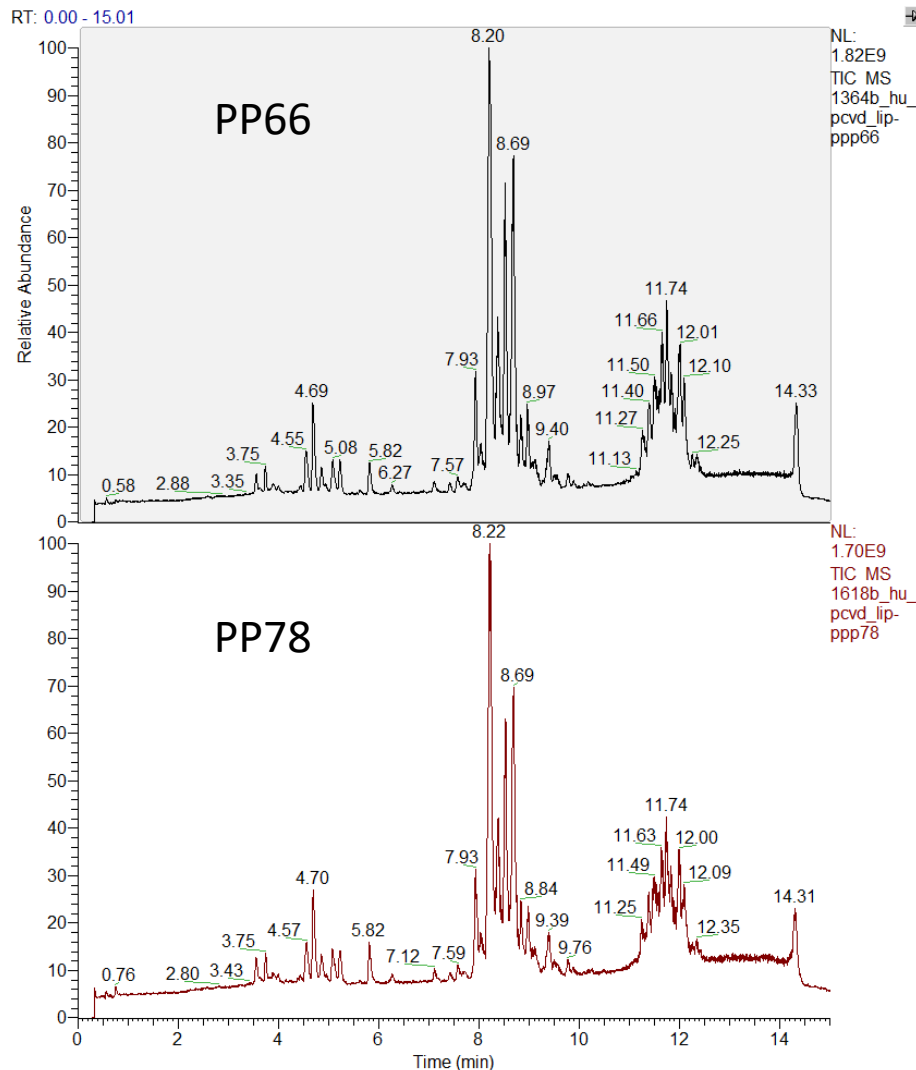
Panel 11

a. Total Ion Chromatograms

b. Extracted Ion Chromatograms

Q12552 m/z 522.5246; RT 10.0

Matches missingness pattern of MetProc

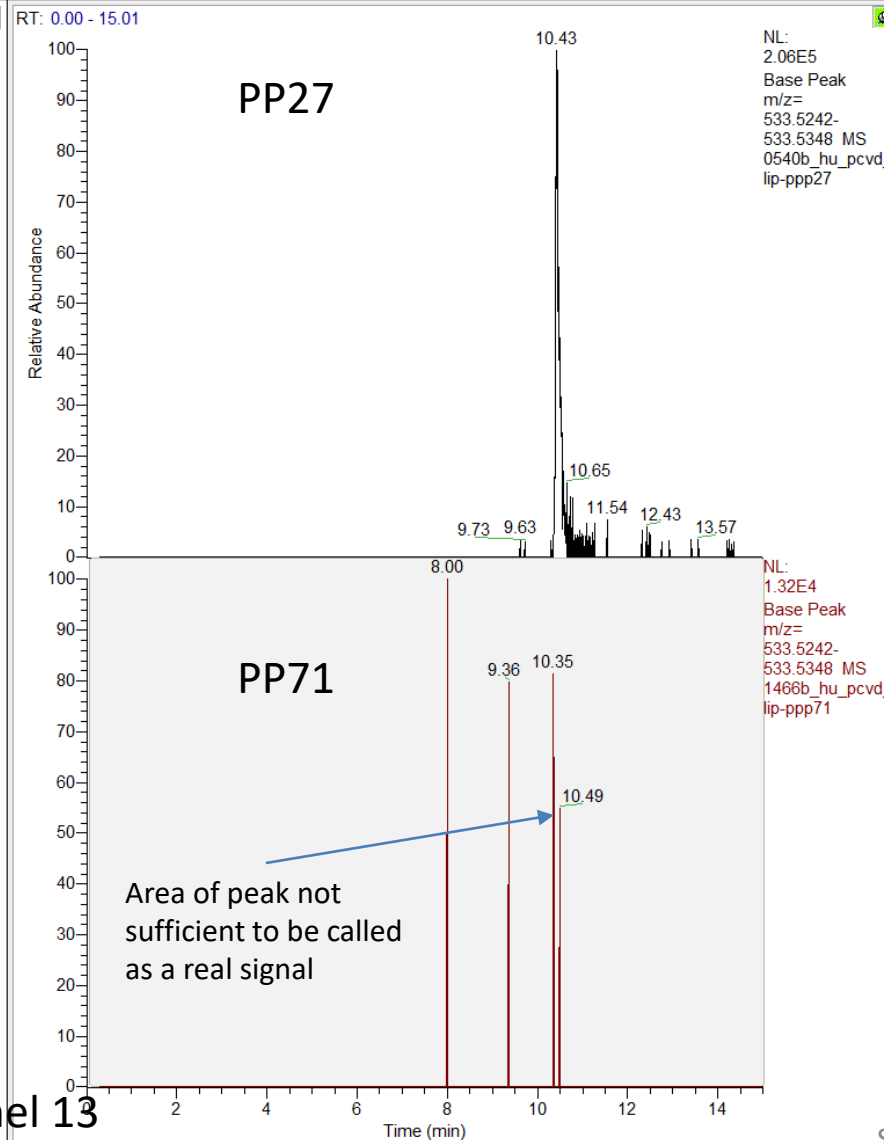
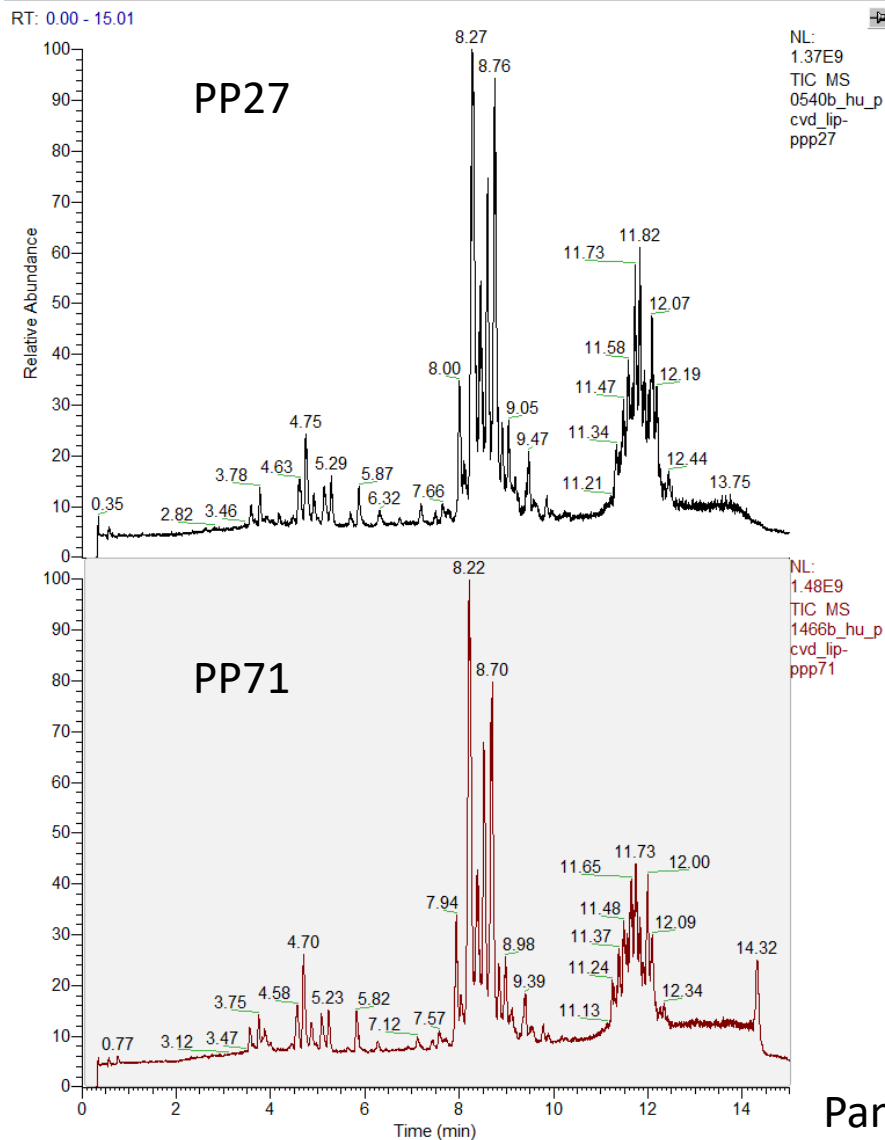


a. Total Ion Chromatograms

b. Extracted Ion Chromatograms

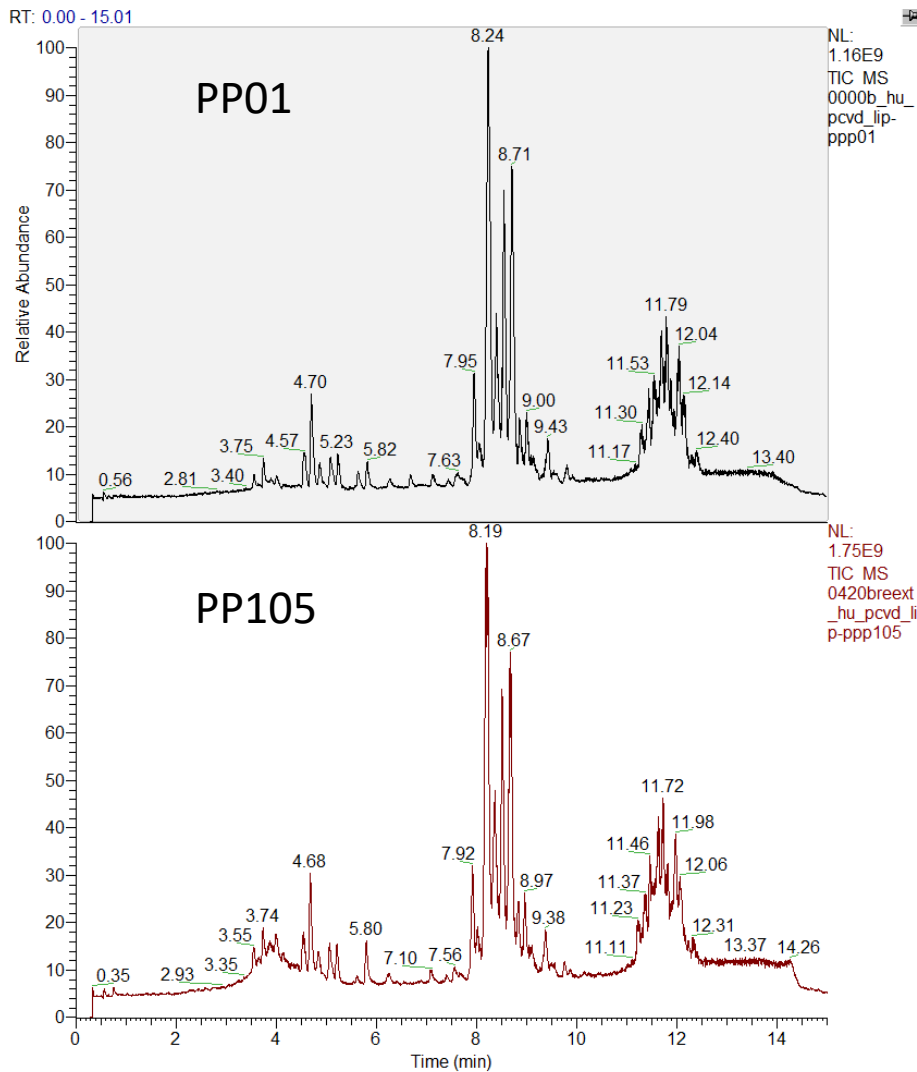
QI2620 m/z 533.5295; RT 10.43

Matches missingness pattern of MetProc



Panel 13

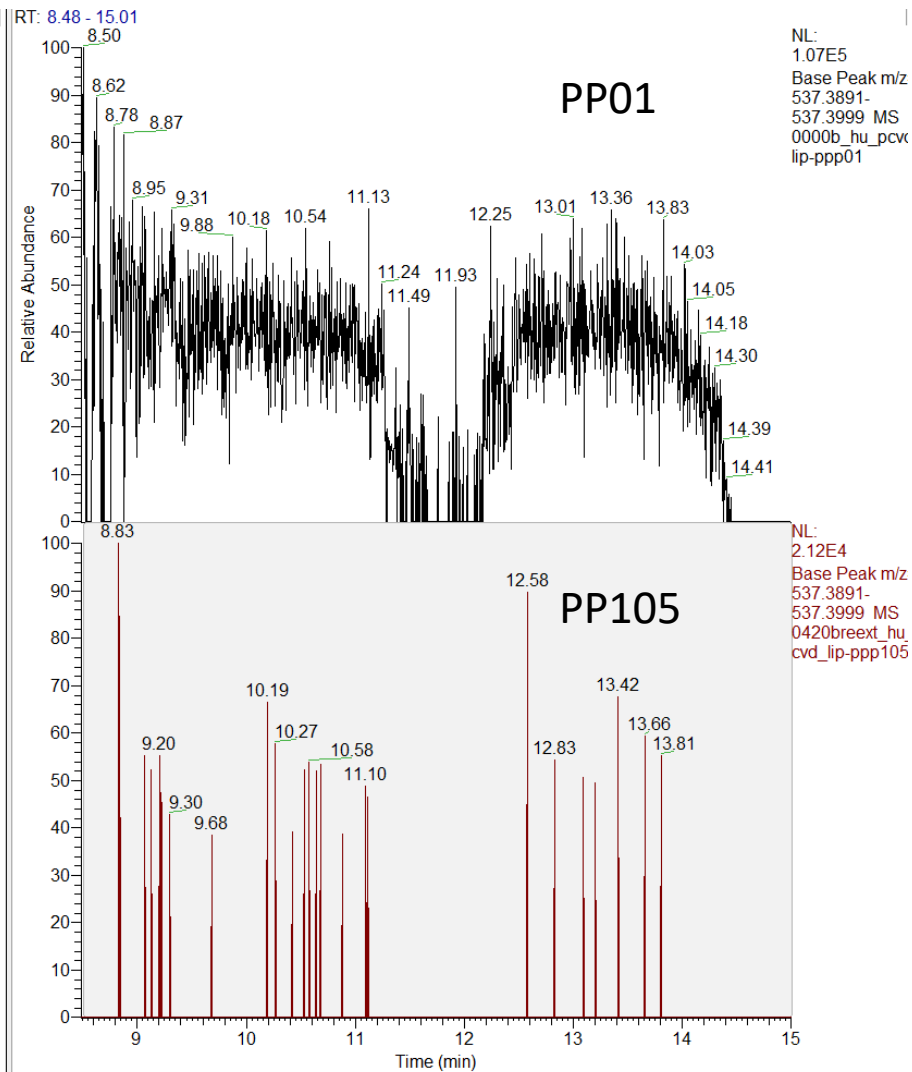
a. Total Ion Chromatograms



b. Extracted Ion Chromatograms

QI2675 m/z 537.3945; RT 12.9

Background noise across both columns

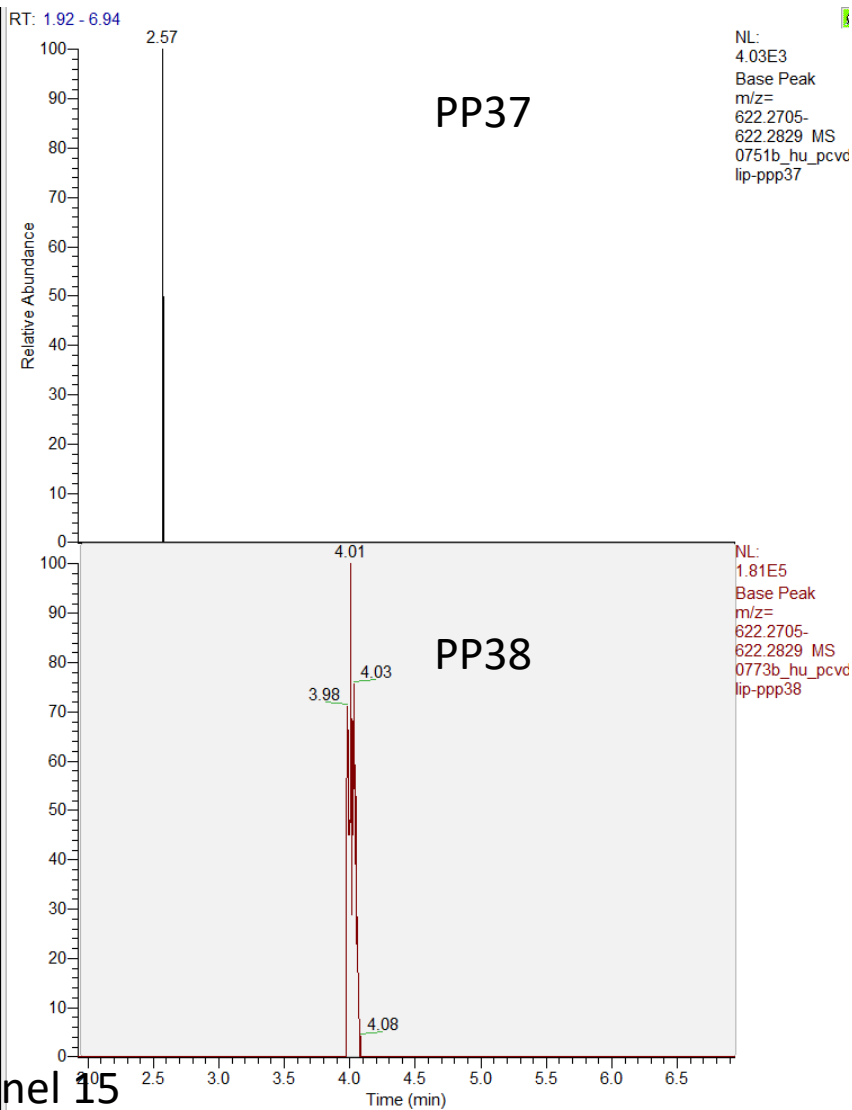
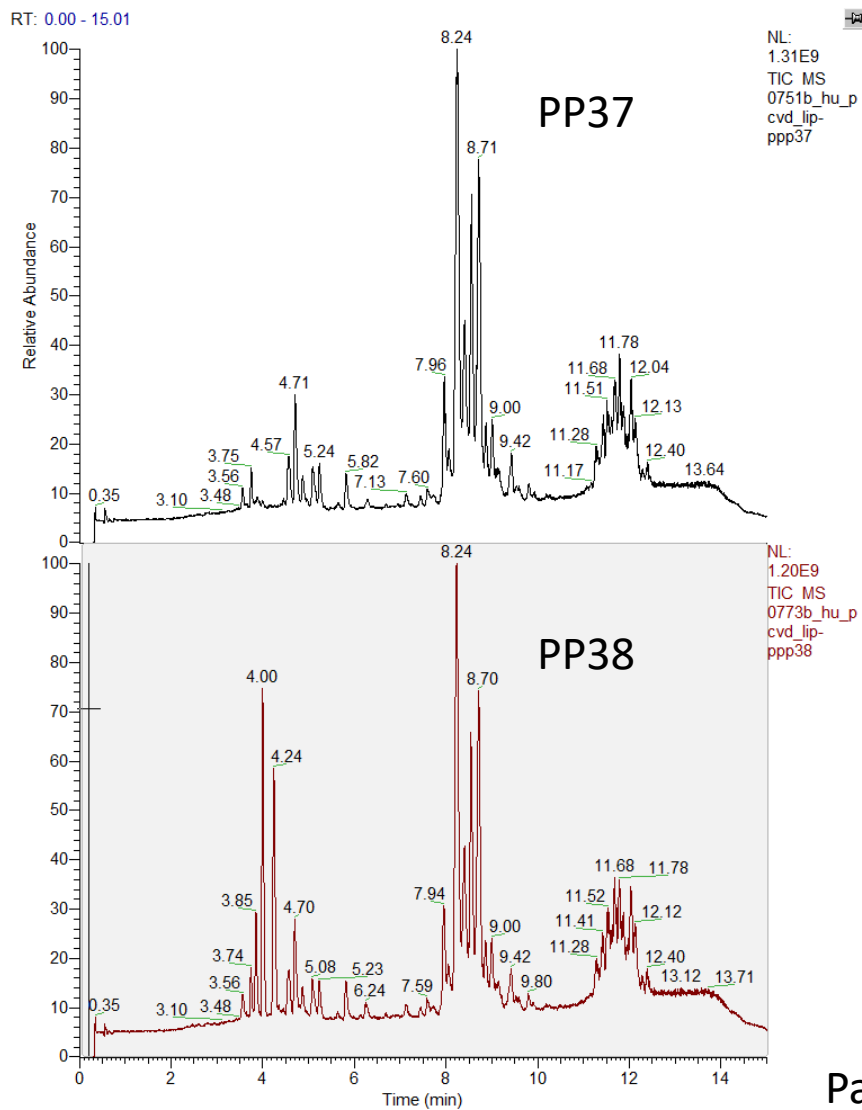


a. Total Ion Chromatograms

b. Extracted Ion Chromatograms

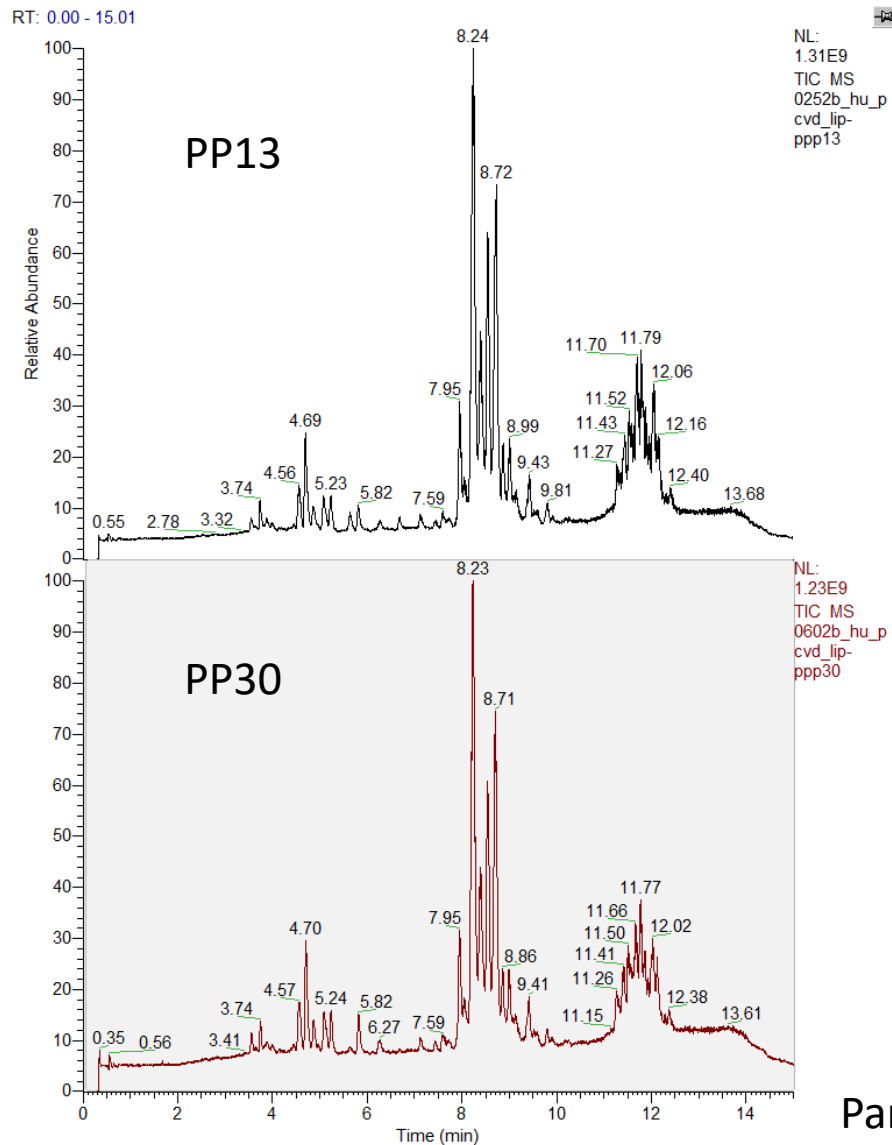
Q13239 m/z 622.2767; RT 4.03

Only present in pool 38



Panel 15

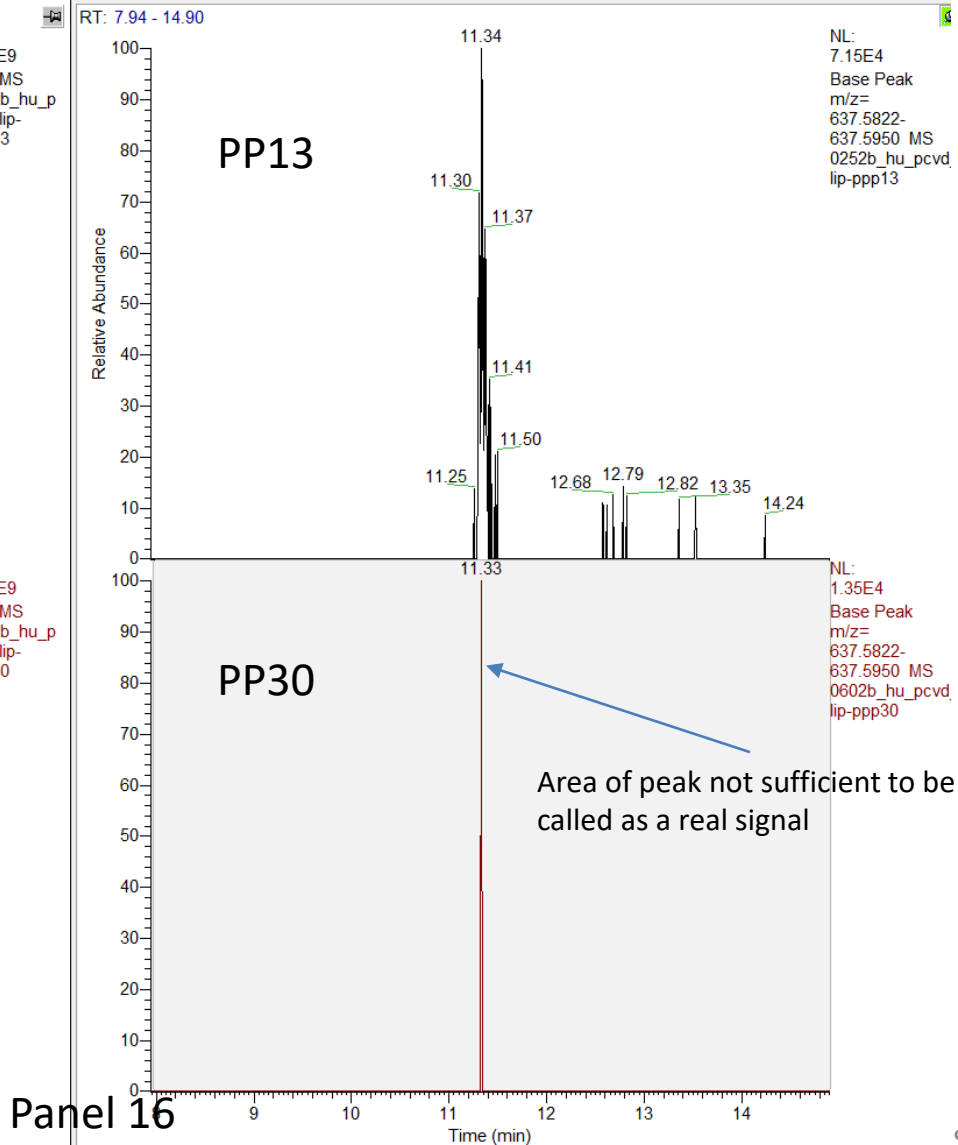
a. Total Ion Chromatograms



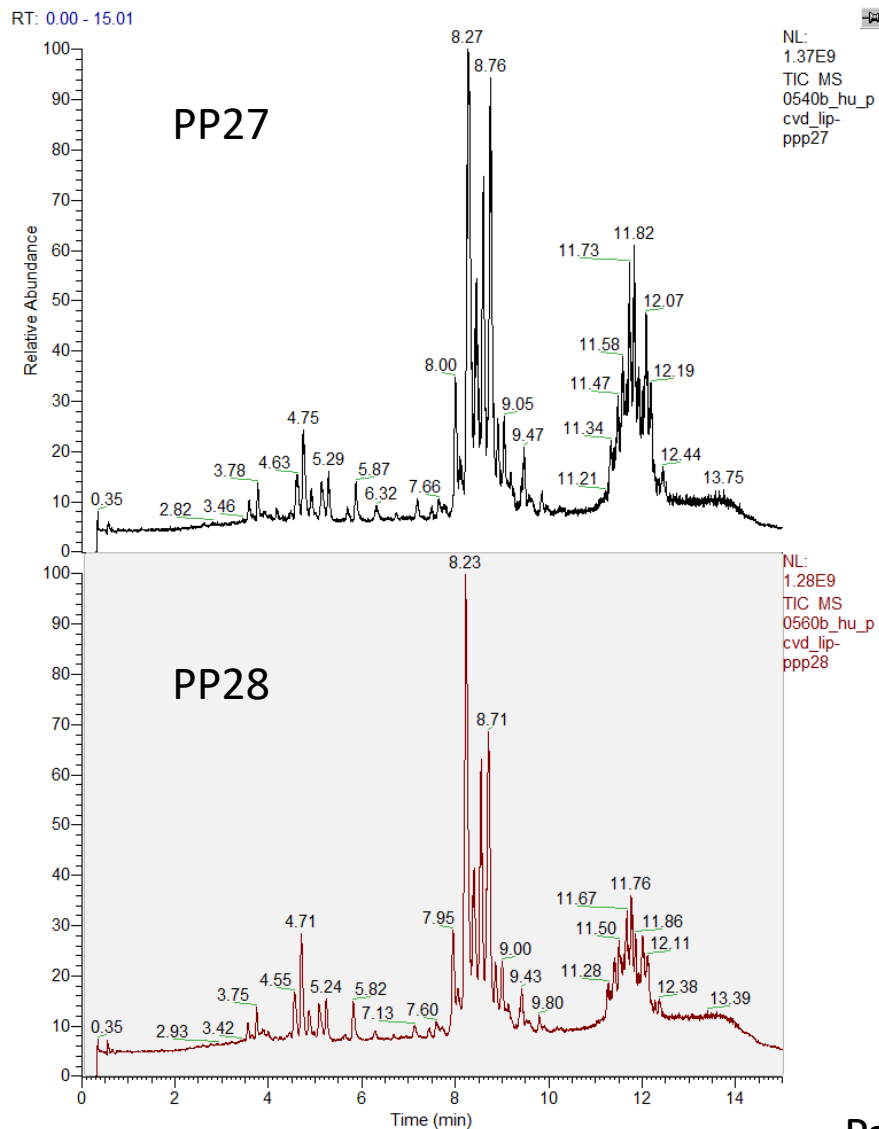
b. Extracted Ion Chromatograms

Q13339m/z 637.5886; RT 11.4

Matches missingness pattern of MetProc



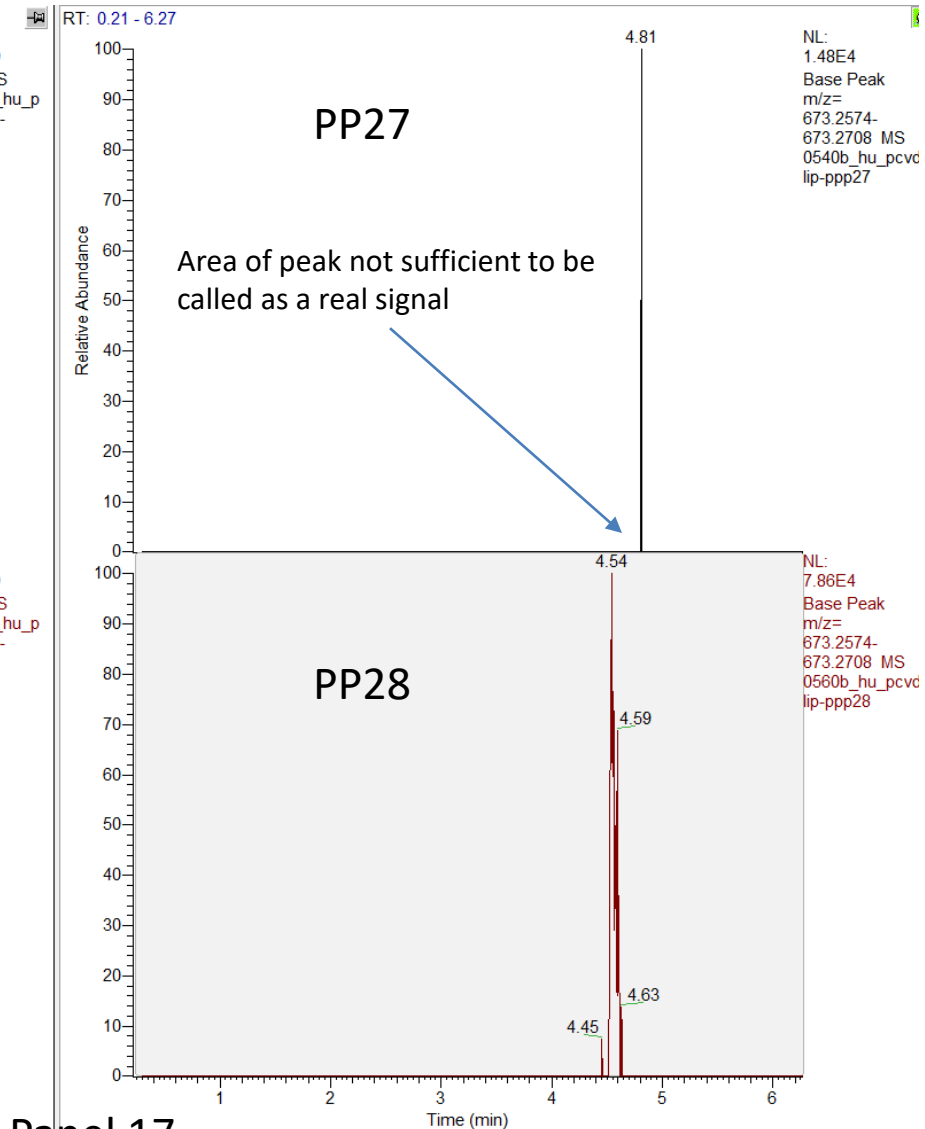
a. Total Ion Chromatograms



b. Extracted Ion Chromatograms

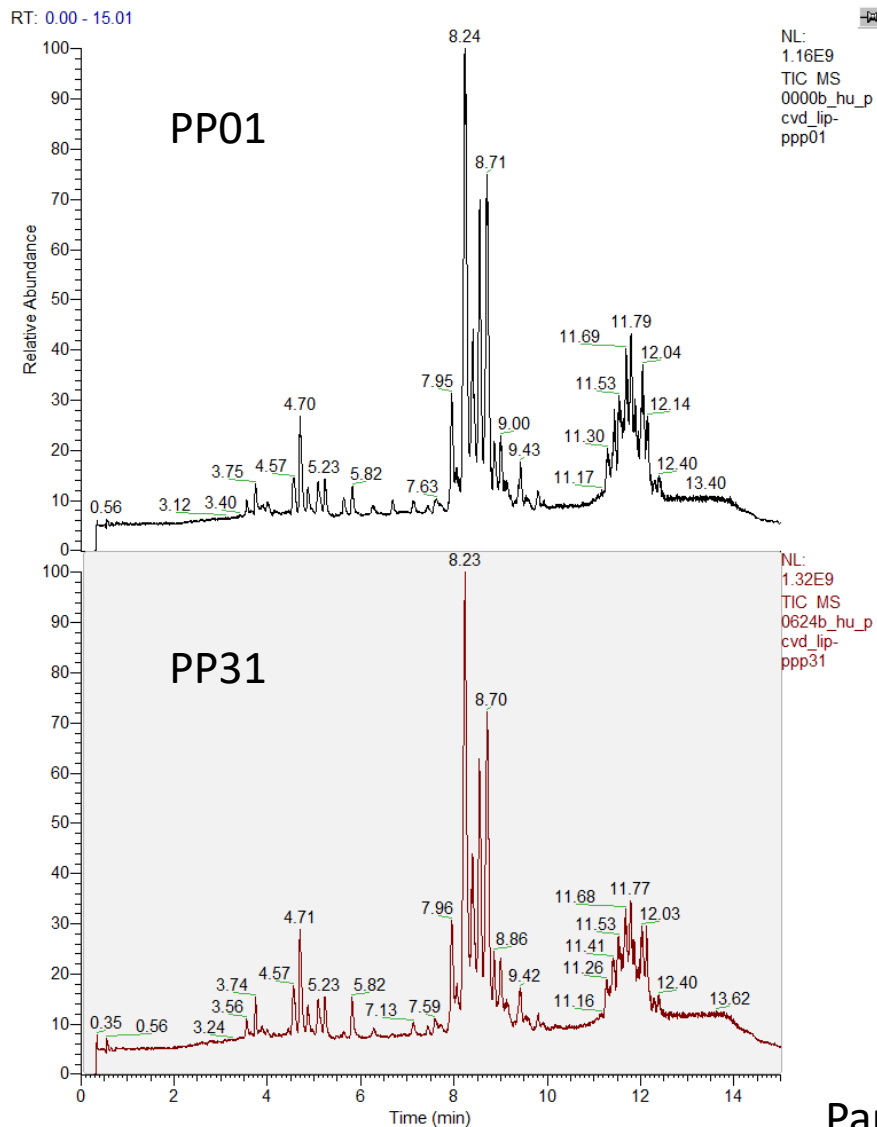
Q13579 m/z 673.2641; RT 4.58

Matches missingness pattern of MetProc



Panel 17

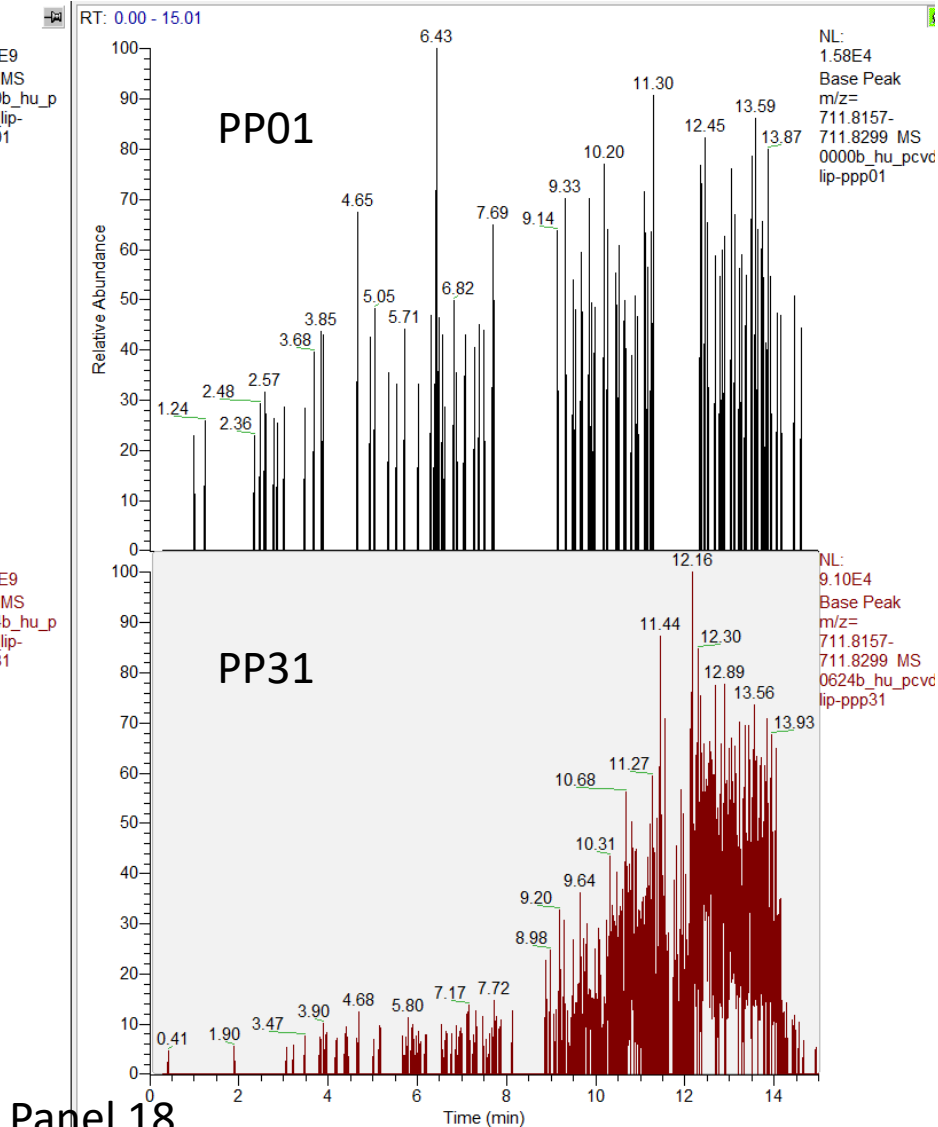
a. Total Ion Chromatograms



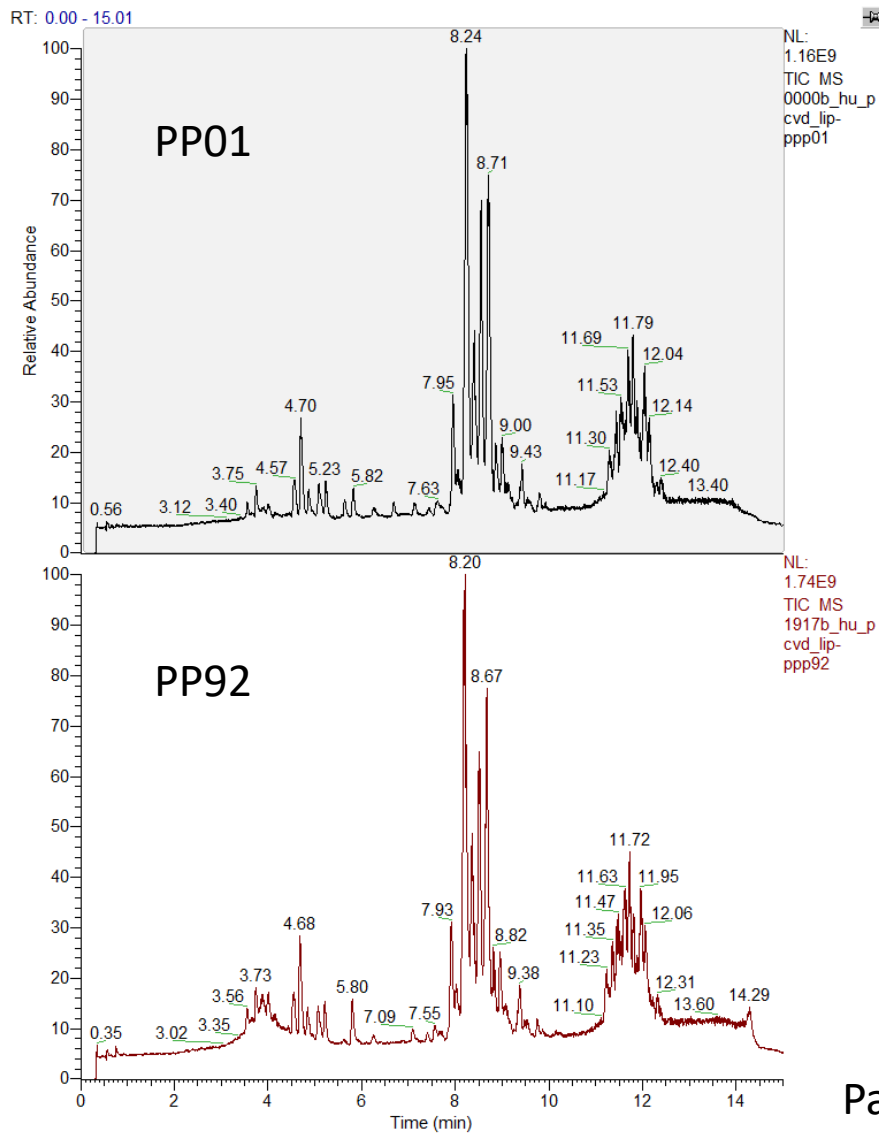
b. Extracted Ion Chromatograms

QI3827 m/z 711.8228; RT 13.03

Peaks detected in QI proved to be background noise



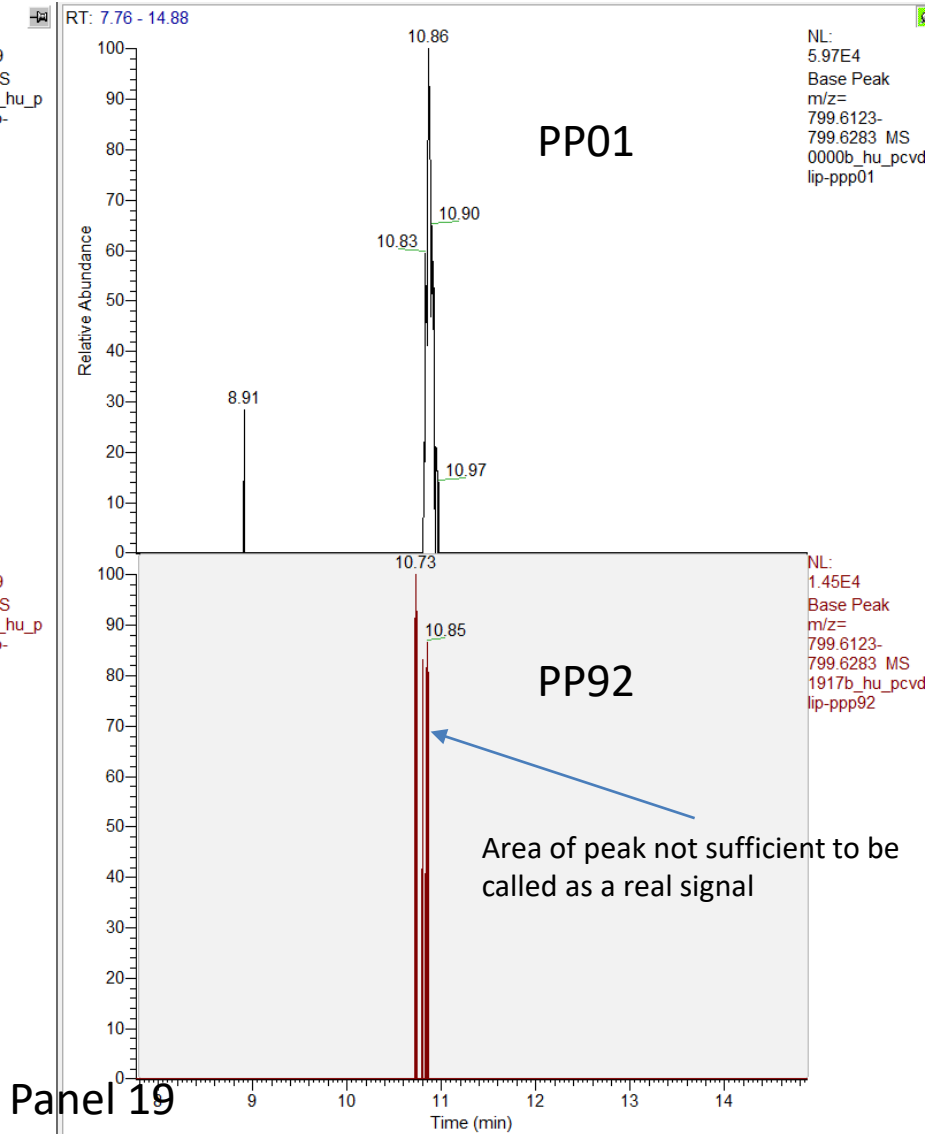
a. Total Ion Chromatograms



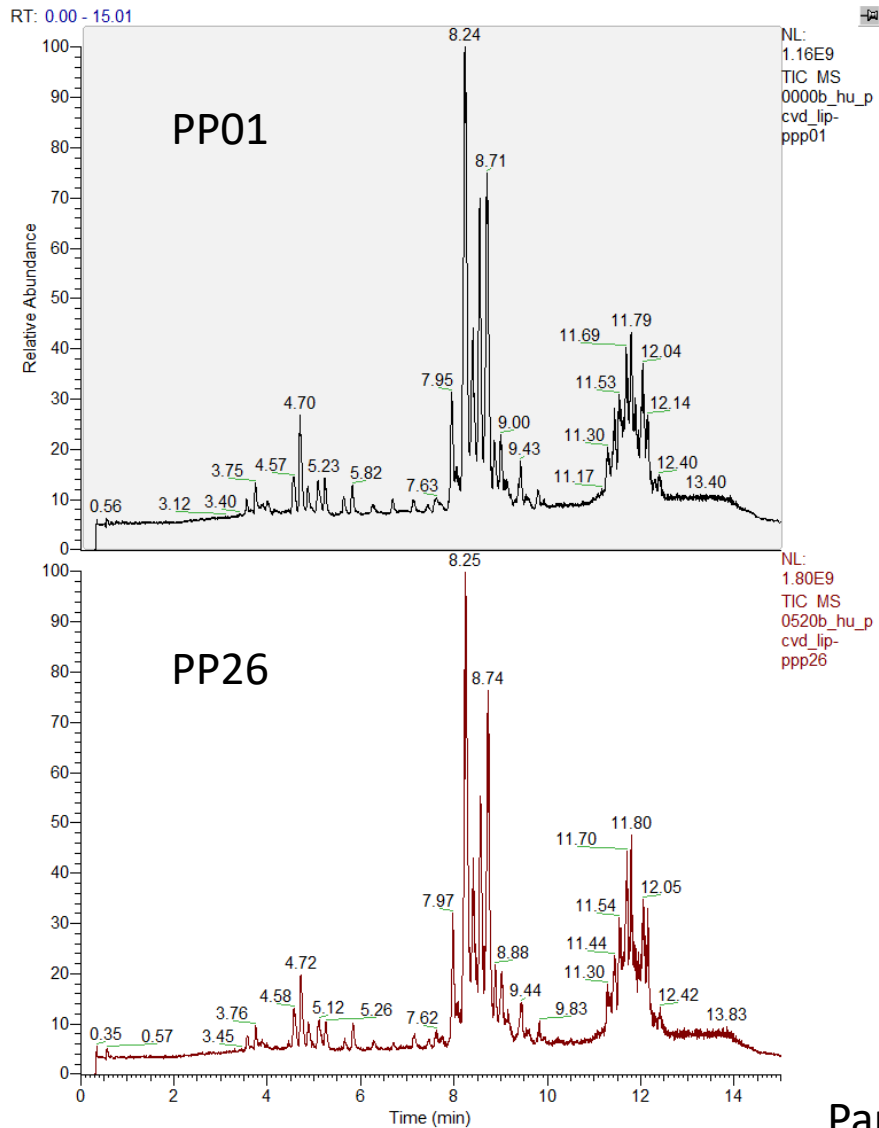
b. Extracted Ion Chromatograms

QI4513 m/z 799.6203; RT 10.9

Missingness pattern matches MetProc



a. Total Ion Chromatograms



b. Extracted Ion Chromatograms

QI6050 m/z 949.7796; RT 13.3

QI detected a peak on Pool 26, and not in Pool1 but both are background

