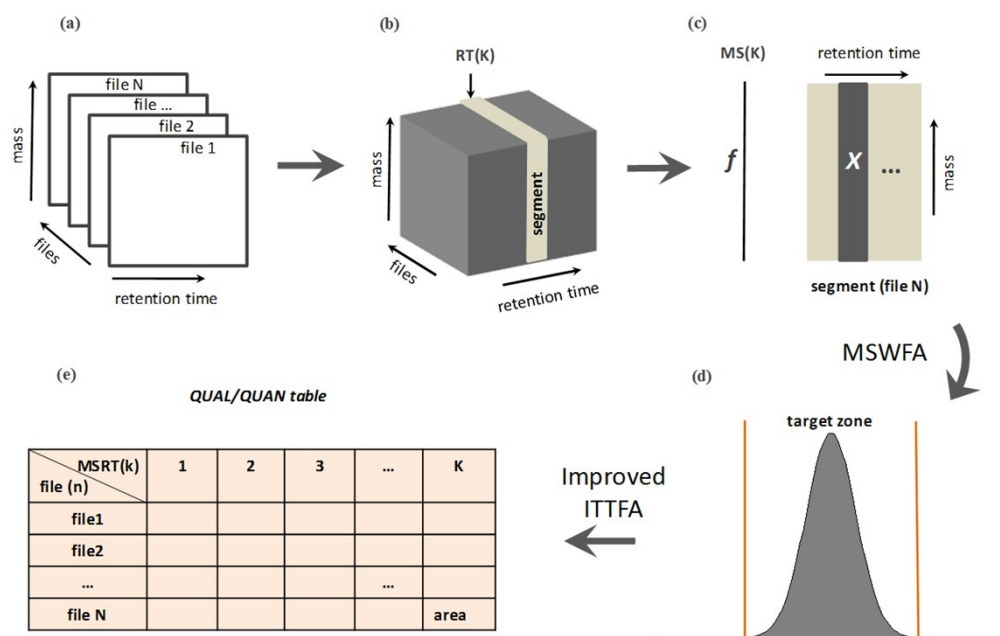# SUPPLEMENTARY MATERIALS

Fig. S-1 Illustration of data analysis workflow of MARS. RT and MS are from MSRT list. (a) a large amount of chromatography mass spectrums; (b) a three-way array of GC-MS data; (c) and (d) graphical representation of sub-window matrix and target zone. (e) QUAL&QUAN table.
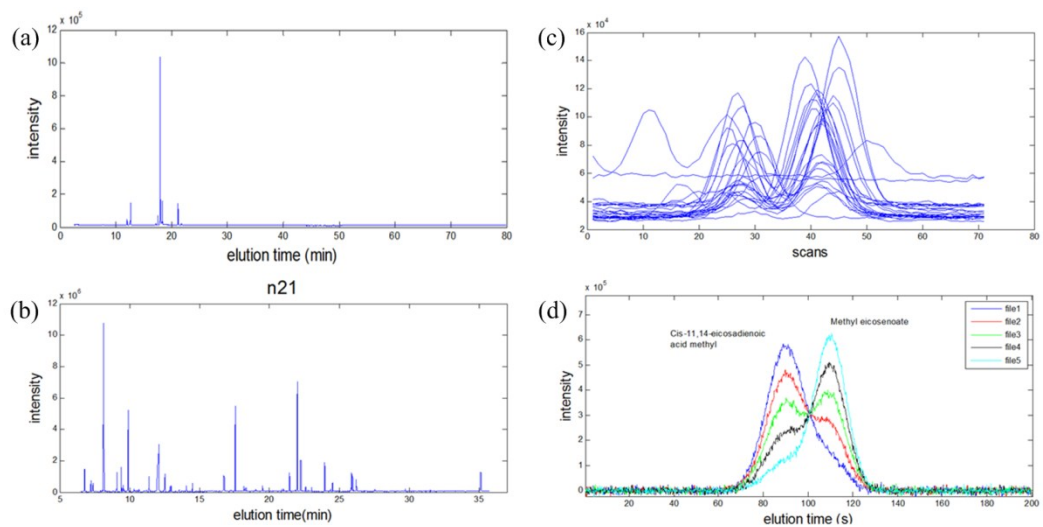
Fig. S-2 Total ion chromatogram of test datasets. (a) peppermint sample, (b) human blood plasma sample, (c) red wine samples, (d) five simulated profiles.
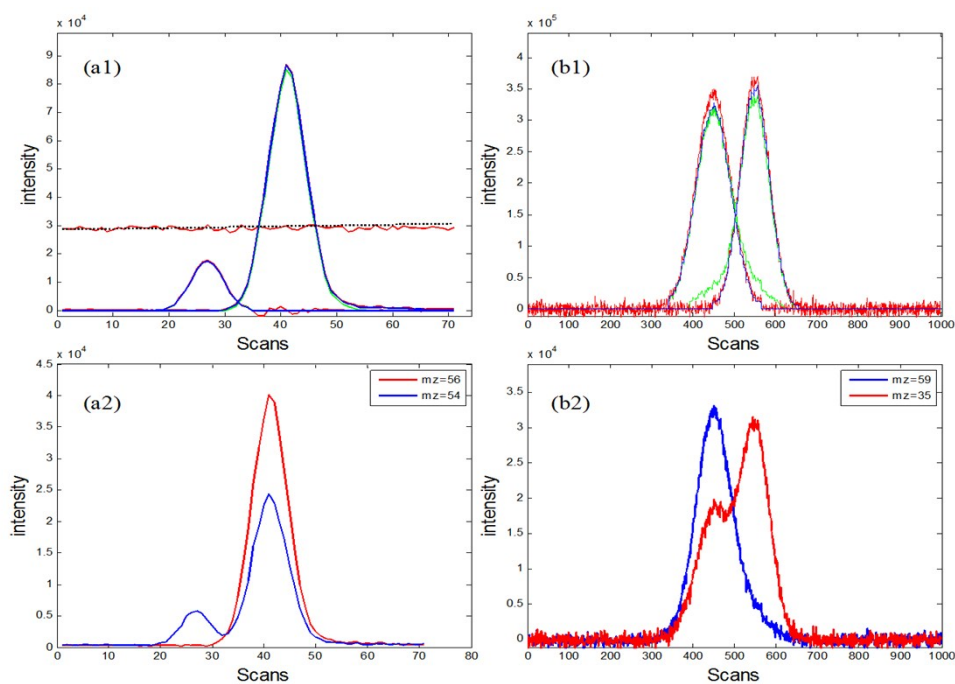
Fig. S-3 Examples of extracted chromatogram by MARS, MCR-ALS, PARAFAC2 and XCMS. Two files are selected from peak cluster A and B respectively for illustration. (a1) and (b1) are resolved chromatographic profiles (Blue: MARS; Green: MCR-ALS; Red: PARAFAC2; Black: baseline by 2D least-squares fitting). (a2) and (b2) are selected ion chromatograms by XCMS.
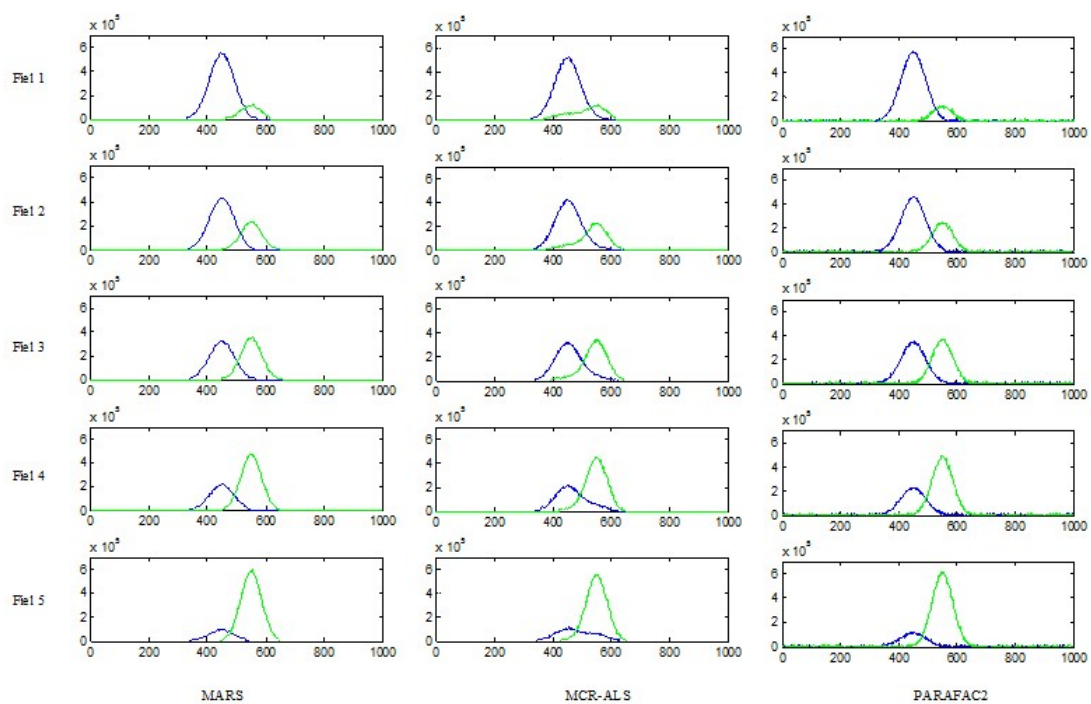
Fig S-4 Resolution of five simulated profiles by MARS MCR-ALS and PARAFAC2. Blue: Cis-11,14-eicosadienoic acid; green: Methyl eicosenoate.
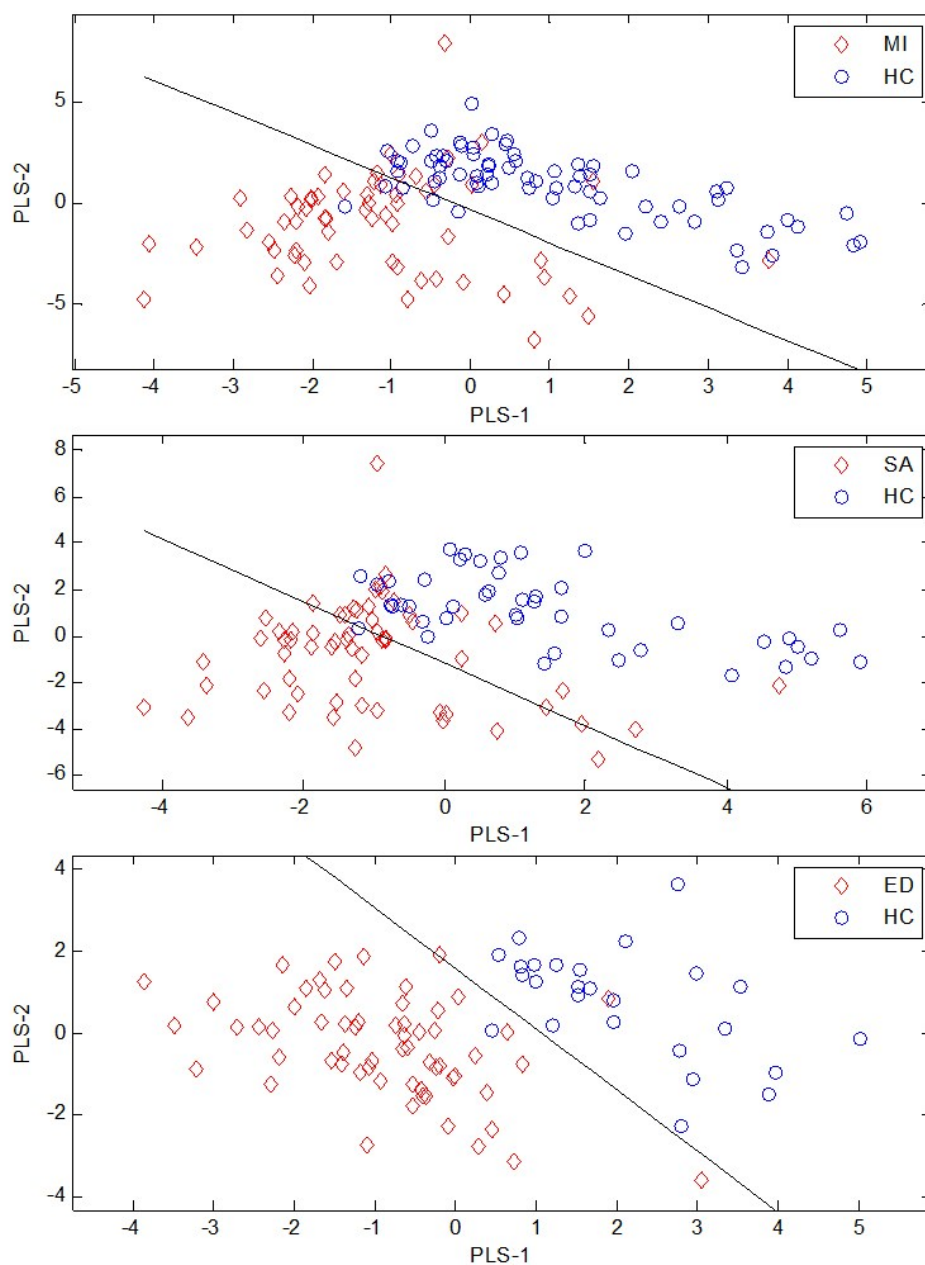
Fig S-5 .

Fig. S-5 PLS-DA score plots of ED, SA, HC and MI. Latent variables used were 2 (MI vs

HC), 5 (SA vs HC), 2 (ED vs HC).

Table S-1 Qualitative and quantitative information of Dataset II, III, IV

QUALQUAN_table_datasetII.zip

QUALQUAN_table_datasetIII.zip

QUALQUAN_table_datasetIV.zip

Table S-2 Five profiles are simulated with different peak height, peak deviation and mass spectral. Firstly, five chromatograms were generated with two Gaussian peaks at different concentration. Then, random noise was added into simulated chromatograms. Next, two mass spectrums of two real standard compounds were obtained from NIST library. Finally, five 2D matrixes were obtained from the outer product of simulated chromatograms and pure mass spectrums.

peaks	file1 (cr)	file2 (cr)	file3 (cr)	file4 (cr)	file5 (cr)	peak position(s)	standard deviation	compounds mass spectral
peak1	1.00	0.80	0.60	0.40	0.20	90	60	Cis-11,14-eicosadienoic acid methyl
peak2	0.20	0.40	0.60	0.80	1.00	110	50	Methyl eicosenoate

cr: concentration ratio.

Table S-3 Comprehensive comparison between MARS, XCMS, MCR-ALS and PARAFAC2 for processing GC-MS dataset.

Methods	Baseline correction	Alignment	Performance on overlapped peaks	Segmentation	Parameters	Speed
XCMS	need	need	worse	not need	parameters of baseline correction, peak detection and alignment[a]	fast, unsupervised, Automatic, batch processing
MCR-ALS	need	not need	good	manual operation	number of component factor	slow, processing part of dataset once
PARAFAC2	not need	not need	good	manual operation	number of component factor	slow, processing part of dataset once
MARS	not need	not need	good	automatic, a method based on MWSFA is applied	moving window size, threshold, optimization	fast, unsupervised, batch processing (with known MSRT)

a: different parameters depends on different methods of baseline correction, peak detection and alignment.