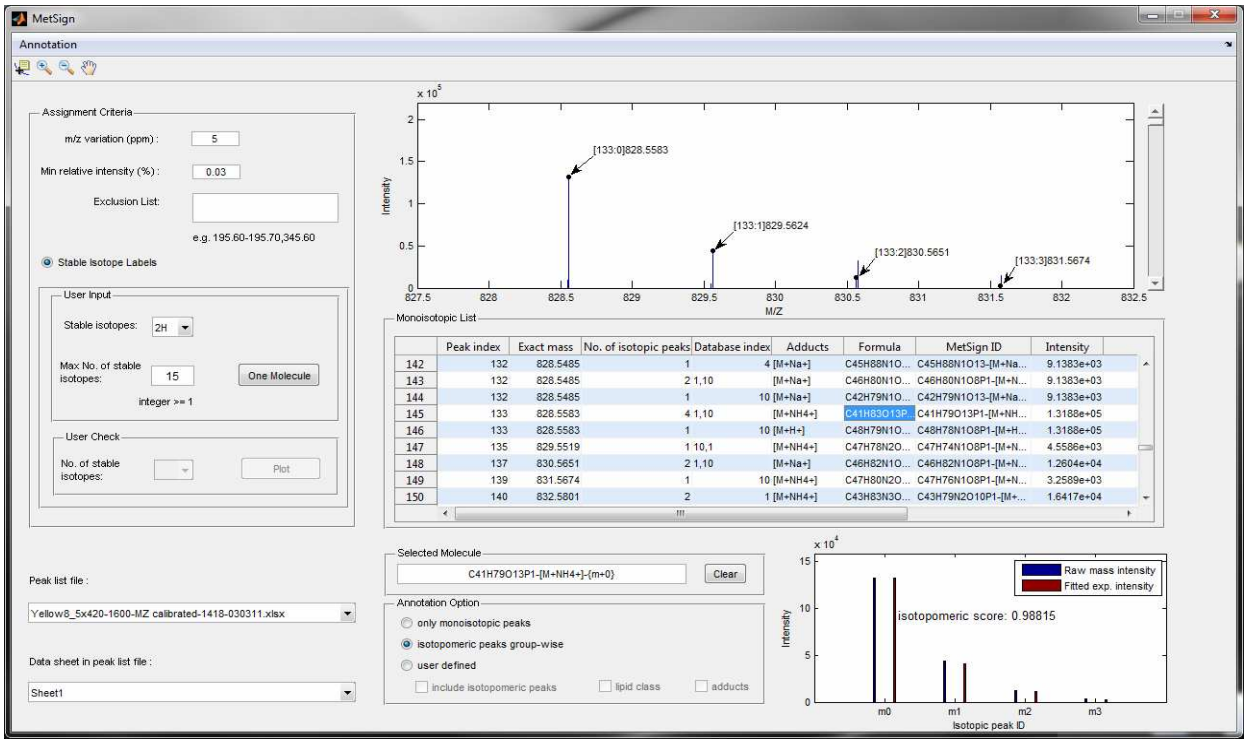
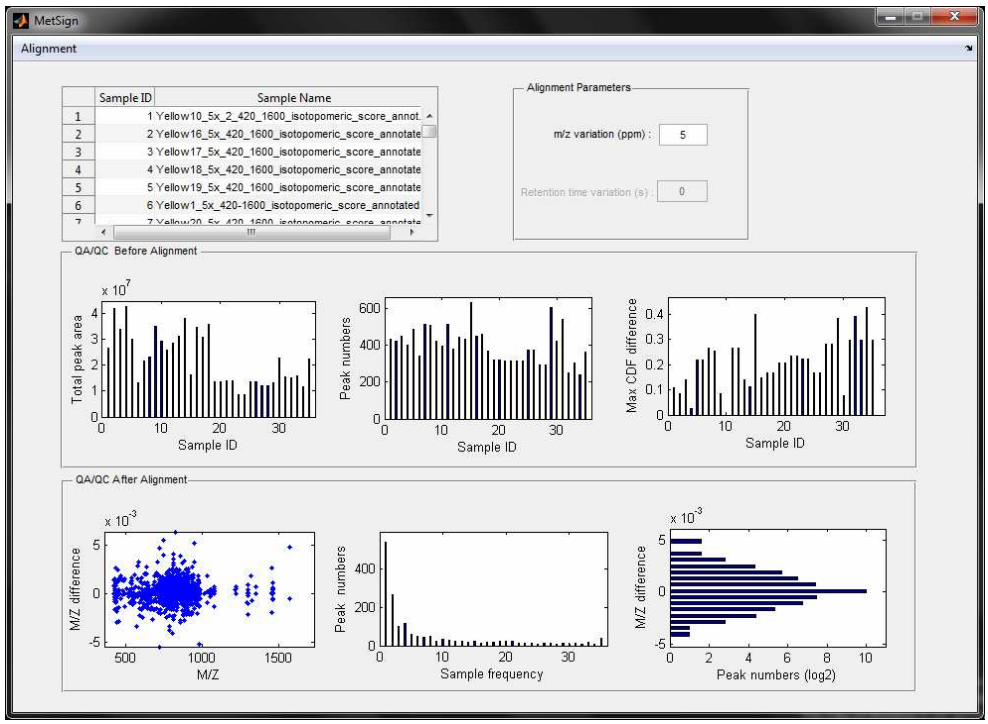


Figure S-1. A screenshot of the metabolite initial assignment window of *MetSign* software. The initial assignment criteria include *m/z* variation (ppm), minimum relative intensity, exclusion list, and stable isotope labels defined by the user. The upper right plot shows the profile of experimentally measured isotopic peaks of metabolite ($C_{41}H_{79}O_{13}P_1$-[M+NH_4^+]-{m+0}). The lower right plot shows intensity analysis result after peak profile deconvolution. The blue line represents the original peak intensity. The red line represents fitted intensity optimized by minimizing root mean square error. The similarity (Pearson's correlation coefficient) between deconvoluted isotopic peaks and theoretical isotopic peaks is 0.98815. A large similarity value, *i.e.*, close to 1, indicates a high probability that this metabolite is present in experimental data.

(A)



(B)

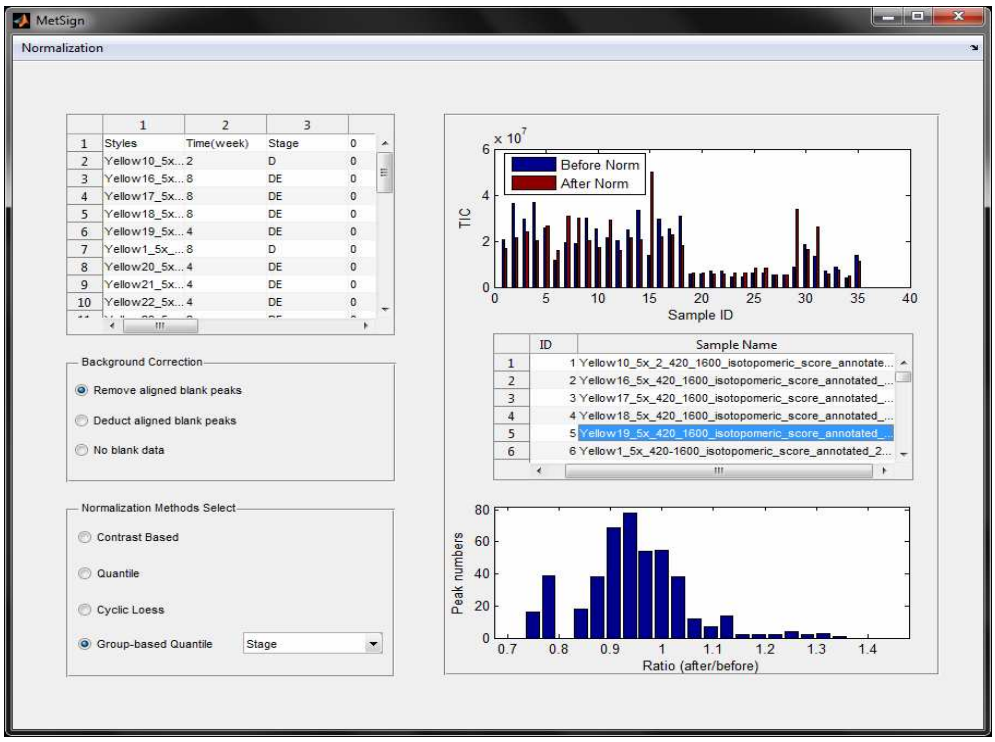


Figure S-2. Example QA/QC plots of peak list alignment and normalization in *MetSign*. (A) is a screenshot of the results of peak list alignment. The upper three plots are results of QA/QC before alignment, while the bottom three plots display the QA/QC results of the alignment. (B) is a screenshot of normalization. The upper right plot shows the distribution of total ion current of all samples before and after normalization. By interactive slection of a sample in the data sheet on the right panel, the lower right plot can display on-the-fly the histogram of normalization factor of all metabolites in the selected sample. The normalization factor is calculated as the peak area after normalization divided by the peak area of the same metabolite after normalization.

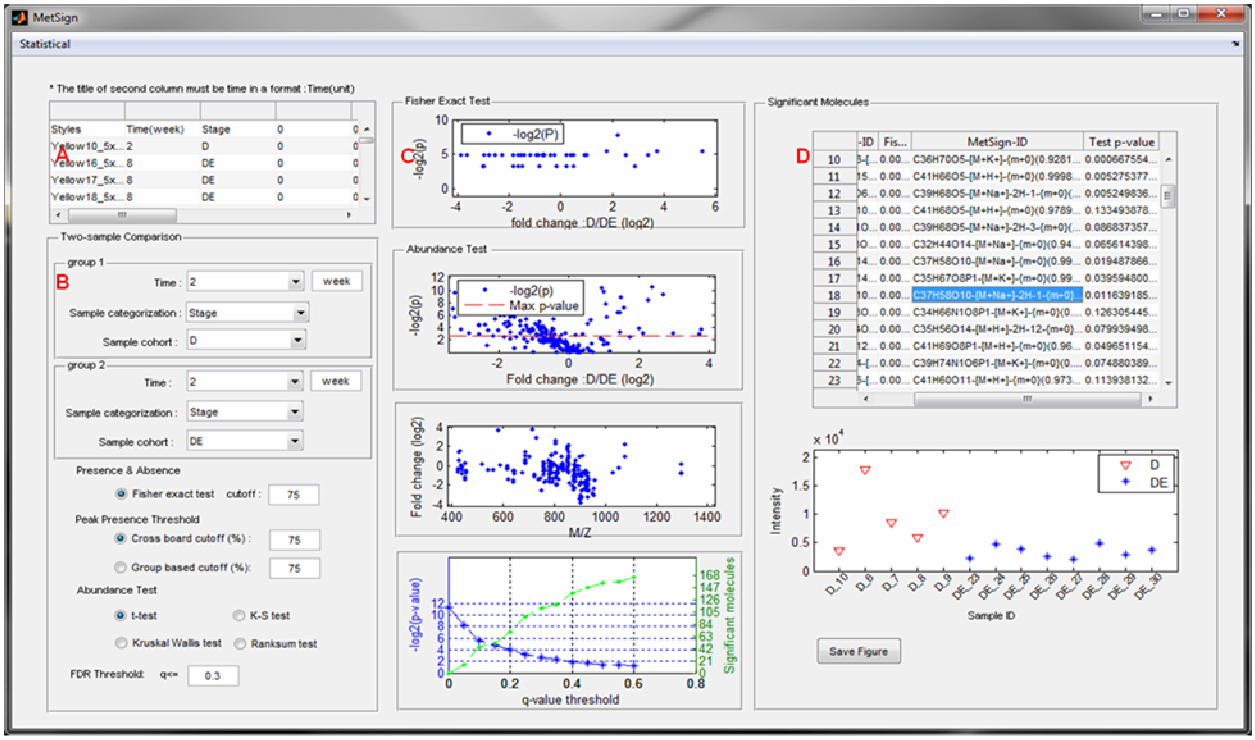


Figure S-3. A screenshot of the statistical significance tests. (A) is a sample information table that contains the sample meta-data including the sample identification number, experimental time, etc. (B) is a parameter setting panel for two-group statistical significance tests. (C) shows the results of presence/absence test, abundance test and false discovery rate. (D) shows the relative abundance of each metabolite in different samples. All plots are automatically generated using an interactive visual data mining approach. D and DE represent two sample cohorts, respectively. The distribution of metabolite abundances shows the regulation changes of a metabolite in between sample cohorts.

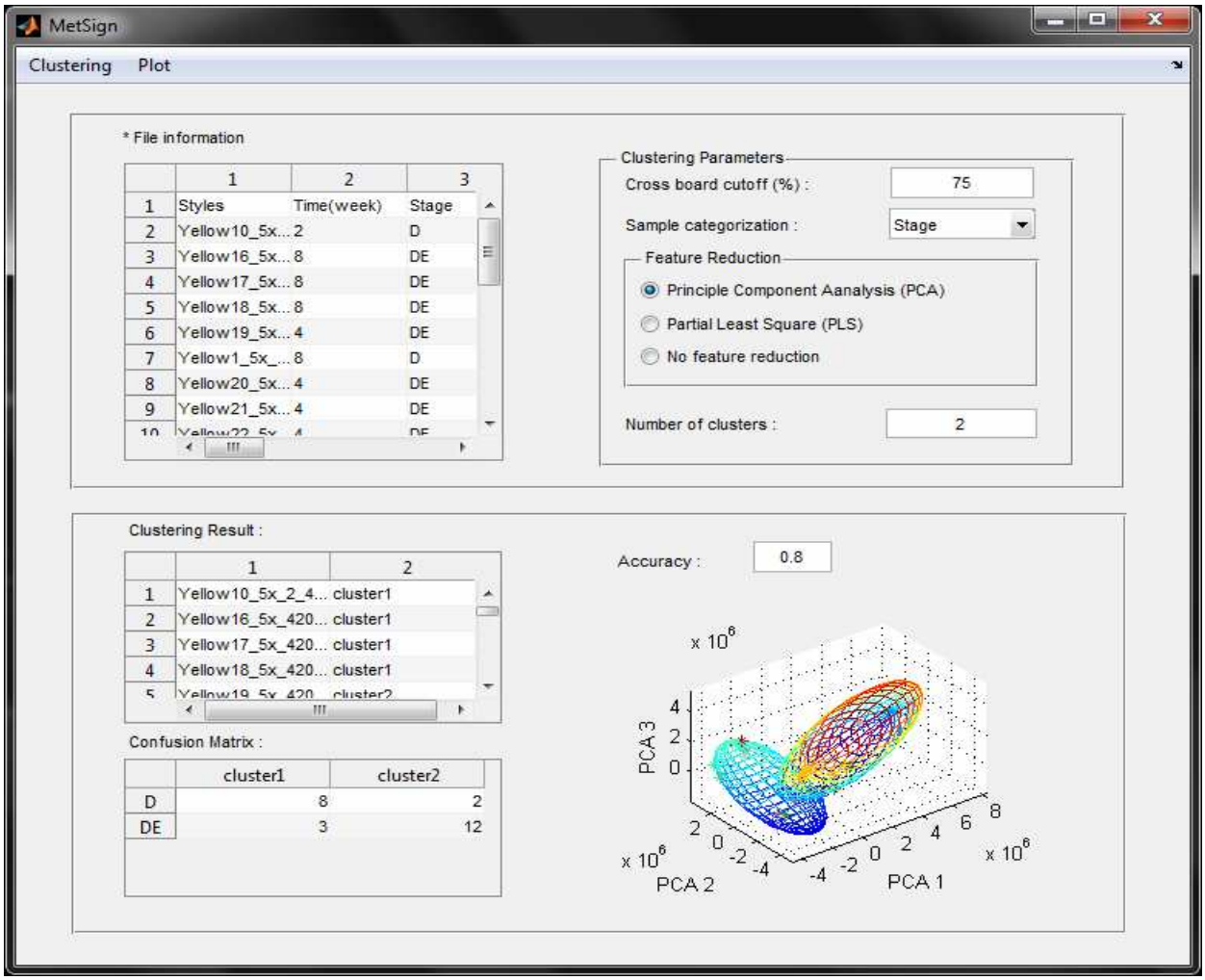


Figure S-4. A screenshot of unsupervised clustering analysis. The upper left panel lists meta-information of all samples. The user can set cluster parameters and select clustering methods in the upper right panel. The clustering results are summarized in the lower left panel. The lower right plot shows the results of three dimension clustering based on principal component analysis (PCA). The visualization is based on the first three components of PCA. Each point represents one sample. All of these samples are clustered by K-means into two groups, D and DE. The ellipsoid mesh is used to represent each cluster. The left bottom table is a confusion matrix that shows the clustering result with a high accuracy of 0.8.

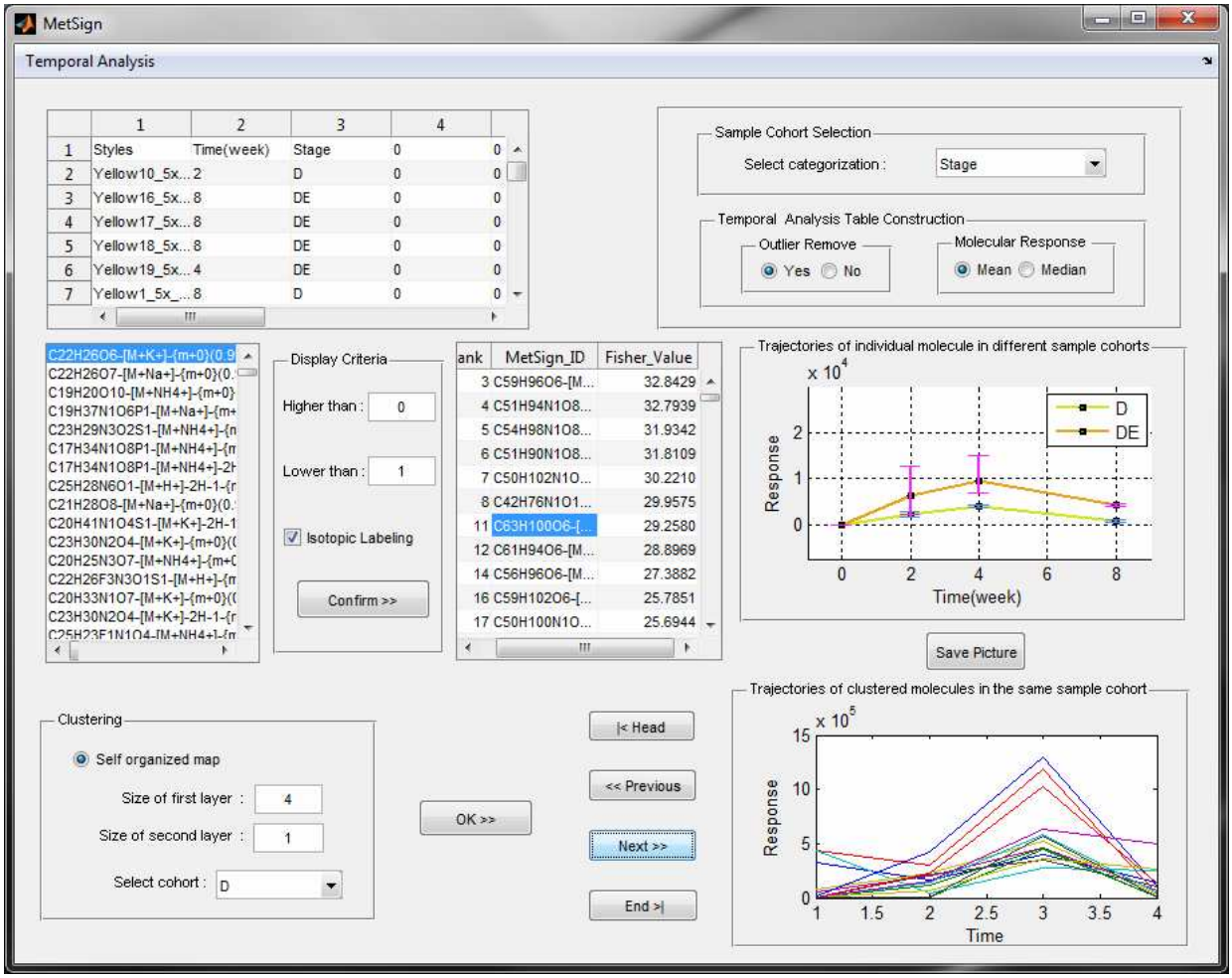


Figure S-5. A screenshot of the temporal analysis. The right middle plot shows the trajectories of a metabolite, prioritized by automatic statistical analysis (see the left middle panel), in two sample cohorts D and DE. This metabolite is labeled with one ^{2}H and has an adduct ion of H^{+} in all samples. The black points, the upper bound and the lower bound of each bar represent the mean, the maximum, and the minimum of the response of this metabolite in all samples at a certain time, respectively. The lower right plot shows the clustering results of all trajectories in the sample cohort D by using a self-organized map (SOM).

Table S-1 15 kinds of acids are used for the spiked-in experiment. Three kinds of acids are detected in the diluted liver extracts. Ten kinds of acids are detected in the spiked-in liver extracts.

All acid	**Liver extract detected**	**Liver extract spike-in detected**
L-Proline		L-Proline
L-Cystine		L-Cystine
L-Histidine	L-Histidine	L-Histidine
L-Phenylalanine		L-Phenylalanine
L-Tyrosine		L-Tyrosine
L-Lysine	L-Lysine	L-Lysine
L-Glutamic acid		L-Glutamic acid
L-Aspartic acid		L-Aspartic acid
L-Leucine		L-Leucine
Nonadecanoic acid	Nonadecanoic acid	Nonadecanoic acid
Heptadecanoic acid		
Heptanoic acid		
Nonanoic acid		
Pentadecanoic acid		
Undecanoic acid		