MsCompare[™] High Resolution Peak Peaking for LC/MS and GC/MS: - A Quick Starting Guide -

Introduction: this quick starting guide teaches you to quickly perform high resolution peak picking for LC/MS and nominal peak picking for GC/MS data. Both techniques have their own peak picking algorithms. Peak Picking, as opposed to Peak Matching, is applied to each individual sample. The results from all samples are grouped / clustered and presented in the final table. Peak Picking can be followed by univariate statistical or multivariate analysis using many of the specialized algorithms from MsCompare. You must be familiar with the MPeaks and MsCompare programs.

- 1. **MsCompare Data Loading:** start MsCompare and load or create a project containing all of your samples. Based on a visual analysis, decide if alignment is necessary. If this is the case, see the tutorial documents on Alignment.
- 2. Start MPeaks and Run LC/MS Peak Picking: MsCompare will use the last used LC/MS Peak Picking parameters from MPeaks. Load one of your samples into MPeaks and run Peak Picking using the criteria and sensitivity you need. For more information on Peak Picking, see the Quick Reference Guide: MPeaks High Resolution Peak Picking.
- 3. Return to MsCompare and Run LC/MS Peak Picking: To run High Resolution LC/MS Peak Picking, select from the menu: Peak Picking > LC/MS settings. A GUI is presented with additional options for multi sample peak picking. These settings will be used after peak picking on individual samples is completed. First select the type of analysis to be applied: Accurate or Nominal Mass Mode. Select the time and mass ranges. Next, you will have the option to remove ¹³C Isotope peaks, delete broad peaks and to specify a new relative or absolute peak height threshold.

Peak Clustering:

Accelerating Data Analysis in LC/MS

Press Done and the settings will be saved. In the next step select: Peak

Picking > Run Peak Picking. You will be asked to enter a value for clustering similar peaks from different samples. Check your data and see how much shift exists between some of the typical peaks. Measure the distance in scans or minutes using the Peak Width Measurement Tool (on the Icon Toolbar). Enter the number of scans so that all peak maxima will

be in the interval (top arrow). A too small clustering value will create additional split clusters; a too large value might miss small peaks. Press "OK" and the Peak Picking &Clustering will be started. If accurate mass mode was selected you will need to enter the mass accuracy and time window for extracting accurate EIC's. Typical values used are an accuracy of 0.01 Da and a time window of 0.5 minutes (lower arrow), depending on the accuracy of your instrument and the width of your peaks.

Press "OK", and the calculations will begin. At the end the table can be saved directly and will

be presented in MsCompare. You will now be ready to apply univariate and multivariate techniques to find peaks that are responsible for discriminating the different groups/classes in your project.





finding what's important...fast...

4. Run GC/MS Peak Picking: MsCompare has a special algorithm for GC/MS Peak Picking. Compared to the LC/MS algorithm, this one is slower, but is more accurate when many peaks are present in one and the same mass chromatogram. To start GC/MS Peak Picking select from the menu: Peak Picking > GC/MS Settings. A window will be displayed from which you can select and experiment with the GC/MS peak picking parameters. The peak picking results for the TIC or individual mass traces will be shown and updated immediately. For detailed description of all parameters and the peak picking interface, see the GC/MS - QC Quant Manual.



Press "Save Settings" to store the GC/MS peak picking parameters. Next, from the menu select: Run Peak Picking. You will have to enter a clustering value to group similar peaks from different samples (see previous page). After the GC/MS Peak Picking has been completed, you will be presented with the final table showing clustered peaks for all samples.

You will now be ready to apply univariate and multivariate techniques to find differences between your groups or classes.

5. Applying Multivariate Analysis methods to Discriminate between Groups: From the Menu select Multivariate Analysis to explore the data set. Use PCA to get an impression of outlying samples, unsupervised grouping and to find peaks that are responsible for class separation. For two groups you can use PLS-DA to apply a supervised analysis. In case of more than two groups use Extended Canonical Variate Analysis (ECVA), a combination of PLS and Linear Discriminant Analysis.

a @ La •	10 10 10 10		ZR	125 BALL	A 🖸 🗛										
		(MD													
MeC	ompare		_				_								67.084 min.
in a c	ompart		Pea	(B)min.)	m/z	s1: Cont	12: Cont	s3. Cont	:4 San.	15: Sam	16: Sam				
Sample List		/0	36.136	1423.7630	39.51	66.31	60.05	70.7	68.09	60.52					
			71	104.940	1296.6835	37.68	36.4	36.31	29.8	30.15	29.68				
No. Nerse	Clerr	Yvar	72	106.104	1555.0593	20.14	20.04	19.58	27.84	25.63	25.35				
Control	Contra		73	85.699	1099.5076	48.21	51.94	52.09	62.97	62.57	53.55				
Control2	Contra		74	/1./30	1161.6528	46.85	47.55	4/22	52.61	51.14	52.22				
Controls	Compa		75	50.045	847.9082	45.08	43.56	43.12	26.05	26.5	25.9				
Sample 1	equec		76	75.264	840.9251	34.14	32.7	32.56	33.84	31.82	34.46				
Sample 2	Sample		77	88.770	935.8063	22.58	22.85	22.77	26.73	27.06	26.26				
Sample 3	Sample		78	72.341	594.8389	43.2	43.85	42.34	45.25	44.2	44.21				
			79	54,590	838.5006	52.67	52.3	55	55.35	54.8	54.65				
			80	56.984	940.9662	11.67	11.48	11.85	12.99	12.11	11.78				
			81	96.149	900.8390	20.42	18.1	19.08	24.28	22.58	22.62				
			82	43.294	/13.0131	31.96	31,45	30.04	33.15	31.37	30.67				
			40	67.380	940,4348	2.23	2.29	2.67	33.69	35.36	33.55				
			84	54,560	419.7567	69.61	67.23	67.54	71.87	70.2	66.7				
			85	87.846	939.9715	11.83	12.45	12.36	13.39	13.41	12.92				
			26	76.299	934.4313	54.18	53.99	51.54	39.1	37.25	37.42				
		>	87	67.034	940,1184	2.67	2.55	2.81	38.77	38.54	36.55				
Lond Dont															
	Cierc	Bullenter						E	tracted lon C	arrents mitz	940 @67.05	nin.			
	Clear	Browser		-				E-	tracted lon 0	Currents eviz	940 @67.08	min.	1	-	
	Clear	Browser Range	1600 -	1	1	T	_	. /	tractied lon C	lurrents eviz T	940 @67.05	min.	1	1	
Chans Ladeel	Clear Label	Browser	1600 -	1	1	1			tractied lon 0	Currents eviz T	940 @67.08	i nin.	,	1	T
Class Label	Clear Label Control	Browser	1600 -	1	T		'	 M	tracted lon 0	Currentic m/lo	940 @67.08	i nin.	 1	,	
Class Label	Clear Label	Browser Range	1600 -	1	1		,	M	tracted lon 0	Currentic m/lz	940 @57.03	l min.	1		
Class Label	Clear Label Control	Browster	1600 -	1	1			M	tracted lon 0	larvents m/s	940 @67.03	i min.	1	,	,
Class Label	Clear Label Control	Browser Range	1600 - 1400 - 1200 -	1	1	,		•	fractied lon C	Larrents m/s	940 @67.03	i min.	1		1
Class Label	Clear Label Control O Top Hitegr.	Browser Range	1600 -	,				M	fractied lon C	Larrents m/z	940 257 03	i min.	,	,	
Class Label (Clear Label Control	Browser Range	1600 - 1400 - 1200 - 1000 -	1				•	fractied lors 0	T	940 257 03	i mn.	,		,
Class Label	Clear Label Control	Browser Plange © Bottom Limits (rein.)	1600 - 1400 - 1200 - 1000 -	1	-			M	fractied lors 0	Currents eviz	940 2267 03	i mn.	 ,		
Class Label	Centrol Label Control Top Hitter.	Browser Plange © Bottom Linits (nin.) Integrate	1600 - 1400 - 1200 - 1000 - 800 -	1	-			E.	tractient lors (Jurrents evic	940 (207.03	i mn.	 ,		
Class Label Class	Clear Label Control Vitegr.	Browser Pange © Bottom Linits (nin.) Integrate	1600 - 1400 - 1200 - 1000 - 800 -	1	-	,		M	tracted ion C	Lerents sviz	940 (2007.03	i min.	T		
Chain Labert Chain Labert Thick Diffic Diffic Diffic Diffic Diffic Set Toth None	Clear Label Control Top Hittp: 	Browser Range Bottom Linds (rein.) Integrate ange	1600 - 1400 - 1200 - 1000 - 800 - 600 -	1	-			E I	tracted ion C	larrents sviz	940 (207 03	i min.	1		
Class Label Class Label Class Label Class Label Class Label Class	Clear Label Control	Browser Range Bottom Umits (min.) Integrate ange S00.2000	1600 - 1400 - 1200 - 800 - 500 -	1				M	tractied lon C	arrents sviz. T	940 (2017) 03	i min.	1		
Chees Labert Chees Labert D TC D BTC D BTC D BTC D BTC D BTC D BTC None Stricts Chees Codors	Clear Label Control Top Hittigr	Bostom Linds (vin.) integrate solution Solution	1600 - 1400 - 1200 - 1000 - 800 - 600 - 400 -		1			······································	tractied lon C	larrente evite T	940 (2017) 03	i min.	1		-
Chars Label Chars Label Chars Label Chars Label Chars Label Chars Codes Chars Codes Chars Codes Chars Codes	Clear Label Cartrol Top Hitegr.	Browster Range Bottom Linits (min.) Irringrato ange 300.2000	1500 - 1400 - 1200 - 1000 - 800 - 600 - 400 -	1	1			M	trached lon C	Serveda aviz	940 (2017) 03	i niti.	1		
Clean Label Clean Label Clean Label Clean Label Clean	Control Control Top Here: w mtz & Same R mtz Same	Browster Plange © Bottom Linits (min.) Integrate Mrge 300.2000	1600 - 1400 - 1200 - 1000 - 800 - 600 - 200 -		-			M	trached lon C	T	940 (2017) 03	80.	1		
Chars Label Chars Label Chars Label Chars Label Chars	Control Control Top Kittyr Ki	Bruwser Range Battom Linkts (HA) Integrate Mage 300 2000 Posits Dette myte	1600 - 1400 - 1000 - 1000 - 600 - 400 - 200 -		1			M	Iraclad Ion C	Samenda mila	940 (2017) 03	1 80.	1		
Chan Labert Chan Labert Chan Labert Chan Labert Chan Labert Chan Chan Chan Chan Chan Chan Chan Chan	Control Control Top Hengy, W Mark Deeps Mark Deeps	Browser Range © Bottom Linits (rein.) Integrato Mign 300:2000 2 Posits Delle m/z 34	1600 - 1400 - 1400 - 1000 - 800 - 400 - 200 - 0					M	tracted tor 0	lamenta este	940 (2017) 03	inn.		_	

6. Applying Univariate Statistical methods to find discriminating Peaks for two Groups: For two groups, select Biomarker/Stats and decide what kind of peaks to search for; up-regulated or down-regulated. Some of the algorithms (using peak ratios) need to know which group is expected to be up-regulated. MsCompare has a large number of powerful statistical test (and graphics) to find discriminating peaks. These tests include: t-test, ratio analysis, uniqueness test, full selectivity, percent up-regulated and the Fisher Discriminant test. MsCompare also has a new feature: Create a Multi Test Statistical Overview. It will calculate all statistics for all peaks and present an overview of the most discriminating peaks. The peaks can be directly plotted from this overview and peaks of no importance can be deleted from the main table.

Statistics Results Overview (Median): Class Sample/Control													
r	43	732.0332	72.30	23.01	1.034	3.11	0.0057	1.7	197	1	100	1.0	
8	51	1402.2258	88.73	14.24	1.189	5.71	0.0047	8.6	296	1	100	3.29	_
9	52	1124.5192	99.78	23.63	0.901	3.08	0.0369	-5.2	-293	-1	0	-1.78	
D	53	1423.269	96.15	29.05	1.17	3.69	0.0211	7.8	599	1	100	2.13	
1	54	1274.1069	82.94	18.47	1.026	3.19	0.0331	1.3	80	1	100	1.84	
2	58	1401.7219	88.73	17.06	1.193	10.78	0.0004	8.8	366	1	100	6.22	
3	59	426.2151	49.19	30.25	11.272	284.5	0	83.7	3713	1	100	164.25	
4	60	1072.0481	74.91	13.6	1.096	4.49	0.0109	4.6	162	1	100	2.59	
5	61	1260.6312	90.58	14.28	0.884	6.28	0.0033	-6.1	-241	-1	0	-3.63	
6	63	645.7847	47.79	20.39	1.048	2.89	0.0446	2.3	159	1	100	1.67	
7	65	626,3206	56.97	12.78	1.043	5.37	0.0058	2.1	81	1	100	3.1	
8	66	973,5335	91.47	43.74	1.273	2.31	0.082	12	1419	1	100	1.33	
9	67	1496.7795	104.5	9.7	0.873	3.88	0.0179	-6.8	-152	-1	0	-2.24	
0	68	830.4195	97.61	17.34	0.88	6.45	0.003	-6.4	-340	-1	0	-3.72	-
1	69	934 4651	88.73	23.57	1 1 2 2	5.88	0.0042	57	420	1	100	3.39	
2	70	1423 7698	96.19	25.7	1 151	7.38	0.0018	7	524	1	100	4.26	-
3	74	1296 6835	104.94	11.11	0.819	14.95	0.0001	-10	-346	4	0	-8.63	-
4	72	1555 0593	106.1	9.77	1 279	7 89	0.0014	12.2	318	1	100	4.55	-
5	73	1099 5076	85.7	22.05	1.204	6.6	0.0027	0.3	548		100	3.84	-1
<i>.</i>		1033.3070	00.1	42.00	1.204	0.0	0.0021	0.0	040		100	3.01	

- 7. Applying Univariate Statistical methods to find discriminating Peaks for more than Groups: To create an overview of the most discriminating peaks for more than two classes select: BioMarkers/Stats > Multi Class Pairwise Ratio Analysis or Pairwise Fisher Discriminant Analysis. Both statistics will be presented for combination of classes (1-2, 1-3, 2-3 etc.). The overview is interactive. You can directly plot the peaks in MsCompare and determine how well separation is between your groups.
- 8. Using Multivariate Analysis on the most Significant Peaks: The methods under 7 and 8 can be used to reduce to original table so that it will contain the most important peaks only. After that you can use the Multivariate techniques for presentation purposes.



Full Document References:

- 1. MsX Differential Analysis for two samples 2011
- 2. MsCompare Univariate and Multivariate Methods for Class Discrimination
- 3. MsX High Resolution Peak Matching using MsCompare