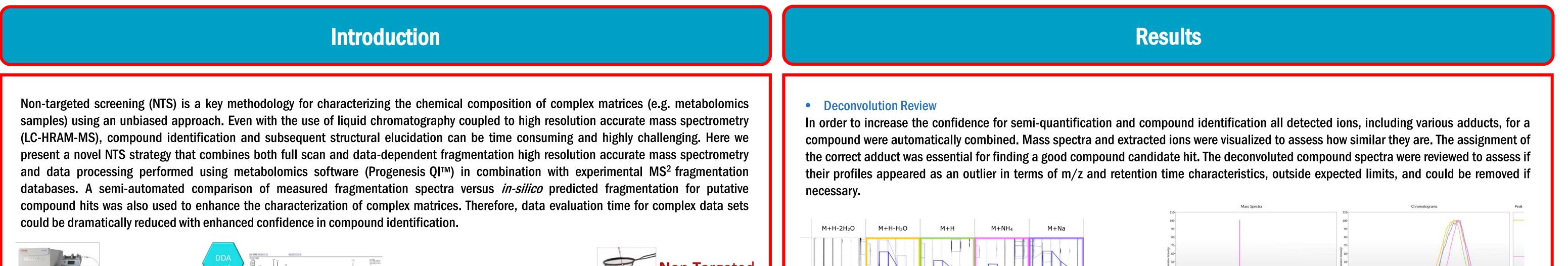
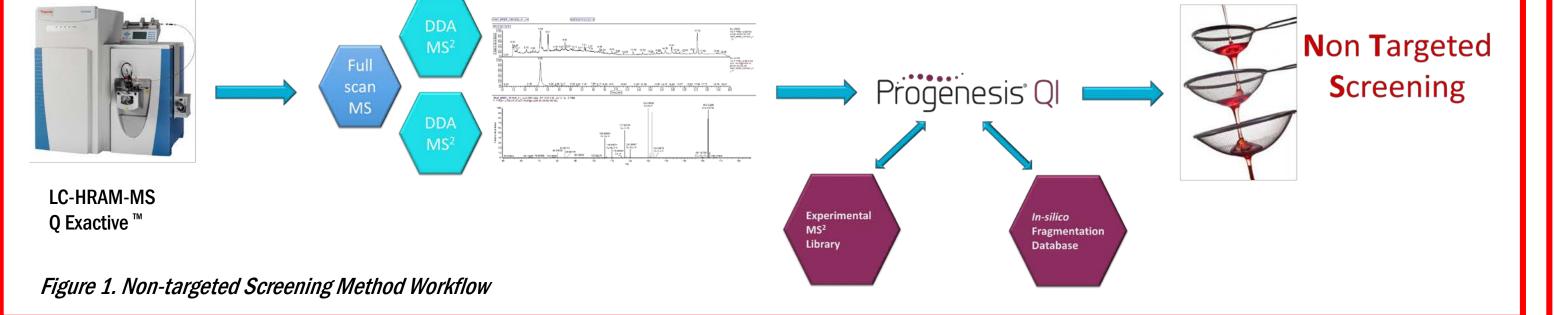
High resolution mass spectrometric data analysis using Progenesis QI software for non-targeted screening (NTS)

Stefania Della Gatta, Christoph Buchholz, Mark Bentley and Daniel Arndt

Philip Morris International R&D, Philip Morris Products S.A., Quai Jeanrenaud 5, 2000 Neuchâtel, Switzerland (part of Philip Morris International group of companies)

stefania.dellagatta@pmi.com



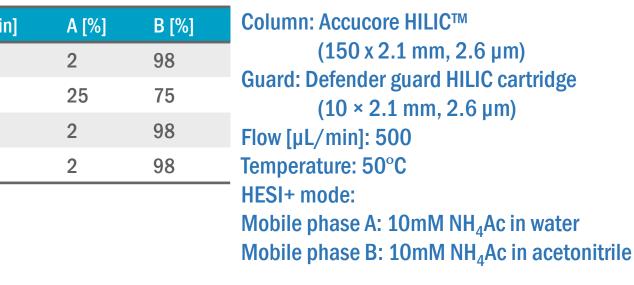


Method

Data generation was performed in full scan mode combined with high-energy collision dissociation (HCD) and stepped normalized collision energy (NCE) applied using a Q ExactiveTM high resolution accurate mass spectrometer (Thermo Fischer Scientific, Bremen, Germany). Sample replicates (n = 5) of 2 or more samples, a blank and a pool sample (equal proportions of each sample including the blank in order to combine all compound features within a single reference sample) were fortified with a set of stable isotope labeled internal standards to enable semiquantification. Analysis was performed using reversed phase (RP) chromatography with positive and negative heated electrospray ionization (HESI+/HESI-) and positive atmospheric pressure chemical ionization (APCI+) to cover a wide range of substances with different ionization properties. In addition, samples were analyzed using hydrophilic interaction chromatography (HILIC) with HESI+.

Table 1. Grad	dient RP N	lodes		Tabl	
Time [min]	A [%]	B [%]	Column: Hypersil GOLD™	Tim	
0	85	15	(150 × 2.1 mm, 1.9 μm) Guard: UHPLC filter cartridge	0	
7.00	10	90	$(10 \times 2.1 \text{ mm}, 0.2 \mu\text{m})$	7.0	
12.80	0	100	Flow [µL/min]: 400	8.0	
18.00	0	100	Temperature: 50°C	15.	
18.10	85	15	HESI+, APCI+ mode: Mobile phase A: 10mM NH₄Ac in water,		
20.00	85	15	Mobile phase B: $1 \text{mM} \text{NH}_4 \text{Ac}$ in methanol		
			HESI- mode:		
			Mobile phase A: 1mM NH ₄ F in water		
			Mobile phase B: methanol		

ble 2. Gradient HILIC Mode



HRAM Detection

Full scan MS was performed at a resolution of 70.000 (FWHM) acquiring a mass range of m/z 80 – 800 in combination with a datadependent MS² Top3 of each scan at a resolution of 17.500 (FWHM) and applied stepped normalized collision energies of 25, 50 and 75 eV and automated gain control of 1 x 10e⁵ in order to generate HCD first order fragmentation (TopN = 3, loop count = 3, dynamic exclusion = 10 s). Vaporizer heater temperature, capillary temperature, spray voltage, sheath gas and auxiliary gas were set at 350° C, 380° C, ±3.00 kV, 60 and 20 arbitrary units respectively for HESI modes. Vaporizer heater temperature, capillary temperature, sheath gas and auxiliary gas were set at 450° C, 380° C, 5.0 µA, 50 and 5 arbitrary units respectively for APCI mode. Figure 9. Mass spectra and chromatograms for the detected adducts of a compound, color coded by adduct

Compound Identification

Figure 8. Montage view showing location of detected adducts for a compound

A revolutionary and major part of the workflow for dealing with high resolution accurate mass first order fragmentation data is the semiautomated process for compound identification. Compound identification was performed using a stepwise approach employing experimental MS² fragmentation databases and *in-silico* predicted fragmentation of chemicals from public databases. In Step 1 all detected constituents were matched and assigned against an in-house database comprising experimental data for approximately 400 reference compounds with accurate mass data, stepped NCE MS² first order fragmentation and retention times (precursor and fragment tolerance 5ppm, retention time tolerance 0.5 min). In Step 2 fragmentation patterns for all detected constituents were compared with *in-silico* predicted fragmentation of putative hits from UCSD (Unique Compounds & Spectra Database, PMI, Neuchâtel, Switzerland)¹, HMDB 3.6 (Human Metabolome Database, University of Alberta, Edmonton, Canada)^{2,3,4} and, via the ChemSpider search plugin, with ChemIDplus (ChemIDplus, SIS, NLM, NIH, Bethesda, MD, USA) and FDA (U.S. Food and Drug Administration, Silver Spring, MD, USA) (precursor and fragment tolerance 5ppm). In Step 3 fragmentation spectra for detected constituents were compared with experimental fragmentation spectra of NIST14 MS/MS library (precursor and fragment tolerance 5ppm) (U.S. National Institute of Standards and Technology, Gaithersburg, MD, USA). All putative hits were scored using Progenesis QITM algorithms, which considered mass similarity, isotope similarity and fragmentation score.

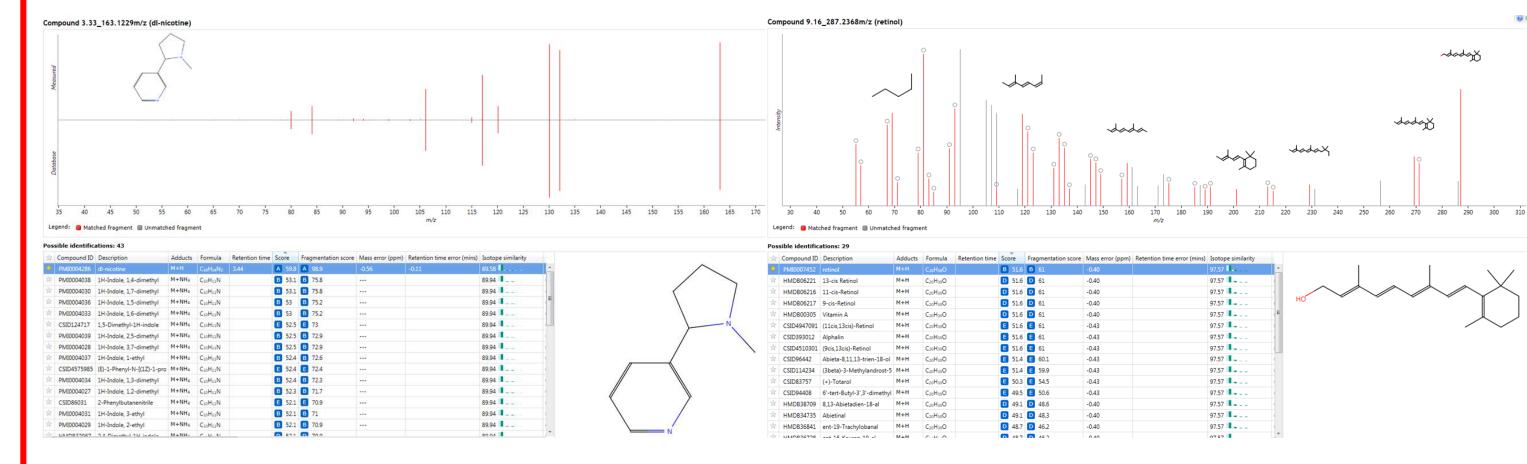


Figure 10. Experimental MS² Database Comparison

Compound Review

A manual review process was performed to ensure the correctness of compound identification. Each putative hit was checked regarding compound abundance, detected adducts, fragmentation score, retention time score, isotope similarity, mass error and overall score. This review step also considered the measured isotopic distribution compared to theoretical. This isotopic distribution match contributed to the overall compound identification score.

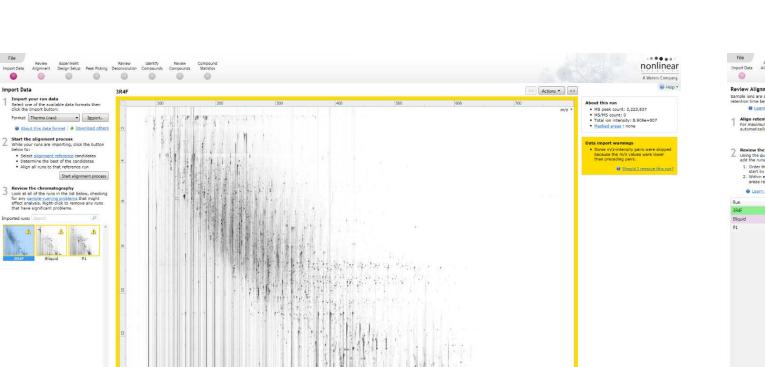
Figure 10. in-silico Fragmentation Database Comparison

Data Processing

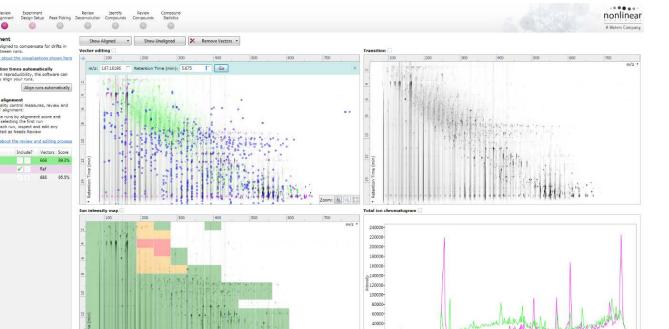
Acquired data were processed using Progenesis QI [™] software (Nonlinear Dynamics, Newcastle upon Tyne, UK), consisting of raw data import, selection of possible adducts, peak set alignment, peak detection, deconvolution, data set filtering, noise reduction, compound identification and normalization with internal standards. In the developed workflow, further data evaluation steps are performed for compound identification.

Fundamental steps within the Progenesis QI™ data evaluation workflow:

- Selection of possible adducts
 - [M+H]⁺, [M+NH₄]⁺ for RP HESI+ and HILIC HESI+
 - $[M+H]^+$, $[M+H-H_2O]^+$ for RP APCI+
 - [M-H]⁻, [M+F-H]⁻ for RP HESI-
- Importing of Thermo.raw profile data
- Visual quality check of each analytical run using ion intensity maps
- Alignment with a selected reference run (one of the pool samples)
- Experimental design setup (defining one or more groups for aligned runs)
- Peak picking
- Normalization versus a set of internal standards
- Automatic deconvolution to enable accurate quantification of each compound



A REAL PROPERTY AND A REAL		the samp	les?		
High resolution mass spec	trometer		•		
elect the runs' data form	at:				
Profile data					
Centroided data					
Resolution (full width	at half maxir	mum): 5	0000		
elect the ionisation polari	ity for this ar	nalysis:			
Positive		-			
Negative					
-					
		P	ack Next	Cancel	
Create New Experiment				··· · · ·	
Adducts in library:	210		Adducts in this experim	ent:	
		22	M+H M+NH4		
		<<	- Contrained		
M+CH3OH+H					
	e (
 M+CH₃OH+H M+K M+ACN+H M+2Na-H 	e (
M+CH ₃ OH+H M+K M+K M+ACN+H M+2Na-H M+ACN+Na					
 M+CH₃OH+H M+K M+ACN+H M+2Na-H M+ACN+Na M+2H 	¥				
 M+CH3OH+H M+K M+ACN+H M+2Ns-H M+2Ns-H M+ACN+Na M+2H M+H+Na 					
 M+CH₃OH+H M+K M+ACN+H M+2Na-H M+ACN+Na M+2H 					
 M+CH₃OH+H M+K M+ACN+H M+2Na-H M+ACN+Na M+2H M+2H M+H+Na M+2Na 					
 M+CH₃OH+H M+K M+ACN+H M+2Na-H M+2AN-H M+2H M+2H M+2H M+2Na M+3H 					
 M+CH₃OH+H M+K M+ACN+H M+2Na-H M+2AN-H M+2H M+2H M+2H M+2Na M+3H 	-	Back	Create experiment	Cancel	
 M+CH₃OH+H M+K M+ACN+H M+2Na-H M+2AN-H M+2H M+2H M+2H M+2Na M+3H 			Create experiment	Cancel	

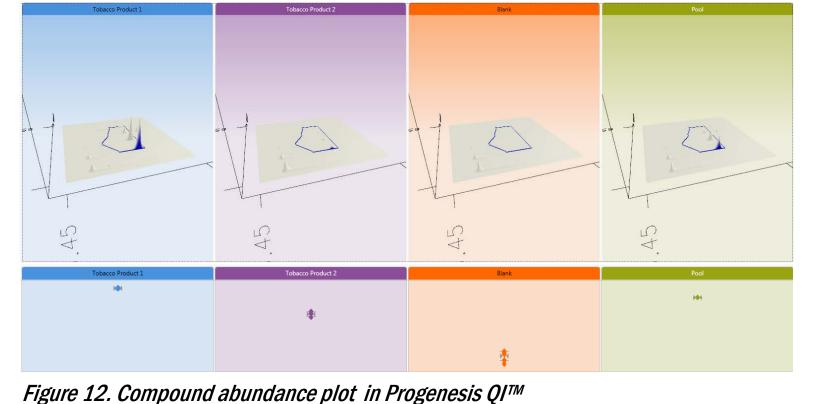


0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 Retention time (mig)

File Review Experiment our Data Algoment Design Setup Peak Picking	Review Identify Review Compo Deconvolution Compounds Compounds Statist							nonline	ang Review Dependent Peak Picking	Review Identify Deconvolution Compound						nonlin A Waters Cor
eview Compounds	Find a compound: Search P	Filter compounds	- Filter is active					9 H	p validate the alignment and peak	O Done						ap layout:
sing this screen, you can find the compounds of terest in your experiment.	Compound Neutral mass m/z	z Retention time Peak V	Width Tag - Accepted ID Iden	fications Anova (p) Max fold cl	hange Highest mean Lowest mean	Isotope distribut	on Max Abundance Min C	% Description	impound's ions. ak picking and alignment	Blank_RPEP_1501932_0	650	1000 C	P_1501932_03_1D 650	Contraction of the second	EF_1501932_03_1E	Lange Competence in Comments
Create a shortlist to review	548.6073m/z <unknown> 648.6073</unknown>	1 17.11 0.90	PM0000109 3	< 1.1E-16 E76	Tobacco Product 1 Blank		120109004/7549 2.86	solanesol	* Iduct in the list below to view its ion		TIL	m/z • 2		m/z * 2	m/z *	
In the table, sort and filter the compounds based on their measurements, to generate a shortlist	• 3.33_163.1229m/z <unknown> 163.1229</unknown>	1 3.33 0.82	PMI0004286 43	< 1.1E-16 248	Tobacco Product 1 Blank	a	87776428.2508 2.58	dl-nicotine	spectrum and chromatogram across I runs:	199		1991		18.6		· hereit /
for further review.	• 3.97_218.0789n 218.0789 236.1127	1 3.97 0.26	PMI0000113 10	<1.1E-16 7.29E+03	Tobacco Product 1 Blank	10	17097246.0307 4.34	triacetin	m/z Charge	a multiple and a mult		S years		S S S		1 PARA
How are the measurements calculated?	9 12.32_429.3728m/z <unknown> 429.3728</unknown>	1 12.32 0.37	PMI0009304 24	< 1.1E-16 482	Tobacco Product 1 Blank	11.4	13781492,4930 5.32	7-Ketositosterol	648.6073 1	17.0		17/0 ime		17.0		and the second second
To sort the table by a given value, simply click the relevant column header.	• 3.99_151.1168m/z <unknown> 151.1168</unknown>	1 3.99 0.68	PM00009939 1	2.53E-12 1.11	Pool Blank	8	11334595.2722 0.96	Myosmine-2,4,5,6-d4		27.5 on T		en T		1 1 10		and the second
	● 5.45_147.1618m/z <unknown> 147.1618</unknown>	1 5.45 0.23	PMI0006924 2	≥.9995 1	Tobacco Product 1 Pool		10128489.1911 0.00	d8-Isophorone	E	18.0	4	38.0	1.1	288		* MI-1 10 M
2 Review the compounds For each compound of interest, inspect the ions'	• 3.51_193.0495m/z <unknown> 193.0495</unknown>	1 3.51 0.80	PMI0000309 43	< 1.1E-16 120	Tobacco Product 1 Blank	8	7332531.9296 2.05	Scopoletin	noound's adduct form is derived	Re 185		Zoom: (2) (3) (2) (2)		Zoom: 12 3 53 2 2	Zoom: 10, 3, 53	The second se
alignment and peak picking:	• 4.84_215.1179m/z <unknown> 215.1179</unknown>	1 4.84 1.95	PMI0002136 72	< 1.1E-16 735	Tobacco Product 1 Blank		6382538.1358 1.85	Harmaline	accepted identification.	Deed 14		Doub 10		050100		111
Review selected compound	• 2.64_177.1022m/z <unknown> 177.1022</unknown>	1 2.64 1.59	PMI0001948 110	<1.1E-16 118	Tobacco Product 1 Blank	10	5837098.5594 2.70	Cotinine	ul rues -	645	650	655	45 650	655	645 650 655	Zoom location = Compound for
You can also double-click to review a compound.	• 9.58_288.2451n 288.2451 289.2524	1 9.58 0.34	PMI0007970 40	< 1.1E-16 642	Tobacco Product 1 Blank	10	5573724.5467 2.18	Isoraimonol	ty problems	3		m/z • 2		m/z • 3	m/z *	
	• 321_191.1178m/z <unknown> 191.1178</unknown>	1 3.21 0.52	PMI0009912 94	< 1.1E-16 1.04E=03	Tobacco Product 1 Blank	18	4384930.3422 1.52	rac-5-Methylcotinine	are misaligned, return to Alignment	20		1		100		Reset zoom to adduct local
3 Choose the correct identifications For each compound, select one of its possible	• 9.16_287.2368m/z unknown= 287.2368 	1 9.16 1.49	PM00007452 29	< 1.1E-16 Z23	Tobacco Product 1 Blank	8	3820499.7640 1.66	retinol	ual alignment vectors.	ST E		3 E		S E		
identifications as the accepted one.	6 5.96_169.0760m/z <unknown> 169.0760</unknown>	1 5.96 0.31	PMI0000439 21	< 1.1E-16 844	Tobacco Product 1 Blank	R	3418795.1911 3.29	norharman	 ing has missed an isotope, either: to Peak Picking to change the peak 	17.0		177.0		17.0		-
To speed this up, you can automatically accept identifications in compounds where only one of	1		Ш					,	parameters, or adduct to add the missing isotope	17.5 on T		17.5	1 1 1 1 1 1 1	D/S		8
the possible identifications has:	Compound 17.11_648.6073m/z:								Edit adduct	tent		tens		18.0		
Score ≥ 50.0 Accept identifications	Compound abundance Possible identifications	3D Montage							Loit adduct	2.84		Zoom: Q Q C a		Zoom: Q Q C a	Zoom: 🔍 🔍 🔀	-
A Export data for further processing	Possible identifications: 3											Mass Spectra		Contraction of the local data		
By exporting your data to external tools, there's	🕸 Compound ID Description	Adducts Form	ula Retention time Score Frag	mentation score Mass error (pp	m) Retention time error (mins) Isotope simila	ń.		04				mass spectra			Chromatograms	
no limit to your analysis.	PME0000409 solanesol	M+NHa CetH	iO 17.11 F 79 F	-0.85			proprog	~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~								
Export to pathways tool	PMI0002014 (2E,6E,10E,14E,18Z,22E,26E,3	02)+3,7,1 M+NH4 CasH2			99.65					600000000-	1			60000000-		(3)
	🖄 CSID4585196 Solanesol	M+NHz CasHz	vO 🚺 43.5 🚺	-0.87	99.65											
										50000000-				50000000-		
										400000000-				400000000-		
										German				(general and general and g		
										\$ 300000000-		1		\$ 300000000-		
										200000000-				20000000-		
										10000000-				10000000-		
														~ ~		
									0.0	0	648 648.5 649	649.5 650 650.5	651 651.5 652	652.5 16.35 16.4	1645 165 1655 166 1665 167 1675 168 1685 169	16.05 17 17.05 171 1715 172 17
									sign a from a different perspective:		and anot and	Adduct m/z	001 001-0 005	0000 1000 104	Retention time (min)	way to stay and 1103 first 1/4
											an Dandwat 4 🕮 Tabanan Dandwat					

Figure 11. Compound Review Window in Progenesis QI™

In the 'Review Compounds' stage the behavior of compound subsets can be examined based on tag filters. Good power for differentiating single compounds between sample groups has been shown and basic statistical evaluations were performed.



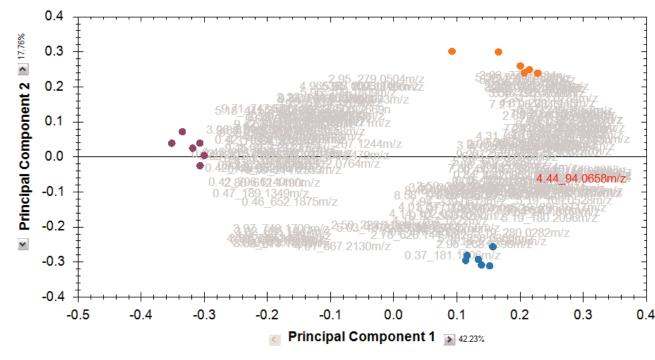


Figure 13. Principal Component Analysis in Progenesis QI™

Definitive compound confirmation was performed using reference standards matched with experimental first order fragmentation, isotopic similarity and retention time.

Conclusions



Figure 4. Data Import Window in Progenesis QI™

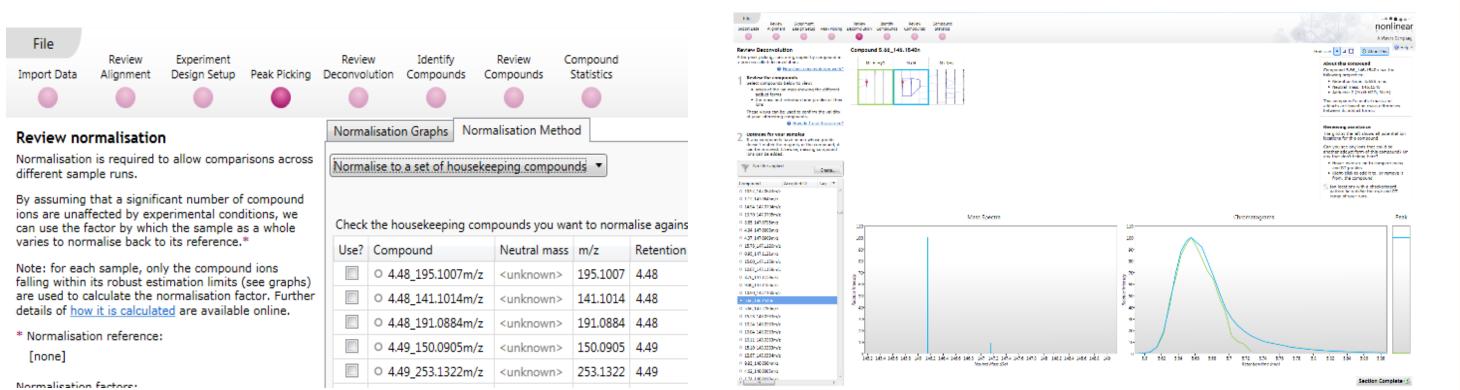


Figure 6. Normalization Window in Progenesis QI™

Figure 7. Review Deconvolution Window in Progenesis QITM

Figure 5. Alignment Window in Progenesis QI™

This non-targeted screening workflow using LC-HRAM-MS, in combination with Progenesis QI™ software, has been demonstrated to be a powerful tool that provides large-scale qualitative and semi-quantitative analysis of analytical datasets with increased processing speed and improved confidence in compound identification for complex matrix characterization. The use of full scan accurate mass data in combination with first order fragmentation spectra enables a robust and efficient process for identifying detected unknown compounds.

References

Martin E., Monge A, et al., Building an R&D chemical registration system, Journal of Cheminformatics 2012 4:11, DOI: 10.1186/1758-2946-4-11
 Wishart DS, Tzur D, Knox C, et al., HMDB: the Human Metabolome Database. Nucleic Acids Res. 2007 Jan;35(Database issue):D521-6
 Wishart DS, Knox C, Guo AC, et al., HMDB: a knowledgebase for the human metabolome. Nucleic Acids Res. 2009 37(Database issue):D603-610.
 Wishart DS, Jewison T, Guo AC, Wilson M, Knox C, et al., HMDB 3.0 – The Human Metabolome Database in 2013. Nucleic Acids Res. 2013. Jan 1;41(D1):D801-7.



NonTarget 2016, Ascona, Switzerlandand

May 29 - June 3, 2016

Competing Financial Interest

The research described in this poster was funded by Philip Morris International