



PMI RESEARCH & DEVELOPMENT

'Next Generation Fingerprinting'

A GC-HR-TOF-MS Method to Semi-Quantify Constituents in Aerosols and Aerosol Fractions

E. Dossin, A. Monge, E. Martin, P. Pospisil, A. Knorr, M.C. Bentley, P.A. Guy

*Metabolomics & Analytical Chemistry Gr.,
Philip Morris Products SA, 2000 Neuchatel, Switzerland*

Outline

1. Description of the existing GC-MS fingerprinting method
 - Pros & cons of the current fingerprinting method
 - Why switching to a 7200 Agilent high resolution MS instrument?
2. How to tackle semi-quantification of the smoke constituents with the help of chemoinformatics tools
 - Curation step for smoke constituents to be monitored
 - Analysis of reference standards
 - Retention time prediction model (QSPR approach)
 - Selection of appropriate internal standards (clustering approach)
3. Semi-quantification from calibration curve of reference standards
 - Linearity
 - Assessment of silylation
4. What about other smoke constituents?
5. Conclusion



PMI RESEARCH & DEVELOPMENT

Existing Fingerprinting Method

- Aerosol sample generated from a smoking machine (ISO)
 - Whole smoke
 - Gas Vapor Phase (GVP)
 - Total Particulate Matter (TPM)
- Compounds list
- GC columns (HP6890 GC)
 - DB-624: HS-SPME-GC-MS
 - DB-FFAP: GC-MS
 - DB-5-MS: HT-GC-MS
- Detection (MSD5973 MS)
 - Electron ionization mode
 - Full scan (low resolution)

} sbPBS



Volatile Chemicals
Non Polar Chemicals
Polar Chemicals (TMS)

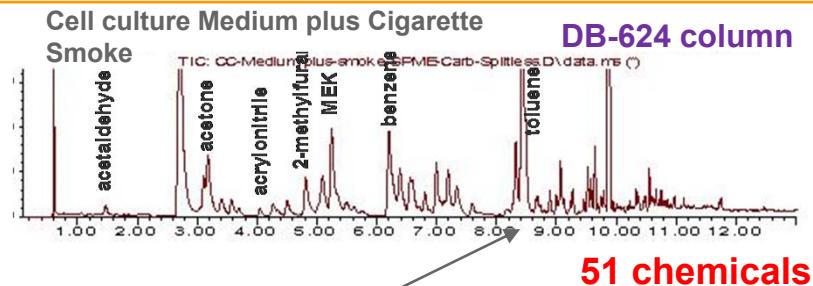
Semi-quantification (d_6 -phenol)



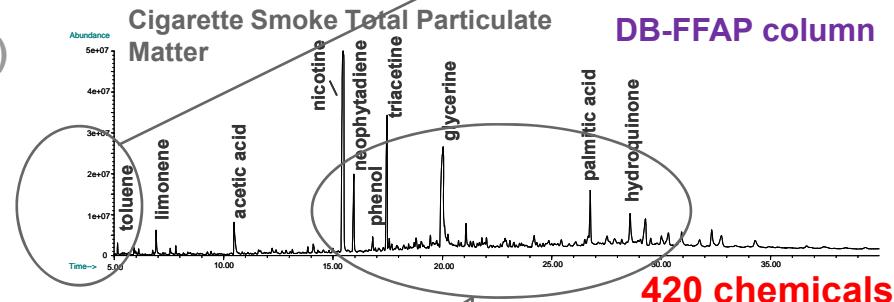
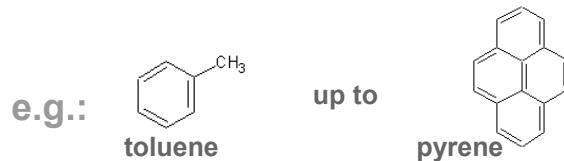
PMI RESEARCH & DEVELOPMENT

Description of the Former GC-MS Fingerprinting Method (HP6890 GC - MSD5973 MS)

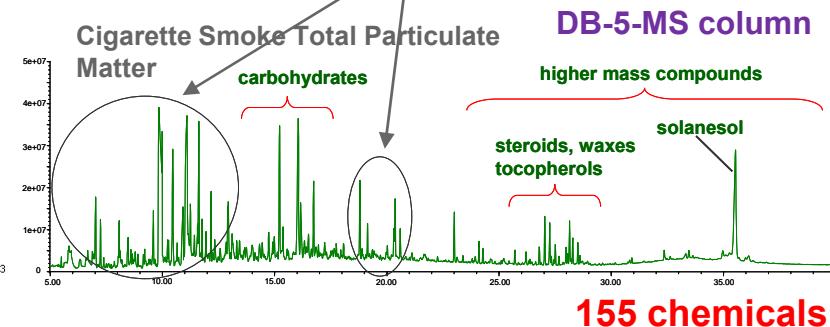
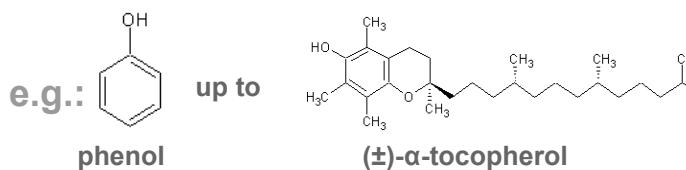
- Headspace(HS)-SPME-GC-MS
high volatile compounds



- GC-MS Fingerprint (Standard Method)
nonpolar compounds



- High Temperature(HT)-GC-MS
derivatized, polar compounds



A total of ~607 smoke constituents are monitored using 3 distinct GC columns
but not all were unambiguously identified



PMI RESEARCH & DEVELOPMENT

Smoke Constituents Targeted in the Previous Method

aliph. Hydrocarbons						
1,3-butadiene	aliphatic hydrocarbon-isom.4	propylbenzene	methyl-fluorene	anethol	benzoic acid	isovaleramide-isom.2
isoprene	unsaturated hydrocarbon-isom.4	C3-alkyl.benzene-isom.1	phenanthrene	methylanisole-isom.1	2-furanecarboxolic acid	4-methyl-pentaneamide
1,3,5-heptatriene	unsaturated hydrocarbon-isom.5	C3-alkyl.benzene-isom.2	fluoranthene	alcoholes (without glycerine)	phenylacetic acid	picolinamide
1,3,5-cycloheptatriene	4,4'-dimethylbiphenyl	C3-alkyl.benzene-isom.3	pyrene	2-methoxyethanol	phenylpropionic acid	glykolamide
dimethyl-methylenecyclohexene	aliphatic hydrocarbon-isom.5	styrene	o-heterocycles	2-methoxypropanol	niacin	phenylacetamide
terpene (type:limonene)	unsaturated hydrocarbon-isom.6	C3-alkyl.benzene-isom.4	furan	propylenglykole	long-chained acids (>C8)	N-methyl-nicotinamide
2,7-dimethyl-1,6-octadiene	unsaturated hydrocarbon-isom.7	2-propenylbenzene	2,5-dimethylfuran	2-furanmethanol	caprylic acid (C8:0)	pyrrole-2-carboxamide
2,6-dimethyl-1,6-octadiene	aliphatic hydrocarbon-isom.6	C4-alkyl.benzene-isom.1	2,4-dimethylfuran	beta-citronellol	decanoic acid(C10:0)	niacinamide
p-menthene (hydrocarbon-terpene)	unsaturated hydrocarbon (ca. C25)	C4-alkyl.benzene-isom.2	2,3,5-trimethylfuran	benzylalcohol	dodecanoic acid(C12:0)	hexadecaneamide
1-undecene	unsaturated hydrocarbon-isom.8	C4-alkyl.benzene-isom.3	vinylfuran	beta-ionol	myristic acid (C14:0)	O-methyl-N-methylcarbamate
trimethylheptatriene	unsaturated hydrocarbon-isom.9	C4-alkyl.benzene-isom.4	2-vinyl-5-methylfuran	farnesol-s		
methyl-methyleneoctadiene	heptacosane (C27H56)	methyl-styrene-isom.1	benzofuran	glycerine-		
beta-myrcene	aliphatic hydrocarbon-isom.7	C3-alkyl.benzene-isom.6	1,3-dihydroisobenzofuran	glycerine-		
limonene	2-methyloctacosane(C29H60)	C4-alkyl.benzene-isom.5	2,3-dihydrobenzofuran	farnesol-is-		
dodecane	aliphatic hydrocarbon-isom.8	methyl-styrene-isom.2	steroids derivatives	geranylina		
dodecene	aliphatic hydrocarbon-isom.9	ethynylbenzene	stigmasteroacetate	cis-phytol		
2,7-dimethyl-1,3,6-octatriene	aliphatic hydrocarbon (ca. C30H62)	C4-alkyl.benzene-isom.6	beta-sitosterolacetate	trans-phyt		
tridecane	triacontane(C30H62)	propenylbenzene	esters + ethers (without triacetine)	farnesol s		
aliphatic hydrocarbon-isom.1	2-methyltriacontane(C31H64)	propenyltoluene-isom.1	3-butenoic acid methylester	isomentho		
1,3,8-p-menthatriene	unsaturated hydrocarbon-isom.11	propenyltoluene-isom.2	2-propenylformate	phytuber		
aliphatic hydrocarbon-isom.2	unsaturated hydrocarbon-isom.12	indene	isoamylbutyrate	sugars,		
aliphatic hydrocarbon-isom.3	hentricontane(C31H64)	methyl-indene-isom.1	glycoldiacetate	dianhydro		
copaene	aliphatic hydrocarbon-isom.10	methyl-indene-isom.2	geranylacetate	3,4-anhyd		
pentadecane	aliphatic hydrocarbon (C32H66)	methyl-indene-isom.3	benzylacetate	levoglucos		
1-pentadecene	aliphatic hydrocarbon-isom.11	dimethylindene-isom.1	dimethylsuccinate	short-ch. + ar. Acids	3-methylbutanenitrile	3-methyl-2-butenal
1-methylcyclooctene	2-methyltriactone(C33H68)	azulene	triacetine	acetic acid	crotonnitrile	2-hexenal
unsaturated hydrocarbon-isom.1	aliphatic hydrocarbon (C33H68)	dimethylindene-isom.2	dimethylsuccinate	formic acid	4-methyl-pentanenitrile	2,4-hexadienal
1-hexadecene	unsaturated hydrocarbon-isom.13	ethylindene	glycerin-diacetate	propionic acid	2-ethylideneamino-propionitrile	furfural
sesquiterpene(unsaturated KW)	beta-caryophyllene	naphthalene	ethylpalmitate	2-methyl-propionic acid	benzonitrile	1-(2-furanyl)-ethanone
heptadecane	alpha-caryophyllene	trimethylindene	methylestearate	butyric acid	tolunitrile-isom.1	benzaldehyde
farnesene	arom. Hydrocarbons	2-methylnaphthalene	6-acetoxy-2,7,11-cembratriene	acrylic acid	tolunitrile-isom.2	methylimidazole-carboxaldehyde
unsaturated hydrocarbon-isom.2	toluene	1-methylnaphthalene	methyl-pyrogulamate	2-/3-methylbutyric acid	nicotinonitrile	5-methyl-2-furfural
delta-cadidene	ethylbenzene	dimethylnaphthalene	benzyl benzoate	isocrotonic acid	benzylcnitrile	tolualdehyde
3,7,11,15-tetramethyl-2-hexadecene	xylene-isom.1	biphenyl	muskalactone	crotonic acid	amides	cinnamaldehyde
neophytadiene	xylene-isom.2	methyl-biphenyl	methyl-salicylate	3-methyl-valeric acid	acetamide	p-ethylbenzaldehyde
unsaturated hydrocarbon-isom.3	isopropylbenzene	trimethylnaphthalene	1,1-dimethyl-2-phenylethylbutyrate	capronic acid (C6:0)	propionamide	acetaldehyde
1,4-eicosadiene	xylene-isom.3	fluorene	triethylcitrate	2-methyl-crotonic acid	isovaleramide-isom.1	trimethoxy-acetophenone(Ga)
				sorbic acid	acrylamide	

Broad chemical diversity
categorized in chemical families
... but
smoke constituents were not
well defined

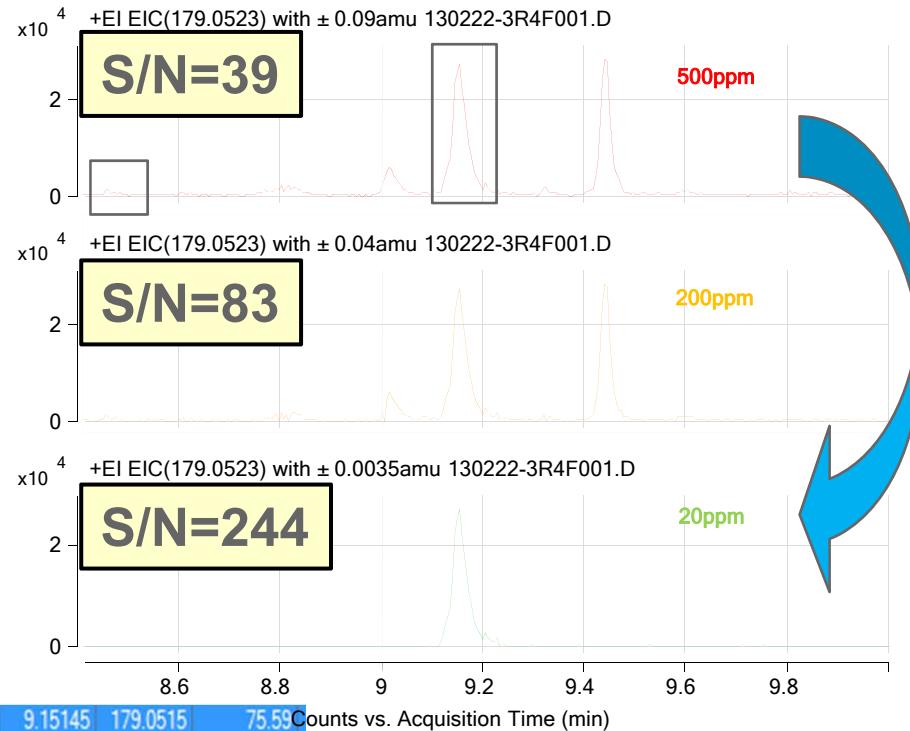
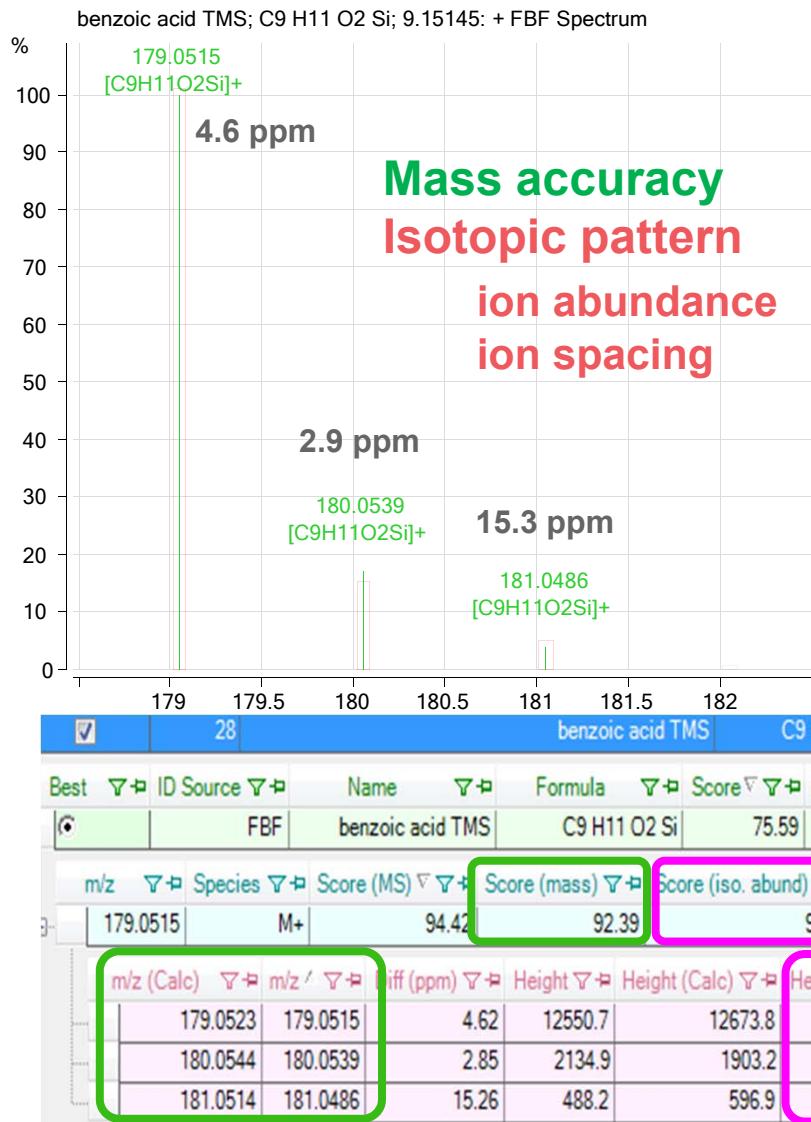
A time consuming curation of smoke constituents was required to register every compound as Unique Compound Spectral Database (UCSD)

Martin et al. Building an R&D chemical registration system, J. Cheminformatics 2012, 4:11.

Next Generation Fingerprinting: Practical Considerations

*FDA: Federal Register / Vol. 77, No. 64 / Tuesday, April 3, 2012 / Notices, 20034-20037

The Impact of use of High Resolution MS Instrument

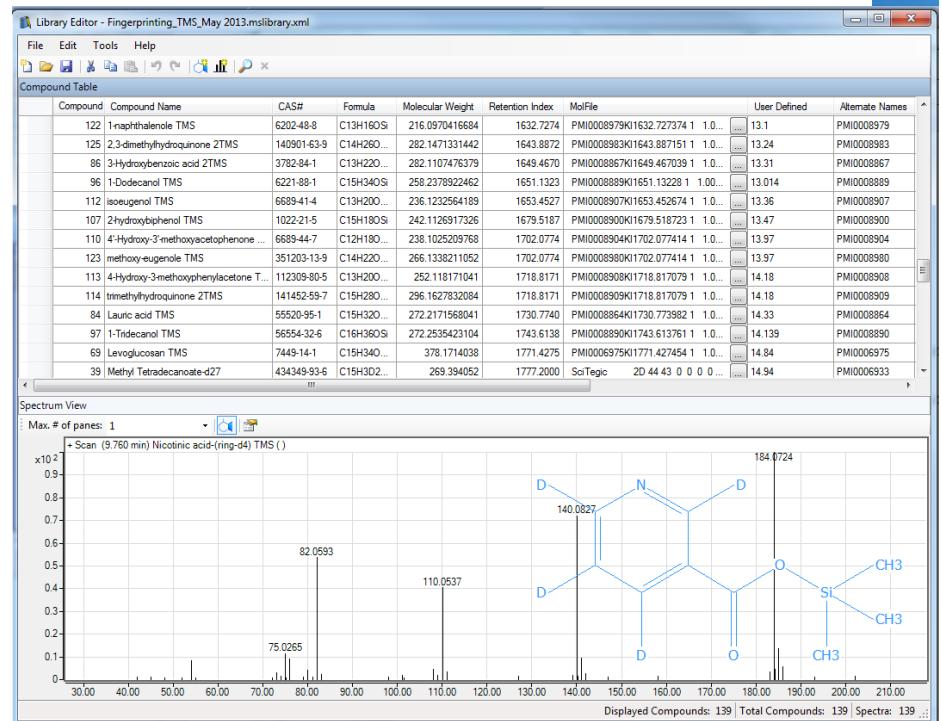


Sensitivity
Confidence level

Analysis of Reference Compounds

Reference standards ordered ($n \sim 600 + 63$ ISs)

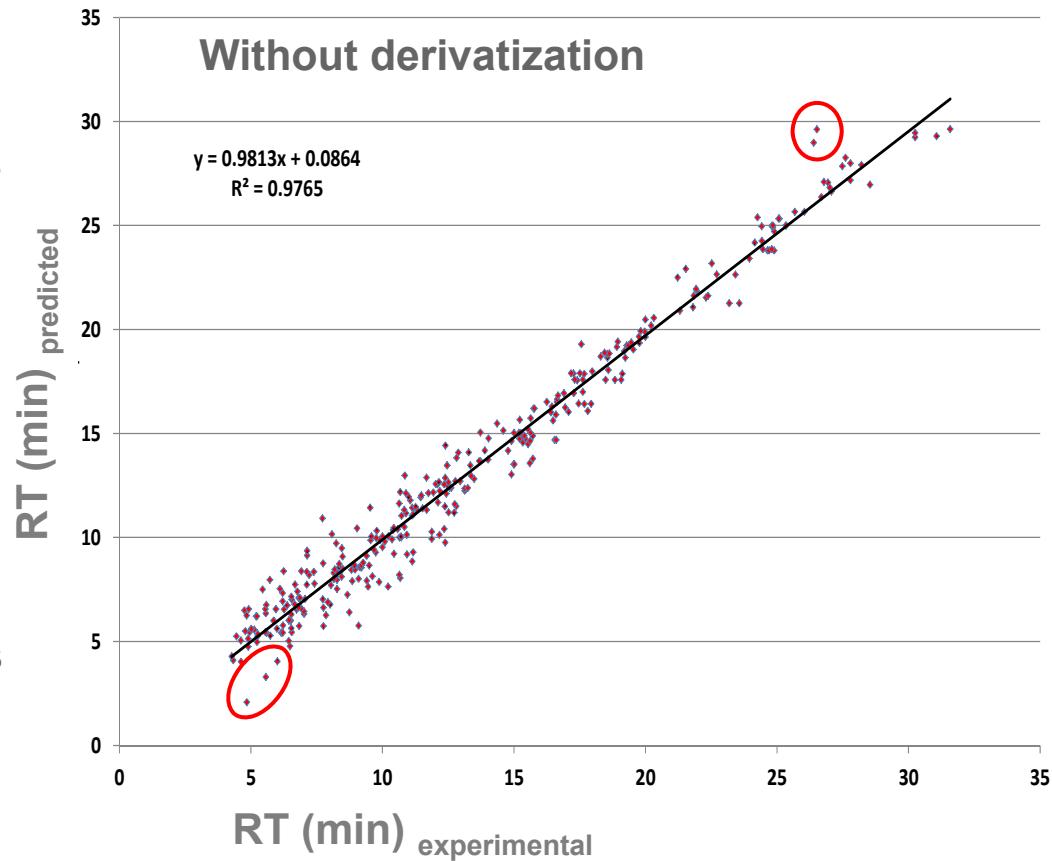
- Chemical registration within UCSD → unique PMI code
- GC-EI-HR-MS (DB-5) with silylation (BSTFA) → n=360 standards
- GC-EI-HR-MS (DB-5) without silylation → n=182 standards
- HS-GC-EI-HR-MS (DB-624) → n= 60 standards
- Assignment of EI HR mass spectrum – reference chemical
- Uploading of HR EI mass spectra in UCSD (unique SpecID)
- Building MS library specific to experimental conditions
 - Absolute Retention Time
 - Retention Index
 - EI mass spectrum
 - Mol file,...



Building Statistical Prediction Model of Retention Time from Reference compounds on a DB-5-MS GC column

QSPR (structure-based) modeling was used to predict RT of compounds (Dragon)

- Training set:
n=116 reference compounds
10 chemical descriptors
 $r^2 = 0.986$
- Test set:
n=56 reference compounds
 $r^2 = 0.979$
- Total external Validation set:
n=346 reference compounds
 $r^2 = 0.9765$
Median dev. = -8 sec
Min = -3.3 min
Max = +3.2 min



**Very good prediction model was obtained from a set of broadly diverse reference compounds.
Same model was used to predict over 900 components.**

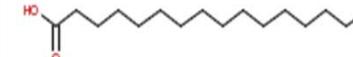
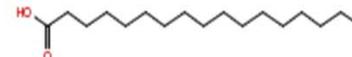
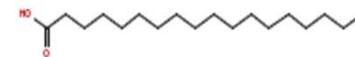
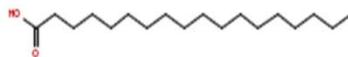
Clustering approaches to select appropriate internal standards

- Fingerprinting method should be high throughput
 - No calibration curve
 - Concentration estimated from an adequate IS (spiking addition)
- Clustering approach was used to select appropriate internal standards (Accelrys Pipeline Pilot 8.5)
- Adequate selection of IS is based on:
 - Similar physicochemical properties (i.e., logP, number of rotational bonds, polar surface area, volatility...)
- Smoke constituents were clustered into chemical families so that
 - At least one isotopically labeled IS (cluster center) was purchased for each cluster

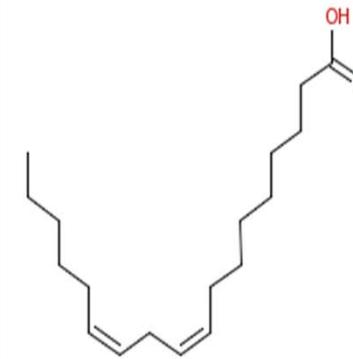
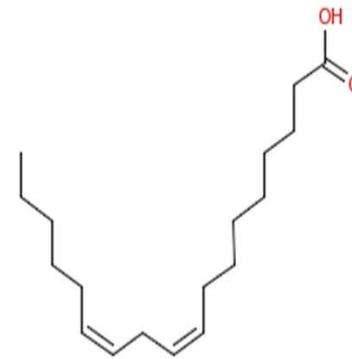
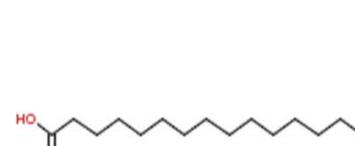
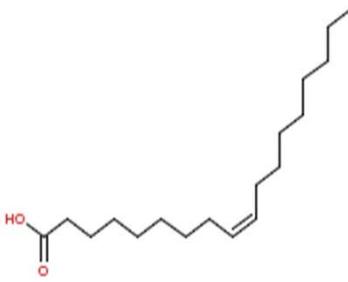


PMI RESEARCH & DEVELOPMENT

Example: Cluster of Long Chain Fatty Acids



Stearic acid-¹³C₁₈



Selection of isotopically labeled internal standard was made from Central
Euclidian distance value & commercial availability



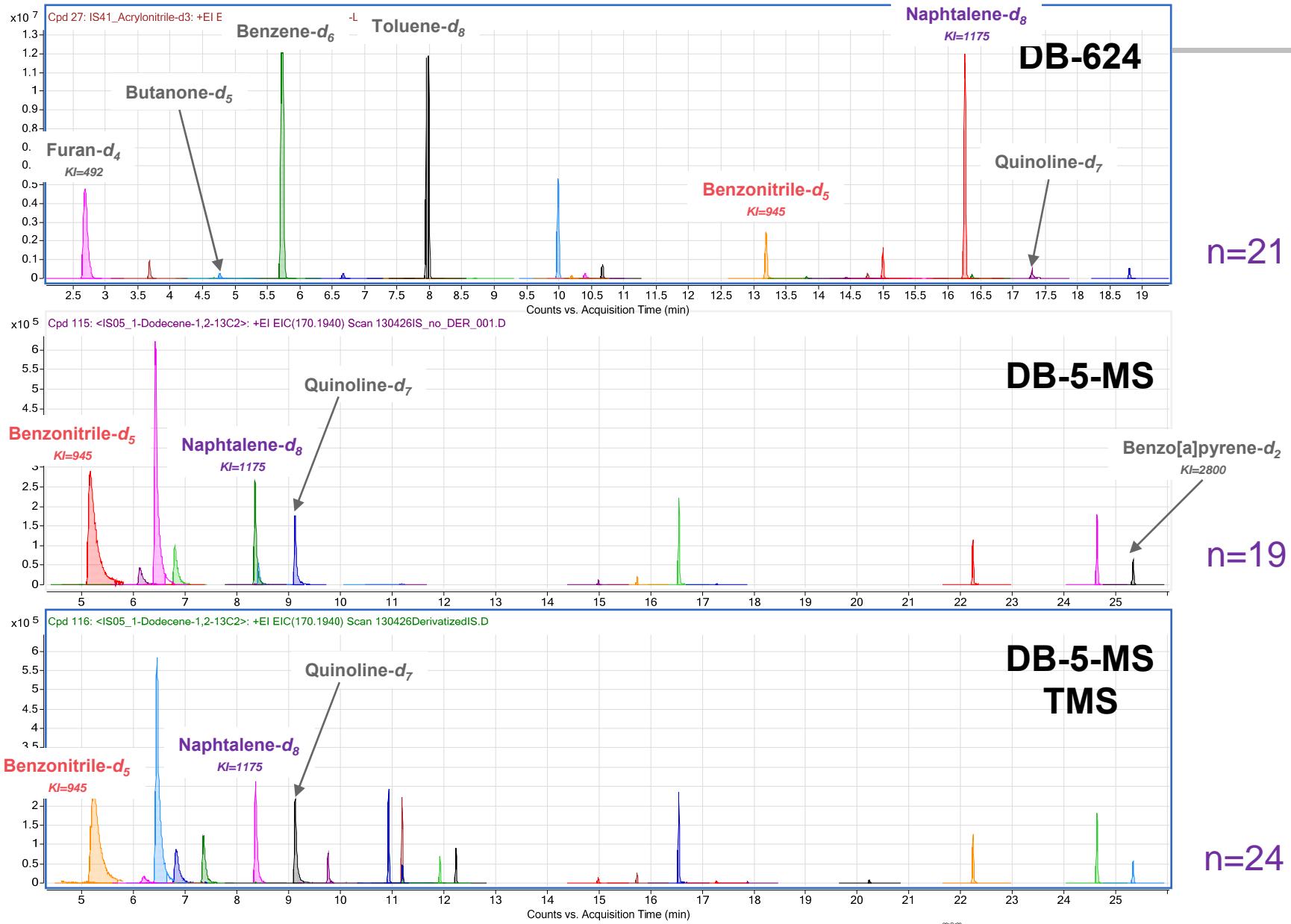
PMI RESEARCH & DEVELOPMENT

Clustering approach to improve semi-quantification

- Selection of 1 or 2 internal standards (IS) per cluster : **63** IS were purchased (from which **19** IS are present in the FDA list)
- Determination of semi-quantitative compound concentrations in relation to their respective internal standard
- Compensation for compound degradation issues (stability, partial silylation,...) was best by selecting adequate IS from the same cluster

A	B	C	D	E	F	G	H	I	J	K
	Name	Class	PMICODE	SMILES	MOLFORMULA	Cluster	DistanceToClosest	ClusterCenter	Predicted RT (min)	ClusterSize
1	1,4-eicosadiene	aliph. Hydrocarbons	PMI0000005	CCCCCCCCCC/C=C/CC=C	C20 H38	3	0.598632979	0	17.72	12
38	38 heptadecane	aliph. Hydrocarbons	PMI0000067	CCCCCCCCCC/C=C/CC=C	C17 H36	3	0.611239758	0	14.62	12
40	36 neonaphthaladiene	aliph. Hydrocarbons	PMI0000070	CC(C)CCCC(C)CCCC(C)CCC(-C)C-C	C20 H38	3	0.630277277	0	15.22	12
645	eicosanoic acid	long chain Fatty Acids (>=C15)	PMI0000329	CCCCCCCCCCCCCCCCC(=O)O	C20 H40 O2	5	0	5	20.9	8
197	stearic acid (C18:0)	long chain Fatty Acids (>=C15)	PMI0000166	CCCCCCCCCCCCCCCC(=O)O	C18 H36 O2	5	0.144818919	0	19.4	8
196	heptadecanoic acid (C17:0)	long chain Fatty Acids (>=C15)	PMI0000165	CCCCCCCCCCCCCCCC(=O)O	C17 H34 O2	5	0.217228378	0	18.04	8
195	palmitic acid (C16:0)	long chain Fatty Acids (>=C15)	PMI0000164	CCCCCCCCCCCCCCC(=O)O	C16 H32 O2	5	0.289637837	0	17	8
198	oleic acid (C18:1)	long chain Fatty Acids (>=C15)	PMI0000167	CCCCCCCC/C=C/CCCCCCC(=O)O	C18 H34 O2	5	0.316505227	0	19.16	8
194	pentadecanoic acid (C15:0)	long chain Fatty Acids (>=C15)	PMI0000163	CCCCCCCCCCCCCCC(=O)O	C15 H30 O2	5	0.362113201	0	15.9	8
199	linolic acid(C18:2)	long chain Fatty Acids (>=C15)	PMI0000168	CCCC/C=C/C=C/C/CCCCCCC(=O)O	C18 H32 O2	5	0.388127838	0	19.23	8
200	linolenic acid (C18:3)	long chain Fatty Acids (>=C15)	PMI0000169	CC/C=C/C=C/C/C=C/CCCCCCC(=O)O	C18 H30 O2	5	0.51606665	0	19.23	8
49	1171 butene	I arom. Hydrocarbons	PMI0000098	lc1cc2ccc3cccc4c3c2fc1)cc4	C16 H10	12	0	12	7.65	54

GC-EI-HR-MS of Internal Standards (n=63)



Good coverage of retention times

Outline

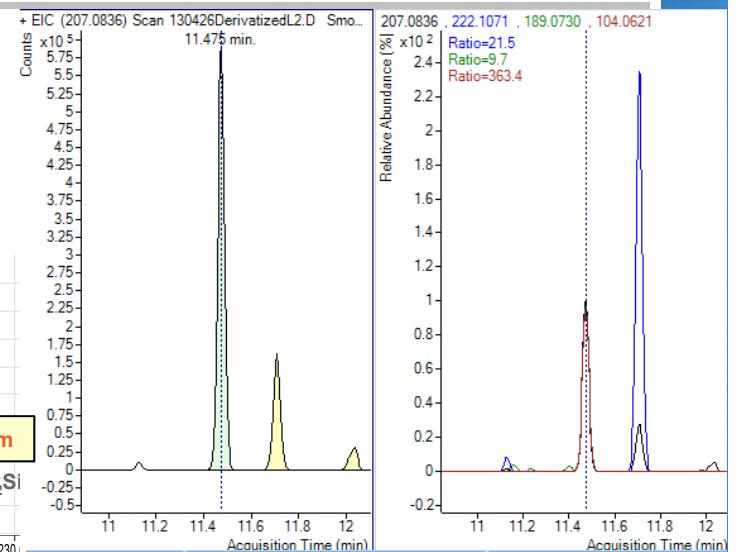
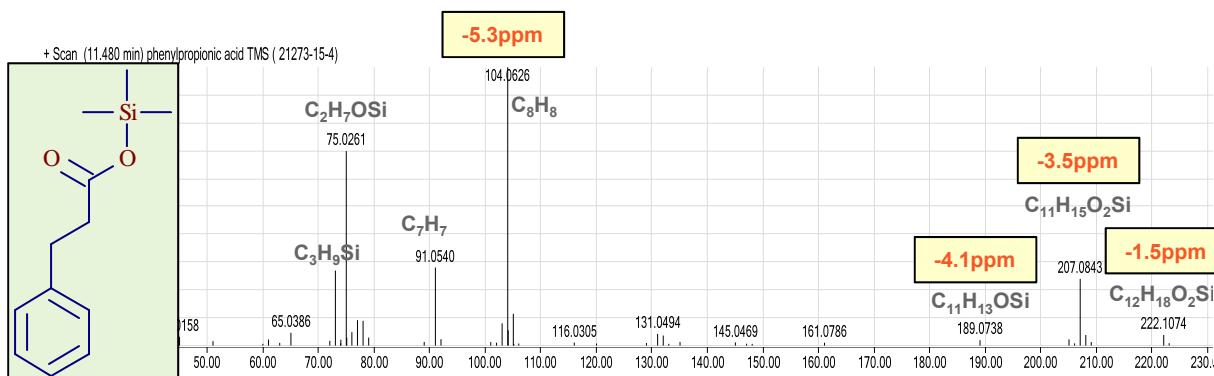
1. Description of the existing GC-MS fingerprinting method
 - Pros & cons of the current fingerprinting method
 - Why switching to a 7200 Agilent high resolution MS instrument?
2. How to tackle semi-quantification of the smoke constituents with the help of chemoinformatics tools
 - Curation step for smoke constituents to be monitored
 - Analysis of reference standards
 - Retention time prediction model (QSPR approach)
 - Selection of appropriate internal standards (clustering approach)
3. Semi-quantification from calibration curve of reference standards
 - Linearity
 - Assessment of silylation
4. What about other smoke constituents?
5. Conclusion



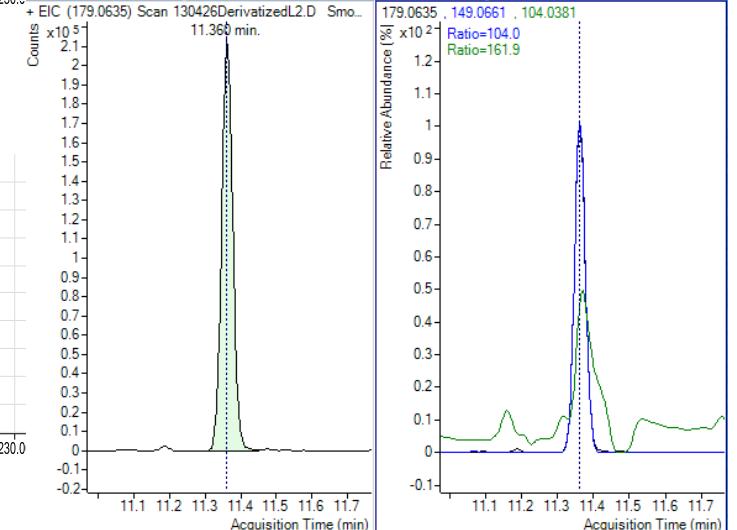
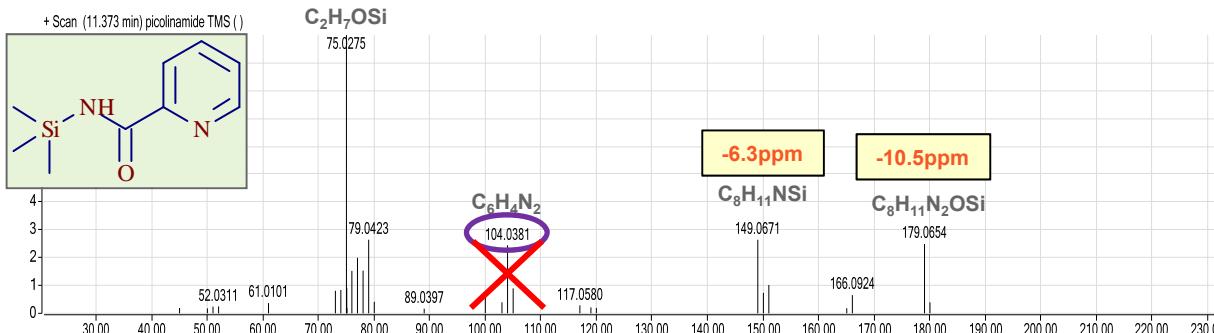
PMI RESEARCH & DEVELOPMENT

Quantitative Aspect: Importance of Qualifier ions

Phenylpropionic acid-TMS



Picolinamide-TMS



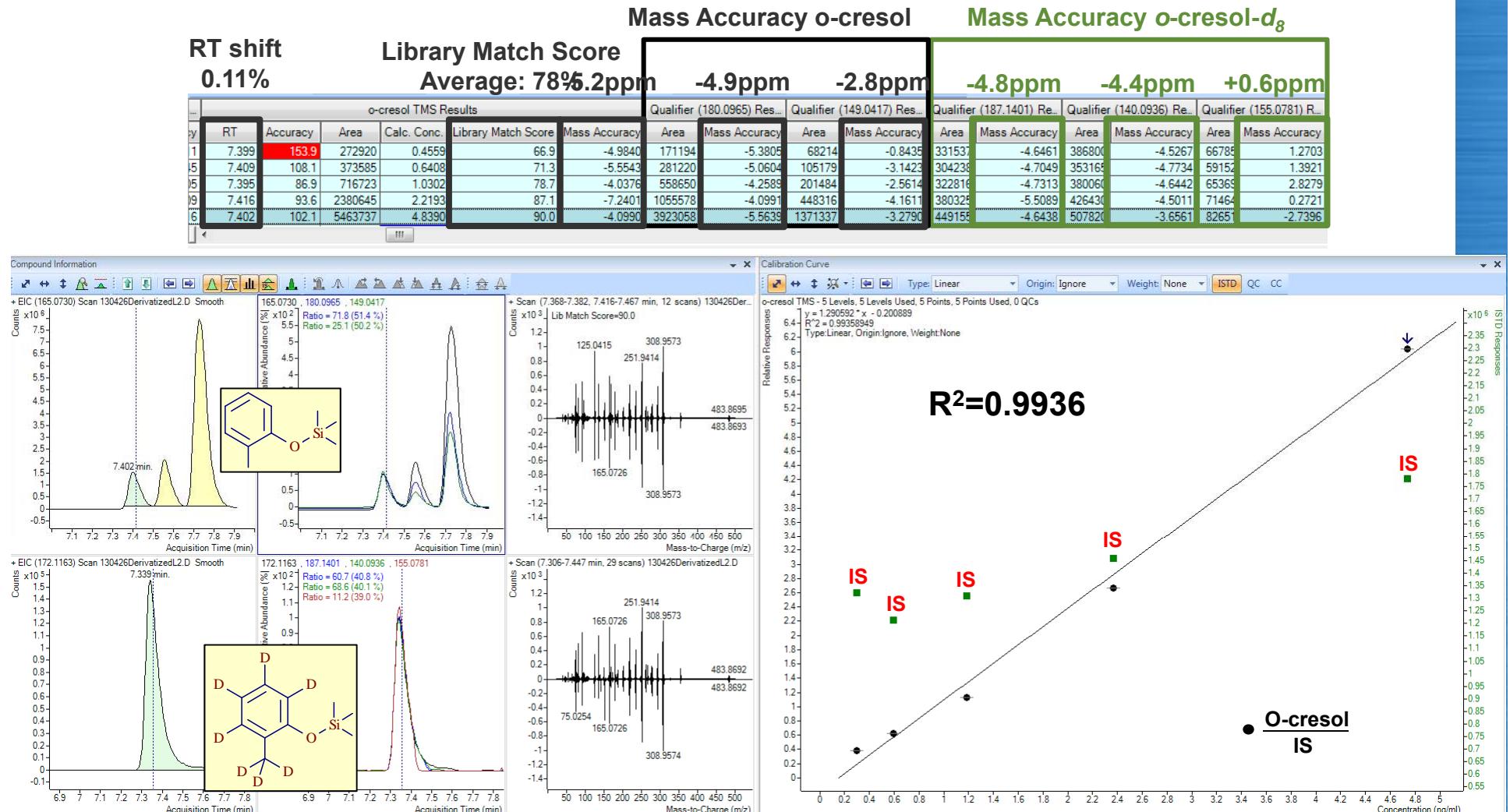
Ions highlighted in yellow were extracted for the semi-quantification



PMI RESEARCH & DEVELOPMENT

Case Study: Cresol Isomers

o-Cresol



Case Study: Cresol Isomers

m-Cresol

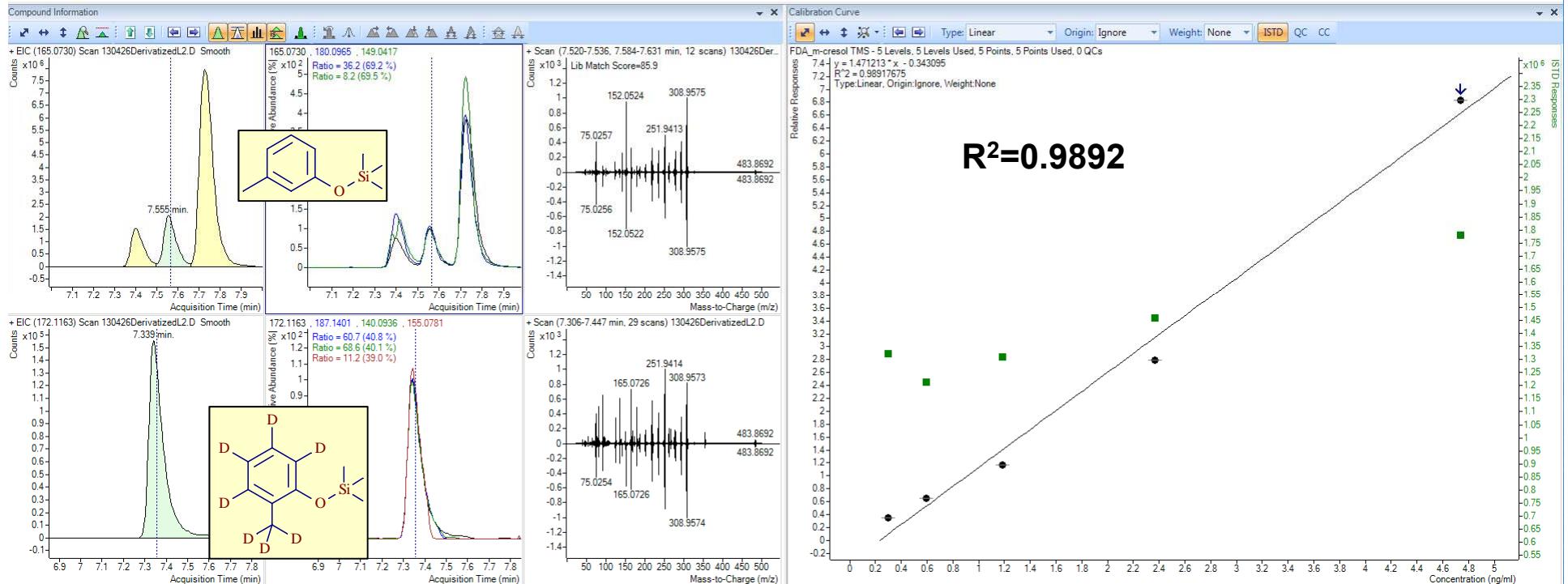
Mass Accuracy

RT shift Library Match Score
0.11% Average: 70%

m-Cresol
-4.6ppm -8.2ppm

o-Cresol-*d*₈

FDA_m-cresol TMS Results						Qualifier (180.0965) Re...		Qualifier (149.0417) Re...		Qualifier (187.1401) Re...		Qualifier (140.0936) Re...		Qualifier (155.0781) R...	
RT	Accuracy	Area	Calc. Conc.	Library Match Score	Mass Accuracy	Area	Mass Accuracy	Area	Mass Accuracy						
7.552	160.5	339547	0.4757	57.7	-5.2242	114394	-5.2844	17729		331537	-4.6461	386800	-4.5267	66785	1.2703
7.567	114.2	563434	0.6767	60.6	-4.7209	190705	-4.5788	37914		304238	-4.7049	353165	-4.7734	59152	1.3921
7.548	86.4	1079822	1.0236	70.1	-3.5166	368591	-4.3676	73235	14.7864	322816	-4.7313	380060	-4.6442	65369	2.8279
7.564	89.8	3032648	2.1291	76.5	-5.5059	832621	-4.2079	203001	5.0035	380325	-5.5089	426430	-4.5011	71464	0.2721
7.555	102.9	8427563	4.8800	85.9	-5.1211	3046810	-4.5390	693679	4.7681	449155	-4.6438	507820	-3.6561	82651	-2.7396



Very good chromatography separation
Reliable automatic peak integration

Case Study: Cresol Isomers

p-Cresol

Mass Accuracy

RT shift
0.11%

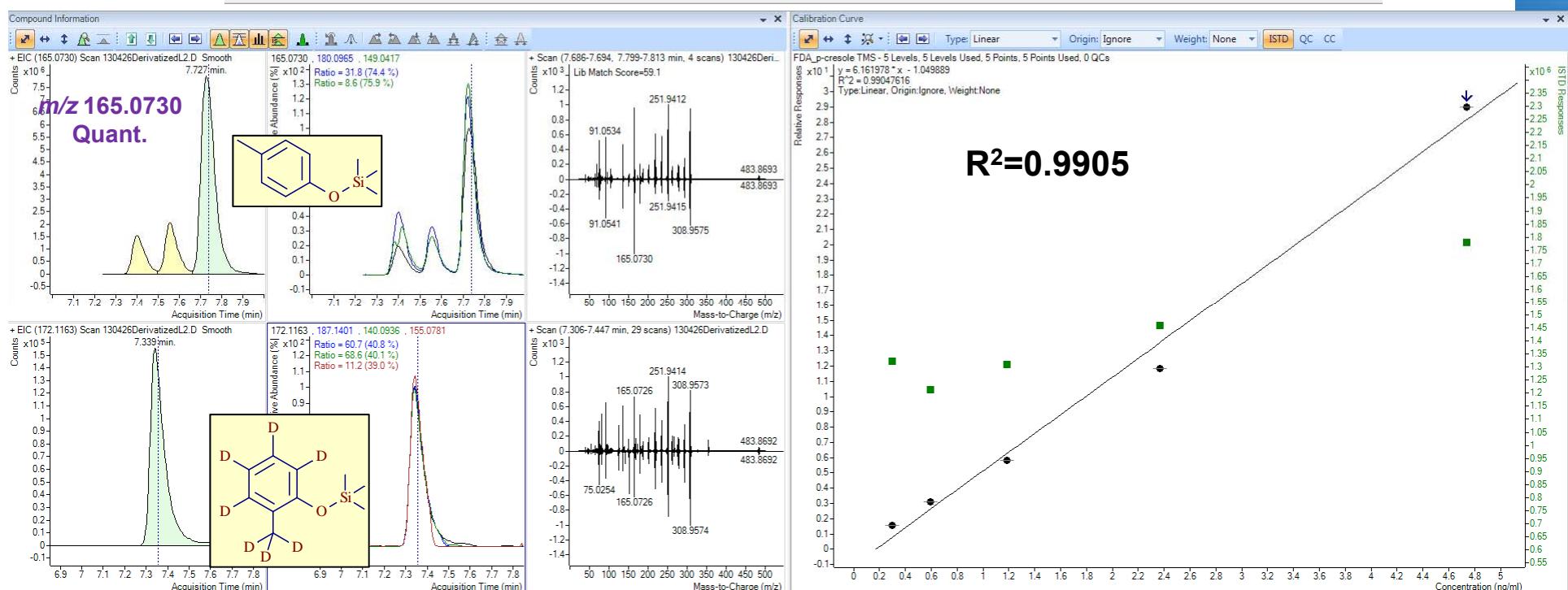
Library Match Score
Average: 57% 5.1ppm

p-cresol

-4.6ppm -3.2ppm

o-cresol-*d*₈

FDA_p-cresole TMS Results						Qualifier (180.0965) Results			Qualifier (149.0417) Results			Qualifier (187.1401) Results			Qualifier (140.0936) Results			Qualifier (155.0781) Results		
RT	Accuracy	Area	Calc. Conc.	Library Match Score	Mass Accuracy	Area	Mass Accuracy	Area	Mass Accuracy	Area	Mass Accuracy	Area	Mass Accuracy	Area	Mass Accuracy	Area	Mass Accuracy			
7.720	144.5	1598784	0.4281	50.9	-6.3749	393860	-5.1960	107200	-2.6070											
7.734	113.1	2905709	0.6702	56.1	-6.2962	657830	-4.5499	175185	-2.1377											
7.720	94.2	5715028	1.1162	60.7	-3.7175	1557736	-4.5404	353181	-3.7319											
7.737	88.2	12508184	2.0909	61.3	-5.0819	3921425	-4.3377	830498	-2.9077											
7.727	102.9	36785631	4.8798	59.1	-3.9370	11711031	-4.1458	3154194	-4.6177											



Very good chromatography separation
Reliable automatic peak integration using MassHunter software



PMI RESEARCH & DEVELOPMENT

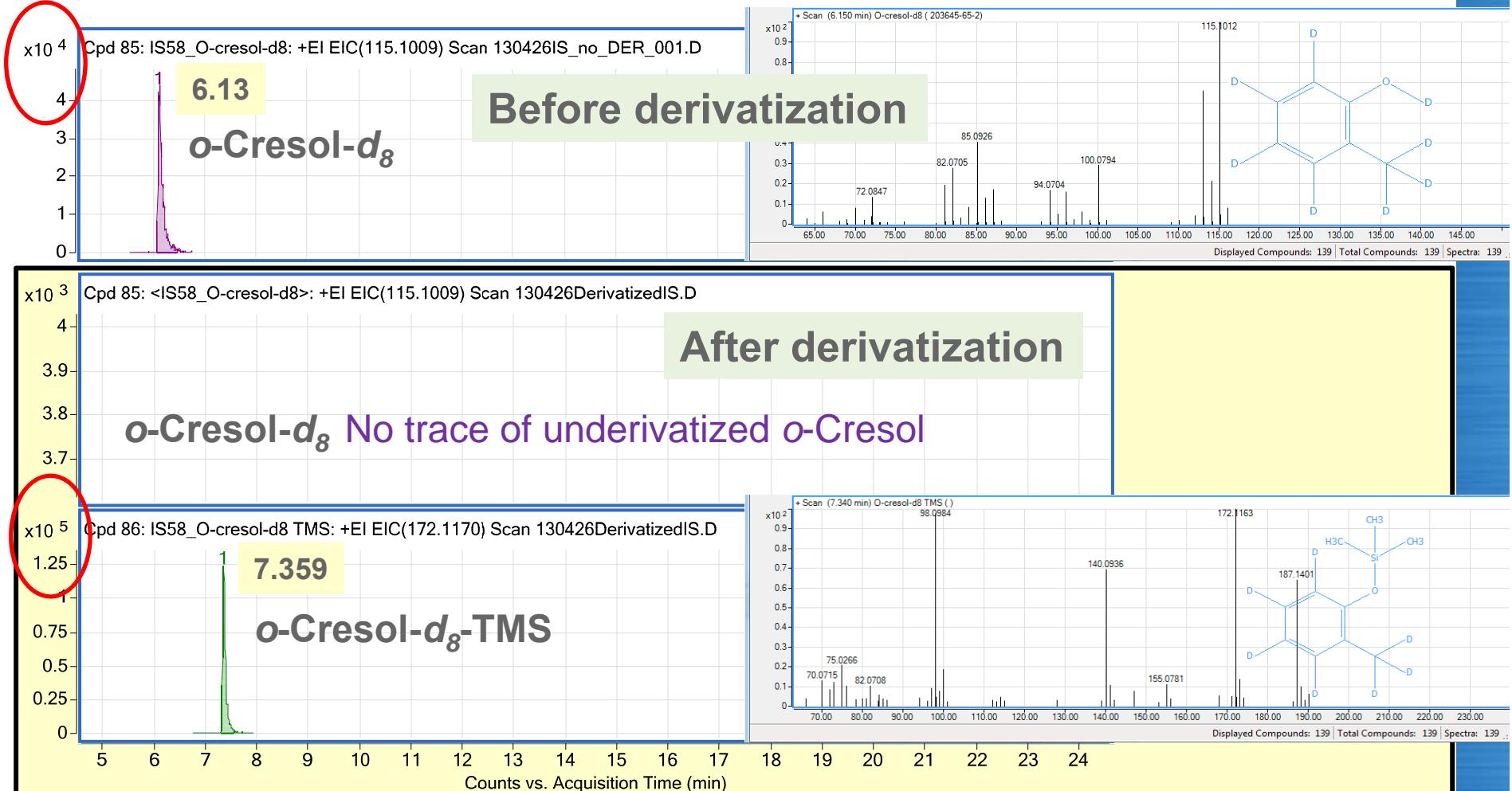
Key Consideration for Semi-Quantification

- Need to have a good chromatographic separation to resolve:
 - Isomers issue
 - Peak deconvolution
- Selection of characteristic for Quantifier & Qualifier ions
- Robustness due to usage of both Retention Times & Retention Index
- Assessment of linearity:
 - Using a 5-points calibration curves [0.29 – 4.8 ng/ μ L]
 - Analysis with and without silylation
 - Average of $R^2=0.9694$ ($n=71$ compounds)
 - $R^2 < 0.9$ only for 3 standards (due to sensitivity)



After silylation

Silylation Efficiency – example of o-cresol



Less fragmentation occurred after TMS derivatization

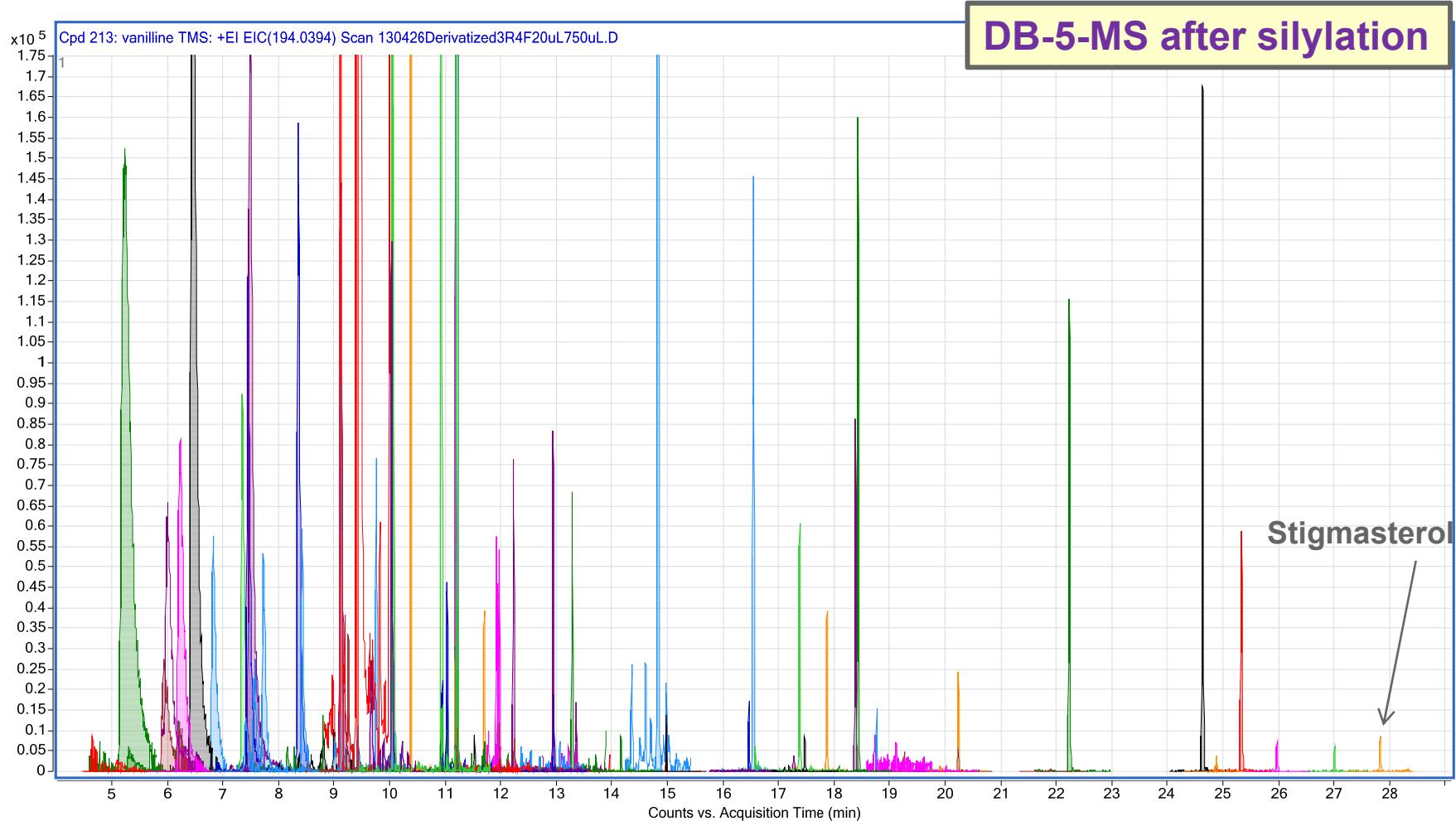


Higher sensitivity



PMI RESEARCH & DEVELOPMENT

Sample of reference cigarette (3R4F TPM) GC-EI-HR-MS after TMS derivatization



The csv file is used for an automatic and systematic search of smoke constituents

Fast data processing

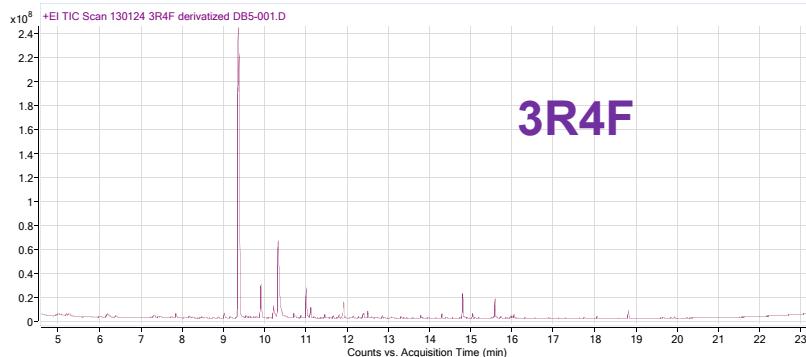


PMI RESEARCH & DEVELOPMENT

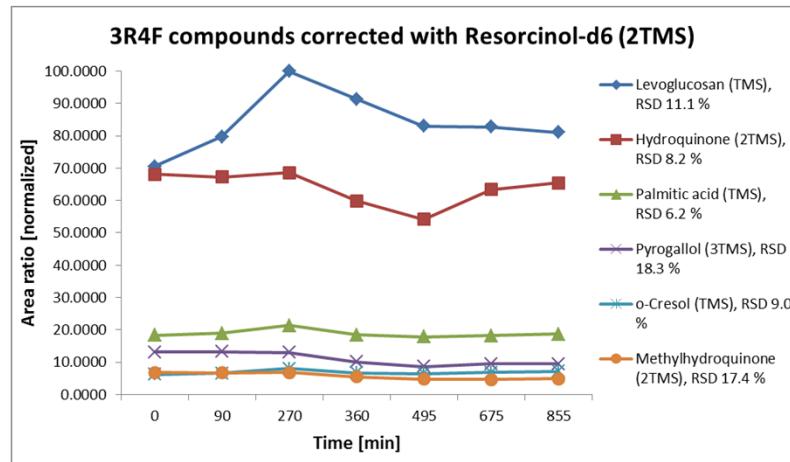
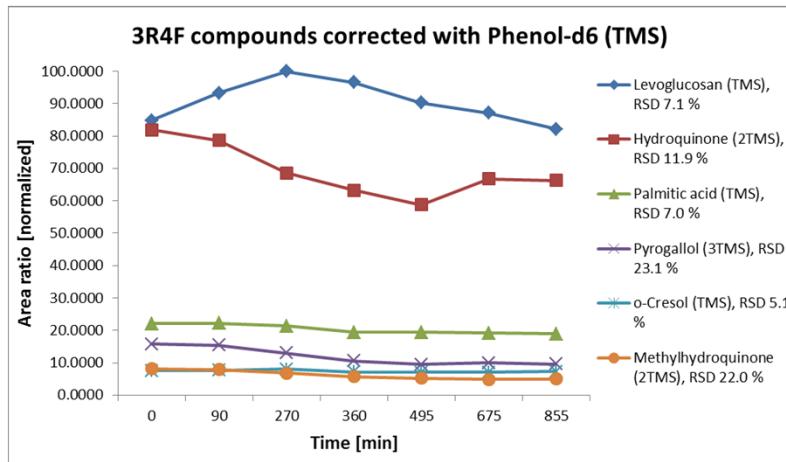
Sample Stability After Chemical Derivatization (Silylation)

10 μ L 3R4F TPM extract (ethanolic) evaporated under nitrogen

Add 100 μ L of IS mix + RI-markers
Add 800 μ L CH_2Cl_2 / Acetone 80:20 v/v
Add 50 μ L of BSTFA and 50 μ L of pyridine
Incubate @ 80°C for different time period



Samples stability was monitored over a period of approximately 15 hours.
The analyzed smoke constituents do not show any significant differences.



**After derivatization, aerosol components are stable over time
Appropriate selection of IS needed**



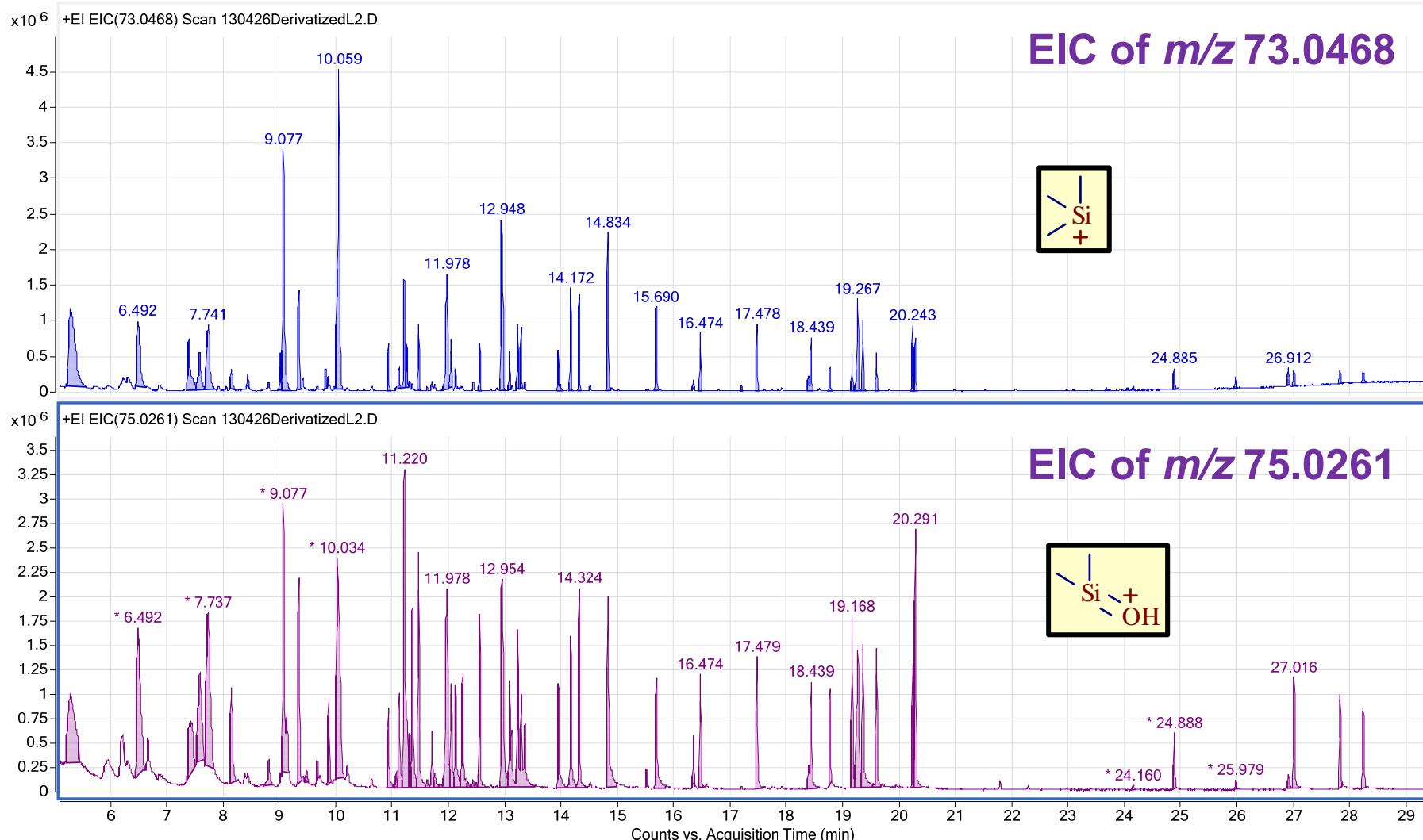
PMI RESEARCH & DEVELOPMENT

What about other Smoke Constituents?

Updating the Fingerprinting Method with New Compounds



Extraction of Silylated compounds can be of great help to identify additional smoke constituents

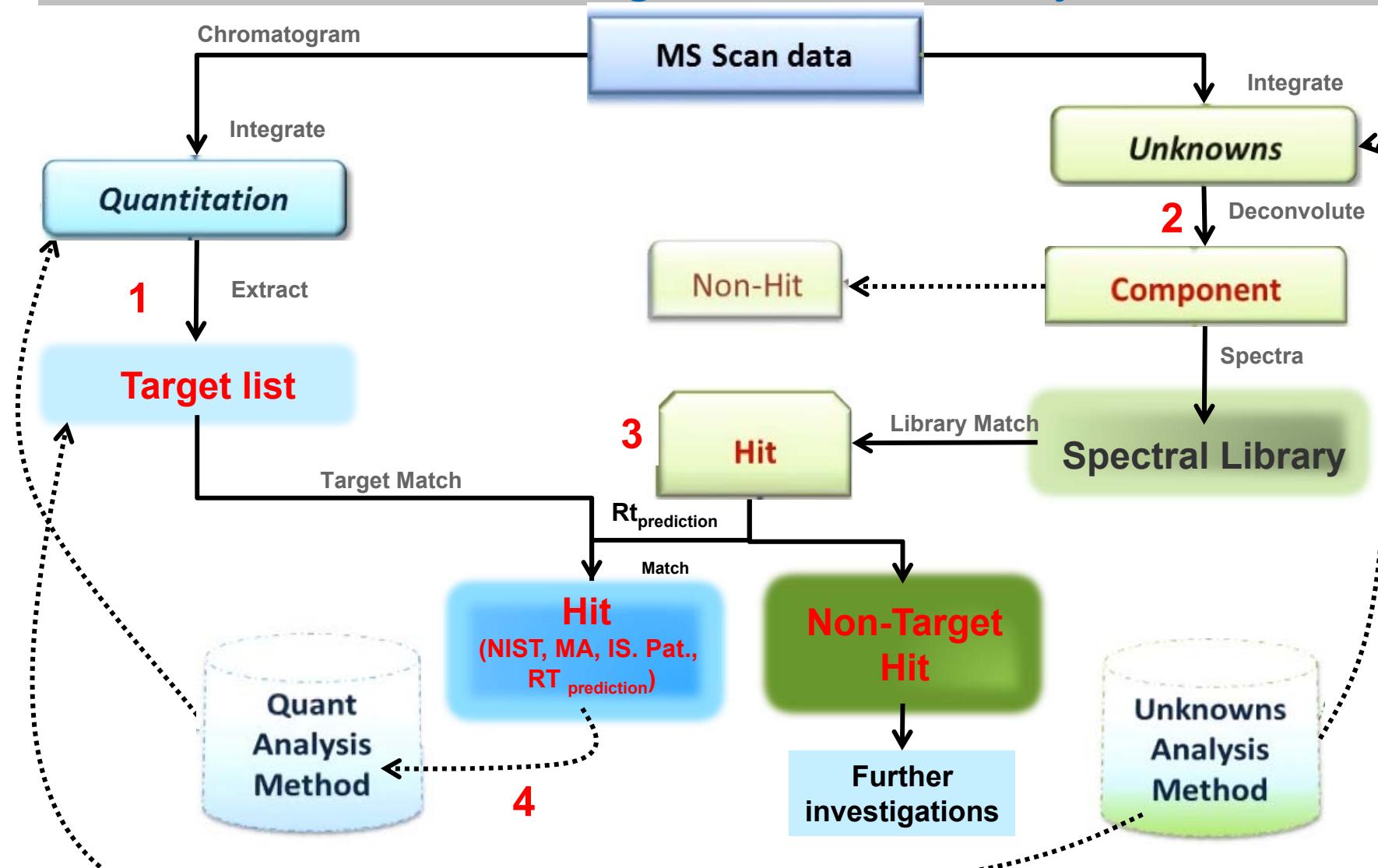


Mass accuracy AND predicted RT are of great interest to postulate new compounds



PMI RESEARCH & DEVELOPMENT

Analytical Workflow for Unknowns: Advantage of Full Scan / High Resolution analysis



Conclusion

- A GC-EI-HR-MS method has been established to semi-quantify smoke constituents in aerosol and aerosol fractions.
- The method has been simplified with:
 - ✓ Two GC columns instead of three,
 - ✓ Removal of Solid Phase MicroExtraction (SPME)
- We have demonstrated the great advantage of high resolution _ mass accuracy:
 - ✓ Improved S/N
 - ✓ Specificity and selectivity
 - ✓ Enhanced sensitivity
- Improved semi-quantification with:
 - ✓ Clustering approach to select adequate internal standards (n=63)
 - ✓ Compensate for compound degradation issues
 - ✓ Very good automatic peak integration (MassHunter software)
 - ✓ Very good linearity: averaged $R^2 = 0.9694$ (n=71)
- Monitoring of compounds can be expanded thanks to:
 - ✓ Full scan analysis at high mass accuracy (< 10 ppm)
 - ✓ Very good prediction model of retention time (median $RT_{shift} = 8$ sec, n=346)



PMI RESEARCH & DEVELOPMENT

Acknowledgements



Eric Dossin



Philippe Guy



Arno Knorr



Martin Almstetter



Mark Bentley



Pavel Pospisil



Elyette Martin



Aurelien Monge



Audrey Debrick



Stefania Della Gatta



Quentin Dutertre



Daniel Arndt

Thank you for your attention.



PMI RESEARCH & DEVELOPMENT