Accelerating the GC/MS Workflow for Non-Target Analysis

Don Kuehl[1], Stacey Simonoff[1], Yongdong Wang[1] and Ben Johnson[2]

[1]Cerno Bioscience
Las Vegas, NV USA

[2]NSF International
Ann Arbor, MI

Testing Drinking Water Equipment

Drinking Water Chemicals(NSF/ANSI 60): Chemicals that may be added to potable water for the purposes of coagulation and flocculation, softening, precipitation, sequestering, pH adjustment, corrosion/scale, and disinfection and oxidation chemicals, as well as miscellaneous treatment and water supply chemicals.

Drinking Water Equipment (NSF/ANSI 61): Anything that delivers a drinking water supply, products such as faucets, pipes, and pumps filtration media. Including water filtration systems to reduce contaminants such as PFOA and PFOS.

Samples are analyzed by GC/MS for the identification and quantitation of target compounds by a modified version EPA 625. Non-target compounds are tentatively identified and quantitated using a single-point calibration against internal standards.

The identification and quantification of the semi-volatile Tentatively Identified Compounds (TICs) was a serious bottleneck in the testing process and will be the focus of this presentation.

The Problem

- 7-8000 samples evaluated annually by a team of 3-4 scientists
- The established work-flow was tedious, with analysts manually combing through spectra and chromatograms, evaluating every chromatographic peak individually. One by one
- “Dirty” samples could sometimes take hours or even days to process with hundreds of unknown peaks
- Reviewing complex matrices for co-elutions or other chromatographic interferences required intensively trained analysts and lengthy data review
- Once reduced, quantitation was also tedious and manual
- Report generation and LIMS integration was done with a complex excel macro that was difficult to maintain
- The sample backlog could get up to 6 months at times

The Challenges

- Average sample processing time estimated at 30-40 minutes
- Limited staff available
- Years of highly skilled training required
- Difficult and costly to replace staff
- Process could not be implemented in other labs/regions
- Assignments often required judgement calls based on skill and experience

Requirements

- Shorten analysis time
- Fully automate the peak mixture deconvolution and quantitation
- Provide a vendor-neutral platform
- Standardize analysis across different labs and instrumentation
- Reduce the expertise required for analysis
- Improve the confidence in compound identification
- Automate reporting and LIMS system integration

Solution

NSF chose Cerno's[1] GC/ID software and contracted NSF specific customizations

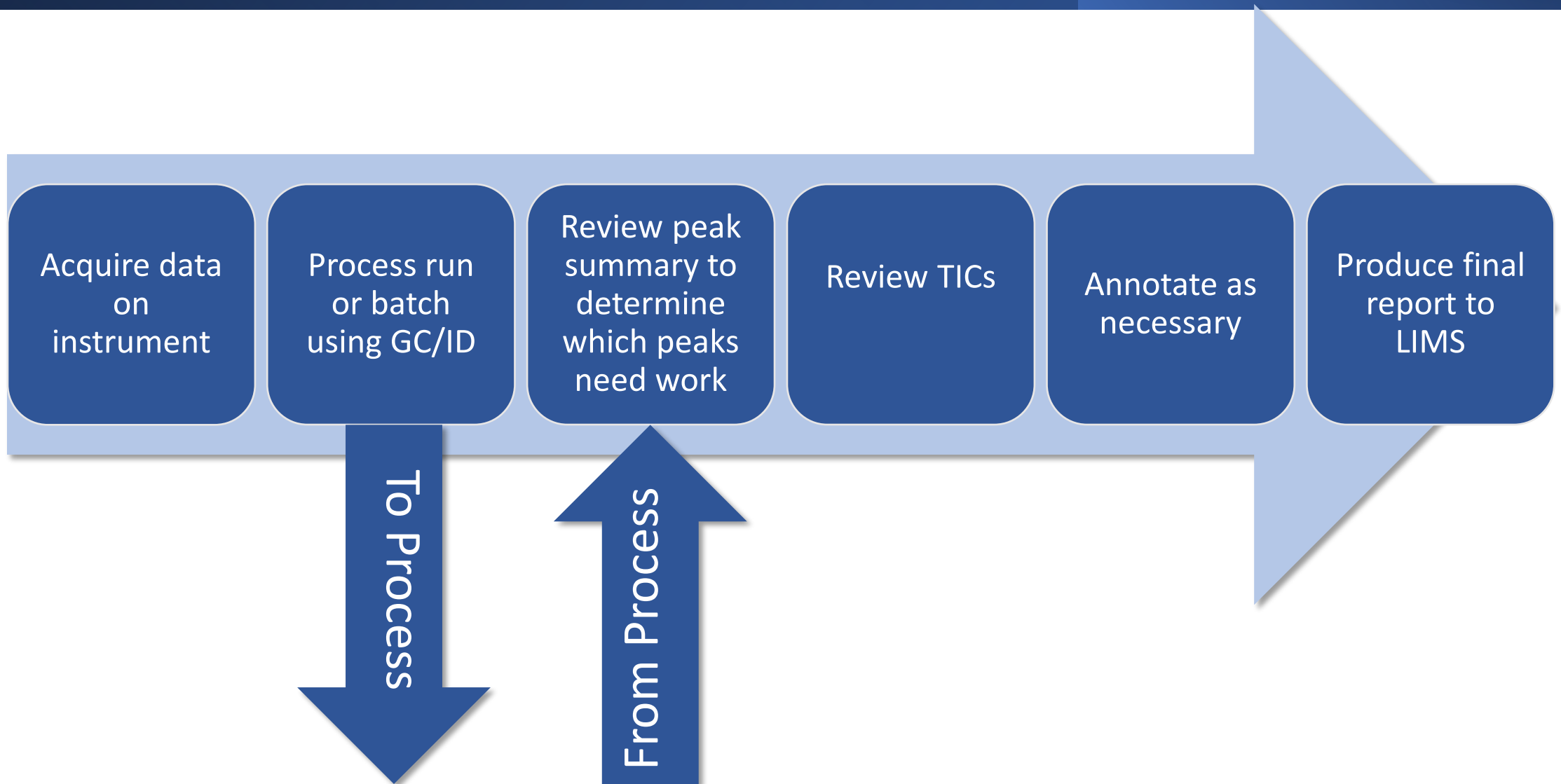
- Is vendor neutral, works with all commercial GC/MS systems
- Fully automates peak deconvolution and quantitation
- Provides superior identification confidence by combining library search, retention index (RI), and accurate mass validation in an automated full run analysis (composite scoring)
- Modern Python scripting allowed for fast and easy NSF specified customizations to be delivered ahead of schedule

1. Cerno Bioscience, Los Vegas, Nevada; cernobioscience.com

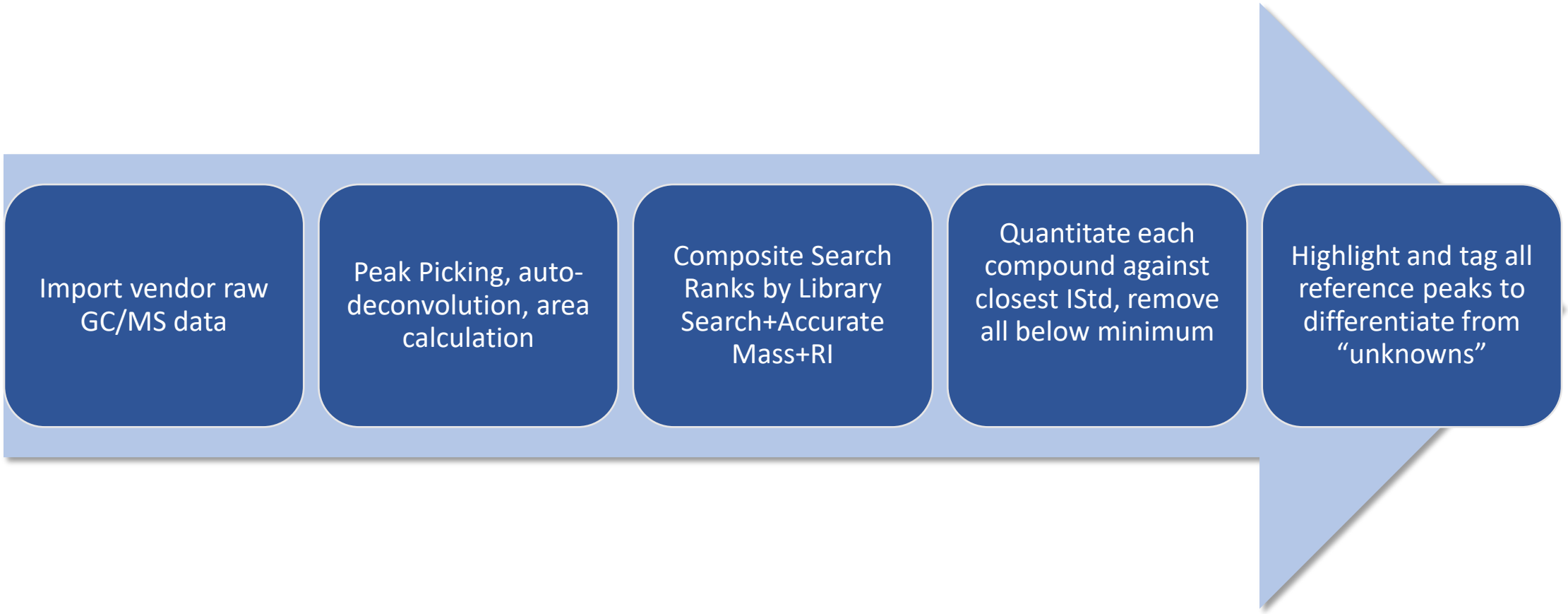
Customizations

- Identify internal standards, surrogates and target compounds so the analyst doesn’t have to spend time evaluating them
- Quantitate unknown compounds based on their response factor relative to the nearest internal standard
- Filter compounds below an established reporting limit so they are not evaluated
- Allow reviewer to easily choose compound “class” for uncertain absolute ID’s

Workflow | Interpretation of the Data



Workflow | Processing of Data



“Composite” Scoring Reduces Ambiguity, Speeds Review

How do we Know?

Because it weighs *MW, Reverse Library Match Score, Forward Library Match Score, Retention Index*

Composite Scoring

Conventional Library results are not definitive!

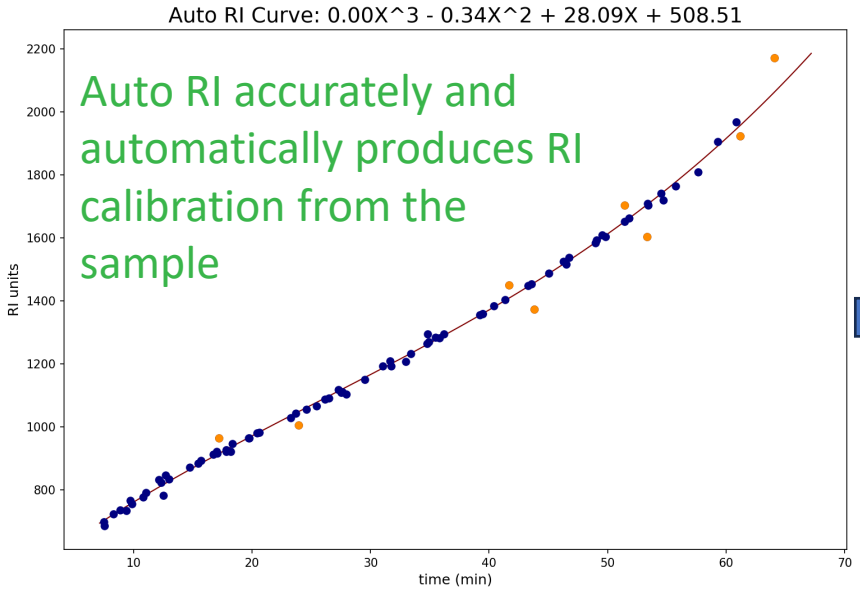
Hit	Match	NIST	RevS	SA-M	SA-F	RIFit	MW	Formula	Name
1	96.96	932	932	99.28	99.18	94.62	95.953	C2H2Cl2	Ethene, 1,2-dichloro-, (E)-
2	92.88	932	932	99.28	99.18	69.38	95.953	C2H2Cl2	1,2-Dichloroethylene
3	88.36	968	968	99.28	99.18	54.38	95.953	C2H2Cl2	Ethene, 1,2-dichloro-, (Z)-
4	86.96	867	876	99.28	99.18	57.62	95.953	C2H2Cl2	Ethene, 1,1-dichloro-
5	76.6	422	423	98.45	99.18	94.62	97.969	C2H4Cl2	Ethane, 1,1-dichloro-
6	64.5	642	644	97.24	99.18	0	95.98	C2H5ClS	Chloromethylmethyl sulfide
7	50.68	787	806	26.28	99.18	0	153.959	C4H4Cl2O2	Dihydro-4,5-dichloro-2(3H)-fu
8	41.97	479	512	14.48	85.65	82.38	93.938	C2Cl2	Ethyne, dichloro-
9	37.42	661	662	-29.09	99.18	0	233.885	C4H5BrCl2O2	Methyl 3-bromo-2,2-dichloro
10	37.09	487	777	-25.77	99.18	0	123.985	C4H6Cl2	Cyclobutane, 1,1-dichloro-
11	34.67	599	600	-26.37	99.18	0	189.936	C4H5Cl3O2	Methyl 2,3,3-trichloropropan

What’s the right answer?

Wrong! It’s this one

Are you sure?

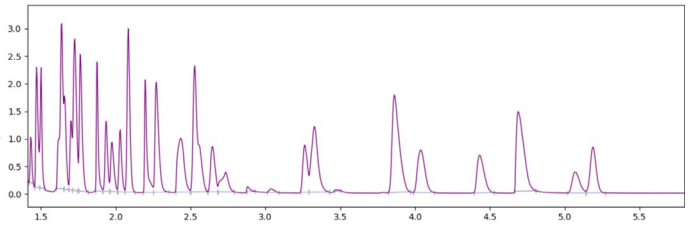
GC Retention Index Calibraton

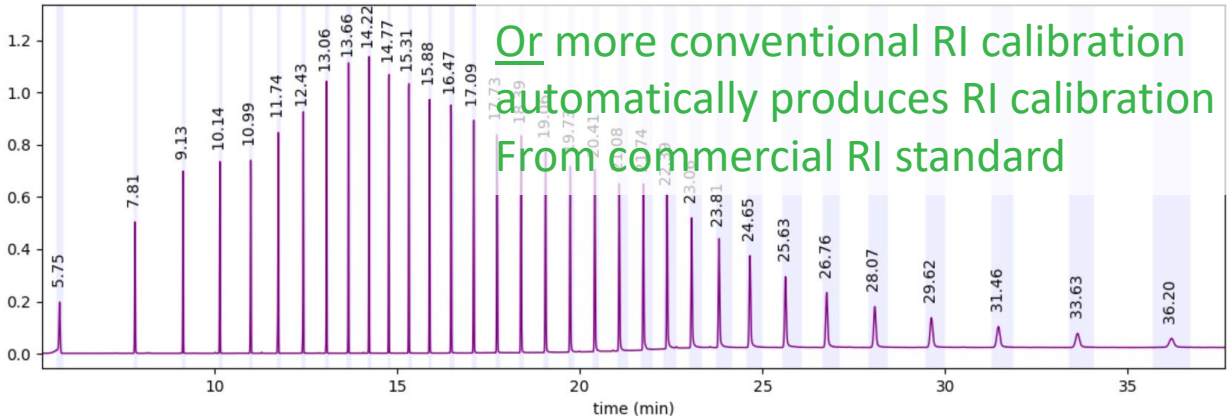


Auto RI accurately and automatically produces RI calibration from the sample

Calibrated GC

Fully RI calibrated chromatogram





Or more conventional RI calibration automatically produces RI calibration From commercial RI standard

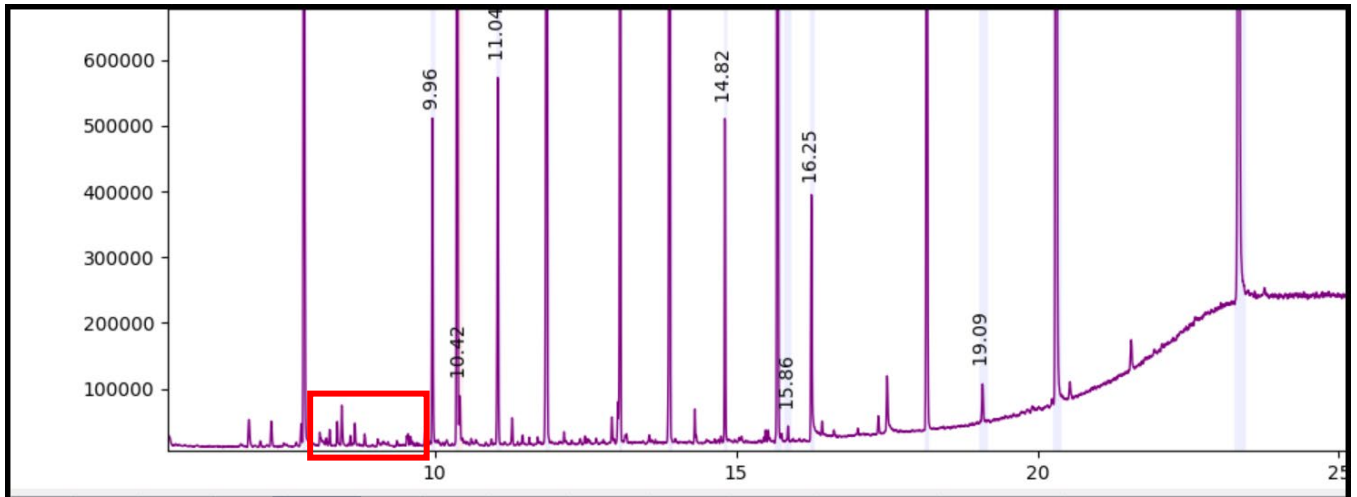
New NIST database with “full” RI coverage

RI match for every compound for superior identification

Hit	Match	NIST	RIFit	Name
1	96.96	932	94.62	Ethylene, 1,2-dichloro-, (E)-
2	92.88	932	69.38	1,2-Dichloroethylene
3	88.36	968	54.38	Ethylene, 1,2-dichloro-, (Z)-
4	86.96	867	57.62	Ethene, 1,1-dichloro-

Typical “Clean” Run

Python scripting enabled cusome, optimized workflow



Peak	Time	Quant	Mix	Match	NIST	RevS	SA-M	SA-F	RI-Fit	MW	CAS	Formula	Name	
1	7.83	28.54	1/1	99.16	833	834	97.28	99.16	81.1	100.089	62238-37-3	C6H12O	3-Penten-1-ol, 2-methyl-	←Unk
2	9.96	7.90	1/1	67.55	876	882	60.36	0.00	98.9	100.080	13127-88-3	C6D6O	S:Phenol-d6-	
3	10.37	40.61	1/1	98.73	961	962	98.73	0.00	82.8	149.994	3855-82-1	C6D4Cl2	I:1,4-Dichlorobenzene-D4	
4	10.42	0.63	2/2	98.60	919	929	-25.65	98.60	97.5	130.136	104-76-7	C8H18O	T:1-Hexanol, 2-ethyl-	
5	11.04	9.72	1/1	98.36	896	901	98.36	0.00	97.5	128.063	4165-60-0	C6D5NO2	S:Nitrobenzene-D5	
6	11.86	40.00	1/1	96.93	921	922	96.93	0.00	72.8	136.113	1146-65-2	C10D8	I:Naphthalene-D8	
7	13.08	14.09	1/1	99.40	960	960	99.40	98.49	98.0	172.069	321-60-8	C12H9F	S:1,1'-Biphenyl, 2-fluoro-	
8	13.89	40.00	1/1	95.73	909	909	88.20	0.00	96.4	164.141	15067-26-2	C12D10	I:Acenaphthene-d10	
9	14.82	4.44	1/1	98.14	955	955	98.14	97.42	96.0	327.773	118-79-6	C6H3Br3O	S:Phenol, 2,4,6-tribromo-	
10	15.69	40.00	1/1+1	92.98	890	890	92.98	0.00	94.6	188.141	1517-22-2	C14D10	I:Phenanthrene-D10	
11	15.86	0.21	1/1+1	99.09	875	897	-8.29	99.09	73.4	278.152	84-69-5	C16H22O4	T:1,2-Benzenedicarboxylic	
12	16.25	3.92	1/1	98.73	927	927	98.23	98.73	94.0	256.240	57-10-3	C16H32O2	n-Hexadecanoic acid	←Unk
13	18.16	19.53	1/1	97.70	913	913	92.65	0.00	94.9	244.197	1718-51-0	C18D14	S:p-Terphenyl-d14	
14	19.09	0.79	1/1+1	99.04	890	890	-78.89	99.04	94.6	312.136	85-68-7	C19H20O4	T:Butyl benzyl phthalate	
15	20.31	40.00	1/1+2	94.03	895	898	91.50	0.00	91.6	240.169	1719-03-5	C18D12	I:Chrysene-D12	
16	23.34	40.00	1/1+1	91.63	873	882	91.63	0.00	90.2	264.169	1520-96-3	C20D12	I:Perylene-D12	

Review time for “unknowns” from hours to minutes for “clean” runs and days to hours for “dirty” runs!

- Batch processes at <4sec/run (avg, multi-threaded)
- Confidently identifies all peaks (composite “Match” score)
- Accurately quantitates all peaks in run against closest IStd even with co-elution
- Filters out all peaks below quant minimum threshold
- Highlights and visually tags Targets, IStd, Surrogates to simplify user review
- Clearly indicates “Unknowns” that need attention
- Creates custom report for LIMS system (both parsable text and hyperlinked PDF)

Summary

- Fully automated tedious manual processing method
- Eliminated 6 month sample backlog, samples analyzed as the are run
- Reduced review time for “clean” samples from hours to minutes
- Reduced review time for “dirty” samples from days to hours
- Produces comprehensive custom report for LIMS integration and customer reports

Cerno Bioscience Booth #2922