Eawag:
Swiss Federal Institute of Aquatic Science and Technology

Creating Reliable Data – a Challenge for Non-target Screening

Juliane Hollender[1,2], Karin Kiefer[1,2], Martin Krauss[3], Jennifer Schollée[1], Emma Schymanski[1,4], Heinz Singer[1], Jaroslav Slobodnik[5], Nikolaos Thomaidis[6]

[1]Eawag, Swiss Federal Institute of Aquatic Science and Technology, Switzerland
[2*]Institute of Biogeochemistry and Pollutant Dynamics, ETH Zurich, [3]UFZ, Helmholtz Centre for Environmental Research, Germany,[4]Luxembourg Centre for Systems Biomedicine (LCSB), Univ. of Luxembourg, [5]Environmental Institute, Slovakia, [6]Univ. of Athens, Greece

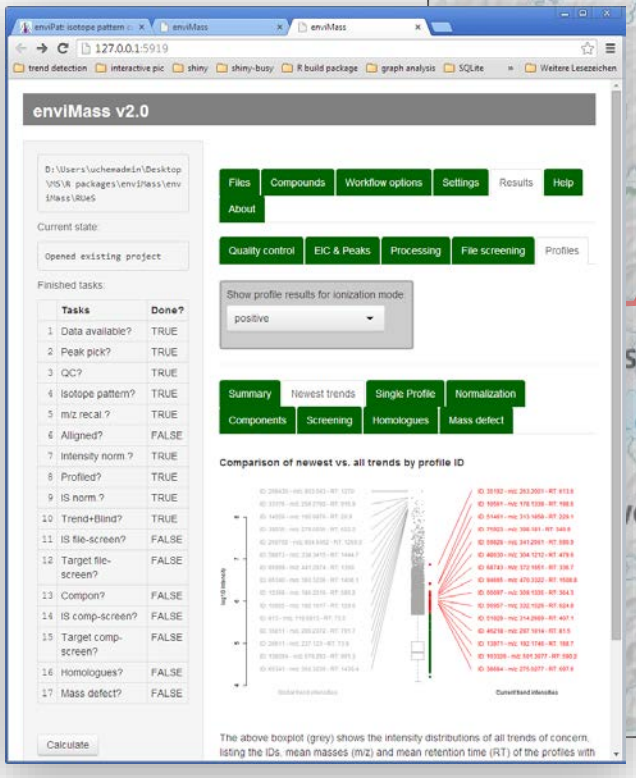
Juliane.Hollender@eawag.ch

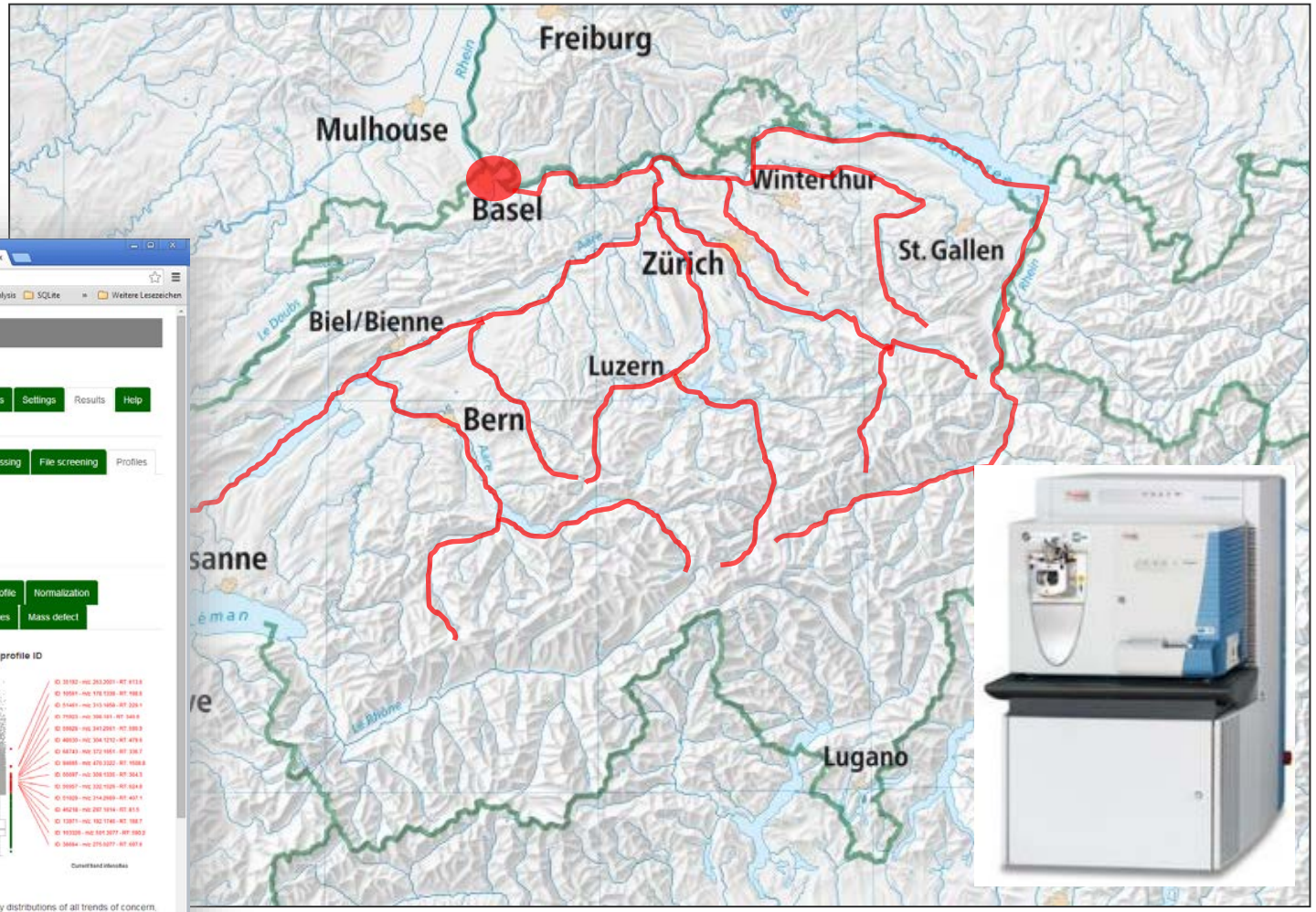
Non-target screening is in daily use to mange chemicals at the International Rhine monitoring station

Departement für Wirtschaft, Soziales und Umwelt des Kantons Basel-Stadt

Amt für Umwelt und Energie

enviMass





http://www.eawag.ch/en/department/uchem/software/

Analytical strategy

Solid phase extraction followed by LC-HRMS/MS (Orbitrap)
High resolved full scan spectra & MSMS fragmentation spectra

Target analysis

50

250

Non-target analysis

remaining masses
∞

Priority substances

Site specific substances

- x calibration
- x exact quantification
- x manual data processing

- x calibration
- x semi-quant. quantification
- x automated processing

- Time profiles
- Statistical analysis
- Changes / trends
- Prioritization

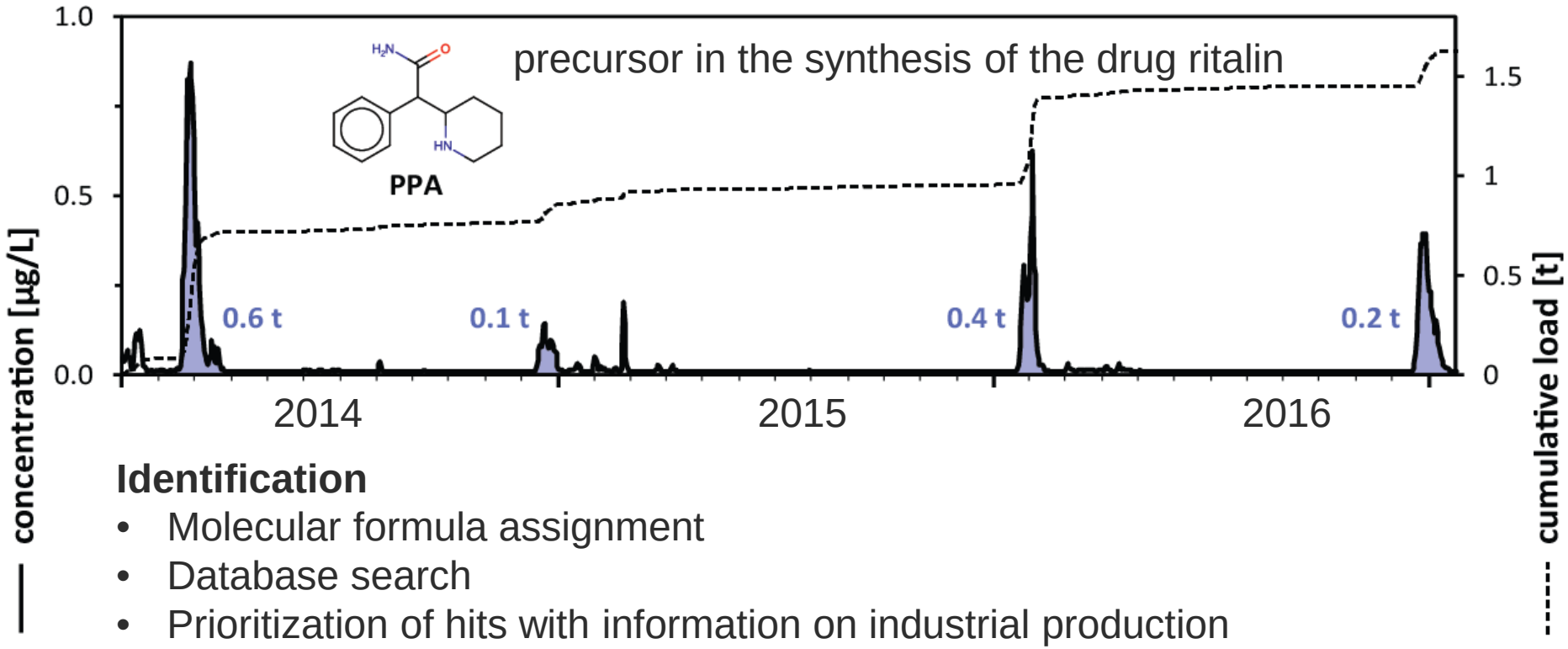
enviMass

Ruff et al., Aqua & Gas 2013, 5: 16-25

Daily non-target screening of the Rhine River

Prioritization using time profiles

Previously unknown chemicals detected due to “stand-out” patterns



Identification

- Molecular formula assignment
- Database search
- Prioritization of hits with information on industrial production
- Confirmation with reference standard

► **10 major spills of non-target compounds in 2014 with > 25 tons of load**

Hollender et al, Env. Sci. Technol. 2017, 51: 11505-11512

Data of the Rhine monitoring station

Content: Reliability - Challenges & Solutions

Experiences/tools of the NORMAN network

Challenges

- Suspect lists
- Quality control
- False positives versus false negatives
- In-source fragmentation
- Identification confidence
- Semi-quantification without standard

Conclusion & points for the discussion

Example
Monitoring of Swiss groundwater

NORMAN network

> 70 organisations
academia, govermental organisations, research centers, industry

NORMAN
Network of reference laboratories, research centres and related organisations for monitoring of emerging environmental substances

The mission of the NORMAN network is to:

- Exchange of information and collection of data on **emerging environmental substances**
- **Validation and harmonisation** of monitoring tools
- Bottom up activities
- Science to policy interface

8 working groups, one on **NTS with > 30 participating organizations**

https://www.norman-network.net/?q=node/252

Cross-Working Group Activity Non-target Screening (NTS)

Non-target screening techniques for environmental monitoring

NORMAN experiences/activities since 2012 in NTS

Comparison & harmonization of NTS methods in Europe

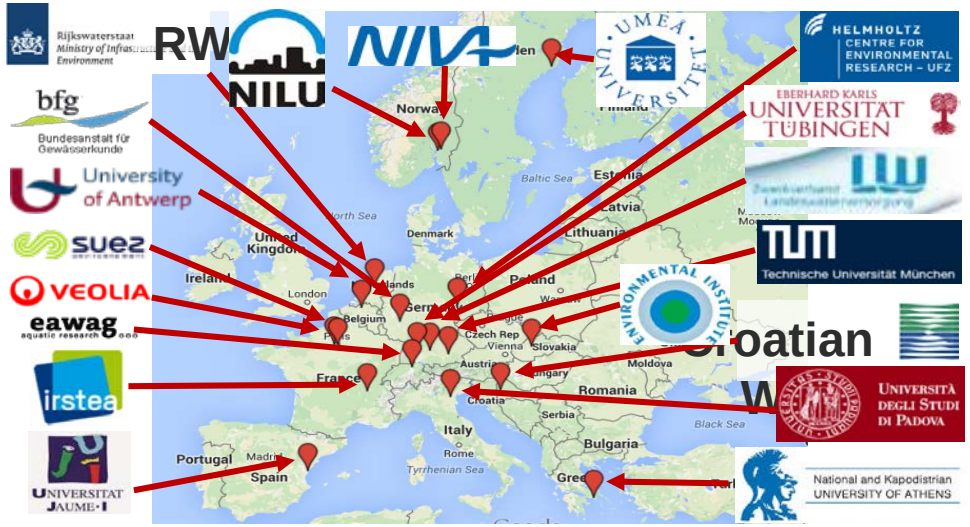
- **Collaborative trials using LC/GC-HRMS/MS:** on water (Schymanski et al. ABC, 2015) and dust (Rostkowski et al, ABC, 2019), in 2019 planned for biota and passive samplers

Databases & Tools

- Home for **MassBank**, **Norman Suspect Lists Exchange**, **Digital Sample Freezing platform** (Alygizakis et al, TrAc 2019)
- **Retention time index** (Aalizadeh, J. Chem Inf Mod 2016 & J Haz Mat 2019)
- **NormaNews exchange (**Alygizakis et al, ES&T 2018)

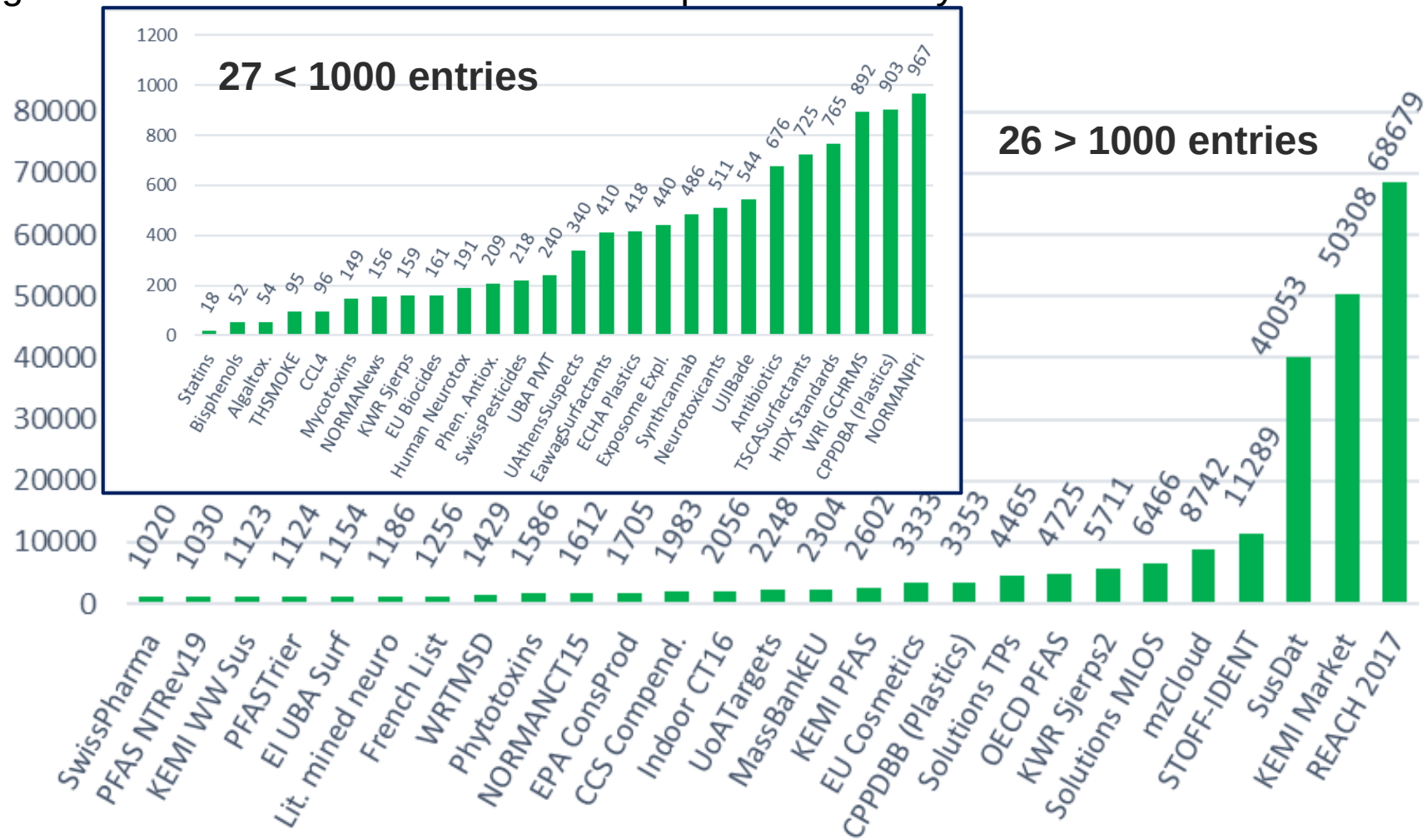
Communication

- NTS **group meetings**
- NTS workshops / **training courses** (e.g. CH 2014, N 2017)
- NTS **workshop for regulators** in Brussels, 2018



NORMAN databases: Suspect List Exchange

- https://www.norman-network.com/?q=suspect-list-exchange
- 53 lists available … specialist collections to market lists
  - Integrated in NORMAN Databases & CompTox Chemistry Dashboard

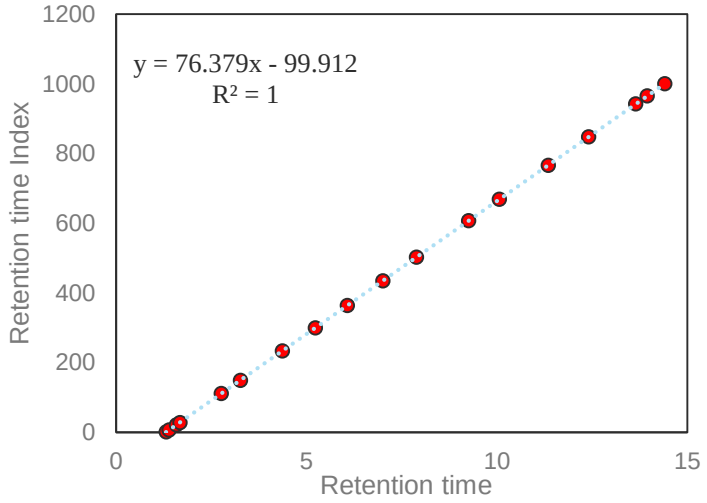


- Schymanski, Aalizadeh *et al*. in prep; https://www.researchgate.net/project/Supporting-Mass-Spectrometry-Through-Cheminformatics

Retention Time Indices for RPLC-(+)ESI-HRMS

Quantitative structure retention relationship models for RTI system using chemical structure information

RTI = 76.379(RT) - 99.912



$$RTI = \frac{(RTx - RTmin)}{(RTmax - RTmin)} * 1000$$

Predicted by QSRR model:
C18 column, H_2O/MeOH formic acid gradient

	Training			Test		
	R2	RMSE	F	R2	RMSE	F
MLR	0.835	92.575	1515.130	0.870	83.184	426.416
SVM	**0.861**	84.869	1838.745	**0.880**	80.029	467.038

Calibrants	RT (Acclaim C18)	RTI
Guanylurea	1.31	1.00
Amitrole	1.39	6.11
Histamine	1.58	20.63
Chlormequat	1.67	27.50
Methamidophos	2.76	110.77
Vancomycin	3.26	148.97
Cefoperazone	4.36	233.00
Trichlorfon	5.23	299.47
Butocarboxim	6.07	363.64
Dichlorvos	7	434.68
Tylosin	7.88	501.91
TCMTB	9.25	606.57
rifaximin	10.06	668.45
Spinosad A	11.34	766.23
Emamectin B1a	12.4	847.21
AvermectinB1a	13.64	941.94
Nigericin	13.94	964.86
Ivermectin B1a	14.4	1000.00

Aalizadeh et al., J. Chem Inf Mod 2016 & J Haz Mat 2019

Online Platform to Calculate Experimental and Predicted Retention Time Indices

Development and Prediction of Retention Time Indices for LC-HRMS (version 2.0.0)

National and Kapodistrian UNIVERSITY OF ATHENS

Select the target ESI:
- (•) +ESI
- ( ) -ESI

Select the RTI versus tR calibration mode:
- ( ) Prediction limits
- (•) Auto-calibrate
- ( ) Manual

Upload RTI calibrants data...

Browse... Calibrants_pos_UOA.csv

Upload complete

Defalut max. file size 1MB (*.csv file)

Click here to build the calibration curve...

Click here to restart the current session...

Trace Analysis & Mass Spectrometry Group

NORMAN Suspect list Exchange website

Retention Time Indices for LC-HRMS (version 2.0.0):

About | Calibrants | Single compound | Batch mode | Comparison of Experimental RTIs | Chemical Curation | NORMAN network | FAQ

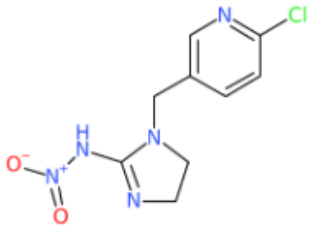
Select the uncertainity measurement:
- (•) OTrAMS
- ( ) Chemical space boundary

Enter the SMILES of a compund here:

[O-][N+](=O)NC/1=N/CCN\1Cc2cnc(Cl)cc2

Enter the tR for specific ESI mode

4.2



Canonical SMILES: O=[N+]([O-])NC1=NCCN1CC2=CN=C(Cl)C=C2
InChIKey: YWTYJOPNNQFBPC-UHFFFAOYSA-N
Experimental tR: 4.2
The ESI selected: +ESI

Estimate RTI & its uncertainty

Status: Processed

Experimental RTI: 220.88
Predicted RTI: 250.83

Experimental tR: 4.2 min
Predicted tR: 4.59 min

Uncertainty: Exp. & Pred. tR are accepted for this candidate (box1)

*RTI vs tR calibration curve: RTI= 76.3788 *(tR) -99.9116 >>> (R^2= 1)*

Save the prediction results...

Open access mass spectra libraries

- **European NORMAN MassBank**:
  currently ~53'000 spectra of ~16'000 compounds
  from 15 main instrument types and 40 institutions

- **MoNa: MassBank of North America:**
  > 200'000 mass spectra including in silico spectra & European MassBank

- **mzCloud** - HighChem:
  17'000 compounds, highly curated, mostly Orbitrap

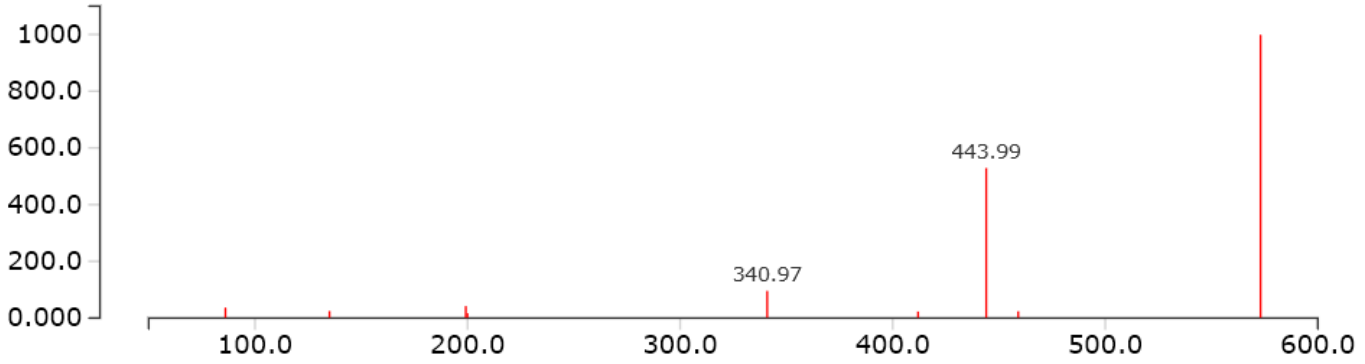
MassBank enables storage of tentative MSMS spectra

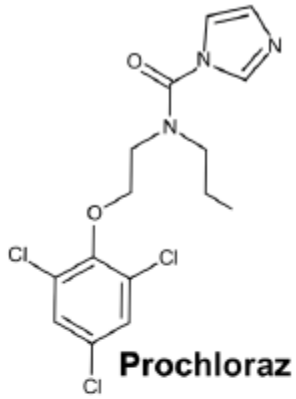
MassBank Record: ET201801

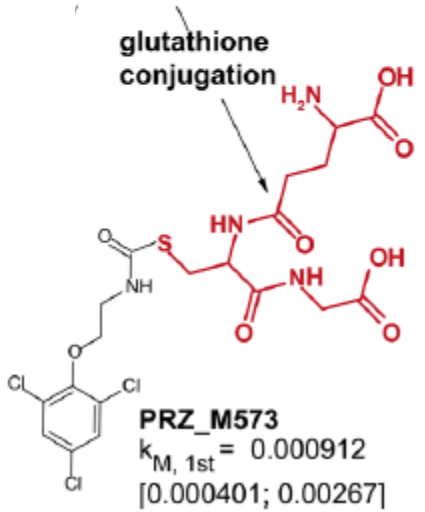
Home | Search | Record Index | Data Privacy | Imprint | MassBank ID: Go

PRZ_M573; LC-ESI-QFT; MS2; CE: 10; R=35000; [M+H]+

Mass Spectrum



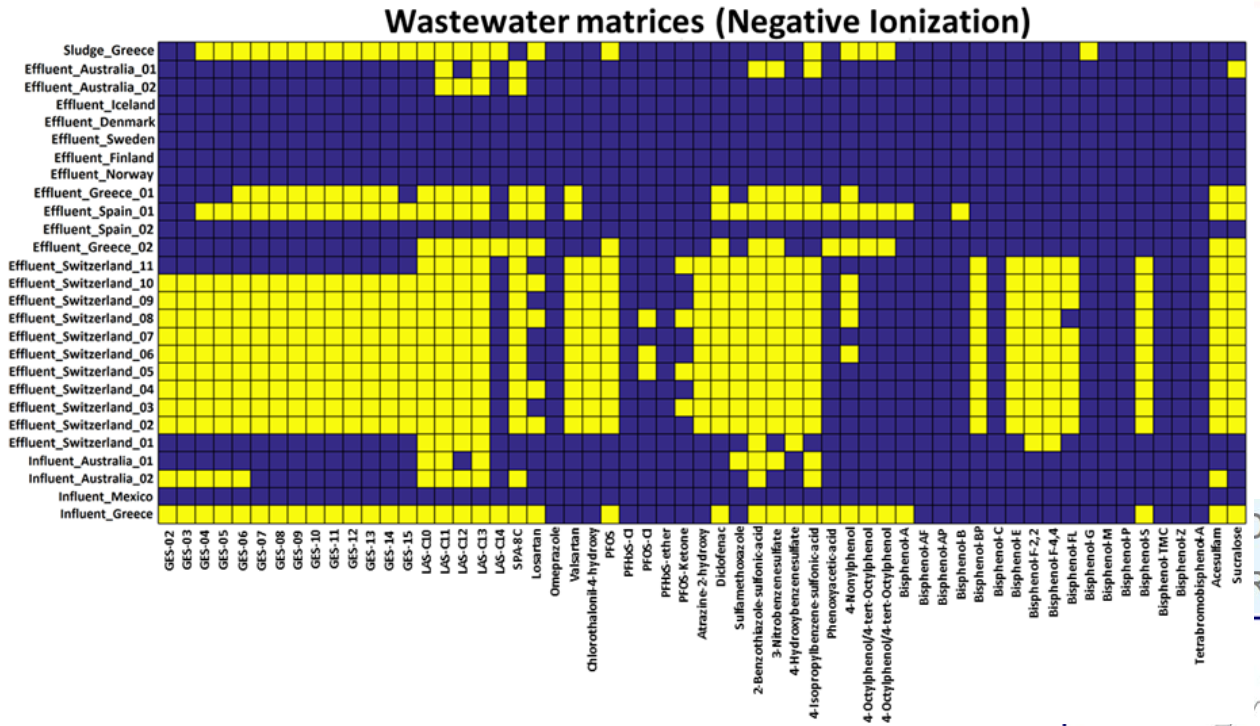





```
ACCESSION: ET201801
RECORD_TITLE: PRZ_M573; LC-ESI-QFT; MS2; CE: 10; R=35000; [M+H]+
DATE: 2016.03.01
AUTHORS: , A. Roesch, E. Schymanski, J. Hollender, Department of Environmental Chemistry, Eawag
LICENSE: CC BY
COPYRIGHT: Copyright (C) 2015 Eawag, Duebendorf, Switzerland
PUBLICATION: Rösch, A.; Anliker, S.; Hollender, J. How Biotransformation Influences Toxicokinetics of Azole Fungicides in
COMMENT: CONFIDENCE Tentative identification only (Level 3)
COMMENT: INTERNAL_ID 2018
```

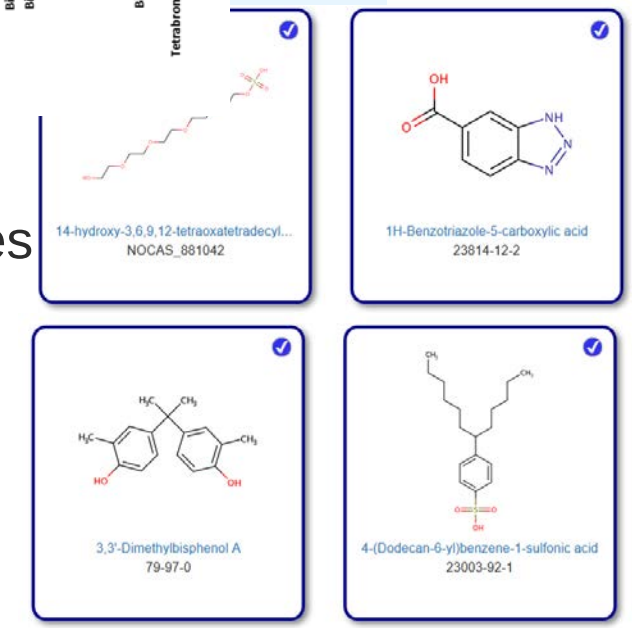
Rösch et al., ES&T 2016

NormaNews: Sharing of emerging contaminants
Retrospective screening of emerging suspects



National and Kapodistrian
UNIVERSITY OF ATHENS

- 150 suspect compounds
- HRMS data from 14 countries
- QA/QC



Alygizakis et al. 2018 ES&T & https://comptox.epa.gov/dashboard/chemical_lists/normanews

Digital Sample Freezing Platform – DSFP
- A digital specimen bank of georeferenced HRMS data

NORMAN Digital Sample Freezing Platform | Results NORMAN-ECHA

Screening of REACH compounds in samples from the Black Sea

Selection of data

Compounds with:

>=4 fragments

Apply changes

Appearance

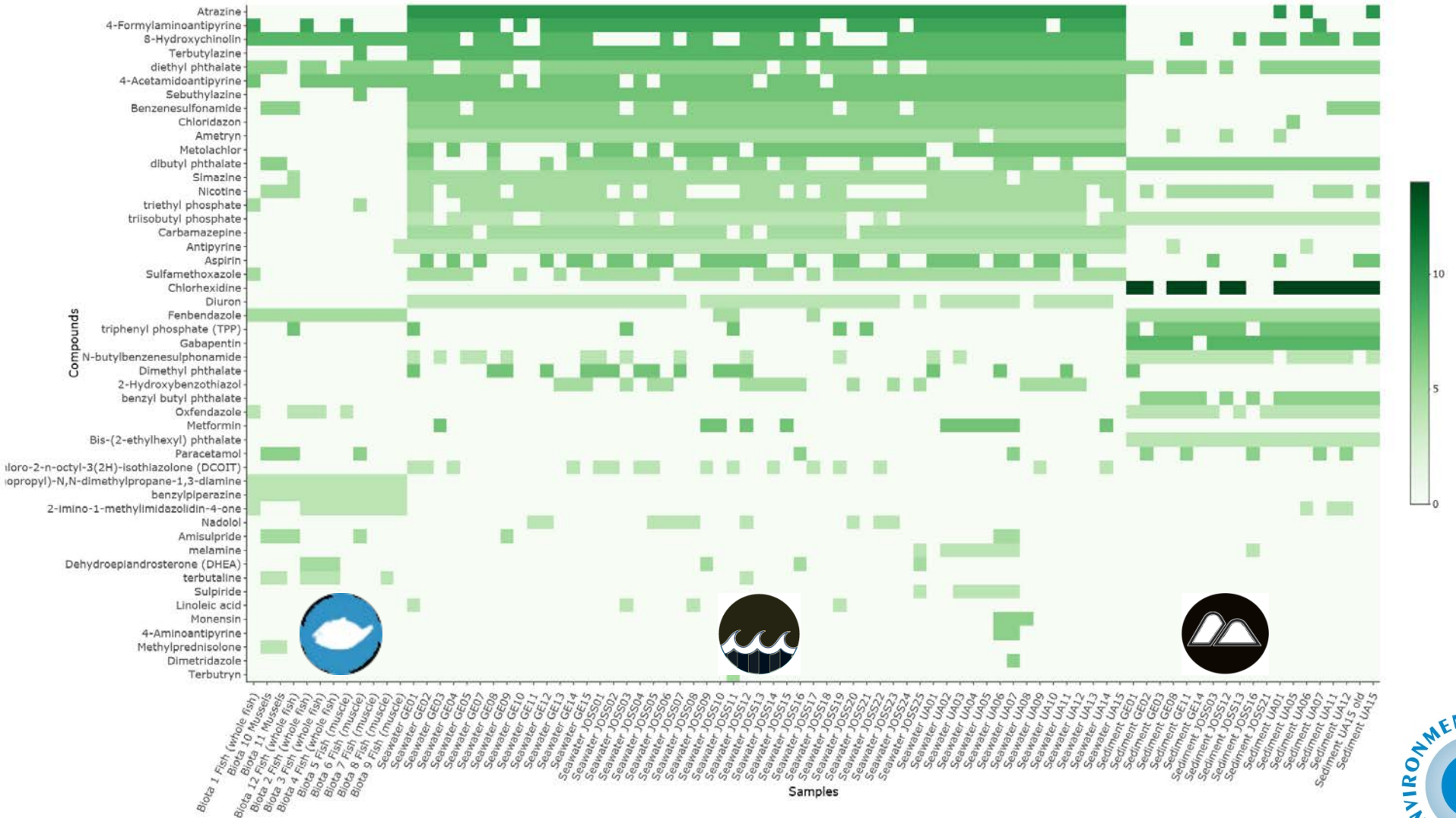
Height: 330

Width: 300

Size Letters Y: 10

Size Letters X: 10

Rotation of labels on x-axis: 65



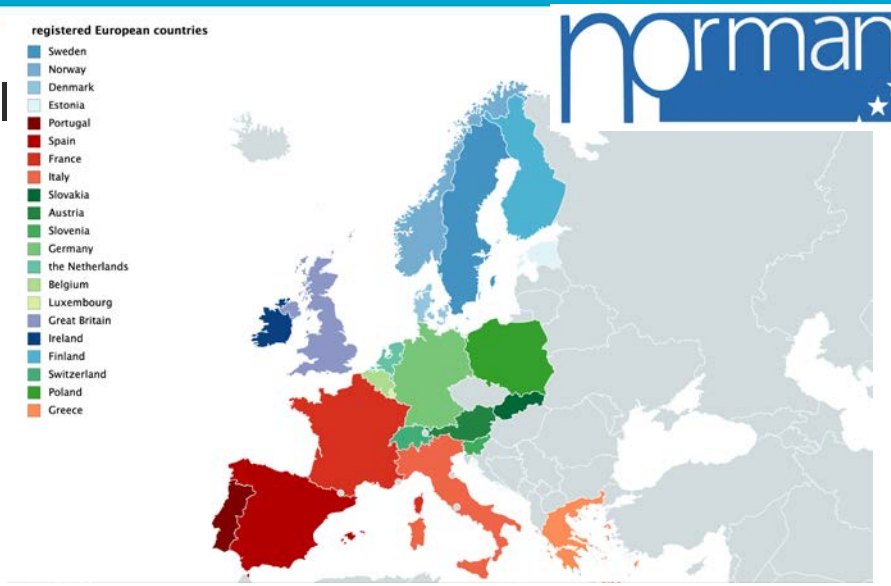
Interactive heatmap available at http://norman-data.eu/NORMAN-REACH, Algyzakis et al., TrAC 2019

NORMAN Workshop for regulators: 25.10.2018, Brussels

How can non-target screening techniques support environmental monitoring and chemicals management?

- NTS can improve the **identification of problematic substances** and **support regulatory processes in environmental and chemical legislation** (e.g. WFD, the Marine Strategy Framework Directive, REACH)
- NTS can be a **first screening step in the exposure assessment** chain but **does not replace target monitoring**.
- **Harmonized NTS protocols** and **minimum quality requirements** should be established.
- **New protocols / infrastructures** are **needed for efficient NTS data** management, evaluation and sharing.
- **Training** would be beneficial to make NTS more widely accessible.
- **Synergies** between NTS and effect-based methods should be strengthened

Hollender et al., submitted to Env. Sci. Europe



> 80 participants from
- Ministries, agencies
- Regulatory institutes
- Research institutes
- Industry
- Academics

responsible for
- Environmental/chemical legislation
- Food/drinking water safety
- Human biomonitoring

NORMAN Non-target screening guidance document

- **Key input:**
  - **experience of NORMAN members** from trials etc.
  - Different approaches for different regulatory questions where NTS can help (monitoring, prioritization,…)
  - Guidance of German Water Chemistry Society on Suspect and Non-target screening in water analysis, draft
  - Important topics: **Number & type of blanks, Replicate samples/ replicate analyses, Compound domains covered vs. LC/GC & MS method setup, data dependent vs. data independent acquisition**
- Presentation of draft version at the SWEMSA 2019 workshop (TU Munich, 21-23/10/19)
- Discussion on general assembly end of November in Milano
- Publish in 2020 with open access in a peer-reviewed journal

Krauss et al., in prepration

Challenges for reliability in non-target/suspect screening

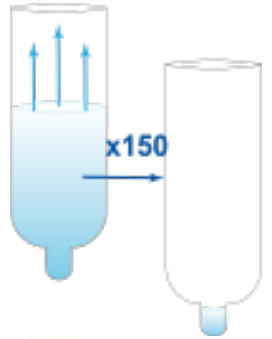
Suspect screening of overlooked pesticide TPs in Swiss groundwater

31 groundwater wells with intensive agriculture

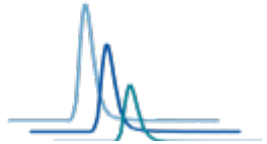
(Swiss national monitoring sites)

Workflow

1. • Enrichment: vacuum evaporation
2. • Analysis: RP-HPLC-ESI-HRMS/MS
3. • Automated screening for exact mass
4. • Prioritization of suspect hits
5. • Confirmation with reference standards

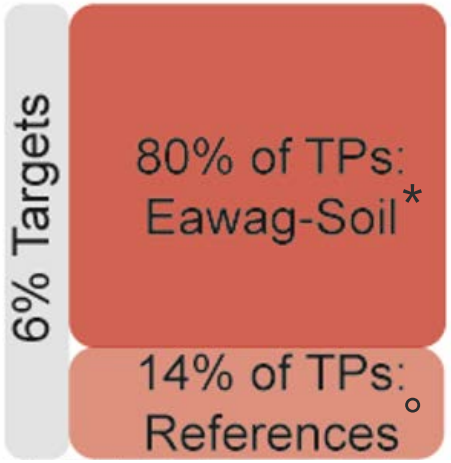


QExactivePlus
R= 140000



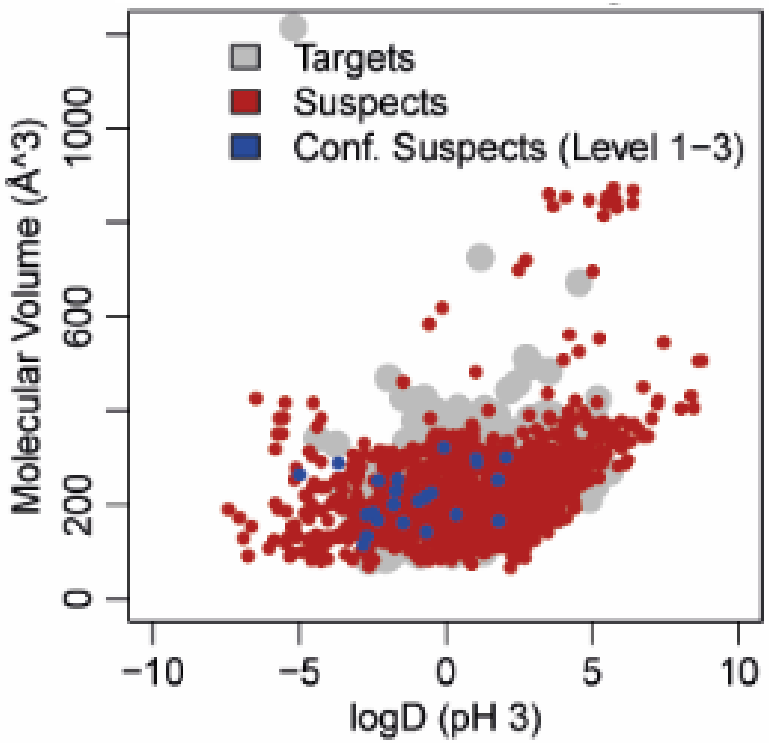
Smart Suspect Screening of pesticide transformation products in groundwater

1120 TPs
of ca. 300 pesticides



*~1000 pesticide TPs, enviPath
from European pesticide registration
(Latino et al. ESPI 2017)
° i.a. PPBD, Lewis et al. HERA 2016,
Reemtsma et al., Wat Res 2013

LC-ESI-MS amenability



95%: similar polarity as targets
99% contain N, O, S atom

- Appropriate suspect list
- Appropriate method for compounds on suspect list

Kiefer et al., submitted

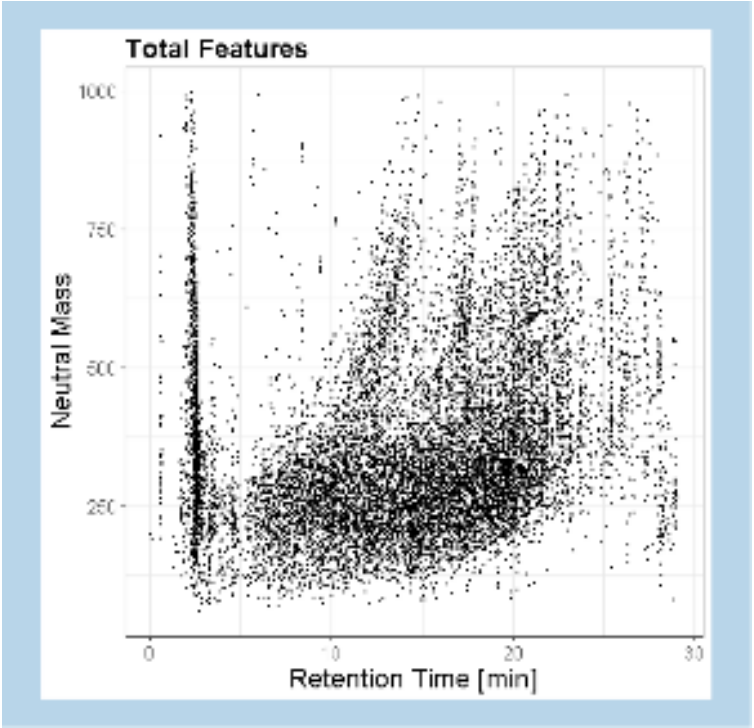
Challenges to achieve reliability – quality control

Quality control is necessary for instrumental analysis, but also data processing!

- Perform **QC samples** in each batch of samples (e.g., composite samples of each matrix, spiked samples, standard mixtures) and a sufficient number of sample processing and instrument blanks

- Use **internal standards** (spiked standards) for QC of analytical performance, but also of peak detection, sample alignment, mass accuracy.

- Take the time to adjust the **peak detection settings to your data**. Automatic parameter optimization is available in some workflows (e.g. XCMS, EnviMass).

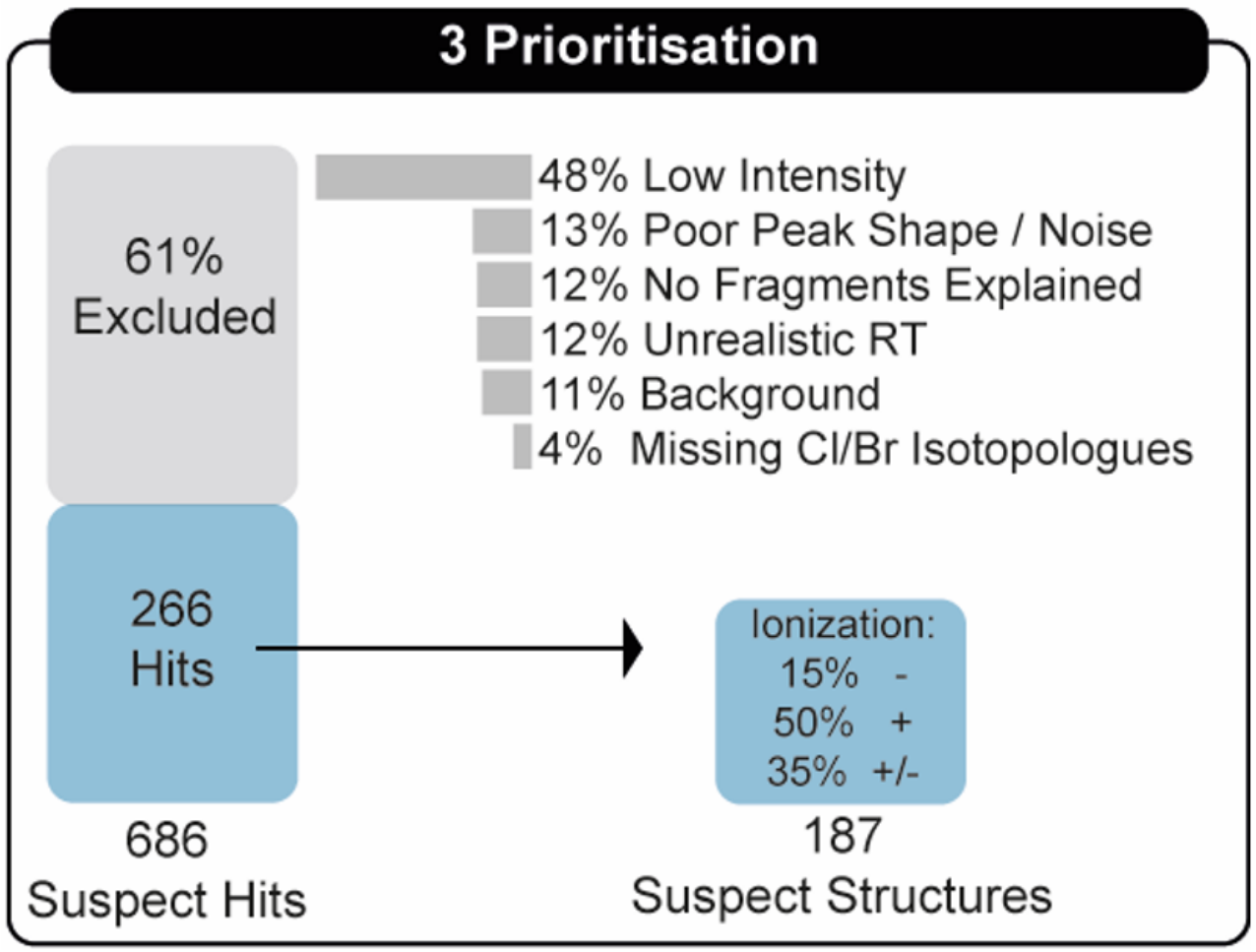


Quality control: False positives versus false negatives

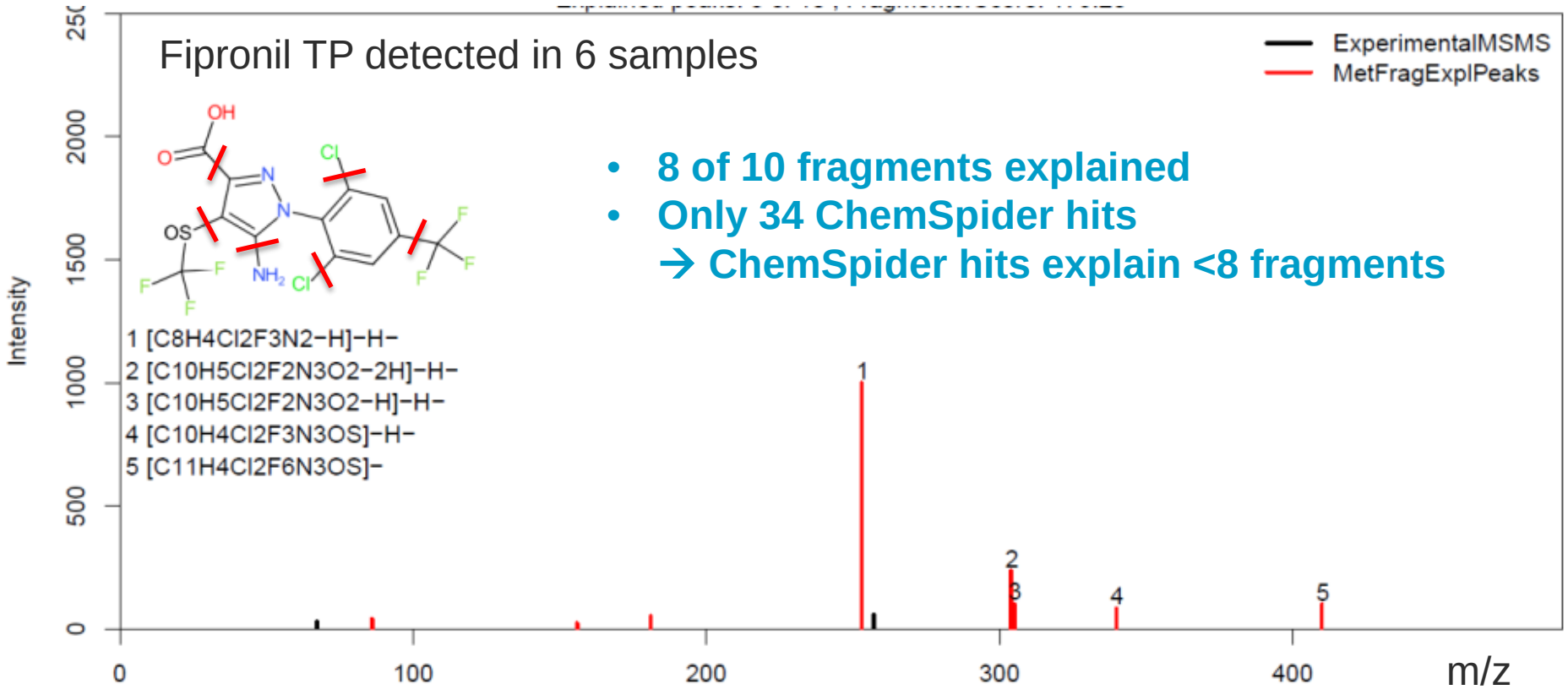
Example: aim to not overlook TPs with low ionization efficiency

Validation of workflow
(peak picking, Rt alignment, isotopologues grouping, background substraction, suspect screening, CD 2.1 Thermo) **with >200 internal standards**

- ► 97% of internal standards detected
- ► but 9300 suspects detected
- ► filtering of noise/background (90%) => 686 hits
- ► further filtering needed



Filtering of data: in silico fragmentation with MetFrag



- **8 of 10 fragments explained**
- **Only 34 ChemSpider hits**
  → ChemSpider hits explain <8 fragments

http://c-ruttkies.github.io/MetFrag/

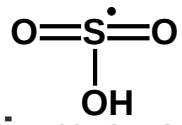
- Compare against other databases
- Check also for in-source fragmentation

Ruttkies et al. J Cheminf 2016

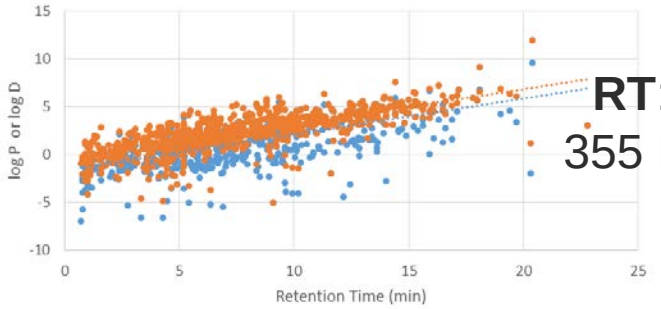
Use of meta information improves identification success

m/z [M-H]$^-$
213.9637
± 5 ppm

Elements: C,N,S
5 ppm
0.001 Da

Substructures: 

Hydrogen deuterium exchange



RT: 4.54 min
355 InChI/RTs

ChemSpider
Search and share chemistry

or

PubChem OPEN CHEMISTRY DATABASE

References
External Refs
Data Sources
RSC Count
PubMed Count

Literature Patents

MetFrag

Suspect Lists

TOFF IDENT

EPA Chemicals

MoNA
MassBank of North America
MassBank.eu

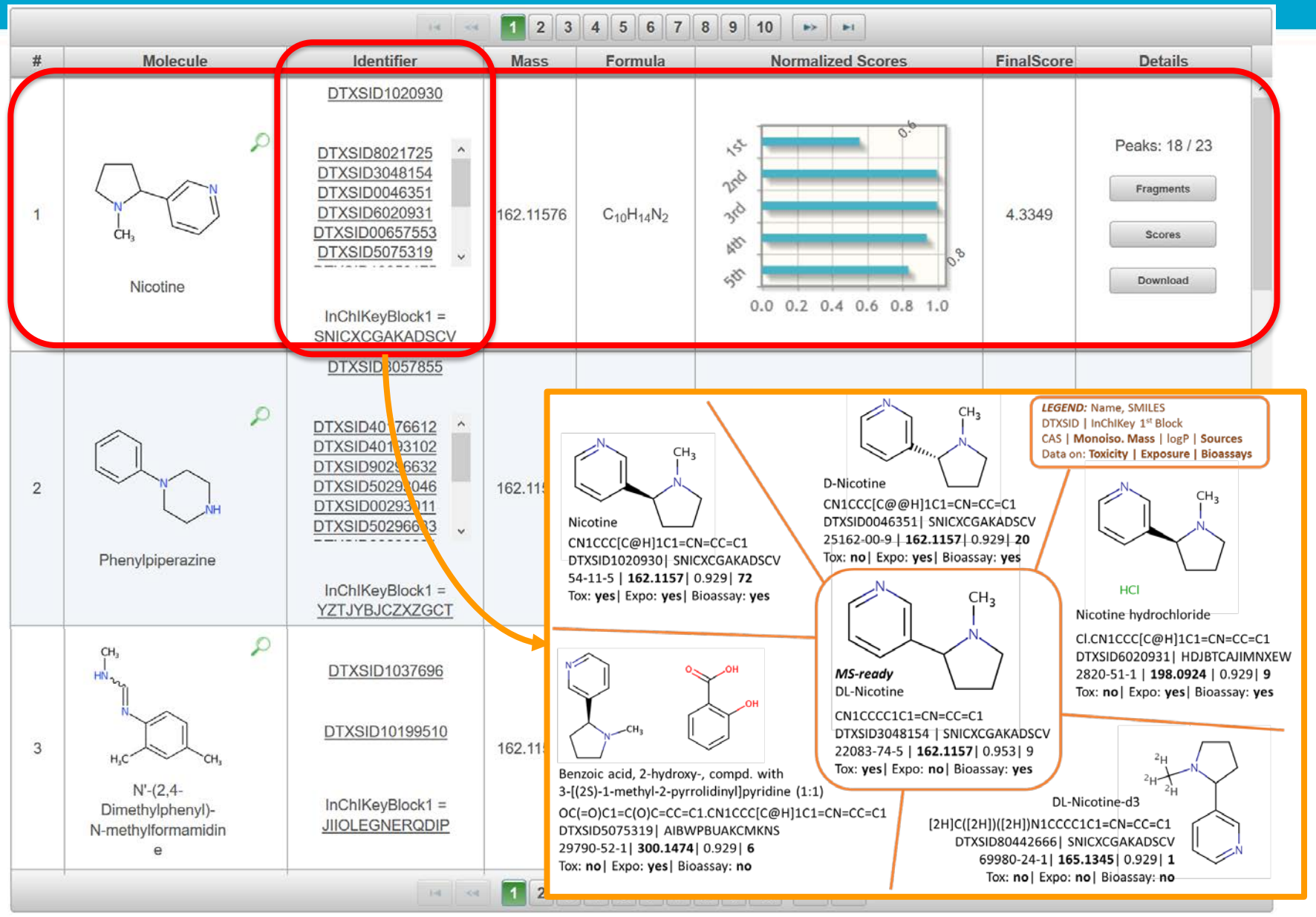
MS/MS

134.0054	339689
150.0001	77271
213.9607	632466

$$Score_{Final} = \omega_{Frag} * score_{Frag} + \omega_{RT} * score_{RT} + \omega_{ref} * score_{ref} + \ldots$$

Ruttkies et al. *J. Cheminf.*, 2016,
Ruttkies et al., ABC 2019

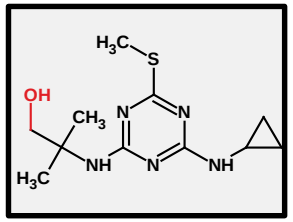
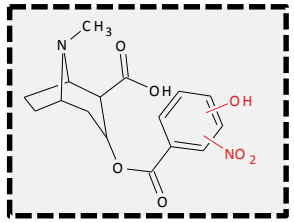
Clear chemical identifiers necessary



“MS-ready” form

McEachran *et al*. 2018, DOI: 10.1186/s13321-018-0299-2; Schymanski & Williams, 2017 *ES&T* DOI: 10.1021/acs.est.7b01908

Communicating confidence of identification (HRMS/MS)

Example	Identification confidence	Minimum data requirements
	*Level 1: Confirmed structure* by reference standard	MS, MS^2, RT, Reference Std.
	Level 2: Probable structure a) by library spectrum match b) by diagnostic evidence	MS, MS^2, Library MS^2 MS, MS^2, Exp. data
	*Level 3: Tentative candidate(s)* structure, substituent, class	MS, MS^2, Exp. data
$C_6H_5N_3O_4$	*Level 4: Unequivocal molecular formula*	MS isotope/adduct
192.0757	*Level 5: Exact mass* of interest	MS

What match value accepted?

How many fragments in data dependent /data independent acquisition?

Which exp. data have how much value?

Schymanski, Jeon, Gulde, Fenner, Ruff, Singer & Hollender (2014) *ES&T,* DOI: 10.1021/es5002105

Communicating confidence of identification: Example pesticide TPs

19 Confirmed (Level 1)

- 3 **chlorothalonil TPs**
- **1 cycloxydim TP**
- 1 dimethachlor TP
- **1 fipronil TP**
- **2 fluxapyroxad/bixafen TPs**
- 1 fludioxonil TP
- 1 metalaxyl TP
- 2 metolachlor TPs
- 2 nicosulfuron TPs
- 1 pinoxaden TP
- 4 **terbuthylazine TPs**

2 Probable (Level 2)

- 1 chlorotoluron TP
- **1 fipronil TP**

3 Tentative (Level 3)

- **2 chlorothalonil TPs**
- **2 cymoxanil TPs**

9 Rejected

- 9 TPs

Kiefer et al., submitted

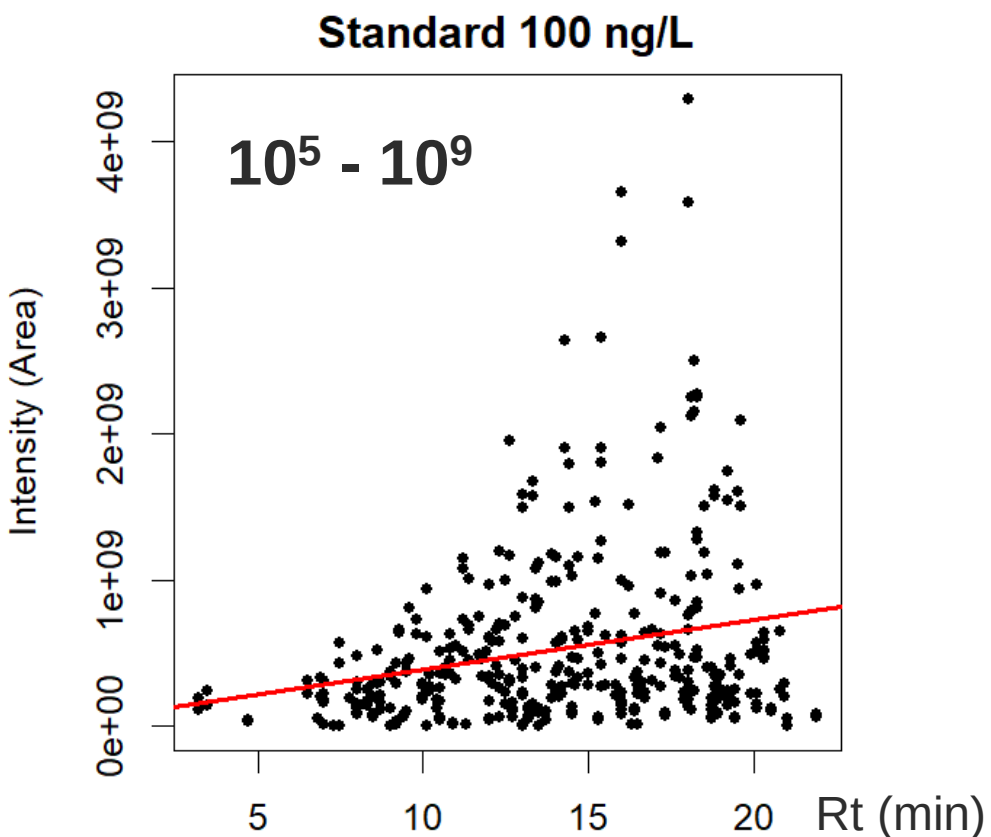
(Semi) Quantification without standards

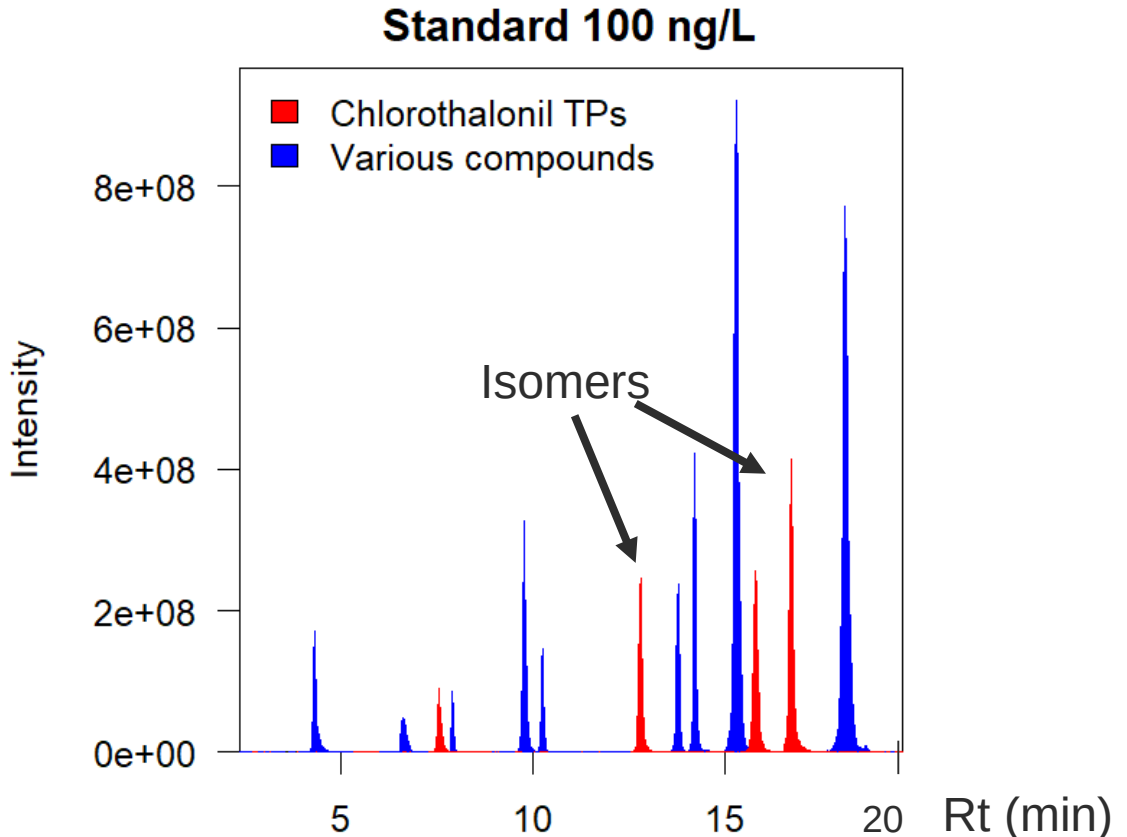
Suspected structure

- with structural similar compounds, e.g. TPs with parent compound
- Quantitative structure response relationship (e.g.Tanimoto coefficient)

Unknown structure

- Concentration range based on standards with different response ratios (Hollender Wat Res X 2018)
- Average response ratio of standards at similar retention time





Conclusions & Points for discussion

- Instruments & tools are available, application in a clever way is important to successfully and reliably identify new compounds
- Optimization of acquisition and data processing is key
- Meta data are very important for successful identification
- Sharing of spectra, data, standards, experiences etc. is useful (NORMAN network)

- Which kind of quality control samples (composites, replicates,…) and how often?
- How to optimize identification workflow?
- How to select smart suspect lists?
- What kind of meta information are most useful?
  Retention time, collision cross section in ion mobility, number of references, exposure index,…?
- What are relevant differences in NTS for different samples (environment, food, human)?
- How to improve communication of identification confidence? DDA or DIA spectra acquisition?
- How to improve data sharing of spectra, suspect lists, etc?

Acknowledgements

Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Bundesamt für Umwelt BAFU

- Cantonal labs, Federal Institute of Metrology for groundwater sampling
- Syngenta, BASF, Dupont, Bayer for reference standards
- Steffen Ruppe, Dorit Griesshaber, Ingrid Langlois, Jan Mazacek, AUE Basel-City
- Tobias Schulze, UFZ
- Reza, Aalizadeh, University of Athens
- Nikiforos Alygizakis, Jaroslav Slobodnik, Environmental Institute
- Saer Samanipour, Kevin Thomas, NIVA
- Christoph Ruttkes, Steffen Neumann, IPB Halle

National and Kapodistrian UNIVERSITY OF ATHENS

Thank you!

eawag
aquatic research

Non-Target Analysis for Environmental Assessment

26–30 May 2020 | Durham, NC, USA

SETAC North America Focused Topic Meeting

Save the Date

26–30 May 2020