

WHAT'S IN THE DUST? GC×GC-MS BASED NON-TARGET SCREENING OF HOUSE DUST

Andriy Rebryk
Peter Haglund
MDCW 2025
3-5/2/2025, Liege



Co-funded by
the European Union

Background



8.9 mio

Gori et al. *Basic Res Cardiol.* 2020



World Health Organization



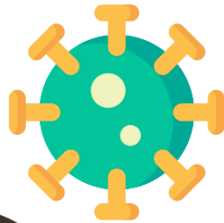
99%

of all people worldwide breathe polluted air in their ambient environment

WHO. 2022



Improving indoor air quality for a healthier home and Europe



Gori et al. *Basic Res Cardiol.* 2020



13–16 h

Schweizer et al. *JESEE.* 2007



>100,000 chemicals

REACH Registration Statistics. 2023



Science of The Total Environment
Volume 957, 20 December 2024, 177639



Comprehensive characterization of European house dust contaminants: Concentrations and profiles, geographical variability, and implications for chemical regulation and health risk

Peter Haglund ^{a, ✉}, Nikiforos A. Alygizakis ^{b, c}, Adrian Covaci ^d, Lisa Melymuk ^e, Pernilla Bohlin Nizzetto ^f, Pawel Rostkowski ^f, Alexandre Albinet ^g, Sylvana Alirai ^b, Dagny Aurich ^h, Stefan Bieber ⁱ, Ana Ballesteros-Gómez ^j, Amanda A. Brennan ^k, Hélène Budzinski ^l, Gabriela Castro ^{m, n}, Fatima den Ouden ^d, Marie-Hélène Dévier ^l, Valeria Dulio ^g, Yong-Lai Feng ^o, Marta Gabriel ^o, Christine Gallampois ^o, ...Xianyu Wang ^g

Picture source: flaticon.com, freepik.com

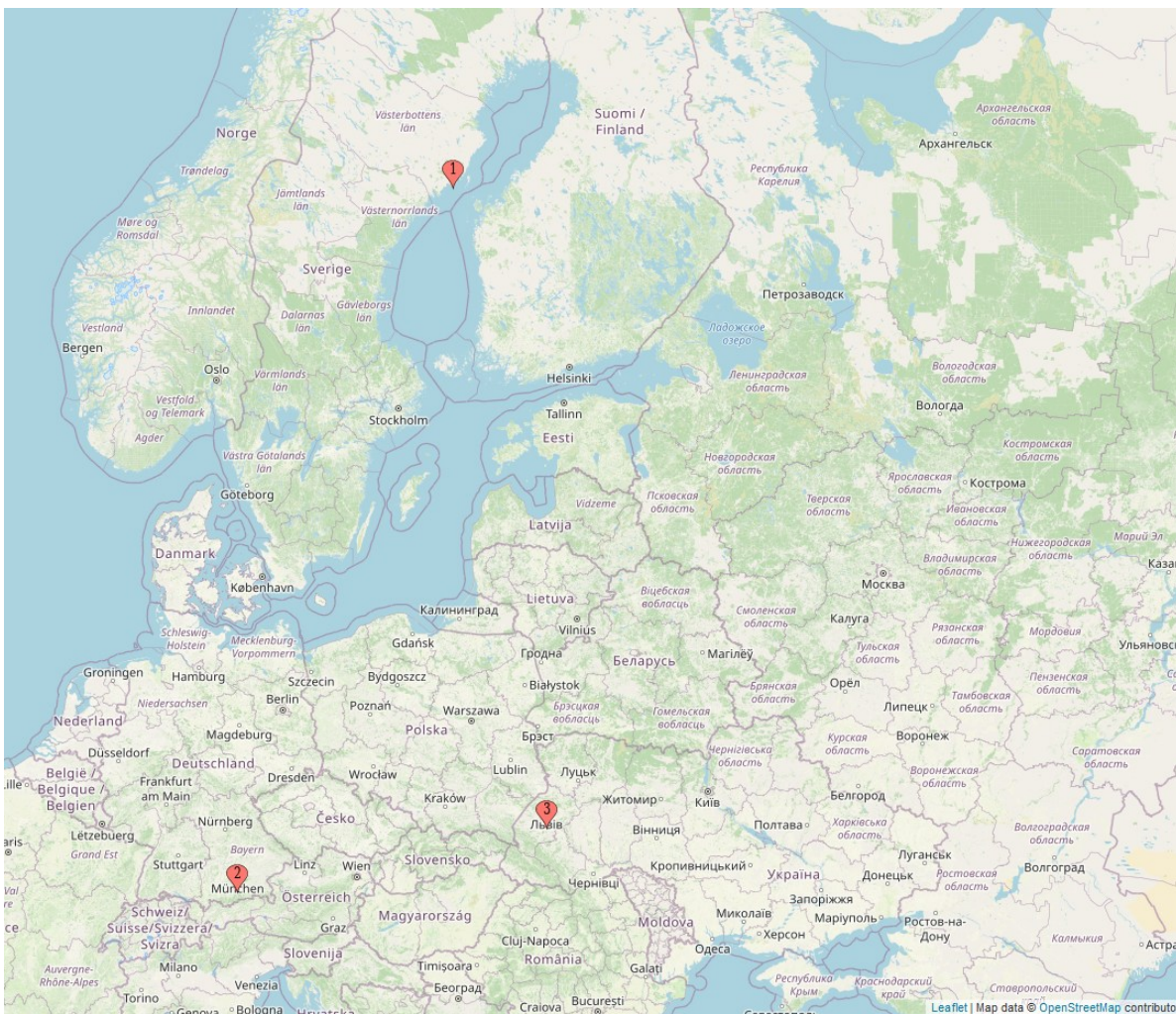
Aim

- Development of experimental workflow:
 - Extraction, clean-up, (possible) fractionation
- Development of data processing workflow:
 - Inter-sample data alignment (incl. various software testing), data filtration & prioritization
- Identification of chemicals of potential concern:
 - Database search & match review, manual identification
 - Classification
- (Semi-)quantification at different levels:
 - Reference standards, internal standards, etc.



Picture source: mapchart.net

Samples



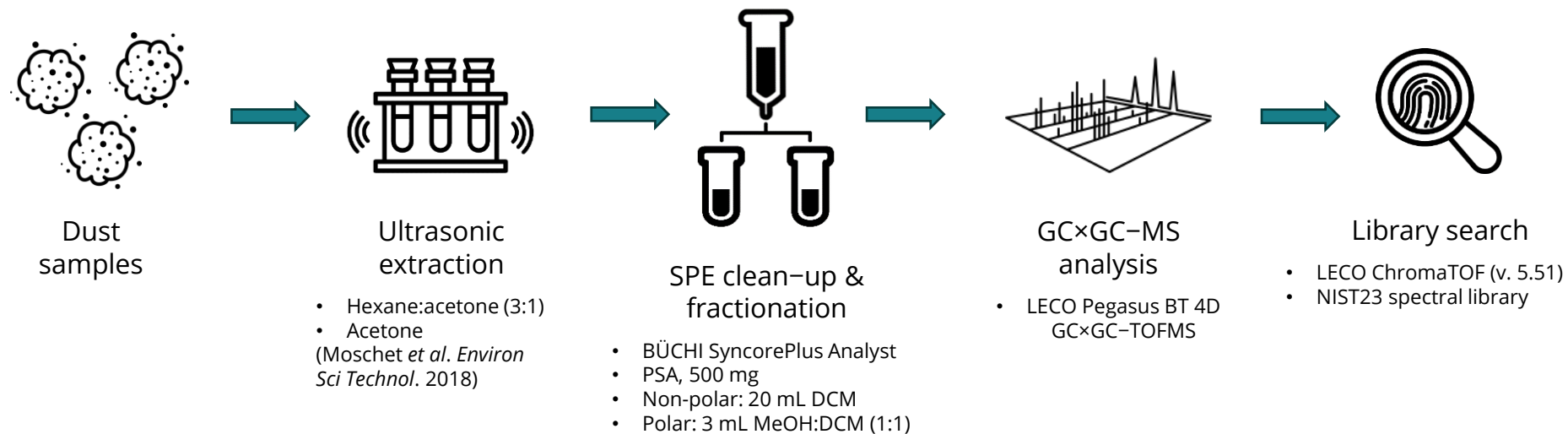
#	City, Country / Other ID	Code	Weight, mg
1	Umeå, Sweden	UMPS	60.9
2	Munich, Germany	MUPS	62.0
3	Lviv, Ukraine	LVPS	62.1

4	NORMAN EU composite sample	EUCS	50.6
5	NIST SRM 2585 (US)	NIST	50.7
6	Laboratory blanks	BLNK	---

#	Code	C1	C2	C3	C4	Standards (152):
1	UMPS	0.05 pg	0.5 pg	5 pg	50 pg	• VOCs
2	EUCS	0.05 pg	0.5 pg	5 pg	50 pg	• SVOCs
3	NIST	0.05 pg	0.5 pg	5 pg	50 pg	• Terpenes
4	BLNK	0.05 pg	0.5 pg	5pg	50 pg	• Phthalates
						• Musks
						• Surfactants

Picture source: mapcustomizer.com

Workflow

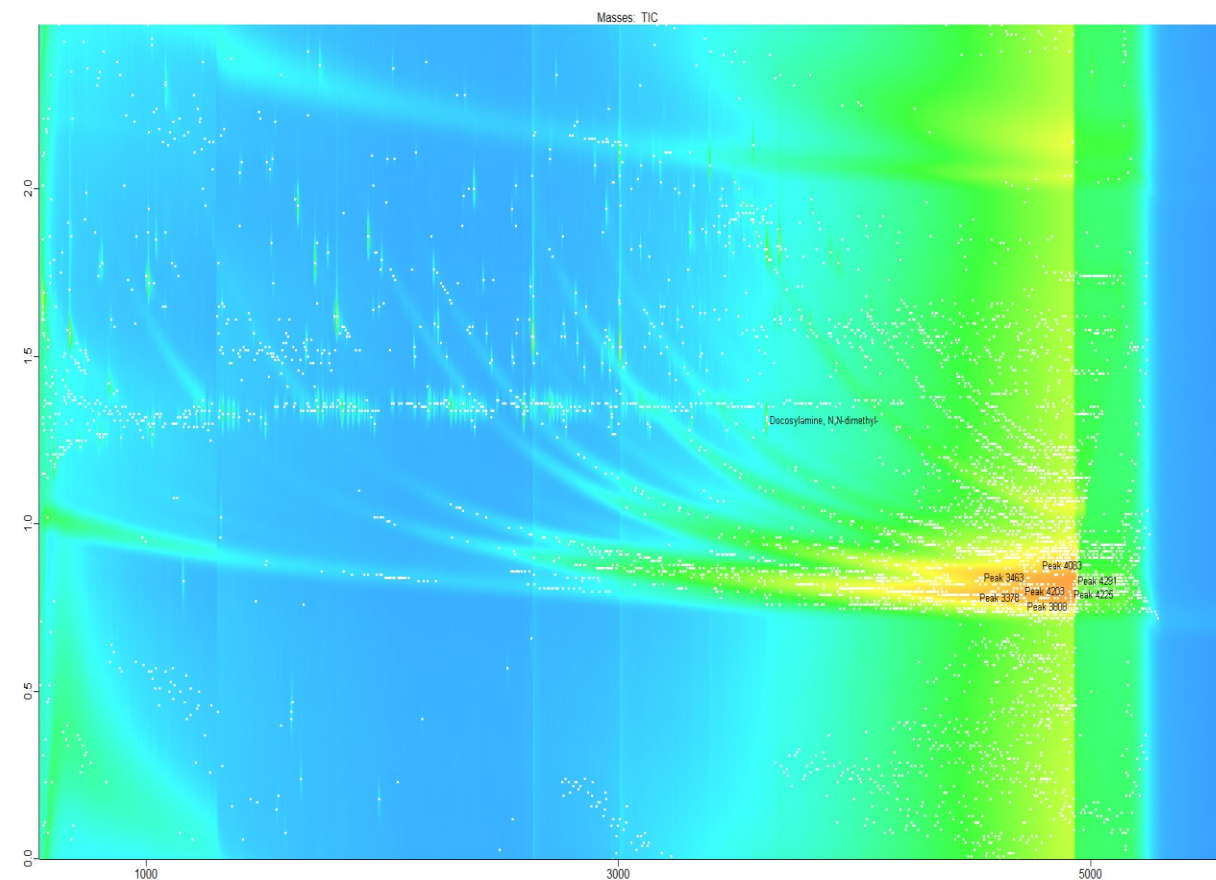
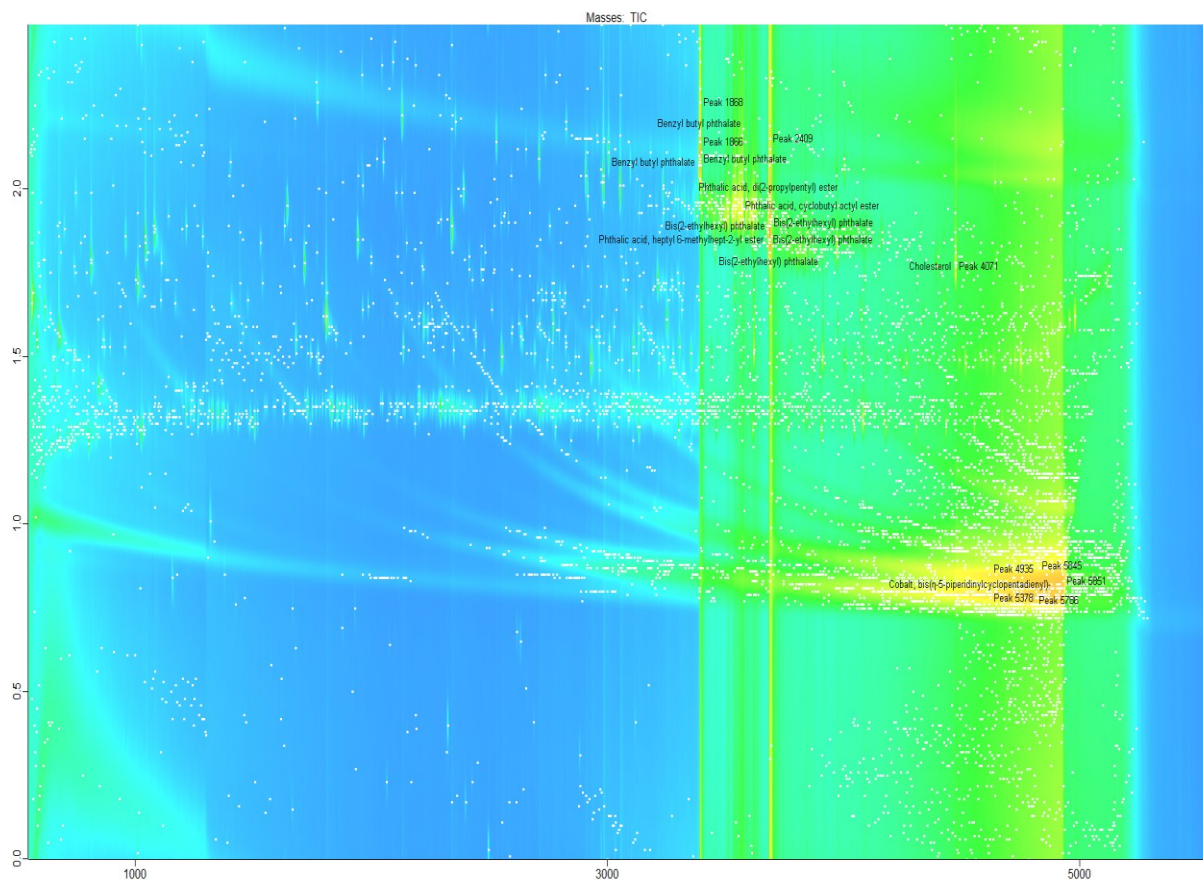


Picture source: flaticon.com, freepik.com

2D GC×GC chromatograms

Umeå, Fr. 1

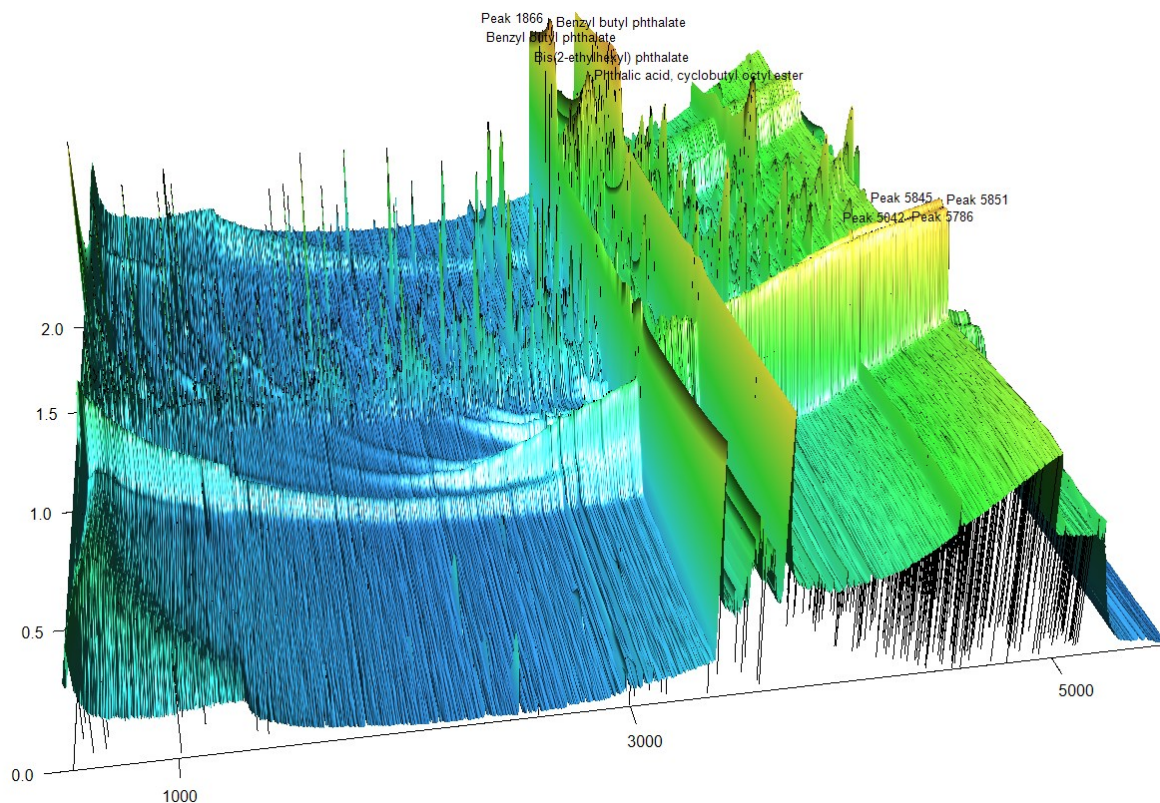
Umeå, Fr. 2



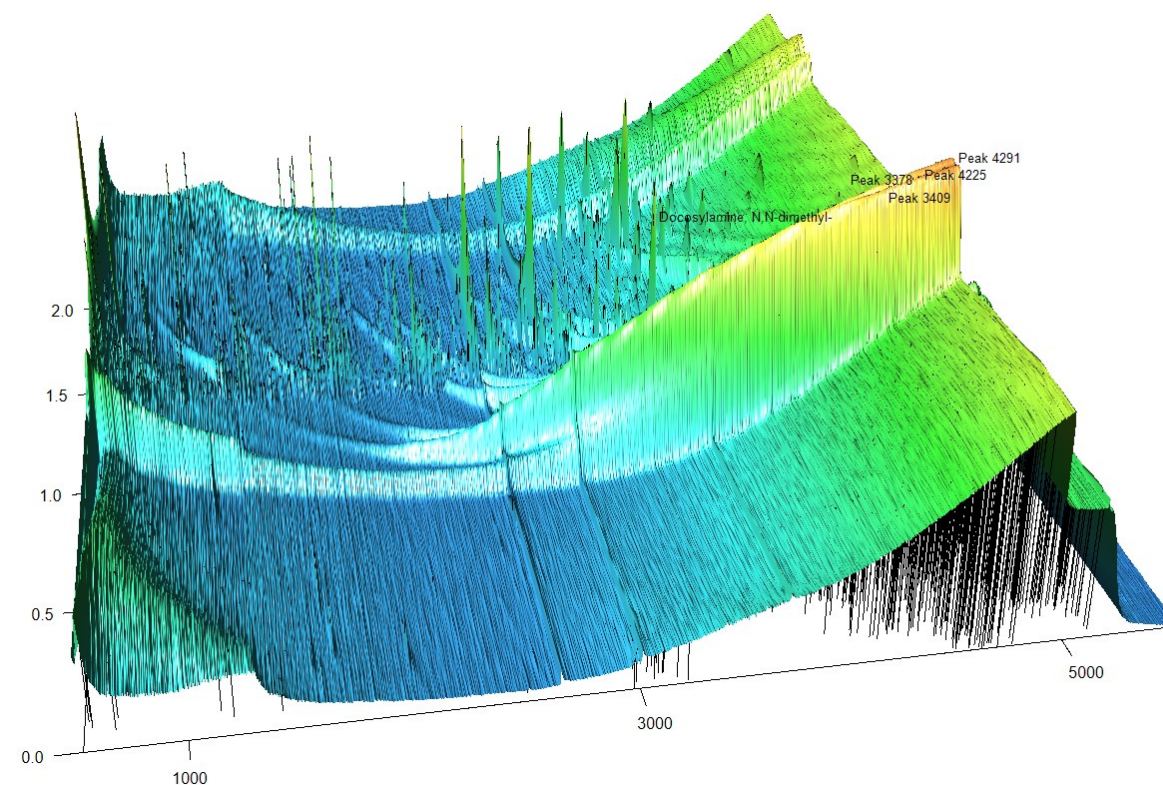
Rebryk and Haglund. *Unpublished data*. 2025

3D GC×GC chromatograms

Umeå, Fr. 1



Umeå, Fr. 2

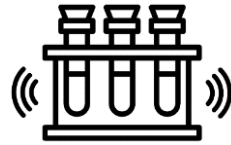


Rebryk and Haglund. *Unpublished data*. 2025

Workflow

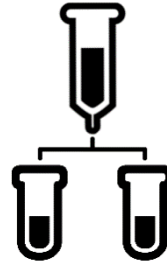


Dust samples



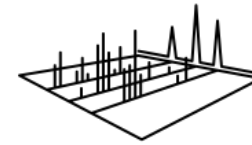
Ultrasonic extraction

- Hexane:acetone (3:1)
 - Acetone
- (Moschet *et al. Environ Sci Technol.* 2018)



SPE clean-up & fractionation

- BÜCHI SyncorePlus Analyst
- PSA, 500 mg
- Non-polar: 20 mL DCM
- Polar: 3 mL MeOH:DCM (1:1)



GCxGC-MS analysis

- LECO Pegasus BT 4D GCxGC-TOFMS



Library search

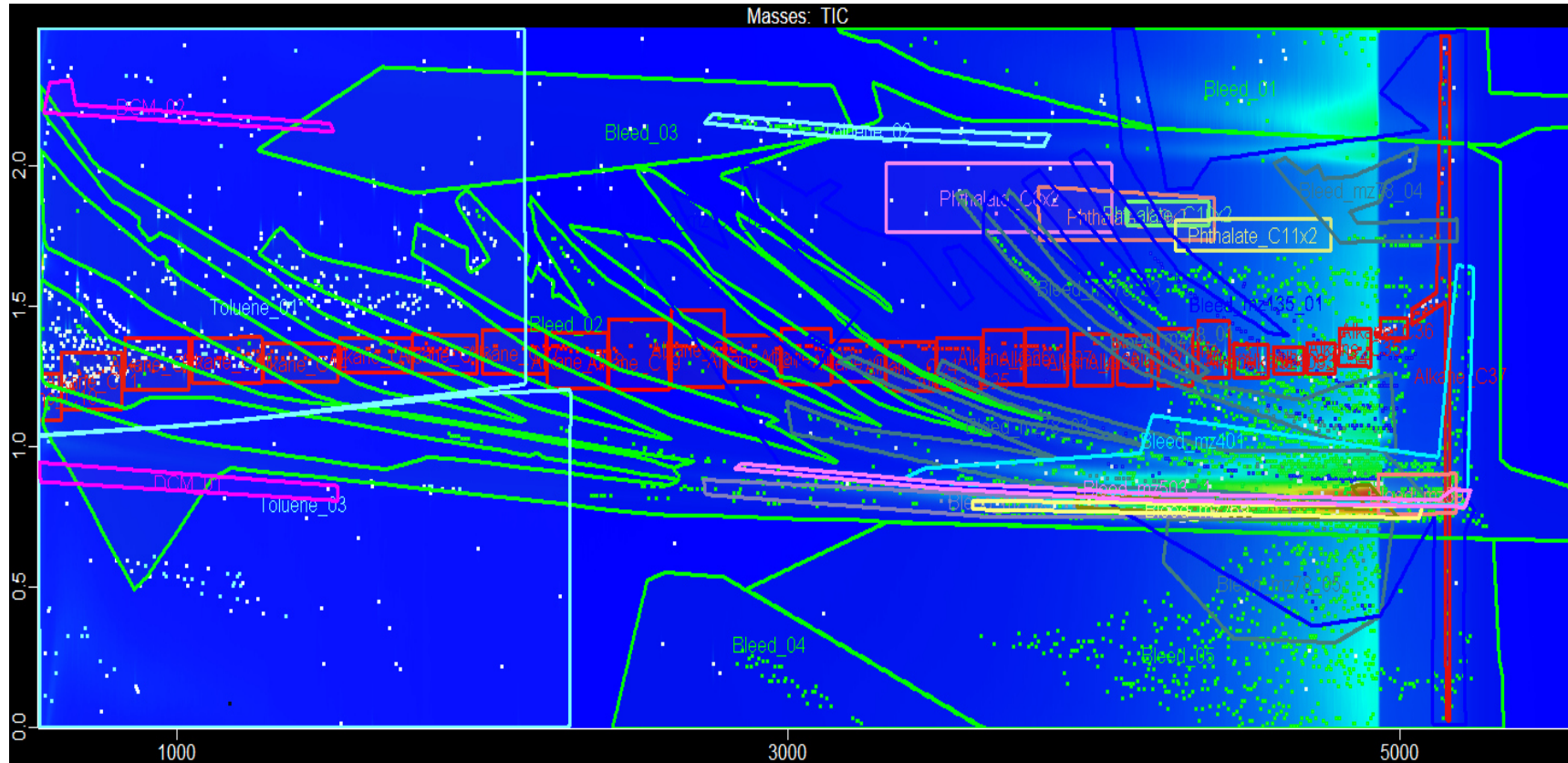
- LECO ChromaTOF (v. 5.51)
- NIST23 spectral library



Column bleed classification

- LECO ChromaTOF (v. 5.51)
- Spectral filters
- Chromatographic regions

Column bleed classification



Peak Filter

Save/Open Filters

Load Filter Set

Save Filter Set

Select Predefined Filter

Select a Template...

Peak Type Peak Properties Chlorine / Bromine Spectral

Build Statements:

Type: Compare by Response Response: Relative Abundance

Mass

See Mass Syntax

Add Replace

Not

Statements:

M/z 73.1±0.11 Da is = Base Mass

OR

M/z 147.05±0.11 Da is = Base Mass

OR

M/z 207.07±0.2 Da is = Base Mass

OR

M/z 267.03±0.15 Da is = Base Mass

OR

M/z 281.07±0.21 Da is = Base Mass

OR

M/z 341.12±0.15 Da is = Base Mass

AND / OR

Edit

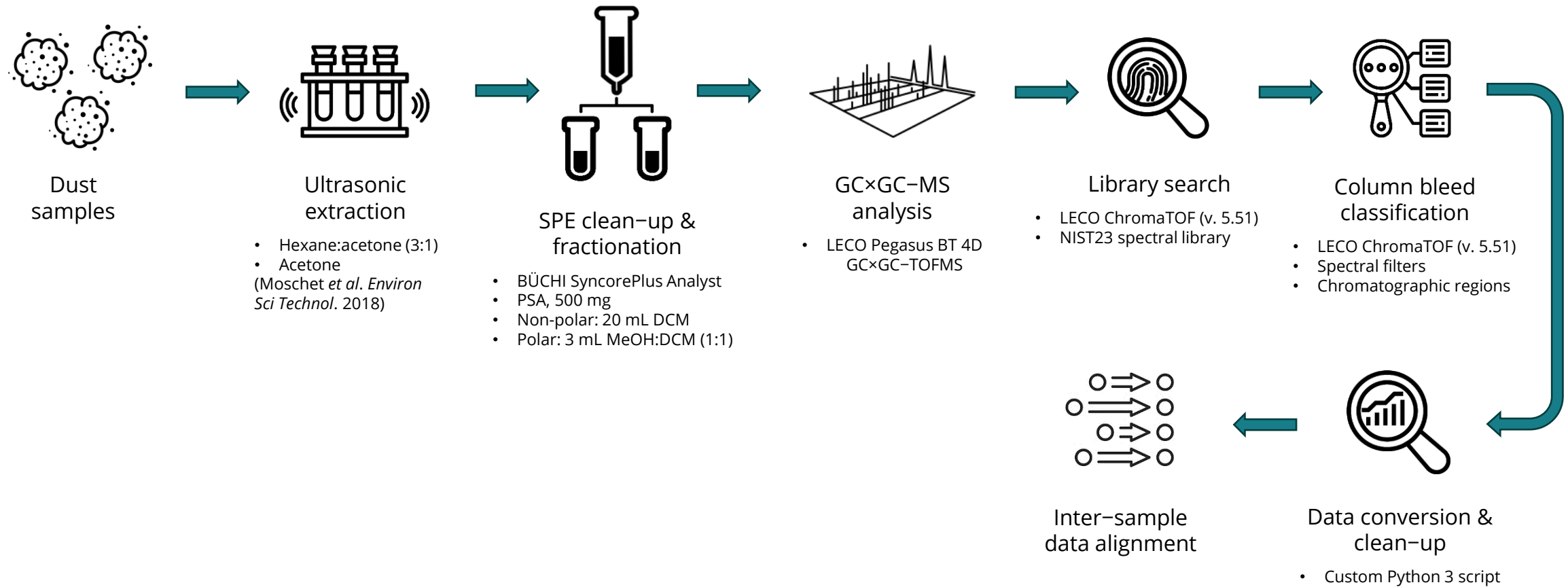
Delete

Promote

Demote

OK Cancel

Workflow



Picture source: flaticon.com, freepik.com

Data alignment

Compound name	RT1, min	RT2, min	RTI	CAS#	Spectrum	Area, S1	Area, S2	Area, S3	Area, S4	...	Area, SN
Triacetin	1378	0.5230	1346	102-76-1	103:1003...	N/A	N/A	0.64	N/A	...	N/A
Tributyl acetylcitrate	3236	2.0082	2260	77-90-7	185:1003...	0.77	9.35	199.31	1.80	...	8.02
Triethyl citrate	2100	2.4418	1660	77-93-0	157:1000...	0.30	0.53	2.48	0.33	...	0.52
Cyphenothrin	3863	2.0310	2668	39515-40-7	123:1000...	N/A	N/A	N/A	0.86	...	N/A
Deltamethrin	4393	2.1760	3068	52918-63-5	181:1000...	N/A	N/A	N/A	9.37	...	N/A
Fenvalerate	4255	2.1660	2959	N/A	125:1021...	1.11	N/A	N/A	N/A	...	N/A

for **feature_X** in sample and **feature_Y** in ref:

if **RT1_X_sample** == **RT1_Y_ref** ± ΔRT and

if **RT2_X_sample** == **RT2_Y_ref** ± ΔRT and

if **Spectrum_X_sample** is 75% similar to **Spectrum_Y_ref**

then **feature_X** = **feature_Y** and

RT1_XY = average(**RT1_X_sample**, **RT1_Y_ref**)

RT2_XY = average(**RT2_X_sample**, **RT2_Y_ref**)

Spectrum_XY = average(**Spectrum_X_sample**, **Spectrum_Y_ref**)

Algorithms:

- DP / NDP (NIST)
- BiPACE 2D
- DISCO
- MSort
- PAM
- etc.

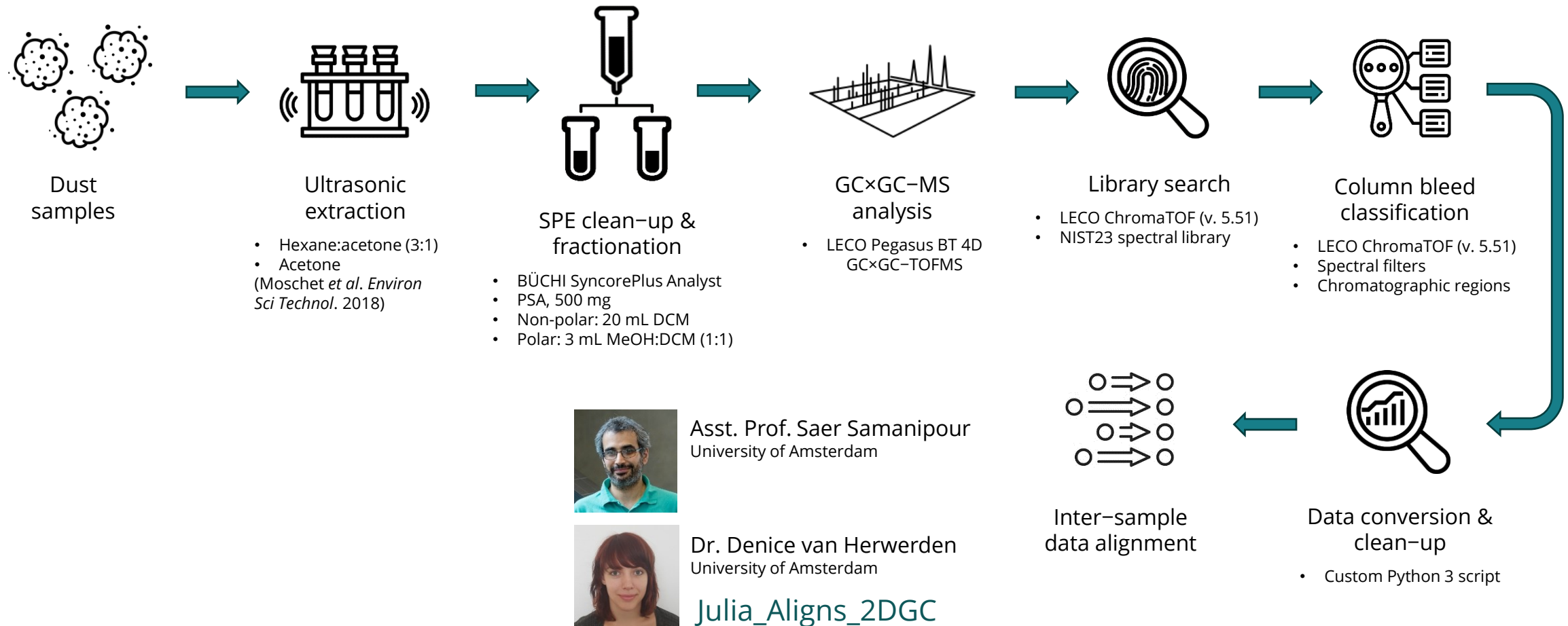
Data alignment software testing

#	Software / script, language	Issues	Reference / vendor
1	DA_2DCHROM , Python 3 (version 3.9.7)	- Alignment is done using template file (target list match). - Separate sample files are created as an outcome but no aligned master table.	Ladislavová <i>et al.</i> 2023. https://doi.org/10.1007/s00216-023-04679-7
2	GasPedal (version 1.1.2), Java	- Imported .cdf or .mzML files (both exported from LECO ChromaTOF v5.51.50.0 for Pegasus BT) look odd and do not resemble initial chromatograms: Since chromatograms are treated as images, this is a problem. - Many functions are unavailable and cannot be evaluated with demo license.	DECODON Software UG, Greifswald, Germany. https://www.decodon.com/support/gaspedal.html
3	GC Automation , Python 3 (version 3.8.4)	- GC Image deconvolution algorithm picks mostly background features which are high in abundance. - Alignment is done using template file (target list match). - Blobs not peaks are detected and aligned.	Wilde <i>et al.</i> 2020. https://dx.doi.org/10.1021/acs.analchem.0c02844
4	GC2MS , Ruby (version 2.1.2p95), Rails (version 4.1.6)	- Does not work even with the test data set provided: The web application returns an error when the analysis is started.	Tian <i>et al.</i> 2016. https://doi.org/10.1021/acs.analchem.6b00755
5	Guineu (version 1.0.3), Java	- Some functionalities got deprecated as things changed in the past 14 years, e.g., Golm database search. - Since search windows/thresholds are user-defined, a candidate peak may be discarded from the possible alignment if the spectral similarity is less than the threshold or the identification of the peak is conflicting with the path.	Castillo <i>et al.</i> 2011. https://doi.org/10.1021/ac103308x

#	Software / script, language	Issues	Reference / vendor
6	MetPP , MATLAB (version 2010b)	- Does not start properly if newer version of MATLAB is used. - Does not work for .txt files exported from LECO ChromaTOF Software for Pegasus BT (version 5.51.50.0): "Performing full alignment" step never concludes and errors appear in the MATLAB environment. - A derivatization method must be specified: Might be the reason why MetPP did not work as derivatization method was not specified.	Wei <i>et al.</i> 2013. https://doi.org/10.1093/bioinformatics/btt275
7	MOREPEAKS , MATLAB (version 2019a)	- Allows alignment for two samples only. - "Automated peak pairing" does not work for .cdf files exported from LECO ChromaTOF Software for Pegasus BT (version 5.51.50.0) or converted .mzXML files.	Molenaar <i>et al.</i> 2021. https://doi.org/10.1016/j.chroma.2021.461922
8	R2DGC , R	- Does not work for .cdf files exported from LECO ChromaTOF Software for Pegasus BT (version 5.51.50.0).	Ramaker <i>et al.</i> 2018. https://doi.org/10.1093/bioinformatics/btx825
9	RGCxGC , R	- Does not work for .cdf files exported from LECO ChromaTOF Software for Pegasus BT (version 5.51.50.0): The software does not "see" the imported files. However, the software works for the test .cdf files.	Quiroz-Moreno <i>et al.</i> 2020. https://doi.org/10.1016/j.microc.2020.104830
10	RMet , RStudio (version 1.1.383), R core (version 3.4.3)	- Does not work for .cdf files exported from LECO ChromaTOF Software for Pegasus BT (version 5.51.50.0): "Reconstructing cdf" step never concludes. However, the software works for the test .cdf files.	Moayedpour and Parastar. 2019. https://doi.org/10.1016/j.chemola.2019.103866

Rebryk and Haglund. *Unpublished data.* 2025

Workflow



Picture source: flaticon.com, freepik.com

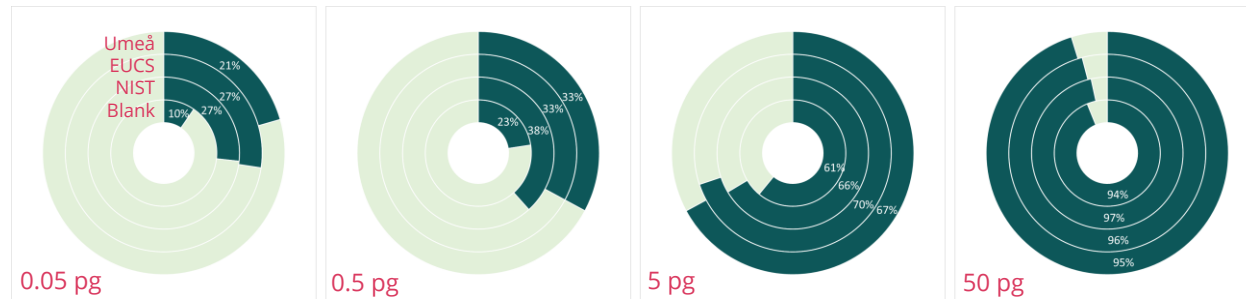
Data alignment software testing

Before review

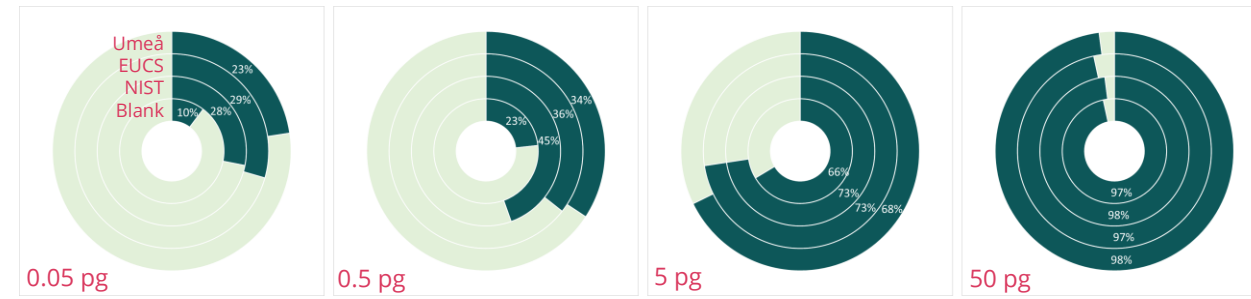
152 spiked compounds
146 detectable compounds

After review

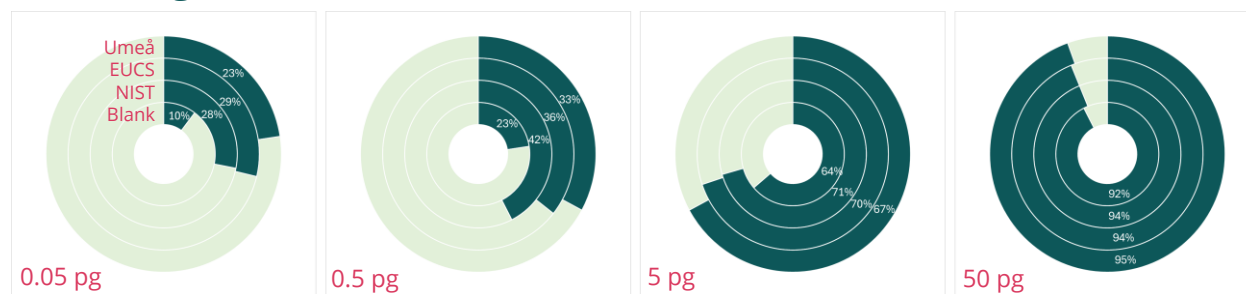
Guineu



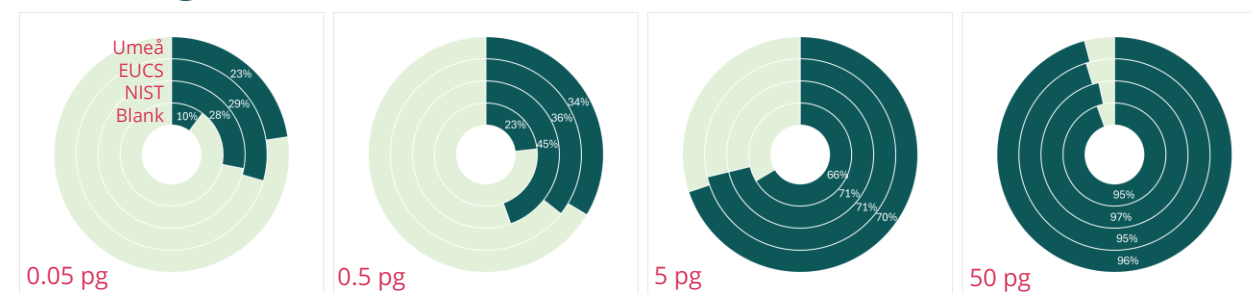
Guineu



Julia_Aligns_2DGC



Julia_Aligns_2DGC



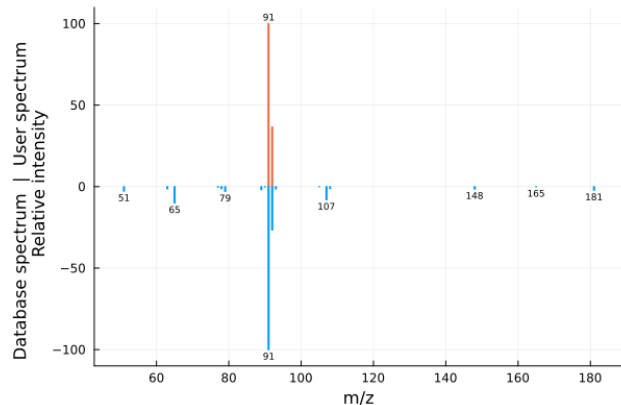
Rebryk and Haglund. *Unpublished data*. 2025

Julia_Aligns_2DGC.jl

Function	Action
TXT reader	Reads TXT (tab delimited) file, e.g., LECO ChromaTOF result file
CSV conversion	Converts TXT files to CSV files
ChromaTOF bleed classification filter	Removes features (rows) for which "Classifications" from ChromaTOF contains 'Bleed', 'Toluene', or 'DCM'
Compound name filter	Removes features (rows) for which "Name" contains unwanted values, e.g., siloxanes & other bleed features
Compound name & base mass filter	Removes features (rows) for which "Name" contains compounds with spectra similar to, e.g., toluene (m/z 91 or 92) if their spectra do not contain other m/z values, like molecular ions that do not belong to toluene (m/z > 91 or 92)
Zero intensity filter	Removes m/z:intensity pairs with intensity values containing '0.' from spectra
Low intensity filter	Removes m/z:intensity pairs with intensity values \leq 3% of the highest intensity value
Alignment	Aligns samples based on specific RT1 & RT2 window and spectra similarity (DISCO algorithm; other algorithms, e.g., NDP, can also be used)
Postalignment merging	Merges features with the same name and base m/z if Δ RI < 5
Library search	Searches library using the same algorithm as for the alignment
Library match visualization	Creates head-to-tail graphs of user vs. library match features

Julia_Aligns_2DGC.jl

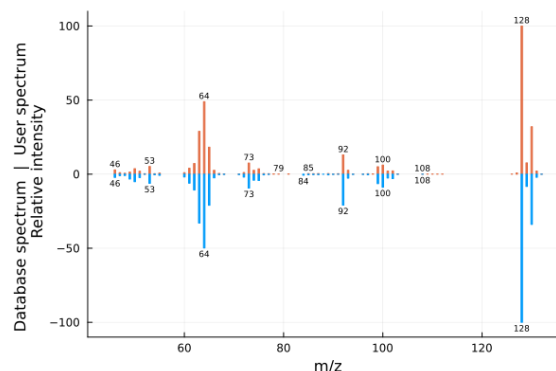
Nr	MinRt1 D	MaxRt1 D	AveRt1 D	MinRt 2D	MaxRt 2D	AveRt 2D	MinR I	MaxR I	Med RI	MinR Ilib	MaxR Ilib	Med RIlib	Det ecti onN um	Det ecti onFr eq	Name	Formula	InChIKey	CAS	BaseM ass	Origin	Class	Simil arity	Prob abilit y	dRI	Scor e	Spectrum _BestMatc h	Spectrum _Consens us	Sample 1	Sample 2	Sample 3	Sample 4	Sample 5
1	645.0	645.0	645.0	1.230	1.231	1.231	1017	1017	1017	1010	1011	1011	2	29	(+)-3-Carene C10H16	BQOFWKZ	498-15-7	93.09	NA	NA	0.77	0.08	0.97	0.61	51:56 52:3	51:27 52:1	0	0	4.33E+06	0	0	
2	4517.8	4517.8	4517.8	2.032	2.032	2.032	3170	3170	3170	3175	3175	3175	1	14	(+)-Sesamin C20H18O6	PEYUIKBAA	607-80-7	149.1	NA	NA	0.87	0.72	0.98	0.85	53:34 63:1	53:34 63:1	0	0	0	2.59E+07	0	
3	1782.6	1782.6	1782.6	1.447	1.447	1.447	1517	1517	1517	1482	1482	1482	1	14	(+)-epi-Bic C15H24	RNDFUOKI	54274-73-	161.2	NA	NA	0.73	0.03	0.86	0.54	91:114 105	91:114 105	0	0	0	0	0	
4	630.0	632.5	631.3	1.395	2.469	1.932	1009	1010	1010	NaN	NaN	NaN	7	100	(1,3,3,3-Te C11H7F7S	ZVZRUXCG	75667-95-	91.09	NA	NA	0.7	0.02	0	0.24	50:120 51:	50:120 51:	0	0	4.57E+07	2.88E+07	0	



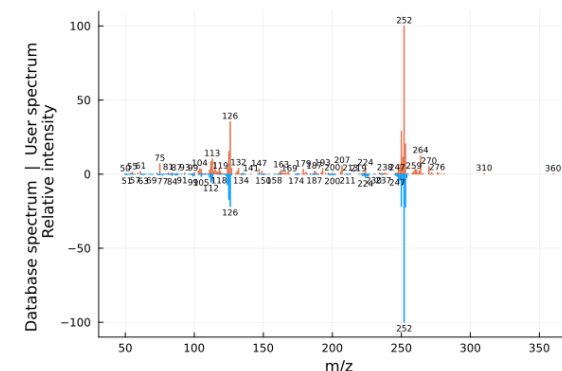
Toluene →

1-Bromo-2-benzyloxybenzene (m/z 107)

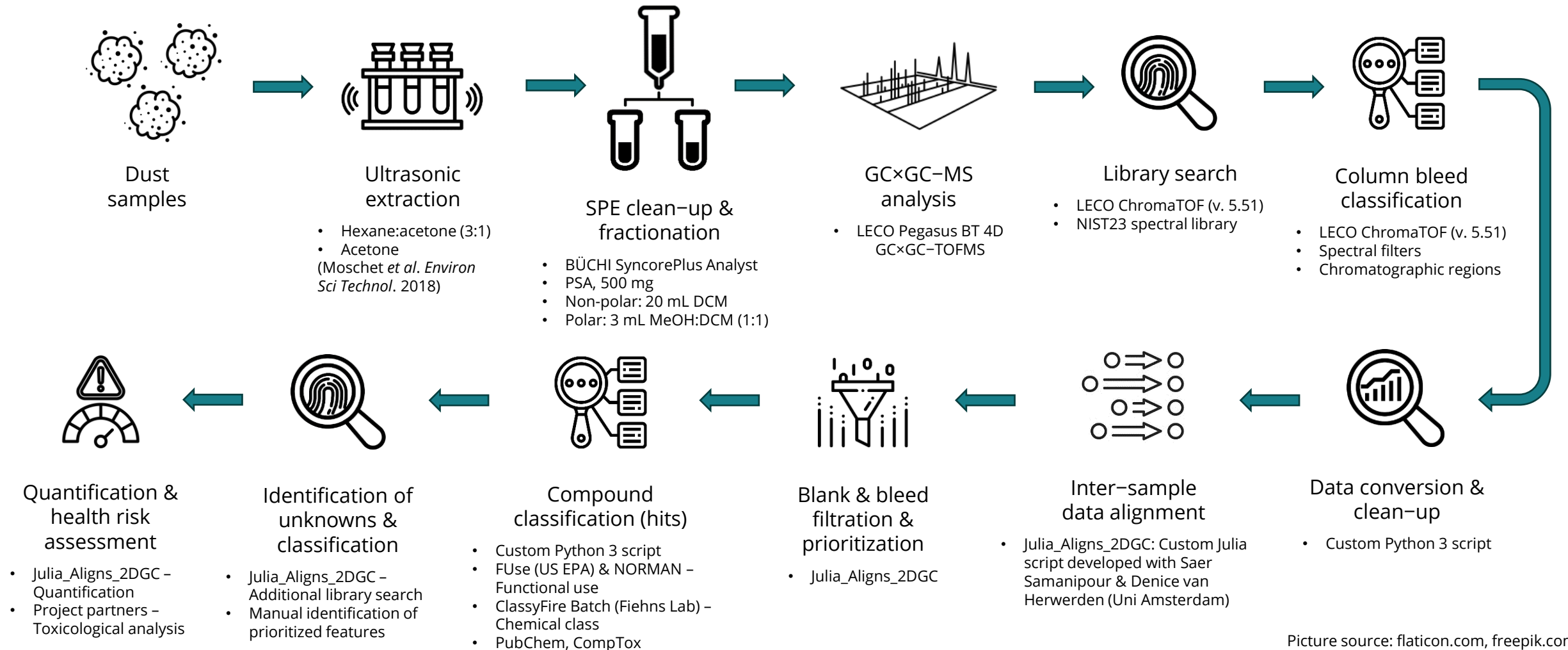
Align_ Num	DB_Num	minRI	maxR I	DB_RI	Dot	Base Mass	Origin	Class	Align_Name	DB_Name	DB_InChIKey	DB_CAS	DB_Formula	Spec	DB_spec	Graph
1	19092	910	1117	969	0.93	93.09	NA	NA	(+)-3-Carene	Cyclopropane, 1,1-di	RESIMHRPGAN	68998-21	C10H16	51:27 52:1	50:20 51:3	file:///F:\Projects\
1	18828	980	1047	1010	0.93	93.09	NA	NA	(+)-3-Carene	(+)-3-Carene	BQOFWKZOCN	498-15-7	C10H16	51:27 52:1	50:100 51:	file:///F:\Projects\
1	19009	980	1047	1028	0.93	93.09	NA	NA	(+)-3-Carene	Cyclohexene, 4-metf	SCWPFSIZUZU	99-84-3	C10H16	51:27 52:1	51:500 53:	file:///F:\Projects\
2	293359	3140	3205	3175	0.98	149.1	NA	NA	(+)-Sesamin	(+)-Sesamin	PEYUIKBAABKC	607-80-7	C20H18O6	53:34 63:1	50:170 51:	file:///F:\Projects\
2	293376	3070	3275	3149	0.9	149.1	NA	NA	(+)-Sesamin	2,6-Bis(3,4-methylen	PEYUIKBAABKC	7076-24-	C20H18O6	53:34 63:1	50:110 51:	file:///F:\Projects\
2	353011	3140	3205	3205	0.79	149.1	NA	NA	(+)-Sesamin	Phthalic acid, decyl	CXYLYGHIEBBL	1000377-	C26H34O5	53:34 63:1	50:40 51:3	file:///F:\Projects\
3	87033	1399	1487	1444	0.98	69.1	NA	NA	(E)-β-Farnesene	cis-.beta.-Farnesene	JSNRRGGBADV	28973-97	C15H24	51:92 53:1	50:100 51:	file:///F:\Projects\
3	87127	1399	1487	1458	0.98	69.1	NA	NA	(E)-β-Farnesene	1,6,10-Dodecatriene	JSNRRGGBADV	77129-48	C15H24	51:92 53:1	45:5 50:14	file:///F:\Projects\
3	87035	1399	1487	1457	0.97	69.1	NA	NA	(E)-β-Farnesene	(E)-.beta.-Farnesene	JSNRRGGBADV	18794-84	C15H24	51:92 53:1	50:50 51:2	file:///F:\Projects\



Phenol, 2-chloro- → Phenol, 2-chloro-

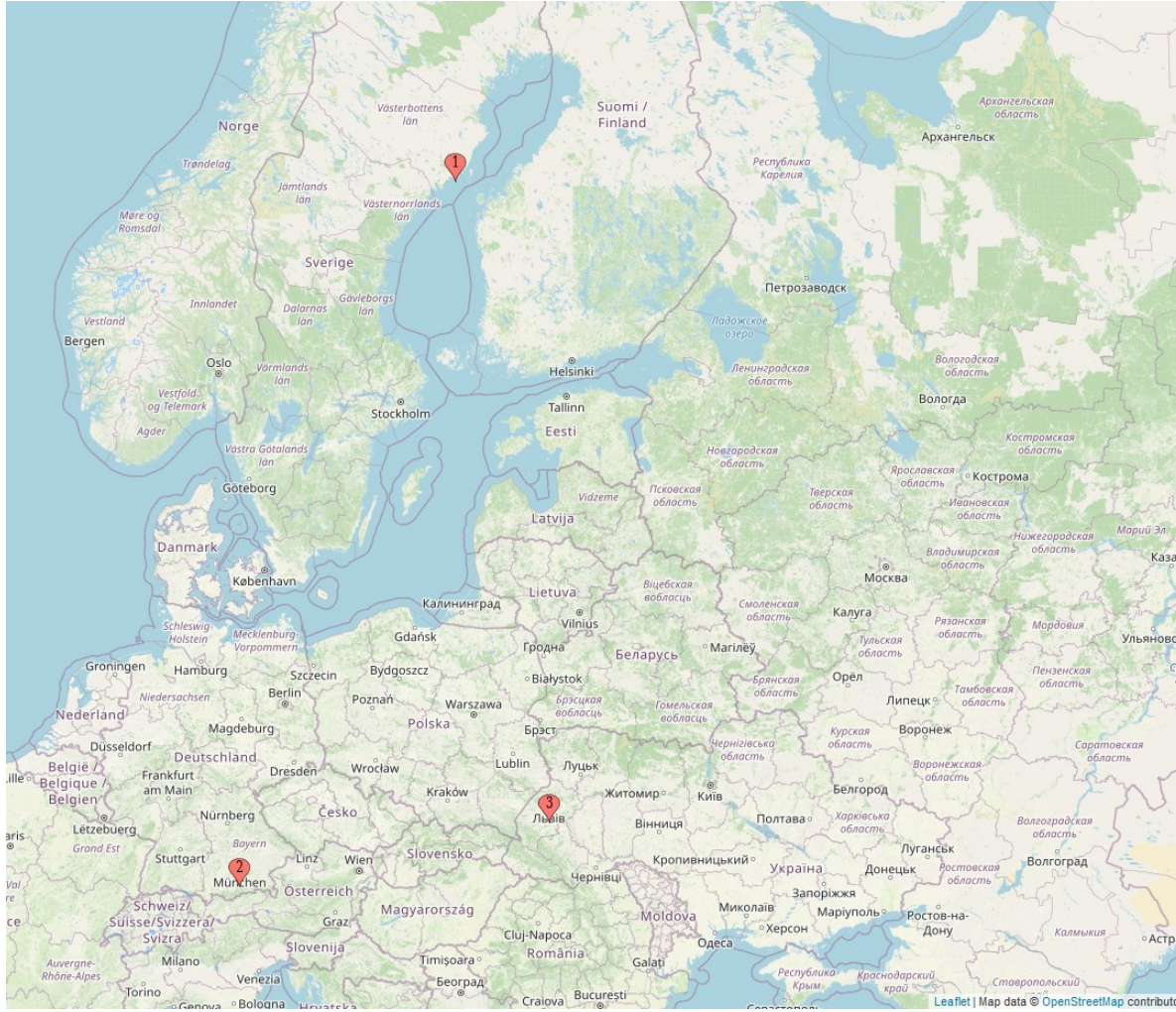


Workflow



Picture source: flaticon.com, freepik.com

Samples



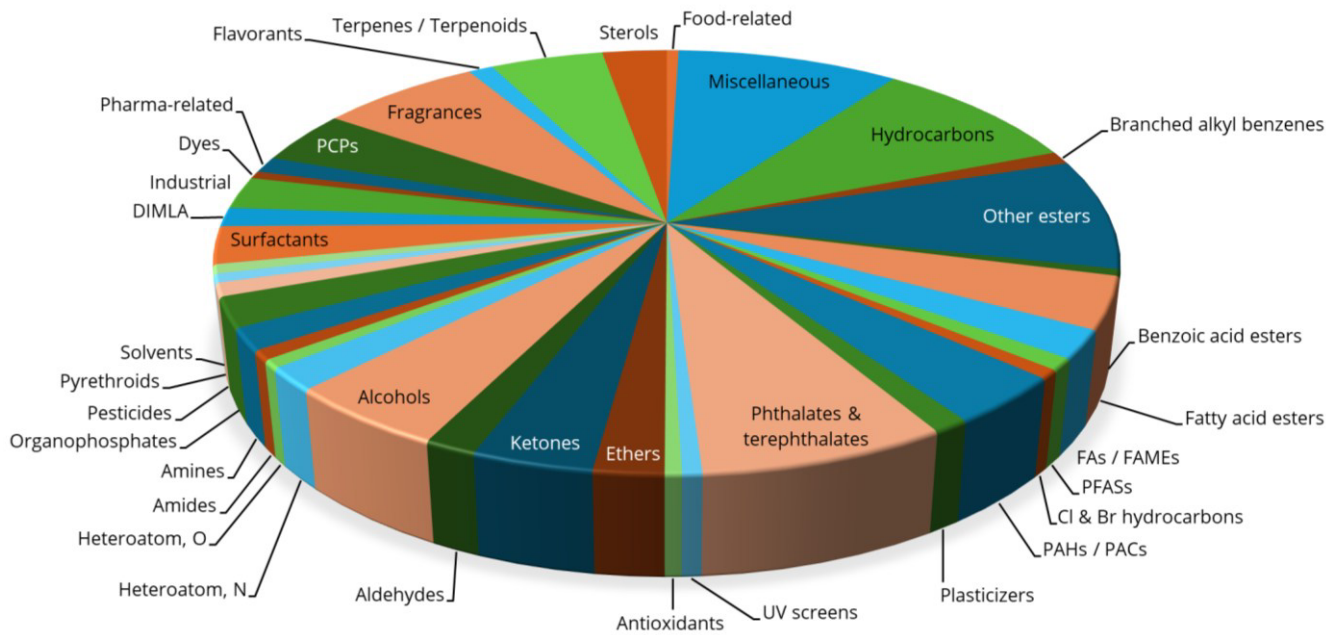
#	City, Country / Other ID	Code	Weight, mg
1	Umeå, Sweden	UMPS	60.9
2	Munich, Germany	MUPS	62.0
3	Lviv, Ukraine	LVPS	62.1

4	NORMAN EU composite sample	EUCS	50.6
5	NIST SRM 2585 (US)	NIST	50.7
6	Laboratory blanks	BLNK	---

Picture source: mapcustomizer.com

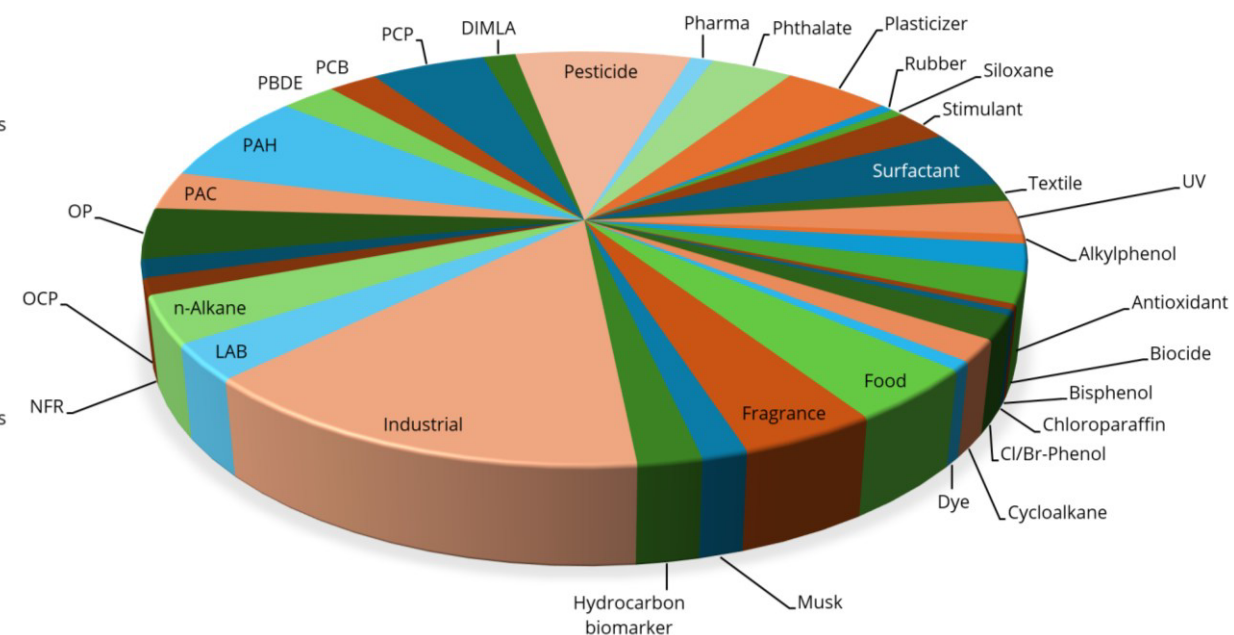
Detected compound classes

GC×GC-MS compounds
(3160 features, excl. unknowns)



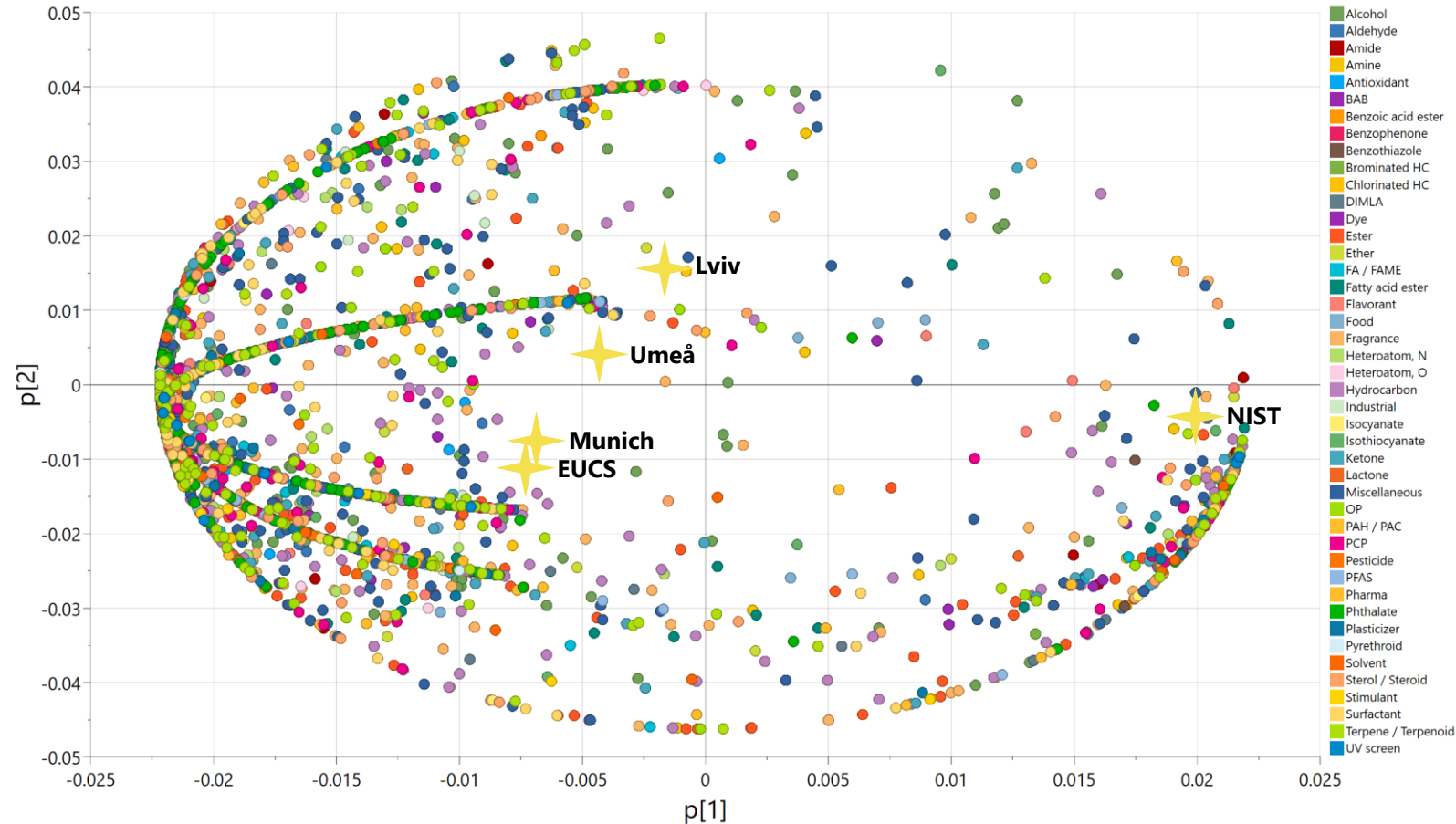
Rebyk and Haglund. *Unpublished data*. 2025

GC / GC×GC-MS compounds
(648 unique features)

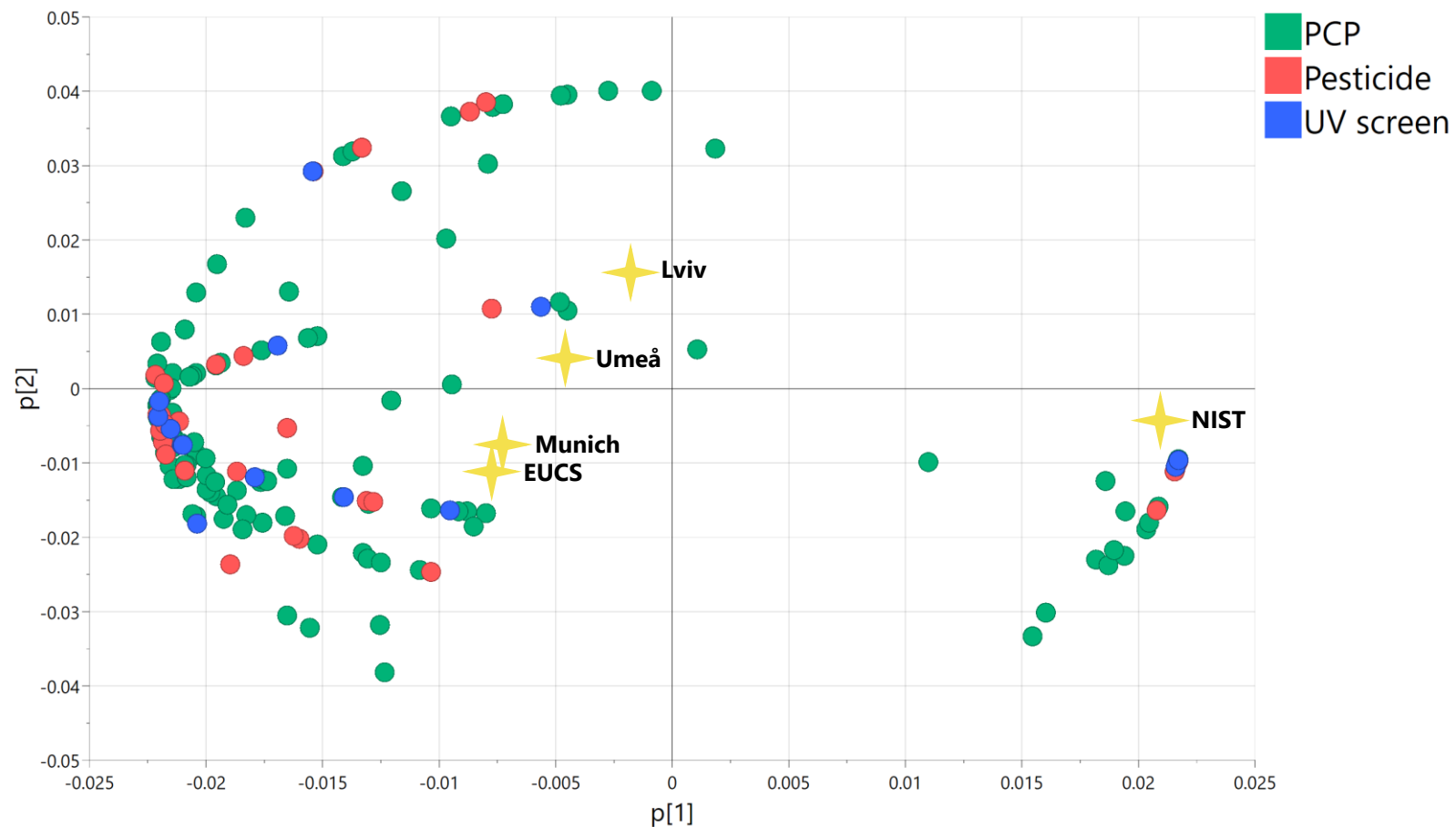


Haglund et al. *Sci Tot Env*. 2024

House dust, Fr. 1 & Fr. 2 – PCA



House dust, Fr. 1 & Fr. 2 – PCA



Conclusions & future work

- A fast highly-automated **data processing workflow for GC×GC-MS data** was created and will be upgraded & developed further
- **GC×GC-MS analysis** revealed presence of incompatibly higher number of features than previous **GC-MS analysis**
- **European house dust** contains enormous number of various compounds:
 - **ca. 3,100 features** among the NIST23-matched features
 - **ca. 6,500 unknown features**
- **Identification** of chemicals of potential concern:
 - Julia_Aligns_2DGC library search & classification
 - Prioritization – toxicology partners
- **Comprehensive (semi-)quantification at different levels:**
 - Reference standards for target compounds
 - Internal standards for suspects

Acknowledgments



Prof. Peter Haglund
Umeå University



Asst. Prof. Saer Samanipour
University of Amsterdam



Dr. Denice van Herwerden
University of Amsterdam

Thank you for your attention!



This work was co-funded by UK Research and Innovation (UKRI) under the UK government's Horizon Europe funding guarantee [grant number 10038689 and 10042425].



Co-funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Health and Digital Executive Agency (HaDEA). Neither the European Union nor the granting authority can be held responsible for them. [grant agreement Project No 101057499 - INQUIRE.]



The work received co-funding support from the Australian National Health and Medical Research Council (NHMRC) [NHMRC; 2022/GNT2017837].



www.inquire-he.eu



Andriy Rebryk
andrew.rebryk@gmail.com

